AGARD-CP-66-70

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

AGARD CONFERENCE PROCEEDINGS No. 66

on

Advanced Radar Systems

NORTH ATLANTIC TREATY ORGANIZATION

AGARD Conference Proceedings No.66

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

ADVANCED RADAR SYSTEMS

Papers presented at the XIXth Meeting of the Avionics Panel of AGARD,
held in Istanbul, Turkey, 25 - 28 May 1970

The material in this pulication has been produced directly from copy set and supplied by each author

Published November 1970

621.396.969

Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte St. London W1P 1HD

PROGRAM CHAIRMAN
EDITOR-COLLATOR

Mr Roger Voles
E.M.I. Electronics Ltd
Blyth Road
Hayes, Middlesex, UK

AVIONICS PANEL CHAIRMAN

Dr Ing. Franco Gnavi
FIAT - Divisione Aviazione
Centro Elettronico Avio
I 10072 Caselle Torinese
Torino, Italy

HOST NATION COORDINATOR

Mr Turgut Ortac
Research and Development Department
Turkish General Staff
ARGE
Genel Kurmay
Baskanligi, Ankara, Turkey

AVIONICS PANEL EXECUTIVE

Cdr C.R.Smith, USN
AGARD

FOREWORD

During the last few years considerable advances have been made over the whole range of components used in radar systems and a number of important new devices have arrived on the scene

These developments have not only led to the realisation of many system techniques that until recently have existed only in theory but they have also introduced a number of important new principles.

It was in order to provide an opportunity to hear about and discuss these new advances that the Avionics Panel of AGARD decided to devote its XIXth Technical Symposium to this topic. 38 papers were presented to 140 participants and a number of lively and informative discussions ensued. The plan of these proceedings follows the conference programme and the records of the discussions appear immediately after the relevant paper.

The getting together and selection of the papers involved a considerable amount of work and I am indebted to the Programme Committee for its effort. In particular, I gratefully acknowledge the help given by Mr J.N.Bloom, Professor D.Bosman, Professor M.Carpentier, Mr J.Freedman, Ir F.Gnavi and Professor Dr G.Ulbricht.

Roger Voles
Programme Chairman and Editor

CONTENTS

SESSION III - SIGNAL PROCESSING

SESSION IV - DATA PROCESSING

SESSION V - SYSTEMS

TECHNICAL EVALUATION REPORT
ON
ADVANCED RADAR SYSTEMS

A.K.Edgar*

1. INTRODUCTION

This XIXth Technical Symposium on "Advanced Radar Systems" was organised by the Avionics Panel of AGARD. It was held at the Grand Hotel, Tarabya, Istanbul, Turkey between 25th and 29th May, 1970. The Program Committee was chaired by Mr R.Voles. This report is an evaluation and assessment of the scientific and technical content of the meeting and aimed at assisting AGARD in planning follow-on activities and future meetings.

The Program Committee is to be congratulated on providing a well balanced and valuable program concerning subjects of vital concern to any radar system designer. The subject matter was covered under the headings of.

(a) Clutter and Targets.

(b) Micro-wave Sub-systems.

(c) Signal Processing.

(d) Data Processing.

(e) Systems.

This report is arranged under the same headings.

2. SESSION I - CLUTTER AND TARGETS

There were 10 papers in this Session, 6 on measurements of returns from clutter (sea, ground and precipitation) and aircraft and 4 theoretical papers on a variety of topics. The opening paper[1] reported on airborne measurements and returns from the sea under high wind and wave conditions. These measurements were made to test the predictions of a model for the radar cross-sections of sea returns and gave good agreement for vertically polarised radiation, although figures greater than predicted were obtained for horizontally polarised radiation. The follow-up paper on results under less extreme conditions will be awaited with interest. Paper 2 discusses results on returns from ground and precipitation clutter with a 10 cm surveillance radar. These results were presented in the form of probability distribution curves for the radar cross-section per resolution cell and for ground clutter showed a log-normal shape. After suitable processing to remove the terrain factor, the probability distributions of the resulting signals from different sites showed similar characteristics. Paper 3 suggested a new model for the calculation of the mean doppler speed and standard deviation of the velocity spectrum of rain clouds. The second part of the paper described a portable instrument for the measurement of clutter spectra and gave some results obtained with it. Paper 4 reported on measurements of radar angels. It gave interesting results with B-scope displays and correlated the pattern of echoes with insect and bird behaviour. Von Schlachta[5] reported on the amplitude and phase shift of echoes over one scan from a variety of targets using a digital recording system and from repeated measurements gave both mean values and fluctuations for different types of targets for both amplitude and doppler frequency, over one scan and over a large number of scans. The last paper[6] from ASWE, Portsmouth, discussed the effects of frequency agility on aircraft using a C-band radar. They measured the root mean square pulse-to-pulse fluctuation for a Sea Hawk and a Comet for different frequency changes. For a Sea Hawk the maximum r.m.s. value was 6dB for a 13 MHz shift, whereas the returns from the Comet were decorrelated at a fixed frequency.

Four theoretical papers followed. The first[7] by Swerling modified the "Swerling models" to include broader families of models whose statistical behaviour is governed by chi-square, log-normal and Rice distributions. It gives a brief discussion on the significance of various target models for various types of receiver processing. Paper 6 deals with amplitude comparison monopulse radar in a cluttered environment. The generalised likelihood ratio test is used to derive the detection and parameter estimation strategy for the radar receiver. Expressions for the detector and angle estimates are derived for a number of cases. For the conventional sum and difference beam-case, an approximation to the likelihood receiver is derived and shown to be a much more easily realisable structure. Paper 7 deals with tracking errors in a frequency agile system. It concludes that frequency agility gives a reduction of tracking errors for a monopulse radar. With conical scan, while the glint error is always reduced the scintillation error may be increased. Paper 10 by Rihaczek discusses the significance of coherent and incoherent bandwidth in radar resolution.

3. SESSION II - MICROWAVE SUB-SYSTEMS

Lemke, Holst and Schmitt[11] produced an interesting paper covering microwave integrated circuits allowing planar integration of magnetic system functions. They illustrated designs for phase shifters, switches and

* Superintendent of the SR and ATC Group of the Royal Radar Establishment of the Ministry of Technology, UK. Assisted by M.P.Warden also of the RRE.

circulators and showed a gun oscillator doppler radar on a 3 cm by 3 cm composite ferrite chip. This paper and a later paper[14] by Sterzer will be of interest to engineers concerned with phased arrays. Sterzer describes solid state microwave sources of P max ~ 30 w at L-band with 40% efficiency and 15 watts at S-band.

Pierrot[15] discussed the computer design of radomes giving very good agreement between theory and practice. The aim is to optimise the design for given electro-magnetic aero-dynamic and mechanical illustrations.

Papers 14 and 15 cover systems using electronic scanning. The former describes a process of space scanning by an array in which the radiation pattern is time modulated using 1 bit ($\theta = \pi$) phase shifters. From the energy point of view there is a loss of about 3dB in the best case. Paper 15 described a radar system with flood-lighting and Nyquist rate scanning at S-band. It gives the possibility of good M.T.I. performance combined with a very high data rate, a combination not possible with non-floodlighting systems. This type of system uses energy efficiently but calls for more complex data processing. Robinson[16] points the way to highly competitive phased array radar designs and suggests that recent technology trends will provide cost effective solutions. He suggests that there is a future for time and frequency sharing of phased arrays between a multiplicity of systems, particularly in aircraft, spacecraft and ships. Simpson[20] also forecasts an increasing impact for phased arrays in lightweight systems for the battlefield and in aircraft. He describes a very compact X-band airborne weather radar for aircraft scanning ±45°, the sector scanned being selected by the pilot.

4. SESSION III - SIGNAL PROCESSING

Papers in Session III on Signal Processing covered pulse compression, correlators and M.T.I. systems.

Forestier and Chevalier[22] described a system combining pulse compression and correlation giving an output requiring narrower video bandwidths than the conventional pulse compression radar. Herring[23] gave an excellent paper on dispersive networks and highlighted the practical problems. He gave examples covering bandwidths up to 500 MHz.

Papers 24-31 covered digital correlators and M.T.I. systems in digital or hybrid form. We make no attempt to summarise individual papers but would point out that the whole session makes valuable reading for workers in the signal processing field. A general observation is that signal processing techniques are aiming at greater efficiency by taking advantage of new technology. The net result is a more complex system which would not have been possible with analogue circuitry.

5. SESSION IV - DATA PROCESSING

Whereas in the past circuits and techniques have been the life-blood of the radar system, data processing is now of equal, or perhaps greater, importance. This change is reflected in Session IV where the emphasis in systems is on the computer and digital processes within the radar. Paper 32 discusses this change in some detail. Digital or computer control may include beam positioning, target acquisition, tracking and the operation of test and diagnostic programmes. It points out that while hardware costs are reducing there has been a large increase in programming, which is now a major item in system development.

In Paper 33 Bijvoet discusses the merits of computer control for a search track radar with full or partial electronic scanning. He concludes that computer control systems must be exploited to the full to provide cost effective radars having acceptable volume cover and shows that very narrow pencil beam radars (typically $\frac{1}{2}^{0} \times \frac{1}{2}^{0}$) with their advantages in tracking accuracy and clutter rejection can be effective in the surveillance mode.

Papers 36 and 37 deal with tracking strategy and logic, mainly from a theoretical viewpoint. Ebert[38] gave a most interesting paper on a digital plot extractor. His system correlates primary and secondary signals and allows display information to be carried on two or three 3 KHz telephone cables. The detection process includes sliding window integration and clutter suppression circuits aimed at giving a substantial rain clutter reduction for a system loss in sensitivity of about 0.6dB.

Paper 39 by Smit describes an experimental noise radar - RUDAR - in which the noise spectrum is emitted via a dispersive antenna to spread the transmission in azimuth. Reception via suitable matched filters provides angular information while range is given by a correlation technique. A range accuracy of the order of 1 meter and a resolution of about 30 meters are claimed.

Trials of a terrain avoidance radar for helicopters are described by Collet[40]. It operates at 35 GHz with a pulse width of 0.2 μS. Large obstacles, hills, pylons, and power cables are indicated, although in the latter case indication is not continuous along the cable length. The display is easy to interpret and enables the pilot to avoid obstacles by change of direction or by alteration of flight level.

In the final paper[42] Kerr, Ministry of Defence, UK, presented a cost effectiveness study of an Airborne Early Warning System. On the assumption that poor low level cover from ground based surveillance systems will be exploited by the enemy he shows that there are significant advantages to be obtained, operationally and economically, by adopting an A.E.W. system. He suggests that enemy action by physical attack or E.C.M. techniques do not negate the significant advantages of such a system.

6. CONCLUSIONS

The Symposium on "Advanced Radar Systems" gave the impression of steady progress rather than spectacular improvements - indeed, this is to be expected at this stage in radar development. While the papers presented covered almost the whole field of radar techniques there was greatest emphasis on the problems that would arise with automatic radar should these become possible. Thus the statistics of clutter and techniques of eliminating unwanted signals due to terrain, precipitation and Angels were always to the forefront, both in signal processing and in data processing. It was clear that the ultimate goal would be reached only by optimising system designs at all stages and ease of maintenance and reliability point to digital processing as the desirable solution.

Several papers emphasised the reduction in size and cost of components for phased arrays and showed that, with their application to radar, systems could be economically viable. Nevertheless, there was the warning that costs for an overall system, since most systems must include some degree of computer control or processing, are in danger of rising because of the increasing associated software costs.

Overall, a very worthwhile Symposium that provided a valuable exchange of information.

REFERENCES

1. Guinard, N.W. Ransome, J.R. Jr	*Radar Sea Return in High Wind/Wave Conditions.*
2. Warden, M.P.	*An Experimental Study of Some Clutter Characteristics.*
3. Kester, T. Einbu, J.	*Methods and Instrumentation for Assessment of Spectral Properties of Radar Clutter.*
4. Fishbein, W. et al.	*Some Measurements of Radar Angels (False Radar Targets).*
5. von Schlacta, K.	*Characteristics of Echoes from Aircraft and Clutter Derived from Digital Recorded Radar Data.*
6. Whitlock, W.S. et al.	*Some Measurements of the Effects of Frequency Agility on Aircraft Radar Echoes.*
7. Swerling, P.	*Recent Developments in Target Models for Radar Detection Analysis.*
8. Hofstetter, E.M.	*The Theory of Amplitude-Comparison Monopulse Radar in a Cluttered Environment.*
9. Guarguaglini, P.	*Tracking Errors in a Frequency Agility System.*
10. Rihaczek, A.W.	*The Role of Coherent/Noncoherent Bandwidth in Radar Resolution.*
11. Lemke, M. et al.	*Integrated Subsystems on Composite Substrates.*
12. Pierrot, M.R.	*Computer Aided Design of Airborne Radome.*
14. Drabowitch, S. Aubry, C.	*Pattern Compression by Space-Time Binary Coding of an Array Antenna.*
15. Edgar, A.K. Jones, I.L.	*Flood-Lighting with Nyquist Rate Scanning.*
16. Robinson, A.S.	*Phased Array Radars for Aerospace Applications.*
17. Reitzig,	*The Linear Array for Beam Steering.*
18. Sterzer, F.	*Microwave Solid State Power Sources for Airborne Phased Arrays.*
20. Simpson, M.	*Lightweight Phased Array Systems for Airborne and Ground Based Applications.*
22. Forestier, H.G.P. Chevalier, M.	*High Range Resolution Technique for Tracking Radar.*

23. Herring, P.G. et al.	*Low Distortion Dispersive Networks for Wideband Signal Processing Systems.*
24. Carre, M.R.	*Digital Correlation.*
25. Harris, D.S.	*Switched Capacitor Storage Arrays for A.M.T.I. and Bandwidth Compression.*
26. Jagger, C.E.	*Hybrid Analog - Digital Correlator.*
27. Reiss, A.	*A Digital M.T.I. Canceller.*
28. Forestier, H.G.P.	*A Digital Processing System for Monopulse Radar Signals in a Clutter Environment.*
29. Sirven, M.J.	*Digital Doppler Processor for Range and Radial Speed Measurement in a Coherent Low P.R.F. Radar.*
30. Wirth, W.D.	*Detection of Doppler Shifted Radar Signals with Clutter Rejection.*
31. Gold, B. Muehe, C.E.	*Fast Fourier Transforms for Range-Doppler Radar.*
32. Cokas, G. Parsons, H.	*Data Processing Considerations in Advanced Phased Array Radar Systems.*
33. Bijvoet, J.A.	*Operational Employment and Detection Performance of Advanced Computer Controlled Radar Systems.*
34. Cochran, D.J.	*Digital Plot Extraction of Primary Radar.*
36. van Keuk, G.	*Description of Tracking Programs for Phased-Array Radar.*
37. Both, W.A. Fleckner, H.	*Some Statistical Results Comparing Different Tracking Logics.*
38. Siegenthaler, U. Ebert, H.	*Digital Radar Plot Extractor - A System Module of a Semi-Automatic Air Traffic Control System.*
39. Smit, J.A.	*RUDAR - An Experimental Noise-Radar System.*
40. Collot, M.	*SAIGA - Radar d'Evitement d'Obstacle pour Helicoptere.*
42. Kerr, T.H.	*A Cost/Effectiveness Study of the Airborne Early Warning Systems as an Integrated Part of Air Defence Systems.*

RADAR SEA RETURN IN HIGH WIND/WAVE CONDITIONS

by

Norman W.Guinard and John T.Ransone, Jr

Naval Research Laboratory
Washington, D.C. 20390

RADAR SEA RETURN IN HIGH WIND/WAVE CONDITIONS

Norman W. Guinard & John T. Ransone, Jr.
Naval Research Laboratory
Washington, D.C. 20390

SUMMARY

Recently, scientists working in both the United States and the Soviet Union have developed a new model for the RCS of the sea which is based on the scattering from a composite surface. The model predicts an upper bound (saturation value) for the RCS and also yields expressions for its variation with frequency, polarization and grazing angle. To verify the predictions of the model under extreme conditions, an extensive measurement program was conducted over the North Atlantic in February 1969 using the NRL Four-Frequency Radar System installed in an EC-121 aircraft. Data were acquired in sea states with winds which varied from Beaufort 0 to 8 with a maximum observed sea state condition of a 48-knot wind with 26-foot SWH. The results of the measurement have verified the existence of the saturation condition for wind speeds above 10 knots as well as the variation of the RCS with frequency, polarization and grazing angle with the proviso that the tilting of the scattering surface by the swell be accounted for in the interpretation.

1. INTRODUCTION

In February 1969, the Naval Research Laboratory based its Four-Frequency Radar (4FR) System, which is a laboratory type measurement radar installed in an EC-121 (Super Constellation) aircraft, in Shannon, Ireland, for the purpose of collecting radar return data in the high sea states which are characteristic of the North Atlantic Ocean in the winter months. The experiment was extremely successful. In the first three weeks of the month, nine missions were flown to Ocean Stations India (59°N, 19°W) and Juliet (52.5°N, 20°W) and on seven of these, winds in excess of 20 knots were observed with a maximum wind/wave condition of 48 knots windspeed combined with a 26-foot significant wave height occurring on 11 February. There were three primary reasons for undertaking an experiment of this magnitude to document the behavior of radar return from high seas. The first was to satisfy the requirement generated by the continuing development of long-range ocean surveillance systems which have made the description of clutter backgrounds in terms of typical and worst case conditions, an increasingly important feature of system specification and design. The second was the desire to measure the variation of the Radar Cross Section (RCS) of the sea with increasing roughness to determine the range over which active radar could be expected to operate effectively in fog, haze and sea clutter. The final reason was the continuation of a study of the new composite surface model for sea clutter whose predictions required verification under the extreme scattering conditions found in the higher sea states. Since the new model provides a theoretical basis from which both the worst case clutter condition and the behavior of the RCS with increasing roughness may be inferred, it offers a convenient starting point in the discussion of the high sea state measurements.

The model which has been formulated by Wright (1) and independently by Bass, Fuks et al (2)* treats the energy backscattered from the sea surface at angles where reflection from facets oriented nearly normally to the incident radiation is negligible. The model views the surface as a composite similar to that studied by Basinov (3) and, as shown in Figure 1, it consists of large swells on which are superimposed smaller wave structure. The smaller structure, in turn, is made up of short gravity and capillary waves which when taken together, comprise the scattering surface while the effect of the underlying swell is to produce a "tilting" of the scattering surface and thus, modify the apparent direction of the incident signal. In the determination of the cross section, the presence of the swell is temporarily neglected and the energy backscattered from the small scale structure calculated. This procedure used by Wright (1) parallels the work of Rice (4), Peake (5) and Valenzuela (6) in that it utilizes the concept of scattering from a slightly rough surface, i.e., one whose height variations are small compared to the incident wavelength and whose slopes are small compared to unity. The energy scattered by the surface is determined by perturbation techniques and is found to be directly proportional to the two-dimensional energy density spectrum of the surface height variations and, also, functionally related to the wavelength, polarization and direction of the incident energy, the scattering direction and the complex dielectric constant. To model the sea return using this approach, the energy density spectrum of the height variations of a wind-generated sea is specified or measured and the energy scattered in the incident direction calculated for the desired transmission parameters and dielectric constant. The RCS is, then, determined in the conventional manner from the ratio of the scattered and incident powers.

The calculation is simplified, since, as Rice (4) has shown, if a random slightly rough surface is decomposed by Fourier techniques, only those components whose wavelength satisfies the Bragg scattering condition established by the wavelength and direction of the incident radiation are effective in the backscattering process. Wright (7) has confirmed this result in wave tank measurements of radar scatter from mechanically generated sinusoidal water waves. A further simplification results if an upper bound on the RCS is desired. Phillips (8) has defined the Equilibrium Range Spectrum as that portion of the sea spectrum in which waves have been developed to a height which is limited by gravitational constraints. If the Phillip's spectrum (there are other options) is used as a specification of surface roughness, then the existence of the maximum waveheight condition results in the determination of an upper limit in the RCS of the sea for a given set of illumination parameters.

The Sea Clutter Model which results from these considerations has three interesting characteristics which may be verified experimentally; first, it predicts naturally the variation of the normalized RCS of the sea with respect to the wavelength, polarization and direction of the incident energy; secondly, it relates the scattering cross section directly to ocean spectra whose growth and decay rates have been

*The bibliography of Reference 2 contains an excellent compilation of Russian efforts in the development of the model.

widely studied, and, thirdly, for situations in which the scattering is produced by waves located in the Equilibrium Range of the Ocean Spectrum, it predicts an upper limit or saturation condition for the backscattered energy similar to that observed by Kerr (9). Consequently, a study was undertaken to verify and expand the model by collating its predictions with cross section data obtained by the 4FR System under a wide variety of illumination parameters under various sea conditions. Before continuing with the results of the collation, a brief review of the 4FR System and its calibration procedures will be presented since its characteristics and the accuracy of the calibration impose constraints on the cross section measurements and their interpretation.

2. THE MEASUREMENT SYSTEM

The 4FR System is an airborne coherent pulsed radar which is capable of transmitting a sequence of four frequencies, X band (8910 MHz), C band (4455 MHz), L band (1228 MHz), and P band/UHF (428 MHz) alternately on horizontal and vertical polarization. The dual polarized operation is achieved by switching the transmitters between horizontally and vertically polarized feeds in the antenna system. The 4FR antenna system is composed of four antennas which are mounted in pairs back to back. One pair, the X- and C-band antennas, are circular parabolas which have a common boresite and have equal beamwidths so that they illuminate equal areas. While the other, the L- and P-band antennas, are intermixed crossed dipole arrays with a common boresite but unequal beamwidths so that the area illuminated by the P-band antenna includes that illuminated by the L-band antenna. In operation, the 4FR System alternates its transmissions in the sequence shown in Figure 2. The minimum period of the alternations, τ, is 170 microseconds while the period, T, for any frequency/polarization combination is 1.27 milliseconds.

The receiving system of the radar is designed to detect both the directly polarized return, i.e., the horizontally polarized return from a horizontally polarized transmission and the cross polarized return, i.e., the vertically polarized return from a horizontally polarized transmission. Consequently, 16 variables are present in the receiving sequence, four for each frequency. These four are the complex components of the scattering matrix and offer a nearly complete description of the reflection characteristics of targets over the parameter range to a limit determined by the resolution cell of the radar. The receiver has two channels, one for horizontally polarized returns and the other for vertically polarized. These function to preserve the amplitude and phase of each component and are time-shared by the transmitters so that calibration is simplified by eliminating relative drift errors. Amplitude and phase are measured independently in each channel. In the amplitude channel, the signal is logarithmically amplified to compress the dynamic range and then incoherently detected. In the phase channel on the other hand, the signal is amplified and hard limited at IF to eliminate amplitude fluctuations and is coherently detected using a stored reference. All signals are gated in range by an operator, digitalized and recorded in flight for later analysis. Some of the system characteristics of the 4FR System are given in TABLE I. A more detailed description of the system is given by Guinard (10).

One of the major problems involved in the measurement of cross section from an airborne platform is the calibration of the radar system. In general, two calibrations are required to eliminate system constants. The first is an internal calibration which is accomplished by measuring the receiver transfer function from the antenna output terminals to the recording medium by means of standard signal generators. The second is an external calibration to determine the constants related to the antenna gain, radome losses and radiated power. In this latter measurement, 3-inch aluminum spheres are used as reference targets of known cross section. These are dropped from the aircraft and tracked manually. Figure 3 is an example of sphere returns. The ordinate in this Figure is relative amplitude of the return in decibels and the abscissa is the logarithm of the range. The Figure is a compilation of several sphere tracks made consecutively prior to the measurement of the sea return in the North Atlantic. Each point represents the upper decile value of a thirty-two pulse sample of the return. The upper decile is used since all errors in sphere tracking tend to lower the observed values of the cross section. The theoretical value of the return shown in this Figure is computed from the radar equation using measured transmitter power, the value of the cross section, and antenna gain appropriate for the frequency transmitted and an estimate of the line and radome losses. As can be seen in the X-band horizontal summary, the theoretical value provides an excellent measure of the calibration value for the return. On the other hand, the L-band vertical summary shows a 7-decibel loss between the theoretical value and the sphere return indicating the need for a 7-decibel correction to the observed values of cross section. In order to use this calibration procedure effectively, all cross section values are referenced to the calibration sphere cross section in the following manner. The RCS of the sphere is given by

$$\sigma_S = \frac{(4\pi)^3}{G^2\lambda^2} \frac{P_{rS}}{P_{TS}} R_S^4 \qquad (1)$$

where P_{rS} and P_{TS} are the received and transmitted powers; λ is the radar wavelength; G, the maximum antenna gain and R_S is the range of the sphere. Applying the same definition to the target cross section and taking the ratio

$$\sigma_T = \frac{P_{rT}}{P_{rS}} \left(\frac{R_T}{R_S}\right)^4 \sigma_S \qquad (2)$$

where σ_T is the target cross section and P_{rT} is the corresponding radar return. In addition to these calibrations, in view of the many switching operations involved in the 4FR System, a constant fraction of the transmitted power, termed the re-entrant signal, is inserted into each of the microwave input lines and observed at the operator's console to monitor system performance. The re-entrant is periodically recorded with the data to assess system stability. As a result of these calibration procedures,

an average accuracy of ±2 dB of cross section is achieved. This value is obtained by monitoring receiver drift and noting the scatter of measurements of the reference sphere when illuminated under optimum conditions.

3. THE COMPOSITE SEA CLUTTER MODEL

In order to arrive at an expression for the RCS of the sea which embodies the characteristics previously discussed, the conditions which specify a slightly rough surface in the sense of Rice (3) must be stated. These are

$$\left|\beta\, f(x,y)\right| \ll 1 \qquad \left|\frac{\partial f(x,y)}{\partial x}\right| \ll 1 \qquad \left|\frac{\partial f(x,y)}{\partial y}\right| \ll 1 \tag{3}$$

for the heights and slopes of the surface respectively. With this specification, several authors (1,2, 4,5) have determined that the RCS of the slightly rough surface for the directly polarized returns (linear) are to first order given by

$$[\sigma_o]_{HH} = 4\pi\beta^4 \sin^4\gamma\, \alpha_{HH}\, W(K_x,K_y)$$
$$[\sigma_o]_{VV} = 4\pi\beta^4 \sin^4\gamma\, \alpha_{VV}\, W(K_x,K_y) \tag{4}$$

where β is the wavenumber of the incident radar energy; γ is the depression angle; K_x, K_y are the wavenumbers in the x and y direction on the infinite surface, and W is the two-dimensional energy density spectrum of the surface height variations. The α terms are given by

$$\alpha_{HH} = \left|\frac{(\epsilon - 1)}{(\sin\gamma + \sqrt{\epsilon - \cos^2\gamma}\,)^2}\right|^2$$
$$\alpha_{VV} = \left|\frac{(\epsilon-1)\left[\epsilon(\cos^2\gamma+1) - \cos^2\gamma\right]}{(\epsilon\sin\gamma + \sqrt{\epsilon - \cos^2\gamma}\,)^2}\right|^2 \tag{5}$$

where ϵ is the complex dielectric constant.

To proceed further, following Wright (1) the spectrum of the surface height variations must be evaluated. As previously noted, Phillips (8) had investigated the growth mechanisms of water waves and had concluded that an upper limit exists for the height of a wave of fixed length whether in the gravity or capillary range. In the former, the height is limited by the extraction of energy in surface detachment or breaking while in the latter, it is limited by the development of sharp troughs and energy loss through release of bubbles. Using a dimensional argument, Phillips found that in these portions of the Ocean Wave Spectrum, i.e., the Equilibrium Range, the spectrum is of the form,

$$W(K) = BK^{-4}$$
$$K = \sqrt{K_x^2 + K_y^2} \tag{6}$$

and the constant B is averaged over the entire surface. There is some uncertainty in the value of the constant; however, using data from several sources, Phillips has determined that the value of B should be 6×10^{-3} for the Equilibrium Range Spectrum for gravity waves and 1.5×10^{-2} for the capillary wave spectrum. Substituting equation (6) into (4) and evaluating at wavenumbers satisfying the Bragg scattering condition, i.e.,

$$K_x = 2\beta\cos\gamma$$
$$K_y = 0 \tag{7}$$

the limiting form of the cross section for angles away from the surface normal is found to be:

$$[\sigma_o]_{HH} = 1.5\pi \times 10^{-3} \alpha_{HH} \tan^4 \gamma$$

$$0 \leq \gamma < \frac{\pi}{2} \qquad (6)$$

$$[\sigma_o]_{VV} = 1.5\pi \times 10^{-3} \alpha_{VV} \tan^4 \gamma$$

where B has been set equal to 6×10^{-3} and the α's are as defined previously.

One striking feature of these relationships besides their simplicity, is that they yield the accepted shape for the RCS vs grazing angle variation (sigma-theta curve) and also the observed polarization dependence (for flat surfaces) without invoking the interference phenomenon which had previously been thought responsible for these effects. Another is the absence of an explicit dependence on the transmitted wavenumber. While this independence has been observed in vertically polarized returns, from extremely rough surfaces and from the higher sea states, the horizontally polarized returns have generally tended to follow a wavelength power law varying from $-\frac{1}{2}$ to a -1. This discrepancy probably indicates that a slight modification should be made in the form of the Equilibrium Range Spectrum.

In view of the several assumptions required to reach the limiting value of the cross section, it is valuable to compare the value predicted with that obtained from measurements. Figures 4 and 5 contain data acquired with the L- and C-band radars operating with vertical polarization on various days of the Puerto Rico experiment. These data consist of the median value of the normalized cross section for a fixed grazing angle, obtained by analysis of approximately 30 seconds of pulse samples. Pulse-width used during the mission was 0.25 microseconds for each of the frequency polarization transmissions. A detailed description of the processing may be found in the work of Daley et al (11). TABLE II lists the various wind/wave conditions under which the data were acquired together with the range of grazing angle used in each condition. The 15th and 19th of July were relatively "rough" days, while the 16th, 20th and 21st were "moderate" and the 27th and 28th, "calm." No C-band data were acquired on the 28th.

Returning to Figures 4 and 5, it can be seen that the theoretical curve not only provides a realistic quantitative upper bound for the data, but also exhibits the correct qualitative variation with grazing angle. In addition, the data acquired on the rougher days, generally, are closer to the limit which is to be expected since it indicates the greater development of the Bragg resonant wave, i.e., the wave of length,

$$L = \frac{1}{2} \lambda_r \sec \theta \qquad (9)$$

where λ_r is the radar wavelength.

The situation is more complex when the comparison is made with horizontally polarized data. In Figures 6 and 7, the comparison between the theoretical limit curve and the cross section is shown. The data were acquired with horizontally polarized signals by the L- and C-band radars. Here, the agreement is good at the larger grazing angles; however, wide discrepancies between the model and the data are apparent at the smaller grazing angles. The explanation for the lack of agreement between these data and the limiting values at the lower grazing angles may be found by considering the cross section of a composite surface. Wright (1) has noted that the effect of large swells on the smaller scattering surfaces is that of changing the direction of the normal to the mean surface with respect to the propagation vector of the incident radiation. This angular change may be resolved into two components, one a rotation in the plane of the incident radiation which will alter the apparent grazing angle and the other a rotation normal to the plane of the incident radiation. Valenzuela (12) has analyzed this latter component and concluded that significant variation in the cross section can be produced by rotations in the scattering surface comparable to those produced by swell. Such a "tilting" of the scattering surface would be expected to produce significantly larger variations in the horizontally polarized cross section in the range from 5 to 30 degrees than in the vertically polarized since the rate of change of cross section with both grazing angle and the angle normal to it is much larger for the horizontally polarized RCS than for the vertically polarized in this region. Below 5 degrees, however, the vertically polarized cross section could be enhanced because of the composite surface slope in a similar manner. Unfortunately, this latter statement is difficult to verify by further measurements both because of the power limitations of the 4FR System and also because of the effect of surface shadowing at these grazing angles. However, the composite surface model as well as cross section data indicate the iradvisability of the assumption of a vanishing cross section at shallow grazing angles for either polarization except in the case of a locally flat surface with slight roughness.

Returning, momentarily, to the question of an upper bound or worst case condition for the sea RCS, the failure of the computed limit for horizontally polarized cross section to bound the values at low angles can be rectified by utilizing the limit curve for vertically polarized signals as an upper bound for both polarizations. The validity of this use of the limit is shown in Figure 7 in which the vertical limit has been plotted with the horizontal limit and horizontally polarized values of the RCS. As inspection of the data for other frequencies will show, a similarly effective bound is achieved over the range from UHF to X band by the use of the vertical limit curve.

Although significant success has been achieved in matching the predictions of the composite surface model with the values of RCS obtained with various polarizations and frequencies in the relatively low sea states encountered in Puerto Rico, the applicability of the upper bound for RCS required further verification in higher sea states where, intuitively, the full development of the Bragg resonant waves

would be expected. Moreover, the effect of such high sea state phenomena as blowing spray, non-linear wave interactions, and severe shadowing which were neglected in the previous analysis need to be assessed to extend the utility of the model to the general case.

4. THE HIGH SEA STATE MEASUREMENT

In order to obtain radar return data from the higher sea states, a site was required which offered a high probability of occurrence and also provided a means for the acquisition of surface truth to specify the wind/wave conditions. Such a site was found in the North Atlantic Ocean. During the winter months, as shown in Figure 8, a series of storms proceed south of Labrador, Greenland and Iceland across the ocean and along the European coast. As a result of these disturbances, during January and February, there is a 30% probability of waveheights in excess of 12 feet and a 10% probability of waveheights in excess of 20 feet. In the path of these storms, as shown, are Ocean Stations India (59°N, 19°W) and Juliet (52.5°N, 20°W), on which Oceanographic Surface Vessels, staffed and maintained by the British, Dutch and French weather services, are located on a rotating schedule. These vessels not only provide oceanographic and meteorological observations to the international weather services, but also are equipped with radio beacons to serve as checkpoints for transoceanic flights and consequently, are almost ideal terminal points for the measurement program. Consequently, the 4FR System was deployed to Shannon which is 405 miles from Juliet and 510 miles from India.

As was previously mentioned, nine missions were flown to Ocean Stations India and Juliet. The choice of location was made daily depending on the sea state reported by each station and the meteorological forecast. At the Station, the OSV was used as a reference point and the flight plan shown in Figure 9 was initiated. The system transfer function and the sphere tracking runs had been carried out in transit to the measurement site. The flight plan consisted of three legs flown with respect to the wind, up, down and cross at each of three altitudes and a tenth leg flown on the return. The use of three altitudes in the flight program is necessitated by the difficulty of encompassing the entire dynamic range of sea return in the radar receiver since gain settings were held constant to maintain calibration. The altitudes shown in the Figure are nominal, the lower altitude was primarily determined by visibility of the ocean surface required by the aerial cameras and laser altimetry. On the first of the low altitude legs, data were taken with the aircraft flying upwind and the antennas oriented fore and aft. To measure the cross section at the shallow grazing angles, the antenna depression angle was set successively at 5°, 10°, 20°, and 30°. At each position, the radar return was measured for approximately 40 seconds. The antennas were then rotated 180° in azimuth and the procedure repeated. This was done to acquire data in the upwind and downwind direction on the same leg. On the downwind and crosswind legs 2 and 3, a similar procedure was followed and the same range of angles used. At the middle altitudes, the aircraft was repositioned with respect to the OSV using the aircraft direction finder in conjunction with the shipborne beacon and the crosswind leg flown. On this leg, data were collected at 20°, 30°, 45°, and 50° with the same antenna reversal procedure. Overlapping of the angles was performed to assess the effect of both altitude and time changes on the observed return. The upwind and downwind legs at this altitude were flown similarly. At the high altitude, the aircraft was again positioned and the 45°, 60°, 75° and 90° angles set and the return measured. The entire pattern required 2 to 3 hours to finish depending on wind strength and generally began at 1300Z. The tenth leg shown on the Figure was flown on the return flight at constant altitude to evaluate the variation of the cross section in a changing sea condition.

TABLE III lists on an hourly basis the wind and wave conditions which were reported at the Ocean Stations during the missions. As can be seen, on seven of the nine missions, data were acquired in sea states in which the wind exceeded 20 knots which was the largest wind condition observed in the Puerto Rico experiment. The maximum sea state condition was encountered on 11 February when a 26-foot significant waveheight was combined with a 46-48 knot wind. As will be noted at a later point, the maximum sea condition which occurred under blue skies does not provide the maximum values of the cross section. These occurred on 6 February when there were intermittent snow falls over the measurement area. Unfortunately, little data have been obtained to document the wind variation on this day beyond that listed in the Table. Verbal reports from the site, however, indicated gusty conditions which could have produced a stronger wind field later in the day. In view of the variation of the cross section with wind exhibited during the remainder of the experiment, the increase in cross section is primarily attributed to the roughening of the surface by precipitation.

Figures 10 through 17 contain the collation of the RCS determined from the composite surface theory, equation (8), and the data collected in the upwind direction by the 4FR System in the North Atlantic experiment. The values of the cross section are coded according to wind speed. The numbers in parentheses refer to the dates on which the data were collected. The RCS data were obtained by taking the median value of the calibrated signal returns acquired on the upwind portions of legs 1,2,5,6,7 and 8 at the prescribed angle; consequently, 120 seconds of data were used in the sample. The sample rate (radar PRF) was 683 and the radar pulsewidth was 0.5 microseconds. In Figure 10, the X-band vertical returns show an excellent agreement between the predictions of the composite surface model and the observed values of the RCS in the sea states characterized by winds in excess of 22 knots, in which the development of the Bragg resonant wavelengths would be expected to be nearing the maximum. In view of the ±2 dB accuracy of the data, this rather graphically indicates the existence of the saturation condition of the RCS. The lack of development of the Bragg resonant wavelength is shown in the low wind, 5 knots, encountered on 17 February. The overall rate of the cross section between these two limiting conditions is seen to be in the order of 20 decibels. Figure 11, the X-band horizontal returns, on the other hand, give further evidence of the "tilting" of the surface by the large swell structure. Again, the closeness of the measured values of the RCS obtained in the wind fields above 22 knots show the existence of saturation. Figures 12 and 13, the returns on C band for both polarizations in general confirm the predictions of the composite surface theory in a manner entirely similar to X band. The L-band returns shown in Figures 14 and 15 offer a similar agreement between the predicted and observed values. However, the variation of the RCS with wind from 5 knots to 48 knots is less for L band than for the previous two frequencies. Turning to the last two Figures in the series, Figures 16 and 17, a similar agreement in shape is found for P band but here the reduction

in the spread of the RCS with wind is far less than for any of the preceding frequencies. The overall behavior of the RCS over the 20 to 1 frequency range with wind is quite consistent with the concept of scattering from Bragg resonant water waves. As is well known, the effects of local wind are much more pronounced in the growth and decay of capillary waves than in the gravity wave portion of the spectrum. Since the Bragg resonant scatterers for X-band energy are of the order of 1 to 2 inches wavelength over the angular range, they are strongly influenced by wind. The Bragg resonant scatterers in P band, on the other hand, are in the order of 2.2 feet to 4.5 feet over the angular range and are in the small gravity wave region, and consequently, they are not greatly influenced by local wind. One last point, of the nine missions flown, only seven days are represented in the data. The 9th of February was omitted from the collation because the lack of steady surface wind made the returns highly non-stationary. Slight wind puffs created strong patchiness in the high frequency returns while the lower frequencies had returns near the noise level of the system. Data from 10 February was omitted because of the failure of the C-band transmitter on this mission. The data on the other three frequencies for this wind speed were entirely similar to that presented for the other missions in the experiment.

Additional data are available beyond that presented here which deals with the behavior of the returns in the other wind directions, cross polarized values of the cross section and wave spectral data acquired by the laser altimeter. These will be presented in later reports on the mission, hopefully, in the summer and fall of 1970. In addition, further analyses are being conducted to determine models for the statistical distributions of the returns and also to model the growth of the RCS with wind in a tractable manner. Lastly, further measurement missions are planned to determine the variation of the RCS with wind in the critical region between 5 knots and 22 knots.

5. CONCLUSIONS

The NRL 4FR System has been used to collect return data in the high sea states characteristic of the North Atlantic in the winter months. The purpose of the experiment was threefold: first, to determine a worst case condition for the RCS of the sea; second, to determine the variation of the RCS with increasing roughness; and third, to verify and expand the utility of the composite surface model in conditions which would test the linear assumptions of the theory. The overall results of the study have shown the existence of a saturation condition for the RCS with increasing roughness which had been postulated by some investigators and is one of the predictions of the composite surface theory. The value of the RCS in the saturation condition specifies a worst case clutter condition. Inspection of the data has shown that the variation of the RCS with wind is greatest for X band, least for the P-band returns. The saturation condition is closely approached by the RCS in winds exceeding 22 knots in the North Atlantic. However, as data previously obtained in moderate wind speeds in Puerto Rico have shown, the saturation condition is closely approached by wind speeds in excess of 10 knots. It is valuable to combine the data collected in both the Puerto Rico and North Atlantic experiments at a fixed angle of incidence to determine the variation of RCS with wind. This is done in Figure 18 for a 30-degree depression angle for the X-band vertically polarized returns. The data in the Figure have been fitted by power laws. In the low wind region, a cube law has been used and in the higher wind regions, a one-half power law has been fitted to the data. Defining the saturation region as that where the one-half power law applies, then the critical wind speed is shown to be in the vicinity of 10 knots.

The data obtained in the North Atlantic experiment have also been used to verify the predictions of the composite surface model in conditions where mechanisms such as spray, non-linear wave interactions, and shadowing, which were neglected in the model, would be expected to be dominant. The results of the comparison have shown a good agreement between the RCS measured in the high sea states for the vertically polarized transmissions and the predicted values over the 20 to 1 frequency range, UHF to X band. The comparison of the horizontally polarized data with the predicted values indicates again the composite nature of the scattering surface, in that, at the lower antenna depression angles, the "tilting" of the scattering surfaces by the swell results in an apparent enhancement of the cross section. This effect indicates the inadvisability of the assumption of a vanishing cross section at shallow depression angles of the radar antenna, except on a flat surface with slight roughness.

6. ACKNOWLEDGEMENTS

The authors wish to acknowledge the assistance and advice of Dr. G. Valenzuela and Dr. J. Wright in the treatment of the composite surface model, to Mr. Lionel Moskowitz for his assistance in planning the experiment and providing surface truth, to Mr. J. Daley for preparation of the data, and lastly, to the 4FR team without whom there would be no data to report.

7. REFERENCES

1. J.W. Wright, "A New Model for Sea Clutter," IEEE Trans. on Antennas and Propagation, Vol. AP-16, pp. 217-223, March 1968.
2. F.G. Bass, I.M. Fuks, A.I. Kalmykov, I.E. Ostrovsky, and A.D. Rosenburg, "Very High Frequency Radiowave Scattering by a Disturbed Sea Surface," IEEE Trans. on Antennas and Propagation, Vol. AP-16, September 1968.
3. B. Semenov, "An Approximate Calculation of Scattering on the Perturbed Sea Surface," IVUZ Radiofizika (USSR), Vol. 9, pp. 876-887, 1966.
4. S.O. Rice, "Reflection of Electromagnetic Waves from Slightly Rough Surfaces," Comm. Pure Appl. Math., Vol. 4, Nos 2/3, pp. 361-378, 1951.
5. W.H. Peake, "Theory of Radar Return from Terrain," 1959 IRE National Conv. Rec., Vol. 7, Pt. 1, pp. 27-41.
6. G.R. Valenzuela, "Depolarization of EM Waves by a Slightly Rough Surface," IEEE Trans. on Antennas and Propagation, Vol. AP-15, No. 4, pp. 552-557, July 1967.
7. J.W. Wright, "Backscattering from Capillary Waves with Application to Sea Clutter," IEEE Trans. on Antennas and Propagation, Vol. AP-14, pp. 749-754, November 1966.

8. O.M. Phillips, The Dynamics of the Upper Ocean, London: Cambridge University Press, pp. 109-119, 1966.
9. D.E. Kerr, Propagation of Short Radio Waves, New York: McGraw-Hill, pp. 512-514, 1951.
10. N.W. Guinard, "The Four-Frequency Radar System," Report of NRL Progress, pp. 1-10, May 1969.
11. J.C. Daley, J.T. Ransone, Jr., J.A. Burkett, and J.R. Duncan, "Sea-Clutter Measurements on Four Frequencies," NRL Report 6806, 29 November 1968.
12. G.R. Valenzuela, "Scattering of Electromagnetic Waves from a Tilted Slightly Rough Surface," Radio Science, Vol. 3 (new series), No. 11, pp. 1057-1066, November 1968.

TABLE I

4 Frequency Radar System Parameters

Parameter	P - BAND		L - BAND		C - BAND		X - BAND	
Polarization	Horiz	Vert	Horiz	Vert	Horiz	Vert	Horiz	Vert
AZ. Beamwidth	12.3°	12.1°	5.5°	5.5°	5°	5°	5°	4.7°
El. Beamwidth	40°	41°	13.8°	13°	5°	5°	5.3°	5.0°
AZ. Minor Lobe	14.5 db	14.5 db	13.4 db	14 db	23.2 db	23.2 db	23.6 db	23.6 db
El. Minor Lobe	30 db	26. db	16 db	14 db	24.5 db	24.5 db	23.5 db	24.2 db
Cross Pol.	25 db	28 db	25 db	25 db	>20 db	>20 db	>20 db	>20 db
Antenna Gain	17.4 db	17.4 db	25.9 db	26.2 db	31.4 db	31.4 db	31.2 db	31.2 db
Peak Power	25 Kw		25 Kw		35 Kw		25 Kw	
Ave Power	140 Watts		140 Watts		100 Watts		160 Watts	
Pulse Width	.25 - 2.0 μs		.25 - 2.0 μs		.1 - 2.0 μs		.1 - 2. 0 μs	
P R F	100 - 1463 PPS		100 - 1463 PPS		100 - 1463 PPS		100 - 1463 PPS	

TABLE II

Gross Surface Conditions, Puerto Rico 1965

DATE	LOCATION	WIND VELOCITY (knots)	WAVE HEIGHT (feet)	ANGLES SAMPLED (degrees)	QUALITATIVE OBSERVATIONS
July 15,1965	10 mi N.W.	15-20	4-5	4-20	Rain
	San Juan	10-15	3-5	30-90	Rain
July 16,1965	10 mi N.W.	10-12	2-2.5	4-15	
	San Juan	7-8	2-3	20, 30	
		8-10	3-4	45-90	
		15-18	3-4	80	Rain squall
July 19,1965	15 mi N.W.	10-12	3-4	4-45	
	San Juan	10-15	3-4	70-90	
July 20,1965	15 mi N.W.	7-9	2-3	5-45	
	San Juan	10-12	2-3	60-90	
		10-12	2-3	4 (cross-wind only)	
July 21,1965	20 mi N.W. San Juan	8-11	4-6	4 in X and C; 10 in L and P	
		8-11	3-5	10-90	Gusts to 12 knots
July 22,1965	20 mi N.W.	10-12	5-7	4-60	
	San Juan	10-12	3-4	70-90	5-8 ft swells
July 23,1965	20 mi N.W.	8-10	2-3	30-60	Sea "Choppy"
	San Juan	8-12	4-6	4-20, 70-90	white caps
July 27,1965	off	2-3	0.5-1	4-15	
	Mayaguez	5-8	1-2 & 4-5	15-66	"Confused"
		0-1	0.5-1	60-90	sea-squall
July 27,1965	off	2-3	0.5-1	4-20	Natural slicks
	Mayaguez	0-1	0,5-1	60-90	
July 28,1965	off	0-1	0.5-1	4-18,60-90	2-ft swells
	Mayaguez	3-4	0.5-1	4, 34	Rain squall

TABLE III

Gross Surface Conditions, High Sea State Program

	1200Z		1300Z		1400Z		1500Z		1600Z	
	WAVES	WIND	WAVES	WIND	WAVES	WIND	WAVES	WIND	WAVES	WIND
6 FEB. Sea 'I' Swell			14.8',9sec 23',13sec 330°	340/40						
8 FEB. Sea 'J' Swell	2',3sec 11.5',9sec 330°	160/09	2',3sec 11.5',9sec 330°	080/07	2',3sec 11.5',9sec 330°	CALM			2',3sec 14.8',9sec	110/06
10 FEB. Sea 'I' Swell			11.5',7sec 18',10sec 260°	240/31	13.1',8sec 18',10sec 240°	240/33	13.1,8sec 18',10sec 240°	240/30	13.1',8sec 18',10sec 240°	240/35
11 FEB. Sea 'I' Swell	21.3',10sec 26.2',12sec 280°	290/46	21.3',10sec 26.2',12sec 280°	290/46	21.3',10sec 26.2',12sec 280°	290/48	21.3',10sec 26.2',12sec 280°	290/48	23',10sec 29.5',12sec 280°	290/52
13 FEB. Sea 'J' Swell	23',12sec	360/39	23',12sec	360/35	23',12sec	360/36	23',12sec	350/39	23',12sec	360/37
14 FEB. Sea 'J' Swell	23',10sec	010/38	26.2',10sec	010/40	26.2',10sec	010/40	23',12sec	360/37		
17 FEB. Sea 'I' Swell	3',3sec 9.8',7sec 240°	290/08	3',3sec 11.5',7sec 240°	240/05	3',1sec 11.5',7sec 240°	260/05	3',1sec 11.5',7sec 240°	250/05	0 Sea 9.8',6sec 240°	CALM
18 FEB. Sea 'J' Swell	9.8',10sec	110/22	9.8',10sec	110/22	11.5',10sec	100/22	11.5',10sec	120/25	11.5',8sec	110/22
20 FEB. Sea 'J' Swell			16.4',10sec	060/29	16.4',10sec	060/29	16.4',10sec	050/28		

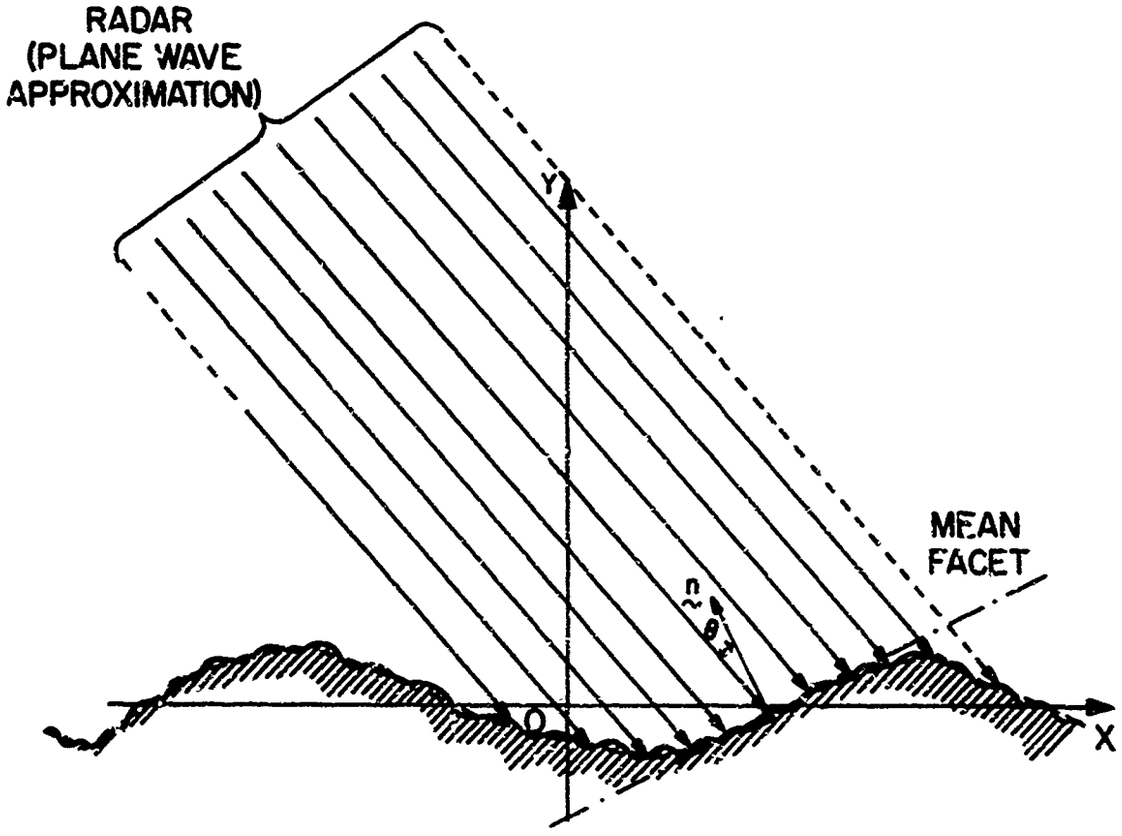
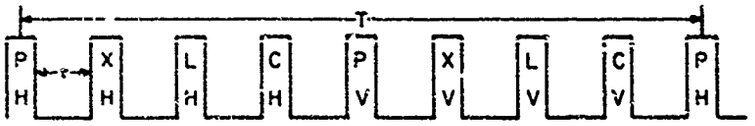
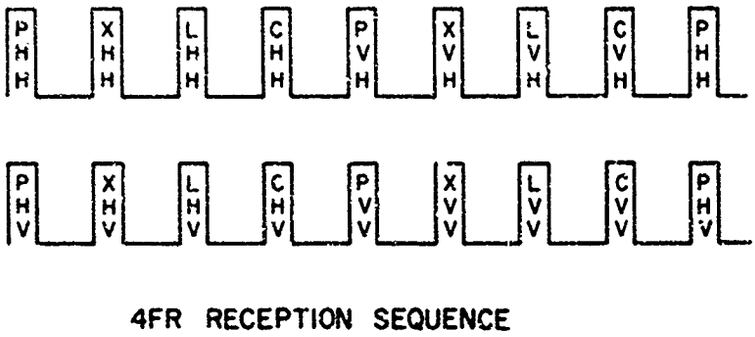


Fig.1 Composite rough surface



4FR TRANSMISSION SEQUENCE



4FR RECEPTION SEQUENCE

Fig.2 4FR transmission/reception sequences

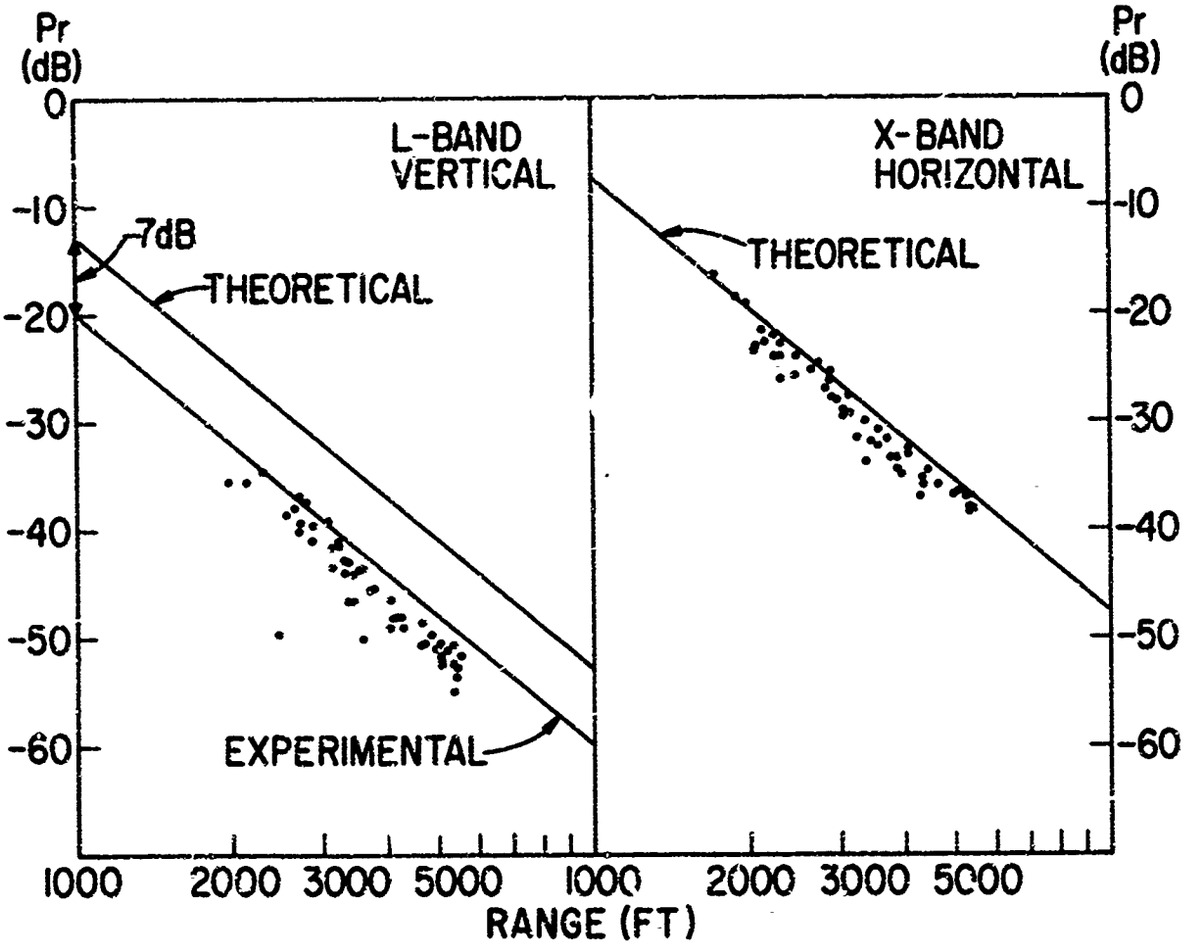


Fig.3 Typical sphere drops

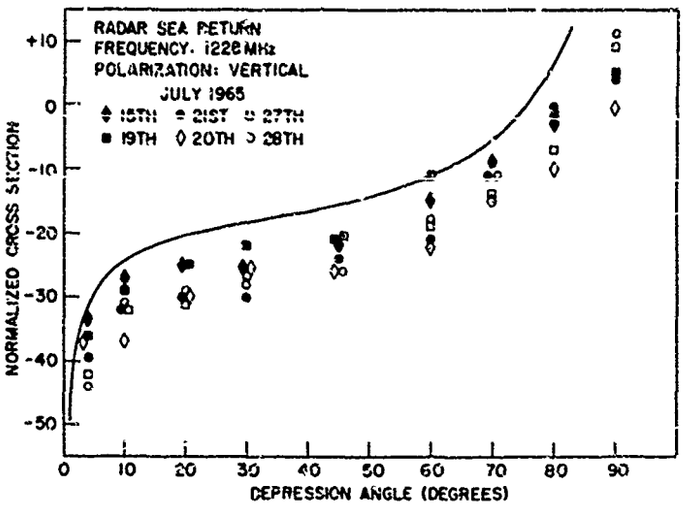


Fig.4 The variation of RCS with grazing angle $W(K) = C \times 10^{-3} K^{-4}$, $\epsilon = 73\text{-}85i$

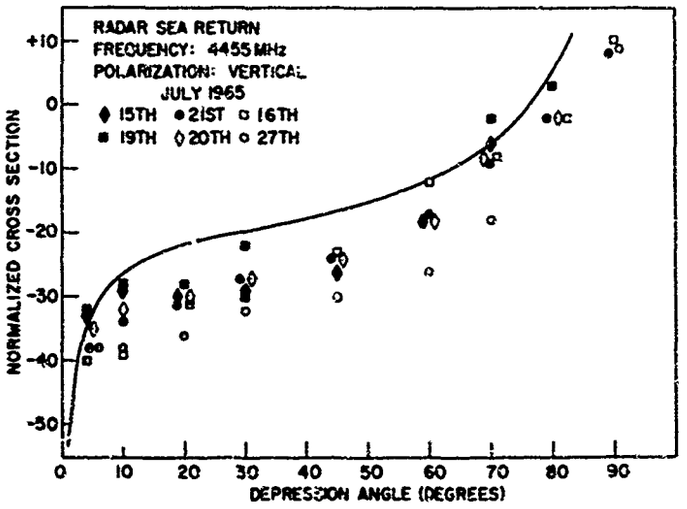


Fig.5 The variation of RCS with grazing angle $W(K) = 6 \times 10^{-3} K^{-4}$, $\epsilon = 57.1\text{-}36.3i$

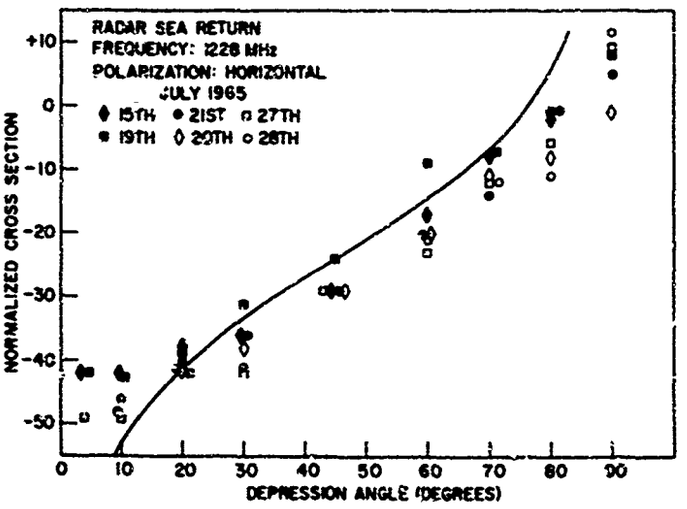


Fig.6 The variation of RCS with grazing angle $W(K) = 6 \times 10^{-3} K^{-4}$, $\epsilon = 73\text{-}85i$

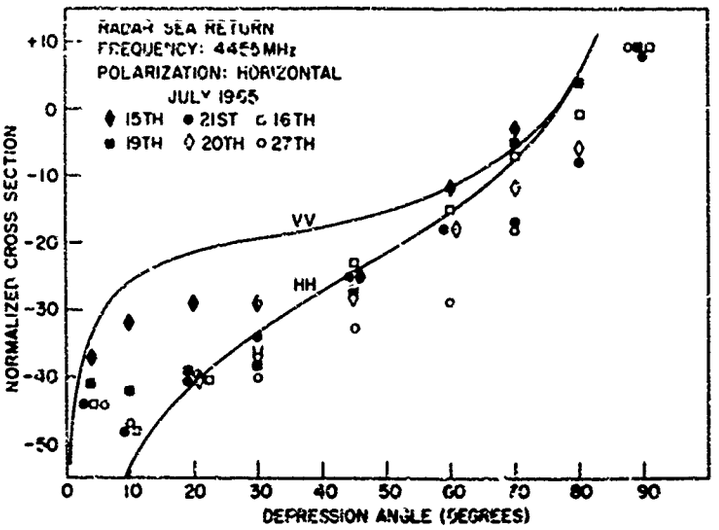


Fig.7 The variation of RCS with grazing angle $WK1 = \epsilon \times 10^{-3}K^{-4}$, $\epsilon = 57.1-36.3i$

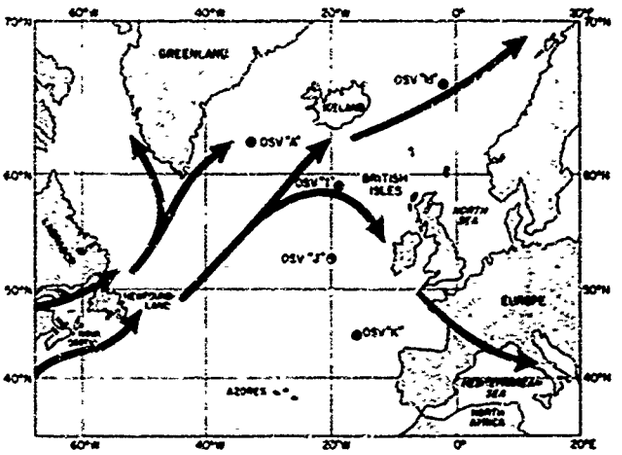


Fig.8 Prevailing storm tracks during the winter season over the North Atlantic Ocean

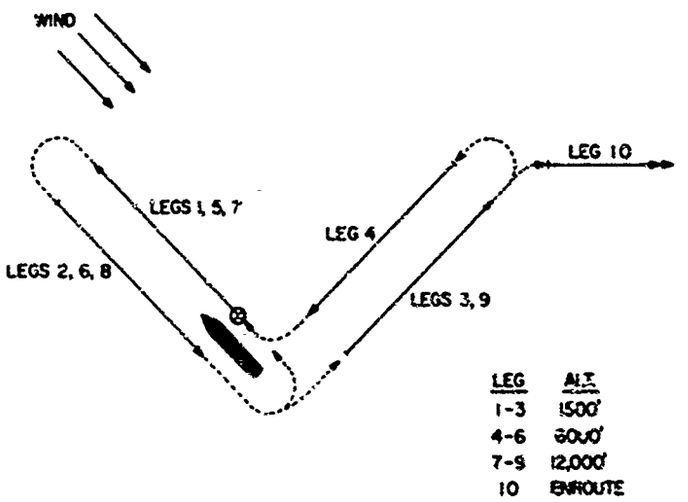


Fig.9 Flight pattern

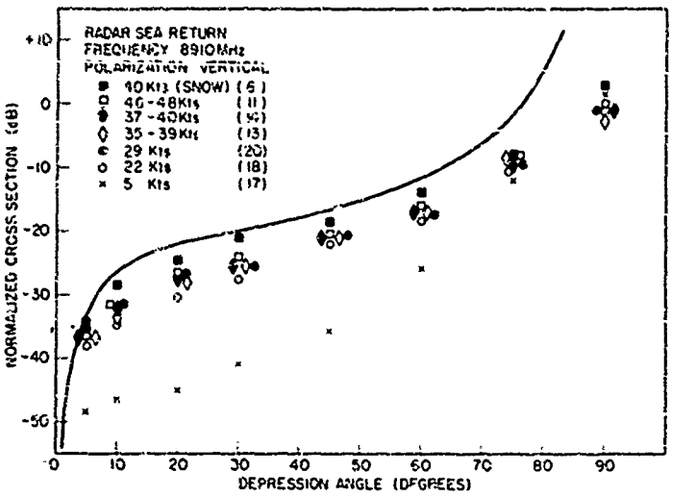


Fig.10 The variation of RCS with grazing angle and wind speed $W(K) = 6 \times 10^{-5}K^{-4}$, $\epsilon = 48.3-34.9i$

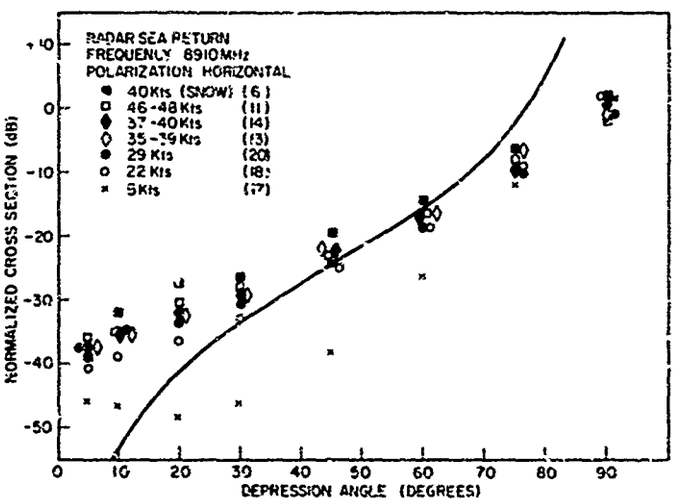


Fig.11 The variation of RCS with grazing angle and wind speed $W(K) = 6 \times 10^{-5}K^{-4}$, $\epsilon = 48.3-34.9i$

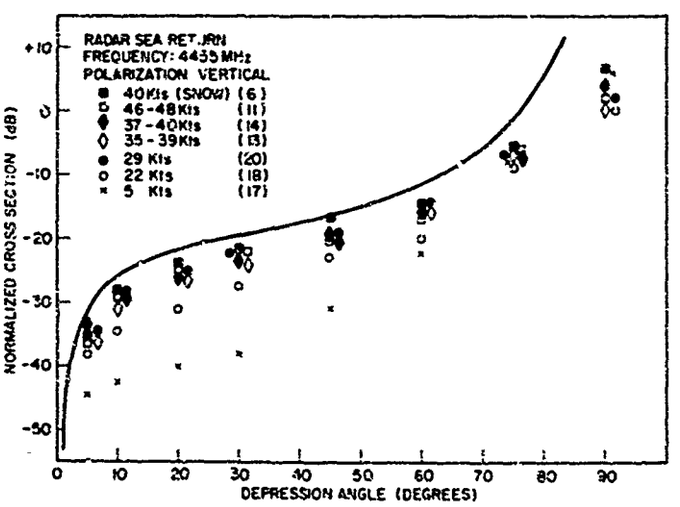


Fig.12 The variation of RCS with grazing angle

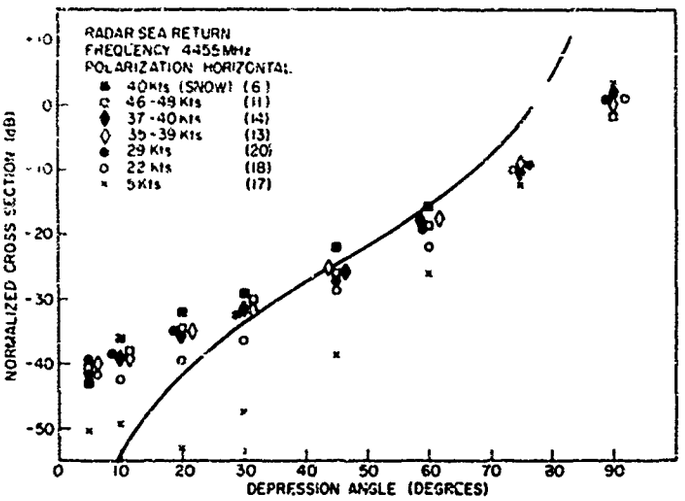


Fig. 13 The variation of RCS with grazing angle

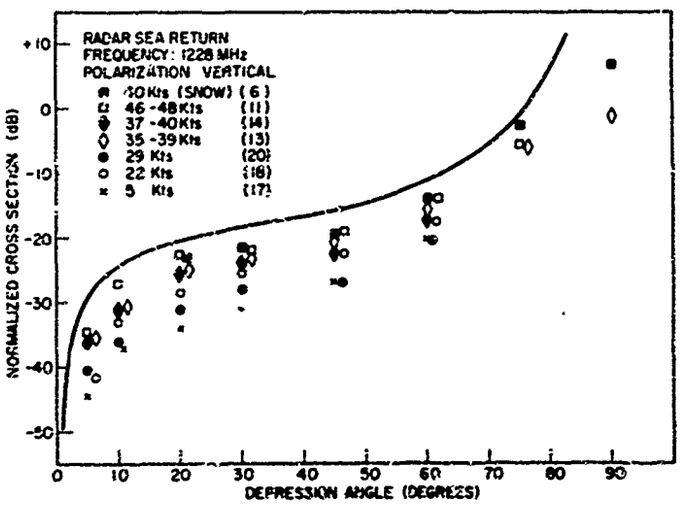


Fig. 14 The variation of RCS with grazing angle

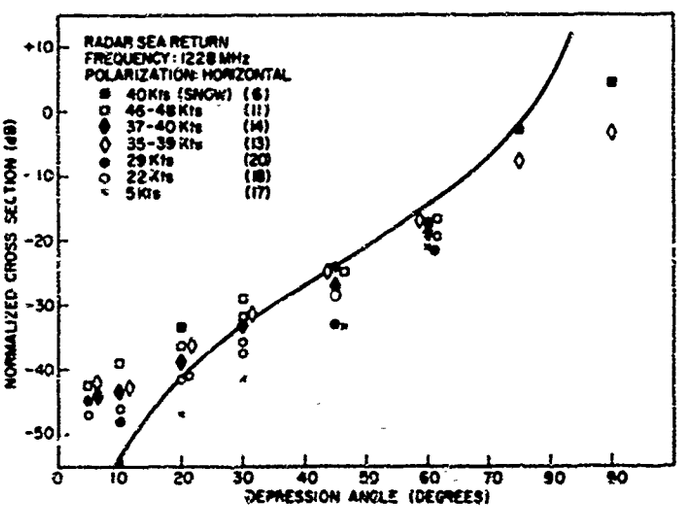


Fig. 15 The variation of RCS with grazing angle

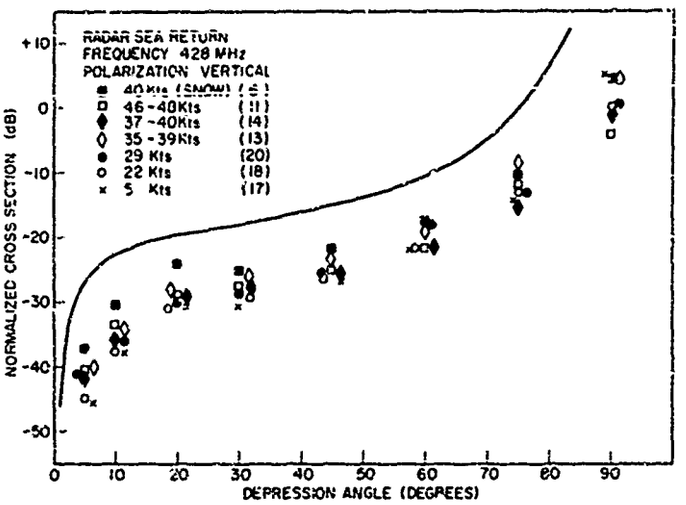


Fig. 16 The variation of RCS with grazing angle and wind speed $W(K) = 6 \times 10^{-2} K^{-4}$, ϵ = 73-165i

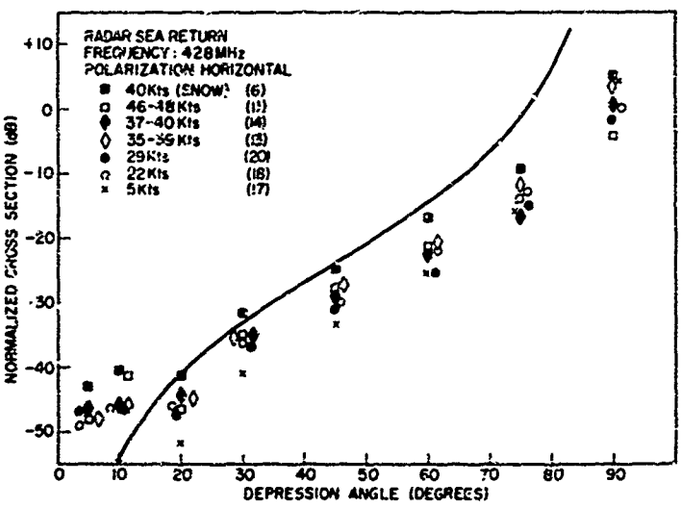


Fig. 17 The variation of RCS with grazing angle and wind speed $W(K) = 6 \times 10^{-3} K^{-4}$, ϵ = 73-165i

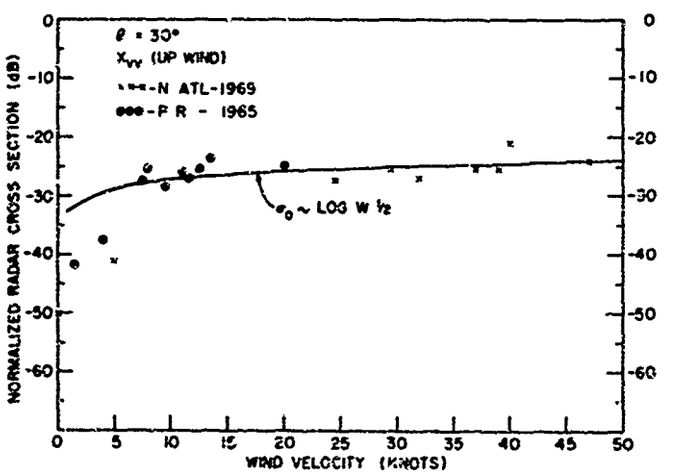


Fig. 18 The variation of RCS with wind

AN EXPERIMENTAL STUDY OF SOME CLUTTER CHARACTERISTICS

by

M.P.Warden

Royal Radar Establishment
Malvern, Worcs., U.K.

AN EXPERIMENTAL STUDY OF SOME CLUTTER CHARACTERISTICS

M. P. Warden
Royal Radar Establishment
Malvern, Worcs., U.K.

SUMMARY

The paper reports on some results based on a statistical analysis of clutter echoes using a 10 cm surveillance radar. Probability distributions of clutter radar cross-sections have been computed for ground, precipitation and bird 'angels' and changes in the distributions for different pulse lengths and polarizations and as a function of time are reported.

Ground clutter cross-section distributions of signals obtained over an arc at constant range and over a defined area show a log normal character with a standard deviation of about 25 dB. Over a period of several months the standard deviation remains unchanged while the mean value fluctuates by ± 2.5 dB. A 10:1 change in pulse length produces an 18 dB shift in the distribution.

Measurements on rain support the theory based on an assembly of a large number of independent scatterers contained in the resolution cell. Using circular polarization cancellation figures of 21 dB have been measured.

A possible ground clutter model has been suggested and its predictions give a good fit to experimental results.

1. INTRODUCTION

Clutter may be defined to be those targets that on illumination produce unwanted signals at the radar display or processing equipment. In this paper we shall concern ourselves with three types of clutter - ground, precipitation and bird 'angels', in particular with measurements of the radar cross-section when illuminated with a 10 cm. surveillance radar with different pulse lengths and polarizations.

We shall not concern ourselves with the phase behaviour of the signal.

1.1 The most casual examination of an A-scope presentation of clutter signals reveals the random nature of the clutter signal fluctuations, suggesting a statistical treatment of the signals. We will characterise the clutter by the probability distribution of its radar cross-section per resolution cell. Because of the uneven nature of the terrain we do not intend to use the derived parameters of cross-section per unit area or cross-section per sine of the angle of incidance.

We consider the two cases:

(i) Stationary aerial looking at a single resolution cell where the clutter echo fluctuates as a function of time.

(ii) Rotating aerial inducing fluctuations in the clutter signal by reason of its rotation.

1.2 The clutter signal is sampled at the output of the log receiver at the radar prf or some submultiple of it with a narrow strobe which holds the amplitude level for one recurrence period during which time the level is digitized into one of 128 levels and printed out on 7 hole paper tape. This data tape is used in conjunction with programmes written for the R.R.E. digital computer to produce the required characteristics.

The clutter cross-section is calculated from the digitized amplitude as follows. From the radar equation we can write, if we use the same radar to look at all targets,

$$\frac{\sigma_1}{\sigma_2} = \frac{P_1}{P_2}\left(\frac{R_1}{R_2}\right)^4\left(\frac{G_2}{G_1}\right)^2 \qquad \ldots\ldots\ldots\ldots (1)$$

where

σ_1 is the clutter cross-section,

σ_2 is the cross-section of some standard target such as a sphere,

P_1, P_2 are the received powers,

R_1, R_2 are the ranges of the targets,

G_2, G_1 are the aerial gains at the appropriate elevations.

We may rewrite P_1/P_2 in terms of the squares of the equivalent voltages v_1, v_2 and taking logarithms on both sides of equation 1, we have:

$$10 \log_{10} \frac{\sigma_1}{\sigma_2} = 20 (\log_{10} v_1 - \log_{10} v_2) + 40 \log_{10}\left(\frac{R_1}{R_2}\right) + 20 \log_{10}\left(\frac{G_2}{G_1}\right) \quad \ldots\ldots\ldots\ldots (2)$$

The digitized voltage level v is related to v by the log receiver characteristic

$$y = \text{const.} \log_{10} (\text{const.}\ v)$$

$$= c \log_{10} v + \text{const.} \quad \ldots\ldots\ldots\ldots (3)$$

where c is the slope of the receiver characteristic.

Substituting equation (3) into equation (2) we have:

$$10 \log_{10} \sigma_1 = \frac{20}{c} (y_1 - y_2) + 40 \log_{10}\left(\frac{R_1}{R_2}\right) + 20 \log_{10} \frac{G_2}{G_1} \quad \ldots\ldots\ldots\ldots (4)$$

or

$$\sigma_1(\text{dB}) = \frac{20}{c} (y_1 - y_2) + 40 \log_{10}\left(\frac{R_1}{R_2}\right) + 20 \log_{10}\left(\frac{G_2}{G_1}\right) + 10 \log_{10} \sigma_2 \quad \ldots\ldots\ldots\ldots (5)$$

where we have expressed σ_1, the cross-section of the clutter area of interest in terms of a known cross-section, that of a sphere, σ_2 and a number of measurable parameters. σ is then given in dB w.r.t. 1 m^2.

2. INSTRUMENTATION

2.1 Data logging equipment (Fig. 1)

Targets within defined bearings and ranges are sampled with a 0.1 μsec. strobe, which is positioned with a range gate, at the p.r.f. or some sub-multiple of it. The held output is digitized in an analogue-to-digital converter and, depending on the sampling frequency, is punched out on paper tape directly or stored in a buffer store during the collection of data and then punched out.

The data tape is headed with information, either set up on a keyboard or from certain recorded functions, relevant to the experiment - data, time, p.r.f., range, attenuation, etc.

All the circuits were designed using TTL integrated circuits and mounted on printed boards; each board is designed to perform a specific function.

2.2 Calibration

To calculate σ from equation 5 we need to know the radar cross-section of the standard target, a metal sphere, its range and digitized level y_2, the slope of the overall receiver characteristic and the aerial gain in the direction of the sphere.

A metal sphere of diameter 1 ft. is used as the reference target with a radar cross-section equal to its projected area. It is suspended 10 ft. below a hydrogen filled balloon tethered 1,000 ft. above the ground so that it lies in the centre of the beam. The site chosen, 17 n.m. from the radar, lies in the shadow of a hill to remove the possibility of ground clutter returns from the same range and also lies on the bearing of a prominent permanent echo for identification purposes. Measurement of the sphere echo using a fixed aerial gave agreement to within 1 dB of the calculated return using the radar equation and measured figures for losses in the waveguide feeds.

The aerial gain, G_2, when looking at the sphere, is the centre of the beam gain. The convention adopted when looking at ground clutter is to use the one way 3 dB beamwidth gain for G_1, and add 1.6 dB for averaging across the beam. The term $20 \log_{10} \frac{G_2}{G_1}$ in equation 5 then becomes (6 + 1.6) dB.

Before each experiment the receiver noise figure is measured and the receiver characteristic is checked for linearity, gradient and intercept using an S-band signal generator.

3. GROUND CLUTTER MEASUREMENTS

Because of the proximity of the Malvern Hills and the town of Malvern to the west of the radar site, the transmitter is only switched on during the N-E-S sector. Starting from the north the area of ground clutter returns is bounded by the Clent Hills (600 ft.), the Lickeys (660 ft.), the high ground (200 ft.) running down towards Evesham ending at Wood Norton, and then south of Evesham, Bredon Hill (800 ft.) and the Cotswolds (800 ft.), a boundary of between 15 and 20 nm from Malvern. Contained within this boundary is typical English rural country - farms with barns and silos, hedges, small woods, hills both wooded and bare, towns and the ever present pylons.

The radar, a 10 cm. surveillance set (see Appendix for parameters), is sited at the foot of the Malvern Hills; the aerial is 50 ft. above ground level at the site which is 200 ft. above sea level.

3.1 Cross-section map giving median value and interquartile range for each resolution cell

Operating the radar with a reduced p.r.f. of 63 pps and a rotational speed of 5.26 r.p.m. one pulse per beamwidth (0.5°) is transmitted over an arc of 90° at constant range. The area to be mapped is defined by 'start' and 'stop' bearing gates and by an adjustable range gate. At a selected range we record the data from 31 revolutions of the aerial, i.e. 31 amplitudes from each resolution cell in the arc. We order these values in a digital computer to obtain the median value and interquartile range for the cross-section for each resolution cell. The measurements are repeated at different ranges spaced by the pulse length.

The constancy of the aerial rotation rate was measured by counting the number of pulses received over the measurement arc for a number of revolutions and was good to one pulse interval (16 m.sec.).

Table 1 shows the median and interquartile range for cross-sections of resolution cells in an arc of 7.5° containing the sphere bearing (section 2.2). The echo from the sphere appears at 5.0 and 5.5 degrees with very little fluctuation. At this range there is very little ground clutter.

Table 1

Angle (degree)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5
Median (dB)	n	n	n	n	-45	-42	-42	-46	n	-40	-19	-14	-31	-40	-37	n
Interquartile range (dB)					9	19	5	2		4	2	1	5	4	2	

All median measurements quoted in dB with respect to 1 m^2 n indicates that the median value is below a noise threshold set so that there is a 10% probability of exceeding it. At the range of the sphere this is equivalent to -48 dB.

A part of the general map is shown in Fig. 2, where the first figure in each column is the median value of the cross-section in dB with respect to 1 m^2 and the second figure is the interquartile range in dB. This section covers a wooded hill with some buildings and a microwave link on top of the hill. The buildings are easily identified by the absence of fluctuations in the resolution cell giving a zero or very low interquartile range.

3.2 Probability distributions

Although clutter cross-section maps are of general interest in highlighting strong echoes in the area of interest and their fluctuations, it is of more interest to the radar designer to know the probability of clutter signals exceeding certain values over the area of surveillance when using the radar operationally. More specifically, he needs to know the probability distribution of the clutter cross-section in the area, the reliance he can place on this distribution under different seasonal weather and polarization changes, the variation with pulse length and is there a mathematical model that will allow him to represent clutter behaviour analytically. We have made a number of measurements on clutter echoes in an attempt to answer these questions.

We operate the radar in its normal mode of a p.r.f. of 252 pps, a 4 r.p.m. rotational rate and a pulse length of 5 μsec. Clutter signals at a constant range are recorded on paper tape over an arc of 200° defined by two bearing gates. By changing the range the procedure may be repeated until the area of interest has been surveyed. Probability distributions are computed using the R.R.E. digital computer. Any probability distribution so computed is only strictly valid for the particular radar, terrain and weather conditions existing at the time of measurement.

3.2.1 Probability distributions as a function of time

Using the above procedure we have recorded probability distributions for two different ranges over a period of several months under different seasonal and weather conditions. The distributions and dates are shown in Figs. 3, 4 and 5.

In each case the slope of the distribution remains roughly the same while the intercept on the cross-section axis varies. The maximum shift at the .25 probability point is of the order of 5 dB for both ranges with a maximum shift of 3 dB occurring on adjacent days. Measurements spaced by thirty minutes over a period of several hours have shown a maximum spread of 1.5 dB which is the resolution of the distribution. The same spread was obtained by varying the range at which the measurement is made by ± 0.3 nm.

An attempt to find a correlation between the variation in the distribution curves and meteorological conditions existing at the time of the measurement has failed. Similarly there appears to be no seasonal variation in the distribution. This series of measurements will continue for another twelve months.

3.2.2 Probability distributions as a function of pulse length

The above measurements were repeated with a pulse length of 0.5 usec. (a reduction of ten in pulse length). The resulting probability distributions for the same ranges for both pulse lengths are shown in Figs. 6 and 7. It is seen that the distributions are similar in shape and gradient but are displaced from one another at the 0.25 probability point by 22 dB at 17.4 nm and 18 dB at 15.2 nm.

In order to develop mathematical models of clutter distributions it is usual to make certain simplifying assumptions about the nature of ground clutter.

We note first of all that the distributions for ground clutter may be approximated by a lognormal distribution with a standard deviation of 25 dB. This figure is supported by ground clutter area measurements reported in 3.2.4 below. We now assume that each resolution cell contains a large number of independent scatterers uniformly distributed throughout the cell and that the resultant vector from each cell varies in such a manner that the cross-section distribution is log normal with a standard deviation of 25 dB, fig. 8 curve 1. A reduction of ten in the pulse length reduces the resolution cell area by ten and all the echoing areas are reduced by 10 dB, curve 2. We now make the assumption that the larger resolution cell contains only one scattering element and that once again the distribution is log normal. If the pulse length is now reduced by a factor of ten, the probability of a given echoing area is reduced by a factor of ten, curve 3. The difference between the curves is 40 dB at the 0.05 probability point and 25 dB at the 0.01 point.

In general ground clutter will lie between the two extreme cases with fields or large woods approximating to the first case and buildings or farm silos to the second case. The observed difference in the distributions, 20 dB, lies between the curves 2 and 3 supporting this view.

3.2.3 Probability distribution as a function of polarization

The use of circular polarization as a means of removing precipitation echoes from the display is fairly common. It is of interest to know the extent of the cancellation, if any, of ground clutter echoes. Fig. 9 shows the distributions when vertically and circularly polarized radiation are transmitted and received over an area of ground clutter extending from 8 to 17 nm over an arc of 197°. The 'circularly polarized' returns are 3 dB lower than the 'linear' returns.

Long[2] gives an expression for the radar cross-section σ_{11} when transmitting and receiving circularly polarized radiation

$$\sigma_{11} = \frac{\sigma_{VV}}{4} + \frac{\sigma_{HH}}{4} + \sigma_{VH}$$

where σ_{HH}, σ_{VV} are the cross-sections for linearly polarized radiation and σ_{VH} is the cross-polarization cross-section for the case of vegetation where the elements in the scattering matrix may be assumed to be statistically independent.

Long, quoting Cembell[3], says σ_{HH} and σ_{VV} are on average equal at X-band and the ratio σ_{VV}/σ_{VH} is between 3 and 10 dB. If we apply these results to the above expression we have $\sigma_{11} = \frac{\sigma_{VV}}{2}$ or a 3 dB difference as computed. We interpret this to mean that the same ground echoes on average will be obtained when transmitting and receiving with either linear polarization and that there is very little cross-polarization.

3.2.4 Probability distribution over the total area surveyed

Figs. 10, 11 show the probability distribution of echoing areas for the area of interest. It should be emphasised that the echo cross-sections are not normalized to unit area or corrected for elevation angle. Fig. 10 covers an area defined by 5 to 20 mile range rings and 354° - 197° bearing gates. Records were taken at 1 mile intervals with a 5 μsec. pulse on 8.8.69. The curve is linear out to the 5% probability point with a standard deviation for the distribution of 27 dB. Below the 5% point the curve falls off more rapidly until the probability of exceeding 37 dB w.r.t. 1 m^2 is 0.3%. On the 29.9.69 the measurements were repeated over a smaller sector of 45° - 135° and a range interval from 7.5 - 18 nm at 0.5 mile intervals, again with a 5 μsec. pulse. Fig. 11 shows the resulting distribution which has a standard deviation of 25 dB. The two curves are almost identical in shape.

When comparing the ground clutter from various sites we must remember that the spatial clutter distribution is unlikely to be the same. We therefore select a threshold of 0.1 m^2 that removes the weaker and less important echoes and compute the probability distributions for cross-sections greater than this threshold. The new distribution, shown in Fig. 12, has a standard deviation of 17 dB compared with 25 dB before applying a threshold.

4. RAIN CLUTTER MEASUREMENTS

4.1 Probability distributions

Fig. 13 shows a number of experimentally determined probability density distributions of rain echoes. Each distribution is computed from consecutive returns from a single resolution cell (5 μsec. long) seen with a fixed aerial and range gate. The sampling frequency is 252 pps. If we assume the cell to be filled with a large number of independent scatterers of roughly equal scattering cross-section then

it can be shown that the probability distribution is a Rayleigh distribution which transforms to a log Rayleigh distribution for our processing equipment. The log Rayleigh distribution is shown as the full curve in Fig. 13. It is seen that the experimental points are a reasonable fit to it.

Similar measurements were made with a 0.5 μsec. length pulse and a comparison made between the probability distributions for the two pulse lengths. The distributions are shown in Fig. 14. We see that compared with the predicted 10 dB for a 10:1 change in pulse length and hence resolution cell volume we have a mean value of 11.7 dB with a spread of ± 1 dB.

An attempt was made to check the measured cross-sections against calculations based on the rainfall rate. This failed, probably because the rain gauge was not recording the rainfall seen by the radar. Attempts to repeat these measurements have been hampered by the lack of rain in England this summer!

4.2 Cancellation of rain echoes with circularly polarized radiation

Measurements were made on the degree of cancellation obtained with circularly polarized radiation compared with vertically linear polarized radiation. The polarizer was of the conventional three layer wire grid type. In all the measurements, the median value of the distribution obtained was used as the average figure. Fig. 15 shows a graph of cancellation (dB) against scattering cross-section in dB w.r.t. 1 m^2 for a constant volume resolution cell. For comparison is shown the cancellation of 24 dB obtained with a metal sphere suspended in the centre of the radar beam. Examination of the rain echo cancellation gives a median value of 20.7 dB for all results with an interquartile range of 2.1 dB. Attempts to correlate the poorer cancellation figures at higher cross-section values with unusual weather features have failed.

5. CLEAR AIR TURBULENCE

Fig. 16 shows a photograph of a PPI display taken on the afternoon of Sept. 9th, 1968. The inner 20 nm (the range rings are spaced by 10 nm) is predominantly ground clutter but from twenty to thirty miles we see a number of parallel streaks spaced by some 2.5 nm and merging into noise around 33 nm which implies they are running out of the bottom of the beam at a height not less than 1,000 ft.

On this particular day ten hours of sunshine was recorded with no perceptible cloud. The wind at ground level was 15 knots from a bearing of 160°, parallel to the streaks. A slight inversion was reported between 3,000 and 4,000 ft. capping a layer of clear air convection from the ground

We suggest that the rising air trapped under the inversion layer produces a humidity gradient which results in refractive index inhomogeneities at the inversion layer. The echoes are obtained by scattering from these inhomogeneities.

6. ANGELS

A series of measurements have been made on angel activity during the height of the bird migration seasons in England. As described in 3.2 we computed probability distributions for angel cross-sections over an arc of 90° at a range chosen to be free of ground clutter. Observations made during the Spring migration in 1969 produced a log normal probability distribution with a mean at 140 cm^2 and a standard deviation of 7 dB. This may be contrasted with a mean figure of 9 cm^2 for one night the previous Autumn. Whether the difference is due to small and large birds or small birds widely spaced and small birds closely spaced in flight is not easy to infer and we do not propose to do so.

Using the technique outlined in 3.1 the angel activity was monitored over a period of several hours from 19.40 BST to 01.10 BST one night - most migrating movements take place at night. At 19.40 the median value of the distribution was 10 cm^2 with 70% of the observed resolution cells containing signals greater than noise peaks with a .01% probability of occurrence. This noise level has an equivalent cross-section of 2.5 cm^2. By 23.10 the median value had fallen to 3.5 cm^2 with 55% of the cells above noise.

7. GROUND CLUTTER MODEL

One of the difficulties in constructing a model of ground clutter is that large scale features of the terrain such as mountains, man made objects, etc., which produce the low frequency fluctuations, vary so much with different radar sites. We attempt to remove these variations from site to site by subtracting out an 'average' signal and examining the probability distribution of the deviation from the average signal, namely the higher frequency component which is of importance in M.T.I. design. We do this by taking the difference between the instantaneous signal and running mean of twenty-seven consecutive signals - equivalent to five beamwidths - using the experimental observations and computing the probability distribution of those fluctuations whenever the running mean is greater than noise. Fig. 17 shows the resulting distribution. The distribution is log normal with a standard deviation of six decibels.

We propose the following model to describe the fluctuation signal. The clutter signal is assumed to arise from a number of uniformly distributed scattering centres. The signal for each scattering centre is taken from a population which has a log normal probability distribution of power and random phase. The vector sum of n signals gives the resultant for each resolution cell. As the aerial, with a rectangular polar diagram, scans across the clutter area so scatterers are discarded and new scatterers added to the illuminated area. When a standard deviation of 6.9 dB is assigned to the scattering centre power distribution, the computed probability distribution for 1,000 signals is as shown in Fig. 18. The standard deviation is 5.2 dB and we compare it with the computed distribution based on experimental measurements made at Malvern. The line is the log normal curve with a standard deviation of 5.2 dB.

The very good agreement between the curves indicates that the model is a valid one. Fig. 19 is a graph of the standard deviation of the resultant signal distribution as a function of the standard deviation of the original distribution and the number of summations in each resolution cell.

We now apply this model to predict the probability distribution of clutter with pulse lengths differing by a factor of ten. The scatterers distribution has a standard deviation of 12 dB and there are 12 scatterers per resolution cell and four pulses per beamwidth. Using Fig. 19 the standard deviation of the resultant distribution is 6 dB. By reducing the number of scatterers to 1.2 the resultant standard deviation is increased to 9 dB. Fig. 20 shows curves computed from measurements made at Malvern with 5 μsec. and 0.5 μsec. pulse lengths. The standard deviations are the same as those obtained from the clutter model above.

The proposed model shows considerable promise in predicting the behaviour of fluctuating clutter signals. More work will have to be done in extending the model to deal with clutter signals at a number of different sites to ensure its generality.

8. CONCLUSIONS

Ground clutter echoes seen with a grazing incidence 10 cm. radar have been measured and shown to have a log normal probability distribution. The variation of the probability distribution as a function of time, pulse length and polarization has been examined with the following conclusions:-

(i) Time

The variation of the probability distribution shows itself in a constant slope for the cumulative curve, that is, a constant standard deviation, and a shifting intercept with a maximum spread of 5 dB. An attempt to find a correlation with meteorological conditions and seasonal variation has failed.

(ii) Pulse length

A change of 10:1 in pulse length produces a 20 dB reduction in the clutter cross-section at the 0.25 probability point. By moving to shorter pulse lengths a considerable reduction in clutter signal strength is achieved.

(iii) Polarization

The same echoes on average are obtained when transmitting and receiving with any plane of linear polarization. There is very little cross-polarization.

Measurements on rain have supported the theoretical predictions of a log Rayleigh distribution for the signals from a resolution cell as a function of time and a signal strength reduction in the ratio of the pulse lengths. For a 10:1 reduction in pulse length a 11.7 $\pm$ 1 dB reduction in signal strength.

An outline of a possible ground clutter model which removes the terrain variations and models the higher frequency components of the signal is given. Measurements with different pulse lengths are compared with the predictions of the model and good agreement has been obtained.

REFERENCES

1. Beckman, P. "Probability in Communication Engineering" Harcourt, Brace and World Inc, New York. 1967.

2. Long, M.W. "Backscattering for circular polarization" Electronic Letters, Sept., 1966.

3. Cambell, J.P. "Backscattering characteristics of land and sea at X-band" Transactions of the Symposium on Radar Returns, University of New Mexico, 11th/12th May, 1959.

ACKNOWLEDGMENT

Contributed by permission of the Director R.R.E.

APPENDIX

Radar	pulsed, surveillance
Operating frequency	10 cm.
Prf	252 pps
Pulse length	5 μsec.
Plane of polarization	vertical
Aerial rotation rate	4 rpm (normal)
Receiver I.F. bandwidth	200 kHz
Video receiver	logarithmic
3 dB Beamwidth, azimuthal	0.5^{o}
elevation	1^{o}
Centre of beam	$\pm 0.5^{o}$
Detector	envelope
No integration	

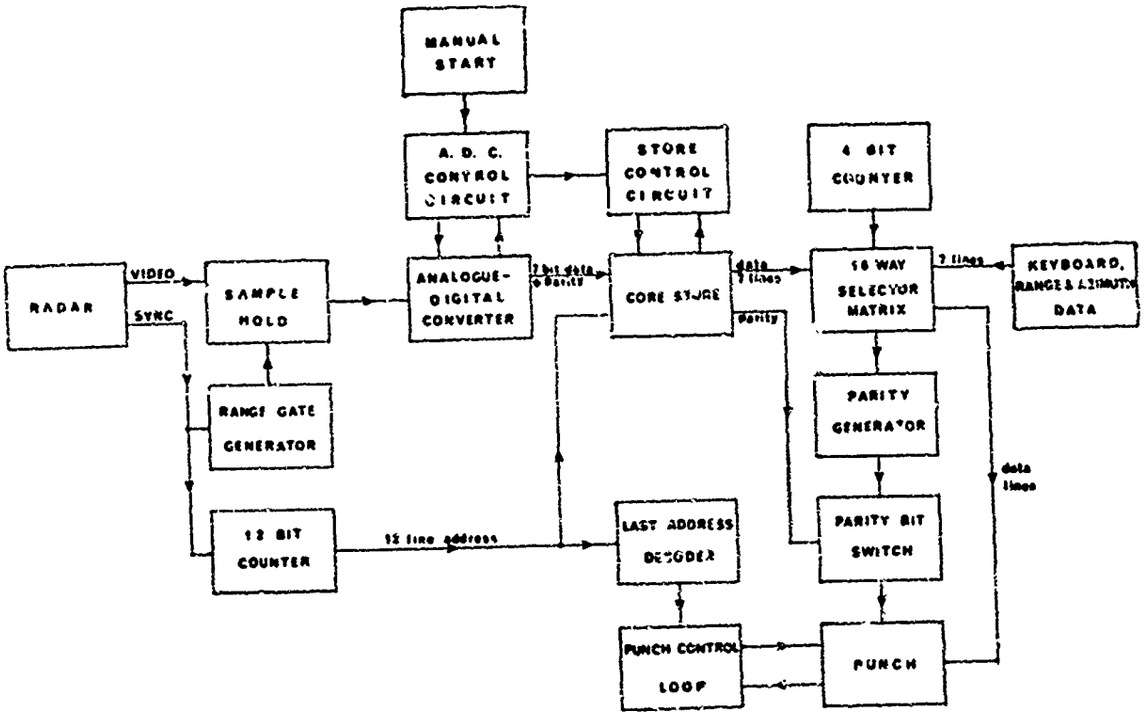


Fig. 1 Block diagram of equipment

Angle (deg)	Range (nm) 10.6	11.1	11.6	12.1	12.6	13.1	13.6
78.0	- 8.4	+ 4.10	15.6	26.9	15.7	5.6	-19.9
78.5	- 4.8	12.3	22.5	36.2	10.5	5.4	-12.8
79.0	1.8	23.10	22.6	37.2	13.6	14.10	-10.6
79.5	8.10	29.5	23.6	37.0	23.4	21.4	5.10
80.0	18.7	21.8	22.6	37.1	20.9	6.11	- 4.8
80.5	25.3	24.6	23.6	33.6	28.6	5.7	- 1.10
81.0	16.6	29.8	20.9	30.5	30.5	13.5	-12.10
81.5	8.6	24.6	30.4	31.5	30.5	17.5	-22.5
82.0	3.4	20.9	34.2	30.1	29.6	-3.13	-24.8
82.5	- 8.6	12.6	33.7	22.6	30.6	-11.6	-21.10
83.0	-16.9	3.7	19.7	14.9	34.2	-4.7	-21.8
83.5	-12.4	0.6	11.9	4.8	33.5	9.7	-25.7
84.0	-12.4	0.10	1.6	-8.6	34.5	8.7	- 4.15

Fig. 2 Radar cross-section map of Wood Norton hill area

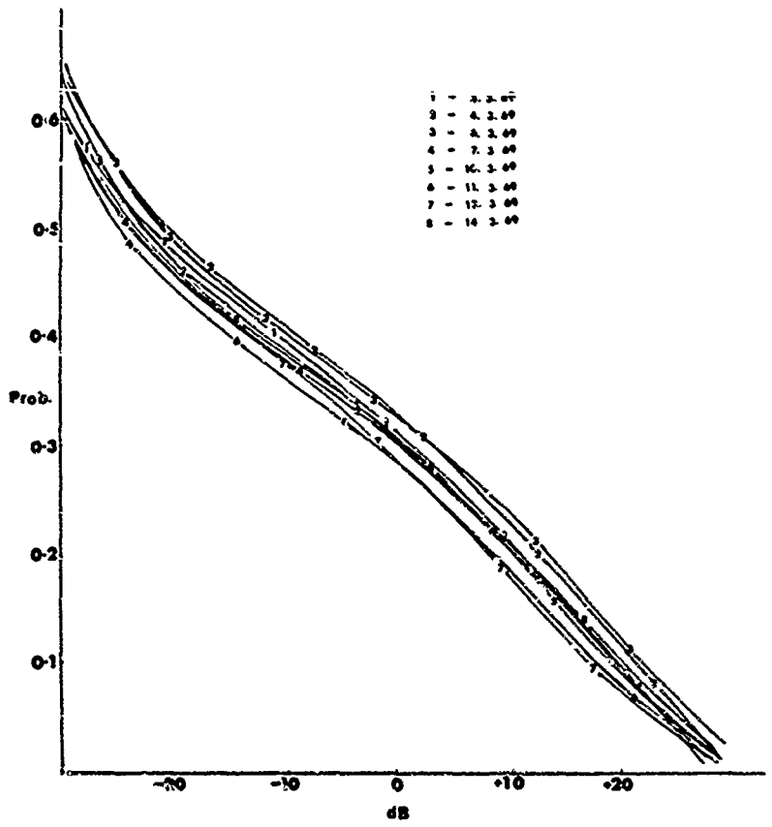


Fig. 3 Probability distribution of clutter cross-section as a function of time

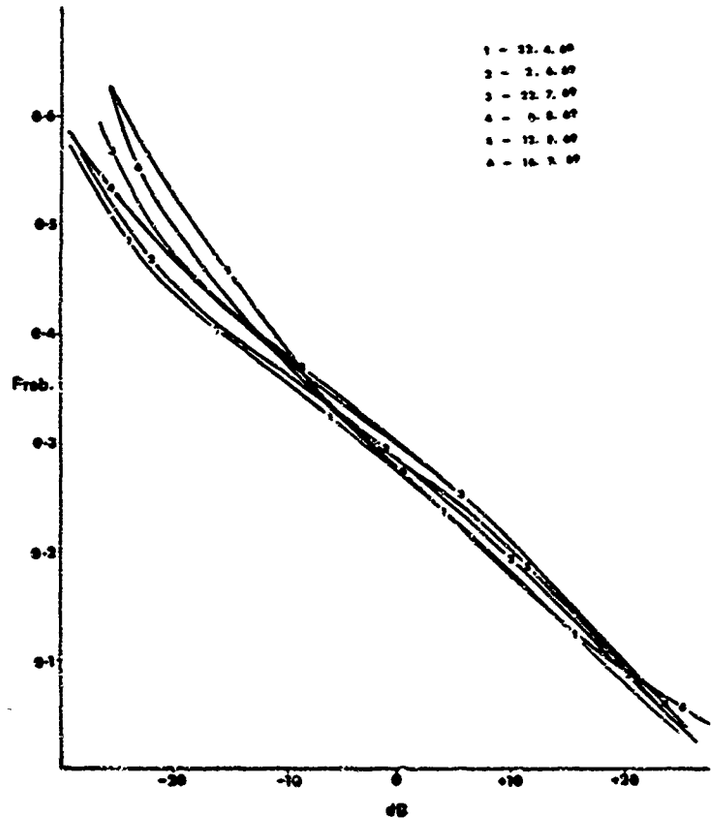


Fig. 4 Probability distribution of clutter cross-section as a function of time

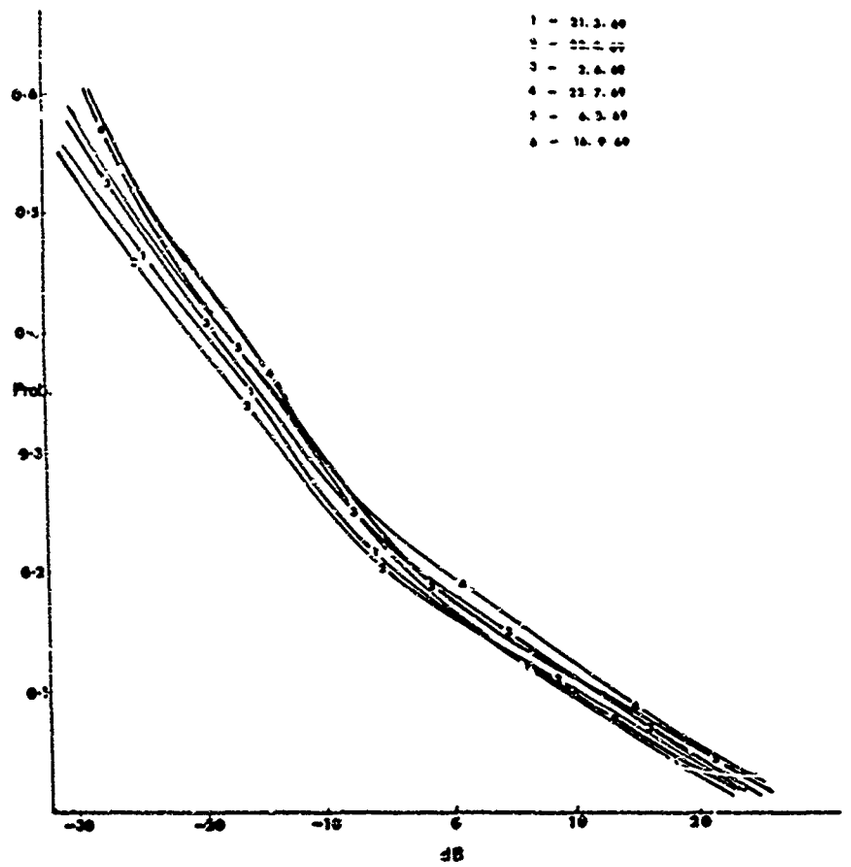


Fig. 5 Probability distribution of clutter cross-section as a function of time

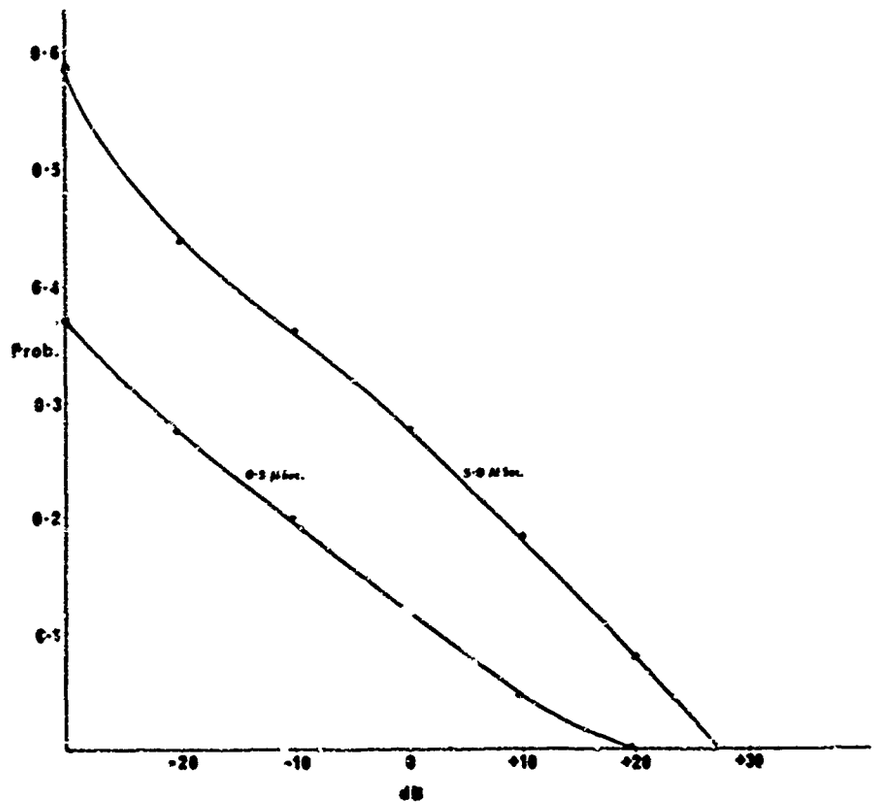


Fig. 6 Probability distribution as a function of pulse length

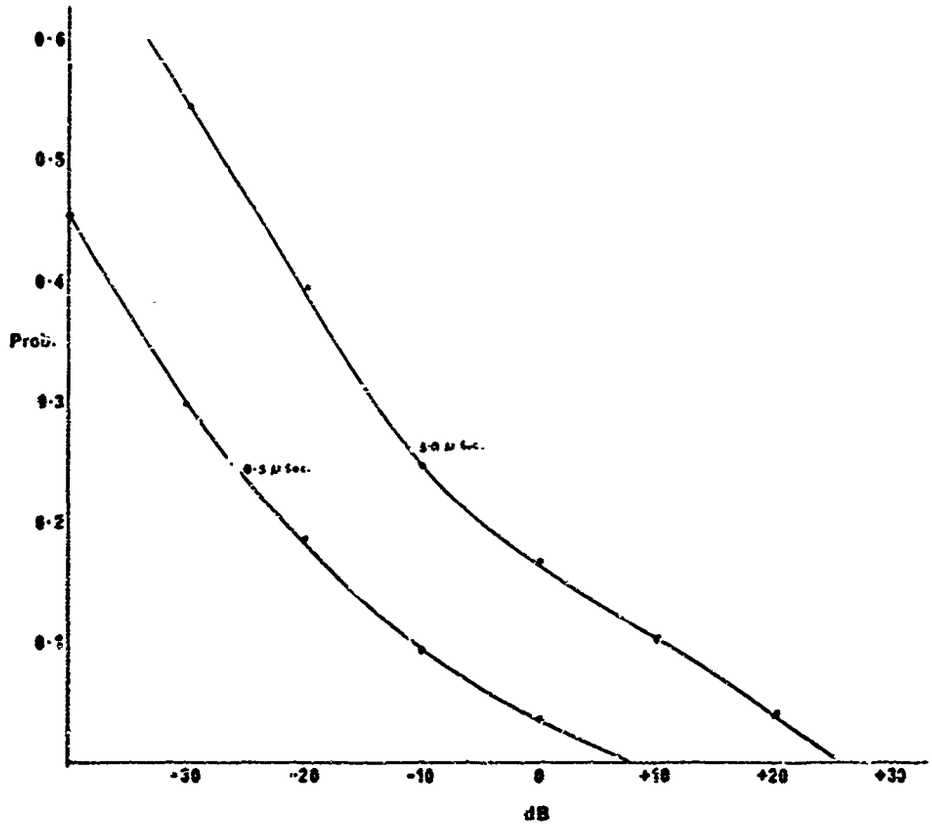


Fig. 7 Probability distribution as a function of pulse length

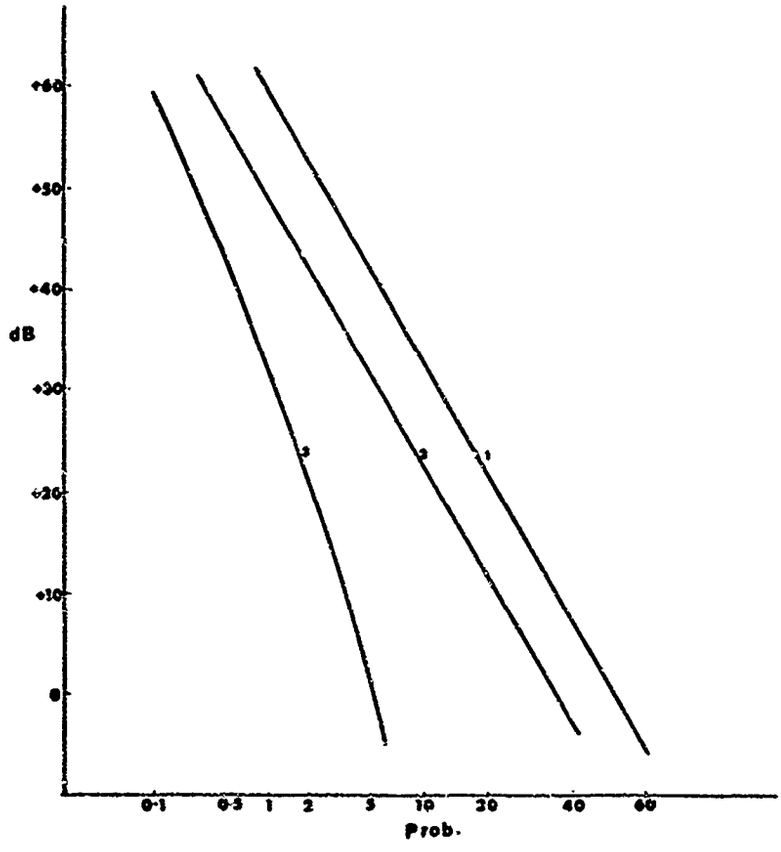


Fig. 8 Theoretical probability distributions for various scattering configurations

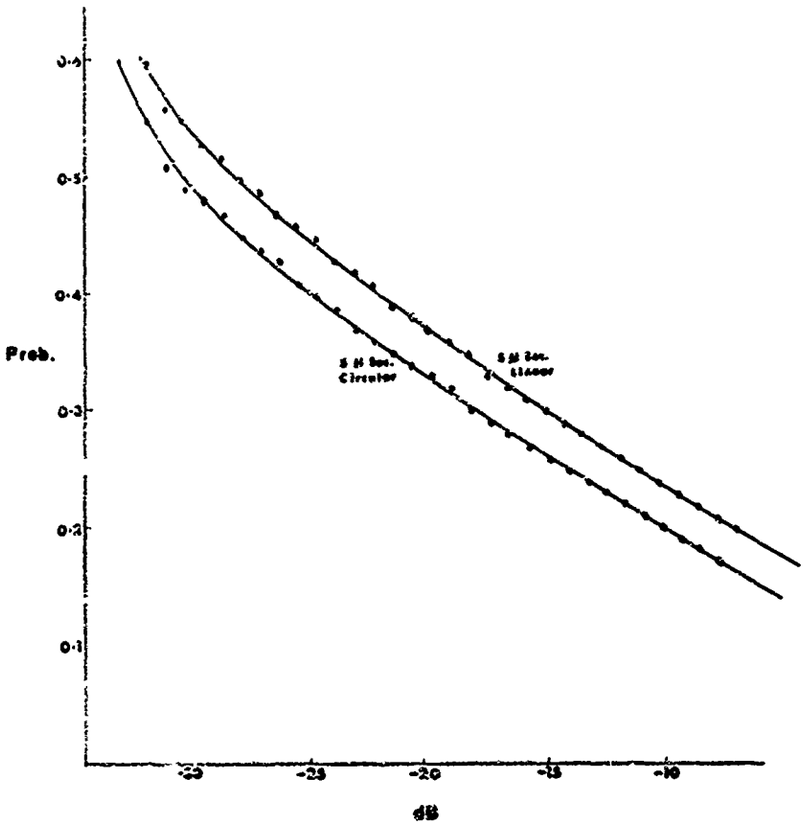


Fig. 9 Probability distributions of clutter cross-sections for linear and circular polarization

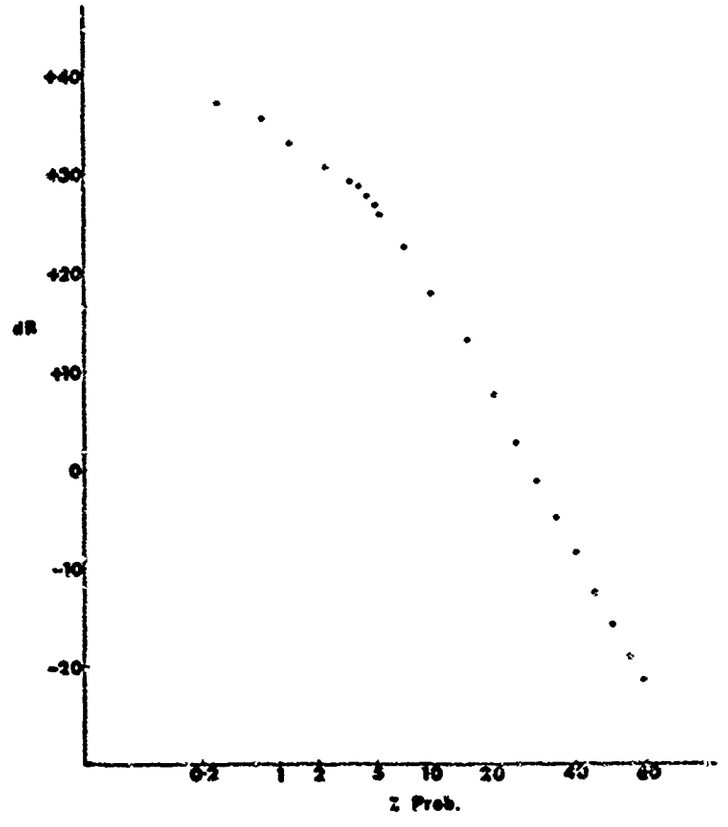


Fig. 10 Probability distribution of clutter cross-section over an area

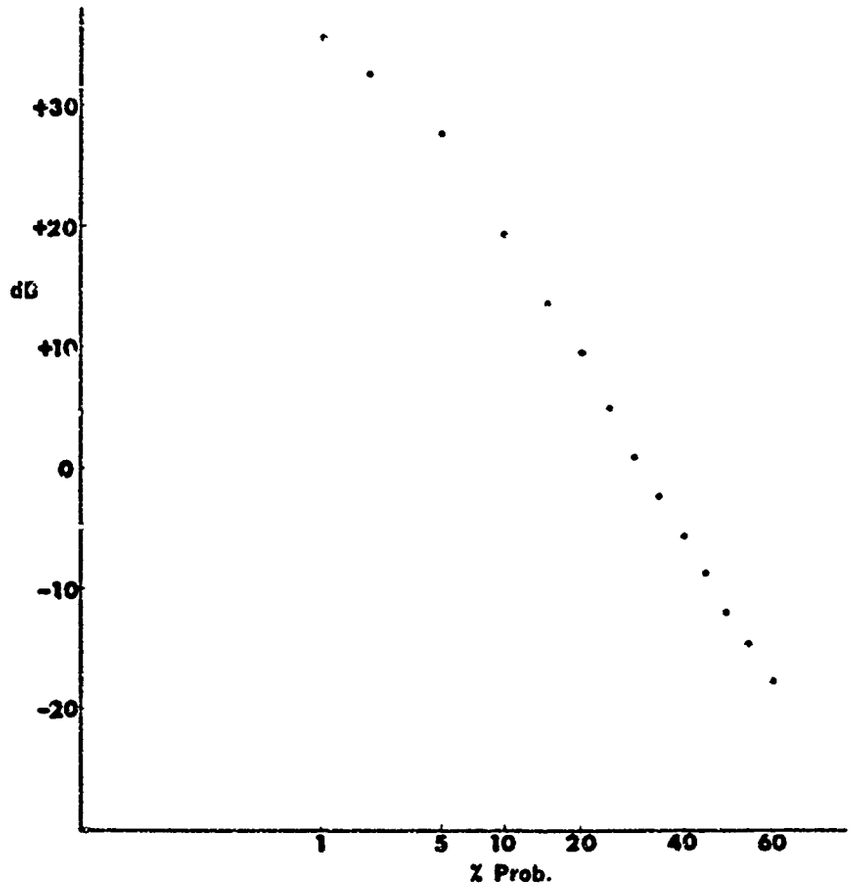


Fig. 11 Probability distribution of clutter cross-section over an area

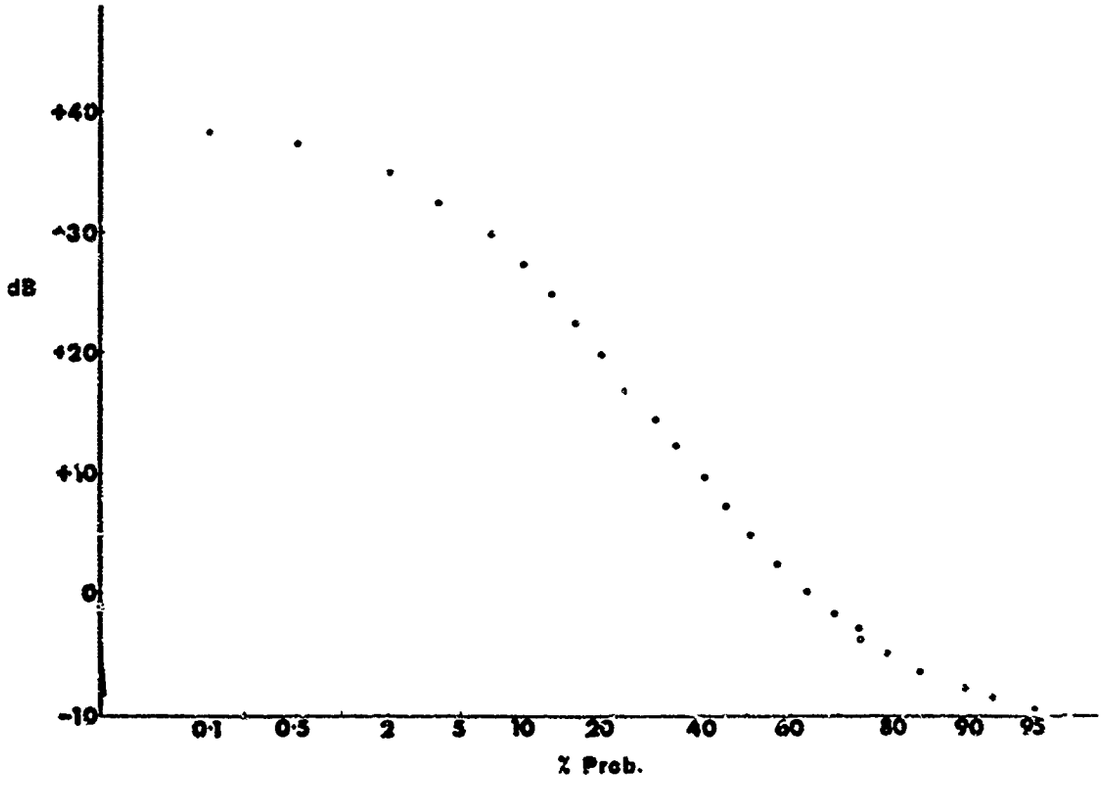


Fig. 12 Probability distribution of clutter cross-section over an area with a threshold of 0.1 m^2

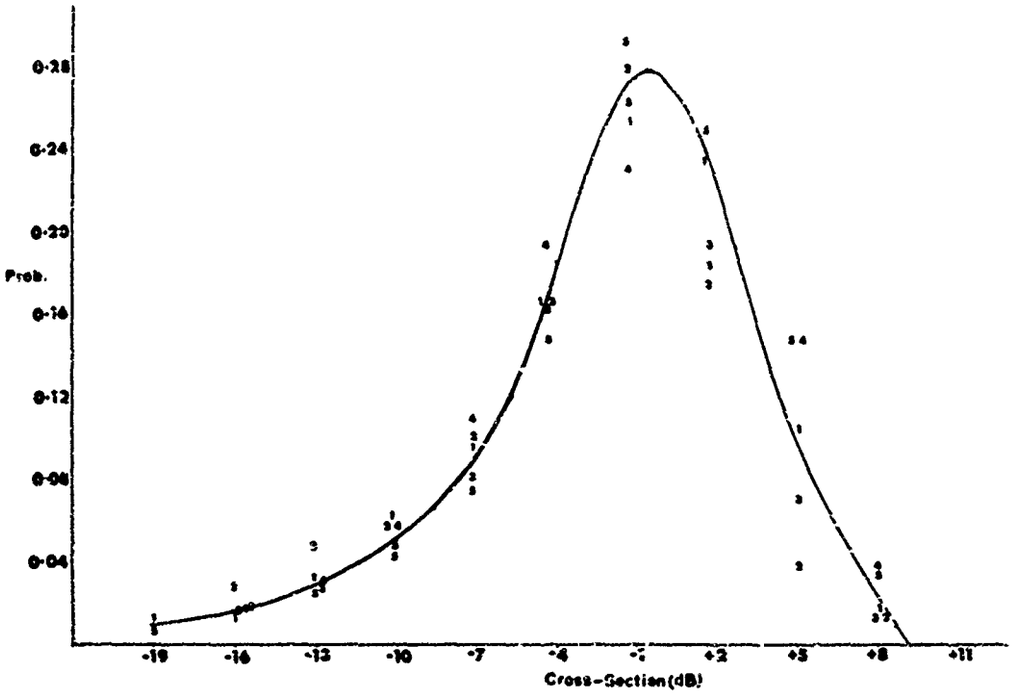


Fig. 13 Probability distribution of rain cross-section

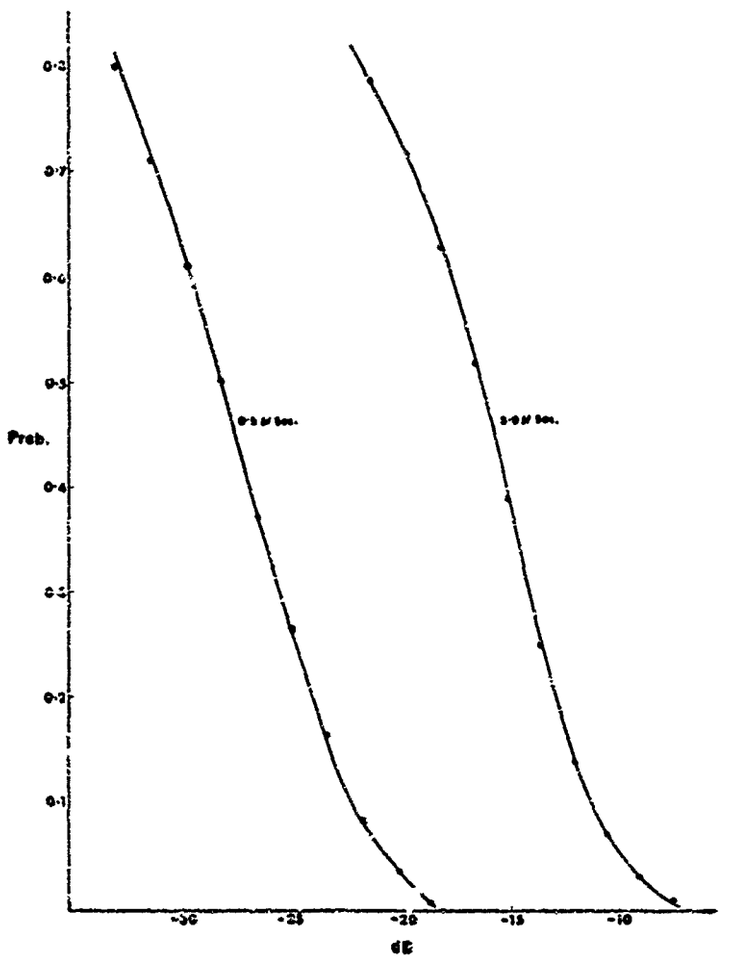


Fig. 14 Probability distribution of rain cross-section as a function of pulse length

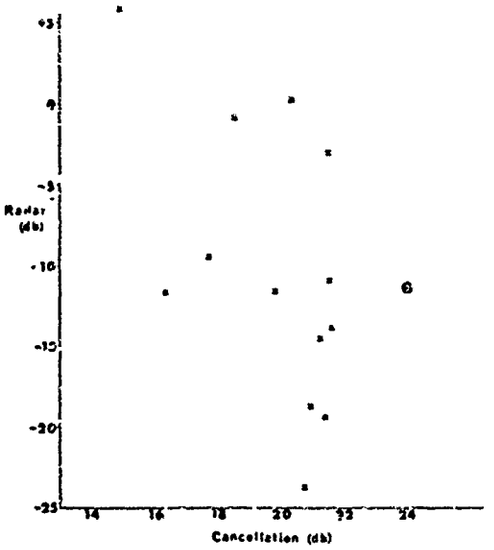


Fig. 15 Cross-section of rain versus cancellation with circular polarization

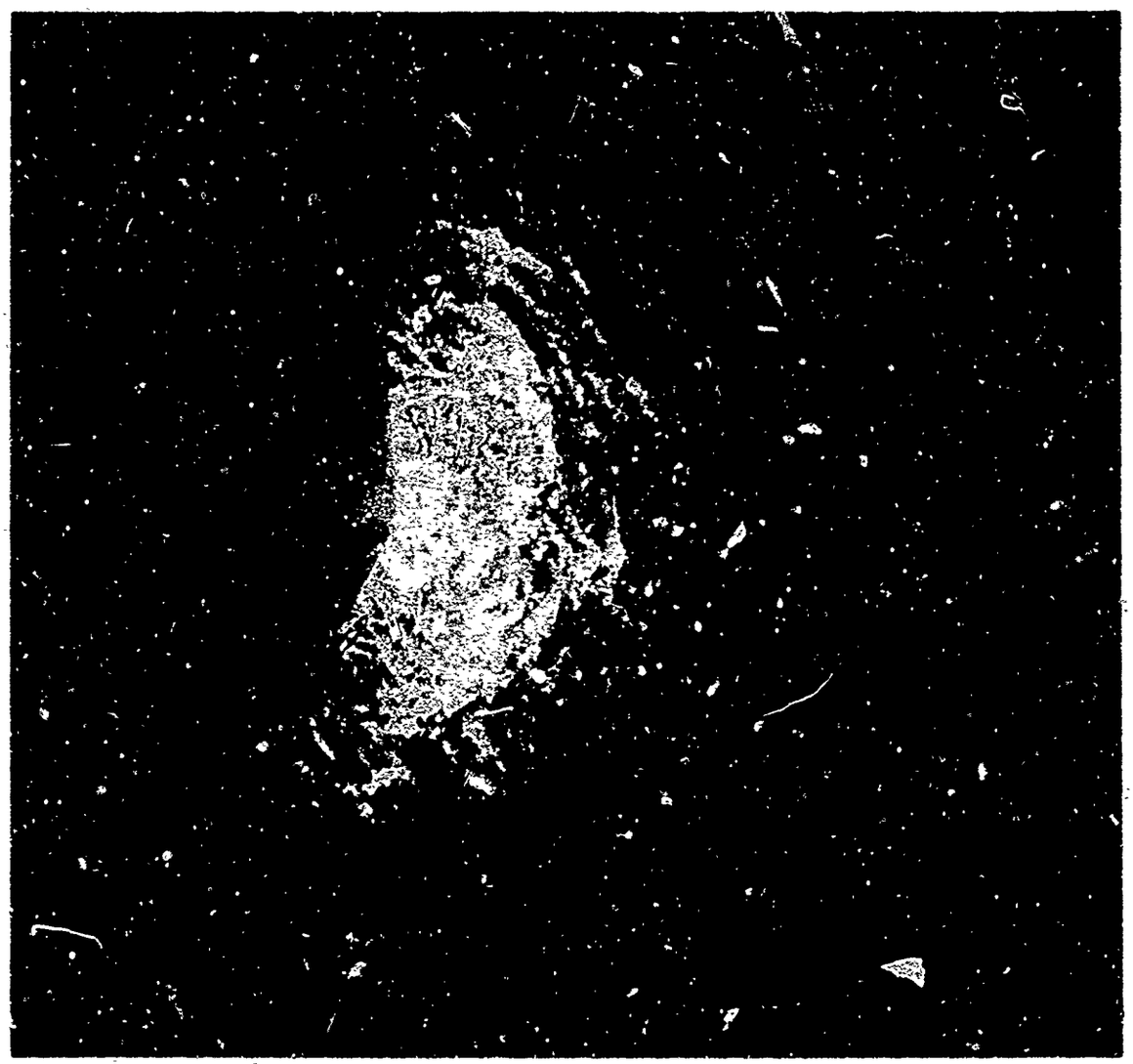


Fig. 16 P.P.I. photograph of clear air turbulence

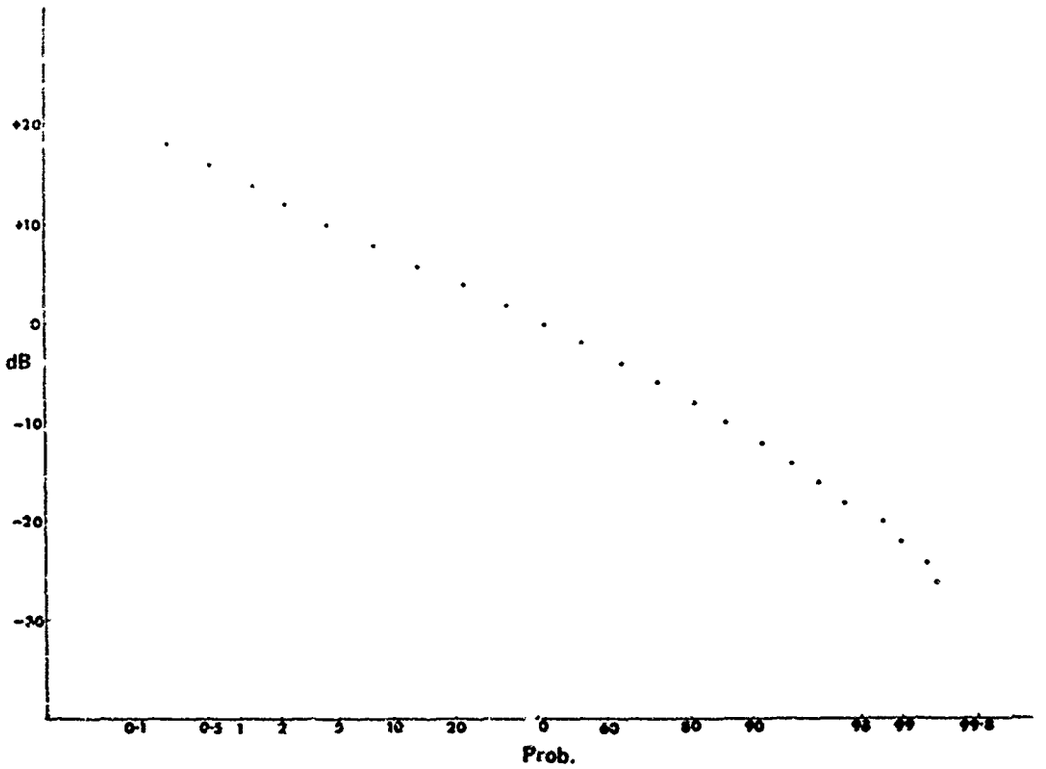


Fig. 17 Probability distribution of signal fluctuations

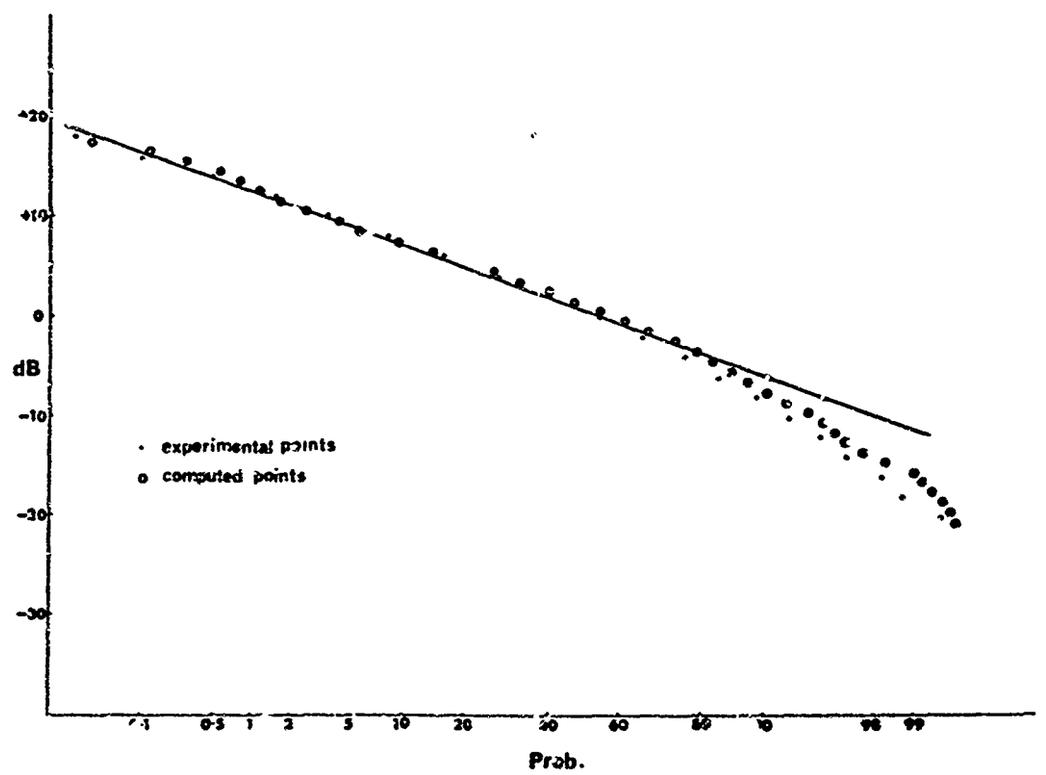


Fig. 18 Probability distributions derived from theoretical model compared with log normal distribution and experimental results

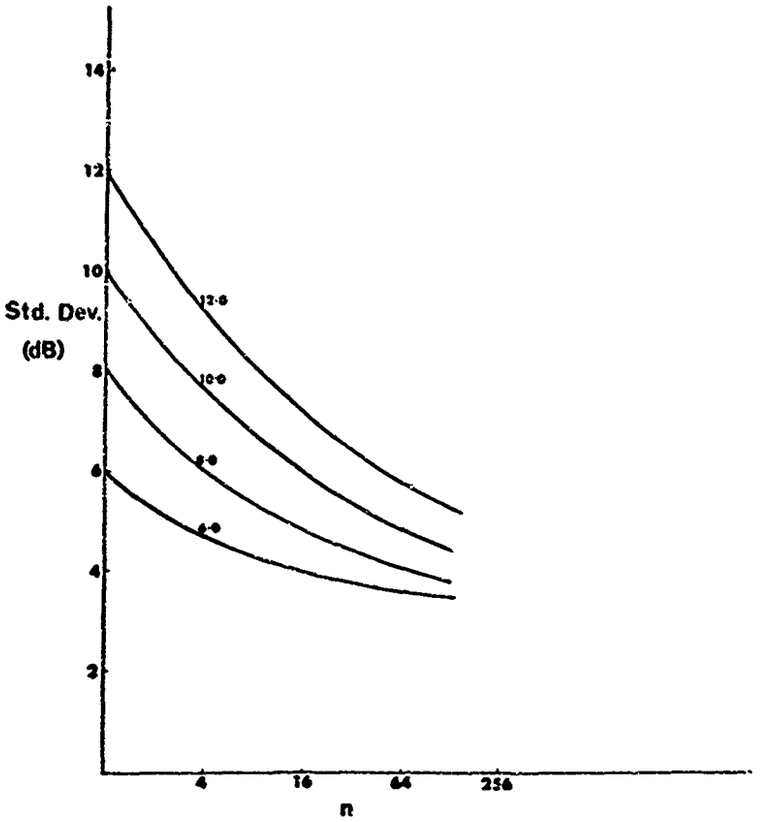


Fig. 19 Standard deviation of resultant distribution as a function of the number of summations and the initial standard deviation

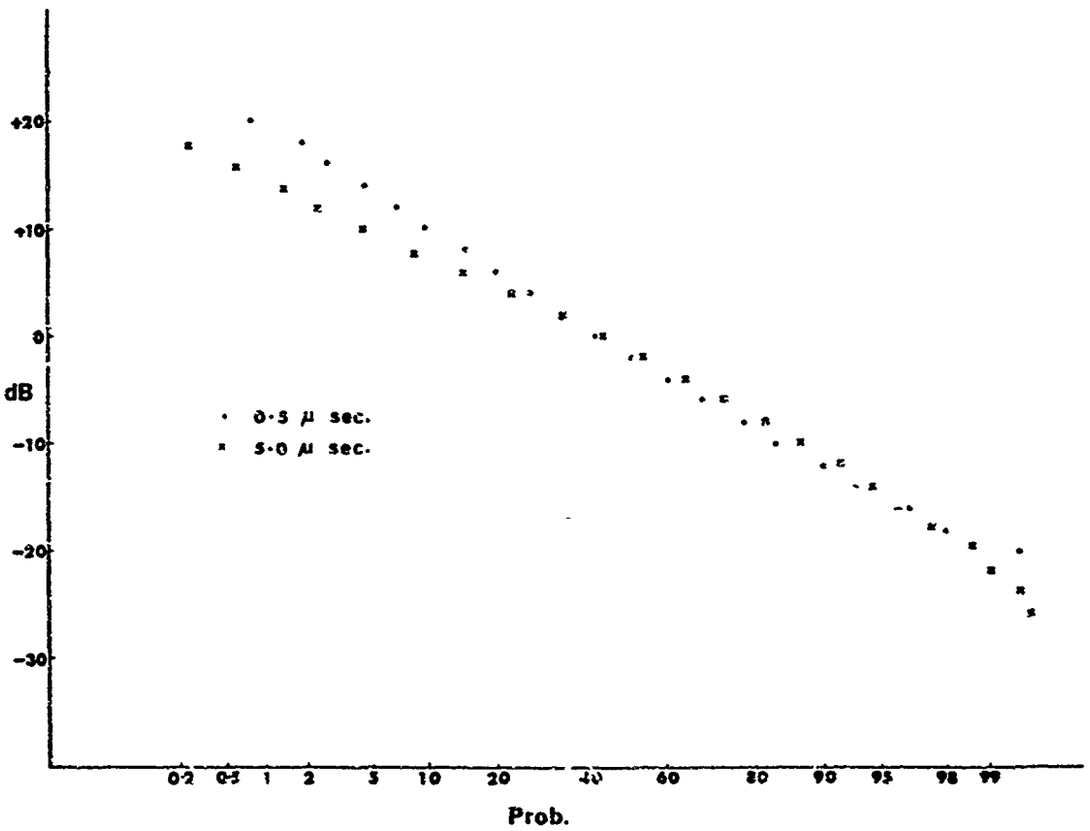


Fig. 20 Probability distributions of signal fluctuations measured at Halve. with 5 μsec. pulse lengths

METHODS AND INSTRUMENTATION FOR ASSESSMENT OF THE SPECTRAL PROPERTIES OF RADAR CLUTTER

by

T.Kester and J.Einbu

SHAPE Technical Centre (STC), The Hague, Netherlands

METHODS AND INSTRUMENTATION FOR ASSESSMENT OF THE SPECTRAL PROPERTIES OF RADAR CLUTTER

T. Kester
Systems Research Division

J. Bintz
Mathematical Consultant and Computing Group

SHAPE Technical Centre (STC), The Hague, Netherlands

SUMMARY

A knowledge of the frequency spectra of radar clutter is essential for the design of new radar systems and the assessment of the performance of existing Moving Target Indicators (MTI). This paper consists of two parts:

1. the theoretical prediction of the spectral parameters of rain clouds;
2. the instrumentation for measurement of clutter spectra.

In the first part, a new model for calculation of the mean doppler speed and the standard deviation of the velocity spectrum of rain clouds is described. These values are computed as a function of different parameters using numerous measured wind profiles. Their statistical evaluation is followed by a comparison of our model with the linear wind shear model.

In the second part, equipment for the measurement of the fine structure of radar doppler spectra is described. Some spectra, obtained with the help of a Fast Fourier Transform and sample weighting, are shown as well.

1. INTRODUCTION

If frequency discrimination techniques are used, the ability of an MTI radar system to reject unwanted clutter signals depends to a large extent on the short-term power spectrum of the clutter. For both the design of new radar systems and the evaluation of existing ones, the availability of a complete "clutter spectrum catalogue" would be highly appreciated. Already in 1949 Barlow (Ref. 1) reports Gaussian power spectra for different kinds of clutter, characterized by experimentally determined coefficients. These results, which are quoted in a current text book on radar (Skolnik, Ref. 2) seem to be used frequently in MTI design and evaluation.

Since in the meantime considerable progress has been made in signal processing techniques as well as in the design of appropriate analysis equipment, new efforts are being made to obtain short-term power spectra for radar clutter (see for instance Ref. 3 and Ref. 4).

However, the establishment of a "clutter spectrum catalogue" has of necessity to be based on numerous observations. On the one hand it is difficult to find a clutter environment or situation which could be called "typical", while on the other most of the measurements obtained will depend on the actual meteorological conditions obtaining at the time. In many cases, therefore, the statistical evaluation of individual measurements could lead to meaningful results.

Moreover, since clutter spectra probably vary according to the frequency, polarization and configuration of the radar, the analyst is faced with a large number of different parameters. Owing to these further complications, a joint effort on the part of several organizations is required, in order to obtain a universal answer to the problem.

In the first part of this paper we have described a theoretical study of the spectral properties of rain clouds, in which measured wind data are used to calculate the parameters of interest.

In the second part, we give the description of a piece of equipment which we designed, for the measurement of short-term doppler spectra in respect of clutter or useful signals, with a high frequency resolution and dynamics.

2. PART I - THE THEORETICAL PREDICTION OF THE SPECTRAL PARAMETERS OF RAIN CLOUD ECHOES

In order to calculate the spread of the velocity spectrum of rain clouds Nathanson and Reilly (Refs. 5 & 6) considered, in a recent paper, four different effects. For reasons of statistical independence, the variance of the resulting spectrum may be written:

$$\sigma^2 = \sigma^2_{shear} + \sigma^2_{beam} + \sigma^2_{turb} + \sigma^2_{fall}$$

The terms σ^2_{index} represent the contributions from wind shear, beam broadening, turbulence and fall velocity, respectively.

Nathanson and Reilly show that for ground-based surveillance radars, only wind shear and turbulence have to be considered. The term σ_{turb} is found to be independent of range and equal to approx. 1 m/sec.

In their study the standard deviation of the spectrum, due to wind shear and represented as σ_{shear}, was calculated for a Gaussian antenna beam using a linear wind shear model.

These assumptions - via relatively simple formulae - lead to σ_{shear}-values in excess of 1 m/sec for ranges from the radar from 10 kilometres up to some one hundred kilometres, depending on the wind shear gradient and the height of the upper cloud ceiling.

Although the average wind shear gradient used in these calculations was derived from balloon tracking measurements at White Sands, New Mexico, we felt that the use of individual wind profiles for calculation of the spectral parameters, followed by a statistical evaluation of the results, would give more generally reliable values and show up seasonal variations, as well as the degree of dependence on the direction of the radar beam.

Moreover since the approximation in respect of the wind profile, which was obtained by means of a linear function is only valid for the lower levels up to the height of the jet stream, clouds with upper ceilings as high as 13 kilometres (Ref. 7), could not be taken into account in the linear model.

In order to obtain representative results for European conditions, we decided to use wind profiles recorded in Vienna (Ref. 8), these being typical for Central Europe, at least as far as the higher layers which interest us most (Ref. 9) are concerned.

2.1 MODEL DESCRIPTION

The basic formula for the received power of a clutter signal from a unit volume is:

$$P = \frac{P_o'}{R_\varphi^4} f_1(\varphi)^2 \cdot f_2(\theta)^2$$

where:

- R_φ = range in beam direction, km
- φ = vertical angle from central line of beam, radians
- $f_1(\varphi)$ = vertical one-way antenna gain function
- θ = horizontal angle from central line of beam, radians
- $f_2(\theta)$ = horizontal one-way antenna gain function
- P_o' = constant, representing a set of parameters related to the antenna and to the reflecting medium.

The actual value of P_o' does not need to be determined, since it will be cancelled out in the final expression for the standard deviation.

The power received from the volume element described when the plane figure dS in Fig. 1 is moved $d\theta$ will be:

$$dP = \frac{P_o'}{R_\varphi^4} f_1(\varphi)^2 \cdot f_2(\theta)^2 \cdot dS \cdot R_\varphi \cdot d\theta$$

$$= \frac{P_o'}{R_\varphi^3} f_1(\varphi)^2 \cdot f_2(\theta)^2 \cdot dS \cdot d\theta$$

If v_θ is the wind speed in the beam direction we obtain (by multiplying dP by v_θ and v_θ^2) an infinitesimal contribution to the first and second moment, respectively, from the received power relative to the wind speed.

Let M_1 denote the first and M_2 the second moment and let P be replaced by M_o. We then have the following formula:

$$dM_i = v_\theta^i \frac{P_o'}{R_\varphi^3} f_1(\varphi)^2 \cdot f_2(\theta)^2 \cdot dS \cdot d\theta \; ; \; i = 0, 1, 2$$

For the cloud region C, the respective moments will be

$$M_i = \iiint_C v_\theta^i \frac{P_o'}{R_\varphi^3} f_1(\varphi)^2 \cdot f_2(\theta)^2 \cdot dS \cdot d\theta \; ; \; i = 0, 1, 2 \qquad (1)$$

and σ, the standard deviation from the mean relative to the wind speed, is obtained from the equation:

$$\sigma^2 = \frac{1}{M_o} \cdot M_2 - \left(\frac{1}{M_o} \cdot M_1\right)^2$$

The wind conditions at any given point are given by the absolute wind speed and the direction relative to the meridian. The model assumes no vertical wind component and no alterations in the wind vector in the horizontal plane. For a given direction of the antenna beam, v_θ will then be the scalar product of the wind vector and a unit vector parallel to the beam.

From Fig. 2 we see that the angle β varies with the range. This means that v_θ will be a function, not only of the beam direction and the height, but also of the range.

Let the beam extend from $-\theta_1$ to θ_2 in the horizontal direction, where θ_1 and θ_2 represent positive angles. This means that $f_2(\theta)$ is equal to 0, outside the interval $\langle -\theta_1, \theta_2 \rangle$.

The model assumes this interval to be so small that $v_\theta = v_o$ for $-\theta_1 < \theta < \theta_2$ where v_o is the wind speed for $\theta = 0$. Furthermore, it assumes that the cloud region C, always includes the integration path from $-\theta_1$ to θ_2. Hence (1) can be rearranged as follows:

$$M_i = \iint_S v_o^i \frac{1}{R_\varphi^3} f_1(\varphi)^2 \cdot dS \cdot P_o' \int_{-\theta_1}^{\theta_2} f_1(\theta)^2 \cdot d\theta$$

where S is the intersection between the cloud region and a vertical plane through the antenna position.

Now, since $P_o' \int_{-\theta_1}^{\theta_2} f_2(\theta)^2 \cdot d\theta$ is independent of i (as well as of height and range) it will be cancelled out in the expression for σ. If we then represent this constant by P_o and further replace $f_1(\varphi)$ by $f(\varphi)$, we have:

$$M_i = P_o \iint_S v_o^i \frac{1}{R_\varphi^3} f(\varphi)^2 \cdot dS \qquad (2)$$

where:

$$P_o = P_o' \int_{-\theta_1}^{\theta_2} f_2(\theta)^2 \cdot d\theta$$

Since the integral (2) has to be determined numerically, it was found most convenient to express R_φ, φ and dS as functions of the height H, together with the range along the earth's surface, R. This led to the following expressions:

$$R_\varphi = \sqrt{H^2 + R^2 \cdot (1 + H/A)}$$
$$\varphi = u - w - \varphi_{el}$$
$$dS = (1 + H/A) \cdot dH \cdot dR$$

where:

A = apparent earth radius ($= \frac{4}{3} \cdot 6370$ km)
$u = \cos^{-1}\{(R_c^2 + R_\varphi^2 - H^2) / (2 \cdot R_c \cdot R_\varphi)\}$
$w = R / (2 \cdot A)$
φ_{el} = angle of elevation
$R_c = R - \frac{1}{6}\{R^3 / (2 \cdot A)^2\}$

The integral (2) over the area determined by the ranges R_1 and R_2, the heights h and H_t and the tangent line through the antenna position on Fig. 3, will then be:

$$M_i = P_o \int_h^H \int_{R_1}^{R_t(H)} v_o(R, H)^i \cdot \frac{1}{R_\varphi(R, H)} f\{\varphi(R, H)\}^2 \cdot (1 + H/A) \cdot dR \cdot dH$$

where the variables R and H are added to the function names to clarify the dependency.

For $H < H_t$ (see Fig. 3)

$$R_t(H) = \sqrt{2 \cdot H \cdot A}\,(1 - \tfrac{5}{12} \cdot \tfrac{H}{A})$$

and for $H \geq H_t$

$$R_t(H) = R_2$$

If the antenna beam is of specifically gaussian shape, $f(\varphi)$ may be set equal to:

$$f(\varphi) = \exp(-\varphi^2 / 4\sigma_A^2)$$

where:

$\sigma_A = 0.297\ \varphi_A$
φ_A = vertical one-way antenna beam width (radians)

In our wind shear model we assume that the clouds under consideration are homogeneous and that no attenuation of the radar wave takes place within the area of interest to us, while in addition the thickness and height of the clouds must be defined in accordance with the cloud situation with which the investigation is linked.

A computer programme based on this model has been developed. The following data may be employed as input data for the computer programme:

a. the tabulated wind data;
b. the vertical beam width;
c. the angle of elevation of the beam;
d. the upper and lower cloud ceilings;
e. the upper and lower radar range.

The vertical antenna gain function can be represented, as an elementary function, or as a numerical table.

The mean value of the velocity spectrum, v, which is obtained as a by-product of the calculation of the standard deviation, is printed out as well.

2.2 RESULTS

In order to obtain statistical information with regard to the standard deviation and the mean value of the velocity spectrum, we calculated the corresponding σ and v values, as a function of range and beam direction, for the whole of the year 1956, using two wind profile measurements per day (Ref. 8). The contribution of turbulence to the spectra must be added, of course, to the final result of our study. Our calculations are for $\Delta R = R_2 - R_1 = 1$ km.

In the first case reported here, the cloud layer chosen was from four to eight kms in height. The calculation relates to a Gaussian vertical antenna beam (3^0 half power one-way antenna beam width, angle of elevation, 1.5^0), the resulting σ and v values being grouped into two monthly periods which were subsequently averaged. The results are shown in Figures 4, 5, 6 and 7. In Fig. 4 the standard deviation, $\bar{\sigma}$, for the velocity spectrum, is calculated on the basis of approx. 120 values each; it is given as a function of the season, for a radar range of R = 100 km. The four different lines represent four different beam directions, spaced at 45 degrees. It should be noted that the $\bar{\sigma}$-values for opposite beam directions are equivalent, and therefore not represented on the figure. Owing to the fact that the wind vector changes direction in accordance with its height, only small differences in $\bar{\sigma}$-values for different beam directions are observed. In Fig. 5 we have plotted the $\bar{\sigma}$-values for more than one radar range. In order to simplify the graph, the spread for different beam directions is indicated by the limit sign, i.e., all the $\bar{\sigma}$-values for the different beam directions are to be found within the limits indicated. Depending on the indication obtained from the linear wind shear model, $\bar{\sigma}$ increases as a function of R and reaches a saturation point which in its turn is dependent on the level of the upper cloud ceiling. As a result of truncation of the antenna beam, the $\bar{\sigma}$-values approach zero in the case of extreme distances from the radar. The absolute values obtained will later be compared with those derived from the linear model. Fig. 5 shows a considerable change in $\bar{\sigma}$-values from Jan.-Febr. to May-June. This is probably due to the fact that, in Central Europe, the westerly winds are generally less strong in springtime.

In Figures 6 and 7, for Jan.-Febr. and May-June respectively, the mean wind speed in the beam direction is given as a function of azimuth angles for three different radar ranges. The graphs show that the mean velocity is higher in winter and that the mean wind direction seen by the radar, changes from WNW in winter to W in the spring.

To obtain Fig. 8, we used a cosecant antenna beam (approx. 5 degrees one-way half power beam width, 2 degrees angle of elevation) and left the remaining parameters unchanged. Whereas the general behaviour of the $\bar{\sigma}$-values vis-à-vis the seasonal ones remains the same, an increase in the absolute value and a decrease of the range dependence factor are noted.

Our statistics for the σ-values show that they represent a confidence level of approx. 60%.

As far as the linear shape of the individual wind shear velocity spectrum is concerned, we found that it is not Gaussian. At a beam width and elevation of 3 degrees however we found that more than 70% of the total power was concentrated at the mean velocity v within the interval $\langle v - \sigma, v + \sigma \rangle$. Two of the line shapes we calculated are shown on Fig. 9. It should be noted that in reality these lines will be modulated by the turbulence effect.

2.3 COMPARISON WITH THE LINEAR MODEL

If we use (in our model) a linear wind profile instead of a measured profile we obtain the same σ-values as those obtained with the help of the Nathanson and Reilly model. The use of measured wind profiles, however, leads to appreciably different results. These differences are shown in Fig. 10, where σ is plotted as a range function for the linear model [wind shear gradient K = 5.7 m/sec-km (see Ref. 5)] and the $\bar{\sigma}$-values for our model. Except in the case of the wind profiles, no other parameters have been changed.

2.4 CONCLUSION

In this part of the paper we have shown how our wind shear model is used to obtain statistical information concerning the parameters of rain cloud spectra. These parameters depend largely on the radar antenna configuration. Whereas the mean velocity $\bar{v}$ seen by the radar varies as a function of azimuth direction, the standard deviation $\bar{\sigma}$ of the velocity spectrum can be considered as being independent of it. For a particular season and a given radar range the study shows that in order to achieve reasonable cancellation factors for rain clouds, an azimuth dependent velocity compensation is necessary, while a variation of the form of the pass band of the MTI filter is not needed.

3. PART II - INSTRUMENTATION FOR THE MEASUREMENT OF CLUTTER SPECTRA

3.1 DESCRIPTION

Since the entire clutter spectrum seen by a coherent impulse radar is the result of convolution of an individual clutter spectrum with the discrete spectrum of the transmitted pulse train, the clutter spectrum is accurately reproduced at every pulse repetition frequency line (prf line), always provided that the pulse repetition frequency is high enough to prevent overlapping of the spectra. We therefore decided to filter out a small frequency band of the spectrum, covering at least two pulse repetition frequency lines. In practice, a band of 3 kHz was chosen, to allow for pulse repetition frequencies of up to almost 3 kHz. In order to prevent duplication of the Doppler lines on either side of the prf lines, a single side band filter is used in an intermediate frequency stage.

Figure 11 shows the block diagram of the system we designed for assessing the fine structure of the short-time power spectra of pulse radar signals.

The intermediate frequency signal available at the output of the radar receiver is gated to select the sector and range ring of interest. The gates are digitally generated, using the north marker, the azimuth change pulses (ACP) and the radar trigger. The window is superimposed on the radar video signal and thus appears on the PPI. The gated IF signal enters the down converter together with the COHO signal and is coherently mixed down to zero frequency in several steps. The single side band filter, which has a band width of 3 kHz, is used in a 100 kHz stage. The band-limited low frequency signal at the output of the down converter can be monitored, recorded or analyzed in real time. In cases where no computing facility is available near the radar station, the signal is recorded on analogue tape. In the play back mode the analogue signal is sampled at a rate of 6 kHz and converted from analogue to digital format. The digital word has a length of 10 bits, in order to guarantee dynamics of at least 60 dB.

The digital time domain signal is sent to a digital computer, where the calculation of its power spectrum is performed, using a Fast Fourier Transform (FFT) method (Ref. 10). Depending on whether the signal being studied is continuous or not, it is weighted prior to computation of the spectrum. In cases where a "Hanning" function is used, the first side lobe of the spectral window is decreased to about 2% of the amplitude of the main lobe (Ref. 11).

The calculated power spectrum is automatically plotted by the computer in accordance with the instructions concerning the range of power and frequency axis.

3.2 PERFORMANCE AND LIMITATIONS

Figure 12 shows the spectrum of a pulsed 60 MHz signal, as recorded and analyzed with the help of our system. The pulse length was 10 μsec and the pulse repetition frequency 400 Hz. Since 4,096 samples were used for the computation of the power spectrum, the effective record length of the analogue input signal was approx. 0.7 sec. The coefficients calculated are spaced at 1.4 Hz. The advantage of side lobe suppression, achieved by the use of the weighting function, is paid for in the shape of a slight degradation in frequency resolution. In this example the resolution is in the order of 2 Hz, while the spectral noise density is about 60 dB below the peaks of the prf lines. This relatively high dynamics factor is due to the performance of the down converter, the input-output characteristics of which are shown in Fig. 13.

Although the frequency resolution capability of the Fourier transform technique is unlimited, there is no requirement for a higher resolution than that corresponding to approximately the inverse of the signal length. The useful range on the power density axis, however, should be as high as possible in order to reveal the spectral content of the signal. To achieve this, the word length of the digital samples has to be high enough to reduce digital noise, while the dynamic range of the down converter must likewise be adequate. Using a 10 bit A/D converter, we obtain a range in spectral power density of about 50 dB for a pulsed signal, with a duty cycle greater than one thousand.

3.3 APPLICATION

The system described above has been designed to measure the short-term power spectra of clutter, but it can equally well be used to perform measurements of target and transmitter spectra. In the latter case, a sample of the transmitter pulse has to be delayed by means of a delay line or, alternatively, the reflection of a fixed target has to be used, always assuming that the spectrum of such a target is narrower than that produced by the transmitter instabilities under investigation.

Figures 14 and 15 show the spectra of two different radar transmitters. For both spectra an isolated target has been selected and the range gate width set to 5 μsec. Whereas the first spectrum consists mainly of prf-lines, the second shows a strong phase modulation of approx. 50 Hz.

3.4 CONCLUSION

Owing to the fact that the power spectrum measurement system described is relatively small and thus transportable, it can easily be used for recordings of clutter spectra on different radar sites. The addition of a small computer, together with optimization of the FFT programme could allow real-time computation of the spectra and consequently make the system useful for the on-site evaluation of same.

ACKNOWLEDGEMENTS

We would like to express our thanks to Dr. Weiss and Dr. Roux of the Luftwaffenamt, Porz-Wahn, Germany, for useful discussions and for the release of wind profile measurements. We are also grateful to Mr. Kampstra, STC, for continuous encouragement, and to Mr. van Slingerland, STC, for his assistance in building the power spectrum analysis equipment.

REFERENCES

1. Barlow, E.J. "Doppler Radar", Proceedings of the IRE, Vol. 37 pp 340-355 (April 1949).

2. Skolnik, M.I. "Introduction to Radar Systems", McGraw-Hill Book Comp., Inc., New York (1962).

3. Baars, E.P. "Digitale Registrierung von Radar-Echosignalen mit dem Radicord-Gerät", NTZ 1968, H 12, p 756.

4. Wirth, W.D. "Clutter- und Signalspektren aus Radicord-Aufnahmen und die Bewegzielerkennung durch Impuls-Radar", NTZ 1968, H 12, p 759.

5. Nathanson, F.E. and Reilly, J.P. "Clutter statistics which affect radar performance analysis", Supplement to IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-3, No. 6, November 1967.

6. Nathanson, F.E. and Reilly, J.P. "Radar Precipitation Echoes", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-4, No. 4, July 1968.

7. Karstensen, H. "Quantitative Untersuchungen der Radarreflektivität in bezug auf die Niederschlagintensität", Deutsche Versuchsanstalt für Luft- und Raumfahrt, März 1968.

8. "Aerologische Berichte. Radiosondenaufstiege und Höhenwindmessungen", Zentralanstalt für Meteorologie und Geodynamik in Wien, 1952-1956.

9. Steinhauser, F. and Cehak, K. "Der Höhenwind über Wien bis 30 km (1952-1956)", Archiv für Meteorologie, Geophysik und Bioklimatologie, Series A, Vol. 16 (1967), Springer-Verlag Wien-New York.

10. Cochran, W.T. "What is the Fast Fourier Transform?", Proceedings of the IEEE, Vol. 55, No. 10 (1967).

11. Blackman, R.B. and Tukey, J.W. "The Measurement of Power Spectra", Dover Publications, INC., New York (1958).

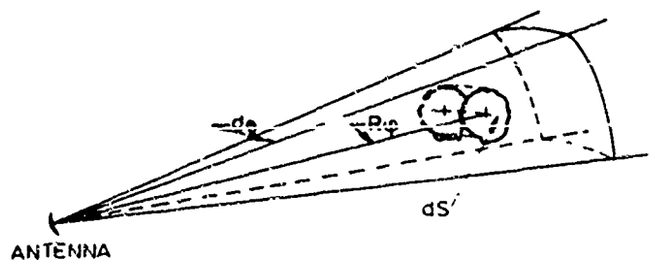


Fig. 1 "Infinitesimal volume element of reflecting medium".

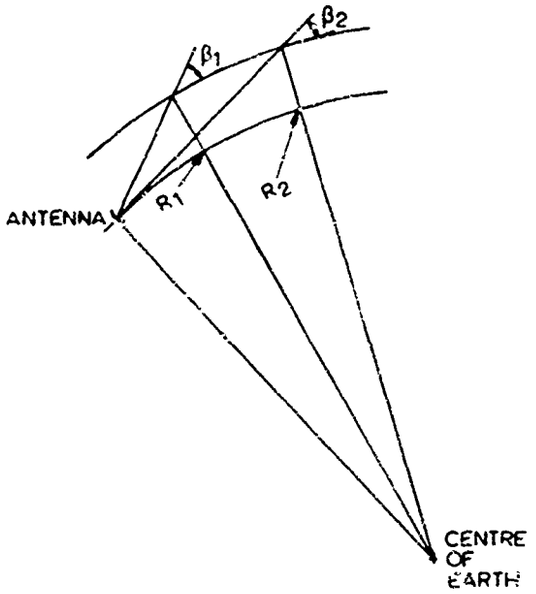


Fig. 2 "Variation of vertical angle with range".

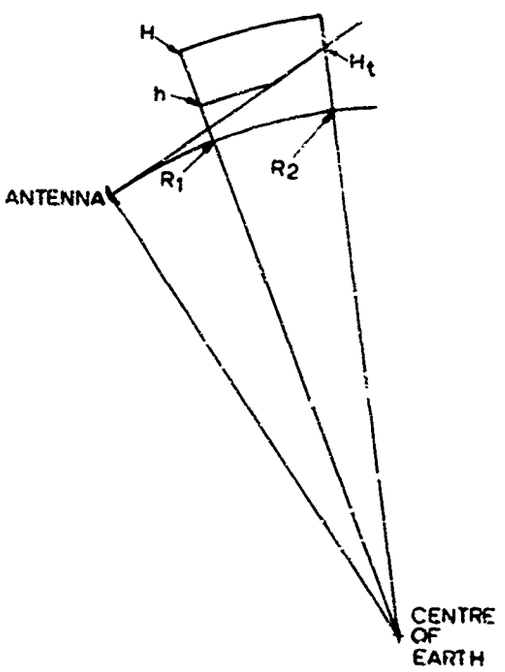


Fig. 3 "Typical effective cloud region".

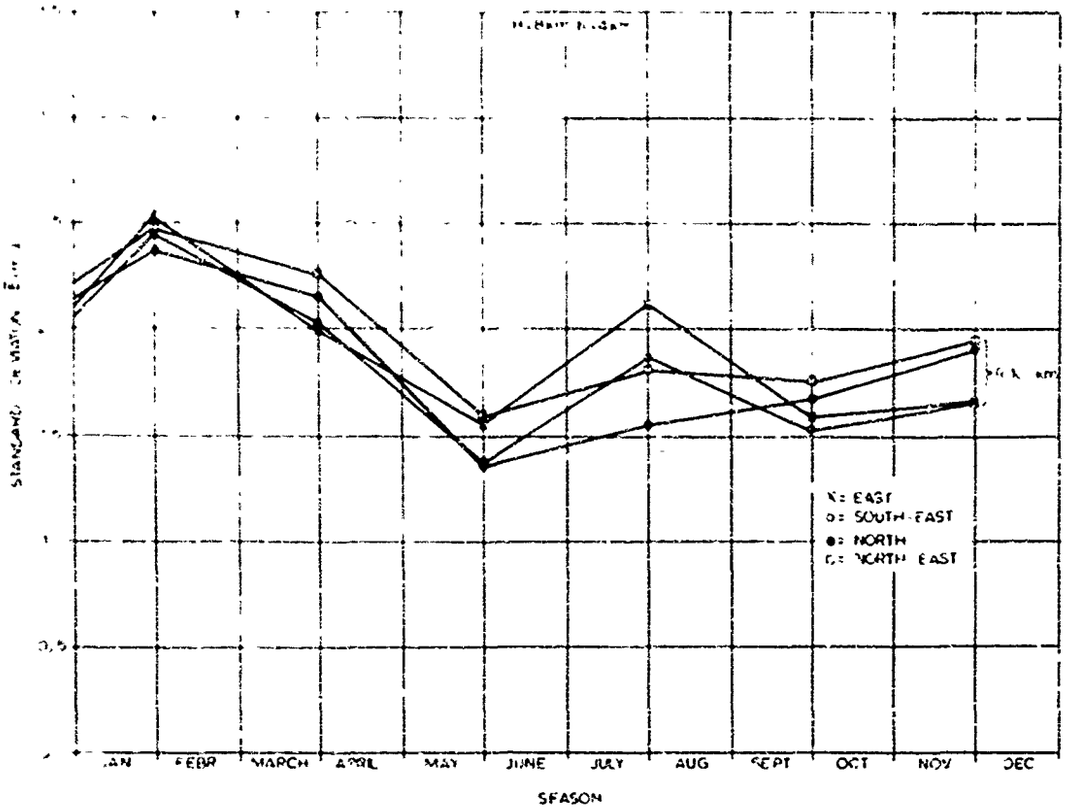


Fig. 4 "Standard deviation $\bar{\sigma}$ as a function of the season with the beam direction as a parameter (h = 4 km; H = 8 km)".

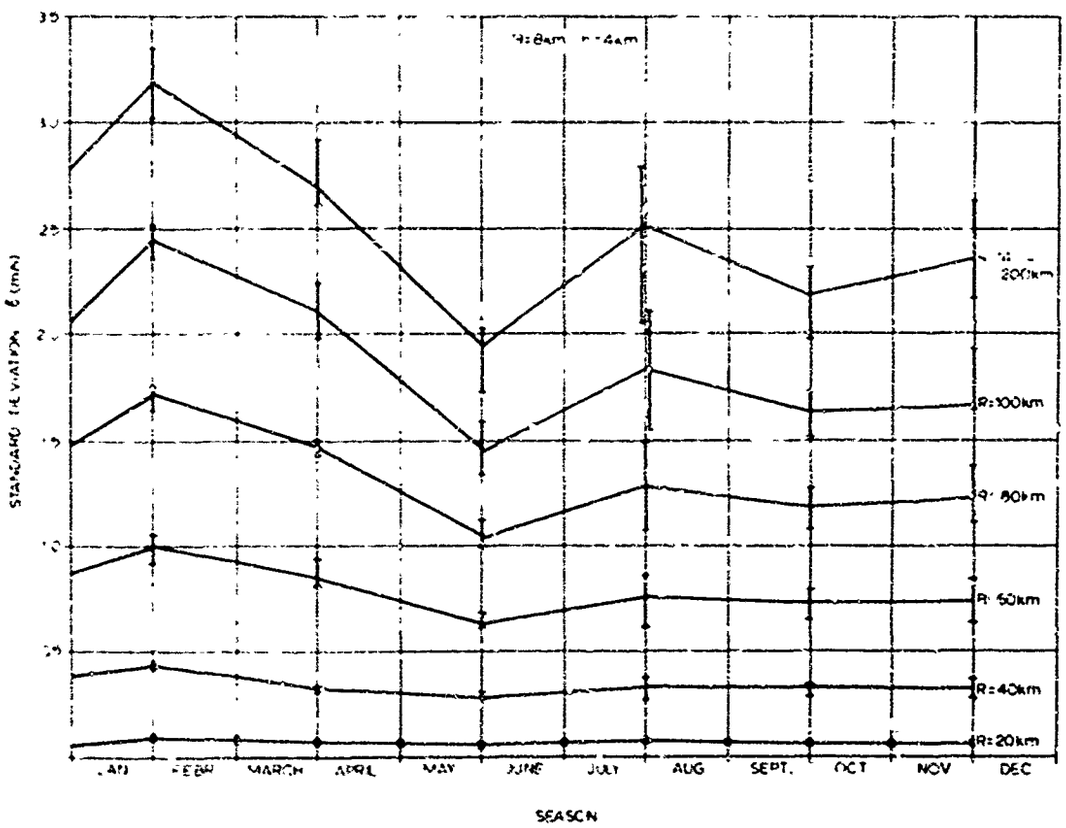


Fig. 5 "Standard deviation $\bar{\sigma}$ as a function of the season for different ranges from the radar (h = 4 km; H = 8 km)".

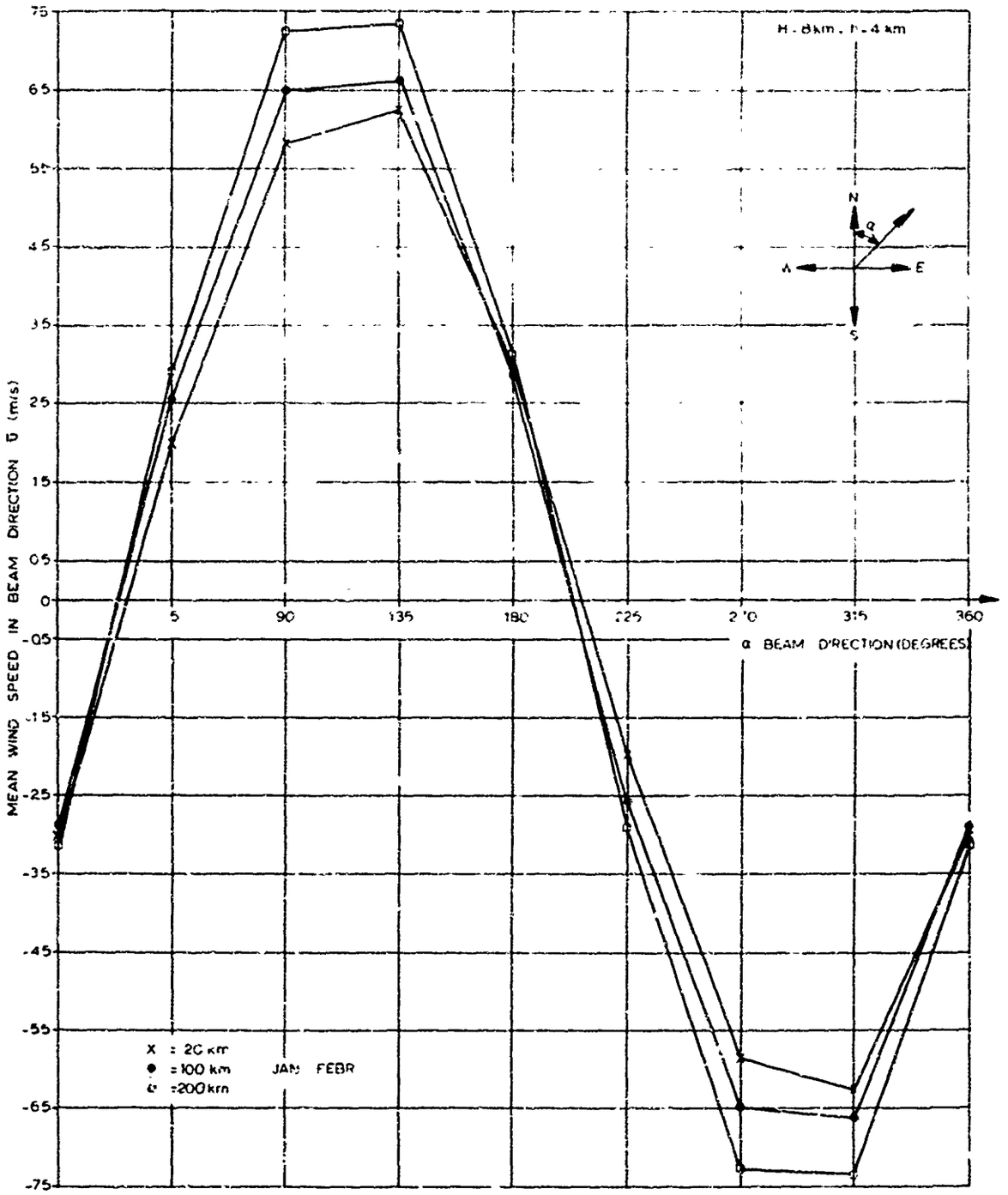


Fig. 6 "Mean wind speed in beam direction $\bar{v}$ versus beam direction in Jan.-Febr. for different ranges from the radar (h = 4 km; H = 8 km)".

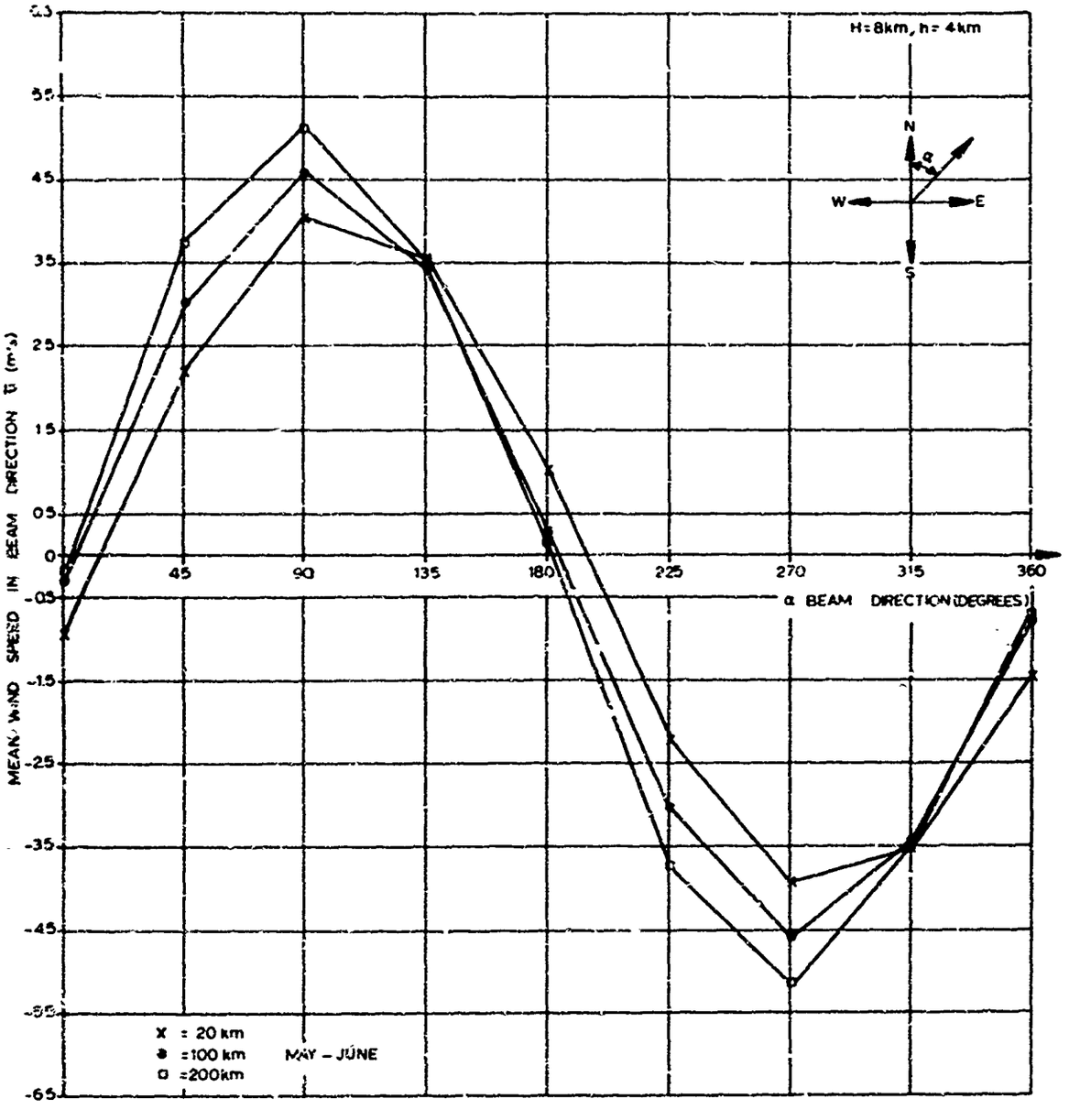


Fig. 7 "Mean wind speed in beam direction $\bar{v}$ versus beam direction in May-June for different ranges from the radar (h = 4 km; H = 8 km)".

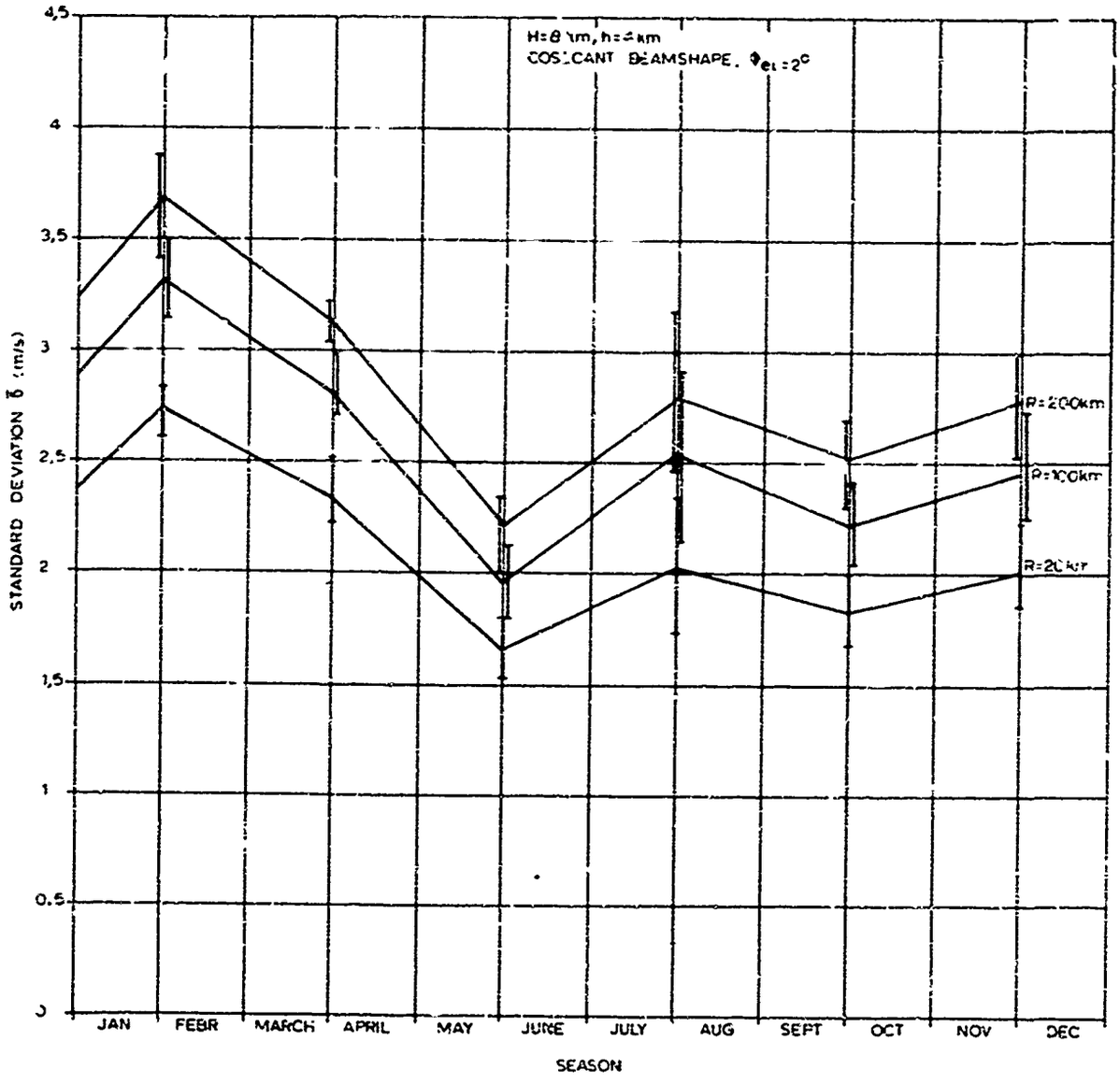


Fig. 8 "Standard deviation $\bar{\sigma}$ as a function of the season for different ranges from the radar (h = 4 km; H = 8 km; cosec² beam shape)".

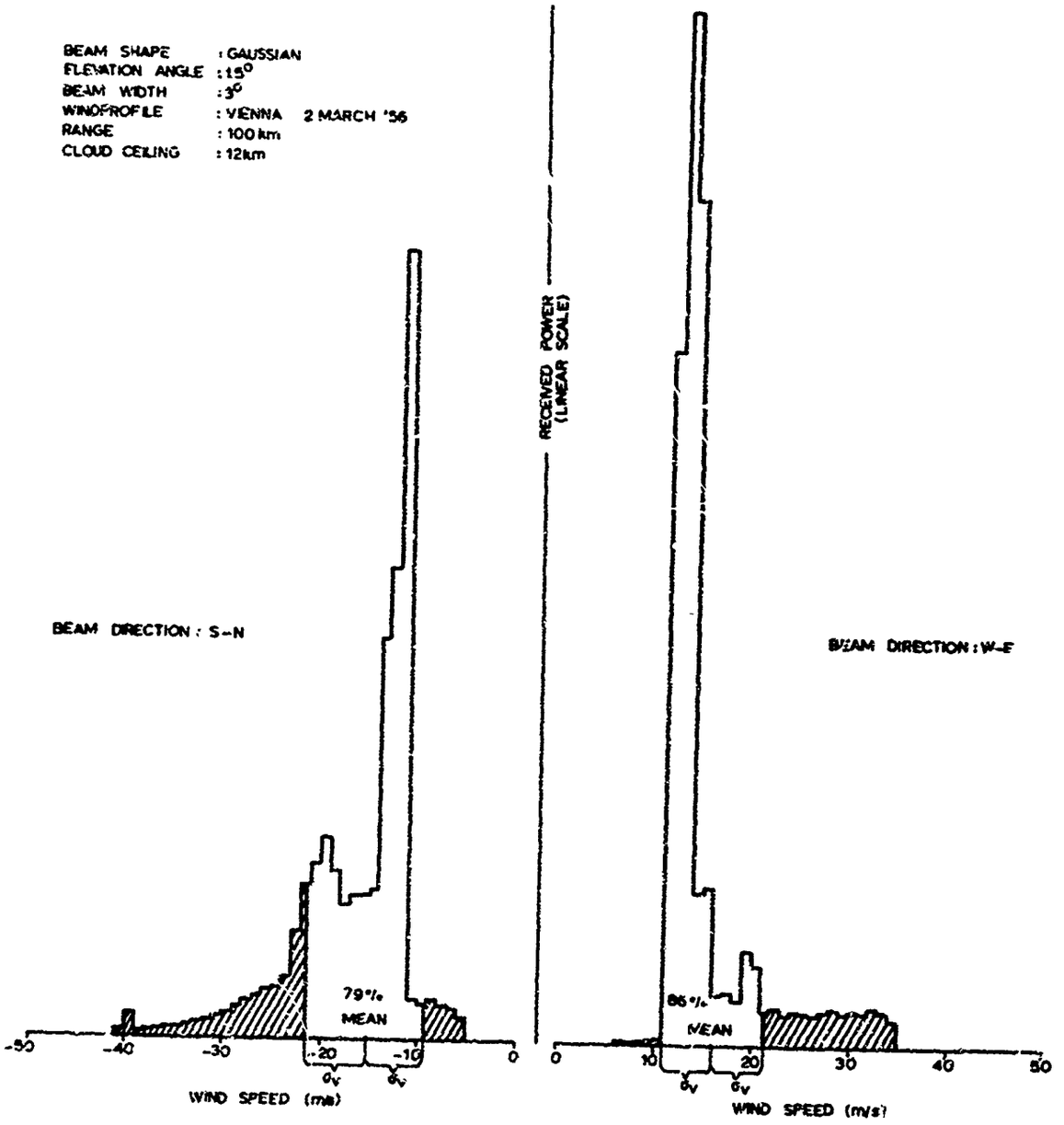


Fig. 9 "Examples of line shapes of the velocity spectrum obtained with the new wind shear mod[illegible]."

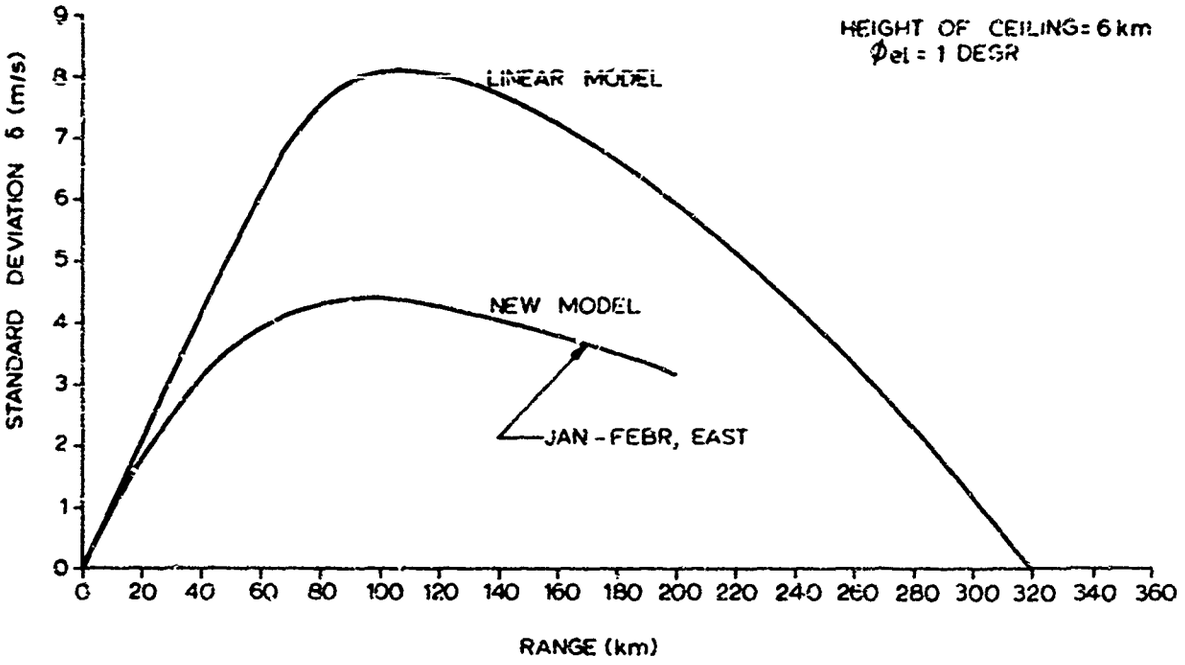


Fig. 10 "The results of the new model composed to those obtained using the linear wind shear model".

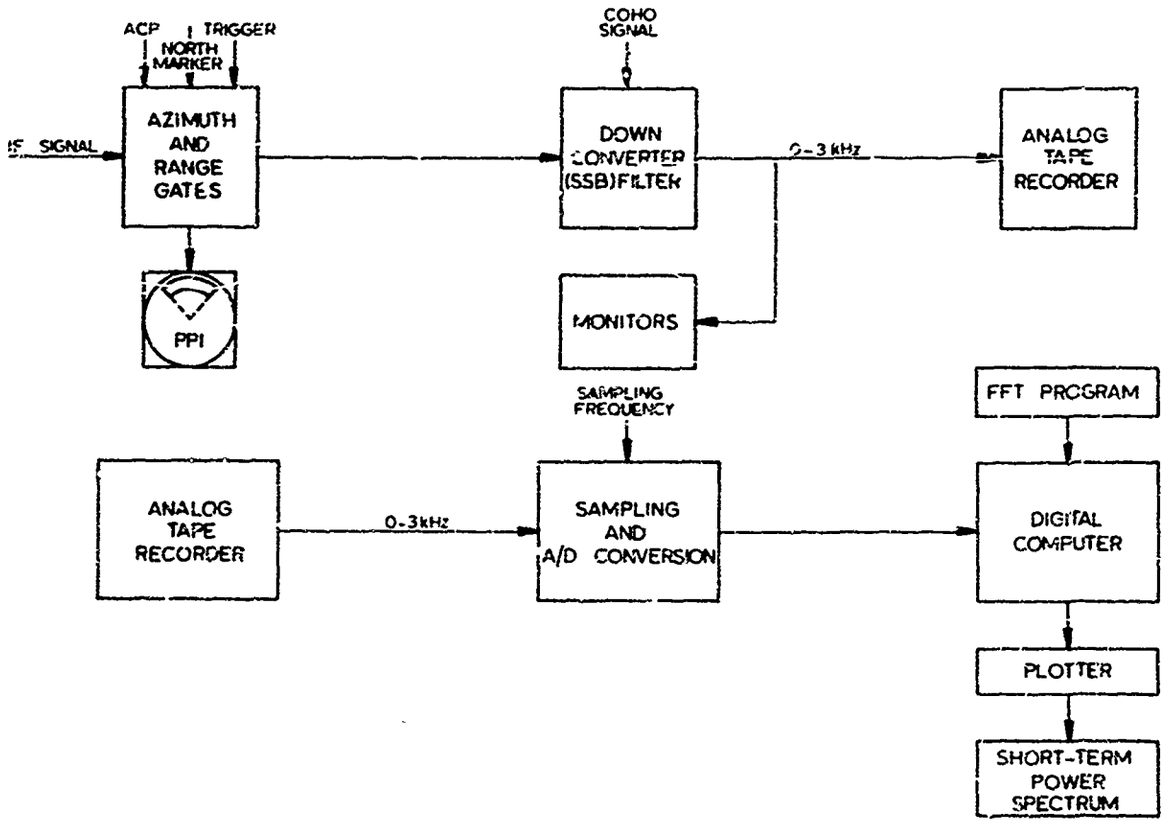


Fig. 11 "Block diagram of the system for measurements of short-term power spectra of radar signals".

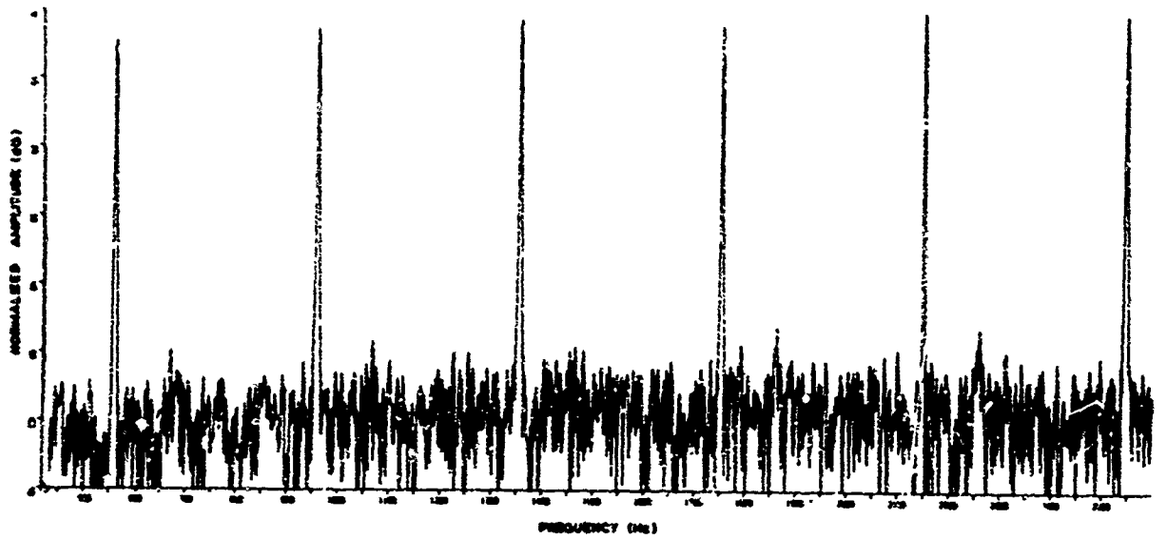


Fig. 12 "Spectrum of a pulsed 60 MHz signal from a frequency synthesizer".

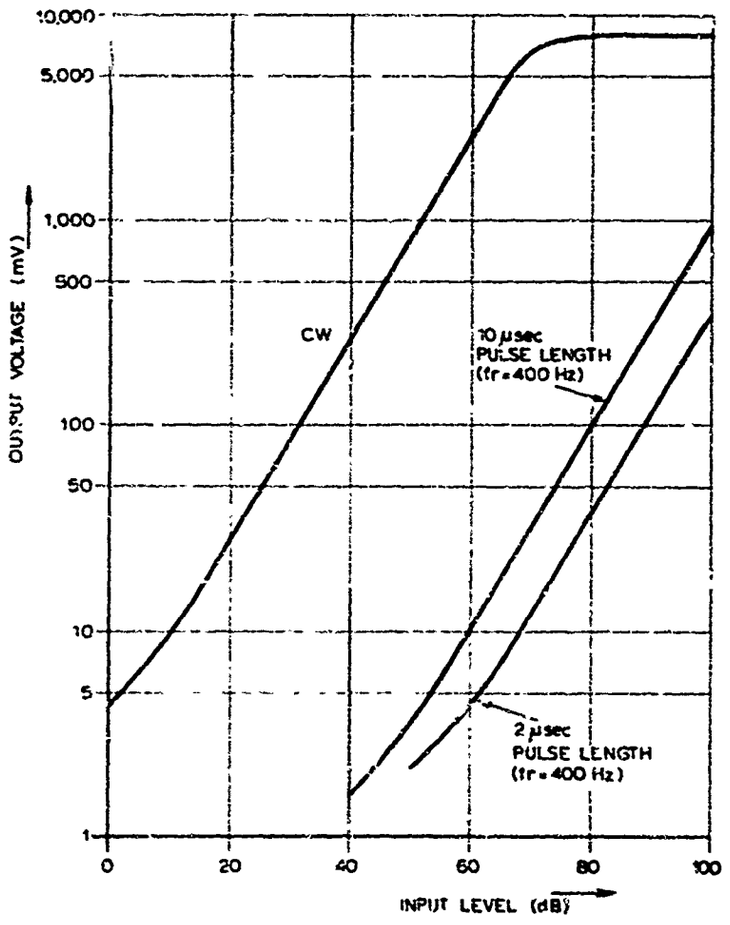


Fig. 13 "Input-Output characteristic of the down converter".

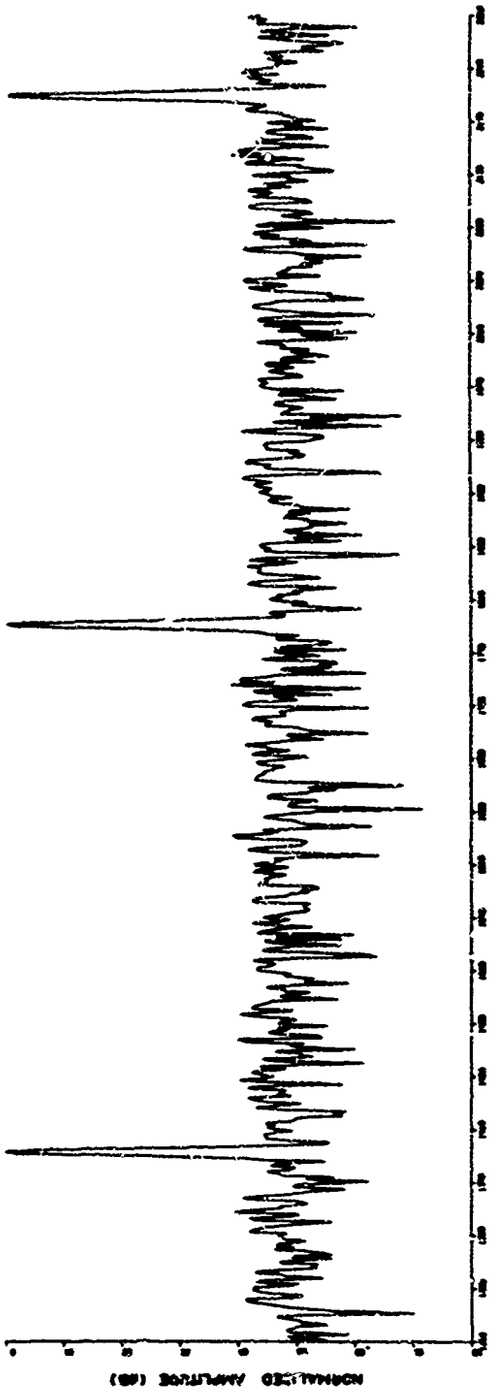


Fig. 14 " Transmitter spectrum of a coherent radar".

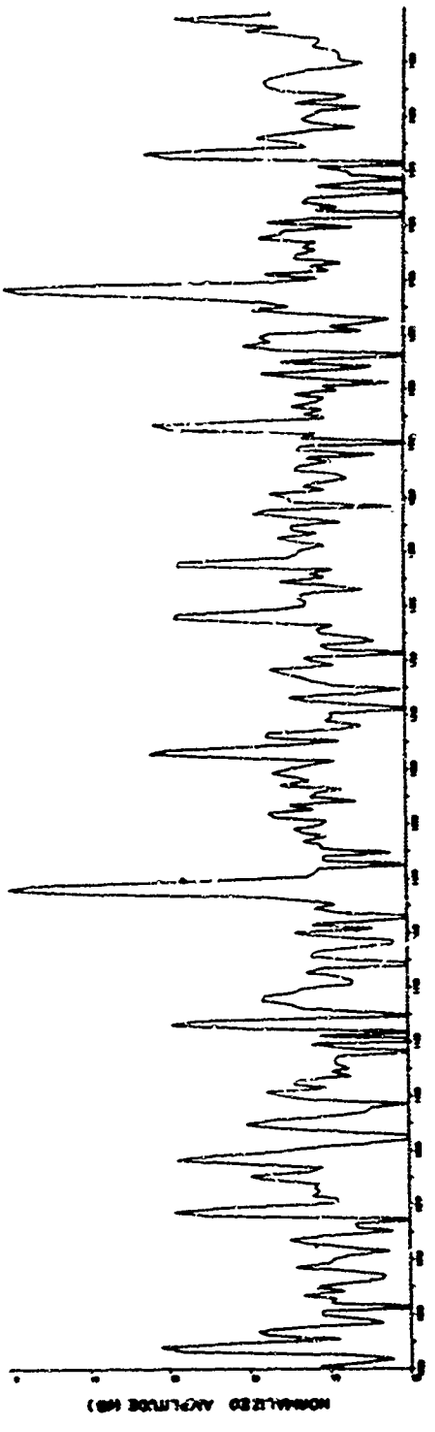


Fig. 15 "Spectrum of a phase-modulated radar transmitter signal".

SOME MEASUREMENTS OF RADAR ANGELS

(FALSE RADAR TARGETS)

by

W. Fishbein, E. Frost and W.E. Vander Meer

US Army Electronics Command
Fort Monmouth, New Jersey, USA

SOME MEASUREMENTS OF RADAR ANGELS
(FALSE RADAR TARGETS)

W. Fishbein
E. Frost
W. E. Vander Mee
U S Army Electronics Command
Fort Monmouth, New Jersey, USA

SUMMARY

There have been many reports of unexplained targets on radar displays. These targets have been given a variety of names, "radar angels", "dot angels", and clear air echoes, to name a few. The density of "radar angel" echoes can often be high enough to cause serious interference to radar operation.

An investigation was conducted to determine some of the detailed characteristics of these echoes. The investigation consisted of observation and analysis of "angel" echoes detected on the "B" display of a 16 gigahertz (GHz) radar and gathering of doppler frequency signatures with a 10 GHz coherent radar.

The results of the investigation show that nearly all the "angels" observed were radar reflections from birds and insects. These results are consistent with observations of other investigators.

The conclusion drawn from the investigation is that in the design of a sensitive radar, the problem of "sky clutter" caused by birds and insects must be considered. Due to its doppler frequency characteristics, "sky clutter" presents a more difficult discrimination problem than conventional "ground clutter."

1. INTRODUCTION

1.1 Sensitive radars can detect many airborne targets in an apparently clear atmosphere. The majority of these targets are birds and insects. Their density can be high enough to cause serious degradation in the performance of a radar. The problem of "sky clutter" caused by birds and insects must be considered in the design of a sensitive radar.

2. BACKGROUND DATA

2.1 In recent years, a number of investigators have attempted to discover the origin of "radar angels". Most have concluded that "angel" echoes are radar returns from birds and insects. Limited data on radar cross section of some birds in flight were obtained by Konrad[1] and Edwards[2]. A summary of their results is shown in Figure 1. In addition, Konrad[3] has shown that the cross sections follow log normal statistics. A log normal statistic is characterized by a large dynamic range and variance[4]. Cross section data of some insects were measured by Hajovsky[5]. His results, shown in Figure 2, illustrate the spread of cross section for various size insects at X-band frequencies. The subscripts refer to the orientation of polarization with respect to the insect's body; L for parallel and T for transverse to the body length. These data indicate that insect cross sections are sensitive to polarization and that the polarization sensitivity is more pronounced for greater body length to diameter ratios. Konrad[6] and Glover[7] have measured both bird and insect cross sections at several wavelengths. The effect of wavelength is shown in Figure 3. The plotted points represent averages for several types of both birds and insects and the spread about each point can easily exceed plus or minus an order of magnitude. The slope of the insect curve for wavelengths greater than 10 centimeters is not exact, but an upper bound of a λ^{-3} dependence has been shown[8].

2.2 Since the detection of targets by radar in a clutter environment normally makes use of the doppler principle, the velocity characteristics of birds and insects are of interest. Figure 4 shows the recorded flying speeds of some birds, as compiled by Houghton[9]. The range of speeds in most cases is broad. Fowler[10] has verified that when the wind speed is greater than the insect's maximum air speed, the velocity characteristics of the insect will be mainly controlled by the wind.

3. DISCUSSION OF EXPERIMENTS CONDUCTED

3.1 The basic raw data for the investigation was obtained from observations and photographs of the B-scope display of a 16 gigahertz (GHz) radar set and doppler recordings from a coherent 10 GHz radar set. The B-scope display data will be discussed first.

3.2 The series of photographs shown in Figure 5 were taken the evening of 10 September 1959 at Sandy Hook, New Jersey. The times listed below each photograph are in Local Daylight-Saving Time (LDT). The radar set was capable of detecting a 10 square centimeter target at a range of 10 kilometers (Km). The B-scope displays a 25° (400 mils) by 10 Km sector and the elevation of the radar beam from horizontal is about 3° (50 mils). The sequence of pictures illustrates an "angel" echo density versus time relationship that was observed rather consistently at several different locations. The large echoes beyond 2 Km in the first photograph are from fixed ground targets. They may be seen in the same location in each of the subsequent photographs. The echo density increases rapidly between an hour and an hour and a half after sunset, peaks within the next two hours, and diminishes slowly for the remainder of the evening. Such a pattern correlates well with known insect behavior.

When the radar beam was elevated to 11° (200 mils), a marked decrease in the number of echoes beyond 2 Km was observed indicating a discontinuity in "angel" density gradient as a function of altitude. The photograph taken at 2100 hours shows vividly the degradation in operational capability that can result because of "radar angels".

3.3 When observing "angel" echoes on the B-scope, motion was discernible and a rough displacement in range and azimuth with time could be measured. This effect was investigated in some detail on 23 October

1969 at the Evans Laboratory of Fort Monmouth, New Jersey. The azimuth of the radar beam was adjusted until the "angel" echoes on the display exhibited only range displacement. That azimuth was then given a reference value of zero. Multiple exposure photographs were produced by taking 10 successive photographs of the B-scope on a single photographic plate, each photograph being 5 seconds apart. A series of four such composite photographs, where the azimuth of the radar was rotated 400 mils clockwise after each composite photograph, is shown in Figure 6. The range at the base of the photograph is 4 Km and that of the bright horizontal line is 6 Km. Measurements of the velocity of several of the echoes were made from the photographs and were found to correlate well with wind data verifying the results obtained by other investigators[11-12]. However, not all targets in the photographs are so completely influenced by the wind, as can be seen by the target trace below and on the right hand side of the 6 Km range marker in the 1200 mil azimuth photograph.

3.4 Doppler recordings of "radar angels" were obtained by instrumenting a coherent 10 GHz radar, as shown in Figure 7. The radar beam was held at a constant azimuth and elevation angle. Using one of the oscilloscope channels as an A-scope display and the other to display the delayed trigger, targets were selected for processing in the Range Gated Doppler Filter (RGDF) by observing the oscilloscope and adjusting the pulse generator delay. Processing the recorded doppler through a low frequency analyzer produced a hard copy doppler frequency versus time plot, such as the one shown in Figure 8. The abscissa represents a 10 second doppler sample. The ordinate represents frequency; a scale is shown on the right. The horizontal lines at the bottom were generated in the tape recorder and spectrum analyzer and are due to the fundamental and harmonics of the prime power source. Harmonics of the doppler signal can be seen faintly and were caused by overdriven amplifier stages in the RGDF on large amplitude peaks. The doppler signal itself is seen to have an erratic variation of frequency with time. For the first two seconds of the sample, the spectrum exhibits a period broadening at an 11.4 Hz rate. It is felt that a doppler signature such as this could only be caused by something possessing self-locomotion. The target is probably a bird.

3.5 The doppler signature of Figure 9 shows two lines that are remarkably sinusoidal in shape if folding about the zero frequency line is taken into account. Such a signature would result if two birds were to fly at a constant still-air speed over a circular path in a constant velocity wind having a component in the direction of the radar. The spacing between similar peaks of the sinusoids is a measure of the separation between the two birds if they have the same flight path. It is believed that this frequency-time plot is indeed the doppler signature of two birds flying a 25 meter diameter circular path at a still-air speed of 7.5 meters/second. Their velocity relative to the ground is modified by a wind that has a 3.7 meter/second component in the direction of the radar. If their flight paths are approximately the same, the birds are separated by a 71° arc or an arc length of 16 meters. Figure 10 shows a flock of between 15 and 20 birds flying a circular path 35 meters in diameter. The still-air speed of the birds is 9 meters/second, which is equal to the wind component in the direction of the radar, as shown by the tangency of the doppler signatures with the zero frequency line.

3.6 In Figure 11, the doppler signature spectrum again shows a periodic broadening. Here the rate is about 12.5 Hz. The spectrum broadening has a duration of about 1/2 to 1 second. Changes in doppler frequency are the greatest during periods of spectrum broadening. The periods of narrow doppler spectrum exhibit a drop in doppler frequency. This is unquestionably the doppler signature of a bird. An examination of the doppler signal waveform indicated that the spectrum broadening was due to amplitude modulation caused by the bird's wing motion. When the bird ceases to flap his wings and glides, wing modulation ceases and the bird's velocity decreases due to drag, as indicated by the negative slope of the doppler signature. If the bird does not fly in a straight line, positive slopes of the doppler signature can be obtained during periods of no wing motion.

3.7 Target signatures from wind borne objects are vividly displayed in Figure 12 by a 187 Hz band centered at 625 Hz. Also shown are two birds whose wing beat rate is 8.8 Hz. The band structure and signal intensity of the wind borne target signatures indicates that the targets are large in number, small in radar cross section, and have some statistical velocity distribution. Rain has similar characteristics, with the exception that the radar cross section is larger. Figure 13 is a doppler signature for wind blown rain. The figure represents a 20 second doppler sample; note the change in the frequency scale. Rain data have been gathered on several occasions and have consistently exhibited a double band structure. The cause of such a structure is not completely understood. Because of the two-band structure of rain, one may conclude that the wind blown targets of the previous figure are something other than rain.

3.8 Another frequency-time plot of birds and wind borne targets is shown in Figure 14. There is no sign of a double band structure in the signature. Figure 15 is a 20 second doppler sample where the frequency scale has been expanded. Several doppler signatures can clearly be seen within the band around 300 Hz. These signatures could be from birds, but the strong possibility exists that they are from large insects, since the doppler frequency variations are more erratic than any observed outside a band type of signature. The remaining targets that contribute to the band doppler signature could be small insects.

4. SUMMARY

4.1 In summary, the data presented has shown that radars having sufficient sensitivity to detect targets of fractional square centimeter radar cross sections can detect birds and insects. The operational capability of such radars can be severely impaired by the presence of "radar angels". "Radar angels" were observed to be, on the whole, wind correlated and to have density versus time characteristics that correlate well with known insect and bird behavior. Analysis of the doppler signatures obtained led to the conclusion that almost all the observed "radar angels" were either birds or insects.

5. CONCLUSIONS

5.1 From the results obtained, it can be concluded that the design of sensitive radars must consider the problem of eliminating "sky clutter" caused by birds and insects. In addition, radar can be an effective tool for both entomologists and ornithologists in the study of the habits of insects and birds.

REFERENCES

1. Konrad, T.G. et. al., "Radar Characteristics of Birds in Flight", Science, Volume #159, 19 January 1968, pp. 274-280.

2. Edwards, J. and Houghton, E., Nature, Volume #194, 1959, p. 1059.

3. Konrad, op. Cit.

4. Heidbreder, G.R. and Mitchell, R.L., "Detection Probabilities for Log-Normally Distributed Signals", IEEE Transactions on Aerospace and Electronic Systems, Volume #AES-3, Number 1, January 1967, pp. 5-13.

5. Hajovsky, R.G. et. al., "Radar Reflections from Insects in the Lower Atmosphere", IEEE Transactions on Antennas and Propagation, Volume #AP-14, Number 2, March 1966, pp. 224-227.

6. Konrad, op. Cit.

7. Glover, K.M. et. al., "Radar Observations of Insects in Free Flight", Science, Volume #154, Number 3752, 25 November 1966, pp. 967-972.

8. Konrad, op. Cit.

9. Houghton, E.W., "Detection, Recognition, and Identification of Birds on Radar", Proceedings of the 1964 World Conference on Radio Meteorology, Boulder, Colorado, 14-18 September 1964, pp. 14-21.

10. Fowler, M.S. and LaGrone, A.H., "Comparison of Insects' Flight Characteristics with Observed Characteristics of Radar Dot Angels", Journal of Applied Meteorology, Volume #8, Number 1, February 1969, pp. 122-127.

11. Browning, K.A. and Atlas, D., "Velocity Characteristics of Some Clear-Air Dot Angels", Journal of Atmospheric Sciences, Volume #23, Number 5, September 1966, pp. 592-604.

12. Chernikov, A.A., "Radar Studies of Echoes from a Clear Sky" (in Russian), Tsentral'naia Aero-logicheskaia Observatoriia Trudy, Number 48, 1963, pp. 56-97 (English Translation: USAF Cambridge Research Laboratories, Bedford, Mass., Technical Report TR-480, July 1965).

BIRD TYPE	$\bar{\sigma}(cm^2)$
SPARROW	2
STARLING	10
GRACKLE	16
PIGEON	30
ROOK	250*

*BROADSIDE ONLY

Fig.1 Average radar cross sections of birds at X-band (from Konrad et al. and Edwards and Houghton)

	BODY LENGTH (mm)	BODY DIAMETER (mm)	σ_L (cm^2)	σ_T (cm^2)
BLUE WINGED LOCUST	20	4	9.60	0.96
ARMY WORM MOTH	14	4	1.22	0.12
HONEY BEE (WORKER)	13	6	1.00	0.30
ALFALFA CATERPILLAR BUTTERFLY	14	1.5	0.65	0.02
RANGE CRANE FLY	13	1	0.30	0.02
GREEN BOTTLE FLY	9	3	0.25	0.10
TWELVE SPOTTED CUCUMBER BEETLE	8	4	0.14	0.05
SPIDER (UNIDENTIFIED)	5	3.5	0.10	0.06
CALIFORNIA HARVESTER ANT	13	6	0.04	0.02
CONVERGENT LADY BEETLE	5	3	0.02	0.01

Fig.2 Radar cross sections of insects at X-band (from Hajovsky et al.)

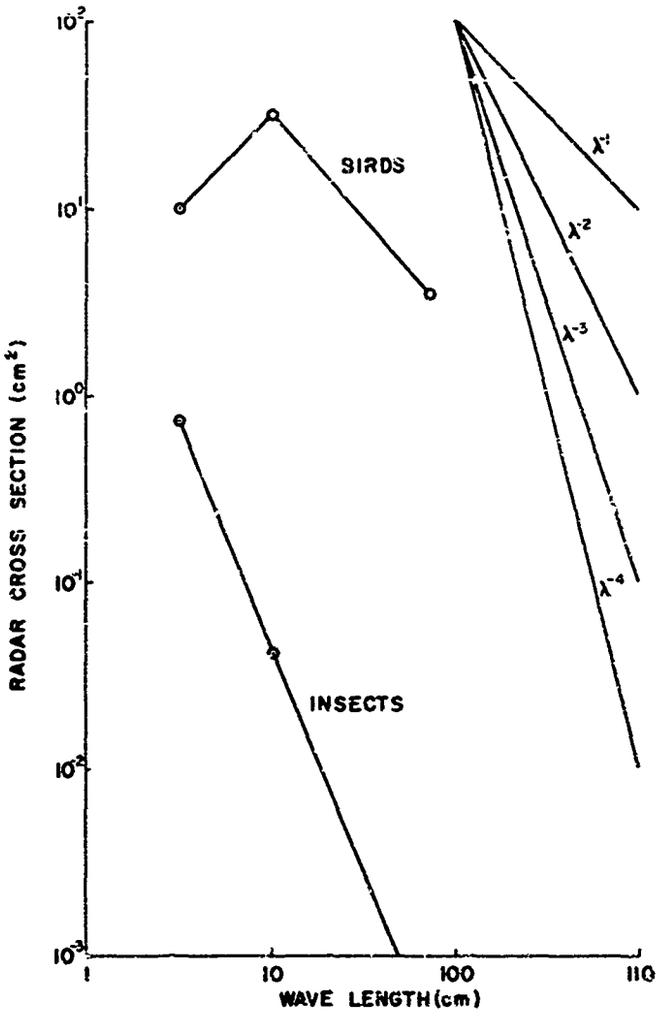


Fig.3 Radar cross section vs wavelength for birds and insects (from Konrad and Glover)

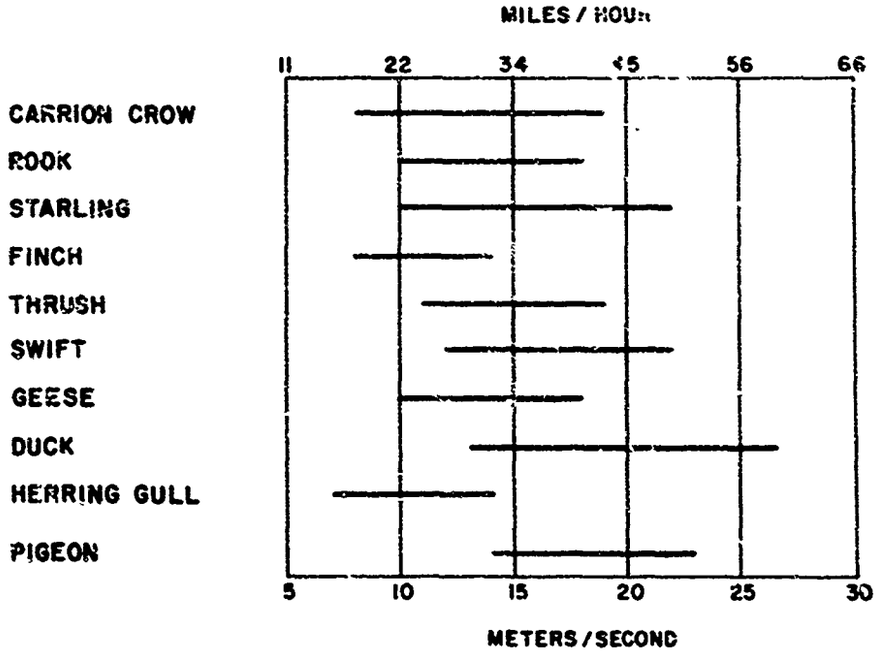
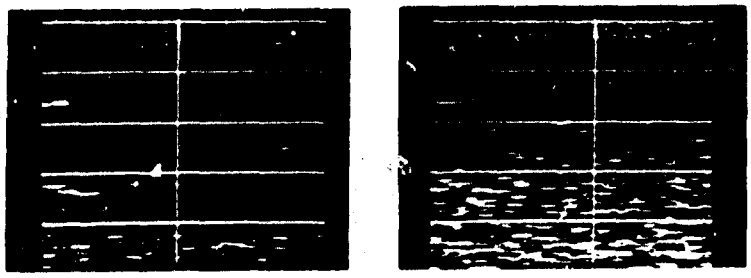
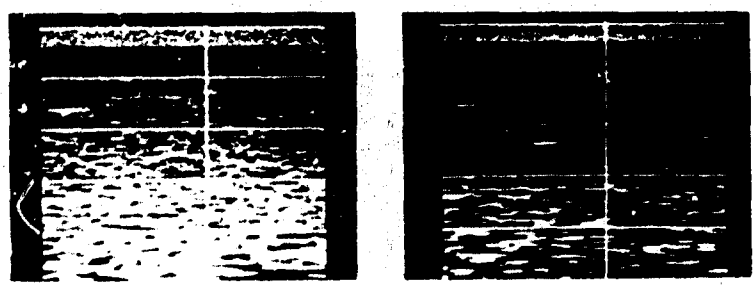


Fig.4 Recorded flying speeds of some birds (from Houghton)



1800 Hours (LDT)

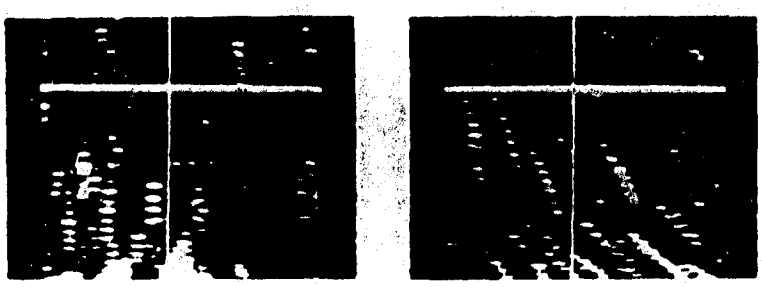
2000 Hours (LDT)



2100 Hours (LDT)

2300 Hours (LDT)

Fig.5 [illegible] "Angel" activity, 10 September 1969, Sandy Hook, New Jersey



AZIMUTH 0000 MILS

AZIMUTH 0400 MILS

AZIMUTH 0800 MILS

AZIMUTH 1200 MILS

Fig.6 "Angel" wind correlation, 2315 hours (LST), 23 October 1969, Fort Monmouth, New Jersey

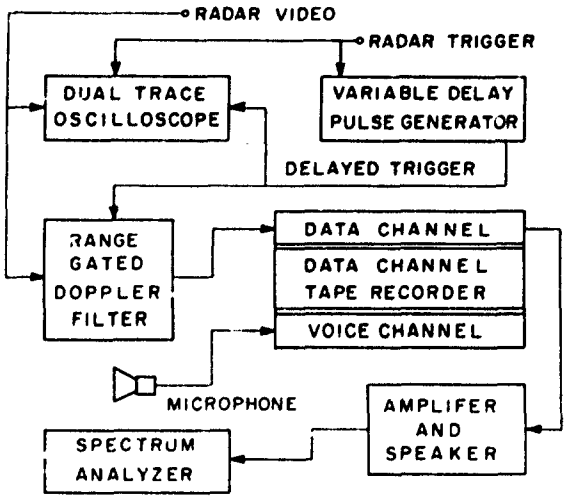


Fig.7 Instrumentation block diagram

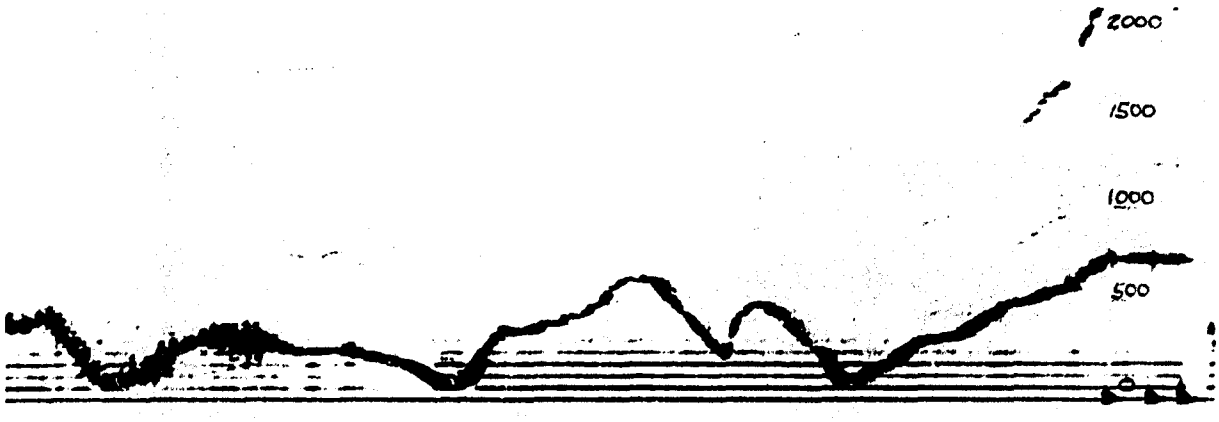


Fig.8 Doppler frequency vs time plot

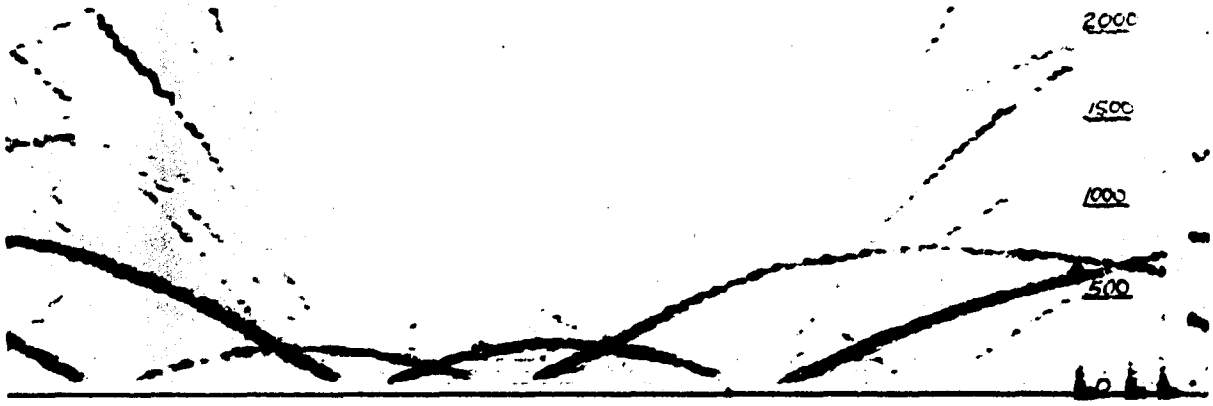


Fig.9 Doppler signature of circling birds

Fig.10 Doppler signature of a flock of circling birds

Fig.11 Doppler signature showing wind beat modulation

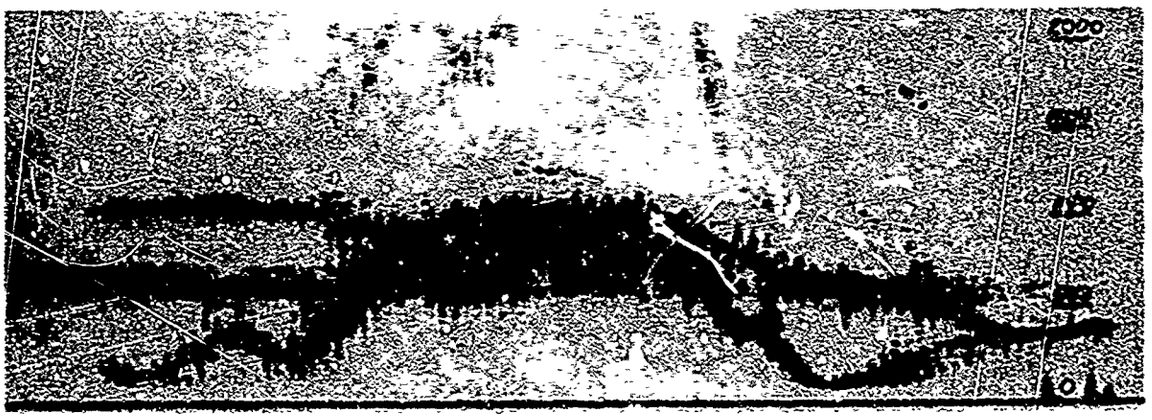


Fig.12 Doppler signature of wind borne objects

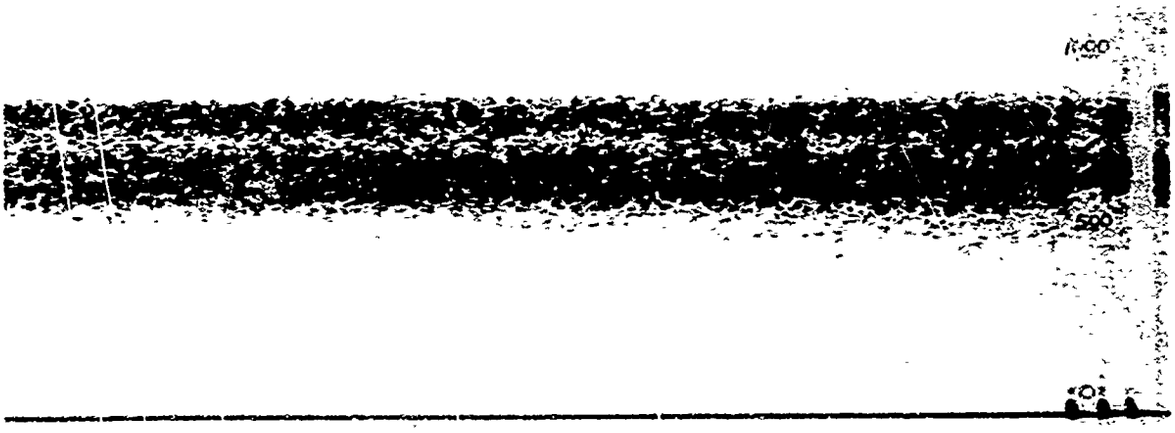


Fig.13 Doppler signature of wind blown rain

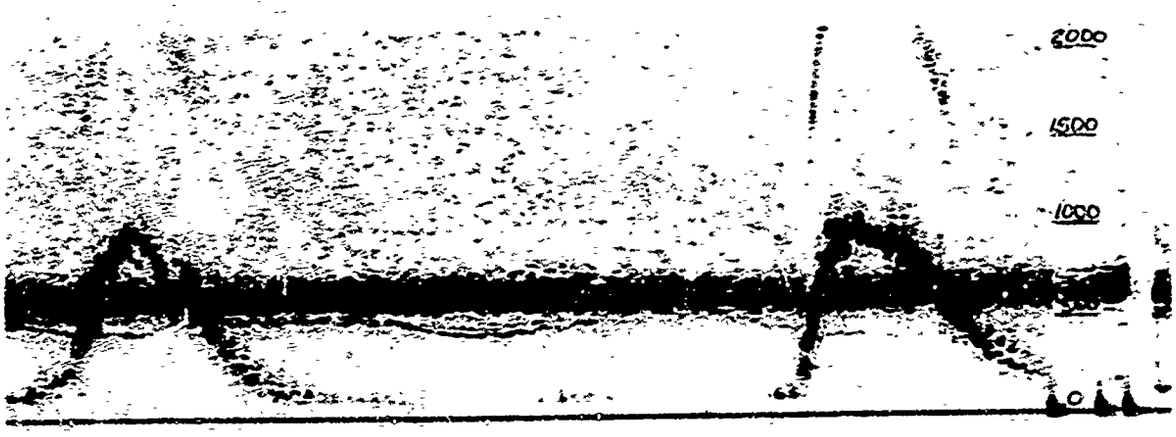


Fig.14 Doppler signature of wind borne targets

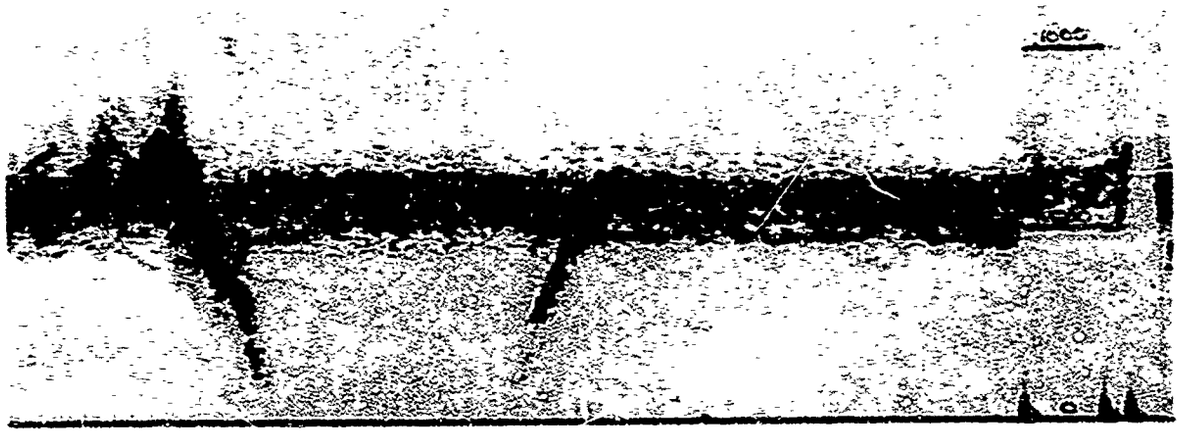


Fig.15 Wind borne targets Doppler signatures with frequency scale expanded

CHARACTERISTICS OF ECHOES FROM AIRCRAFT AND CLUTTER DERIVED FROM DIGITAL RECORDED RADAR DATA

by

Karl von Schlachta

Forschungsinstitut für Funk und Mathematik
5307 Wachtberg-Werthhov :, Königstrasse

CHARACTERISTICS OF ECHOES FROM AIRCRAFT AND CLUTTER DERIVED FROM DIGITAL RECORDED RADAR DATA

Karl von Schlachta
Forschungsinstitut für Funk und Mathematik
5307 Wachtberg-Werthhoven, Königstraße

SUMMARY

The coherent echo signal from different surveillance radars was digitized and stored on magnetic tape with a special recording device (RADICORD).

The evaluation with a computer gives knowledge about the statistical behaviour of different targets and various kinds of clutter.

Echoes from several types of aircraft were recorded. The evaluation gives results about the amplitude and phase shift of the echosignal measured during one scan, and about the fluctuation of the echovector during the flight. One may derive conclusions for the construction of plotextractors and MTI-systems out of this statistical evaluation. The behaviour of amplitude and doppler shift of aircraft echoes along the tracks are of special interest for the problems of automatic tracking.

The evaluation of clutter data gives a survey about the distribution of the power spectral density for several kinds of clutter as ground clutter, wheather clutter, phantom clutter and bird echoes. It follows, that the methods of clutter suppression depend on the kind of the clutter and have to be different for fixed clutter and moving clutter.

1. PURPOSE OF THE INVESTIGATION.

The automatic observation of the air space uses information about the positions of aircraft from radar echoes. The link between the observing surveillance radar and the following tracking computer is the digital detector or plot extractor, which is used for detection and positioning of radar echoes from signals mixed with clutter or noise.

The decision logic of the digitsl detector has to be chosen in such a way that targets will be detected with high probability. The false alarm rate shall be low. The

number of wrong decisions of the detector shall be minimized. This problem can be solved if there is enough statistical information about target and clutter signals.

The fluctuation models of cross section, well known as Swerling case 1 - 4, only take account of the amplitude signal. Moreover the Swerling cases do not base on fully realistic assumptions.

The evaluation of experimental data of coherent radar-echosignals by a computer was possible after development of the RADICORD system (RAdar-DIgitizer and reCORDer) /1/.

The signals of aircraft and different kinds of clutter were measured at several L-band and S-band radar sets. The digitized signals of amplitude and phase were stored on magnetic tapes and evaluated under several aspects.

The following paper shall give a survey about some methods of evaluation and the results of the analysis.

2. DETECTION OF A SINGLE TARGET AND DETERMINATION OF THE CENTRE.

The target-information selected with the gate of the RADICORD fills a two-dimensional data matrix with m range elements (rings) and n azimuth elements (sweeps). This arrangement corresponds to the area scanned with the radar beam.

Each ring contains a sequence of signals:

$$x(k) = a(k)\, e^{j\varphi(k)} \qquad (k = 1 \ldots N)$$

From this sequence a target has to be identified. It is favourable to test the amplitude a(k) because of its modulation by the antenna. Its symmetry is useful for the determination of the centre. It has to be expected that the amplitude sequence a(k) will deviate from the antenna function of a stable fixed target by fluctuation or receiver noise.

If a series of echosignals is expected, it seems reasonable to perform the decision test /2/ with a set of weights. The following test is used:

An odd weighting function b(τ) is stepwise shifted along the sequence a(k) (fig. 5-1) until, at $\tau_o = K_o$,
the value of C

$$C = \sum_{k=-K}^{K} a(k+\tau_o) \cdot b(k), \qquad (1)$$

where $\tau_o = K \ldots (N-K)$,

changes sign or becomes zero. Simultaneously a test is made with the amplitude threshold S as follows:

$$C_1 = \sum_{k=-K}^{K} a(k+K_o) > S \qquad (2)$$

If this condition is satisfied, K_o will be the centre of the detected target. The test is made for every ring of the data matrix. Depending on the ratio "transmitter pulse width to sampling interval" the target is detected on several adjoining rings. The ring with the maximum mean targetamplitude will be selected for the following evaluation.

3. STATISTICAL CHARACTERISATION FOR THE PULSE TRAIN OF A SCANNED ISOLATED TARGET.

We now try to find a method to compare the pulse sequences of different antenna scans and different trials.

3.1. Mean and deviation of the amplitude.

The modulation of the scanning antenna is taken into account by using the mean pattern of a stable fixed target (fig. 5-2) as standard.

$$a_2(k) = a_1(k) \;/\; f(k) \qquad (3)$$

$$\text{with } a_1(k) = a(k+K_o)$$

Then the mean value $\bar{a}$ and the deviation σ_a are calculated for the characterization of the sequence $a_2(k)$ within the 3-db-beamwidth B, B = 2K+1. The deviation σ_a will be normalized with $\bar{a}$ because of the fluctuation of $\bar{a}$.

3.2. Mean and deviation of the doppler shift.

Since the signal is coherently measured the phase information $\phi_1(k) = \phi(k+K_o)$ is available. $\phi_1(k)$ is the phase angle between the transmitted and the received signal. The difference $\Delta\phi(k)$ between adjoining sweeps along a distance ring gives the doppler phase shift and is a criterion for the radial speed of a target.

The vectors z(k) from all 2K+1 values $\Delta\phi(k)$ within the 3-db-width will be calculated next

$$z(k) = \{x(k);\ y(k)\} = \{\cos \Delta\phi(k);\ \sin \Delta\phi(k)\}.$$

In the case of constant pulse repetition time (unstaggered mode) the characteristic phase shift $\Delta\phi$ is derived from the arithmetic mean of the vector components.

$$\bar{z} = \frac{1}{2K+1} \sum_{k=-K}^{K} \{x(k);\ y(k)\} = \{\bar{x};\ \bar{y}\}$$

$$\overline{\Delta\phi} = \arctan\ (\bar{y}/\bar{x}) \qquad (4)$$

In order to obtain a measure of the fluctuation of the phase shift over all k, the

following equation is used /3/ as an index of the relative scatter of the directional vector.

$$\sigma_r = \frac{\sigma_z}{|\bar{z}|} = \frac{\sqrt{\overline{(x(k) - \bar{x})^2} + \overline{(y(k) - \bar{y})^2}}}{|\bar{z}|} \qquad (5)$$

The scatter of the components is added quadratically and referred to the length of the mean directional vector. In the following paper δ is used from

$$\delta = \arctan \sigma_r$$

δ figures as a scattering angle, yielding information on the deviation of phase.

The four calculated values $\bar{a}$, σ_a, $\overline{\Delta\phi}$, δ characterize the echo signal of an antenna scan and therefore may be called the scan characteristic. The scan characteristic is in a certain sense a target signature.

4. THE RELATION BETWEEN THE RADIAL SPEED AND THE DOPPLER PHASE SHIFT.

To explain the later discussed results a well known relation may be mentioned between the radial speed v_r of a moving target and the doppler phase shift $\Delta\phi$

$$v_r = \frac{\lambda}{4\pi} f_i \cdot \Delta\phi \qquad (6)$$

Using a radar set with fixed parameters, for instance the wave length λ and the pulse repetition frequency f_i, one has a fixed relation between v_r and $\Delta\phi$. Certainly there is an ambiguity in v_r because of the ambiguous phase measurement.

The following ambiguous values of the velocity v_r correspond to the range $\Delta\phi = 0 \ldots 2\pi$ for the two analyzed radars:

1. L-band radar (λ = 23 cm, f_i = 500 hz)
   v_r = (0 ... 200) + n · 200 (km/h)
2. S-band radar (λ = 10 cm, f_i = 250 hz)
   v_r = (0 ... 40) + n · 40 (km/h)

5. RESULTS OF THE MEASUREMENTS.

THE BEHAVIOUR OF THE SCAN CHARACTERISTIC DURING A TRACK.

A few examples shall demonstrate the behaviour of the values of the scan characteristic during the tracks of aircraft and birds. Some remarks concerning the presentation on the plots:

The polar coordinates of the moving targets measured with RADICORD are plotted into a rectangular system. The extreme courses such as tangential and radial flight direction with respect to the radar are parallel to one of the axes by this kind of

plot. Mean and deviation of amplitude and phase shift are plotted over the antenna turn.

Example 1: Fig. 5-3a shows the track of a propeller driven aircraft on several airlanes measured with an L-band radar. This example is suitable to demonstrate how the signals depend on the flight direction.

Fig. 5-3b shows the mean amplitude $\bar{a}$ and its deviation $\sigma_a/\bar{a}$ during the track, and fig. 5-3c shows the mean doppler shift $\overline{\Delta\phi}$ and its scattering angle δ along the track.

The behaviour of the mean doppler shift depends on the radial speed of the aircraft as follows:

1. Antenna turn 1 - 20: Almost radial flight shows constant phase-shift.
2. Antenna turn 21 - 82: The reduction of the radial speed is shown in the decreasing of the phase shift, crossing zero at a blind speed of about 200 km/h (turn 46).
3. Antenna turn 83 - 152: After changing flight direction the dopplershift increases in the same manner.
4. Antenna turn 153 - 197: The curve at the antenna turn 152 is very narrow, so that the antenna speed of 6 rpm is too slow for measuring the exact progress of the phase shift. At antenna turn 176 the aircraft comes to tangential flight direction. The phase shift is almost zero.

The scattering angle $\bar{\delta}$ depends on the flight direction and is low if the aircraft is scanned from aside.

Fig. 5-3b shows the dependence of the distance law in the radar equation for $\bar{a}$. The fluctuation between successive antenna scans is nearly 10 db, depending on the fluctuation of the cross section.

The deviation of the amplitude $\sigma_a/\bar{a}$ is low, if there is no influence by the propeller. This is only the case at tangential flight direction (turn 176).

Example 2: Fig. 5-4a shows the track of an approaching small helicopter, flying from 30 km to 5 km distance of the antenna of an L-band radar. During the flight there are small corrections of the flight direction. This can be seen in the plot of the mean phase shift (fig. 5-4c). The mean radial speed of 150 km/h varies between $\pm$ 5 km/h.

The mean amplitude $\bar{a}$ (fig. 4b) increases quadratically in accordance to the decreasing of distance during flight. The fluctuations of $\bar{a}$ do not exceed the plotted limits of $\pm$ 3 dB.

Example 3: Fig. 5-5a. Here is a track of a medium fast aircraft, probably a turboprop aircraft, measured with an S-band radar.
Fig. 5-5b shows the typical fast change of the phase shift because of the low value of the first blind speed (nearly 40 km/h). As shown in example 1 the scattering angle $\bar{\delta}$ is low during tangential flight.

Example 4: A radially flying small jet aircraft was observed with the fixed antenna of an S-band pulse radar (fig. 5-6a, 5-6b). The antenna scanning rate was 1/sec and therefore it was 15 times faster than that of example 3. The mean values $\bar{a}$, $\overline{\Delta\phi}$ are computed from 35 pulses contrary to 7 used pulses at example 3 or 15 pulses at example 1+2.

Consider that the amplitude shows strong fluctuations. The highest frequency of fluctuation lies in the range of the prf at almost some hundred hertz (it cannot be seen from this plot). The amplitude $\bar{a}$ changes during one second up to 14 dB.

The doppler shift changes steadily as a consequence of corrections of the flight direction. The scanning rate of 1/sec is high enough to measure all variations of the doppler shift during this radial flight.

Example 5: Fig. 5-7a shows the tracks of birds measured with an S-band radar. The plots fig. 5-7b, 5-7c of amplitude $\bar{a}$ and phase shift $\overline{\Delta\phi}$ show no difference to the plots of the aircraft (fig. 5-5). But the deviation of the amplitude can be large because of superposition of several birds to a quasi single target.

The doppler shift can show a high scattering angle $\bar{\delta}$ if several birds with different radial speeds are superposed in the resolution cell of the antenna.

6. MEASURING THE VELOCITY FOR THE PURPOSE OF TRACKING TARGETS.

The usual plot extractors for tracking purposes use the Normal - or MTI-video signal and attach the position of the distance and azimuth counter to each recognized target /4/.

For each target an expectation area will be layed out within the tracking program of the computer /5/. The size of the area depends on the precision of measuring distance and azimuth. The usual extractors use a resolution of distance of about 2 µs so that the distance will be measured with a precision of $\pm$ 0,15 km. The covered distance of a jet aircraft at normal speed gives 2,5 km during 10 sec of an antenna turn. This gives an error in measuring of distance up to $\pm$ 6 %. At lower velocities, for example during the approach to the airport, the error can grow to $\pm$ 15 %.

The above mentioned examples have shown, that the doppler shift can be measured with an amazing precision using the few sampled values within the antenna 3-db-beamwidth. By simple conversion (Equ.(6)) the radial component of the velocity can be calculated very exactly.

The precision depends on the accuracy of the phase measuring. The error in case of an approaching flight (400 km/h) can be reduced from $\pm$ 15 % to $\pm$ 1,5 %. The practical method is as follows: Approximate calculation of the velocity from the coordinates and then correction with the results from doppler information.

Fig. 5-8 shows the velocity of a target a) calculated from coordinates and b) corrected with doppler information.

7. COMPARING DIFFERENT TYPES OF TARGETS.

The main purpose of the following investigation is to get characteristic values for single moving targets such as aircraft and helicopter. Nevertheless the evaluation is extended to clutter echoes. It shall be pointed out that all the evaluated data are selected from the data matrix with a detection criterion according to Equ. (1) and (2). This method permits to specify more typical properties of clutter, compared with aircraft, than other realized investigations.

The results of similar trials have been compared to get a statistical estimation of the signals from different targets. Therefore the mean values $\overline{\bar{a}}$, $\overline{\overline{\delta\phi}}$, $\overline{\sigma_a/\bar{a}}$, $\overline{\delta}$ and the deviations $\sigma(\bar{a})$, $\delta(\overline{\delta\phi})$, $\sigma(\sigma_a/\bar{a})$, $\delta(\delta)$ of the scan characteristic have been computed. Some significant results are given in table 5-1 and table 5-2.

To compare amplitude signals with each other the dependence on the distance given by the radar equation has to be eliminated. So the power reduction with distance is canceled by conversion of the signal amplitude.

Table 1 presents the mean amplitudes $\overline{\bar{a}}$ for aircraft and clutter. The value of aircraft, used as reference, was taken from different radially flying aircraft. The small reflection of birds and phantoms is remarkable. This kind of clutter only was measured near the radar.

There is a dependence of $\overline{\bar{a}}$ on the direction of flight. Measurements were made with two different targets. A helicopter and a small jet aircraft gave 6 db (7 db) more amplitude at tangential flight than at radial direction.

The variation of the mean amplitude $\sigma(\bar{a})$ of a traffic aircraft during a track on several airlanes with different flight directions (200 antenna scans) is almost 6 db.

7.1. Mean and variation of the mean variation of amplitude and phase shift.

Some remarks on the interpretation of table 5-2. The value $\sigma_a/\bar{a}$ is low for targets that resemble to the antenna pattern of a single fixed target. Echoes that spread beyond the resolution cell of the radar such as clouds, flocks of birds, phantoms give a high value of $\sigma_a/\bar{a}$. The scattering angle δ of the doppler phase shift indicates the width of the doppler spectrum.

Conclusions from table 5-2.

rst column: 1. The mean variation $\overline{\sigma_a/\bar{a}}$ is very low for jet aircraft and is high for helicopters and propeller driven aircraft.

2. The evaluated ground clutter is composed of many single targets, $\sigma_a/\bar{a}$ is low.

3. Weather clutter shows relative low amplitude dynamics, $\overline{\sigma_a/\bar{a}}$ is high.

4. Phantom clutter probably induced from metereologic layers shows low amplitude dynamics in direction of antenna scanning. $\sigma_a/\bar{a}$ is high. As cannot be seen from table 5-2, the signal extends only over the width of the transmitted pulse in distance direction.

5. Bird clutter shows high $\sigma_a/\bar{a}$ because of superposition of many scatterers.

Second column: $\sigma(\sigma_a/\bar{a})$ indicates the deviation for the values given in column 1.

Third column: 1. $\bar{\delta}$ is low for the fixed target as expected

2. Low $\bar{\delta}$ for jet aircraft, increased $\bar{\delta}$ for helicopter and propeller

driven aircraft. $\overline{\delta}$ of helicopters lies between both kinds of aircraft.

3. Low $\overline{\delta}$ for fixed clutter as expected.
4. High $\overline{\delta}$ for weather clutter because of internal movement of the reflecting centre.
5. High $\overline{\delta}$ for phantom clutter because of internal movement of the layer.
6. $\overline{\delta}$ for birds depends on the number of birds in the resolution cell and on the direction of flight of the single bird. $\overline{\delta}$ is small if birds fly in the same direction. The shown value is a mean value for different situations.

Fourth column: $\delta(\delta)$ indicates the deviation for the values given in column 3.

7.2. Probability density.

Some examples shall be given for the probability density of the parameters within the scan charactaristic.

1. Fig. 5-9 shows the probability density of the normalized amplitude deviation $\sigma_a/\overline{a}$ and the scattering angle $\overline{\delta}$ for two types of aircraft as mentioned above. Top: jet aircraft. Bottom: propeller aircraft. Left side: $P(\sigma_a/\overline{a})$ right side: $P(\delta)$.
2. A good characterization of a target can be given by the probability density of the mean doppler shift. Fig. 5-10 shows typical results for different types of signals.

2.1. Fixed target: The probability density has its maximum for $\overline{\Delta\phi} = 0$, the deviation of $\overline{\Delta\phi}$ is small.

2.2. Phantom clutter: The probability density has its maximum for $\overline{\Delta\phi} = -0{,}17$ (Radians). That is the explanation for the low suppression by MTI-Systems.

2.3. Weather clutter: The curve shows two maxima in the negative part of the doppler shift scale. The measured radial velocities are 45 and 70 km/h. These two speeds are due to spatial distributed zones of different doppler shift. The record was made with a gate of 30 range elements and 70 azimuth elements and was situated always amidst the slowly moving cloud. The typical feature of weather clutter is, that the power of the signal mostly takes a small part of the doppler range and that the speed of clouds is slow compared with aircraft. The above mentioned example is a special case of superposition of two clouds.

2.4. Signal from aircraft. The plot shows the probability density of the mean doppler shift for signals from traffic aircraft during a track on several airlanes with different flight directions. The typical attribute of signals

from aircraft is, that $\overline{\Delta\phi}$ can change between succeeding antenna turns within all values between $-\pi$ and $+\pi$

$$\overline{\Delta\phi_n} - \overline{\Delta\phi}_{n-1} = -\pi \ldots \pi$$

were n = number of the antenna turn.

8. CONCLUSIONS.

The measurements described above go beyond the usual in the literature described investigations of analyzing aircraft on turn tables or of measuring approaching aircraft with a single resolution cell in range.

The described measurements are made with the signals of real flying aircraft. The recorded signal has the full bandwidth of the reflected echo signal.

The presented example of evaluation uses an amplitude detector that is conceived for the detection of single targets in noise. The results show that complex targets as the most clutter signals give values that are different from that of aircraft.

It was tried to classify the targets using the results. The following list shall give criteria, that can be used for separation of different types of targets

1. Mean amplitude with regard to the distance of target (and the vertical antenna diagram).
2. Deviation of the mean amplitude between successive scans.
3. Shape of the amplitude signal
4. Deviation of the amplitude signal σ_a within scan
5. Value of the mean doppler shift
6. Scattering angle δ of the doppler shift within scan
7. Deviation of the mean doppler shift between successive scans.
8. Velocity of the target by comparing coordinates of successive scans.

ACKNOWLEDGMENTS

The RADICORD project was performed according to a contract with the European Organisation for the Safety of Air Navigation, EUROCONTROL. The Bundesanstalt für Flugsicherung and the Royal Radar Establishment kindly made available their radars to the RADICORD measurements. I have to thank Dr. Wirth for discussion of the evaluation methods.

REFERENCES

/1/ Baars, E.P. Digitale Registrierung von Radar Echosignalen
Nachrichtentechnische Zeitschrift NTZ 21 (1968) Vol. 12 P. 756-58

/2/ Stors, W., Wirth, W.D. Automatische Auswertung digitalisierter Radarsignale
Nachrichtentechnische Zeitschrift NTZ (1963) Vol. 12 P. 643-656

/3/ Wirth, W.D. Clutter und Signalspektren aus Radicord Aufnahmen
Nachrichtentechnische Zeitschrift NTZ 21 (1968) Vol. 12 P. 759-765

/4/ Wirth, W.D. Ein Digitaldetektor zur automatischen Radarzielerkennung
Telefunken Zeitung 38 (1965) Vol. 2

/5/ Springer, H. Digitalradar - automatische Luftraumbeobachtung mit einer elektronischen Rechenmaschine
Umschau 1963 Vol. 3 P. 81-85

TABLE 1

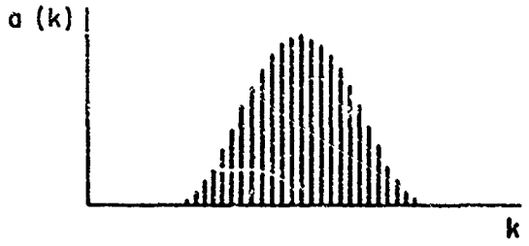
Comparison of the Mean Amplitude $\overline{a}$ for Several Types of Signals

	$\overline{a}$
jet aircraft radial	0 db
single target (tower)	12 db
ground clutter	2 db
weather clutter	- 7 db
bird clutter	- 28 db
phantom clutter	- 39 db

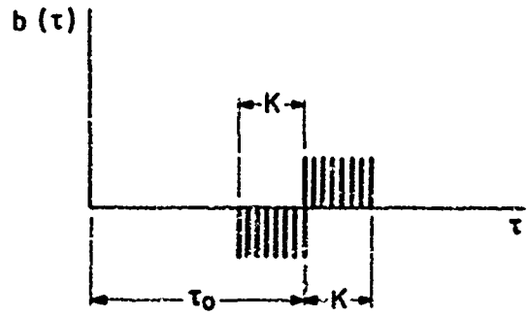
TABLE 2

Mean and Deviation of the Deviation Values in the Scan Characteristic (Radians)

	$\overline{\sigma a/\overline{a}}$	$\sigma(\sigma a/\overline{a})$	$\overline{\delta}$	$\delta(\delta)$
	radians			
single target (tower)	.09	.04	.10	.03
jet aircraft	.15	.07	.23	.18
helicopter	.30	.15	.42	.26
propeller aircraft	.30	.12	.80	.31
ground clutter	.25	.20	.13	.08
weather clutter	.51	.17	.88	.19
phantom clutter	.45	.17	.43	.23
bird clutter	.35	.15	.52	.32



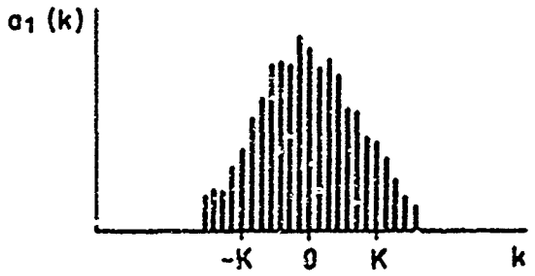
amplitude pattern
of a single target



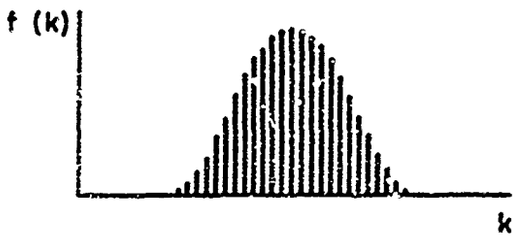
weighting function

3 - db - beam width ≃ 2K + 1

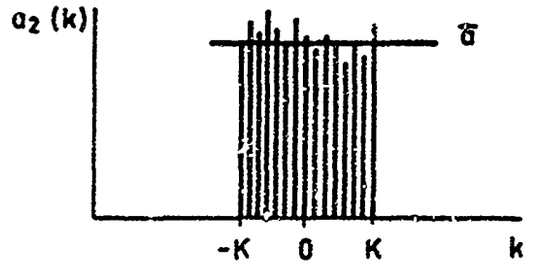
Fig.1 Determination of the centre for a single target



signal from
aircraft



signal from
fixed target



normalized
signal

$$a_2(k) = \frac{a_1(k)}{f(k)}$$

Fig.2 Normalizing of amplitude

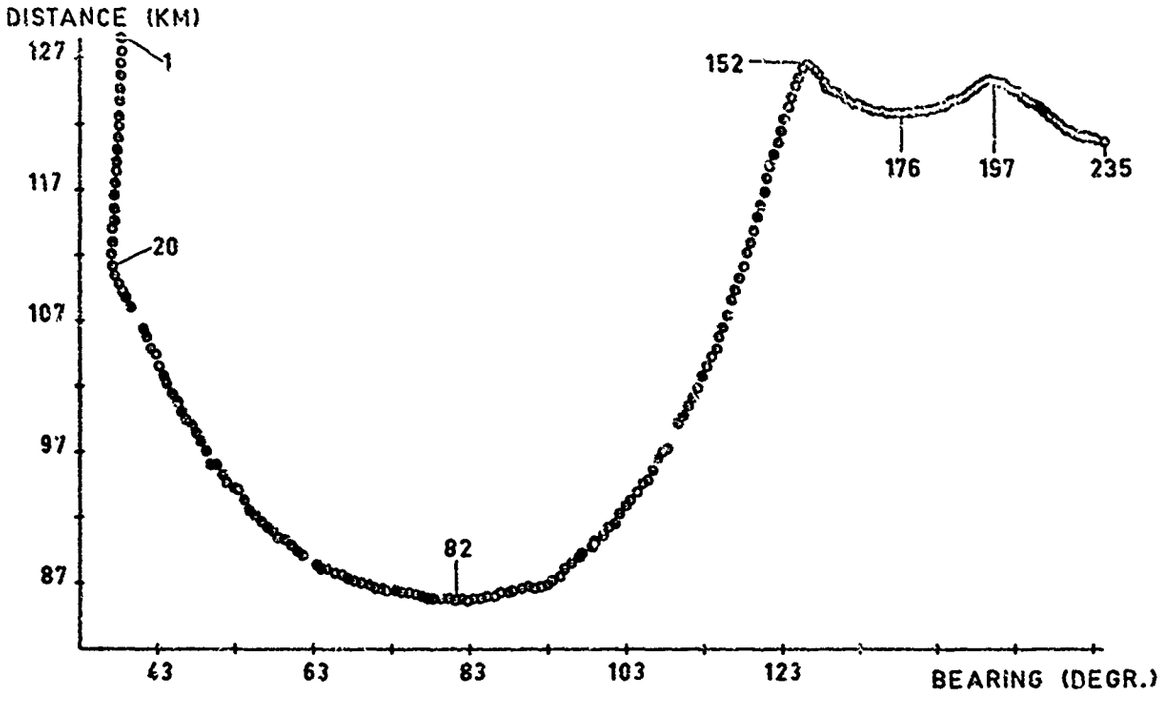


Fig.3a Track of a propeller aircraft L-band-radar

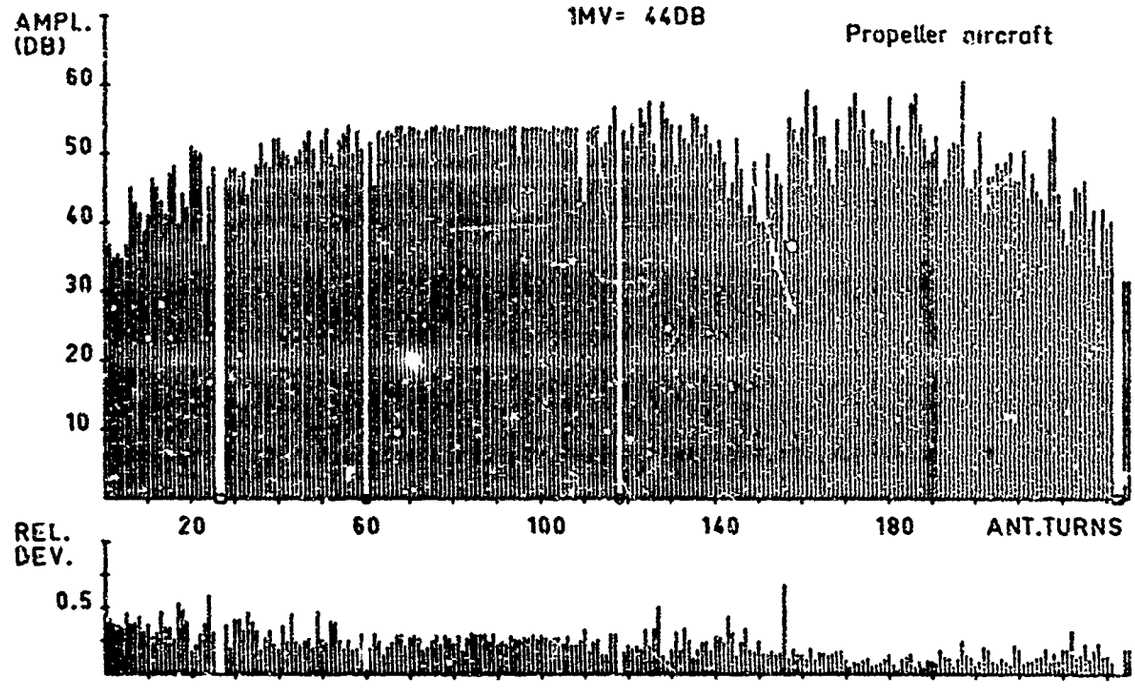


Fig.3b Mean and deviation of target amplitude along the track

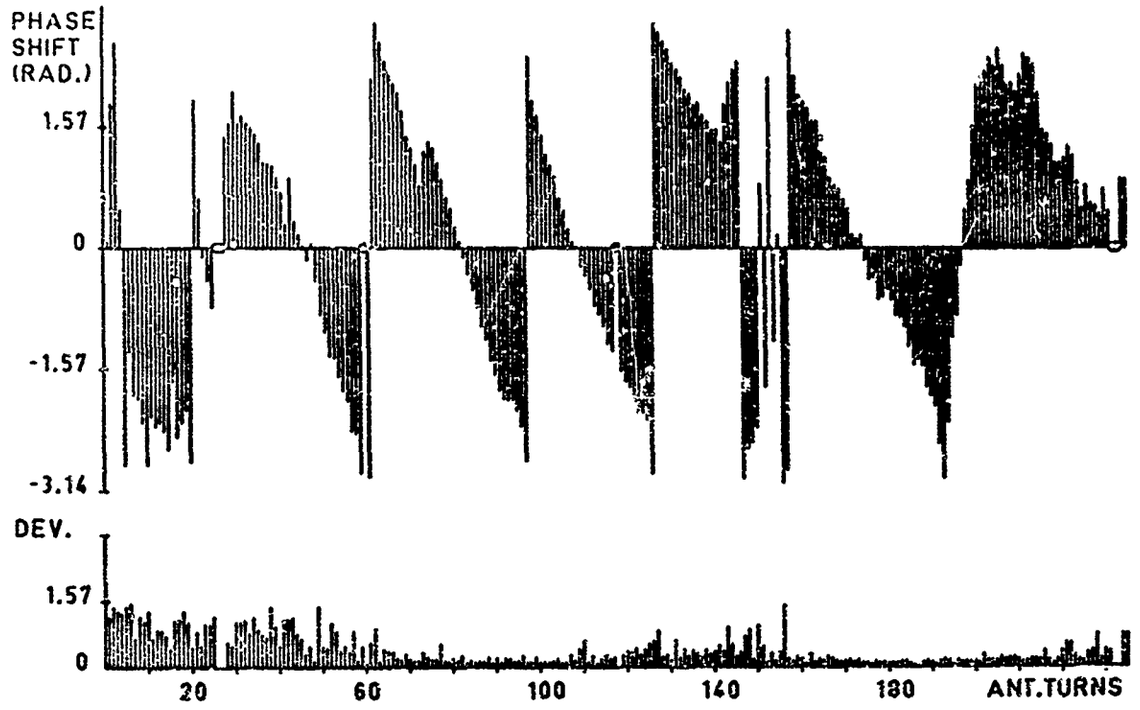


Fig.3c Mean and deviation of the doppler shift along the track

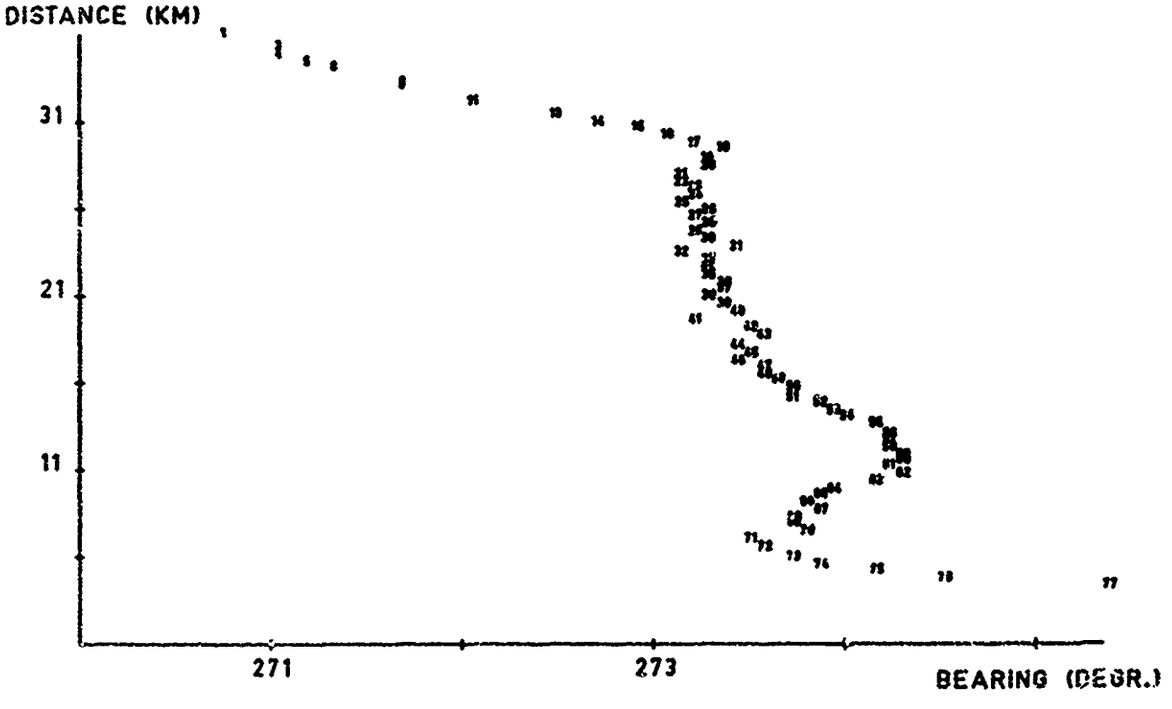


Fig.4a Track of a helicopter L-band-radar

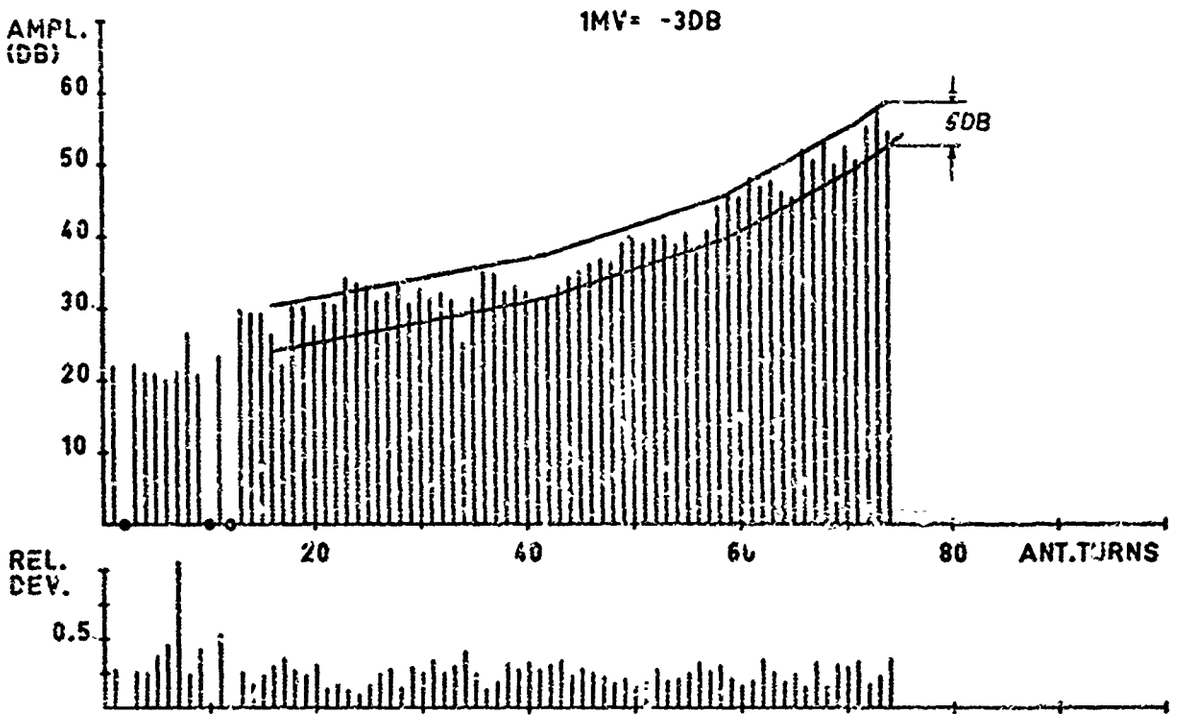


Fig.4b Mean and deviation of target amplitude along the track

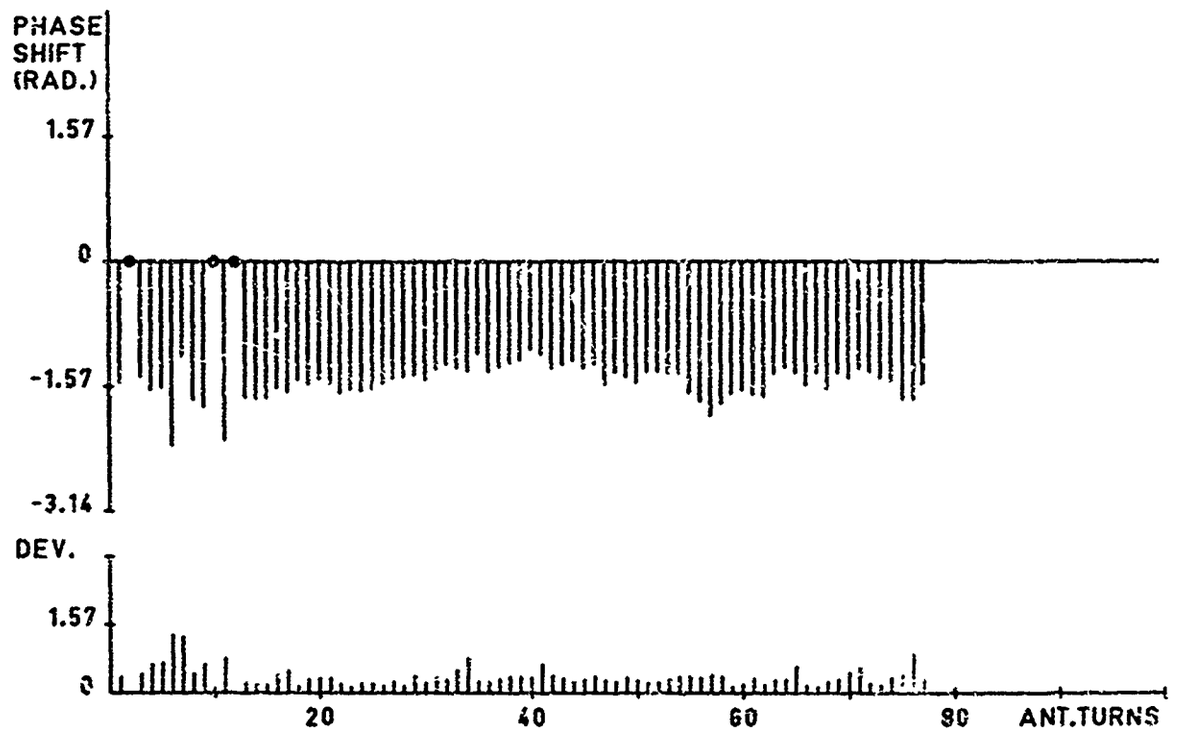


Fig.4c Mean and deviation of doppler shift along the track

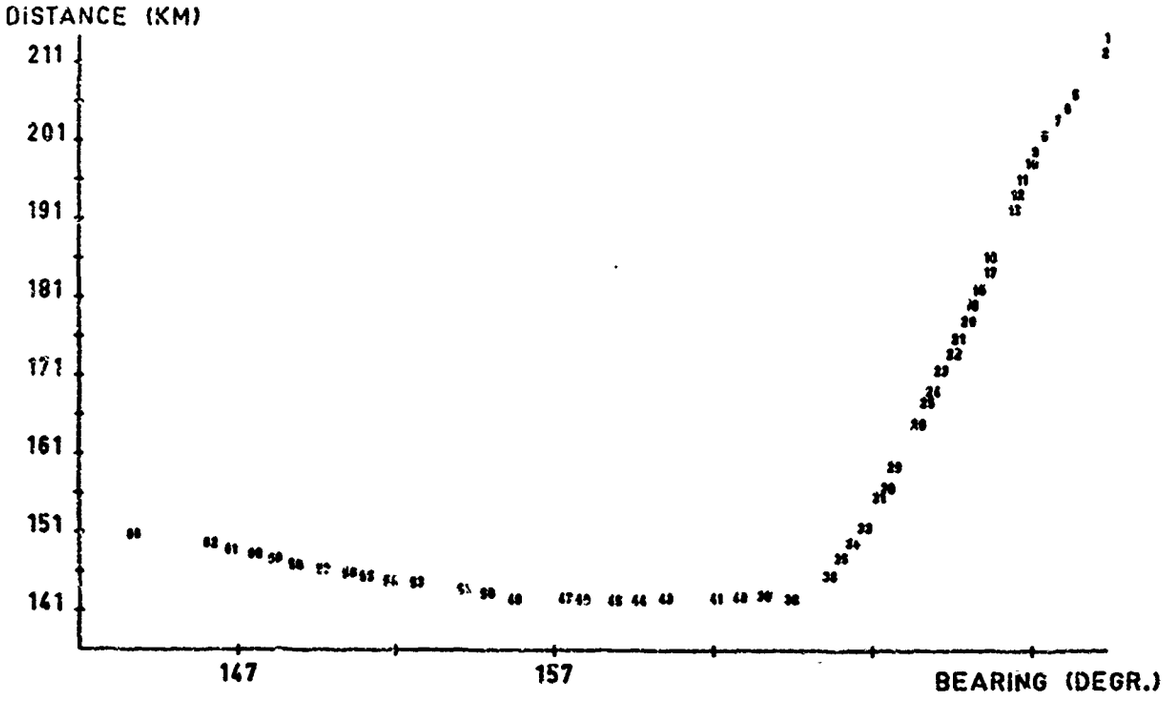


Fig.5a Track of a propeller aircraft S-band-radar

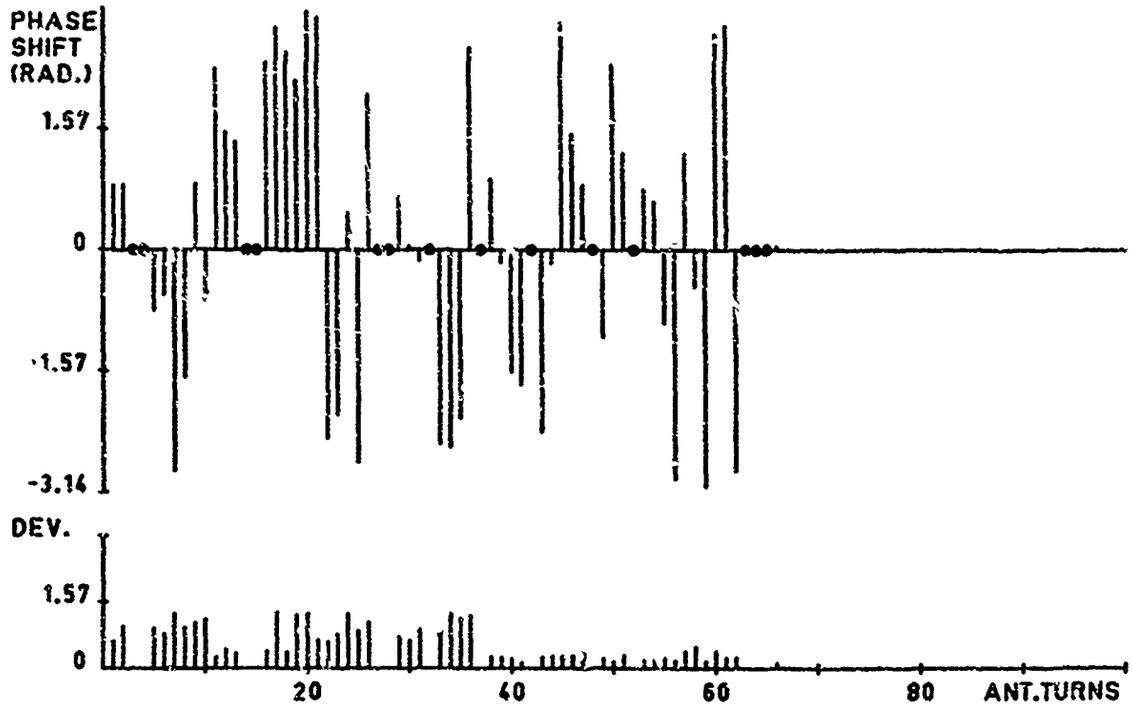


Fig. 5b Mean and deviation of doppler shift along the track

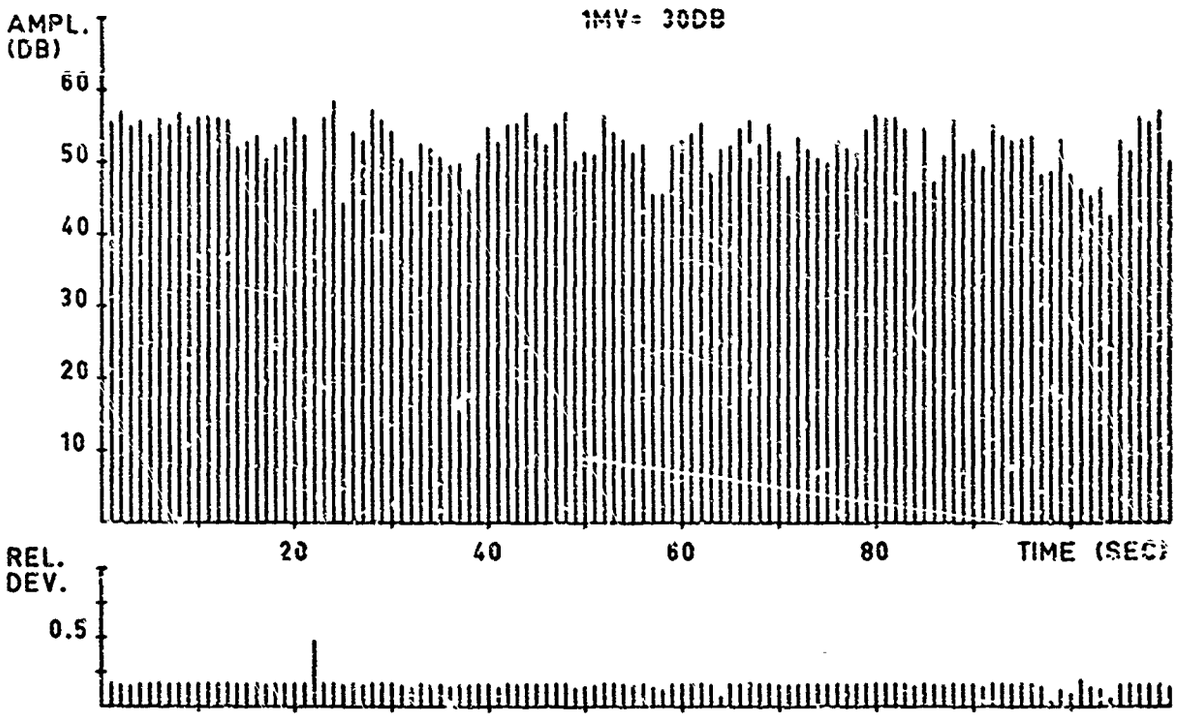


Fig.6a Radial flying jet aircraft scanned with fixed antenna S-band-radar.
Mean and deviation of target amplitude

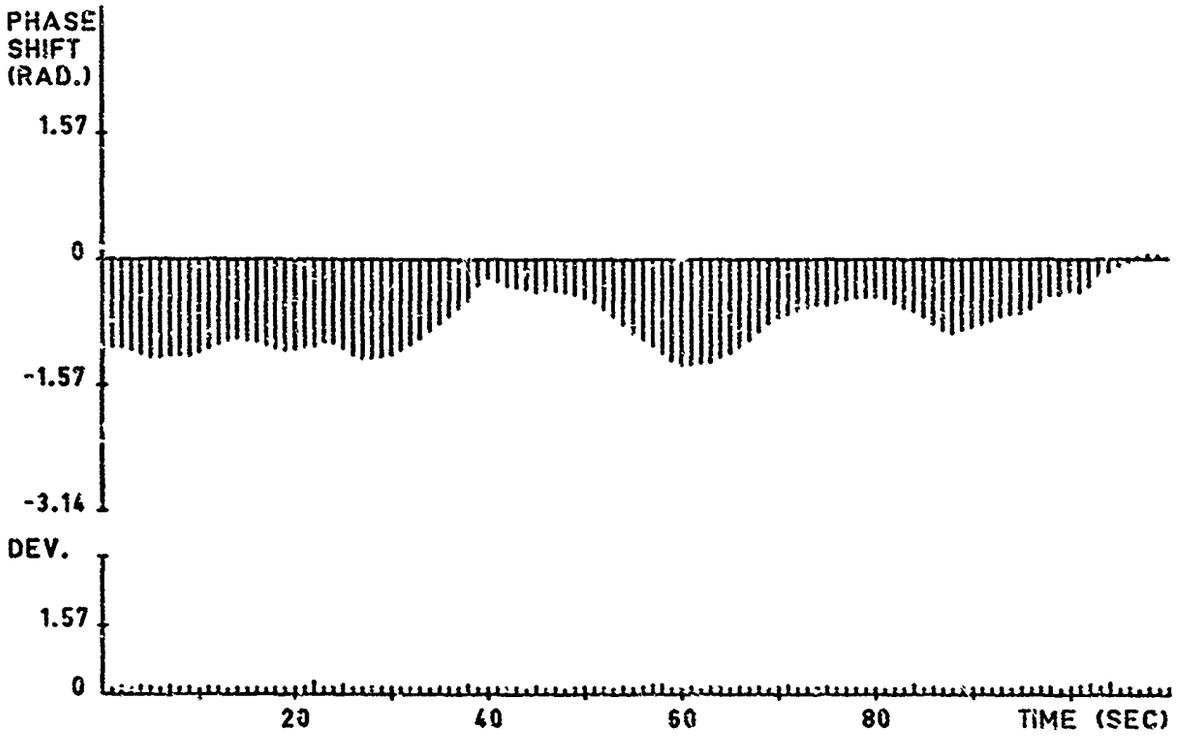


Fig.6b Mean and deviation of doppler shift

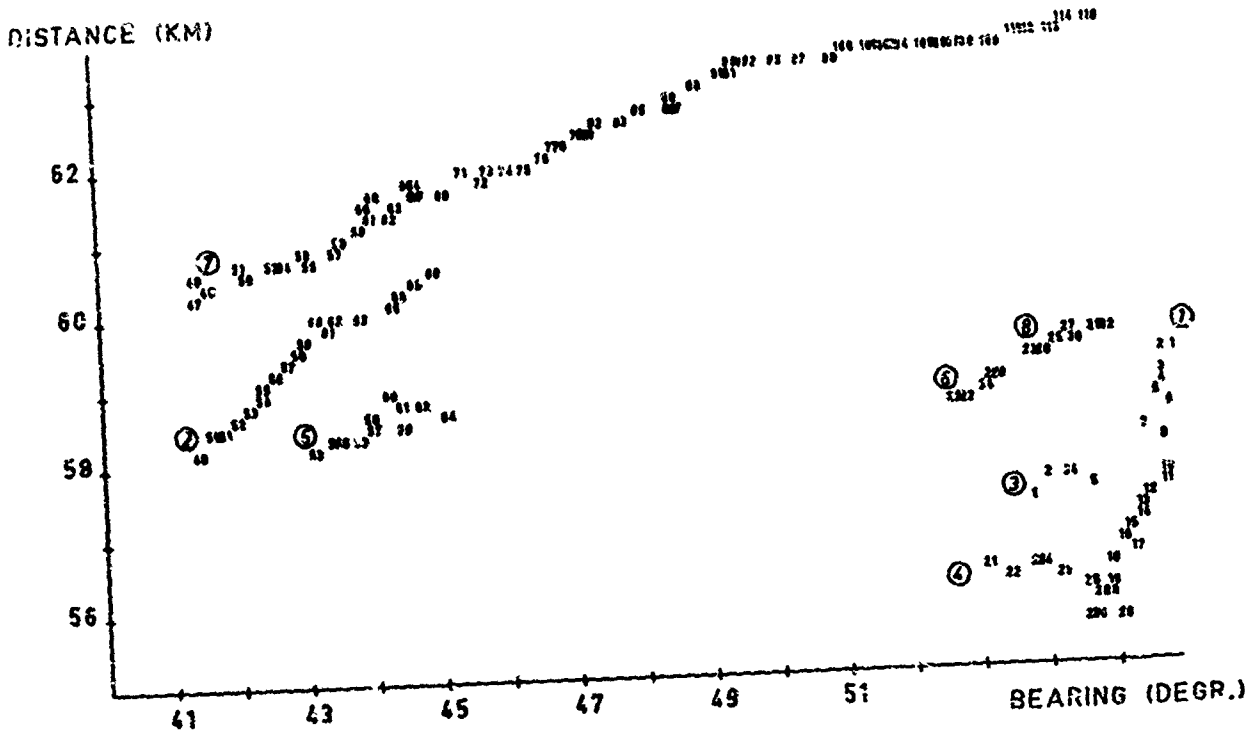


Fig.7a Tracks of birds

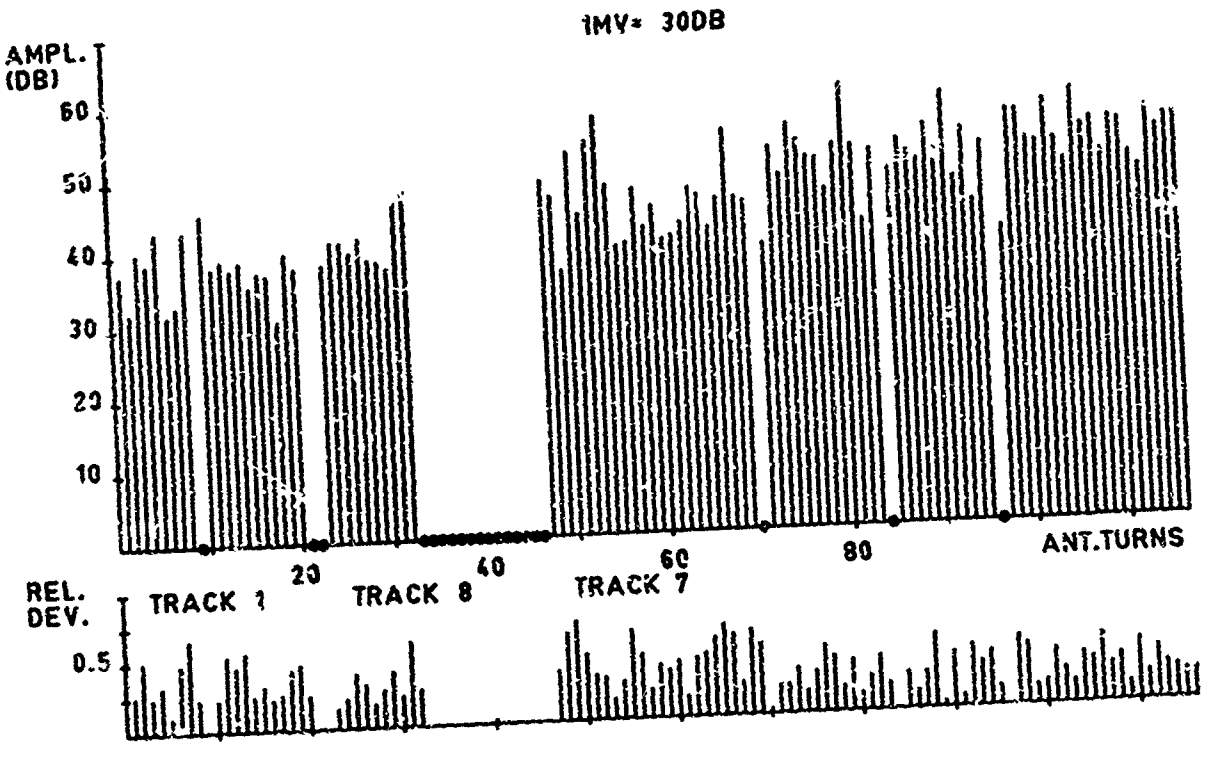


Fig.7b Mean and deviation of target amplitude

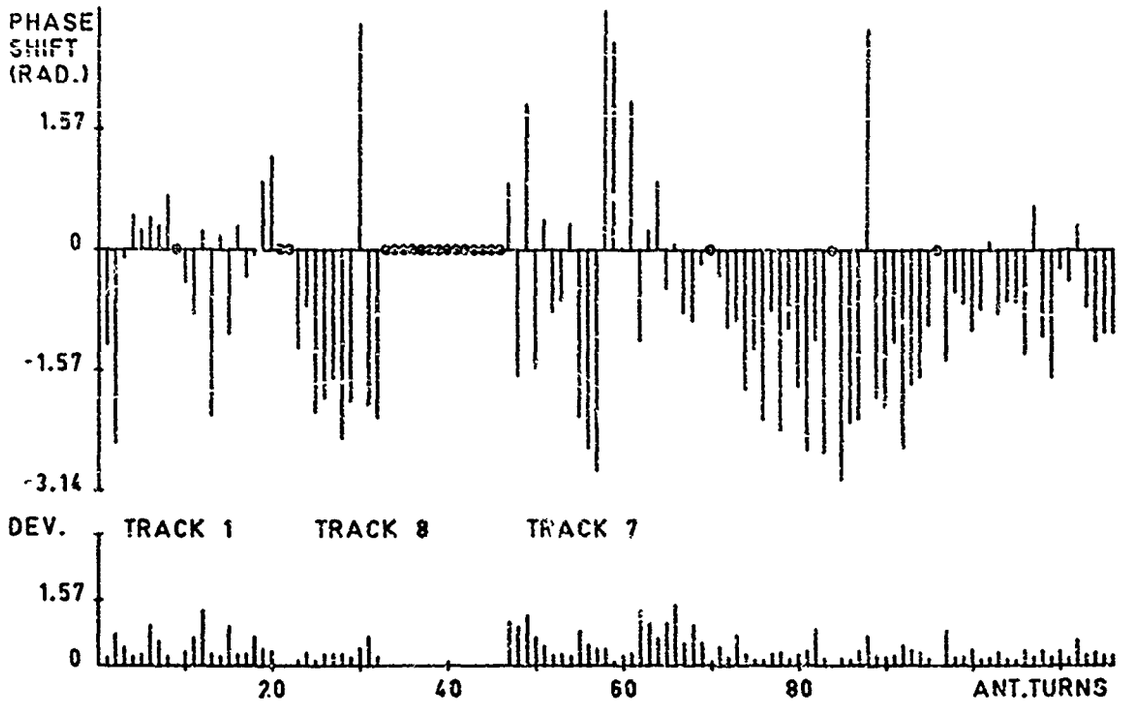
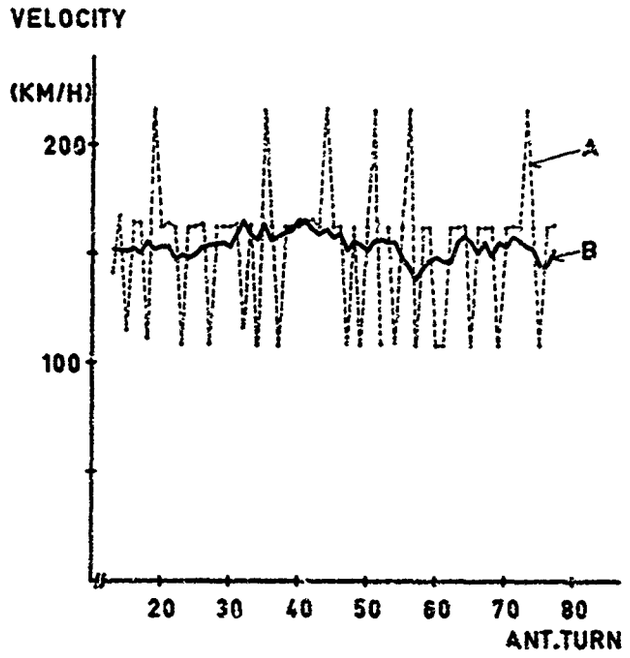


Fig.7c Mean and deviation of doppler shift



VELOCITY OF HELICOPTER (RN: 35003)

A) FROM RADICORD COORDINATES
(RESOLUTION: ± 0.15 KM)

B) FROM DOPPLERSHIFT
(RANGE FROM A))

Fig.8 Velocity of helicopter

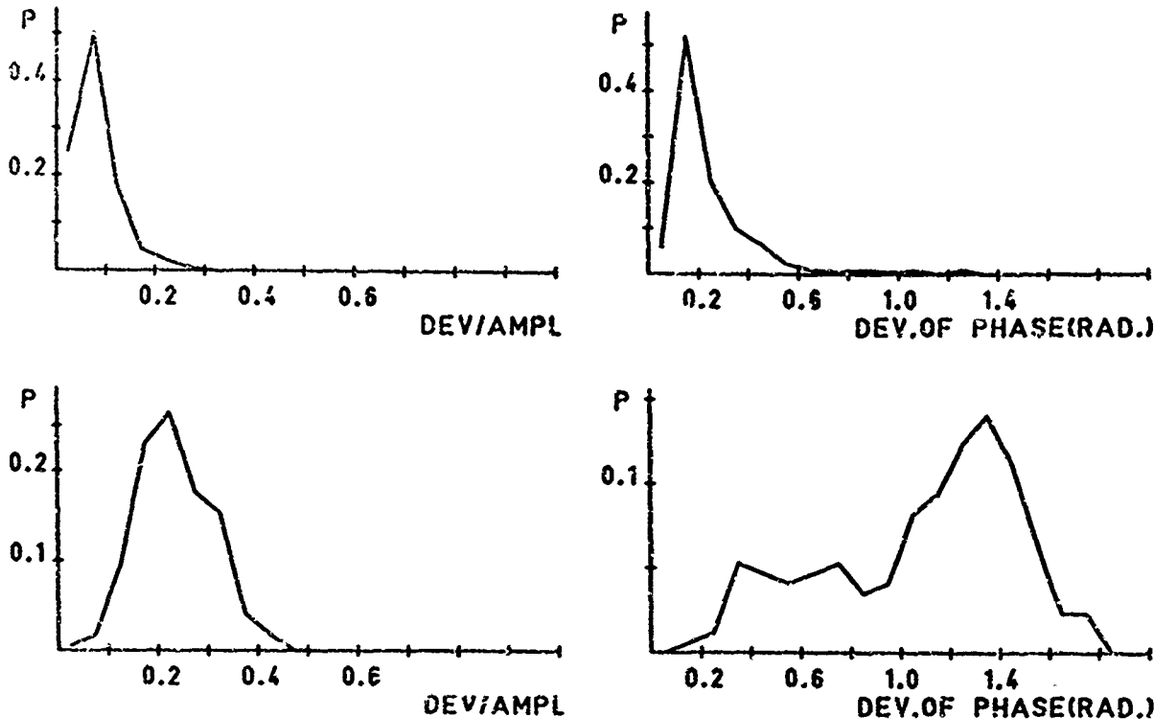


Fig.9 Probability density. Top: jet aircraft, bottom: propeller aircraft

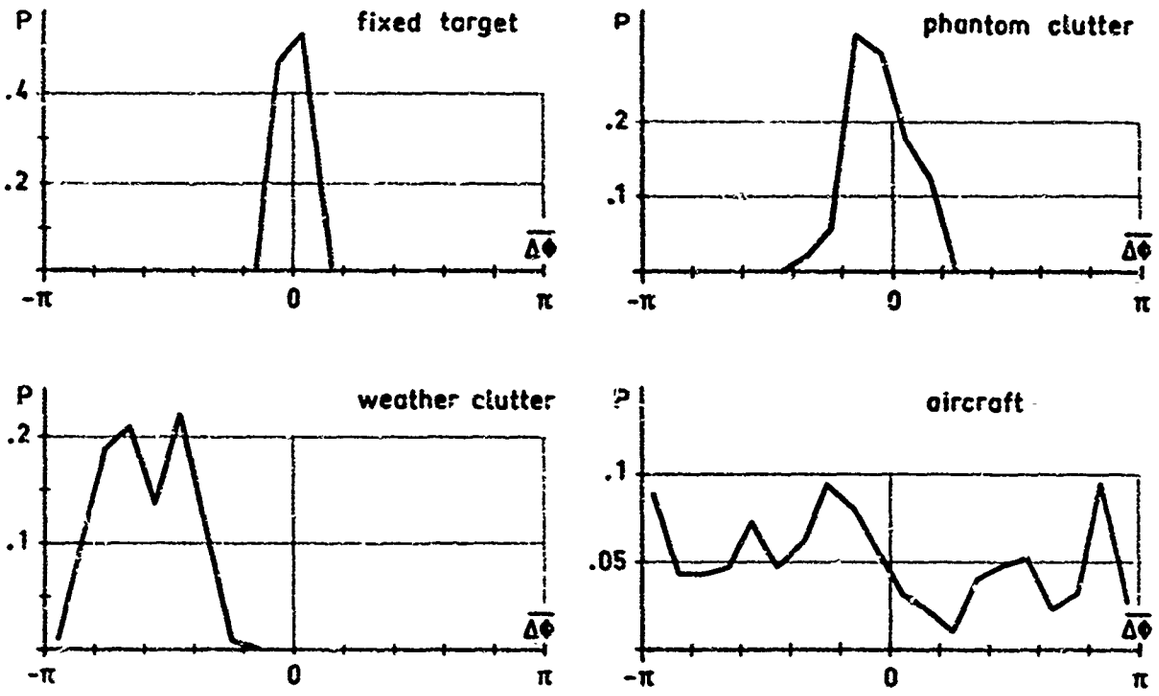


Fig. 10 Probability density P for the mean doppler shift $\overline{\Delta\phi}$ of a scan

SOME MEASUREMENTS OF THE EFFECTS OF FREQUENCY AGILITY ON AIRCRAFT RADAR RETURNS

by

W.S.Whitlock, A.M.Shepherd and A.L.C.Quigley

Ministry of Defence (Navy), A.S.W.E., Portsmouth, England

SOME MEASUREMENTS OF THE EFFECTS OF FREQUENCY AGILITY ON AIRCRAFT RADAR RETURNS

B.S. Whitlock, A.K. Shepherd and A.L.C. Quigley
Ministry of Defence (Navy), A.S.W.E., Portsmouth, England.

SUMMARY

A C-band radar with pulse-to-pulse frequency agility is used to investigate the effects of frequency change on aircraft echo amplitude and scanned echo patterns. Decorrelation effects of pulse-to-pulse frequency changes from zero to 198 MHz have been measured using Sea Hawk and Comet aircraft. The aircraft's radar cross-section changed by as much as 20 dB for a frequency change of 20 MHz. The frequency cross-correlation coefficient is zero for a frequency change of 20 MHz and is negative for larger changes. No significant difference between the frequency correlation coefficient was found for the two aircraft studied. Some effects of frequency agility on scanned echo patterns and echo fading are presented.

1. INTRODUCTION

The radar cross-section of reflecting objects is, in general, dependant on the frequency of the incident radiation, hence a change in radar transmitted frequency may result in a change of echo amplitude. This dependance of radar cross-section on frequency may be exploited in a number of ways; for example it may be advantageous to induce changes in the aircraft echo amplitude by appropriate changes in radar frequency in order to reduce echo fading. As experimental results relating aircraft radar cross-section with frequency change are of particular value to the radar designer, some appropriate measurements have been carried out using a radar with a pulse-to-pulse frequency agile facility. The results of these measurements are detailed in this paper.

2. THE RADAR AND DATA ACQUISITION SYSTEM

The salient parameters of the radar system are as follows:-

Transmitted Frequency	-	5.35 to 5.85 GHz
Transmitted Peak Power	-	1 MW
Pulse Length	-	2 microseconds
P.R.F.	-	196 p/s
Antenna Gain	-	38 dB
Antenna Beam Width	-	1.5° x 1.5° (3 dB, two-way)
Polarisation	-	Horizontal
Receiver	-	Logarithmic Response
I.F.	-	30 MHz
Noise Figure	-	6 dB

A block diagram of the system is shown in Fig. 6.1.

The radar system is fully coherent and the transmitted frequencies are produced by multiplication from a frequency synthesiser operating in the range 84 to 92 MHz. This facility is backed by a special purpose control computer giving a choice of pulse-to-pulse frequency increments, at C-band, between zero and 198 MHz. Random or pattern frequency jumping is possible since all frequencies in the band are defined by any one of 128 different 7-digit binary numbers.

Two modes of pulse-to-pulse frequency agility were used, viz:-

(a) 'Alternate' mode in which only two frequencies were transmitted in sequence f_1, f_2; f_1, f_2; etc., see Figure 6.2. The frequency change, $\Delta f = f_1 - f_2$, could be varied between zero and 198 MHz.

(b) 'Staircase' mode in which successive pulses were transmitted at decreasing frequencies in equal steps for 8, 16 or 32 contiguous pulses before reversing back to the original frequency. As before the frequency change per step could be varied between zero and an upper value which was limited by the number of steps and the 500 MHz system bandwidth.

When carrying out frequency agility experiments it is essential to ensure that the radar performance is not frequency sensitive. For example, it was known that the main t.w.t. output power could vary by $\pm$ 4 dB over the frequency band. In order to reduce this power variation an electronically controlled p-i-n diode attenuator was introduced in the t.w.t. transmitter chain so that the transmitted power levels could be equalised at 80 points in the frequency band. This equalisation is automatic with frequency selection.

It was recognised that a further source of error could arise from the gain/frequency characteristics of the complete receiving channel. This characteristic was found to be cyclic across the 500 MHz band with a gain variation of $\pm$ 2 dB and the technique used to compensate for this gain variation is described below.

The radar data was acquired on-line and transferred to paper tape for subsequent analysis and storage. The video output from a logarithmic receiver was sampled at the range of interest before conversion into a 7-digit binary signal. Up to 10,000 binary video samples could be stored in a computer before transfer to paper tape. In order to compensate for changes in receiver gain the actual signal/noise ratio was measured for each video signal. Firstly, the signal amplitude was measured, then an automatic measurement was made of the mean noise level over a period of 3 ms following the receipt of the signal but before the next frequency change was initiated. The computer was programmed to store the difference between the two logarithmic values.

3. DESIGN OF EXPERIMENTS

3.1 Basic Effects of Frequency Agility

The average value of an aircraft's radar cross-section generally varies little with frequency over the band 1 GHz to 10 GHz, Reference 1. Nevertheless it can be demonstrated that a small change in transmitted frequency can produce a significant change in aircraft cross-section. This apparent anomaly arises because the reflectivity of an aircraft shows fine structure variations with frequency which are normally smoothed in gross measurements. Consideration of a simple model illustrates some of the salient features of this frequency/cross-section relationship.

Let the aircraft be represented by two identical spherical scatterers spaced a distance 'd' apart. Other aircraft models, Reference 2, have been proposed but are not necessarily more realistic. The amplitudes of the two reflected components received by an illuminating radar may be expressed in the form:

$$a_1 = K \sin \left(\omega t + \beta + \frac{2\pi d \cos \theta}{\lambda}\right)$$

$$a_2 = K \sin \left(\omega t + \beta - \frac{2\pi d \cos \theta}{\lambda}\right)$$

Where K is a constant derived from the radar equation, β an arbitrary phase angle and θ the angle between the line joining the reflectors and the radar line of sight, i.e. the aspect angle. The resultant received power $P_{(r)}$ is proportional to:

$$\left[\cos \left(\frac{2\pi d \cos \theta}{\lambda}\right)\right]^2$$

or expressed explicitly in terms of frequency:

$$P_{(r)} = \left[\cos \left(\frac{2\pi df \cos \theta}{c}\right)\right]^2 \qquad (1)$$

It will be noted that $P_{(r)}$ varies periodically with θ. This produces the well-known scintillation characteristics of aircraft echoes when changes in θ of 0.1° may change the aircraft cross-section by 20 dB or more. Further, the received power varies periodically with frequency; in fact, for a given aspect angle the aircraft cross-section varies as the product of the separation of the scatterers and the frequency. Using this simple model it can be shown that for a small change in frequency the maximum cross-section change occurs when $\Delta f \approx 10$ MHz.

This simple model has been used to highlight two important effects:

(a) The magnitude of the change in aircraft cross-section due to a change in frequency is aspect dependent.

(b) The maximum change in cross-section should occur when the change in frequency is about 10 MHz.

Initial experimental results did indicate, in fact, that for a fixed change of frequency the cross-section change of a particular aircraft was not reproducible unless the aspect remained virtually constant. In the design of the radar experiments it was necessary to decide on the parameters which would express the quantitative effects of frequency agility in a form which was meaningful and had value to system designers. The choice of such parameters is discussed in the next section.

3.2 Pulse-to-Pulse Amplitude Changes

Given time, aircraft echoes will become decorrelated even at fixed frequency due to random aspect changes. Alternatively the aircraft echoes may be decorrelated by rapid frequency changes. In practice time and frequency decorrelation may both operate on the received signal. These effects may be expressed quantitatively through the use of correlation coefficients, but without computational assistance the derivation of these coefficients is laborious. In the early stages of the experimental studies a computer facility was not available so some alternative parameter was sought which could be derived simply and which expressed quantitatively the decorrelating effects of frequency agility.

Consider a transmitter operating in the alternate frequency agile mode, then two interlaced signal pulse trains will be received, each train corresponding to one of the two transmitted frequencies. Let these pulse trains be identified as p and q, where $p = q = 1$ to n, where n is the total number of pulses in each train. Let Δa be the change in amplitude between successive pulses, then the mean square of Δa is given by:

$$\overline{\Delta a^2} = \overline{(a_p - a_q)^2}$$

where a_p and a_q are the amplitude deviations from the mean. Hence

$$\overline{\Delta a^2} = \overline{a_p^2} + \overline{a_q^2} - \overline{2a_p a_q} = 2\left[\overline{a^2} - \overline{a_p a_q}\right] \text{ if } \overline{a^2} = \overline{a_p^2} = \overline{a_q^2}$$

$$= 2\left[\overline{a^2} - \rho \overline{a^2}\right]$$

where ρ is the cross-correlation coefficient. Therefore,

$$\frac{\overline{\Delta a^2}}{\overline{a^2}} = 2\left[1 - \rho\right] \qquad (2)$$

so that the mean square pulse-to-pulse amplitude difference, which is simple to compute, is related to the cross-correlation coefficient through Equation (2) so long as the basic assumptions are valid. The correlation coefficient may be, of course, both time and frequency dependent.

It is of interest to note that when pulse samples are taken from a source of Gaussian noise and uncorrelated then the r.m.s. pulse-to-pulse difference after processing through a narrow-band logarithmic receiver, can be theoretically derived, Appendix I, and is given by:

$$\left(\overline{\Delta a^2}\right)^{\frac{1}{2}} = \frac{10\kappa}{\sqrt{3}\log_e 10} = 7.88 \text{ dB}$$

The standard deviation of pulses which have a log-Rayleigh distribution of amplitude has also been derived, Reference 3, viz:

$$\left(\overline{a^2}\right)^{\frac{1}{2}} = \frac{10\kappa}{\sqrt{6}\log_e 10} = 5.57 \text{ dB}$$

Therefore for pulses originating from Gaussian noise and uncorrelated then $(\overline{\Delta a^2})/(\overline{a^2}) = 2$, which agrees with Equation 2 as $\rho = 0$.

The object of this analysis is to show that $\overline{\Delta a^2}$ is a simply computed parameter which may be used as a measure of decorrelation of a pulse train, even when only two discrete frequencies are transmitted. Where the pulse train is noise-like and decorrelated from pulse-to-pulse and the train is processed by an ideal logarithmic receiver then $(\overline{\Delta a^2})^{\frac{1}{2}} = 7.88$ dB. It is important to note that few logarithmic receivers have ideal characteristics so far as their response to noise-like signals is concerned, Reference 4.

3.3 Methods of Aircraft Echo Measurement

When frequency agility data from aircraft was required it was usual to operate the radar in a 'search-light' mode with manual range gating and tracking. The antenna was elevated to 1.5° and the aircraft were directed to fly up or down the stationary beam from a maximum range of 30 miles to a minimum range of 30 miles. Continuous data runs of over five minutes at a nominally head-on aspect were possible although in practice each run was sub-divided into 40 second (2048 pulses) periods, each period corresponding to a particular value of Δf.

When scanned echo patterns were required then both manual range gating and azimuth gating were applied simultaneously with the antenna rotating at 6 rev/min. When the gates were following correctly a computer was switched to acquire data and the scanned echo pattern was passed into the store for subsequent print-out by a line printer.

4. RESULTS

4.1 Initial Experiments

Early work using frequency agility was directed at land clutter and Figure 6.3 shows values of $\sqrt{\overline{\Delta a^2}}$ obtained when the resolution cell contained agricultural land. Figure 6.3 shows that $\overline{\Delta a^2}$ reaches a maximum when $\Delta f \geq 1/\text{pulse length}$. Using the same data, it was found that the auto-correlation coefficient for one pulse lag reached a minimum when $\Delta f = 1/\tau$.

Initial aircraft experiments using unidentified aircraft targets of opportunity yielded results similar to those also shown in Figure 6.3. For all aircraft trials the alternate frequency jumping mode was used. It was found that the curve of $\sqrt{\overline{\Delta a^2}}$ against Δf was very variable from run to run even for the same aircraft, presumably due to random changes in aspect. It became apparent that, for consistent data, it would be necessary to make observations against controlled aircraft whose aspects would be better defined.

4.2 Sea Hawk and Comet Measurements

Using aircraft controlled from A.S.W.E., two experimental runs were carried out; one using a Sea Hawk GA6 and the other a Comet 4. The object of choosing these very different aircraft types was to try to resolve any dependence of frequency agility response on aircraft size. The Sea Hawk wing-span is 12m whilst that of the Comet is 35m. Figure 6.4 shows the values of $\sqrt{\overline{\Delta a^2}}$ plotted against the frequency separation, Δf, between transmitted pulses for the Sea Hawk in both opening and closing flight. Each point on the graph corresponds to 2048 pulses, i.e. about 10s of record. Maximum decorrelation appears to occur at about 13 MHz. Amplitude probability density functions were computed about a zero mean and Figure 6.5 shows a case of a bi-modal distribution due to a shift in mean cross-section of 6 dB as the frequency is changed by 128 MHz.

The Comet experiments produced results very different from those obtained using the Sea Hawk for the Comet returns were found to be largely decorrelated even at fixed frequency. Rain clutter (15 dB above noise) corrupted part of the recordings although the Comet return was generally 30 dB above noise. However, it is believed that the rapid modulation on the Comet return, corresponding to a head-on aspect, could be due to engine turbine-blade motion. Figures 6.6 and 6.7 illustrate the effect of this intrinsic decorrelation process which tends to hide the effects due to frequency agility.

A penalty of using a relatively low P.R.F. (196 p/s) for frequency agile experiments is that there is generally a significant degree of time decorrelation superimposed on any decorrelation due to frequency change. This problem is clearly demonstrated in Figure 6.8 which shows values of the auto-correlation coefficient derived from the Comet data using an alternate frequency jumping mode. The value of coefficient falls rapidly for one pulse lag due to the time-decorrelating process intrinsic to the aircraft, then the coefficient alternates between two values due to changes in frequency. The upper dotted line in Figure 6.8 corresponds to the auto-correlation curve for the fixed frequency case (every other pulse is at the same frequency). This upper curve, therefore, reflects only time decorrelation, $\rho(\tau)$. However, the lower curve traces values of the composite correlation coefficient, $\rho(\tau, f)$. In an attempt to isolate $\rho(f)$ it is assumed that:

$$\rho(\tau, f) = \rho(\tau) \cdot \rho(f) \qquad (3)$$

Having made this assumption, then using all the aircraft data tapes, the auto-correlation coefficients corresponding to all pulse lags between 1 and 10 were computed for the various values of Δf. Using this reduced data the mean values of $\rho(\tau, f)$ and $\rho(\tau)$ were calculated and hence the values of $\rho(f)$ using Equation 3. The results of this approach are plotted in Figure 6.9 and show no appreciable difference between the Comet and Sea Hawk so far as their frequency dependence is concerned despite their different fuselage lengths. This analysis also shows that $\rho(f) = 0$ for $\Delta f \approx 20$ MHz and that the correlation coefficient in the frequency domain is negative for larger frequency changes. This is in accord with the concept that an aircraft consists of a few dominant scatterers.

4.3 Scanned Aircraft Echo Amplitude Patterns

Further experiments were performed in order to simulate aircraft echo amplitude patterns as would be obtained when using a frequency scanned antenna. In fact a mechanically azimuth-scanned antenna was used. The data acquisition technique is described in Section 3.3 for targets which were unidentified aircraft. Using fixed frequency, the ideal signal pattern should only reflect the two-way antenna radiation pattern, as shown in Figure 6.10(a) which is a facsimile of a computer print-out. However, Figure 6.10(a) depicts a rare occasion for it is more probable that a fade will coincide with the time the beam is incident on the aircraft, as is shown in Figure 6.10(b). When alternate frequency jumping is used the echo patterns show considerable modification. For the no-fading case, Figure 6.11(a), the two patterns corresponding to the two frequencies can be resolved clearly. When fading is concurrent with scanning the effect of the two frequencies is clearly shown in Figure 6.11(b) where the long fades have been eliminated. From pulse-to-pulse it can be seen that a change in frequency of 64 MHz can change the amplitude of the return by as much as 25 dB. Other records show numerous examples of large changes in amplitude for relatively small changes in frequency. Where a staircase form of frequency jumping was used it was found that the shape of the echo pattern was distorted by the incremental frequency changes so that if beam shape is used as a detection criteria the effect of frequency agility may degrade detection performance.

5. CONCLUSIONS

Experimental investigations have confirmed that the magnitude of an aircraft's radar cross-section is a sensitive function of transmitted frequency. Cross-section changes greater than 20 dB have been found to occur for a pulse-to-pulse frequency change as small as 20 MHz whereas the average change in cross-section was found to be about 6 dB for a 20 MHz change in frequency. For aircraft, a complex periodic cross-section/frequency relationship exists which is analogous to the cross-section/aspect relationship; therefore a change in frequency may result in the decorrelation of the aircraft return, as would a change in aspect. Frequency correlation coefficients (subtracted mean) have been calculated from data corresponding to nominally head-on aspects of a Sea Hawk and Comet aircraft. It is found that the coefficient is zero for a pulse-to-pulse frequency change of about 20 MHz but for frequency changes between 30-200 MHz the coefficients are negative. Some theories suggest that the frequency change necessary for maximum decorrelation should be inversely proportional to aircraft size, or scatterer separation, but no significant difference was found between the frequency changes necessary to decorrelate the returns from a Sea Hawk (wing span 12m) and a Comet 4 (wing span 35m) aircraft.

The main advantage of frequency agility as an aid to aircraft detection derives from the considerable reduction in the occurrence of long fades of the target echo. This advantage may be obtained using only two frequencies separated by 20 MHz; then, on average, successive returns will be decorrelated and fade duration will be confined to one pulse interval. However, in many cases, for example, where the natural fading rate is high and the signal/noise ratio large, the use of frequency agility will not necessarily improve the probability of detection.

6. REFERENCES

(1) Olin, I.D., and F.D. Queen; "Dynamic Measurement of Radar Cross Sections", Proc. I.E.E.E. 53(8), 954-961, August 1965.

(2) May, M.; "Improving Radar Range and Angle Detection with Frequency Agility", Microwave Journal, 9, No. 5, May 1966 p.63.

(3) Smith, M.E.A.; "The Rayleigh and Log-Rayleigh Probability Density Functions", A.S.W.E. Laboratory Note XRE-67-8, August 1967.

(4) Reiss, A., and G.W. Carver; "Response of Logarithmic Amplifiers to Noise Signals", Electronic Letters, 4, No. 13, June 1968.

(5) Bromwich, T.J.I'A; "An Introduction to the Theory of Infinite Series", Macmillan, London, 1947, 2nd Edition, p. 498.

(6) Dwight, H.B.; "Tables of Integrals and other Mathematical Data", Macmillan, N.Y. 1961, 4th Edition, p. 13.

APPENDIX I

PULSE-TO-PULSE DIFFERENCES OF SAMPLED NOISE AT THE OUTPUT OF A LOGARITHMIC RECEIVER

M.H.A. Smith, A.S.W.E.

1. DERIVATION OF THE PULSE-TO-PULSE DIFFERENCE PROBABILITY DENSITY FUNCTION (P.D.F.)

Let the amplitude of the logarithmic receiver output sample be 'y' dB (with respect to a datum which will be defined later) and let the probability density for this amplitude be f(y). By definition, therefore, the probability that the output of the receiver will lie within the range 'y' to 'y + δy' is f(y).δy. Similarly, the probability that the output will lie within the range 'y + D' to 'y + D + δy' is f(y + D).δy.

It is assumed that successive samples of the receiver output are uncorrelated. This assumption is true for a noise waveform if the sampling rate is much less than the receiver bandwidth. The probability after a pulse has occurred in the amplitude range 'y' to 'y + δy' that the next pulse will occur in the range 'y + D' to 'y + D + δy' is now simply the product of the probabilities of the two individual occurrences, i.e. f(y).δy.f(y + D).δy.

If the amplitude difference between two successive pulses is D dB, then the first pulse can take on any value whatsoever in the permissible range of 'y', provided only that the second pulse, which is specified to be the larger in amplitude by D dB, also lies within the permissible range of 'y'. Let the p.d.f. of 'D' be denoted by p(D), then the probability that the difference between successive pulses lies between 'D' and 'D + δy' is

$$p(D).\delta y = \sum_y f(y).\delta y.f(y + D).\delta y$$

where the summation is to be performed over all permissible values of 'y'. Since δy can be considered as a factor this equation can be rewritten.

$$p(D) = \sum_y f(y).f(y + D).\delta y$$

or, in the limit,

$$p(D) = \int_y f(y).f(y + D).dy \qquad \ldots\ldots (1)$$

Now Reference 3 shows that the p.d.f. of the noise output from a logarithmic receiver can be written

$$f(y) = K.z.\exp(-z), \qquad \ldots\ldots (2)$$

where $z = 10^{(y/10)}$,

and $K = \ln(10)/10$,

giving $y = 10 \log z = \ln(z)/K$.

The datum from which 'y' is measured is 2.50 dB above the mean logarithmic noise level and is in fact the modal value of the logarithmic noise output p.d.f.

For an ideal logarithmic receiver 'y' can take on any value from $-\infty$ to $+\infty$.

Thus

$$p(D) = \int_{-\infty}^{\infty} f(y).f(y + D).dy \qquad \ldots\ldots (1\text{ bis})$$

$$= \int_0^{\infty} (K\,z\exp(-z))(K\,z^2\exp(-z^2))(1/Kz).dz$$

where $z^2 = 10^{(y + D)/10} = 10^{(y/10)}.10^{(D/10)}$

If we now denote the linear ratio of the two successive pulse amplitudes as 'r', then

$$r = 10^{(D/10)}$$

$$z^2 = r\,z$$

and $p(D) = K\,r\int_0^{\infty} z.\exp(-z(1 + r)).dz$

therefore $p(D) = K\,r/(1 + r)^2$, $\qquad \ldots\ldots (3)$

where $r = 10^{(D/10)}$

and $K = \ln(10)/10$

The mode of this p.d.f. occurs for r = 1 or D = 0 dB, that is, the most probable pulse-to-pulse difference is zero, and the modal value is seen to be K/4 (i.e. 0.05756). If the most probable pulse-to-pulse differences were not zero then there would be an upward or downward trend in the mean noise level, and it is tacitly assumed that this is not so.

The total area under the p.d.f. is

$$\int_{-\infty}^{\infty} p(D).dD = \int_{0}^{\infty} (r/K).p(r).dr$$

$$= \int_{0}^{\infty} 1/(1 + r)^2 \, dr$$

$$= 1$$

This value is to be expected as it includes all possible values of 'D'.

2. STANDARD DEVIATION OF THE PULSE-TO-PULSE DIFFERENCE P.D.F.

The p.d.f. of the pulse-to-pulse difference, D, is given by

$$p(D) = Kr/(1 + r)^2 \qquad \ldots\ldots (3, bis)$$

where $r = 10^{(D/10)}$

If the standard deviation of the p.d.f. is denoted by σ then

$$\sigma^2 = \int_{-\infty}^{\infty} p(D).D^2.dD$$

$$= \int_{-\infty}^{0} p(D).D^2.dD, \text{ from considerations of symmetry}$$

Now $D = \ln(r)/K$

and $dD = (1/Kr)dr$

so that $\sigma^2 = (2/K^2) \int_{0}^{1} (\ln r)^2/(1 + r)^2 dr$

$$= (2/K^2)I \qquad \ldots\ldots (4)$$

where I denotes the definite integral.

This integral is not easily evaluated since the logarithmic term becomes infinite at the lower limit and the power series expansion for the denominator is not valid at the upper limit. However, in Reference 5 if we substitute $(\ln r)^2$ for $\ln x$ and set the parameter 'p' to zero we obtain

$$I = 2 \sum_{n=0}^{\infty} (-1)^n/(n + 1)^3$$

Therefore $I/2 = 1/1^3 - 1/2^3 + 1/3^3 - \ldots$

$$= \pi^3/12 \text{ by Reference 6}$$

Substituting back into Equation (5)

$$\sigma^2 = \pi^2/(3K^2)$$

i.e. $\sigma = \pi/(\sqrt{3}K)$

$$= 10\pi/(\sqrt{3}.\ln 10) \qquad \ldots\ldots (5)$$

$$= \underline{7.877 \text{ dB}}$$

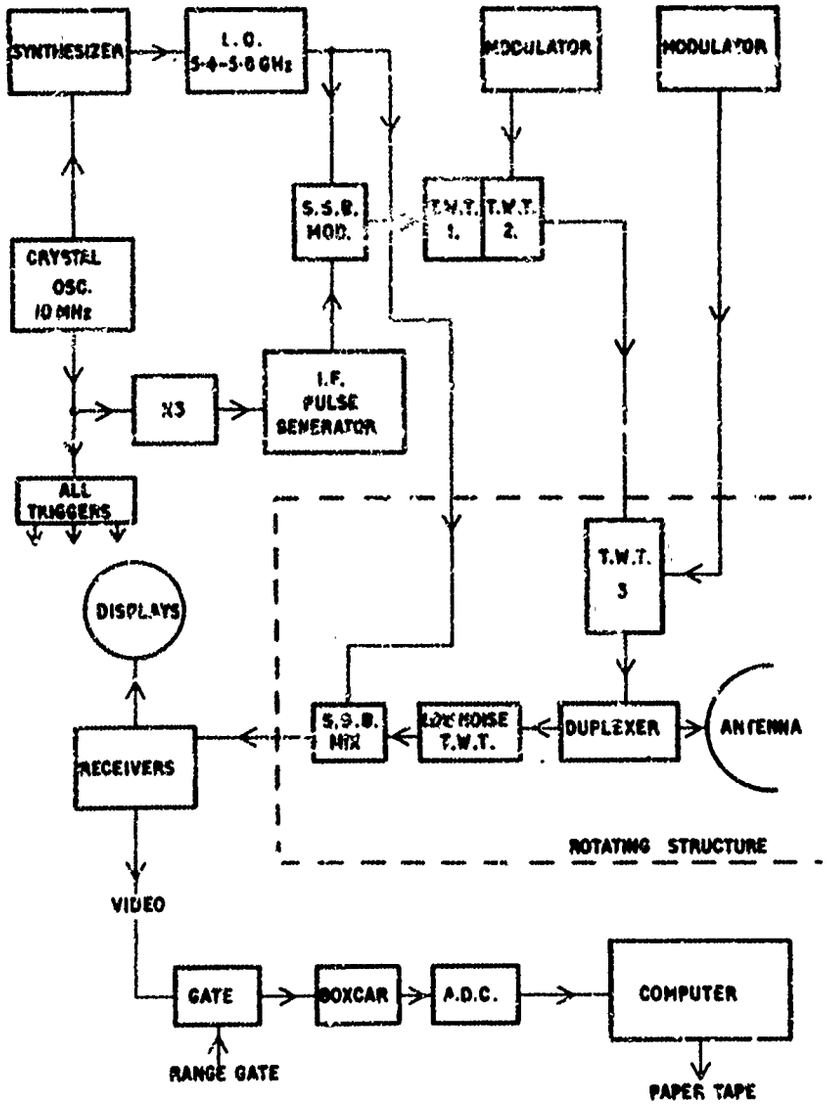


Figure 1 Radar System Block Diagram

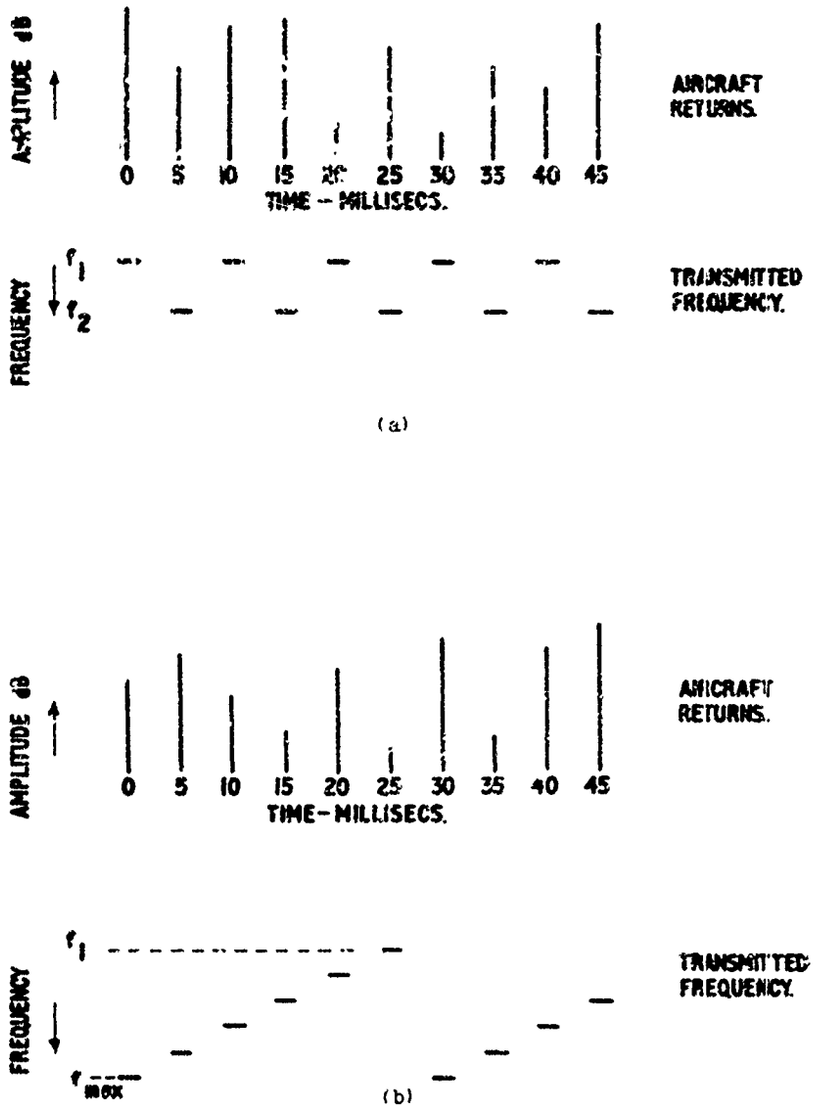


Figure 2 Frequency Agile Modes (a) Alternate Mode; (b) Staircase Mode

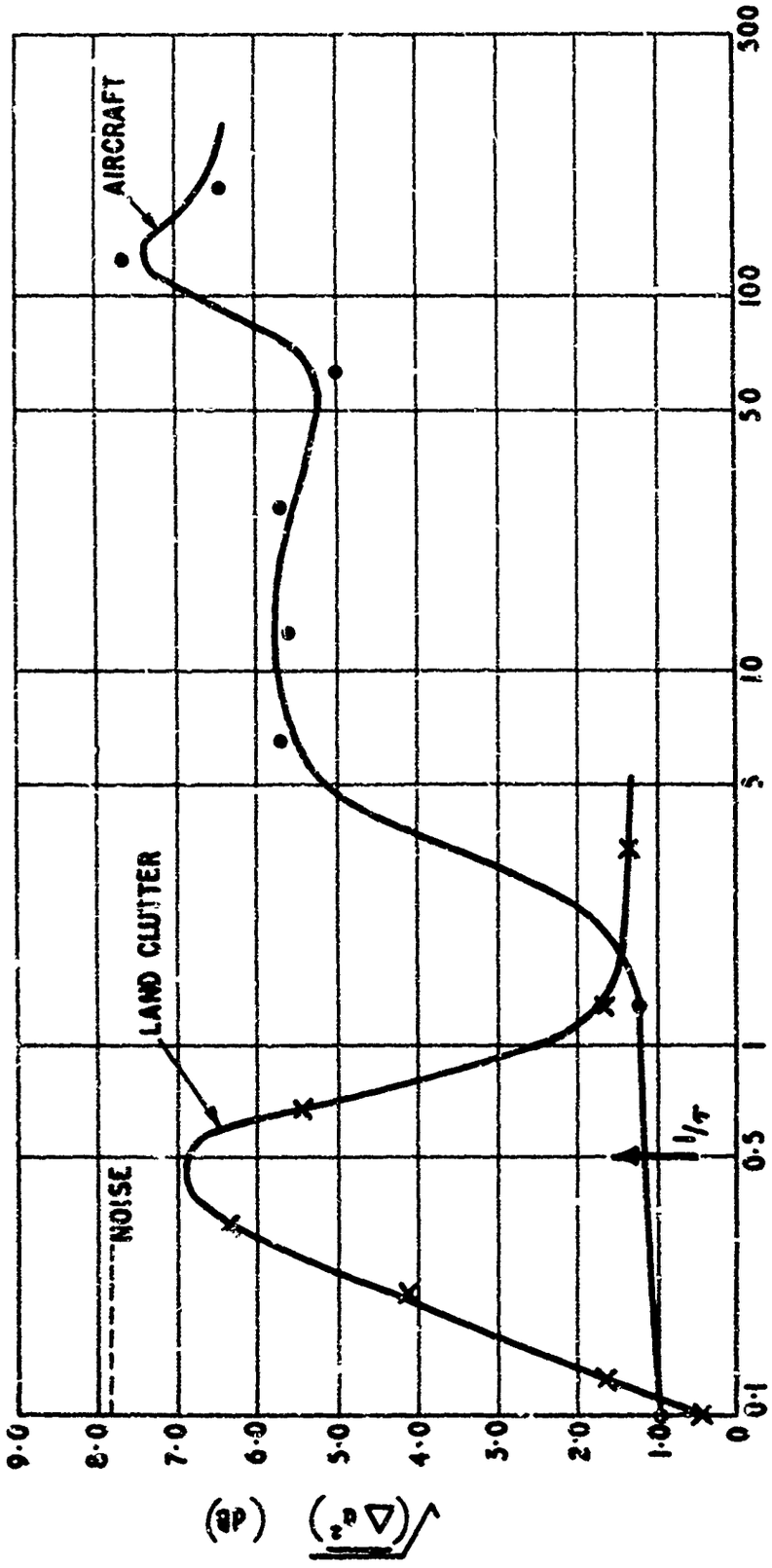


Figure 3 R.M.S. Pulse-to-Pulse Amplitude Change versus Frequency Change for Land Clutter and Aircraft

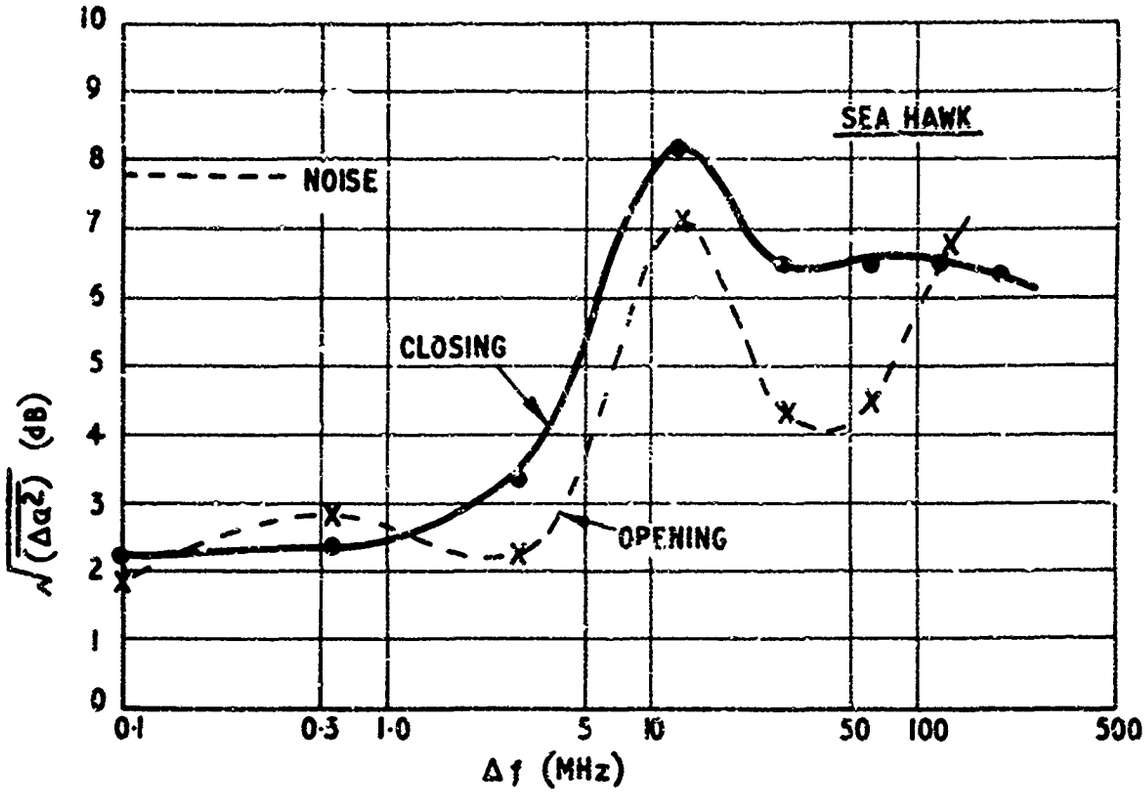


Figure 4 R.M.S. Pulse-to-Pulse Amplitude Change for a Sea Hawk Aircraft

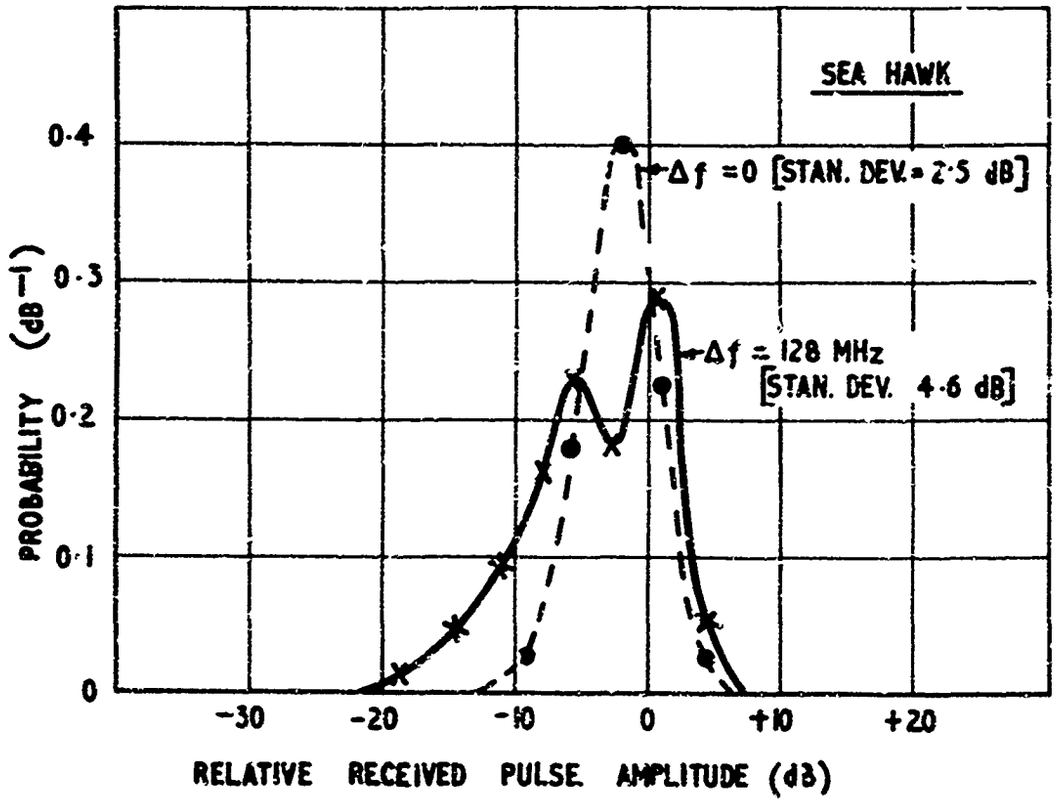


Figure 5 Echo Amplitude Probability Density Function for a Sea Hawk Aircraft

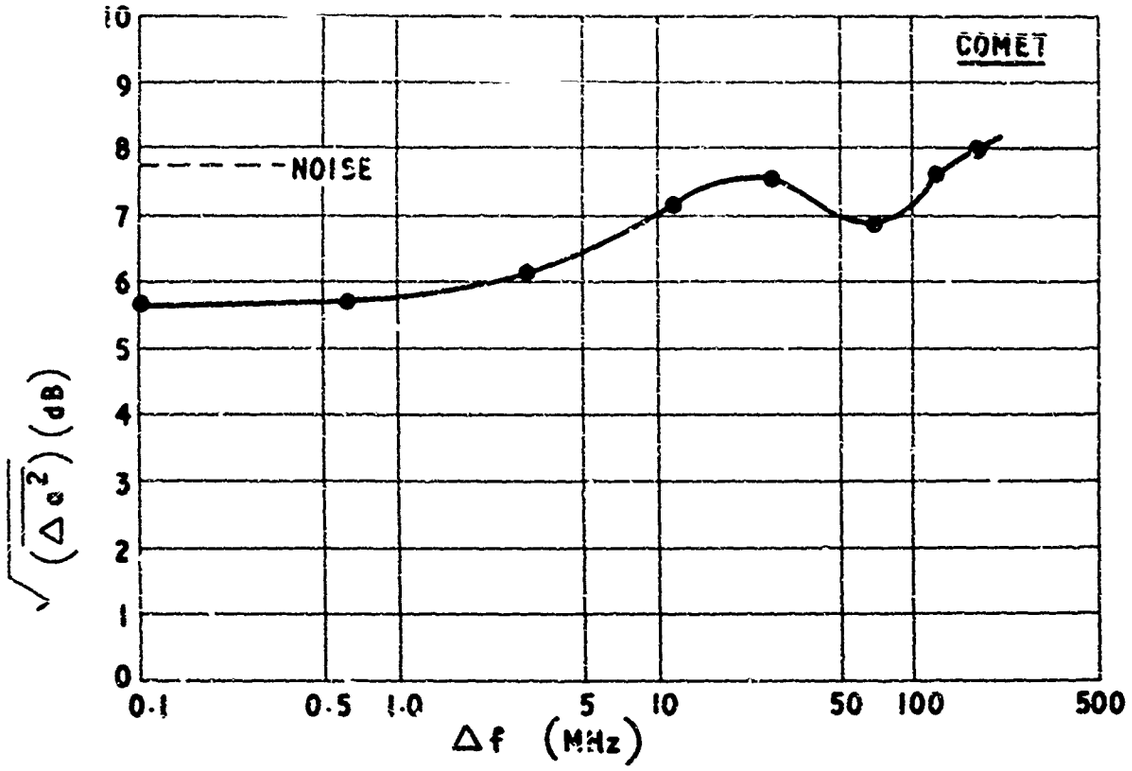


Figure 6 R.M.S. Pulse-to-Pulse Amplitude Change for a Comet Aircraft

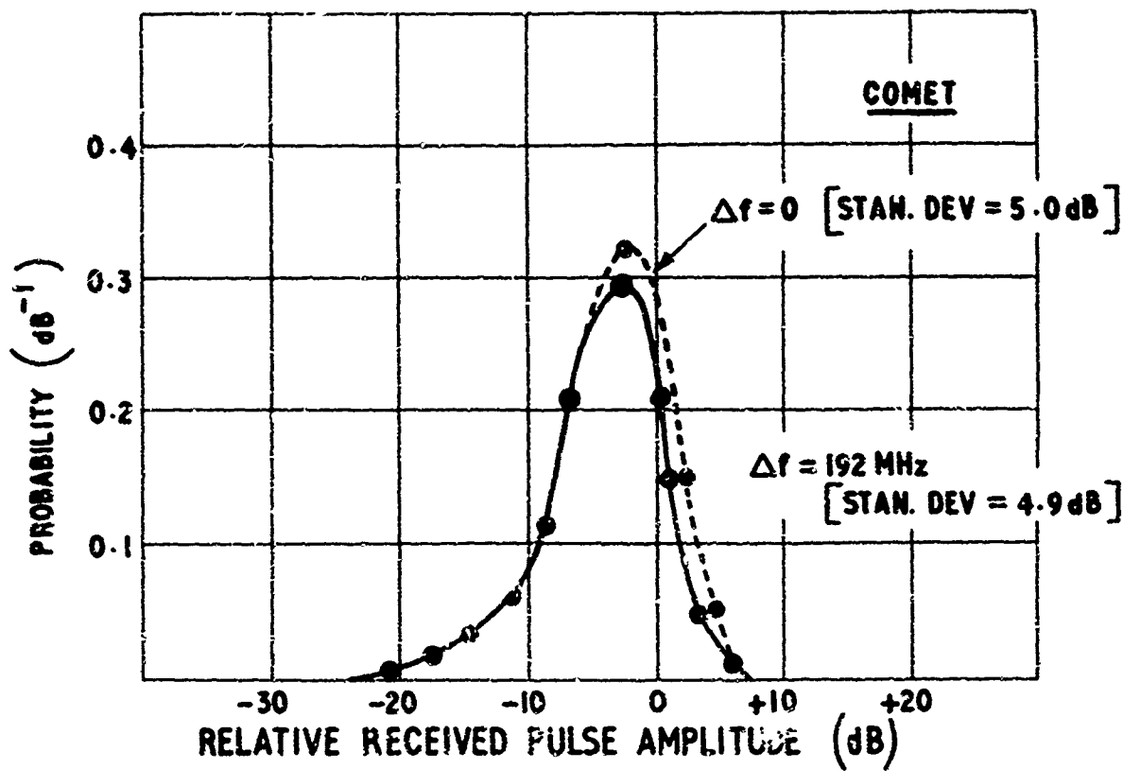


Figure 7 Echo Amplitude Probability Density Function for a Comet Aircraft

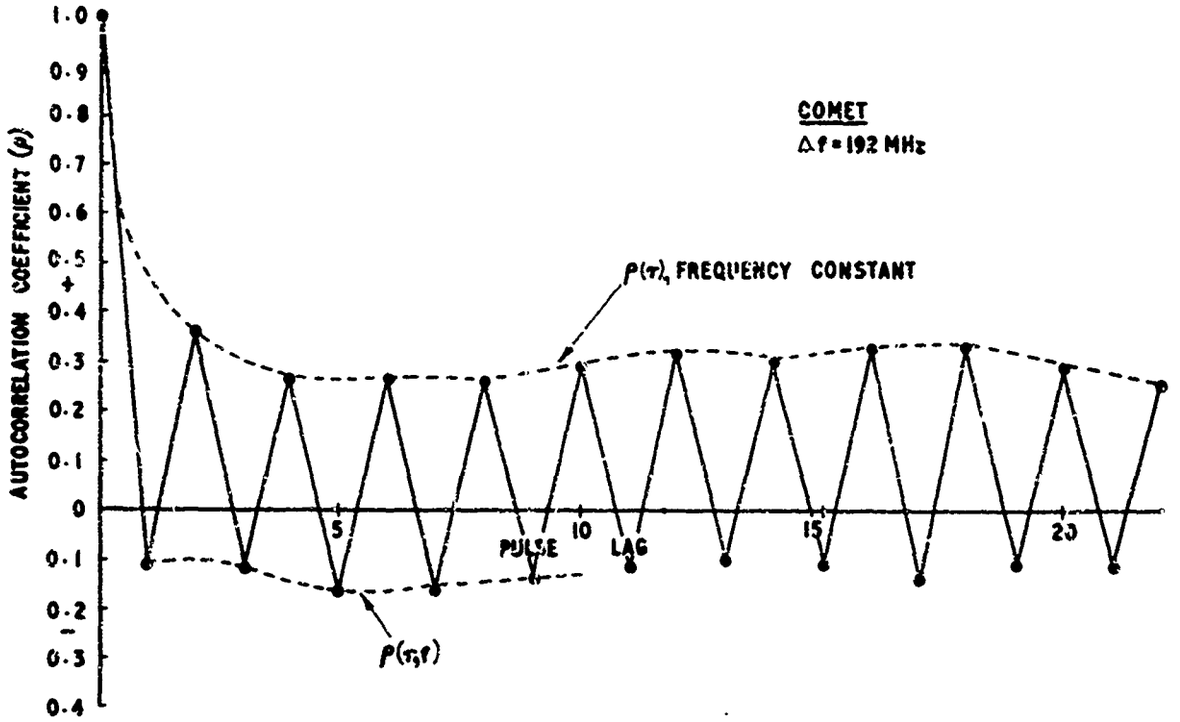


Figure 8 An Auto-Correlation Function of the Comet Aircraft Echoes with Frequency Agility

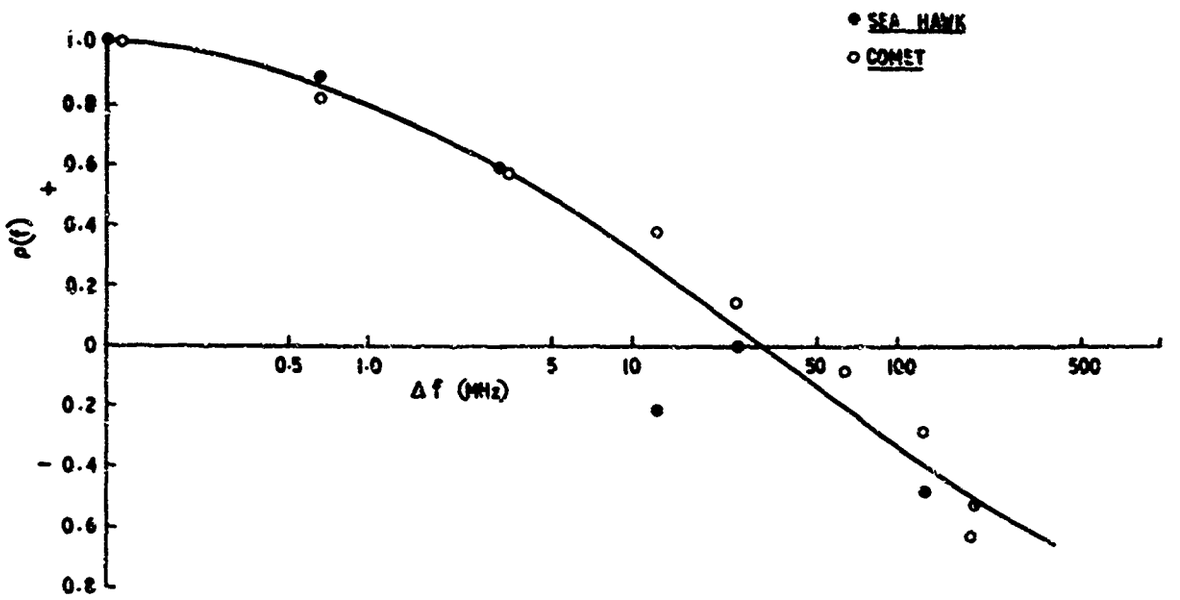


Figure 9 Derived Frequency Correlation Coefficients

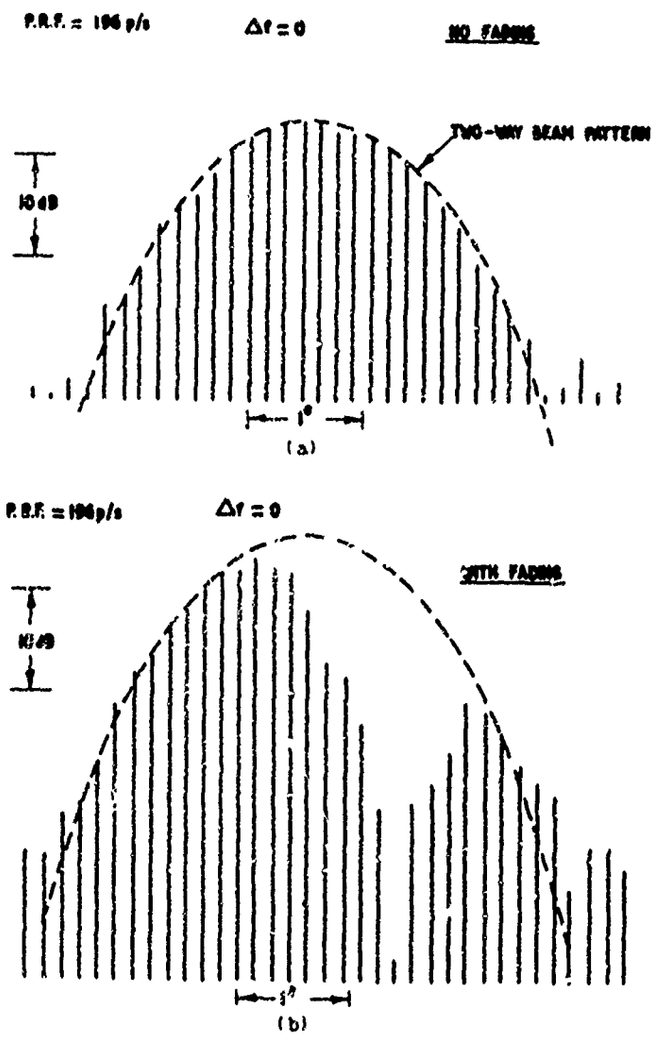


Figure 10 Scanned Echo Amplitude Patterns - Fixed Frequency

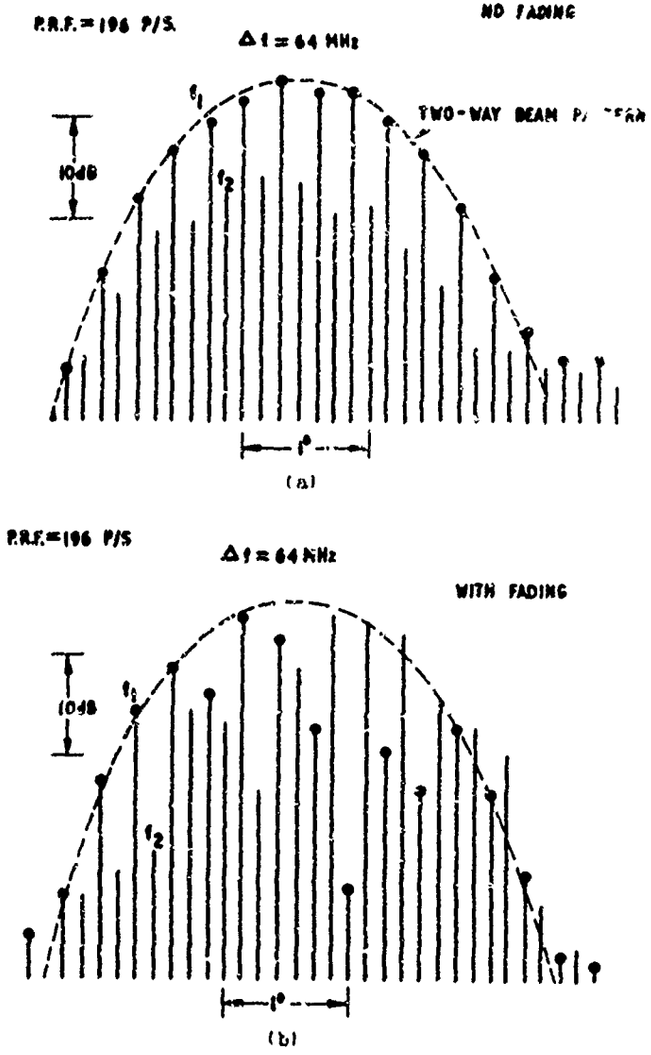


Figure 11 Scanned Echo Amplitude Patterns - Frequency Change (4 MHz

RECENT DEVELOPMENTS IN TARGET MODELS FOR RADAR DETECTION ANALYSIS

by

P.Swerling

Technology Service Corporation,
225 Santa Monica Boulevard,
Santa Monica,
California, U.S.A.

RECENT DEVELOPMENTS IN TARGET MODELS FOR RADAR DETECTION ANALYSIS

P. Swerling
TECHNOLOGY SERVICE CORPORATION
225 Santa Monica Boulevard, Santa Monica, California, U.S.A.

SUMMARY

In the fifteen or so years since publication of detection probability curves for the so-called "Swerling" target fluctuation models, analysis of target signature data for many different target populations has indicated that the Swerling models do not necessarily describe, or even bracket, the radar cross section (RCS) fluctuation behavior of all targets of interest. For this reason, much broader families of target RCS models have been developed, of which the Swerling models are special cases, and detection probability curves have been generated for these broader families.

This paper will review several of these broader families of target signature models, including fluctuations whose statistical behavior is governed by chi-square, log-normal, and Rice distributions; will briefly describe possible physical characteristics of targets which tend to give rise to these distributions; and will review detection probability results for these families. Partially correlated cases will be included.

Several methods will be given whereby published detection probability curves can be interpolated or extrapolated in a simple manner to partially correlated cases as well as to other probability distributions.

Most of the results to be discussed assume conventional radar receiver models of the type assumed by Marcum and Swerling in their early work. However, a brief discussion will be given of the consequences of various target models for more intricate types of receiver processing such as Constant False Alarm Rate (CFAR) methods, frequency diversity methods, and high resolution (short pulse) methods.

1. INTRODUCTION.

1.1. Purpose and Outline of the Paper.

Oversimplified descriptions of the reflective properties of radar targets--for example, the description of a target simply as having a radar cross section of so many square meters--can lead to incorrect performance predictions. For this reason, the "classical" target fluctuation models (the Swerling cases, Ref. 1) were introduced, and the effect of such fluctuations on detection performance were analyzed.

However, since that time it has become apparent that these conventional models may still be inadequate as descriptions of target behavior; in fact, the original Swerling models do not necessarily describe or even bracket the fluctuation behavior of all target populations of interest. Various examples will be given below.

In the years since these models were originally formulated, methods have been worked out for computing detection probability for essentially arbitrary types of fluctuation, and specific numerical results have been developed for a great many cases. The present paper will summarize many of these results, and refer to the literature where further such results can be found.

In Section 2., three families of fluctuation models will be described which have been found to describe, to a reasonable approximation, many target populations of interest--though even these three families are not co-extensive with all cases of practical interest. Possible scattering mechanisms which may give rise to these models will also be discussed.

In Section 3., general methods will be stated by which detection probability can be calculated for arbitrary fluctuations. Also in this section, a number of methods will be stated by means of which the actual task of calculation can be greatly simplified.

In Section 4., detection probability results will be given (or appropriate literature referred to) for the three families introduced in Section 2. The methods by which these results were obtained in some cases afford examples of the simplified (but still very accurate) methods stated in Section 3.

Section 5. gives a very brief discussion of some related topics, chiefly the relation of target fluctuation models to more intricate types of signal processing.

1.2. Assumptions for Radar Detection Analysis.

The radar receiver model upon which the results of this paper are based is the one conventionally employed for detection analysis, as in Figure 7-1.

The input to the detector can be written as

$$\text{detector input} = v(t)\exp(j\omega t) \tag{1}$$

where

ω = IF frequency (radians/sec)

$v(t)$ = complex envelope of detector input

Let

v_i = complex envelope of detector input at the time epoch of the i^{th} sample.

Then, it is assumed the detector is a square law envelope detector, with outputs w_i where

$$w_i = \frac{1}{2} |v_i|^2 \tag{2}$$

(The constant 1/2 is for analytic convenience and involves no loss of generality. It is also the case that the ensuing results are not sensitive to the specific detector law assumed.)

The integrated signal V is

$$V = \sum_{i=1}^{N} \alpha_i w_i \tag{3}$$

where α_i are positive constants.

The signal-to-noise ratio for the i^{th} sample will be denoted x_i, i.e.,

x_i = ratio of signal energy to noise power density for the i^{th} sample.

The noise will be assumed to be Gaussian and independent from sample to sample. This does not necessarily restrict consideration to thermal noise in the receiver.

Some additional comments and assumptions are as follows.

(a) The term "sample" is used here instead of "pulse", although in most cases each pulse may in fact be a separate sample and vice versa. However, the analysis applies to some cases when the concepts of pulse and sample may not be identical. For instance, a sample may consist of a matched filter response (at a specified time instant) to a pulse burst, or to a coherent pulse train.

As an example, suppose the radar is fully coherent for the duration of a target illumination time, i.e., all pulses during the target illumination time are coherently processed. In that case, N = 1, i.e., the number of samples per illumination time is one, although the number of pulses may be greater than one. Moreover, the assumptions of the analysis are satisfied.

If the radar coherently processes several non-overlapping pulse trains (e.g., pulse bursts) during the illumination time, then the assumptions are still satisfied but each sample corresponds to several pulses. If the receiver coherently processes overlapping sets of pulses (as in a pulse-cancellation MTI receiver), the assumption of sample-to-sample noise independence is no longer satisfied but it may still be possible to apply the analysis approximately.

(b) It is assumed that the target signal strength does not significantly fluctuate within the coherent processing time of the radar receiver. This assumption is usually true unless the coherent processing time extends over more than one pulse.

If the assumption is violated, the signal phase fluctuations are generally of greater significance to the resulting performance than are the amplitude fluctuations.

1.3. Signal-to-Noise Ratio Fluctuations--Effects on Detection Performance.

In general, due to a number of factors, particularly but not exclusively the fluctuation of target radar cross section (RCS), the signal-to-noise ratios x_i are themselves random variables. Subject to the assumptions just stated, the effect of fluctuations on detection performance can be entirely specified by the joint probability density function $p(x_1,\ldots,x_N)$ of the signal-to-noise ratios for the N samples.* Thus, the required target model must be capable of predicting this joint probability density.

There is one condition under which the required specification of target fluctuation is much simpler.

Specifically, if the α_i's all have the same value, then it can be shown (Section 3.; also Ref. 2) that the effects of fluctuation on detection are entirely determined by the probability distribution of X, where

$$X = \sum_{i=1}^{N} x_i \tag{4}$$

We shall use this fact extensively in what is to follow. In this connection it is convenient to state the following definitions: let

$$\bar{X} = \text{mean of } X = \sum_{i=1}^{N} \bar{x}_i \tag{5}$$

where $\bar{x}_i$ are the ensemble means (over the signal fluctuations) of x_i.

Define

$$\bar{x} = \text{average signal-to-noise ratio (SNR) per sample} = \frac{\bar{X}}{N} \tag{6}$$

We may also define the second moment of X in terms of the second order moments of x_i, i = 1,...N:

*Specific formulas to be given in Section 3.

$$E(X^2) = \sum_{i=1}^{N} \sum_{j=1}^{N} F(x_i, x_j) \tag{7}$$

$$\text{variance of } X = \sigma_X^2 = E(X^2) - (\bar{X})^2 \tag{8}$$

Discussion of the Assumption That α_i = Constant (Independent of i).

The assumption that α_i = constant (independent of i), i.e., that the video integrator forms a uniformly weighted sum of detector outputs, is almost universal in radar detection analysis and is the basis for virtually all published detection curves.

Such an assumption is physically true in some cases (e.g., in many automatic radar video processors) but not in others (e.g., when integration is performed on a CRT).

The probability distribution of the integrated signal-plus-noise V can be obtained without this assumption (e.g., Ref. 3). Extensive calculations based on the results of Ref. 3 have not been carried out to the author's knowledge.

A limited amount of analysis of detection probability for the case of non-constant α_i (Refs. 4 and 5) indicates that the results for non-constant α_i can be closely approximated by those for constant α_i if in the latter case, the "effective" value of N is appropriately chosen. However, more work remains to be done in this area.

2. CATEGORIES OF FLUCTUATION MODELS.

A major portion of the results to be stated will be devoted to treatment of three families of RCS fluctuation models which have been observed to characterize target populations of interest: the chi-square family, the Rice family, and the log-normal family. These families are much more comprehensive than the conventional fluctuation models (which are special cases of the chi-square family). It is <u>not</u> contended that even these more comprehensive families describe all targets which may be of practical interest. However, they describe, at least to a good approximation, a wide variety of target populations.

2.1. The Chi-Square Family.

2.1.1. Definitions.

A random variable v will be said here to have a chi-square distribution if its probability density function (p.d.f.) is of the form

$$w_k(v,\bar{v}) = \frac{1}{(k-1)!}\frac{k}{\bar{v}}\left(\frac{kv}{\bar{v}}\right)^{k-1}\exp\left(\frac{-kv}{\bar{v}}\right), \quad v > 0 \tag{9}$$

$$= 0, \quad v \leq 0.$$

where

$$\bar{v} = \text{mean of } v \tag{10}$$

$$k = \frac{\bar{v}^2}{\text{variance of } v}$$

Some notes on terminology: in statistics texts, the chi-square density functions are defined as in Eq. (9) but with the additional requirement that $\bar{v} = 2k$; thus the conventional statistics text definition is that of a one-parameter family. Our chi-square family is a two-parameter family with $\bar{v}$ and k independent parameters; it is the same as the conventionally defined chi-square family up to an additional scale factor $\bar{v}/2k$. (Our family is also equivalent to the gamma distributions of statistics (Ref. 6).)

Also, in statistics texts, 2k is called the "number of degrees of freedom" and is an integer. In our usage, 2k is <u>not</u> required to be an integer, i.e., we will allow k to be any positive real number, although we will sometimes borrow the conventional terminology and refer to k as "half the degrees of freedom."

When the individual signal-to-noise ratios x_i have chi-square distributions, we shall refer to the fluctuation model as "wide-sense chi-square." The p.d.f.'s of the individual x_i are then obtained by substituting, in Eq. (9), $v = x_i$, $\bar{v} = \bar{x}_i$, $k = k_i$ (the individual samples may conceivably have different values of k_i, although most of our treatment will be devoted to cases in which $k_i = k$ is not dependent on i).

When $X = \sum x_i$ also has a chi-square distribution, we shall say that the fluctuation model is "strict-sense chi-square" and the parameter representing half the degrees of freedom of X will be written:

$$w_K(X,\bar{X}) = \frac{1}{(K-1)!}\frac{K}{\bar{X}}\left(\frac{KX}{\bar{X}}\right)^{K-1}\exp\left(\frac{-KX}{\bar{X}}\right), \quad X > 0 \tag{11}$$

$$= 0, \quad X \leq 0$$

Thus, k will denote the degree-of-freedom parameter for the individual samples and K the parameter for the sum variable X.

Also, as defined in Eq. (6), the average single-sample SNR $\bar{x}$ is by definition $\bar{X}/N$, or in other words $\bar{X} = N\bar{x}$.

2.1.2. Relation of k to K.

There are three special cases in which X is exactly chi-square distributed if the individual samples are chi-square distributed:

(a) Rapid fluctuation with x_i drawn from a common chi-square distribution. "Rapid fluctuation" is defined to mean that the values of x_i, $i = 1,\ldots,N$, are all statistically independent.

(b) Slow fluctuation with the samples having a common k, but not necessarily common means. "Slow fluctuation" is defined to mean that the values of $x_i/\bar{x}_i$ are all identical, $i = 1,\ldots,N$.

(c) Block-correlated fluctuation with samples from a common distribution. "Block-correlated" fluctuation is defined to be a case where the N samples can be subdivided into F equal-sized groups, within each group the samples are all equal, but they are statistically independent from group to group. F is called the number of independent sample groups. This model would characterize a case in which a target would be slowly fluctuating in the absence of frequency diversity, and in which F different frequencies (sufficiently separated) are used, with N a multiple of F and an equal number of samples at each frequency. (The samples in each block need not be successive samples.) The relation of k to K in each of these cases is:

Type of Fluctuation	Relation of k to K
Rapid	$K = Nk$
Slow	$K = k$
Block-correlated	$K = Fk$

For more general types of correlation, X will no longer be exactly chi-square even if the individual samples are. However, X will be well approximated by a chi-square distribution having the same first and second moments as X, i.e., the distribution of X will be approximately chi-square with*

$$\bar{X} = \sum_{i=1}^{N} \bar{x}_i \tag{12}$$

$$K = \frac{(\bar{X})^2}{\text{variance of } X} \tag{13}$$

The variance of X is given by Eqs. (7) and (8), Section 1.

The approximation of a distribution by a chi-square distribution having the same first and second moments as the given distribution is equivalent to using the first term in the Laguerre series expansion (Ref. 1, p. 189). Thus, we are saying that if the individual samples have chi-square distributions, the distribution of X can be well approximated by the first term in its Laguerre series expansion, for general correlation properties.**

In the block-correlated case, if the F blocks are not all of equal size but are of approximately equal size, then X is not exactly chi-square distributed but it is approximately chi-square and to a very good approximation $K \approx Fk$. For example, this holds if there are F blocks and N is not an exact multiple of F, but the left-over samples are assigned to blocks as evenly as possible. (E.g., if there are 23 samples, and five blocks, with two blocks containing four samples and three blocks containing five samples.

2.1.3. Relation to Swerling, Marcum, and Weinstock Cases.

Swerling and Marcum cases are special cases of strict-sense chi-square cases:

Model	Value of k	Value of K
Non-fluctuating	∞	∞
Swerling I	1	1
Swerling II	1	N
Swerling III	2	2
Swerling IV	2	2N

In addition, Weinstock (Ref. 7) investigated RCS fluctuation models for a class of objects consisting of long cylinders and cylinders with fins with the objective of seeing whether their fluctuation statistics could be approximated by chi-square statistics. When the RCS patterns of such objects are viewed at randomly selected aspects, he found that the best-fitting chi-square distributions have values of k less than unity, specifically, k could fall in the range $1/3 \leq k \leq 1$.*** Thus, cases in which the individual samples are chi-square with $k < 1$ will be called Weinstock models.

---

*Provided the individual samples are chi-square distributed.

**The precise accuracies achieved have not been systematically explored for all conceivable correlations of x_i. This would form an interesting Master's Thesis. A possible approach is to calculate the second or higher terms in the Laguerre series.

***It is also true that Weinstock found that even the best-fitting chi-square distributions did not give a very good fit; this and the discussion of Section 2.4 suggests that perhaps a better fit could have been obtained with log-normal distributions.

This brings up an important point: the fluctuation models of physical interest are not bracketed by the Swerling and Marcum models, nor is the probability of detection behavior bracketed by that for Swerling and Marcum models (see further discussion, Section 4.).

2.2. The Rice Family.

2.2.1. Definition.

A random variable v will be said to have a Rice power distribution if v has p.d.f.*

$$w(v,s,\psi_o) = \frac{1}{\psi_o}\exp\left[-x-\frac{v}{\psi_o}\right] I_o\left(2\sqrt{sv/\psi_o}\right) \quad , \qquad v > 0 \tag{14}$$

$$= 0 \quad , \qquad v \leq 0$$

where I_o = modified Bessel function of order zero (Ref. 1, p. 167).

It is also convenient to define a parameter d by

$$s = \frac{d^2}{2\psi_o} \tag{15}$$

in terms of which the Rice power p.d.f. is

$$w(v,d,\psi_o) = \frac{1}{\psi_o}\exp\left[\frac{-v-(d^2/2)}{\psi_o}\right] I_o\left(\frac{d\sqrt{2v}}{\psi_o}\right) \tag{16}$$

When applied to the distribution of individual samples x_i of SNR, the p.d.f. is given by simply substituting x_i for v in Eq. (16).

2.2.2. Relation to Chi-Square Family.

The Rice Power distribution is well approximated by the first term in its Laguerre series expansion, i.e., by the chi-square distribution with the same first and second moments.

The first two moments of the Rice Power distribution are

$$\text{mean} = \bar{v} = \psi_o(1+s) = \bar{x} \tag{17}$$

$$\text{variance} = \psi_o^2\,(1+2s) \tag{18}$$

Thus, the value of k for the approximating chi-square distribution is

$$k = 1 + \frac{s^2}{1+2s} \tag{19}$$

(When s = 0, the Rice distribution is exactly the chi-square with k = 1.)

In this connection, we may note that Scholefield (Ref. 3) treated the Rice model of RCS fluctuation, doing two things:

(a) Calculating detection probability of Rice-fluctuating targets.

(b) Stating correspondence between Rice and chi-square fluctuations.

The first he did correctly but the second he did incorrectly. For example, he states that Swerling Case III (with k = K = 2) corresponds to s = 1. However, from Eq. (19), we see that k = 2 actually corresponds to s = 2.4, while k = 4/3 corresponds to s = 1. This is confirmed by detection probability calculation (see Section 4.).

2.3. The Log-Normal Family.

2.3.1. Definition.

A log-normally distributed random variable is one whose logarithm has a normal distribution. For a random variable v, the p.d.f.'s for the log-normal family are

$$w(v,\alpha,\sigma) = \frac{1}{\sigma\, v\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}\left[\ln\left(\frac{v}{\alpha}\right)\right]^2\right\} \quad , \qquad v > 0 \tag{20}$$

$$= 0 \quad , \qquad v \leq 0$$

where

α = median of v

σ = standard deviation of $\ln\left(\frac{v}{\alpha}\right)$

*Often in the literature the symbol ψ_o denotes average receiver noise power; our use of this symbol is different, i.e., it is a parameter of a distribution of SNR or of RCS.

The parameter σ is sometimes in the literature replaced by another, ρ, defined as

$$\rho = \exp(\sigma^2/2) \tag{21}$$

Interpretation of ρ

It is best to interpret ρ as being defined by Eq. (21) and as being simply an alternate form of stating the parameter σ. The parameter σ determines the slope of the straight line representing the cumulative log-normal distribution on normal probability paper.

For an exactly log-normal distribution of v, we may also say that

$$\rho = \frac{\bar{v}}{\sigma} = \text{ratio of mean to median of } v \tag{22}$$

$$\rho = \sqrt{v_2}/\bar{v} \quad (\text{where } v_2 = \text{second moment of } v) \tag{23}$$

However, for a random variable v which is approximately but not exactly log-normal, the value of ρ for the best-approximating log-normal distribution does not necessarily satisfy either of the relations (22) or (23) in terms of the actual moments of the original random variable v or of the data.

2.3.2. Distribution of X.

When the individual x_i come from a common log-normal distribution and are slowly fluctuating, X is also log-normal with the same value of σ, and median = N_o.

In other cases, X is not exactly log-normal. There is some evidence that for a moderate number of samples, the distribution of X is well approximated by a log-normal distribution. However, the moments of the best-approximating log-normal distribution are not the same as the actual moments of X, and at present, no general method is known of determining the parameter σ (or ρ) for the best-approximating log-normal distribution to that of X.

2.4. Physical Scattering Mechanisms Associated with the Three Families.

2.4.1. Motivation.

It is important for purposes of radar systems analysis to have some understanding of physical scattering mechanisms which lead to given RCS statistics, not just in the sense that a specific target is observed to have given statistics, but in a theoretical sense. Here, "theoretical" is not used in opposition to "practical"; rather, it refers to a case in which given RCS statistics can be mathematically derived from a scattering model, in the sense that Rayleigh amplitude statistics (i.e., chi-square with k = 1 power statistics) are mathematically derivable from a many-scatterer mechanism. Of course, the majority of real targets are only approximations to the ideal scattering mechanisms which theoretically give rise to given distributions; nevertheless, it is necessary to understand what ideal mechanisms may produce given statistics.

The requirement for such understanding arises from those cases (which are more common than any others) in which measurement data are fragmentary or have significant gaps and in which no exact physical description of the target or target population is available. In such cases, we may wish to (a) extrapolate statistical models to parameter regions where measurement data are lacking, or (b) hypothesize a variety of possible statistical target models.

The second point is particularly important. When data and detailed physical descriptions are lacking, the standard method of approach is to postulate a variety of statistical models which are thought to bracket the behavior of the target population of interest, and then perform radar performance analyses for a variety of such models, noting the performance sensitivity to the assumed model, attempting to choose design configurations whose performance is relatively insensitive to the statistical model etc.

However, it is essential that the postulated models be: (a) self-consistent, (b) consistent with possible physical scattering mechanisms, and (c) consistent with such measurement data as has been observed. It is very difficult to postulate consistent models of this type without some understanding of ideal scattering mechanisms which can produce given RCS statistics.

Section 5. gives a more detailed discussion of this, with several specific examples: postulation of models for fluctuation behavior of a target with respect to frequency diversity, and postulation of models for detection analysis of targets which are extended in range compared to a range resolution cell.

Each of these examples is discussed with respect to models calling for Rayleigh RCS statistics, Rice RCS statistics, and log-normal RCS statistics. These examples illustrate why a lack of understanding of theoretical scattering mechanisms, which can give rise to given statistics, can be a stumbling block in postulation of consistent models.

2.4.2. Physical Models for Rice and Chi-Square Distributions.

It is of course well known that the Rayleigh amplitude distribution, i.e., the exponential power distribution (chi-square with k=1 or Rice with $s = 0$) arises from a many-scatterer mechanism, which is defined to be a mechanism with a very large number of scatterers, no one of which ever contributes more than a small fraction of the returned power, and which have random relative phases.

A Rice distribution of radar cross section with parameters σ_o and s arises from a scatterer with a many-scatterer component whose average RCS is σ_o, plus a single non-fluctuating scatterer having RCS equal

to $s\sigma_n^2$. The relative phases of all scatterers are assumed random. Thus, s represents the ratio of RCS of the single scatterer to that of the Rayleigh component. This is often called a "single predominant scatterer" model, although we do not require that s be greater than unity. Possibly "single significant scatterer" would be a better terminology.

The chi-square distributions have been observed empirically to give a good fit to the RCS distributions of many types of targets. However, except for $k = 1$, there is no theoretical scattering mechanism which exactly gives rise to chi-square distributions.* We previously noted, however, that the chi-square distributions for $k > 1$ are good approximations to Rice distributions; thus, the chi-square distributions can be theoretically explained by a single-predominant plus Rayleigh model. The chief reasons for consideration of the chi-square family are that: (a) probability of detection curves are available and easily calculated for the complete chi-square family (both strict sense and wide sense), and (b) some chi-square cases of interest have empirically been found to correspond to $k < 1$ (Ref. 7) which does not correspond to any Rice distribution.

2.4.3. Physical Models for Log-Normal Distributions.

A log-normal distribution theoretically arises from a product of statistically independent random variables in the limit as the number of factors increases. The difficulty in explaining log-normal amplitude statistics of radar targets on this basis lies in the difficulty of identifying a product of independent or approximately independent random variables with radar targets. Progress along these lines has been made recently. The results to be stated are due to Drs. R. L. Mitchell and G. E. Pollon.

A first step in the direction of understanding the scattering mechanisms underlying approximately log-normal amplitude statistics is the observation that the amplitude distribution of the gain patterns of directive antennas, viewed at randomly chosen aspects, tends to be approximately log-normal. This has been noticed by deriving empirical distribution functions for such gain patterns. This also suggests that the radar cross section patterns of highly directive scatterers such as flat plates viewed at random aspects may give rise to log-normal RCS statistics, approximately (Ref. 10). This, however is still an empirical observation rather than a theoretical explanation.

Dr. Gerald Pollon has recently developed (Ref. 11) an approach to the formulation of a theoretical explanation of why sufficiently directive scatterers give rise to approximate log-normal statistics. This approach succeeds in identifying the RCS of directive scatterers with a product of approximately independent random variables. Thus, this approach results in a theoretical model to explain why directive scatterers give rise to such statistics. This method also has promise of being able to yield quantitative, albeit approximate, values for the parameters of such distributions.

A great deal of further work remains to be done relative to physical scattering mechanisms underlying targets whose RCS statistics are observed to be log-normal. Some additional avenues of investigation would include:

(a) RCS statistics for the phasor sum of several directive scatterers.

(b) RCS statistics for the phasor sum of a directive scatterer and a Rayleigh mechanism.

It may also be noted that actual scattering patterns have a finite peak value and hence the RCS statistical distribution cannot be exactly log-normal. Possibly truncated log-normal statistics will give a better fit to some observed distributions.

3. DETECTION AND ACQUISITION PROBABILITY.

3.1. General Exact Formulas.

Recalling Section 1., let V be the final integrated signal-plus-noise variable which is subjected to a threshold test for detection on any one scan.

Let

$$P(V|x_1,\ldots,x_N) = \text{probability density function of } V \text{ conditional on a specific set of SNR's } x_1,\ldots,x_N \tag{24}$$

and

$$C(t|x_1,\ldots,x_N) = \int_0^\infty P(V|x_1,\ldots,x_N)e^{-Vt}\,dV \tag{25}$$

$$= \text{characteristic function}^{**} \text{ of } V \text{ conditional on a specific set of SNR's } x_1,\ldots,x_N.$$

Finally, let $p(x_1,\ldots,x_N)$ denote the joint p.d.f. of $x_1,\ldots,x_N$.

Then, the unconditional p.d.f. P(V) and its characteristic function (i.e., Laplace transform) C(t) are given by

---

*Mathematically, a chi-square distribution with $2k$ degrees of freedom, $2k$ an integer, results from the sum of squares of $2k$ independent variates drawn from a common Gaussian distribution. For $k \neq 1$, one cannot exactly identify a physical scattering mechanism with such a sum of squares.

**For positive random variables, the characteristic function can, if we wish, be defined with t real, i.e., as the Laplace transform of P.

$$P(V) = \int P(V|x_1,\ldots,x_N)\, p(x_1,\ldots,x_N)\, dx_1 \ldots dx_N \tag{26}$$

$$C(t) = \int_0^\infty P(V)\, e^{-Vt}\, dV \tag{27}$$

$$= \int C(t|x_1,\ldots,x_N)\, p(x_1,\ldots,x_N)\, dx_1 \ldots dx_N$$

It is well known (assuming uniform weights in the video integrator, i.e., α_i = constant in Figure 7-1) that $C(t|x_1,\ldots,x_N)$ is (Refs. 1 and 2)

$$C(t|x_1,\ldots,x_N) = C(t|X) = (1+t)^{-N} \exp\left[\frac{tX}{1+t}\right] \tag{28}$$

where

$$X = \sum_{i=1}^{N} x_i \tag{29}$$

and

$$P(V|x_1,\ldots,x_N) = P(V|X) = \left(\frac{V}{X}\right)^{\frac{N-1}{2}} e^{-V-X}\, I_{N-1}(2\sqrt{VX}) \quad , \qquad V > 0 \tag{30}$$

These are the exact formulas whereby the p.d.f. of V can be determined for arbitrary joint distributions of $x_1,\ldots,x_N$. By combining Eqs. (28) and (27), another convenient form for C(t) can be stated:

$$C(t) = (1+t)^{-N}\, \varphi\left(\frac{t}{1+t}\right) \tag{31}$$

where $\varphi(t)$ is the characteristic function of X:

$$\varphi(t) = \int_0^\infty e^{-tX}\, p(X)\, dX \tag{32}$$

p(X) = p.d.f. of X. From Eq. (32) the first n moments of V can easily be determined in terms of the first n moments of X.

3.2. The Representation Problem.

One of the problems in using the above-stated formulas is in representing the joint p.d.f. of $x_1,\ldots,x_N$ for general partially correlated cases.

For fast, slow, or block-correlated cases, the representation is simple: it is merely the product of the p.d.f.'s of the independent samples, regardless of what form these p.d.f.'s have.

For more general types of partial correlations, we can still give the representations for any cases in which the samples can be expressed as functions of Gaussian random variables. These include the chi-square, Rice, and log-normal families defined in Section 2.

The general method is:

(a) Express $x_1,\ldots,x_N$ in terms of a set of Gaussian random variables which may be denoted by a vector $\underline{u}$.

(b) Express the integrals in Eqs. (26) - (27) as integrals with respect to a joint Gaussian density function $p(\underline{u})$, i.e., replace $p(x_1,\ldots,x_N)\, dx_1 \ldots dx_N$ by $p(\underline{u})\, d\underline{u}$ and replace $x_1,\ldots,x_N$ by the appropriate functions of the components of $\underline{u}$.

The general joint Gaussian distribution for $\underline{u}$ is

$$p(u) = (2\pi)^{-n/2}\, \Delta^{1/2} \exp\left\{-\frac{1}{2} \sum_{i,j=1}^{n} \xi_{ij}(u_i - \bar{u}_i)(u_j - \bar{u}_j)\right\} \tag{33}$$

where

ξ_{ij} = inverse covariance matrix of $\underline{u}$

Δ = determinant of (ξ_{ij})

Here n is in general larger than N.

The application to chi-square, Rice, and log-normal families is as follows.

3.2.1. Chi-Square (2k = integer).

If the individual x_i are drawn from a common chi-square distribution, then we can write

$$x_i = \sum_{j=1}^{2k} u_{ij}^2 \quad , \qquad i = 1,\ldots,N \tag{34}$$

where u_{ij} are Gaussian, zero-mean and have correlation properties expressed by

$$E(u_{ij}\, u_{i'j'}) = \eta_{iji'j'} \tag{35}$$

$$\eta_{iji'j'} = 0 \quad \text{if} \quad i = i' \quad \text{and} \quad j \neq j' \tag{36}$$

$$\eta_{ijij} = \eta_{ij'ij'} \quad \text{for all} \quad i,j,j' \tag{37}$$

Conditions (36) and (37) are required if x_i are to be exactly chi-square distributed. If these conditions are omitted, x_i will still be approximately chi-square, in general with 2k non-integral. Otherwise, $\eta_{iji'j'}$ is an arbitrary correlation matrix with i = 1,...,N, j = 1,...,2k, and similarly for i' and j'. Thus, $\underline{u}$ has components u_{ij} and n = 2kN in Eq. (33).

Reference 3 has applied these representations to determine C(t) even when the video integrator weights (α_i in Figure 7-1) are not constant.

3.2.2. Rice.

If x_i are Rice-distributed, then they can be represented

$$x_i = \frac{1}{2}\left[(u_{i1} + d)^2 + u_{i2}^2\right] \tag{38}$$

where u_{i1} and u_{i2} are zero-mean Gaussian and

$$E\left[u_{ij}\, u_{i'j'}\right] = \eta_{iji'j'} \tag{39}$$

where i,i' = 1,...,N and j and j' take on only the values 1 and 2. $\eta_{iji'j'}$ is a correlation matrix satisfying

$$\eta_{iji'j'} = 0 \quad \text{if} \quad i = i' \quad \text{and} \quad j \neq j' \tag{40}$$

$$\eta_{i1i1} = \eta_{i2i2} \quad , \quad \text{all } i \tag{41}$$

Then, $\underline{u}$ has 2N components, u_{i1}, u_{i2}, i = 1,...,N.

3.2.3. Log-Normal.

For partially correlated log-normal samples $x_1,\ldots,x_N$, it is not necessary to replace $p(x_1,\ldots,x_N)$ by $p(\underline{u})$; the joint p.d.f. of $x_1,\ldots,x_N$ can be expressed directly as

$$p(x_1,\ldots,x_N) = (2\pi)^{-N/2}\, \Delta^{1/2} \prod_{i=1}^{N} \left(\frac{1}{x_i}\right) \tag{42}$$

$$\exp\left\{-\frac{1}{2}\sum_{i,j=1}^{N} \xi_{ij}\, \ln\left(\frac{x_i}{\alpha_i}\right) \ln\left(\frac{x_j}{\alpha_j}\right)\right\}$$

where ξ_{ij} is an arbitrary inverse covariance matrix and α_i are arbitrary positive constants.

As a technical point, it may be noted that the above methods do not generate all possible joint p.d.f.'s of variables which individually have chi-square, Rice, or log-normal distributions, because it is possible for random variables to be individually Gaussian but not jointly Gaussian, i.e., $p(\underline{u})$ given by Eq. (33) is the most general joint Gaussian p.d.f. but does not represent all possible joint p.d.f.'s of variables which are individually Gaussian.

3.3. Power Fluctuation Correlation.

Often the observed data consist of amplitude-only or power-only observations and not observations of phase and amplitude (or of sine and cosine components). In such cases, it may be possible to derive only power correlation functions directly from the data, so it is useful to state relations between the power fluctuation correlation and the correlation of the Gaussian variables by which Rice, chi-square, or log-normal variables may be represented.

3.3.1. Rice and Chi-Square.

The general method of relating the power-correlation functions for Rice and chi-square variables to the correlations for the components of $\underline{u}$ in Eq. (33) is as follows: simply express the Rice or chi-square variables as in Eqs. (38) or (34) and then apply the well-known formulas relating fourth-order moments of Gaussian distributions to second-order moments. The relation of power correlation to the correlation properties of $\underline{u}$ is of course not one-to-one: many different correlation matrices of $\underline{u}$ can give rise to the same power correlation properties.

The following examples are not for the most general cases given in Section 3.2., but they illustrate the method.

(a) Rice (including Rayleigh if d = 0). Let u_{i1} and u_{i2} of Eq. (38) have correlation properties

$$E(u_{i1}\, u_{j2}) = 0 \quad \text{all } i,j \tag{43}$$

$$E(u_{i1}\, u_{j1}) = E(u_{i2}\, u_{j2}) = \psi_o\, \rho_{ij} \tag{44}$$

where ρ_{ij} is a correlation matrix with $\rho_{ii} = 1$ but otherwise arbitrary. In Eqs. (43) and (44), we have changed the subscript notation somewhat from Eqs. (39) - (42): the indices i and j take on the values i = 1,... ,N and j = 1,...,N, i.e., i' is replaced by j. Then the power fluctuation correlation is

$$E(x_i x_j) - \overline{x_i}\;\overline{x_j} = \psi_o^2\, \rho_{ij}^2 + d^2\, \psi_o\, \rho_{ij} \tag{45}$$

The mean value of x_i is

$$\overline{x_i} = \psi_o + \frac{d^2}{2} = \psi_o(1+s) \quad , \quad \text{all } i \tag{46}$$

(b) Chi-Square with k = 1/2. Suppose $x_i = u_i^2$ with

$$\overline{u_i\, u_j} = \psi_o\, \rho_{ij} \tag{47}$$

where ρ_{ij} = correlation matrix with $\rho_{ii} = 1$.

Then

$$E(x_i x_j) - \overline{x_i}\;\overline{x_j} = 2\psi_o^2\, \rho_{ij}^2 \tag{48}$$

(c) Log-Normal. Let $x_i = e^{u_i}$, where u_i are jointly Gaussian, with

$$u_i = \alpha \quad , \quad \text{all } i \tag{49}$$

$$E(u_i^2) = \sigma^2 \quad , \quad \text{all } i \tag{50}$$

$$E(u_i u_j) = \sigma^2\, \rho_{ij} \tag{51}$$

Then

$$E(x_i x_j) - \overline{x_i}\;\overline{x_j} = d^2\, e^{\sigma^2} \left\{\exp(\sigma^2 \rho_{ij}) - 1\right\} \tag{52}$$

$$x_i = \alpha \exp(\sigma^2/2) \tag{53}$$

3.4. Exact Distribution of V When X is Chi-Square Distributed.

Another useful exact formula is that for the p.d.f. of V when X is chi-square distributed. In such a case, the chi-square p.d.f. given by Eq. (11) can be combined with Eqs. (27) and (28) and the resulting integral evaluated exactly. The inverse transform can be determined from no. 581.1 of Ref. 12, the result being that the probability density function of V is

$$P(V) = \frac{1}{(N-1)!}\left[1 + \frac{N\bar{x}}{K}\right]^{-K} V^{N-1}\, e^{-V} \cdot {}_1F_1\left[K,\, N,\, \frac{V}{1 + K/N\bar{x}}\right] \quad , \quad V \geq 0 \tag{54}$$

where ${}_1F_1$ is a confluent hypergeometric function (see Refs. 1 and 13). Despite its formidable appearance, Eq. (54) can be easily used for purposes of calculation since many standard handbooks (e.g., Ref. 13) give series expansions of ${}_1F_1$ (see Ref. 14 for the case K = 1/2).

3.5. General Principles Leading to Simple Interpolation Formulas.

From the basic results expressed above, it is possible to state some general principles which lead to methods for calculating results in many specific cases which are exact, or are good to a very accurate approximation, but are much simpler than the evaluation of the integrals appearing in Eqs. (26) or (27).

These principles are predicated on the assumption α_2 = constant in Figure 7-1.

3.5.1. First General Principle.

For any type of fluctuation, the detection probability results are exactly the same as for a slowly fluctuating target in which the individual samples have the same probability distribution as that of X/3.

3.5.2. Second Principle--Method of Obtaining Results for Block Correlations from Results for Rapid Fluctuations.

The next principle, derivable from the first, applies to any set of fluctuations and enables one to determine detection probability for block-correlated fluctuations if the results are known for rapid fluctuations.

For any distribution of the individual samples x_i, and for any given detection probability P_D, false alarm number n, number of integrated samples N, and number of independent groups of samples F, let

$\bar{x}(N,F)$ = average per-sample SNR required to achieve P_D

$L_{INT}(N,F)$ = incoherent integration loss for N integrated samples, for a slowly fluctuating target whose fluctuation distribution is the distribution of the average of F independent samples of the original distribution of x_i.

I.e., it is, $L_{INT}(N,F)$ is the incoherent integration loss for integration of N samples when the target fluctuation model is fixed and corresponds to X having the same distribution as the sum of F independent samples from the original distribution (except for the scale factor N/F).

Both $\bar{x}$ and L_{INT} also depend on P_D and n, but this dependence has been omitted for notational convenience.

Then, expressing all terms in db, we have exactly

$$\bar{x}(N,F) = \bar{x}(F,F) - \frac{N}{F}\ (db) + L_{INT}(N,F) - L_{INT}(F,F) \tag{55}$$

In Eq. (55), N/F (db) means $10 \log_{10}(N/F)$. The integers F and N satisfy the relation $1 \leq F \leq N$.

Now, $\bar{x}(F,F)$ is the required per-sample SNR for a rapidly fluctuating target with F integrated pulses. The quantities L_{INT} can be determined exactly from the results for a slowly fluctuating target of the type indicated in the definition of L_{INT}.

For all the families considered in Section 2., it has been found that, to high accuracy (order of .5 db or better), the integration loss L_{INT} is approximately equal to that for non-fluctuating targets (Ref. 1, p. 225).

Thus, to a good approximation,

$$\bar{x}(N,F) = \bar{x}(F,F) - \frac{N}{F}\ (db) + L_{INT}(N) - L_{INT}(F) \tag{55a}$$

where $L_{INT}(N)$ is the non-fluctuating integration loss for integration of N samples, and $L_{INT}(F)$ is the non-fluctuating integration loss for integration of F samples.

When applying this to completely different types of fluctuations than those which have been considered one would have to re-check the replacement of L_{INT} by that for non-fluctuating targets.

However, it should also be noted that the accuracy of the results based on Eq. (55a) is increased since L_{INT} enters only as the difference between L_{INT} at two values of N, and this is even more accurately represented by the corresponding difference for non-fluctuating targets. In most cases that have been checked, the error thus introduced is actually of the order one- to two-tenths of a db.

3.5.3. An Interpolation Formula for Chi-Square Distributed Fluctuations.

Another formula derivable from the first principle stated above, which is of great use in deriving results for the chi-square family, can be stated as follows: for any given detection probability P_D and false alarm number n, let

$\bar{x}(N,K)$ = average per-pulse SNR required when the fluctuation distribution of X is chi-square with parameter K

$L_{INT}(N,K)$ = incoherent integration loss for integration of N samples, when the distribution of X is chi-square with parameter K

($\bar{x}$ and L_{INT} also depend on P_D and n.)

Then, for any two values of N, say N and N' (either of which may be the larger)

$$\bar{x}(N',K) = \bar{x}(N,K) - \frac{N'}{N}\ (db) + L_{INT}(N',K) - L_{INT}(N,K) \tag{56}$$

where all quantities are expressed in db. Equation (56) is exact.

For the chi-square family, it can be verified that the difference of integration losses is equal to that for a non-fluctuating target within one- to two-tenths of a db, so that we can write to a high degree of approximation

$$\bar{x}(N',K) = \bar{x}(N,K) - \frac{N'}{N}\ (db) + L_{INT}(N') - L_{INT}(N) \tag{56a}$$

where $L_{INT}(N)$ = incoherent integration loss for non-fluctuating targets.

4. SPECIFIC DETECTION PROBABILITY RESULTS.

4.1. Detection Probability for the Chi-Square Family.

4.1.1. Statement of Results.

Figures 2, 3, and 4 present probability of detection results for cases when X is chi-square distributed with $1/2 \leq K \leq \infty$, N = 1, 10, 100, and $n = 10^{10}$. The results can easily be extended to other values of N and n (see below). The right-hand vertical scale is K; the horizontal scale is average per-pulse SNR $\bar{x} = \bar{X}/N$ in db.

These results can also be utilized whenever X has a distribution which can be well-approximated by a chi-square distribution. For example,

(a) When the individual samples x_i are chi-square distributed, for any type of correlation, X is approximately chi-square with K given by Eqs. (13), (7), and (8).

(b) When the individual x_i have Rice distributions, X will also be approximately chi-square distributed with K given by Eqs. (13), (7), and (8). The means and variances of x_i are given by Eqs. (17) and (18). An equivalent result is obtained by approximating the x_i distributions by chi-square distributions with k given by Eq. (19), and henceforth treating them as chi-square distributions.

(c) Section 2.1.2. states cases in which X is exactly chi-square distributed.

4.1.2. Method of Obtaining Results and Estimated Accuracy.

The results stated in Figures 7-2 to 7-4 for $1 \leq K \leq \infty$ were obtained very easily from previously published results for the non-fluctuating and Swerling cases (Ref. 1), aided by the interpolation formula Eq. (56a). Also refer to Section 2.1.2. and Section 2.1.3. The points for K = 1/2 are the only ones for which new computations were required.

(a) For N = 1:

The point K = 1 is Swerling Case I

The point K = 2 is Swerling Case III

The point K = 1000 is within .1 db of non-fluctuating

The points K = 10 and K = 20 were obtained by first determining $\bar{x}(10,10)$ and $\bar{x}(10,20)$ from the published results for Swerling Cases II and IV for N = 10, and then utilizing Eq. (56a) with N' = 1.

(b) For N = 10:

K = 1 is Swerling Case I

K = 2 is Swerling Case III

K = 10 is Swerling Case II

K = 20 is Swerling Case IV

K = 1000 is within .1 db of non-fluctuating

The points K = 100 and K = 200 were obtained by determining $\bar{x}(100,100)$ and $\bar{x}(100,200)$ from published results for Swerling Cases II and IV, N = 100, and then applying Eq. (56a) with N' = 10.

(c) For N = 100:

K = 1 is Swerling Case I

K = 2 is Swerling Case III

K = 100 is Swerling Case II

K = 200 is Swerling Case IV

K = 1000 is written .1 db of non-fluctuating

The points K = 10 and K = 20 were obtained by determining $\bar{x}(10,10)$ and $\bar{x}(10,20)$ from Swerling Cases II and IV, N = 10, and then applying Eq. (56a) with N' = 100.

In all cases, the point K = 1/2 was determined by applying the exact formula Eq. (54) (see Ref. 14).

The results presented are estimated to be accurate to within .2 to .3 db. The error budget can be estimated as follows:

(a) The points for K = 1/2 were calculated from the exact formula, Eq. (54); however, a series approximation for ${}_1F_1$ was used, which could introduce errors of .2 to .3 db.

(b) The points which are taken directly from published results for the Swerling and non-fluctuating cases have whatever accuracy those results have, taking into account also errors in plotting and reading the curves in Ref. 1.

(c) For those points where Eq. (56a) was used, the error introduced by using L_{INT} for non-fluctuating targets is estimated to be .1 to .2 db at most.

4.1.3. Extension to Other N and n.

To extend to other N, use Eq. (56a).

To extend to other values of n, the false alarm number, the same procedure as stated above can easily be used to determine results for $n = 10^4$, 10^8, 10^6 or any other value for which published results exist for Swerling Cases I-IV and non-fluctuating.

Actually, for the higher values of K (e.g., $K \geq 4$), the correction to different n is almost exactly the same as for non-fluctuating targets. For $1/2 \leq K \leq 1$, the correction is nearly that for $K = 1$ (Swerling I). For $2 \leq K \leq 4$, the correction is nearly the same as for Swerling III.

In fact, for all values of K, and for $P_D \geq .10$, a correction accurate to within less than .5 db can be obtained by subtracting 1.0 db from the given values of $\bar{x}$ ($n = 10^{10}$) to obtain results for $n = 10^8$, and subtracting 2.5 db from $\bar{x}$ ($n = 10^{10}$) to obtain results from $n = 10^6$.

4.1.4. Interpreting Results for Cases in Which Frequency Diversity is Employed.

Figures 7-2 to 7-4 are a convenient representation for picturing the gains due to frequency diversity. For example, suppose that without frequency diversity, X would be slowly fluctuating chi-square with parameter K. If F different frequencies were used, separated by sufficient increments, there would be F independent target samples so that according to the rules given in Section 2., X would then be chi-square with parameter KF. A glance at Figures 7-2 to 7-4 immediately tells how much in detection sensitivity is gained for any given detection probability.

4.1.5. Detection Probability for the Rice Family.

Good approximations to detection probability for the Rice family can be obtained most simply by approximating the Rice distributions of the individual samples by the appropriate chi-square distributions having the same means and seconds moments.

An illustration of the results is shown in Figure 7-5, where the right-hand vertical scale is detection probability. The solid curve is the exact detection probability for a Rice-distributed target with $s = 1$, i.e., with equal power in the predominant scatterer and the Rayleigh component, and for $N = 10$, $N = 10^6$, and slowly fluctuating. From Eq. (19), $k = 4/3$ for the best-fitting chi-square distribution and since the target is slowly fluctuating, $k = K$. The circled points are detection probability for a chi-square fluctuation with $K = 4/3$, $N = 10$, $n = 10^6$ (taken from Figure 7-3, Section 4., corrected to $n = 10^6$).

4.2. Detection Probability for the Log-Normal Family.

4.2.1. Statement of Results.

Heidbreder and Mitchell (Ref. 15) have presented comprehensive results for detection probability of targets when the individual samples have log-normal distributions and are either slowly or rapidly fluctuating. Figures 7-6 to 7-8 show a replot of their results for $n = 10^{10}$ in the slowly fluctuating case (i.e., in the case where X is also log-normal with the same value of σ, hence the same value of $\rho = \exp(\sigma^2/2)$, as x). Thus, the curves are labelled "detection probability when X is log-normally distributed." The dotted portions of the curves in Figure 7-6 indicate a range of uncertainty in reading and replotting the results from Ref. 15. However, extrapolation of the remaining portion of the curves, and comparison with those of Figures 7-7 and 7-8, makes it clear that the solid curves are the correct ones.

Incidentally, by comparing results for N = 1, 10, and 100 shown in Figures 7-6, 7-7, and 7-8, we can verify that the incoherent integration loss is nearly the same as for non-fluctuating targets.

Figure 7-9 is a replot of the results of Ref. 15 for the rapidly fluctuating case, $n = 10^{10}$, $\rho = 4$.

The reader is referred to Ref. 15 for other values of ρ in the rapidly fluctuating case and for other values of n.

These results can readily be extended to obtain detection probabilities for block-correlated log-normal cases with the aid of Eq. (55a).

Results are not yet developed for generally correlated log-normal models (they can be obtained by means of Eqs. (26) and (42)).

4.3. A Further Word on Approximation.

Two examples have been given where good approximate results can be obtained by approximating a fluctuation distribution by the first term in its LaGuerre Series expansion, i.e., by the chi-square distribution having the same first and second moments. These were: approximation of wide-sense by strict-sense chi-square cases, and approximation of Rice cases by chi-square cases.

However, it must be emphasized that this is <u>not</u> a generally applicable approximation method. Many examples can be given in which such an approximation would give extremely inaccurate results.

For instance, it is known (Refs. 9 and 14) that for certain ranges of the parameters, good to fair chi-square approximations may be found for log-normal distributions and correspondingly, the detection probability results approximate each other. However, the chi-square distributions which give the good approximations are <u>not</u> those having the same first and second moments as the log-normal distributions to which the approximations are being made. An attempt to base the approximations on moments would lead to large errors.

In addition, there are many cases in which no good approximations can be found within the chi-square family. Here too, examples are provided by the log-normal family, for other ranges of the parameters (see Refs. 9 and 14).

5. RELATED TOPICS.

5.1. CFAR Processing and Target Models.

The foregoing results and techniques can be applied to SNR fluctuations regardless of their physical source (target RCS fluctuations or other), provided they are due to fluctuations of the signal and not of the average noise power. The reason is that so long as the SNR fluctuations are due to signal power fluctuations, the relation between the false alarm rate and the detection threshold is not affected; however, if mean noise power also fluctuates, the detection threshold necessary to achieve a given false alarm probability is affected. In fact, such a situation generally leads to the use of adaptive thresholds to achieve constant false alarm rate (CFAR) at the least possible cost in detection sensitivity.

A large number of CFAR methods have been devised and their performance evaluated; unfortunately the bulk of work accomplished is still in the form of reports which have not appeared in "regular" journals.

Here our intention is merely to cite some examples of the effect of target fluctuations on CFAR detection. This effect depends on what type of CFAR method is being used. To illustrate:

A common CFAR technique, sometimes termed "cell averaging," involves estimation of average noise level by averaging received power in a set of resolution cells known or assumed to contain noise-only of the same average power as the noise in the cell being examined for presence of a target. Analysis of the detection performance of such methods for various fluctuating target models shows that the relative performance against different kinds of fluctuating (or non-fluctuating) targets is roughly the same as where average noise power is known a priori, although the absolute performance for all target models shows a decreased sensitivity, the magnitude of which depends on the number of noise cells available for the estimation of average noise power.

Another category of CFAR tests, which might be termed "single-cell" tests, involves essentially tests for whether a sequence of envelope samples arises from a Rayleigh distribution with statistical independence of samples. The effects of the target fluctuation model may be quite different for these tests than for the cell-averaging tests; also the optimum design of the tests themselves depends sensitively on the class of target models against which the radar is assumed to operate. More extensive analysis may be found in Ref. 16.

5.2. Effects of Frequency Diversity.

Use of frequency diversity together with pulse integration to smooth out fluctuations is a technique of great potential benefit in modern radar, both for detection and tracking. Thus, it becomes of interest to postulate fluctuation models with respect to the RCS frequency signature. Among the important questions that may be asked for particular target populations are:

(a) Does frequency diversity always cause fluctuations of RCS for targets which fluctuate with respect to aspect?

(b) When both frequency and aspect fluctuations occur, can it always be consistently postulated that the samples come from the same statistical distribution as when the fluctuations are due to aspect changes?

(c) In general, how should one postulate models of the p.d.f. and correlation properties of target fluctuations with respect to frequency, given the information that the fluctuations with respect to aspect have specified statistical models?

Some illustrations will suffice to show that the answers to questions (a) and (b) may be negative, and that the answer to question (c) may not always be obvious.

An example of a type of target which will fluctuate with respect to aspect changes but will be non-fluctuating with respect to frequency changes is provided by a prolate or oblate spheroid satisfying the condition that for all frequencies under consideration, the radii of curvature are much greater than the wavelength. The RCS of such a target at any aspect is equal to $\pi r_1 r_2$, where r_1 and r_2 are the radii of curvature at the point on the target such that the tangent plane is normal to the line of sight to the radar. Consequently, there will be no fluctuation with frequency (so long as $\lambda \ll r_1$ and $\lambda \ll r_2$). On the other hand, $r_1 r_2$ changes as aspect changes, so that there will be RCS fluctuations with aspect.

A converse example is that of a sphere in the resonance region ($\lambda \approx 2\pi a$); here, there will be RCS fluctuations with respect to frequency but not with respect to aspect.

Turning to the major categories of statistical target model:

(a) Rayleigh and Rice. A target which has Rayleigh fluctuations with aspect is consistent with a many-scatterer mechanism which will also fluctuate with respect to frequency, with the samples of RCS also coming from an exponential p.d.f., in fact from the same p.d.f. as the aspect fluctuations.

The correlation properties with respect to frequency will of course be calculated or modelled differently than those with respect to aspect.

For Rice targets, assuming the predominant scatterer is non-fluctuating with frequency (a consistent assumption although not a necessarily true one), the RCS fluctuations with frequency can also be consistently postulated to be drawn from the same Rice distribution as the fluctuations with respect to aspect (i.e., they will have the same p.d.f.).

(b) Log-Normal. If a target is observed to have log-normal fluctuation statistics with respect to aspect changes, it would not be clear what statistics could consistently be postulated for fluctuations with respect to frequency unless a physical mechanism for log-normal statistics is known.

Although more work remains to be done, the directive scatterer model for log-normal statistics indicates that it is at least consistent to postulate models for which questions (a) and (b) stated above can be answered affirmatively. This is because of the duality between small frequency changes and small aspect changes on the sidelobes of directive scatterers. The directive scatterer model also suggests the method of postulating correlation properties with respect to frequency.

5.3. Models for Range-Extended Targets.

With increasing range resolution capabilities of modern radars, it becomes of interest to conduct detection analyses for targets whose electromagnetic range extent exceeds the range resolution cell width. Such analyses are aimed not only at predicting performance for specific designs but also at questions such as: (a) how small should the resolution cell be made to optimize detection performance, or (b) how should signals in contiguous range cells be processed to optimize detection performance? These questions become especially pertinent in clutter-dominant situations, where decreasing the resolution cell size decreases the average clutter power.

In performing detection analyses in such cases it becomes clear that a generalized type of target model is required. Specifically, the target model must describe the statistical fluctuation properties of those portions of the target in each resolution cell, and how such fluctuation properties change as th cell size is varied.

Since actual data, on which such models can be based, are rather scarce, it is of interest to consider the question of postulating such models in such a way as to be consistent. Illustrations follow.

(a) Rayleigh Targets. Suppose it is known or assumed that a certain population of targets have Rayleigh fluctuation statistics when the resolution cell size equals or exceeds the target electromagnetic extent. In this case, it is simple to postulate consistent models for smaller resolution cells. Rayleigh statistics for the target as a whole are consistent with the following models (for example):

(i) A model in which the portion of the target in each resolution cell is also a Rayleigh target (i.e., a many-scatterer mechanism).

(ii) A model in which the portion of the target in some or all cells are Rice or chi-square targets when the cells become small.

(iii) A model in which, for sufficiently small cell size, the portion of the target in each cell is non-fluctuating.

In any of these cases, it would be consistent to postulate that the rate of fluctuation of the contents of each range cell is less than or equal to the rate of fluctuation for larger cells.

(b) Rice Targets. If the target as a whole is known to have Rice fluctuation statistics, consistent models could be postulated similar to those stated above, except that at least one cell would always contain a Rice target. Moreover, the ratio of predominant to Rayleigh components in that cell would increase with decreasing cell size.

(c) Log-Normal Targets. Suppose the whole target has log-normal RCS statistics. The question of how to model the target fluctuation properties when the resolution cell becomes smaller than the target in such cases requires further investigation. The directive scatterer model for log-normal statistics indicates an approach, but the problem remains to be solved.

6. REFERENCES.

1. Marcum, J. I., and P. Swerling, "Studies of Target Detection by Pulsed Radar," IRE Trans. Info. Theory, Vol. IT-6, No. 2, April 1960.

2. Swerling, P., "More on Detection of Fluctuating Targets," IEEE Trans. Info. Theory, Vol. IT-11, No. 3, July 1965.

3. Swerling, P., "Detection of Fluctuating Pulsed Signals in the Presence of Noise," IRE Trans. Info. Theory, Vol. IT-3, No. 3, September 1957.

4. Palmer, D. S., and D. C. Cooper, "An Analysis of the Performance of Weighted Integrators," IEEE Trans. Info. Theory, Vol. IT-10, No. 4, October 1964.

5. Di Franco, J. V., and W. L Rubin, "Radar Detection," Prentice Hall, 1968.

6. Mood, A. M., "Introduction to the Theory of Statistics," McGraw-Hill Book Co., New York, 1950.

7. Weinstock, W. W., "Target Cross Section Models for Radar Systems Analysis," Ph.D. Dissertation, Univ. of Pennsylvania, Philadelphia, Pa., 1956.

8. Scholefield, P. H. R., "Statistical Aspects of Ideal Radar Targets," IEEE Proceedings, Vol. 55, No. 4, April 1967.

9. Swerling, P., "Lecture Notes on Radar Target Signatures: Measurements, Statistical Models, and Systems Analysis," Technology Service Corporation, Santa Monica, Calif., August 1968.

10. Mitchell, R. L., "Radar Cross Sections of Randomly Oriented Discs and Rods," Proj. Memo. 014-10, Technology Service Corporation, Santa Monica, Calif., November 26, 1968.

11. Pollon, G. E., "Statistical Parameters for Scattering from Randomly Oriented Arrays, Plates, and Cylinders," IEEE Trans. Antennas and Propagation, February 1970.

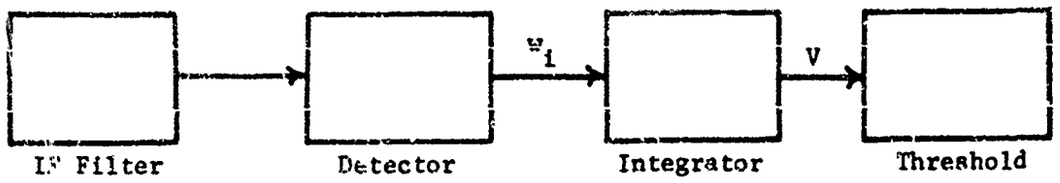
12. Campbell, G. A., and R. M. Foster, "Fourier Integrals for Practical Applications," D. Van Nostrand, New York, 1947.

13. Abramowitz, M., and I. A. Stegun, "Handbook of Mathematical Functions," Applied Math. Series 55, National Bureau of Standards, Washington, D. C., June 1954.

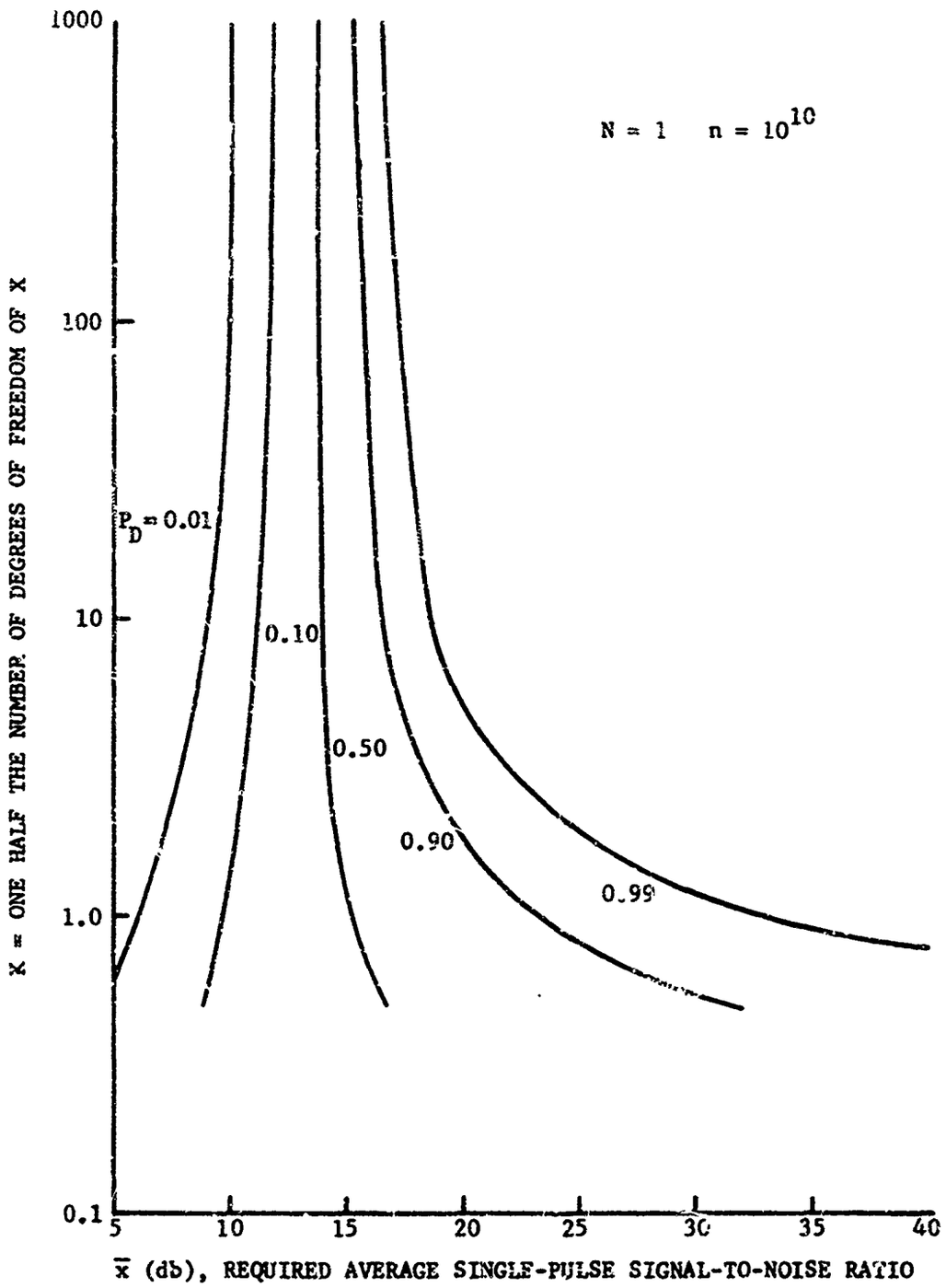
14. Swerling, P., "Probability of Detection for Some Additional Fluctuating Target Cases," Report No. TOR-669(9990)-14, Aerospace Corporation, Los Angeles, Calif., March 1966.

15. Heidbreder, G. R., and R. L. Mitchell, "Detection Probabilities for Log-Normally Distributed Signals," IEEE Trans. Aerospace and Electronic Systems, Vol. AES-3, No. 1, January 1967; also Report No. TR-669 (9990)-6, Aerospace Corporation, Los Angeles, Calif., April 1966.

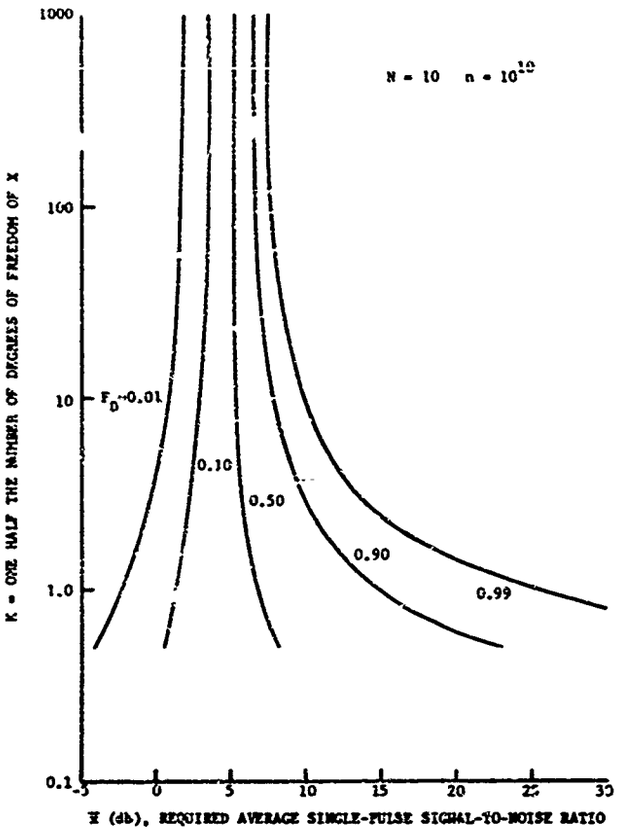
16. Swerling, P., L. E. Brennan, and I. S. Reed, "Optimum CFAR Detection Tests for Fluctuating and Non-Fluctuating Signals," TSC-PD-005-6, Technology Service Corporation, Santa Monica, Calif., Prepared for Lincoln Laboratory, M.I.T., under Purchase Order C-673, Prime Contract AF 19(628)-5167), December 12, 1967.



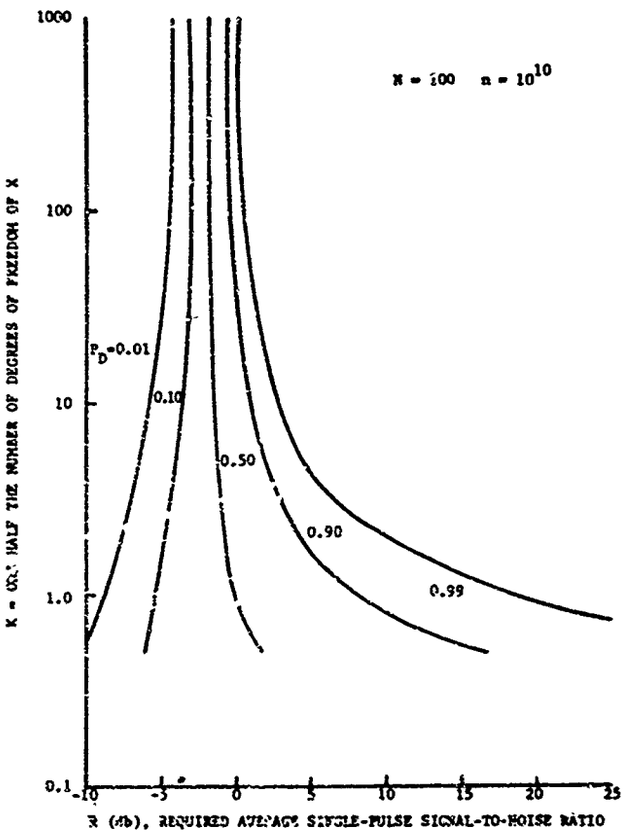
7-1. Receiver Model



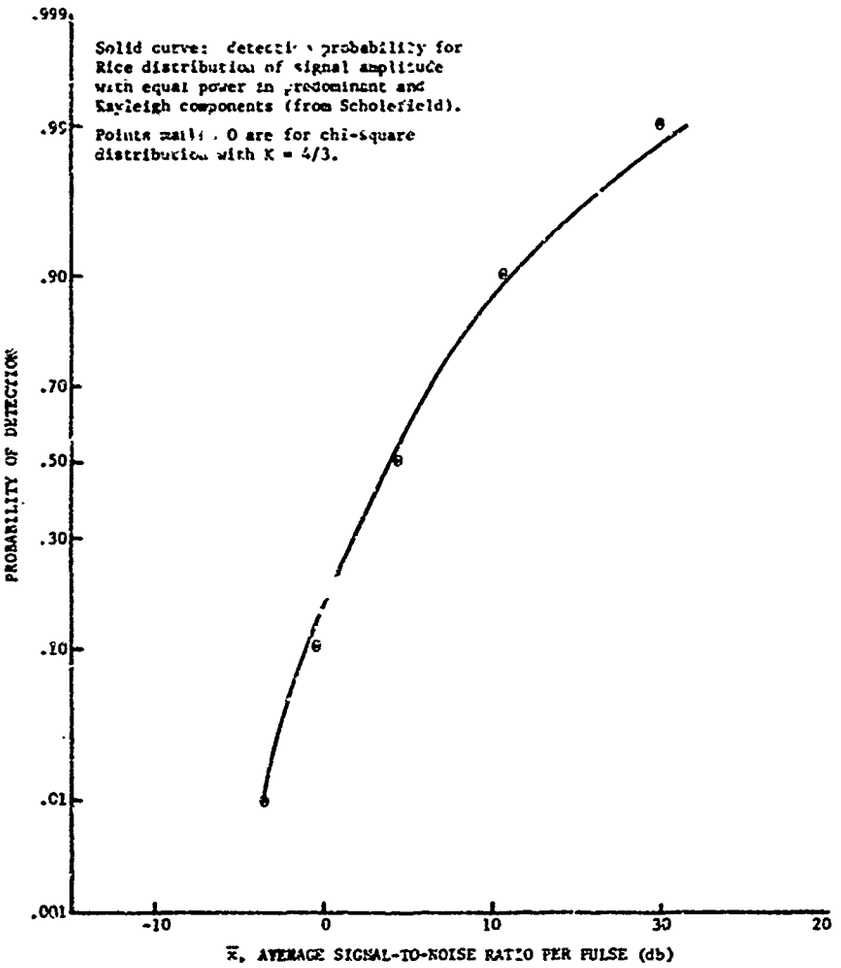
7-2. Detection Probability When K is Chi-Square (N = 1, n = 10^{10})



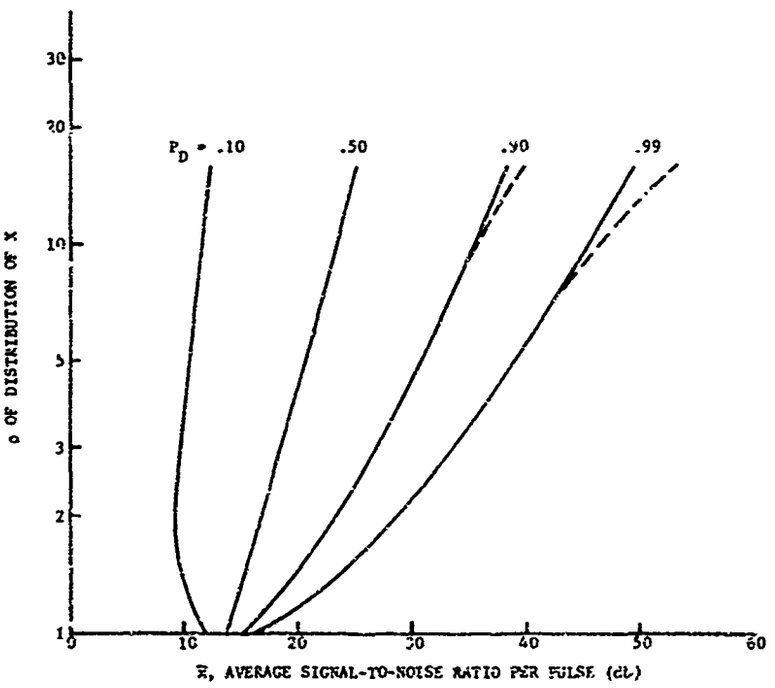
7-3. Detection Probability When X is Chi-Square (N = 10, n = 10^{10})



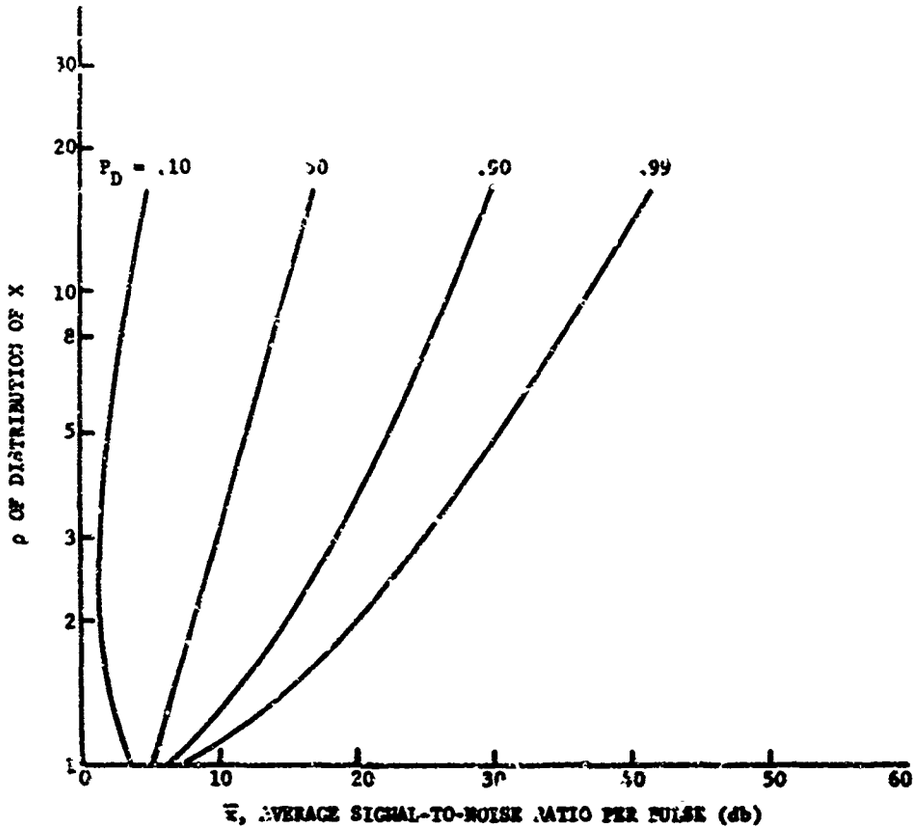
7-4. Detection Probability When X is Chi-Square (N = 100, n = 10^{10})



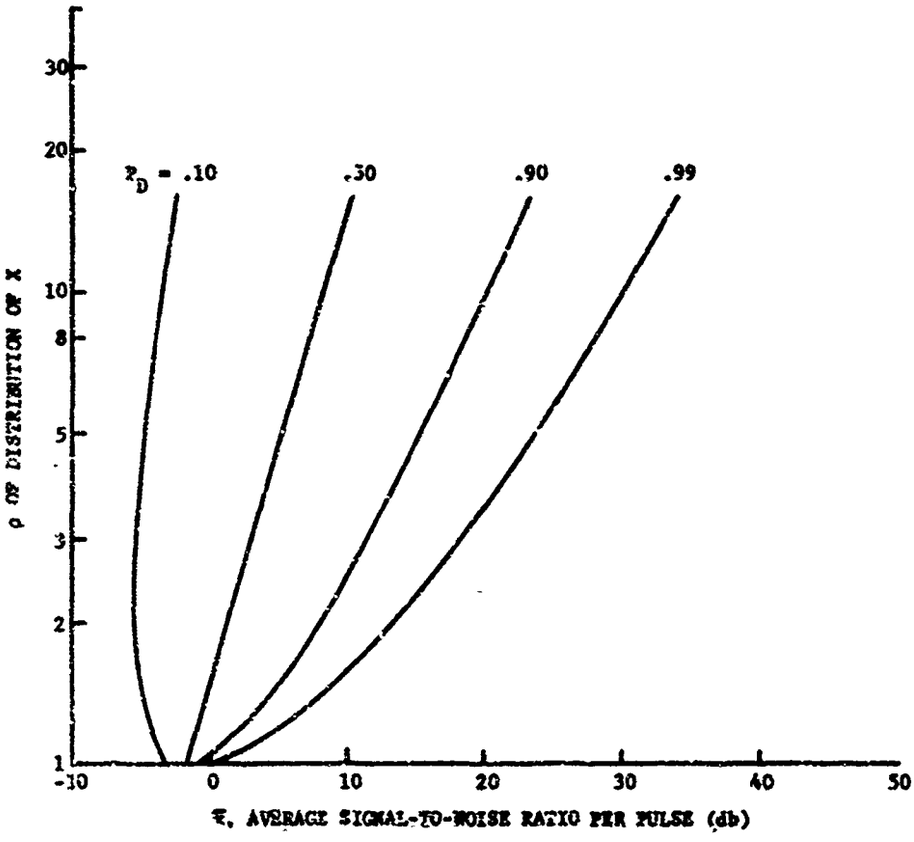
7-5. Detection Probability for Rice Distribution (n = 10, n = 10^6, Slow Fluctuation)



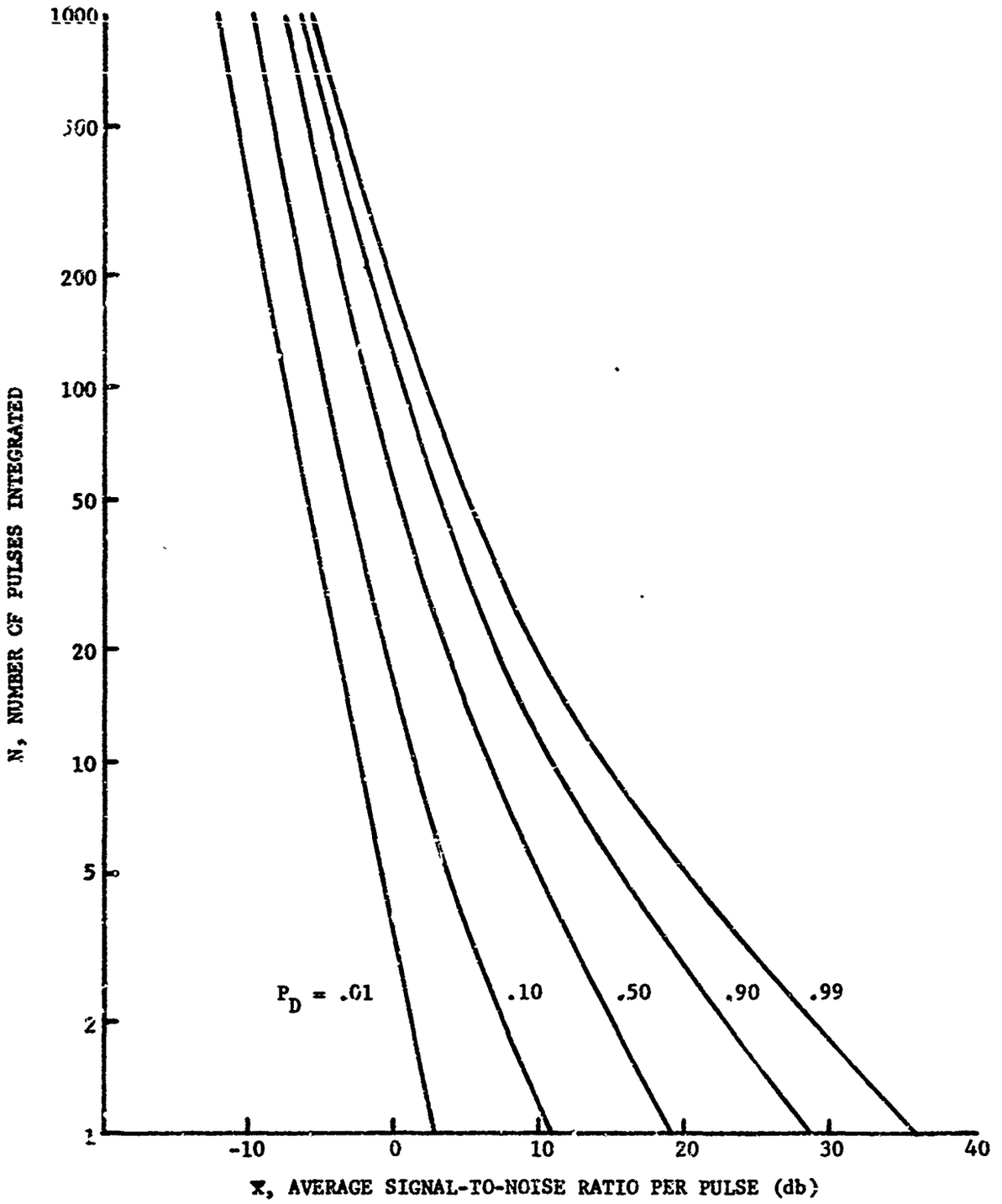
7-6. Detection Probability When X is Log-Normally Distributed (N = 1, n = 10^{10})



7-7. Detection Probability When X is Log-Normally Distributed (N = 10, n = 10^{10})



7-8. Detection Probability When X is Log-Normally Distributed (N = 100, n = 10^{10})



7-9. Detection Probability for Rapidly Fluctuating Log-Normal Target ($\sigma = 4$, $n = 10^{10}$)

THE THEORY OF AMPLITUDE COMPARISON MONOPULSE RADAR IN A CLUTTERED ENVIRONMENT

by

E. M. Hofstetter

M. I. T. Lincoln Laboratory
Lexington, Massachusetts, U. S. A.

THE THEORY OF AMPLITUDE COMPARISON MONOPULSE RADAR IN A CLUTTERED ENVIRONMENT *

E. M. Hofstetter
M. I. T. Lincoln Laboratory
Lexington, Massachusetts
U. S. A.

ABSTRACT

This paper considers the problem of detecting and estimating the unknown angular location of a target that is observed simultaneously by a number of antenna beams. The amplitude of the signal received by a particular antenna is assumed to depend on the angular location of the target, but the time of arrival of the signal is assumed to be the same at all of the antennas. The received signals are assumed to be corrupted by a combination of receiver front-end noise and clutter. The covariance matrix of the clutter signal appearing on the various beams either may be assumed known from previous measurements or else assumed to be an unknown parameter that must be estimated along with the angular coordinates of the target. In the latter case, multiple coherent or incoherent observations of the target-clutter complex must be available for processing at the receiver.

The generalized likelihood ratio test is used to derive the detection and parameter estimation strategy for the radar receiver. Expressions for the detector and the angle estimates are derived in a number of important cases and used to demonstrate several interesting properties of the likelihood ratio receivers. Finally, the case where the beams are conventional sum and difference beams is considered. An approximation to the likelihood receiver is derived for this case and shown to have a much more easily realizable structure than the exact likelihood receiver.

1. INTRODUCTION.

Most previous theoretical studies of amplitude comparison monopulse systems have been based on the premise that the target returns in the various beams were corrupted only by thermal noise generated in the receiver's front-end [1], [2], [3], [4]. The usual mathematical embodiment of this premise is the assumption that the signals on the various beams have added to them independent white, Gaussian noises of equal power. Starting from this mathematical model it is possible to derive the generalized likelihood-ratio strategy for detecting the presence or absence of a target and for estimating its angular location in space. As indicated in the above cited references, the usual hardware implementations of amplitude comparison monopulse are, under certain assumptions, reasonable approximations to the maximum likelihood receiver.

The purpose of the present paper is to extend previously obtained results to the case where the target returns are masked by clutter returns as well as receiver noise. The basic assumption underlying the clutter model will be that there are a large number of independent clutter scatterers per radar resolution cell so that the clutter return approximates Gaussian noise. The case where only a few scatterers are present in each resolution cell will not be considered. The clutter noise power received in each beam as well as the correlation coefficients of the clutter noise from beam to beam will be left arbitrary. Stated in mathematical terms, the covariance matrix of the noise on the various beams will be left arbitrary in distinction to the receiver-noise-alone case where the beam covariance matrix is assumed to be diagonal with equal elements on the main diagonal.

Two important cases of the monopulse problem will be considered. One where true monopulse operation is desired and detection and angle estimation must make use of the results of only one transmission and the other where the coherent or incoherent returns from multiple transmissions are available for these purposes In the former case, the beam covariance matrix must be known (or otherwise have been estimated) in advance whereas the latter case allows the additional possibility of estimating the unknown beam covariance matrix along with other unknown parameters such as signal amplitude and carrier phase. Both cases will be considered in detail.

The analysis of the problems sketched above unfortunately cannot always be carried out as explicitly as in the case where receiver noise alone corrupts the target returns. The final results usually are in the form of computational algorithms which may or may not be practically realizable depending on the particular parameters of the specific problem at hand. Nevertheless, the performance of these algorithms can be evaluated by means of computer simulations and used as benchmarks against which to compare more easily realizable systems in order to determine to what extent the theoretically possible performance is being achieved. Work to accomplish this end is already in progress but, unfortunately, definitive results are not available at this time.

An important reference for the general problem at hand is Mosca and Minguzzi, [5]. These authors attack the problem of maximum likelihood angle estimation in the presence of receiver noise and clutter. Their analysis differs from the one presented in this paper in that they assume a less general form for the beam covariance matrix, they consider only coherent observations and they assume that the time

---

* This work was sponsored by the Department of the Army.

interval between observations is short enough so that the clutter is completely correlated from observation to observation. The present paper assumes that the clutter is independent from observation to observation, a situation that can be achieved either by using a sufficiently low PRF or by jumping the carrier frequency from transmission to transmission. The assumption of independent clutter returns also allows consideration of the case where the beam covariance matrix is unknown and must be estimated along with the other unknown parameters. Mosca and Minguzzi do not consider this case.

2. MATHEMATICAL FORMULATION OF THE PROBLEM

The complex envelope of the signal received by the i^{th} antenna beam when the signal s(t) is transmitted will be taken to be of the form

$$r_i(t) = Ae^{j\phi}G_i(\theta,\varphi)s(t) + n_i(t), \quad i=1,\ldots,m \tag{1}$$

where A denotes the unknown received signal amplitude, ϕ denotes the unknown carrier phase and $n_i(t)$ denotes the interference (receiver noise plus clutter) appearing on the i^{th} beam. Equation (1) assumes that the target is located at zero range and has zero velocity. This has been done for convenience; the formulation can be changed in an obvious way to handle the situation where the target is located at an arbitrary range and velocity assuming that the target's range and angular location do not change appreciably during the time the observations are being processed. The real function $G_i(\theta,\varphi)$ denotes the two-way, voltage beam pattern of the i^{th} beam measured with respect to the angular coordinates θ and φ. These coordinates may be thought of as azimuth and elevation angles, but any two other coordinates capable of defining a direction in space could be used. Thus, the term $AG_i(\theta,\varphi)$ denotes the amplitude of the target return received on the i^{th} beam when the angular coordinates of the target are θ and φ. The assumption that the G_i are real reflects the fact that the arrival times of the received waveform at the different antennas are assumed to be identical regardless of the angular location of the target. This is never exactly the case in a practical amplitude-comparison monopulse system, but careful antenna design can yield a system for which it is a reasonable assumption.

The analysis will be greatly simplified by assuming that the signals on the various beams are first passed through filters matched to the assumed target range and velocity before becoming available for further processing. Stated in mathematical language, the observables available at the receiver are assumed to be the complex voltages given by†

$$y_i = \int r_i(t)s^*(t)dt = Ae^{j\phi}G_i(\theta,\varphi) + n_i, \quad i=1,\ldots,m \tag{2}$$

where

$$n_i = \int n_i(t)s^*(t)dt, \quad i=1,\ldots,m \tag{3}$$

and it has been assumed that the transmitted signal is normalized so that $\int |s(t)|^2 dt = 1$. Equation (2) can be rewritten in an equivalent vector form as follows:

$$\underline{y} = Ae^{j\phi}\underline{G}(\theta,\varphi) + \underline{n} \tag{4}$$

where

$$\underline{y}^T = [y_1,\ldots,y_m], \quad \underline{G}^T(\theta,\varphi) = [G_1(\theta,\varphi),\ldots,G_m(\theta,\varphi)], \quad \underline{n}^T = [n_1,\ldots,n_m]$$

This vector equation will prove to be much more convenient to work with in the subsequent analysis than its equivalent form given by Equation (2).

The next set of assumptions that must be made in order to complete the mathematical model of the received signal concerns the statistical description of the complex interference vector $\underline{n}$. It will be assumed that $\underline{n}$ is a complex Gaussian random vector having zero mean and covariance matrices given by††

$$E(\underline{n}\,\underline{n}^*) = \Lambda, \quad E(\underline{n}\,\underline{n}^T) = 0 \tag{5}$$

where Λ is a real, symmetric, positive definite matrix. The second assumption in Equation (5) is made so that the statistical properties of $\underline{n}$ will be consistent with the fact that $\underline{n}$ was derived from the random processes $n_i(t)$, $i=1,\ldots,m$ via Equation (3). The processes $n_i(t)$ are complex envelopes and thus have the well-known property $E[n_i(t)n_k(t')] = 0$ (see [6], p. 156) from which the second equation in (5) follows immediately.

---

† The asterisk denotes complex conjugate of a scalar or conjugate transpose of a matrix.

†† The superscript T denotes matrix transpose.

The reason for assuming that Λ is a real matrix instead of allowing it to be a more general complex covariance matrix comes from the following argument. The interference is the sum of noise produced by clutter returns and receiver front-end noise. These two noise sources are statistically independent which means that Λ can be written in the form $\Lambda = \Lambda_c + \Lambda_r$ where Λ_r is the covariance matrix of the receiver noise. Since the receiver noise is independent from beam to beam, Λ_r is a diagonal matrix with (necessarily) real elements on the main diagonal. Thus, all that remains to be shown is that Λ_c is real.

It is reasonable to assume that the clutter return is the sum of the returns from a large number of independent scatterers distributed in angle. Mathematically this means that the clutter return vector $\underline{n}_c$ can be written in the form

$$\underline{n}_c = \sum_k a(\theta_k, \varphi_k)\underline{G}(\theta_k, \varphi_k) \tag{6}$$

where $a(\theta_k, \varphi_k)$ denotes the complex amplitude of the return from the k^{th} scatterer which is located at the angular coordinates θ_k, φ_k. The assumptions that the returns from different scatterers are statistically independent and that the phases of these returns are uniformly distributed from 0 to 2π radians lead to the conclusion that

$$E[a(\theta_i, \varphi_i)a^*(\theta_k, \varphi_k)] = \begin{cases} E|a(\theta_i, \varphi_i)|^2 = \sigma(\theta_i, \varphi_i), & i=k \\ 0, & i \neq k \end{cases} \tag{7}$$

where $\sigma(\theta_i, \varphi_i)$ is the average cross-section of the scatterer located at coordinates θ_i, φ_i. It now follows that the clutter covariance matrix Λ_c can be written in the form

$$\Lambda_c = E[\underline{n}\,\underline{n}^*] = \sum_i \sigma(\theta_i, \varphi_i)\underline{G}(\theta_i, \varphi_i)\underline{G}^T(\theta_i, \varphi_i) \tag{8}$$

Since $\underline{G}$ has already been assumed to be a real vector, it follows that Λ_c and thus Λ must also be real.

The probability distribution of the received beam voltages $\underline{y}$ can now be written down in the form[†]

$$p(\underline{y} \mid A, \psi, \theta, \varphi, \Lambda) = \frac{1}{\pi^m |\Lambda|} \exp[-(\underline{y} - Ae^{j\psi}\underline{G})^*\Lambda^{-1}(\underline{y} - Ae^{j\psi}\underline{G})] \tag{9}$$

More generally, the probability distribution for the sequence of observations $\underline{y}_1, \ldots, \underline{y}_n$ resulting from n transmissions can be written in the form

$$p(\underline{y}_1, \ldots, \underline{y}_n \mid A, \underline{\psi}, \theta, \varphi, \Lambda) = \frac{1}{\pi^{nm} |\Lambda|^n} \exp[-\sum_{k=1}^{n} (\underline{y}_k - Ae^{j\psi_k}\underline{G})^*\Lambda^{-1}(\underline{y}_k - Ae^{j\psi_k}\underline{G})] \tag{10}$$

where $\underline{\psi}^T = [\psi_1, \ldots, \psi_n]$ and it has been assumed that the interference voltages are statistically independent and identically distributed from observation to observation.

Now that a statistical description of the observables has been given, the generalized likelihood ratio test can be used to derive a strategy for determining the presence or absence of a target and for estimating the unknown parameters of the target if one is present. If the beam covariance matrix Λ is assumed to be known (presumably from previous measurements made in a target-free environment), then the generalized likelihood test requires the receiver to compute the quantity

$$L = \frac{\max_{A, \underline{\psi}, \theta, \varphi} p(\underline{y}_1, \ldots, \underline{y}_n \mid A, \underline{\psi}, \theta, \varphi, \Lambda)}{p(\underline{y}_1, \ldots, \underline{y}_n \mid A=0)} \tag{11}$$

$$= \max_{A, \underline{\psi}, \theta, \varphi} \exp \sum_{k=1}^{n} [-(\underline{y}_k - Ae^{j\psi_k}\underline{G})^*\Lambda^{-1}(\underline{y}_k - Ae^{j\psi_k}\underline{G}) + \underline{y}_k^*\Lambda^{-1}\underline{y}_k]$$

and to compare it with a preset threshold λ. If $L < \lambda$, the decision "target absent" is made. If $L \geq \lambda$, the decision "target present" is made and the values of $A, \underline{\psi}, \theta, \varphi$ that achieved the maximum in Equation (11) are used as estimates of these unknown parameters.

The maximization with respect to $\underline{\psi}$ that appears in Equation (11) must be performed in different ways depending on whether the case of coherent or incoherent target returns is to be considered. For incoherent returns, the maximization must be carried out over all possible values of the ψ_k's, $0 \leq \psi_k \leq 2\pi$, $k=1, \ldots, m$ whereas for coherent returns, the ψ_k's are all set equal to a common value ψ and the maximization is carried out over all possible values of ψ.

---

† The functional dependence of $\underline{G}$ on θ and φ will not always be indicated explicitly in the sequel. The symbol $|\,|$ denotes the determinant of a 2×2 or larger matrix or the absolute magnitude of a scalar.

The generalized likelihood-ratio test for the case where Λ is unknown has the same general structure just described with the exception that the quantity L is to be calculated from the formula

$$L = \frac{\max\limits_{A,\underline{\psi},\theta,\varphi,\Lambda} \left\{ |\Lambda|^{-n} \exp \sum_{k=1}^{n} - (\underline{y}_k - Ae^{j\psi_k}\underline{G})^{*}\Lambda^{-1}(\underline{y}_k - Ae^{j\psi_k}\underline{G}) \right\}}{\max\limits_{\Lambda} \left\{ |\Lambda|^{-n} \exp \sum_{k=1}^{n} - \underline{y}_k^{*}\Lambda^{-1}\underline{y}_k \right\}} \tag{12}$$

where the maximization with respect to Λ is to be carried out over all real, symmetric, positive definite matrices. The maximization with respect to $\underline{\psi}$ is to be carried out in the manner described above.

The next sections of this paper will carry out the calculations required by Equations (11) and (12) as explicitly as possible. The case of known Λ will be treated first because it is mathematically much simpler than the case where Λ is unknown.

3. DERIVATION OF THE RECEIVERS

3.1. Known Beam Covariance Matrix.

Equation (11) can be rewritten in the form

$$\frac{1}{n}\ln L = \max_{A,\underline{\psi},\theta,\varphi} \left\{ 2A\underline{m}^{T}\Lambda^{-1}\underline{G} - A^{2}\underline{G}^{T}\Lambda^{-1}\underline{G} \right\} \tag{13}$$

where

$$\underline{m} = \mathrm{Re}\left[\frac{1}{n} \sum_{k=1}^{n} \underline{y}_k e^{-j\psi_k} \right] \tag{14}$$

The maximization on A will be carried out by differentiating the bracketed term in Equation (13) with respect to A and setting the result equal to zero. This yields the following expression for $\hat{A}$, the maximizing value of A,

$$\hat{A} = \frac{\underline{m}^{T}\Lambda^{-1}\underline{G}}{\underline{G}^{T}\Lambda^{-1}\underline{G}} \tag{15}$$

Substitution of Equation (15) into Equation (13) now yields

$$\frac{1}{n}\ln L = \max_{\underline{\psi},\theta,\varphi} \left\{ \frac{(\underline{m}^{T}\Lambda^{-1}\underline{G})^{2}}{\underline{G}^{T}\Lambda^{-1}\underline{G}} \right\} \tag{16}$$

The maximization on $\underline{\psi}$ can be carried out by writing (16) in the form

$$\frac{1}{n}\ln L = \max_{\underline{\psi},\theta,\varphi} \left\{ \frac{(\frac{1}{n} \sum_{k=1}^{n} \mathrm{Re}\, \underline{y}_k^{*}\Lambda^{-1}\underline{G}e^{j\psi_k})^{2}}{\underline{G}^{T}\Lambda^{-1}\underline{G}} \right\} \tag{17}$$

In the incoherent case, the maximum in (17) is obviously achieved when $\psi_k = -\arg(\underline{y}_k^{*}\Lambda^{-1}\underline{G})$ and is given by

$$\frac{1}{n}\ln L = \max_{\theta,\varphi} \left\{ \frac{(\sum_{k=1}^{n} |\underline{y}_k^{*}\Lambda^{-1}\underline{G}|)^{2}}{\underline{G}^{T}\Lambda^{-1}\underline{G}} \right\} \tag{18}$$

The corresponding maximum for the coherent case is achieved when $\psi_k = \psi = -\arg \sum_{k=1}^{n} \underline{y}_k^{*}\Lambda^{-1}\underline{G}$, k=1,....m and is given by

$$\frac{1}{n}\ln L = \max_{\theta,\varphi} \left\{ \frac{|\sum_{k=1}^{n} \underline{y}_k^{*}\Lambda^{-1}\underline{G}|^{2}}{\underline{G}^{T}\Lambda^{-1}\underline{G}} \right\} \tag{19}$$

No way of explicitly performing the remaining maximization required by (18) or (19) has been found; however, additional insight into the nature of the angular location estimates can be obtained in certain important special cases.

The best way to obtain these results is to return to Equation (16). Application of the Schwarz inequality to the bracketed term in this equation yields the inequality

$$\frac{(\underline{m}^T\Lambda^{-1}\underline{G})^2}{\underline{G}^T\Lambda^{-1}\underline{G}} \le \underline{m}^T\Lambda^{-1}\underline{m} \tag{20}$$

and the information that equality is achieved in inequality (20) if, and only if, $\underline{G} = \alpha\underline{m}$ for some constant α. If this condition can be met, it follows that

$$\frac{1}{n}\ln L = \max_{\underline{\psi}} \underline{m}^T\Lambda^{-1}\underline{m} \tag{21}$$

and

$$\underline{\hat{G}}(\hat{\theta},\hat{\varphi}) = \alpha \mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n} y_k e^{-j\hat{\psi}_k}\right] \tag{22}$$

where the $\hat{\psi}_k$ are those ψ_k that achieve the maximum in (21). Equation (22) still does not give explicit estimates of θ and φ in general but it does yield the important result that $\hat{\theta}$ and $\hat{\varphi}$ depend on Λ only through the estimates $\hat{\psi}_k$ of the carrier phases. This information is useful in deriving practically realizable approximations to the likelihood-ratio receiver as will be done later.

The assumption made in arriving at Equations (21) and (22) is that, for any $\underline{m}$, a constant α and angles $\hat{\theta},\hat{\varphi}$ can be found such that $\underline{G}(\hat{\theta},\hat{\varphi}) = \alpha\underline{m}$. These conditions can only be met if the vector $\underline{G}(\theta,\varphi)$ assumes all possible directions in m-space as θ and φ vary through their respective ranges. This is obviously not possible if $m > 3$ but it is true if $m = 3$ and the beam gains have the property that the equations

$$\frac{G_1(\theta,\varphi)}{G_3(\theta,\varphi)} = x_1$$

$$\frac{G_2(\theta,\varphi)}{G_3(\theta,\varphi)} = x_2 \tag{23}$$

have a solution for all x_1 and x_2. (A similar statement holds if $m = 2$ and if the two beam gains are functions of only one angle.) This assumption is identical to the one previously made in reference [3], and will be referred to "assumption I" whenever it is needed in the sequel.

There is one case for which the maximizations required above can be explicitly performed and that is the case where there is only a single observation, $n = 1$. Equation (21) now can be manipulated as follows:

$$\begin{aligned}\frac{1}{n}\ln L &= \max_{\psi_1} \mathrm{Re}\left[y_1^* e^{j\psi_1}\right]\Lambda^{-1}\mathrm{Re}\left[y_1 e^{-j\psi_1}\right] = \max_{\psi_1}\mathrm{Re}\left[\underline{r}^* e^{j\psi_1}\right]\mathrm{Re}\left[\underline{r}e^{-j\psi_1}\right] \\ &= \tfrac{1}{2}\underline{r}^*\underline{r} + \max_{\psi_1}\tfrac{1}{2}\mathrm{Re}\left[\underline{r}^T\underline{r}e^{-j2\psi_1}\right] = \tfrac{1}{2}\underline{r}^*\underline{r} + \tfrac{1}{2}\left|\underline{r}^T\underline{r}\right| = \tfrac{1}{2}y_1^*\Lambda^{-1}y_1 + \tfrac{1}{2}\left|y_1^T\Lambda^{-1}y_1\right|\end{aligned} \tag{24}$$

where $\underline{r} = \Lambda^{-\frac{1}{2}}y_1$. The maximum occurs when ψ_1 assumes the value $\hat{\psi}_1$ given by

$$\hat{\psi}_1 = \tfrac{1}{2}\arg \underline{r}^T\underline{r} = \tfrac{1}{2}\arg y_1^T\Lambda^{-1}y_1 \tag{25}$$

In arriving at Equation (24), use was made of the easily verified identity

$$\mathrm{Re}\,\underline{u}^*\,\mathrm{Re}\,\underline{v} = \tfrac{1}{2}\mathrm{Re}\,\underline{u}^*\underline{v} + \tfrac{1}{2}\mathrm{Re}\,\underline{u}^T\underline{v} \tag{26}$$

which is valid for any complex vectors $\underline{u}$ and $\underline{v}$. The results given by Equations (24) and (25) are generalizations of the results given in [3] which assumed that the beam covariance matrix was a constant times the identity matrix.

The results of this section are summarized in Fig. 1. The notation I next to an equation means that assumption I was used in deriving the equation.

3.2. Unknown Beam Covariance Matrix.

The starting point for this analysis is Equation (12). Denoting the numerator and denominator of this equation by N and D respectively, it is seen that

$$\frac{1}{n}\ln L = \frac{1}{n}\ln N - \frac{1}{n}\ln D \tag{27}$$

where

$$\frac{1}{n}\ln N = \max_{A,\underline{\psi},\theta,\varphi,\Lambda} \left\{ -\frac{1}{n}\sum_{k=1}^{n} (\underline{y}_k - Ae^{j\psi_k}\underline{G})^*\Lambda^{-1}(\underline{y}_k - Ae^{j\psi_k}\underline{G}) - \ln|\Lambda| \right\} \tag{28}$$

and

$$\frac{1}{n}\ln D = \max_{\Lambda} \left\{ -\frac{1}{n}\sum_{k=1}^{n} \underline{y}_k^*\Lambda^{-1}\underline{y}_k - \ln|\Lambda| \right\} \tag{29}$$

Introducing the notation $Q = \Lambda^{-1}$ and

$$C = \frac{1}{n}\mathrm{Re}\left[\sum_{k=1}^{n} (\underline{y}_k - Ae^{j\psi_k}\underline{G})(\underline{y}_k - Ae^{j\psi_k}\underline{G})^*\right] \tag{30}$$

Equation (28) can be written in the form

$$\frac{1}{n}\ln N = \max_{A,\underline{\psi},\theta,\varphi,Q} \{-\mathrm{tr}CQ + \ln|Q|\} \tag{31}$$

where the maximization on Q is to be taken with respect to all real, positive definite symmetric matrices. The term in brackets appearing in Equation (31) can be maximized with respect to Q by taking its variation on Q and setting it equal to zero with result

$$\mathrm{tr}(Q^{-1} - C)\delta Q = 0 \tag{32}$$

where use has been made of the formula

$$\delta\ln|Q| = \mathrm{tr}Q^{-1}\delta Q \tag{33}$$

A derivation of Equation (33) can be found in reference [7] on page 35. Since Equation (32) must be valid for all sufficiently small real symmetric δQ, it follows that $\hat{Q}$, the maximizing value of Q, is given by $\hat{Q} = C^{-1}$ or equivalently, $\hat{\Lambda} = C$. Substitution of this result in Equation (31) yields the result,

$$\frac{1}{n}\ln N = \max_{A,\underline{\psi},\theta,\varphi} \{-m - \ln|C|\} \tag{34}$$

Setting A=0, in Equation (34) also yields

$$\frac{1}{n}\ln D = \{-m - \ln|R|\} \tag{35}$$

where

$$R = \mathrm{Re}\,\frac{1}{n}\sum_{k=1}^{n} \underline{y}_k\underline{y}_k^* \tag{36}$$

from which it follows that

$$\frac{1}{n}\ln L = \max_{A,\underline{\psi},\theta,\varphi} \{\ln|R| - \ln|C|\} \tag{37}$$

It is possible to perform the maximization on A required by Equation (37) explicitly; however, it seems that additional assumptions are necessary in order to proceed further than this point. For this reason, assumption 1 will be invoked now. This assumption states that the vector $\underline{G}(\theta,\varphi)$ can assume any orientation in m-space and this means that the vector $\underline{a} = A\underline{G}(\theta,\varphi)$ can be set equal to an arbitrary m-vector by suitable choices of the parameters A, θ and φ. In other words, the maximization on A, θ and φ required by Equation (37) is equivalent to maximization with respect to $\underline{a}$. Thus, Equation (37) can be written in the form,

$$\frac{1}{n}\ln L = \max_{\underline{\psi},\underline{a}} \left\{\ln|R| - \ln\left|\mathrm{Re}\,\frac{1}{n}\sum_{k=1}^{n} (\underline{y}_k - e^{j\psi_k}\underline{a})(\underline{y}_k - e^{j\psi_k}\underline{a})^*\right|\right\} \tag{38}$$

Equation (38) can be rewritten as follows,

$$\frac{1}{n}\ln L = \max_{\underline{\psi},\underline{a}} \left\{ \ln|R| - \ln|R - \underline{m}\underline{m}^T + (\underline{m}-\underline{a})(\underline{m}-\underline{a})^T| \right\}$$

$$= \max_{\underline{\psi},\underline{a}} \left\{ \ln|R| - \ln|R-\underline{m}\underline{m}^T| - \ln|I + (R-\underline{m}\underline{m}^T)^{-1}(\underline{m}-\underline{a})(\underline{m}-\underline{a})^T| \right\} \tag{39}$$

where $\underline{m}$ was defined by Equation (14).

Next, making use of the following identity (reference [7], page 36) which is valid for an arbitrary matrix W and arbitrary vectors $\underline{u}$ and $\underline{v}$,

$$|W + \underline{u}\,\underline{v}^T| = |W|[1 + \underline{v}^T W^{-1}\underline{u}], \tag{40}$$

Equation (39) can be put in the form

$$\frac{1}{n}\ln L = \max_{\underline{\psi},\underline{a}} \left\{ -\ln|I - R^{-1}\underline{m}\,\underline{m}^T| - \ln[1+(\underline{m}-\underline{a})^T(R-\underline{m}\underline{m}^T)^{-1}(\underline{m}-\underline{a})] \right\}$$

$$= \max_{\underline{\psi},\underline{a}} \left\{ -\ln[1-\underline{m}^T R^{-1}\underline{m}] - \ln[1+(\underline{m}-\underline{a})^T(R-\underline{m}\underline{m}^T)^{-1}(\underline{m}-\underline{a})] \right\} \tag{41}$$

The maximization on $\underline{a}$ required by Equation (41) is performed by setting the variation of the bracketed term with respect to $\underline{a}$ equal to zero. This results in the equation

$$\delta\underline{a}^T (R-\underline{m}\underline{m}^T)^{-1}(\underline{m}-\underline{a}) = 0 \tag{42}$$

which must be satisfied for all $\delta\underline{a}$. It follows that $\hat{\underline{a}}$, the maximizing value of $\underline{a}$, is given by $\hat{\underline{a}} = \underline{m}$ and Equation (41) becomes

$$\frac{1}{n}\ln L = \max_{\underline{\psi}} \left\{ -\ln(1-\underline{m}^T R^{-1}\underline{m}) \right\} = -\ln[1 - \max_{\underline{\psi}} \underline{m}^T R^{-1}\underline{m}] \tag{43}$$

No way of explicitly performing the maximization required by Equation (43) has been found; however, if the maximizing phases are denoted $\hat{\psi}_k$, then the amplitude and angle estimates are given by

$$\hat{A}\,\underline{G}(\hat{\theta},\hat{\varphi}) = \hat{\underline{a}} = \mathrm{Re}\,\frac{1}{n}\sum_{k=1}^{n} \underline{y}_k e^{-j\hat{\psi}_k} \tag{44}$$

Note that this equation has the same form as Equation (22) which was derived for the case of a known beam covariance matrix. The only difference is that now the phases are determined by Equation (43) instead of Equation (21).

Although the maximization appearing in Equation (43) cannot be performed explicitly, the problem can be transformed in a way that makes the resulting receiver structures look similar to the ones derived for the case of a known beam covariance matrix and given by Equations (18) and (19). The key to this transformation is the following general†

Lemma: If $\underline{v}(\underline{x})$ is a real, vector-valued function of a vector parameter $\underline{x}$ then

$$\max_{\underline{x}} \|\underline{v}(\underline{x})\| = \max_{\underline{e},\,\|\underline{e}\|=1} \left\{ \max_{\underline{x}} \underline{e}^T \underline{v}(\underline{x}) \right\}$$

and $\hat{\underline{e}}$, the maximizing value of $\underline{e}$ is given by

$$\hat{\underline{e}} = \frac{\underline{v}(\hat{\underline{x}})}{\|\underline{v}(\hat{\underline{x}})\|}$$

where $\hat{\underline{x}}$ is the maximizing value of $\underline{x}$.

Proof: The truth of this lemma is obvious on geometric grounds but a formal proof can be given as follows. The Schwarz inequality implies that,

---

† The author is grateful to J. R. Johnson, consultant to this Laboratory, for pointing out this lemma to him.

$$\underline{e}^T \underline{v}(\underline{x}) \leq \|\underline{v}(\underline{x})\| \tag{45}$$

for all $\underline{e}$ satisfying $\|\underline{e}\| = 1$ and for all $\underline{x}$. It follows that,

$$\max_{\underline{e},\ \|\underline{e}\|=1} \left\{ \max_{\underline{x}} \underline{e}^T \underline{v}(\underline{x}) \right\} \leq \max_{\underline{x}} \|\underline{v}(\underline{x})\|. \tag{46}$$

Next, it is obvious that,

$$\max_{\underline{x}} \|\underline{v}(\underline{x})\| = \|\underline{v}(\hat{\underline{x}})\| = \hat{\underline{e}}^T \underline{v}(\hat{\underline{x}}) \tag{47}$$

where $\hat{\underline{e}}$ is defined by $\hat{\underline{e}} = \underline{v}(\hat{\underline{x}})/\|\underline{v}(\hat{\underline{x}})\|$. But,

$$\hat{\underline{e}}^T \underline{v}(\hat{\underline{x}}) \leq \max_{\underline{x}} \hat{\underline{e}}^T \underline{v}(\underline{x}) \leq \max_{\underline{e},\ \|\underline{e}\|=1} \left\{ \max_{\underline{x}} \underline{e}^T \underline{v}(\underline{x}) \right\} \tag{48}$$

therefore,

$$\max_{\underline{x}} \|\underline{v}(\underline{x})\| \leq \max_{\underline{e},\ \|\underline{e}\|=1} \left\{ \max_{\underline{x}} \underline{e}^T \underline{v}(\underline{x}) \right\} \tag{49}$$

The truth of the lemma now follows from inequalities (46) and (49).

Applying this lemma to the problem of finding $\max_{\underline{\psi}} \underline{m}^T R^{-1} \underline{m} = \max_{\underline{\psi}} \|R^{-\frac{1}{2}} \underline{m}\|^2$ results in the equation

$$\max_{\underline{\psi}} \underline{m}^T R^{-1} \underline{m} = \max_{\underline{e},\ \|\underline{e}\|=1} \left\{ \max_{\underline{\psi}} (\underline{e}^T R^{-\frac{1}{2}} \underline{m})^2 \right\} \tag{50}$$

Now, it is easily seen that

$$\max_{\underline{\psi}} (\underline{e}^T R^{-\frac{1}{2}} \underline{m})^2 = \max_{\underline{\psi}} \left(\frac{1}{n} \sum_{k=1}^{n} \mathrm{Re}(\underline{e}^T R^{-\frac{1}{2}} \underline{y}_k e^{-j\psi_k}) \right)^2 = \left(\frac{1}{n} \sum_{k=1}^{n} |\underline{e}^T R^{-\frac{1}{2}} \underline{y}_k| \right)^2 \tag{51}$$

in the incoherent case. The maximizing phases are given by $\hat{\psi}_k = \arg \underline{e}^T R^{-\frac{1}{2}} \underline{y}_k$. Thus, for the incoherent case,

$$\max_{\underline{\psi}} \underline{m}^T R^{-1} \underline{m} = \max_{\underline{e},\ \|\underline{e}\|=1} \left(\frac{1}{n} \sum_{k=1}^{n} |\underline{e}^T R^{-\frac{1}{2}} \underline{y}_k| \right)^2 \tag{52}$$

The corresponding result for the coherent case is given by

$$\max_{\underline{\psi}} \underline{m}^T R^{-1} \underline{m} = \max_{\underline{e},\ \|\underline{e}\|=1} \left| \frac{1}{n} \sum_{k=1}^{n} \underline{e}^T R^{-\frac{1}{2}} \underline{y}_k \right|^2 \tag{53}$$

The maximizing vectors $\hat{\underline{e}}_i$ and $\hat{\underline{e}}_c$ for the incoherent and coherent cases are given by

$$\hat{\underline{e}}_i = \alpha_i R^{-\frac{1}{2}} \mathrm{Re} \frac{1}{n} \sum_{k=1}^{n} \underline{y}_k e^{-j\hat{\psi}_k} \tag{54}$$

and

$$\hat{\underline{e}}_c = \alpha_c R^{-\frac{1}{2}} \mathrm{Re} \frac{1}{n} \sum_{k=1}^{n} \underline{y}_k e^{-j\hat{\psi}} \tag{55}$$

respectively. The constants α_i and α_c appearing in these equations are picked to make $\|\hat{\underline{e}}_i\| = 1$ and $\|\hat{\underline{e}}_c\| = 1$ respectively.

Comparison of Equation (54) or (55) with Equation (44) shows that the amplitude estimate $\hat{A}$ and the angle estimates $\hat{\theta}$ and $\hat{\varphi}$ are given by the equation

$$\hat{A} \underline{G}(\hat{\theta}, \hat{\varphi}) = \alpha^{-1} R^{\frac{1}{2}} \hat{\underline{e}} \tag{56}$$

where the α, $\hat{\underline{e}}$ pair corresponding to either the incoherent or coherent case are to be used. The angle estimates and the detection logic given by Equations (52), (53) and (56) can be cast into a more illuminating form by recalling that because of assumption I, the vector $\underline{G}(\theta, \varphi)$ assumes all possible directions in 3-space (or 2-space if $\underline{G}$ only depends on a single angle) as θ and φ are varied through their ranges. This means that the vector $R^{-\frac{1}{2}} \underline{G}$ also assumes all possible directions in 3-space and, therefore, that the vector

$$R^{-\frac{1}{2}} \underline{G} / [\underline{G}^T R^{-1} \underline{G}]^{\frac{1}{2}}$$

can be set equal to an arbitrary 3-space unit vector. In other words, a vector $\underline{e}$ appearing in either

Equation (52) or (53) can be replaced by $R^{-\frac{1}{2}}\underline{G}/[\underline{G}^T R^{-1}\underline{G}]^{\frac{1}{2}}$ and the maximization performed on θ and φ instead of on $\underline{e}$. This leads to the following equations

$$\max_{\underline{\psi}} \underline{m}^T R^{-1}\underline{m} = \max_{\theta,\varphi} \frac{\left[\frac{1}{n}\sum_{k=1}^{n} |\underline{G}^T R^{-1} y_k|\right]^2}{\underline{G}^T R^{-1}\underline{G}} \tag{57}$$

and

$$\max_{\underline{\psi}} \underline{m}^T R^{-1}\underline{m} = \max_{\theta,\varphi} \frac{\left|\frac{1}{n}\sum_{k=1}^{n} \underline{G}^T R^{-1} y_k\right|^2}{\underline{G}^T R^{-1}\underline{G}} \tag{58}$$

pertaining to the incoherent and coherent cases respectively. Moreover, it now follows from Equation (56) and the fact that $\underline{e} = R^{-\frac{1}{2}}\underline{G}/[\underline{G}^T R^{-1}\underline{G}]^{\frac{1}{2}}$ that the angle estimates θ and φ are those values of θ and φ that achieve the maximum in Equation (57) or (58). Note that the form of Equations (57) and (58) is the same as that of Equation (18) and (19) which were derived for the case where the beam covariance matrix was known. The main difference between these sets of equations is that the known covariance matrix Λ appears in one set whereas the other set uses the matrix R which is derived from the observations via Equation (36). Another difference is that Equations (57) and (58) are only valid when assumption I is in force: Equations (18) and (19) were derived without making this assumption.

The results of this section are summarized in Fig. 1 along with the results derived previously for the case where Λ was known. There are no unknown-Λ results when only a single observation is available because in this case there are more unknown parameters than observed quantities which means that estimation of these parameters is impossible.

4. SOME PROPERTIES OF THE LIKELIHOOD-RATIO RECEIVERS.

One interesting property of the receiver structures that just have been derived is that they are invariant to linear transformations of the beam patterns. To illustrate this notion, consider the case of a two-beam radar whose beam patterns are identical except that they are skewed in angle with respect to one another, i.e.

$$G_1(\theta) = G(\theta-\theta_0), \qquad G_2(\theta) = G(\theta+\theta_0) \tag{59}$$

Most monopulse systems do not work directly with beams G_1 and G_2 but rather with sum and difference beams

$$H_1(\theta) = G_1(\theta) + G_2(\theta), \quad H_2(\theta) = G_1(\theta) - G_2(\theta) \tag{60}$$

and process the signals from these beams in the receiver. It is natural to ask whether there is any difference in either structure or performance between the likelihood-ratio receiver based on the signals derived from G_1 and G_2 and the likelihood-ratio receiver based on the signals obtained from H_1 and H_2. The answer to this question is, no; both receivers have identical structures and perform equally well. This statement is a special case of the invariance principle referred to above and which now will be proved.

To fix ideas, consider the case when the beam covariance matrix is known and the observations are incoherent. The receiver structure pertinent to this case is given by Equation (18) and the angle estimates θ and φ are those values of θ and φ that achieve the maximum required by this equation. Suppose now, that instead of starting with the beam vector $\underline{G}$ and the corresponding received voltages y_k, the beam vector $\underline{H} = B\underline{G}$ and the corresponding received voltages $\underline{y}_k = By_k$, where B denotes an arbitrary, real, non-singular matrix, had been used. The beam covariance matrix for this case is given by $K = BAB^T$ and the corresponding likelihood-ratio receiver is defined by the equation

$$\frac{1}{n}\ln L = \max_{\theta,\varphi}\left\{\frac{\left(\frac{1}{n}\sum_{k=1}^{n} |\underline{H}^T K^{-1}\underline{y}_k|\right)^2}{\underline{H}^T K^{-1}\underline{H}}\right\} \tag{61}$$

Using the definitions of $\underline{H}$, K and $\underline{y}_k$, this equation can be rewritten in the form,

$$\frac{1}{n}\ln L = \max_{\theta,\varphi}\left\{\frac{\left(\frac{1}{n}\sum_{k=1}^{n} |\underline{G}^T B^T (B^T)^{-1}\Lambda^{-1}B^{-1}B\underline{y}_k|\right)^2}{\underline{G}^T B^T (B^T)^{-1}\Lambda^{-1}B^{-1}B\underline{G}}\right\} = \max_{\theta,\varphi}\left\{\frac{\left(\frac{1}{n}\sum_{k=1}^{n} |\underline{G}^T\Lambda^{-1}\underline{y}_k|\right)^2}{\underline{G}^T\Lambda^{-1}\underline{G}}\right\} \tag{62}$$

The last equation is identical to Equation (18) and this establishes the fact that the receiver defined by Equation (61) and that defined by Equation (18) are equivalent in every respect. This argument has established the invariance principle for the case where the beam covariance matrix is known and the observations are incoherent. The proofs for the other cases under consideration are almost identical and will not be given here.

Another property of considerable interest is that the likelihood-ratio receivers derived for the case where the beam covariance matrix is unknown exhibit a CFAR (constant false alarm rate) behavior. More precisely, it will be shown that these receivers have a false alarm probability that is completely independent of the true beam covariance matrix. This fact can be established by consideration of Equation (57) which defines the likelihood-ratio receiver for the case of incoherent observations. Introduction of the change of variables defined by $\underline{\eta}_k = \Lambda^{-\frac{1}{2}}\underline{y}_k$, where Λ denotes the true beam covariance matrix, results in the equation,

$$\max_{\theta,\varphi}\left\{\frac{(\frac{1}{n}\sum_{k=1}^{n}|\underline{G}^T R^{-1}\underline{y}_k|)^2}{\underline{G}^T R^{-1}\underline{G}}\right\} = \max_{\theta,\varphi}\left\{\frac{(\frac{1}{n}\sum_{k=1}^{n}|\underline{G}^T\Lambda^{-\frac{1}{2}}M^{-1}\underline{\eta}_k|)^2}{\underline{G}^T\Lambda^{-\frac{1}{2}}M^{-1}\Lambda^{-\frac{1}{2}}\underline{G}}\right\} \tag{63}$$

where the matrix M is defined by

$$M = \mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{\eta}_k\underline{\eta}_k^*\right] \tag{64}$$

Next, Equation (63) plus the recollection that assumption I (which was used in deriving this equation) implies that the vector $\Lambda^{-\frac{1}{2}}\underline{G}$ ranges over all possible directions in m-space as θ and φ vary through their ranges, results in the equation,

$$\max_{\theta,\varphi}\left\{\frac{(\frac{1}{n}\sum_{k=1}^{n}|\underline{G}^T R^{-1}\underline{y}_k|)^2}{\underline{G}^T R^{-1}\underline{G}}\right\} = \max_{\underline{H}}\left\{\frac{(\frac{1}{n}\sum_{k=1}^{n}|\underline{H}^T M^{-1}\underline{\eta}_k|)^2}{\underline{H}^T M^{-1}\underline{H}}\right\} \tag{65}$$

where the last maximation is to be taken over all possible, real vectors, $\underline{H}$. It is easily seen that, when $A = 0$,

$$E(\underline{\eta}_k\underline{\eta}_k^*) = I \quad , \quad E(\underline{\eta}_k) = 0 \tag{66}$$

which means that Equation (65) is completely independent of the true beam covariance matrix Λ when no target return is present. This means that the probability that the right-hand side of Equation (65) exceeds a given threshold (the false alarm probability of the receiver) is independent of Λ thus establishing the CFAR behavior asserted earlier. A similar argument, leading to the same conclusion easily can be carried out for the coherent case.

5. AN APPROXIMATION TO THE LIKELIHOOD RATIO RECEIVER.

The case to be considered here is that of a monopulse receiver using conventional sum and difference beams. For simplicity, attention will be restricted to a two-beam system where the beams are functions of only one angular coordinate. It has been shown above that when assumption I holds, the angle estimate $\hat{\theta}$ is always given by an equation of the form

$$\underline{G}(\hat{\theta}) = \alpha\,\mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{y}_k e^{-j\hat{\phi}_k}\right] \tag{67}$$

regardless of whether the observations are coherent or incoherent or if the beam covariance matrix is known or unknown. The only difference between these various cases is the manner in which the carrier phase estimates $\hat{\phi}_k$ are obtained. For the two-beam case under discussion, Equation (67) is equivalent to the equation,

$$\frac{G_1(\hat{\theta})}{G_2(\hat{\theta})} = \frac{\mathrm{Re}\left[\sum_{k=1}^{n} y_{1k}e^{-j\hat{\phi}_k}\right]}{\mathrm{Re}\left[\sum_{k=1}^{n} y_{2k}e^{-j\hat{\phi}_k}\right]} \tag{68}$$

where $G_1(\theta)$ will be understood to be the difference-beam pattern and $G_2(\theta)$ the sum-beam pattern. Similarly, y_{1k} denotes the signal received on the difference beam after the k^{th} transmission and y_{2k} denotes the signal received on the sum beam after the k^{th} transmission.

One way of obtaining an approximation to the angle estimate $\hat{\theta}$ is to first decide on a reasonable approximation to the carrier phase estimates $\hat{\phi}_k$ and then to use these approximate estimates in Equation (68). For the incoherent case, it seems reasonable to set $\hat{\phi}_k$ equal to the phase of the k^{th} observation made by the sum beam, i.e., $\hat{\phi}_k = \arg y_{2k}$. With this approximation, Equation (68) becomes,

$$\frac{G_1(\theta)}{G_2(\hat{\theta})} = \frac{\sum_{k=1}^{n}|y_{1k}|\cos(\theta_{1k}-\theta_{2k})}{\sum_{k=1}^{n}|y_{2k}|} \tag{69}$$

when $\theta_{1k} = \arg y_{1k}$ and $\theta_{2k} = \arg y_{2k}$. Finally, making the usual assumptions that the difference beam is linear in θ and the sum beam is constant, leads to the equation,

$$\hat{\theta} = \alpha \frac{\sum_{k=1}^{n} |y_{1k}| \cos(\theta_{1k} - \theta_{2k})}{\sum_{k=1}^{n} |y_{2k}|} \tag{70}$$

where α is a constant that can be determined from the known beam shapes.

For the case of one observation, n=1, the angle estimate given by Equation (70) is identical to the one used by many practical monopulse systems. The generalization to n observations is a straightforward incoherent combination of the conventional monopulse data available after each individual transmission as indicated by Equation (70).

A similar analysis applicable to the coherent case leads to the equation,

$$\hat{\theta} = \alpha \frac{\sum_{k=1}^{n} |y_{1k}| \cos(\theta_{1k} - \hat{\psi})}{\left| \sum_{k=1}^{n} y_{2k} \right|} \tag{71}$$

where the approximate carrier phase estimate $\hat{\psi}$ is taken to be the phase of the coherently integrated sum beam returns,

$$\hat{\psi} = \arg \sum_{k=1}^{n} y_{2k} \tag{72}$$

Once again, Equation (71) dictates a straightforward combination of the conventional monopulse data available at the end of each transmission but this time coherent rather than incoherent integration is involved.

The next job at hand is to derive easily realizable approximations to the detection algorithms defined by Equations (18), (19), (57) and (58). The simplest way of accomplishing this end seems to be to assume that only the outputs from the sum beam will be used for target detection. Given this assumption, the next task is to derive the likelihood-ratio receiver based on the observations $y_{21}, \ldots, y_{2n}$ in the hope that its structure will be simple enough to be easily realizable. This indeed turns out to be the case as now will be demonstrated.

The details of the derivation of the likelihood-ratio receiver based on the observations $y_{21}, y_{22}, \ldots, y_{2n}$ will not be presented because they are similar in every respect to the derivation of the likelihood-ratio receiver based on $y_1, y_2, \ldots, y_n$ that has been given above. There are again four cases to consider depending on whether the target returns on the sum beam are coherent or incoherent and whether the variance of the sum-beam interference signal is known or unknown. When this variance is known, the receiver is defined by the equation

$$\sum_{k=1}^{n} |y_{2k}| \begin{array}{l} \text{target present} \\ > \\ < \\ \text{target absent} \end{array} \lambda \tag{73}$$

in the incoherent case and by the equation

$$\left| \sum_{k=1}^{n} y_{2k} \right| \begin{array}{l} \text{target present} \\ > \\ < \\ \text{target absent} \end{array} \lambda \tag{74}$$

in the coherent case. When the variance of the interference is unknown, the corresponding receivers are defined by the equations

$$\frac{\left(\sum_{k=1}^{n} |y_{2k}| \right)^2}{\sum_{k=1}^{n} |y_{2k}|^2} \begin{array}{l} \text{target present} \\ > \\ < \\ \text{target absent} \end{array} \lambda \tag{75}$$

and

$$\frac{\left| \sum_{k=1}^{n} y_{2k} \right|^2}{\sum_{k=1}^{n} |y_{2k}|^2} \begin{array}{l} \text{target present} \\ > \\ < \\ \text{target absent} \end{array} \lambda \tag{76}$$

valid for the incoherent and coherent cases respectively. Note that these receivers, which are only approximations to the exact receivers defined by Equations (57) and (58), still have the CFAR property, i.e., their false alarm probabilities are independent of the true value of the variance of the interference.

The results of this section are summarized in Fig. 2.

6. CONCLUSION.

The main results of this paper are the receiver structures summarized in Figs. 1 and 2. From an inspection of these figures, it is apparent that the exact likelihood-ratio receivers are much more complex and, therefore, much harder to implement than the approximations to these receivers shown in Fig. 2. In view of this fact, it is very important to compare the performance (angular accuracy and detector characteristic) of the exact and approximate receivers in order to determine whether the increased complexity of the exact receiver really achieves any worthwhile gains. Computer simulations as well as theoretical studies are now in progress to accomplish this goal. The results of this work are not available at this time but they will be presented in the near future.

7. REFERENCES.

[1] McGinn, J. W. "Thermal Noise in Amplitude Comparison Monopulse Systems," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-2, pp. 550-556, September 1966.

[2] Urkowitz, H. "The Accuracy of Maximum-Likelihood Estimates in Radar and Sonar," IEEE Transactions on Military Electronics, Vol. MIL-8, pp. 39-45, January 1964.

[3] Hofstetter, E. M. and DeLong, D. F., Jr. "Detection and Parameter Estimation in an Amplitude Comparison Monopulse Radar," IEEE Transactions on Information Theory, Vol. IT-15, No. 1, pp. 22-30, January 1969.

[4] Mosca, E. "Angle Estimation in Amplitude Comparison Monopulse Systems," IEEE Transactions on Aerospace and Electronics Systems, Vol. AES-5, No. 2, pp. 205-212, March 1969.

[5] Mosca, E. and Minguzzi, B. "Optimum Monopulse Data Processing in a Clutter Plus Noise Background," Alta Frequenza, Vol. XXXVII, No. 8, pp. 768-777, August 1968.

[6] Turin, G. L. "On Optimal Diversity Reception," IRE Transactions on Information Theory, Vol. IT-7, pp. 154-166, July 1961.

[7] Bodewig, E. Matrix Calculus, Amsterdam: North Holland Publishing Co., 1956.

	COHERENT OBSERVATIONS	INCOHERENT OBSERVATIONS	ONE OBSERVATION
KNOWN Λ	$\max_{\theta,\omega}\left\{\frac{\left\lvert\sum_{k=1}^{n} \underline{y}_k^* \Lambda^{-1}\underline{G}\right\rvert^2}{\underline{G}^T\Lambda^{-1}\underline{G}}\right\} \gtrless \lambda$ $\hat{\theta},\hat{\omega}$ maximizing values above or $\underline{G}(\hat{\theta},\hat{\omega}) = \alpha\,\mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{y}_k e^{-j\hat{\psi}}\right]$ (I) where $\hat{\psi}$ determined by $\max_{\psi}\left\{\underline{Y}^T\Lambda^{-1}\underline{Y}\right\}$	$\max_{\theta,\omega}\left\{\frac{\sum_{k=1}^{n}\left\lvert\underline{y}_k^*\Lambda^{-1}\underline{G}\right\rvert^2}{\underline{G}^T\Lambda^{-1}\underline{G}}\right\} \gtrless \lambda$ $\hat{\theta},\hat{\omega}$ maximizing values above or $\underline{G}(\hat{\theta},\hat{\omega}) = \alpha\,\mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{y}_k e^{-j\hat{\psi}_k}\right]$ (I) where $\hat{\psi}_k$ is determined by $\max_{\{\psi_k\}}\left\{\underline{Y}^T\Lambda^{-1}\underline{Y}\right\}$	$\frac{1}{2}\underline{y}_1^*\Lambda^{-1}\underline{y}_1 + \frac{1}{2}\left\lvert\underline{y}_1{}^T\Lambda^{-1}\underline{y}_1\right\rvert \gtrless \lambda$ (I) $\underline{G}(\hat{\theta};\hat{\omega}) = \alpha\left[\mathrm{Re}\ \underline{y}_1 e^{-j\hat{\psi}}\right]$ (I) where $\hat{\psi} = \frac{1}{2}\arg \underline{y}_1{}^T\Lambda^{-1}\underline{y}_1$
UNKNOWN Λ	$\max_{\theta,\omega}\left\{\frac{\left\lvert\sum_{k=1}^{n}\underline{y}_k^* R^{-1}\underline{G}\right\rvert^2}{\underline{G}^T R^{-1}\underline{G}}\right\} \gtrless \lambda$ (I) $\hat{\theta},\hat{\omega}$ maximizing values above or $\underline{G}(\hat{\theta},\hat{\omega}) = \alpha\,\mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{y}_k e^{-j\hat{\psi}}\right]$ (I) where $\hat{\psi}$ determined by $\max_{\psi}\left\{\underline{Y}^T R^{-1}\underline{Y}\right\}$	$\max_{\theta,\omega}\left\{\frac{\left(\sum_{k=1}^{n}\left\lvert\underline{y}_k^* R^{-1}\underline{G}\right\rvert\right)^2}{\underline{G}^T R^{-1}\underline{G}}\right\} \gtrless \lambda$ (I) $\hat{\theta},\hat{\omega}$ maximizing values above or $\underline{G}(\hat{\theta},\hat{\omega}) = \alpha\,\mathrm{Re}\left[\frac{1}{n}\sum_{k=1}^{n}\underline{y}_k e^{-j\hat{\psi}_k}\right]$ (I) where $\hat{\psi}_k$ is determined by $\max_{\{\psi_k\}}\left\{\underline{Y}^T R^{-1}\underline{Y}\right\}$	NOT APPLICABLE

Fig. 1 Exact Likelihood-Ratio Receivers

	COHERENT OBSERVATIONS	INCOHERENT OBSERVATIONS
KNOWN Λ	$\left\| \sum_{k=1}^{n} y_{2k} \right\| \underset{<}{\overset{>}{}} \lambda$ $\hat{\theta} = \alpha \dfrac{\sum_{k=1}^{n} \lvert y_{1k} \rvert \cos(\theta_{1k} - \hat{\psi})}{\left\| \sum_{k=1}^{n} y_{2k} \right\|}$ where $\hat{\psi} = \arg \sum_{k=1}^{n} y_{2k}$	$\sum_{k=1}^{n} \lvert y_{2k} \rvert \underset{<}{\overset{>}{}} \lambda$ $\hat{\theta} = \alpha \dfrac{\sum_{k=1}^{n} \lvert y_{1k} \rvert \cos(\theta_{1k} - \theta_{2k})}{\sum_{k=1}^{n} \lvert y_{2k} \rvert}$
UNKNOWN Λ	$\dfrac{\left\| \sum_{k=1}^{n} y_{2k} \right\|^2}{\sum_{k=1}^{n} \lvert y_{2k} \rvert^2} \underset{<}{\overset{>}{}} \lambda$ same $\hat{\theta}$ as for known Λ	$\dfrac{\left(\sum_{k=1}^{n} \lvert y_{2k} \rvert \right)^2}{\sum_{k=1}^{n} \lvert y_{2k} \rvert^2} \underset{<}{\overset{>}{}} \lambda$ same $\hat{\theta}$ as for known Λ

Fig. 2 Approximations to the Likelihood-Ratio Receivers

TRACKING ERRORS IN A FREQUENCY AGILITY SYSTEM

by

P.F.Guarguaglini

SELENIA S.p.A.
Via Tiburtina km 12.4, Roma, Italy

TRACKING ERRORS IN A FREQUENCY AGILITY SYSTEM

P.F.Guarguaglini
SELENIA S.p.A.
Via Tiburtina Km 12.4,Roma,Italy

SUMMARY

The present study evaluates the influence of frequency agility on tracking errors in a radar system. The use of frequency agility modifies the received spectrum. Therefore the magnitude of the glint and scintillating errors changes. In order to find the signal spectrum, the correlation, between succeeding pulses, has been evaluated. The correlation is dependent on frequency agility bandwidth, on target dimensions and on the law, by which the transmitted frequency is chosen.

Transforming the signal autocorrelation function, the received power spectrum has been obtained; a broadening of the spectrum due to the reduction of the correlation among pulses, occurs. Expressions of glint and scintillation errors have been obtained. For some specific cases, errors have been evaluated and plotted as a function of the correlation coefficient.

From the analysis of the results, it is possible to draw this conclusions: the use of frequency agility gives a reduction of tracking errors for a monopulse radar. For a conical-scan system, while the glint error is always reduced, the scintillation error may increase.

1. INTRODUCTION.

The angular accuracy has been completely evaluated for radars, using a fixed transmission frequency [1]. The main causes of error, for both a monopulse and a conical-scan system, are [2]: a) amplitude fluctuation, b) angle scintillation or glint, c) receiver noise, d) servo noise. While the latter two errors are independent of the transmitted signal spectrum, the former two are dependent on how a target looks when a pulse hits it.

Therefore in a frequency agility system, where the transmitted frequency is changed at every pulse, the tracking errors, due to angle and amplitude fluctuations, assume different values that in the conventional radar.

The purpose of this paper is to carry out a complete evaluation of the tracking errors for monopulse and conical scan radars, both employing frequency agility.

2. ANGULAR ERRORS IN A MONOPULSE RADAR

The system, here considered, is a traditional monopulse radar. A simplified block diagram of such a radar may be found in [1] or [2].

The errors, introduced by target scintillation, depend mainly on the AGC and servo characteristics and on the noise power spectra at the input of these two circuits. This requires to compute the received signal spectrum.

2.1. Received signal spectrum.

The signals, both in the sum and in difference channels, have spectra of equal shape. When frequency agility in not used, the echo spectrum in skechted in fig. 1.

The broadening in the lines is caused by target scintillation and its shape is dependent on target characteristics. A good fit for this shape is assumed,[1]:

$$W(f)=\frac{1}{\pi f_0}\,\frac{f_0^2}{f^2+f_0^2} \qquad (1)$$

where f_0 , the half-power frequency, depends on the maximum target dimension and on radar wavelength.

Without taking into consideration the carrier frequency, the spectrum is given by:

$$S(f) = \left[\frac{\sin(\pi\tau f)}{\pi\tau f}\right]^2 \sum_{n=-\infty}^{+\infty} W(f - nF_r) \tag{2}$$

where τ is the pulse length and F_r is the pulse repetition frequency.

In a frequency agility radar, the target echo fluctuates not only because of motions of the target but also because the trasmitted frequency is different from pulse to pulse. Therefore the r.f. power spectrum of the received signal is different from the one in the previous case.

The power spectrum is related to the correlation function by:

$$S(f) = 2\int_0^\infty R(\tau)\cos(2\pi\tau f)\,d\tau \tag{3}$$

Therefore, by evaluating the correlation function it is possible to obtain the signal spectrum.

Let $\Phi_{ij}(f_i, f_j)$ be the correlation between the i-th and the j-th pulses when f_i and f_j are the transmitted frequencies.

Considering every possible case, the correlation factor is given by:

$$\Phi_{ij} = \iint \Phi_{ij}(f_i, f_j)\,p(f_i, f_j)\,df_i\,df_j \tag{4}$$

where $p(f_i, f_j)$ is the joint probability density function of f_i and f_j.

The criterion by which at every pulse the transmitted frequency is chosen depends on operative considerations. The criterion does not change during a single operative condition. Therefore the relation is valid:

$$p(f_q, f_{q+n}) = p(f_r, f_{r+n}) \qquad r \neq q \tag{5}$$

Eq. 4 may be written:

$$\Phi_n = \iint \Phi_n(f_i, f_{i+n})\,p(f_i, f_{i+n})\,df_i\,df_{i+n} \tag{6}$$

Let

$$p_s(t) = 2\int_0^\infty W(f)\cos\omega t\,dt \tag{7}$$

be the autocorrelation function of the internal fluctuations of the target. Taking into account eq. 6, except for a scaling factor, the envelope of the autocorrelation function of an infinite train of pulses is given by:

$$R_s(t) = s_0(t)\left[\sum_{n=-\infty}^{+\infty} \Phi_n\left(1 - \frac{|t - nT_r|}{\tau}\right)\cdot \mathrm{rect}\left[\frac{t - nT_r}{2\tau}\right]\right] \tag{8}$$

where rect(t) is given by [5]:

$$\mathrm{rect}(t) = \begin{cases} 1 & \text{for } |t| \leq \tfrac{1}{2} \\ 0 & \text{for } |t| > \tfrac{1}{2} \end{cases} \tag{9}$$

The presence of Φ_h in eq. 8 makes $R_s(t)$ decrease more rapidly than in the case of a single transmitted frequency. This means that the spectrum around the lines presents a broadening.

As eq. 6 shows, a complete evaluation of Φ_h is possible only if $\Phi_h(f_t, f_{t+n})$ and $p(f_t, f_{t+n})$ are known. There is no published experimental work on $\Phi_h(f_t, f_{t+n})$. A theoretical expression has been often used [4], [5]. In order to evaluate $p(f_t, f_{t+n})$ it is necessary define the criterion, employed for choosing the transmitted frequency. In a frequency agile radar, usually the frequency is randomly chosen within a band. Therefore the frequency of any two pulses are independent and the relation

$$p(f_t, f_{t+n}) = p(f_t)\,p(f_{t+n}) \tag{10}$$

holds.

Taking into consideration eq. 10, eq. 6 gives:

$$\Phi_h = \begin{cases} 1 & \text{for } h = 0 \\ \Phi & \text{for } h \neq 0 \end{cases} \tag{11}$$

The autocorrelation is now given by:

$$\begin{aligned} R_s(t) = \; & \varrho_s(t) \left\{ \left(1 - \frac{|t|}{\tau}\right) \operatorname{rect}\left(\frac{t}{2\tau}\right) + \right. \\ & \left. + \Phi \cdot \sum_{\substack{n=-\infty \\ n \neq 0}}^{n=+\infty} \left(1 - \frac{|t - nT_r|}{\tau}\right) \operatorname{rect}\left(\frac{t - nT_r}{2\tau}\right) \right\} = \\ = \; & \varrho_s(t) \left\{ (1-\Phi)\left(1 - \frac{t}{\tau}\right) \operatorname{rect}\left(\frac{t}{2\tau}\right) + \right. \\ & \left. + \Phi \cdot \sum_{n=-\infty}^{n=+\infty} \left(1 - \frac{|t - nT_r|}{\tau}\right) \operatorname{rect}\left(\frac{t - nT_r}{2\tau}\right) \right\} \end{aligned} \tag{12}$$

Usually the pulse length is much shorter than the target correlation time and it may be assumed, without loss of generality $\varrho_s(\tau) = 1$. Therefore

$$R_s(t) = (1-\Phi)\left[1 - \frac{|t|}{\tau}\right] \operatorname{rect}\left(\frac{t}{2\tau}\right) + \Phi \cdot \varrho_s(t) \sum_{n=-\infty}^{n=+\infty} \left(1 - \frac{|t - nT_r|}{\tau}\right) \operatorname{rect}\left(\frac{t - nT_r}{2\tau}\right) \tag{13}$$

Using eq. 3 it is possible to obtain the spectrum of the received signal:

$$S(f) = (1-\Phi) \cdot \left[\frac{\sin(\pi f \tau)}{\pi \tau f}\right]^2 + \frac{\Phi}{T_r} \left[\frac{\sin(\pi \tau f)}{\pi \tau f}\right]^2 \sum_{n=-\infty}^{n=+\infty} W\left(f - nF_r\right) \tag{14}$$

As eq. 14 shows, the use frequency agility divides the spectrum in two terms. One of them has a shape equal to the spectrum of a convential radar, but is reduced by a factor Φ, which depends on target characteristics and on the width of the transmitted frequency band. The other term gives rise to a broad-band spectrum. If complete uncorrelation between any two pulses is obtained, only the broad-band term is present.

Fig. 2 shows a typical spectrum of the received signal for a frequency agile monopulse radar.

2.2. Error evaluation.

In order to evaluate the filtering effect of AGC and servo loops on scintillating noise, it is necessary to consider the presence of detectors both in the sum and difference channels. As mentioned before, the signal spectrum shape is equal in both channels. Therefore the signal spectrum at the output of a phase-sensitive detector is given by the convolution of the i.f. spectrum with itself, as for a square-law detector.

Regarding the envelope detector, the shape of detected spectra are slightly different for square-law and linear detection [6]. Therefore only the square-law case will be considered.

In Appendix the signal spectrum after detection has been computed. This spectrum is given by:

$$S_e(f) = k_1^2 \left[\Phi^2 \frac{F_r}{\pi f_g} \sum_{n=-\infty}^{n=+\infty} \frac{f_g^2}{f_r^2 + (f - nF_r)^2} \cdot \frac{2\pi n F_r \tau - \sin(2\pi n F_r \tau)}{(2\pi n F_r \tau)} + \right. \tag{15}$$

$$\left. + (1 - \Phi^2) \frac{2\pi f \tau - \sin(2\pi f \tau)}{(2\pi \tau f)^2} \right]$$

where $f_g = 2f_0$.

If k_1 is given in meter per sec or mils per sec, $S_e(f)$ givesdirectly the error spectrum.

The servo and AGC filters are low-pass; therefora, only the error spectrum components around $f = o$ are of interest.

This part, $S_e'(f)$, is given by:

$$S_e'(f) = k_1^2 \left\{ \Phi^2 \frac{\alpha}{\pi} \frac{f_g^2}{f_g^2 + f^2} + (1 - \Phi^2) \right\} \tag{16}$$

where f_g.

It is now possible to evaluate errors.

The error introduced by amplitude scintillation, is proportional to the bias error, θ_0 . It is null when the target is on the boresight.

The spectrum of this error is [1]:

$$W_n(f) = S_e'(f) \left| Y_s(f) \right|^2 \theta_0^2 \tag{17}$$

where $Y_s(f)$, the scintillation error factor, is defined by:

$$Y_s(f) = \frac{Y_c(f)}{\left| 1 + Y_a(f) \right|} \tag{18}$$

$Y_a(f)$ is the open-loop voltage gain transfer function of the AGC loop and $Y_c(f)$ is the closed-loop transfer function of the servo.

Integrating $W_n(f)$ from $-\infty$ to $+\infty$ and taking the square root of the integral

the error, σ_ξ ,is obtained:

$$\sigma_{\varepsilon_1} = \vartheta_0 \left\{ k_2 \bar{\Phi}^2 \frac{\alpha}{\pi} \int_{-\infty}^{+\infty} \frac{f_g^2}{f_g^2 + f^2} \left| Y_s(f) \right|^2 df + k_2 (1-\bar{\Phi}^2) \int_{-\infty}^{+\infty} \left| Y_s(f) \right|^2 df \right\}^{1/2} \tag{19}$$

The error, when frequency agility is not used, is:

$$\sigma_\xi = \vartheta_0 \left\{ k_2 \frac{\alpha}{\pi} \int_{-\infty}^{+\infty} \frac{f_g^2}{f_g^2 + f^2} \left| Y_s(f) \right|^2 df \right\}^{1/2} \tag{20}$$

Therefore eq. 17 may be written:

$$\sigma_{\varepsilon_1} = \sigma_\xi \left\{ \bar{\Phi}^2 + \frac{\pi (1-\bar{\Phi}^2)}{\alpha \sigma_\xi} B_n \left| Y_s(0) \right|^2 \right\}^{1/2} \tag{21}$$

where B_n is the noise bandwidth.

Eq. 21 shows that, for $\bar{\Phi}$ constant, the error reduces as B_n and $Y(0)$ increase. It reduces by a lesser amount as α and σ_ξ increase. The angular noise spectrum is like the scintillation spectrum. Therefore the glint error, σ_ε , when AGC is absent, is expressed by eq. 21, where $|Y_s(0)|^2$ is substituted by $|Y_c(0)|^2$.

To compute the glint error, σ_{ε_1} , when a fast AGC is present, is very complex. However it is possible to use an approximation. It has been found, [6] [7], that, because of a correlation between the echo amplitude and the angular error, the presence of a fast AGC doubles σ_{ε_1} . Therefore, considering this approximation valid, it is possible to use eq. 21.

3. ANGULAR ERRORS IN A CONICAL-SCAN RADAR.

In a conical-scan radar another error must be added to the two errors, evaluated for monopulse systems. This error is produced by high frequency components within the passband which is centered at the scan frequency.

This error is [1] :

$$\sigma_c^2 = 0.45\, \vartheta_B^2\, S_c'(f)\, B_n \tag{22}$$

where ϑ_B is the 3 dB beamwidth and B_n is the servo bandwidth.

Taking into consideration eq. 16 we have:

$$S_c'(f) = k_2 \frac{\alpha}{\pi} \frac{f_g^2}{f_g^2 + f^2} \qquad \text{for} \quad \bar{\Phi} = 1 \tag{23a}$$

$$S_1'(f_0) = k_2 \left[\frac{\pi}{\pi} \Phi^2 \frac{f_g^2}{f_g^2 + f_c^2} + (1-\Phi^2) \right] \qquad \text{for } \Phi \neq 1 \tag{23b}$$

Therefore

$$\sigma_{\varepsilon_1} = \sigma_\varepsilon \left\{ \Phi^2 + \frac{\pi}{\alpha'} (1-\Phi^2) \frac{f_g^2 + f_0^2}{f_g^2} \right\}^{1/2} \tag{24}$$

for small values of Φ and $f_g/f_0 \ll 1$:

$$\sigma_{\varepsilon_1} \simeq \sigma_\varepsilon \frac{f_c}{f_g} \sqrt{\frac{\pi}{\alpha'}(1-\Phi^2)} = \sigma_\varepsilon \frac{f_0}{\sqrt{f_g F_c}} \sqrt{\pi(1-\Phi^2)} \tag{25}$$

Therefore the use of frequency agility may increase this error, and the total error could be worse than that of a conventional radar, without frequency agility.

4. EXAMPLES.

In order to give an idea of the effect of frequency agility, errors have been evaluated, assuming for the servo and AGC loops, the transfer function, chosen in [1]. The servo band-width and the target spectrum bandwidth are 5 cps, while three different AGC characteristics are taken into consideration (no AGC, slow AGC and fast AGC). A slow AGC usually in employed in a conical-scan radar, while a fast AGC is used in a monopulse radar.

By graphical integration, it is shown that:

$$\sigma_{\varepsilon_1}^2 = \sigma_\varepsilon^2 \left[\Phi^2 + \frac{5.4}{\alpha} (1-\Phi^2) \right] \qquad \text{for no AGC} \tag{25a}$$

$$\sigma_{\varepsilon_1}^2 = \sigma_\varepsilon^2 \left[\Phi^2 + \frac{6}{\alpha} (1-\Phi^2) \right] \qquad \text{for slow AGC} \tag{25b}$$

$$\sigma_{\varepsilon_1}^2 = \sigma_\varepsilon^2 \left[\Phi^2 + \frac{0.8}{\alpha} (1-\Phi^2) \right] \qquad \text{for fast AGC} \tag{25c}$$

In figs 3,4 and 5 $\sigma_{\varepsilon_1}^2 = \sigma_{\varepsilon_1}^2/\sigma_\varepsilon^2$ has been plotted versus Φ for three different values of α.

Comparing the results, the use of frequency agility reduces the conical-scan radar more than the monopulse error.

Eq. 26a allows us to obtain the glint error. It is possible to evaluate how a typical situation is modified by frequency agility.

The situation considered is shown in [8] and in [1]. Assuming $\Phi = 0.4$, the glint

error is reduced by a factor 10. In fig. 6 the noise components vs.relative range are plotted, when the bias error Θ_o is zero, for a monopulse radar. The amplitude noise is not present because $\Theta_o = 0$. Curve A represent the total error for a conventional radar, while curve B is the total error for a frequency agile radar.

The reduction of glint error is great and the total error is reduced for low values of R. The advantage of frequency agility disappears as R becomes large.

In the case of a conical-scan radar, assuming $f_s/f_{fa} = 6$, we have:

$$\begin{aligned} \sigma_{\varepsilon_1} &= 1.5\,\sigma_\varepsilon && \text{for} \quad \alpha = 50 \\ \sigma_{\varepsilon_2} &= 1.06\,\sigma_\varepsilon && \text{for} \quad \alpha = 100 \\ \sigma_{\varepsilon_3} &= 0.75\,\sigma_\varepsilon && \text{for} \quad \alpha = 200 \end{aligned} \tag{27}$$

Then for $\alpha = 50$ the amplitude noise error increases.

In fig. 7 curve A shows the total error for a conventional conical-scan radar, curve B represents the total error for a frequency agile conical-scan radar when the amplitude fluctuation does not vary. Curve C represents the case of $\alpha = 50$. Because the amplitude fluctuation increases, curve C is not always lower than curve A. Therefore there are some value of R, where the use of frequency agility worsens the tracking performance.

5. CONCLUSION.

The use of frequency agility reduces the glint error. This reduction may be great. The amplitude scintillation error, which is present only in a conical-scan radar, may increase or decrease. For small values of Φ, amplitude scintillation error increases as $\sigma_\varepsilon^2/(f_s \sigma_\varepsilon)$ increases.

Therefore the use of frequency agility increases monopulse system performance. For a conical-scan radar it is not possible to reach an univocal conclusion.

APPENDIX

Signal spectrum at the detector output

Apart a continuous term the autocorrelation function at the output of a square-law detector is:

$$R_u(t) = k\,R_\varepsilon^2(t) \tag{A1}$$

If $R_\varepsilon(t)$ is given by eq. 12, we have

$$R_u(t) = k\left\{\sigma_\varepsilon^2(t)\,\bar{\Phi}^2 \sum_{n=-\infty}^{n=+\infty}\left[1-\frac{|t-nT_r|}{\tau}\right]^2 \mathrm{rect}\left(\frac{t-nT_r}{2\tau}\right) + \left(1-\bar{\Phi}^2\right)\left(1-\frac{|t|}{\tau}\right)^2 \mathrm{rect}\left(\frac{t}{2\tau}\right)\right\} \tag{A2}$$

The spectrum corresponding to the second term is:

$$S_{\varepsilon_2}(f) = 2k\left(1-\bar{\Phi}^2\right)\int_0^\infty \mathrm{rect}\left(\frac{t}{2\tau}\right)\left(1-\frac{t}{\tau}\right)^2 \cos(\omega t)\,dt = 2k\left(1-\bar{\Phi}^2\right)\int_0^\tau \left(1-\frac{t}{\tau}\right)^2 \cos(\omega t)\,dt \tag{A3}$$

Putting $t' = 1 - t/\tau$

$$S_{e_2}(f) = 2k(1-\bar{\Phi}^2)\int_0^1 t'^2 \cos(\pi f \tau (1-t'))dt' = 4k(1-\bar{\Phi}^2)\tau \frac{2\pi f\tau - \sin(2\pi f\tau)}{(2\pi f\tau)^3} \tag{A4}$$

Considering that:

$$\rho_c^2(t) = \exp[-4\pi f_0|t|] = \exp[-2\pi f_g|t|] \tag{A5}$$

the spectrum corresponding to $\rho_c^2(t)$ is:

$$W'(f) = \frac{1}{\pi f_g}\,\frac{f_g^2}{f_g^2 + f^2} \tag{A6}$$

Taking into account eqs. A4 and A5 and that:

$$\sum_{-\infty}^{+\infty} u(t-nT_r) \Longleftrightarrow F_r \sum_{-\infty}^{+\infty} U(nF_r)\,\delta(f-nF_r) \tag{A7}$$

we have:

$$S_{e_2}(f) = 4k\frac{\bar{\Phi}^2 F_r}{\pi f_g}\sum_{-\infty}^{+\infty}\frac{f_g^2}{f_g^2+(f-nF_r)^2}\,\frac{2\pi nF_r\tau - \sin(2\pi nF_r\tau)}{(2\pi nF_r\tau)^3} \tag{A8}$$

The total error spectrum is

$$S_e(f) = k_1\left[\bar{\Phi}^2\frac{F_r}{\pi f_g}\sum_{-\infty}^{+\infty}\frac{f_g^2}{f_g^2+(f-nF_r)^2}\,\frac{2\pi nF_r\tau - \sin(2\pi nF_r\tau)}{(2\pi nF_r\tau)^3} + (1-\bar{\Phi}^2)\frac{2\pi f\tau - \sin(2\pi f\tau)}{(2\pi f\tau)^3}\right] \tag{A9}$$

REFERENCES

[1] Barton D.K., "Radar System Analysis", Prentice-Hall, 1964.

[2] Skolnik M.I., "Introduction to Radar Systems", Mc Graw-Hill Book Co., 1962.

[3] Woodward P.M., "Probability and Information Theory, with Applications to Radar", Pergamon Press, 1960.

[4] Ray H., "Improving Radar Range and Angle Detection With Frequency Agility", 11th Annual East Coast Conf. on ANE, Oct. 1964.

[5] Grasso G.C. and P.P. Guarguaglini, "Valutazione delle prestazioni di un Sistema Radar a Diversità di Frequenza", Alta Frequenza vol. 36, n.9, Sept.1967 pp. 799-805.

[6] Davenport W.B. and W.L. Root, "Random Signals and Noise", Mc Graw-Hill Book Co.,1958.

[7] Dunn J.H. and D.D. Howard, "The Effects of Automatic Gain Control Performance on the Tracking Accuracy of Monopulse Radar Systems", Proc. IRE Vol. 47, pp. 430-435, March 1959.

[8] Dunn J.H. and al. "Phenomenon of Scintillation Noise in Radar Tracking Systems" Proc. IRE Vol. 47, pp. 855-863, May 1959.

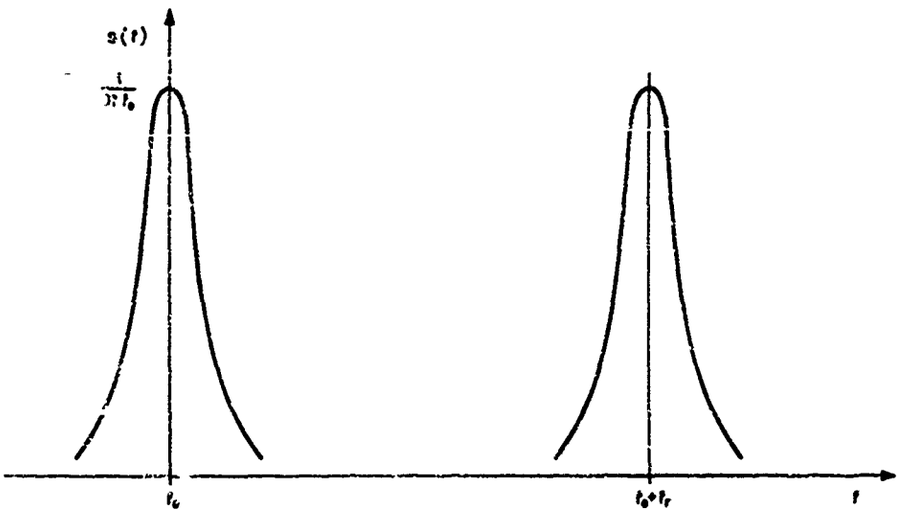


Figure 1

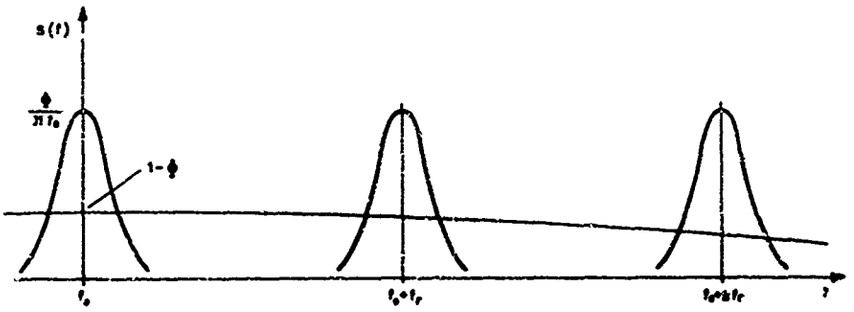


Figure 2

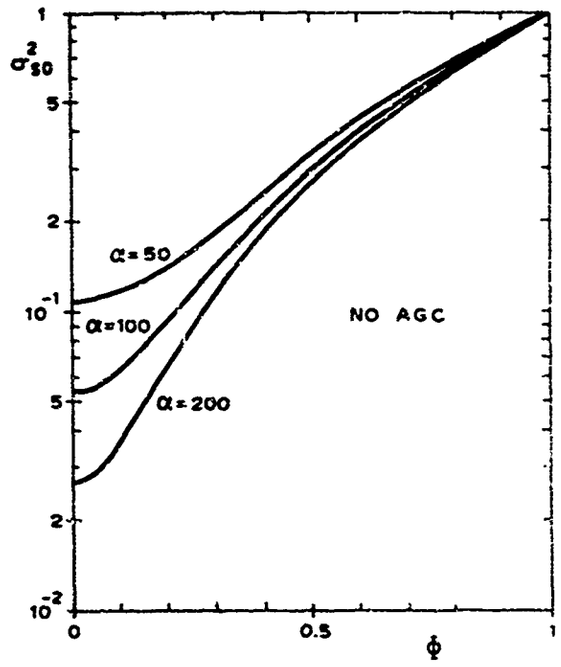


Figure 3

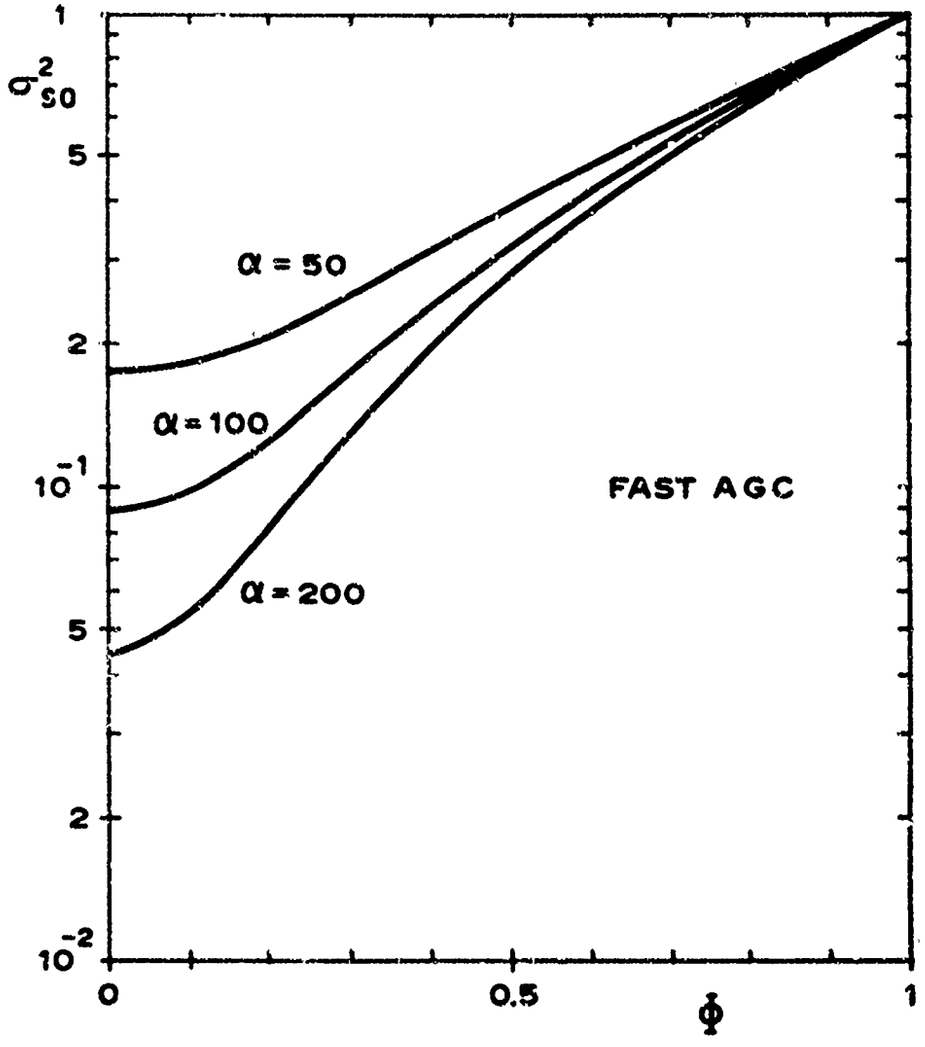


Figure 4

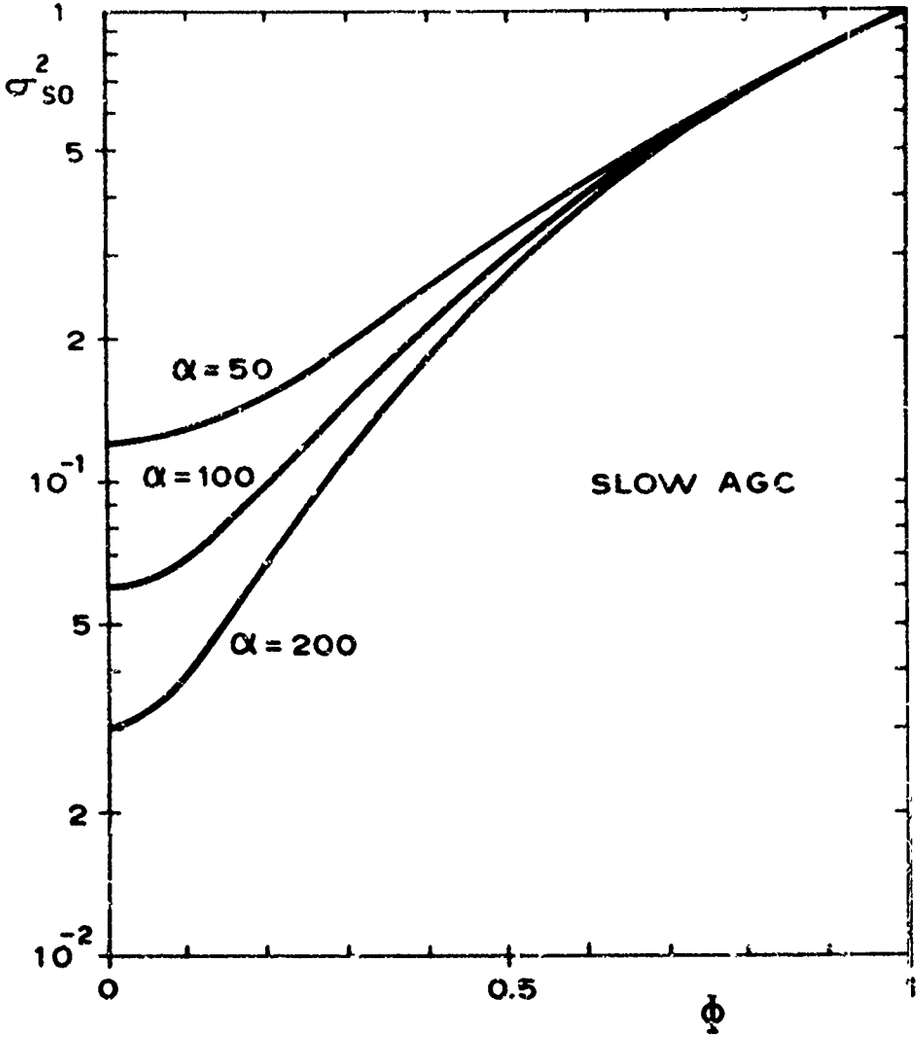


Figure 5

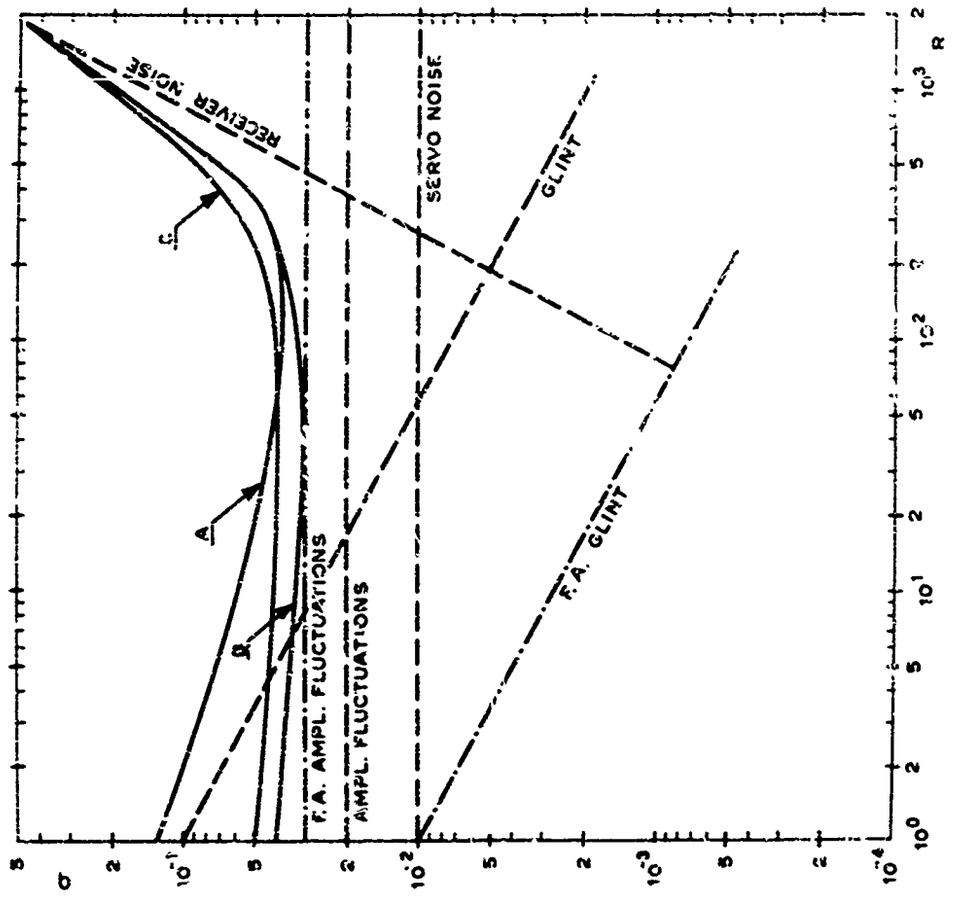


Figure 7

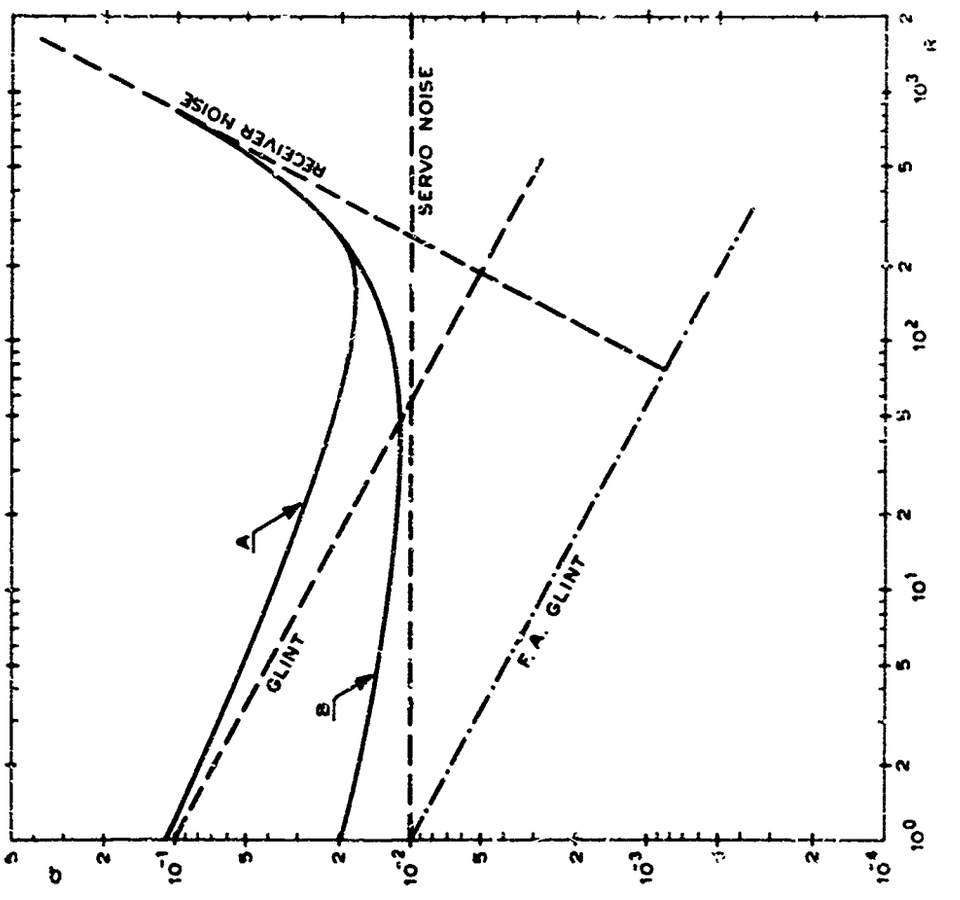


Figure 6

COHERENT AND NONCOHERENT BANDWIDTH IN RADAR RESOLUTION

by

August W.Rihaczek

Technology Service Corporation
Santa Monica, California, USA

COHERENT AND NONCOHERENT BANDWIDTH IN RADAR RESOLUTION

August W. Rihaczek
Technology Service Corporation
Santa Monica, California, U.S.A.

SUMMARY

Coherently used signal bandwidth permits the energy concentration in time that is the basis for much of the resolution potential of radar. There still remain problems such as target fluctuations or delays not representative of target range. Noncoherent use of bandwidth, be it in the form of single-pulse or pulse-to-pulse frequency diversity, does not improve nominal resolution. However, it smooths the fluctuations of clutter, for improved detection performance, or the variations in target cross sections or cross-section distributions. This paper investigates the significance of coherent and noncoherent bandwidth for high-performance radar. The utility of the two methods of bandwidth use is considered for such representative tasks as the measurement of target cross sections and high-resolution range profiles, signature analysis, target detection in clutter, and resolution based on the synthetic aperture approach. The principles are illustrated by several examples of the processing of simulated target returns.

1. INTRODUCTION

Target resolution is the task of recognizing the target of interest in the presence of other reflecting objects. These objects may be other targets, or such undesirable scatterers as rain drops or the earth background. Although the problem then is commonly designated as target detection in clutter, it is basically a resolution problem. In applications concerned with radar imagery, the interfering objects may be other scattering centers on the same target, which fact may make it difficult to measure the true cross-section distribution in range or some other coordinate.

In the early development of pulse radar, the parameter that defined resolution performance for targets within the same radar beam was pulse length. Modern radar resolution theory [1,2,3,4] has shown that range resolution is determined by the signal bandwidth. Moreover, radar performance can be enhanced by also utilizing resolution on the basis of Doppler or more general forms of target motion. This has led to the use of rather complicated radar waveforms, with coherent or noncoherent integration of many consecutive pulses, possibly with a change of the carrier frequency from pulse to pulse. An interesting question then becomes exactly what the role of signal bandwidth is in these high-performance systems.

This paper investigates the significance of signal bandwidth in, basically, matched-filter or correlation radar. It is an examination of the utility of bandwidth, both in coherent and noncoherent form, for the typical tasks of modern radar systems. After a discussion of the various methods of utilizing the available frequency band, we consider such problems as the detection of moving or fluctuating targets with high and low resolution systems, high-resolution approaches to target identification, target detection in clutter, and synthetic aperture radar. The purpose is to examine the manner in which bandwidth enters into the performance of these systems.

2. FORMS OF BANDWIDTH USE AND TERMINOLOGY

In the simplest system, the receiver processes coherently a single pulse at a time by passing the signal through a matched filter or a correlator. The process effectively subdivides the target space into range bins whose width ΔR depends on the pulse bandwidth B as

$$\Delta R = c/2B \tag{1}$$

where c is the speed of light. Ideally, only targets within the same range bin can interfere with each other, and by increasing the pulse bandwidth to where the range bin is so narrow that it contains only the target of interest, we obtain perfect resolution. The radar is said to make coherent use of the pulse bandwidth.

A closer examination of the situation shows that mutual interference will occur between separate range bins, unless the pulse spectrum is tapered toward the edges so that it is roughly bell-shaped. However, the tails of the spectrum, since they contain little energy, do not significantly contribute to range resolution. A given available frequency band thus can be utilized either for obtaining a narrow range bin or for eliminating the interference between different range bins, depending on the practical requirements.

Suppose, now, that the single-pulse signal-to-interference ratio is inadequate even though the pulse bandwidth was chosen at the largest allowable value. In order to enhance the signal-to-interference ratio, we then must integrate many pulses. However, an integration gain is obtained only if the interference is decorrelated from one pulse to the next, as would be the case if the scatterers contributing to the interference have sufficient differential motion with respect to the radar. In the absence of such motion, or if it is not strong enough, the interference must be artificially decorrelated by changing the carrier frequency from pulse to pulse. The process is referred to as pulse-to-pulse frequency diversity.

There is the option of integrating the pulses coherently or noncoherently, the latter referring to integration after envelope detection. A radar with coherent pulse integration may be said to use coherent frequency diversity, as opposed to noncoherent frequency diversity when pulse integration is noncoherent. In the coherent system, the signal-to-interference ratio is enhanced by a factor N from the single-pulse value, where N is the number of pulses integrated. Also, the range bin width decreases in inverse proportion to the bandwidth widening due to frequency diversity, and the long coherent integration time means an improved Doppler resolution capability. With noncoherent frequency diversity, there is only an enhancement in the signal-to-interference ratio, but no change in range or Doppler resolution from the single-pulse case.

The noncoherent integration gain is only somewhat smaller than for coherent integration when N is small, but approaches $\sqrt{N}$ when N is very large, which then means a large loss compared with coherent integration.

In a radar with frequency diversity, the spacing of the pulses is immaterial since decorrelation of the interference from pulse to pulse is due to the change in the carrier frequency. Hence, if more practical, the N pulses can be transmitted contiguously in the form of a single pulse. This means that we have the option to use the bandwidth of a pulse coherently and obtain a range bin in accordance with the entire pulse bandwidth, or to subdivide the pulse bandwidth into N segments, process each segment coherently, but combine the outputs noncoherently. With this type of single-pulse frequency diversity, the range bin width is N times as large as for coherent use of the entire pulse bandwidth. Single-pulse frequency diversity is of interest primarily when the bandwidth needed in order to obtain an adequate signal-to-interference ratio is so large that the target is broken up into its individual scattering centers with coherent use of the entire pulse bandwidth, and if this is undesirable from the point of view of target detectability.

An important question is how large a frequency step is needed in order to decorrelate the interference. It follows from matched-filter theory [4] that decorrelation between returns from scatterers separated by some range interval Δr occurs if the phase between the two returns is changed by at least π. Now, the relation between phase shift ϕ between the two returns and the range separation Δr of the scatterers is

$$\phi = \Delta r(2\pi/\lambda) = \Delta r(2\pi/c)f_o \tag{2}$$

where λ is the wavelength and f_o is the carrier frequency. With a carrier change of Δf_o, we have

$$\phi + \Delta\phi = \Delta r(2\pi/c)(f_o + \Delta f_o) \tag{3}$$

and with $\Delta\phi = \pi$, we find a required carrier shift of

$$\Delta f_o = c/2\Delta r \tag{4}$$

This shows that the required frequency shift decreases with increasing range separation of the scatterers. The most demanding requirement thus is due to the scattering objects within the same range bin that contains the target. The minimum frequency step Δf_o hence is found by substituting for Δr in Eq. (4) the range separation corresponding to the range bin width ΔR. By comparing Eq. (4) with Eq. (1), we then see that the frequency step must be at least in the order of the pulse bandwidth. In practical terms, frequency diversity requires frequency steps sufficiently large to avoid significant overlap of the individual pulse spectra. With N-fold frequency diversity and the most conservative use of the frequency band, the system bandwidth equals N times the single-pulse bandwidth. It follows from the fact that the maximum integration gain equals N that the detection performance in the interference improves proportionally with system bandwidth if the processing is fully coherent, and somewhat less than proportionally if the system bandwidth is utilized in part noncoherently. It is immaterial whether the bandwidth is concentrated within a single pulse or spread over many pulses.

3. FLUCTUATING TARGETS AND RANGE PROFILES

We consider a system in which the range bin is large compared with the target dimensions. A target whose dimensions also are small compared with one wavelength would have a cross section that is independent of the aspect angle and, assuming that the reflective properties of the target material are not strongly frequency dependent, also does not change with the carrier frequency. Actually, even with the often practical assumption that the target can be modeled as a fixed set of scattering centers, the apparent cross section will be strongly frequency and aspect dependent. In detection calculations, the assumption of a "fluctuating" target then is necessary.

The reason for the cross-section fluctuations is that the returns from the individual scattering centers of a target superpose with relative phases which depend on aspect angle as well as wavelength. Let us assume a target consisting of N scattering centers returning signals with amplitudes a_n and delays t_n. If the range bin is large compared with the target dimensions, we might as well consider the target response to the cw signal $\exp(j2\pi f_o t)$, which is given as

$$\psi(t) = \sum_{n=1}^{N} a_n e^{j2\pi f_o(t-t_n)} = e^{j2\pi f_o t} \sum_{n=1}^{N} a_n e^{-j2\pi f_o t_n} \tag{5}$$

The target cross section is proportional to the square of the signal magnitude

$$|\psi(t)| = |\psi(f_o)| = \left| \sum_{n=1}^{N} a_n e^{-j2\pi f_o t_n} \right| \tag{6}$$

Even for a fixed carrier frequency, the signal magnitude changes with aspect angle since the relative time delays t_n of the components vary.

The frequency dependence of the return of Eq. (6) can be utilized to obtain, for fixed aspect angle, the average target cross section. This is to say that noncoherent frequency diversity will smooth out the cross-section fluctuations. As long as the target presents the same aspect, the timing of the measurements on different carrier frequencies is immaterial, and other considerations govern the choice of single-pulse

or pulse-to-pulse frequency diversity. The interesting questions are how large the frequency step must be to yield an independent sample, and over what band the carrier frequency must be varied.

The answers can be found simply by considering the superposition of the sine waves returned from the N scattering centers. It follows from the π phase shift criterion used earlier that sample independence, or a decorrelated measurement, is ensured if the frequency change produces a phase shift of at least π between the returns from the (significant) closest and farthest scattering center. Hence, if the range spread of the significant scattering centers is ΔR_T, the required frequency shift is given by Eq. (4) when ΔR_T is written instead of Δr. From Eq. (1), it then is clear that the frequency step must at least equal the bandwidth needed to obtain, with coherent processing, a range bin equal to the target range spread ΔR_T.

To determine the overall frequency band, we must distinguish between targets consisting of so many scattering centers of comparable strength that they must be considered statistical in nature, and those where a few strong scattering centers predominate. In the first case, the number of measurements, and hence the required frequency band, depends on the desired measurement accuracy. Our interest here is with the second type of target. Proper averaging of the cross section simply requires that a phase shift of π must be introduced even between the returns from the most closely spaced significant scattering centers. However, this means that the noncoherent frequency diversity must extend over a frequency band that, when used coherently, would allow resolution of the individual scattering centers. Whether processing is noncoherent for cross-section smoothing, or coherent for the measurement of the cross-section distribution in range, the so-called range profile, the required bandwidth is the same.

The same conclusions can be reached from the point of view of the target transfer function. Here the target is taken as equivalent to a filter that modifies the radar signal in some characteristic manner. The transfer function of this filter is obtained by measuring amplitude and phase of the return signal as a function of frequency and, for our purposes, provides a complete description of the reflective properties of the target. Since the transfer function is the Fourier transform of the impulse response, we are dealing for the assumed target model with the transform of a set of lines, giving a sum of cosine waves with different amplitudes and "frequencies."

In the above picture, the most closely spaced scattering centers will produce the highest and the most widely separated scattering centers will give the lowest "beat frequency." Cross-section averaging requires that the frequency band extend over at least one half-period of the lowest beat frequency. Thus the required frequency band again is determined by the closest spacing of significant scattering centers.

The preceding conclusions are verified by Fig. 1. It is a plot of the return of Eq. (6) as a function of the carrier frequency for a target for which the closest spacing of scattering centers was taken as one time unit (of round-trip delay) and the total target spread was 20 time units. The smooth curve in Fig. 1 gives the average over the oscillating curve over the interval from the origin to the particular point on the abscissa. It is seen that a good measurement of the average cross section is obtained when the carrier frequency is varied over a band of unit width, which is indeed the inverse of the assumed closest time separation of the individual returns. Over this frequency band, the oscillating curve shows 20 half-cycles, corresponding to the ratio of target range spread and minimum spacing of the scattering centers.

Although the discussion was given in terms of the cross section of the entire target, it can evidently also be applied to a high-resolution radar measuring the range profile of the target and the fluctuations of the individual peaks of the range profile. Suppose that the coherently processed bandwidth is high enough to provide a detailed range profile, but not so high that all scattering centers of the target are resolved. A given peak in the range profile then may be due to the returns from several adjacent scattering centers, and slight changes in the aspect angle will produce changes in the shape of the range profile in the same manner as they cause cross-section fluctuations in a low-resolution radar. In order to obtain an average, or smoothed, range profile, noncoherent frequency shifting must extend over the same frequency band that, if processed coherently, would resolve the individual scattering centers contributing to a given peak in the range profile.

4. HIGH-RESOLUTION TARGET SIGNATURES

The primary utility of high-detail range profiles is for target classification and identification. However, aside from the difficulty of relating measured target cross-section distributions to the physical features of the target, several problems are associated with this particular approach to identification. First, if the available coherent bandwidth is insufficient to allow resolution of neighboring scattering centers, the shape of the range profile will vary from one measurement to the next, as discussed earlier. Second, since radar actually measures delay rather than range, creeping waves and artificially delayed returns will appear to originate from a different point in range and falsify the range profile. Third, internal target motion such as tumbling may interact with the measurement of the range profile and additional resolution on the basis of Doppler. In this section, we consider these questions in more detail.

For purposes of illustration, we assume a target model consisting of a number of point scattering centers spread over some range interval but with negligible spread in azimuth and elevation. The target is to perform a tumbling motion about its tip. The relation between the tumbling rate and the duration of the radar pulse is to be such that the change in the target aspect is noticeable from one pulse to the next but not during a single pulse.

In Fig. 2, we have plotted a succession of range profiles for this type of target. It is seen that the range profile changes with time, that is, as the target presents different aspects. Those peaks of the range profile that do not change are caused by single predominant scattering centers. Where several scattering centers contribute to a composite lobe of the range profile, the aspect variation causes typical fluctuations. The sequence of range profiles thus permits us to analyze the nature of the lobes that constitute the range profile.

The consecutive recording of the range profiles is a form of Doppler or, more generally, motion resolution. When motion resolution is on the basis of the carrier, the requirement for two scattering

centers to be resolved is that their relative range to the radar change by at least $\lambda/2$ over the coherent signal duration [4]. When consecutive range profiles are recorded, the combined amplitude from the returns associated with the two scattering centers thus will go through one full fluctuation cycle, which will be observable in the record. The fact that a complete Doppler capability is realized without coherence from pulse to pulse may appear puzzling. However, since the amplitude of the combined return depends on the relative phase of one component with respect to the second one scattering center essentially acts as a reference over all consecutive pulses. This is adequate when the number of interfering scattering centers is small, as is of interest here.

The plot of Fig. 2 corresponds to the situation where the change in the target aspect is so small, or range resolution so crude, that no worthwhile capability is obtained of measuring the tumbling motion through a sequence of range measurements on the individual scattering centers. In other words, the gross structure of the range profile remains the same for all curves of Fig. 2, and only the deformations of some of the lobes of the range profile indicate the changing aspect. The case of increased range resolution can be illustrated simply by continuing the recording of consecutive range profiles for a longer period.

In Fig. 3, we extend the record of Fig. 2 over ten times as large an interval, except that the time between two range profiles now is increased to keep the number of measurements the same. The tumbling motion of the target is evident from the figure. The range profile is being compressed as the aspect angle changes more toward broadside. However, while the lobes of the range profile generally converge toward the same point as we move toward the top of Fig. 3, this is not the case for the last lobe. It follows along a curve corresponding to a return from a point much closer to the tip of the target. This means that the last return includes some constant time delay, either because it is a creeping wave or because retransmission occurs with some delay.

The implications of such measurements on target analysis cannot be discussed here in any detail. It is clear that the coherent use of large bandwidths is the key to signature analysis in that it restricts the number of interfering scattering centers to those within the same range bin and, moreover, allows range tracking of the scattering centers in accordance with the particular target motion.

It will not always be possible in practice to observe a target over a sufficient spread of aspect angle. The variations in the aspect angle may be so small that the cross-section fluctuations they cause cannot be utilized for target analysis. Instead, we may want to average over these fluctuations by means of noncoherent frequency diversity and obtain a smoothed range profile. This was discussed earlier. As an illustration of the method, we show in Fig. 4 such averaging for the same target as used for Fig. 3. The bottom curve gives the range profile for one particular value of the carrier frequency, for which the first lobe of the range profile appears due to a single scattering center. The second curve is the average over the first range profile and another one taken at a different frequency. The third curve is the average over three measurements on three different carrier frequencies, and so forth, with the top curve giving the average over 8 range profiles.

The figure illustrates two points. First, the averaging introduces features that were not visible in the bottom curve, such as the break in the first lobe on the left. Second, there is an asymptotic approach to a smoothed range profile as the number of measurements on decorrelated frequencies is increased. In the case shown, the smoothed range profile is approached rather rapidly because the number of interfering scattering centers is small. In Fig. 4a, we repeat the plot of Fig. 4, except that the initial carrier frequency (for the bottom curve) was different. Comparison of the bottom curves for the two figures shows the pronounced differences for those parts of the range profiles where several scattering centers interfere. These differences represent the fluctuations in the range profile from one frequency to the next. Again, however, the smoothed range profile is approached rapidly for the case of Fig. 4a, and with only a few measurements the differences between Figs. 4 and 4a become insignificant.

5. TARGET DETECTION IN CLUTTER

Detection of a target in a clutter background is a special form of the resolution problem. The designation clutter implies that the interfering scatterers are of no interest, and also that they are so dense that their resolution would be impractical. For the following discussion, the coherently processed signal is first assumed so short that no Doppler resolution can be utilized.

In order to obtain an adequate signal-to-clutter ratio, two requirements must be met. First, the clutter originating within the same range bin as the target must be small, which means a coherent pulse bandwidth as determined by the allowable range bin width. This is signal bandwidth in the sense that every part of the frequency band contains high signal energy. Second, clutter interference contributed from range bins other than the one containing the target must be suppressed. This is accomplished by further widening of the signal bandwidth, but with relatively little energy in the added bandwidth. We refer to the fact that the frequency spectrum must be smoothly tapered toward the edges. In either case, it is seen that clutter performance improves with increasing coherent bandwidth. Once the mutual interference between different range bins has been eliminated, the signal-to-clutter ratio is proportional to the signal bandwidth.

Suppose that the available pulse bandwidth is limited to a value for which the range bin is larger than the target spread in range. A further enhancement of the signal-to-clutter ratio then requires pulse integration. However, with the assumption that the total integration time is too short to achieve a worthwhile Doppler capability, pulse-to-pulse frequency diversity must be used to decorrelate the clutter. As was shown earlier, the minimum frequency step equals the pulse bandwidth. Since the coherent integration gain equals the number of integrated pulses, and hence the number of frequency steps, clutter performance is proportional to coherent bandwidth even if it is spread over many pulses.

Pulse-to-pulse coherent frequency diversity is difficult to implement. Furthermore, the total integration time often becomes so large that the radar is sensitive to Doppler, which in some cases may be undesirable. Thus it may be preferable to employ noncoherent, rather than coherent, frequency diversity. As far as signal detection in the clutter is concerned, the only difference then is the smaller gain from integrating N pulses noncoherently.

Another limit on the coherent bandwidth if the target breakup that occurs if the range bin is made smaller than the target spread. Since the cross section of the individual scattering centers will be smaller than the combined cross section of the target as a single unit, the breakup will be undesirable if the problem is target detection rather than signature analysis. We may elect to use a large coherent bandwidth, causing the target to spread over K range bins, but then noncoherently superpose the returns from K consecutive range bins. If the target cross section is uniformly distributed in range, the signal-to-clutter ratio per range bin is independent of K, so that the approach leads to a target detectability improved by a factor equal to the gain from integrating noncoherently K pulses. The use of noncoherent bandwidth thus has resulted in further clutter smoothing. The process is entirely equivalent to noncoherent pulse-to-pulse frequency diversity, and it was shown earlier that it does not matter whether the frequency band is spread over several pulses or is concentrated within a single pulse. It matters only whether it is used coherently or noncoherently. The number of effective integrations that determine the noncoherent gain equals the number of integrated pulses times the factor K by which the pulse bandwidth is subdivided into noncoherently combined parts.

It is worthwhile to note that the conclusions apply also if range ambiguities are included in the considerations. If one attempts to increase the integration gain by raising the pulse repetition rate beyond the unambiguous value, the added integration gain is offset by the clutter contributions through the range ambiguities. Suppression of this clutter requires that the ambiguities be "resolved" by pulse-to-pulse frequency shifting, which is the same as frequency diversity.

We now consider the case where Doppler effects are significant. If the target has a relative Doppler shift with respect to the clutter, the problem of clutter suppression is one of achieving adequate Doppler resolution, with signal bandwidth being of secondary concern. If the Doppler spread of the clutter encompasses the target Doppler, Doppler resolution simply becomes an additional means of clutter suppression. The Doppler spread $\Delta\nu$ of the clutter defines the clutter correlation time as $T_c = 1/\Delta\nu$, and if the pulse repetition period is made no smaller than T_c, an integration gain is obtained without the use of frequency diversity. In other words, Doppler resolution takes the place of frequency diversity and conserves system bandwidth. However, the clutter performance still can be improved by decreasing the repetition period below T_c and using frequency diversity for clutter decorrelation. The situation is the same as discussed above, except that the initial detection performance before the incorporation of frequency diversity is improved through Doppler resolution.

6. BEAM-SHARPENING RADAR

The system to be discussed first is the conventional terrain-mapping radar, where the beam is oriented at a right angle to the flight vector and the PRF is chosen so as to prevent range ambiguities. With respect to the problem of detecting a particular target on the ground, the task is one of achieving adequate signal-to-clutter ratio, as discussed in the preceding section. A large pulse bandwidth serves to subdivide the ground swath into narrow ground-range bins, limiting the clutter space within which the target must be seen. Further clutter suppression is obtained through coherent pulse integration.

The agreement with earlier discussions can be verified by calculating the PRF from the requirement of pulse-to-pulse decorrelation of the clutter. Since the clutter is due to the terrain, the Doppler spread of the clutter equals the variation of the Doppler over the azimuth beamwidth θ_{az}. If the platform flies with velocity v, the range rate at the edge of the beam is given as $\dot{R} = v \sin \theta_{az}/2 \approx v\theta_{az}/2$ for small bandwidths. With the beamwidth in terms of the real antenna length L given as $\theta_{az} = \lambda/L$, the Doppler spread can be found as

$$\Delta\nu = 2(2\dot{R})/\lambda = 2v/L \tag{7}$$

The clutter correlation time then is

$$T_c = 1/\Delta\nu = L/2v \tag{8}$$

Since the lateral beamwidth at range R is $\theta_{az}R = \lambda R/L$, the time it takes the target to traverse the beam, or the total integration time, is

$$T = \lambda R/vL \tag{9}$$

From Eqs. (7) and (8), the number of integrated pulses becomes

$$N = T/T_c = (\lambda R/L)/(L/2) \tag{10}$$

Since L/2 is the azimuth resolution of the synthetic beam, the last result shows that the number of pulses integrated, and hence the integration gain, equals the beam compression ratio. The signal-to-clutter ratio thus is given by the value in the absence of pulse integration (as determined by the pulse bandwidth) times the gain from coherent integration. A further increase of the PRF above the unambiguous value requires the use of frequency diversity, either coherent or noncoherent. In the first case, the enhancement of the signal-to-clutter ratio is proportional to the bandwidth increase through diversity, whereas with noncoherent diversity the gain is smaller.

The differences between coherent and noncoherent pulse integration are more important when the radar is to obtain a map of the terrain. Whereas it is immaterial where the interference originates when the problem is target detection, it is not so with terrain mapping. With coherent pulse integration, the interference is reduced by the formation of a range-Doppler resolution cell, which with the system geometry under consideration translates into a range-azimuth cell. This fact permits the mapping of the cross-section

distribution in range and azimuth. With noncoherent integration, azimuth resolution is that of the real beam, and the enhancement of target detectability is through smoothing of the interference, but one that comes from an area corresponding to the azimuth width of the real beam and the range resolution cell.

As a last system, we consider synthetic aperture radar in the "inverse" mode, where a stationary ground radar observes a target as it passes through the beam. If a target can be modeled as a set of fixed scattering centers, the operation is analogous to the ground mapping radar, and beam compression occurs along the target track. However, the system is of particular interest for the smooth, man-made objects for which the apparent scattering centers often shift with a change in the aspect.

The problem can be illustrated on the basis of a perfect conducting sphere. Since it does not matter whether the radar is assumed stationary and the target moving, or vice versa, we can represent the situation as in Fig. 5. The sphere is stationary and the radar is moving along a straight line, with measurements taken at times $t_1, t_2, \ldots, t_n$. As seen from the figure, the scattering center shifts along the surface of the sphere, and the measured range change does not correspond to the motion of the object as a whole.

In the case under consideration, synthetic aperture processing must be adapted to take into account the shifting of the scattering center. One of the difficulties of so doing is to devise a processor that can economically determine which of the possible motions fits the observed signal sequence best. This practical question will be ignored here. We are interested in the theoretical limitations of the method.

Resolution performance for a more general (nonconstant range rate) motion of the scattering center can be analyzed by means of the trajectory diagram for the reflecting point [4]. Since the only motion-related quantity that can be measured is the reflector range as a function of time, we draw the suitably scaled range R(t) as a function of time, as indicated by the solid curve in Fig. 6. We refer to this curve as the reflector trajectory in the range-time plane. The radar signal now may be represented by photons emitted at each zero crossing of the rf signal, or at a multiple thereof. The photons move on trajectories that are straight lines, reversing the direction upon reflection. The timing of the photons arriving at the receiver represents the law for the zero crossings of the received signal. Thus they indicate the phase "modulation" imposed on the radar waveform by the target motion.

In a matched-filter or correlation receiver, the phase function of the filter, or of the reference waveform in the correlator, must match the spacing of the descending lines of the grid in Fig. 6. However, it is clear that such a receiver also will be matched to motions along the dashed curves in Fig. 6 since a target moving along one of these curves also will impose the same phase modulation on the radar waveform. As follows from the spherical geometry of the assumed target, the entire family of possible trajectories (if the lines of the grid in Fig. 6 are chosen more densely) corresponds to the motion of the reflecting point on the surface of a sphere of arbitrary diameter. This includes the limiting case of a point reflector at the center of the sphere.

The interpretation of these findings is that synthetic aperture processing, although it could determine the fact that the reflecting center is moving on the surface of a sphere, could not measure the diameter of the sphere. It takes high range resolution, and hence a large coherent signal bandwidth, to resolve this ambiguity and measure the diameter of the sphere to an accuracy commensurate with the signal bandwidth.

A large signal bandwidth also is essential when coherent processing is to be performed on a target with translational as well as rotational motion. As an example, we consider the ground-based radar with a spinning target moving through the beam. This time, the target is to be modeled as a set of fixed point scatterers. As shown in Fig. 7, the translational motion can be accommodated by assuming the radar to move, so that a given scattering center on the target rotates about a fixed point. Let the distance from the center of rotation be r and the angular velocity be ω.

The parameter that determines the target-imposed phase modulation is range as a function of time, and from Fig. 7 we have

$$R^2 = (R_o + r \sin \phi)^2 + (vt - r \cos \phi)^2 \tag{11}$$

where $\phi = \phi(t) = \omega t + \phi_o$, and v is the target speed. If the beamwidth, or observation angle, is modest, the quantity vt/R_o is small enough to allow the approximations $\sin x \approx x$, $\cos x \approx 1$. Expansion of the squares in Eq. (11), substitution for ϕ, and use of the approximations gives

$$R^2(t) = R_o^2 + r^2 + (vt)^2 + 2R_o r \sin\left(\omega t - \frac{vt}{R_o} + \phi_o\right) \tag{12}$$

Since both vt and r will be small compared with R_o, we can simplify Eq. (12) into

$$R(t) = R_o + \frac{r^2}{2R_o} + \frac{(vt)^2}{2R_o} + r \sin\left(\omega t - \frac{vt}{R_o} + \phi_o\right) \tag{13}$$

Examination of this result shows the following. There are two time-dependent terms. The first is quadratic in t and is due to the translational motion of the target. It introduces a linear FM into the signal and allows beam compression along the target track. There also is a sinusoidal variation in the range, but it depends on both the rotational and the translational motion. The problem is to separate these two components in the processor.

The solution again requires use of an adequate signal bandwidth. A large signal bandwidth permits determination of R_o to the required accuracy. This allows the elimination of the linear FM term without an

expensive search for the correct FM slope. In addition, it permits the isolation of the translational component in the sinusoidal variation of range, so that the component remaining is due to the rotation. This is even more important, and amounts to more than an equipment simplication, when the processor must resolve many different scattering centers on the target. The use of high range resolution means that resolution on the basis of motion can be performed for each range bin at a time, eliminating the type of interference produced if the only criterion for target resolution is poor correlation. A filter that is not matched to the received waveform may not yield a high output peak, but it will pass all the energy falling within its passband. This means that signal bandwidth must be used for background suppression.

7. CONCLUSIONS

The preceding discussions have shown the fundamental significance of bandwidth for all types of resolution problems. Coherent use of signal bandwidth restricts the number of interfering scatterers to those within a given range bin, although difficulties may appear when combined resolution in range as well as Doppler is attempted. Noncoherent bandwidth use results in the smoothing of fluctuations, be it those of the target cross section, high-resolution range profiles, or of the clutter and interference. It has little significance on radar performance whether the bandwidth is provided within a single pulse or spread out over many pulses, in the form of frequency diversity.

REFERENCES

10-1. Woodward, P. M., "Probability and Information Theory, with Applications to Radar," Pergamon Press, London, 1953.

10-2. Berkowitz, R. S. (ed.), "Modern Radar: Analysis, Evaluation, and System Design," Wiley, New York, 1965.

10-3. Cook, C. E. and Bernfeld, M., "Radar Signals, an Introduction to Theory and Application," Academic Press, New York, 1967.

10-4. Rihaczek, A. W., "Principles of High-Resolution Radar," McGraw-Hill Book Company, New York, 1969.

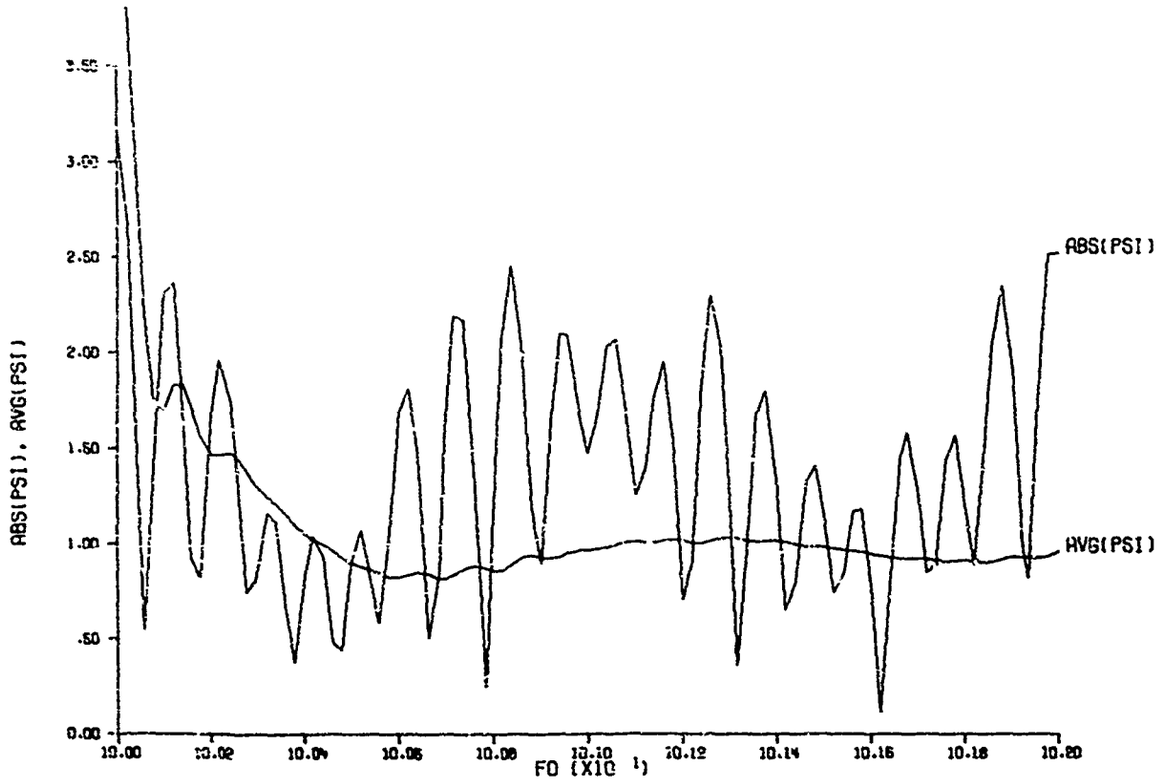


Fig.1 Target cross section as a function of frequency, and averaged cross section

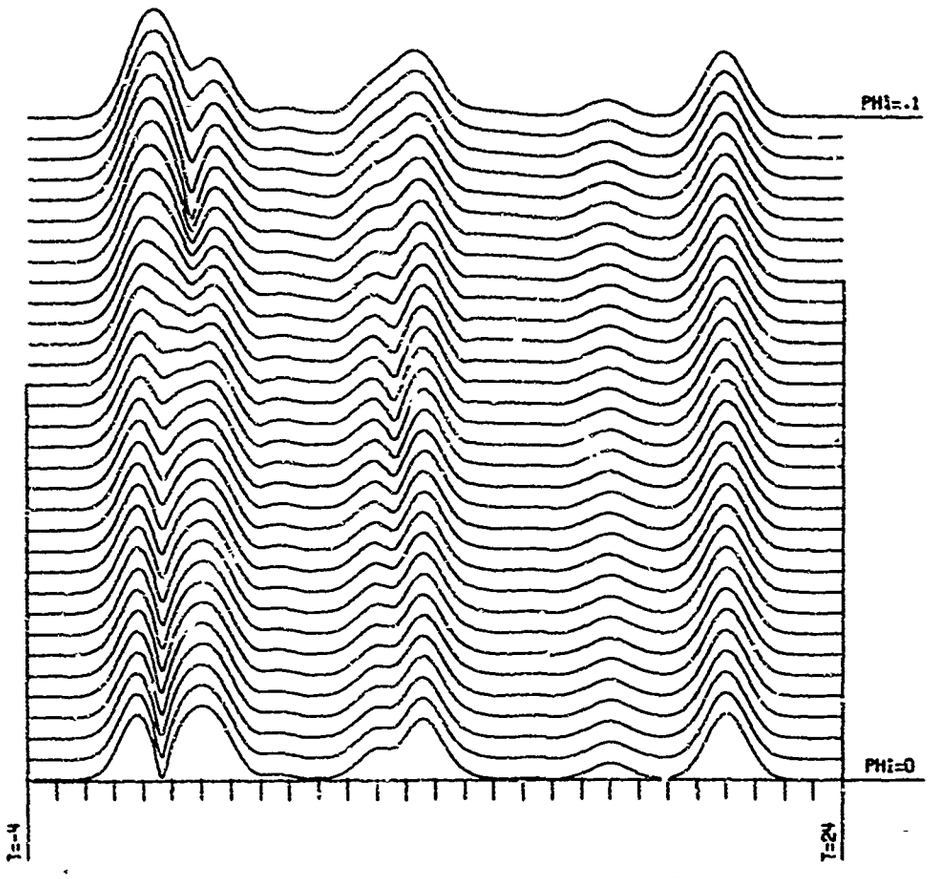


Fig.2 Succession of range profiles for a tumbling target

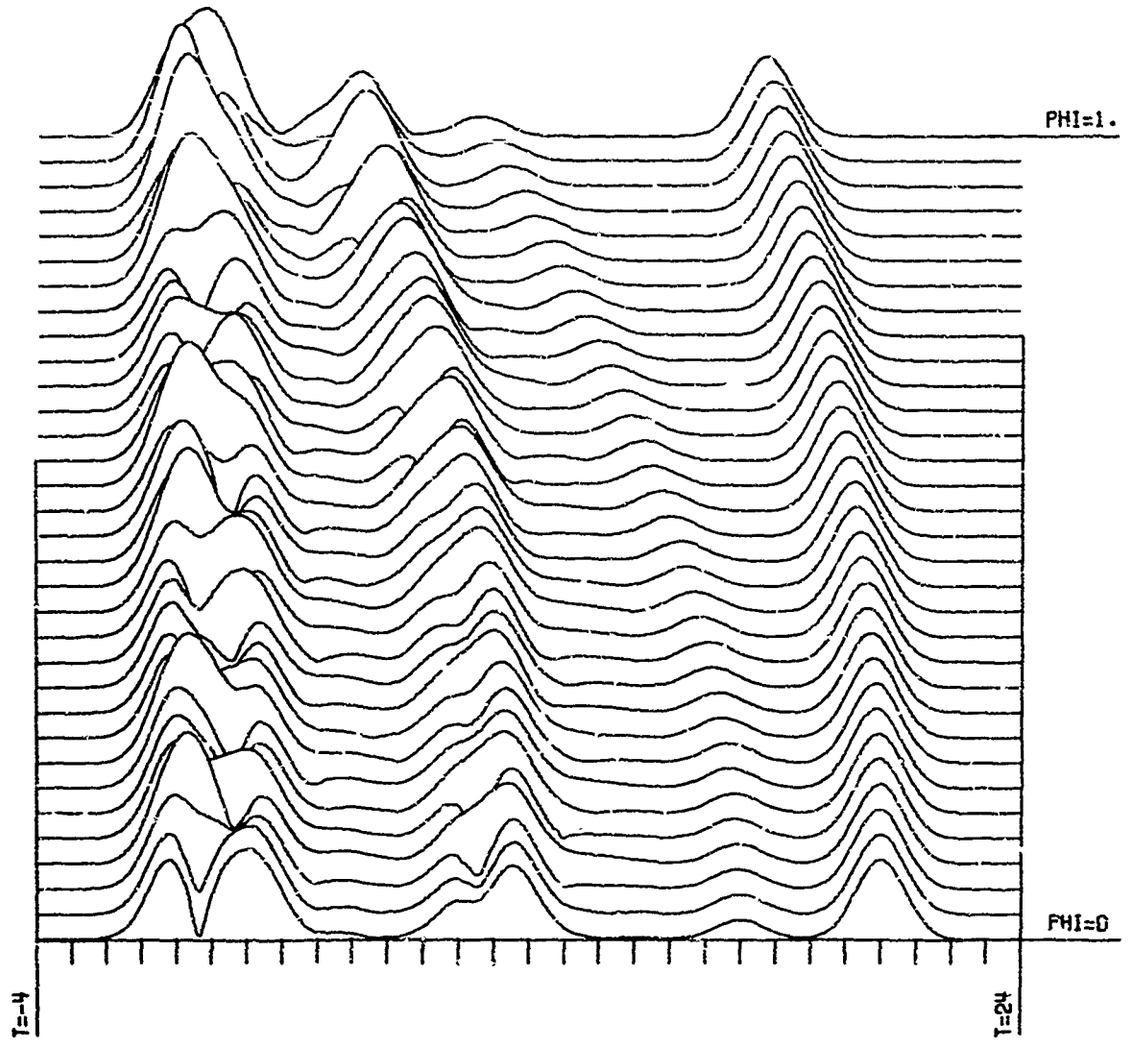


Fig.3 Succession of range profiles recorded over a larger aspect angle change

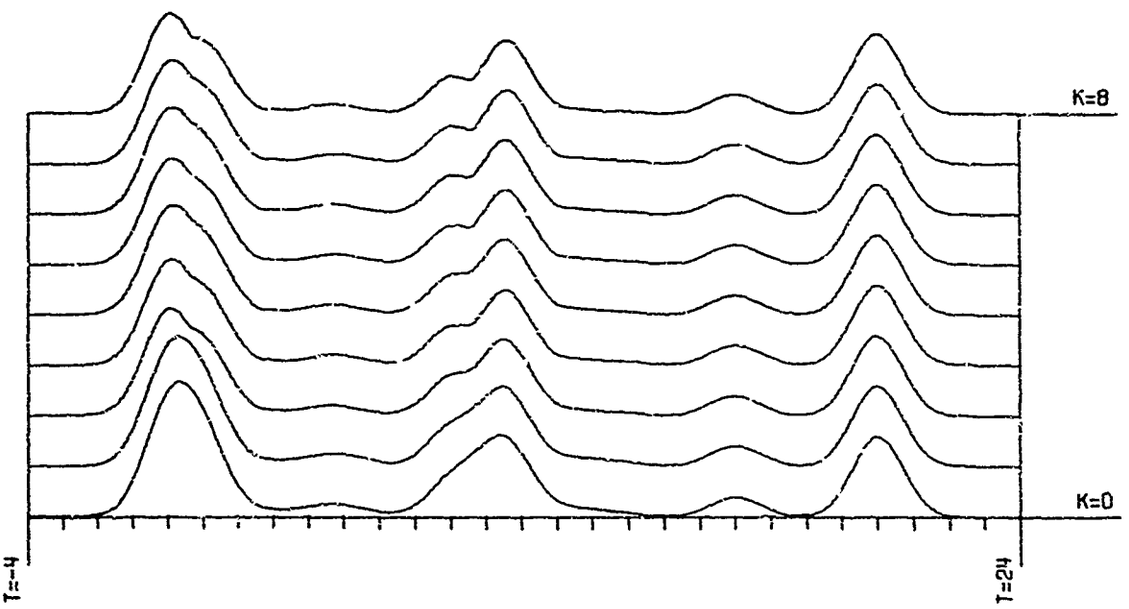


Fig.4 Range profile (bottom curve) and the average over several range profiles measured at different frequencies

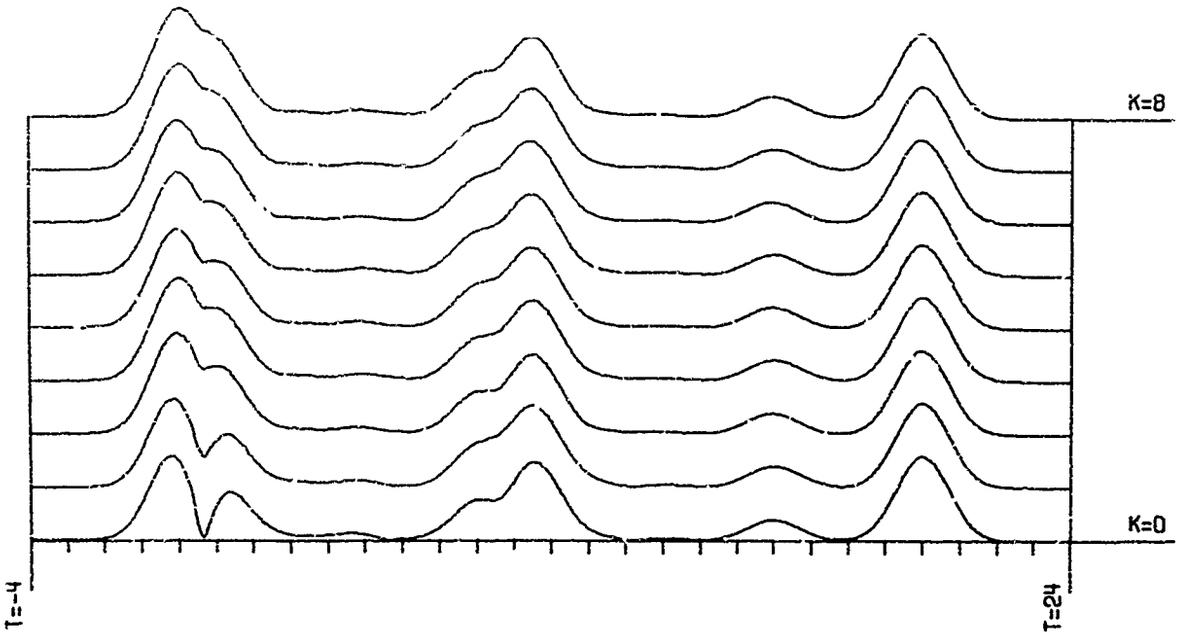


Fig.4a Same as Figure 4, except that the initial range profile (bottom curve) was taken at a different frequency

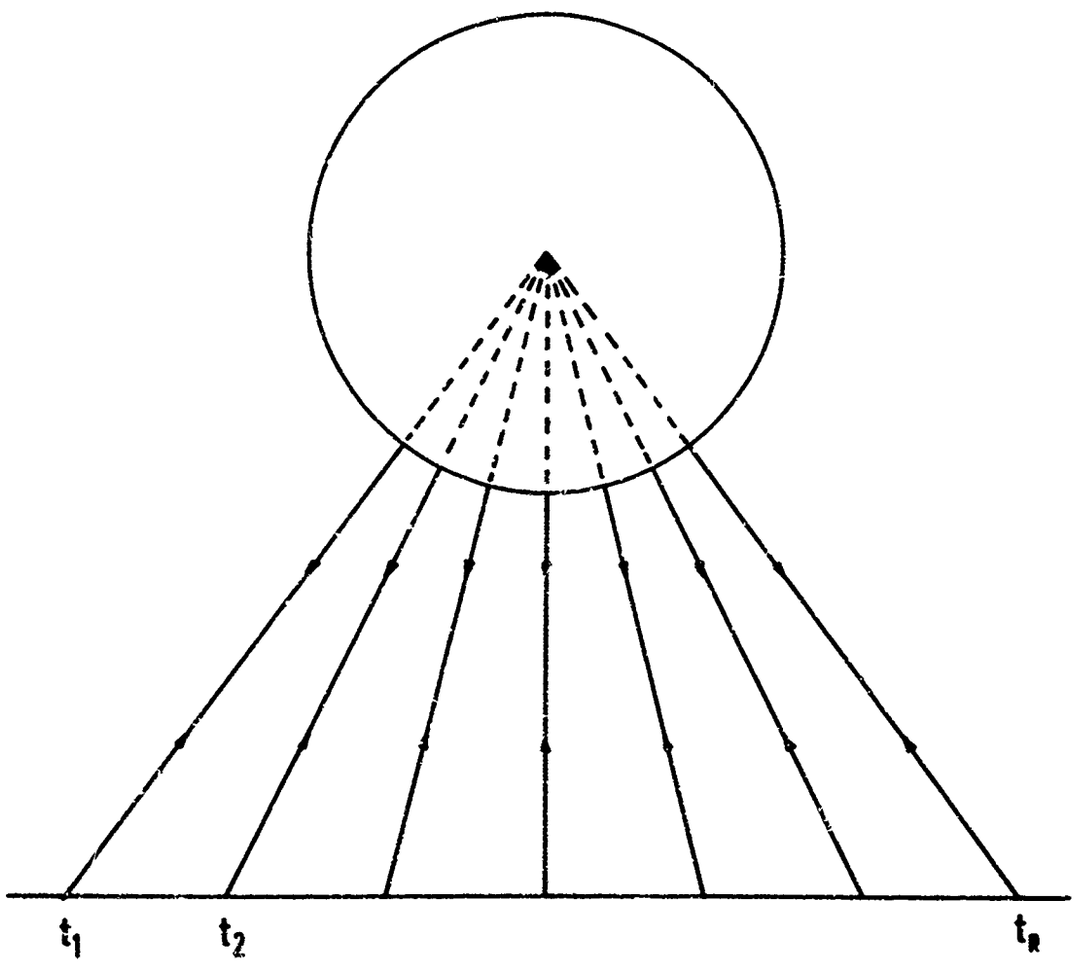


Fig.5 Geometry for synthetic aperture resolution on a conducting sphere

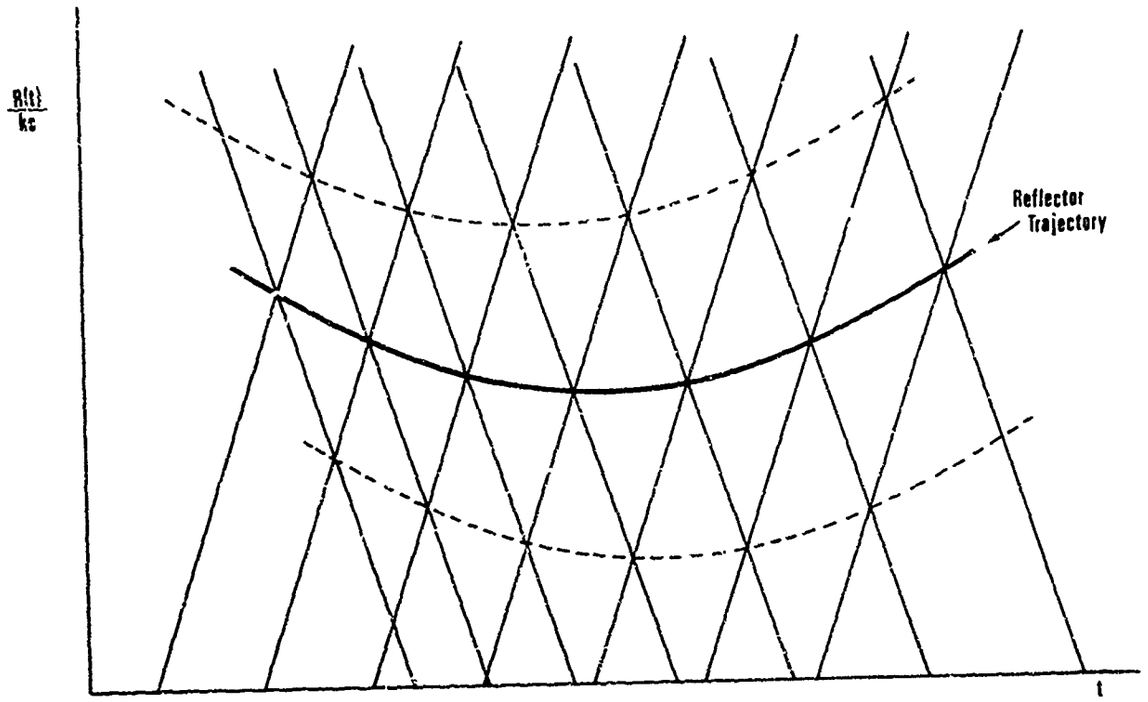


Fig.6 Trajectory diagram for the case of Figure 5

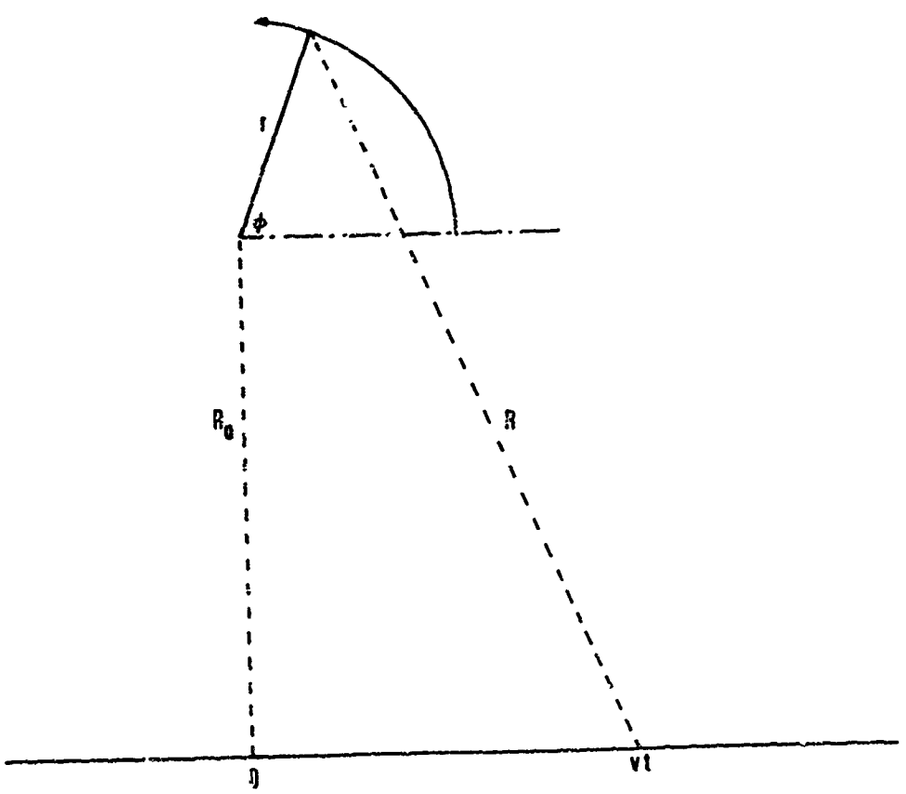


Fig.7 Geometry for a synthetic aperture radar operating on a spinning target

INTEGRATED SUBSYSTEMS ON COMPOSITE SUBSTRATES

by

M.Lemme, P.Holst and H.J.Schmitt

Philips Zentrallaboratorium GmbH, Laboratorium Hamburg
Hamburg, Germany

INTEGRATED SUBSYSTEMS ON COMPOSITE SUBSTRATES

M. Lemke, P. Holst, H. J. Schmitt
Philips Zentrallaboratorium GmbH, Laboratorium Hamburg
Hamburg, Germany

SUMMARY

A new substrate technology for microwave integrated circuits is described, which allows a planar integration of magnetic system functions. The approach uses the fact that some substituted ferrites, with Curie points sufficiently below operating temperature of the system, provide an excellent, low-loss nonmagnetic host substrate which can be sintered together with magnetically active zones of similar but unsubstituted ferrites. This technique offers the advantage of reduced coupling and system losses compared to the technique using an all-magnetic ferrite substrate, and avoids the mechanical complexity of conventional dielectric substrates with attached or plug-in ferrite devices.

Composite ferrite substrates have been successfully employed for miniaturized latching ferrite phase shifters for phased arrays. The technique has also been used in the development of miniature X-band radar front ends with integrated isolators and circulators and for a pocket-size Doppler radar.

1. INTRODUCTION

Today's conventional microwave circuitry and plumbing is increasingly being superseded by integrated subsystems in planar technology, using microstrip or microslot lines on a dielectric substrate combined with hybrid integration of active solid state devices. The integrated system approach has already been used to reduce space, weight and costs in several radars, such as MERA [1] and other airborne systems as well as in various other equipment for low-power communication gear, ECM circuits etc.

A typical subsystem, such as a radar front end may include 10-20 passive elements in the form of microstrip line sections arranged on a single dielectric substrate of a few square inches of Al_2O_3, sapphire or quartz, and typically 2-5 active elements for the mixer, local oscillator and the AFC circuit. Since the entire passive circuitry is fabricated in a photolithographic process large numbers of identical circuits may be produced in an economical way thus making this technique particularly attractive for phased array systems with large numbers of individually fed radiators, for standardized receiver front ends, or for compact expendable warning radars.

All these systems in general also require a number of ferrite elements, such as circulators, isolators and phase shifters etc. The integration of such ferrite elements into a conventional integrated system on alumina substrate poses problems:

i) Ferrites cannot readily be combined with alumina substrates. To accommodate circulators etc. in a complex integrated system, these elements either have to be attached as separate "subsystems" with additional plug connections, and thus some of the advantages of the integration technique are lost, or alternatively, a hole must be provided in the hard ceramic substrate for the ferrite "puck". In practice small air gaps or cracks can hardly be avoided in this process, causing electrical discontinuities and mismatches.

ii) If the whole system is built up on the ferrite, one faces the difficulty that the choice of material properties cannot be simultaneously optimized for the various system functions. These are precisely the same problems as encountered for monolithic microwave IC's on semiconducting materials which ultimately led to the development of the hybrid approach [2]. Also additional losses occur due to the magnetic loss tangent, and an undesirable and generally intolerable temperature sensitivity is introduced due to the temperature dependence of magnetic parameters. Moreover, the circuit properties of the passive elements may be seriously perturbed by the magnetic flux extending from static or switched fields applied to adjacent ferrite elements.

Recent progress in materials technology now offers a solution to these problems through the successful development of composite ferrite substrates, which allow a complete integration of most ferrite elements for nonreciprocal devices, switches and steering functions etc. in integrated microwave systems. For these composite substrates "nonmagnetic" ferrites, i.e. substituted ferrites with Curie points sufficiently below operating temperature of the system, are used as a low-loss host substrate instead of alumina. These "nonmagnetic" ferrites have such thermal properties that they can be sintered together with magnetically active zones of similar, but unsubstituted, or less substituted ferrite material to form smooth, vacuum-tight composite substrates which may be metallized and etched as a unit to provide maximum reliability and reproducibility of the circuit [3].

Fig. 1 shows the application of this integration technique to a miniaturized multibit latching ferrite phase shifter for an X-band scanning array antenna. The upper left picture shows for comparison the "very old" technique of a 4-bit waveguide phase shifter, the upper right the miniaturized integrated microstrip version of the

same device with "plug-in" ferrite pucks in an Al_2O_3 substrate. The small holes in the ferrite pucks are to take up the individual address wires. Here considerable difficulties are encountered in the transition from the host substrate to the ferrite pucks: if the pucks are inserted before metal deposition some metal is likely to the deposited in cracks and gaps. If the pucks are separately prepared and inserted after the patterns have been etched, 8 bonding steps are required to connect the devices. The lower left picture shows a composite substrate, after polishing, with the 4 magnetically active zones sintered together in a "nonmagnetic" ferrite host substrate. On the lower right the completed device with the etched meander line pattern is shown.

A micrograph of the transition zone in a composite garnet substrate is shown in Fig. 2. Nonmagnetic and magnetic zones show up on account of the slightly different grain size in an otherwise perfect transition. In particular, the discontinuity in the dielectric constant is also small, due to the similar dielectric characteristics of the dense materials.

2. TECHNOLOGY

The technology of fabrication of composite substrates has been reported before [3]. In principle, dry milled powders both of the "nonmagnetic" ferrite and the magnetic ferrite are mixed with a small amount of polyester resin and separately filled into a die and then pressed and sintered in the final geometry. The density achieved in this process is comparable to that of bulk material. So far, composite substrates have been made in the spinel system from Al,Zn substituted MnMgFe-spinels, and in the garnet system from Al,Ca,Si substituted yttrium iron garnet as host substrates, both with Curie temperatures below -40 °C, combined with different magnetic ferrites of their respective family. There still exists a certain limitation in the selection of materials suitable for composite substrates because the sintering behavior, especially the thermal expansion must be reasonably well matched over a wide range of temperatures. To a certain degree this can be influenced by altering the prefiring temperature of the powdered materials, and further progress with respect to available compositions is likely in the near future.

A different and potentially more elegant method starts from extruded plasticized ferrites. These are rolled into foils, with magnetic zones (also as a plastic material) placed at the desired locations. These are then sintered together after slowly burning out the plasticizer. While useful for simple "nonmagnetic" or simple "magnetic" substrates, this method so far has not resulted in vacuum-tight transitions in composite substrates.

Further processing of the composite chips follows the steps applied for the fabrication of conventional thin film integrated microwave circuits. A particular advantage arises from the fact that fine polish (to <0,1 µ roughness) is much more easily achievable for ferrite substrates than for alumina, due to the low porosity and hardness of ferrites. This is important because losses of microstrip circuits tend to rise rather strongly for a surface roughness of the order of a skin depth (~0,5 µ at X-band frequencies). No mechanical defects have been found on composite substrates during the following vacuum deposition cycle, with temperatures going up to about 300 °C. Mechanical adhesion of deposited copper layers (after "chrome flash" or 200 Å Vd intermediate layer) is excellent on all properly cleaned ferrites tested so far.

3. SUBSTRATE PROPERTIES

The electrical and environmental properties of the "nonmagnetic" ferrite host substrates are similar to those of alumina. Dielectric loss tangents of $<10^{-4}$ are consistently achieved in bulk material by refined ferrite technology. The unloaded Q_0 of test resonators is about 350 at X-band frequencies and thus compares well with similar values for resonators on Al_2O_3 [4]. With normal copper circuit patterns the measured attenuation at X-band is about 0.07 dB/cm in a 50 Ohm line. The attenuation as a function of frequency is shown in Fig. 3. Also shown is a theoretical result for this geometry which gives rise to the expectation that losses can be even further reduced. However, the origin of the remaining discrepancy is unresolved at present.

The dielectric constant ε_r of the host substrate is slightly higher (~11 for spinels, ~16 for YIG) than that of Al_2O_3 (~9.7). The dielectric constant is substantially frequency-independent in the microwave range, at least below 18 GHz. The slight dispersion shown in Fig. 3 for the normalized guide wavelength is entirely due to the propagation characteristics of the microstrip line. This is also evident from the theoretical plot of the guide wavelength versus frequency, calculated for a constant value of ε_r.

The microstrip guide wavelength is not very sensitive to the environmental temperature. In the temperature range from -40 to +60 °C the variation is less than 1%, similar to the behavior of alumina, whereas for magnetic ferrites the propagation is much more strongly affected by temperature due to changes in the permeability, Fig. 4.

Peak power handling capability of the "nonmagnetic" ferrite substrate is also of the same order as that of other dielectrics. Thus in composite substrates the maximum power level will be governed by the high-power behavior of the magnetically active zones. However, within the range of usual microstrip applications with peak

power levels up to 1-2 kW magnetic instabilities will not normally arise in the garnets and spinels used in composite substrates. Moreover, if required the peak power level can be increased still further by known techniques of coping with relaxers. High average power, on the other hand, leads to a rise in temperature of the devices at a rate depending on the unavoidable insertion losses. For circulators, phase shifters etc. the limitation is of the order of 10 Watts CW at X-band frequencies, for temperature rises of about 50 °C.

In addition to their higher brittleness, the only minor disadvantage of "nonmagnetic" ferrite substrates is the poor thermal conductivity, which is about an order of magnitude lower than for Al_2O_3. Quantitative data on both mechanical shock or thermal shock behavior are not yet available, however.

4. INTEGRATED RADAR SUBSYSTEMS

For radar, the composite substrate technology offers interesting possibilities especially for phasing systems of electronically scanned antenna arrays. The X-band 4-bit digital phase shifter as shown in Fig. 4 with dimensions of only 3x3x0.06 cm^3 is readily integrable with the antenna system at locations near the antenna elements. This holds for systems with directly fed radiating elements and retroreflective types of antennas with "optical" power distribution from a common source. In the particular device shown the meander lines are designed for phase shifts of 12, 24, 48 and 96° respectively, offering 16 phasing positions between 0 and 180° differential phase shift. Since only the two remanent magnetization stages of the toroidal ferrite zones (clockwise and counterclockwise remanent magnetization) are used, no holding power is required. This is a clear advantage over PIN-diode phasing schemes. The remanent magnetization of an individual stage is reversed by application of a short (<1 µs) pulse of about 10 A and appropriate polarity. The switching time for the ferrite zones is potentially less than 1 µs with an energy expense of <1 µJoule. In this geometrical arrangement the differential phase shift results from the interaction of the atomic spins in the ferrite with a locally circularly polarized magnetic field of the microwave within the substrate. The resulting phase shift is nonreciprocal. In a monostatic radar system this requires a separate reverse switching operation between transmit and receive operation, to obtain identical antenna bearings for both cases. Because of the fast switching performance, this should present no particular problem. In principle also reciprocal phase shifters can be fabricated by a wider spacing between adjacent meander lines. However, the figure of merit for reciprocal devices is only 40...140° phase shift/dB attenuation and hence much lower than the figure of merit for nonreciprocal devices (200-300°/dB). The temperature drift of the multibit phase shifter can be maintained at less than ±5% of the differential phase shift over a range from -20 to +60 °C by selection of the composition of the magnetic zones. In the garnet system this may be achieved by using Gd,Al substituted YIG for which the saturation magnetization is substantially independent of temperature over a fairly broad range. The phase change, which depends primarily on the remanent magnetization, is then only slightly affected by temperature via the anisotropy field, Fig. 5.

An alternative form of the digital phase shifter consists of a single meandering line per antenna element. The desired number of phase steps are generated electronically by partial switching of the magnetization with short current pulses of discrete amplitudes, but below those required for saturation of the material. This design simplifies the address wiring at the cost of increased complexity of the electronic driver circuitry. The figure of merit is comparable to that of the multibit version.

In continuously variable microstrip ferrite phase shifters the adjustment of phase delay is more economically achieved with transverse magnetic fields between the poles of a separate magnet. In this case the magnetically active zone should fill the area between the poles to avoid nonmagnetic gaps.

Two 90° latching phase shifters (possibly on one magnetically active zone) can be combined with a power divider and a branch line coupler to act as a miniaturized latching single pole double throw switch for microwave power. A module for X-band use is shown in Fig. 6. In the present configuration it provides about 20 dB discrimination between the symmetrical output arms of the coupler in a frequency band of 10%. However, the discrimination and the bandwidth may be improved by further optimization. As a single switch it may be used for example to separate transmit and receive operations. In principle, a tree of such modules may be used as a switchable addressing system for multi-input circuits, e.g. for connection to a Butler matrix [5] in an electronically scanned array.

Another important use of composite substrates for radar systems involves its application for miniaturized integrated receiver front ends or frequency converters etc., which generally require a number of isolators and circulators for decoupling and matching purposes. Fig. 7 shows two versions of solid state X-band front ends. The systems differ in the manner in which frequency of the local oscillator is stabilized for use in low noise radar and communication systems. In both circuits the basic part includes a singly balanced mixer with 2 Schottky barrier diodes mounted at the output arms of the branchline coupler. These are followed by low-pass i.f. filters. The mixer noise figure is of the order of 8 dB, if measured with a 1.5 dB i.f. noise. The signal input of the mixer contains microstrip isolators, built up of 3-port Y-circulators with one arm terminated reflectionfree by a vacuum-deposited

thin-film resistor (Cr-Ni). The isolators essentially reduce reradiation of local oscillator power and improve the match as seen from the branch line coupler and from the antenna. In the system shown in Fig. 7a a Gunn effect device is used as local oscillator, and an APC loop is included in the subassembly. For this purpose part of the local oscillator power is coupled to a discriminator built up from a meandering delay line as the frequency-sensitive element and a separate balanced mixer. Any frequency deviation from the center frequency of the discriminator gives rise to a rectified voltage which, after amplification, is superimposed on the Gunn element driving voltage in such a phase as to cancel the frequency deviation. The entire chip with dimensions of 3x4 cm^2 contains 14 passive components, including the circulator, and 5 active devices on one composite substrate. The front end shown in Fig. 7b contains instead provisions to connect a separate quartz stabilized rf signal derived e.g. from a step recovery diode. The filter circuit essentially serves to reject harmonics and spurious signals from the multiplier chain used to transpose the 50 MHz quartz signal into the X-band region. Alternatively a low-power stable rf signal may be used for injection locking of a separate Gunn oscillator. With an injection power of 0.1 mW a pulling range of about 150 MHz has been achieved for an 11 mW Gunn element.

The composite substrate technology has also been used for the development of small all solid state radars, which find a variety of applications either as warning devices, electronic fences, as boat radars or in vehicle traffic control systems. A pocket-size Doppler radar for the detection of moving objects is shown in Fig. 8. The entire rf circuitry is arranged on a 3x3 cm^2 composite ferrite chip with an integrated magnetically active zone for the circulator element. The systems 15 mW of X-band power are generated by a free-running Gunn effect device. A small part of that energy is fed directly into a balanced mixer to act as the local oscillator. The remaining energy passes through the 3-port circulator and is radiated by the attached dielectric rod antenna. Microwaves received by the antenna pass the circulator towards the signal input of the balanced mixer, again formed by the branch line coupler and two Schottky barrier diodes. A small permanent magnet (ferroxdure) placed beneath the circulator disk serves to magnetise the magnetically active zone. Received signals, which are Doppler shifted, are amplified and indicated acoustically. With a simple transistorized amplifier, an antenna gain of about 13 dB and a beam width of 40° a moving man can be detected at a range of about 40 meters. The range can be readily extended by means of a refined electronic trigger circuit and improved frequency stabilization of the main oscillator.

5. CONCLUSION

The use of composite ferrite substrates instead of conventional dielectric substrates for microwave integrated circuits offers the advantages of increased system reliability coupled with potentially lower production costs because of the easier mechanical handling of ferrites. So far, the spectrum of materials which can be combined to form a monolithic substrate with magnetic functions is somewhat limited and leads to necessary compromises where different system operations have to be performed. It is expected, however, that within the next few years through further progress in molecular architecture a much wider choice of compositions will be available to the system designer.

6. REFERENCES

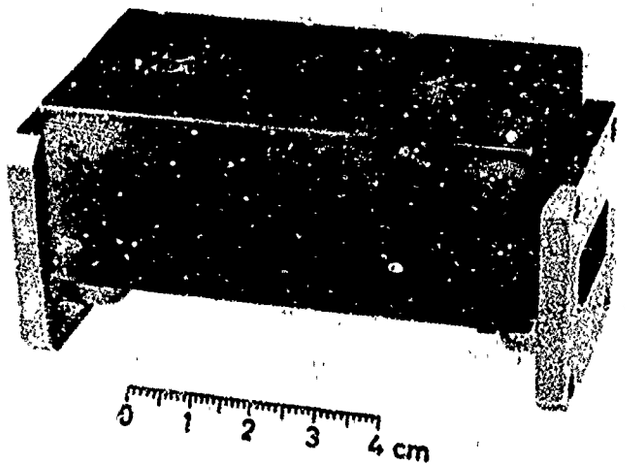
[1] Hyltin, T.M. "Microwave integrated electronics for radar and communication systems" Microwave J., 11, p. 51-55, 1968.

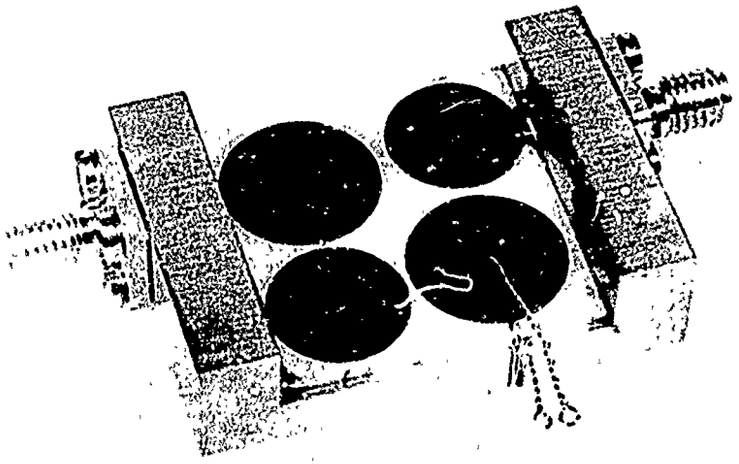
[2] Morony, W.J.: Microwave IC's: part 1 "New problems, but newer solutions" Electronics, p. 100-108, June 1968.

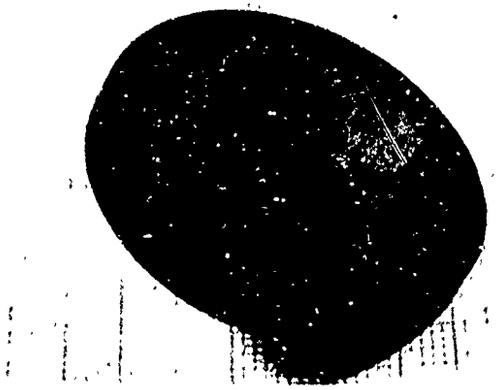
[3] Holst, P. and Lemke, M. "Ferrite Substrates for Microwave Integrated Systems" IEEE Trans. MAG 5, 3, p. 478-480, 1969.

[4] Meister, F.Z. "An evaluation of materials and processes for integrated microwave circuits" IEEE Trans. MTT 16, p. 469-475, 1968.

[5] Butler, J.L. "Digital, matrix, and intermediate-frequency scanning" in: Microwave Scanning Antennas, vol. 3, R.C. Hansen, Ed., Academic Press, New York, 1966.







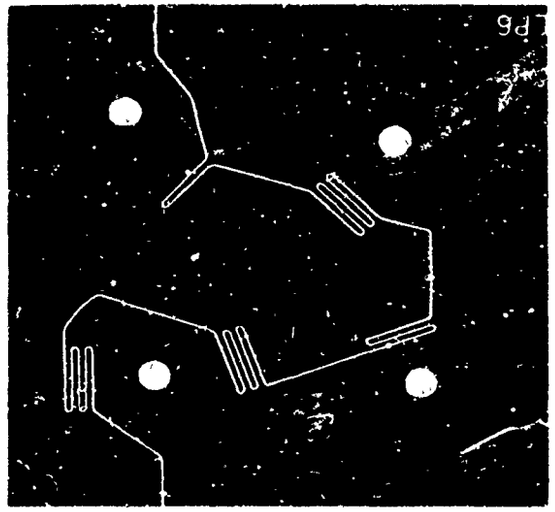


Fig.1 4-bit versions of latching phase shifters

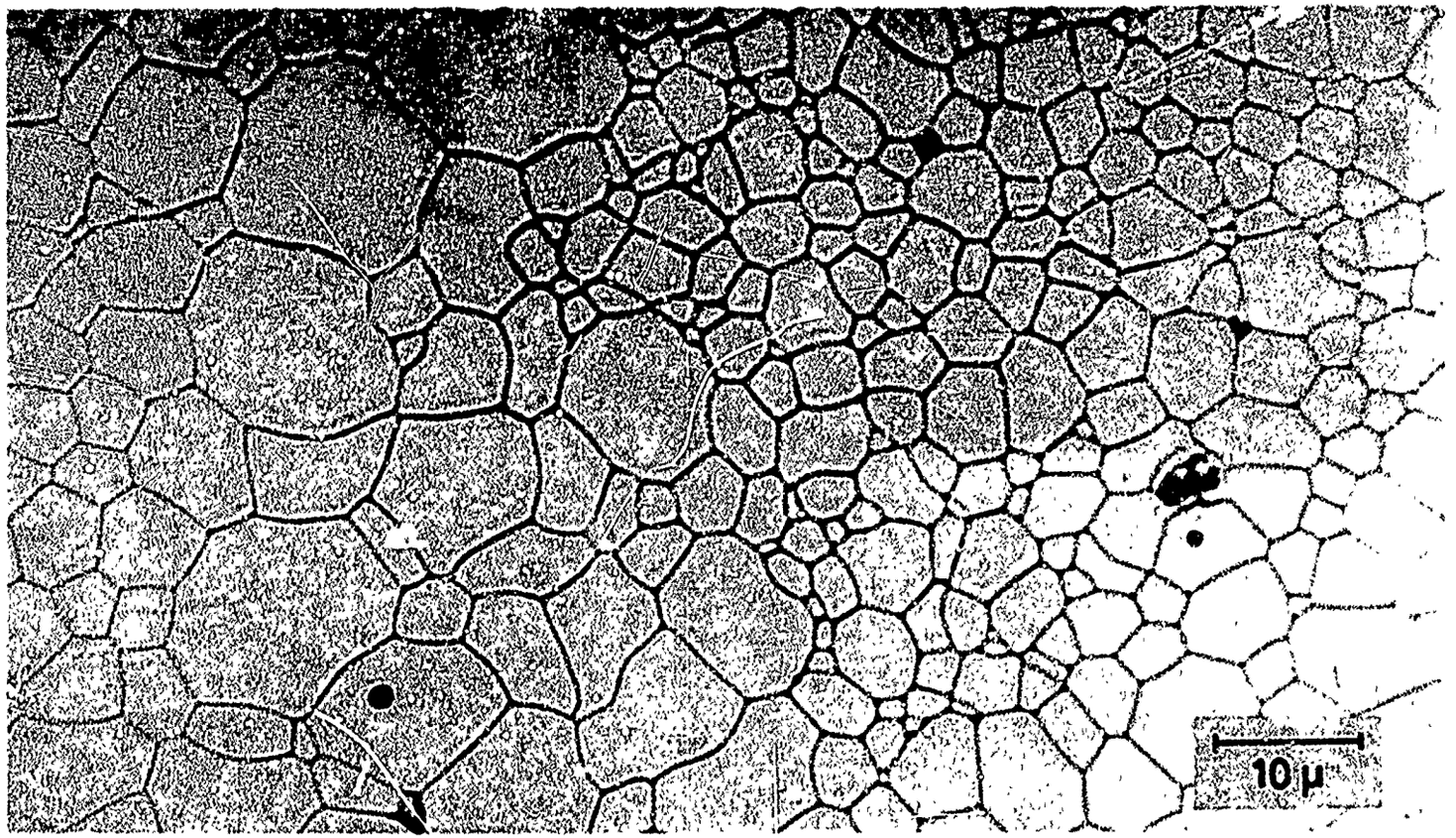


Fig.2 Micrograph of the transition zone in a composite garnet substrate

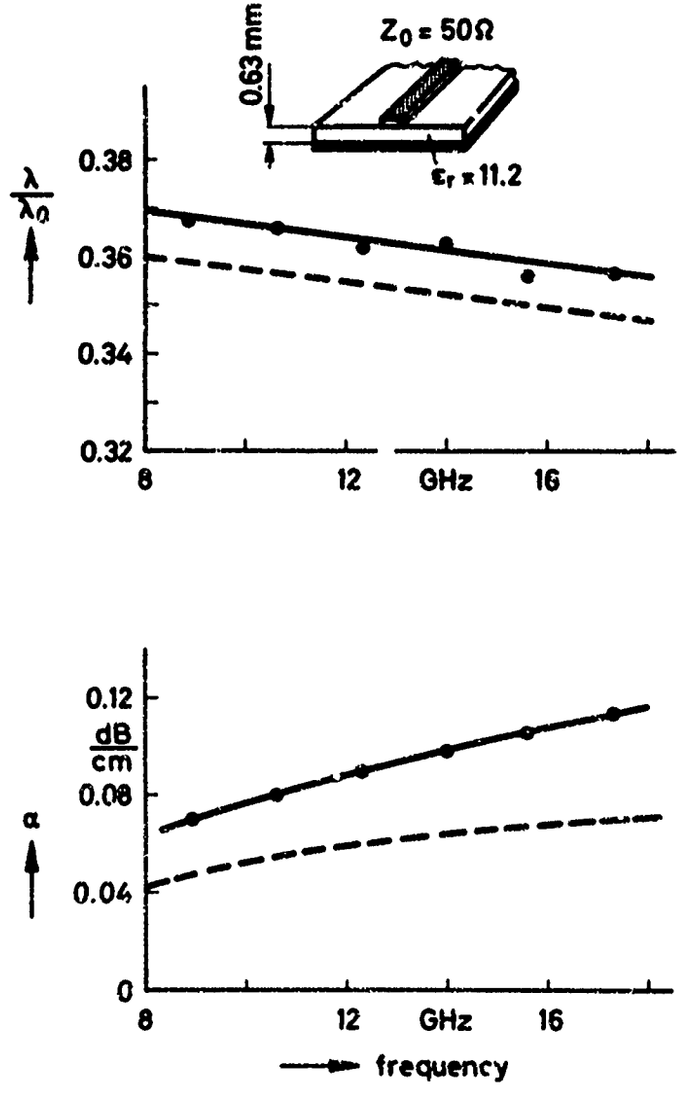


Fig.3 Dispersion and loss characteristic for microstrip-lines
.-.-. measured, --- theoretical

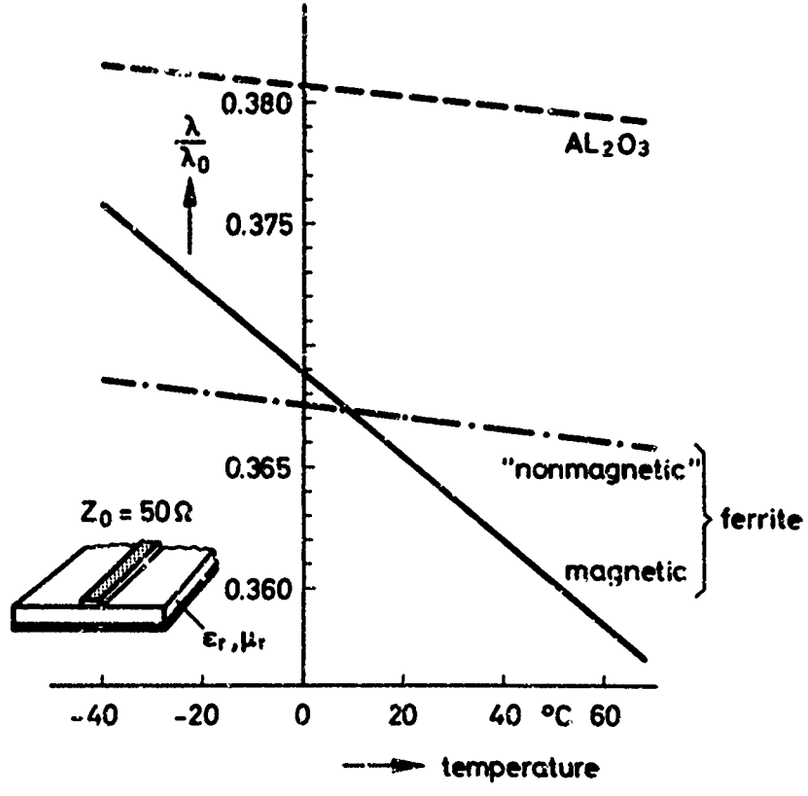


Fig.4 Temperature dependence of microstrip wavelength on different substrates

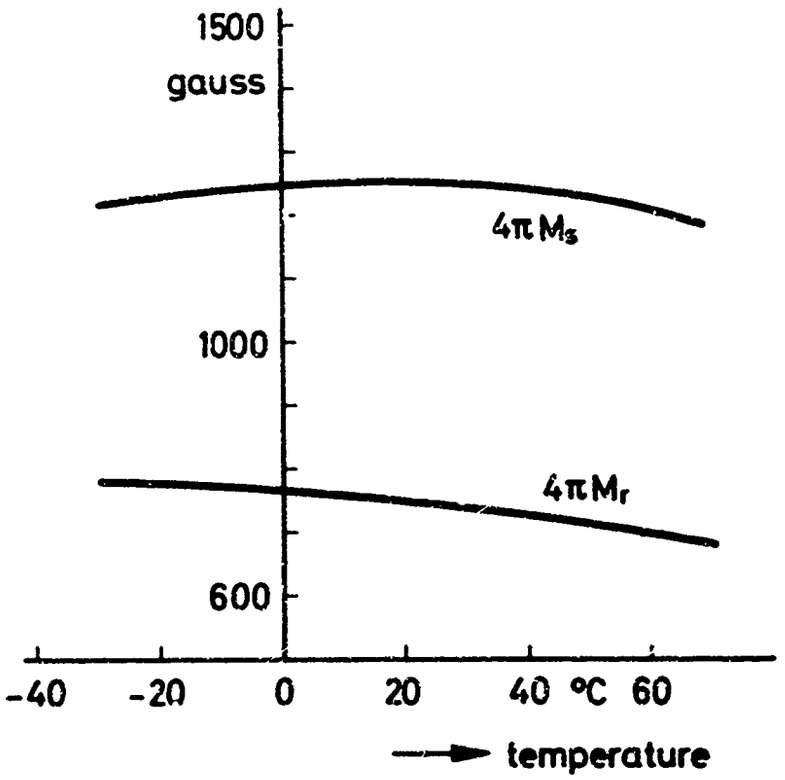


Fig.5 Magnetization of temperature-compensated YIG $Y_{1.89}Gd_{1.0}Tb_{0.02}Sm_{0.09}Fe_5O_{12}$

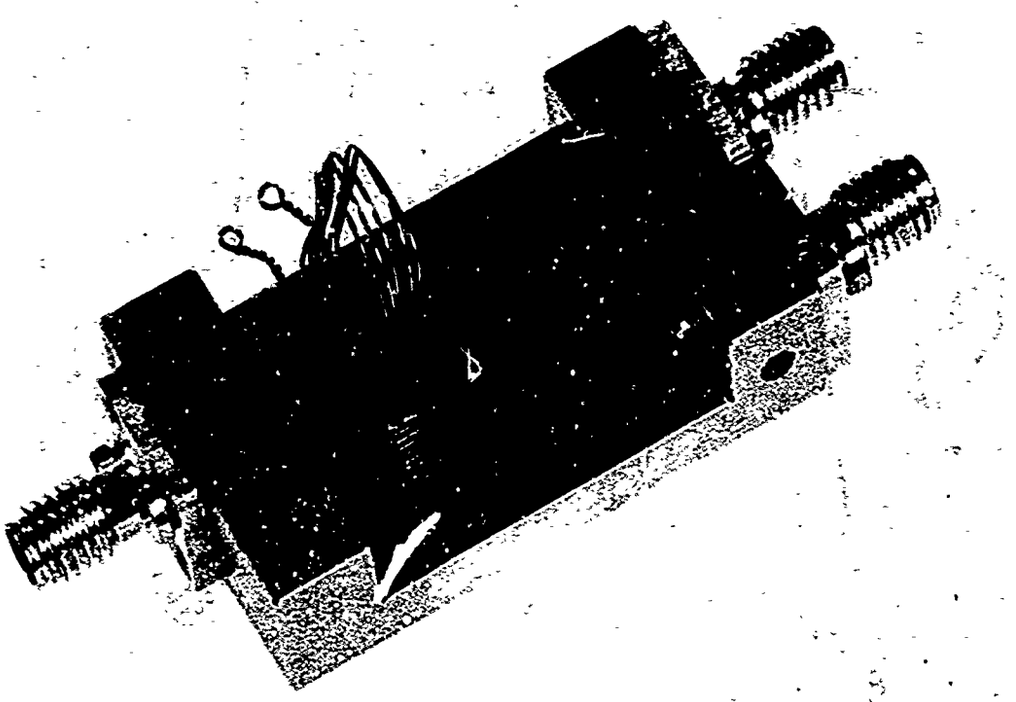
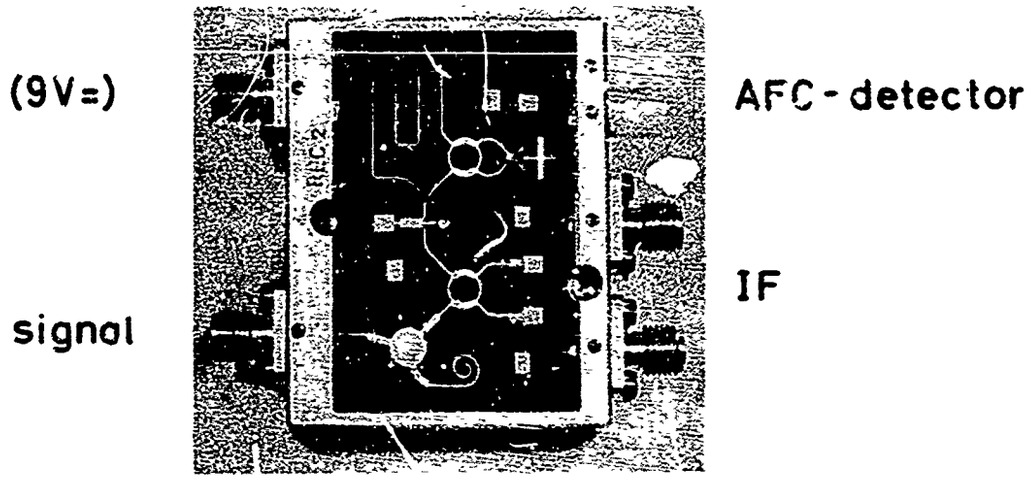
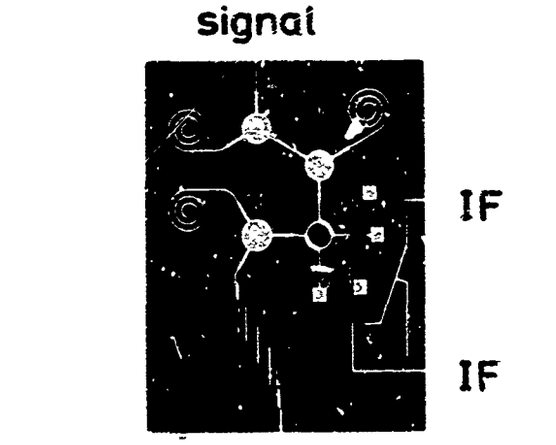


Fig.6 X-band microstrip switch



(a)



(b)

Fig.7 X-band front ends based on "nonmagnetic" ferrite

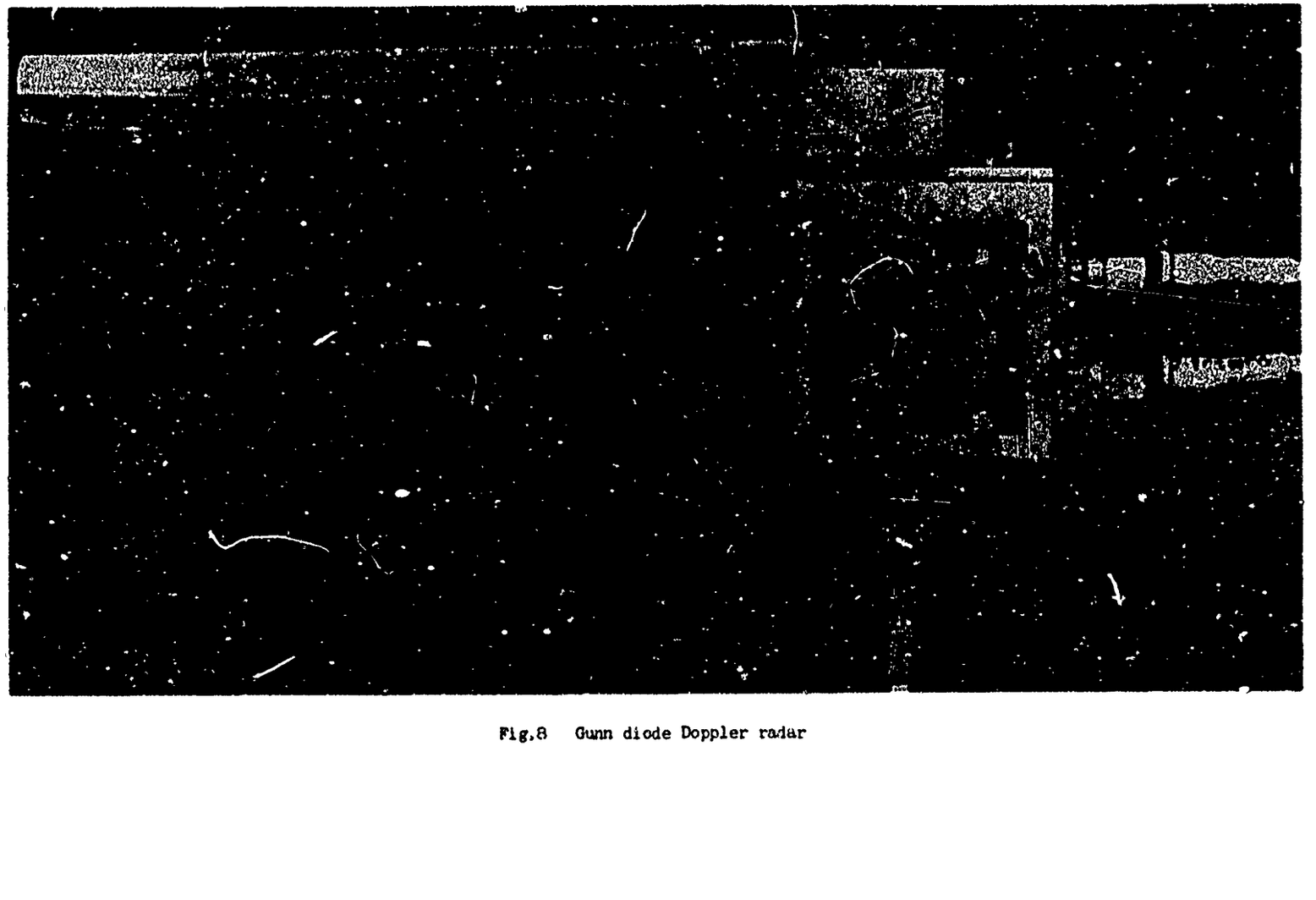


Fig.8 Gunn diode Doppler radar

OPTIMISATION SUR ORDINATEUR DES RADOMES AEROPORTES

par

R.Pierrot

Thomson - C.S.F
92 - Malakoff
France

OPTIMISATION SUR ORDINATEUR DES RADOMES AÉROPORTÉS

par R. PIERROT Ingénieur AM et ESE
Chef de laboratoire à la

THOMSON - C.S.F
92 - MALAKOFF

SOMMAIRE

Une programmation sur ordinateur du calcul des déformations d'un faisceau d'antenne produites par un radôme de pointe avant a été mise au point.
L'utilisation des programmes de calcul de la déviation d'axe, de la perte de transmission et des remontées des lobes secondaires, permet de déterminer :

- la combinaison paramétrique optimale répondant à un problème posé (forme générale du radôme, permittivité et loi d'épaisseur de la paroi)
- les performances radioélectriques théoriquement réalisables
- l'importance des effets perturbateurs dus aux imprécisions de réalisation du radôme et aux variations de température dans les conditions de vol.

1 - INTRODUCTION

Cette communication est une synthèse des études effectuées depuis plusieurs années par la THOMSON-CSF sur les radômes aéroportés. Elle a pour but de montrer les possibilités de la méthode de calcul des performances radioélectriques mise au point au cours de ces études.
Les caractéristiques d'un radôme aéroporté constituant la pointe avant d'un avion ou d'un engin interviennent beaucoup sur les performances du système radar qu'il est destiné à protéger. L'optimisation de ces caractéristiques en fonction du problème posé est donc de première importance.
Le problème général consiste à rechercher le meilleur compromis entre les exigences radioélectriques, aérodynamiques et mécaniques. Généralement, les considérations aérodynamiques et mécaniques définissant la forme générale du radôme et l'épaisseur minimale de la paroi sont prépondérantes, et le problème se ramène à l'optimisation de la loi d'épaisseur de la paroi (en limitant l'investigation aux technologies classiques "Monolithe", bien que la méthode de calcul présentée s'applique aussi aux parois plus complexes de type multicouche), pour obtenir les déviations d'axe, pertes de transmission et remontée des lobes secondaires souhaités dans une bande passante donnée.
L'objet de cette communication est de montrer que ces caractéristiques sont accessibles par le calcul, ce qui permet non seulement de définir la loi d'épaisseur optimale, mais aussi d'effectuer des études paramétriques générales permettant de choisir en connaissance de cause le meilleur compromis entre exigences aérodynamiques et radioélectriques.

2 - PRINCIPE DES CALCULS

Le calcul exact des déformations apportées par le radôme au faisceau rayonné par l'antenne est pratiquement inextricable. Des approximations sont nécessaires et entraînent inévitablement des limitations au domaine de validité des résultats.

L'approximation principale consiste à effectuer des raisonnements d'optique géométrique. Elle est valable en champ proche de l'antenne jusqu'à une distance de l'ordre de 0,1 à 0,2 d^2/λ si d est le diamètre d'antenne et λ la longueur d'onde.

En effet, la distribution de champ dans un plan parallèle à l'ouverture rayonnante de l'antenne est pratiquement la même que sur l'antenne si cette condition de distance est respectée (1). Cette condition limite généralement l'utilisation des calculs à des antennes de diamètre supérieur à dix longueurs d'onde.

La figure 1 montre le principe du calcul. Le raisonnement est effectué en réception : un front d'onde plan équiamplitude arrive sur le radôme dans la direction d'une source émettrice située à grande distance. Deux cas de rayons sont à considérer : le premier concerne un rayon qui après la traversée du radôme arrive directement sur l'antenne, et le second un rayon qui n'atteint l'antenne qu'après une ou plusieurs réflexions. A chaque traversée ou réflexion au niveau de la paroi, les coefficients de transmission ou de réflexion sont calculés en amplitude et phase au moyen des relations classiques des lames à faces parallèles. Le signal total reçu par l'antenne est l'intégrale des signaux reçus par l'intermédiaire de chacun de ces rayons. En effectuant ce calcul à plusieurs fréquences pour chaque position d'antenne par rapport au radôme et pour chaque position du front d'onde incident par rapport à l'antenne, on obtient une vue complète des effets du radôme sur le faisceau d'antenne.

Ce type de calculs répétitifs se prête très bien à l'exploitation sur ordinateur. Cette exploitation est realisée au moyen de trois programmes principaux :

- le premier a pour but le calcul de la déviation d'axe en fonction de l'orientation de l'antenne par rapport au radôme (en calculant la direction du zéro d'écartométrie d'un système monopulse), il permet également le calcul de la perte de transmission.
- le second a pour but le calcul de la remontée des lobes secondaires de l'antenne
- le troisième a pour but l'optimisation de la loi d'épaisseur de la paroi afin de minimiser la déviation d'axe.

3 - RESULTATS OBTENUS

La figure 2 montre un diagramme de rayonnement calculé, comparé à un diagramme mesuré. Le calcul comme l'expérimentation font bien apparaître le "lobe de réflexion" bien connu.

La figure 3 montre une courbe de déviation d'axe, en fonction du dépointage de l'antenne par rapport au radôme, calculée et mesurée avant d'effectuer ces corrections d'épaisseur, et la figure 4 la même courbe après application de la correction d'épaisseur calculée au moyen du programme d'optimisation.

Ces quelques exemples montrent que la corrélation entre les résultats de calcul et les résultats expérimentaux est bonne. Elle apparaît encore meilleure sur des résultats de synthèse que sur les calculs élémentaires.

Par exemple, la courbe de la figure 5 montre la variation calculée de la pente d'aberration (maximale en fonction du dépointage) en fonction de la fréquence pour un radôme d'allongement 1,8 (l'allongement est défini comme le rapport entre la distance de la pointe du radôme à l'antenne et le diamètre du radôme au droit de l'antenne) avec une paroi de permittivité 5, comparée à une courbe déduite de résultats expérimentaux. Ce type de courbe met bien en évidence la sélectivité du radôme. Cette sélectivité n'est pas modifiée par les corrections d'épaisseurs qui provoquent seulement une translation de la courbe suivant l'axe des ordonnées. Sur la même figure, les courbes tracées pour des permittivités de 3 et 9 montrent qu'un accroissement de permittivité provoque une diminution de la bande passante (résultat normal) et une diminution des aberrations à la fréquence d'accord pour un radôme à épaisseur constante (résultat plus inattendu).

Sur la figure 6, les courbes tracées pour une permittivité de 4,5 et des allongements de 1,4 1,8 et 2,5, montrent de même que la sélectivité augmente avec l'allongement (et que les aberrations à la fréquence d'accord diminuent).

La figure 7 montre une représentation des aberrations en fonction du dépointage relatif de l'antenne par rapport au radôme en site et en gisement. Chaque vecteur représente en échelle dilatée la déviation d'axe au point de site et de gisement correspondant à son origine. Cette figure montre l'évolution des aberrations pour des balayages de direction quelconque par rapport aux axes. Ces résultats sont d'un grand intérêt pour l'étude par simulation de l'effet des perturbations sur l'ensemble du Système Radar. Sur cette figure il apparaît nettement que les composantes d'aberration pour des dépointages combinés en site et gisement n'ont pas de relation simple avec les composantes trouvées pour les mêmes dépointages dans les plans principaux.

Un autre résultat important est la mise en évidence de la proportionnalité de la déviation d'axe à la largeur du faisceau pour une géométrie de radôme donnée. Cette propriété est mise à profit pour tracer les courbes en coordonnées réduites et obtenir ainsi des résultats plus généraux. Ainsi sur les figures 3, 4, 5 et 6, l'échelle des ordonnées est graduée en aberration ou en pente d'aberration multipliée par le rapport d/λ .

Il résulte de cette propriété que la facilité de réalisation d'une déviation d'axe donnée dépend de la largeur du faisceau : une déviation d'un dixième de largeur de faisceau est très facile à obtenir, alors qu'une déviation d'un centième de largeur de faisceau est très difficile à réaliser.

4 - POSSIBILITES DES CALCULS

Quelques exemples des possibilités d'étude paramétrique et d'optimisation de la loi d'épaisseur ont été donnés précédemment. D'autres possibilités très intéressantes en résultent : l'étude des effets perturbateurs provoqués par les variations de température et les imprécisions de réalisation.

Une variation de température provoque une variation de permittivité et une variation d'épaisseur de la paroi. En simulant les conditions de température réelles du vol (variation continue de température le long et dans l'épaisseur de la paroi) par des variations quantifiées de permittivité et d'épaisseur, le calcul est possible.

Les courbes de la figure 8 montrent que l'effet de la température provoque une translation des courbes suivant l'axe des abcisses (cet effet est dû à la variation de la température moyenne de la paroi) et une translation suivant l'axe des ordonnées (effet dû au gradient de température le long de la paroi qui provoque une détérioration partielle de la loi d'épaisseur correctrice). Par ce procédé, il est possible d'une part de prévoir le radôme pour ses conditions d'utilisation exactes, et d'autre part de déduire les performances du radôme en vol des mesures effectuées au sol à température ambiante.

L'influence des imprécisions de réalisation est étudiée de manière semblable. Ces imprécisions portent sur la permittivité et l'épaisseur de la paroi. Une imprécision sur la valeur moyenne provoque une translation de la courbe suivant l'axe des abcisses (figure 8) et les variations autour de cette valeur moyenne provoquent une translation suivant l'axe des ordonnées en détruisant partiellement l'effet de la correction d'épaisseur optimale. Contrairement aux effets de la température qui sont systématiques, ces translations sont aléatoires (les lois de répartition dépendent de la technologie utilisée pour la réalisation), leur exploitation est donc faite par des procédés probabilistes. Les résultats permettent de connaître a priori les effets des dispersions de fabrication et d'en déduire le pourcentage de rebuts prévisible. Si le pourcentage trouvé est inacceptable, il faut admettre de retoucher les radômes après un premier test radioélectrique (ce cas est fréquent pour les radômes en stratifié verre résine). Le programme de calcul d'optimisation de la loi d'épaisseur permet alors de calculer les corrections à effectuer d'après les résultats de mesure.

5 - CONCLUSION

L'utilisation des programmes de calcul est utile au cours des différentes étapes d'une réalisation :

- au cours de la prétude, elle permet d'élaborer les informations nécessaires au choix du meilleur compromis entre les exigences aérodynamiques, radioélectriques et mécaniques (ce choix aboutit à la définition de la forme du radôme et de la permittivité de la paroi)
- au cours de l'étude, elle permet de connaître a priori les performances et les rebuts prévisibles en fonction des tolérances de fabrication, et de définir la loi d'épaisseur à réaliser.
- au cours de la réalisation, elle permet de définir les corrections à effectuer pour compenser les dispersions de fabrication.

L'existence de ces programmes de calcul a permis d'éclaircir considérablement le problème de la définition des radômes, problème dont la difficulté est connue depuis longtemps et qui provoque des limitations aux performances des systèmes radar aéroportés.

Réf. : - (1)HANSEN R.C. "Microwave Scanning Antennas" VOL.1 Apertures Academic Press 1964.

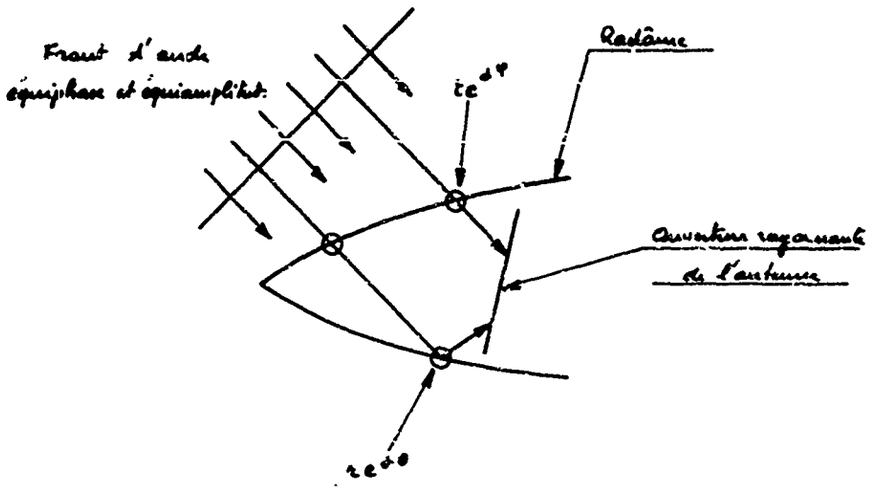


Fig.1 Principe du calcul

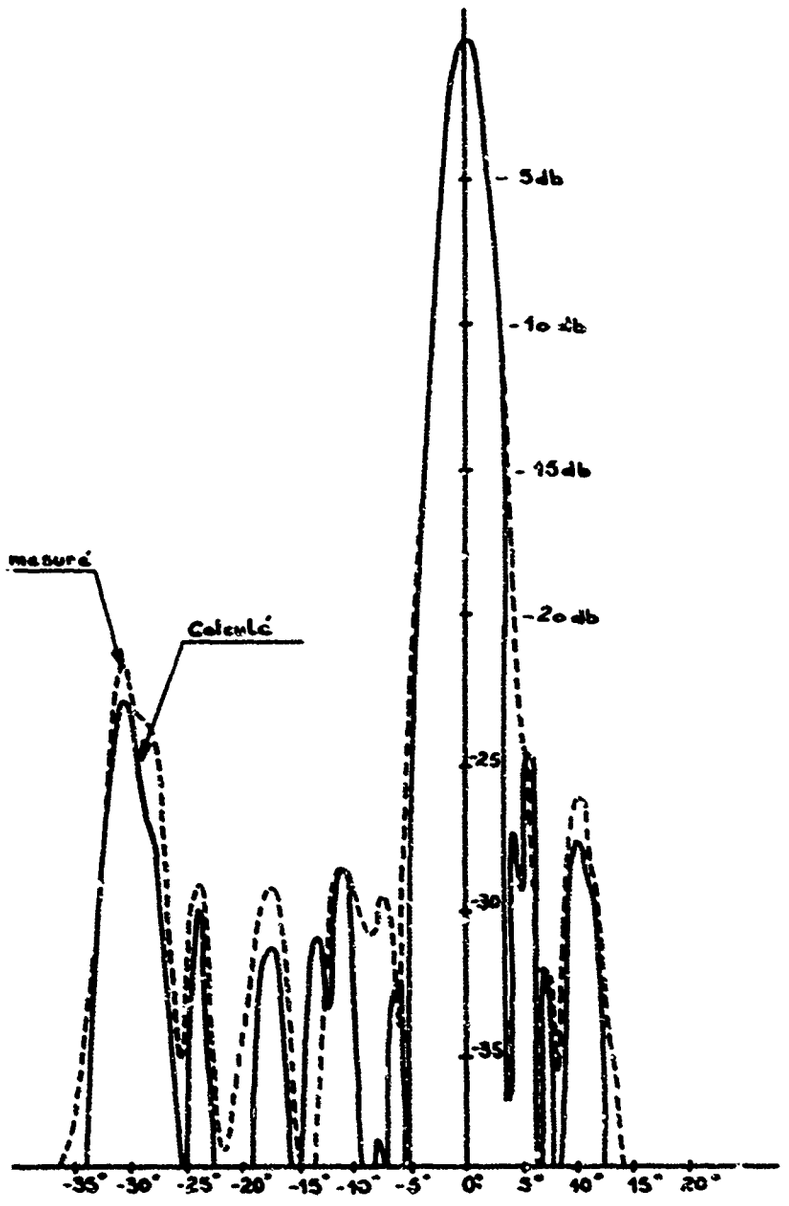


Fig.2 Diagramme de rayonnement

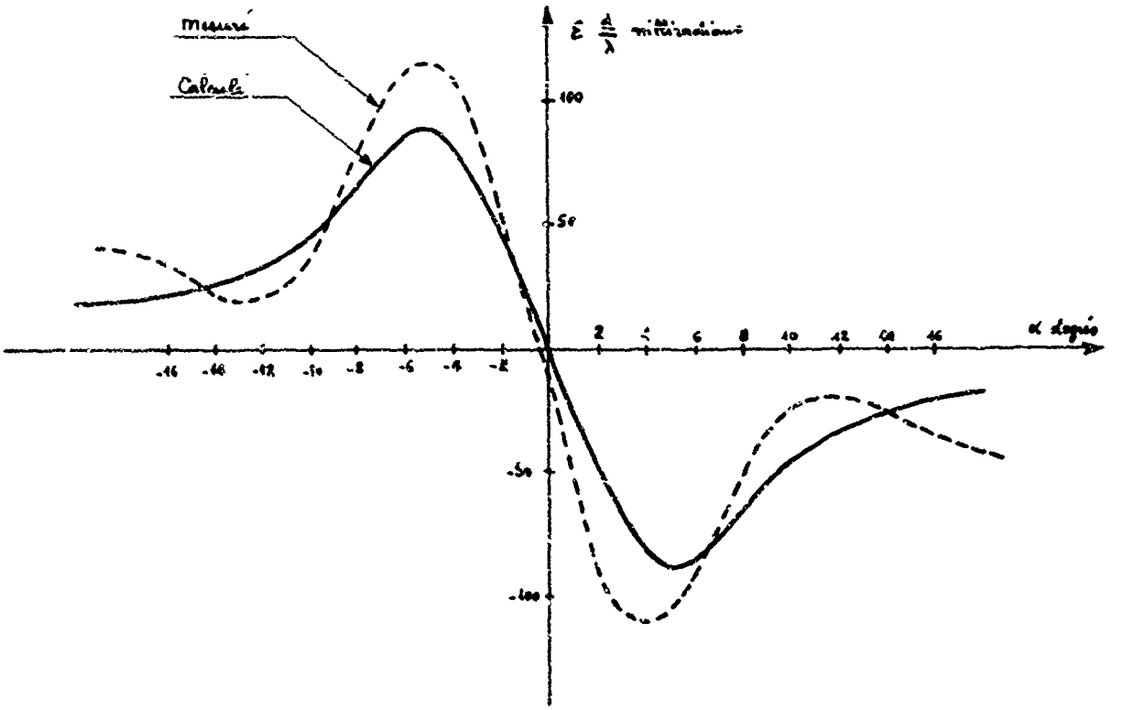


Fig.3 Comparaison calcul-experience avant correction

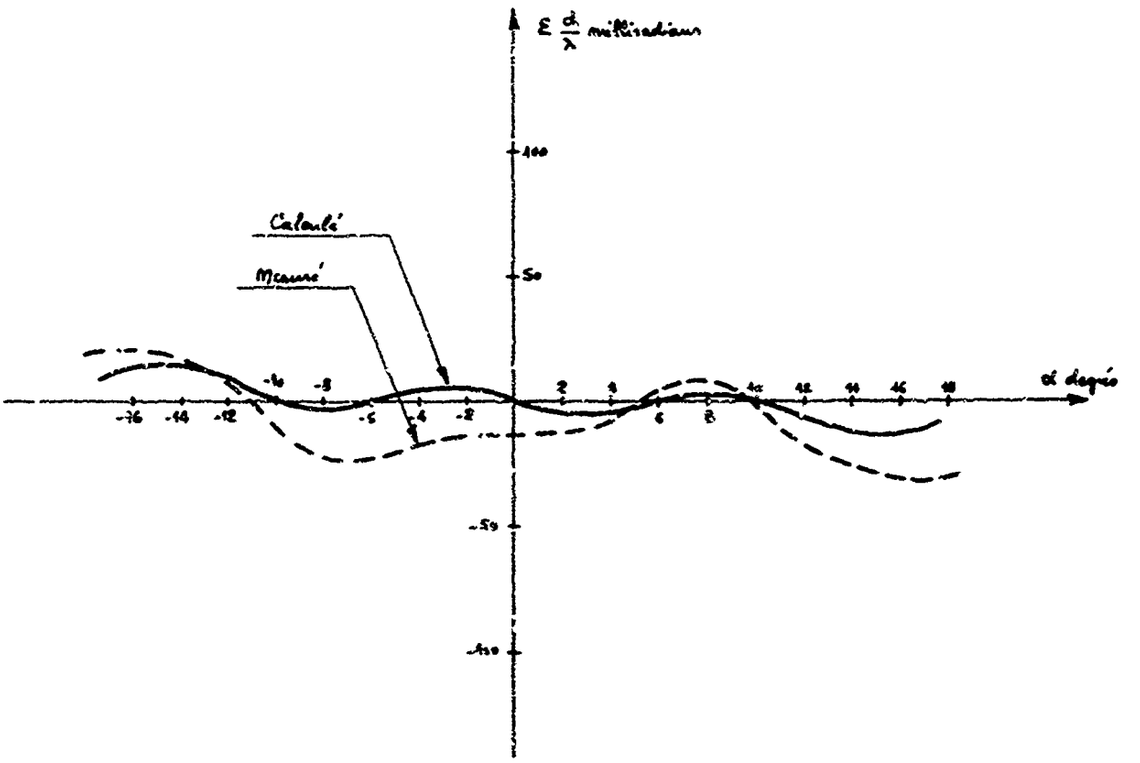


Fig.4 Comparaison calcul-experience apres correction

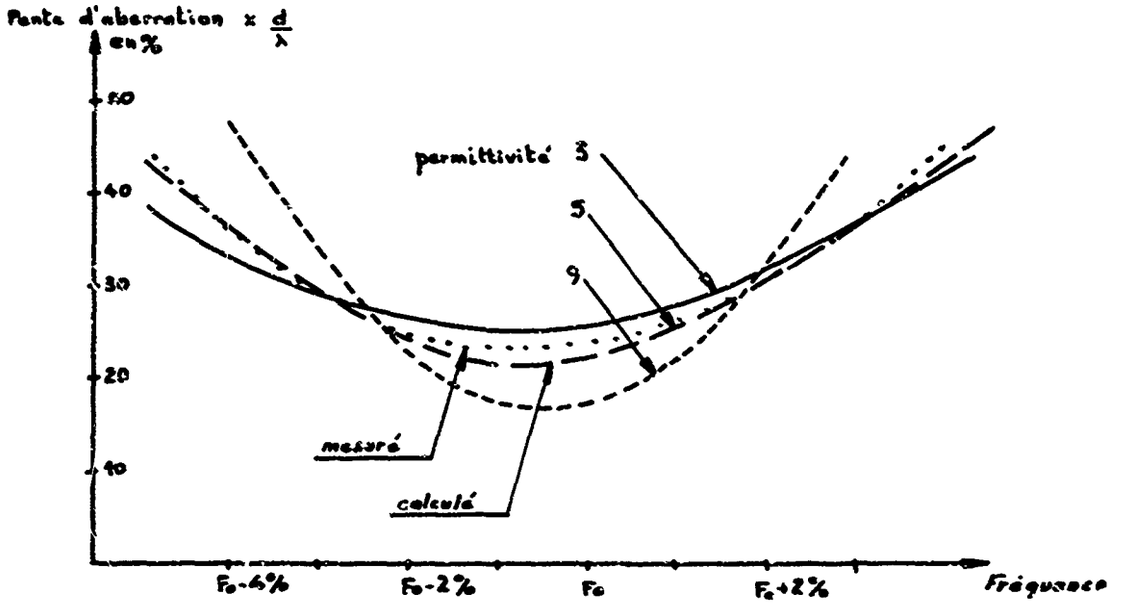


Fig.5 Pente d'aberration en fonction de la frequence (allongement = 1,8)

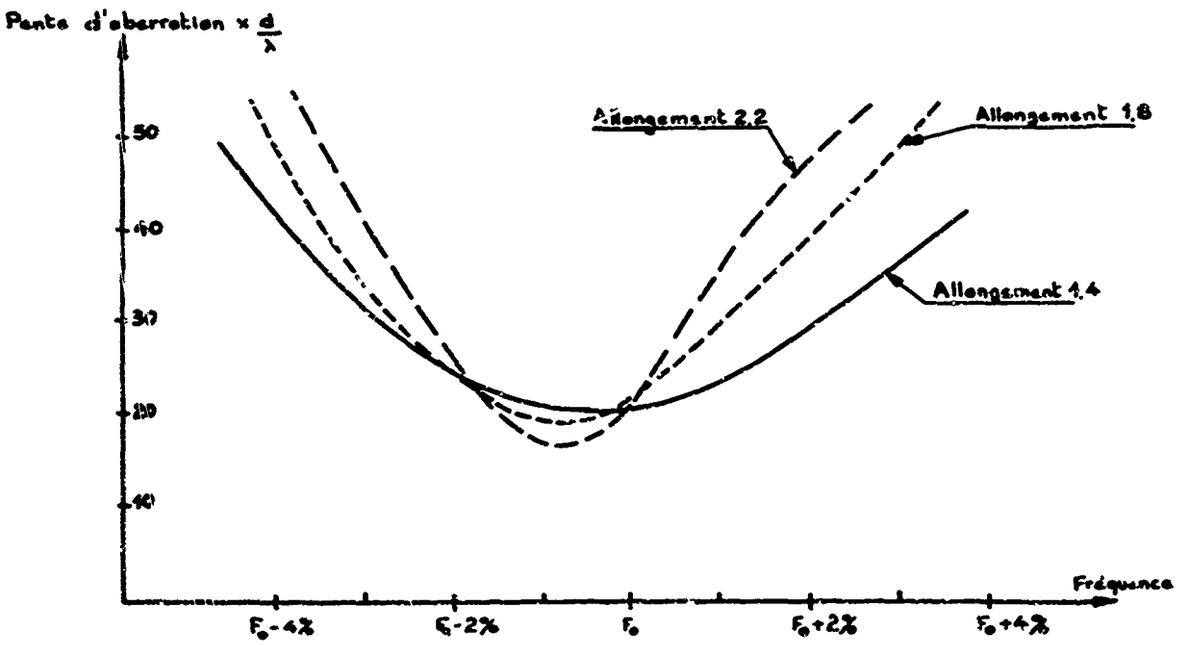


Fig.6 Pente d'aberration en fonction de la frequence (permittivité = 9)

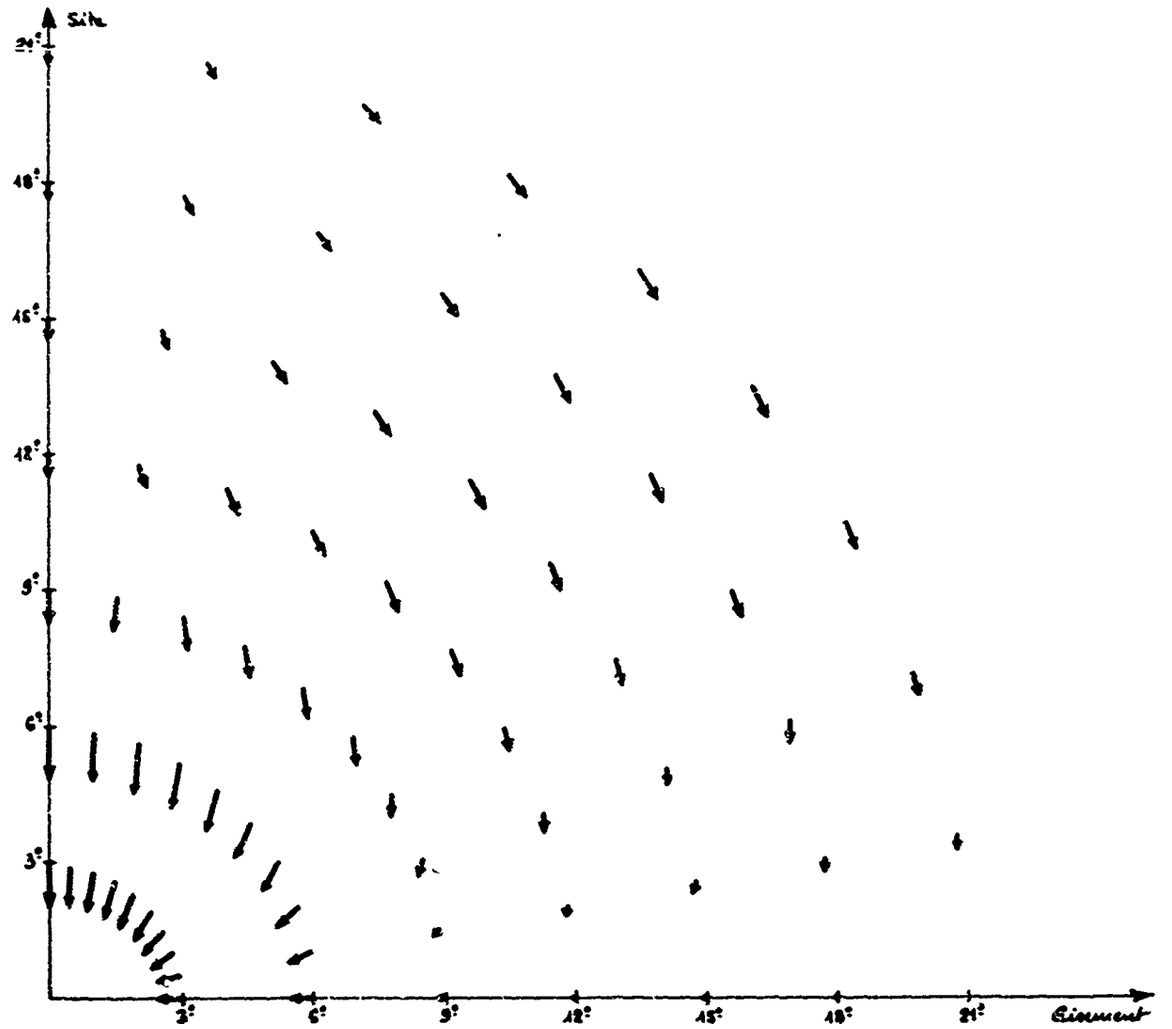


Fig.7 Carte d'aberration

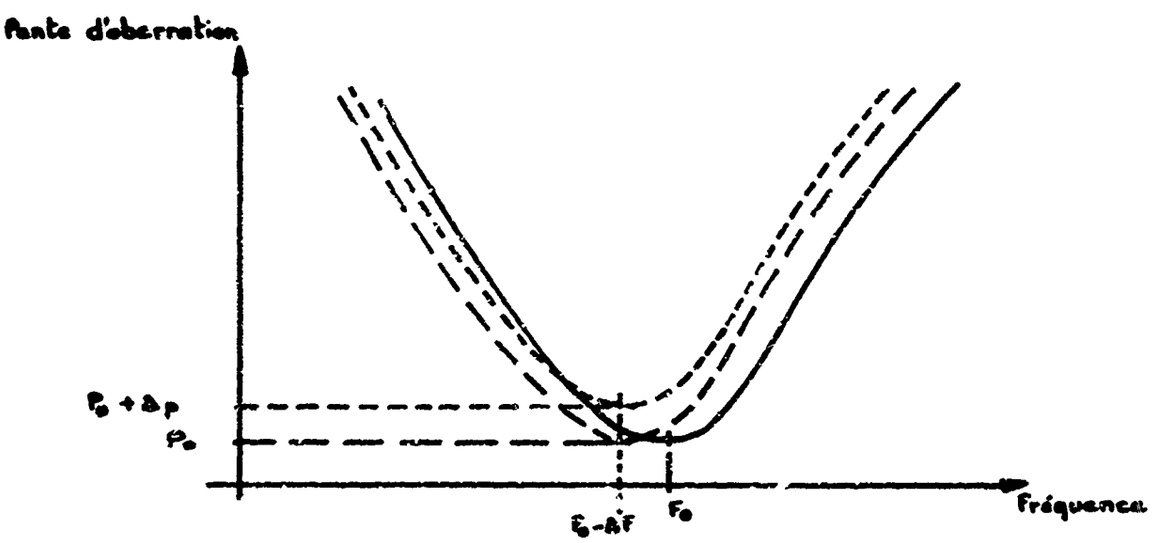


Fig.8 Influence des perturbations

PATTERN COMPRESSION BY SPACE-TIME BINARY CODING OF AN ARRAY ANTENNA

by

S. Drabowitch and C. Aubry

Thomson - CSF
92, Bagneux, France

PATTERN COMPRESSION BY SPACE-TIME BINARY CODING OF AN ARRAY ANTENNA

S. Prabowitch and C. Aubry
Thomson - CSF
92, Bagneux. France

SUMMARY

The purpose of this paper is to describe a process of space scanning by means of an array antenna, the radiation pattern of which is time-modulated. This modulation is obtained by associating a 1 bit phase-shifter ($0 - \pi$) to each radiating element of the array. The received signals are processed by means of adaptive filters or of correlators corresponding to various direction samples. Three coding types are considered : reception coding, transmission-reception coding and periodic coding. The main advantages of the space-time binary coding results from the simplicity and economy of the phase-shifters and their control elements. These advantages are obtained at the expense of an energy loss of at least 3 dB. This disadvantage is not necessarily redhibitory and the process constitutes undoubtedly an additional solution available to the engineer.

1. INTRODUCTION

This communication is made within the frame of R. and S. studies on the signal processing antennas. Its purpose is to describe a process of space scanning by means of an array antenna, the radiation pattern of which is time modulated. As compared with conventional arrays, the proposed system offers a certain number of advantages :

- Simplicity and economy of the phase-shifters and their control elements ; reduction of the insertion losses (advantages over electronically scanning antennas)
- Use of only one receiver (an advantage over multi-beams antennas)

These advantages are obtained at the expense of a certain loss in the signal-to-noise ratio, and they require a certain number of correlators or adaptive filters.

In section 2, we present the general principle of the space-time coding of an array antenna. The following sections deal exclusively with phase coding, and specially the binary phase coding allowing to obtain the above mentioned advantages. We shall examine the cases in which the coding is made at reception (section 3), then at transmission and reception (section 4). Further we shall consider the highly interesting case of the periodic coding (section 5). For each type of coding we shall try to evaluate the integrated energy at reception during the measurement time T. For a target located in the direction referenced by , this energy is proportional to

$$E = \int_T G_E(\theta,t)\; G_R(\theta,t)\, dt$$

$G_E(\theta,t)$ and $G_R(\theta,t)$ are the gains at transmission and reception in direction θ and at the moment t. The proportionality factor depends on a certain number of parameters which are : the transmitter power, the target distance, its cross-section, the various losses the integration gain etc ... it will be assumed that it does not depend on the antenna type used. The results obtained are discussed in the last part of this paper (section 6).

2. GENERAL CONSIDERATIONS

2.1 Angle time analogy

There is a great analogy between a time signal and its frequency spectrum on the one hand, and the radiation pattern of an array antenna and the illumination law of this array on the other hand (1). The definitions being those of fig. 1 one will write :

$$\begin{array}{ccc} t & f & S(t) = \int F(f)\; e^{j2\pi ft}\, df \\ \updownarrow & \updownarrow & \updownarrow \\ \theta & \nu & D(\theta) = \int \Phi(\nu)\; e^{j2\pi\nu\theta}\, d\nu \end{array}$$

This analogy strongly suggests to transpose into the angular field, the radar signal processing techniques used in range and velocity measurements. Thus, as the range resolution only depends on the autocorrelation function of the transmitted signal (or on its spectrum intensity), in the same way, the angular resolution only depends on the autocorrelation function of its radiation pattern (or on the intensity of the illumination law) (1), (2), (3). An array illuminated with a "messy" phase law, has a "messy" pattern, as well in phase as in amplitude, but the autocorrelation function of this pattern - which characterises the antenna possibilities as regards the angular resolution - may present a very important peak value. Here we recognize, in angular form, the principle of the pulse compression used in range measurements. This pattern compression possibility, together with the possibilities provided by the use of quantized electronically phase-shifters, is the basis of the process we intend to describe.

2.2 Principle of the space-time coding

Fig. 2 shows the principle of the space-time coding of the array. This array consists of N radiating elements, each one of which associated to a modulator. The SPACE CODING is obtained by choosing, in the time interval T_1, the complex-valued amplitude a_{in} of each array element ($n = 1, 2 \dots N$). The time succession of a number M of illumination laws, with respective durations T_i ($i = 1, 2 \dots M$), defines the TIME CODING. The coding is then obtained by a matrix with complex valued factors, including N lines and M columns. To each space direction a time signal $F(\theta,t)$ is thus associated, received by the array, characteristic for this direction. In the case where the array operates at reception, $F(\theta,t)$ is proportional to the modulated pattern $D(\theta,t)$ (fig. 3). In the case where it operates at transmission-reception, $F(\theta,t)$ is proportional to $D^2(\theta,t)$. $\{\theta_1, \theta_2, \dots \theta_k\}$ representing a set of direction samples in the space under scan, the processing of the received signals will be made either by means of filters adapted to the signals $F(\theta_k,t)$, or by means of correlators.

2.3 Phase coding. $0 - \pi$ coding

We only consider the case of the phase coding, and specially the binary phase coding. The modulators represented on fig. 2 are simple $0 - \pi$ phase-shifters with all the advantages implied. It shall be noted that the electronically scanning process, by means of a beam of conventional shape, corresponds in fact to a particulary simple space time coding, derived from mechanical scanning antennas :

- The space coding consists in setting a linear or quasi-linear phase law along the array
- The time coding is defined by the different phase laws successively set, i.e. by the successive positions of the antenna beam (fig. 5)

2.4 Conservation of the phase information

Since the illumination law of the array consists of a succession of terms with phase 0 and π, according to a conventional property of the Fourier Transform of a real function, the resulting pattern has an even amplitude law and an odd phase law. It can then be realized that, to make the distinction between the left and right parts of the sector under scan, it is necessary to keep the phase information contained in the pattern, i.e. in fact, in the received modulated signal.

3. RECEPTION CODING

3.1 Expression of the correlators output signals

We assume that the transmission gain is uniform in the sector to be watched ; the signal reflected by a target located in the θ_o direction is modulated by the antenna pattern in this direction : $D(\theta_o,t)$ (fig. 3). Simultaneously, a code generator elaborates the IF reference signals $D(\theta_k,t)$ concerning the various direction samples. After a frequency change, the received signal is compared to the signals $D(\theta_k,t)$ in correlators producing the signals

$$\Psi(\theta_o,\theta_k) = \int_T D(\theta_o,t)\, D^*(\theta_k,t)\, dt$$

In the time interval T_i, the illumination law of the array is expressed by

$$\sum_{n=1}^{N} a_{in}\ \delta(\vartheta - n\Delta\vartheta) \qquad \text{(N isotropic sources)}$$

Where Δ represents the interval, expressed in space frequencies, between two successive array elements. In order to simplify the calculations, it will be assumed that $a_{in} = \pm 1$. The corresponding pattern is expressed by :

$$D_i(\theta) = \sum_n a_{in}\ \exp[j2\pi n\,\Delta\vartheta\,\theta]$$

Then one can write :

$$\Psi(\theta_o,\theta_k) = \sum_{n=1}^{N}\sum_{m=1}^{N} \exp[j2\pi(n\theta_k - m\theta_o)\Delta\vartheta] \sum_{i=1}^{M} a_{in}\, a_{im}$$

Let us assume that the coding matrix is such that

$$\sum_{i=1}^{M} a_{in}\, a_{im} = M\,\delta_{mn}$$

Then $\Psi(\theta_o,\theta_k)$ is reduced to

$$\Psi(\theta_o,\theta_k) = M \sum_{n=1}^{N} \exp\left[j 2\pi n (\theta_n - \theta_o) \Delta\psi \right]$$

$\Psi(\theta_o,\theta_k)$ appears as the autocorrelation function of the pattern, i.e. $C(\theta_o-\theta_k)$. It is the Fourier transform of the illumination intensity. With a tolerance of one factor, one has

$$\Psi(\theta_o,\theta_k) = C(\theta_o-\theta_k) = \frac{\sin\left[N\pi\Delta\psi(\theta_o-\theta_k)\right]}{\sin\left[\pi\Delta\psi(\theta_o-\theta_k)\right]}$$

REMARKS

(a) In order to correctly perform the correlation, it is necessary to start the reference signal generator at the exact time when the pulse reflected by the target is received by the array. It is obvious that this instant is not known, not even whether the reflected pulse exists. To eliminate this difficulty, the coding is carried out in a cyclical manner, with a period equal to the length of the transmitted pulse. The reference signal corresponding to a given direction sample is then identical, as regards its envelope, to the modulated signal eventually received from a target located in this direction (fig. 6).

(b) The Doppler effect due to the target velocity has been fully neglected. This effect introduces spurious terms in the expression of $\Psi(\theta_o,\theta_k)$. In particular, $\sum a_{in}\, a_{im}$ becomes $\sum a_{in}\, a_{im} \exp[j2\pi f_d i\frac{T}{M}]$ where f_d is the Doppler frequency. In fact if we consider the following orders of magnitude : $\lambda = 10\,cm$, $T = 4\,\mu s$, $V_R = 1800\ Km/h$ (target radial velocity), it appears that $2\pi f_d T < 0.25\ rad.$. Under these conditions the influence of the exponential term can be neglected, and it can easily be seen that this also applies for the other spurious terms.

(c) Comparing the received signal with the reference signal, one obtains in fact the real part X of the autocorrelation function. To obtain the imaginary part Y of $C(\theta_o-\theta_k)$, it is sufficient to shift the reference signal phase by $\frac{\pi}{2}$. Then one can calculate

$$|C(\theta_o-\theta_k)|^2 = X^2 + Y^2$$

(d) We could simplify the expression of the correlator output signals, assuming that the following relations were verified

$$\sum_{i=1}^{M} a_{in}\, a_{im} = M\,\delta_{mn}$$

If A is le coding matrix (M lines, N columns), these relations are summarized by the matrix relation $A^t A = I$ (I is the N x N unity matrix). In the case where A is a square matrix, i.e. when the number of patterns used is equal to the number of radiating elements, A is orthogonal. Since the Fourier transform keeps the scalar product invariant, it results that the patterns are orthogonal to one another.

3.2 Energy considerations

The notations will be the following

- Θ angular amplitude of the observed sector
- L number of direction samples
- T measurement time (not to be confused with the pulse width)
- G maximum gain of the array (determined by the array dimensions)

Let us first examine the case of conventional arrays operating at transmission-reception. In case of SCANNING ANTENNAS (electronical or mechanical) the antenna pattern can be regarded as a sectoral beam with width $\frac{\Theta}{L}$ and with gain G. If T corresponds to the scanning time of the observed sector, the target is illuminated during the time $\frac{T}{L}$. The value of the integrated energy is :

$$E = \frac{T G^2}{L} = E_o$$

In case of MULTI-BEAMS ANTENNAS, the receiving pattern includes L beams with a width $\frac{\Theta}{L}$ and a gain G, the transmission pattern is quasi uniform, with a gain $\frac{G}{L}$ in sector Θ . The transmission and reception is continous during the time T and the value of the integrated energy is

$$E = G \cdot \frac{G}{L} \cdot T = E_o$$

From the energy point of view, these two cases are theorically equivalent.

In case of $0 - \pi$ CODING ANTENNAS, the accurate calculation of E depends on the choice of the codes used. However it is possible to evaluate this energy by processing the pattern as a complex-valued random function of the angular variable θ. In the hypothesis where the number of patterns is great, the mathematical expectations can be identified as time averages. As an example,

$$\langle D(\theta) \rangle = \frac{1}{M} \sum_{i=1}^{M} D_i(\theta) = \frac{1}{T} \int_T D(\theta, t)\, dt$$

Further we make the following hypothesis :

- The process is stationary. For instance $\langle D(\theta) \rangle$ is independant of θ
- The process is ergodic : the average values with respect to θ are equal to the statistical averages. For instance

$$\langle D(\theta) \rangle = \frac{1}{\Theta} \int_\Theta D_i(\theta)\, d\theta$$

Finally if we assume that the array includes a great number of sources, the Central Limit theorem can be applied. The real part $X(\theta)$ and imaginary part $Y(\theta)$ have independant, stationnary, zero-mean normal distributions, having the same variance $\langle X^2 \rangle = \langle Y^2 \rangle = \sigma^2$. $|D(\theta)|$ has then a Rayleigh distribution

$$p(|D|) = \frac{|D|}{\sigma^2} \exp\left[|D|^2 / 2\sigma^2 \right]$$

The ergodicity hypothesis allows to evaluate σ^2 :

$$\langle |D|^2 \rangle = \langle X^2 + Y^2 \rangle = 2\sigma^2 = \frac{1}{\Theta} \int_\Theta |D(\theta, t)|^2\, dt$$

If the patterns are suitably normed, so that $|D(\theta,t)|^2 = G(\theta,t)$ (gain function), the integral represents the average power radiated in the sector of scan, i.e. $G \frac{\Theta}{L}$. It results that $\sigma^2 = \frac{G}{2L}$

In case of RECEPTION CODING, considered in this section, we have $G_E(\theta,t) = \frac{G}{L}$ so

$E = \frac{G}{L} \int_T |D(\theta,t)|^2\, dt$ but $\frac{1}{T} \int_T |D(\theta,t)|^2\, dt$ represents the second order moment of the Rayleigh distribution, i.e. $\langle |D(\theta)|^2 \rangle = 2\sigma^2 = \frac{G}{L}$. Then one finds $E = \frac{E_0}{L}$. The loss is measured by the number of angular samples.

4. TRANSMISSION-RECEPTION CODING

4.1 Description of the process

If the coding at transmission-reception is to be done, the pulse coding described above has to be given up. In fact it can be realized that it becomes impossible to associate a unique reference signal to a given direction (the reference signal would depend on the target range). Then one is led to consider a pulse-to-pulse coding defined as follows (fig. 7).

The transmission is made according to a given pattern $D_i(\theta)$ corresponding to a given space coding of the array, and the reception is made according to the same pattern during a time τ corresponding to the radar range. Then one transmits according to $D_{i+1}(\theta)$ and one receives with this pattern during τ and so on. The characteristic signal of a given direction, i.e. the signal received when a target is in this direction, consists of a pulse train with width d, with repetition period τ, and with complex amplitudes $[D_i(\theta)]^2$ $(i = 1, 2 \ldots M)$. This signal is sent through a set of filters, each adapted to a direction sample. The filter adapted to the Θ_k direction has an impulse response which is a succession of M pulses of width d, shifted by τ, with amplitudes $[D_i^*(\theta_k)]^2$ $(i = M, M-1 \ldots 2, 1)$. The so called Turin process (4) allows to obtain such an impulse response Assume that a target is in the θ direction. The Θ_k filter output signal consists of $2M - 1$ pulses with amplitudes

$$b_k(\theta, h) = \sum_{i=1}^{M} \left[D_i(\theta) \right]^2 \left[D^*_{M-h+i}(\theta_k) \right]^2 \qquad \text{for } h = 1, 2 \ldots M$$

$$b_k(\theta, h) = \sum_{p=1}^{2M-h} \left[D_{M-p+1}(\theta) \right]^2 \left[D^*_{2M-h-p+1}(\theta_k) \right]^2 \qquad \text{for } h = M+1 \ldots 2M-1$$

After a quadratic detection one obtains $2M - 1$ pulses with amplitudes

$$|b_k(\theta, h)|^2 \qquad (h = 1, 2 \ldots 2M-1)$$

This function of two variables θ and t defines the angle-range ambiguity corresponding to the θ-filter. The set of ambiguity functions corresponding to the different direction samples characterises the system ability as regards its range and angular resolution. Fig. 9 represents such a surface corresponding to the following data :

- The array includes 8 isotropic sources
- The number of patterns used is 30
- The considered direction sample is defined by $\theta_K = \frac{1}{4\Delta\vartheta}$

For the sake of simplicity, we have only represented two sections of the ambiguity surface on fig. 9

1) - The $\theta = \theta_K$ section, which shows the range ambiguity for a target in the θ_K direction

2) - The $t = 0$ section, which shows the angular ambiguity for a target at "distance" $t = 0$.

4.2 Choice of the codes

An array having N radiating elements associated to $0 - \pi$ phase shifters allows to radiate 2^N patterns. The problem is to choose a certain number of patterns in this set. This choice has to be made according to criteria concerning the ambiguity function corresponding to the different direction samples. It is sure that the obtained results depend on the number N of these codes, on the choice of the N codes, as well as on the order in which they run. We have made some calculations of the ambiguity surface. The obtained results do not offer a sufficient general basis to make the best choice between all possible codes. However, it seems the choice of patterns is by itself far more important than the number of these patterns. As a conclusion, when a precise operational problem will be raised, one will be led to perform a certain number of trials on sets of N more or less randomly choosen patterns and to select the combination leading to the best results.

4.3 Energy considerations

In case of transmission-reception coding, one has

$$E = \int_T |D(\theta,t)|^4 \, dt = T \langle |D(\theta)|^4 \rangle$$

The 4 th order moment of the Rayleigh distribution being $8\sigma^4$, one finds $E = \frac{2}{L} E_0$. Thus one has gained 3 dB in relation to the preceeding case. It can be thought to improve E by a particular feature of the coding, e.g. by selecting codes being always even or odd. The patterns $D_i(\theta)$ then have a constant phase, and their form is $D_i(\theta) = d_i(\theta) \exp[j\varphi_i]$. According to the Central Limit theorem, $d(\theta,t)$ is a stationary, zero mean, gaussian function, with variance $2\sigma^2$ (energy conservation). Then one obtains

$$E = \int_4 [d(\theta,t)]^4 \, dt = T \langle [d(\theta)]^4 \rangle$$

but $\langle d^4(\theta) \rangle = 3 \left[\langle d^2(\theta) \rangle \right]^2$ (property of the gaussian variables) so $E = 3 \frac{E_0}{L}$, i.e. an improvement by approximately 2 dB with respect to the preceeding case.

It shall be noted that, with this type of coding, there is a total ambiguity between the left and right parts of the area. In fact, considering the "round-trip", the signals received by two targets being symmetrical versus the antenna axis, will be equal, both in phase and amplitude (the Fourier transform of a real function, even or odd, is itself even or odd).

5. TRANSMISSION-RECEPTION CODING BY MEANS OF EVEN OR ODD PERIODIC CODES

5.1 Principle of the periodic coded scanning

Consider an array of length ν_0, the illumination law of which consists of periodic square waves with a period $\Delta\vartheta_K$ (such a law can be obtained by combining in phase or anti-phase a certain number of E - plane horns). According to the property of the Fourier transform of the product of two functions, the corresponding pattern is obtained through the convolution of the two following patterns :

- The pattern relative to the uniformly illuminated antenna $\frac{\sin \pi \nu_0 \theta}{\pi \nu_0 \theta}$
- The pattern relative to the considered periodic law (antenna with infinite length). This pattern consists of rays with angular positions $\frac{2n-1}{\Delta\vartheta_K}$, with amplitudes $\frac{\pi}{2n-1}$ (n takes all integer values from $-\infty$ to $+\infty$)

The resulting pattern exhibits two main lobes, with a $\frac{1}{\nu_0}$ 3 dB width, arranged on both sides of the antenna, in the $\pm\theta_K = \pm\frac{1}{\Delta\vartheta_K}$ directions (fig. 10). These two lobes are in phase if the code is even ; they are in anti-phase if the code is odd. It is clear that, by varying the illumination period, one obtains an actual scanning of the space to be watched, each value of $\Delta\vartheta_K$ being relative to a direction θ_K as well as to the symmetrical $-\theta_K$.

5.2 Left-right separation

The periodic code considered being even or odd, the signals reflected by two identical targets located in the θ and $-\theta$ directions are equal. We shall now describe a process allowing to cancel this "left-right" ambiguity. To each even pattern relative to the $\pm \theta_k$ directions, is associated the odd pattern relative to the same directions. This pattern is obtained by shifting the even illumination law by a half period (fig. 10). These two patterns are processed as follows :

- The pulse transmitted in the θ_K direction (and in the symmetrical direction) is phase modulated by means of the successive illuminations EVEN, ODD, -EVEN, -ODD, EVEN ... The reception code is even. Thus the field reflected by a right located target varies as 1, j, -1, -j, 1 ... (the even and odd pattern are in quadrature), whereas if the target is left located, one receives 1, -j, 1, j, 1 ... (fig. 11). Therefore one has pratically made a frequency shift the sign of which (increase or decrease of the frequency) determines a portion (right or left) of the plane. Then it is clear that if the spectra of the pulses relative to two symmetrical targets are sufficiently far away on both sides of the transmitted spectrum, it will be possible, by use of two suitable filters, to cancel the left-right ambiguity. If the pulse, with length T, is made up of p sequences {even, odd, -even, -odd} the significant parts of these two spectra are separed by $\frac{2p}{T}$. Their 3 dB width is $\frac{4}{T}$; this shows that, despite the phase coding, the transmitted pulse is not compressed.

5.3 Energy considerations

The radiation pattern exhibits two main lobes symmetrical with respect to the antenna axis, each one having approximately a gain of $\frac{G}{2}$. 3 dB are lost each way ("Round trip") at each passage on the target, but there are twice as many passages during the same time T, as there would be with a single lobe. It results :

$$E = \frac{G^2}{4} \frac{2T}{L} = \frac{E_0}{2}$$

The loss, as compared to conventional arrays is about 3 dB ; it is independant from the number of direction samples.

6. CONCLUSION

In this paper, we have shown that it is possible to perform angular measurements by means of the space-time coding of an array antenna including 1-bit phase-shifters. We have successively examined :

- The case of reception coding (coding in the pulse by means of orthogonal codes, use of correlators)
- The case of transmission-reception coding (pulse by pulse coding, use of adaptive filters)
- The case of periodic coding (simultaneous scanning of the two portions of the plane by means of two symmetrical lobes)

From the energy point of view, periodic coding provides the best results. However, even in this favourable case, as compared to conventional electronically scanning or multi-beams antennas, the loss is about 3 dB. The interest of the space-time $0 - \pi$ coding appears then to be rather limited. In fact, the cost of phase-shifters and their control systems seems to be a key factor in the total cost of an electronically scanning antenna ; so it is quite certain that an array including 3N ou 4N 1 - bit phase-shifters is more economical than a classical array using N 3 or 4 bits phase-shifters. Furthermore, as compared to the multi-beams process, periodic coding offers the same advantages as electronics scanning, as regards the sharpness of the main lobe and the side lobe level (one takes advantage of the "round-trip"). It can be therefore expected that, in some operational cases, the type of coding we have described could be profitably used to replace the existing processes.

ACKNOWLEDGMENT

The authors wish to thank La Direction des Recherches et des Moyens d'Essais, which contributed to launch this study, and which has authorized the publication of this paper.

REFERENCES

(1) Drabowitch, S. "Application de la Théorie du Signal aux Antennes". L'Onde Electrique, mai 1965

(2) Carpentier, M. "Radars- Concepts nouveaux". Dunod, Paris, 1966

(3) Urkowitz, Hauer, Koval "Generalized Resolution in Radar System". PIRE, Oct 1962

(4) Turin, G.L. "An Introduction to Matched Filters". IRE Transactions on Information theory, vol IT.6, Jan 1960

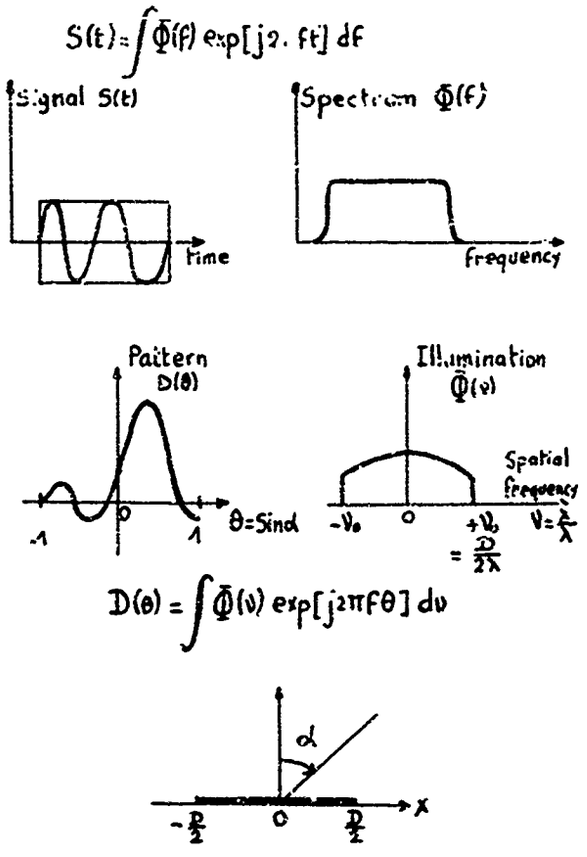


Fig. 1 Angle time analogy

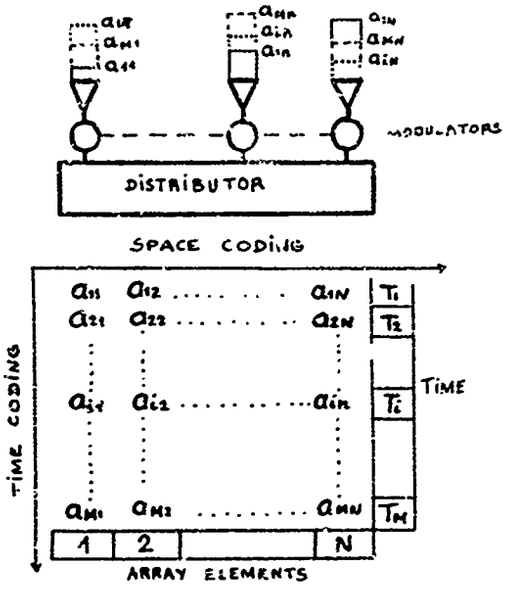


Fig. 2 Space time coding of an array antenna

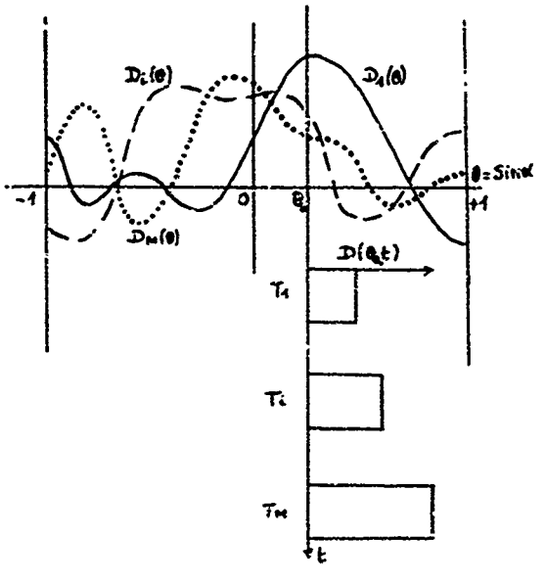


Fig. 3 Radiation pattern of a space time coding antenna

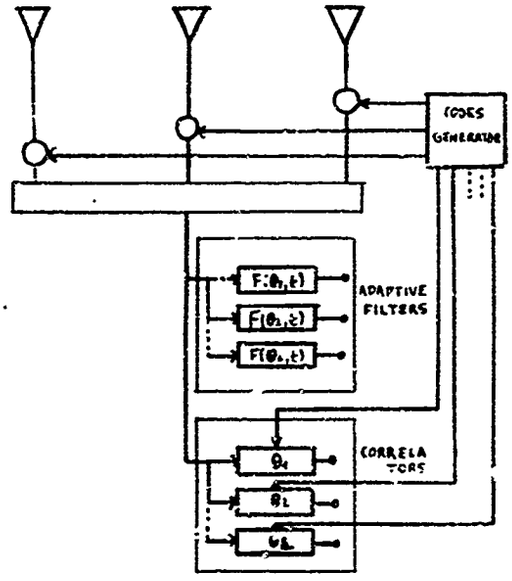


Fig. 4 Signal processing

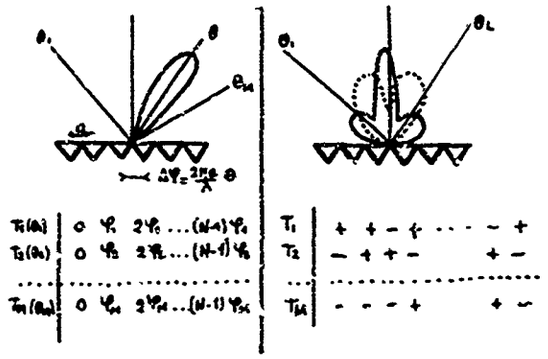


Fig.5 Time space phase coding

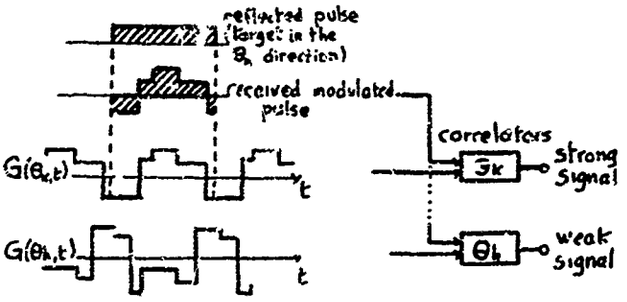


Fig.6 Reception coding

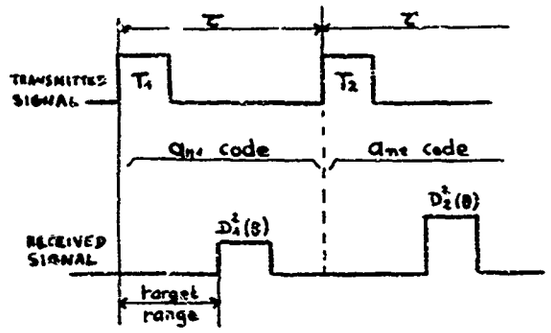


Fig.7 Transmission reception coding

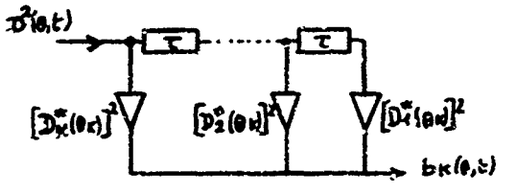


Fig.8 Turin filter adapted to the θ_k direction

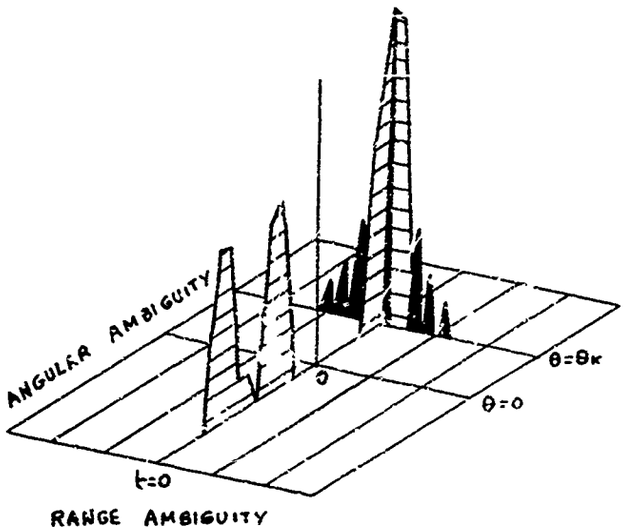


Fig. 9 Range angular ambiguity

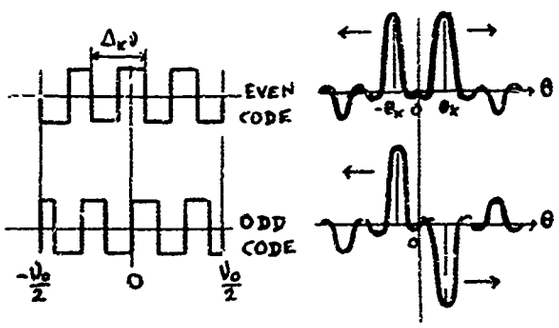


Fig. 10 Periodic coding

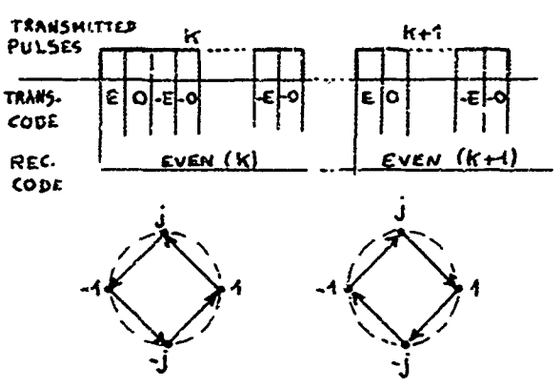


Fig. 11 Left-right separation

FLOOD-LIGHTING WITH NYQUIST RATE SCANNING

by

A.K.Edgar and I.L.Jones

Ministry of Technology
United Kingdom

FLOOD-LIGHTING WITH NYQUIST RATE SCANNING

A.K. Edgar and Dr. I.L. Jones
Ministry of Technology
United Kingdom

SUMMARY

The principles of flood-lighting with Nyquist rate scanning as applied to radar systems are given, followed by a description of a practical embodiment covering an experimental 10 cm. radar.

The experimental radar uses a 63 element receiver array, scanning at the Nyquist rate (100 kHz) and a klystron transmitter of 200 watts mean power. The system is designed to scan over 90° in azimuth.

An outline is given of the main features of the equipment, together with a discussion of experimental results. It is shown that theoretical predictions on efficiency are borne out, and that in particular, very good M.T.I. performance is possible.

1. INTRODUCTION

With surveillance radars, various subterfuges are adopted to overcome the constraints imposed by the velocity of propagation. It can be argued, for example, that only the perimeter need be searched and that targets within the perimeter can be dealt with by looking relatively infrequently at the predicted target positions. The main problem is one of time. It can be shown that a typical S-band 3-D radar which looks in turn at each angular resolution cell cannot have a period between looks at any one target much less than a minute. Furthermore, any attempt to improve on this leads to a lowering of performance in some direction; for example, in angular discrimination or in anti-clutter properties.

Ideally, we wish to see all targets as often as is necessary to give an accurate track and to detect new targets as early as possible. In many circumstances, a typical case being that of air traffic control at a busy terminal, data rate constraints may be unacceptable.

Within-pulse-scanning reduces the restrictions imposed by the velocity of propagation. We may examine all angular resolution cells in parallel, and in principle the data interval minimum is the time required for the energy to return from maximum range. In practice, it is usually necessary to integrate several returns from maximum range targets in order to keep the transmitter mean power at an economic level.

The paper describes a variant of a within-pulse-scanning system proposed by G.F. Clarke of the Ministry of Technology, in connection with a defence project in 1957. Security restrictions prevented earlier publication. We include a short theoretical description of the system, together with a description of some of the practical aspects. Experimental results are given of the range and M.T.I. performance obtained with the experimental equipment.

2. TECHNICAL DETAILS

It is not always recognised that transmission and reception have different aims. The former calls for maximum power in a given direction, while the latter is required to receive information with maximum efficiency. We know from data processing theory that samples of information taken at the Nyquist rate, that is, at periods $T = \frac{1}{2B}$, where B is the communication bandwidth, will provide all the information present in the channel. There is no lower limit to the duration of the sample, but the sampling process may cause contamination of the information unless suitable precautions are taken. The effect, often described as "aliasing", results from the introduction of additional unwanted information if the bandwidth being sampled is greater than that required for signal communication.

Within-pulse-scanning is essentially a sampling technique, although the circuit configuration does not make it obvious.

In our system, an S-band linear array of 66 primary elements, electronic processing is effected by the well-known method of modulation scanning - Fig. 1.

Consider a signal V sin ωt arriving at an angle θ to the normal to the array. It is clear from the geometry that the signal phase at the r^{th} element is

$$\phi_r = \frac{2\pi\, rd}{\lambda} \sin\theta = r\phi \qquad r = \text{integer} \qquad \ldots\ldots\ldots\ldots (1)$$

and that the phase is progressive across the array. For the scanning beam to point in the direction θ all elements must be equalised in phase and therefore a phase correlation $-\phi_r$ must be applied. Let the correction be $r\omega_m t$ that is a phase linear with time, which is equivalent to a frequency shift $r\omega_m$ in the appropriate feeder. The frequency shifts across the array are harmonically related and must be coherent in the sense that all phases must pass through zero at the same time. Thus

$$r\omega_m t = -\phi_r = -r\phi = \frac{-2\pi\, rd}{\lambda} \sin\theta \qquad \ldots\ldots\ldots\ldots (2)$$

and

$$\theta = \sin^{-1}\frac{\lambda}{d}\frac{t}{Tm} \qquad Tm = \frac{2\pi}{\omega_m} \qquad \text{.....(3)}$$

When all element channels are added the array output is, for a signal arriving from a direction normal to the array

$$V_T = \sum_{r=-M}^{+M} Vr \sin\left[(\omega + r\omega_m)t + \phi r\right]$$

$$= Vr \sin \omega t \frac{\sin N\,(\omega_m t/2 + \phi/2)}{\sin (\omega_m t/2 + \phi/2)} \qquad \text{..... (4)}$$

where $N = 2M + 1$.

V_T is maximum when $\omega_m t/2 = \phi/2$ giving a pulsed signal in which ϕ and therefore θ is a function of time - equation (2).

Consider a scan varying from $-\frac{Tm}{2}$ to $+\frac{Tm}{2}$. The maximum angle of scan - from (2) is

$$\pm\theta max = \sin^{-1}\frac{\lambda}{d}\pm\frac{Tm}{2Tm} = \sin^{-1}\pm\frac{\lambda}{2d} \qquad \text{..... (5)}$$

For $\theta max = 45^\circ$, the value chosen for the system $d = \frac{\lambda}{\sqrt{2}}$. The phase variation over the scan is from $-\pi$ to $+\pi$, and only one beam can exist between $-\theta m$ to $+\theta m$. However, it is apparent that "spurious" beams exist outside this arc, as with any linear array of spacing greater than $\lambda/2$. In our case they occur at approximately 90 degree intervals when from (1)

$$\theta = \sin^{-1}\left(\frac{\lambda}{d}\frac{\phi}{2\pi} \pm \frac{n\lambda}{d}\right) \qquad \text{..... (6)}$$

As the arc of interest is 90 degrees only, beams outside this region are suppressed by designing the primary antenna polar diagram to cover only $\pm 45^\circ$ to the 3 dB points. In practice the aperture is chosen to give a filled array, with the primary aperture equal to the antenna spacing d.

From (2)

$$\dot{\theta} = \frac{K}{\cos\theta} \qquad \text{..... (7)}$$

As with any fixed array, the beam width is inversely proportional to cos θ thus

$$\text{Beam dwell} = \frac{\text{beamwidth}}{\text{rate of scan}} = \text{constant.}$$

We have therefore a linear array scanning over 90° with a polar diagram of the form $\frac{\sin x}{x}$ and constant beam dwell. The rate of scan is unlimited and it repeats at periods $Tm = \frac{2\pi}{\omega_m}$. The scan is clockwise or counter-clockwise according to the arrangement of the modulating frequencies across the array, and continuous in the sense that as a scanning beam leaves at $+45^\circ$ another beam enters at -45°. At this point in the scan cycle beams exist at $\pm 45^\circ$ simultaneously.

In practice the modulation is carried out at I.F., - Fig. 2, rather than R.F. as shown in Fig. 1, and circuits are included in each I.F. channel to allow the adjustment of gain and phase across the array. It is usual for aperture taper to be applied for side-lobe control and this is allowed for in the design of the gain control circuit.

Turning now to array efficiency, let us examine antenna channel r. Normally the R.F. bandwidth is some MHz so that the external and circuit noise bandwidth greatly exceeds that of the signal. The spectrum of signals and noise after scanning modulation has been applied is as shown in Fig. 3. If we simply add the outputs then all channels will contribute noise at any frequency $(\omega + \omega_m)$ and the output signal to noise power ratio at that frequency will be N times worse than that in channel r alone. For optimum efficiency we must include a filter whose bandwidth is matched to that of the signal before combining the outputs. The system output for an input signal to noise ratio of $\left(\frac{Vi}{Vn}\right)^2$ is

$$\frac{Vi^2}{Vn^2} = \frac{(NVi)^2}{N(Vn)^2} = N\frac{Vi^2}{Vn^2}$$

since the signals add coherently, and the aerial gain is N, the number of primary aerials in the linear array

This is exactly what would be required if a linear array could be rotated, very fast, mechanically and with θ varying as in equation 7. The individual element signals would be shifted in frequency precisely as in the electronically scanned arrangement, and filters in each channel would be required for optimum efficiency. In practice mechanical rates are such that the frequency shifts imparted are very much less than the signal bandwidth, consequently producing a negligible effect on efficiency.

We noted in Section 2 that the scanning period is $Tm = \frac{2\pi}{\omega_m}$, and that we must scan at the Nyquist rate to extract all available information. The basic modulation frequency ω_m must therefore be equated to the signal bandwidth - we prefer it to be somewhat greater in practice. The scanning, in effect, samples any signal from a direction $-\theta max$ at sc · time $-\frac{fm}{2}$ and so on to $+\theta max$ at $+\frac{fm}{2}$. All signals are sampled perfectly and the array is able to resolve signals from N different directions, providing N simultaneous information channels. The price we pay is in post scanning bandwidth which is necessarily N times the signal bandwidth, i.e. $N\omega_m$ as shown in figure 1.

3. APPLICATION TO RADAR

While our application is to radar it is worth noting that the array described may be used for any communication system. For radar Tm is matched to τ the transmission pulse length and the transmitting array must floodlight the required surveillance volume. All targets are scanned once per pulse length so that all targets at all ranges are sampled during one repetition period (i.e. from zero to maximum range). A convenient type of display is shown in Fig. 4 where range increases vertically and angular direction is shown horizontally. In this raster display the horizontal scan period is equal to the radar pulse length and the vertical is the radar pulse period.

At first sight it may appear that the overall system is deficient in performance by the conventional transmitter gain. In fact this is not so because a return echo is received for every transmitted pulse, whereas in normal radar systems we receive returns only when the transmitter array is pointing in a given direction. With within-pulse-scanning, however, we must integrate the received signal to keep the transmitter power to a reasonable value for maximum range targets. In theory the gain from integration exactly matches the transmitter array gain of the equivalent mechanically rotated array if we match data periods. In practice because we must integrate after detection there is some loss of efficiency in the integration.

A more basic explanation is that in any data period radars of equal mean power illuminate targets with equal energy: in within-pulse-scanning it is at low beam power over the whole data (integration) period; with a rotating array it is at high beam power for only a fraction of the data (scan) period.

The loss in performance caused by integrating after detection is counterbalanced by the averaging of target returns over the integration period. The slow fading target for a mechanically rotated radar is equivalent to a fast fading for electronic scanning so that the lower integration efficiency equates, more or less, with the fading loss and the overall performance is the same for both systems.

It is perhaps appropriate, at this point, to consider the properties of a within-pulse-scanning radar and to see what advantages or disadvantages it gives. First of all it provides simultaneous information on all targets of interest. The data rate can be set at will but if we set a limit on mean transmitter power we must integrate for long range targets. As the target range reduces the amount of integration necessary for a given probability of detection decreases rapidly and we may use very short data periods, down to the limit of one pulse period at very short range. The integration time as a function of range for 50% probability of detection and 10^{-6} false alarm probability is given in Fig. 5.

The M.T.I. performance will be very good. Electronic scanning, as with rotating arrays, can impart target echo modulation, but by proper design of the scanning modulation, that is by making the radar p.r.f. an integral number of electronic scanning periods, we can avoid any increase in spectrum of clutter signals. By so doing we cause the angle sample for a given target to be taken at precisely the same point on the pulse envelope and the antenna appears to be stationary as far as clutter and target are concerned.

We pay a price in signal processing bandwidth and in multi-channel receivers, but the cost of the latter is very much less, being at small signal level, than in any system having combined receiver transmitter electronic scanning. The transmitter and its antenna are very simple. With the ability of now building very cheap integrated circuits we feel that such a system is economically viable.

The antenna array being stationary can be made large and angular resolution need no longer be related to wavelength. We may have microwave resolution characteristics at metric wavelengths if required, the area of the site being the only penalty.

4. SIGNAL AND DATA PROCESSING

In the experimental equipment the radar pulse length selected was 10 μS to keep the post scanning bandwidth at a value appropriate to techniques available at that time. Even so with 63 aerial channels it is 6.3 MHz.

The first problem with Nyquist rate scanning is to integrate signals received over the data period to provide reasonable radar efficiency for a given transmitter mean power. In practice, this means integration over the period during which a target can be expected to remain within one resolution cell. It follows that we must have an integrator capable of working at about 6.3 MHz, and with sufficient linearity to cover the integration gain expected. In the experimental installation the radar characteristics were:-

Pulse length - 10 microseconds,

P.R.F. - 500 Hz.

The antenna beamwidth allowing for amplitude taper for sidelobe control is about 1.8° or 50 beams per 10 μSecs. The integrator adopted in the first place was a conventional direct view storage tube with the circuits adjusted for linear integration over a period of 8 seconds, that is, an integration over 4000 points. The requirement of 50 resolvable cells in azimuth and 200 in range is met easily, but the video bandwidth and dynamic range requirements were quite severe. Experimental results suggest that there is about 6 dB loss of efficiency in the integrator compared with theoretical.

Integration on the D.V.S.T. necessarily brings the tube background to some low intensity and the signal when integrated appears as a light spot on this white background. Visual acuity sets the limit on detection. The relatively low efficiency is due in some measure to the poor contrast and to some extent to the fact that, in practice, it was not possible to drive the tube hard enough to produce false alarms without at the same time causing distortion which caused even greater loss of efficiency.

In view of the low efficiency of the D.V.S.T. an electronic integrator using ultra-sonic delay line stores was developed. A block diagram of the device is shown in figure 6. It consisted of a quartz delay line of 200 microseconds operating at about 25 mega-bits per second. The video signal is digitised at the input to the delay line and re-constituted at the output. This effectively removes any distortions imparted during one circulation through the delay line and permits indefinite storage of signals, if arrangements are made for re-circularisation. We see from the block diagram that the delay line contents are controlled by a master clock. The angle scan and the p.r.f. of the radar system are also controlled by the same clock so that the position of any sample of video information is accurately fixed in relation to the angle and range information provided by the radar system.

It will be noted that the angle scan period is equal to the pulse-width of 10 microseconds, and that during the pulse-width 50 beams are scanned. We require, therefore, 50 video samples every 10 microseconds The mode of operation is as follows. A trigger pulse delayed on the transmitter pulse by a given amount opens the integrator. Video signals are digitised and entered in the delay line over a period of 200 microseconds (20 range quanta). The information contained in the delay line is re-circulated for the remaining 7.8 mS (9 x 200 μSecs), when the next video input is again digitised over a period of 200 microseconds and added to the information circulating in the delay line. This process is repeated for a period of 8 seconds when the delay line contents are read out. It will be seen that the integrator accepts signals over any 20 mile range interval, each repetition period being precisely related in time.

It can be shown that hard clipping - two level quantisation - introduces a loss of integration efficiency of 1.9 dB, a small price to pay for the reduction in system complexity. With two level quantisation 10 bits are required to integrate 4000 separate samples. The ultra-sonic store was capable of operation at 25 mega-bits per second so that two lines in parallel provide the required capacity.

In the experimental work the electronic integrator was used with a variable trigger delay so that an assessment could be made of the system performance at any range. In practice, twenty integrators would be required and it would be relatively simple to set the integration according to range to give the optimum data rates.

5. RANGE PERFORMANCE

The range performance of a floodlit system depends on some post-detector integration with its associated integration loss, which, as has been stated in section 3 is counterbalanced by the averaging of target returns over the integration period. The object of the range performance measurements was to confirm that this averaging of target returns did in fact make up the integration loss.

A series of flight trials was carried out with two aircraft having scattering cross-sections between 2.5 and 40 sq.m. head and tail-on. Radial flights were carried out at heights consistent with the vertical radiation pattern of the aerials. The presence or absence of a target return was recorded at 30 sec. intervals by an observer looking at a direct view storage tube which had a known integration time constant. The results are presented in histogram form in Fig. 7 where the percentage of detections in given range intervals are plotted for the four different scattering cross-sections. The maximum possible number of observations in any range interval are also listed.

In order to assess these results it is necessary to compare with the theoretically expected performance. Most of the radar parameters can be determined with reasonable degree of accuracy by measurements on the actual radar and values inserted into the radar equation to calculate maximum range for a given probability of detection and false alarm rate. It was not however possible to make a direct measurement of the signal to noise ratio after integration by the storage tube, but comparison with measurements on the digital integrator indicated that when a target detection was made the signal to noise ratio was consistent with a 95% probability of detection and a very low false alarm rate. It was assumed in the calculations that the fading loss was zero, that is it was assumed that the signal did not fade.

The expected maximum range for 95% pd and false alarm time of 60 min. is plotted on the histograms. These show that the range performance of the radar is consistent with the theoretically predicted value, and consequently that the integration loss factor is counter balanced by the averaging of target returns over the integration period.

6. MOVING TARGET INDICATOR PERFORMANCE

For a Nyquist rate scanner where the angle scan and prf are frequency locked, the MTI performance is not limited by scan modulation as with a common transmit receive system, and the removal of this limitation becomes important for a radar having high angular resolution and short data interval. A range

gate filter (pulse doppler) system was utilised with the experimental radar to demonstrate the MTI performance and attempt to establish the areas which limited the MTI performance.

The range gate filter system covered 100 resolution cells, sufficient for experimental purposes, and the filter response was such that at near zero frequency the attenuation was about 70 dB decreasing very rapidly to zero attenuation at about 65 Hz which is just over one tenth of the unambiguous frequency band (0 to 500 Hz).

The performance tests show that the MTI capability of the radar is not limited by scan modulation and clutter attenuation of 45 dB or better has been measured on clutter returns within the scan sector with the receiver beam scanning at 100 kHz. This limiting value was not fundamental, and not imposed by the scanning frequencies but was associated with random phase noise introduced at the pulse modulator of this experimental radar. There is no reason to suppose that the limitations to performance on this radar are not similar to those of a conventional radar but with its beam stationary. The removal of the scan modulation limitation is extremely effective. For a conventional system having a dwell time of 10 m.sec. corresponding to 5 pulses per target the maximum clutter attenuation would be limited by scan modulation to about 15 dB for the particular filter characteristic described here.

7. CONCLUSIONS

The techniques described give the possibility of high data rates without loss of efficiency in the utilisation of energy. In practical applications such a system should be much cheaper than electronically scanned radars using common transmit/receive principles.

There is no loss of information in the electronic processing and the designer is presented with a different set of constraints. The techniques may be applied with any type of transmitter modulation - including frequency modulation for pulse compression applications. It offers very good M.T.I. in conjunction with high angular discrimination.

There are, of course, penalties to pay. It is a one way side-lobe system and special attention must be given to their control. The high data rate is obtained at the expense of signal bandwidth, which present difficult but not impracticable signal processing problems.

8. ACKNOWLEDGMENT

Contributed by permission of the Director R.R.E.

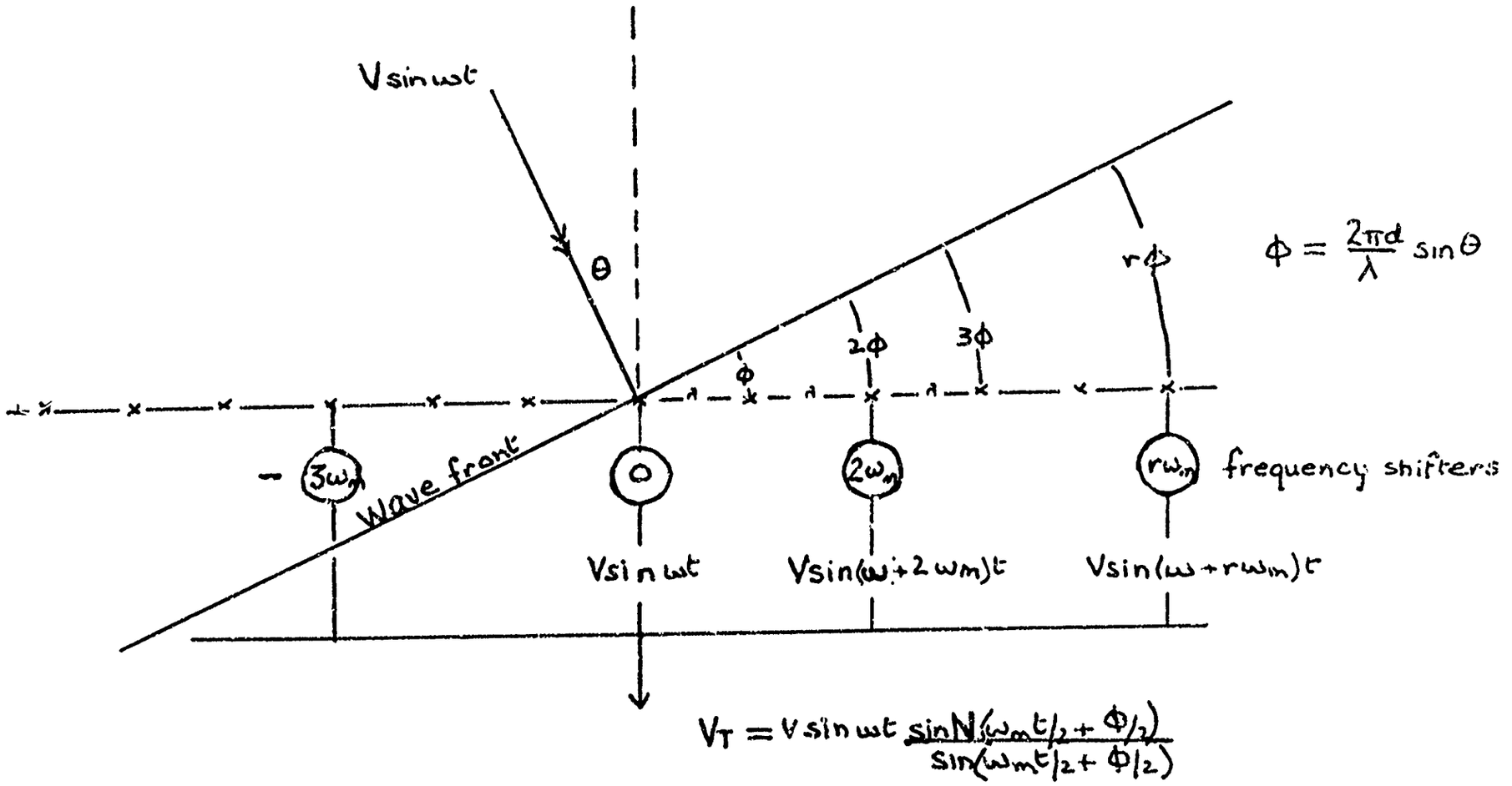


Fig.1 Electronic scanning with a linear array

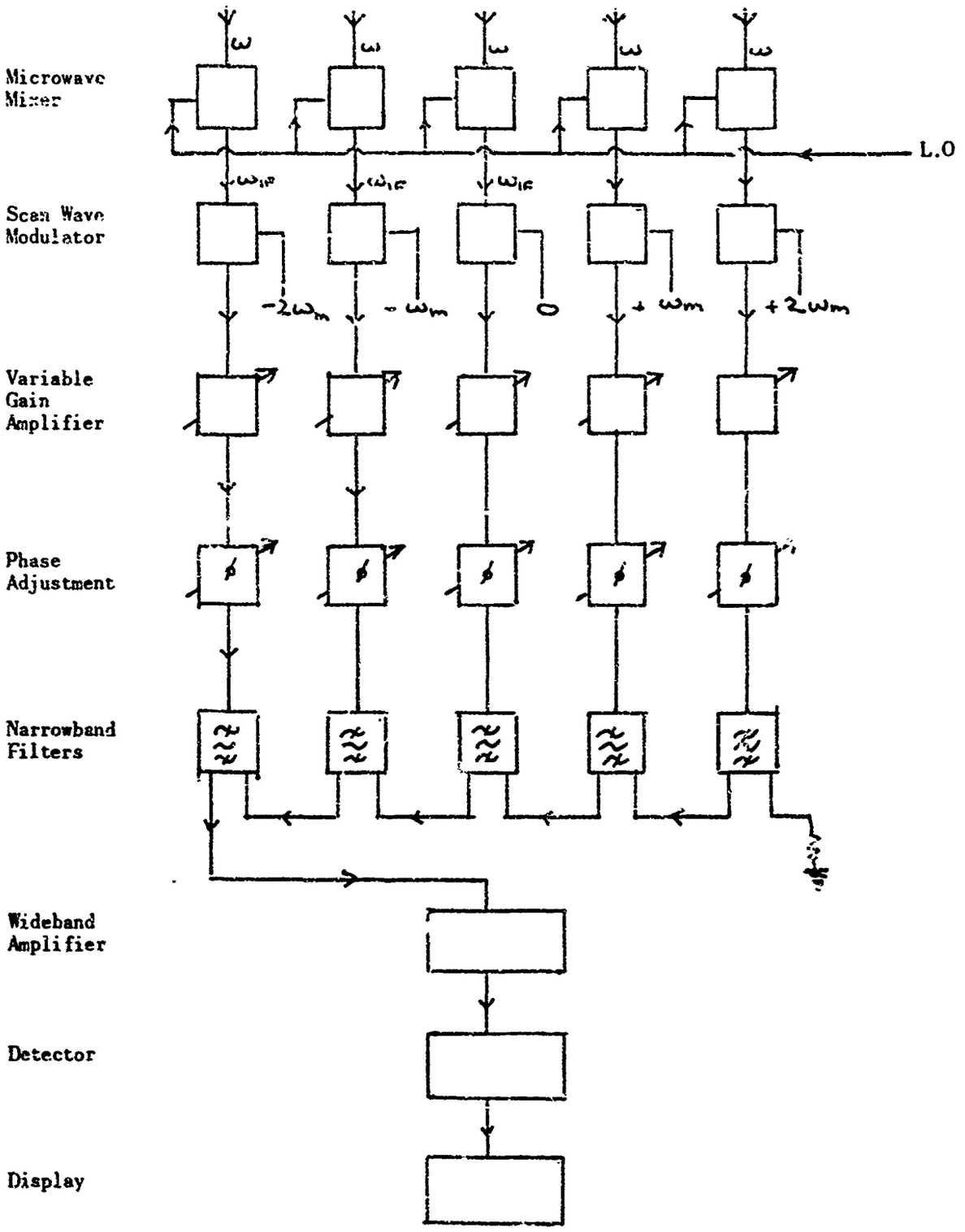
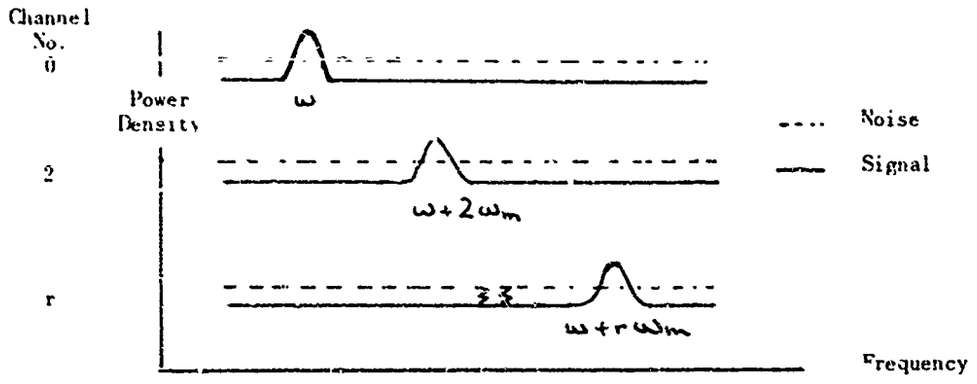
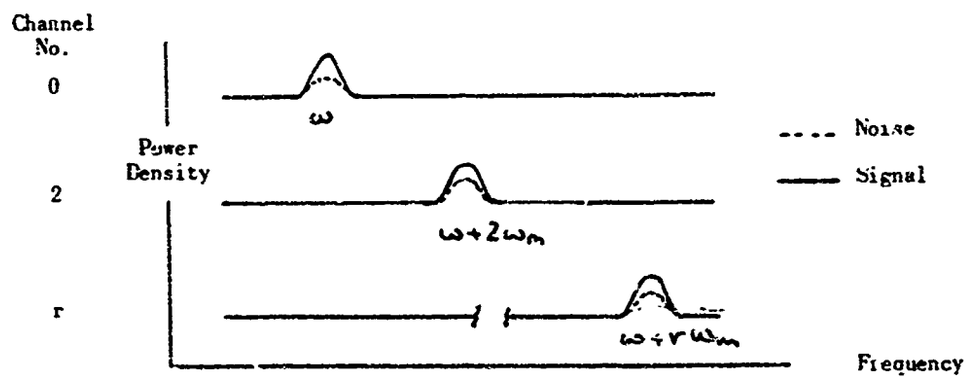


Fig.2 Electronics scanning - basic system



Spectrum before filtering



Spectrum after filtering

Fig.3 Primary array spectra before and after filtering

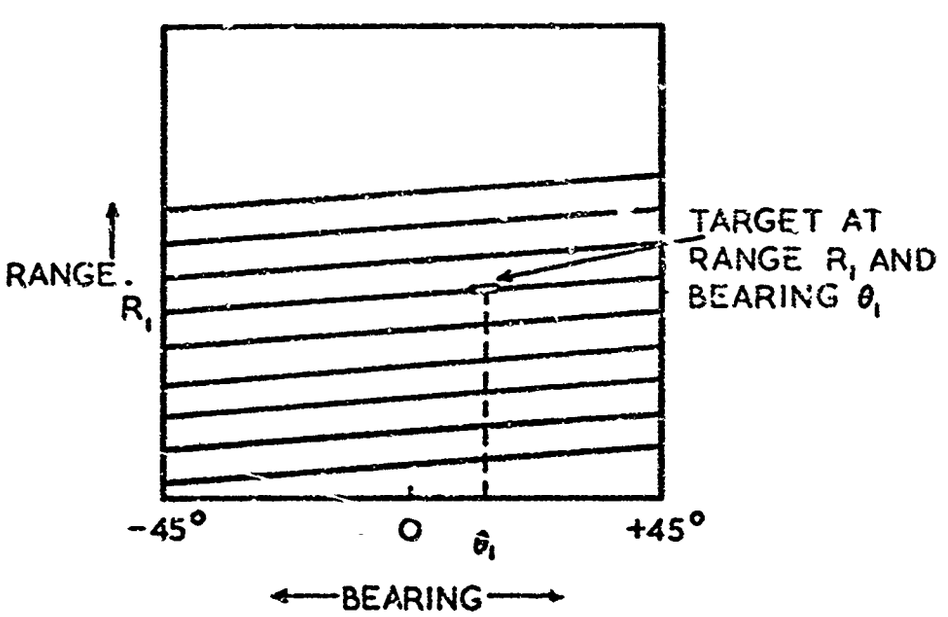


Fig.4 Display format

Integration time as a function of range for $P_D = 90\%$ and $P_F = 10^{-6}$ and correlation over 50 m sec.

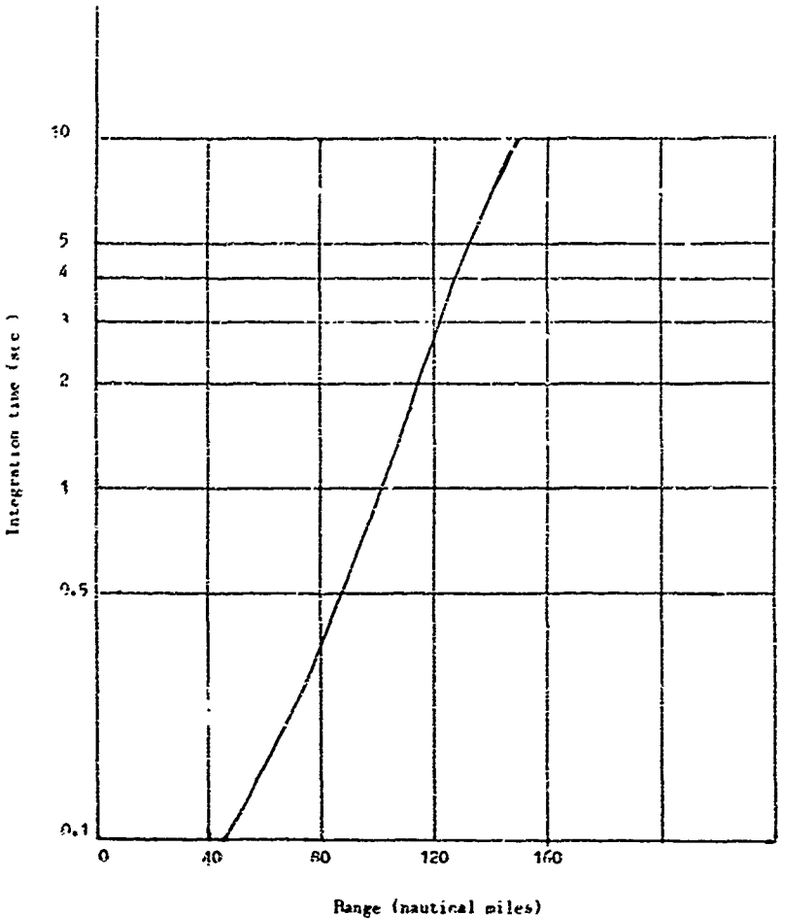


Fig.5 Integration as a function of range

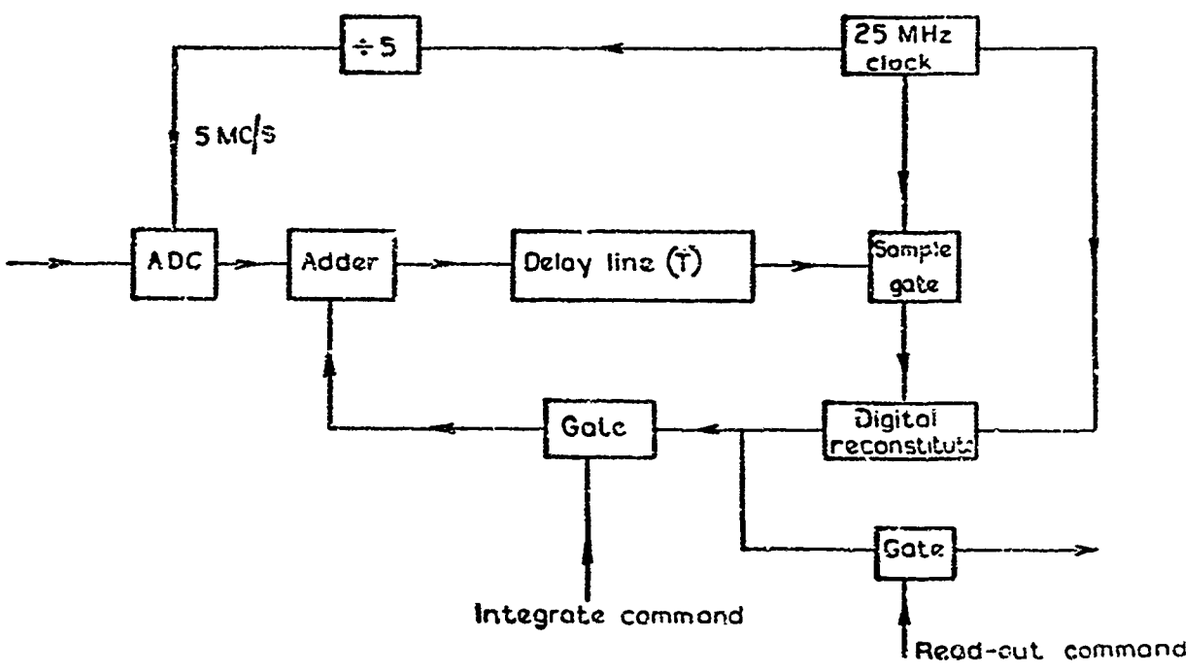
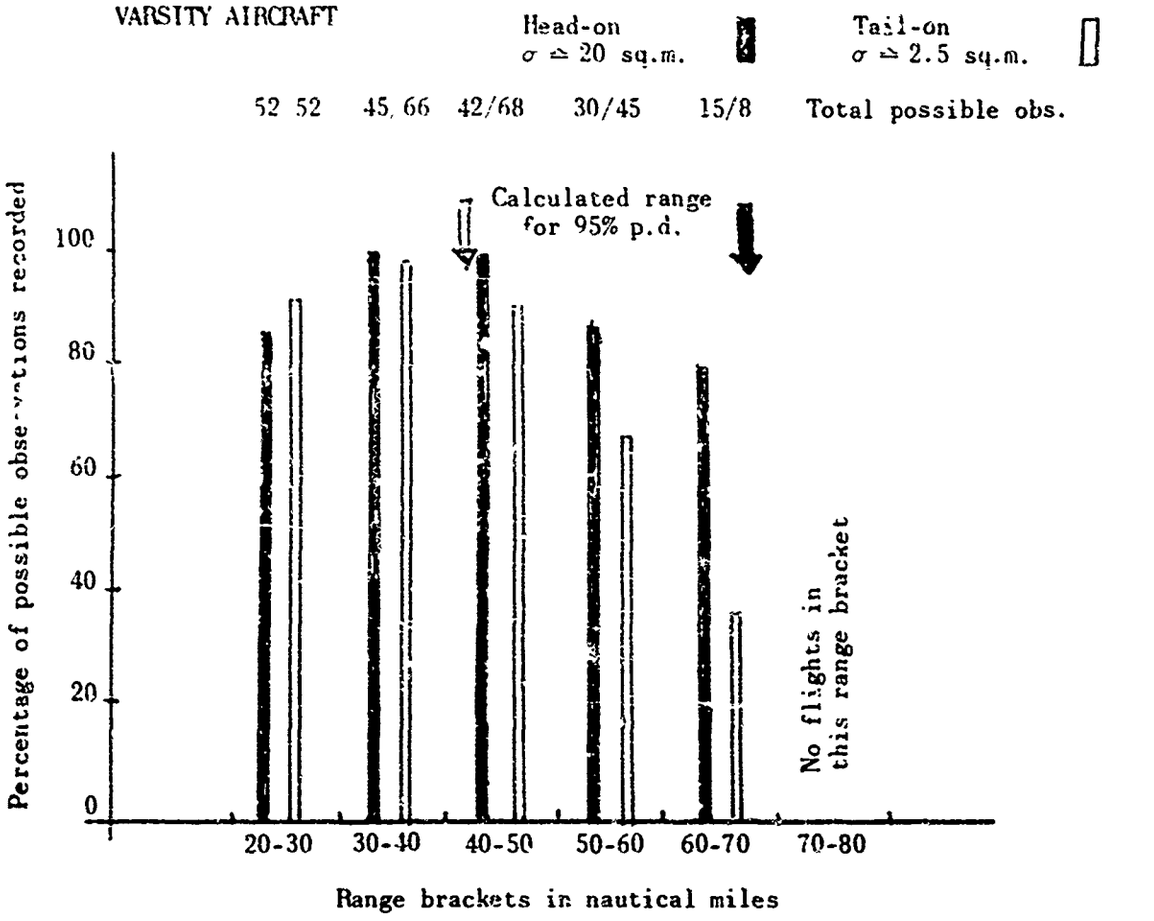


Fig.6 Recirculating digital integrator



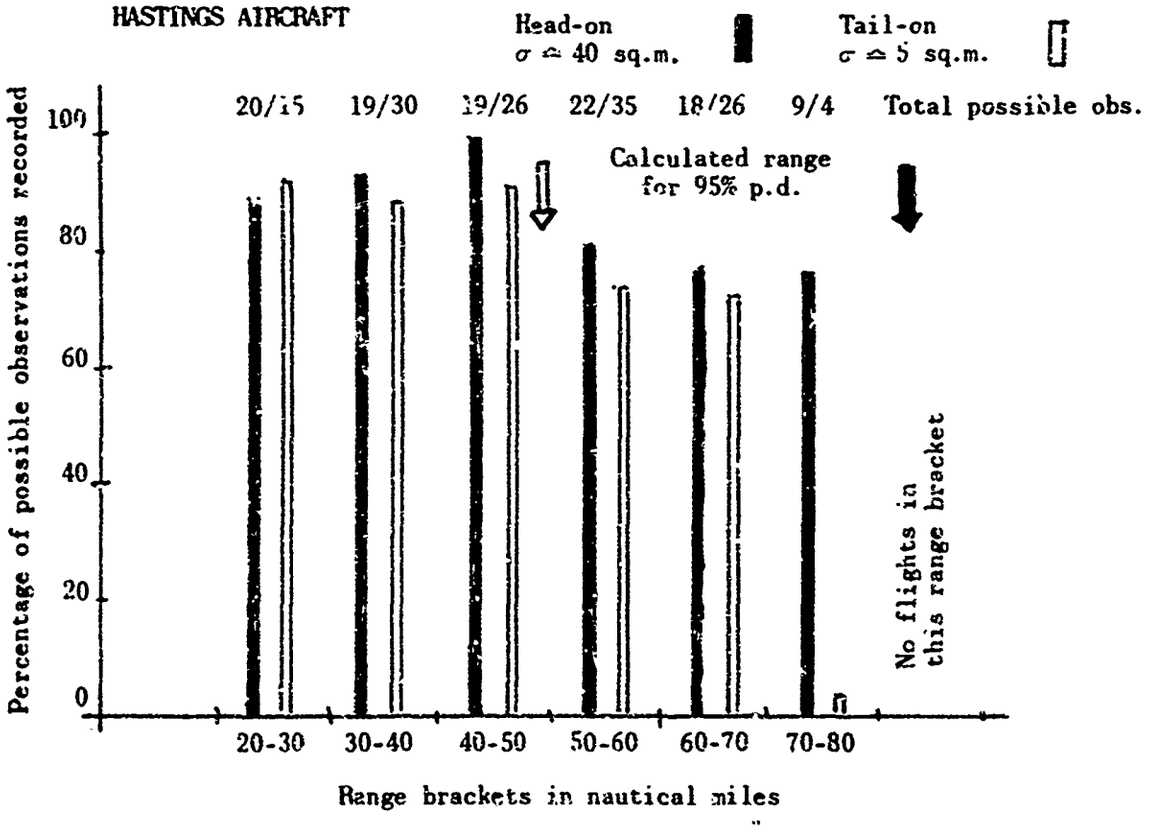


Fig.7 Range performance

PHASED ARRAY RADARS FOR AEROSPACE APPLICATIONS

by

Arthur S.Robinson

RCA
Morrestown, New Jersey 08057, USA

PHASED ARRAY RADARS FOR AEROSPACE APPLICATIONS

Dr. Arthur S. Robinson
RCA
Moorestown, New Jersey 08057, USA

SUMMARY

Phased array radars have been contenders for aerospace applications for many years. Until recently, the conversion of this potential into competitive equipment designs has been severely limited by technology and cost considerations.

A number of major trends have now emerged that point the way to highly competitive aerospace phased array radar designs. This paper highlights these trends by exploring the relative roles of operational requirements, system designs, science, and technology, in achieving cost-effective phased array radar designs.

The potential advantages of phased array radars for aerospace applications have been widely recognized for many years. Unfortunately, the conversion of this potential into competitive equipment designs has been limited by relatively slow rates of progress in a number of critical areas of technology.

During recent years several major technology trends have emerged that overcome these former limitations, so that phased array radars can and are now being developed as cost-effective solutions to a broad spectrum of aerospace problems. This paper provides a system perspective of these current trends in phased array radar design, and of additional trends to be expected in the future.

Let us first consider the role of life-cycle costing in comparing alternative systems for solving defense problems. A simplified representation of the process used to select a minimum cost system is suggested in Figure 1.

Problems to be solved can appear initially in many forms - ranging from a formally documented operational requirement to informal knowledge of unsatisfied needs. Early in the game the nature of these needs have to be defined more precisely - usually in terms of system performance requirements against a "threat" model. Initial temptations, seldom resisted, tend toward establishing very high levels of performance against worst case levels of threat, that is, the number, distribution and speed regimes of targets; clutter, chaff and ECM levels, etc. In response to the performance and threat, so defined, various fundamental systems for coping with the problem are postulated, and for each such system, alternative competing equipment embodiments are conceived. All of the resulting equipment designs are then weighed in the light of current and projected technology, including information gained in prior related equipment and system developments. From these considerations, life-cycle-cost models are evolved to define the full costs to the military for developing, producing, operating, maintaining, supporting, and protecting such a system in the field throughout its life, or for some fixed period of time.

Initial temptations having been succumbed to, these costs are usually substantially greater than the bounds established by budget constraints, and so compromises in system performance and/or in the level and complexity of threat become essential. While some elements of the threat, and some performance requirements, cannot be changed, other elements often represent desirable, rather than essential, constraints. Perturbation of these variable elements of performance and threat begins the second turn around the loop, with iterations continuing until a suitable threat/performance/system/equipment/cost compromise is found, or until it becomes apparent that the performance and level of threat that can be met, does not merit the cost.

Sometimes the judgments involved in these decisions become a matter of national debate. More often, necessary compromises are made with full knowledge that the system selected is not the ultimate answer to the problem, but is simply the best answer within the constraints dictated by technology, the national economy, and national and military priorities.

During the past twenty years, phased array radars have appeared as candidates for the solution of many different problems. During most of this period, technology limitations have worked strongly against the success of these systems. Some of the phased array programs initiated during the period foundered on the rocks of overly optimistic estimates of technology capabilities and overly complex system designs - made necessary by unrealistically high levels of assumed threats. Other programs accepted the limitations imposed by technology, but then suffered from a level of performance that was not really adequate in certain critical problem areas. One of the most popular compromises was the use of frequency scanning, a technique intrinsically limited in bandwidth and susceptible to modest levels of electronic countermeasures.

During recent years, the phased array technology/performance gap has slowly closed. As a result, high-performance phased array radars have emerged as cost-effective sensors for a number of new defense systems, and as strong contenders for many future defense applications. Many of these new radars are based on variations to the design approach depicted in Figure 2.

In this type of system a single aperture is used for both transmit and receive. One or more transmitter tubes are used to generate RF power, with the energy from each tube distributed through an RF feed to a multiplicity of high-power phase shifters, each in series with a radiating element. Each phase shifter is controlled by a phase shift driver, with the driver in turn controlled by commands from a central beam-steering controller.

In radar designs requiring large apertures and wide bandwidths, the overall radar array is subdivided into a number of smaller subarrays, with each subarray driven from a single transmitter tube. The phase shifters in the main array face then serve to steer the relatively broad subarray patterns to the commanded pointing angle, while time-delay units in series with the coherent exciter signal feeding each transmitter serve to fine steer the narrow overall array beam to the precise commanded angle.

The function of the time-delay units is to compensate for the differential path lengths encountered by energy emanating from each subarray. If technology could give us low-cost, high-power, low-loss, time-delay units, we would use them in wide bandwidth systems instead of phase shifters, and not bother with subarrays. Since such units are not yet within the state-of-the-art, we use available, relatively high-loss, high-cost, designs in modest quantities and at low-power levels to achieve reasonable bandwidth capabilities, even with large apertures.

The subarray approach offers advantages even in moderate bandwidth applications, where fine steering phase shifters can be used instead of time-delay units. The advantages stem from the fact that comparably located phase shifters in each subarray receive the same steering command. This feature can be used to minimize the cost and complexity of the beam-steering controller and the subarray phase shift drivers.

In some smaller aperture systems the entire radar can consist of a single tube, feed and phase shifter array. In these applications the bandwidth requirements are low enough and the aperture small enough that neither time-delay steering nor subarrays are needed.

I will not attempt here to show all of the variations on this basic theme. The point to be made is that tube-driven radars using high-power, low-loss phase shifters following the transmitter are an important class of phased arrays that have only recently broken the cost/performance barrier, and that will, therefore, represent a major trend in phased array design for at least the next decade. We can expect to see arrays of this type implemented for land, sea, air, and space applications in the years ahead.

Science had a major role to play in this achievement - through improvements in tube elements, making possible superior transmitter designs, and even more important, through materials research that pointed the way to economical, efficient, high-power phase shifters. Technology gave us the actual design of the transmitters and phase shifters, as well as economical hybrid digital/analog phase shifter drivers; compact, efficient feeds, and superior radomes and radome phased array design techniques. Figure 3 is an artist's concept of several of these advances, combined in a portion of a phased array radar subarray that is particularly well suited to aerospace applications. RF energy is distributed through a very shallow depth feed to phase shifters, radiating elements, and finally, the radome. Figure 4 is a photograph of an advanced shallow depth phased array feed that has been recently developed at RCA. In the design shown, energy is distributed to 32 phase shifters through a power divider that is only 1 inch in depth. This compares to conventional design techniques, requiring in the order of 28 inches for this function. Figure 5 shows a single phase shifter, one of the 32 that fit into the feed. The radiating element that the phase shifter feeds is also illustrated in this view. Figure 6 shows how the phase shifters are mounted into the power divider. In this case 8 of the 32 phase shifters are in position. Figure 7 presents the complete 32-element subarray, the housing into which it fits (from the back), and the radome that covers the radiating face.

Technology of this type has made it possible for us to move from the large, fixed phased arrays of today to the transportable tactical phased arrays of tomorrow. These improvements are by no means complete. The years ahead will see continuing refinements and enhancements of this class of system, with particular emphasis on system producibility and on techniques for decreasing both hardware and software costs. Signal processing and data processing have been and will continue to be significantly advanced as part of this technological progress package. More about this later.

Nature, having been relatively unkind to phased arrays for twenty years, now appears to have decided to make up for this oversight by inundating us with a burst of new technology - some of it directly benefiting tube-driven phased arrays, some applicable only to new phased array design concepts.

Some of the technology pacing the new array design trend is suggested in Figures 8 and 9. Figure 8 highlights progress in solid state RF power sources and receivers, while Figure 9 shows typical devices, packaged with their associated circuitry in hybrid microwave integrated circuit form. Power curves are based on average power from a single active chip. Typical power amplifier modules will combine in the order of 2 to 10 such active elements on a single substrate.

In comparing transistor, bulk and avalanche power sources, it is important to remember that in many applications transistors cannot be operated at their full average power ratings, due to system duty cycle constraints. A derating factor of 5 to 10 is often necessary, depending on the permissible transmitter duty cycle. Bulk and avalanche sources, on the other hand, generate substantially higher peak than average power, so no such derating is necessary.

While it is very difficult to tie down a broad view of costs, order of magnitude figures have been included to provide a rough insight into the problem. The point here is that if the cost of RF power alone determined the cost of a system, solid state would not currently be in the running. But the fact is that completely solid state phased array designs are becoming increasingly competitive in many (but certainly not all) applications - when compared on a life-cycle-cost basis.

Figure 10 presents a typical solid state phased array radar design, illustrating some of the factors that tend to counterbalance the higher initial cost of solid state RF power. Notice first that each element is driven by a phase shifter in series with a power amplifier. Some tube designers feel that they will be able to compete with solid state power amplifiers in this type of duplexed array, but the trend here seems strongly toward solid state. In a conventional tube-driven phased array, in the order of half of the RF power is lost on the way to the radiating element - in the feed and in the high-power phase shifter. In the solid state design, the power amplifier drives the radiating element directly. Losses in the feed and phase shifter occur at low power and therefore have negligible effect on efficiency. Lightweight, low cost, compact, feeds and phase shifters become feasible in exchange for losses at a point in the system that does not significantly effect radar performance. In the conventional tube design the high-power phase shifters require relatively high-power drivers to set them on command, while substantially lower-power, higher-speed, drivers do the job in the solid state design. This increased switching speed is of particular importance in high-density tracking applications.

In the conventional tube design, all of the power lost in the phase shifters has to be carried away by an integral cooling system, above and beyond the transmitter cooling requirements. The solid state design requires cooling only for the solid state power amplifiers. In lower frequency applications, net solid state power amplifier efficiency will be substantially better than the efficiency of a competing tube transmitter. On receive, the solid state system amplifies the signal before it passes through a phase shifter. This minimizes the effect of phase shifter losses, and introduces a really exciting new flexibility in our handling of received RF signals. Because the signals are amplified immediately, they can drive a variety of lossy microwave circuits without significantly affecting system noise performance. Thus, lossy microwave time-delay devices such as microwave acoustic-delay lines become serious long-term contenders to replace phase shifters; multiple "receive" beams can be formed simultaneously; and circuits can be so configured that pilot signals received from anywhere within the antenna field-of-view automatically point the antenna beams at the pilot tone sources.

Think about the rich variety of antenna designs that have emerged in the past using essentially only passive components. As these new active RF devices evolve, we can expect many new and exciting improvements in the design of solid state antennas.

But this is only part of the tube vs. solid state story. Figure 11 gives an overview of the relative distributions of costs in a typical trade-off between two such systems. It has become clear that the actual life-cycle costs of many defense systems are strongly affected by the ease with which they can be operated, maintained, supported, and protected in the field. Reduction in the number of military personnel required to perform these functions is an area of great potential cost savings for each individual so reduced proportionately decreases the service personnel and facilities needed to maintain that man and his functions in the field. The most intriguing aspect of solid state phased arrays is their potential ability to meet high levels of threat, while simultaneously requiring substantially less field support due to their increased reliability, compared to tube-driven systems. In some applications these cost savings more than compensate for high initial solid state system costs - but not in all applications. Solid state is not a panacea; for example, it is not yet a serious contender for high-power applications at or above S-band. It is a contender for high-power applications at L-band and below, and for low-power applications in all microwave frequency bands. The only way to really know whether a solid state design is sensible for a given application, is to go through preliminary designs of the best competing solid state and tube approaches for that use, and to then compute their relative life-cycle costs.

A technology trend of importance for both solid state and tube applications is illustrated in Figure 12. We can expect the price break in medium and large scale digital arrays that is currently being projected by the semiconductor industry to have a significant impact on phased array radar signal and data processing. In signal processing, more and more processing functions will be implemented in digital form, while the analog-to-digital conversion function will move progressively up the RF chain. Highly adaptive techniques will thus become an economic reality, as signal processor action comes increasingly under programmable software control.

The structure of a typical high-speed-digital signal processor is suggested in Figure 13. For applications involving substantial bandwidths and time-bandwidth products, conventional digital time-shared data buses will give way to non-time-shared networks, switchable under computer control. Extensive paralleling and pipeline processing will compensate for data rate limitations of individual conversion and computing elements. Many variations on this theme are to be expected. For example, in applications requiring processing of only a limited number of range cells, it will be possible to minimize digital signal processor functions by loading analog-to-digital converter outputs into high-speed memory, and then using the system data processor to operate on these data at lower speeds. At the other end of the spectrum, in some applications it should prove feasible to digitize data directly after amplification at the antenna receiving elements, and to thereby include antenna beam forming as part of the digital signal processing function.

The flow of data and commands between the signal processor, radar controller, system data processor and the overall radar system is illustrated in Figure 14. In essence, the radar controller performs all functions necessary to match the signals flowing to and from the data processor to the rest of the radar system.

Data processing remains a major problem in many phased array radar applications. The systems of which these radars are a part often require reaction times that make almost completely automatic operation essential. For demanding applications, establishing the computer architecture best suited to the threat is a major problem. Once the architecture is established, developing software to cope with the constantly shifting spectrum of targets, clutter, chaff, jamming, decoys, interference, etc. is a really monumental undertaking. We can expect this area to assume increasing importance over the years, with a greater share of system investment going into improving real-time programming, as well as into developing the programs themselves.

For complex applications, data processing architecture is moving toward the simultaneous utilization of a multiplicity of processors. Figures 15 and 16 present one of many competing architecture concepts[1] in this area, in order to provide some insight into the level of complexity of these machines. The computing modules shown in Figure 15 all operate under control of an executive computer. The executive allocates portions of the data processing task to individual modules, and establishes appropriate data routing, consistent with these tasks. Adjacent modules communicate through local data exchanges, while all modules can exchange data and instructions with one another and the executive over time-shared data and instruction buses. It is possible to so configure hardware and software for this system that the executive function can be assumed by any module. This doesn't do much for the ego of the executive, but it does quite a bit for the reliability of the data processor.

Each computing module consists of a multiplicity of digital processors, as shown in Figure 16. These processors can communicate to their adjacent neighbors within the module and to their neighbors in adjacent modules. In the example shown, a total of 16 processors are used in a computing module, while 16 computing modules are combined in the data processor - a total of 256 computers operating simultaneously on a common problem.

There is a final trend that I would like to call to your attention - the need for multiple function, multiple frequency, phased array antenna complexes, capable of servicing all of the antenna requirements for a given vehicle. This problem is particularly acute in aircraft, spacecraft, and ships, and in certain tactical Army applications where the available space for well situated antennas is extremely limited, while the number of subsystems requiring antenna access is constantly growing. Expect a trend, then, toward unified designs of the phased array antennas for such systems, consistent with the frequency and time conflicts intrinsic in any time-shared operation. Such conflicts will make the utilization of a single antenna for all functions extremely difficult when a multiplicity of systems are involved - a small cluster of arrays is a more likely possibility. There is every reason to believe, however, that time and frequency sharing of phased array antennas between a multiplicity of systems will come into widespread use in the years ahead.

REFERENCE

1. This approach is closely related to the University of Illinois, ILLIAC IV.

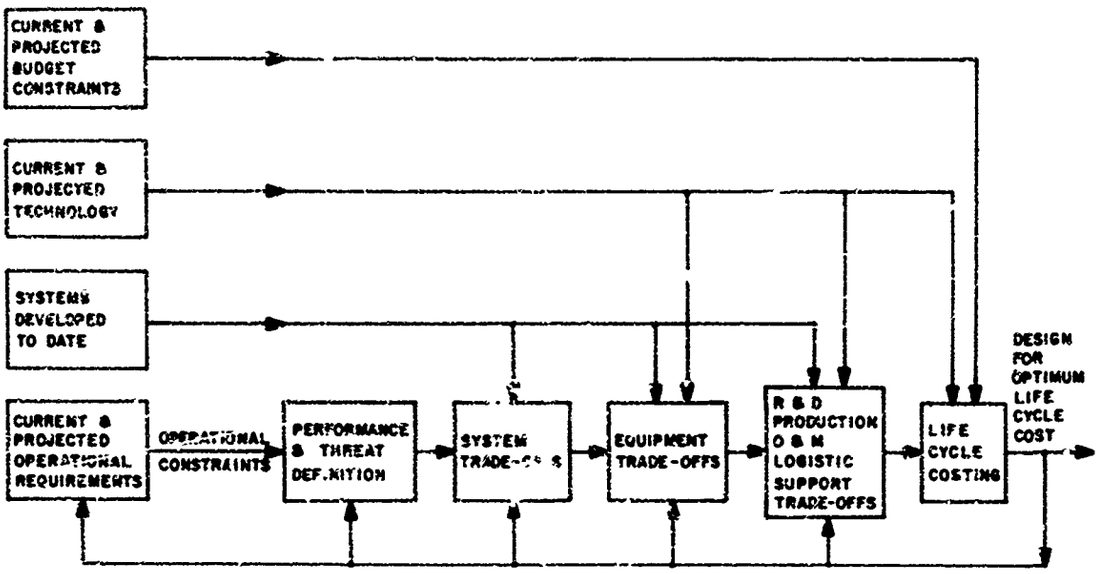


Fig.1 Some fundamental factors affecting the design of next generation radars

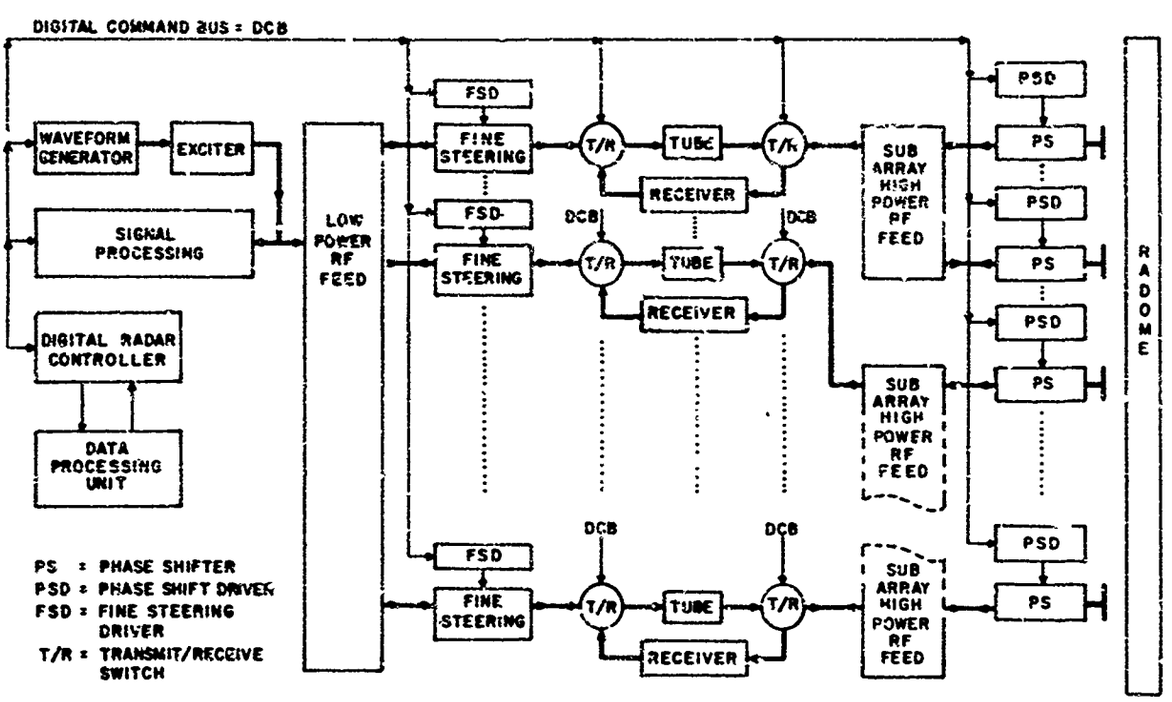


Fig.2 Simplified block diagram - tube-driven phased array radar

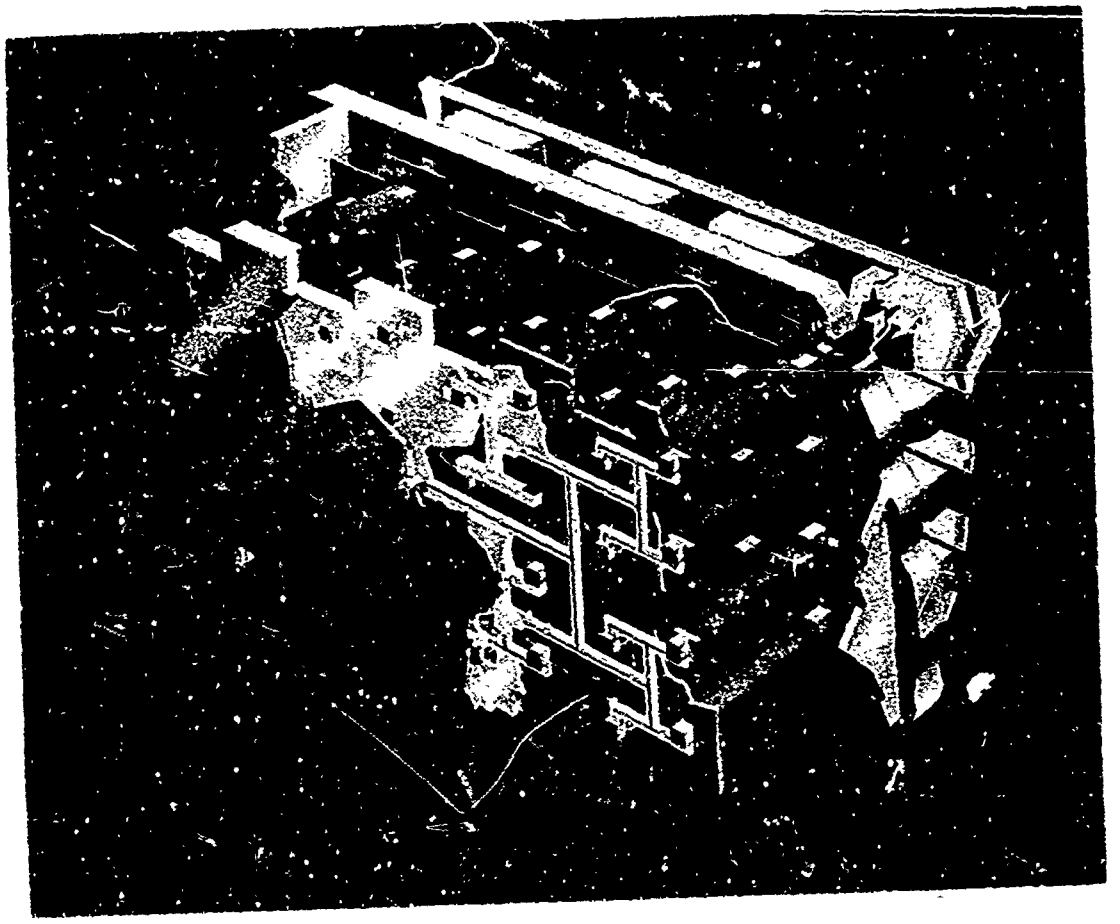


Fig.3 Phased array radar subarray - artist's concept

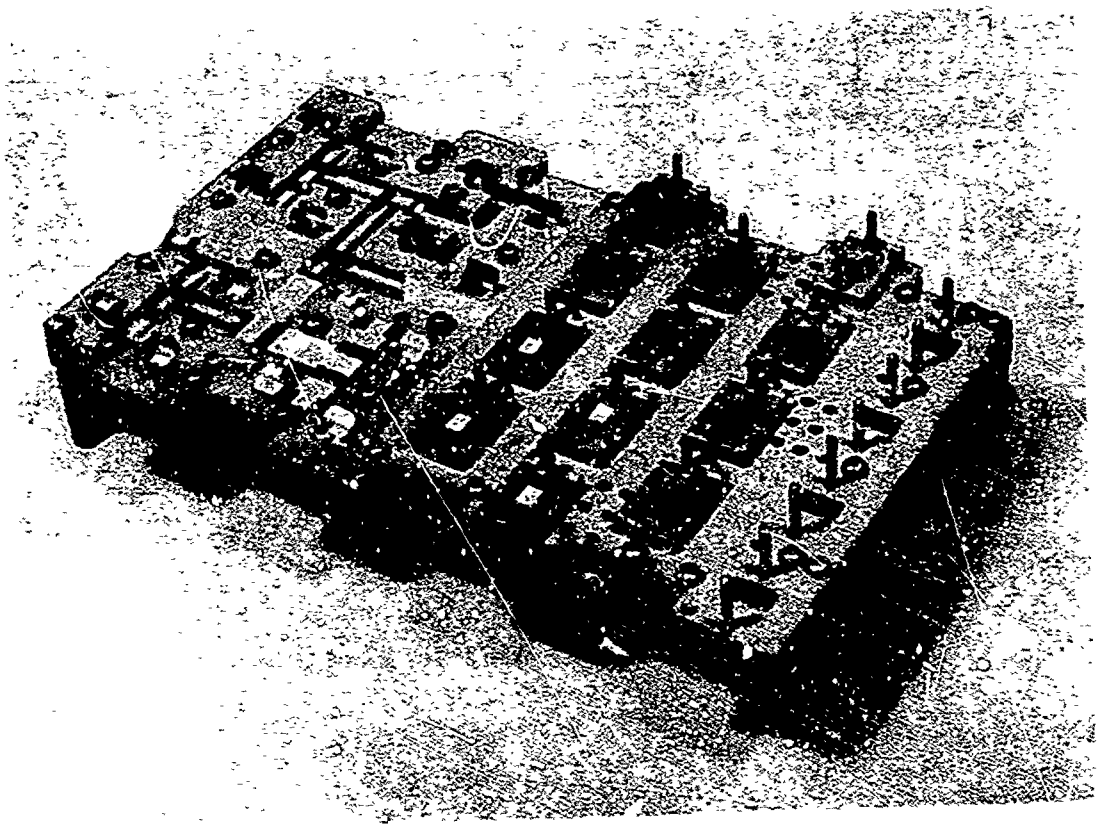


Fig.4 Shallow depth phased array feed

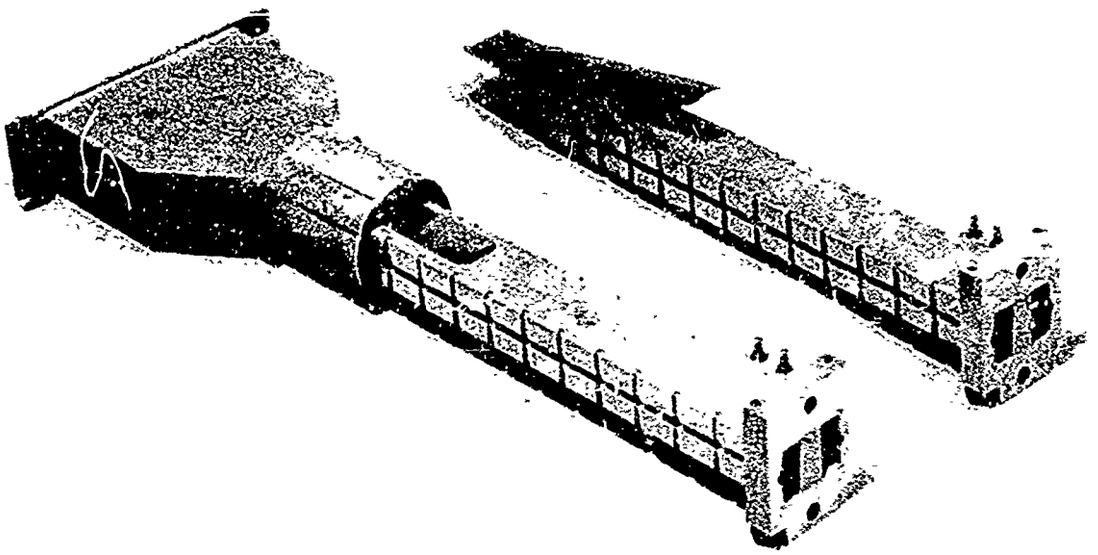


Fig.5 Phase shifter and radiating element

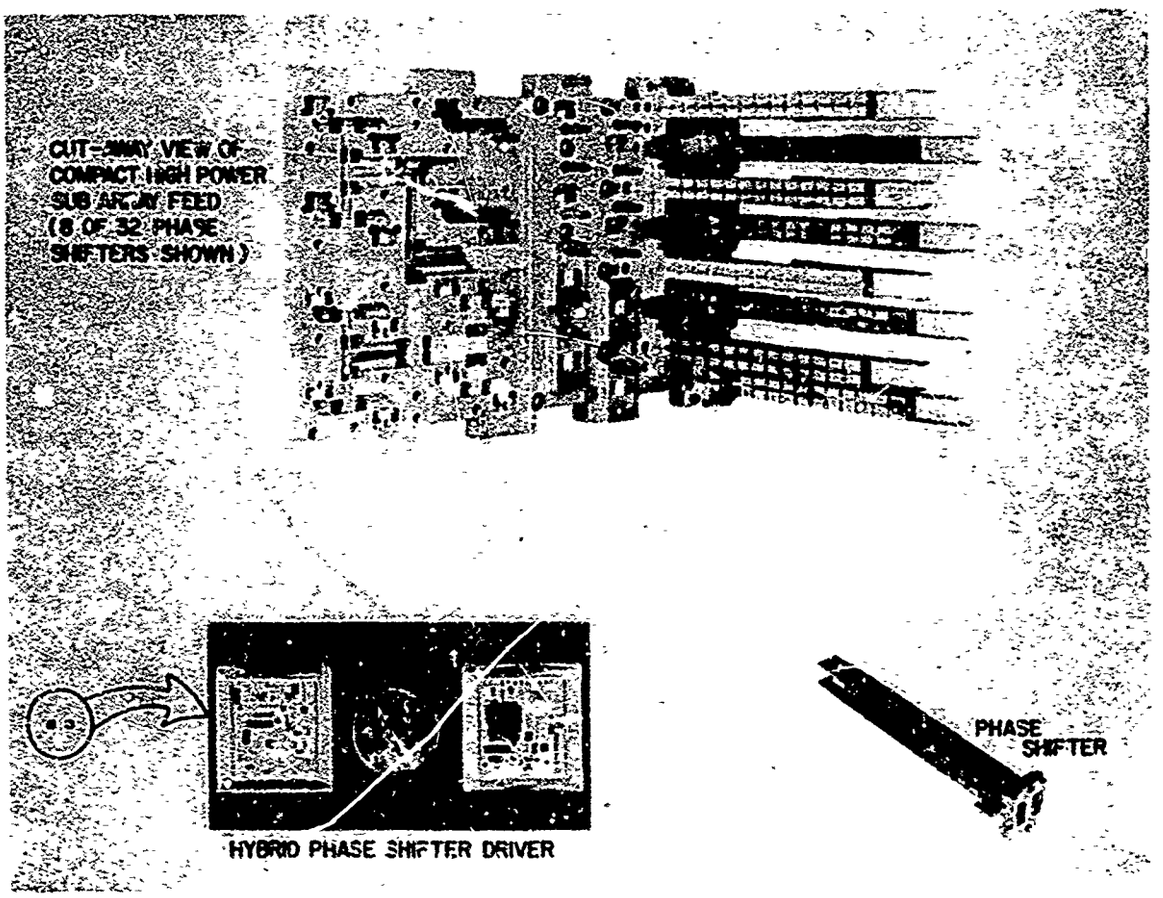


Fig.6 Phase shifters mated to shallow depth feed

Fig.7 Subarray face, mating phase shifters and feed

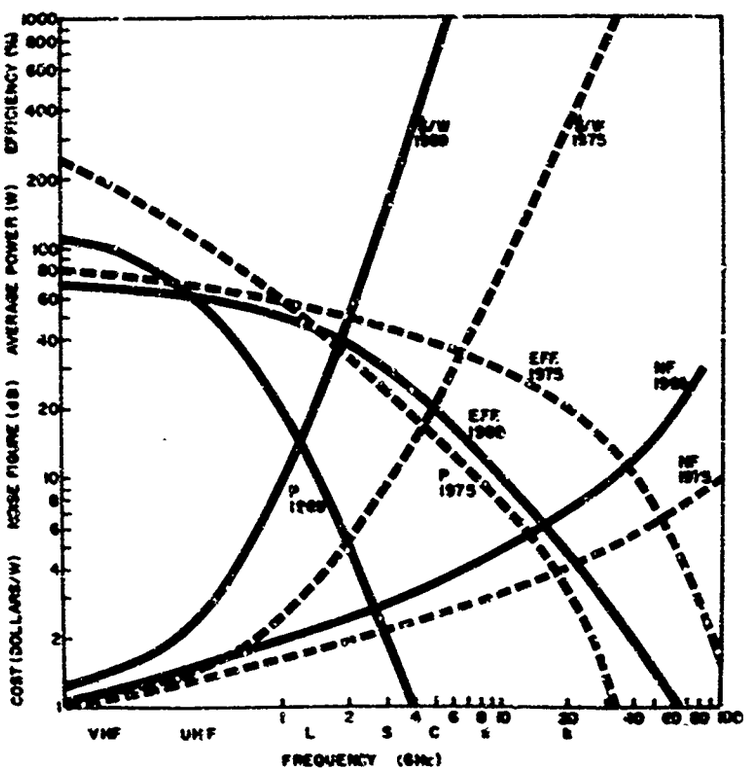


Fig.8 Solid state device trends

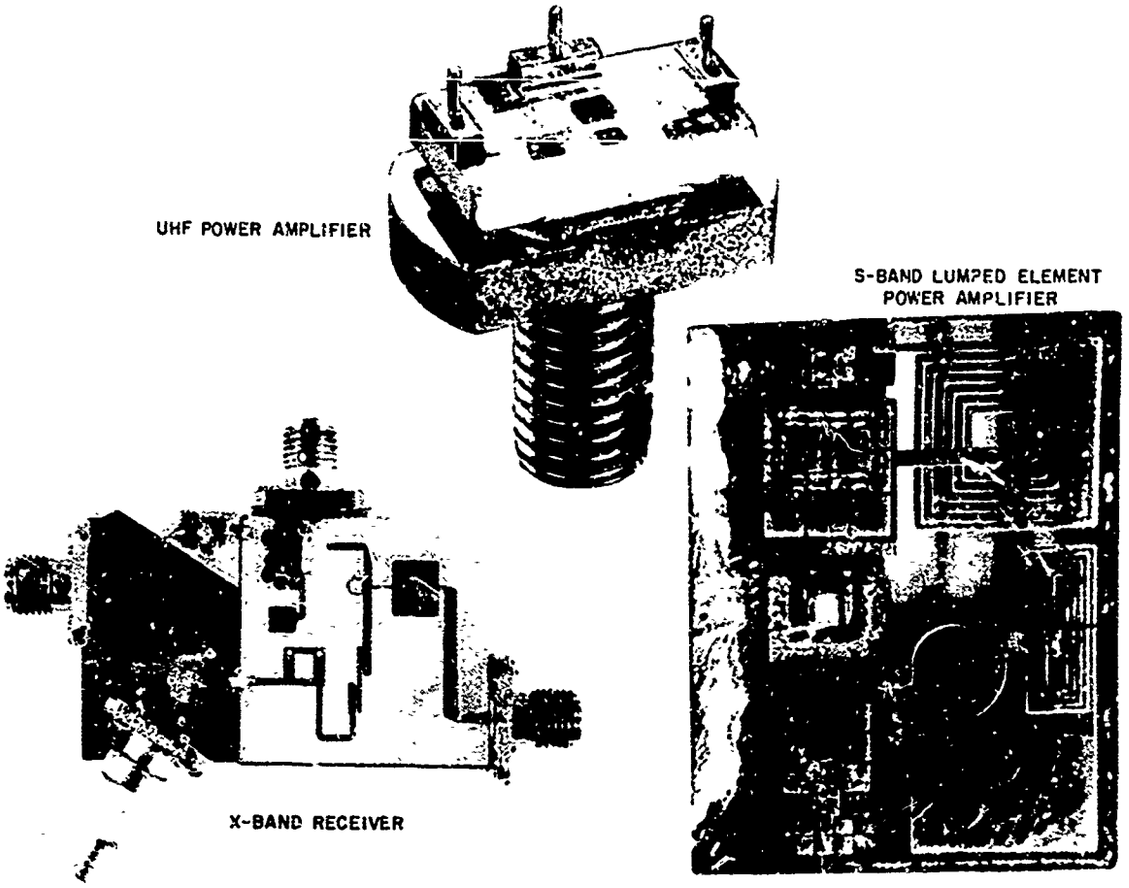
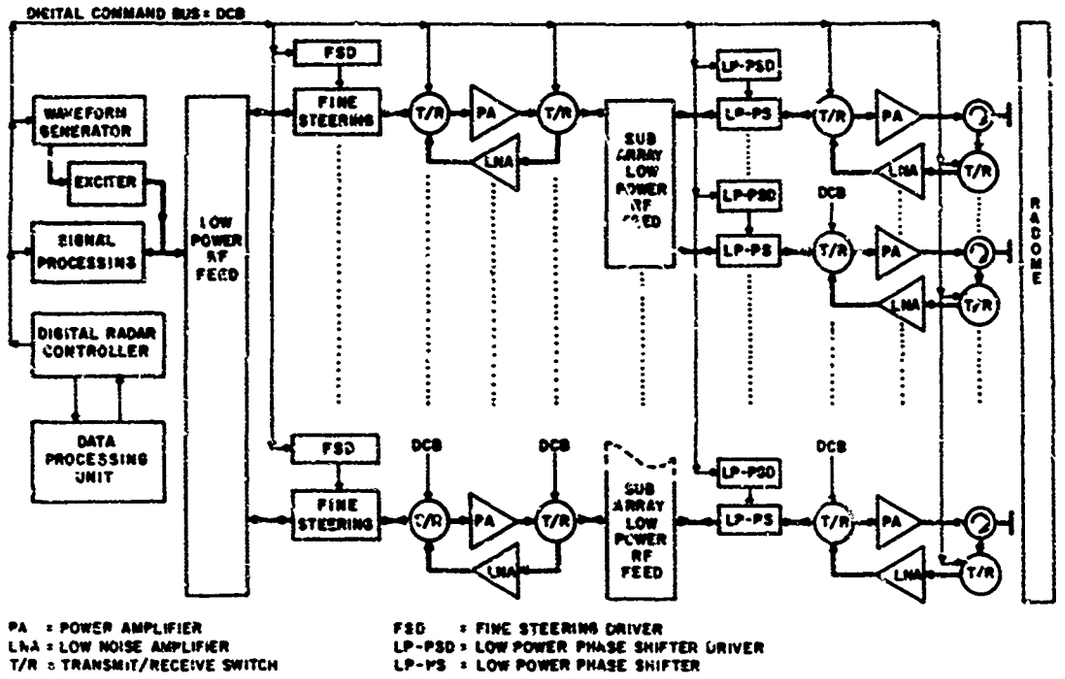


Fig.9 Typical solid state microwave integrated circuits and devices



PA = POWER AMPLIFIER
LNA = LOW NOISE AMPLIFIER
T/R = TRANSMIT/RECEIVE SWITCH
FSD = FINE STEERING DRIVER
LP-PSD = LOW POWER PHASE SHIFTER DRIVER
LP-PS = LOW POWER PHASE SHIFTER

Fig.10 Simplified block diagram - solid state or tube per element phased array radar

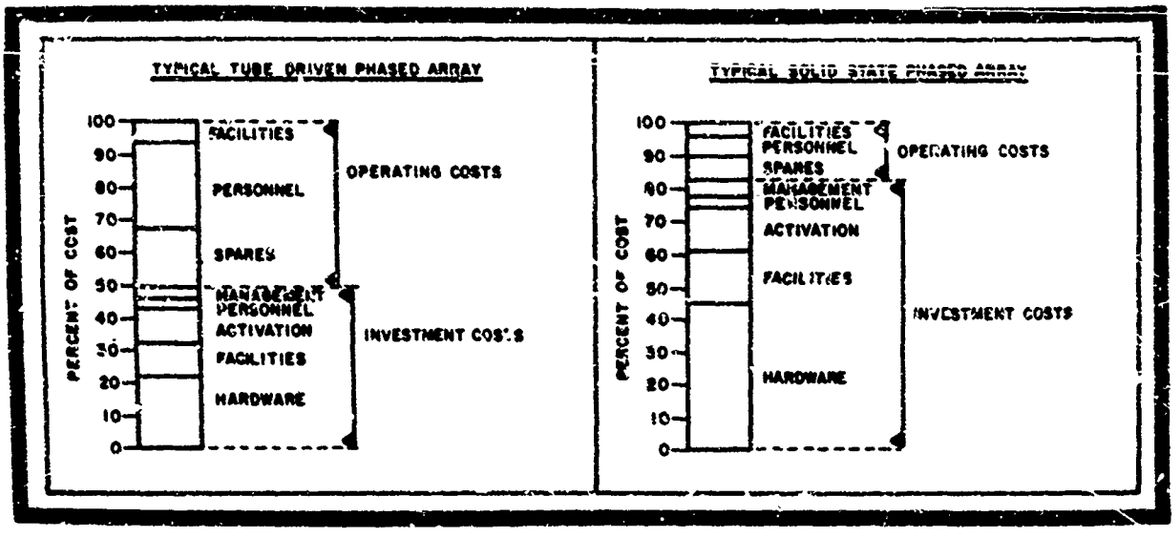


Fig.11 Typical distribution of costs - solid state and tube-driven phased arrays

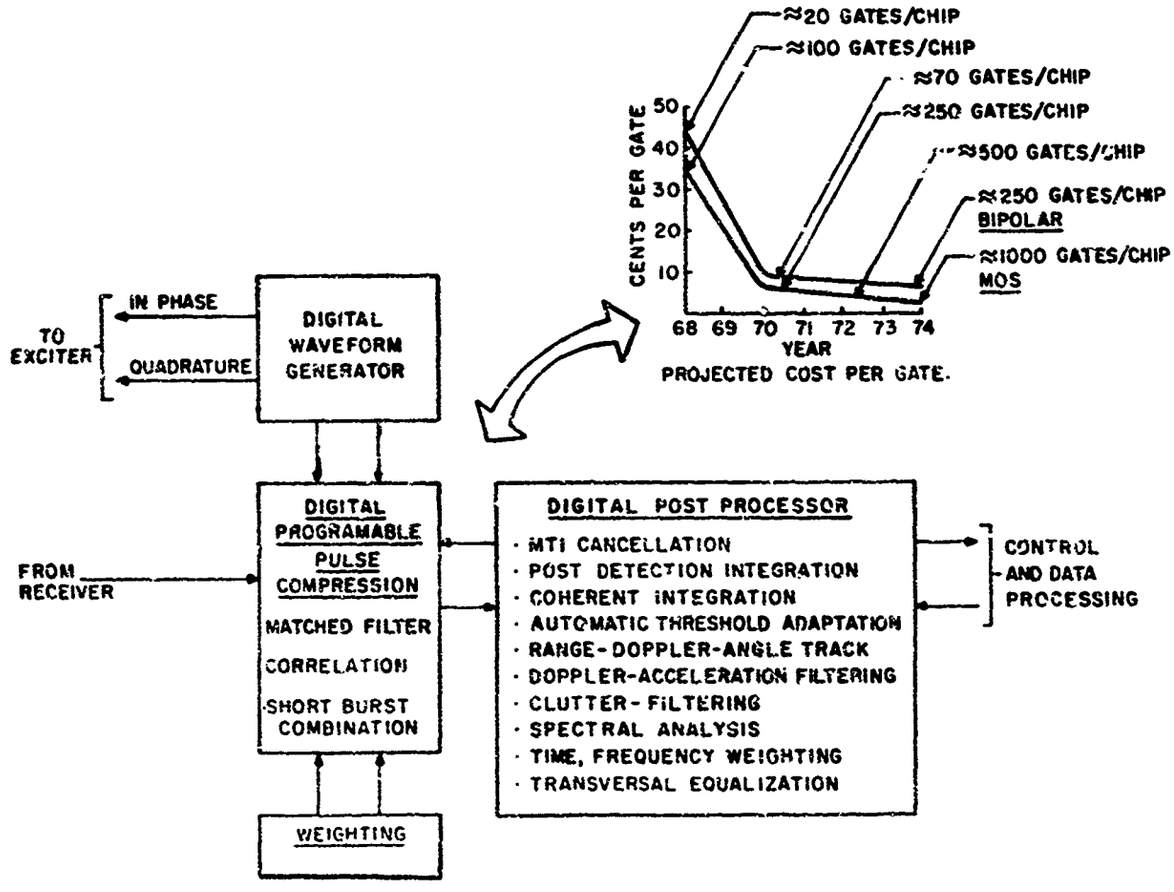


Fig.12 Digital signal processing functions

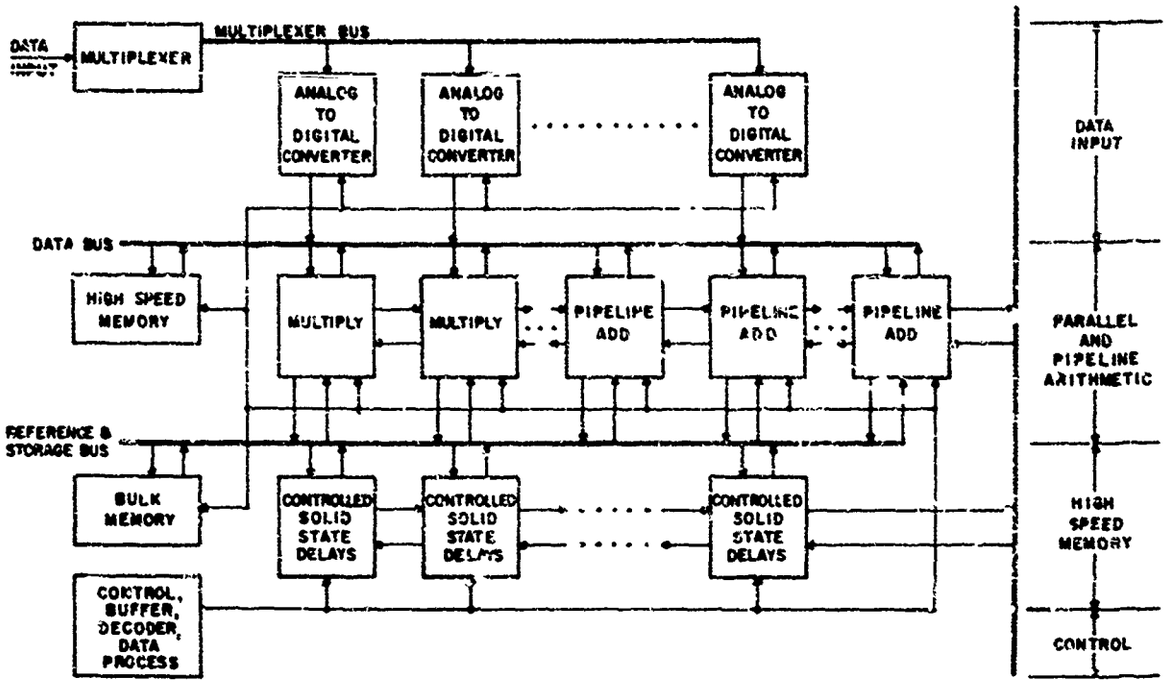


Fig. 13 Parallel, pipelined, digital signal processor

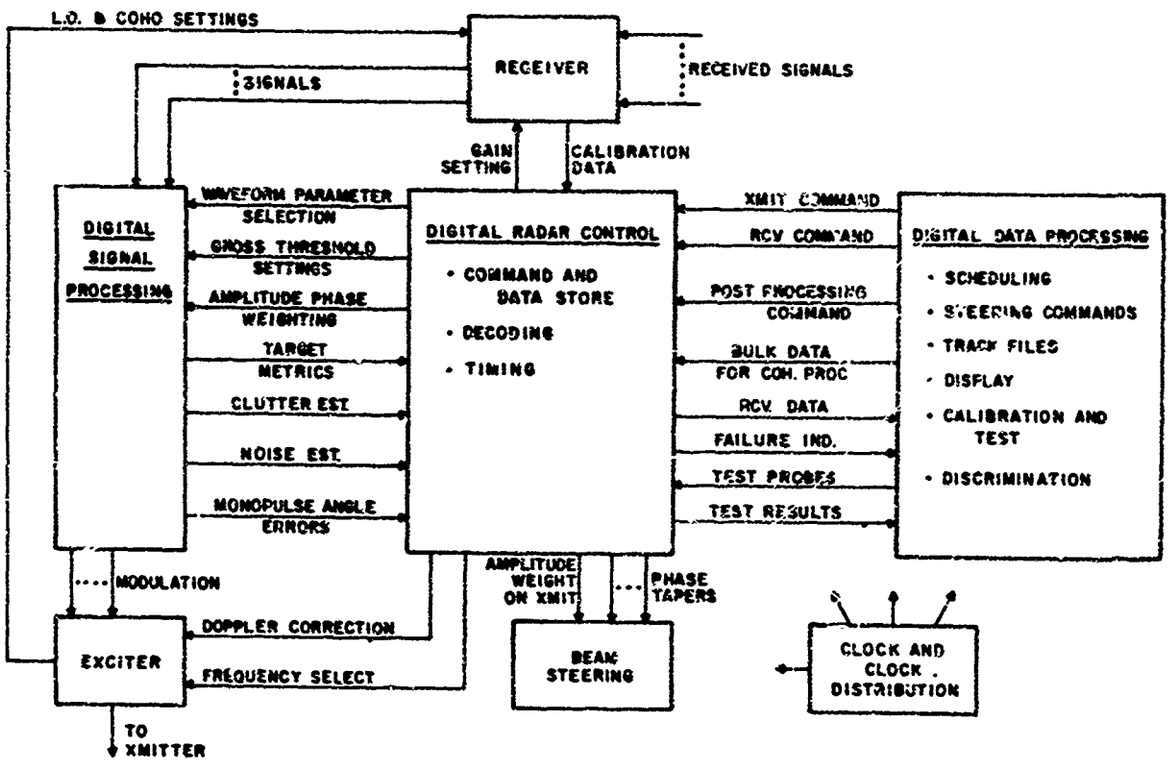


Fig. 14 Inter-relationship between digital subsystems

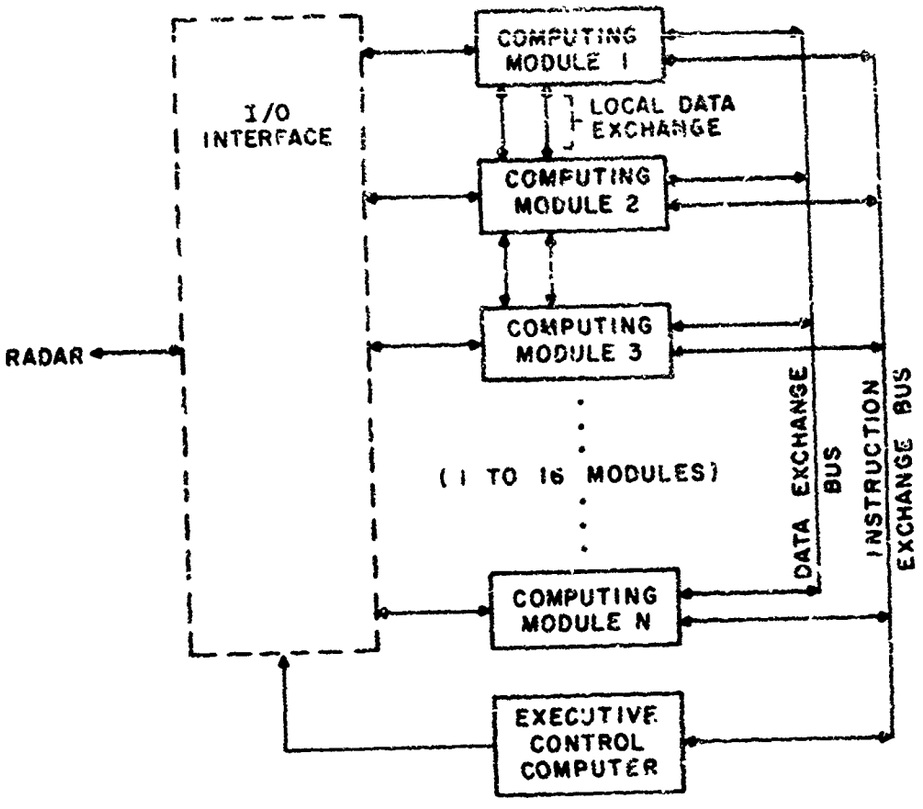


Fig.15 Modular array data processor

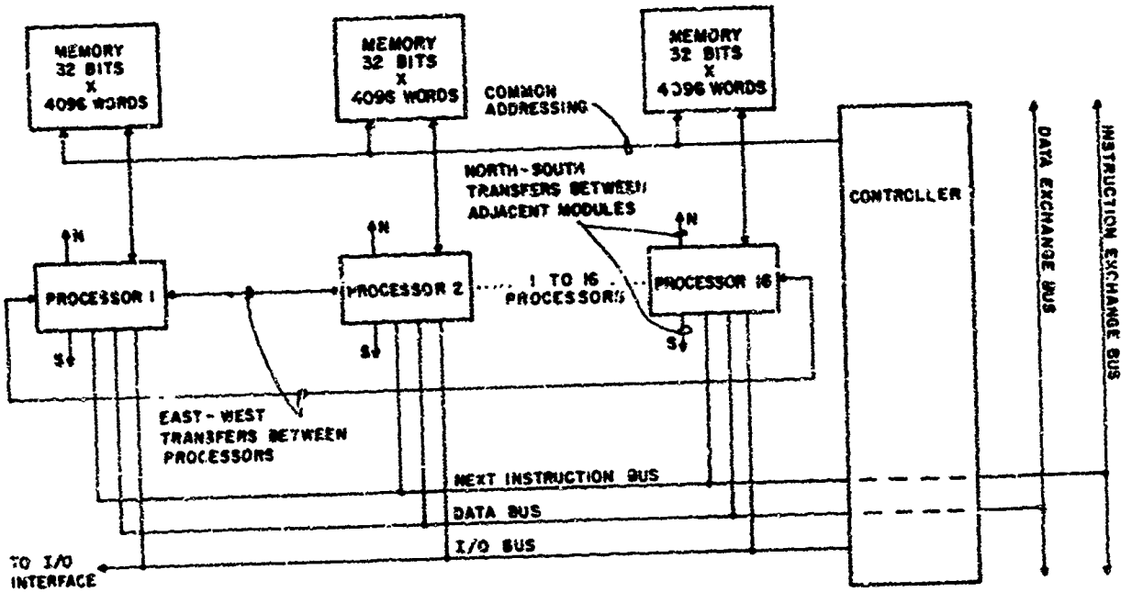


Fig.16 Typical computing module

THE LINEAR ARRAY FOR BEAM STEERING

by

R.Reitzig

SIEMENS AG
81, Hofmannstrasse, 8 München 25, Germany

THE LINEAR ARRAY FOR BEAM STEERING

R. Reitzig
SIEMENS AG
51, Hofmannstraße, 8 München 25, Germany

SUMMARY

The radiation performance of a phased 8-dipole array with uniform and Dolph-Tchebycheff excitation and a 16-dipole array with a cosine aperture distribution is theoretically studied on models which exclude the element coupling. The computed results are then compared with measurements thus giving some information on the deteriorating effects of mutual coupling. The paper deals also with the influence of systematic phase-errors as they are usually obtained by digitally approximating a linear phase progression along the array aperture.

1. INTRODUCTION.

The paper presents some theoretical and experimental results obtained on the study of linear phased array antennas with a uniform, cosine and Dolph-Tchebycheff aperture excitation, respectively. The study was initiated to acquire some first experience in the radiation and impedance performance of an electronically scanned array.

To prevent the appearance of grating lobes in wide-angle scanning applications, the radiating elements must generally be small in size -- approximately a half-wavelength in cross-section -- and closely spaced to each other, again in the order of a half-wavelength. The elementary radiators have a range of impedance, pattern and polarization characteristics. However, due to the close element spacing energy is parasitically transferred from the neighbouring elements thus causing the radiation pattern and the impedance characteristics of the element in the array environment to differ markedly from that of the isolated element. In addition, the mutual coupling changes as the array antenna is scanned through different angles. This change implies that the antenna radiation efficiency varies as the scan-angle changes. Hence, the influence of mutual coupling has to be taken into account if a true picture of array performance is to be obtained. In general, a theoretical array model greatly simplifies the problem by neglecting the mutual coupling. The electrical characteristics of the elements are assumed to be constant during the scanning period and set equal to the isolated element, independent of its location within the array configuration. The computed results thus obtained will certainly differ from the experimental results thus giving some account on the degrading influence of coupling on the radiation pattern.

Apart from the radiation characteristics of the coupled array elements the electrical properties of the feeding network of the array antenna have also an essential influence on the overall performance of the antenna system. Due to the varying mismatch of each formerly matched isolated element when put in a phase-scanned array, fractions of energy are reflected back toward the feed structure, thus altering the phase and amplitude distribution among the antenna elements. In turn, the radiating properties of the coupled antennas will be affected, unless provisions are made to eliminate the element coupling through the array feed system.

In an array of a small number of elements the laws of a statistical distribution and, hence, cancellation of amplitude and phase errors of the aperture distribution do not apply. As pointed out, such errors are due to mutual coupling. Phase errors, however, are also introduced due to a phase quantization by digitally altering the phase-progression along the aperture. Using n-bit digital phase-shifters the phase-progression along the array aperture is correct only for 2^n+1 scanning directions. In order to increase the number of beam directions the desired phase function may then be approximated as closely as possible with the number and fineness of steps that are available. It is, therefore, of vital interest for an antenna designer to conclude from the kind and the degree of pattern distortion whether amplitude or phase errors of the array excitation are prevailing. Countermeasures have, then, to be developed to minimize these errors.

Answers to these problems of mutual coupling and array excitation are sought in a theoretical and experimental analysis of an eight-element linear array of uniform and Dolph-Tchebycheff aperture distribution, fed by a corporate microwave network. The radiating elements were dipoles printed on a low-loss dielectric material. The phase-shifting was performed by coaxial trombones. A linear array of sixteen dipole elements was also analyzed and built. However, a modified space-feed system was used to energize the radiators. Ferrite phase shifters were applied to alter the phase progression along the aperture.

2. THE 8-ELEMENT LINEAR PHASED ARRAY WITH UNIFORM AMPLITUDE EXCITATION.

Provided coupling is neglected the amplitude pattern of a linear array of n directive radiating elements is given by

$$C(\vartheta) = F(\vartheta) \sum_{\nu=1}^{n} |A_\nu| \; e^{j(\nu-1)(\varphi_0 + \frac{2\pi}{\lambda} d \sin\vartheta)} \qquad (1)$$

where $F(\vartheta)$ describes the pattern of the elementary radiator of the array. A_ν is the amplitude of the excitation coefficient, d the spacing between radiating elements and λ the wavelength in free-space. ϑ specifies the angular position of observation. φ_0 is the phase difference between adjacent equispaced radiators in an array with linear phase progression along the aperture. By altering the phase delay the main beam is deflected. The variation of the pattern with the scan angle is studied.

Although the measured module of the coupling coefficient between two dipoles a half-wavelength apart is about 10 dB in the E-plane of the radiated field vector and 14 dB in the H-plane, it was preferred for the experimental verification of the computations to line up the dipoles in the E-plane rather than in the H-plane in order to obtain more striking results due to the higher amount of coupling. In addition, the radiation pattern of a dipole is about twice as much directional in the E-plane than in the H-plane. Thus, any beam distortion, side-lobe level and beam-pointing error due to the higher directivity of the dipole pattern and the higher element coupling in the E-plane will be more easily demonstrated. The elementary pattern is well defined by

$$F(\vartheta) = \frac{\cos(\frac{\pi}{2}\sin\vartheta)}{\cos\vartheta} \qquad (2)$$

as plotted in Fig. 1 along with the experimental pattern.

For n=8 dipoles each $d=\frac{\lambda}{2}$ apart and uniform amplitude excitation the broadside radiation pattern was computed, Fig. 1. This pattern was experimentally well verified. The half-power beamwidth is 12.5°, the side-lobe is about 13 dB below the maximum.

The comparison between theory and experiment was made for various scan-angles. The experimental array was fed by a corporate feed structure. In a first attempt the power splitting was performed by means of a reactive two-way power divider. However, such a three-port junction cannot be matched simultaneously when looking into all arms. Thus, due to the mismatch of the coupled radiators in a phased array antenna multiple reflections arose, which in turn deteriorated the formerly uniform amplitude excitation and the linearly progressive phase distribution. The agreement between theory and experiment was not too good. An improvement, especially of the side-lobe level, was found with a corporate feed using strip-line hybrids with an isolation of more than 40 dB to perform the power division. In Fig. 2 the deflected beams at $\vartheta = 29^\circ$ and $\vartheta = 52^\circ$ corresponding to $\varphi_0 = 60^\circ$ and $\varphi_0 = 150^\circ$, respectively, are plotted. The half-power beamwidth increases up to about 19°. The radiation pattern becomes very asymmetrical. The side-lobe level increases considerably relative to the peak-level of the deflected beam. The gain drops by about 6 dB. Due to the directivity of the radiating elements the beam maximum is shifted with respect to that of an array of isotropic radiators. For $\varphi_0 = 150^\circ$ the angular displacement amounts to 21 % with respect to the half-power beamwidth.

Although mutual coupling is present, the uniformly excited array shows a good agreement between the computed and the measured performance, apart from the far-off scan angle due to the difference between the theoretical and experimental element pattern. Probably amplitude and phase errors induced by coupling do not seem to have a too strongly deteriorating effect. In fact, calculations have shown that small random phase and amplitude errors are of minor importance. More serious influence on the radiation performance of the phased array have systematic phase errors. For the theoretical analysis a systematic phase distortion of the kind

$$\varphi(\nu) = (\nu-1)\varphi_0 - \Delta\varphi \sin\,(\nu-1)\,\frac{\pi}{2} \qquad (3)$$

was introduced in eq. 1, where $\Delta\varphi$ defines the maximum phase error. This phase variation describes an alternating zigzag-curve around the linear phase progression φ_0. As an example, in Fig. 3 the undisturbed pattern of $\varphi_0 = 90^\circ$ is compared with the corresponding patterns for $\Delta\varphi = 10^\circ$ and $\Delta\varphi = 20^\circ$, respectively. The deteriorating effect of a systematic disturbance of the phase progression is chiefly restricted to the side-lobe level, especially in the angular region opposite to the scanning direction. For small phase errors as $\Delta\varphi = 10^\circ$ the decrease of the gain is negligible. For a phase error of $\Delta\varphi = 20^\circ$ a gain drop of about 0.3 dB only was found. This behaviour was experimentally well verified.

In general, a linear array with uniform aperture excitation has very little practical application as the side-lobe level of about 13 dB is too high. An array with a tapered amplitude excitation of the aperture has better chances. As an example an 8-element array with a Dolph-Tchebycheff pattern was chosen for the analysis.

3. THE 8-ELEMENT LINEAR ARRAY WITH DOLPH-TCHEBYCHEFF RADIATION PATTERN.

It was expected that for an aperture distribution whose properties are optimum in the sense that it will produce the narrowest beamwidth for a given side-lobe level and vice versa, the radiation properties of the linear array are more liable to be degraded by a variation of the amplitude excitation due to mutual coupling or due to a misalignment of the power distributing network. In fact, a random amplitude error of the aperture excitation up to 1 dB did not cause any essential pattern deterioration as shown in Fig. 4. The pattern was computed and measured on an 8-dipole linear array with a half-wavelength element spacing and 26 dB side-lobe attenuation. A severe pattern deterioration, however, is introduced by phase errors of the aperture excitation. Before the radiation pattern in Fig. 4 was experimentally obtained, a pattern with shoulders on either side of the main beam was measured, Fig. 5. This principal defect is due to a quadratic phase error induced by coupling. An evaluation of the magnitude of the quadratic phase error, however, did not improve the radiation pattern. Again, mutual coupling deteriorated this phase compensation. By a trial and error method a deliberately induced linear phase variation of about 25° symmetrically to the array normal was found to compensate for mutual coupling effects on the phase progression along the linear dipole array aperture.

This phase compensation had not to be altered within a scanning section of about ± 40°. Hence, it may be concluded that the phase error introduced by a varying mismatch during a scanning period due to mutual coupling is of small order. In Fig. 6 the computed and measured patterns for two phase progressions $\varphi_0 = 90^\circ$ and $\varphi_0 = 120^\circ$, respectively, are plotted. The scanning angle and the side-lobe level agree well, however, the power level of the computed diagram is increasingly lower than the measured one. In the theoretical analysis the elementary pattern $F(\vartheta')$ has been assumed to be identical for all elements. However, as measurements have shown it differs considerably dependent on its position along the array, Fig. 7. The isolated dipole has a higher directivity than the dipole centrally positioned in an 8-element array. The radiation pattern of the edge element is completely distorted, especially the side-lobe level is strongly asymmetrical.

A systematic phase error of the zigzag form described in the previous section results in a severe pattern deterioration. For reasons of comparison with the uniform illumination the same systematic phase variation of the aperture excitation was introduced. Again, for $\varphi_0 = 90^\circ$ the pattern is plotted in Fig. 8. The gain variation is of about the same order very small as for uniform illumination, however, the side-lobe level is most severely asymmetrically distorted, even by small phase increments. These results agreed also very well with the measurements.

4. THE 16-ELEMENT LINEAR ARRAY WITH A COSINE AMPLITUDE EXCITATION.

For a linear array of small element number, say n = 8, a corporate type of feed structure is to be preferred. The two-way power division network, however, becomes very complex for a higher element number. An alternative solution is the application of the space-feed principle. In order to efficiently illuminate the collector side of the antenna, the space-feed had to be modified such that the energy leaving the primary feed propagates in a sectoral-horn closed environment. However, such an oversized waveguide system is very sensitive to the excitation of high-order modes by any kind of asymmetries like the varying input impedance of the elementary radiator during a scanning period. The danger of high-order modes is an increase of the coupling to even far-off positioned array elements. In addition, at certain frequencies and scan-angles the radiation performance of an array fed by such a structure will be drastically altered due to trapped-mode resonances. By the insertion of mode-absorbers which have a least effect on the dominant waveguide mode the amount of high-order modes might be greatly reduced. The amplitude excitation of such a feed system for a 16-element linear phased array with all the output terminals of the emitter side of the antenna system terminated by a matched load was measured, Fig. 9. It is fairly well approximated by a cosine type of aperture distribution. In contrast to the two previous array antennas described where the theoretical analysis was based on ideal models of the amplitude excitation and of the elementary pattern, now the measured amplitude distribution of Fig. 9 and the measured H-plane radiation pattern of the isolated dipole (Fig. 10) were fed into a computer and the radiation pattern of the array antenna evaluated for various scan angles and systematic phase errors, using eq. 1. The calculated broadside pattern of a 16-dipole linear array with an element spacing of λ is presented in Fig. 10 and compared with the measured one. The agreement of the main-lobe pattern is good, however, even after the compensation of an initial quadratic phase error due to mutual coupling it was not possible to lower the measured side-lobe level to the theoretical value. Due to the lack of a high isolation between the collector elements in a closed-environment overmoded waveguide feed structure mutual

coupling greatly affects the power rating, especially of the edge elements. In fact, calculations have proved the increased measured side-lobe level being essentially due to an edge illumination of about 10 dB, a value of about the same order as measured without the mode-absorber in the horn.

For a phase progression of $\varphi_0 = 90^0$ and $\varphi_0 = 112.5^0$ corresponding to $\vartheta = 30^0$ and $\vartheta = 39^0$, respectively, the deflected beams were computed and plotted in Fig. 11. The comparison with the experimental patterns discloses as a most striking novel feature an angular displacement between the measured and calculated main-lobe pattern. This performance may be explained by a variation of the "effective" element pattern and/or of the linear phase progression along the array aperture due to mutual coupling during the scanning period. The beamwidth and the gain agree well with the computations. However, the side-lobe level varies considerably, the degree and character of variation being entirely different for two beam directions only a few degrees apart. The increase of the power-level of the side-lobes can be contributed to phase errors which seem to have some systematic character. In Fig. 12 the deteriorating effect of a systematic phase error of the type defined by eq. 3 with $\Delta\varphi = 11.5^0$ is well demonstrated on the computed pattern for $\varphi_0 = 56.25^0$.

5. CONCLUSION.

An agreement between the experimental and computed radiation pattern of a linear phased array of small element number was found for a tapered aperture excitation only after the compensation of a quadratic phase error. This phase error is introduced to the linear phase progression along the array aperture by mutual coupling between the radiating elements. An array feeding structure with a high isolation between the output terminals is to be used to eliminate any additional coupling through the feed. A systematic phase error of the linearly progressive phase excitation of the array aperture has a severely deteriorating effect on the side-lobe level.

6. ACKNOWLEDGEMENT.

Grateful acknowledgement is made to Mrs. Degler, Mr. Hoermann and Mr. Dieges for their valuable contribution and helpful cooperation in preparing this paper.

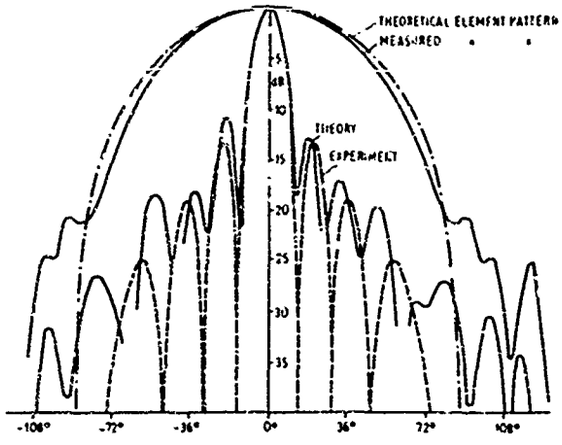


Fig.1 Pattern of 8-element dipole array with uniform excitation

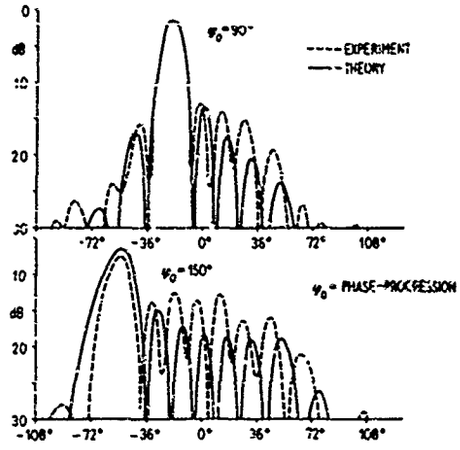


Fig.2 Uniform excitation of 8-dipole array

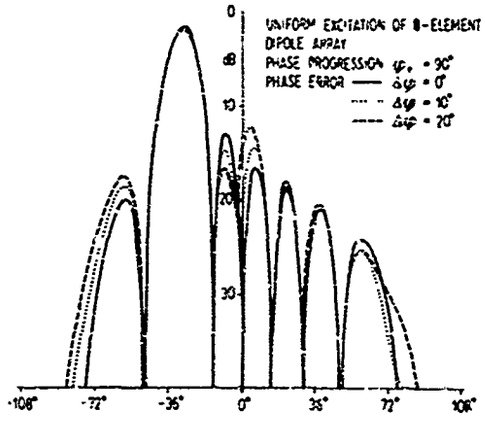


Fig.3 Influence of systematic phase errors

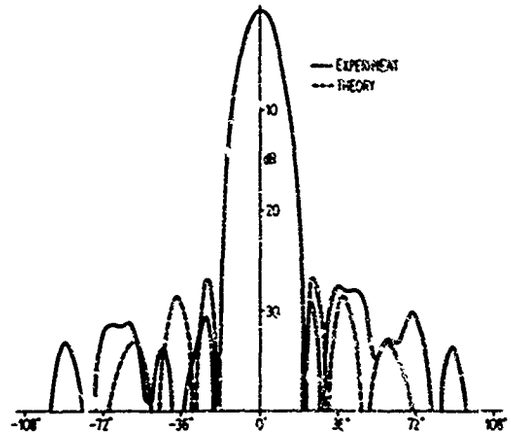


Fig.4 Dolph-Tchebycheff pattern of 8-element dipole array

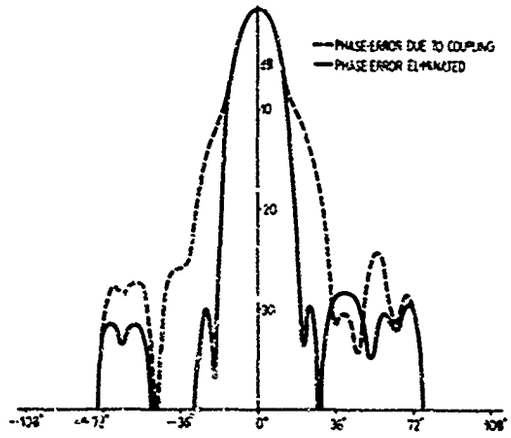


Fig.5 Dolph-Tchebycheff pattern of 8-element dipole array

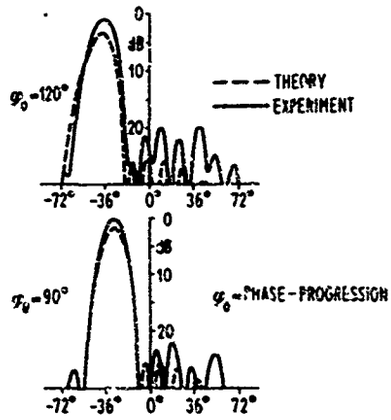


Fig.6 Dolph-Tchebycheff excitation of 8-dipole array

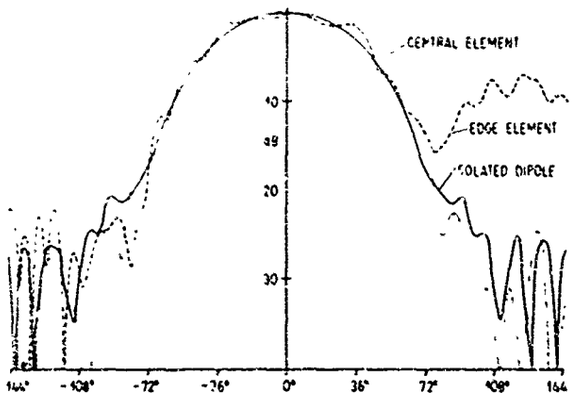


Fig.7 Influence of coupling on dipole pattern

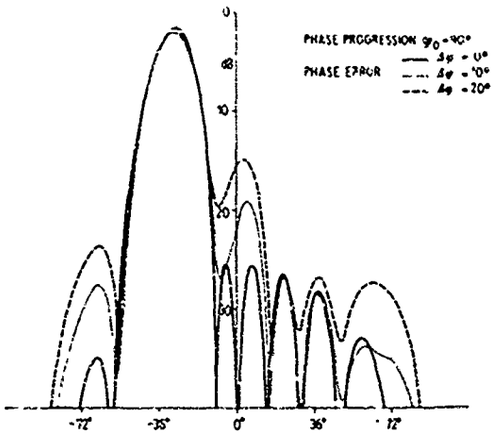


Fig.8 stematic phase errors on Dolph-Tchebycheff pattern

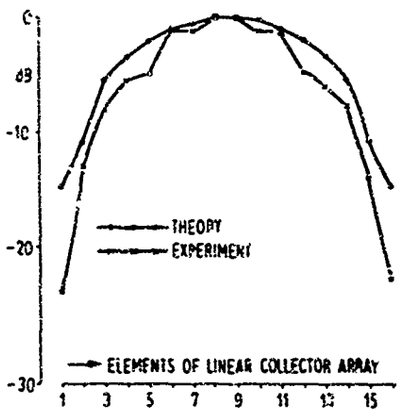


Fig.9 Amplitude excitation of collector array

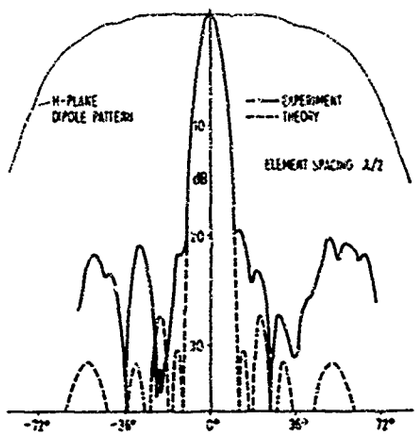


Fig.10 Cosine excitation of 16-dipole array

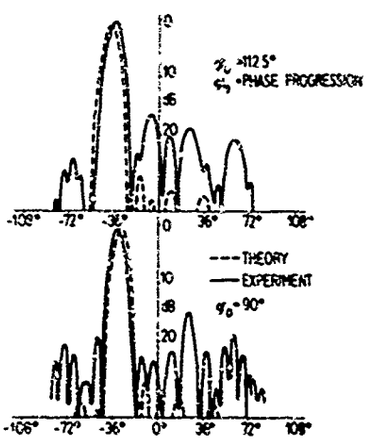


Fig.11 Cosine excitation of 16-dipole array

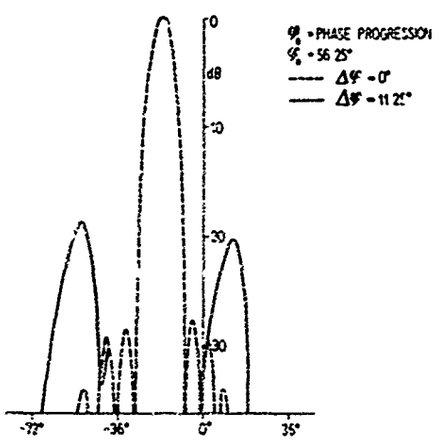


Fig.12 Influence of systematic phase error

MICROWAVE SOLID STATE POWER SOURCES FOR AIRBORNE PHASED ARRAYS

by

Fred Sterzer

RCA Electronic Components
Princeton, New Jersey

MICROWAVE SOLID STATE POWER SOURCES FOR AIRBORNE PHASED ARRAYS

Fred Sterzer
RCA Electronic Components
Princeton, New Jersey

SUMMARY

Several cw solid-state power modules for airborne phased array applications are described: (1) Hybrid integrated L-band modules using transistor amplifiers ($P_{max} \sim 30$ watts). (2) Hybrid integrated S-band modules using transistor amplifiers followed by varactor doublers ($P_{max} \sim 15$ watts). (3) Coaxial C-band modules using circulator-coupled negative resistance transferred electron amplifiers ($P_{max} \sim 1$ watt). (4) Coaxial X-band modules using circulator coupled negative resistance transferred electron and Impatt amplifiers ($P_{max} \sim 0.5$ watt).

1. INTRODUCTION

The amount of microwave power that can be generated with solid state devices is now sufficiently high so that practical solid state power modules can be built that are suitable for a variety of airborne microwave phased array applications. This paper will describe several such sources currently being developed at RCA's Advanced Technology Laboratory in Princeton, New Jersey.

2. TRANSISTOR AMPLIFIERS

High-power L-band modules using microstrip transistor amplifiers are being developed. These modules are much smaller than comparable coaxial modules, yet their electrical performance is similar to the best that can be achieved with coaxial modules.

The transistor chips used in these modules are mounted in carriers that are specifically designed for high-power microstrip circuits. These carriers provide excellent heat conduction from the chip to the metal ground plane of the microstrip, they act as impedance transformers between the chip and the microstrip circuit, and finally, they are so small that several carriers can be placed side-by-side for paralleling of several transistors in one circuit. An example of an amplifier using three transistors in parallel is shown in Fig. 1. In multi-stage modules, the following performance has been obtained at L-band frequencies: Gain = 23 dB, 1 dB bandwidth = 10%, CW power output = 30 watts, efficiency = 40%. We expect that experimental quantities of transistors suitable for multiwatt power levels at S-band frequencies will become available by the end of 1970.

3. TRANSISTOR DRIVEN VARACTOR DOUBLERS

The L-band hybrid integrated modules described in the previous section are being used to drive hybrid integrated varactor doublers to produce cw power outputs of 15 watts at S-band frequencies. A module of this type is described in Fig. 2.

4. TRANSFERRED ELECTRON AND IMPATT AMPLIFIERS

For frequencies above S-band, transferred electron and impatt amplifiers are currently the only solid state amplifiers that can produce cw powers in the watt range. The construction of a typical transferred electron device that can produce approximately 1 watt cw at C- and X-band frequencies is shown in Fig. 3.

Most transferred electron and impatt amplifiers are circulator-coupled negative-resistance amplifiers. A schematic diagram of a transferred electron amplifier (TEA) of this type is shown in Fig. 4.

Coaxial modules using transferred electron and impatt devices are being developed for C- and X-band frequencies. Some preliminary results are as follows:

(a) Single-stage C-band TEA: Gain = 8 dB, bandwidth = 3 GHz, CW power output (1 dB gain compression) = 250 mW, CW power output (5 dB gain compression) = 1 watt.

(b) Two-stage C-band TEA: Gain = 22 dB, bandwidth = 2 GHz, CW output power (1 dB gain compression) = 250 mW.

(c) X-band TEA: Gain = 5 dB, bandwidth = 2 GHz, CW power output (2 dB gain compression) = 500 mW. (See Fig. 5)

(d) Single-stage X-band impatt amplifier: Gain = 5 dB, bandwidth = 1 GHz, CW output power = 150 mW.

The efficiencies of these amplifiers is currently only a few percent. However, it seems likely that by the end of 1970, average power outputs exceeding 1 watt will be achieved with efficiencies exceeding 20%.

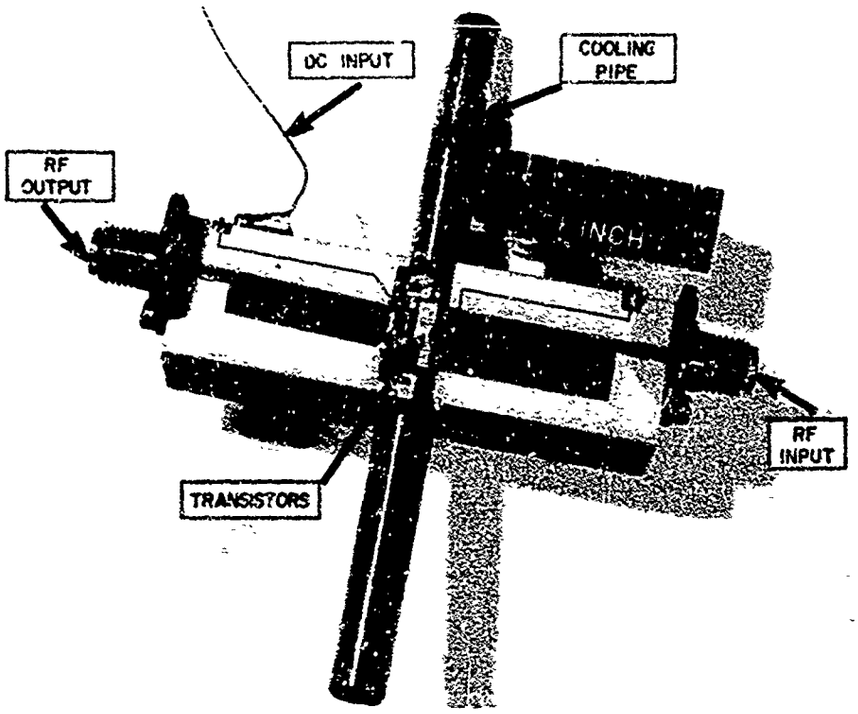


Figure 1 - Photograph of a hybrid integrated L-band transistor amplifier. This amplifier has a gain of 8 dB, a bandwidth of 200 MHz, a cw power output of 22 watts, and an efficiency ranging from 40 to 48%. (Courtesy Dr. E. T. Belohoubek and Dr. D. M. Stevenson, RCA Electronic Components, Princeton, New Jersey.)

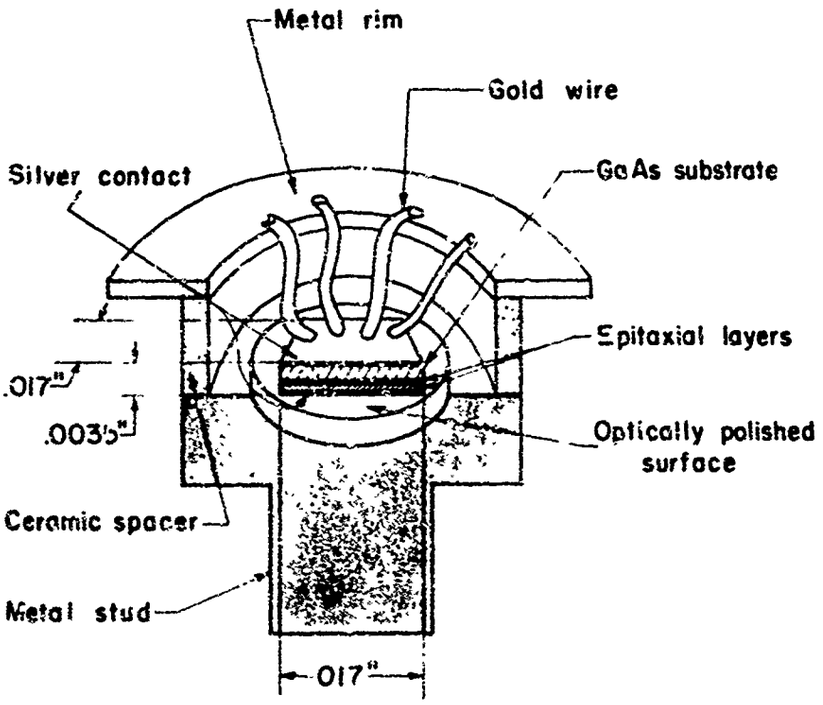


Figure 3 - Schematic drawing of a typical transferred electron device mounted in a metal-ceramic package. Devices of the type shown in the drawing can produce power outputs of the order of 1 watt cw at C- and X-band frequencies. (Courtesy Dr. S. Y. Narayan and B. S. Perlman, RCA Electronic Components, Princeton, New Jersey.)

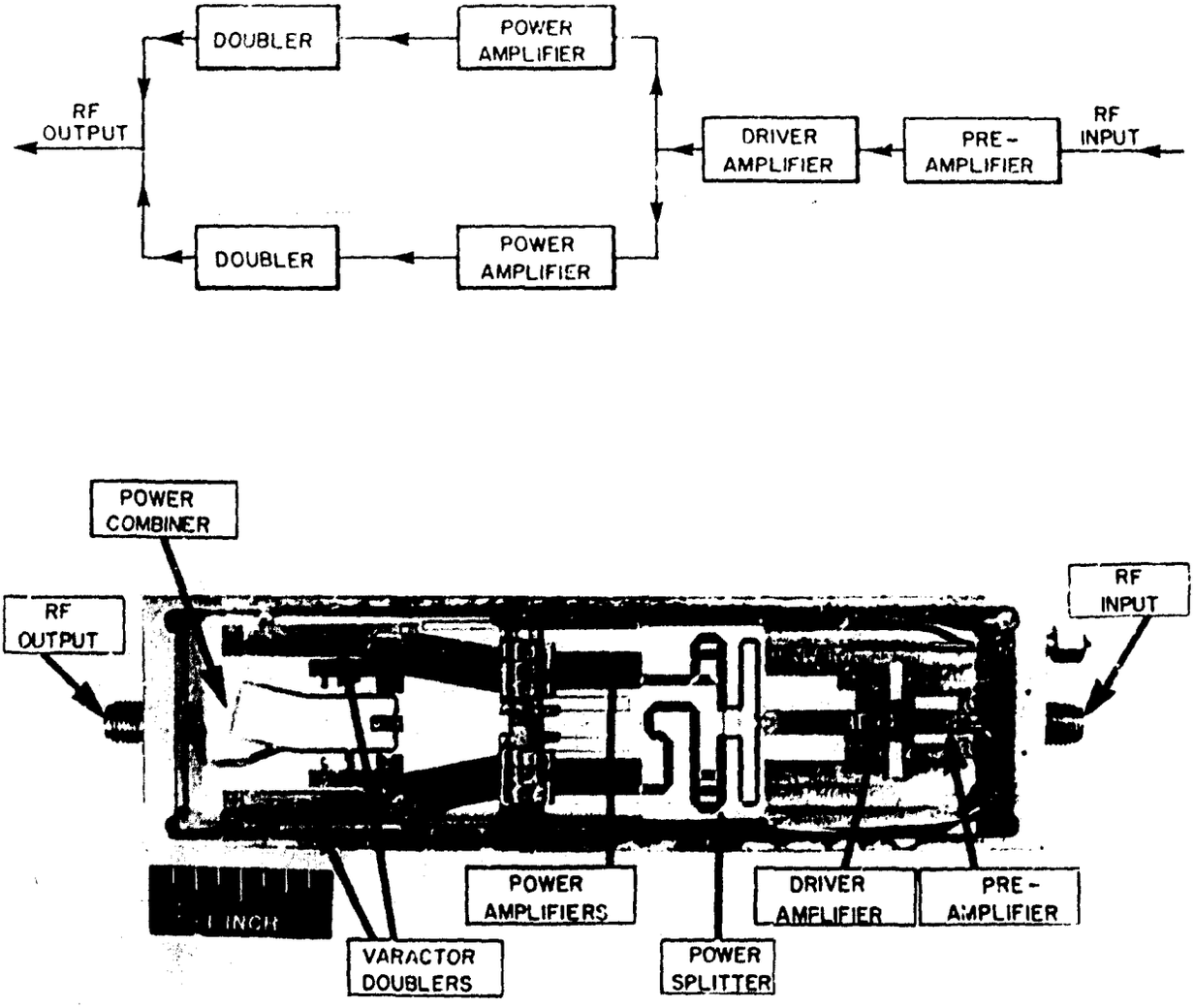


Figure 2 - Hybrid integrated S-band transistor amplifier-varactor doubler power module. With a power input of 100 mW at L-band frequencies, this module produces a power output of 15 watts cw at S-band frequencies. The bandwidth of the module is 12%, and its efficiency ranges from 11 to 14%. (a) Block diagram of module. (b) Photograph of module. (Courtesy of Dr. E. J. Belohoubek, A. Presser, and Dr. D. M. Stevenson, RCA Electronic Components, Princeton, New Jersey.)

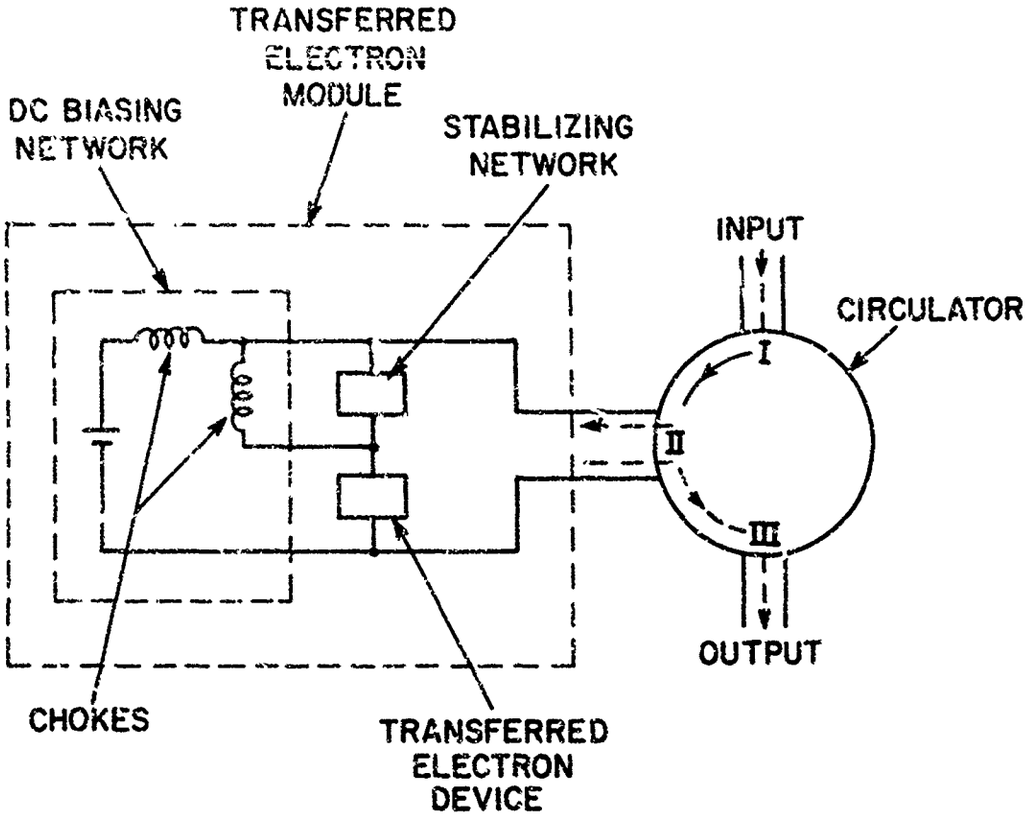


Figure 4 - Schematic diagram of a circulator-coupled negative resistance transferred electron amplifier. The signal to be amplified enters the circulator at port I, leaves at port II, is amplified by the negative resistance of the transferred electron module, is reflected back to the circulator, and leaves the circulator at port III.

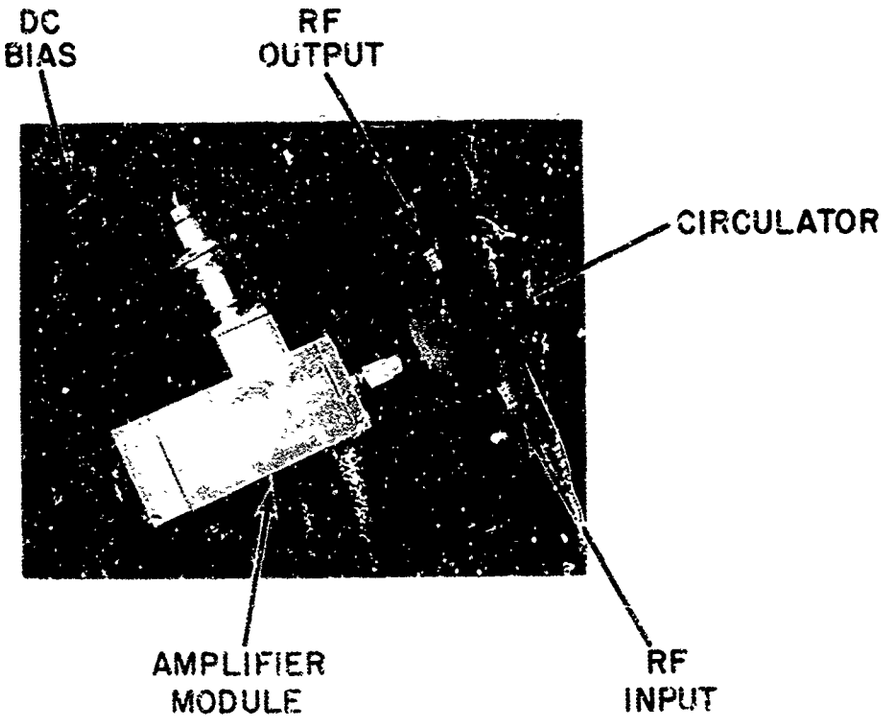


Figure 5 - Photograph of an X-band transferred electron module. This module has a gain of 3 dB, a cw power output of 500 mW, and a bandwidth of 2 GHz. (Courtesy B. S. Perlman, RCA Electronic Components, Princeton, New Jersey.)

LIGHTWEIGHT PHASED ARRAY SYSTEMS

FOR AIRBORNE AND GROUND BASED APPLICATIONS

by

M. Simpson

SEDCO SYSTEMS INC.
Farmingdale, New York, USA.

LIGHTWEIGHT PHASED ARRAY SYSTEMS FOR AIRBORNE AND GROUND BASED APPLICATIONS

M. Simpson
SEDCO SYSTEMS INC.
Farmingdale, New York, U.S.A.

SUMMARY

This paper discusses the design and performance capabilities of several different types of Lightweight Phased Array Antenna Systems which may have application in a number of areas such as forward area battlefield surveillance radar, air traffic control radar, lightweight mortar locator systems, airborne weather, terrain clearance and terrain avoidance radars, among others. The systems to be discussed fall within the microwave frequency band of 1 to 18 GHz. Specific design and component selection parameters are discussed for linearly scanned phased array antenna systems optimized for operation in various portions of the overall frequency band. Antenna models in the S and X frequency bands were fabricated and tested. Results indicated very close correlation with theoretical predictions and were in accordance with specification requirements. A discussion is presented of the critical components used in these antenna systems such as phase shifters (ferrite, nonreciprocal and reciprocal, and diode), and antenna feed systems. The allowable errors for these critical components are analyzed. It is shown that error tolerances are considerably more severe than those which may be accepted on large phased array systems. It is also shown that antenna performance for small phased arrays can be substantially superior to that obtained with conventional dish type mechanical antenna systems. A description is given of the electronic beam control and processing system used with these antennas in order to provide scan and track capability.

1. INTRODUCTION.

A program was initiated by Sedco Systems Inc. in May 1964 under contract with the U. S. Army Electronics Command at Fort Monmouth, New Jersey to investigate the design and fabrication of lightweight, minimal complexity phased array antenna systems for use as part of forward area surveillance radar systems. A number of radar systems in this category are presently in use, operating in various parts of the microwave frequency band from L-band through K-band. It was therefore determined that a design study would be made in order to establish the optimum designs for various portions of the overall frequency band. Following the completion of the study, antenna models were to be fabricated at X-band and S-band in order to demonstrate actual performance. The objective specifications which were to be used as a guideline in the proposed designs were as follows:

a) Electronic Scan range -	± 60 degrees in azimuth, fixed in elevation.
b) Beamwidth -	Azimuth - 1 to 5 degrees Elevation - 5 to 20 degrees.
c) Sidelobe Level -	20 db or less in both planes.
d) Frequency Bandwidth -	10% tunable, 10 MHz instantaneous.
e) Power Level -	100 Kw peak, 100 watts average.
f) Beam Pointing Accuracy -	1/30 of a beamwidth.
g) Automatic Tracking Accuracy -	1/20 of a beamwidth.
h) Weight -	20 pounds objective.
i) Scan Rate and Power -	0 to 15 cps; 10 watts.

The phased array antenna systems described below represent some of the specific results of this program. These systems have been fully tested and delivered. They have been found to be generally in accordance with the objective specifications. The weight of the experimental model antenna systems has been in the range of 60 pounds. Approximately one-third of this weight may be eliminated as a result of productization and the use of lighter weight materials. The other major specifications have been achieved.

2. PRINCIPLE OF DESIGN

The basic design of the phased array antenna systems developed in this program involves the use of a parallel fed phase shifter scanned array operating at essentially a fixed (but tunable) RF frequency. A schematic representation of this type of system is shown in Figure 1. This type of design was chosen over the so-called frequency scanned system since it may be readily adapted to conventional fixed frequency radar receivers and transmitters and provides maximum efficiency and minimum use of the already crowded RF spectrum.

A quick examination of the diagram given in Figure 1 indicates that the antenna system contains several major elements, namely, the antenna feed, the electronic phase shifter group, the radiator group, and the electronic control system. The latter element is actually a composite of several subsystems such as the beam steering control system, the programmer and logic system, the operator interface providing such functions as antenna scan sector control, scan rate control, frequency selection, and mode control, i.e. manual, scan or track. In addition, the electronic control system may include an indicator providing beam position in either CRT or digital output format.

One of the most important considerations in the design is the efficiency of the system. For a given aperture efficiency, determined primarily by the required near in sidelobe level, the overall efficiency will then be determined by the additional losses in the system between the transmitter output connection and the antenna radiators. All of the RF elements in the antenna contribute to this loss. However, the one element which frequently provides the

greatest contribution to reduction in antenna efficiency is the electronic phase shifter. In addition, this element also dissipates power during the course of antenna scanning. Furthermore, it is used in relatively large quantity, generally equal to the number of radiators, and contributes substantially to the overall cost of the antenna system. It therefore must be given special consideration in the course of the antenna design.

The design study indicated that considerations of maximum efficiency and minimum size and weight dictate differences in design for different parts of the microwave frequency band. In particular, the design of the antenna feed system and the electronic phase shifter is definitely determined by the particular frequency band of operation. As an example, at the higher end of the microwave frequency band (C-band and above), waveguide feeds and ferrite phase shifters tend to provide maximum efficiency and minimum weight. On the other hand, in the lower end of the microwave spectrum, stripline feeds and digital diode phase shifters provide a greater net improvement in the efficiency-weight factor. The antenna systems described below provide definitive implementation of lightweight, low cost, high efficiency, phased array antennas suitable for use in portable ground radar systems, as well as airborne radar systems.

3. DESCRIPTION OF X-BAND AIRBORNE WEATHER RADAR ANTENNA SYSTEM

This antenna was developed by Sedco Systems Inc. under the aforementioned contract with the U. S. Army Electronics Command, Fort Monmouth, New Jersey, for the purpose of being used with an airborne weather radar system. A photograph showing the antenna, the electronic beam steering programmer and the pilot's control unit is shown in Figure 2. This antenna has an aperture dimension of 24" x 10" and provides a 3 db beamwidth of 4 degrees in azimuth and 9 degrees in elevation. The antenna is capable of electronic scanning in the azimuth plane within a pilot adjustable sector of $\pm$ 45 degrees. The antenna may be mechanically oriented in the elevation plane within an angle of $\pm$ 15 degrees by a servo actuated pilot control. The electronic scan rate is adjustable within the range of 1 to 12 cps.

The basic design of this antenna is similar to that shown in Figure 1. A novel parallel plate feed system is used to deliver the input signal to 32 phase shifter-radiator elements and provide the desired cosine amplitude taper. Each phase shifter radiator assembly consists of a J-shaped waveguide section incorporating a flux-transfer, latching ferrite phase shifter. The output of this member is coupled through a coax transition to a 10-way stripline unequal power divider whose outputs are in turn coupled to a 10-element radiator assembly. A photograph of the phase shifter-radiator assembly is shown in Figure 3. This photograph also shows the electronic driver which provides the desired phase shift control signals to the ferrite phase shifter. This driver is designed as a printed circuit plug-in assembly which may be replaced by removing the top cover of the antenna assembly.

Phase and amplitude errors across the antenna aperture affect the RMS sidelobe level, the beam pointing accuracy and to a smaller extent the antenna gain. The specific relationships that apply are as follows:

$$\frac{G}{G_0} \approx \frac{1}{1 + \frac{3\pi}{4}\left(\frac{d}{\lambda}\right)^2 \epsilon^2}$$

where $\frac{G}{G_0}$ = ratio of actual gain to error-free gain

$\frac{d}{\lambda}$ = element spacing

ϵ^2 = total mean-square error $(\Delta^2 + \delta^2)$

Δ^2 = mean-square amplitude error

δ^2 = mean-square phase error

The RMS sidelobe level is given by:

$$S.L._{R.M.S.} = \epsilon \frac{\Sigma (A_n)^2}{(\Sigma A_n)^2} = \frac{1.05\epsilon^2}{N} \qquad \text{(for cosine illumination)}$$

where A_n = relative voltage at the nth radiating element

N = total number of elements

The RMS pointing error is given by:

$$\sigma_\theta = \sigma_\phi \frac{2}{\pi} \frac{(BW)}{\sqrt{N_s}}$$

where σ_θ = rms pointing error

σ_ϕ = rms phase error

Ψ = 3 db beamwidth

N_e = effective number of elements (0.81 N for cosine distribution)

It can be seen that the acceptable RMS phase error is directly proportional to the square root of the number of antenna elements for a given sidelobe and beam pointing accuracy requirement. This implies a substantial increase in performance capability for phase shifters used on small phased arrays.

Typical performance characteristics of this antenna are shown in Figures 4 and 5. Figure 4 shows a composite of a series of azimuth antenna patterns, electronically scanned within the range of ± 45 degrees from boresight at intervals of approximately 2 beamwidths. It may be seen that the peak sidelobe recorded on any of these beam patterns is 24.5 db down. This demonstrates the extremely low value of phase error across the aperture throughout the electronic scan range. Figure 5 shows an elevation beam pattern. In this case the sidelobe level is 26 db down.

The antenna gain is 28 db at 9.375 GHz. The total loss in the antenna is less than 2.5 db. Of this loss 1 db is contributed by the 10-way elevation stripline power divider and slightly less than 1 db by the ferrite phase shifter. The ferrite phase shifter used in this antenna is a flux-transfer, nonreciprocal, latching phase shifter developed by Sedco. It is a highly efficient unit having a loss of less than 1 db and automatic temperature compensation. The phase shift is obtained by applying a constant amplitude, variable width pulse signal. This is based on the fact that the differential phase shift is directly proportional to the differential flux, which, by the application of Faraday's Law, is as follows:

$$\Delta\Phi = \int E\,dt$$

For constant voltage this may be written as:

$$\Delta\Phi = E\Delta t$$

The differential phase shift may therefore be continuously varied in an analog fashion by applying constant amplitude, variable width pulses from the driver circuit. Moreover, variations in insertion phase between phase shifters may be eliminated by the application of an initial bias signal to the phase shifter driver.

A curve showing the linear relationship between phase shift and pulse width is shown in Figure 6. The loss characteristic of this phase shifter is given in Figure 7. The high degree of temperature stability is obtained by operating the phase shifter below the region of ferrite saturation. Thus the normal variation of ferrite saturation magnetization with temperature will not affect the differential phase shift of the unit even with relatively inexpensive ferrite materials. A plot of differential phase shift vs. temperature is given in Figure 8.

The weight of this antenna is approximately 40 pounds. It is expected that productization and use of lighter weight materials can bring the weight down to approximately 25 pounds. The unusually high performance capabilities of this antenna make it extremely attractive for various airborne applications, such as weather radar, terrain clearance, and ground target surveillance and tracking.

4. DESCRIPTION OF X-BAND FORWARD AREA, INERTIALESS SCAN, SURVEILLANCE RADAR ANTENNA

This antenna is designed to provide inertialess scanning in the azimuth plane within a controllable sector up to ± 60 degrees. Since the radar system must operate at very short ranges, it is not desirable to use nonreciprocal phase shifter elements which require switching between the transmit and receive mode. Consequently the phase shifters used in this antenna are based on a novel Sedco design of a reciprocal latching ferrite phase shifter. This unit was developed by Sedco as part of a separate R&D contract with the U. S. Army Electronics Command at Fort Monmouth, New Jersey.

A photograph of this antenna is given in Figure 9. The design of this antenna is similar to that of the airborne antenna described previously, with the exception that this antenna contains 62 phase shifter radiator elements, providing an aperture width of 42" and an azimuth half-power beamwidth of 2 degrees. The aperture height is 10", providing an elevation half-power beamwidth of 9 degrees. A parallel plate feed, similar to that of the airborne antenna, is used, having appropriately larger dimensions corresponding to the larger number of elements. In order to maintain the required rigidity in the feed without the need for excessive stiffening elements, use is made of aluminum honeycomb panels which have a very high stiffness to weight ratio.

As indicated above, the phase shifters used in this antenna are reciprocal latching devices. These phase shifters are incorporated in a J waveguide section similar to that used in the airborne antenna described above. A photograph of the reciprocal latching phase shifter, similar to that used in this antenna, is shown in Figure 10. The overall dimensions of this device are similar to those of the nonreciprocal phase shifter. However, it contains a completely different type of ferrite element, namely, a double toroid, longitudinally magnetized element. In view of the complete electrical symmetry in this device, reciprocal differential phase shift is obtained. The latching characteristic of the phase shifter results in the need for electronic drive power for the purpose of beam scan which is directly a function of the beam switching rate and is independent of the radar pulse repetition frequency.

The phase shifter drivers are similar to those used in the airborne antenna, namely, printed circuit plug-in assemblies which are inserted by removing the top cover of the antenna array.

This antenna has a provision for automatic tracking in addition to continuous scan. In the automatic tracking mode the antenna, upon detection of a target, will automatically lock on and maintain track to an accuracy of at least 1/20 of a beamwidth, or 4 minutes of arc. A photograph of the tripod mounted antenna and the electronic beam control unit is given in Figure 11. In addition to the adjustable scan sector width controls provided by thumbwheel dials, the beam control unit contains two types of readouts. One is a 3-digit Nixie tube display which provides azimuth beam position within a ± 60 degree sector to an accuracy of 0.1 of a degree during the track mode of operation. In addition, a track error meter provides an additional order of magnitude of accuracy in beam position, namely, to an accuracy of 0.01 degrees. This system is designed primarily for detection and tracking of moving ground targets and is normally used with an MTI type radar system.

The antenna feed is provided with a two-channel monopulse input which is connected to a conventional duplexer monopulse radar receiver. The monopulse error output of the radar receiver and the sum channel output are connected to the error and detection inputs of the beam control unit. This completes the feedback loop and permits the antenna to provide automatic target tracking. The azimuth scan rate of the antenna is adjustable between 0 and 13 cps within any part of the azimuth sector of ± 60 degrees.

A complete set of tests has been conducted on this antenna prior to delivery. The results of these tests have shown that the antenna rms sidelobe level, beam pointing and tracking accuracies are in conformance with the specification. The tripod mount permits mechanical adjustment of the antenna within 360 degrees of azimuth and approximately ± 30 degrees in elevation. This type of antenna system may be used for many ground based applications with existing radar systems while providing a significant enhancement of capability in multiple-target surveillance and tracking accuracy.

5. DESCRIPTION OF S-BAND FORWARD AREA INERTIALESS SCANNED SURVEILLANCE RADAR ANTENNAS

An antenna designed for S-band operation was also fabricated as part of this program. A photograph of this antenna is shown in Figure 12. Although the general characteristics of this antenna are similar to those of the X-band antenna previously described, there are a number of significant differences in the design of the two systems. The S-band antenna consists of 30 elements providing a horizontal aperture of 5' and an azimuth half-power beamwidth of 5 degrees. Each element consists of an aluminum dip brazed assembly of 4 radiators having a height of 15" and providing a vertical half-power beamwidth of 20 degrees. A stripline board containing a 4-bit diode phase shifter and a 4-way unequal power divider is mounted alongside each radiating element. The inputs to the diode phase shifters are connected by a stripline unequal split corporate feed assembly. Two such feed networks are used, one for each half of the antenna. The inputs of these corporate feeds are connected to a stripline hybrid in order to provide a sum and difference monopulse output which may be used with a conventional monopulse radar receiver-transmitter for the purpose of automatic tracking.

It may be seen in Figure 12 that the electronic beam control unit is the same as that used in the X-band antenna previously described. In fact it is the same unit, the design of which is such as to operate satisfactorily with either the ferrite scanned X-band antenna or the diode scanned S-band antenna. The beam control unit therefore is in general capable of operating with the two major types of phased array antenna systems.

A schematic representation of the 4-bit diode phase shifter and 4-way power divider is given in Figure 13. The phase shifter contains 4 stripline hybrids and 8 PIN diodes. Each hybrid and diode pair comprise one of the phase shifter bits and in effect switch the corresponding differential phase shift in and out of the system. Since the minimum bit is 22.5 degrees, the quantization error will reach the maximum value of 11.25 degrees. Since the antenna contains a total of 30 phase shifter-radiator elements, the effect of this quantization error limits the rms sidelobe level to a value of 29 db. If a smaller number of bits were used in the phase shifter, this effect would be substantially more severe and would seriously deteriorate the antenna performance.

The 4-way stripline unequal power divider is designed to provide a Tchebycheff distribution in the elevation plane. This distribution provided an elevation sidelobe level of 23 db.

The extensive use of stripline networks permitted the attainment of a very efficient form factor. The overall antenna dimensions are 5' wide x 15" high x 6" deep. The weight of the antenna is approximately 60 pounds. An additional technique that was used in the mechanical design permits the antenna to be separated in two parts across the long dimension. This permits the antenna to be easily carried since each half has the dimensions of a relatively small sized suitcase.

The performance characteristics of this antenna are similar to those of the X-band antenna described previously. It has the same capability for electronic scan and automatic track within a sector of ± 60 degrees. Since the same electronic control system is used, all of the controls and readouts apply equally to the S-band antenna. Thus the beam pointing and tracking accuracies which are specified as 1/30 and 1/20 of a beamwidth respectively, will result in equivalently larger errors for the S-band antenna.

This antenna has been completely tested and delivered. Results of the tests indicate the attainment of rms sidelobe level and tracking accuracy in accordance with specifications. The tripod mount used for this system permits mechanical orientation within 360 degrees in azimuth and +30 degrees and -60 degrees in elevation.

6. CONCLUSIONS

The antenna systems developed during the course of this program have demonstrated the capability for the design of relatively light weight, moderate complexity and low cost, field type airborne and ground based phased array antenna systems in the frequency range of S through X-band. Additional work carried on at Sedco on critical components such as phase shifter and radiator elements permit the extension of similar design techniques over the frequency range of L through K-band. An important part of the program was the attainment of an optimum design in each frequency band from the point of view of performance, efficiency, size and weight. Although the specified power level for these antennas was moderate, similar designs can be used for power levels as high as an order of magnitude greater. The antenna performance achieved in terms of sidelobe level and beam pointing and tracking accuracy is in general greater than is attainable with conventional parabolic dish antenna systems. The electronic beam control system developed during this program is an extremely versatile unit which is capable of controlling two general classes of phased array antennas, namely, ferrite scanned or diode scanned systems. Based on the results achieved during this program, it may be expected that antennas similar to those developed here could be placed into early operational use as part of field type airborne and ground based radar systems.

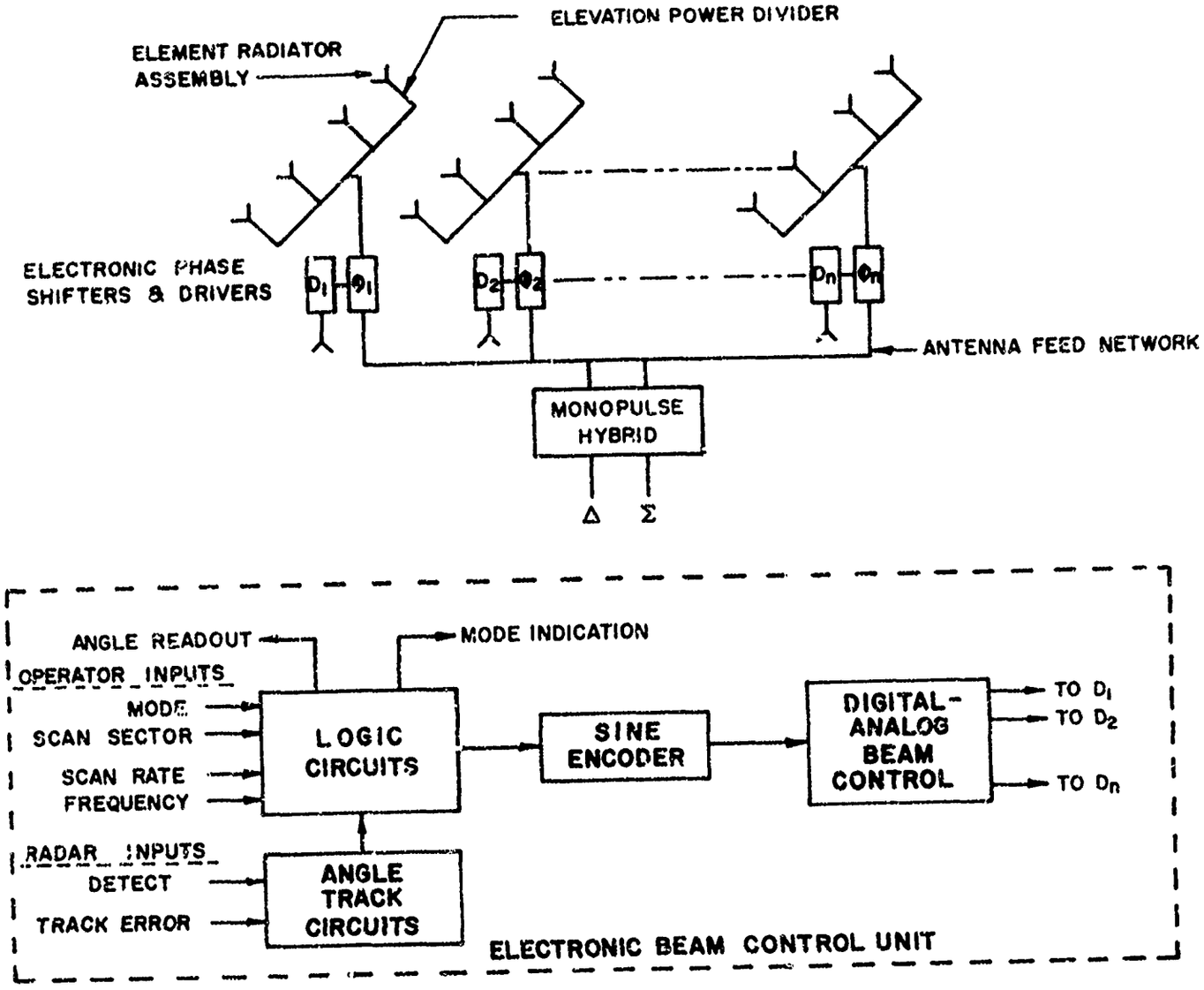


Fig. 1 Schematic diagram of phased array antenna system

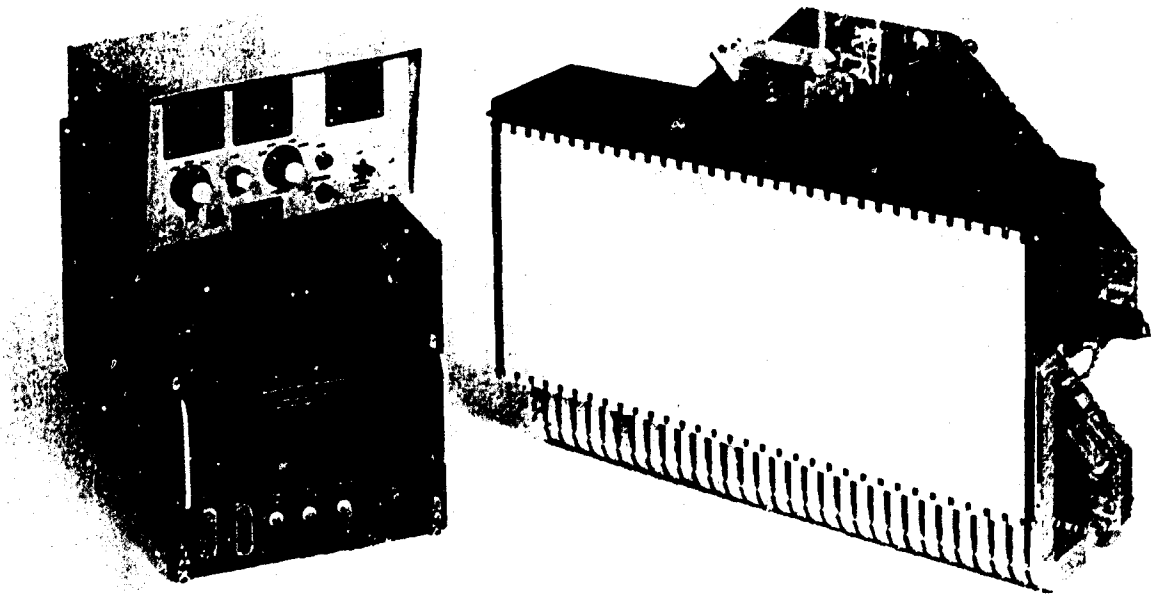


Fig. 2 X-band inertialess-scanned, weather-radar, antenna system

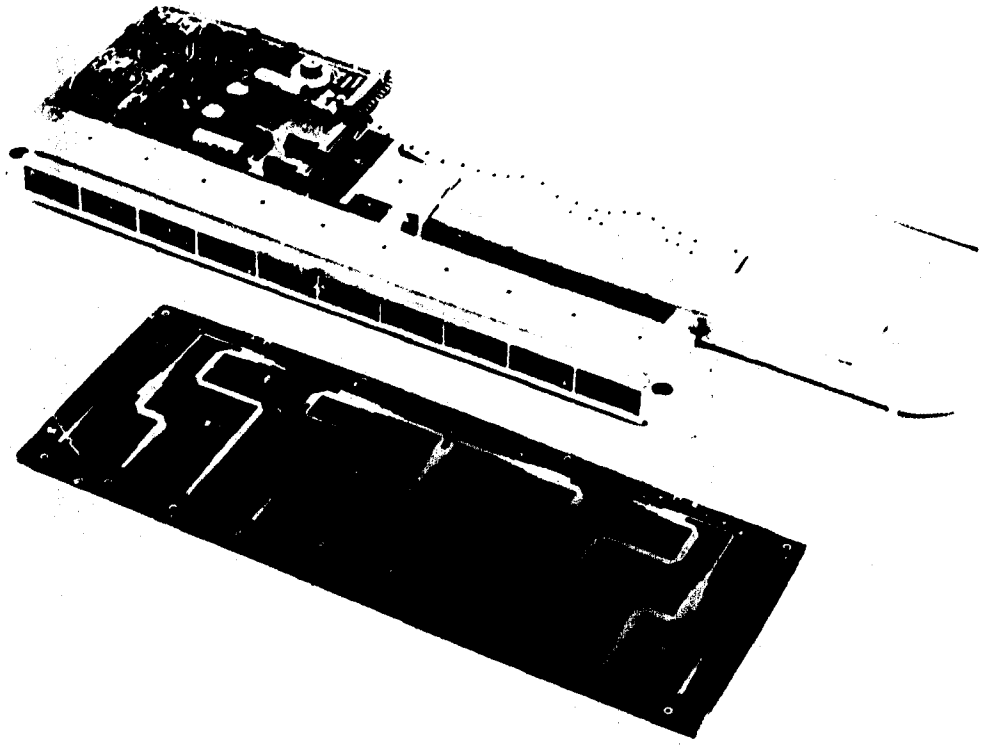


Fig. 3 Radiating element including power divider, phase shifter and driver

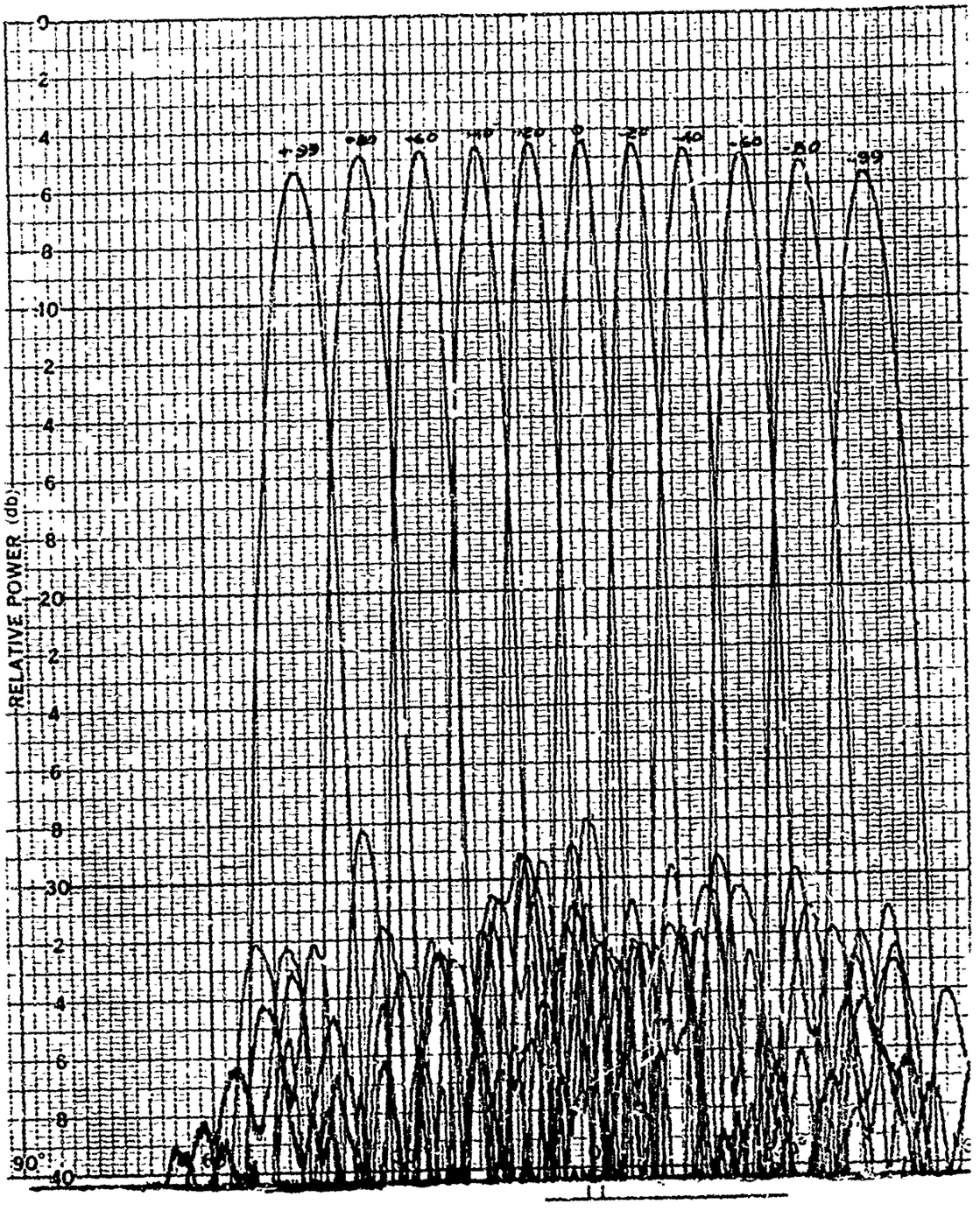


Fig. 4 Electronically scanned azimuth beam patterns, airborne weather radar antenna (true angle scale at bottom of figure)

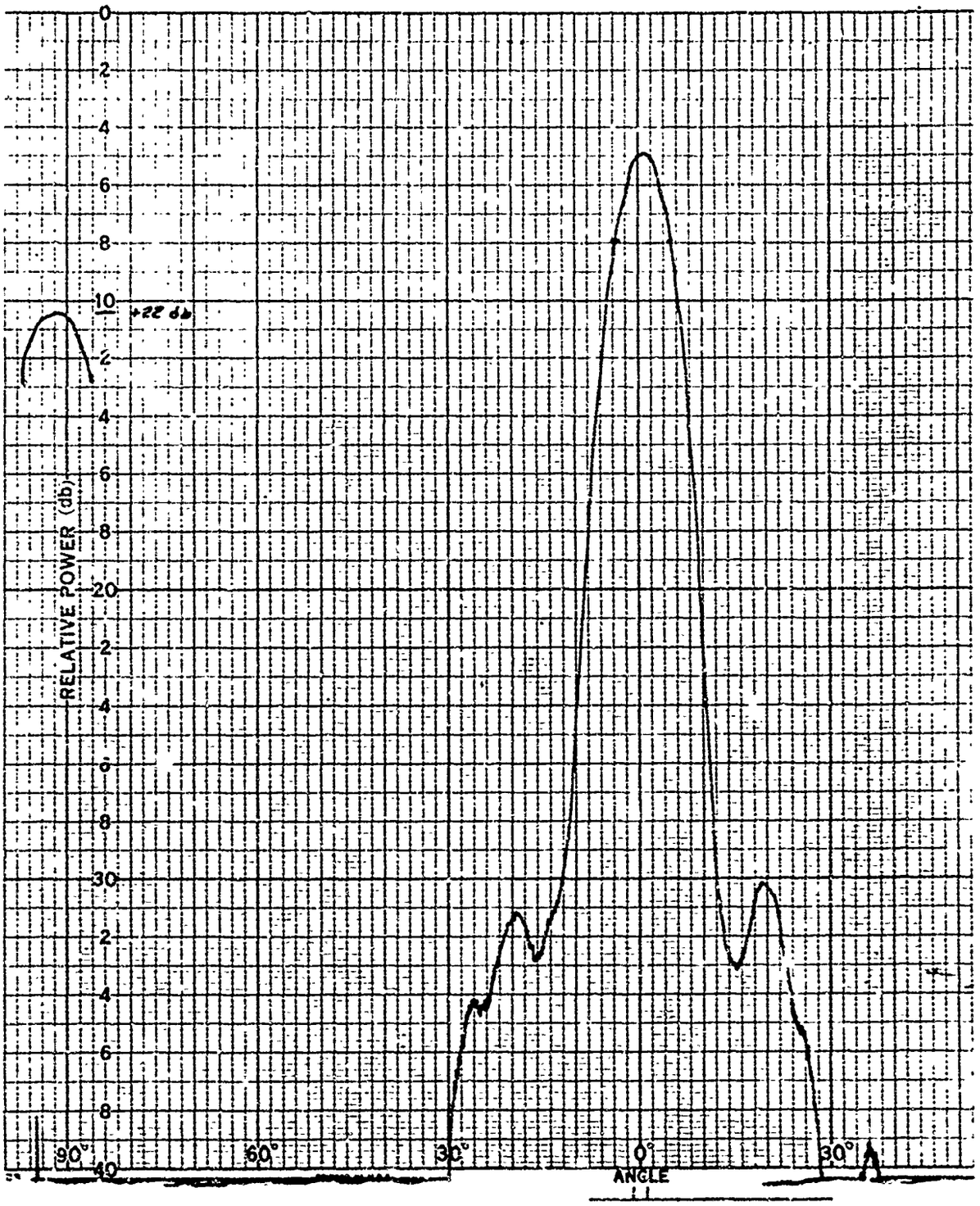


Fig. 5 Elevation beam pattern, airborne weather radar antenna

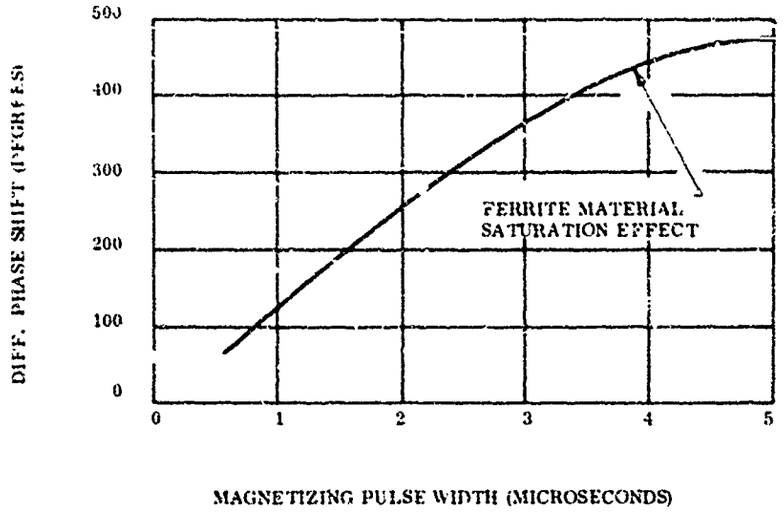


Fig 6 Differential phase shift vs. magnetizing pulse width

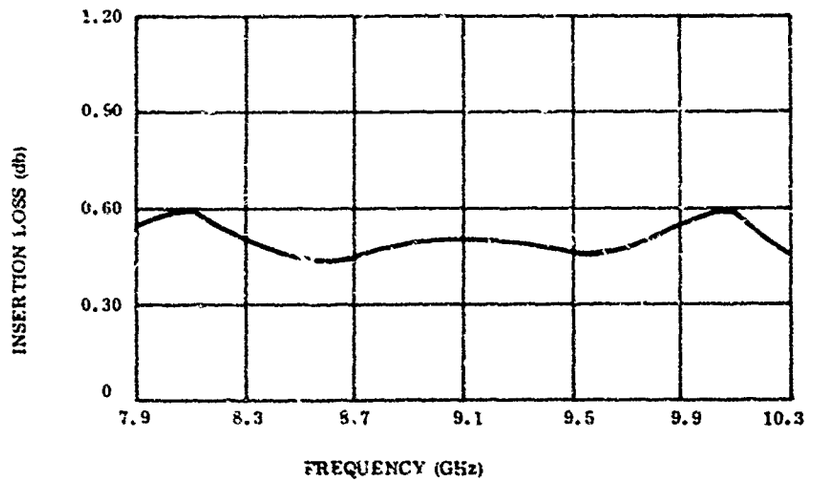


Fig. 7 Insertion loss vs. frequency

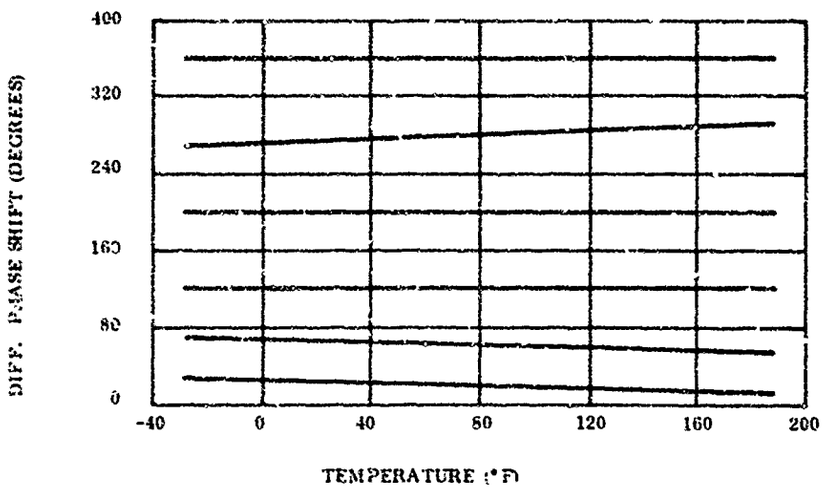


Fig. 8 Differential phase shift vs. temperature

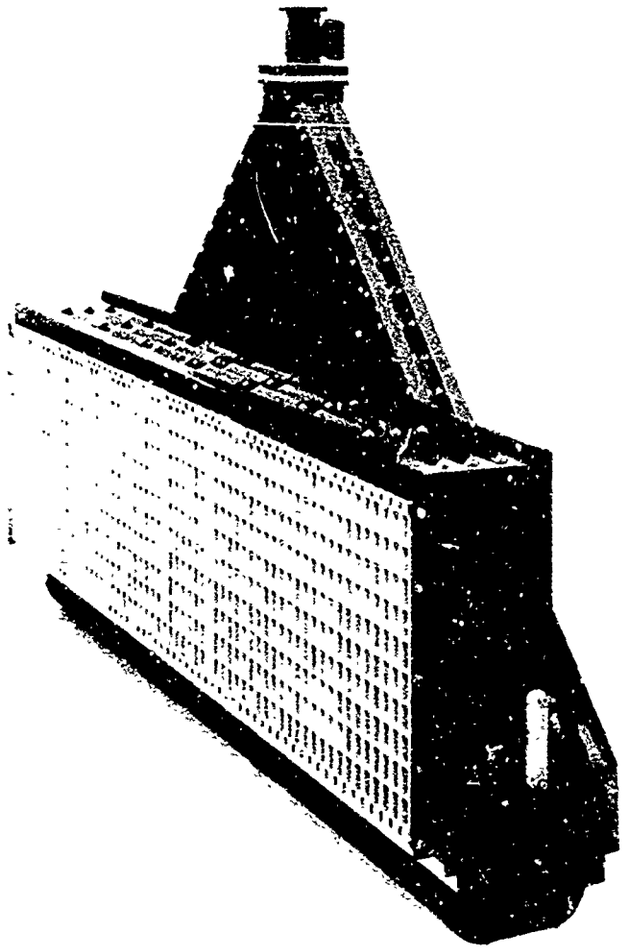


Fig. 9 X-band inertialess-scanned surveillance-radar antenna

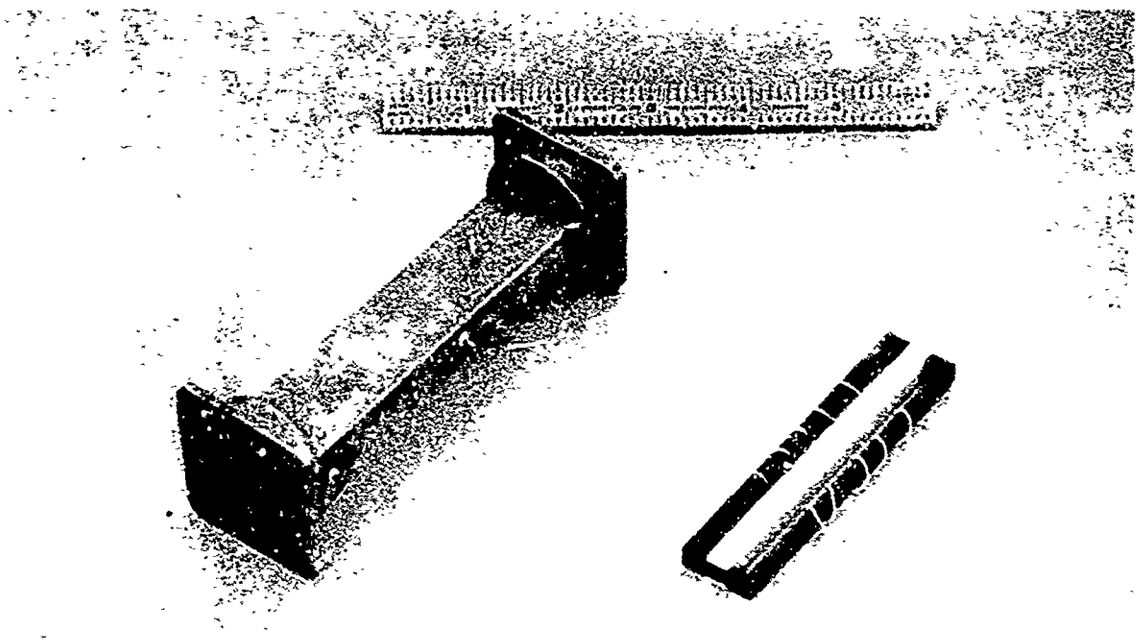


Fig. 10 Reciprocal latching ferrite phase shifter

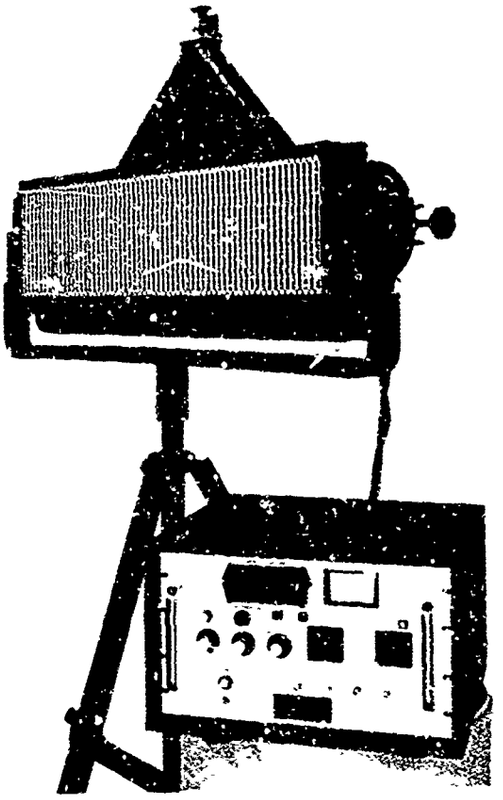


Fig. 11 X-band inertialess-scanned surveillance-radar antenna system

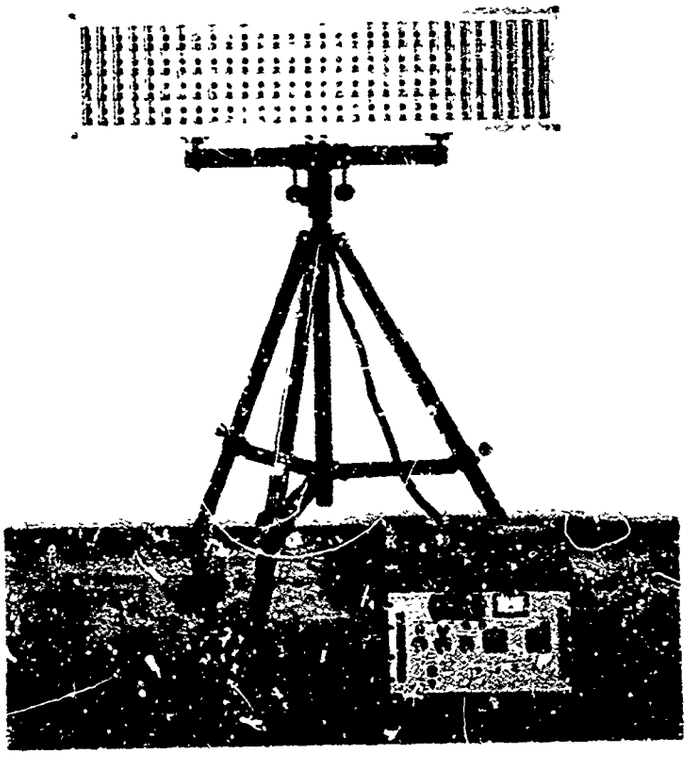


Fig. 12 X-band inertialess-scanned surveillance-radar antenna system

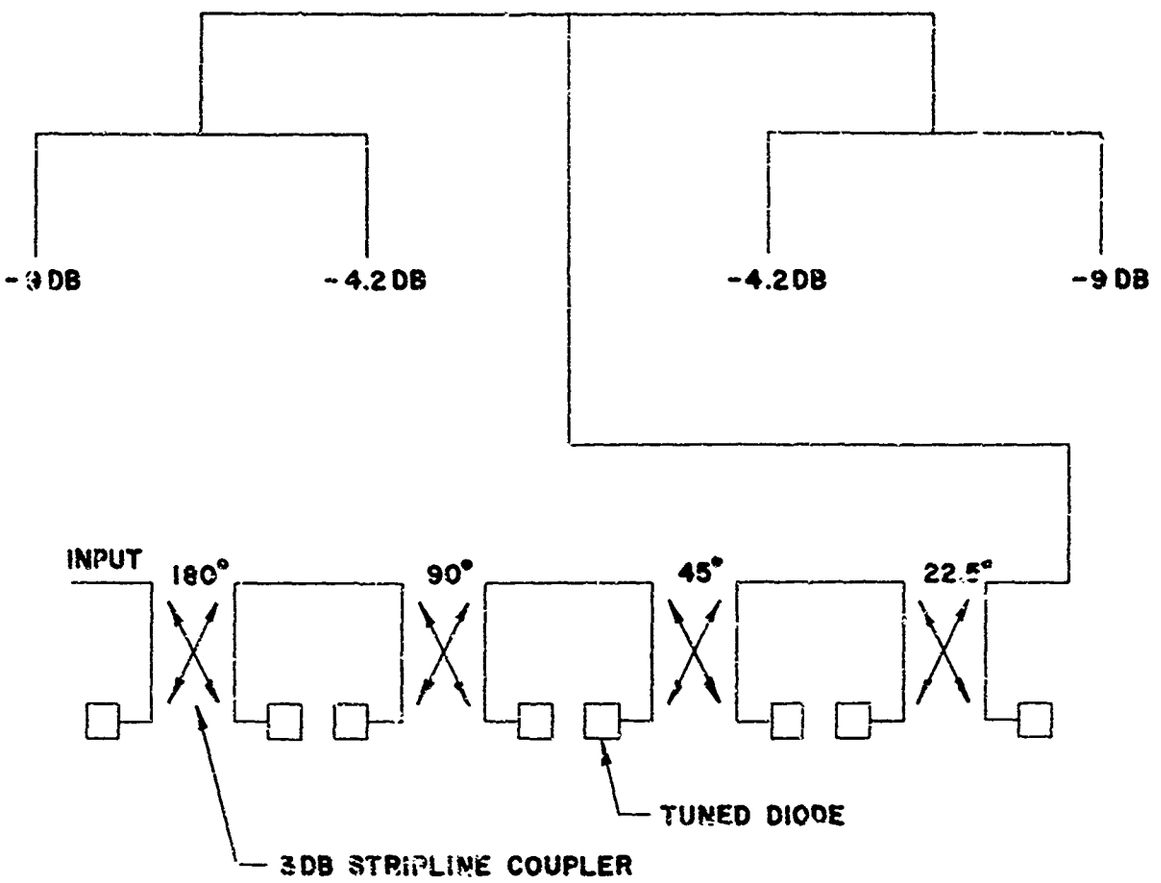


Fig 13 Schematic representation X-band elevation power divider and 4 bit diode phase shifter

RADAR DE POURSUITE
A GRAND POUVOIR SEPARATEUR EN DISTANCE

Par

Henri C.P. FORESTIER et Michel CHEVALIER

THOMSON - CSF
Division Radar de Surface
92, Bagneux, FRANCE

La Recherche présentée dans cet exposé est partiellement supportée par la

DELEGATION MINISTERIELLE POUR L'ARMEMENT, Direction des Recherches et Moyens d'Essais FRANCE

RADAR DE POURSUITE A GRAND POUVOIR SEPARATEUR EN DISTANCE

Henri G.P. FORESTIER et Michel CHEVALIER
THOMSON - CSF
92 Bagneux, France.

SOMMAIRE

Pour obtenir un grand pouvoir séparateur en distance, un radar doit émettre un spectre de fréquence très large. Par contre, les techniques classiques d'acquisition de cible et d'exploitation des signaux d'écartométrie (distance et angles) s'accomodent mal de la très grande bande passante des signaux reçus ou des faibles durées d'impulsions comprimées correspondant au grand pouvoir séparateur en distance recherché.

Pour pallier ces inconvénients, une certaine compression de bande passante, malgré une dilatation corrélative de l'échelle des temps, semble hautement souhaitable au cours du traitement du signal avant son exploitation, pour en tirer les signaux d'écartométrie utilisables. Un tel traitement est possible dans un radar de poursuite pour lequel les signaux exploités ne correspondent qu'à une faible portion de la récurrence radar.

On envisage donc une telle forme de traitement du signal, dans lequel on fait appel à la fois aux techniques de corrélation et de compression d'impulsion, pour doter un radar de poursuite d'un grand pouvoir séparateur en distance, correspondant à un spectre de plusieurs dizaines de Megahertz, exploité sous forme d'impulsions de quelques dizaines (au lieu de centièmes) de Microsecondes. Ceci permet d'envisager un pouvoir séparateur en distance de quelques mètres, exploitable avec des circuits videofréquences radar classiques.

1. INTRODUCTION.

Pour obtenir un grand pouvoir séparateur en distance, un radar doit émettre un spectre de fréquence très large. L'emploi d'une modulation de fréquence à l'intérieur d'une impulsion d'émission de durée relativement grande pour obtenir de larges spectres de fréquence est à l'heure actuelle une technique radar largement répandue et les diverses façons d'exploiter une telle forme de signal à la réception ont été largement décrites pour réaliser ce qu'il est convenu d'appeller la compression d'impulsion.

En pratique il existe deux grandes méthodes possibles pouvant servir de base à la réalisation d'un récepteur radar :

- l'emploi d'un corrélateur effectuant la corrélation du signal reçu avec un signal identique au signal émis et convenablement retardé dans le temps, et

- l'emploi d'un filtre adapté au signal émis.

Chacune des deux méthodes est actuellement bien connue ainsi que leurs avantages et inconvénients respectifs.

L'emploi simultanné des deux méthodes offre des particularités intéressantes notamment dans le domaine des radars de poursuite à grand pouvoir séparateur en distance ainsi que nous allons le montrer en décrivant, à titre d'exemple, une étude de recherche effectuée en France à la THOMSON - CSF.

2. PRINCIPE DE FONCTIONNEMENT.

Il repose essentiellement sur le choix d'un signal émis par le radar, avec des caractéristiques permettant un grand pouvoir séparateur en distance, qui puisse être traité à la réception de façon efficiente.

2.1. Signal émis (Fig.1.).

Le radar émet un signal de durée assez grande T_E (de l'ordre de la dizaine de microsecondes), la fréquence d'émission variant linéairement pendant la durée d'émission entre les fréquences $F_E - \frac{\Delta F}{2}$ et $F_E + \frac{\Delta F}{2}$.

P est grand et correspond au pouvoir séparateur en distance recherché de l'ordre de quelques dizaines de Megahertz pour un pouvoir séparateur de quelques mètres). Le produit $T_E \cdot \Delta F$ est grand (de l'ordre de plusieurs centaines) de même que la pente de modulation de fréquence $\frac{\Delta F}{T_E}$ (de l'ordre de plusieurs Megahertz par microseconde).

2.2. Signal reçu-traitement à la réception.

Après réflexion sur une cible, le signal reçu est très semblable au signal émis (à l'atténuation et à l'effet Doppler près). Le signal reçu sera traité en deux étapes successives :

- Une corrélation imparfaite,

- Une compression d'impulsion par dispersion dans un filtre adapté au résultat de la corrélation partielle.

2.2.1. Corrélation imparfaite. (Fig.2)

La corrélation s'effectue à l'aide d'un signal d'oscillation locale approprié, réplique partielle du signal émis, convenablement retardé dans le temps. Un oscillateur local déclenché avec un certain retard T_D par rapport à l'émission génère un signal de durée T_R, supérieure à T_E, et dont la fréquence varie linéairement pendant la durée T_R entre les fréquences $F_R - \frac{\Delta F'}{2}$ et $F_R + \frac{\Delta F'}{2}$. $\Delta F'$ est supérieur à ΔF et voisin, mais non égal à $\Delta F . \frac{T_R}{T_E}$ soit :

$$\left.\begin{array}{l} \Delta F' = K\ \Delta F\ \dfrac{T_R}{T_E} \\ K \neq 1 \qquad \text{et} \quad (1-K) \ll 1 \end{array}\right\} \qquad (1)$$

La pente de modulation de fréquence du signal d'oscillation locale est $\frac{\Delta F'}{T_R} = K\ \frac{\Delta F}{T_E}$.

La démodulation du signal reçu par le signal d'oscillation locale fournit (pour autant que le signal reçu existe entre les instants T_D et $T_D + T_R$) un signal de durée T_E et dont la fréquence varie linéairement pendant cette durée entre les fréquences $F_I - \frac{\Delta F''}{2}$ et $F_I + \frac{\Delta F''}{2}$. La pente de modulation de fréquence du signal démodulé $\frac{\Delta F''}{T_E}$ a pour valeur absolue la différence des pentes de modulation entre le signal émis et le signal d'oscillation locale soit.

$$\left.\begin{array}{l} \dfrac{\Delta F''}{T_E} = (1-K)\ \dfrac{\Delta F}{T_E} \\ \text{ou} \qquad \Delta F'' = (1-K) \cdot \Delta F . \end{array}\right\} \qquad (2)$$

Cette pente est positive ou négative selon que F_R est inférieur ou supérieur à F_E.

Si le signal d'oscillation locale était une réplique exacte (au lieu de partielle) du signal émis, c'est-à-dire si on avait K = 1 ou ce qui revient au même $\Delta F' = \Delta F \frac{T_E}{T_R}$ le résultat de la démodulation du signal reçu par l'oscillation locale serait un signal de durée T_E (sans modulation de fréquence de la porteuse soit $\Delta F'' = 0$), à la fréquence F_I. Dans le cas de la corrélation parfaite ou totale c'est la fréquence porteuse F_I du signal de durée T_E qui contient l'information d'écartométrie temporelle ou en distance. Pour une distance de cible $D_o = \frac{c}{2}(T_D + \frac{T_R}{2} - T_E)$ on a $F_I = |F_E - F_R|$, tandis que pour une distance de cible $D = D_o + \Delta D$ on a $F_I = \left|F_E - F_R - \frac{2K\Delta D}{c}\frac{\Delta F}{T_E}\right|$. De façon analogue dans le cas de la corrélation imparfaite c'est la fréquence médiane du signal résultant de la corrélation partielle qui contient l'information d'écartométrie temporelle. A une distance de cible D_o correspond une fréquence médiane $F_I = |F_E - F_R|$ et à la distance de cible $D = D_o \pm \Delta D$ correspond la fréquence médiane $F_I = |F_E - F_R - \frac{2K\Delta D}{c}\frac{\Delta F}{T_E}$.

2.2.2. Compression d'impulsion. (Fig.3)

Le signal partiellement corrélé résultant de la démodulation du signal reçu par le signal d'oscillation locale subit un filtrage cohérent dans un réseau dispersif dont la caractéristique temps de propagation en fonction de la fréquence a une pente de valeur absolue égale à la pente de modulation de fréquence du signal partiellement corrélé mais de sens opposé. Autrement dit le réseau dispersif comprime de façon adaptée le signal partiellement corrélé et ceci sous forme d'un signal de durée approximativement égal à l'inverse de $\Delta F'$ (soit de l'ordre de $\frac{1}{(1-K)\Delta F}$) positionné à un certain temps T_C tel que:

$$\left.\begin{array}{l} T_C = T_D + \dfrac{T_R}{2} + T_o + \dfrac{2K\cdot\Delta D}{c} \cdot \dfrac{\Delta F}{T_E} \cdot \dfrac{T_E}{(1-K)\Delta F} \\ \text{ou}\ T_C = T_D + \dfrac{T_R}{2} + T_o + \dfrac{2\Delta D}{c} \cdot \dfrac{K}{1-K} \end{array}\right\} \qquad (3)$$

T_o étant le temps de propagation du réseau dispersif pour la fréquence $|F_E - F_R|$. Dans l'opération de compression d'impulsion l'écartométrie temporelle de cible dont l'information était contenue dans la fréquence médiane du signal partiellement corrélé est redevenue une information temporelle sous la forme d'un retard pour lequel toutefois le facteur de proportionalité avec les distances n'est plus donné par la vitesse C de propagation des ondes électromagnétiques mais par cette vitesse multipliée par $\frac{1-K}{K}$, sensiblement égal à 1 - K.

2.3 Interprétation du résultat du traitement du signal.

Nous avons vu qu'à la sortie du récepteur les signaux provenant d'échos de cibles apparaissent sous forme d'impulsions de largeur $\frac{1}{(1-K)\Delta F}$ et que deux cibles distinctes séparées en distance de ΔD apparaissent sous forme de deux impulsions distantes de $\frac{2\Delta D}{c}\frac{K}{1-K}$. On peut donner de ce résultat diverses interprétations équivalentes permettant de mieux saisir l'intérêt de la combinaison simultanée de la corrélation avec la compression d'impulsion.

Une première interprétation tirée directement du facteur de proportionalité entre les distances et les retards substitue à la vitesse de propagation des ondes électromagnétique une vitesse de propagation fictive $\frac{K}{1-K}$ fois plus petite d'où possibilité d'exploiter un pouvoir séparateur en distance donné avec des circuits de bande passante réduite dans les mêmes proportions.

Une seconde interprétation revient à considérer une compression d'impulsion normale de taux de compression $T_E \cdot \Delta F$ suivi d'une dilatation de l'échelle des temps de $\frac{K}{1-K}$ fois comme si on avait utilisé une "loupe" de grossissement $G = \frac{K}{1-K}$ pour observer les signaux sur un oscilloscope. Cette dernière interprétation de la dilatation de l'échelle des temps implique une conséquence importante pour le traitement du signal reçu en "temps réel" : en effet si le grossissement est de G on ne peut l'effectuer en temps réel qu'au maximum sur $\frac{1}{G}$ de la résidence radar. Ceci exclut l'application d'un tel traitement pour un radar de veille ou de surveillance mais s'applique généralement bien au cas du radar de poursuite pour lequel les signaux exploités ne correspondent qu'à une faible fraction de la période de récurrence.

L'application à un radar de poursuite de traitement du signal décrit permet (grâce à la compression de bande passante corrélative à une dilatation possible de l'échelle des temps sur une partie de la période de récurrence) d'utiliser des circuits video radar classiques pour l'acquisition de cibles et l'écartométrie appliquée à des radars à grand pouvoir séparateur en distance.

2.4. Limitations du procédé.

Outre la limitation déjà citée de l'emploi à un radar de poursuite ou d'atterrissage, la nécessité d'effectuer correctement la corrélation imparfaite et la compression d'impulsion pour des cibles situées dans un certain domaine spatial d'étendue non nulle applique certaines contraintes ou limitations d'ordre technologique.

Nous avons déjà signalé que la durée T_R de l'oscillateur local devait être supérieure à la durée T_E du signal émis, la quantité $\frac{c}{2}(T_R - T_E)$ définit la profondeur de la zone dans laquelle peuvent se situer les cibles. Pour pouvoir exploiter entièrement cette zone il faut encore que le réseau dispersif effectuant la compression d'impulsion ait une bande passante suffisante : la valeur minimale de cette bande passante est de $(1-K+K\frac{T_R}{T_E}-1)\Delta F = \frac{T_R - T_E}{T_E} K \Delta F$.

Le traitement du signal à la réception s'effectuant en deux étapes dont la dernière consiste en une compression d'impulsion par un réseau dispersif il importe pour que cette dernière opération s'effectue dans des conditions satisfaisantes que le taux de compression $(1-K)\,T_E\,\Delta F$ soit suffisamment élevé ce qui implique à la fois T_E pas trop faible, si l'on ne veut pas avoir une impulsion comprimée trop brève, et $(1-K)\cdot\Delta F$ suffisamment grand vis à vis de l'inverse de T_E. Ceci peut limiter le grossissement $G = \frac{K}{1-K}$ à une valeur pas trop élevée.

De toutes façons ces diverses limitations théoriques imposent des contraintes dans le choix des divers paramètres d'une réalisation particulière et font apparaître la nécessité de la recherche du meilleur compromis pour chaque application particulière envisagée.

Enfin, outre les limitations théoriques citées, certains des problèmes liés à un grand pouvoir séparateur en distance restent entiers en ce qui concerne les limitations tehcnologiques et les imperfections de réalisation. A ce titre nous ne citerons que le plus important à savoir la réduction des "jitters" dans tous les générateurs de base de temps et de télémétrie.

3. REALISATION ET MISE EN OEUVRE DU PROCEDE.

Afin de mettre en oeuvre les principes précédemment décrits il a été réalisé une maquette destinée à être associée à un radar dont la partie émission, tête HF et aérien présente une bande passante instantannée suffisante. Cette maquette comporte (Fig.4) un générateur de signal d'émission, un générateur de signal d'oscillation locale, un récepteur à corrélation imparfaite, un ensemble de compression d'impulsions et une télémétrie.

Les paramètres retenus pour la maquette de mise en oeuvre sont résumés dans le tableau I.

3.1. Générateurs de signal d'émission et d'oscillation locale.

La génération du signal d'émission et d'oscillation locale est du type actif et fait appel à un montage original qui confère à l'ensemble une très grande souplesse permettant de faire varier dans certaines limites les paramètres du signal émis.

Chacun des deux générateurs est basé sur le même principe seuls certains paramètres ont des valeurs différentes, nous ne décrirons sommairement que le générateur de signal d'émission. Il est constitué par un générateur de fréquence en marche d'escalier à commande digitale : le saut de fréquence élémentaire est de $\frac{\Delta F}{32}$ et un saut de fréquence de $\frac{\Delta F}{32}$ intervenant tous les $\frac{T_E}{32}$.

Dans le passé de tels générateurs de fréquence en marche d'escalier ont été réalisés par commutation d'oscillateurs synchronisés par un standard de fréquence et ont toujours donné lieu à des difficultés liées aux discontinuités de phase lors des commutations de fréquence.

Nous avons délibérément choisi une approche différente (Fig.5). Nous sommes partis de 5 oscillateurs commandés pouvant chacun prendre les deux fréquences $(F_o - \frac{\Delta F}{2^{n+1}})$ et $(F_o + \frac{\Delta F}{2^{n+1}})$ n prenant les valeurs 1, 2, 3, 4 ou 5. Le saut de fréquence de $\frac{\Delta F}{2^n}$ est obtenu en branchant ou non en parallèle sur le circuit oscillant de l'oscillateur une réactance asservie de grandeur appropriée, le branchement s'effectuant par commande digitale à l'aide des bits successifs d'un compteur binaire à 5 étages.

Les sorties des cinq oscillateurs controlés sont ajoutés ensuite dans une chaîne de transposition de fréquence classique du type de celles utilisées dans les synthétiseurs de fréquence.

Moyennant certaines précautions sur le branchement des réactances asservies nous avons pu faire disparaître les discontinuités de phase lors des commutations de fréquence. Par contre nous avons échangé la stabilité de fréquence d'oscillateurs synchronisés par un standard de fréquence contre la stabilité intrinsèque d'oscillateurs non synchronisés. Nous avons pu lever cet inconvénient en effectuant un asservissement sur les valeurs des réactances en utilisant un C.A.F. à ligne à retard avec battement de fréquence et discriminateur de fréquence (Fig.6) pour controler les sauts de fréquence $\frac{\Delta F}{2^n}$. Il a même été possible d'utiliser en temps partagé un seul discriminateur de fréquence venant séquentiellement controler et asservir les différents sauts de fréquence $\frac{\Delta F}{2^n}$ et ceci en dehors des instants d'émission.

La figure 7 donne le schéma de principe d'un oscillateur commandé avec emploi de transistors à réactance dans un montage symétrique. La valeur de la réactance équivalente est obtenue en contrôlant le gain des transistors au moyen de l'asservissement du générateur de courant.

La fig. 8 montre la réalisation d'un oscillateur commandé.

En retournant à la Fig. 1 on peut voir que les instants de commutation d'échelons de fréquence $\frac{\Delta F}{32}$ sont commandés par une horloge, en faisant varier la fréquence de l'horloge du compteur linaire à 5 étages il est possible de faire varier la pente de la modulation de fréquence sans faire varier l'excursion de fréquence ΔF. Enfin une sélection de largeur variable permet de régler la durée d'émission jusqu'à 32 pas d'horloge du compteur linaire.

3.2. Récepteur à corrélation imparfaite et compression d'impulsion. (Fig. 9)

Le récepteur comprend un préamplificateur, un mélangeur de signal d'oscillation locale, un amplificateur de fréquence intermédiaire, un réseau dispersif de compression d'impulsion, un détecteur d'enveloppe et un filtre passe bas en videofréquence.

Le préamplificateur comporte un filtre accordé sur la fréquence de réception dont la caractéristique de transfert amplitude fréquence permet d'effectuer une pondération en amplitude du signal reçu. Pour ce faire on profite de ce que le produit $T_E \cdot \Delta F$ est grand pour effectuer une pondération amplitude fréquence qui, du fait de la loi de variation linéaire du signal émis et de la rectangularité du spectre transmis, est équivalent à une pondération amplitude temps.

Sur le signal reçu ainsi pondéré, le mélangeur de signal d'oscillation locale effectue la corrélation imparfaite et transforme la pondération amplitude fréquence en une pondération amplitude temps tout en effectuant une transposition en fréquence du signal pour l'amener (en fréquence intermédiaire) dans la bande passante du réseau dispersif.

Le réseau dispersif est du type à constantes réparties et utilise des ondes de surface de Rayleigh sur un barreau de quartz revêtu d'un transducteur dispersif.

Afin d'obtenir un domaine de distance suffisant des cibles tout en ayant des réseaux dispersifs réalisables il est possible d'utiliser des combinaisons série ou parallèle de plusieurs réseaux dispersifs. En fait nous avons travaillé soit avec un seul réseau dispersif dans une zone de distance étroite soit avec deux réseaux dispersifs en montage parallèle pour une zone de distance plus profonde.

Après dispersion les impulsions comprimées sont détectées par un détecteur d'enveloppe à large bande et la bande passant vidéo fréquence est réduite à environ $\Delta F' = (1-\kappa)\Delta F$, valeur suffisante pour extraire les signaux d'écartométrie distance.

Dans le cas d'une exploitation monopulse du radar de poursuite les voies écart site et écart gisement seront traitées en parallèle de façon analogue en ce qui concerne la corrélation imparfaite, l'écartométrie angulaire étant alors traitée simultanément avec la compression d'impulsion de façon classique.

3.3. Télémétrie.

La télémétrie assurant le retard au déclenchement de la génération du signal d'oscillation locale en fonction de la distance de la cible poursuivie a été réalisée en technique digitale ce qui a permis de réduire au minimum les problèmes de jitter aux grandes distances ceci au prix d'une quantification en distance qui a été limité à un mètre soit environ le quart du pouvoir séparateur en distance.

4. RESULTATS EXPERIMENTAUX.

Une série d'expériences a été effectuée sur la maquette réalisée pour vérifier la validité du procédé décrit. Les résultats de ces expériences sont illustrés par les oscillogrammes de signaux de la Fig. 10. Les délais de publication des "pré prints" nous ont obligé à présenter les toutes premières photographies mettant en évidence certains défauts qui ont été corrigés depuis.

La Fig. 10 a présente le signal d'émission ou impulsion primaire qui dure 10 µ sec. soit 5 cm sur un balayage de 2 microsecondes / cm relevé au tektronix 585. La forme trapézoïdale de l'impulsion résulte de la courbe de réponse de l'oscilloscope qui chute avec la fréquence (entre 70 et 110 MHz).

La Fig. 10 b présente le signal après corrélation imparfaite on peut noter la pondération d'amplitude obtenue par la courbe de réponse du préamplificateur,

La Fig. 10 c représente une impulsion en sortie du milieu dispersif de compression d'impulsion. Sur le balayage à 0,5 microseconde / cm on peut mesurer une largeur d'impulsion à 3 dB de 0,25 microseconde (0,5 cm) correspondant à $\Delta F' \approx 4$ MHz. Les lobes latéraux sont assez mauvais et se situent vers -16 dB ceci est dû partiellement à une pondération dissymétrique et légèrement insuffisante (cf la Fig 10 b) mais aussi et surtout à la modulation de fréquence en marche d'escalier et à des distorsions de phase mal compensées dans la chaîne de combinaison des oscillateurs contrôlés.

La Fig 10 d représente la sortie du milieu dispersif de compression d'impulsion en présence des échos de deux cibles ponctuelles séparées par 4 m 50 et dont la première est à 10 dB en dessous de la seconde ceci a permis de vérifier le pouvoir séparateur de 4 mètres prévus.

La Fig. 10 c est identique à la précédente mais pour deux cibles distantes de 10 m 30 la séparation complète des cibles est nettement apparente. et le grossissement a pu être mesuré sur Fig. 10 e les deux cibles sont distantes de 1,5 cm soit 0,75 μ sec. et le grossissement est de

$$\frac{0,75 \times 150}{10,3} = \frac{112,5}{10,3} = 10,9$$ qui correspond bien à 11 du tableau I dans le cas du fonctionnement à K > 1. (K = 1,10).

Les Figures 10 c à 10 g font également bien apparaître sous forme d'un léger flou les difficultés dues à un léger jitter des bases de temps qui devient visible du fait du grossissement de 11. Ce jitter a depuis été corrigé par l'emploi d'une télémétrie digitale.

5. DISCUSSION DES RESULTATS ET CONCLUSIONS.

L'emploi combiné des techniques de corrélation et de compression d'impulsion pour le traitement d'un signal à large spectre modulé linéairement en fréquence nécessaire pour obtenir d'un radar à grand pouvoir séparateur en distance s'est révélé très prometteur dans le cas d'un radar de poursuite dont les signaux d'erreur peuvent être élaborés à partir de circuits video à largeur de bande classique de quelques Megahertz seulement.

La théorie a pu être vérifiée expérimentalement et on a pu mettre en évidence la dilatation de l'échelle des temps similaire à un grossissement qui accompagne la réduction de bande passante des signaux video.

Les premiers résultats ont mis en évidence les limitations technologiques du système qui ont nécessité des efforts à la fois dans le domaine de la réduction des jitters et dans celui de la correction des défauts de phase dans les diverses parties de la maquette. Après correction de ces diverses imperfections la phase suivante de l'étude doit consister à monter l'ensemble du dispositif sur un radar existant en vue de faire des expériences en vraie grandeur sur le terrain.

Les développements futurs du procédé décrit doivent permettre d'améliorer sensiblement les performances des radars de poursuite à grand pouvoir séparateur en distance.

6. REMERCIEMENTS.

Les auteurs tiennent à remercier les membres de la Division Radar de la Thomson CSF et plus particulièrement M. H. Carpentier Directeur Technique pour les encouragements et l'assistance qui leur ont été apportés au cours de l'Etude qui vient d'être présenté.

La recherche ici rapportée et ses développements sont actuellement partiellement supportés par la Délégation Ministérielle pour l'Armement - Direction des Recherches et Moyens d'Essais.

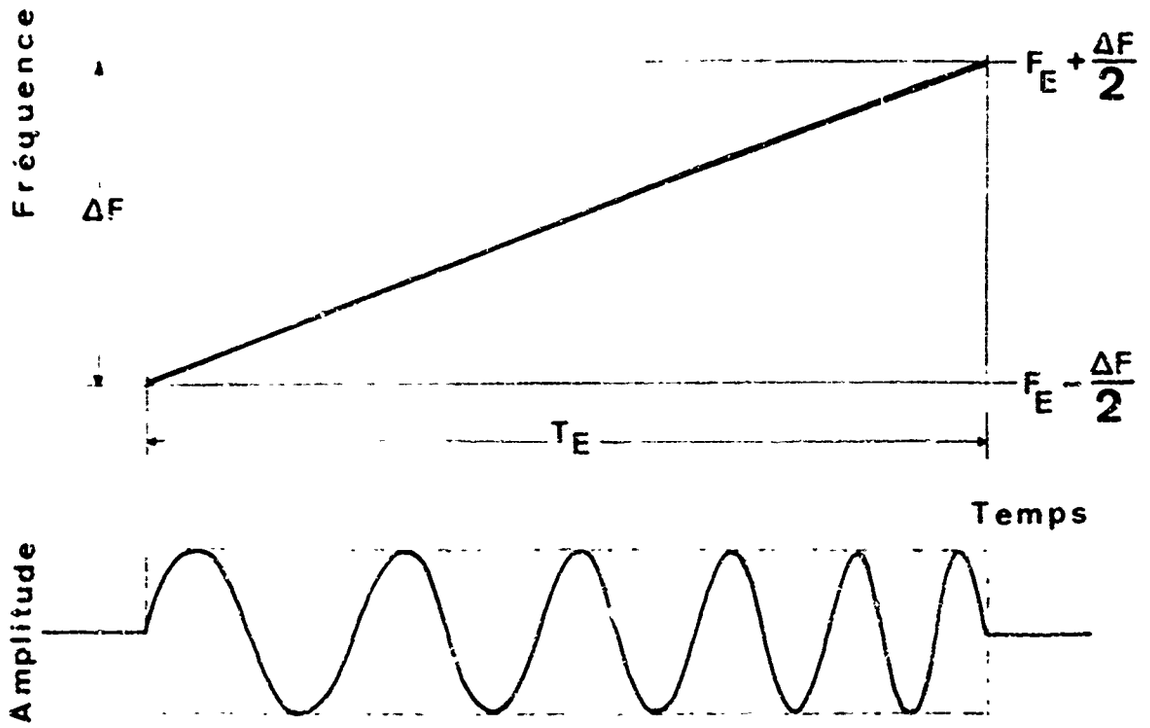


Fig.1 Signal d' Emission

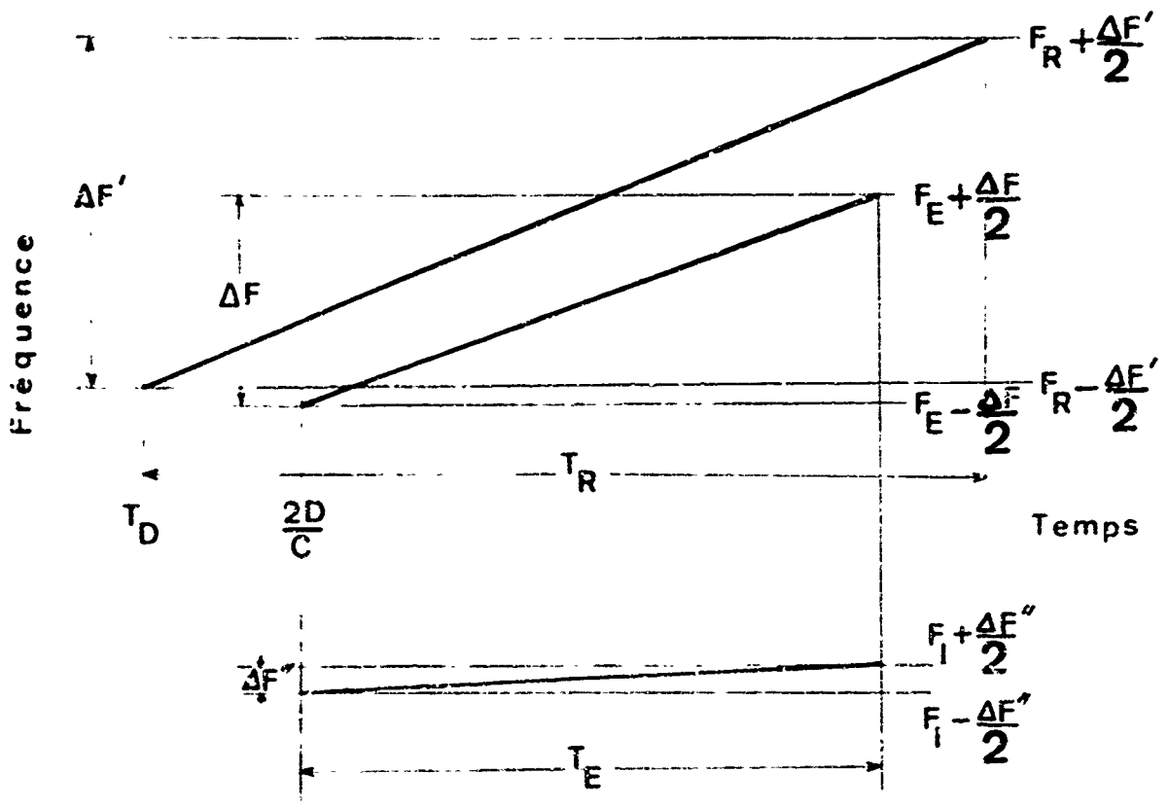
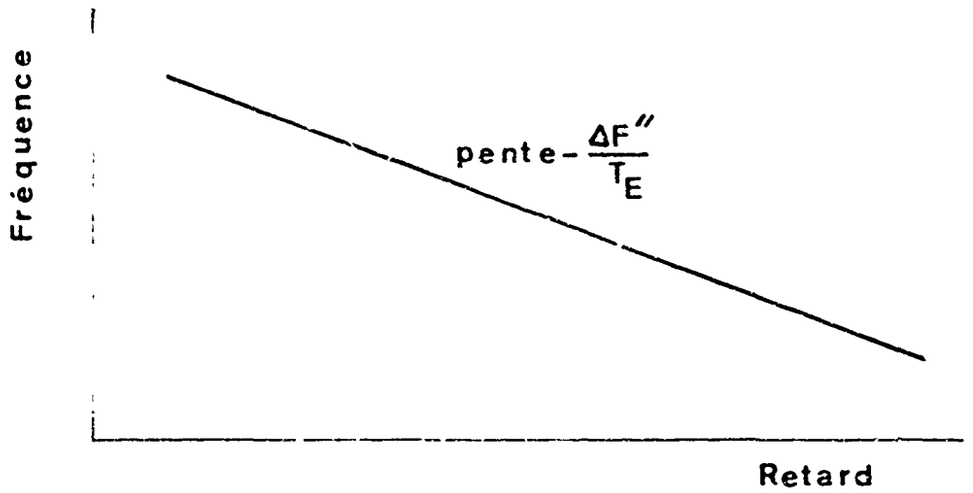


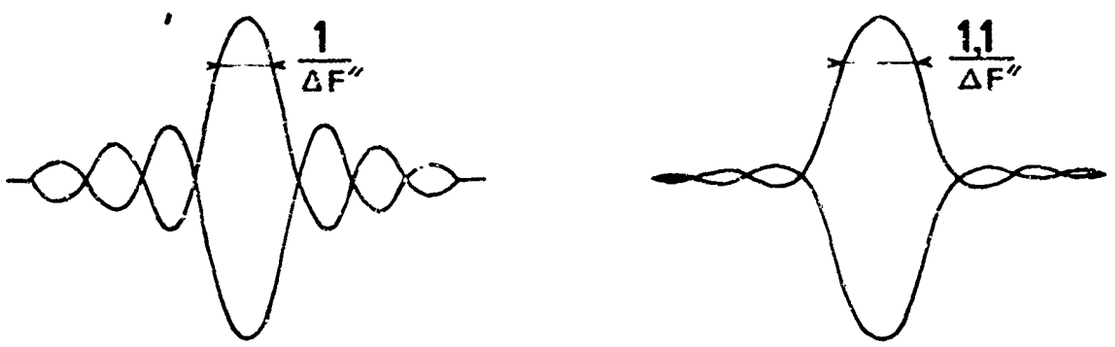
Fig.2 Corrélation imparfaite

Non pondérée Pondérée

a. Impulsion d'entrée



b. Transfert du réseau dispersif



Non pondérée Pondérée

c. Impulsion comprimée

Fig. 3 Compression d'impulsion

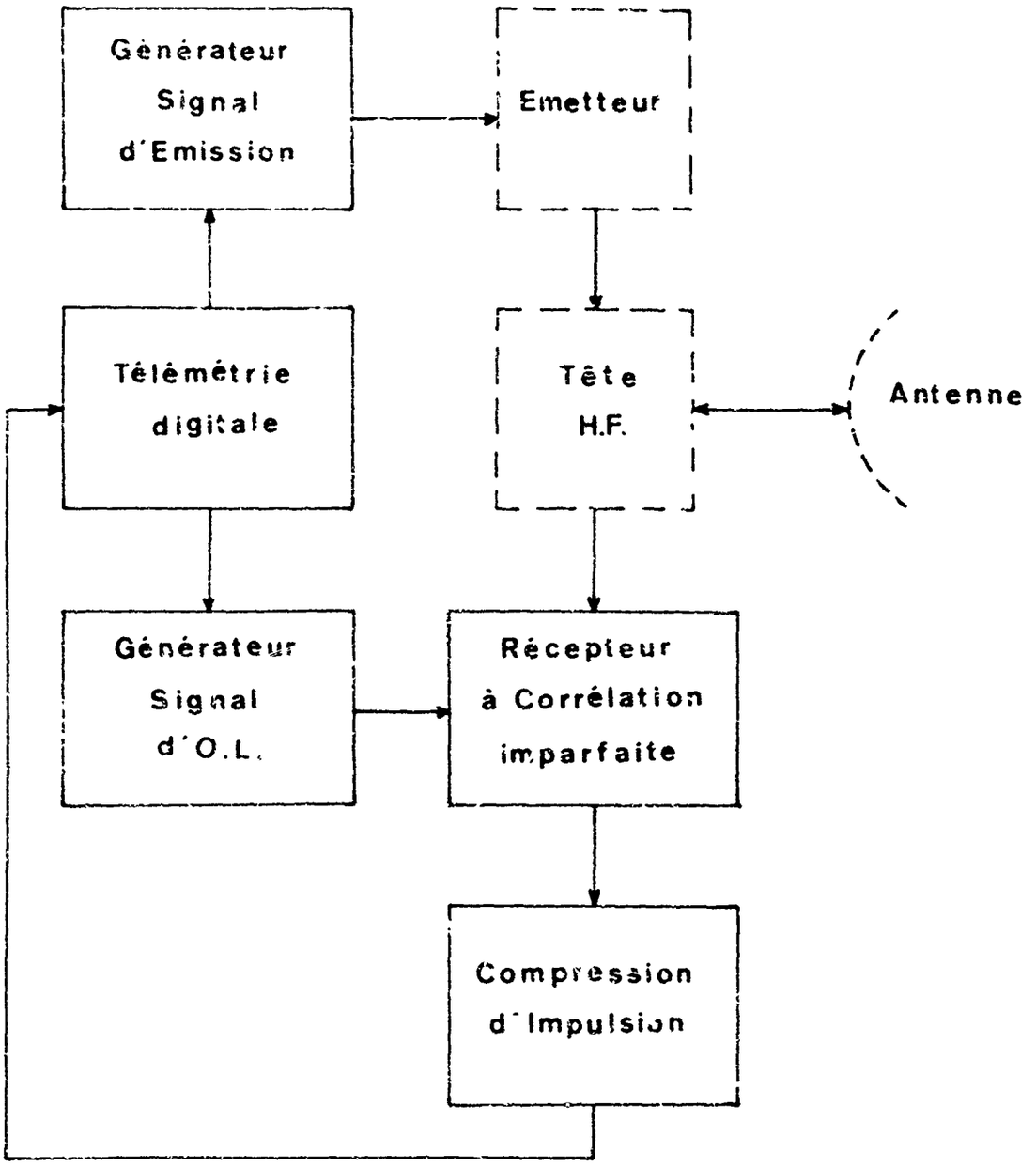


Fig.4 Schéma synoptique d'ensemble

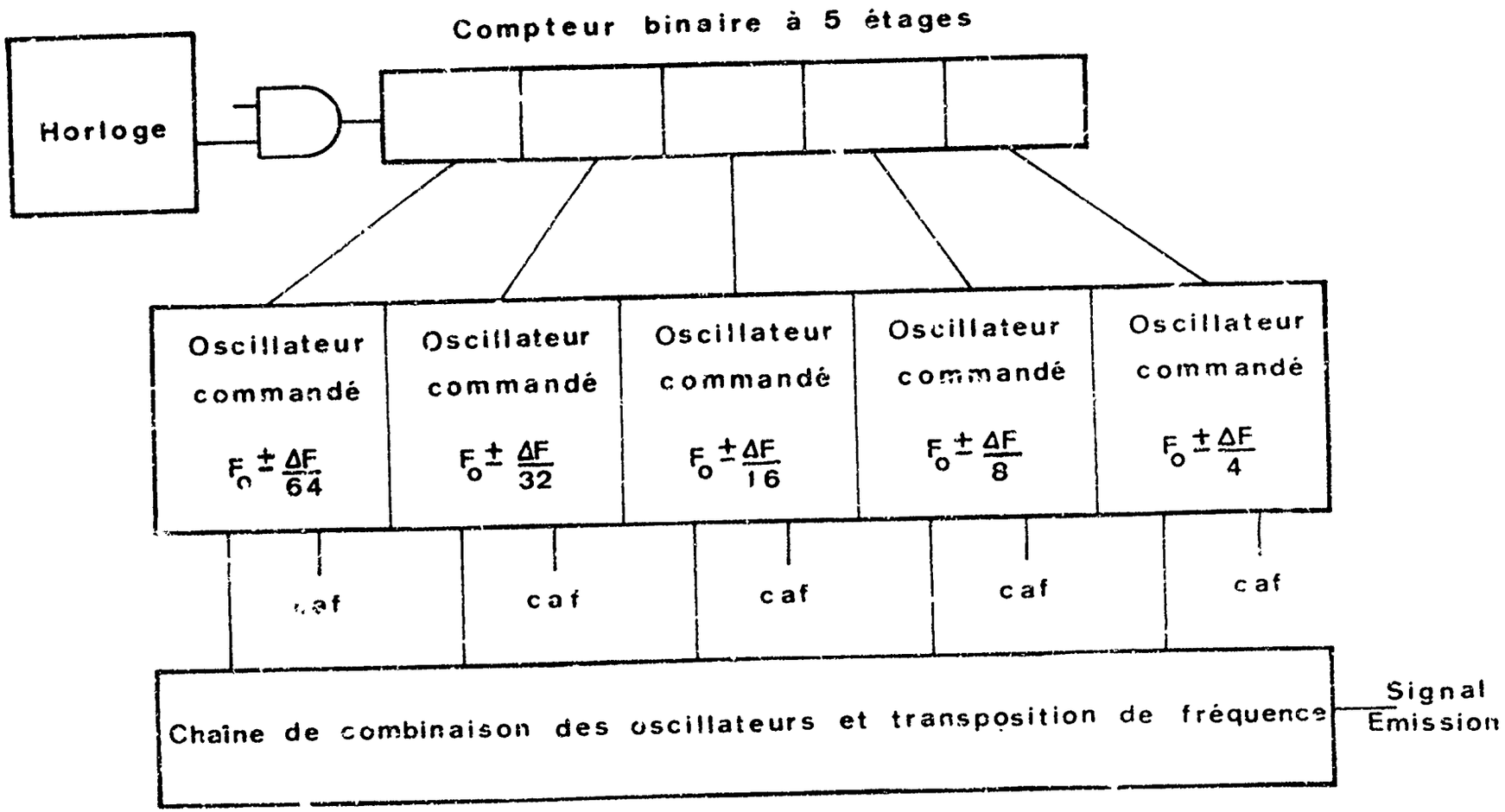


Fig.5 Générateur de Signal d'Emission

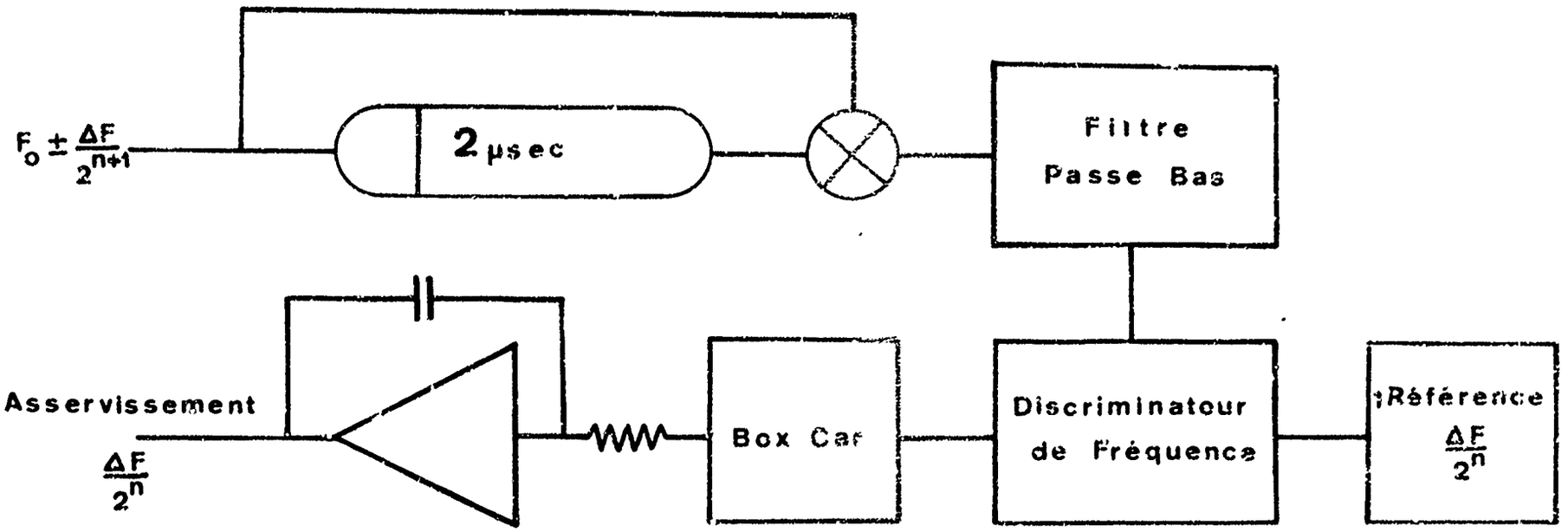


Fig.6 Controle automatique de Fréquence

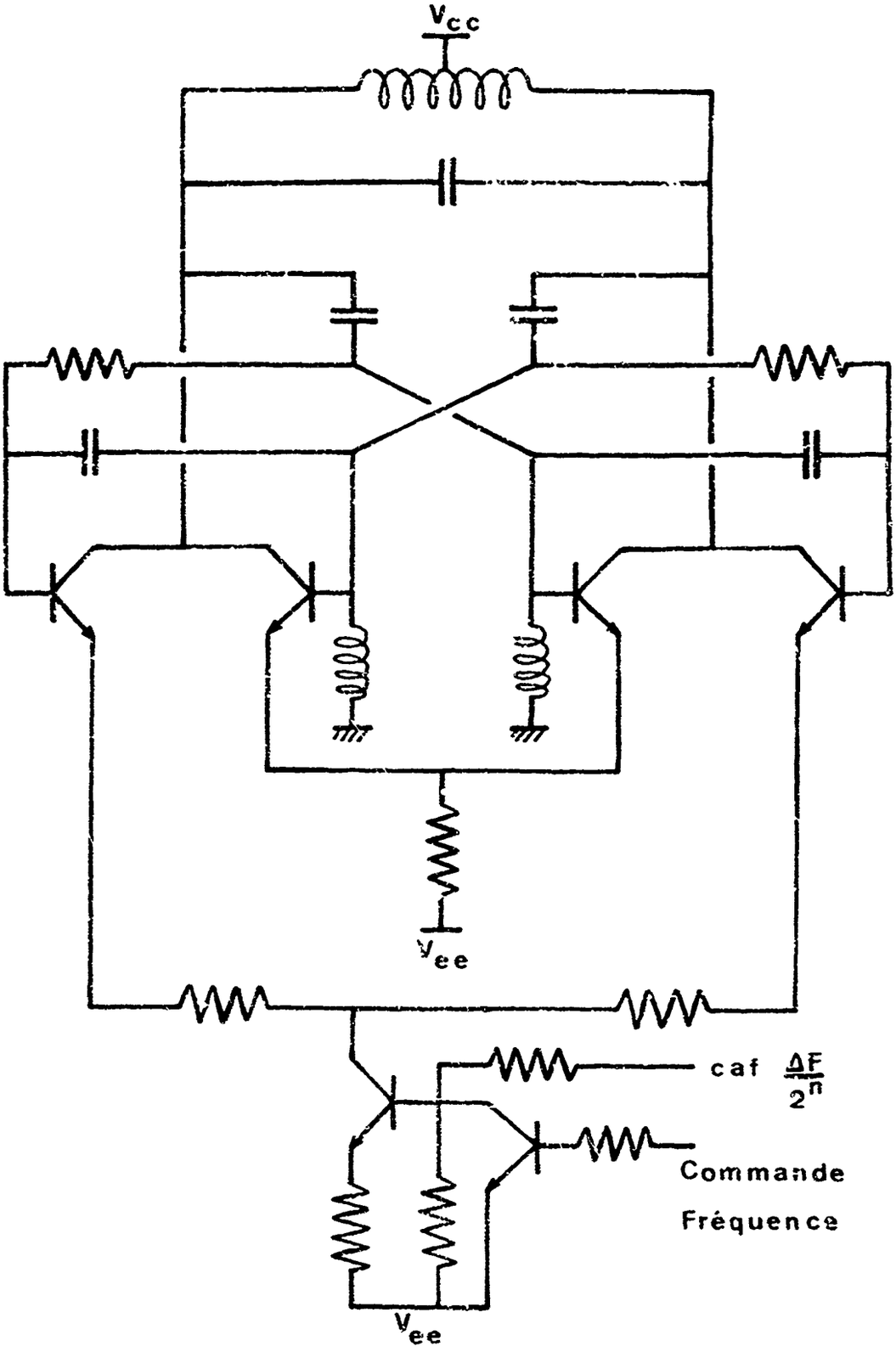


Fig. 7 Oscillateur commandé

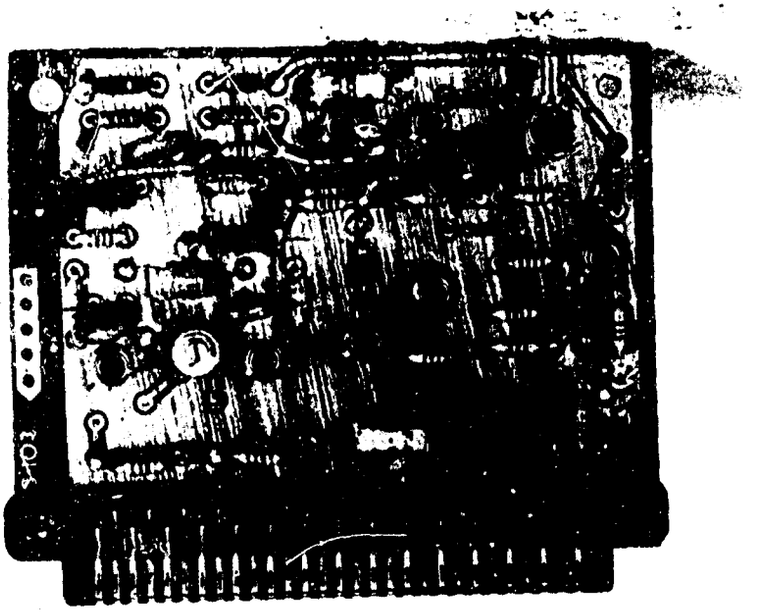


Fig.8 Oscillateur commandé

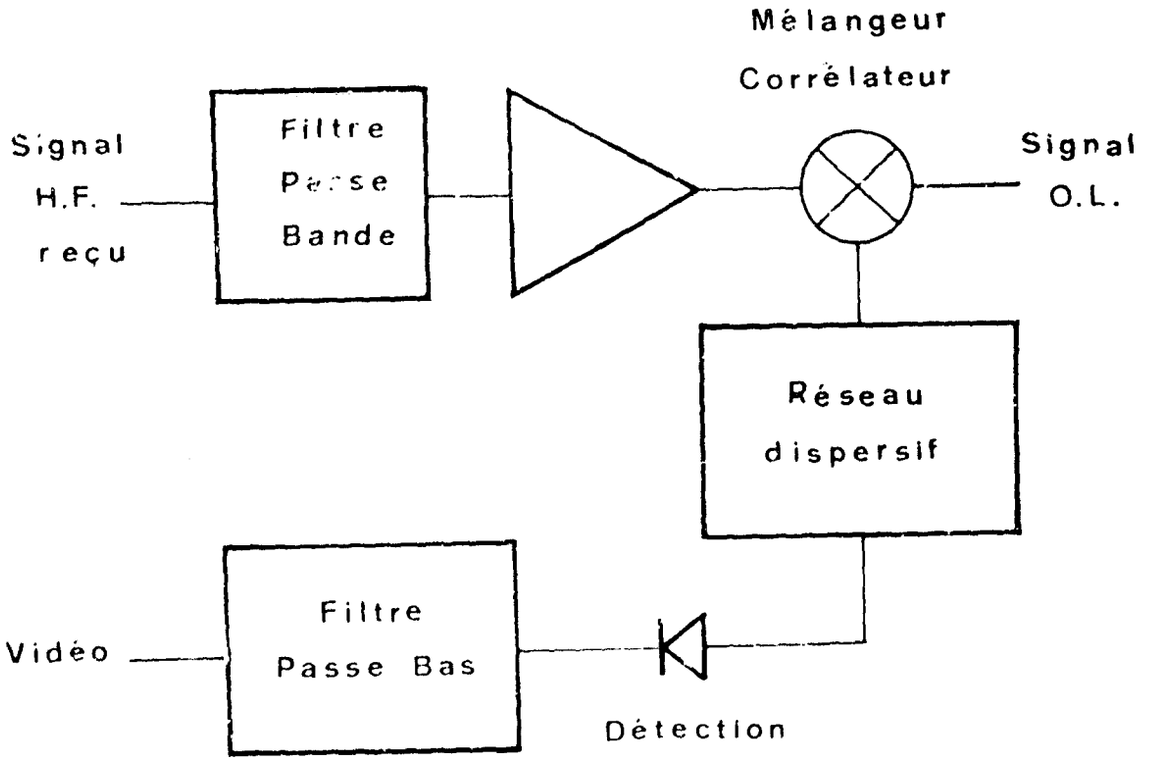
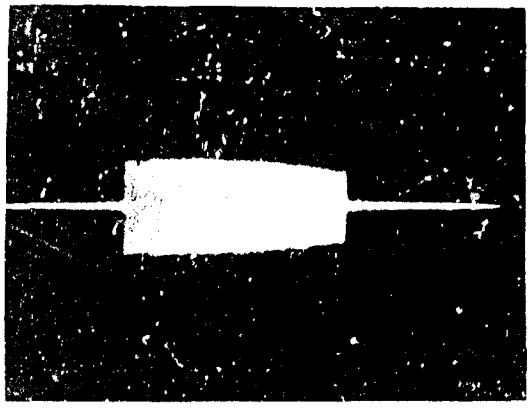
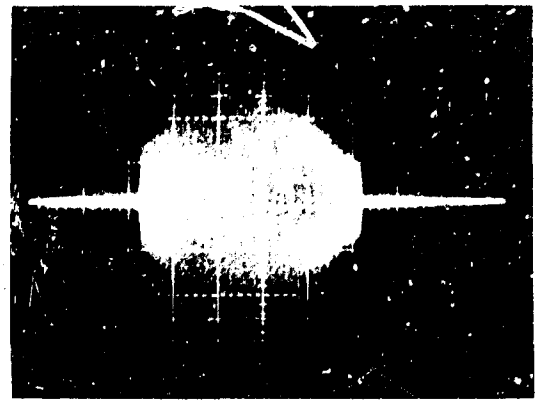


Fig.9 Récepteur à Corrélation imparfaite et Compression d'Impulsion

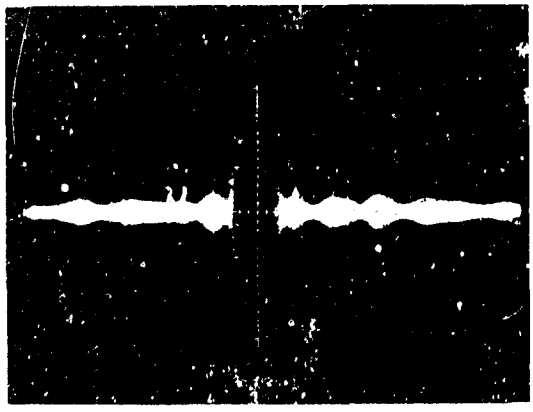


a

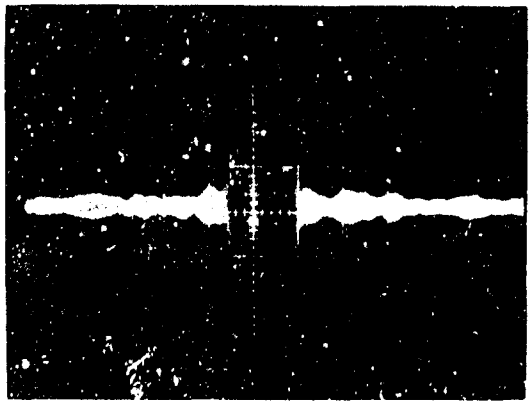


b

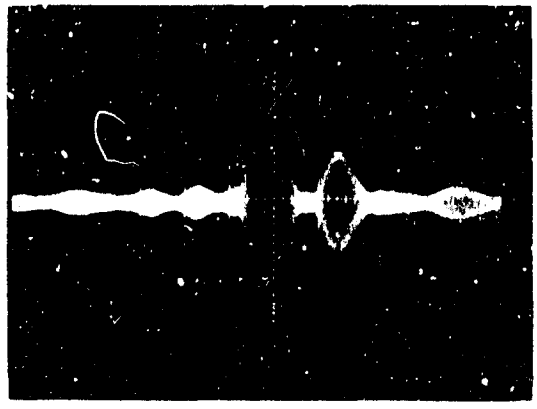
Balayage 2 µsec/cm



c



d



e

Balayage 0,5 µsec/cm

Fig. 10 Oscillogrammes de divers signaux

TABLEAU I PARAMETRES DE LA MAQUETTE

EMISSION

Durée d'impulsion:	T_E	10 µsec
Excursion de Fréquence:	ΔF	40 MHz

OSCILLATION LOCALE ET CORRELATION

Excursion de Fréquence		
Initiale:	$\Delta F'$	52 MHz
Résiduelle:	$\Delta F''$	4 MHz
Durée d'O.L.:	T_{R1} 11,8 µsec	T_{R2} 14,4 µsec

COMPRESSION D IMPULSION

Durée d'impulsion		
Initiale:	T_E	10 µsec
Finale:	$\frac{1}{\Delta F''}$	0,25 µsec
Taux de compression:	$T_E \Delta F''$	

COEFFICIENTS

De corrélation:	K1	1,100	K2	0,903
De modulation:	$\lvert 1-K1 \rvert$	0,100	1-K2	0,097
Grossissement:	G1	11	G2	9,3
Pouvoir séparateur				
En temps:	$\frac{1}{\Delta F}$		25 nsec	
En distance:			3,75 m	

LOW DISTORTION DISPERSIVE NETWORKS FOR WIDEBAND SIGNAL PROCESSING SYSTEMS

by

F. G. Herring, P. M. Krencik and A. J. Axel

Hazeltine Corporation
Little Neck, New York, USA

LOW DISTORTION DISPERSIVE NETWORKS FOR WIDEBAND SIGNAL PROCESSING SYSTEMS

F. G. Herring, P. M. Krencik and A. J. Axel

Hazeltine Corporation
Little Neck, New York, U. S. A.

SUMMARY

This paper describes the synthesis, design, and implementation of wideband, low distortion, lumped and distributed constant, dispersive networks. These networks are important in modern signal processing systems to implement (radar) pulse expansion and compression, instantaneous spectrum analysis (real-time Fourier transformer), antenna beam steering and electronically variable delays.

Three types of dispersive networks, designated as Type I, II and III, are described. Each type was developed to satisfy an operational requirement; together, the three types permit the realization of a wide range of dispersive characteristics. Measured amplitude and phase data are presented for each type, and oscilloscope photographs are shown which demonstrate the network performance in a (pulse compression) system application.

1. INTRODUCTION

This paper is concerned with the realization of lumped and distributed constant dispersive networks. These networks are useful in modern signal processing systems to implement (radar) pulse expansion and compression, instantaneous spectrum analysis (real-time Fourier transformer), antenna time-delay beam steering and electronically variable delays. Three types of dispersive networks, designated as Type I, II and III, are described.

The Type I network was originally developed to implement the dispersive function for high bandwidth-time product systems. The network has subsequently been used in several instrumentation radars for pulse expansion, compression and for instantaneous spectrum analysis. These applications require the standard deviation of the residual phase and amplitude errors to be in the order of 1.0 degrees and 0.2 dB. These errors are commensurate with a -40 dB time sidelobe level.

The network synthesis techniques are described and theoretical error curves presented. The second-order bridged-tee network is described; this network includes a first-order correction for all major parasitic elements to obtain a theoretically all-pass characteristic. This correction is responsible for the low phase distortion obtained. Measured amplitude and phase data are presented for six networks, each having a bandwidth-time product of 264. Fourier analysis of the phase and amplitude errors are shown to predict a paired-echo sidelobe level below 40 dB.

The Type I network is particularly applicable to large, ground based, instrumentation-type radar systems in which network size and initial cost are subservient to electrical performance and reliability. The network is useful for dispersive bandwidths up to about 25 or 30 MHz and bandwidth-time products up to several thousand. Oscilloscope photographs are shown which demonstrate the performance of the Type I network in several pulse compression systems.

The Type II network was developed as part of an in-house IR and D program to satisfy the need for a low cost, minimum-volume dispersive network. It is useful for compression ratios up to several hundred and for dispersive bandwidths up to about 100 MHz. The Type II dispersive network is realized by the application of microelectronic techniques to a lumped-constant first-order bridged-tee. The use of these techniques has resulted in a significant size reduction and an increase in the attainable dispersive bandwidth. The fabrication and alignment processes and methods are described; performance data is given for representative networks.

The Type III network is the microwave distributed constant equivalent of the Type II lumped-constant all-pass network. This network consists of a cascade of shorted, quarter-wave, coupled-transmission lines. The coupled-line cascade is realized in dielectric stripline. Each line in the cascade has identical dimensions (except for length) and a multiplicity of lines are contained on a single printed circuit board. The simplicity of this construction results in a highly reproducible, low cost dispersive network. The device is useful to synthesize dispersive bandwidths up to several gigahertz.

The methods which were developed to synthesize, design and construct the dispersive networks using quarter-wave lines are described. Design examples are given for several fractional bandwidths. These basic designs can be frequency-scaled to obtain bandwidths from 100 to 2000 MHz and cascaded to achieve compression ratios of several hundred.

The three types of networks are subsequently described in detail and measured performance data presented. Prior to this description, however, some useful concepts and techniques concerned with lumped and distributed constant dispersive network synthesis are presented in Part 2 below.

2. DISPERSIVE NETWORK SYNTHESIS CONSIDERATIONS

A cascade of all-pass bridged-tee networks can be used to synthesize any reasonable dispersive characteristic. The number of bridged-tees required is determined by the delay area to be filled, the maximum delay slope required, the permissible ripple of the delay characteristic, and the fractional bandwidth of the network. For most dispersive networks, the delay area (which is proportional to the bandwidth-time product) and the fractional bandwidth of the network are the governing considerations. Specific synthesis procedures are available in the literature [1] [2].

All-pass networks are characterized by a constant (frequency-independent) and real characteristic impedance. This property prescribes that the total time delay of the cascaded network must be equal to the sum of the time delays of each bridged-tee section in the cascade. The time delay contributed by each bridged-tee depends only on the location of the poles and zeros of the network transfer function. Thus, the problem of synthesizing the dispersive characteristic may be viewed as that of locating, in the complex-frequency plane, a suitable distribution of pole-zero pairs. In general, this distribution will have a density which is proportional to the desired time delay.

The pole-zero location in the complex-frequency plane of a band-pass, first-order, bridged-tee network is shown in Fig. 1(b). In addition to conjugate symmetry, the constellation also has symmetry about the $j\omega$-axis for the lossless case. Figure 1 (a) shows the lattice and the equivalent unbalanced form of the lattice (bridged-tee) all-pass network which will realize this pole-zero configuration. The network time delay, which is the derivative of the transfer phase characteristic, is given by:

$$t_d(\omega) = -\frac{\partial \Phi}{\partial \omega} = \frac{2}{\sigma_i}\left[\frac{1}{1+\left(\frac{\omega-\omega_i}{\sigma_i}\right)^2} + \frac{1}{1+\left(\frac{\omega+\omega_i}{\sigma_i}\right)^2}\right]$$

where

σ_i and ω_i are the real and imaginary parts of the root locations in radians per second.

The second term in the expression is the delay contribution of the negative half-plane pole-zero pair and can usually be neglected.

The location of the pole-zero constellation in the s-plane determines the shape of the time delay characteristic. Poles and zeros which are in the vicinity of the $j\omega$-axis exhibit delay characteristics which have steep slopes and high peak values of time delay at $\omega = \omega_i$; conversely, large values of σ_i produce a characteristic which is broad and has a low peak value at ω_i. This effect is illustrated by the curves plotted in Fig. 2.

The maximum value of the time delay function is $2/\sigma_i$. The area under the delay curve is 2π radians and is independent of the pole-zero location. (The area under the time delay versus cyclic frequency curve is, thus, unity.) This relationship is important because it specifies the complexity of the bridged-tee cascade required to synthesize a time response having a prescribed bandwidth-time product. Specifically, for a linear dispersion, the number of bridged-tees required cannot, theoretically, be less than one-half of the bandwidth-time product.

2.1 Selection of a Root Contour

The first step in the synthesis problem is the selection of a root contour in the complex frequency plane for the cascaded network. Once the root contour has been selected, the poles and zeros of each bridged-tee section in the cascade are constrained to points which lie on this contour. For a small sacrifice in flexibility, the root contour method permits an orderly, iterative procedure to be used to determine the network pole-zero locations.

A logical root contour is the constant-Q contour shown in Fig. 1 (c). This contour results in a constant ratio of ω_i to σ_i. The network Q, Q_n, is defined as,

$$Q_n = 2\omega_i/\Delta\omega = \omega_i/\sigma_i$$

where ω_i is the radian frequency at which the delay curve peaks and $\Delta\omega$ is the bandwidth at which the time delay is down by 50% from the peak value. The selection of this contour has the following two advantages:

(1) For constant (frequency-independent) component Q's, the percent error in peak time delay due to the finite component Q is a constant.

(2) The L and C component values vary as 1/f. (This situation may be compared to that

resulting from the selection of a constant sigma (σ_i) contour in which two components remain fixed and two vary as $1/f^2$.)

The selection of the value of Q_n is an important consideration. Generally, a high value is desirable because the dispersive bandwidth of the networks in the cascade is reduced and more of the time-delay area of the tee will be contained in the band of interest. Thus, the network efficiency is increased. However, several limitations exist which restrict the extent to which Q_n may be increased. One limitation is the increase in peak (and rms) time delay error that occurs for an individual tee when the network component quality factor, Q, does not significantly exceed Q_n. Figure 3 is a plot of this error as a function of Q/Q_n for a matched source and load impedance. The error is not a simple inverse function and increases rapidly as Q/Q_n decreases. Figure 3 also shows the increase in the peak attentation of the tee which is accompanied by a corresponding increase in the input VSWR. Another limitation is the impractical component values which result from a high Q_n. From an examination of the equations which determine the component values (Fig. 1(d)), the ratio of C1 to C2 and L2 to L1 is seen to vary appropriately as Q_n^2.

2.2 Network Efficiency

A useful parameter in dispersive network synthesis is the network efficiency, η. This efficiency is the ratio (expressed in percent) of the minimum area (delay pedestal equals zero) under the desired delay function, T(f), to the number of bridged-tee networks, m, needed to synthesize T(f). Thus, if T(f) is defined in the interval f_1 to f_2, then:

$$\eta \text{ (in percent)} = \left(\int_{f_1}^{f_2} T(f)\,df/m \right) 100$$

The network efficiency can only approach 100 percent in the limit, because a single bridged-tee network can only contribute unit area when the delay curve is integrated from zero to infinite frequency.

The network efficiency which can be obtained in practice is a function of the fractional bandwidth at which the synthesis is performed and the Q_n selected. The efficiency increases with an increase in each of these factors. The limitations on Q_n have previously been defined. The maximum value of fractional bandwidth is limited to about unity. The limitation is imposed by the requirement for subsequent translation to a lower percent bandwidth in the transmitter or receiver signal processor. This translation requires the selection of the desired sideband by a filter, and the order of the filter, for a prescribed undesired sideband rejection, is determined by the fractional bandwidth of the spectrum prior to translation. These filters must be delay equalized, and the complexity of the equalization becomes excessive for the high order filters which accompany a high fractional bandwidth.

Typical values for network efficiency range from 30 to 60 percent.

2.3 Dispersive Subsection

A representative dispersive characteristic is shown in Fig. 4. The time delay versus frequency curve is characterized by the dispersive bandwidth, B, the dispersion, T, the network center frequency, f_o, and the delay pedestal, t_p. The bandwidth-time product of the network is BT; this product may range from 20 to several thousand with values of several hundred being particularly common. The minimum area beneath the curve (for $t_p = o$) is $0.5 \cdot BT$.

Since the bandwidth-time product of the network is usually large, it is expedient to synthesize a subsection having a lower BT product (but retaining the required bandwidth) and to then achieve the total dispersion by a cascade of these subsections. The BT product of the subsection should be in the range of 15 to 30 in order to secure the following advantages:

(1) The theoretical inter-tee time delay ripple will be negligible even when many subsections are cascaded.

(2) Some types of individual-tee time delay errors will have a tendency to cancel. In particular, those errors which are not associated with parasitic elements have a zero mean, and the major contribution of the error occurs within a dispersive bandwidth, $\Delta\omega$. Thus, if the tee tuning frequencies, ω_i, are closely spaced with respect to $\Delta\omega$, similar type errors will tend to cancel. This closer spacing is obtained as the BT product of the subsection increases.

2.4 Network Losses

Dissipative losses occur in the dispersive networks due to the finite component Q (primarily the coil) of the bridged-tee circuit. The dB magnitude of this loss is proportional to the frequency, the time delay and the component Q. A good approximation to the dB loss in a dispersive network at a frequency, f_i, for which a time delay, t_i, exists is:

$$\text{Loss in dB} = 27.3 \cdot f_i \cdot t_i / Q_i$$

where Q_1 is the shunt or series component Q (assumed equal) at f_1.

Both the differential and absolute losses are important. Differential losses can usually be compensated by constant resistance amplitude equalizers which, in turn, may require phase equalization. High absolute losses result in the dispersive network appearing as a band-reject filter. This effect occurs because the time delay, and thus the loss, rapidly goes to zero outside of the band of interest. Additional phase-equalized, band-pass, filters may be required to recover the signal-to (out-of-band)-noise ratio.

The dissipative losses can be minimized by: (1) synthesizing the network at the highest practical fractional bandwidth, (2) synthesizing the delay characteristic with the peak time delay at the low-frequency end of the band, and (3) selecting high-Q components.

3. TYPE I NETWORK

The Type I network is comprised of a cascade of lumped constant second-order bridged-tee networks. This second-order network was originally developed under Air Force contract AF30 (602)-2106 with the Rome Air Development Center, Rome, N.Y., to realize pulse compression networks having near theoretical performance for bandwidth-time products up to several thousand. It is suitable for dispersive bandwidths up to about 30 MHz.

The first-order network (previously shown in Figure 1(a)) is useful to implement moderate compression ratio (up to several hundred) pulse compression systems. This network is less satisfactory to implement high compression ratio (several thousand) systems for several reasons. The first is the deleterious effect of the stray (parasitic) parameters. The major parasitics consists of the stray capacity across the two inductances (L1 and L2) and the lead inductance of the two series capacitors (C1). These parasitics cause perturbations in the time delay function; additional errors will arise when this network is cascaded, because the input impedance is not a constant resistance and reflections will occur. These effects are aggravated in the synthesis of high compression ratio systems for the following two reasons:

(1) The permissible network time delay error for a prescribed distortion-echo level is a function only of the system bandwidth. The delay error, expressed as a percent of the absolute delay, decreases directly as the compression ratio (or bandwidth-time product) increases.

(2) The network absolute and differential losses are proportional to the peak time delay (and thus compression ratio) and to frequency. In order to preclude large absolute and differential losses, the designer is forced to increase the fractional bandwidth of the network. High fractional bandwidths accentuate the deleterious effects of parasitic elements.

3.1 Second-order (Compensated) Bridged-tee

The second-order bridged-tee network and pole-zero plot is shown in Fig. 5. The inclusion of three compensating inductances, designated as L2 and L4, correct for the stray capacities, C1 and C3, and the lead inductance of C2 and C4. This compensation restores the constant resistance characteristic of the network and eliminates the time delay error due to these parasitic sources. The compensation introduces an additional pole-zero pair which is usually removed from the frequency band of interest. The position of this pole-zero pair is prescribed by the parasitic elements and its effect on the in-band delay, although small, is known and may be included in the dispersive network synthesis.

The values of the compensating inductances depend only on the network impedance level and the stray coil capacity. The values of the series compensating inductance (L2) and the shunt compensating inductance (L4) are given by:

$$L2 = (C_{3/2})R_o^2 \text{ henries}$$

$$L4 = (C_1\text{-}C3/4)R_o^2 \text{ henries}$$

If the total capacity across each coil can be designed or adjusted to some constant value, then the compensating inductances will be the same for all tees in the cascade. If the primary pole locations (ω_1 and σ_1) are specified and the values for C1, C3 and R_o are known, then the remaining circuit elements (L1, C2, C4, and L3) and the location of the secondary pole-zero pair are determined.

3.2 Implementation of Second-Order Tee

The second-order bridged-tee schematic is given in Fig. 6. Element values are given in the accompanying table for bridged-tees tuned to 7, 18 and 27 MHz. A trimmer capacitor is used to set the total capacity in parallel with each coil to 6.0 $\mu\mu$f. Thus, C1, C3 and L2 and L4 of Fig. 6 are the same for all tees in the cascade.

The impedance level selected for the network is 93 ohms. This impedance is somewhat below the impedance for optimum coil Q. However, the effect of stray capacity to ground, between and within the bridged-tee network, is reduced. In addition, 93 ohms is a standard co-axial cable impedance and

the interconnection of the chassis which contain the tees is simplified. (Note: Higher impedance levels are satisfactory at lower frequencies.) The bridged-tees are interconnected by a plug-in connecting coil. This coil, together with the bridged-tee input and output stray capacity, forms a pi section that is matched to the tee impedance. The value of the loading coil was determined empirically by viewing the swept input impedance of a large number of cascaded tees and selecting a coil value which resulted in a minimum input VSWR.

3.2.1 Series and Shunt Coil

The performance of the compensated bridged-tee network is largely determined by the characteristics of the series (L1) and shunt (L3) coils. The coil Q determines the network dissipative losses and effects the time delay error. The long term and temperature stability of the network are also primarily determined by the coil. The coil developed is shown in Fig. 7. The core material is Siferrit K-12 which is manufactured by Siemens-Halske in Germany. The intrinsic core permeability is 20 and the temperature coefficient is 0 to 200 ppm/^{0}C. The core is gapped to an effective permeability of about 3 with a resultant reduction in the temperature coefficient to 0 to 30ppm/^{0}C. All windings are single layer; optimum Q is obtained with a space between adjacent windings equal to the wire size. The wire size is made as large as possible within these constraints. Q values over 300 were obtained up to 20 MHz.

The coil assembly is completely shielded by a drawn metal case. The assembly also contains a piston trimmer capacitor which is used to adjust the coil shunt capacity to the design value of 6.0 $\mu\mu$f.

3.2.2 Construction Details

A photograph of a representative bridged-tee assembly is shown in Fig. 8. The tee is 4" long, 3" wide and 2 1/2" deep. The shunt and series coils are mounted on the top side and the coil adjusting slug is accessible from this side. The two terminals of the coil extend through a cut-out in the support board and are connected through supporting legs to a Y board assembly.

The Y board assembly is a glass-epoxy printed circuit board that contains the remaining components of the bridged-tee network. These components are the series (C2) and shunt (C4) silvered-mica capacitors and the three compensating inductances (L2 and L4). The Y board is a convenient sub-assembly that is fabricated and aligned prior to its assembly to the main board.

3.3 Error Analysis

An error analysis was performed on the second-order bridged-tee network. This analysis determined the time delay and input VSWR caused by the finite circuit Q and deviations of the primary and compensating elements from the theoretical value. The time delay errors were calculated assuming the tee to be driven by a resistive source impedance equal to 93 ohms. In addition, it was assumed that the series and shunt inductances were tuned (after the circuit perturbation had been introduced) in a manner identical to that prescribed by the normal production-line tuning procedure. This tuning procedure results in the series (L1) and shunt (L3) coils being detuned from the design value to compensate for the original error. The procedure minimizes the input VSWR and time delay error. Figure 9 shows the peak time delay error for a tee tuned to 18 MHz versus the percent error for C1, C3, L2 and L4. L1 and L3 have been tuned to minimize the input VSWR while simultaneously obtaining the correct phase shift at the resonant frequency of the tee (18 MHz). These errors may be compared to the finite Q error of about 60 picoseconds for a shunt and series coil Q of 250.

The effect of unequal coil Q on the time delay error is shown in Fig. 10. The error curve has even symmetry and a zero mean error. Note the large increase in the error for the unequal Q case. To avoid this large error, the shunt and series Q are set equal during the production alignment of the tee by selecting the Q-equalizing resistor (see Fig. 6) which shunts L3. The equal Q condition is recognized by observing the swept input impedance of the completed tee assembly. The input VSWR has the same shape as the delay error curve and the VSWR is a minimum when the Q's are equal

3.4 Applications and Measured Electrical Performance

The second-order tee was used initially to implement two PC systems having compression ratios over 1000 and bandwidths of 0.6 and 33 MHz. This original experimental work was sponsored by the Rome Air Development Center, Rome, N. Y. Subsequently, this network was used to implement the pulse compression subsystems for the ALTAIR and ALCOR instrumentation radars on Kwajalein Atoll and the Ramsite instrumentation radar at the White Sands Missile Range. The ALCOR subsystem was produced under subcontract to Lincoln Laboratory and production records available from this program are indicative of the control of the dispersive characteristic which can be achieved.

A basic dispersive subsection for the ALCOR system consists of 48 bridged-tees. This network had a dispersion of 3 μs over a bandwidth of 17.6 MHz. The BT product of a subsection is 52.8 and the network (synthesis) efficiency is 55%. The network Q_n is 10 and the delay pedestal is about 20% of the dispersion. The network has unity fractional bandwidth.

A total of thirty-two subsections were produced and the phase characteristic of each network was measured. This phase characteristic was fitted to that quadratic curve which resulted in the least mean square error fit over the bandwidth of 17.6 MHz. This quadratic defined the network dispersion and the rms phase error of the network.

The rms phase error and the dispersion for the 32 networks are shown in Table I of Fig. 11. The average rms error is about 1.0^{0}; the average dispersion is 3.006 μs. Table II in Fig. 11 shows the results

of a Fourier analysis of the residual phase error of a typical network. The rms error is 0.83°. The phase error shown in the fourth column can be converted by the paired-echo theory of Wheeler [3] to an equivalent error-echo sidelobe level. This sidelobe level is shown in column five and the largest paired echo sidelobe results from the sixth harmonic term which is 46 dB down.

The 3 μs networks were subsequently cascaded in groups of five to obtain six separate 15 μs networks. The overall phase error of the network was then measured and a 15-tee corrector designed to provide the final residual phase correction. The data is summarized in Table III of Fig. 11. The rms phase error for five of the six networks is less than 1.0°.

The six networks were subsequently delivered to Lincoln and thence to the radar site at Kwajalein. On site phase measurements on the six networks were performed approximately one year after the initial measurements. This data is also shown in Table III of Fig. 11. In the worse case, the rms error had increased to 1.42°.

A representative phase error curve for a 15 μs, 17.6 MHz network (BT = 264) is shown in Fig. 12(a). This network is #3 of Table III and has a rms error of 0.95 degrees. A Fourier analysis of this error was made and the largest equivalent error-echo sidelobe level is 47 dB down. The amplitude response of this network is shown in Fig. 12(b). This degree of amplitude correction was obtained by appropriately "Q-loading" the 15-tee final phase corrector. The error sidelobe level associated with this degree of amplitude distortion is commensurate with the network phase errors.

In addition to the amplitude equalizers, the 15 μs network contained a sharp, phase-equalized band-pass filter. The impulse response of the dispersive network and this band truncating filter is shown in Fig. 12(c). The 6 db expanded pulse width is 15 μs.

3.4.1 Oscilloscope Photographs

Figure 13 shows some oscilloscope photographs from another pulse compression system implemented with the Type I network. The expanded and compressed pulses are shown for a system having a BT product of 210 and a bandwidth of 7 MHz. A transversal equalizer was used to provide the residual amplitude and phase correction needed to obtain the -43 dB sidelobes shown in Fig. 13(c).

4. TYPE II DISPERSIVE NETWORK

The Type II dispersive network is realized as a cascade of miniaturized, first-order bridged-tees. This network was developed, as part of an in-house IR and D program, to fulfill a need for small, lightweight, pulse compression systems intended for mobile or airborne use. In addition, the following secondary objectives were generally achieved:

(1) Retention, to the extent possible, of the electrical performance of the Type I network.

(2) Adaptability to mass production at low unit cost.

4.1 Electrical and Mechanical Design

A schematic of the first-order network is shown in Fig. 14(a) and an exploded view showing the tee mechanical configuration is shown in Fig. 14(b). The capacitor micro-package consists of three chip capacitors, a lead frame and a ceramic die. A ceramic sheet 0.020 in. thick contains 63 0.375 x 0.25 in. dice with vacuum evaporated copper film conductor patterns. The lead frame (which is an etched 0.010 in. thick copper structure) and the capacitors are soldered to the dice. A molded shell, within which one such chip is encapsulated, comprises the support structure for the individual tee.

The capacitors are NPO ceramic chip and are reflow-soldered, 189 at a time, before the substrate is separated into dice. The lead frame is attached to the die perimeter and functions both as the external electrical connection and the mechanical support for the tee assembly. The lead frame contains tabs (not shown in the exploded view) to accept the inductor wires. The capacitors are then trimmed to the desired value by air abrasion. A capacitor trimming fixture was designed and constructed for attachment to the S. S. White, Model G, Industrial Airbrasive Unit. A microscope is used to view each capacitor under the nozzle in order that the operator can monitor the mechanical erosion. The capacitance value is monitored by a General Radio 1680 Automatic Capacitance bridge.

The cleaned capacitor substrate is encapsulated in a molded shell with a low loss, low dielectric, silicon rubber potting compound commercially available as Emerson and Cumming Eccosil 4640. The use of this material for the capacitor encapsulation results in negligible change in measured capacities and no apparent effect upon capacitor Q.

The coils are attached to the sides of the completed capacitor package with a high strength silicon rubber adhesive. The coils are adjustable pot core inductors (11 x 7 mm) for frequencies up to about 30 MHz. From 30 MHz to 150 MHz these coils are replaced by an air coil having a powdered iron slug.

A cascade of 52 tees is shown in the photograph of Fig. 15. Each chassis contains 13 tees and the cover of one chassis has been removed to show the internal configuration. The individual chassis is 1" x 1 1/4" x 9 3/2" and the resultant volume per tee is 0.9 cubic inches. For network efficiencies of about 50%, the size of the compression or expansion can be quickly estimated since the volume in cubic inches will be approximately equal to the compression ratio (or bandwidth-time product).

The chassis developed is a compact rigid structure which holds 13 bridged-tees and has recessed

co-axial connectors. The recessed connector feature eliminates interconnecting cables by permitting the chassis to be plugged together as shown in Fig. 15.

4.2 Electrical Performance

Three different dispersive subsections were designed, constructed and measured. Each network had 26 tees and a BT product of about 26. The networks differed in bandwidth and center frequency. The design and measured parameters are summarized in Table I of Fig. 16. Table II of Fig. 16 shows the equivalent error-echo sidelobe level resulting from a Fourier analysis of the three sets of measured phase data.

4.2.1 Network #1

This network had a bandwidth of 17.6 MHz and a dispersion of 1.5 μs. It was designed at unity fractional bandwidth and can, thus, be compared with the 17.6 MHz, Type I network previously described. The theoretical RMS phase error for this design is 0.09 degrees. The measured phase error was 0.96 and the measured dispersion was in error by 2.3 ns. The largest paired-echo sidelobe is 49 dB down. The unit has an equalized loss of 6 dB.

4.2.2 Network # 2

The network has a 60 MHz dispersive bandwidth centered at 60 MHz. The network was designed at a 50 ohm impedance level to reduce the effects of stray parameters and to be compatible with test equipment (used for alignment) in this frequency range. The initial tuning of the network resulted in an RMS phase error of 8.3°. A subsequent retuning reduced the error to 2.2°. It is probable that this retuning is equivalent to a calibration of the alignment fixture. If this is true then the anticipated error in production would be about 2 degrees.

4.2.3 Network # 3

This network was designed for a dispersive bandwidth of 60 MHz centered at 120 MHz. This lower fractional bandwidth (0.5) resulted in a decrease in network efficiency and resultant BT product of 20. With the exception of the decreased network efficiency, the performance was generally comparable to the previous network. The RMS phase error after the initial tuning of the network was 4.4°; a subsequent retuning reduced the error to 2.7°.

5. TYPE III NETWORK

For pulse compression systems with bandwidths above 100 MHz, the usefulness of the lumped-constant, bridged-tee network rapidly decreases. This is particularly true for high compression ratio systems because, while the network alignment and component tolerances become more difficult to achieve, the mathematical model (a lumped constant circuit) of the network is less valid.

To replace the bridged-tee circuit, for high compression ratio, wide-band systems, a network having the following properties is desired:

(1) A distributed-constant network.

(2) A theoretically all-pass network. (Narrow-band approximations are not sufficient, because a high compression ratio forces the designer to use high fractional bandwidths to avoid excessive dissipative loss.)

(3) A dispersive characteristic which results is a reasonable network efficiency.

(4) A high Q network. (A trade-off should also exist between the physical size of the network and the network Q. This permits the designer to choose a Q commensurate with the compression ratio and realize a smaller size.)

(5) Compact, rugged and inexpensive to fabricate.

A network which has these properties is a quarter-wave, all-pass, coupled transmission line. This line is the distributed constant equivalent of the lumped-constant, all-pass, bridged-tee network.

5.1 Theory of All-pass, Quarter-wave Coupled Lines

The basic microwave structure consists of two closely spaced parallel conductors (lines) located either between ground planes or above one ground plane, and shorted at one end. The lines are usually uniform and symmetrical with respect to the ground plane. If a dielectric is used, it should be homogeneous throughout the structure. The line is driven and terminated (loaded) with respect to ground at the conductor ends which are not shorted. The dispersion peaks at the frequency at which the line becomes a quarter of a wavelength and thereafter at odd multiples of this frequency.

A schematic representation is shown in Fig. 17(a) and a practical implementation is shown in Fig. 17(b). The theory of coupled lines is described in detail in the literature [4] [5] [6] and the following is presented without proof.

For the structure under consideration, the characteristic impedance of the line to ground Z_o, is the geometric mean of the odd, Z_{oo}, and even, Z_{oe}, mode impedance,

$$Z_{oo} = \sqrt{Z_{oo} \cdot Z_{oe}}$$

Two useful parameters of the line are the coupling ratio, ρ,

$$\rho = Z_{oe} / Z_{oo}$$

and the coupling factor, k,

$$k = \frac{\rho-1}{\rho+1} = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}}$$

From the relations that define them, Z_{oo} and Z_{oe} are independent quantities determined by the geometry. Thus, the characteristic impedance of the line and the coupling ratio can be independently specified.

5.1.1 Transfer Function

The transfer function, the location of the singularities in the complex frequency plane, and the expression for the time delay for all-pass, coupled lossless lines have been derived by Steenaart [6] and is presented below:

The transfer function of the line is

$$\frac{V_{out}}{V_{in}}(\theta) = \frac{\sqrt{\rho} - j\tan\theta}{\sqrt{\rho} + j\tan\theta}$$

where $j\theta = j\frac{\pi}{2}(\omega/\omega_o)$

l = the length of the line

$\theta=\beta l$ = the electrical length of a uniform line of length l and phase constant β

ω_o = the radian frequency at which the line length is equal to a quarter wave-length and $\theta=\pi/2$

By introducing and substituting the complex frequency variable,

$$s = \sigma + j\theta = \frac{\pi}{2\omega_o} \cdot (\alpha + j\omega)$$

the singularities in the s-plane can be shown to exist at:

$$\omega = \omega_o (1 + 2k), \quad k = 0, 1, 2\ldots$$

$$\sigma = \omega_o \left(\pm \frac{1}{\pi} \ln \frac{\sqrt{\rho}+1}{\sqrt{\rho}-1}\right)$$

The singularities are periodic and have a period of $2\omega_o$. The pole-zero plot for a coupled line is shown in Fig. 17 (c). ω_o is determined by the physical length of the line.

For a given ω_o, α is determined by the coupling ratio ρ. As ρ increases, by increasing the coupling between the lines, the pole-zero pairs move closer to the $j\omega$-axis and the dispersive bandwidth associated with a pole-zero pair decreases with a corresponding increase in the peak time delay. For efficient dispersive network synthesis, ρ should be as high as possible.

The phase shift, Φ, and time delay $T_D(\omega)$ are given by:

$$\Phi(\omega) = 2\tan^{-1}\left[\frac{\tan\frac{\pi}{2}\left(\frac{\omega}{\omega_o}\right)}{\sqrt{\rho}}\right]$$

$$T_D(\omega) = \frac{\pi}{\omega_o}\left[\frac{\sqrt{\rho}}{\rho\cos^2\frac{\pi}{2}\left(\frac{\omega}{\omega_o}\right) + \sin^2\frac{\pi}{2}\left(\frac{\omega}{\omega_o}\right)}\right]$$

For high values of ρ, the time delay is nearly identical to that for a lumped constant, first-order, all-pass, bridged-tee network. This similarity is shown in Fig. 18 which shows the phase and time delay in the interval from 0 to $2\omega_o$ for a practical line which has a ρ equal to 15. Also plotted is the time delay which would occur from a single pole-zero pair which characterizes a lumped-constant, first order, network. This single pair is located at ω_o and has the real part of the pole location (σ_1) adjusted to give the same peak time delay as the periodic curve. The curves are alike in the region where the time delay is significant.

5.2 Dispersive Network Synthesis

A dispersive characteristic can be synthesized by a cascade of all-pass, quarter wave, coupled transmission lines. In general, the delay contribution of the first pole-zero pair (which occurs at ω_o) of each line in the cascade is used to realize the desired time-delay function. The desired delay function is usually synthesized within the interval from ω_o' to $2\omega_o'$, where ω_o' is the quarter-wave frequency of the longest line in the cascade. Thus, we avoid most of the delay contribution from all the remaining (periodic) pole-zero pairs of all the lines in the cascade. The fractional bandwidth must be less than one and, in practice, the fractional bandwidth is about 0.5 for an efficient design. The time delay function is synthesized with the maximum time delay at the lowest frequency. This reduces the absolute and differential dissipative losses of the network.

5.3 Network Efficiency

Network efficiency may be defined for quarter-wave lines in the same manner as for the lumped constant networks. For practical values of ρ ($\rho < 15$), the efficiency will, in general, be lower by almost a factor of 2.

A network Q, Q_n, is a useful parameter. It is defined in the same manner as for the lumped constant circuit and is the ratio of the quarter-wave radian frequency, ω_o, to the real part of the complex

frequency variable, α Q_n is a function of the coupling ratio as shown below:

$$Q_n = \omega_o/\alpha = \pi/\ln\left(\sqrt{\rho}+1/\sqrt{\rho}-1\right)$$

The following simplified expression for Q_n may be used with less than 10% error, when $\rho > 4$:

$$Q_n = \frac{\pi}{2}\cdot\sqrt{\rho} = 1.57\sqrt{\rho}$$

Figure 19 shows the network efficiency as a function of several network parameters. Figure 19 (a) plots efficiency versus fractional bandwidth for several values of the coupling ratio. Figure 19 (b) plots the efficiency versus fractional bandwidth for two networks, which have different time-bandwidth products. The plot shows that the efficiency is almost independent of the BT product for BT > 10. Figure 19 (c) is for a fractional bandwidth of 0.5 and shows the variation in efficiency over a practical range of coupling factors. The curves plotted in Fig. 19 were obtained by computer-synthesizing networks to obtain the data points. The networks were synthesized for the best efficiency with a delay error commensurate with a 40 dB sidelobe level. However, the synthesis procedure was not optimized in a strict sense, and the curves should be interpreted as a "good estimate" rather than as a limit to the obtainable efficiency.

5.4 Practical Implementation of Cascaded Coupled Lines

The best practical configuration for cascaded lines is obtained by realizing the coupled lines in dielectric stripline as broadside coupled plates (or strips) parallel to the ground plane. This configuration is selected for two reasons: (1) close coupling can be obtained (ρ's in the range from 10 to 25), and (2) it is easy to cascade the lines because the interconnecting strips are co-planar with the coupled lines. The parameters of a broadside coupled line are the ground plane separation, b, the coupled line separation, s, the coupled line width, w, and the thickness of the coupled lines, t. These dimensions are shown in Fig. 20 (a) and define the Z_o and ρ of the line [7] [8].

It is desirable, for ease of construction, for all lines in the cascade to have identical dimensions (except for length). It is also desirable, for reasons previously discussed, to perform the network synthesis with the first pole-zero pairs constrained to a constant Q_n contour. The following shows how this is realized. Consider a cascade of lines wherein each line has identical dimensions except for length. The lines are physically separated by at least a ground plane spacing to avoid coupling between adjacent lines. The lines are shown in Fig. 20 (b). The lines are shorted at one end (not shown) and are interconnected by strips (also not shown) which are co-planar with the coupled lines. The interconnecting strips have the same Z_o as the coupled lines. Since the line separation, s, is constant, the entire cascade and interconnecting strips can be realized on a single double-sided printed circuit board. Also, since each line in the cascade has identical dimensions, the coupling ratio, ρ, for each line is the same. The network Q, Q_n, of the first pole-zero pair is also the same. Thus, the root contour for the first pole-zero pair of each line in the cascade is a straight line having a slope Q_n and an intercept at the origin. This root contour is shown in Fig. 20 (c).

5.5 Dispersive Subsections

Several dispersive subsections have been synthesized on a constant Q contour. A multiplicity of these identical subsections can be cascaded to increase the bandwidth-time product. The results of the network synthesis for three representative designs are shown in Table I of Fig. 21. The bandwidth-time product of two designs is 10 and the third is 15. A Fourier analysis and the resultant sidelobe level of the theoretical time delay error is shown in Table II for Design No. 1. These (theoretical) errors are significantly below the errors which will occur in the physical realization of the lines. Although the designs are shown for a specific center frequency and bandwidth, the designs may be (frequency) scaled and are generally useful.

5.6 Construction Details

Techniques to realize coupled lines in stripline are covered in detail in the microwave literature. In particular, the coupled line is in widespread use as a quadrature, backward wave, hybrid. One method of implementing a cascade of lines is shown in Fig. 22. This particular design is for 1/2 inch ground plane spacing. The dielectric is Tellite 3B or an equivalent irradiated polyolefin. Item 1 consists of 1/4 inch thick dielectric backed on one side by a ground plate. These sheets are commercially available. Item 2 is a printed circuit board with the coupled lines printed on opposite sides. For the design shown, the printed circuit board has a thickness of 0.023 inches. The dielectric thickness is 0.020 inches (s) and the 1 ounce copper strips are 0.0015 inches (t). The resultant ground plane spacing is 0.523 inches (b). The coupled lines are separated by two ground plane spacings (about one inch) to reduce the coupling between adjacent pairs to a negligible amount. The coupled line width (w) is 0.278 inches. The design results in a line impedance of 30 ohms and a coupling ratio of 15.2. The strips which interconnect the lines also have a Z_o of 30 ohms and are about 0.816 inches wide. The coupled lines are shorted at one end of the printed circuit board by a copper tab which extends through the dielectric. A cascade, as shown in Fig. 22, usually constitutes a dispersive subsection. These subsections are connected together with 30 ohm cable. A quarter-wave transformer is used at the input and output of the composite network to convert to 50 ohms.

Any network design can be physically scaled in two dimensions; the line length, of course, must remain the same. A scaled network retains the same impedance, coupling ratio and (thus) transfer function. In Figure 23 the approximate network volume is plotted versus bandwidth for three values of ground plane spacing for a network efficiency of 25 percent.

5.7 Network Losses

Dissipative losses occur in the coupled conductors and in the dielectric. These losses have been

analyzed by Kolker [Reference 9 ; this analysis was performed at Hazeltine under contract AF30(602)-3576]

For a given impedance level and coupling ratio, the losses are primarily controlled by the ground plane spacing. The large spacing of the previous example (0.523 inches) results in low peak and differential losses. The theoretical losses are plotted in Fig. 24 for this ground plane spacing as a function of the dispersive bandwidth of the network. The curves are for a network with a BT product of 100 and a delay pedestal of 40 percent of the dispersion. The dB loss scales directly with the BT product. The dielectric losses were assumed to be zero; this is a reasonable assumption for an irradiated polyolefin dielectric for the bandwidths shown. Theoretical Q versus frequency is shown in Fig. 25 for three values of ground plane spacing for the case of a lossless dielectric. These curves may be used with the loss equation of Section 2.4. to compute the network loss for any bandwidth and center frequency for the design (or scaled design) given in paragraph 5.6 above. Measured Q data, for 0.5 inch ground plane spacing, from 200 to 400 MHz is also shown in Fig. 25.

5.8 Experimental Results

Experimental results are presented to demonstrate the feasibility and performance of the proposed stripline technique. One design which was constructed is Design No. 1 shown in Fig. 21. This design has a bandwidth of 500 Mhz and a center frequency of 1000 MHz. The basic dispersive subsection has a BT product of 10 and is comprised of 20 cascaded lines. The ground plane spacing used was 0.25 inches and the 20 coupled lines were contained on a single printed circuit board. The dimensions of the basic subsection was 0.25" x 4" x 15". A photograph of the basic subsection is shown in Fig. 26.

A total of ten subsections was constructed. These subsections were then cascaded to realize a network having a BT product of 100. Measurements were made on the individual subsections and on the cascaded network. The results confirm the technique and are summarized in Fig. 27. The RMS phase error for the subsection was about 2 degrees. The RMS error for the entire cascade was 6.5 degrees. The individual errors do not correlate but add in a root-mean-square manner. The equalized loss of the subsection was almost 40 dB and is plotted in Fig. 27 (d).

ACKNOWLEDGMENT

The initital work on the Type I network was supported by the Rome Air Development Center (RADC), Rome, N.Y. under Air Force contract AF-30(602)-2106. Subsequent work on the Type III network was supported by the RADC under contract AF30(602)-3372 and by the Advanced Research Projects Agency under contract AF30(602)-3576. The authors wish to acknowledge the support and assistance of of Mr. Clarence Silfer (EMATS) of the Rome Air Development Center who was the project monitor for these programs.

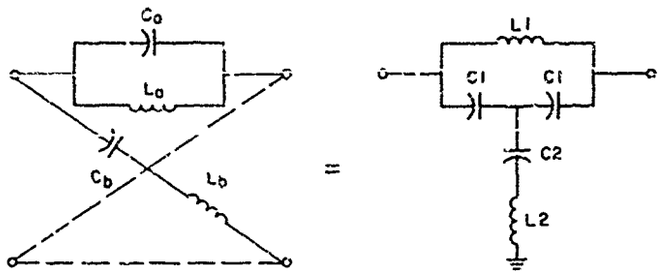
The development of the Type II network and additional work on the Type III network was supported by Hazeltine Corporation as part of an in-house IR and D program.

The production data and oscilloscope waveforms presented for the Type I network resulted from programs either direct [illegible] or under the cognizance of Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts. The authors wish to express their acknowledgement and appreciation of the technical contributions, support and stimulation provided throughout these programs by Jerome Margolin of Lincoln Laboratory.

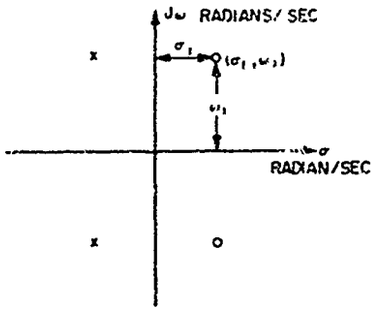
The work described has been performed over a period of years and is the result of the efforts of many people in research, engineering and production. In particular, the authors wish to acknowledge the significant contributions and direction provided by Messrs. John J. Strong and Charles E. Page of the Hazeltine Research Laboratories and the technical contributions of Mr. Sheldon Kerbel.

REFERENCES

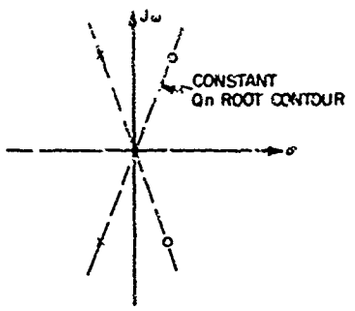
1. T. R. O'Meara, "The Synthesis of Band-Pass, All-Pass Time-Delay Networks with Graphical Approximation Techniques", Hughes Aircraft Co., Research Laboratories, Malibu, California, Report No. 114, February 1962.
2. P. S. Brandon, "The Design Methods for Lumped-Constant Dispersive Networks Suitable for Pulse Compression Radar", Marconi Review, Volume XXVII, Number 159, pp 225-253, 1965.
3. H. A. Wheeler, "The Interpretation of Amplitude and Phase Distortion in Terms of Paired Echoes", Proc. IRE, Vol. 27, pp 359-438, June 1939.
4. Bernard M. Oliver, "Directional Electromagnetic Couplers", Proc. IRE, Vol. 42, pp 1686-1692, Nov. "54.
5. E.M.T. Jones and J. T. Bolljahn, "Coupled-Strip-Transmission-Line Filters and Directional Couplers", IRE Trans. on Microwave Theory and Techniques, Vol. MTT-4, pp 75-81, April 1956.
6. W.J.D. Steenaart, "The Synthesis of Coupled Transmission Line All-pass Networks in Cascades of 1 to n", IEEE Trans. on Microwave Theory and Techniques, Vol. MTT-11, pp 23-29, Jan. 1963.
7. S. B. Cohn, "Characteristic Impedance of Broadside Coupled-Strip Transmission Lines", IRE Trans. on Microwave Theory and Techniques, Vol. MTT-8, pp 633-637, Nov. 1966.
8. S. B. Cohn, "Thickness Corrections for Capacitive Obstacles and Strip Conductors", IRE Trans. on Microwave Theory and Techniques, Vol. MTT-8, pp 638-644, Nov. 1960.
9. R. A. Kolker, "The Amplitude Response of a Coupled Transmission Line, All-Pass Network Having Loss", IEEE Trans. on Microwave Theory and Technique, Vol. 15, pp 438-443, August 1969.



(a) LATTICE AND EQUIVALENT UNBALANCED FORM (BRIDGED-TEE) OF ALL-PASS NETWORK



(b) POLE-ZERO LOCATION



(c) CONSTANT Qn CONTOUR

$$L1 = \frac{4\sigma_i R_o}{\sigma_i^2 + \omega_i^2} \qquad L2 = \frac{R_o}{4\sigma_i}$$

$$C1 = \frac{1}{2\sigma_i R_o} \qquad C2 = \frac{4\sigma_i}{R_o(\omega_i^2 - 3\sigma_i^2)}$$

(d) NETWORK ELEMENT EQUATIONS

Fig. 1 Uncompensated (first-order) all-pass network

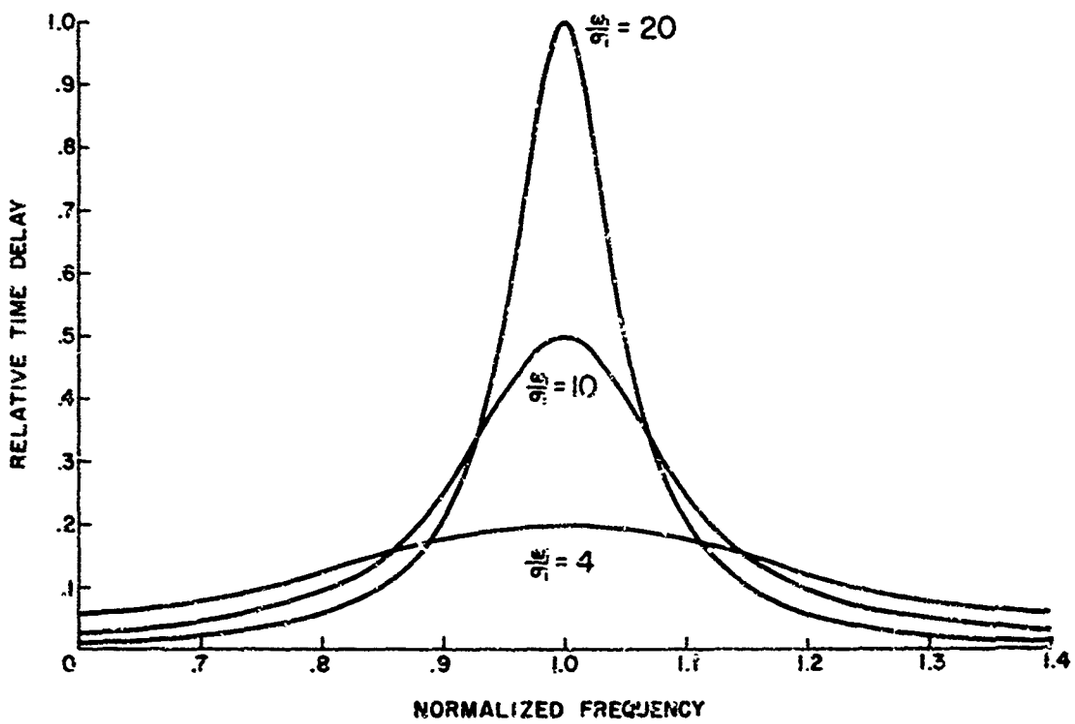


Fig. 2 Bridged-tee delay curves

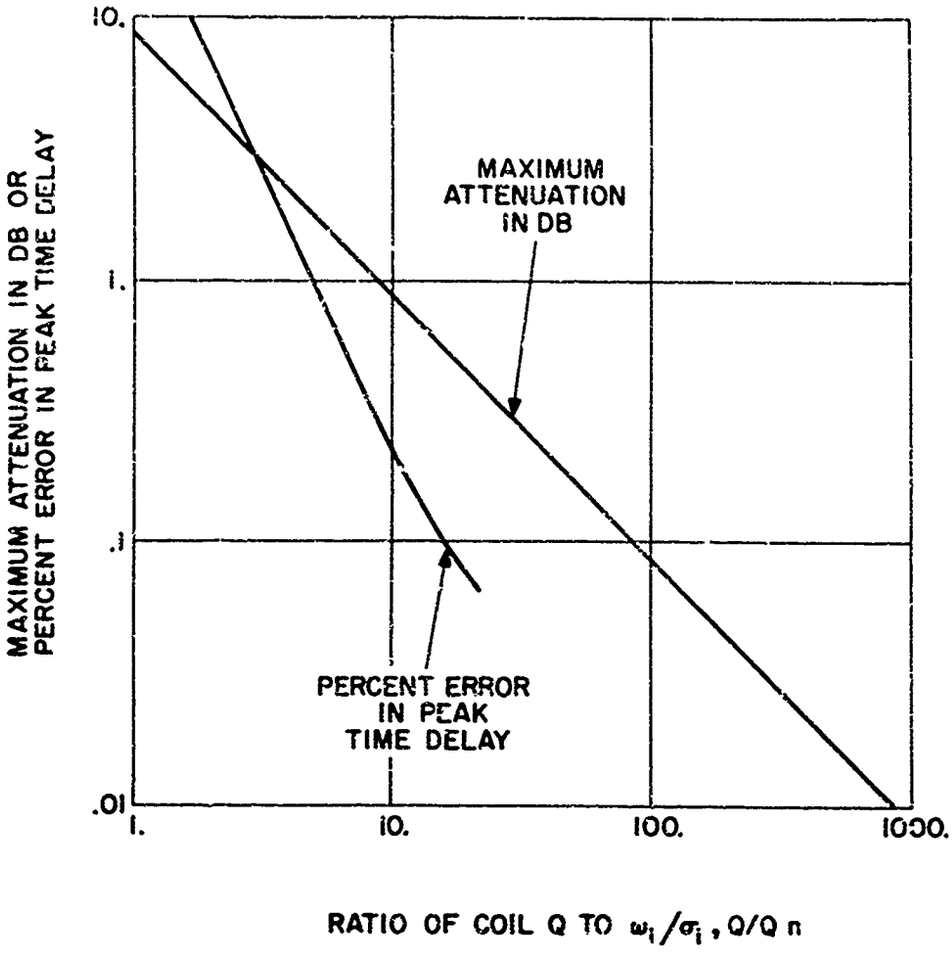
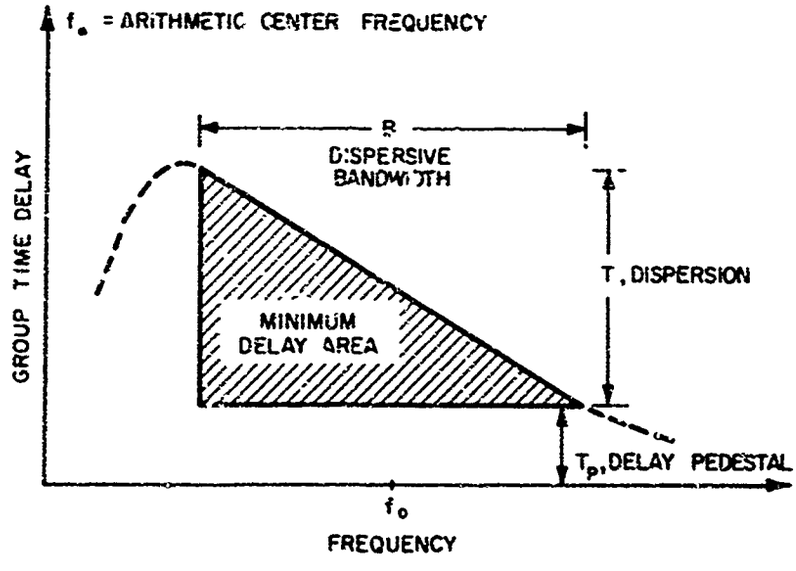


Fig. 3 Finite Q time delay error and attenuation

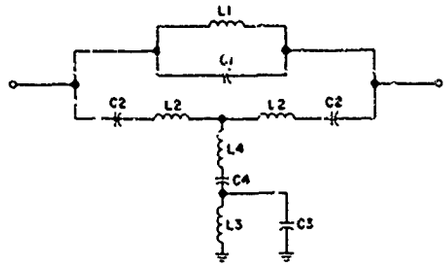


NETWORK FRACTIONAL BANDWIDTH = $\frac{B}{f_o}$

BANDWIDTH – TIME PRODUCT = BT

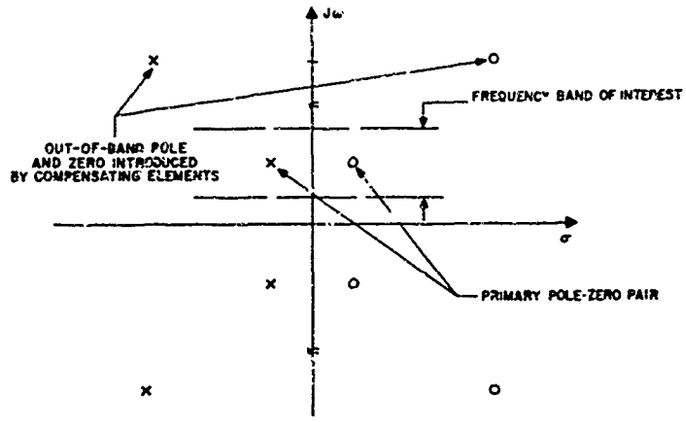
MINIMUM DELAY AREA = 0.5 • BT

Fig. 4 A linear dispersive characteristic



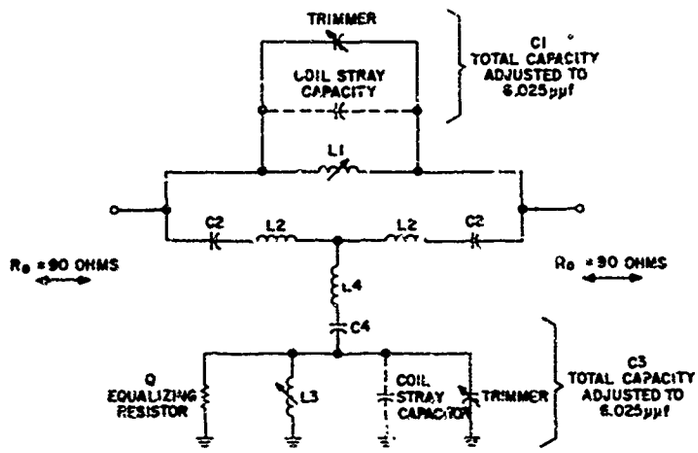
L2 = SERIES COMPENSATING COIL
L4 = SHUNT COMPENSATING COIL
C1 AND C3 ARE LARGELY COMPOSED OF THE INTERWINDING CAPACITY OF L1 AND L3

(a) COMPENSATED BRIDGED-TEE NETWORK



(b) POLE-ZERO PLOT

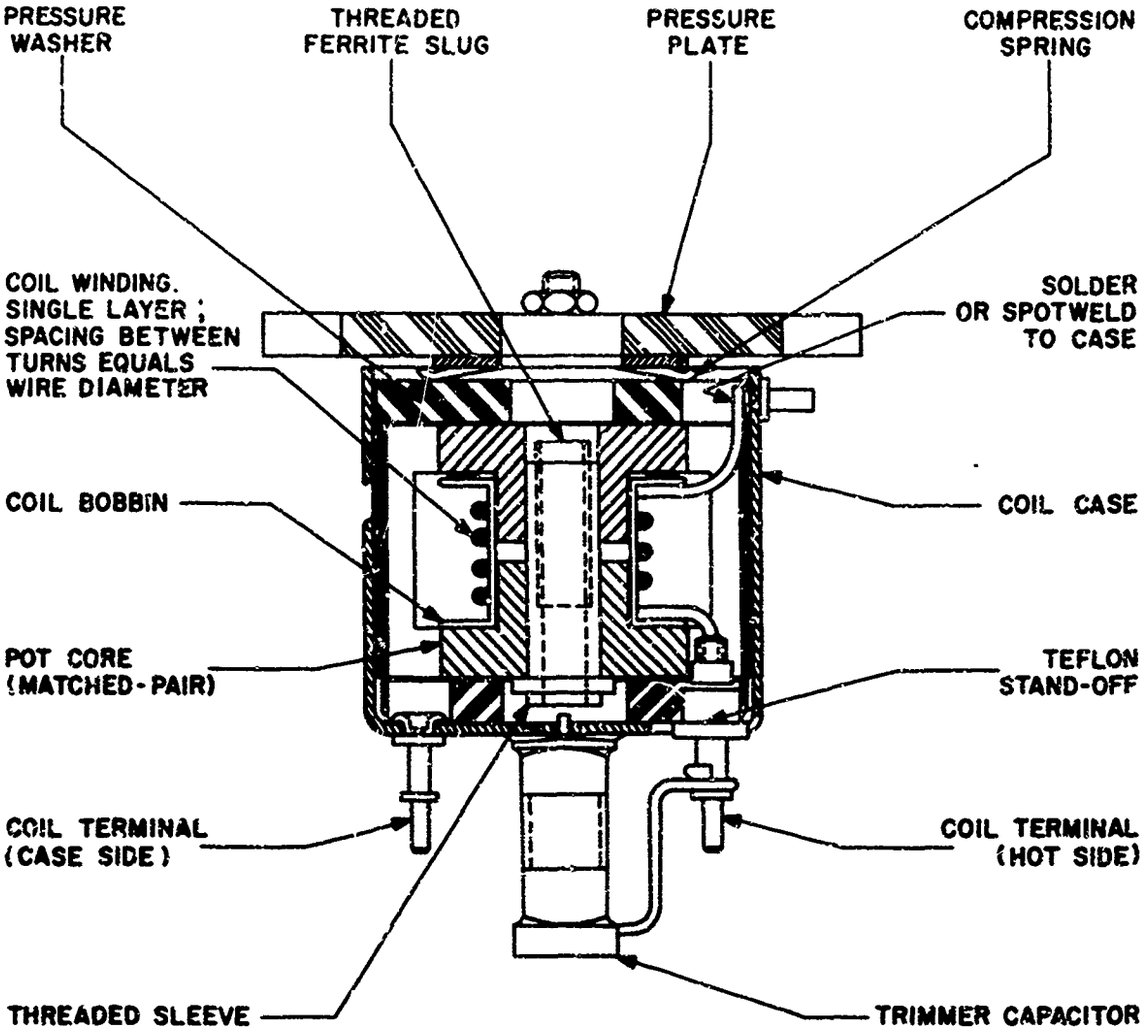
Fig. 5 Second order bridged-tee schematic and pole-zero plot



FREQUENCY	L1	C1	L2	C2	L3	C3	L4	C4
7MHz	0.85	6.0	0.024	1123	4.55	6.0	0.036	111
18MHz	0.36	6.0	0.024	375	1.52	6.0	0.036	47
29MHz	0.13	6.0	0.024	325	1.32	6.0	0.036	17

INDUCTANCE IN μh
CAPACITY IN μμf

Fig. 6 Bridged-tee element values



DATA FOR TYPICAL COIL	
INDUCTANCE	1.52 μh
FREQUENCY	18MC
COIL Q	315
STRAY CAPACITY	3.0μμf
NUMBER OF TURNS	6
WIRE SIZE	24

NOTE: SCALE IS 2:1

Fig. 7 Bridged-tee network coil

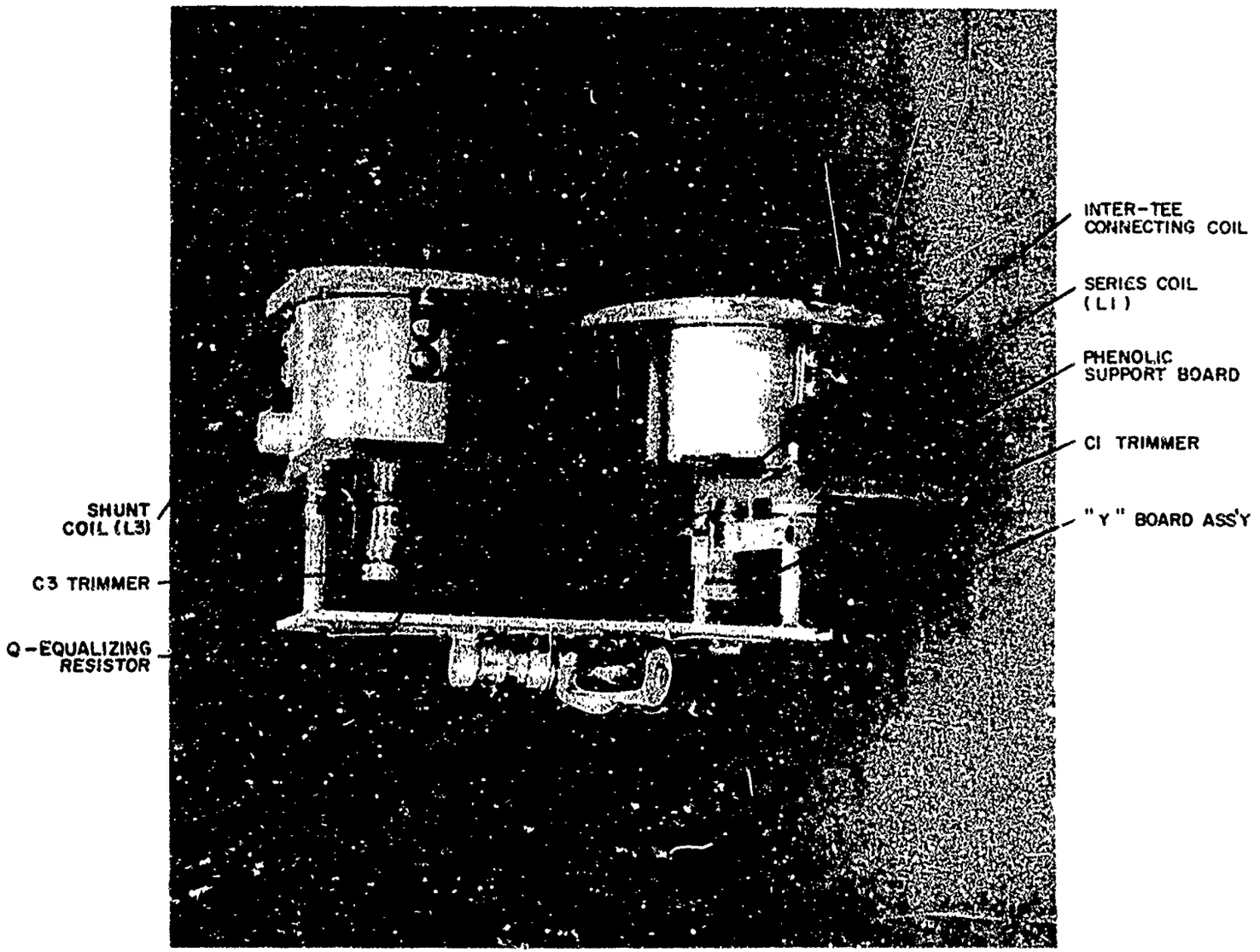
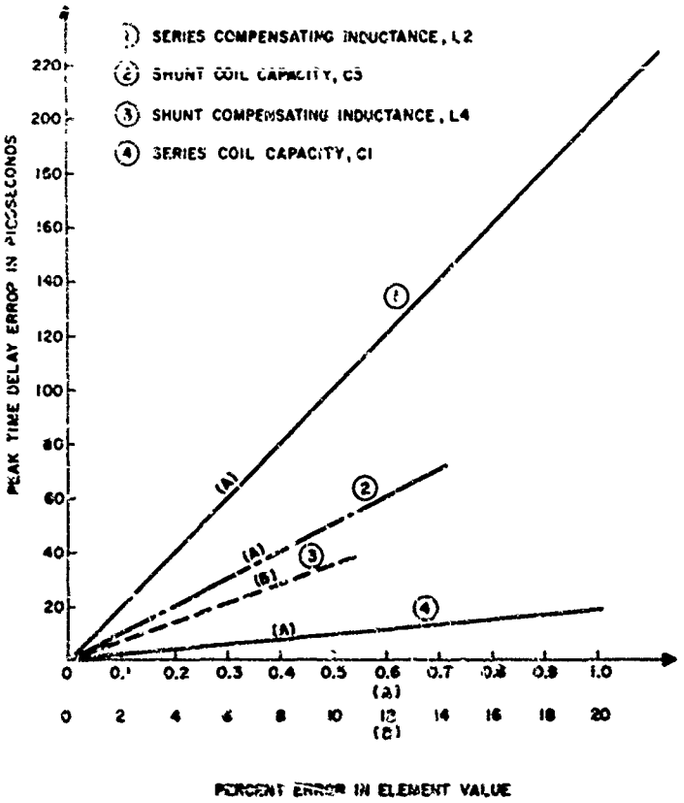


Fig. 8 Bridged-tee assembly



THE ABOVE CURVES SHOW THE PEAK TIME DELAY ERROR FOR ERRORS IN L2,L4,C1 AND C3 FOR A BRIDGED-TEE TUNED TO 18MC. THE PEAK TIME DELAY OF THE TEE IS 176.8nsec L1 AND L3 HAVE BEEN RETUNED TO COMPENSATE THE ELEMENT ERROR

Fig. 9 Element error curves

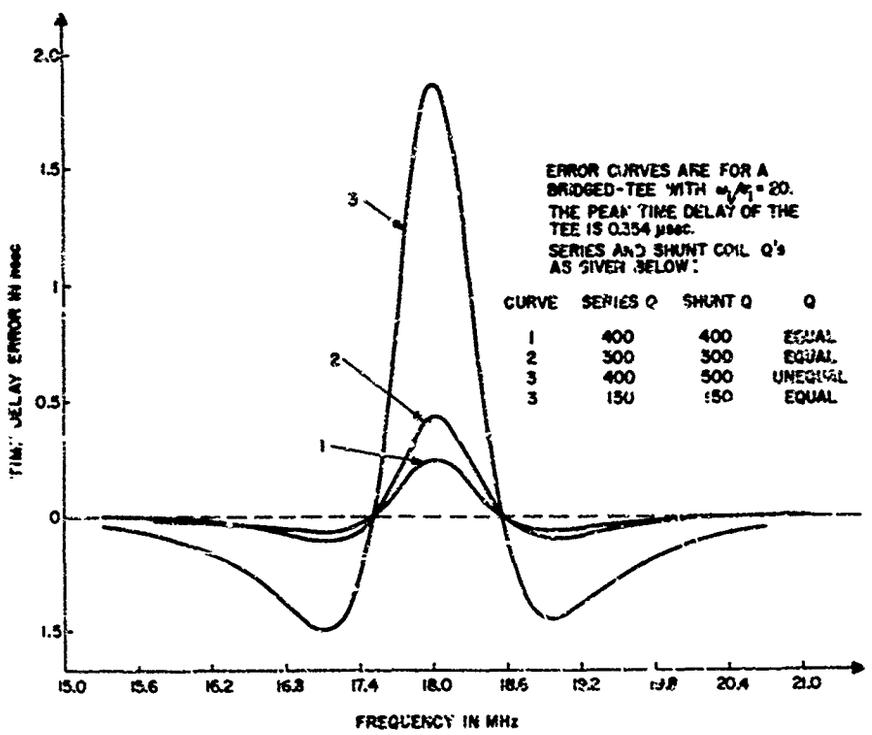


Fig. 10 Effect of unequal Q on time delay error

TABLE I

RMS Phase Error and Dispersion for 3 μs, 17.6 MHz Network

Network Number	RMS Phase Error in Degrees	Dispersion in μs	Network Number	RMS Phase Error in Degrees	Dispersion in μs
1	0.78	2.999	17	0.62	3.008
2	0.83	2.999	18	0.79	3.006
3	1.34	3.001	19	0.94	3.012
4	0.75	3.004	20	0.89	3.009
5	0.81	3.007	21	0.44	3.010
6	1.30	3.007	22	0.46	3.004
7	1.23	3.007	23	1.01	3.004
8	0.93	3.007	24	0.88	3.005
9	0.94	3.005	25	0.82	3.009
10	0.77	3.005	26	0.83	3.005
11	1.06	3.012	27	1.33	3.009
12	1.55	3.002	28	0.60	3.008
13	0.67	3.004	29	0.34	3.004
14	1.25	3.010	30	1.02	3.013
15	0.77	3.004	31	1.33	3.006
16	0.76	3.007	32	1.37	3.006

Avg. RMS Phase Error = 1.06°; Avg. Dispersion = 3.006 μs

TABLE II

Fourier Analysis of Phase Error for Network No. 26

Order of Harmonic	Phase Error in Degrees	Paired-Echo Sidelobe Level
1	0.77	-44 dB
2	0.55	-47 dB
3	0.56	-47 dB
4	0.18	-58 dB
5	0.15	dB
6	0.24	-54 dB
7	0.12	-60 dB
8	0.09	-63 dB
Remaining	< 0.1	< -61 dB

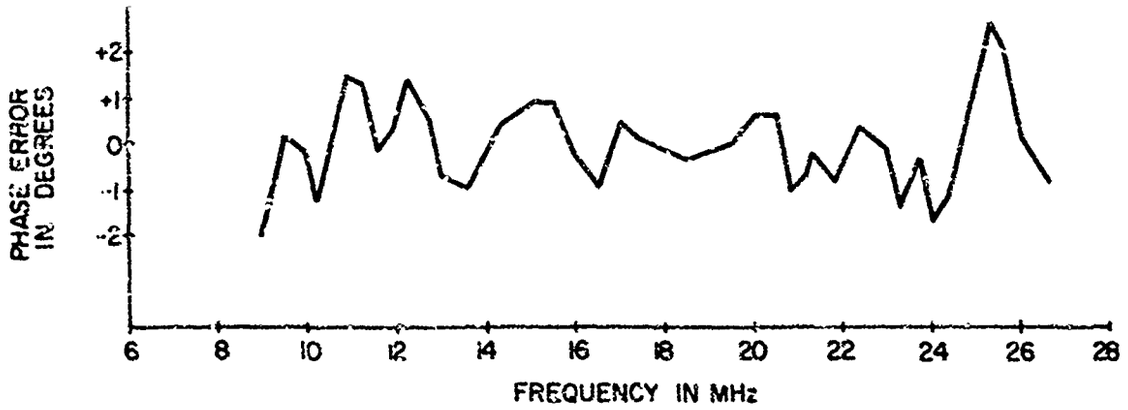
TABLE III

Phase Error and Dispersion for 15. μs, 17.6 MHz Network

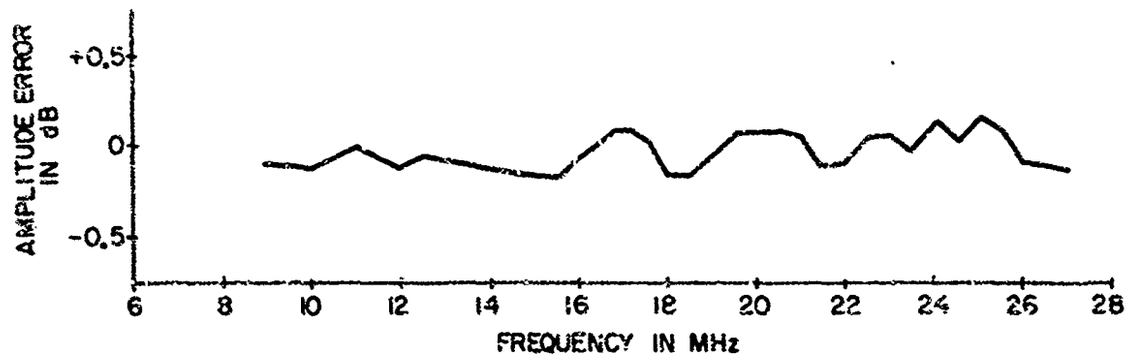
15 μs Network No.	RMS Phase Error in Deg.			Dispersion in μs
	Initial	Final (Hazeltine Nov. 1967)	Remeasured (Kwajalein Nov. 1968)	
1	5.00	1.53	0.86	14.992
2	3.44	0.47*	1.42*	14.969
3	1.63	0.35	1.39	14.996
4	2.09	0.75	0.94	14.992
5	1.86	0.89	1.20	15.004
6	2.08	0.42*	0.93*	14.989

* Phase Error for 12.0 MHz Bandwidth

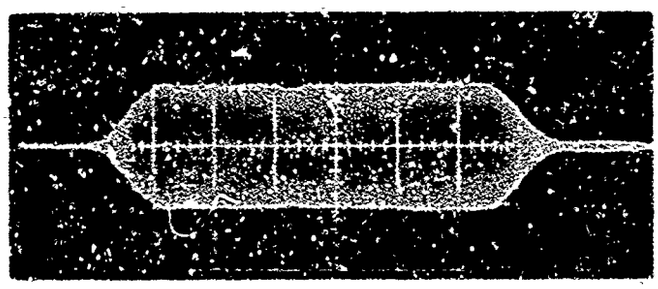
Fig. 11 Summary tables of type I network performance



(a) RESIDUAL PHASE ERROR OF 15 μs, 17.6 MHz NETWORK

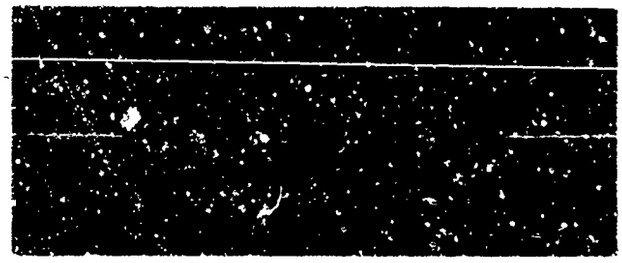


(b) RESIDUAL AMPLITUDE ERROR OF 15 μs, 17.6 MHz NETWORK

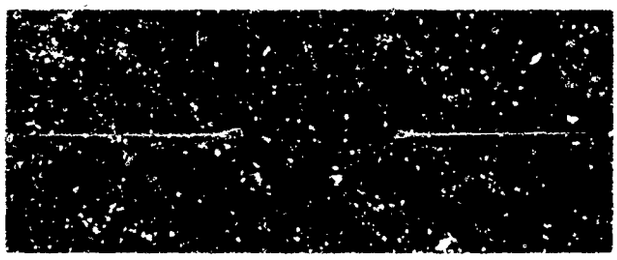


(c) NETWORK IMPULSE RESPONSE

Fig. 12 Amplitude and phase error of a dispersive network having a BT product of 264



(a) Expanded Pulse, 5 μsec/cm



(b) Compressed Pulse, 200 nanoseconds/cm

(c) Compressed Pulse, 500 nanoseconds/cm
Sidelobes shown are 43 db down

Fig. 13 Oscilloscope photographs of a pulse compression system having a bandwidth of 7 MHz and a compression ratio of 210

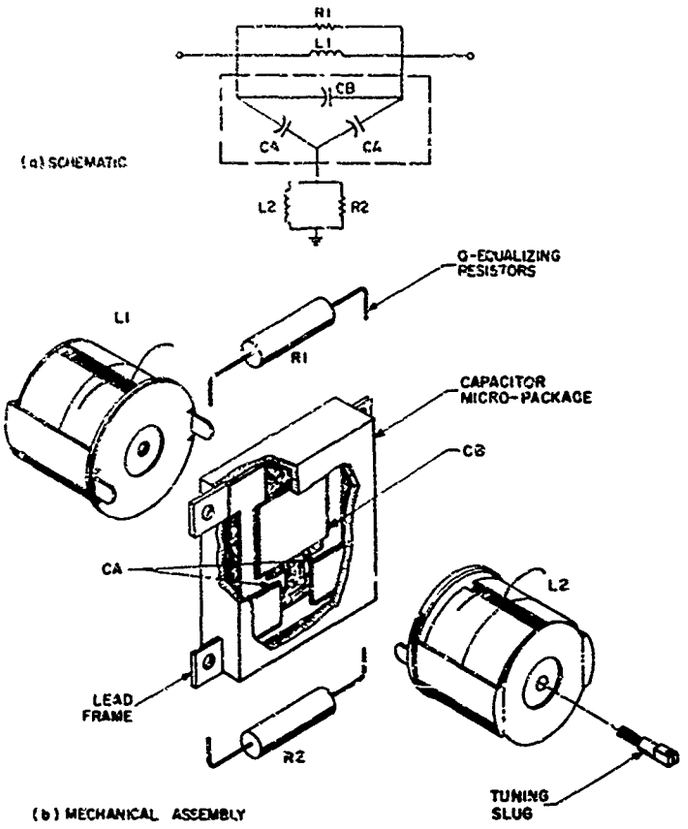


Fig. 14 Exploded view of miniature tee (type II) configuration

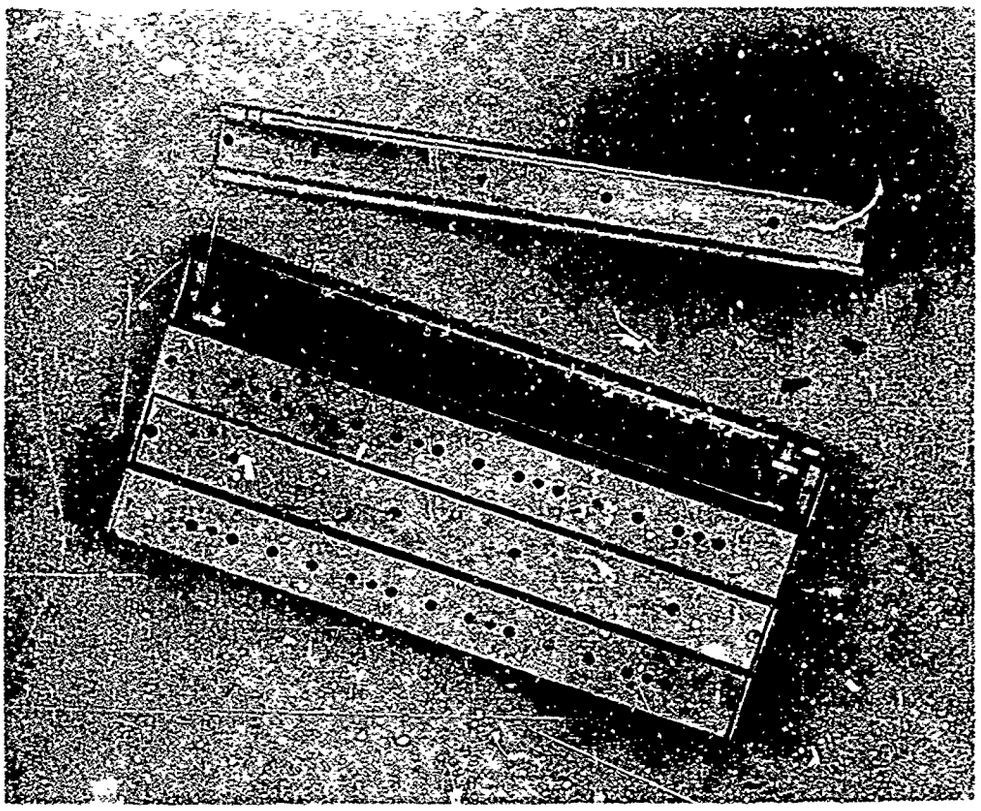


Fig. 15 Photograph showing dispersive delay unit

TABLE I

NETWORK PARAMETERS

Parameter	Network 1	Network 2	Network 3
Number of Tecs	26	26	26
BT Product	26.4	25.7	20
Bandwidth in MHz	17.6	60	60
Center Frequency in MHz	17.6	60	120
Dispersion in μs (Design)	1.50	0.428	0.333
Dispersion in μs (Measured)	1.50	0.427	0.327
Delay Pedestal in μs	0.267	0.075	0.13
RMS Phase Error (Design)	0.08°	0.2°	0.3°
RMS Phase Error (Measured)	0.96°	2.1°	2.7°
Equalized Loss in dB	7	7.5	11.5
Volume in Cu. In.	23	23	23

TABLE II

FOURIER ANALYSIS OF MEASURED PHASE ERROR

Order of Harmonic	Network 1		Network 2		Network 3	
	Phase Error in Deg.	Sidelobe Level in dB	Phase Error in Deg.	Sidelobe Level in dB	Phase Error in Deg.	Sidelobe Level in dB
1	0.41	-49	0.77	-44	1.04	-41
2	0.42	-49	0.15	-58	1.79	-37
3	0.35	-50	1.49	-38	1.33	-39
4	0.20	-55	0.93	-42	0.92	-42
5	0.27	-53	0.06	-66	1.48	-36
6	0.26	-53	0.21	-55	0.59	-46
7	0.35	-50	0.34	-50	0.83	-43
8	0.08	-63	0.27	-53	0.47	-48

Fig. 16 Summary tables of type II network performance

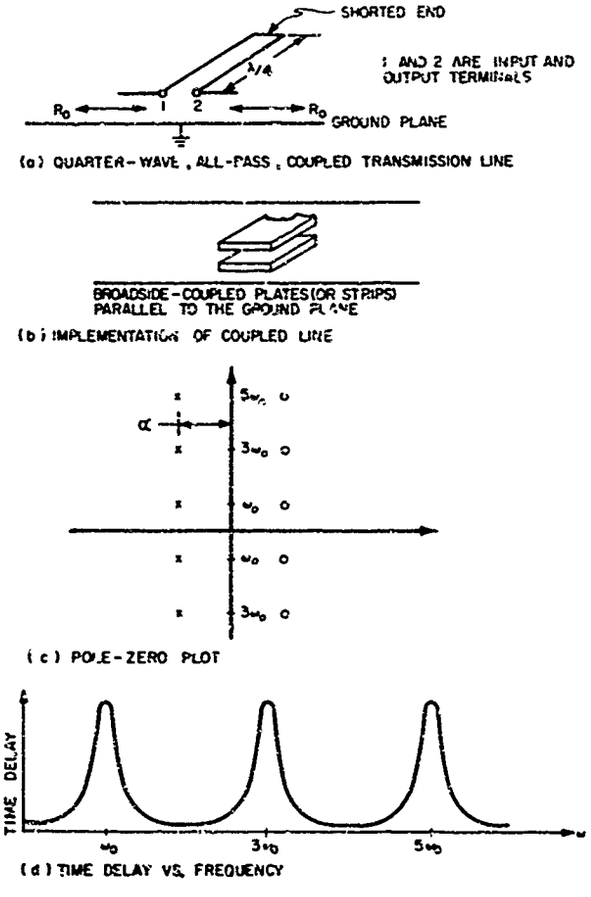


Fig. 17 All-pass quarter-wave coupled line

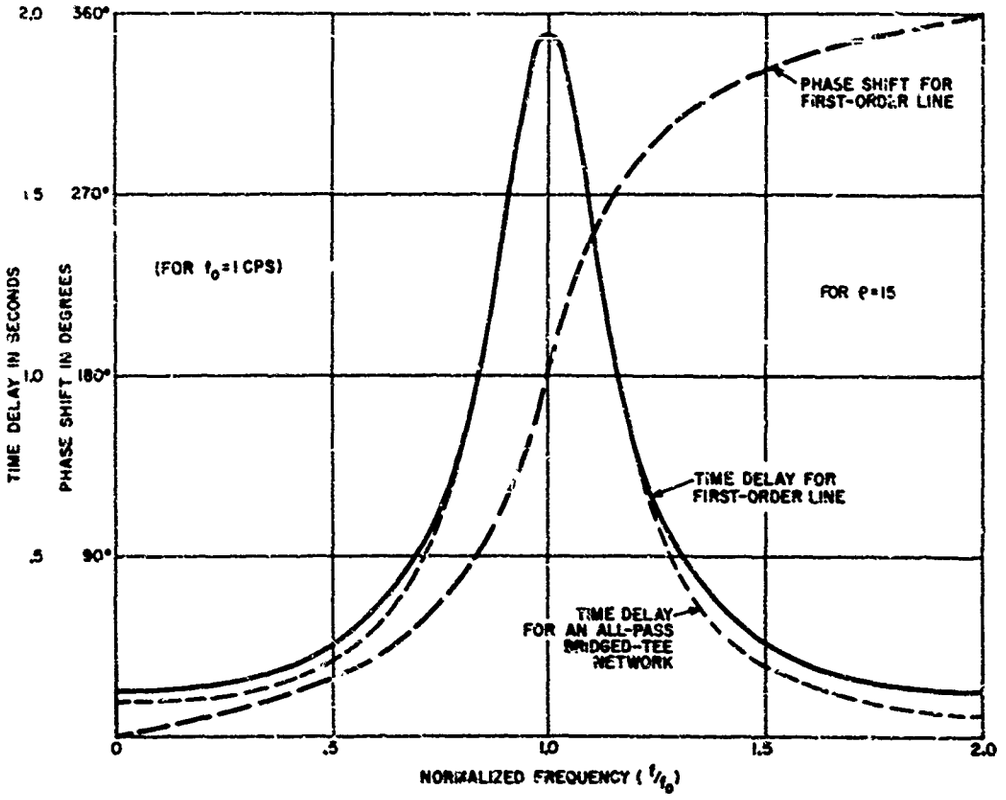


Fig. 18 Time delay and phase shift for a coupled line

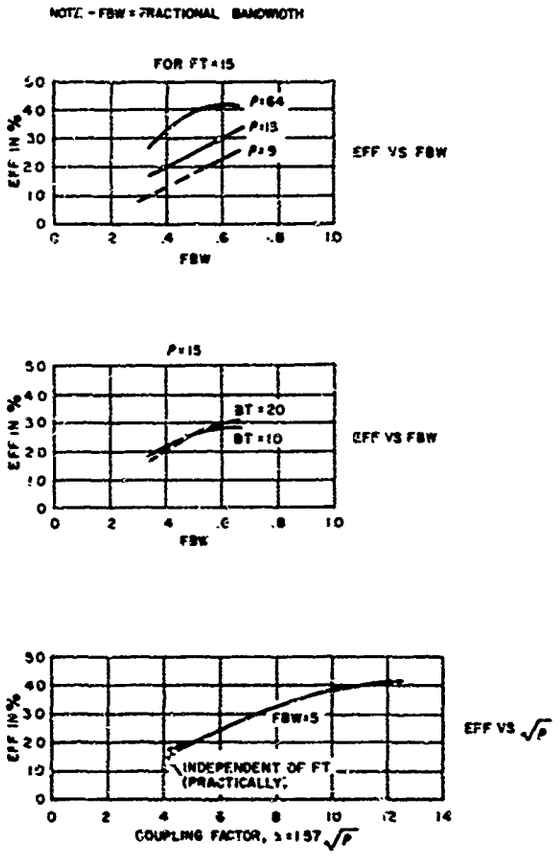


Fig. 19 Type III network efficiency curves

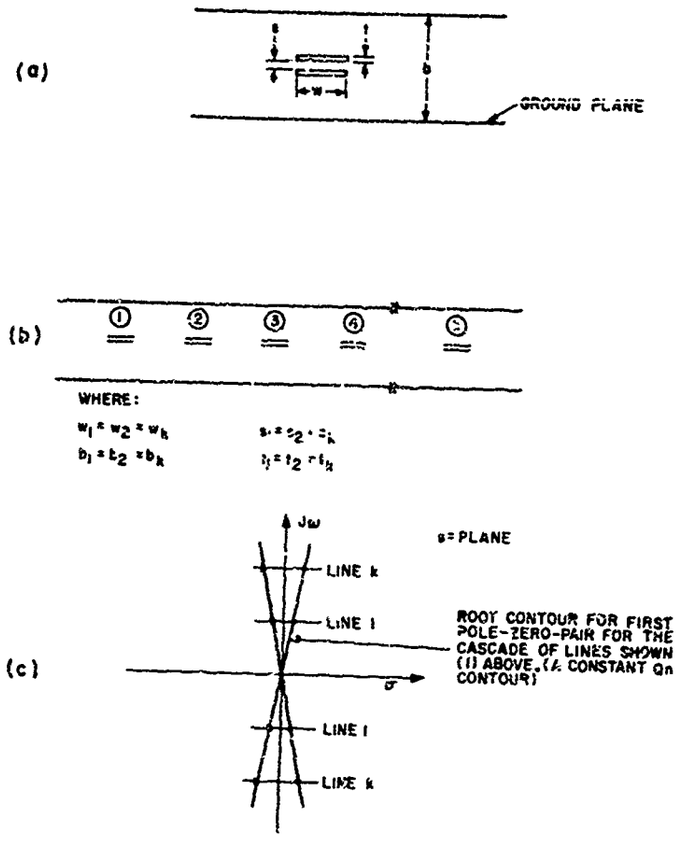


Fig. 20 Broadside-coupled stripline

TABLE I

DESIGN PARAMETERS FOR REPRESENTATIVE DESIGNS

Parameter	Design Number		
	1	2	3
BT Product	10	10	15
Number of Lines	20	14	30
Center Freq. in GHz	1.0	1.5	1.0
Bandwidth in GHz	0.5	2.0	0.5
Fractional Bandwidth	0.5	0.666	0.5
Dispersion in ns	20	10	30
Delay Pedestal in ns	6.1	3.5	12.1
Coupling Ratio (ρ)	15.2	24.5	15.2

TABLE II

FOURIER ANALYSIS OF PHASE ERROR FROM 0.75 GHz TO 1.25 GHz FOR DESIGN NO. 1

Order of Harmonic	Phase Error in Deg.
1	0.034
2	0.093
3	0.108
4	0.058
5	0.028

Fig. 21 Results of network synthesis

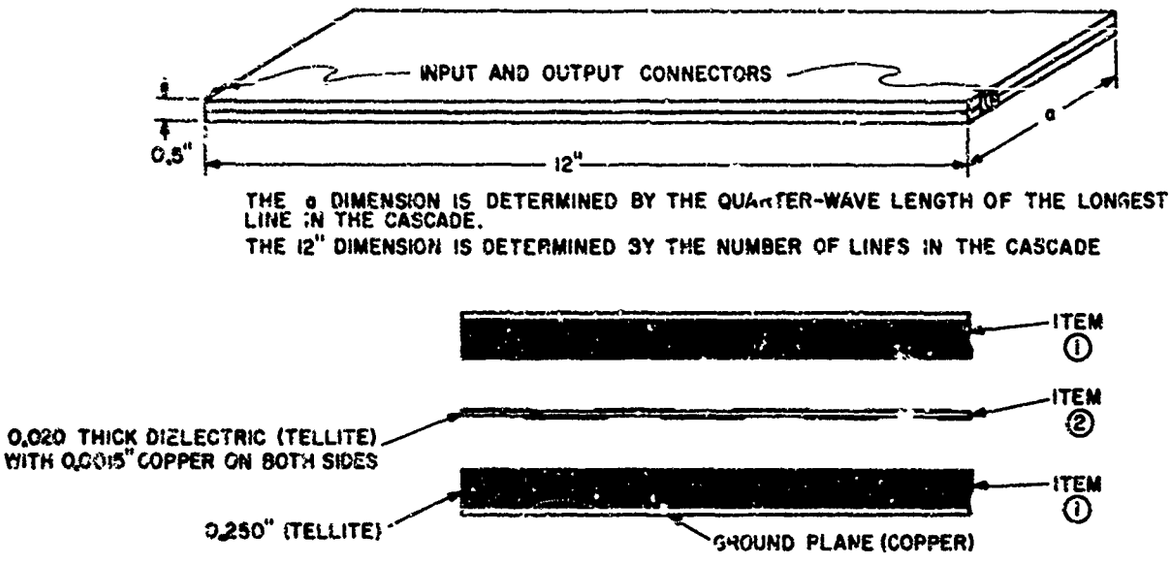


Fig. 22 Implementation of a cascade of coupled lines in stripline

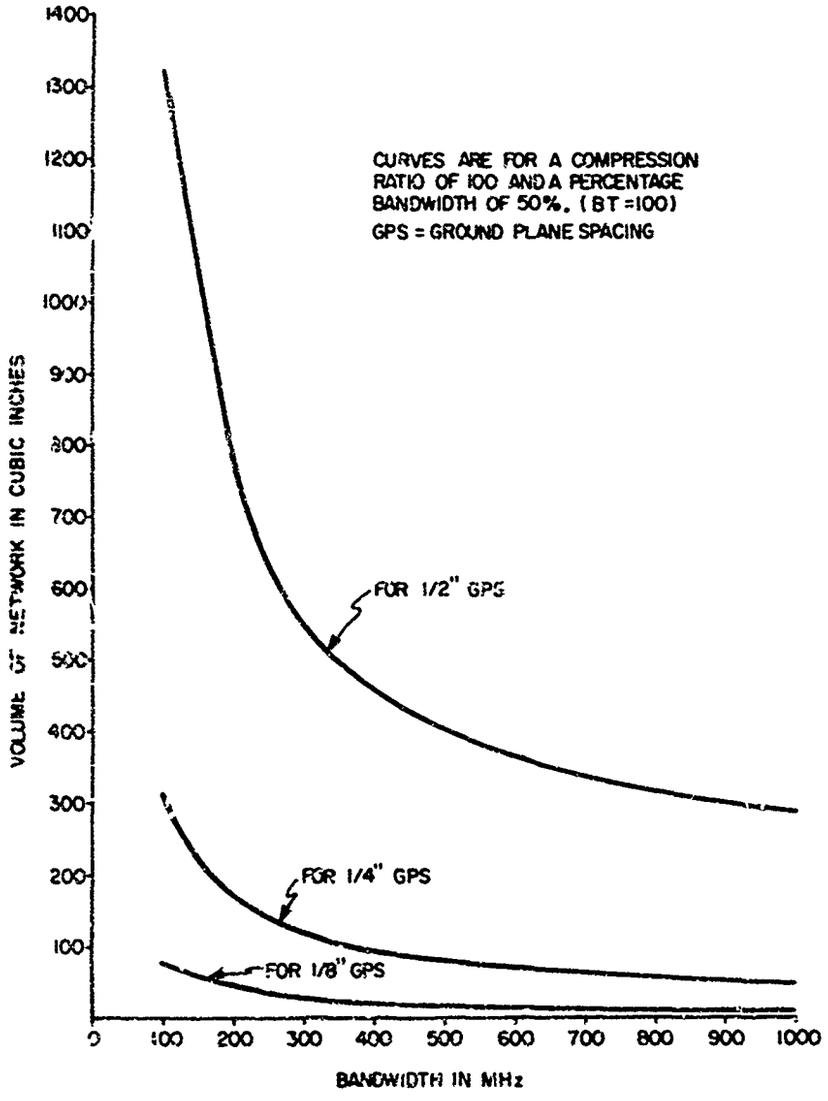


Fig. 23 Volume versus dispersive bandwidth

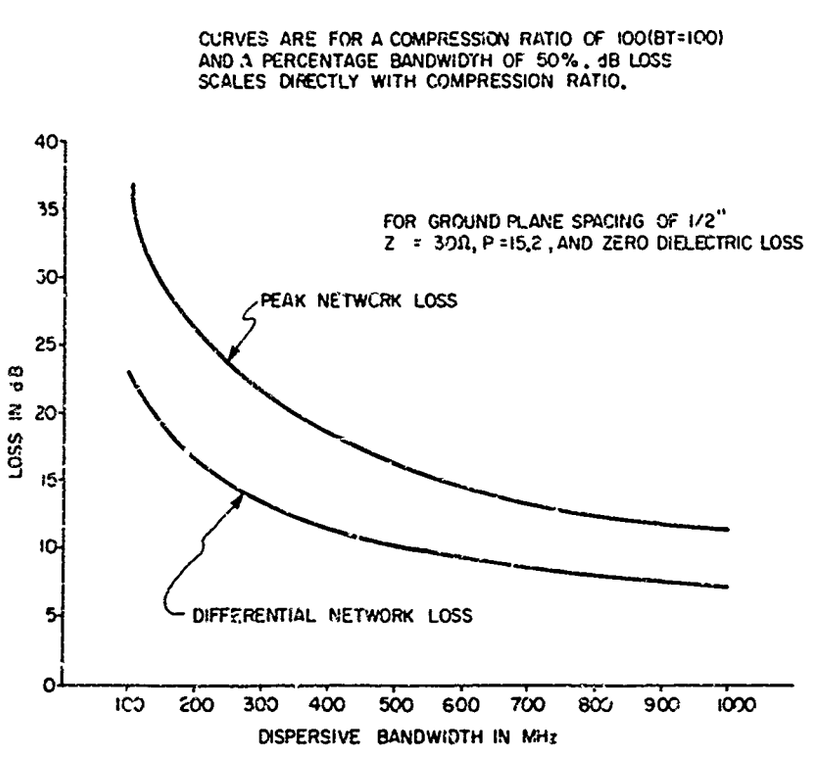


Fig. 24 Peak and differential network loss versus dispersive bandwidth

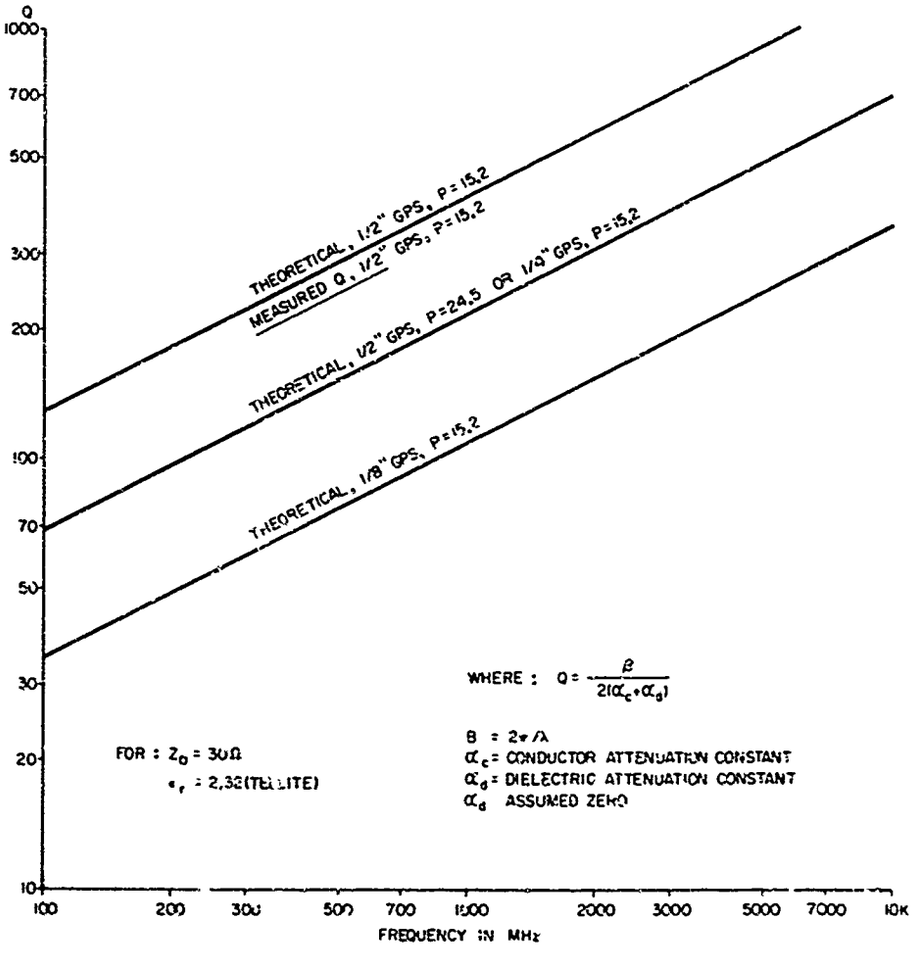


Fig. 25 Q versus frequency for broadside coupled stripline

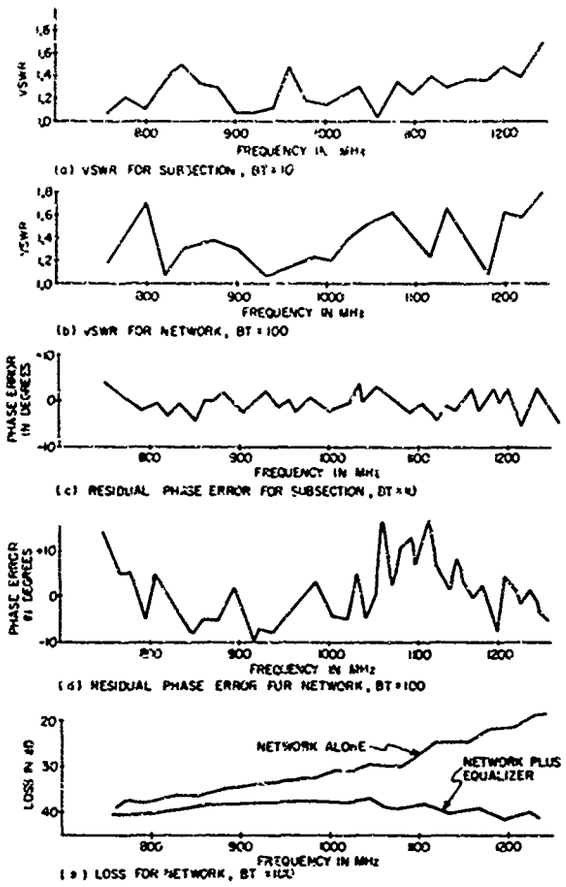


Fig. 26 Type III network performance

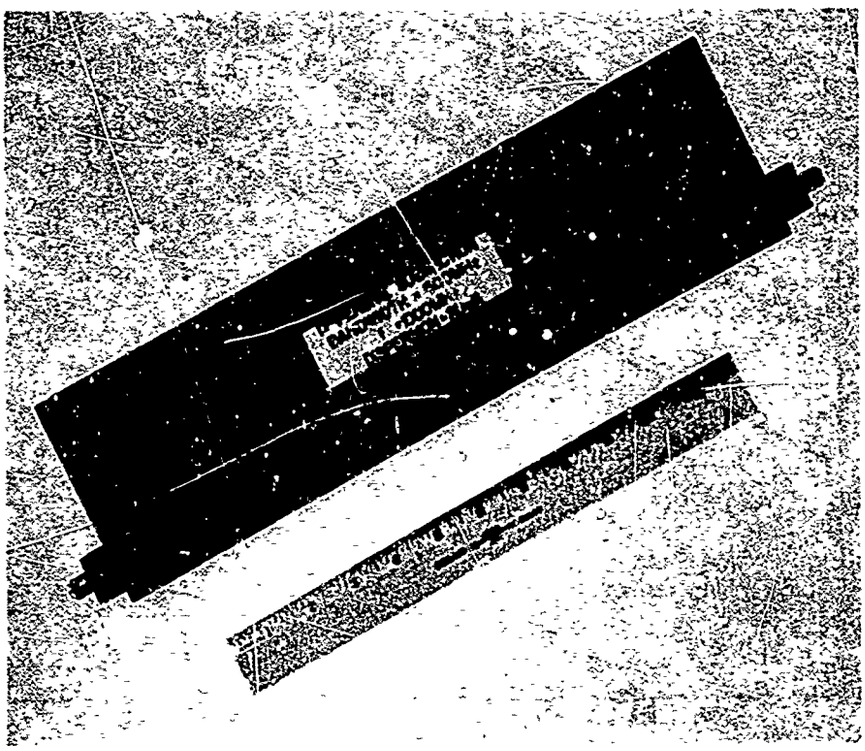


Fig. 27 Photograph of basic subsection, BT = 10

CORRELATION DIGITALE

par

R. Carre

THOMSON-C. S. F
92 - Malakoff
France

CORRELATION DIGITALE

PAR R. CARRE

THOMSON-C.S.F
92 - MALAKOFF

- SOMMAIRE

Les applications de la corrélation sont très nombreuses. L'une des plus importantes est la détection des signaux noyés dans le bruit ou le brouillage que l'on trouve dans les systèmes suivants :

Radars, Sonars, transmissions codées.

Dans cette application elle doit être en général réalisée en temps réel. Les progrès récents des circuits logiques et des mémoires intégrées (M.O.S) permettent de realiser cette fonction par des techniques digitales. Ces techniques sont les seules qui permettent d'obtenir une véritable agilité du code de modulation des signaux émis. Cette agilité se révèle très utile pour effectuer des levées d'ambiguïté et pour assurer la sécurité des transmissions.

Les corrélations digitales qui sont employées actuellement sont basées sur les résultats d'étude de modulations spéciales qui conduisent à des traitements faciles (codes binaires de phase par exemple).

L'étude qui a été réalisée s'est fixée comme but de définir un procédé applicable à toutes les formes de modulation de fréquence ou de phase. Ce procédé consiste à utiliser une réplique approximative que l'on obtient par quantification de la phase de la loi de modulation.

Les résultats suivants ont été obtenus :

1°/ Le traitement est presqu'aussi simple qu'avec les codes binaires

2°/ La dégradation de la fonction de corrélation apportée par ce procédé est faible. Elle peut être facilement évaluée en fonction des paramètres du codage.

L'étude comportait une analyse théorique et des simulations numériques de contrôle. Elle a été réalisée à THOMSON-CSF avec la collaboration technique et financière de la D.R.M.E (D.R.M.E : Direction des recherches et moyens d'essais du Ministère des Armées).

2 - PRINCIPES GENERAUX

2.1. Choix de la modulation à l'émission

Le procédé s'applique à toute modulation à amplitude constante. Avant de définir les paramètres du codage, il est nécessaire de choisir la loi de modulation la mieux adaptée au problème à résoudre.

En radar la modulation linéaire de fréquence est la forme la plus employée. Elle correspond à une modulation parabolique de la phase en fonction du temps.

En Sonar on utilise soit la même loi, soit une modulation de phase logarithmique.

Ces formes de modulation sont en général préférées aux codes binaires de phase en raison de l'excellente qualité de l'impulsion comprimée qu'elles permettent d'obtenir même lorsque l'effet Doppler décale la fréquence des signaux reçus.

2.2. Approximation de la fonction de modulation

Deux modes peuvent être utilisés. Le corrélateur est le même dans les deux cas.

Mode 1

La fonction émise n'est pas modifiée, elle est donc identique à celle que l'on utilise lorsqu'on veut réaliser une corrélation analogique. Les signaux sont échantillonnés et codés en valeurs numériques à la réception. Ils sont mis en mémoire sur la durée du signal émis et corrélés avec la réplique approximative. Cette réplique est obtenue par une quantification en valeurs discrètes de la loi de phase de la modulation, suivie d'un échantillonnage à fréquence égale à celle des signaux reçus. Un exemple est donné à la figure 1. La fonction est représentée par des multiples de $\frac{2\pi}{m}$.

m : Nombre de moments de phase de la quantification.

En général m sera égal à 4 ou à 8.

Mode 2

On remplace à l'émission la loi de phase originale par la loi de phase quantifiée. La réplique est identique à celle du mode précédent.

2.3. Corrélateur

Le bloc diagramme est représenté figure 2.

Les signaux transposés et amplifiés en fréquence intermédiaire sont appliqués à deux détecteurs synchrones qui reçoivent deux références à même fréquence déphasées entre elles de $\pi/2$. On obtient deux vidéos bipolaires qui représentent la loi de modulation des signaux décomposée sur deux axes perpendiculaires.

Ces deux voies sont échantillonnées et codées en amplitude et signe. Les échantillons successifs glissent dans deux registres à transfert.

La corrélation consiste à effectuer une somme algébrique des signaux contenus dans les registres à chaque instant d'échantillonnage suivant un programme défini par la réplique. La méthode en est exposée ci-dessous.

2.4. Principe théorique

A chaque instant d'échantillonnage le signal reçu peut être représenté par un vecteur dont les projections sont les échantillons des deux voies.

Soit x_{S1} , x_{S2} , x_{S3} ..... x_{SN} les N échantillons successifs d'une voie qui se trouvent en mémoire et soit y_{S1} , y_{S2} ..... y_{SN} les N échantillons correspondants de l'autre voie.

Les vecteurs correspondants sont :

$$\rho_{S1} e^{j\phi_1} , \rho_{S2} e^{j\phi_2} , \text{---------} \rho_{SN} e^{j\phi_N}$$

On a alors les relations suivantes :

$$x_{Si} = \rho_{Si} \cos \phi_i$$

$$y_{Si} = \rho_{Si} \sin \phi_i$$

La réplique du signal émis peut être aussi représentée par une suite de vecteurs à amplitude constante soit :

$$e^{j\theta_1} , e^{j\theta_2} , e^{j\theta_3} \text{ ..... } e^{j\theta_N}$$

La réplique approximative quantifiée devient alors :

$$e^{j\theta q_1} , e^{j\theta q_2} \text{ ..... } e^{j\theta q_n}$$

Cette réplique quantifiée peut aussi être représentée par les deux projections de ses vecteurs soit :

$$x_{ci} = \cos \theta q_i$$

$$y_{ci} = \sin \theta q_i$$

La corrélation consiste à former le vecteur résultant $\overrightarrow{V_R}$ suivant :

$$\overrightarrow{V_R} = \sum_{i=1}^{i=N} \rho_{Si} e^{j\phi_i} \cdot e^{-j\theta q_i} = \sum_{i=1}^{i=N} \rho_{Si} e^{j(\phi_i - \theta q_i)}$$

Cette opération est répétée à chaque instant d'échantillonnage. Le corrélateur calcule les deux composantes du vecteur $\overrightarrow{V_R}$ soit :

$$x_R = \sum_{i=1}^{i=N} \rho_{Si} \cos(\phi_i - \theta q_i) = \sum_{i=1}^{i=N} x_{Si} \cdot x_{ci} + y_{Si} \cdot y_{ci}$$

$$y_R = \sum_{i=1}^{i=N} \rho_{Si} \sin(\phi_i - \theta q_i) = \sum_{i=1}^{i=N} y_{Si} \cdot x_{ci} - x_{Si} \cdot y_{ci}$$

On calcule finalement le module de $\overrightarrow{V_R}$ qui représente l'amplitude de l'échantillon de la fonction de corrélation

$$Z_R = \sqrt{x_R^2 + y_R^2}$$

Les valeurs θ_q sont multiples de $\frac{2\pi}{m}$, x_c et y_c n'ont que m valeurs différentes.

C'est ainsi que pour m = 4 les calculs de x_R et y_R se réduisent à des sommes algébriques des échantillons x_s et y_s .

Les valeurs de x_c et y_c pour m = 4 et m = 8 sont données dans le tableau ci-dessous :

ϑ_q	x_c	y_c
0	+1	0
$\pi/4$	$+\frac{\sqrt{2}}{2}$	$+\frac{\sqrt{2}}{2}$
$\pi/2$	0	+1
$3\pi/4$	$-\frac{\sqrt{2}}{2}$	$+\frac{\sqrt{2}}{2}$
π	-1	0
$5\pi/4$	$-\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$
$3\pi/2$	0	-1
$7\pi/4$	$+\frac{\sqrt{2}}{2}$	$-\frac{\sqrt{2}}{2}$

m = 4 — m = 8

2.5. Organisation des calculs

Un exemple d'organisation est donné à la figure 3 pour m = 4.

Seul le calcul de x_R est représenté, le terme y_R se calcule de la même façon.

Tous les termes qui reçoivent le même coefficient sont sommés ensemble. L'aiguillage sur les lignes communes se fait au moyen de portes commandées par la réplique.

Les sommes peuvent être digitales ou analogiques suivant les vitesses à réaliser.

2.6. Etablissement de la réplique

Les angles θ_q sont d'abord déterminés au moyen d'une quantification au plus près comme il a été montré à la figure 1. La fonction quantifiée est ensuite échantillonnée avec un intervalle d'échantillonnage $\Delta t = \frac{1}{2\Delta F}$

ΔF : largeur moyenne du spectre émis.

La fonction ainsi obtenue est ensuite modifiée au moyen d'un critère de transformation. Cette opération a pour but de rompre les périodicités qui se produisent dans la quantification au plus près. Ces périodicités produiraient en effet des réponses parasites en dehors du pic principal.

Plusieurs critères ont été étudiés au cours de l'étude, le plus simple est décrit ci-dessous.

2.6.1. Critère de transformation

Ce critère est basé sur le principe suivant :

En l'absence de quantification de phase l'amplitude du vecteur résultant $\overrightarrow{V_R}$ est pratiquement nul en dehors du pic de corrélation, il est la somme de vecteurs élémentaires de même module qui réalisent un contour fermé. On peut démontrer qu'on a un pic parasite minimal avec la fonction quantifiée si la somme algébrique des erreurs de quantification de phase est nulle.

L'application du critère se déduit de ce principe. On effectue la somme algébrique En des erreurs de quantification à partir d'une des extrémités de la réplique soit :

$$E_n = \sum_{i=1}^{i=n} \left(\theta q_i - \theta_i\right)$$

Ce calcul est effectué pour tous les échantillons de la réplique. Chaque fois que l'on obtient $|E_n| > \frac{\eta}{m}$ on modifie le codage de l'échantillon correspondant de $2\eta/m$ avec le signe convenable. On maintient ainsi la somme algébrique En entre deux bornes tout le long de la réplique soit :

$$-\frac{\eta}{m} < E_n < +\frac{\eta}{m}$$

Les changements effectués sont peu nombreux mais ils suffisent à rompre les périodicités.

3 - RESULTATS DE L'ETUDE

L'analyse théorique et les simulations numériques ont donné des résultats pratiquement identiques.

3.1. Analyse théorique

Les points suivants ont été analysés :

1°/ L'influence de la quantification sur l'amplitude du pic principal

2°/ Le niveau de bruit apporté par la quantification en dehors du pic principal.

Soit un signal de durée T et de largeur de spectre ΔF et une approximation à m moment de la phase.

On peut définir les deux paramètres suivants :

- le taux de dispersion D (taux de compression en radar) $D = \Delta F.T$
- L'erreur efficace de quantification $\sqrt{1-e_m^2}$

$$e_m = \frac{m}{\eta} \sin \frac{\eta}{m}$$

La perte Pc introduite par la quantification sur le pic principal est donnée ci-dessous en décibels pour les deux modes d'approximation (voir 2.2)

Pc	dB
Mode 1	Mode 2
$20 \log \varepsilon_m$	0

La puissance du bruit de quantification est maximale au voisinage du pic principal elle décroît ensuite linéairement jusqu'à l'extrémité de la fonction de corrélation.
Le rapport Rp entre le pic principal et la valeur efficace maximale du bruit de quantification est donnée ci-dessous en décibels pour les deux modes d'approximation :

Rp	dB
Mode 1	Mode 2
$10 \log D \frac{\varepsilon_m^2}{1-\varepsilon_m^2}$	$10 \log \frac{D}{1-\varepsilon_m^4}$

Le bruit de quantification est 3 dB supérieur en mode 2 qu'en mode 1.

Note :

Ces valeurs ne comprennent pas le bruit supplémentaire provenant de l'application du critère de transformation. Les simulations ont montré qu'il faut compter sur une dégradation de Rp de 2 dB environ.

3.2. Simulation

Les résultats sont sensiblement identiques que l'on utilise une modulation parabolique ou logarithmique de phase.

Afin de ne pas cacher les effets de la quantification, la corrélation a été complétée par une pondération de Hamming du 1er ordre. L'effet de cette pondération est de réduire les lobes naturels de la fonction de corrélation. Elle consiste à effectuer la somme de trois fonctions de corrélation décalées entre elles de $\frac{1}{\Delta F}$, soit :

$$S_p(t) = S_{(t)} + \lambda \left[S\left(t - \frac{1}{\Delta F}\right) + S\left(t + \frac{1}{\Delta F}\right) \right]$$

λ est un coefficient qui a pour valeur optimale 0,44.

Pour simplifier les calculs on a pris $\lambda = 0{,}375 = \frac{1}{4} + \frac{1}{8}$

Le tableau ci-après donne quelques résultats typiques qui ont été obtenus avec une modulation parabolique et m = 8.

Des décalages de fréquence ont été introduits entre l'émission et la réception pour simuler l'effet Doppler.

Le coefficient α exprime le rapport entre la fréquence Doppler et le spectre ΔF du signal.

Pc représente la perte sur le pic principal par rapport à une corrélation sans quantification et à Doppler nul

R'p représente le rapport entre le pic principal et le plus élevé des lobes secondaires.

D	α	Sans quantificat.		Mode 1		Mode 2	
		Pc : dB	R'p: dB	Pc : dB	R'p: dB	Pc : dB	R'p : dB
100	0	0	36	0,5	27,4	0	23,4
	5 %	0,2	34	0,7	26,2	0,9	22
	10 %	0,3	29,6	0,9	25,5	1,1	20,4
200	0	0	35,4	0,4	31	0	26,2
	5 %	0,2	35,4	0,6	31,4	0,9	25,5
	10 %	0,3	30,5	0,8	29,6	1,2	24,6

3.2.1. Codage en amplitude

Le bruit provenant du codage en amplitude des signaux reçus a déjà été abondamment traité dans la littérature technique. On s'est donc contenté de vérifier par simulation que ce bruit n'apportait pas d'autre dégradation que celle que prévoie la théorie. C'est ainsi qu'avec un taux de dispersion D $\geqslant$ 100, on a bien constaté que le codage pouvait se faire à 3 bits et signe pour chaque voie x et y. Les modifications de la fonction de corrélation étaient négligeables.

4 - CONCLUSION

Le procédé qui consiste à remplacer une fonction de modulation continue par une fonction quantifiée en phase permet de réaliser la fonction de corrélation par des méthodes digitales. On peut ainsi envisager une véritable agilité des codes d'émission tout en conservant des fonctions d'ambiguïté proches des modulations continues.

La simplification du traitement apportée par ce procédé est d'autant plus importante que le taux de compression est grand.

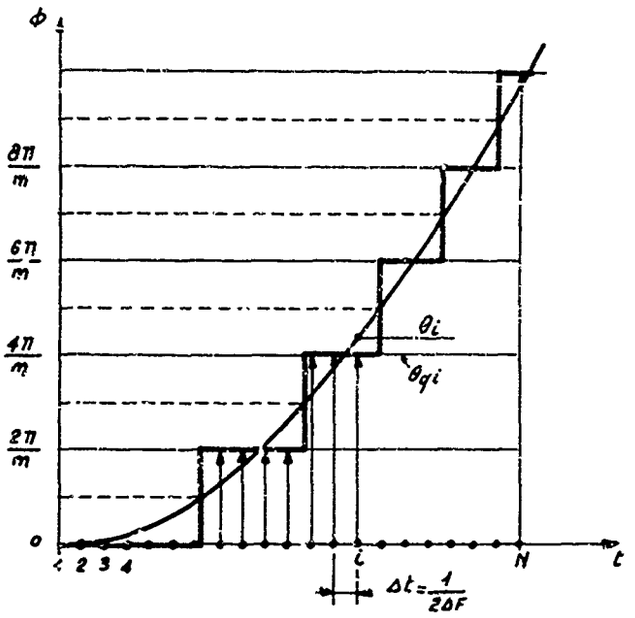


Fig. 1 Quantification et echantillonnage de la loi de phase

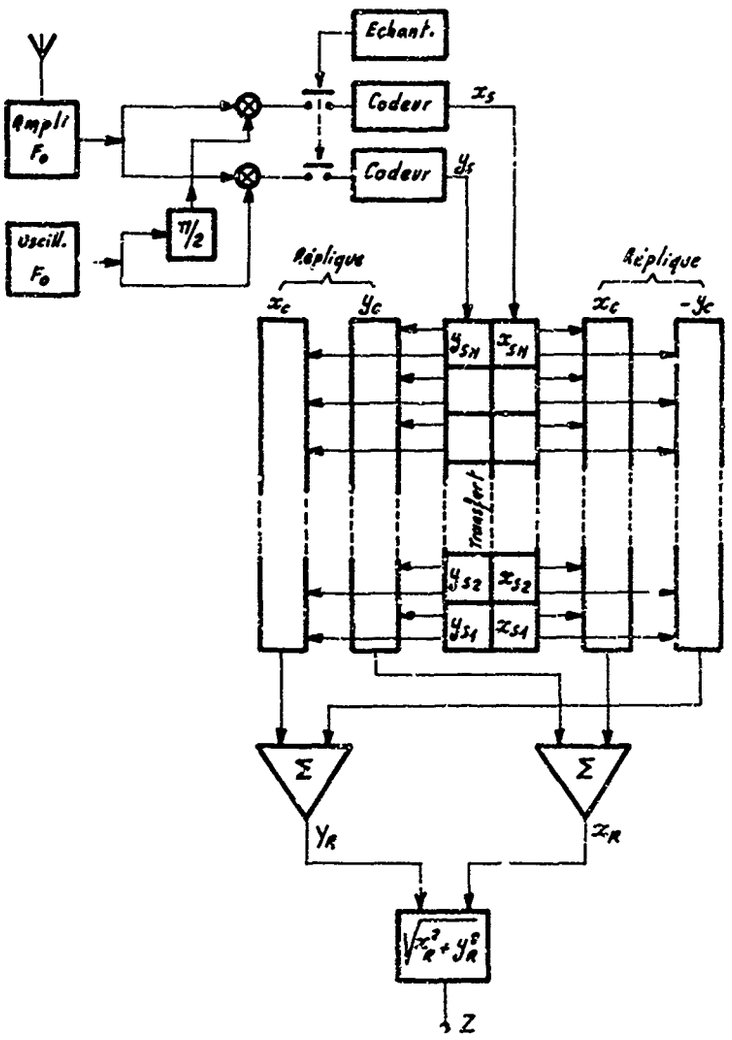


Fig. 2 Correlateur

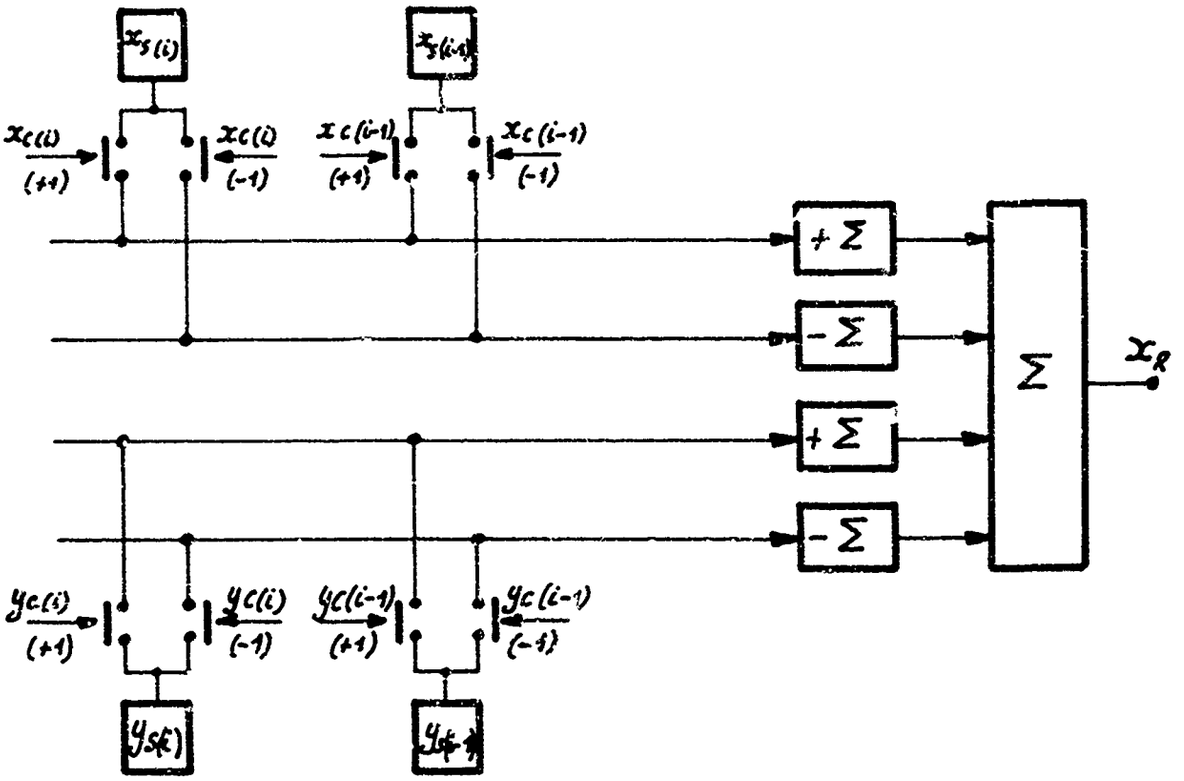


Fig.3 Organisation des calculs

SWITCHED CAPACITOR STORAGE ARRAYS FOR A.M.T.I. AND BANDWIDTH COMPRESSION

by

D.S.Harris

Elliott Automation Radar Systems Limited
Borehamwood, Hertfordshire, UK

TCHED CAPACITOR STORAGE ARRAYS FOR M.T.I. AND BANDWIDTH COMPRESSION

D. J. Harris

ELLIOTT-AUTOMATION RADAR SYSTEMS LIMITED
Borehamwood, Hertfordshire, U.K.

A Member of G.E.C.-Marconi Electronics Limited

SUMMARY

In non-coherent M.T.I. systems, fully coherent M.T.I. systems and clutter locked Stalo-Coho systems, moving targets cause amplitude modulation of either clutter or a reference signal at the detected receiver output. This amplitude modulation is usually separated from the clutter either by bandpass filtering or by subtraction of successive returns in the amplitude-time plane.

A commonly used means of subtraction of successive returns is by the use of an ultrasonic delay line. This method however necessitates accurate control of the transmitted pulse spacing to maintain identity with the store delay time. One of the main sources of cancellation error is usually due to time jitter in the transmitter firing system.

An alternative to the ultrasonic delay line canceller is the sampled range gate canceller described here. In this system two sets of capacitors are used in place of the convertional delay line. Each outgoing transmitter pulse is used to start a precision range clock which is stopped after the maximum range delay has elapsed. This range clock operates a number of video sampling switches to connect each sample of a radar signal in turn to a different capacitor in the store. On successive returns, the storage capacitors are reversed relative to the flow of video current, thus performing subtraction. Being asynchronous, the system has the advantage of being insensitive to pulse-to-pulse jitter and is thus directly applicable to the use of multiple P.R.F.'s. The read-out of subtracted differences can be performed at any time after the delay and before the next transmitted pulse. Each read-out sequence is followed by an erase sequence to prepare the capacitance store for the next period of radar returns.

By adding rather than subtracting successive radar returns, and reading the capacitance store more slowly than writing into it, almost identical circuitry will perform bandwidth compression.

1. INTRODUCTION

This paper describes a multiple capacitor storage technique that can, in principle be used either for integration (i.e. Bandwidth Compression for example, putting a radar picture down a telephone line), or for Moving Target Indicator by the subtraction of successive radar returns.

The moving target indicator processor described here uses the principle of cancellation or subtraction of the video returns from successive pulses within each range gate. In this system the total radar receive time of 63 microseconds, corresponding to a maximum radar range of 5 nautical miles, is divided up into 630 range cells, each 100 nanoseconds wide, corresponding to range elements of 15 metres each. Each "range cell", or "slot" is a separate signal processing system in which subtraction of successive video returns takes place, radar targets which are stationary relative to the radar, return identical video signals from successive transmitter pulses at any particular range, and thus the processor gives zero output after subtraction of these returns. Range cells containing true movement return non-identical signals on successive pulses, and yield finite difference signals.

The canceller uses a capacitor storage system to store the radar video signals prior to the subtraction, one capacitor store for each range cell or slot. Each capacitor store is connected, and then disconnected, to the returning video signals in sequence by diode switching circuits, controlled by a clock.

Three main operational sequences are carried out by these diodes: the feeding of returning video signals into the slots in sequence, called the "write" period; the formation of the difference signals resulting from successive returns and their extraction, called the "read period"; and the erasing from the multiple range slot capacitor store of video signals in readiness for the next radar return, called the "erase" period.

In order to connect the 630 capacitor stores quickly and efficiently to their respective write, read and erase circuits, the capacitor stores are arranged in a "matrix" consisting of rows and columns, such that each capacitor store is only activated for the required period when both its row and its column are addressed. Two capacitors are used in each capacitor store or range slot to enable the video signal to be subtracted from the previous video signal. Thus, each new signal produces subtraction information, as opposed to every new pair of signals in a single capacitor storage system.

A relatively simple re-arrangement of the logic described above allows successive radar returns to be integrated instead of subtracted and this by arranging that the reading sequence be performed more slowly than the writing sequence, bandwidth compression may be achieved.

2. OPERATION OF THE BASIC RANGE SLOT IN THE M.T.I. MODE

The term "range slot" refers to a group of components comprising two storage capacitors and a number of diode switches, in which the amplitude of radar signals, corresponding to a particular range element, is temporarily stored until the next interpulse period. The amplitude difference derived from

successive returns is then "read" out.

The range slots are arranged for ease of addressing in rows and columns to form a matrix. The circuit of the basic range slot is shown in Figure 25-1(a), while Figure 25-1(b) shows the basic range slot connected to the various function switches.

Switches S1 and S3 are the column commutation switches and Switches S4 and S5 are the row commutation switches, designated Y and X respectively in Figure 25-1(b).

Switches S1 and S5, and S3 and S4 work together to direct the video current through the slot capacitors in opposite directions.

Switch S6 is the video switch, and is the only high speed switch in the system, operating for 100 nanoseconds at a time.

Switch S7 is the equivalent to the video switch S6, and operates only during the read period.

Switch S2 is the read-out/erase switch.

Switches S8, S9 and S10 are the erase switches, and operate with Switches S1 and S3 to complete the erase switch system.

Consider now the operational sequence of this single range slot during one interpulse period, and assume that the previous radar return video signals, corresponding to this range slot, of amplitude (A) has been stored on capacitor C2. Also assume that the charge on capacitor C1 due to previous signals has been erased.

"Write" Sequence

In order to convey the incoming video signal, of amplitude (B), into the capacitors, the Y commutation switch S1, the X commutation switch S5 and the video switch S6 are all closed as shown in Figure 25-2(a). The video signals therefore flow via S1, D1, C1, C2, D7, S5 and S6, producing a voltage (+B) on capacitor C1, and a voltage (A - B) on capacitor C2.

"Read" Sequence

To read out the video signal (A - B) from capacitor C2, (See Figure 25-2(b)) the right hand plate of the capacitor is switched to a constant potential, i.e. +12 volts, less the voltage drop across S3 and D7, due to the current flowing down the chain S3, D8, D7, S5, R4 and S7 to earth. The voltage appearing at the left hand plate of the capacitor is then read out relative to the voltage on the right hand plate, via D4 and S2.

"Erase" Sequence

To erase the video signal (A - B) from the capacitor C2, both plates of the capacitor are taken to similar potentials, (See Figure 25-2(c)) i.e. the right hand plate is connected to the potential formed by the current flowing from the +12 volt supply through S3, D8, D6, R3 and S9 to earth, while the left hand plate is connected to the potential formed by +12 volts across S2, D4, D5, R2 and S8. The capacitor may be left with a small error voltage due to the difference of the forward voltage drops across the diodes D4 and D8.

The next (i.e. third) radar return signal applied to this range slot, of amplitude C, say, is switched into the store by closing switches S3 and S4, as shown in Figure 25-3(a), producing a voltage (C) on capacitor C2, and the required difference signal (B - C) on capacitor C1. This difference is read via S2 and erased by closing S1, S2, S9 and S10 as shown in Figures 25-3(b) and 25-3(c) respectively.

The erase period lasts for 25 microseconds.

This description completes the sequence, and the processes shown above for signals A, B and C and are repeated over and over again.

3. OPERATION OF THE MATRIX STORE SYSTEM

3.1. Switching and Storing

The basic range slot of the A.M.T.I. Canceller shown in Figure 25-1 is connected into the matrix store in such a way that it can be addressed by the row and column switching systems. The anodes of the diodes D1, D4 and D8 are connected to a column switching system, while the cathodes of the diodes D2 and D7 are connected to a row switching system. The resistors R1, R2 and R3 are connected to the three erase lines which run through the whole matrix store.

The matrix store is both physically and electrically in two halves. Physically so that, for ease of construction and access, it opens out like a book, and electrically, so as to allow relatively slow switching techniques to be used for the Y column commutation switches. This is done by arranging that when one half of the store is being written into (or read out of), the other half is primed at least 1 microsecond early, thus allowing plenty of time to change over columns.

The store comprising a matrix of 28 columns and 24 rows has a total of 672 slots of which 630 are required to cover 5 n.m. Each column has its own commutation and read out switches. The rows, or X lines, are connected to the X commutation switches S4 and S5, which are in turn driven from the fast video switch S6 during the write period. The video switches S6 are driven from the fast two phase

clock waveforms $\phi 1$ and $\phi 2$.

In order to write video signals into Slot 1 for instance, commutation switches S1 and S5 are closed. When the video switch S6 closes for 100 nanoseconds (controlled by the two phase clock waveform $\phi 1$), the video current charges up the capacitors C1 and C2. To write video signals into Slot 2, Switch S6 in the next X line is closed for 100 nanoseconds. The timing of switch S6 is controlled by the two phase clock waveform $\phi 2$. Waveform $\phi 1$ controls all the odd numbered Y lines, while waveform $\phi 2$ controls all the even numbered lines. The selection of the next X line is controlled by a shift register contained within the X drive modules. This shift register selects the X lines in sequence (X lines numbered 1 to 24) and then back to X line number 1. This process is repeated until all the slots in the column have been written into.

To ensure that the write or read operation is a continuous process, the preceding half column is "primed", i.e. switched on by Switch S1, a half column early. This arrangement ensures that the operational sequence is continuous in going, say, from slot 24 in column 1, to slot 25 at the top of column 2. Smoothness in switching from one slot to the next is essential to preserve constant sensitivity and cancellation with range, and to prevent the generation of repetitive spurious pulses when, for example, switching between columns takes place.

In the following interpulse period, the commutation waveforms change over. Thus the video signals can be written into the capacitors C1 and C2 in the opposite direction, i.e. switches S1 and S5 are open, and switches S3 and S4 are closed. Signal subtractions are therefore performed as outlined in section 2 on the basic slot operation.

The matrix store operation as a Bandwidth Compressor system is similar to that described above for the A.M.T.I. Canceller system and operates under a column and row addressing sequence to fill the store with video information. However as the range slot is a much simpler unit, no commutation switches are required, and thus the system logic is far less complex.

3.2. The Complete Matrix System

All the high speed waveforms and circuits are contained within the matrix store so that the interconnection leads are kept short, thus helping to prevent "crosstalk" problems. Apart from the storage and switching circuits, three other circuit systems are also required within the matrix:

(a) The Video Drive Circuits

These video circuits convert the radar signals applied as a voltage (in 50 Ω.) into a current drive for the capacitor stores. Compensation circuits are also included to ensure that the video amplifier working point is stabilised against temperature changes.

(b) The Clock Pulse Generator

The row or X drive circuits require two phase clock pulses to achieve a continuous switching process by activating switch S6 in the matrix addressing system. The waveforms $\phi 1$ and $\phi 2$ are the two phase clock pulse waveforms. This generator uses a 'primed' oscillator which achieves its full oscillatory amplitude within two or three cycles. High speed switch transistors and T.T.L. integrated circuits assist in the generation of the two phase clock pulses.

(c) Modulus Output Circuit

The capacitors in the individual range slots are arranged in pairs so that the difference signals formed by subtracting successive radar returns may be erased from one capacitor after read-out, whilst leaving unaffected the most recent radar return on the other capacitor. The 'read-out' point is most conveniently taken from the junction of the two capacitors and the voltage stored on either capacitor is read out by connecting a reference potential to its other plate.

This arrangement, whilst resulting in several advantages, such as simple switching circuits and the ability to obtain good cancellation has the disadvantage that any difference signals formed by moving targets are read in antiphase on alternate read-out periods. These antiphase or bipolar signals are converted into a unipolar output by suitable amplifiers and diode detectors.

By their very nature, the switching signals used to drive the diodes at the beginning and end of the read period of each range slot, generate spikes which are comparable with, and in some cases, larger than the required video signals. These switching spikes must be prevented from being displayed, and this is achieved by using a "sample and hold" circuit based on a balanced diode bridge and storage capacitor.

By means of the 5 MHz clock driving a narrow pulse generator to open the bridge, a 20 nanosecond slice is taken from the centre or "quiet" region of each of the range elements during the read period.

In addition to the essentials of the circuit described above, a number of amplifiers and emitter followers are used for buffering, gain adjustment and impedance transformation, as for example, from the high impedance sample and hold circuit to the 50 ohm coaxial output.

The overall gain of the system for a doppler frequency equal to half the p.r.f. (i.e. the frequency of maximum sensitivity) is made equal to unity.

Figure 25-4 shows in the upper portion of the photograph, the bipolar signals as produced at the output of the matrix. This is a photograph taken from an experimental 48 slot breadboard canceller, and shows, due to lead inductances and slow switches, the row and column switching spikes. The column switching spikes are those larger positive spikes clearly visible.

The lower portion of the photograph shows the signal after being processed by the modulus output [illegible]. The row scanning spikes have been eliminated, and the column switching spikes are only just visible. In the completed larger scale canceller, 650 slots, these column spikes due to a slow speed switch are also eliminated.

The A.M.T.I. Canceller system as described in this paper has given cancellation ratios of between 40 - 50 db.

3.3. Subsidiary Switching Sequences

To complete the canceller system operation many other circuits and waveforms are needed to produce the correct timing of the various operations, and to ensure that the "write", "read" and "erase" processes occur with the correct sequence within an interpulse period. These circuits are generally of a much slower nature than those used within the high speed matrix store system and are therefore accommodated on conventional double-sided printed circuit boards.

4. THE PHYSICAL DESIGN CONSIDERATIONS

The physical problems of designing an A.M.T.I. Canceller to the described system are two fold. Firstly there was a need to package all the high speed circuits close together to achieve short, low inductive leads and reduce the problems of cross-talk. Secondly, it was necessary to assemble approximately 11,500 conventional miniature components into a convenient package suitable for airborne equipment use.

It was decided, therefore, in the experimental model to use welded wire techniques in encapsulated 'cordwood' modules, to achieve the reliability and packing density. Indeed, the storage or slot modules contain 156 components at a packing density of approximately 200 components per cubic inch. The terminations to the various modules are by wrapped wire joint to copper bus-bars to ensure accessibility for servicing and ease of removal.

The copper bus-bars stretch in both directions across the matrix store to provide the many connections to the modules both for the power lines and the high speed waveforms. These bus-bars are spaced within a "mother board" and achieve reasonably low capacity one with another.

Conventional printed circuit design and power unit complete the A.M.T.I. Canceller equipment, the whole unit is contained in a ½ A.T.R. long size weighing approximately 25 lbs. and consuming approximately 50 watts of power.

Acknowledgement

Acknowledgement is made to the Ministry of Technology for their support.

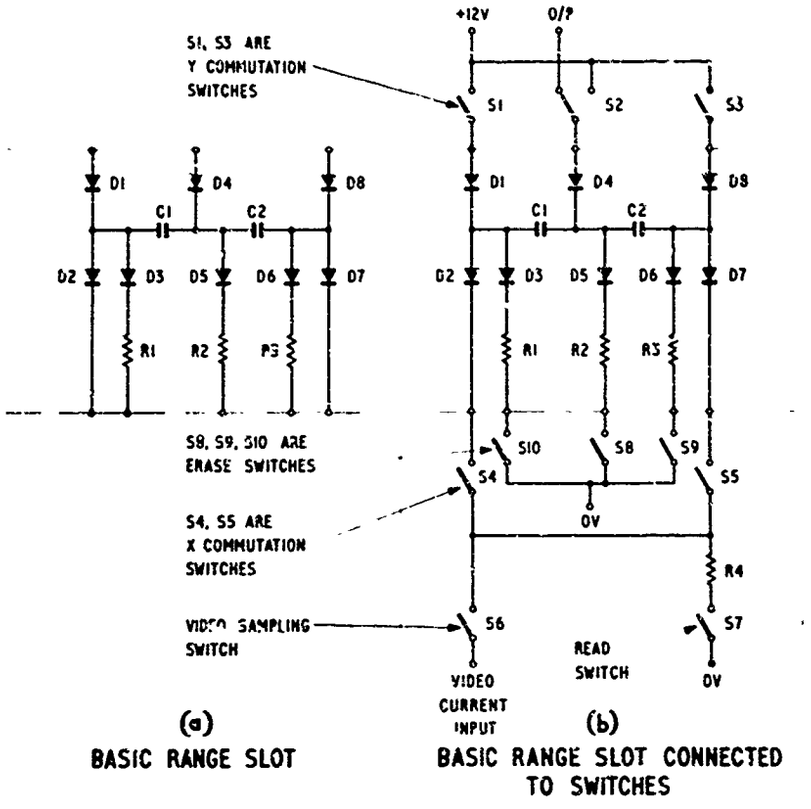


Figure 1

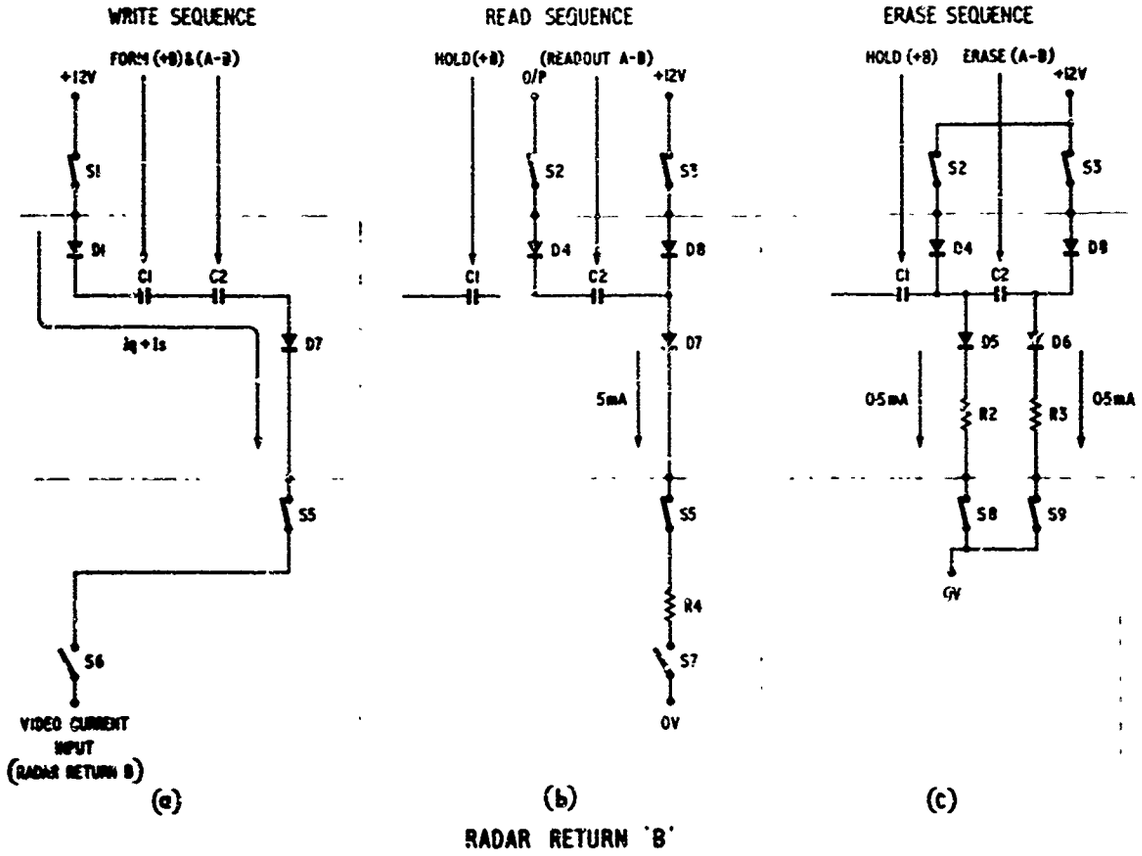


Figure 2

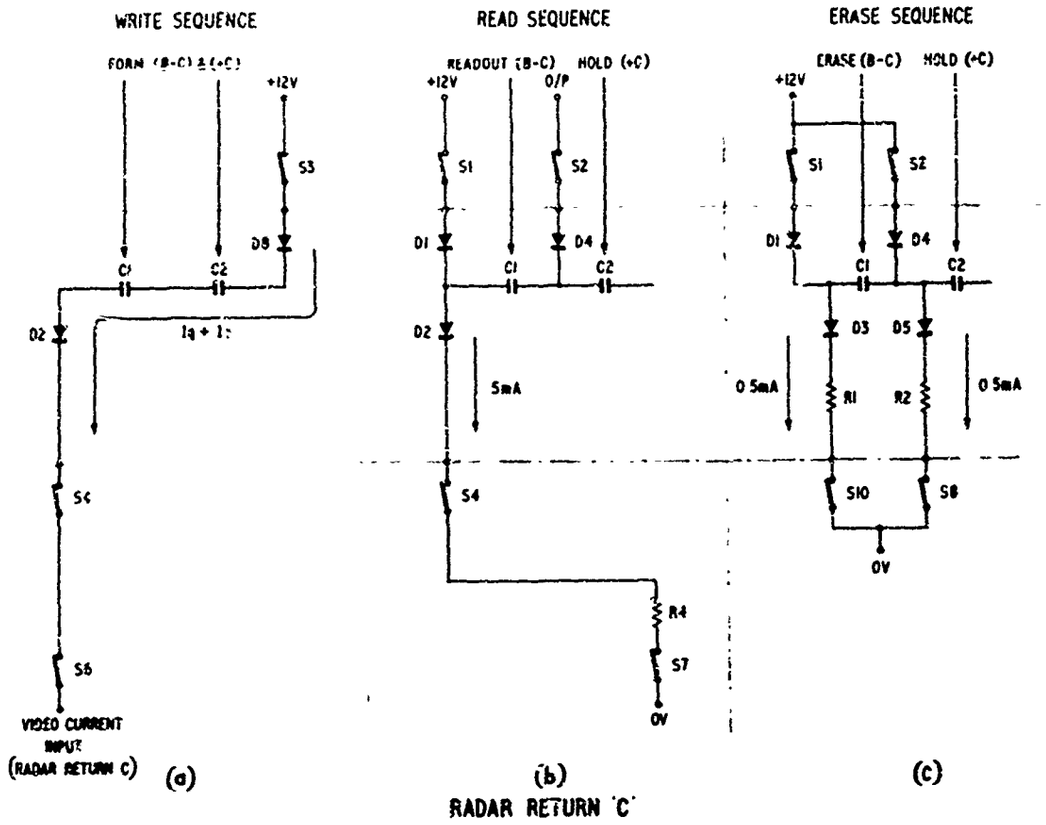


Figure 3

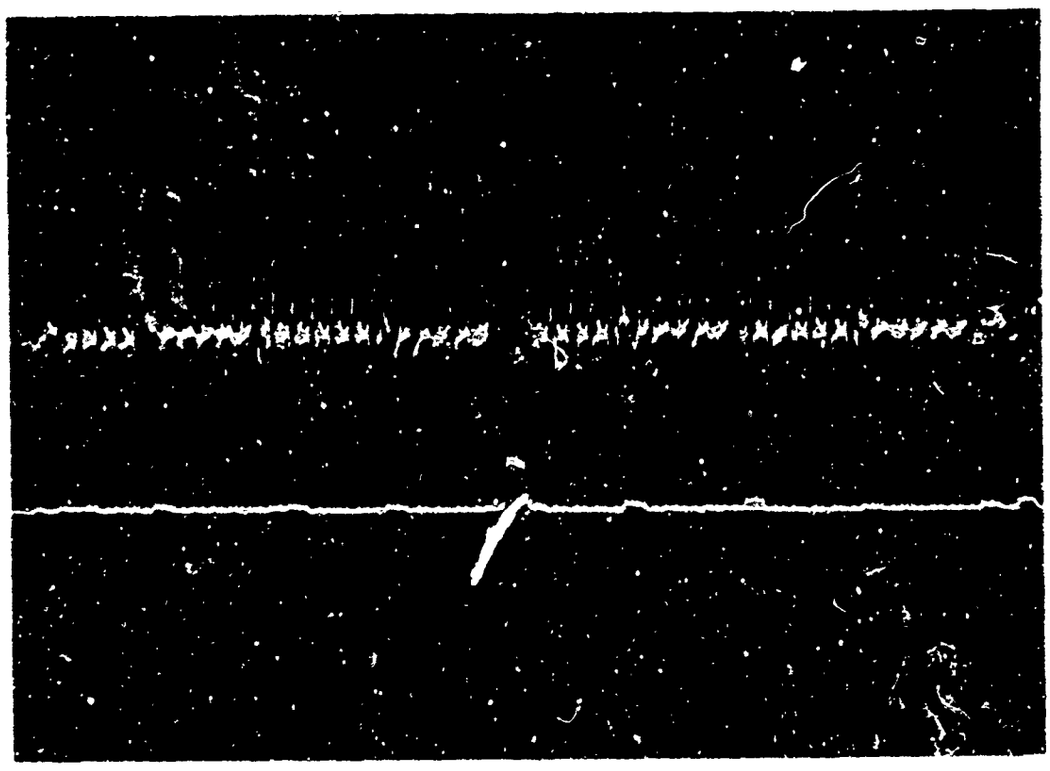


Fig.4 Matrix store output. Modulus circuit output

HYBRID ANALOGUE/DIGITAL CORRELATOR

by

C. E. Jagger

Canadian General Electric Company
Toronto, Canada

HYBRID ANALOGUE/DIGITAL CORRELATOR

C. E. Jagger
CANADIAN GENERAL ELECTRIC COMPANY
Toronto, Canada

ABSTRACT

A hybrid technique is presented which possesses the desirable features of delay precision inherent in digital filters and dynamic range of analogue devices.

Such an ideal seems to be offered by the N-path filter technique, and the "bucket-brigade" delay line, both of which make use of the zero order hold circuit for information storage.

To demonstrate the potential of these devices, elementary matched filters have been implemented at bandpass to operate on signals with time-bandwidth products up to 15.

INTRODUCTION

Literature over the past two decades has disclosed a plethora of methods by which matched filters or correlators may be implemented. For the needs of high resolution radar, these methods usually involve analogue circuitry most often in the form of dispersive delay lines.

These devices, although compact and relatively economical, are useful only for a restricted class of signals; those of the frequency modulated variety such as chirp, logarithmic F. M. and cubic phase law frequency modulation.

Because analogue circuitry is involved, practical limitations on component stability and precision determine the degree of realization of the desired dispersive characteristic and therefore set upper bounds on the complexity of signals that may be handled.

To a lesser extent, other analogue realizations have employed the transversal filter or tapped delay line, wherein, through the sampling theorem, any realizable band-limited impulse response can be constructed from the weighted summation of equi-spaced tap outputs, where the tap spacing is the Nyquist interval.

Signal complexity for this device is easily defined in terms of the spacing and number of taps contributing to the overall sum. Thus, if the delay line is T seconds long and the Nyquist interval or tap spacing $\frac{1}{2W}$ seconds, where W is the low-pass bandwidth of an elemental sample of the desired impulse response, then the time-bandwidth product of the overall sum or impulse response is 2TW.

Realization of the flat delay characteristic necessary for the transversal filter is equally difficult as with the realization of dispersive delay lines, the progressive increase of waveform distortion with delay setting upper limits on signal time-bandwidth products at less than 1000.

Since the problem with analogue devices lies in the manner in which delay errors, timing errors and distortion accumulates, it would seem that digital methods with all the benefits of precise timing from a common clock should offer the requisite panacea.

Such techniques [1] have indeed found wide use at much lower bandwidths than are suitable for radar. In addition, the environment (sonar) in which they have been used, is such as to give rise to relatively low losses in S/N resulting from the restricted dynamic range of binary devices.

For the radar case, however, consideration of dense target environments leads to dynamic range requirements sometimes in excess of 40 db. The equivalent digital transversal filter for this application may be envisaged to consist of a cascade of an analogue to digital converter of 8 bit capability (7 bits plus a sign bit) followed by a compound line of 8 shift registers in parallel. A problem arises in the weighting or multiplying circuitry at each tap. This process must be fast; less than or equal to $\frac{1}{2W}$ seconds. For radar requirements such multiplying speeds are not feasible, and even if they were, the cost of such a device for a high number of taps would be prohibitive.

With the above difficulties in both the analogue and the digital methods in evidence, it is of interest to examine a hybrid of the two techniques for possibly superior but at least more economical realizations. Ideally, one would like to acquire the precision of delay of the digital technique and the dynamic range of the analogue methods.

(i) Hybrid Techniques

One answer to the multiplier speed limitation of the digital delay line is that given in [2] and Figure 1. Here a digital to analogue conversion is performed at each equivalent tap of the compound

delay line and weighting performed in the analogue manner. Undoubtedly, with increasing advances in large scale integration, such a scheme will be the basis of future processors, however at present, the need for such a multiplicity of converters is still somewhat costly, in addition to being bulky.

The desirability of such a device nevertheless exists in its basic universality. Clearly, a simple change of clock frequency accommodates a range of signal bandwidths limited at present only by the rate one can convert from analogue to digital signals.

As an alternative, the properties of capacitor storage in zero and first order hold circuits offer interesting possibilities for implementing processors that obviate the need for A/D or D/A converters.

Although by no means a new idea this method of information storage has recently been made attractive by the rapid advances that have taken place in MOS FET technology, particularly in the developments related to integrated switching arrays.

A well known scheme, employing a first order hold or integrate and dump circuit, is the cross-correlator of Figure 2, here shown in an elementary form adapted to the processing of a Barker code (2TW=3).

Correlators require as many local references, in the form of replicas of the transmitted waveform when matched filtering is desired, as there are independent samples making up the waveform to be processed, i.e. 2TW. Such waveforms, in the most general form are easily produced with 4T W length binary transversal filters, these being shift registers with analogue weighting at the taps.

A variation of the correlator of Figure 2 is the "convolver" [3] of Figure 3, so termed because the local replicas applied as shown are effectively convolved rather than correlated with the incoming signal. Of course, when the replicas are negative time copies of transmitted waveform, the result is a correlation operation identical with that of Figure 2. The operation of this device on a Barker 3 code is also depicted in Figure 3 and is self-explanatory.

A closer approximation to the transversal filter avoiding thereby the need for local reference generators, is the "bucket-brigade" [4, 5, 6, 7] transversal filter shown in Figure 4.

Here, as with the convolver or N-path filter of Figure 3, the properties of the zero-order hold circuit are employed to effect the memory function.

It is with the "bucket-brigade" and N-path filter realizations that this paper is concerned, implementation of both having been made to process a code length of 2TW=13 with $\frac{1}{2W} = 200 \times 10^{-9}$ seconds.

(ii) "Bucket-brigade" Storage Principle

Referring again to Figure 4, two unit delay elements are depicted wherein all even numbered switches, and all odd numbered switches operate respectively on alternate phases of the clock. Thus in the state shown S_1 has connected C_1 to the input, S_3 has connected C_3 to receive a charge proportional to the voltage across C_2, similarly C_5 through S_5 charges to some voltage proportional to the voltage on C_4, etc.

In the alternate phase of the clock all odd numbered switches open and the even ones close and C_2 charges to a voltage proportional to that across C_1 and so on in this manner the input signal is transferred along the delay line. Ideally the amplifier-switch combination should possess a gain of unity, infinite input impedance and zero output impedance.

To the extent that such an ideal is obtainable, one should be able to construct lossless delay lines of arbitrary length possessing the properties of analogue dynamic range and a digital accuracy of delay. Also, as with the digital devices, simple alteration of the clock frequencies allows filtering to be adapted to a broad range of signal bandwidths.

Practically such ideals are component limited, restricting the allowable range of clock frequencies that may be employed. Achievable performance at present has been reported to be: 30 MHz for a hybrid integrated device announced in [5] and 800 kHz to 50 Hz for a MOS FET integrated circuit [6]. Both designs give upper to lower sampling frequency ratios of about 1500: 1.

From Figure 4 it will be recalled that two capacitors are required to supply one unit = $\frac{1}{2W}$ seconds of delay. One therefore requires 4TW storage capacitors for a signal dimensionality of 2TW. When transversal filtering is required, the need for access to independent taps allows no simplifications, however where only a delaying operation is involved, more efficient use of the capacitor and switch elements is possible by use of a poly-phase clock.

In effect, what is desired is that each capacitor hold or store a given sample of information for an interval approaching $\frac{1}{2W}$ seconds as opposed to the $\frac{1}{4W}$ seconds the sample is held when a two phase clock as above is used.

Thus, instead of a transference taking place, each $\frac{1}{4W}$ seconds, the switch operation between adjacent capacitors can be made to take place at an interval of $\frac{P-1}{P} \times \frac{1}{2W}$ seconds with a P phase clock. Hence the number of capacitors required for a delay T seconds becomes $\frac{P}{P-1} \times 2WT$.

It will, however, be noted that each adjacent capacitor stores an overlap of information over a time $\frac{1}{2WP}$ seconds denying easy access to independent samples for the purpose of transversal filtering, that is except for the case P = 2 where every other capacitor stores an independent sample. Also, a difficulty arises in the reduced sampling time occasioned by an increase in P. Nevertheless, for low frequency applications, a useful reduction in component count is possible with this alternative.

(iii) Basic Circuitry

Although MOS FETs offer the greatest promise for both the unity gain amplification and switching operations, the presently attainable speed, magnitude of gate leak through, and for that matter, cost, fall somewhat short of the requirement for a matched filter operating at a clock rate of 5 MHz.

Instead, the particular device employed was the four-transistor bridge of Figure 5 [7]. This circuit simultaneously satisfies the requirement for fast switching, near unity voltage gain, high "off-state" input impedance of 10^7 to 10^8 ohms and low (about 5 ohms) output impedance.

The form in which a unit delay element was implemented is given in Figure 6. With the power supply voltages indicated, the signal dynamic range is 8 volts peak to peak, loss after 15 cascaded unit delays 0.6 db and linearity in excess of 70 db.

The N-path filter implementation given in Figure 7 is self-explanatory. An interesting feature of this circuit is the economy possible when a restriction is made to use with binary phase codes only. Here, in effect, each channel is comprised of only one 4 transistor integrated array, the remaining circuitry being common to all channels.

(iv) System Considerations

The specific problem in radar ranging is that of estimation of a signal, known except for amplitude and time of occurrence. When the interfering noise processes are white, stationary, and additive, the best estimate results from a cross-correlation between received signal and transmitted waveform [8].

To perform such an operation in real time requires as many delayed replicas of the transmitted signal as there are resolvable range cells within the duration of the signal. This number is of course the time-bandwidth product 2TW of the signal which in turn specifies either the number of multipliers in the correlator or the number of taps on the transversal matched filter.

The previous discussion has dealt only with correlators at low-pass. Radar returns are of course at band-pass and cross-correlation would additionally require a precise knowledge of received carrier phase.

Failing such knowledge, as is generally the case, detection theory [8, 9] requires that the envelope of the cross-correlation between bandpass replicas of the transmitted waveform and the received waveform be determined and compared with some specific decision criteria.

As an illustration of the required processing of band-pass signals and an indication of the means by which these may be processed at low-pass, the following development is useful:

Let the transmitted signal be represented by s(t) where:

$$s(t) = a(t) \cos \left[\omega_o t + \varphi(t)\right] \tag{1}$$

This is a general bandpass signal modulated in amplitude by a(t) and in phase by $\varphi(t)$.

In analytic signal form (1) may be written

$$s(t) = \mathrm{Re}\, z(t) \exp j\, \varphi(t) \tag{2}$$

where

$$z(t) = u(t) + j\, \hat{u}(t) \tag{3}$$

$z(t)$ is termed the pre-envelope of the real signal $u(t)$ which from (1)

$$a(t) \cos \omega_o t \tag{4}$$

$\hat{u}(t)$ is defined as the Hilbert transform of the real signal $u(t)$. For signals for which the carrier term is greater than the low pass bandwidth of the signal, $\hat{u}(t)$ becomes:

$$\hat{u}(t) = a(t) \sin \omega_o t \tag{5}$$

and from (2), (3), (4), and (5), we can write

$$s(t) = \mathrm{Re}\, a(t) \exp j\left[\omega_o t + \varphi(t)\right] \tag{6}$$

Upon reception, the only acquisition for which the receiver does not have a locally stored reference is an unknown phase angle λ (noise, loss and doppler are not included in this development). Thus we can, without loss of generality, ignore the round trip time delay and write the received signal as:

$$y(t) = \mathrm{Re}\, a(t) \exp j\left[\omega_o t + \varphi(t) + \lambda\right] \tag{7}$$

Cross-correlating (6) and (7) gives: [7]

$$R_{ys}(t) = \mathrm{Re}\, \frac{1}{2} \int_{-\infty}^{\infty} a(\tau) \exp j\left[\omega_o \tau + \varphi(\tau) + \lambda\right] \cdot a(\tau - t) \exp j\left[\omega_o(\tau - t) + \varphi(\tau - t)\right] d\tau \tag{8}$$

Equation (8) may be re-written as:

$$R_{ys}(t) = \mathrm{Re}\, \frac{1}{2} (\exp j\lambda) \rho_s(t) = \mathrm{Re}\, \frac{1}{2} (\exp j\lambda)(\exp j\omega_o t) \Phi_s(t) \tag{9}$$

where $\rho_s(t)$ is defined as the auto-correlation of the complex transmitted signal $a(t) \exp j\left[\omega_o t + \varphi(t)\right]$. $\Phi_s(t)$ is the auto-correlation of the complex low pass signal.

The desired quantity is the envelope of the cross-correlation function $R_{ys}(t)$ and this is given as the absolute value of $\rho_s(t)$ (or $\Phi_s(t)$) [9]

Thus

$$\mathrm{Env.}\ R_{ys}(t) = \frac{1}{2}\left|\rho_s(t)\right| = \frac{1}{2}\left|\Phi_s(t)\right| \tag{10}$$

which is independent of carrier phase λ and carrier frequency ω_o.

In a bandpass correlator or matched filter, this result is simply achieved by rectification or detection of $R_{ys}(t)$. However, since bandpass correlators require a delay precision related to the carrier frequency rather than the information bandwidth 2W, translation is often made to a carrier frequency slightly greater than the low pass bandwidth W, or to zero frequency wherein in-phase and quadrature channels are necessary to avoid loss of information. For the latter system the indicated operations may be illustrated by further expansion of $\rho_s(t)$ in (10)

Hence:

$$\rho_s(t) = 2\left[R_s(t) + j\, \hat{R}_s(t)\right] \tag{11}$$

Where

$R_s(t)$ = the auto-correlation of the real signal $s(t)$

and

$\hat{R}_s(t)$ = Hilbert transform of $R_s(t)$ = the cross-correlation between $s(t)$ and $\hat{s}(t)$, and

$$\mathrm{Env.}\ R_{ys}(t) = \sqrt{R_s^2(t) + \hat{R}_s^2(t)} \tag{12}$$

Equation (12) may be interpreted to indicate that the received signal y(t) must be correlated not only with s(t) the real transmitted signal but also with

$$\hat{s}(t) = \text{Im } a(t) \exp j[\omega_o t + \varphi(t)] \tag{13}$$

to yield the desired envelope.

Equivalently the same result obtains if s(t) is convolved with s(T-t) and $\hat{s}(T-t)$ where s(T-t) and $\hat{s}(T-t)$ are the respective in-phase and quadrature matched filter impulse responses.

Further expanding s(t) and $\hat{s}(t)$ and recalling $\hat{R}_s(t) = R_s \hat{s}(t)$ gives the required correlator configuration. Thus:

$$s(t) = a(t) \cos \varphi(t) \cos \omega_o t - a(t) \sin \varphi(t) \sin \omega_o t \tag{14}$$

and

$$\hat{s}(t) = a(t) \sin \varphi(t) \cos \omega_o t + a(t) \cos \varphi(t) \sin \omega_o t \tag{15}$$

Figure 8 depicts the indicated correlator for low-pass processing. In Figure 9 is shown the in-phase and quadrature processing operation of a matched filter on a simple Golomb polyphase code.

At video, this code is defined by the complex waveform 1, j, 1 where the commas represent a contiguous series of rectangular pulses of duration $\frac{1}{2W}$.

The auto-correlation of this code is found through application of Bernfeld's algorithm [10] wherein the above sequence is simply multiplied by its reversed time conjugated replica.

Thus

$$\begin{array}{r} 1\ ,\ j\ ,\ 1 \\ \underline{1\ ,-j\ ,\ 1} \\ 1\ ,\ j\ ,\ 1 \\ -j\ ,\ 1\ \ -j \\ \underline{\qquad 1\ ,\ j\ ,\ 1} \\ 1\ ,\ 0\ ,\ 3\ ,\ 0\ ,\ 1 \end{array}$$

Auto-correlation = Sum =

In Figure 9 the input signal y(t) at bandpass with arbitrary phase shift λ is represented as:

$$y(t) = a(t) \cos(\omega_o t + \lambda) + a(t-\Delta) \sin(\omega_o t + \lambda) + a(t-2\Delta) \cos(\omega_o t + \lambda)$$

where

$$a(t) = \text{rect}(t) = 1, \ \frac{\Delta}{2} < t < \frac{\Delta}{2}$$

As will be seen from Figure 9, the indicated operations yield the desired envelope of the auto-correlation function.

(v) Experimental Filter and Results

Depicted in Figure 10 is a block diagram of a bandpass filter utilizing two identical "bucket-brigade" transversal filters programmed to match a Barker 13 code. The carrier frequency f_o is 30 MHz and the I. F. bandwidth 2W = 5 MHz.

It will also be noted that an alternative to the operations of squaring, summing, and square-rooting, has been employed in the deriviation of the envelope. This scheme termed "double translation" involves re-modulation of the in-phase and quadrature outputs of the respective delay lines on to in-phase and quadrature carriers. Summation of the outputs of this process gives a signal g(t) where:

$$g(t) = \sqrt{I^2 + Q^2} \cos\left[\omega_o t - \tan^{-1}\frac{Q}{I}\right]$$

This signal is identical to the output of a bandpass matched filter at carrier frequency ω_o and the envelope is correspondingly obtained through detection.

Oscilloscope photographs given in Figures 11, 12, 13, 14 and 15 show various performance characteristics of the device.

In particular, Figure (12) indicates the output from the in-phase or quadrature filters at low-pass when the input carrier is slightly offset from the filter local oscillator. This offset gives a cyclically varying unknown phase angle and demonstrates the independence of phase of the detected resultant shown in the upper trace of Figure (12).

Amplitude linearity of the filter is demonstrated in Figure 13 and 14 which are two views of the composite response surface or ambiguity function of this particular code. This surface was generated by deviating the signal source carrier frequency by ± 2.5 MHz.

As a further illustration of the "bucket brigade" delay line linearity, a binary phase, multi-amplitude Huffman impulse equivalent code was matched with the result shown in Figure (15).

Results for the N-path realization are virtually identical except that the particular form (Figure 7), since it does not use a true 4 quadrant multiplier, cannot be programmed for a Huffman code.

(vi) Effect of Sampling

As distinct from a pure analogue transversal filter which is a time invariant filter, sampling devices such as correlators, "bucket brigade" lines and digital filters are time-variant filters and as such suffer an average processing loss when noise and realizable signals are considered.

Because of the sampling operation, it is necessary to band-limit the received signal plus noise to avoid or minimize spectrum folding or aliasing of the noise around the sampling frequency, into the baseband.

On the other hand, band-limiting gives rise to inter-symbol interference especially for a target of unknown position relative to the local clock or sampling epoch. This situation is depicted in Figure (16) in which is shown two possible outputs of the sampler operating on a band-limited pulse input.

Thus, one may expect that excessive band-limiting will deteriorate the S/N ratio through production of inter-symbol interference, whereas inadequate band-limiting increases the aliased noise power in the baseband [11, 12].

Improper sampling epoch as depicted in Figure (16) gives rise to a "range-cusping loss" which varies from 4.5 - 5.0 db midway between samples to less than 1 db, dependent on the band-limiting filter employed, at the optimum sampling epoch. [13]

Average loss over all sampling epochs is about 2.5 db which reduces to 1.2 db if the I and Q filters are made twice as long and the sampling rate doubled. [13]

(vii) Limitations

The sampling speeds presently obtainable with sampling oscilloscopes indicate few limitations for the N-path filter technique in this area. However, replica generation and four quadrant multiplier technique at the present state of the art would appear to place a practical upper bound on $\frac{1}{2W}$ of about 10×10^{-9} seconds.

The "bucket-brigade" delay line, on the other hand, is a cascade device and errors therefore cumulative. Such errors are comprised largely of the effects of distortion in the switch-amplifier circuits and those due to feed-through because of the unavoidable capacitances and parasitic reactances of the switches and circuit layout.

Of the two problems, distortion accumulation is the lesser, this being relatively easily compensated for any practical length of filter or delay line.

Operation at high speeds, i.e. $\frac{1}{2W}$ small, brings up a significant effect from feed-through capacitances, this becoming of ever increasing proportion in relation to the desired signal sample. Neutralization at appropriate points, along the delay line offers some control of this effect. Better devices, forthcoming should also improve the situation.

At present, the highest speed reported is 30 MHz or $\frac{1}{2W} = 33 \times 10^{-9}$ seconds [5].

Recirculation tests, wherein the output of the delay line is fed back to the input, on the circuit reported upon in this paper, have indicated the realizability of 2TW values up to 1000 with little difficulty.

(viii) Conclusion

The methods indicated in this paper are capable of realizing matched filters of sufficient accuracy and speed for a large number of applications in the radar systems of to-day.

The attractiveness of such techniques is further enhanced by the promise offered by integrated circuit technology which for some applications should permit realization of signal processors which are more economical in both cost and bulk than either of the equivalent analogue or digital implementations.

(ix) Acknowledgement

The work herein reported upon has been supported in part by the Directorate of Industrial Research, Defence Research Board of Canada. DIR Project E-18.

REFERENCES

1. McArthur, R. J., "A Hybrid Analog/Digital Pulse Compression Technique". General Electric Technical Information Series No. RQ69ELS-6, May 1969.

2. Allen, W. B., and Westerfield, E. C., "Digital Compressed Time Correlators and Matched Filters for Active Sonar". NATO Advanced Study Inst. on Signal Processing (presented at Grenoble September 14-26, 1964).

3. Franks, L. E., and Sandberg, I. W., "An Alternative Approach to Realization of Network Transfer Functions - the N-Path Filter". Bell Sept. Tech. Journal 1960, 39 pp 1321-1350.

4. Hannan, W. J. et al. "Automatic Correlation of Timing Errors in Magnetic Tape Recorders". IEEE Transactions on Military Electronics, Vol. MIL-9 pp 246-254, July-October, 1965.

5. Sangster, F. L. J., and Teer, K., "Bucket-Brigade Electronics - New Possibilities for Delay, Time-Axis Conversion and Scanning". IEEE Journal of Solid-State Circuits, Vol. SC-4, No. 3, June 1969, pp 131-136

6. Mao, Roger A. et al., "Integrated Mos Analog Delay Line". IEEE Journal of Solid-State Circuits, Vol. SC-4, No. 4, August 1969, pp 196-201.

7. Baker, R. H., McMahon, R. E., and Burgess, R. G., "The Diamond Circuit". Tech. Report No. 300. Massachusetts Institute of Technology, Lincoln Lab., January 30, 1963.

8. Woodward, P. M., "Probability and Information Theory, with Applications to Radar". McGraw-Hill Book Co., Inc. 1953, Chapter 7.

9. Rihazek, August W., "Principles of High-Resolution Radar". McGraw-Hill, 1969, Chapter 3.

10. Cook, C. E., and Bernfeld, M., "Radar Signals - An Introduction to Theory and Application". Academic Press, 1967, p 245, Chapter 8.

11. Luggannani, R., Intersymbol Interference and P[illegible]ability of Error in Digital Systems". IEEE Transactions on Information Theory, Vol. IT-15, [illegible]6, November 1969, pp 682-689.

12. Ackerman, C. L. et al., "Theoretical Basis and Practical Implications of Band-Pass Sampling". Proceedings of the National Electronics Conference, Vol. 18, 1962. pp 1-9.

13. Nathanson, Fred E., "Radar Design Principles". McGraw-Hill Book Co., 1969. pp 477-478 and [illegible]58.

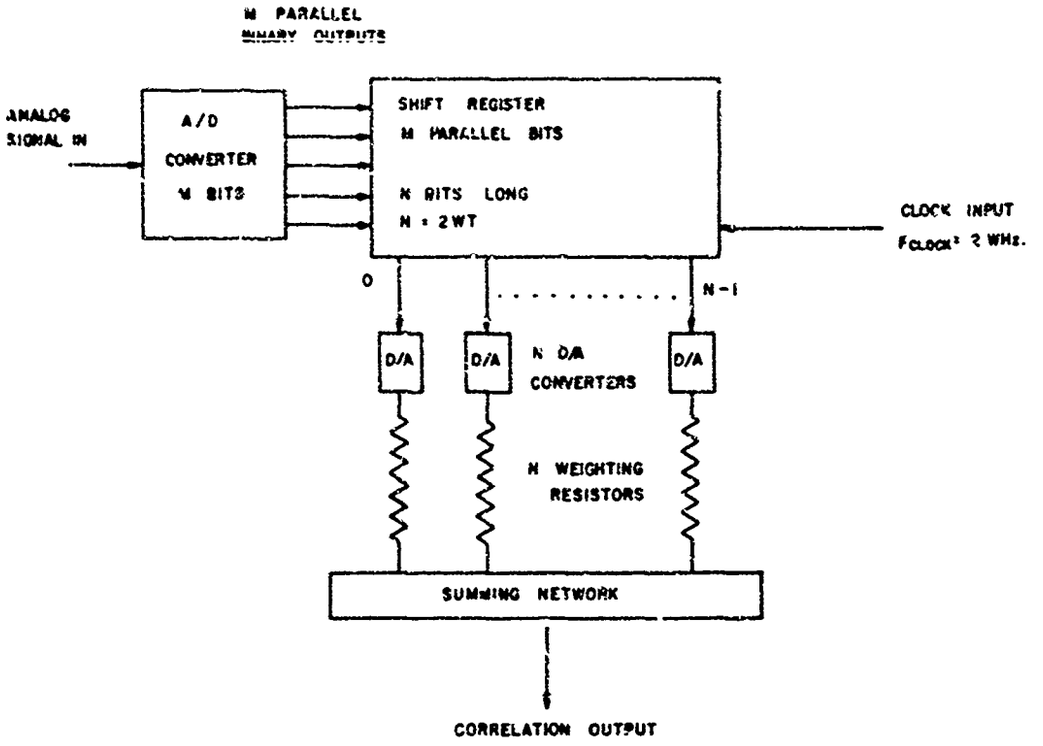


Fig.1 Hybrid analog/digital transversal filter featuring digital storage and analog processing

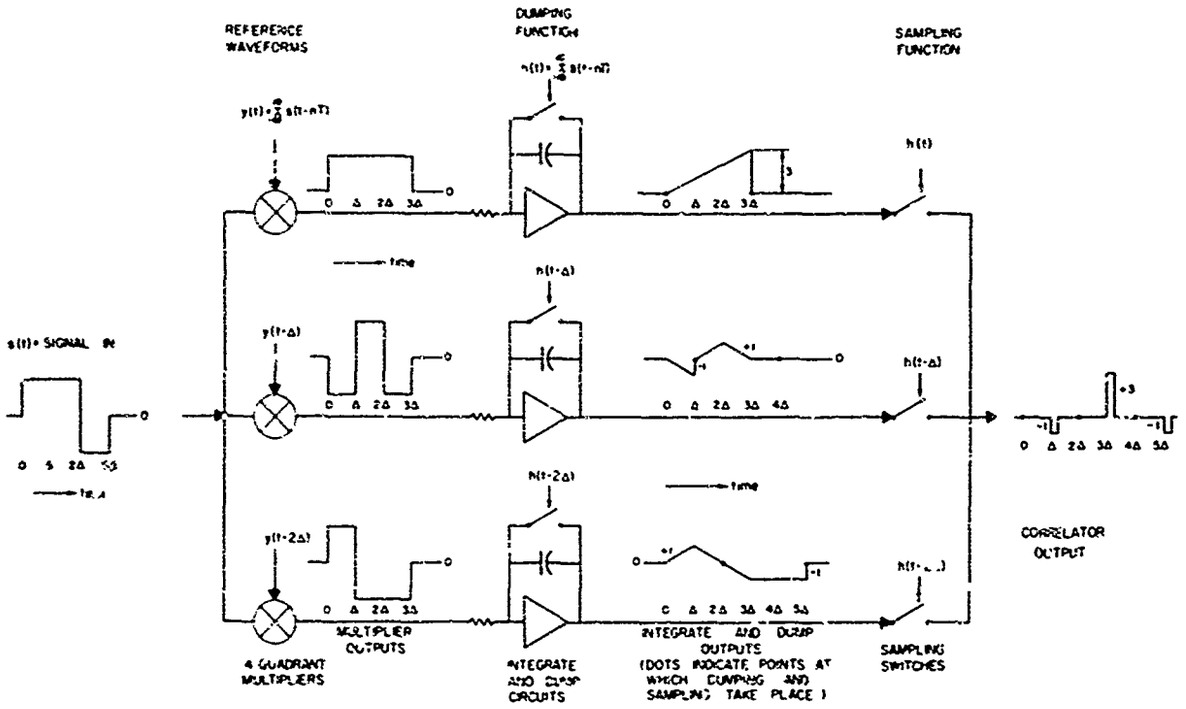


Fig.2 N-path cross correlator configuration for a Barker 3 code

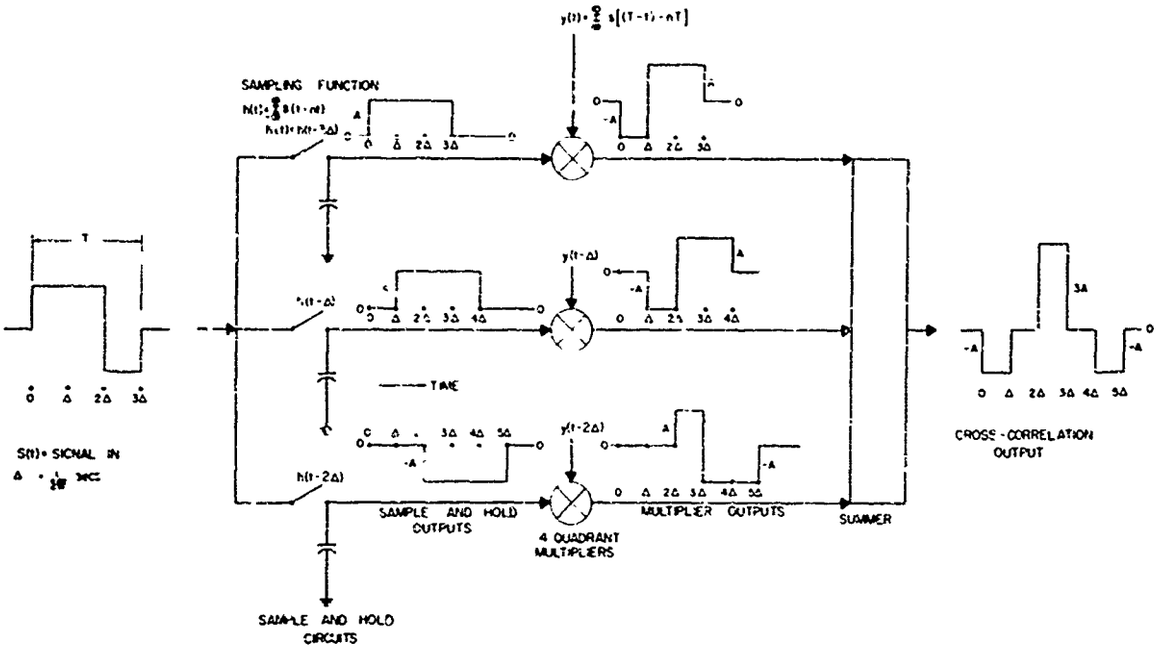


Fig.3 N-path matched filter configuration for a Barker 3 code

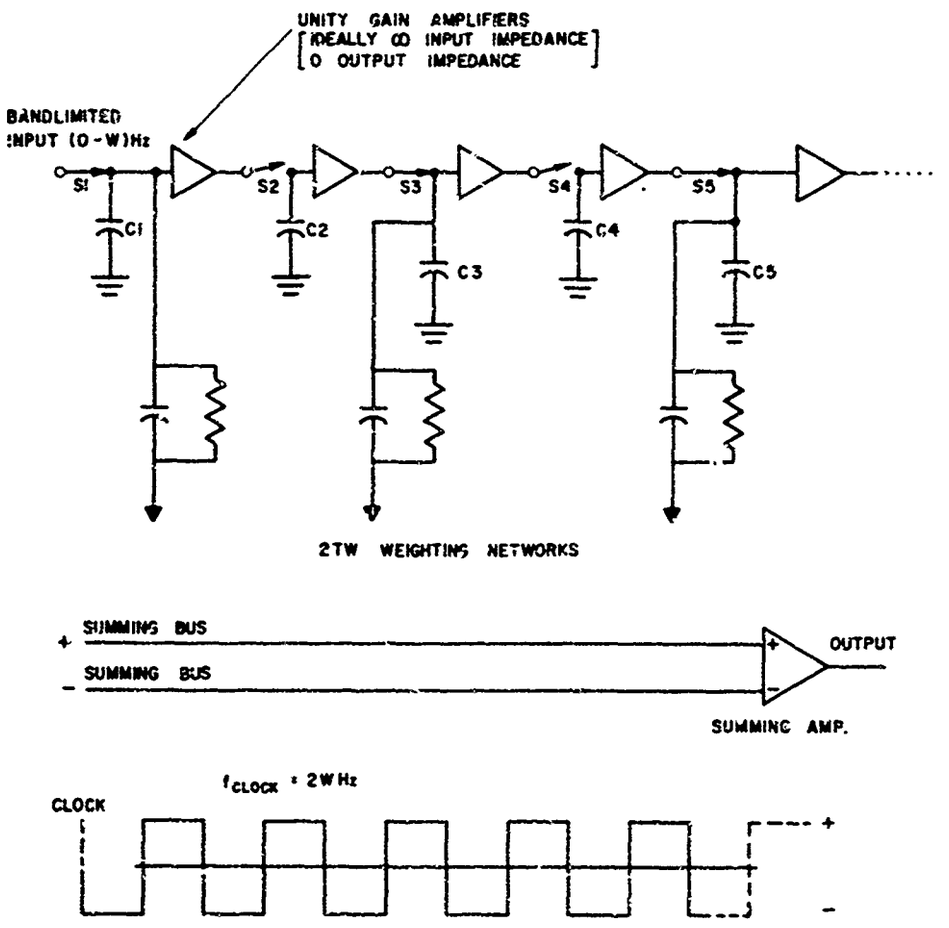


Fig.4 Bucket-brigade transversal filter

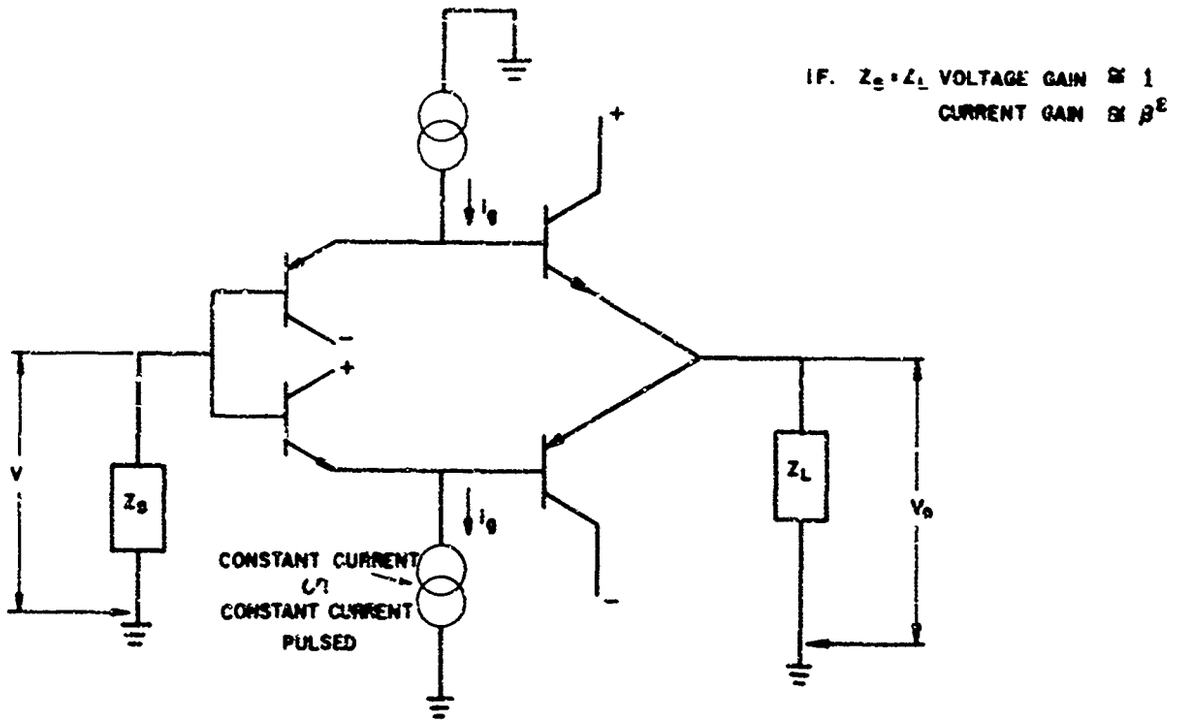


Fig.5 Basic diamond circuit

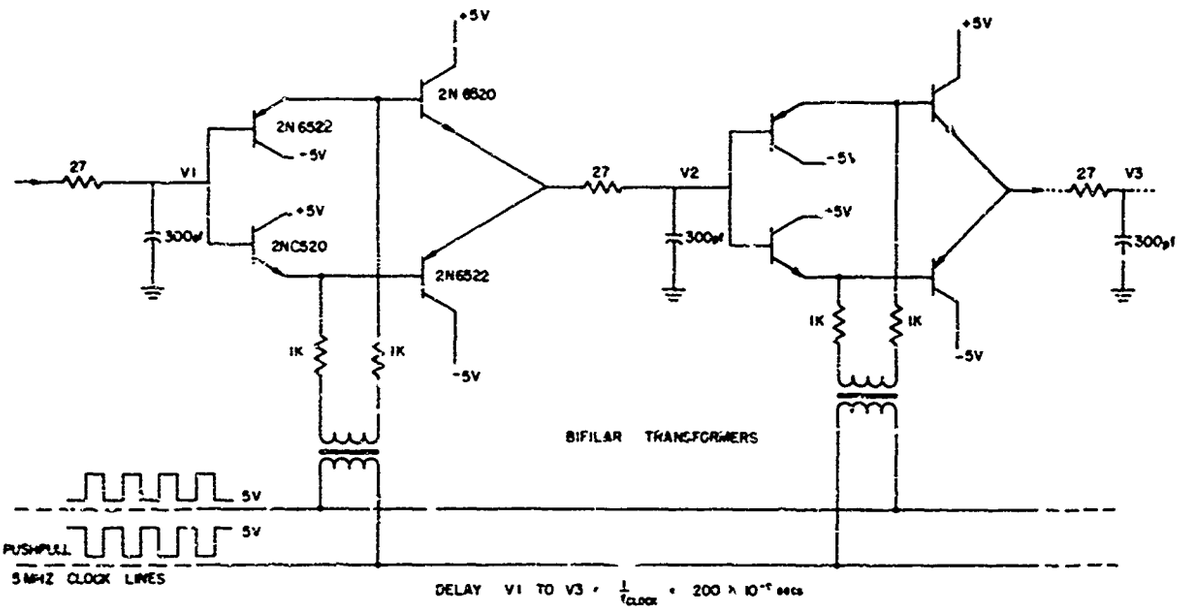


Fig.6 Circuit of a unit delay element of the bucket-brigade delay line

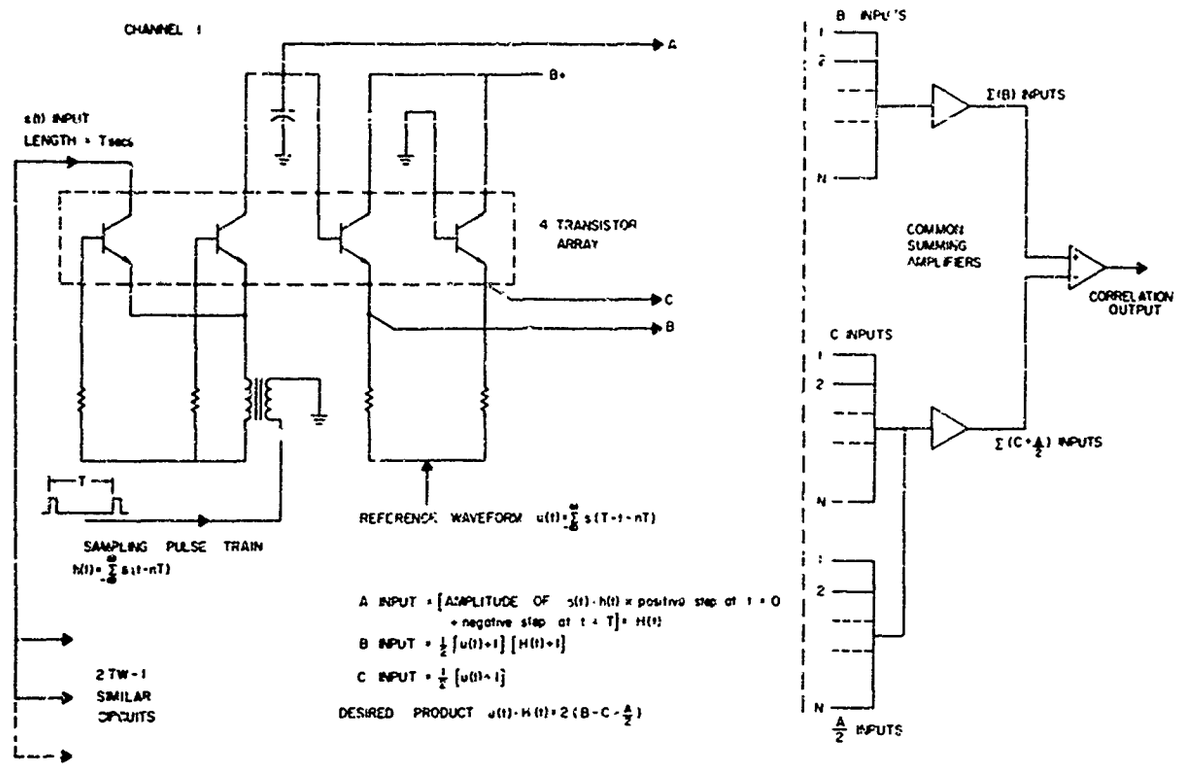


Fig.7 N-path matched filter realization suitable for binary phase codes

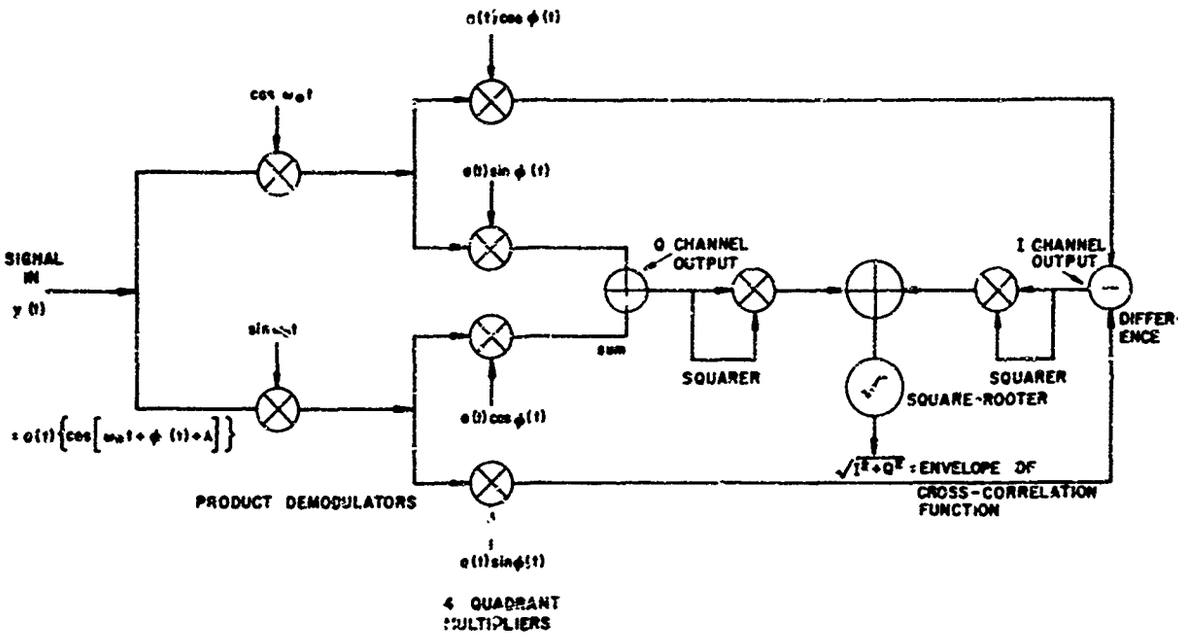


Fig.8 I.F. cross-correlator employing I (inphase) and Q (quadrature) processing of a band pass signal

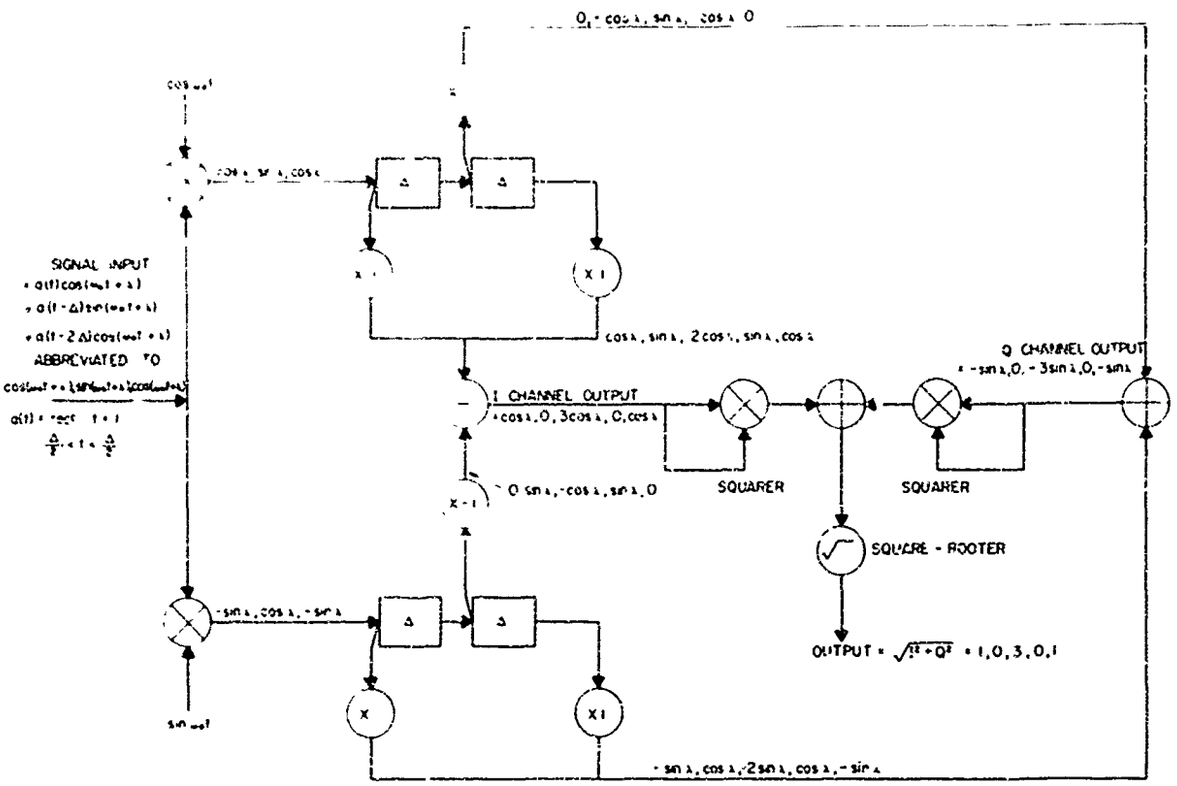


Fig. 9 Inphase and quadrature processing of a Golomb polyphase code with arbitrary phase shift Δ

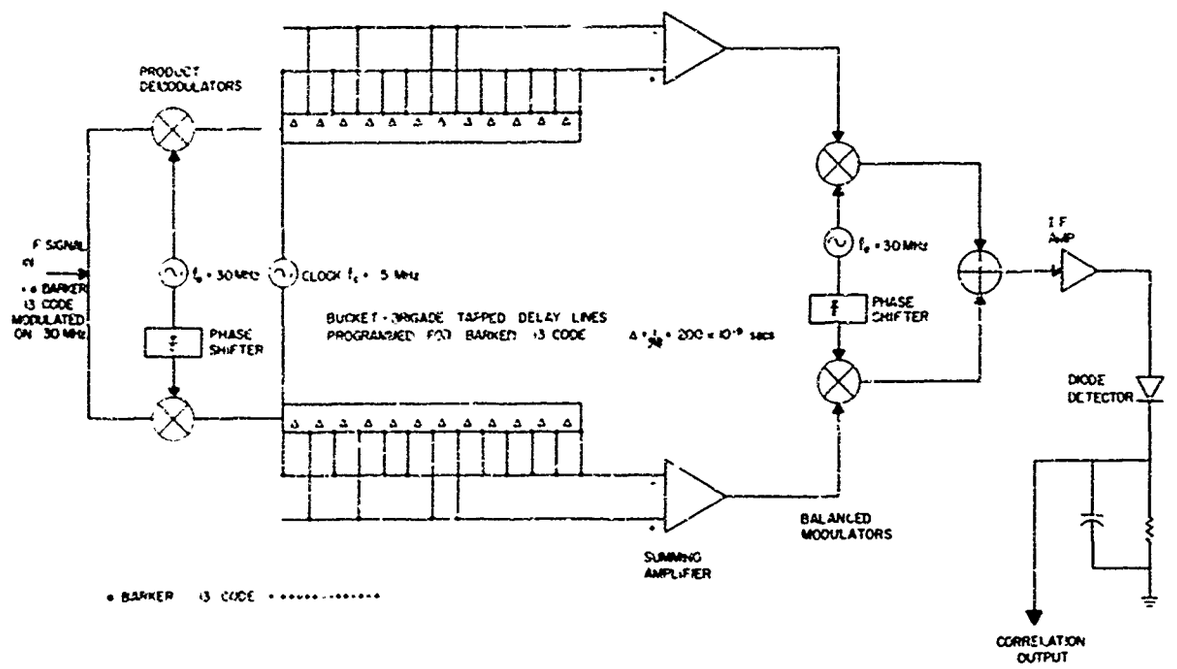


Fig. 10 Block diagram of experimental system using double translation

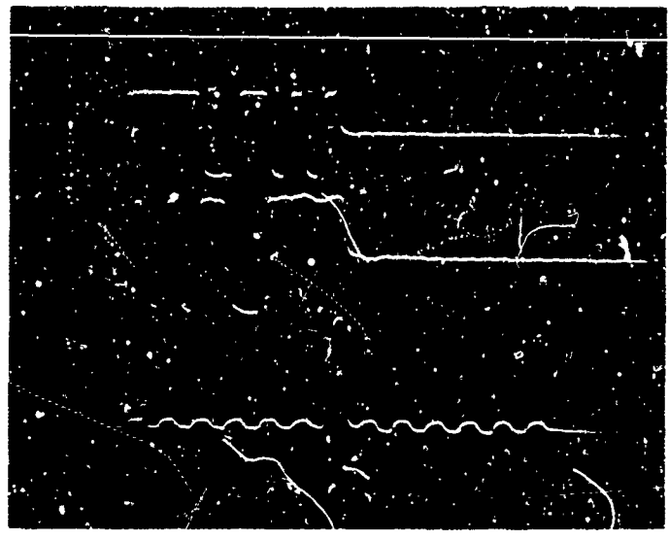


Figure 11

Upper trace : Barker 13 code.
Middle trace: Impulse response of bucket-brigade matched filter programmed for Barker 13 code.
Lower trace: Auto-correlation of Barker 13 code

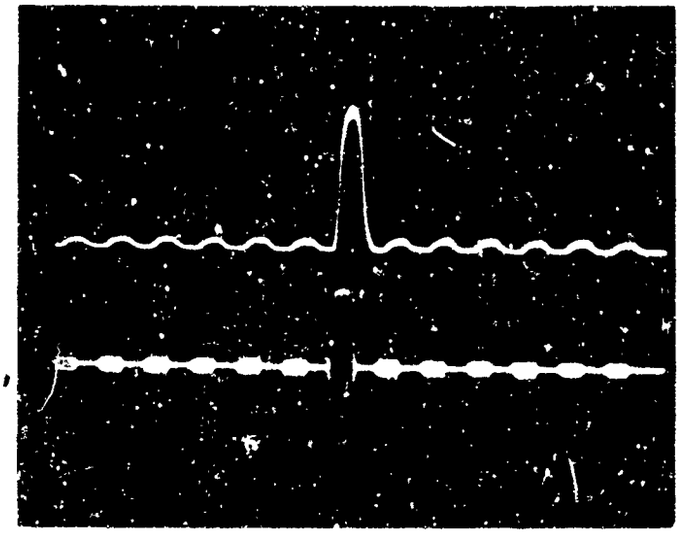


Figure 12

I.F. Matched filter

Lower trace: Output of the I or Q matched low-pass filter in response to a Barker 13 code modulated on a slightly offset carrier frequency.
Upper trace: Envelope of auto-correlation function.

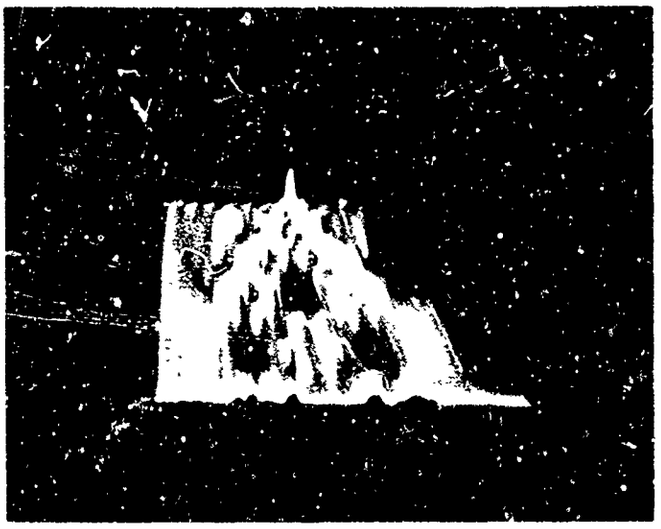


Figure 13

Auto-ambiguity surface of Barker 13 code.
Carrier frequency swept from 30 to 35 MHz)

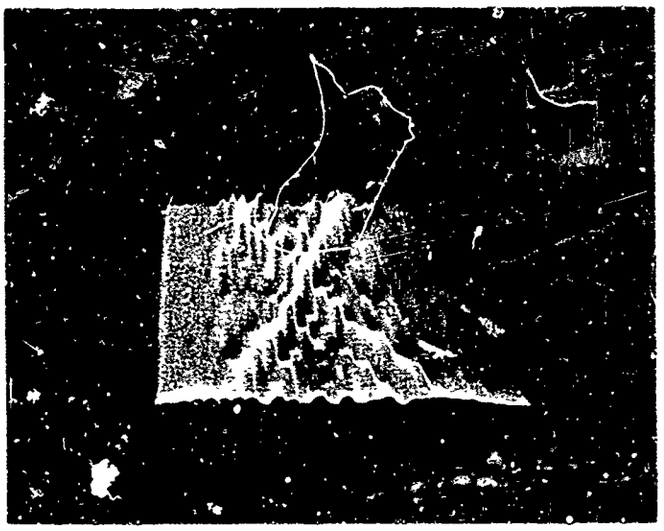


Figure 14

Auto-ambiguity surface of Barker 13 code.
Carrier swept 12.5 MHz about 30 MHz)

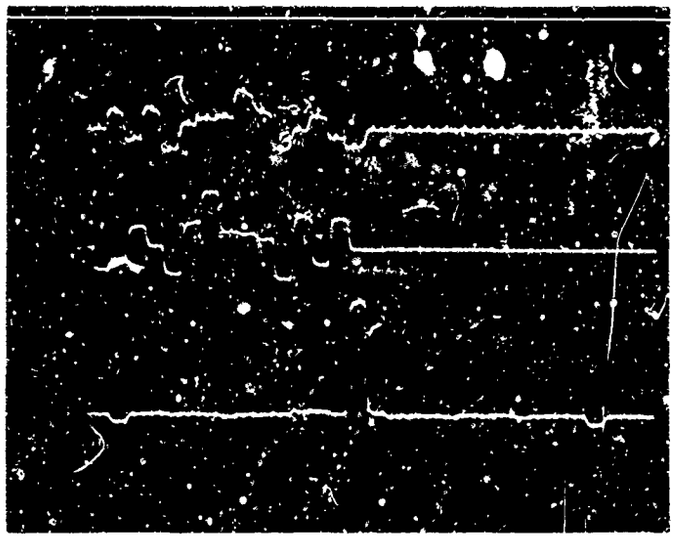


Figure 15
Upper trace: Huffman 14 impulse equivalent code.
Middle trace: Impulse response of bucket-brigade matched filter programmed for Huffman code.
Lower trace: Auto-correlation of Huffman code

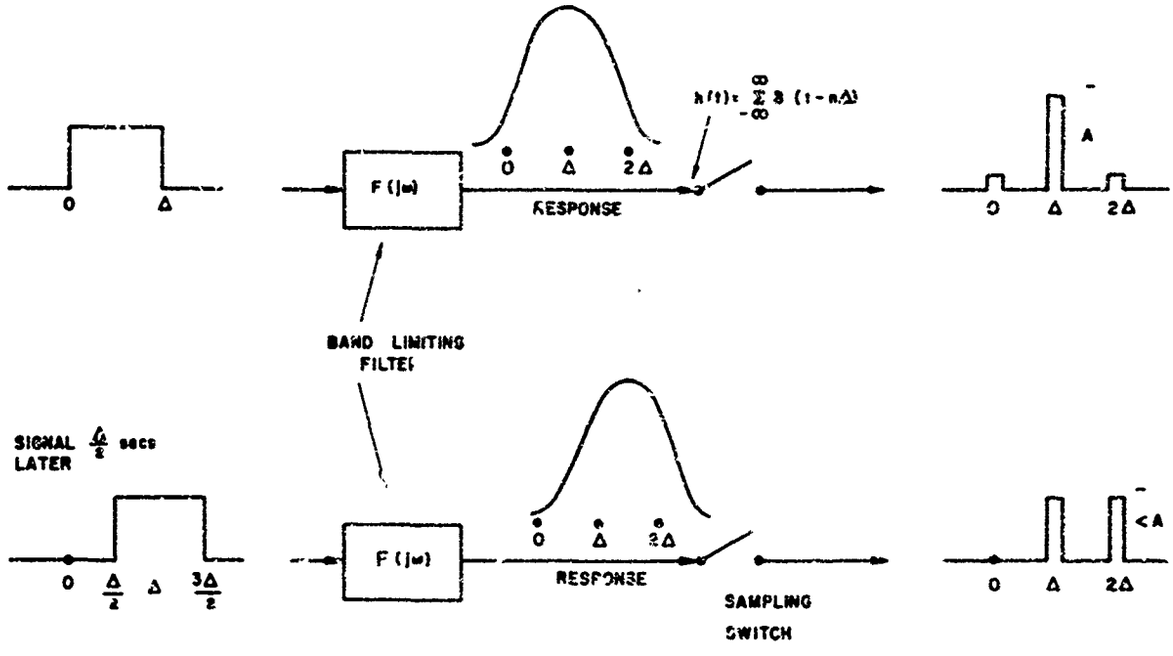


Fig. 16 Illustration of effect of sampling epoch

A DIGITAL M.T.I. CANCELLER

by

A. Reiss

Ministry of Defence (Navy)
A.S.W.E., Portsmouth, England

A DIGITAL M.T.I. CANCELLER

A. Reiss
Ministry of Defence (Navy)
A.S.W.E., Portsmouth, England

SUMMARY

Some practical aspects of an experimental digital M.T.I. canceller are described with particular emphasis on circuits and components for the principal functions of sampling, analogue/digital conversion, the digital canceller which comprises MOST shift register storage and binary subtraction, the combining of in-phase and quadrature channels and, finally, the digital/analogue conversion to obtain video signals for P.P.I. displays. P.R.F. flexibility, off-the-shelf integrated circuits, small volume and low power consumption are some of the advantages in digital canceller systems.

The experimental system, completed at A.S.W.E. in early 1969, was the first model to be demonstrated in the U.K. Features of this model are: 120 range quanta, choice of 1, 2 or 4 microsecond quantum period, 8 bit A.D.C. and twin channel, double cancellation. A second model with a larger store is nearing completion.

1. INTRODUCTION

The recent availability of very long MOST shift registers and the expectation of very low cost per storage element on such registers leads to the consideration of digital signal storage and digital subtraction: compared with the so far conventional analogue storage in quartz delay lines, the digital system has all the advantages of microelectronic systems and also a high flexibility for P.R.F. stagger.

The storage of information on digital storage elements requires the quantisation of radar signals in both range and amplitude; in other words, signals are sampled in each successive range quantum and the sample amplitudes are then converted into binary number form by means of an analogue to digital converter (A.D.C.).

Following a given radar transmission, the quantised signals as they come along from the A.D.C. are shifted into registers, until the registers are completely filled up; shifting then stops until the next radar transmission, when new signal samples are shifted into the register inputs while at the same time the previously stored samples spill out at the other end. The number of memory cells (bits) in a shift register thus determine the number of range quanta available to the radar system and a separate shift register is needed for each output line from the A.D.C., to achieve parallel operation.

As a practical example, a canceller designed for 8-bit numbers and 500 range quanta would require 8 shift registers, each containing 500 bits.

Fresh numbers, entering the registers, are subtracted from the old numbers leaving the registers; this takes place in a parallel binary subtractor the output of which is used for further cancellation, data processing in a computer or for digital-to-analogue conversion and subsequent display on a P.P.I.

The first A.S.W.E. model was made to the following specification:-

Number of range quanta	- 120
Range quantum - choice of	- 1/3, 1/6 or 1/12 n.m.
Binary number	- 8 bits
P.R.F.	- 200-7,000 Hz possible
Cancellation System	- Two-channel, double canceller

The second model has a larger store so as to provide up to 500 range quanta for long range operation.

2. M.T.I. CANCELLATION

The basic circuit for a coherent M.T.I. canceller is shown in Fig. 27-1. Bipolar video signals from a phase sensitive detector (P.S.D.) are fed into some form of storage device - quartz line or shift register - and a subtraction is carried out between the fresh signals entering the store at 'B' and the stored signals emerging at 'A', from the previous radar transmission. The subtractor residue 'R' is the output from the canceller. System timing is arranged such that at any given instant the two signals being subtracted are range-coincident and thus, in principle, stationary or slowly moving targets will be cancelled. The P.S.D. in normal use gives an output $A\cos d\phi$ where A is the signal amplitude and $d\phi$ is the phase difference between signal and the reference carrier; the circuit is therefore truly phase sensitive only for large, limiting signals. Non-limiting, smaller, clutter signals will be cancelled to a degree depending on their pulse-to-pulse phase and amplitude stability and the amount of corruption due to receiver noise.

The transfer function of the single M.T.I. canceller shown in Fig. 27-1 has the well known comb filter characteristic with half sinusoidal lobes. For better rejection of clutter velocity spectra and of radar instabilities, the filter rejection notches are broadened by the use of 2 or 3 cancellers in cascade to obtain lobes with a $\sin^2$ or $\sin^3$ shape respectively: the price paid for this improved clutter rejection is of course the broadening of blind velocity gaps but this can be overcome by appropriate P.R.F. stagger and also to some extent by the somewhat complicated use of feedback in the cancellers, to manipulate the shaping of the filter response.

In the system here described a cascade of two cancellers is used, with no feedback.

Another consideration in M.T.I. cancellers is the loss or breaking up of moving target echoes due to particular pulse-to-pulse phase conditions of the signals when processed by the cosinusoidal response of the P.D.S.: this 'Blind Phase' effect can lead to some embarrassment in automatic detection systems and may also cause the loss of weak targets. One solution which may be attractive with low-cost storage is the provision of a duplicate cancellation channel to deal with the phase-quadrature signals from a second P.S.D.; thus, if blind phase conditions should occur in one channel then a maximum signal will be obtained from the quadrature channel. The final decision must be made on how to combine the two channel outputs: one can either add the two vectorially, or select whichever is the larger signal. The latter solution is much simpler in terms of hardware and does not incur any appreciable signal/noise loss relative to the vector addition method.

3. THE DIGITAL CANCELLATION SYSTEM

A block diagram of major functions is shown in Fig. 27-2. For a given channel, video signals from a P.S.D. are sampled in a sample/hold circuit and are then converted in the analogue to digital converter (A.D.C.) into an 8-bit parallel number for each successive range quantum. The A.D.C. supplies these numbers to the two cascaded cancellers and the cancellation residues from the two separate channels (in-phase and quadrature) are then combined. From the combining unit, in which negative numbers are first complemented and then the larger of the two signals is chosen, the output is reconverted to video in the digital-to-analogue converter (DAC) and can then be used for P.P.I. display or auto-extraction.

Some form of interference suppression is also required because in a double canceller system a single pulse of interference causes a group of three adjacent pulses to be generated.

Other functions shown in Fig. 27-2 cover the generation of suitable clock pulses for the sample/hold, A.D.C. and shift registers circuits.

The number of binary digits needed to express the signal amplitudes depends on the amount of clutter cancellation which must be achieved to suit a particular radar system. A fairly good radar/M.T.I. system may be capable of 40 dB clutter reduction and thus would operate with signals limited, at i.f., to 40 dB above noise so that clutter residues after cancellation would then be equal to receiver noise. Signal quantisation levels must be such that least significant digit uncertainties at the subtractor output are less than the clutter residue. Assuming the uncertainty to be $\pm \frac{1}{2}$ a least significant bit for each of the two numbers subtracted, the worst case residue would be ± 1 bit. The choice of 8-bit numbers for the experimental model thus gives an r.m.s. canceller-generated noise of approx. $\frac{1}{2}$ compared with a maximum signal level of 128 (the 256 amplitude levels cover a range of bipolar signals from maximum positive to maximum negative i.e. +128 to -128). This ratio is equivalent to 48 dB and the canceller should therefore be capable of achieving 40 dB clutter reduction with a suitable radar system. For lower performance systems, with clutter reduction capabilities of say 28 to 30 dB, the binary number need not comprise any more than 6 bits.

In analogue M.T.I. cancellation the residual uncertainties are due:

(a) to the noise generated in the delay line equalising amplifiers, and

(b) to possible scaling errors at the inputs to the analogue subtractor:

in practice, analogue systems can be carefully set up to give about 40 dB of canceller residues but the calibration of all circuits require periodic checks to maintain optimum performance. In the digital system one still requires some setting up of the analogue input circuits, i.e. sample/hold and A.D.C. but these circuits incorporate feedback and temperature compensation to minimise routine setting up.

4. SAMPLE/HOLD FUNCTION

The range quantum is defined by the clock rate in the M.T.I. system. Making the range quantum equal to the radar pulse is the most economical approach as far as storage is concerned, but the price paid in performance is some loss in range resolution and some loss in signal amplitude when the echo straddles two range quanta: the latter effect is offset to some extent due to optical integration, on the P.P.I., of the larger area of the paint. Reducing the range quantum, on the other hand, does improve definition at the cost of increased storage and of increased noise corruption due to the larger bandwidth requirement.

In the A.S.W.E. model, the range quantum equals the radar pulse duration, and the sampling process takes the form of 'box car integration', i.e. the signal is integrated over the whole range quantum period and the result is then held, for A/D conversion, during the next quantum period: the circuit is reset in the third period. Thus, three sample/hold circuits operate in rotation and suitable output gating presents the successive "Hold" periods to the A.D.C. A switching schematic is shown in Fig.27-3.

5. ANALOGUE-TO-DIGITAL CONVERSION (A.D.C.)

At the time of the initial development there were no commercially available A.D.Cs in the U.K. which met the required speed of operation with 8-bit outputs. The design aim for the A.D.C. was a clock rate of 1 MHz, and a circuit capable of higher speeds with minor modifications when faster shift registers became available.

Various possible designs were looked at and a useful experimental circuit was found in an industrial research laboratory, this being a 6-bit A.D.C. operating at 12 MHz clock rate: the circuit produced 3 parallel bits from a 7-level comparator bank and then a further 3 less significant bits from another 7-level comparison. The early digits are delayed so that the complete number is available at the same time.

Since an 8-bit number was desirable, and no one outside A.S.W.E. was prepared to develop and manufacture an 8-bit design at short notice, it was decided at A.S.W.E. to modify and extend the original 6-bit circuit so that 8-bit numbers could be produced at clock rates of up to 2 MHz.

In the A.S.W.E. circuit, two 15-level comparator banks plus the decoding logic produce 4 bits each - see Fig. 27-4.

In the first comparator bank, video signals from the sample/hold circuit are compared with 15 equally spaced voltage levels from a reference chain; the comparator outputs are strobed at the clock frequency and held in 'D' type registers during the clock pulse intervals. The decoding logic then produces a 4-bit parallel output by interpretation of the number of reference levels exceeded by the signal. The next four digits represent the 16-level interpolation between any two of the coarse levels in the first comparator bank, thus a total of 256 levels can be obtained. The interpolation process is carried out by subtracting from the input signal a voltage proportional to the number of reference levels exceeded by this signal in the first comparator bank; this difference is the amount by which the signal has exceeded the nearest reference level, and this amount must be expressed in terms of a finer level structure (16 levels) by the second bank of 15 comparators and its associated registers and logic.

The difference circuit shown in Fig. 27-4 is an integrated circuit differential amplifier with a gain of 16; its output is therefore of a magnitude such that the second bank of comparators and its reference levels can be identical to the first bank.

Measured performance of the A.D.C. indicated that the 4-bit parallel converters could operate at speeds approaching 10 MHz. The critical component is the differential amplifier which goes into saturation from the onset of an input signal step until the appearance of the level summation signal: the amplifier remains saturated for an additional time of about 0.5 μs and then slews to its final correct output level in a further 0.2 μs. Using this type of amplifier, the maximum A.D.C. speed is about 1 MHz. For faster operation another type of difference extraction may be required, or the difference amplifier must be non-saturating and have a very high slewing rate. Development of such an amplifier is now in hand and this should increase the A.D.C. speed to between 2 and 3 MHz.

6. CANCELLATION CIRCUIT (Fig. 27-5)

The two basic functions in a canceller are (a) storage and (b) subtraction.

6.1 Storage is provided here by MOST shift registers; the dynamic type of register stores data on the gate capacitance in each memory cell so that a leakage problem exists which imposes a lower limit on clock rates, i.e. the time for which these minute capacitances can hold their charge.

M.T.I. canceller storage was found to be a rather unusual application for MOST shift registers in that clocking only proceeds while data inputs are shifted along and until all the register cells have been filled up with new data: clocking then stops and the registers must be capable of holding the data until clocking resumes after the next radar transmission.

The problem is very much less severe nowadays due to the availability, on the market, of d.c. type MOST registers with 100 and more bits per circuit and with maximum clock rates of 1 MHz (and possibly more in some cases). These circuits are effectively dynamic registers but each memory cell contains an additional latching circuit which comes into play at low clock rates or when clocking stops.

The first M.T.I. model used a particular 24-bit dynamic shift register which at the time of development was the only available one with sufficiently low leakage: no useful d.c. types were then readily available. In the second M.T.I. model, 128-bit d.c. type registers are used. Each shift register in the first experimental model consists of 5 cascaded 24-bit modules, i.e. 120 range quanta: thus a complete canceller with 8 shift registers contains 40 modules. The number of range quanta has been increased in the second model by using 100 or 128 bit modules in the existing sockets.

The economics of shift register storage depend on the market price of such devices and at the present time it is already possible to get storage for a whole M.T.I. system for a few hundred pounds sterling.

For very short pulse systems one must, however, pay a higher price: pushing the present technique may achieve operation with 0.25 μs pulses (4 MHz clock rate) but beyond this limit one would probably have to use time multiplexing and circuit replication.

6.2 Subtraction (A-B)

8-bit parallel digital subtraction is achieved by using currently available fast adders and the function A+ complement of B, this being equivalent to A + B inverted +1.

Two 4-bit parallel adder modules are used for each 8-bit adder and these can operate up to a 4 MHz rate in the subtractor mode. The inversion of B, previous to addition, is carried out in fast TTL gates and the addition of '1' is conveniently made by holding the adder's input carry in a permanent '1' state.

A 4-bit subtractor is currently under sponsored development in industry.

6.3 Signal Timing

It is of course most important that the two numbers presented to the subtractor at a given time should result from successive responses of the same range quantum. This requirement is met as sketched out below: consider the range quanta to be numbered 1 to n and thus, following a given radar transmission, the consecutive 8-bit signal samples are shifted into the n stages of eight registers until the whole register bank is filled up, i.e. number 1 and number n quantum signals are stored respectively in the nth and the first register cells. Following the next radar transmission, new data is clocked into the register inputs while, at the same time, the old data is clocked out from the register outputs: thus new and old signals from range quantum 1 appear simultaneously, then signals from quantum 2 etc. up to n.

6.4 Interfaces

Logic levels for MOST and for bipolar circuits are generally incompatible: interface circuits are therefore necessary at the inputs and outputs of the register chains. Multiple input interfacing is available in I/C modules and the output interface only requires a transistor and a couple of diodes for each register chain.

Recent work on MOST technology has resulted in lower gate switching levels and the problem of interfacing may be solved by direct MOST/bipolar compatibility. Such low level registers are used in the second canceller model.

6.5 Construction

A canceller is fabricated on a single plug-in card and comprises 40 register modules, interfacing where required, clock pulse buffers, subtraction circuits and output retiming strobes. The experimental card has a copper ground plane on one side and a grid of holes suitable for insertion of anchoring pins.

6.6 The Complete Canceller

The first canceller, in a cascaded pair, handles 8-bit numbers from the A.D.C., those numbers being positive only; the output can be either real positive or a complement for negative numbers, the latter being associated with a state '1' sign bit.

In the second canceller of the cascade one shift register chain is assigned to the sign bit and the other seven are available for the binary numbers: thus the least significant bit from the first cancellation is discarded.

The output number from the second canceller is normally a 7-bit number but an extra bit is generated, when required by the subtraction of two large numbers of opposite sign (the residue being ≥ 128).

A second identical pair of cancellers is provided for the quadrature channel.

7. COMBINING UNIT (FIG.27-6) AND OUTPUT CIRCUITS

Outputs from the cancellation circuits consisting of noise, clutter residues and moving target signals appear as 8-bit numbers which are either positive or complements, if negative: a seperate sign bit indicates the polarity. Two processes need to be carried out before the final output is obtained - firstly, the complementing of negative numbers to obtain the modulus, and then the combination of the modulus signals from the in-phase and quadrature channels by either a vector addition method or by selection of the largest signal from every range quantum: the latter method being simpler is therefore used in the experimental model.

The modulus is obtained from a complementing circuit together with gates for the selection of either the positive number or the complement of the negative number - this selection being determined by the state of the sign bit.

Having thus produced the modulus from each channel, a comparison is then made of the two numbers in a subtractor, the 'carry' output of which indicates which of the two numbers is the greater: the output selection gates are controlled from this carry.

The final result, in digital form, could be used in a digital data processing system; the present aim, however, is to obtain analogue video signals for display on a P.P.I. A digital-to-analogue converter (DAC) is provided, followed by output amplification, despiking and interference blanking. Gating-in of the signals from the radar's conventional receiver is also possible beyond the M.T.I. range coverage.

8. TRIALS

Trials up-to-date have consisted in visual comparisons of two P.P.Is of the digital system and a quartz line analogue double canceller. Clutter reduction of up to 40 dB has been observed with a metric radar.

A number of cine films have been taken frame-by-frame off the P.P.I. to show speeded up movement; both metric and centimetric radars were used for this type of presentation.

Fig. 27-7 shows a metric radar display without and with M.T.I. in photographs A and B respectively: the range ring is at 20 nautical miles.

9. CONCLUSIONS

Some basic building blocks have been described for a twin channel, double cancellation, digital M.T.I. system. Testing and evaluation of circuits has proved the design of individual functions and system tests have shown good agreement of the digital system with the best available analogue system.

Integrated circuits presently available make it possible to design systems with clock rates up to 2 MHz, and these speeds may be doubled very soon. No practical difficulties exist for slower systems associated with normal search radars operating with 2 to 10 μs pulses.

Core storage has not been mentioned so far, due partly to non-familiarity but mainly to the belief that where fully random access storage is not required, the MOST shift register solution has the edge on speed, power dissipation, bulk, and simplicity.

No further work on M.T.I. system development is proposed within A.S.W.E. Industry is by now fully aware of the potential and advantages of digital systems and has by now overtaken the work at A.S.W.E., having derived some benefit from our early developments and results.

There are many possible variants of our basic system, some of which may well be incorporated in new designs. The following is a list of useful variants:-

(a) Time sharing of registers for the I and Q channels, to reduce some of the associated circuitry.

(b) Feedback for the shaping of velocity characteristics.

(c) Multipulse cancellers with weighted summation of delay taps in place of individual subtraction after each delay - this arrangement becomes attractive with digital systems where progressive time errors, due to incorrect P.R.F., are eliminated.

(d) Quantising and delay of non-M.T.I. video signals, to line up with M.T.I. video (M.T.I. video may be 1 or 2 range quanta late due to various delays in the digital system).

(e) Signal integration for additional interference rejection and clutter reduction - this may well be applied to quantised non-M.T.I. video also.

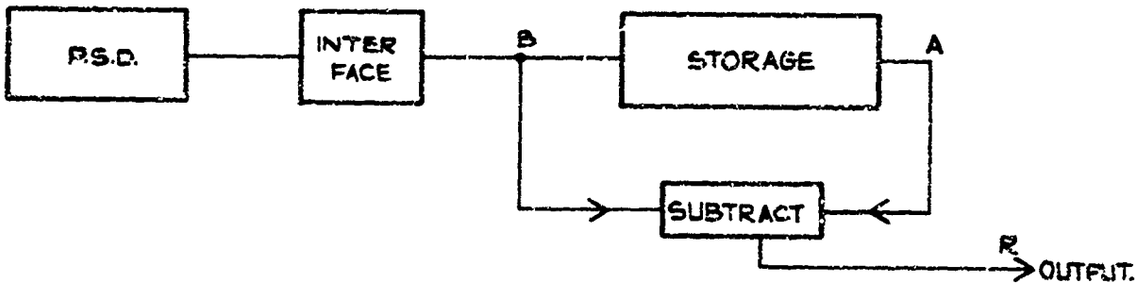


Fig. 1 Basic MTI canceller

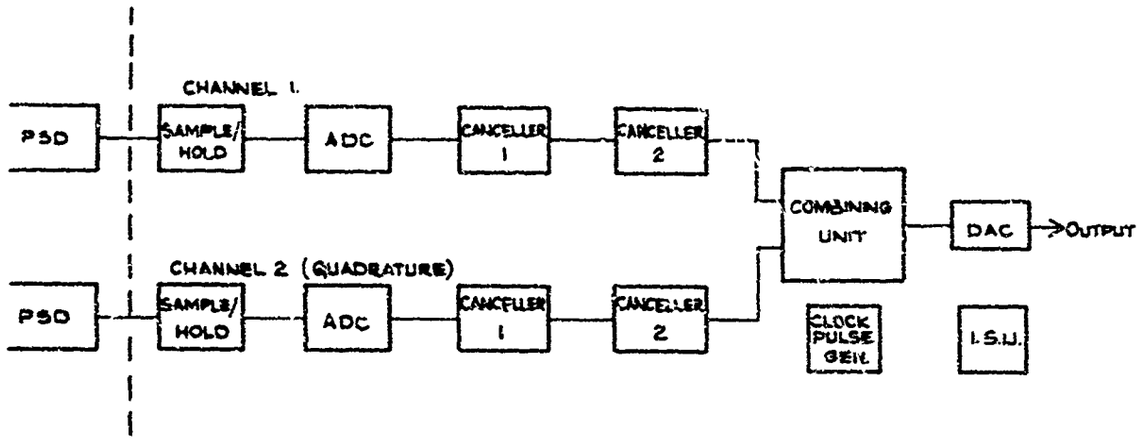


Fig. 2 MTI cancellation system

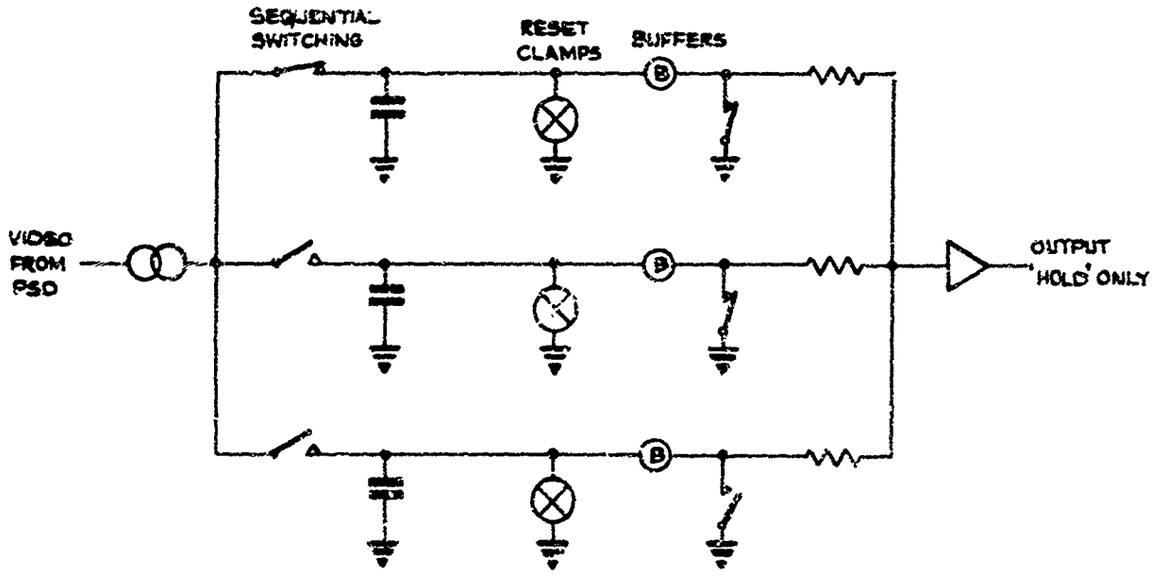


Fig. 3 Sample/hold schematic

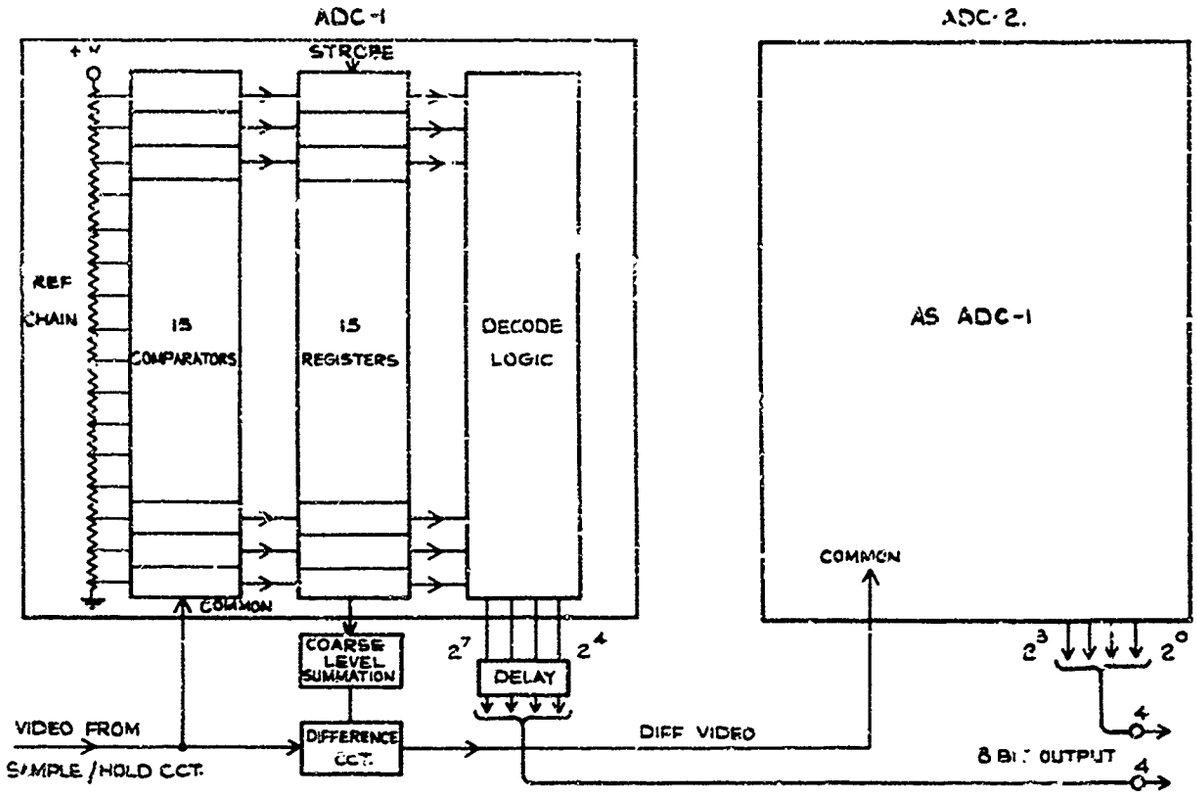


Fig.4 8-bit analogue/digital converter (ADC)

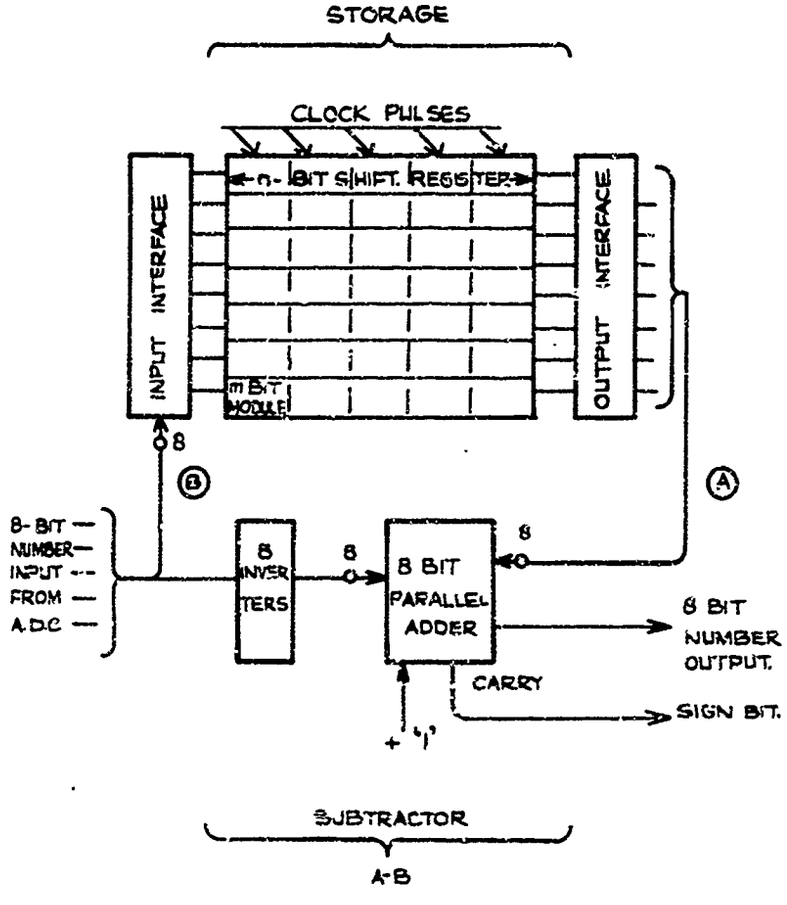


Fig.5 Cancellier circuit

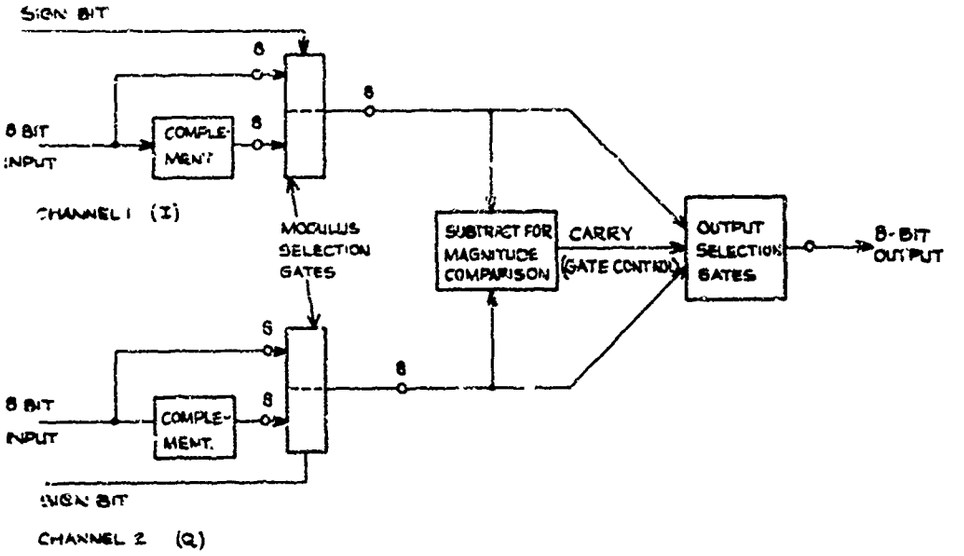
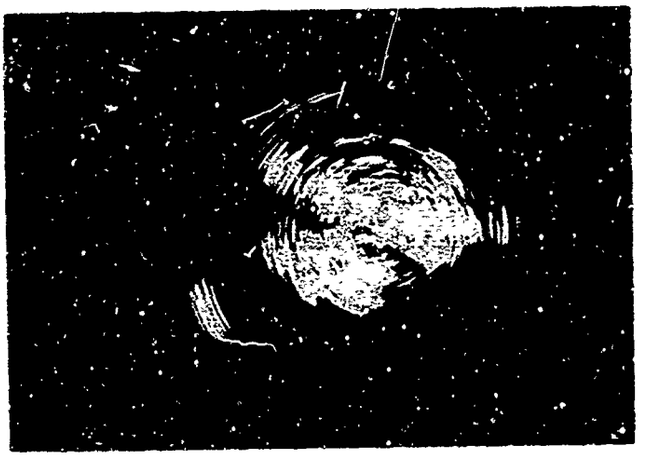


Fig. 6 Combining unit



A - without MTI

B - with MTI

Fig. 7 Metric radar PPI display

THE PLESSEY DIGITAL M.T.I. SYSTEM

by

P.Bradsell

Plessey Radar Ltd
Cowes, Isle of Wight, UK

THE PLESSEY DIGITAL M.T.I. SYSTEM

This discussion is complementary to the previous paper, in that it describes the development of a production digital M.T.I. System incorporating several of the features noted as useful in paper 27.

1. Introduction

The Plessey Digital M.T.I. System was designed from the start as part of an air traffic control radar system, the A.R.5. The cabinet contains the coherent receivers and digitiser, the canceller frame, the digivideo (signal processing) frame and all clock and timing circuits. It takes in i.f. signals and lock pulse and normal video from the transmitter receivers and gives out video to the displays. The transmitter is triggered from the M.T.I. rack either at internally generated p.r.f. or in synchronism with an external lock pulse. The system was completely developed in one year starting November 1968 and is now in production.

2. Initial Requirements

The features considered essential to satisfactory system performance are:-

1. Means of eliminating phase blindness.
2. Staggered p.r.f. giving constant sensitivity over a wide range of speeds.
3. C.F.A.R. circuits after cancellation to deal with uncancelled residue, say from heavy rain.
4. Video integration effective on normal radar and M.T.I.
5. Velocity shaping for good cancellation with low pulses/beamwidth.
6. A comprehensive built-in test system.

3. Design Decisions

A number of design decisions were required early in the project.

These were:-

3.1. Storage Media

In late 1968 MOS stores were cheaper than core stores for 2 uS and faster cycle times, and for the size and word length required. It was apparent that M.O.S. devices would develop in both speed and cost so these were chosen. Events have certainly shown the decision to be correct.

The present system is designed to operate down to 1 μS range bins, 2 μS pulse length.

3.2. Velocity Shaping

The choice here was either a feedback system or a multi-pulse canceller. The S.C.V. was to exceed 27dB with 8.75 pulses per beamwidth. Either method would meet the specification, the feedback system using less storage than the multi-pulse. The comparison in complexity of arithmetic between the two methods is marginal. The main trade-off is between the length of transient response and the cost of storage for a given cancellation performance. Since the cost of stores is the largest item on the materials bill, the feedback system was chosen.

3.3. Number of bits

The number of bits determines the ultimate performance of the cancellation circuits. The Plessey system has 8 bits (7 bits + sign) corresponding to 42dB cancellation. This gives 9dB in hand on an overall cancellation of 33dB (27dB S.C.V.).

If all the bits are kept after subtraction, the number increases by one at each stage of cancellation i.e. from 8 bits to 9 bits to 10 bits. In practice, the most significant bit is dropped after the first subtraction and the least significant bit is dropped after the second stage. This makes a reasonable compromise between loss of sensitivity by dropping lower bits and loss of dynamic range by dropping the upper ones and allows us to maintain 8-bit precision throughout the canceller. The normal radar signals are digitised in 4 bits, and the M.T.I. video reduced to 4 bits either by scaling or logarithmic compression in the later stages of the system.

3.4. Phase Blindness

The system uses the accepted technique of phase and quadrature coherent detectors. The outputs of these are handled in alternate range bins throughout the canceller and combined at the end.

3.5. Constant False Alarm Circuits

The circuit in which the signals are averaged over a group of range bins and the average subtracted from the signal at the centre of the group, was used. It is convenient to implement in digital form. An 8 bit to 4 bit logarithmic converter was developed to go before the C.F.A.R. circuit, so that the signal is some 30dB above noise before it saturates.

3.6. Staggered P.R.F.

The whole canceller works on a staggered clock so that the de-stagger delays are put in after cancellation using 4 bit storage.

Using two blocks of storage in the ratio 2:1 there are 4 possible pulse positions and 7 different periods. Computer programmes were run for various permutations of pulses, looking at the velocity response, the effect on cancellation and disturbance of the transmitter.

3.6. Staggered P.R.F. (contd)

The chosen pattern is wired in as a fixed program in the D.M.T.I. rack. The overall result gives uniform sensitivity from a hundred knots or so up to Mach 2.

3.7. Video Integrator

The system includes a digital accumulator as a two-threshold video integrator. Plessey designed a hybrid system working on this principle some years ago; it is a simple and effective method.

4. Development Program

The design process followed very smoothly from the initial design decisions. The decision was made to standardise on an existing format of p.c. board and hardware. The board design was rationalised so that, for instance, all store boards are the same whether used as canceller, de-stagger or integrator.

The main changes which have been made in the course of development are to include a P.R.F. discriminator, to prevent strong local interference punching through to the display, and an antilog converter to improve the video contrast after the C.F.A.R. circuits.

It was found possible to go directly from the logic diagram to production circuit board layout using T.T.L. logic with a high probability of first-time success.

5. Prototype Testing

The boards were brought together for the first time in an experimental rack containing a word generator, a test highway, and a D/A converter for monitoring.

Boards were plugged in one by one and the logic functions tested as the whole system gradually built up.

During this stage any wiring errors and damaged semi-conductors, were located and dealt with.

Some logic errors also appeared during system test that were not seen at the unit test stage.

Experience during the prototype testing helped with the design of the built-in test system - details of this were not made final until after the prototype was working.

The main development problems were with the analog/digital converter. The amplifier linking the two stages of the converter is subject to overload and hence paralysis - it is not easy to meet the slewing rate specification. These problems were solved and we now have an 8-bit conversion in 650 nS. Some care has been taken with screening around the A/D converter. This unit is in the same frame as the coherent receiver to avoid trouble from earth currents. The output from the converter is at a balanced logic interface which gives enough noise immunity. The prototype D.M.T.I. works quite happily a few feet from the transmitter, without any covers on the rack.

6. Built-in Test System

There is automatic on-line testing of all the storage and arithmetic functions of the canceller using test words normally gated from the input video. Triangular test signals of either fixed or alternating polarity, are available for use during fault finding.

Lamps indicate errors; repeated errors cause alarm. The pattern of lamps which light identify the faulty card which is then replaced. An 8-bit D/A converter allows an oscilloscope to be used to examine all the test points through the system for a conventional trouble-shooting procedure.

7. Performance

The systems shows a regular 29dB sub-clutter visibility in the field with 8.75 pulses/beamwidth, and gives a particularly clean and attractive display, as the slides will show.

8. Future Development

Plessey have designed a flexible set of units which are to be appliedmore widely. The first development is to increase the speed so as to operate with a 1 μS pulse length.

Improved methods of digitising radar signals and of gating normal radar and m.t.i. ouputs will be sought.

New M.O.S. stores will shortly be available which interface directly with T.T.L. This will simplify the store boards and give a useful improvement in propagation time, as we will eliminate the interface logic.

9. Conclusion

A system has been described which, although well within the state of the art, has been consolidated with all the engineering features which make it a thoroughly practical piece of equipment. It is seen as the first of a family of digital signal processing systems from Plessey.

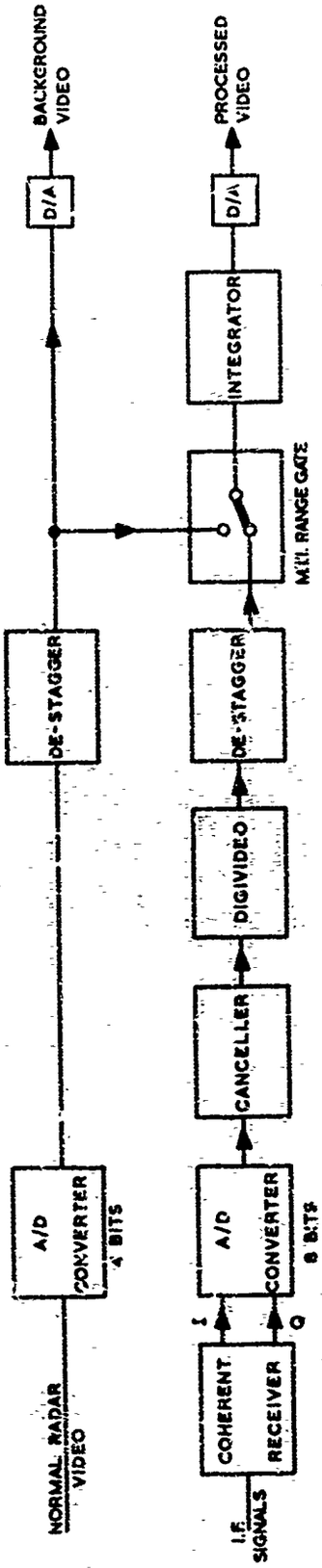


Fig.1 Block diagram, digital MTI system

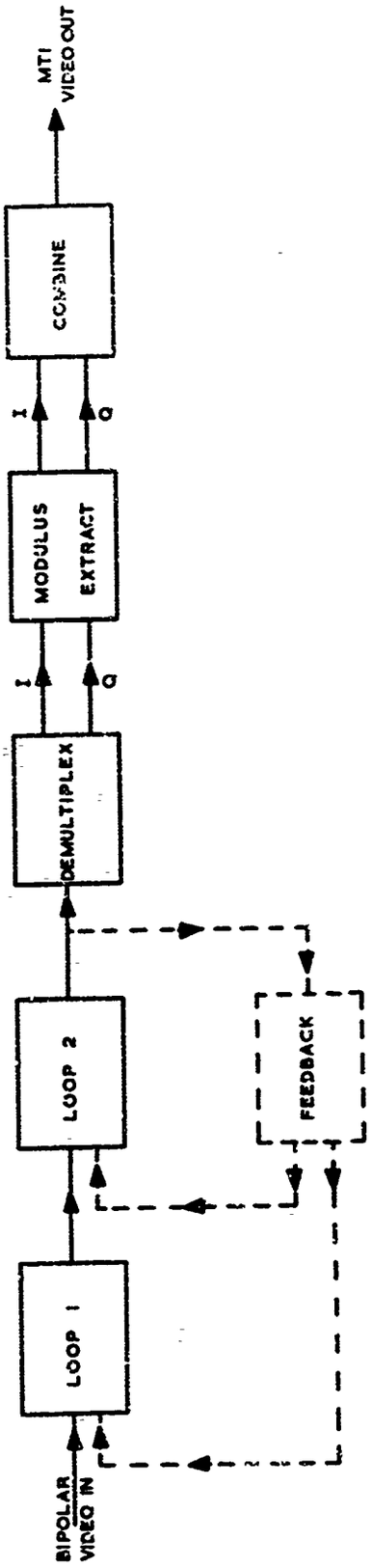


Fig.2 Signal flow, feedback canceller

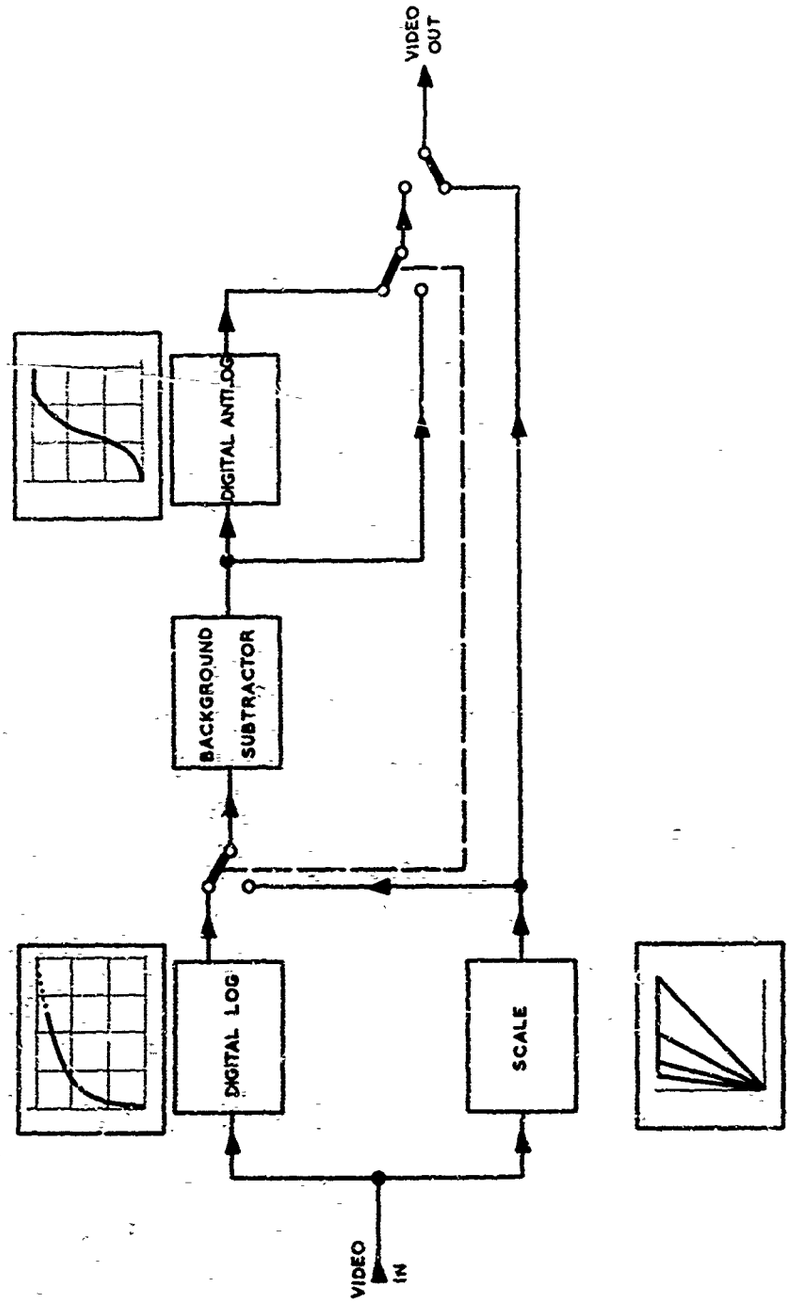


Fig.3 'Digivideo' signal processing

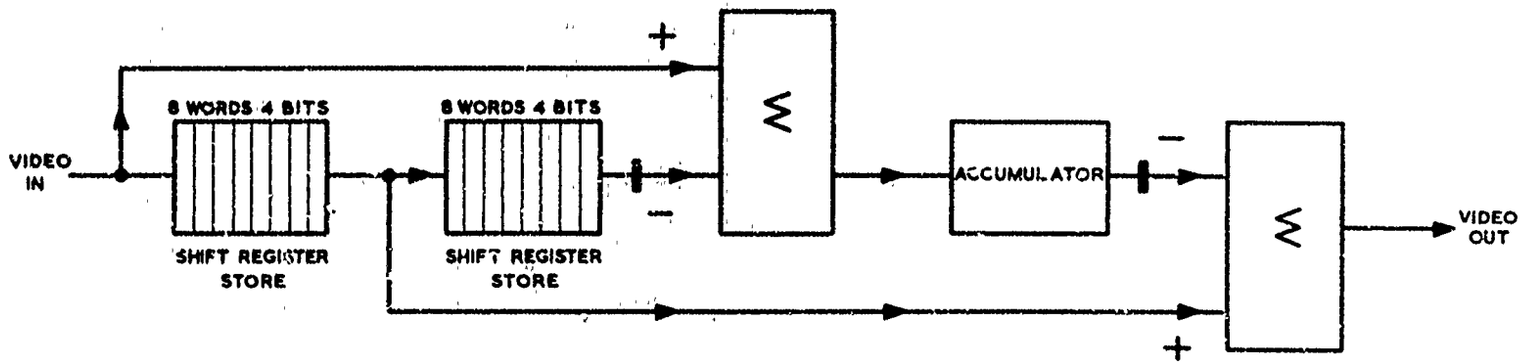


Fig.4 Background subtraction circuits

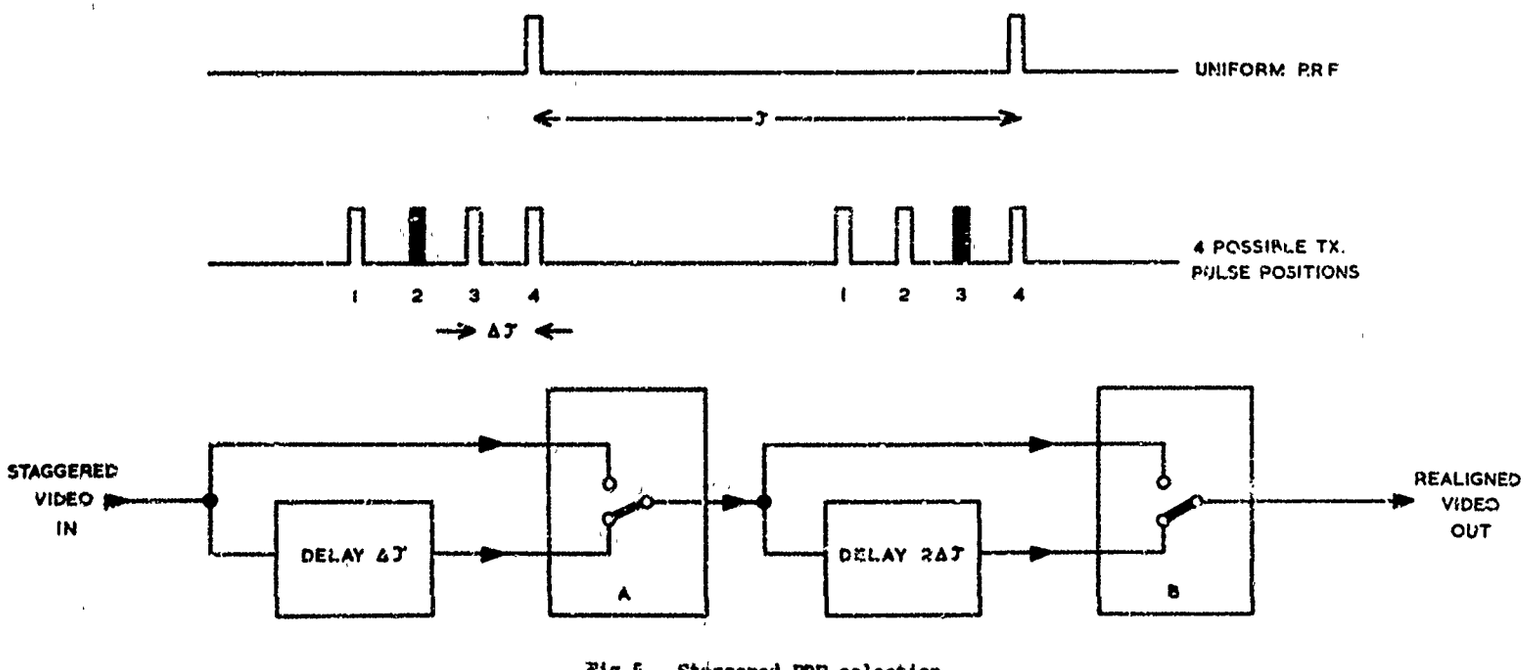


Fig.5 Staggered PRF selection

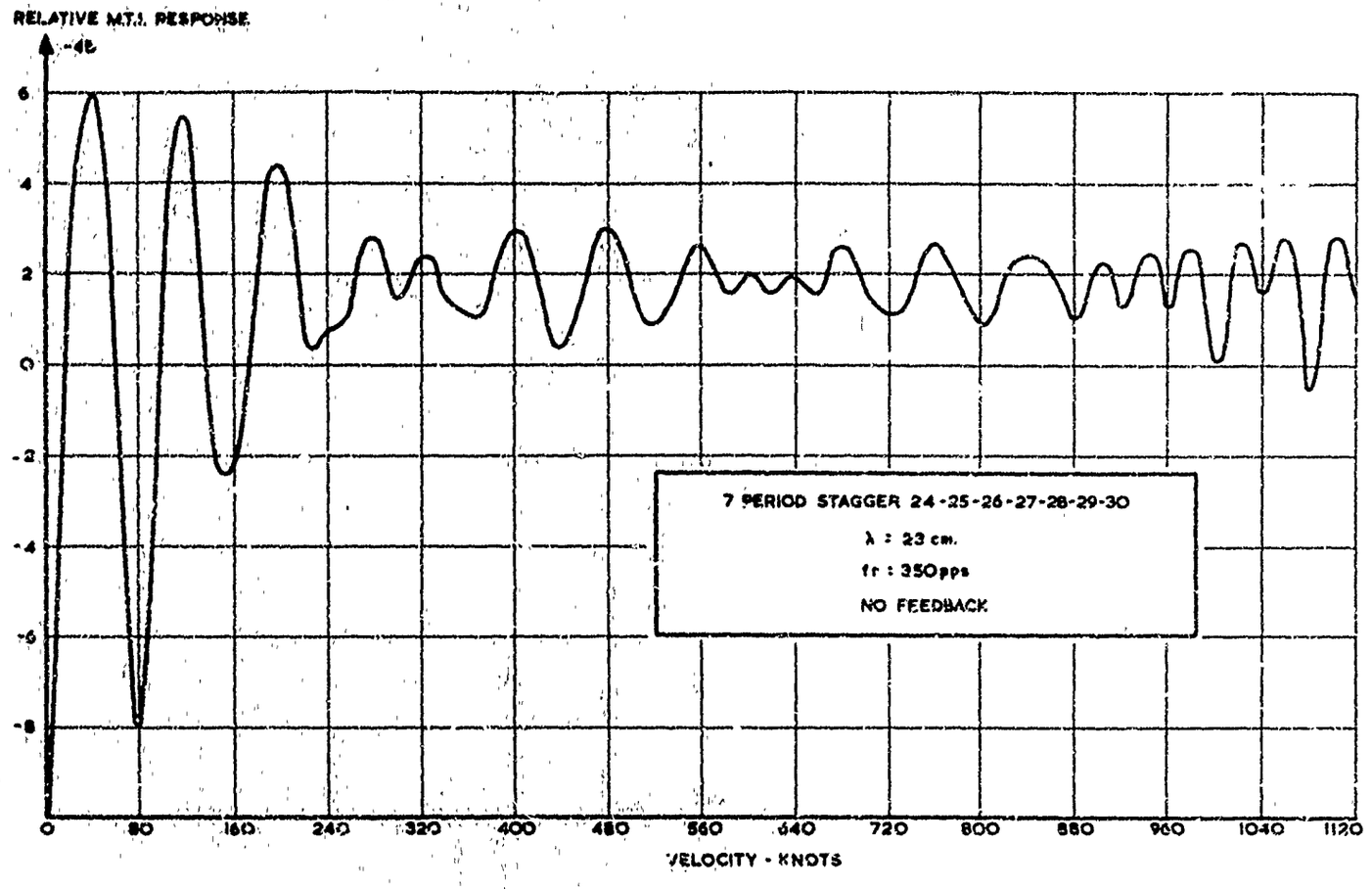


Fig.6 AR5 velocity response

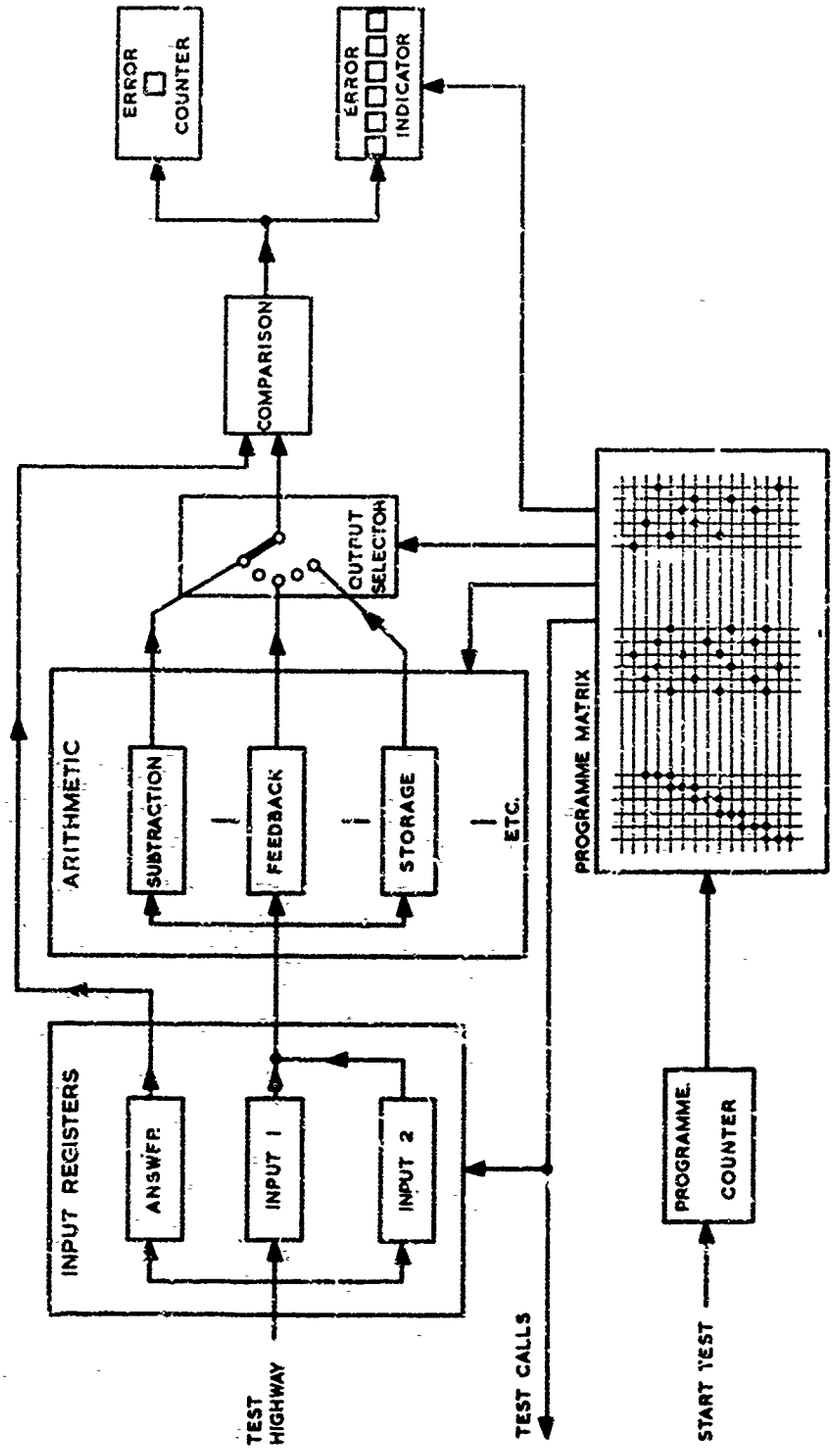


Fig.7 Built-in test system

METHODE DE TRAITEMENT DIGITAL DE SIGNAUX MONOPULSE EN PRESENCE DE CLUTTER

Par

Henri G.P. FORESTIER et Michel GAUDRON

THOMSON - CSF
Division Radar de Surface
92,Bagneux, FRANCE

L'Etude présentée dans cet exposé est partiellement supportée par la

DELEGATION MINISTERIELLE POUR L'ARMEMENT, Direction Technique des Constructions Aéronautiques Service Technique des Télécommunications de l'Air FRANCE

METHODE DE TRAITEMENT DIGITAL DE SIGNAUX MONOPULSE EN PRESENCE DE CLUTTER

Henri C.P. FORESTIER et Michel GAUDRON
THOMSON - CSF
92 Bagneux, France.

SOMMAIRE

On se propose de décrire une méthode possible de traitement digital des signaux radars en présence de clutter et destiné, en particulier, à l'exploitation monopulse d'un radar pour lequel le niveau de clutter peut être très différent dans deux lobes adjacents d'un même aérien.

Dans ce traitement digital, le signal de chaque canal du monopulse est codé :

1°) en amplitude dans le récepteur MF avec une grande dynamique en utilisant une représentation en virgule flottante.

2°) en phase de façon analogue au procédé de M T I dit à "codage de phase" et ceci sur 360° de phase.

L'élimination des échos fixes et du clutter se fait par résolution de triangles dont on montre l'analogie avec le M T I linéaire à simple et double annulation.

Le signal de chaque canal de monopulse ainsi traité est finalement utilisé pour calculer numériquement la direction angulaire de la cible pour en tirer soit l'altitude pour un radar de veille, soit la valeur des erreurs pour un radar de poursuite.

Cette méthode a été appliquée à l'aide d'un petit calculateur pour des mesures d'altitude sur cible désignée ; elle pourrait également être utilisée sur radars de poursuite en permettant une digitalisation toujours plus poussée du traitement du signal radar.

1. INTRODUCTION.

La technique monopulse repose sur la mesure simultanée et la comparaison des énergies reçues par deux diagrammes de rayonnement distincts d'un même aérien. En présence d'une cible ponctuelle cette comparaison permet d'en tirer une mesure angulaire relative et d'effectuer une interpolation entre les directions des maxima des diagrammes de rayonnement. Il en résulte deux applications principales de la technique monopulse :

- la mesure d'altimétrie pour les radars de surveillance tridimensionnels et

- l'écartométrie angulaire pour les radars de poursuite.

En présence de cibles multiples situées à une même distance tel que celà existe dans le cas d'un écho utile noyé dans le clutter il est bien connu que le niveau de clutter reçu peut être très différent entre deux lobes adjacents d'un même aérien et la séparation du signal utile et du clutter dans chaque canal du monopulse peut rendre difficile la mesure précise des énergies de signal utile et même rendre peu fiable la comparaison de ces énergies si le traitement n'a pas été effectué de la même façon dans chaque canal.

Pour assurer l'identité du traitement dans chaque canal monopulse, les méthodes digitales sont particulièrement tentantes du fait que, procédant par calculs numériques, l'identité des transferts dans deux canaux différents peut s'obtenir aisément avec une certaine rigueur.

Comme pour tout traitement de signal radar visant à l'élimination des échos fixes et du clutter on est amené à faire une exploitation cohérente du signal reçu tenant compte de la stationarité caractéristique de la phase du signal reçu par rapport au signal émis propre aux échos fixes et au clutter par opposition à une phase évolutive pour des signaux d'échos mobiles.

Nous nous proposons de décrire une méthode possible de traitement digital des signaux radar en présence de clutter et destinés à l'exploitation monopulse. Cette méthode basée sur la représentation vectorielle de Fresnel des énergies reçues effectue l'élimination des échos fixes et du clutter en résolvant des relations vectorielles à l'aide de la trigonométrie. Commencée d'être étudiée en FRANCE à la THOMSON - CSF cette méthode se développe actuellement selon deux directions correspondant aux deux principales applications de la Technique Monopulse.

2. PRINCIPE DE LA METHODE

La méthode étant basée sur la représentation vectorielle de FRESNEL d'un signal complexe composé d'un signal utile dû à une cible mobile en présence de clutter examinons plus en détail cette représentation.

2.1. REPRÉSENTATION DE FRESNEL D'UN SIGNAL COMPLEXE (Fig.1.)

Considérons un signal complexe composé d'un signal utile dû à une cible mobile accompagné d'un signal parasite dû à un écho fixe situé à la même distance que la cible mobile et illuminé simultanément à la cible mobile.

Par rapport à une référence de phase constituée par le CoHO HF du récepteur le signal reçu sera représenté par un vecteur somme de deux vecteurs composants

Un premier vecteur $\overrightarrow{OF}$ d'amplitude f et de phase Θ_F représente le signal parasite auquel vient s'ajouter le vecteur $\overrightarrow{FM}$ d'amplitude m et de phase Θ_M représentant le signal utile. Par définition Θ_F est constant tandis que Θ_M tourne à la vitesse angulaire $W_d = 2\pi F_d$ (F_d étant la fréquence Doppler associé à la cible mobile).

Considérons maintenant deux récurrences radar successives.

A la première récurrence on a :

- un vecteur $\overrightarrow{OF}$ fixe de phase Θ_F
- un vecteur $\overrightarrow{FM_1}$ de phase Θ_{M1}
- le vecteur résultant $\overrightarrow{OM_1}$ représentant le signal reçu, de phase φ_1 et d'amplitude r_1

A la récurrence suivante on a :

- un vecteur $\overrightarrow{OF}$ fixe de phase Θ_F identique au précédent
- un vecteur $\overrightarrow{FM_2}$ de phase Θ_{M2}
- le vecteur résultant OM_2 représentant le signal reçu de phase φ_2 et d'amplitude r_2

Sur la Fig. 1 on a les relations angulaires suivantes :

$$\Theta_{M2} - \Theta_{M1} = \alpha = W_d \; T_{r\,12} \qquad (1)$$

$$\varphi_2 - \varphi_1 = \varphi_{12}$$

Considérons maintenant le vecteur $\overrightarrow{M_1 M_2}$.

Dans le triangle $FM_1 M_2$ on a :

$$M_1 M_2 = 2m \sin \frac{\alpha}{2} = 2m \sin \frac{W_d T_{r\,12}}{2} \qquad (2)$$

Dans le triangle $O M_1 M_2$ on a :

$$M_1 M_2 = r_1^2 + r_2^2 - 2 r_1 r_2 \cos \varphi_{12} \qquad (3)$$

En comparant (2) et (3) on peut noter :

- que $M_1 M_2$ ne dépend que m et est donc représentatif de la cible utile
- que $M_1 M_2$ est accessible par calcul à partir du signal reçu puisqu'il ne dépend que de r_1, r_2, φ_1 et φ_2

Pour exploiter la méthode il faut

- A chaque récurrence coder le vecteur $\overrightarrow{OM}$ en amplitude pour avoir et en phase pour avoir
- A chaque récurrence effectuer une résolution de triangle (3) pour calculer $M_1 M_2$ représentatif de m

2.2. CODAGE DIGITAL DU VECTEUR $\overrightarrow{OM}$

Le vecteur $\overrightarrow{OM}$ est codé sous la forme polaire en une amplitude $r = (OM)$ et une phase φ

Le codage d'amplitude s'effectue dans un récepteur moyenne fréquence s'apparentant à un récepteur logarithmique (Fig.2) dont chaque étage aurait un gain de 6 dB et aurait une caractéristique linéaire sur au moins 6 dB.

Le codage d'amplitude utilise une représentation en virgule flottante en base 2 : la caractéristique étant donnée par le nombre d'étages pour lequel on a franchi le seuil correspondant à la limite supérieure de la caractéristique de transfert linéaire sur au moins 6 dB. Une fois la caractéristique codée un multiplexeur définit l'étage sur lequel on effectue le codage de la mantisse linéaire.

Le principal avantage de ce codage d'amplitude en virgule flottante réside dans la grande dynamique (de l'ordre de 80 dB) sur laquelle peut s'effectuer le codage.

Le codage de phase s'effectue dans un détecteur d'amplitude phase entre la sortie limitée du récepteur MF logarithmique et la référence délivrée par le COHO MF du récepteur. Le détecteur amplitude phase est du type quatre quadrants avec référence en phase et en quadrature ; il est linéarisé dans la voie en phase et dans la voie en quadrature sin ± 45°. Une logique permet de choisir le quadrant et la caractéristique linéaire à coder de façon à avoir un codage entièrement linéaire sur 360°. En fait il s'agit du codeur utilisé sur les M T I à codage de phase 1 qui a été présenté au XV technical Symposium de l'AGARD en 1968.

2.3. RESOLUTION DU TRIANGLE OM_1M_2

La résolution du triangle OM_1M_2 s'effectue sous forme d'un calcul numérique donnant $|M_1 M_2|^2$ Les grandeurs étant codées en virgule flottante il est tout naturel d'effectuer les calculs en digital dans le format virgule flottante.

La formule classique (3) de résolution du triangle conduit normalement à 5 multiplications et 3 additions. En virgule flottante la multiplication par 2 est triviale et pour la formule $M_1 M_2^2 = r_1^2 + r_2^2 - 2r_1r_2 \cos \varphi_{12}$ (4)

peut être modifiée pour avoir un calcul plus rapide en 3 multiplications et 2 additions.

$$M_1 M_2^2 = (r_1 + ir_2)^2 - 4\, i\, r_1\, r_2 \cos^2 \frac{\varphi_{12}}{2} \qquad (5)$$

$i = \pm 1$ selon que le quadrant de φ_{12}

φ'_{12} est une fonction très simple de φ_{12} soit $\frac{\pi}{2}$

$\varphi'_{12} = |\varphi_{12}|$ pour $|\varphi_{12}|$ compris entre 0 et

$\varphi'_{12} = |\pi - \varphi_{12}|$ pour $|\varphi_{12}|$ compris entre $\frac{\pi}{2}$ et π.

Pour le calcul en virgule flottante il sera commode d'utiliser une table de $4\cos^2 \frac{\varphi'_{12}}{2}$ adressée de façon indirecte par φ_{12}. De plus si φ_n est codé avec n bit la table n'aura que 2^{n-2} de longueur.

2.4. ANALOGIE AVEC LE M.T.I.

En revenant à la Fig. 1 et en projetant les points M_1 et M_2 sur un axe (par exemple l'axe de référence de phase) en M'_1 et M'_2 il est facile de voir l'analogie avec le MTI linéaire dont la sortie du détecteur d'amplitude phase serait un signal proportionnel à OM'_n, la boucle d'annulation effectuant la soustraction $OM'_2 - OM'_1 = M'_2 M'_1$.

Toutefois la méthode de triangulation présente deux avantages :

- Du fait du codage et du calcul en virgule flottante il est possible de travailler dans une grande dynamique.

- Du fait de la représentation vectorielle on évite le problème des phases aveugles (cas du vecteur $M_1 M_2$ perpendiculaire à l'axe de projection) qui n'est généralement résolu qu'en utilisant dans le MTI deux voies en quadrature (projection sur deux axes orthogonaux).

L'extrapolation de l'analogie avec le MTI linéaire au cas de la double annulation peut conduire à un perfectionnement de la méthode initiale de triangulation. Il en résulte une méthode de double triangulation (Fig.3). Dans la méthode de double triangulation on considère 3 récurrences successives avec les trois vecteurs OM_1, OM_2, OM_3. Par résolution de triangle on peut trouver les longueurs e_{12} e_{23} et e_{31} des vecteurs $\overrightarrow{M_1M_2}$, $\overrightarrow{M_2M_3}$ et $\overrightarrow{M_3M_1}$. Par analogie avec le M T I à double annulation il est possible d'envisager le vecteur $\overrightarrow{M_1M_2} - \overrightarrow{M_2M_3}$ pour celà considérons le point M'_3 symétrique de M_3 par rapport à M_2 on a :

$$\overrightarrow{M_1 M_2} - \overrightarrow{M_2M_3} = M_1 M_2 + \overrightarrow{M_2M'_3} = \overrightarrow{M_1 M'_3} \qquad (6)$$

d'où en utilisant les triangles $M_1 M_2 M_3$ et $M_1 M_2 M'_3$

$$|M_1 M'_3|^2 = |M_1 M_2|^2 + |M_2 M'_3|^2 - 2\,|M_1 M_2|\,|M_2 M'_3| \cos \beta$$

$$|M_1 M_3|^2 = |M_1 M_2|^2 + |M_2 M_3|^2 - 2\,|M_1 M_2|\,|M_2 M'_3| \cos (\pi - \beta)$$

Ou en définitive

$$|M_1 M'_3|^2 = 2\left[|M_1 M_2|^2 + |M_2 M_3|^2\right] - |M_1 M_3|^2$$

$$|M_1 M'_3|^2 = 2\,(e_{12}^2 + e_{23}^2) - e_{31}^2 \qquad (7)$$

Dans l'application de cette méthode de double triangulation il faut à chaque récurrence

- Effectuer deux résolutions de triangle comme indiqué précédemment

- Mémoriser un des résultats

- Effectuer le calcul de la relation (7)

L'analogie avec le M T I linéaire à double annulation au moyen de la projection vectorielle sur un (ou deux axes) fait en outre apparaître un cas particulier de phase aveugle lorsque M_1 et M_3 sont symétriques par rapport à M_2 ou lorsque M_1 et M_3 sont confondus . Ce cas est indécelable dans le fonctionnement du M T I tandis que dans le cas du calcul de triangulation il donne lieu à un résultat nul en double triangulation et un résultat non nul pour la simple triangulation qui est en fait un sous programme de la double triangulation un test logique sur les résultats partiels et finaux permet aisément alors le changement de programme pour repasser en simple triangulation.

5. APPLICATION DE LA METHODE DE TRIANGULATION AU MONOPULSE

Dans l'exploitation monopulse d'un radar le niveau de clutter peut être très différent dans les deux lobes de rayonnement décalés d'un même aérien ceci est notamment le cas lorsque le décalage a lieu en site, le faisceau bas risquant de recevoir plus d'échos fixes ou de clutter de sol ou de mer que le faisceau haut tandis que ce dernier risque de recevoir plus de clutter de marge que le faisceau bas. Dans ces conditions le traitement de séparation de l'énergie utile et de l'énergie de clutter doit être effectué de façon aussi identique que possible dans les deux canaux monopulses à comparer. La méthode de triangulation décrite travaillant sur deux mesures à deux récurrences successives dans le cas de la simple triangulation et sur trois mesures à trois récurrences successives dans le cas de la double triangulation paraît bien adaptée au problème monopulse c'est pourquoi des études se poursuivent pour en faire l'application au deux grands cas d'emploi du monopulse.

5.1. APPLICATION A UN RADAR DE VEILLE

Une première application en cours d'étude concerne un radar de veille avec un aérien à couverture haute et couverture basse dans l'exploitation duquel on recherche à pouvoir évaluer l'altitude de cibles mobiles designées à partir d'une console P.P.I.

Sur la console P.P.I. la cible désignée est entourée par une fenêtre qui définit un rectangle en distance et azimuth à l'intérieur duquel on effectue tant pour la couverture haute que pour la couverture basse le codage en amplitude du signal reçu. Par ailleurs les codages en phase sont effectués dans les M T I à codage de phase du type décrit dans [1] qui équipent les deux chaînes de couverture. Les résultats de ces codages sont envoyés en fin de récurrence à la mémoire d'un petit calculateur. Ce petit calculateur 10 010 de la C I I (Compagnie Internationale pour l'Informatique) reçoit des données tant que dure la fenêtre de désignation en azimuth ; lorsque cette fenêtre est refermée le calculateur effectue les résolutions de triangle pour les échos reçus il en résulte deux suites de résultats

- une suite m^2_{nh} des carrés des secteurs $\overrightarrow{M_{n-1}M_n}$ en couverture haute
- une suite m^2_{nb} des carrés des secteurs $\overrightarrow{m_{n-1}M_n}$ en couverture basse

Chacun des termes de ces deux suites est représentatif du carré du vecteur correspondant à l'énergie de la cible mobile dans le canal considéré avec une pondération correspondant au diagramme de rayonnement en azimuth de l'antenne en rotation. Il est alors possible d'éliminer de ces deux suites les termes qui auraient donné lieu à une précision moindre dans l'une ou l'autre des triangulations (ceci à l'aide de critères sur φ_{12} ou l'amplitude des termes comparés à leur moyenne).

La sommation dans chaque suite de tous les termes restants donne à une constante près une moyenne pondérée de m^2_h ou m^2_b la pondération due au diagramme de rayonnement en azimuth donnant plus de poids aux mesures (et donc aux résultats des triangulations) correspondant au centre du diagramme de rayonnement en azimuth de l'antenne. Le rapport de ces moyennes pondérées permet de calculer le site de la cible mobile en passant par une fonction connue du rapport des gains des diagrammes de rayonnement décalés en site. Une fois connu le site de la cible mobile désignée et sa distance il est alors facile d'en calculer l'altitude et de la transmettre en digital à la console.

Du fait de la faible vitesse du calculateur utilisé en virgule flottante et du grand nombre de triangles à résoudre le programme complet dure plusieurs secondes mais reste inférieur à une révolution complète de l'aérien du radar de veille.

5.2. APPLICATION A UN RADAR DE POURSUITE

Une seconde application également en cours d'étude concerne l'écartométrie angulaire d'un radar de poursuite. Il s'agit de calculer récurrence par récurrence la valeur numérique des erreurs angulaires par rapport à une cible mobile en présence de clutter.

Une première difficulté intervient du fait que l'on cherche à annuler cette erreur pour la cible mobile ce qui risque de donner dans le diagramme différence une très faible valeur de m pour une valeur non négligeable de f . Pour éviter ces inconvénients il convient de travailler non pas sur des diagrammes somme et différence comme c'est le cas en général mais sur les diagrammes élémentaires. Nous considérerons donc le cas du monopulse à 4 lobes : deux lobes décalés en site et deux lobes décalés en gisement.

Du fait de la poursuite en distance le codage devra être effectué simultanément sur les 4 lobes et pour préserver le temps réel les triangulations devront s'effectuer en moins d'une récurrence ce qui correspond à 4 résolutions de triangles dans le cas de la simple triangulation et à 8 dans le cas de la double triangulation. Il en résulte qu'une opération de triangulation doit être effectuée entre le sixième et le vingtième de la récurrence. En utilisant deux opérateurs digitaux effectuant

- l'un l'addition ou la soustraction en virgule flottante
- l'autre la multiplication ou la division en virgule flottante

Il est possible d'effectuer en virgule flottante une opération élémentaire d'addition ou de soustraction et une opération de multiplication ou de division en quelques microsecondes ce qui permet de résoudre le problème à condition d'utiliser un programme câblé et une table de fonction trigonométrique en format virgule flottante.

Les résultats des résolutions de triangle donnent les carrés des modules de vecteurs représentatifs de α, soit pour deux lobes désignés X et Y les quantités m_x^2 et m_y^2

L'erreur résultant de la comparaison des lobes X et Y serait normalement

$$\varepsilon = \frac{mx - my}{mx + my} \qquad (8)$$

ou en multipliant le numérateur et le dénominateur du terme de droite de (8) par $m_x + m_y$

$$\varepsilon = \frac{mx - my}{mx + my} = \frac{m_x^2 - m_y^2}{(mx + my)^2} = \frac{m_x^2 - m_y^2}{m_x^2 + m_y^2 + 2m_x m_y}$$

Pour une erreur nulle ou faible : $m_x \approx m_y$ aussi prendrons nous au lieu de ε la quantité

$$\varepsilon' = \frac{m_x^2 - m_y^2}{2(m_x^2 + m_y^2)} \qquad (9)$$

ce qui permet d'éviter d'avoir à extraire des racines carrées.

4. LIMITATIONS DE LA METHODE

Dans l'état actuel de la technique la principale limitation semble se situer au niveau des codages. Il est bien sur évident que les récepteurs doivent avoir la qualité monopulse.

Pour le codage en phase actuellement bien connu et pour des durées d'impulsion de l'ordre de la microseconde le codage à 8 bits pour 360° est chose courante et un codage à 10 bits est du domaine du possible.

Pour le codage en amplitude le codage de la caractéristique ne pose pas de gros problème, celui de la mantisse s'effectue actuellement sur 4 ou 5 bits et un codage sur 8 bits est du domaine du possible.

Lorsque la largeur d'impulsion diminue en dessous de la microseconde le codage rapide devient plus difficile et le nombre de bit devra être réduit.

Le faible nombre de bits de codage significatifs possibles limite la précision des données d'entrée du calcul. Il est à remarquer que le format en virgule flottante délivre des données dont la quantification n'affecte que la précision relative et non la précision absolue.

Dans les calculs eux-mêmes par contre on ne se trouve plus limité en général par la longueur des mots et seuls interviennent les impératifs de vitesse de calcul ce qui conduit à utiliser des tables de fonctions trigonométriques qui devront avoir une précision suffisante donc nettement plus de bits significatifs que pour le codage.

L'étude des erreurs dues à la quantification ou aux calculs permet d'optimiser ces derniers et de trouver des critères qui serviront à éliminer les résultats qui risqueraient d'être par trop erronnés ou ceux dont les erreurs ne pourraient pas être réduites par un filtrage ultérieur des résultats. Ce filtrage ultérieur est du reste toujours souhaitable et nécessaire du fait que les calculs permettent d'obtenir un résultat par récurrence correspondant à un échantillonnage à une fréquence très nettement supérieure à l'utilisation habituelle.

5. CONCLUSIONS

Nous avons décrit une nouvelle méthode de traitement de signaux radars en présence de clutter, méthode adaptée à l'exploitation monopulse d'un radar.

Cette méthode de calculs numériques trigonométrique se développe actuellement selon les grandes applications de la technique monopulse. Elle pourrait ouvrir la voie vers une digitalisation toujours plus poussée du traitement du signal à l'intérieur même des radars.

6. REMERCIEMENTS

Les auteurs tiennent à remercier Monsieur l'Ingénieur de l'Air DARRICAU du S T T A qui le premier leur a donné l'idée de base de la méthode de triangulation ainsi que la Direction Technique de la Division Radar de Surface de la THOMSON - CSF qui n'a cessé de les encourager et de les assister dans l'Etude qui nous a été permis de vous présenter.

Cette étude est partiellement supportée par la Direction Ministérielle pour l'Armement, Direction Technique des Constructions Aéronautique, Service Technique des Télécommunications de l'Air.

BIBLIOGRAPHIE

(1) R. CARRE - M T I à Codage de Phase
XVème AGARD AVIONICS PANEL TECHNICAL SYMPOSIUM
NOV. 1968.

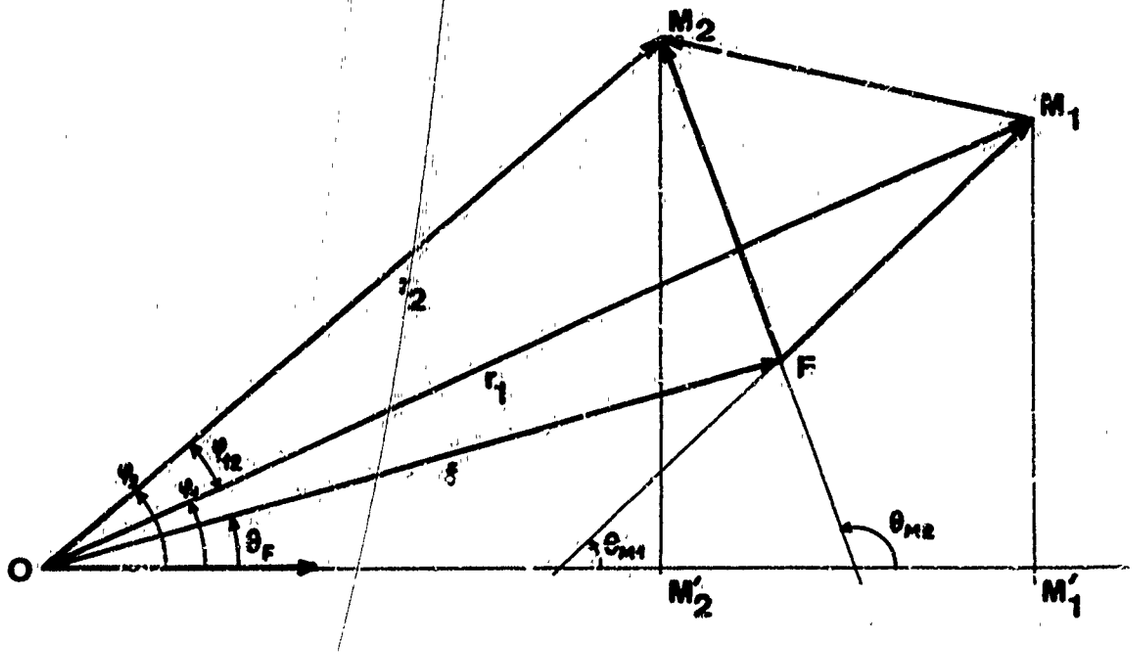


Fig.1 Méthode de triangulation, diagramme de Fresnel

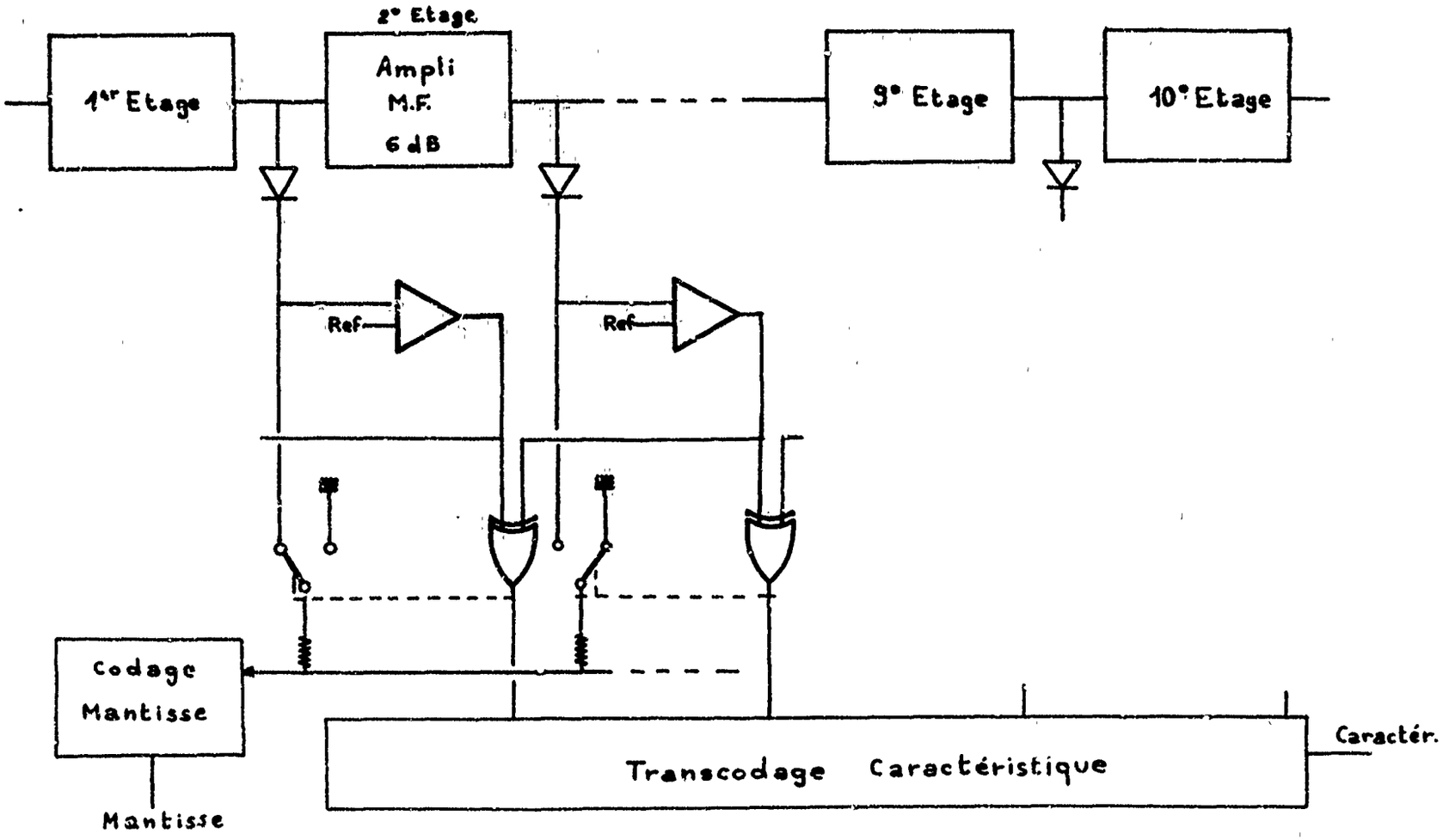


Fig.2 Récepteur Codeur

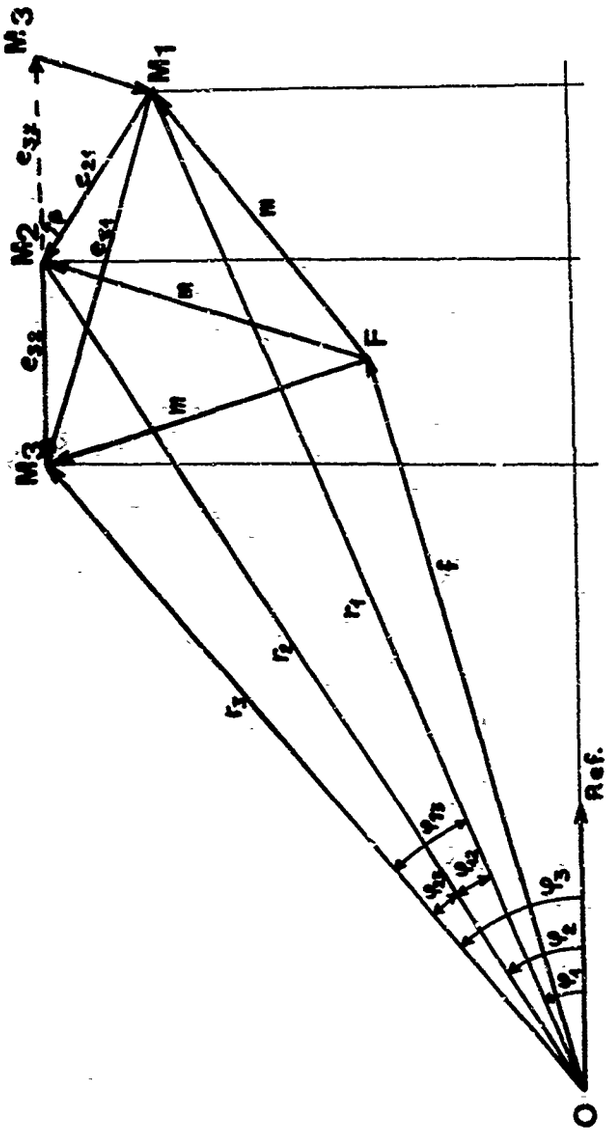


Fig.3 Double triangulation, diagramme de Fresnel

SYSTEME DE TRAITEMENT ASSOCIE A UN RADAR COHERENT POUR MESURER LA DISTANCE ET LA VITESSE RADIALE DES CIBLES

par

J.Sirven

THOMSON-C.S.F
92 - Malakoff
France

SYSTEME DE TRAITEMENT ASSOCIE A UN RADAR COHERENT POUR MESURER LA DISTANCE ET LA VITESSE RADIALE DES CIBLES

PAR J. SIRVEN
THOMSON-CSF
92 - MALAKOFF

1 - SOMMAIRE

Les radars cohérents à basse fréquence de récurrence permettent d'éliminer les échos parasites dont la fréquence Doppler est faible, ils fournissent la distance des cibles utiles, mais en général la vitesse radiale de ces cibles n'est connue qu'avec ambiguïté.

Le système décrit permet de mesurer la vitesse radiale vraie d'un mobile, il est associé à un émetteur cohérent qui module l'onde émise par deux trains périodiques d'impulsions, séparés en temps par un intervalle beaucoup plus faible que la période de récurrence. Le traitement à la réception est un traitement linéaire, ce qui permet d'éliminer les échos parasites, il utilise deux détections synchrones pour conserver l'information de phase jusqu'à la fin du filtrage. L'ambiguïté de distance introduite par la forme de l'onde émise est levée par une opération logique tandis que le lever d'ambiguïté de vitesse est obtenu par mesure du déphasage entre les deux trains d'impulsions qui ont été réfléchis par la cible.

2 - INTRODUCTION

Un radar cohérent à basse fréquence de récurrence permet d'éliminer les échos fixes ; il fournit la distance des cibles mobiles, les vitesses radiales ne peuvent être connues que de manière ambiguë, sauf dans le cas particulier où la longueur d'onde est grande et la portée maximum réduite.

Or, les systèmes de traitement de l'information qui sont associés au radar peuvent avoir besoin de connaître la vitesse radiale.

Le système que nous allons décrire permet d'obtenir la distance, la vitesse radiale des échos mobiles et simultanément il assure un taux de réjection des échos fixes très élevé.

Caractéristiques du signal émis

Le signal émis est constitué d'une onde pure découpée par deux trains d'impulsions identiques, séparés l'un de l'autre par un intervalle de temps τ petit devant la période de récurrence T qui caractérise chaque train d'impulsions (voir figure 1).

La période T du signal émis est fonction de la portée maximum que l'on veut obtenir, comme dans tous les radars sans ambiguïté de distance.

La largeur d'impulsion est fonction du pouvoir de résolution en distance cherché. Et nous montrerons que l'intervalle de temps τ séparant les deux trains d'impulsions définit la plage Doppler Δf_d sur laquelle il est possible de mesurer la fréquence Doppler des cibles :

$$\Delta f_d = \frac{1}{\tau}$$

3 - PRINCIPE DU FILTRAGE DOPPLER

A la réception, le signal est échantillonné ; le pas d'échantillonnage est sensiblement plus petit que la largeur des impulsions.

Les échantillons distants de la période T sont regroupés de façon à être traités dans des filtres numériques. Les deux trains d'impulsions réfléchis par une cible seront donc traités dans des voies séparées.

Il est possible de la sorte d'obtenir une aussi forte réjection des échos fixes que dans un radar cohérent dont le signal émis est constitué par un train d'impulsions non modulé en phase.

Les filtres numériques sont légèrement différents de ceux utilisés couramment pour éliminer les échos fixes ; ils sont caractérisés par le fait que la fréquence de sortie de l'information est beaucoup plus faible que la fréquence de récurrence du radar.

Le temps d'observation est divisé en intervalles jointifs de durée θ ; chaque intervalle constitue un cycle de filtrage ; la période de sortie des informations filtrée est θ .

Chaque train d'impulsions reçu d'une cible au cours du cycle de filtrage est multiplié par une fonction de pondération h(t), le signal ainsi pondéré est démodulé par deux démodulateurs synchrones travaillant en quadrature et le résultat de cette démodulation est intégré dans un intégrateur numérique. La fonction h(t) est la réponse impulsionnelle d'un filtre passe-bas ; la sortie de l'intégrateur numérique dépassera un certain seuil si le produit de la démodulation est un signal dont la fréquence est voisine de la fréquence nulle, ou d'un multiple de la fréquence de récurrence.

Si l'on désigne par H(f) la transformée de Fourier de h(t), on peut considérer que le système est équivalent à un filtre dont la réponse en fréquence est $|H(f - f_1)|^2$, f_1 désignant la fréquence de démodulation.

Pour éviter un phénomène de phases aveugles, il y a deux démodulateurs en quadrature. Plusieurs filtres identiques dont les fréquences centrales f_1, f_2, f_3 ... sont calculées de façon à obtenir un léger recouvrement des plages de fréquence permettent de détecter les signaux sur l'ensemble de la plage Doppler utile. Le schéma de principe du banc de filtres est donné en figure 2.

La figure 3 donne la fonction de pondération h(t) et la figure 4 le module de sa transformée de Fourier.

La figure 5 indique la courbe de réponse en fréquence de l'ensemble des filtres.

4 - MESURE DE LA DISTANCE ET DE LA FREQUENCE DOPPLER

4.1. Mesure de la distance

Le signal émis est composé de deux trains d'impulsions et les filtres qui traitent ce signal sont sensiblement adaptés à un train d'impulsions ; il en résulte qu'un but mobile donne deux réponses dans un filtre Doppler, et ces réponses sont séparées dans le temps par l'intervalle τ.

Le système de décision de présence qui est indiqué sur la figure 6 permet d'extraire la distance de la cible en fournissant l'écart entre le signal émis et le signal reçu.

4.2. Mesure de la fréquence Doppler

En décrivant le filtrage, nous avons indiqué qu'un filtre utilisait deux démodulateurs synchrones. Les signaux de démodulation sont à la même fréquence mais ont une différence de phase de $\frac{\pi}{2}$.

Le premier but recherché est d'éviter un phénomène de phases aveugles. Montrons qu'il est également possible de connaître la fréquence Doppler des cibles.

Le système de filtrage permet de déceler un effet Doppler ; la fréquence mesurée peut cependant différer de celle de la cible d'une quantité $k f_r$ multiple de la fréquence de récurrence, k étant un nombre entier que l'on veut connaître.

Pour cela faisons l'hypothèse que la fonction h(t) est paire, cela n'est pas une nécessité mais simplifie l'exposé.

Considérons un écho situé dans le premier filtre , donc un écho dont la fréquence f est telle que $f - f_1$ est voisin de $k f_r$. Le signal reçu se compose de quatre termes : Pour le premier train d'impulsions, à la fin de l'intégration, nous avons :

$$A = \sum_{i=-n}^{i=+n} h(i T) \cos\left[2\pi(f - f_1)\, i T + \varphi\right]$$

et

$$B = \sum_{i=-n}^{i=+n} h(i T) \sin\left[2\pi(f - f_1)\, i T + \varphi\right]$$

Pour le second train d'impulsions :

$$C = \sum_{i=-n}^{i=+n} h(i T) \cos\left[2\pi(f - f_1)\, (i T + \tau) + \varphi\right]$$

et

$$D = \sum_{i=-n}^{i=+n} h(i T) \sin\left[2\pi(f - f_1)\, (i T + \tau) + \varphi\right]$$

En tenant compte du fait que h(t) est paire, il vient :

$$\begin{cases} A = \cos\varphi \sum_{i=-n}^{i=+n} h(i\,T) \cos 2\pi(f - f_1)\, i\,T \\ B = \sin\varphi \sum_{i=-n}^{i=+n} h(i\,T) \cos 2\pi(f - f_1)\, i\,T \end{cases}$$

$$\begin{cases} C = \cos\left[\varphi + 2\pi(f - f_1)\tau\right] \sum_{i=-n}^{i=+n} h(i\,\tau) \cos 2\pi(f - f_1)\, i\,T \\ D = \sin\left[\varphi + 2\pi(f - f_1)\tau\right] \sum_{i=-n}^{i=+n} h(i\,T) \cos 2\pi(f - f_1)\, i\,T \end{cases}$$

A et B sont les composantes d'un vecteur V_1 de module

$$M = \sum_{i=-n}^{i=+n} h(i\,T) \cos 2\pi(f - f_1)\, i\,T$$

et de phase φ

C et D sont les composantes d'un vecteur V_2 de même module et de phase $\varphi + 2\pi(f - f_1)$

La différence de phase $\Delta\varphi$ entre V_1 et V_2 permet de calculer $k\,f_r$.

On a :

$$k\,f_r = \frac{\Delta\varphi}{2\pi\tau}$$

Ce qui donne la fréquence Doppler f_d

$$f_d = f_1 - F + \frac{\Delta\varphi}{2\pi\tau}$$

F est la fréquence émise par le radar, f_1 la fréquence de démodulation.

Le système est donc capable de lever des ambiguïtés de fréquence sur une plage égale à $\frac{1}{\tau}$.

Avant de comparer ce dispositif à ceux généralement utilisés pour effectuer un filtrage Doppler, il est bon de remarquer que le schéma de la figure 2 n'est qu'un schéma de principe.

Dans les réalisations les plus intéressantes, seuls subsistent deux détecteurs synchrones utilisant la même fréquence de démodulation.

La sortie des démodulateurs est pondérée par la fonction h(t) puis codée et un traitement numérique tout à fait semblable à celui qui sert à calculer une transformée de Fourier permet d'obtenir le signal de sortie des filtres.

5 - CONCLUSION

Il est intéressant de comparer un radar cohérent à basse fréquence de récurrence à un radar réalisé selon le schéma que nous venons d'indiquer.

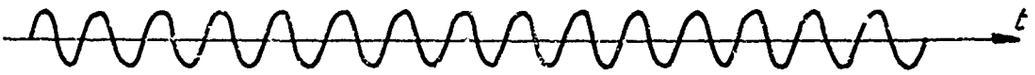
Les possibilités de réjection d'échos fixes sont semblables. Le radar que nous avons décrit introduit une ambiguïté de distance pour les cibles mobiles, ambiguïté qui peut être facilement levée par la logique de détection. Par rapport aux radars connus, il présente l'avantage de permettre la mesure de la fréquence Doppler et cela peut être extrêmement utile pour éliminer des cibles mobiles mais lentes. Si le radar sert à la veille aérienne, il lui est possible d'éliminer tous les véhicules terrestres détectés.

D'autres particularités du système sont liées au mode de filtrage et plus particulièrement au fait de découper le temps d'observation en intervalles de filtrage ne se chevauchant pas. De cette particularité résulte la faible fréquence de lecture des filtres et dans un radar dont l'antenne effectue une rotation continue la faible précision de la mesure angulaire.

Il est par contre possible de profiter de cette particularité pour changer les caractéristiques d'émission d'un cycle de filtrage au suivant ; ce qui permet, de diminuer l'effet des zones de vitesses aveugles par changement de la fréquence de récurrence, et de diminuer l'effet des fluctuations de cible par changement de la fréquence d'émission.

Ces changements des caractéristiques d'émission permettent d'accroître notablement la probabilité de détection du radar.

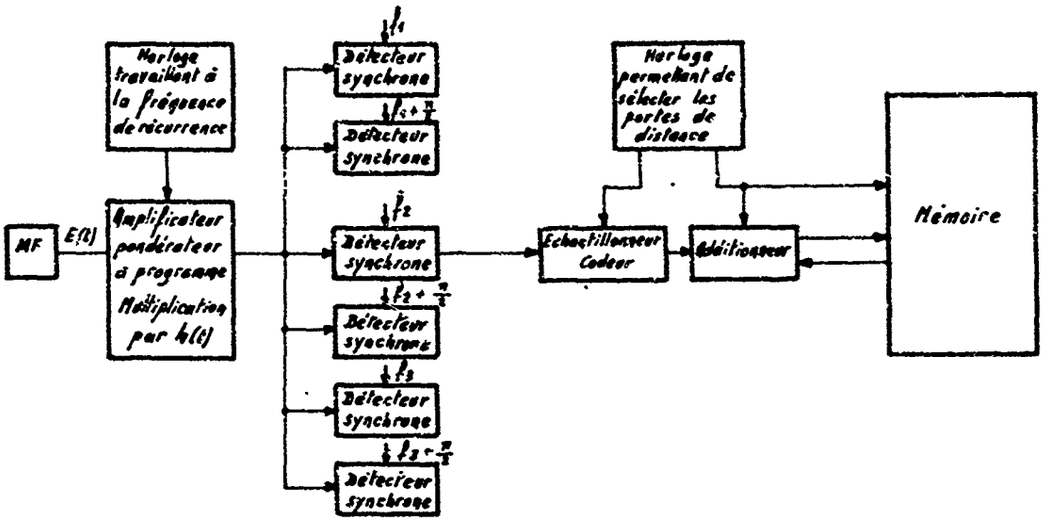
Nota : Le dispositif décrit fait l'objet du brevet français n° 1 555 854.



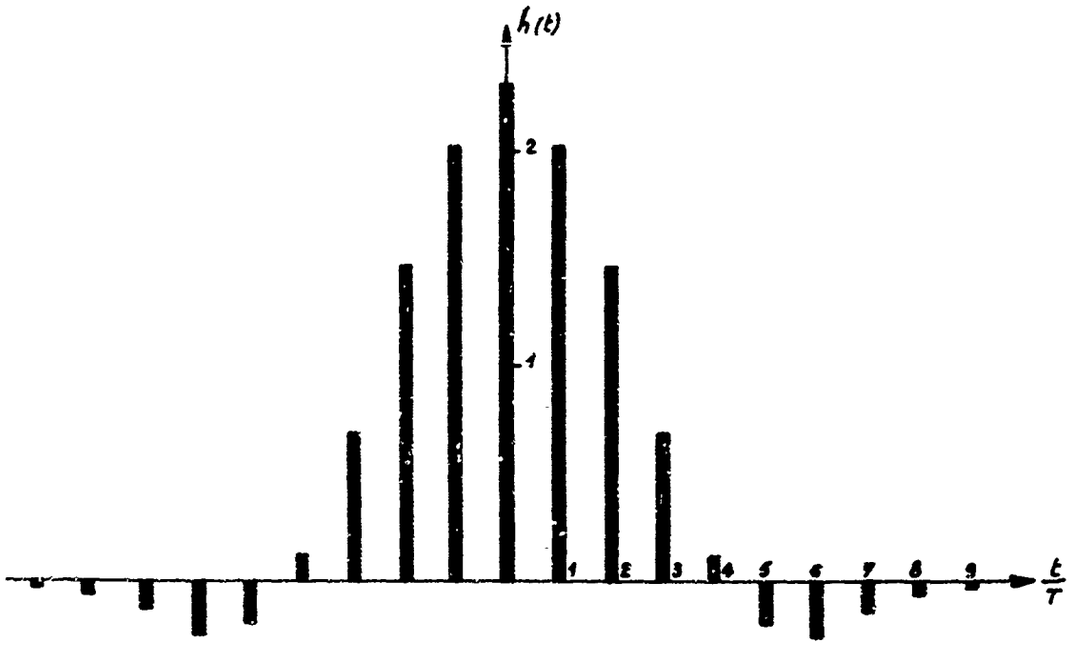
- Onde porteuse cohérente -

- Signal de modulation de l'onde porteuse -

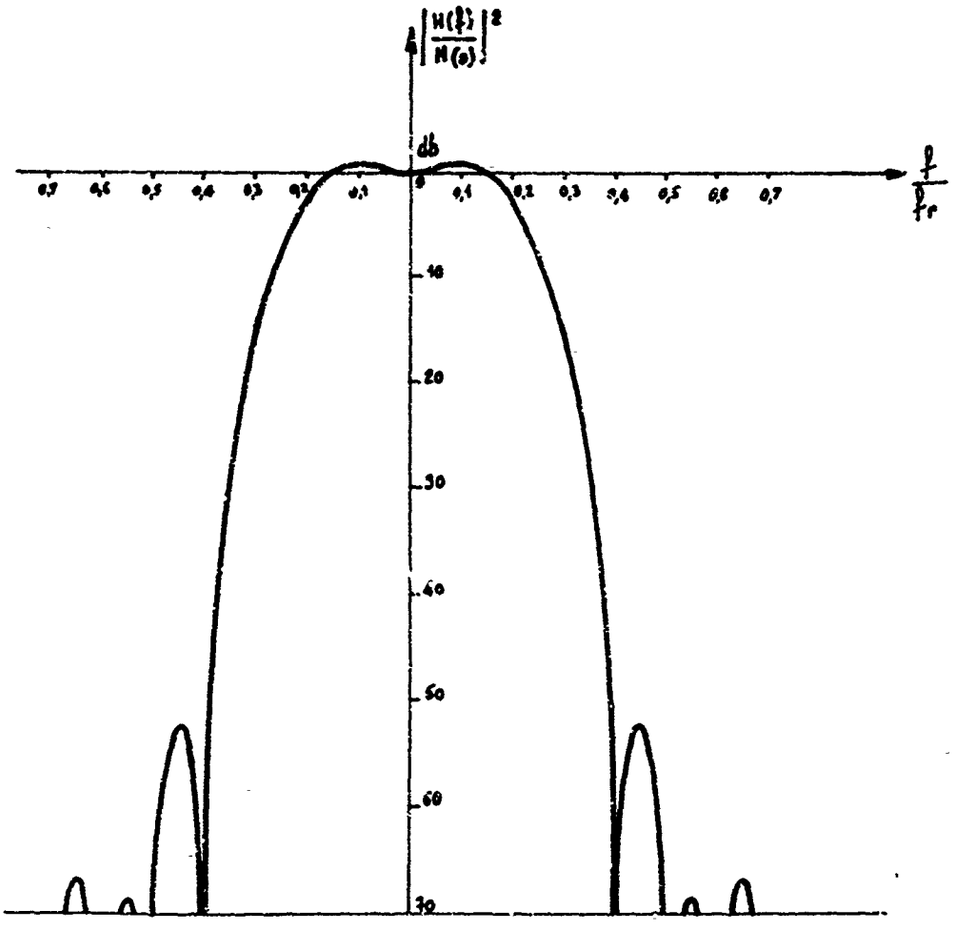
- Signal émis - Fig. 1



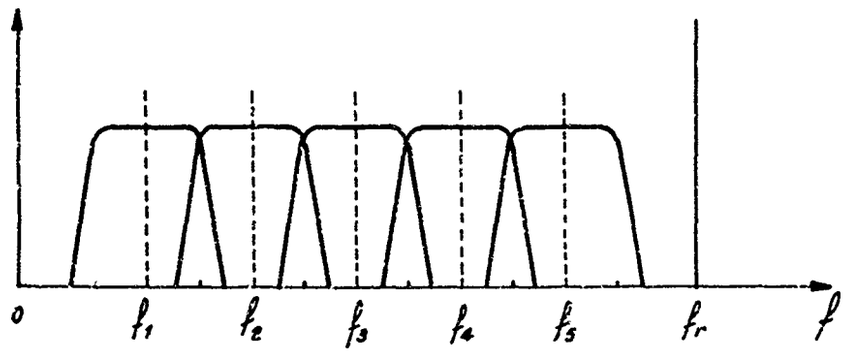
- Schéma de principe du banc de filtres - Fig. 2



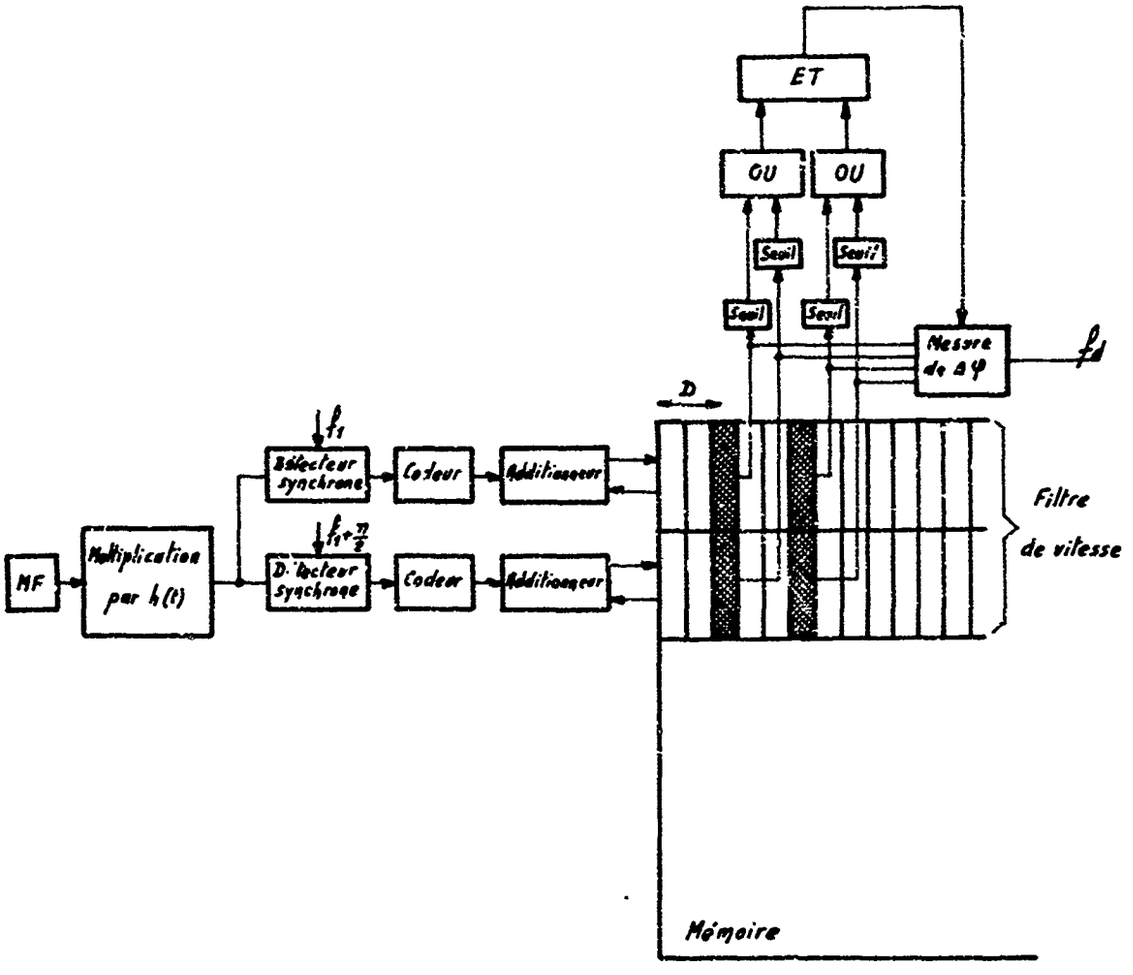
- Fonction de pondération - Fig. 3



Module de la transformée de FOURIER de la fonction de pondération - Fig. 4



- Réponse en fréquence de l'ensemble des filtres - Fig. 5



- Logique d'extraction d'un filtre - Fig. 6

DETECTION OF DOPPLER SHIFTED RADAR SIGNALS WITH CLUTTER REJECTION

by

W.D.Wirth

Forschungsinstitut für Funk und Mathematik
5307 Wachtberg-Werthhoven, Germany

DETECTION OF DOPPLER SHIFTED RADAR SIGNALS WITH CLUTTER REJECTION

W. D. Wirth

FORSCHUNGSINSTITUT FÜR FUNK UND MATHEMATIK

5307 Wachtberg-Werthhoven, Germany

SUMMARY

The problem of detecting doppler shifted radar signals in clutter and noise is treated with the likelihood-ratio test. For the application of this test, which is optimal with respect to the error probabilities, informations about targets and clutter are necessary. A separate measurement and data evaluation program resulted in sufficient knowledge about target and clutter signals to construct the test function. The detection system consists of a set of weighting functions, which form a filter bank to cover the whole doppler range. All individual weight functions perform two tasks: detection of their respective doppler signal, if present, and suppression of the clutter signals. An experimental system combines digital and analog techniques. If the spectrum of the clutter is varying with time, the filter may be divided into an adaptive clutter suppression filter and a doppler filter bank. The weighting function for clutter suppression in this case has to be derived from real time estimates of the clutter correlation.

1. INTRODUCTION.

The problem of detecting moving targets and simultaneously suppressing clutter echoes is common to all surveillance radars, especially to those for the acquisition of low flying aircrafts. The SCV (Sub Clutter Visibility) depends on the frequency stability and other parameters of the radar and on the signal processing system. Our objective was the development of a detection system for moving targets in clutter with some improvements against the conventional MTI. We have to admit an improved system of being more complex. On the other hand, considering the new possibilities of modern integrated circuits for analogue and digital devices, this seems to be no handicap. The desired improvement is achieved by the application of the fundamental results of the statistical decision theory to our special problem, together with an extensive program for the measurement and evaluation of radar signals from different types of targets and clutter. This method allows generally to optimize the signal processing system for a certain radar.

2. SIGNAL REPRESENTATION.

Amplitude and phase of the received signals shall be used. The phase is related to the phase of the transmitted pulse. Furthermore the signals have to be sampled before quantization at a rate of $1/2\tau$, where τ equals the width of the transmit pulse. Thus the whole radar area is divided up into range elements. For each range element we get an infinite sequence of echo vectors at a rate equal to the pulse recurrence frequency (prf). (fig. 1) Series of target echoes belong to their corresponding range elements, hence, the detection procedures have only to process signals belonging to single range elements.

The echo vectors may be represented by their amplitude and phase or by their orthogonal components. We choose for our purpose the representation by components. The reason simply is, that echoes of targets and clutter superpose linearly their components and we can hope to extract them with a linear system. The signal components may be derivated from the IF-signal by multiplication with the coho-signal respectively the 90°-shifted coho signal. After a low pass filter the orthogonal components are sampled.

3. SIGNAL MODEL FOR TARGETS AND CLUTTER.

Before we can apply the statistical decision theory to a special detection problem, suitable models for the target and interference signals have to be chosen.

The model for the echo signals of an aircraft during one scan is a series of stepwise rotated signal vectors. (fig. 2) Their length is determined by the target cross section modulated by the antenna function. The phase difference between successive vectors is given by the doppler frequency shift ω_d, which depends on the radial velocity of the target, multiplied with the radar pulse period T. The series of signal components shows samples of the doppler signal with the antenna function as envelope. This model of an ideal aircraft echo may differ from reality in two points:

The cross section of an aircraft may fluctuate during the scan (Swerling case II) and the phase differences may be more or less random so that there is no predominant doppler signal. Both effects may be caused by propeller rotation and vibration effects.

With our measurement and evaluation program for radar signals (RADICORD) /1/, /2/, /3/, echoes of different types of aircrafts along tracks were recorded. One purpose of the evaluation of these data was to clarify these points. The amplitude series according to the 3 dB-beamwidth is used to compute the standard deviation of the amplitude, normalized to the respective mean amplitude. The influence of the antenna function was elimi-

nated. Fig. 3 shows a typical example of a propeller aircraft. Compared with the steady part, the power of the fluctuating part of the signal is low enough to generally verify our signal model in this point. For jet aircrafts the model fits even better. The phase differences and their deviation were computed and figured in a corresponding manner. Fig. 4 shows an example for a jet aircraft and fig. 5 for a helicopter. Both examples show the mean phase differences for single scans along the track and the deviation of the phase differences within the single scans. Compared with the mean phase differences, the deviation is small enough for both extreme examples verifying our signal model in this point as well.

Single fixed point targets produce a series of echo vectors with constant phase, only the amplitude is modulated by the antenna function. If the antenna scans an area with randomly distributed fixed scatterers, the resultant echovector is slowly varying its amplitude and phase. Slowly means in comparison with the radar pulse period. This is due to the so called "scanning effect": because the antenna beam is shifted only a small part of its beamwidth between successive pulses, there is only a small change of the illuminated clutter area. In most cases the illuminated clutter area contains many independent scatterers, and with the "central limit theorem" we may assume, that the signal components are normally distributed. Furthermore the time behaviour of the clutter signals is influenced by the internal motions of the scatterers, instabilities of the radar and the superposition of receiver noise and is described by their correlation function.

If we know the correlation function of the clutter signals, we can formulate the n-dimensional normal distribution density as the joint distribution density for n successive clutter signals. The correlation function for fixed clutter is closely related to the antenna function, and this relation may be used to compute an idealized fixed clutter correlation. But to cover all other effects cited above, it is necessary to estimate this correlation from recorded clutter signals. Within the above mentioned RADICORA-program clutter signals were recorded and evaluated for several radars and different types of clutter: fixed clutter, terrain clutter and weather clutter. By using the Fourier-transform we get the power spectrum out of the correlation. Both are shown as an example for fixed clutter on fig. 6.

Note the main part of the clutter spectrum around doppler frequency zero and the noiselike part at all higher frequencies in the unambiguous frequency interval. This "clutternoise" is produced by uncorrelated signals due to radar instabilities. The density of this noise determines the improvement factor for the S/C ratio achievable by the signal processing system. Fig. 7 shows the correlation and the spectrum of an example of weather clutter. In this case the correlation function is complex because

the main part of the spectrum is doppler shifted, according to the radial wind velocity. The power spectrum of the weather may assume very different shapes according to the respective weather situation.

4. LIKELIHOOD-RATIO TEST.

The detection of moving targets in clutter may be regarded as a problem of testing two hypothesis H_o and H_1. H_o is the hypothesis that there is only clutter and noise. H_1 is the hypothesis that there is a moving target with a doppler frequency shift. We need a decision rule to select after receiving a signal one of the two hypothesis. There are two decision errors possible:

1. Decision for H_1, but H_o is true. This event is a false alarm, its probability is usually named α.

2. Decision for H_o, but H_1 is true. This event is a miss of a target, its probability is usually named β.

Optimizing the decision rule means minimizing α for a preselected β or vice versa.

The solution of this problem is given by the well known likelihood-ratio test (fig. 8). The set z of received signals has to be inserted into the test function $\lambda(z)$. If $\lambda(z)$ exceeds the threshold η, H_1 is chosen.

Assuming H_o to be correct, $p_o(z)$ is the joint probability density function for z. $p_1(z)$ is the joint probability density function for z if a target is present (hypothesis H_1), averaged with respect to the possible parameter space of the signal.

Fig. 9 shows the test function for the detection of a certain signal s_i, with doppler frequency $\omega_d = \omega_{d_i}$ as parameter, out of correlated clutter signals and noise. z' is the row matrix containing the series of signal components corresponding to one range element. The length n of the series may be chosen equal to the number of pulses within the 3 dB beam width. The elements of the correlation matrix K are the measured correlation values of the clutter. Multiplying K^{-1} and s_i we find the column matrix h_i. The result of the scalar product of z and h_i is y_i. y_i may be regarded as the output of a time domain filter with the weighting function h_i.

If the whole doppler frequency range is covered by M weighting functions h_i, the resulting test function would be the average of all M different $\lambda_i(z)$ (fig. 10). This test function is very difficult to handle with respect to a signal processing system

because of the exponential function. Fortunately a simulation study /4/ resulted in a negligible loss for the following alternative test function:

Take the maximum of all y_i (i = 1...M) for the threshold comparison.

A small signal approach leads to the average of the y_i as the test function. But this test shows very poor results for reasonable signal to noise ratios.

The processing system must have a shifting memory with a shift rate equal to the prf and with n stages for each range element, a weighting network acting as a bank of filters, and a threshold comparator. Because clutter from fixed targets is the most severe interference, the correlation of fixed clutter is used to compute the weighting functions h_i. As an example fig. 10 shows the frequency transfer functions of the single filters. This representation of the filters demonstrates their clutter suppression property:

All single weighting functions have a high attenuation for frequencies with a high clutter spectral density and have their passbands for their respective doppler frequencies. Because of the finite length n of the weighting functions, all filters have sidelobes in the spectral range of the noise. In our example we have chosen n = 15.

Some remarks to our clutter model:

In contrast to our assumption there are sometimes predominating single fixed point targets. But their signals have a spectral density equal to that of clutter areas and are therefore suppressed as well.

The clutter scatterers are distributed non-uniformly in space. Therefore the clutter signals are non-stationary. If the false alarm probability α shall be constant, it is necessary to estimate the clutter power and to control correspondingly the detection threshold.

5. EXPERIMENTAL SYSTEM FOR THE DETECTION OF TARGETS IN FIXED CLUTTER

We developed an experimental model for the realization of the test function described above. This system combines analog and digital techniques to achieve an economic, flexible and reliable solution. Fig. 11 shows the block diagram of the system. The IF signal is converted into the orthogonal components x and y. These are sampled with 1 MHz sample rate and converted into digital form by the two ADCs. The ADC has 8-bit resolution. The delay of the x and y-components is realized by a core memory. Each of the 1760 range

elements is related to a memory address. The shifting operation with 14 delay steps for the echoes of one range element takes place within one memory word. After delaying the signals are converted back into the analog form and weighted with a resistor-network. The output of the summing network is compared with the detection threshold. One weighting network is working for all range elements in real time. The boards with the weighting networks forming the filter bank can be easily changed. Thus the system may be adapted to different radars. The system offers the opportunity to measure the doppler frequency of targets. Fig. 12 shows a front view of the system. The left side contains the core memory with 1 µs cycle time, 1760 addresses and a word length of 250 bits. The right side contains the gain controlled IF amplifier, ADCs, DACs, the weighting network and some logics for the time control of the whole system.

6. IMPROVEMENT OF THE SIGNAL-TO-CLUTTER RATIO

A recommended performance measure for MTI-systems is the improvement factor I for the signal-to-clutter ratio S/C.

For a conventional MTI the improvement factor I is about 20 - 25 dB for a doppler frequency equal to the half prf. For other doppler frequencies I has lower values.

The improvement factor of an optimal test function depends on the power spectrum of the clutter. If the power spectrum of fixed clutter is normalized to one at zero doppler frequency, the power density of the "clutter noise", caused by radar instabilities, is r. The possible improvement for the S/C ratio is then about r^{-1}. The resolution of the analog-to-digital conversion has to be chosen adequately to the clutter noise, to have no significant noise contribution by quantization.

For a first test we combined our system with a magnetron radar. Fig. 14 shows a PPI picture of the output signal of the detection system with clutter suppression for several revolution. There are some aircraft tracks and some false alarms by clutter. With unchanged sensitivities but with clutter signals, realized with an additional filter for zero doppler frequency, the following picture was made (fig. 15).

7. ADAPTIVE CLUTTER SUPPRESSION

If the correlation of the clutter is varying with time, an adaptive clutter suppression is desirable. In this case the filter system should be devided into an adaptive pre-whitening filter and a constant doppler filter bank.

The problem is to find a weighting function for the suppression of the clutter signals.

We assume that the clutter correlation has been estimated by a real time correlator. Then we can derive an estimation function to predict the actual signal out of the recent signals $x_o \ldots x_{n-1}$. The series of differences u between the predicted signal $\hat{x}_n$ and the actual signal x_n is then uncorrelated and therefore contains no clutter signals, if n is great enough and we use the correlation in an optimal manner.

An optimal procedure for the prediction of x_n may be derived with the principle of maximum likelihood estimation:

We have to formulate the joint probability distribution density for the set $x_o \ldots x_n$, substitute the received signals $x_o \ldots x_{n-1}$ and choose $\hat{x}_n$ to maximize the probability density. For normally distributed signals the result is equivalent to the Wiener-filter. The weighting function for getting the difference signal u out of the signals $x_o \ldots x_n$ is proportional to the first column of the matrix K^{-1} of rank n. This function is the desired weighting function for clutter suppression. K^{-1} is the inverse of the correlation matrix K. It is possible to compute the first column of the matrix K^{-1} by a simple iterative algorithm /5/. Therefore the complicated inversion of the correlation matrix is avoided. Fig. 17 shows as an example the spectrum of weather clutter without and with a clutter suppression filter of length n = 15. The power spectral density at the output of the clutter filter is nearly constant or white. The accuracy of the whitening of the clutterspectrum depends of course on the length n of the weighting function.

So it seems possible to realize an adaptive clutter suppression filter, consisting of a correlation estimator, an arithmetic unit and a weighting filter. The implementation within our detection system is prepared and we look forward to the various investigations.

8. REFERENCES

1. Baars, E.P. "Digitale Registrierung von Radar-Echosignalen" Nachrichtentechn. Z. 21 (1968) no. 12

2. Wirth, W.D. "Clutter- und Signalspektren aus Radicord-Aufnahmen und die Bewegtzielentdeckung durch Impuls-Radar" Nachrichtentechn. Z. 21 (1968) no. 12

3. von Schlachta, K. "Characteristics of echoes from aircrafts and clutter derived from digital recorded radar data" XIX Avionics Panel-Symposium Advanced Radar Systems

4. Wirth, W.D. "Die Entdeckung von Radarsignalen mit unbekannter Dopplerfrequenz und der Gewinn durch adaptive Filterung" Nachrichtentechn. Z. 21 (1968) no. 4

5. Burg, J.P. "A new analysis technique for time series data" Nato Advanced Study Institute, Enschede 1968

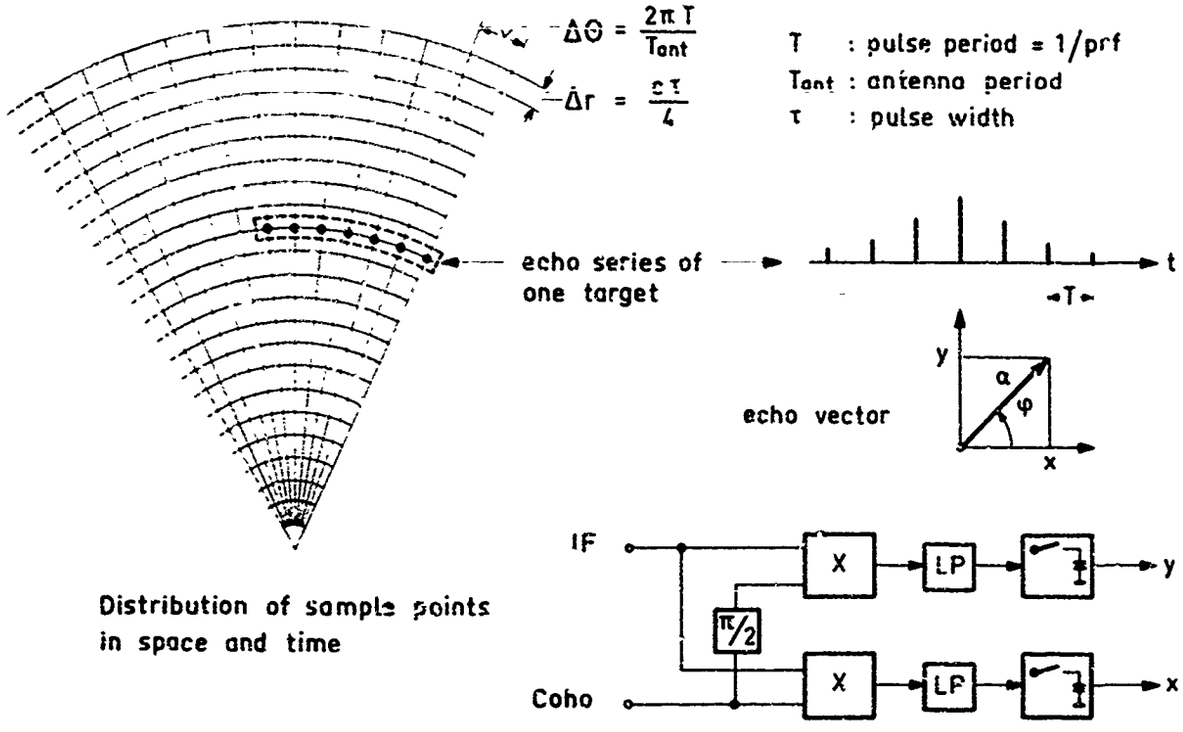


Fig.1 Representation of radarsignals

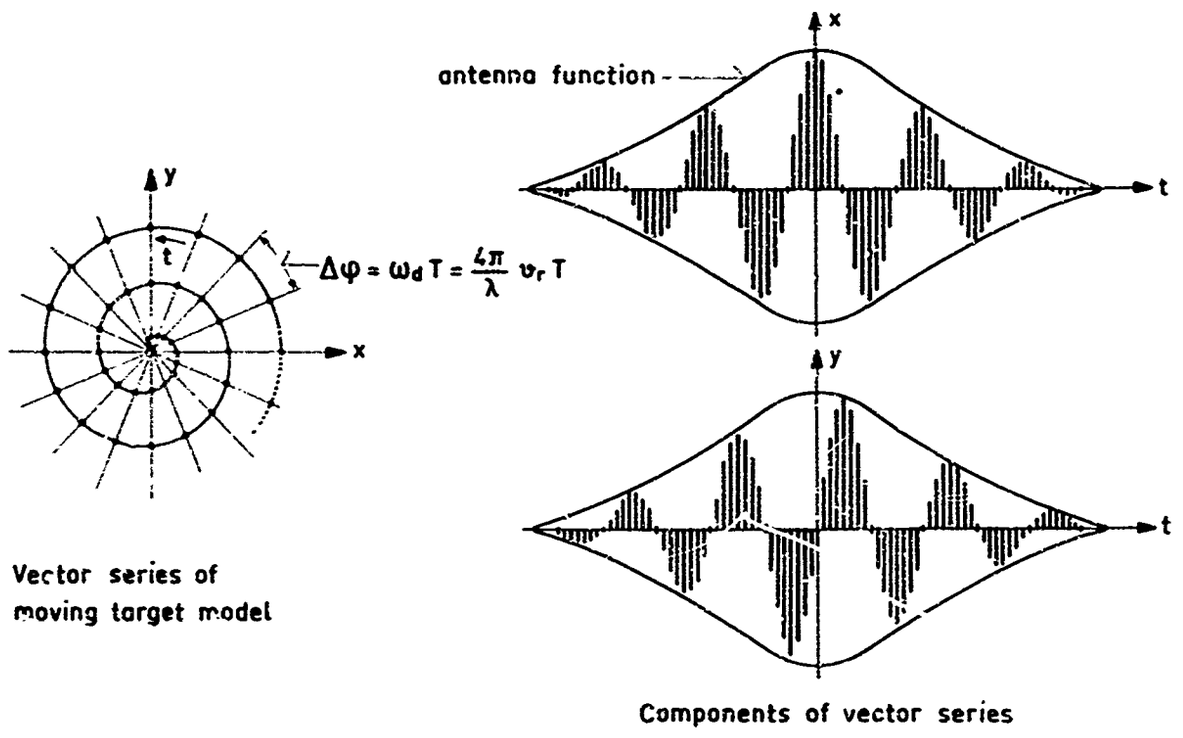


Fig.2 Signal model for moving target

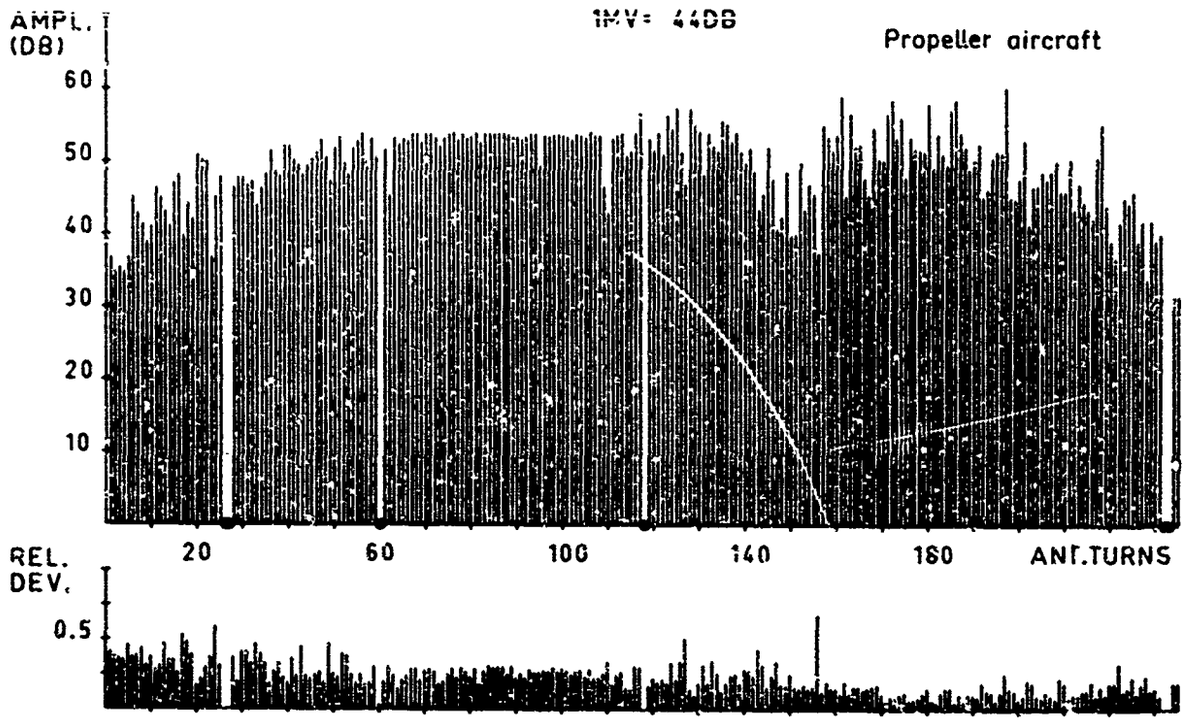


Fig.3 Mean and deviation of target amplitude along a track

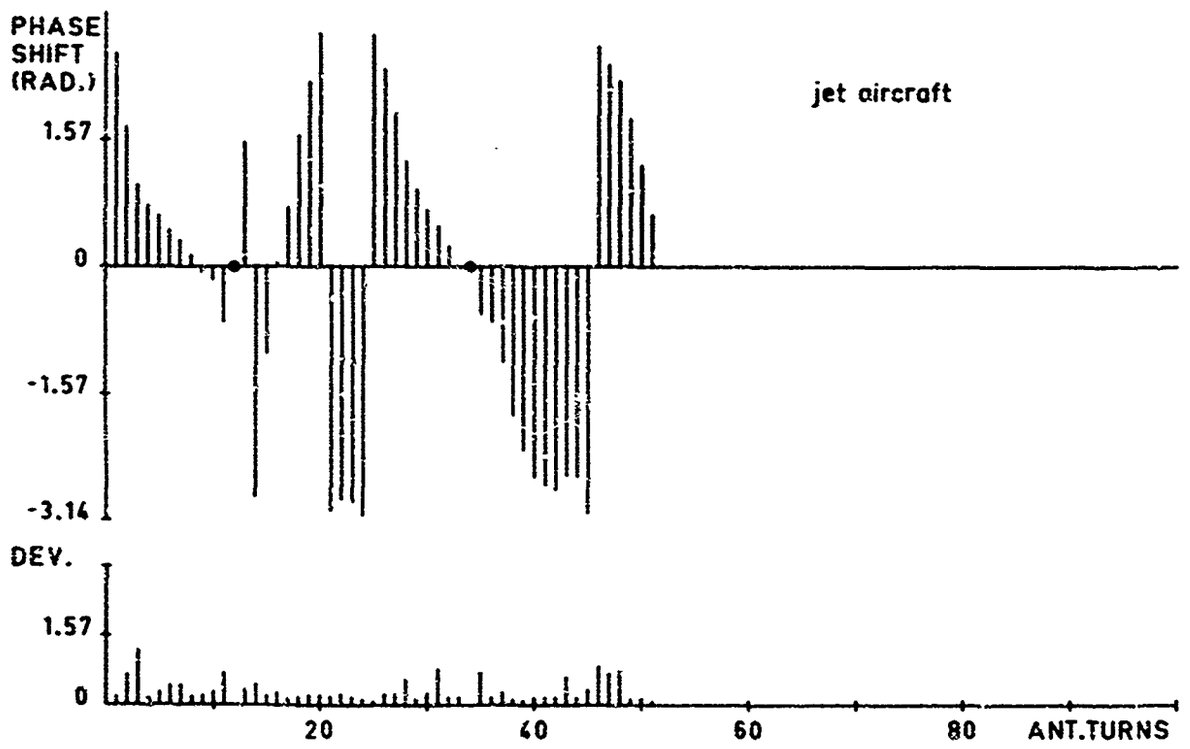


Fig.4 Mean and deviation of dopplershift along a track

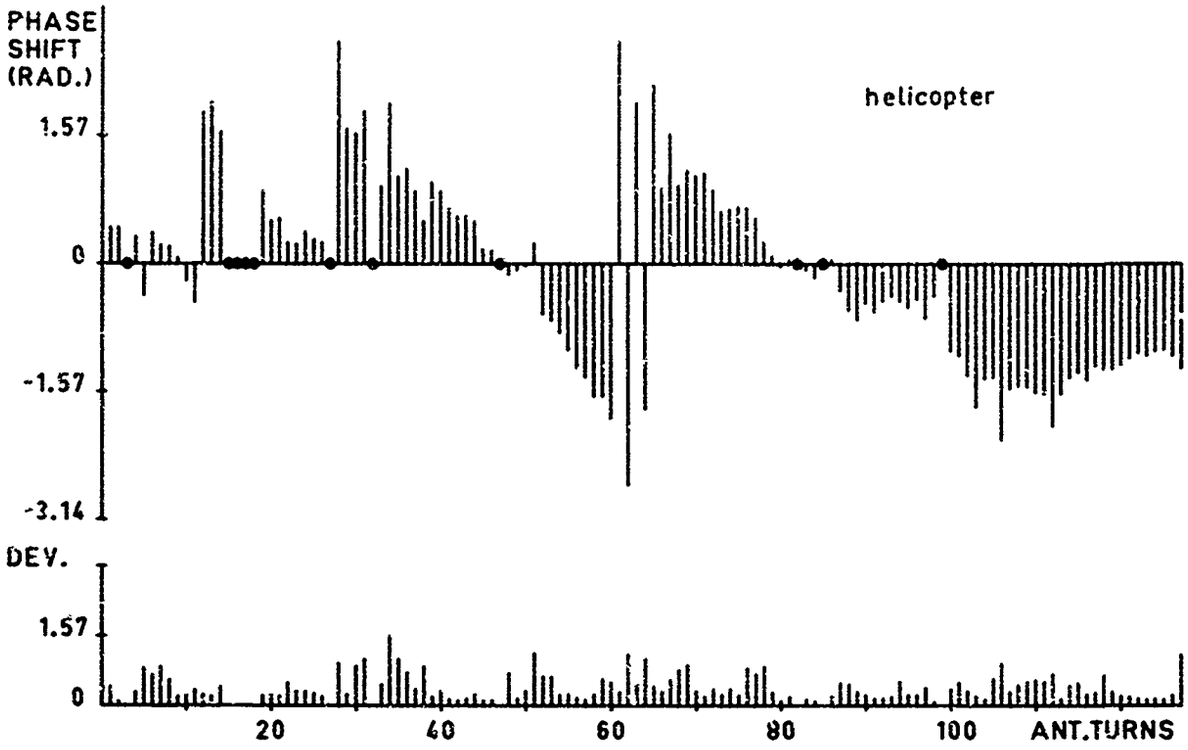


Fig.5 Mean and deviation of dopplershift along a track

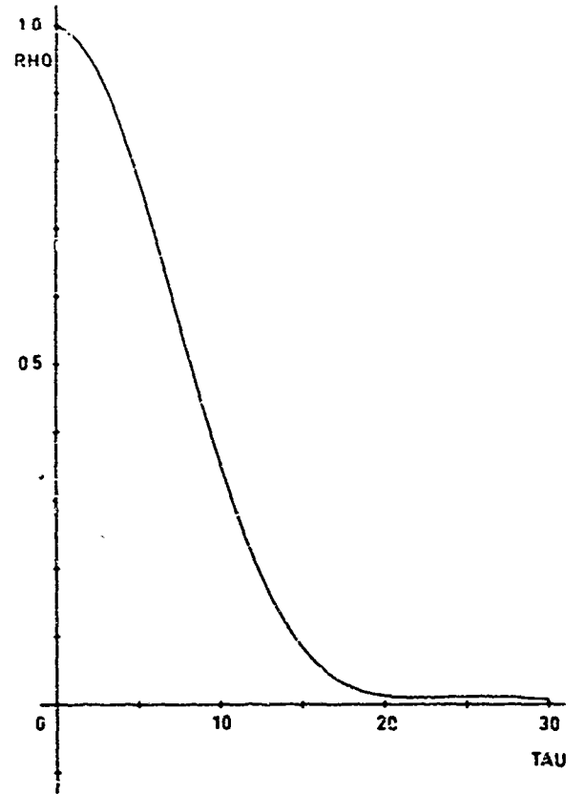


Fig.6(a) Correlation of groundclutter

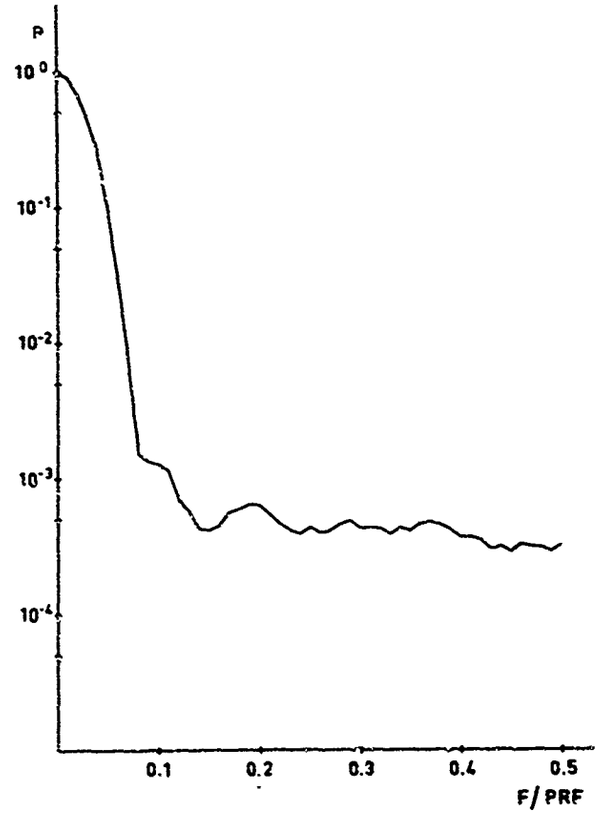


Fig.6(b) Power spectral density P of groundclutter

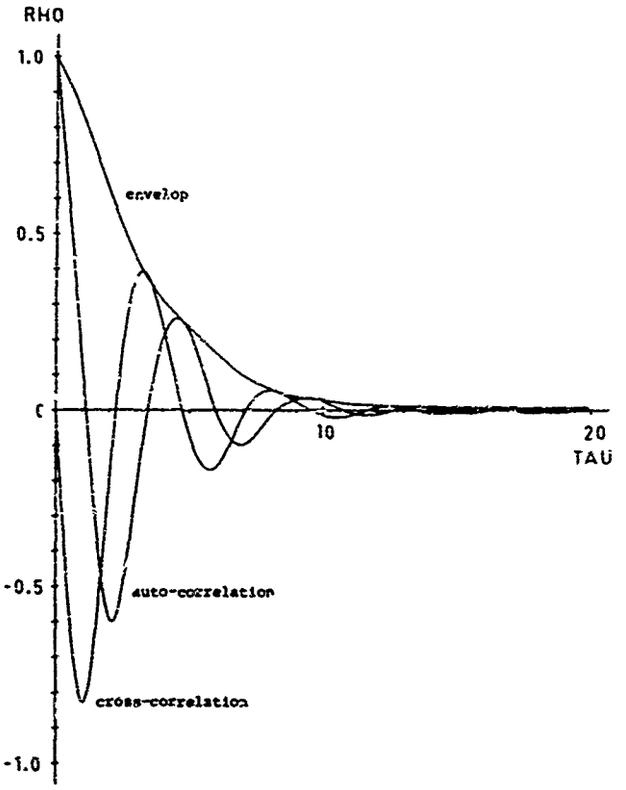


Fig.7(a) Correlation of weatherclutter

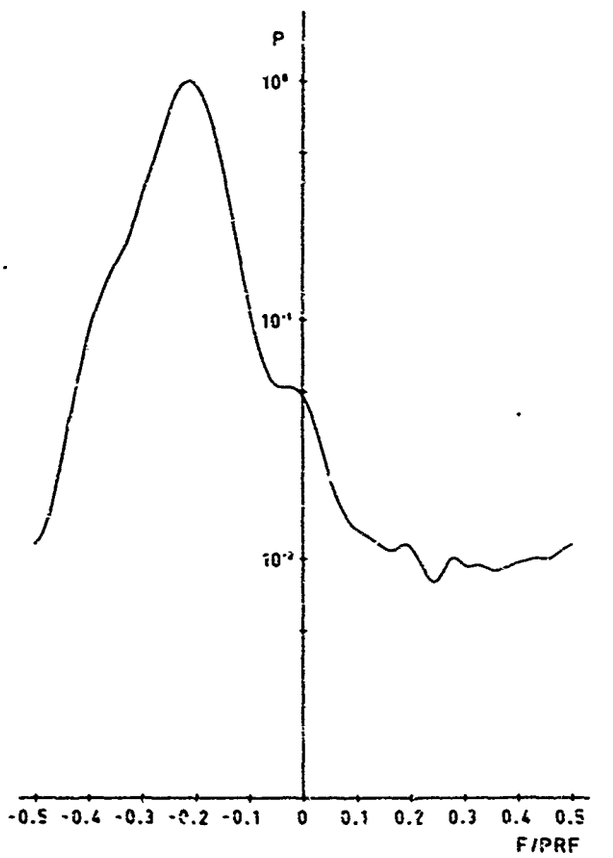


Fig.7(b) Power spectral density P of weatherclutter

Fig. 8 Likelihood ratio test :

$$\lambda\ (z) = \frac{p_1(z)}{p_0(z)} \geq \eta \text{ for } H_1$$

$$p_1(z) = E_\omega \left\{ p_1\ (z/\omega) \right\}$$

z = set of received signals
ω = parameter of target signal

Fig. 9 Test function for detection of signal s_i in correlated normal noise :

$$\lambda_i\ (z) = \exp\ (z'\ K^{-1}\ s_i\ ) = \exp\ (z'\ h_i\ ) = \exp\ (y_i)$$

$$z' = (x_1 \ldots x_n\ \ y_1 \ldots y_n\ )$$

$$s_i' = (a_{i_1} \ldots a_{i_n}\ \ b_{i_1} \ldots b_{i_n}\ )$$

K = correlation matrix

Fig. 10 Test function for M different doppler signals :

$$\lambda\ (z) \triangleq \frac{1}{M} \sum_{i=1}^{M} \exp\ (y_i) \geq \eta$$

Alternative :

$$\max_{1 \leq i \leq M}\ (y_i) \geq \eta_1$$

Small signal approach :

$$\frac{1}{M} \sum_{i=1}^{M} y_i \geq \eta_2$$

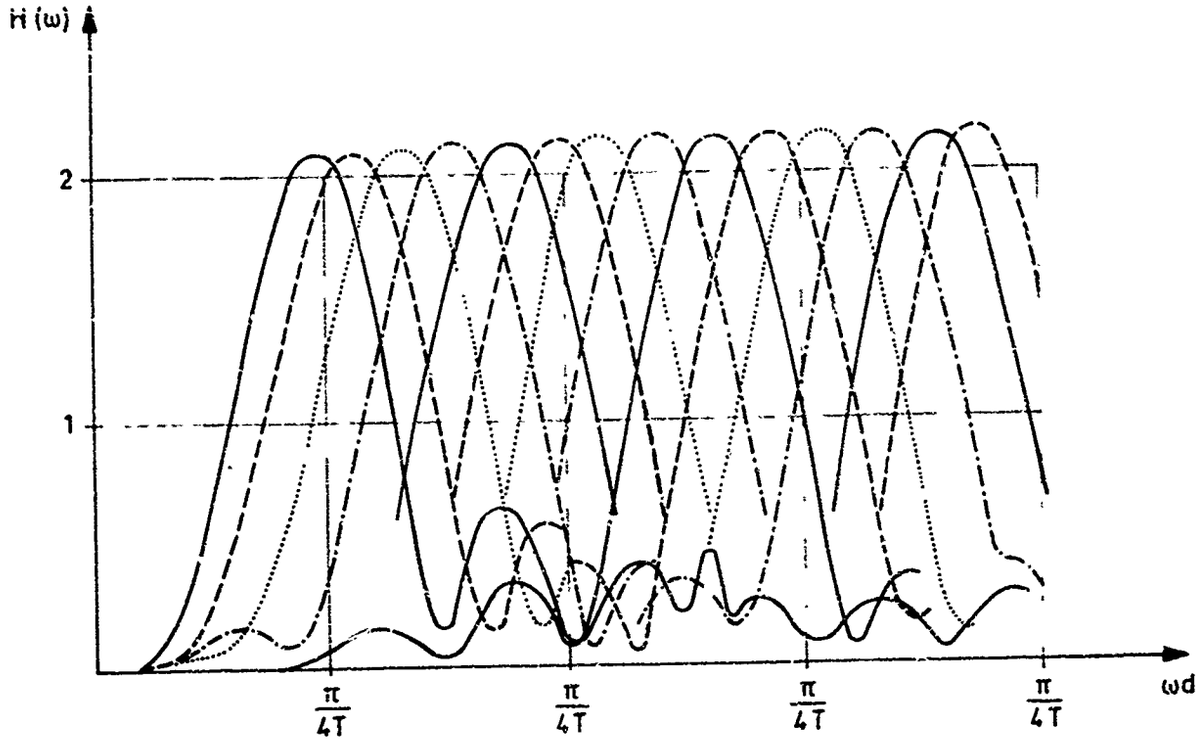


Fig. 11 Filterfunctions of the filterbank

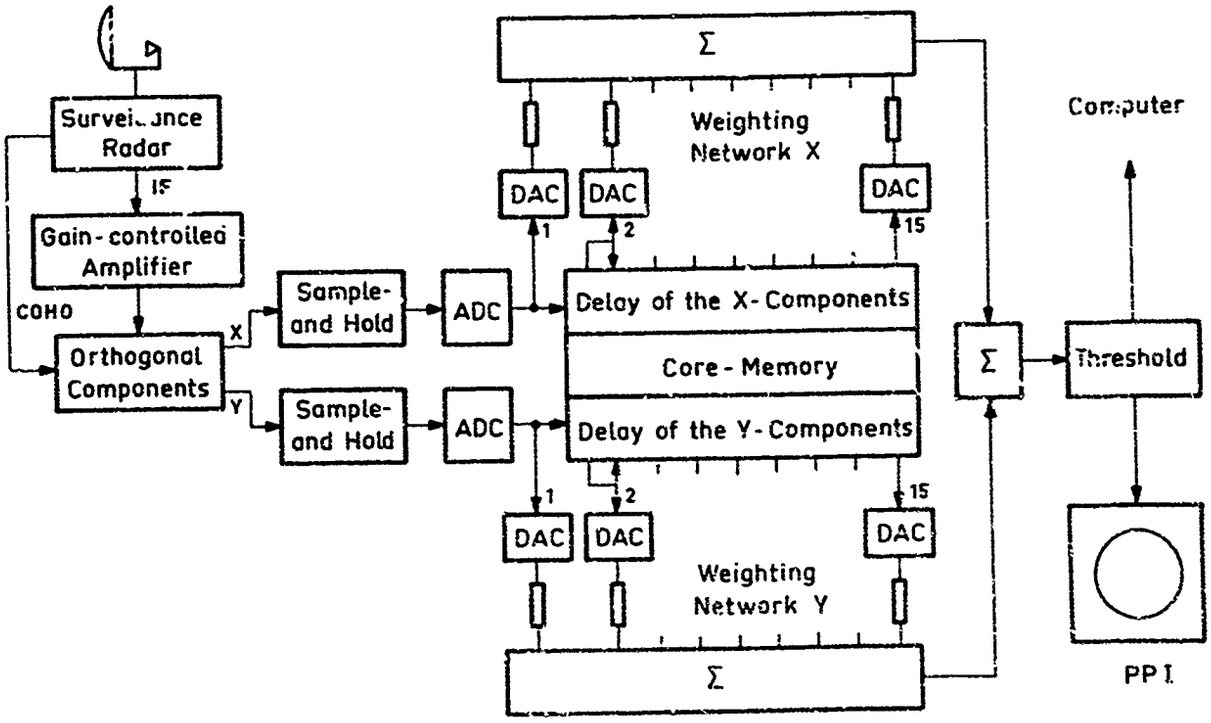


Fig. 12 Doppler-filter-system

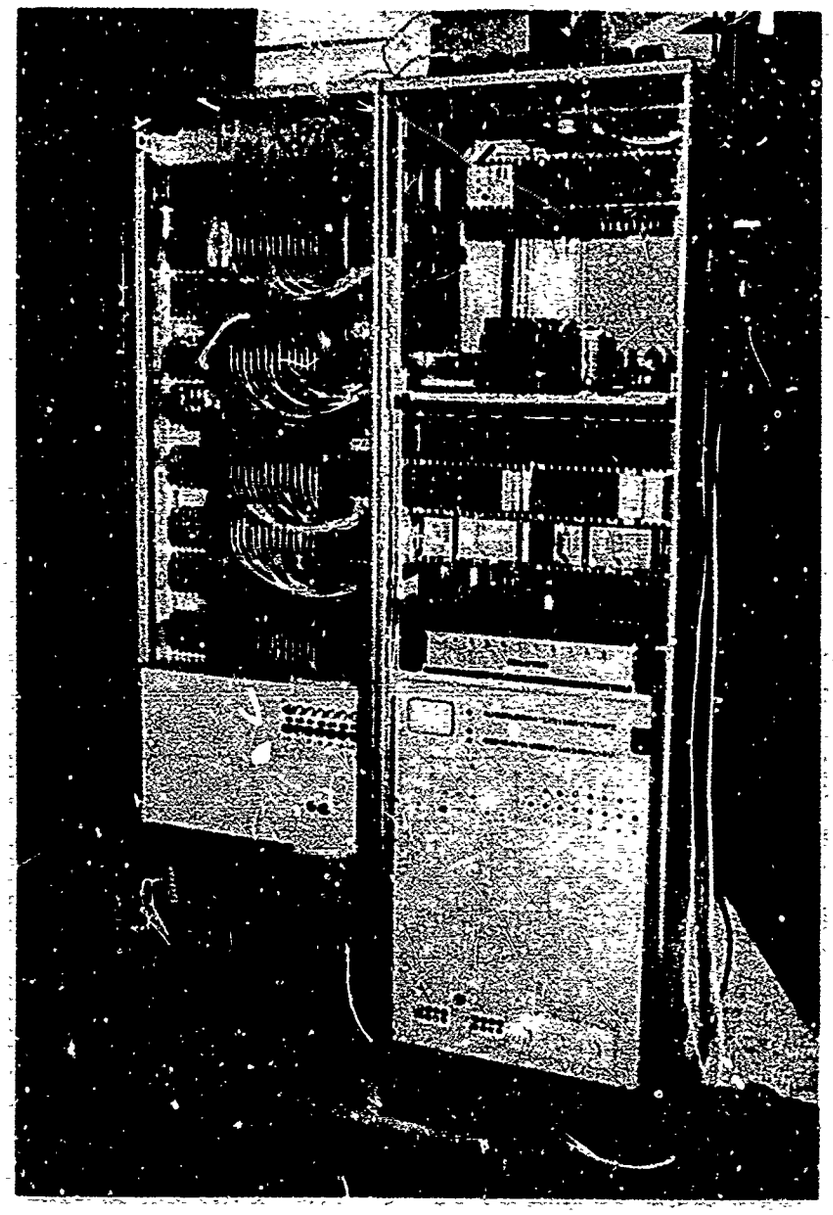
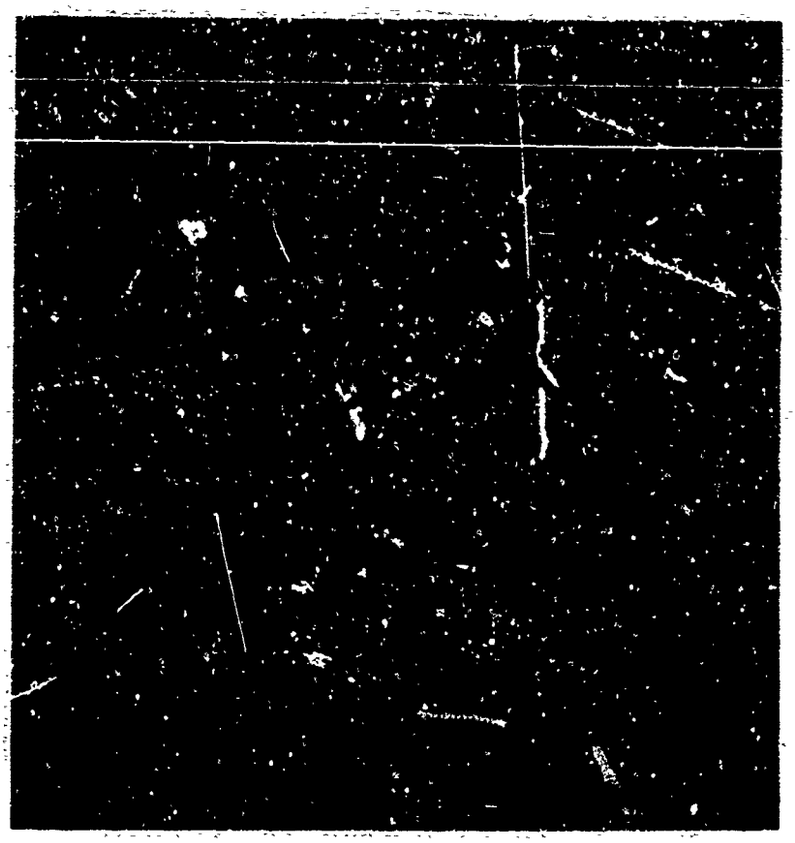
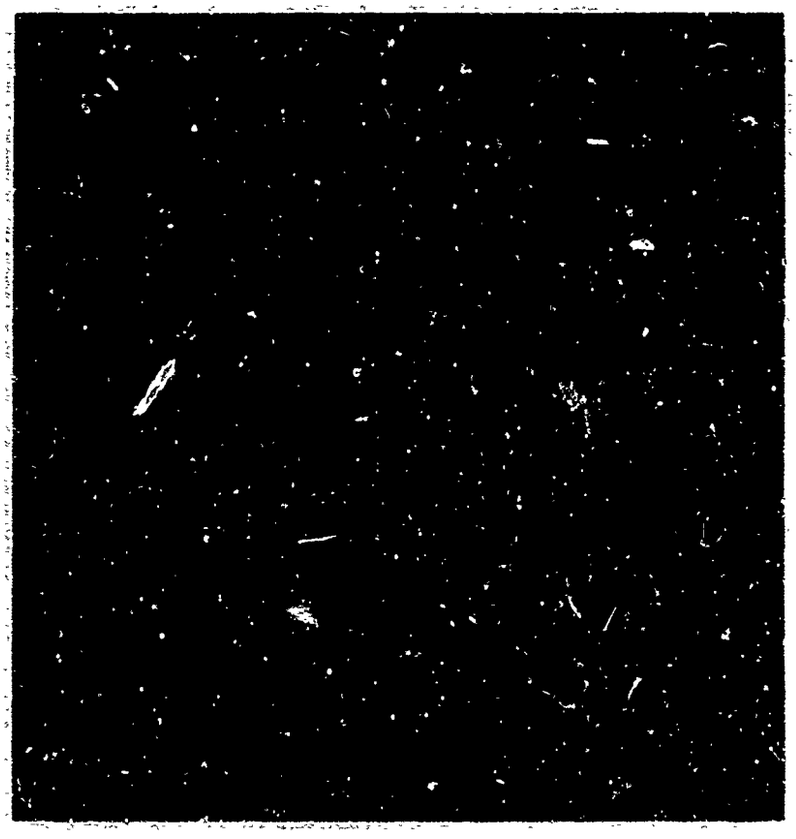


Fig.13 Front view of doppler filter system



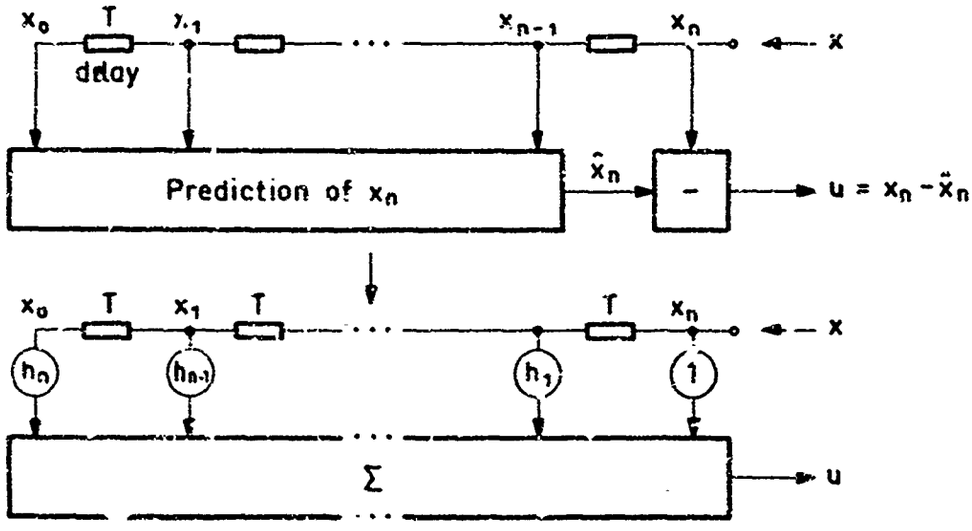
(a) Several antenna revolution



(b) One antenna revolutions

Fig.14 Plots of aircraft on the PPI

Fig.15 Clutter of one antenna revolution



$$K^{-1} = \frac{1}{P_n} \begin{pmatrix} 1 \\ h_1 \\ \vdots \\ h_n \end{pmatrix}$$

P_n : Power of output signal u

Fig.16 Weighting filter for clutter suppression

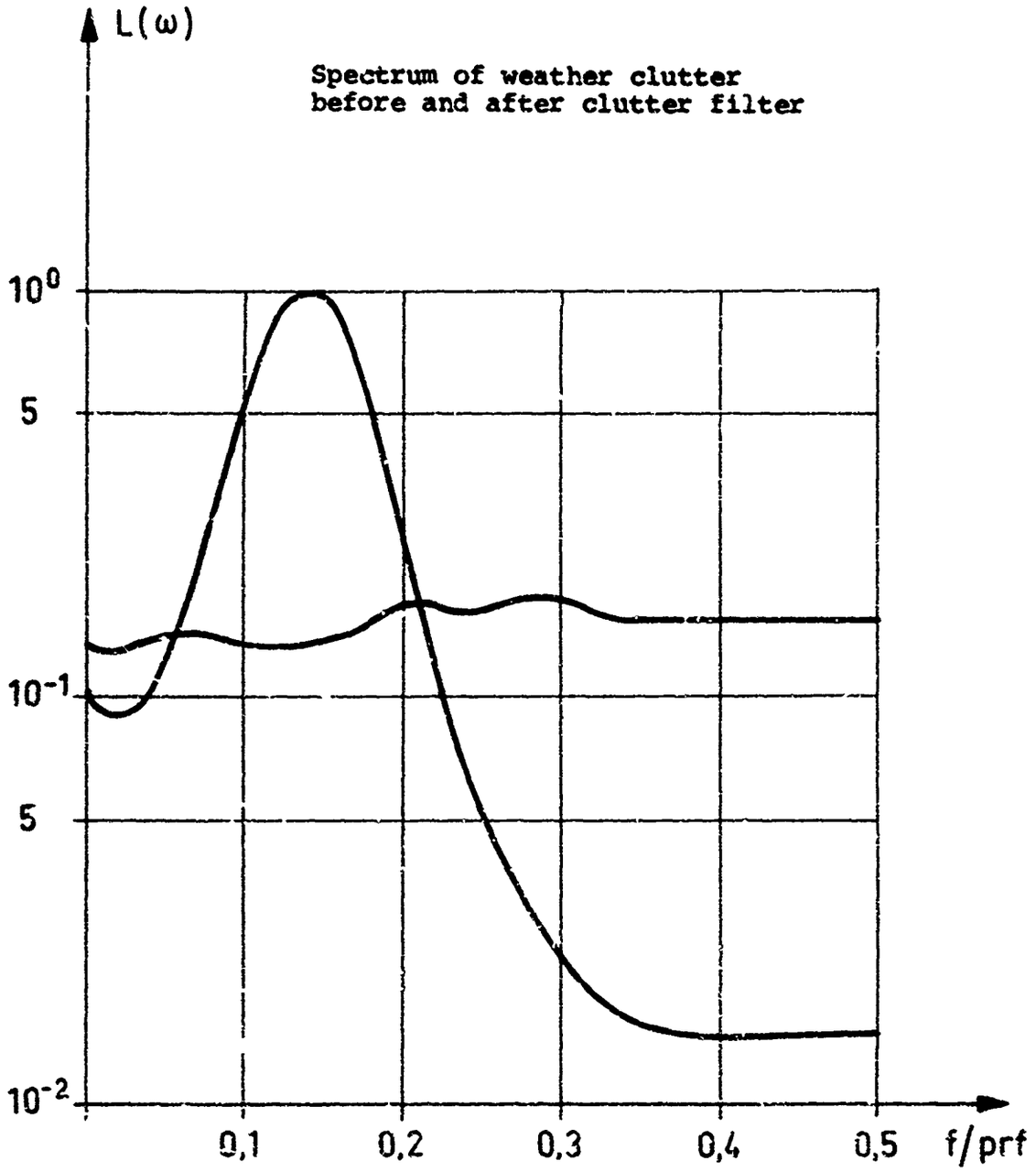


Fig.17 Clutter-filter

DIGITAL SIGNAL PROCESSING FOR RANGE-GATED PULSE DOPPLER RADARS

by

B. Gold and C. E. Muehe

M. I. T. Lincoln Laboratory
Lexington, Massachusetts, USA

DIGITAL SIGNAL PROCESSING FOR RANGE-GATED PULSE DOPPLER RADARS

B. Gold and C. E. Muehe
M.I.T. LINCOLN LABORATORY*
Lexington, Massachusetts, U.S.A.

SUMMARY

For many applications, digital signal processing can increase the effectiveness of range-gated pulse doppler radars. When many doppler filters are desired for each range gate, the fast Fourier transform algorithm is more efficient than digital filtering. In this paper, digital processing methods are briefly reviewed and the signal processing requirements are established in terms of the radar parameters. Equivalence relations are established between the fast Fourier transforms and a bank of velocity filters. Finally, the structure of a general purpose radar digital signal processor is described which should be applicable to a variety of range-gated pulse doppler systems.

I. INTRODUCTION.

In most cases the effectiveness of a range-gated pulse doppler radar is limited by its signal processing capabilities. The gradual introduction of digital signal processing techniques over the last few years[1,2] is a result of the general recognition of the superiority of digital techniques over analog. Digital processing offers filtering free of the usual analog errors due to zero drifts, non-linearities, stray capacities, and component tolerances. It offers exact reproducibility from gate-to-gate with no adjustments to make. The only noise introduced is quantization noise. This noise can be reduced and the dynamic range can be increased in most cases by increasing the number of bits used. The output is in digital form for easy transmission.

Finally, where a large number of range gates are to be processed, digital processing has a cost advantage over analog. The cost of analog filters varies linearly with the number used, costing about $10 a pole. Because a single arithmetic unit of a digital processor can provide filtering for a large number of gates, although the cost for processing a single gate is large, added gates require only the addition of more memory and therefore come cheaply. The crossover point for simple filters (say 5 poles) is several hundred gates. For more complicated filters it is usually more economical to use digital filters when more than about 100 gates are to be processed.

II. DIGITAL PROCESSING METHODS.

Two major processing schemes have been or are being implemented using digital techniques. The first scheme synthesizes high-pass filters by using the digital equivalent of delay lines. To produce high-pass filters these digital filters incorporate feed forward connections to produce zeros and feed-back to produce poles. The second scheme makes use of the discrete Fourier transform (DFT) to digitally synthesize the equivalent of a bank of identical equally spaced filters.

The DFT provides one more dimension to radar target filtering. The others are range, azimuth, and elevation. This extra dimension allows one to eliminate (by blanking) large fixed targets (discretes) of known location. It allows for the reduction or elimination of moving clutter such as rain which is at a different velocity than ground clutter so is not eliminated by the digital high-pass filter approach. Finally, this separate filtering dimension allows a thresholding scheme to be employed wherein the clutter spectrum can be estimated from the DFT outputs. This estimate would be based on returns from the range cell being examined for a target and would not rely on the magnitude of the clutter from other range cells.

The DFT coherently integrates the signal over the time it takes to collect all the samples. The digital high-pass filter must be satisfied with non-coherent integration over this same period resulting in some signal-to-noise or signal-to-clutter advantage for the DFT when the target energy falls principally in one output of the DFT.

In the case of airborne radar, the use of the DFT obviates the need for changing the cut-off frequency of the high-pass filter as the antenna azimuth is changed, or of changing the local oscillator frequency to center the main-beam clutter return in the center of the high-pass filter. Instead, the DFT outputs containing main-beam clutter are either ignored or desensitized as far as thresholding is concerned. This technique is particularly applicable to tactical radars which are interested in close-in targets as well as distant ones. Because of the differences in depression angle (and consequently main-beam doppler frequency) the local oscillator cannot be adjusted simply when one wants to examine all target ranges.

Because the DFT gives a radial velocity sorting using two prf's, a partial sorting of ground targets from airborne targets can be effected. Of course a better method would be to observe motion on a scan-to-scan basis to sort according to speed rather than velocity, but this may require a larger computer capacity to handle all the targets.

The digital high-pass filter is a natural substitute for the analog delay lines previously used. A digital shift register, usually using MOS memory elements, replaces the analog delay lines. The feed-forward and feed-back coefficients are usually chosen to be a positive or negative integral power of two or else the sum of two such coefficients. These are then implemented with simple shifting and adding operations which are hard wired.

---

*This work was sponsored by the Department of the Air Force.

The DFT, on the other hand, uses a large core memory which, at the present time, is cheaper per bit than the MOS memory. It generally uses one or several array multipliers to sequentially process batches of data from various range cells using a fast Fourier transform (FFT) algorithm. Since the designs of the two types of processors are entirely different, one cannot make a general statement as to which is cheaper. In many cases, considering the advantages listed above, it is worth some added expense to employ the DFT approach. Where a large amount of memory is involved the DFT approach may be less expensive. Both approaches involve about the same complexity.

III SIGNAL PROCESSING REQUIREMENTS.

Let M be the number of required range gates, f_d the highest doppler frequency to be measured and Δf the doppler resolution. Then, the sampling rate f_s must be at least twice f_d for unambiguous determination of velocity. If it is not possible to achieve this sufficiently high value of f_s, other techniques (not discussed in this paper) must be used to resolve doppler ambiguity. In performing spectrum analysis, the total number of frequency points is then $2f_d/\Delta f$. The total memory requirements are $2Mf_d/\Delta f$ registers (or twice that number if quadrature detection is assumed). There will be $M(\Delta f)$ spectrum analyses performed every second.

Figure 1 shows a range-gated pulse doppler block diagram. All blocks right of the dashed line represent digital hardware. Typically, the operation of this hardware would be equivalent to many banks of doppler filters, each bank receiving information from a specific region of space, as indicated in Figure 2. The plan of this paper is, first, to relate the signal processing requirements to the range and doppler specifications. Next, some theoretical questions are treated which arise when the fast Fourier transform is used to perform the doppler filtering. Finally, a block diagram of general purpose signal processing hardware will be proposed for performing range-doppler filtering for a variety of differing requirements. If such a proposal were to prove feasible, it would ultimately be possible to buy (rather than spend several years building) the requisite digital processing hardware for a given range-doppler radar.

Digital processing techniques need not be restricted to spectrum analysis to perform the doppler filtering. Instead, digital filters may be used. The trade-off depends primarily on the number of doppler filters desired per range gate; a large number implies that spectrum analysis is more efficient. In this paper, we restrict ourselves to the spectrum analysis approach, although the general purpose hardware described in the latter part of this paper could be programmed to perform either recursive or non-recursive doppler filtering.

Referring to Figure 1, we assume that the M range gates chosen by the operator are one of many possible sets that he could have chosen. Thus, the data rate from the A/D converter is higher than the data rate into the main memory. The preprocessor and high speed buffer is needed to scale down the data rate and, if necessary, to perform predetection integration and/or waveform decoding.

The type of main memory used depends mainly on the required storage size. If this size should be less than .5 to 1 megabits of memory, it would be most economical to use a commercial general purpose computer which could then be used as well for overall systems control, decision making and display. If the memory size is 1-10 megabits, a so called "window screen" memory (2 wires per core, rather than 3 or 4) may be used but beyond this a drum or a disc memory may prove most economical. Such a memory (the circulating type) is feasible because of the repetitive nature of the algorithms.

At this point, it seems logical to estimate speed and memory requirements of the signal processing hardware. Since such results depend greatly on the properties of the fast Fourier transform algorithm, we will divert our path to describe this algorithm, explain the theoretical filtering effects of the discrete Fourier transform, and then return to the hardware considerations.

IV. THE FAST FOURIER TRANSFORM AS A BANK OF SIGNAL DETECTORS.

The discrete Fourier transform (DFT) of a sequence of numbers x(n) is

$$X(k) = \sum_{n=0}^{N-1} x(n) W^{-nk}$$

where $W = \exp(-j2\pi/N)$. In general x(n) is a complex number To compute N values of X(k) requires N^2 complex multiplications and additions.

The fast Fourier transform (FFT) refers to a set of algorithms for more rapid computation of the DFT. The most well known is the so called radix 2 algorithm which computes the FFT by repetitive computation of a "butterfly", defined as:

$$u(m+1) = u(m) + W^{\ell} v(m)$$

$$v(m+1) = u(m) - W^{\ell} v(m)$$

where u(m) and v(m) are complex numbers and ℓ is an integer which varies with m. It is necessary to perform $N/2 \log_2 N$ of these butterflies to compute the FFT, obtaining all N values of X(k).

In this section we will discuss the relations between the DFT, the filter banks of Figure 2 and signal detection. The detection problem is illustrated in Figure 3, which shows the spectrum of a moving target in a large clutter background. The "correct" filtering procedure is pre-whitening of the clutter spectrum followed by a filter matched to the resultant pre-whitened target spectrum, followed by detection (see Figure 4) and integration for as long a time as the target remains in the range gate.

It does not appear to be possible, with a DFT, to obtain the result indicated in Figure 4 if the input samples are contiguous. It is possible, however, to imagine the input appearing in bursts, as shown in Figure 5a. With this model we can now show how to perform matched filtering via the DFT. For simplicity we have set the on and off times and the impulse response duration of the filter equal. (If the filter is recursive, with an infinitely long impulse response, it is immediately apparent that the filter bank cannot be realized via DFT.) Now, the product of the DFT's of the augmented signals shown in Figure 6 is indeed the DFT of y (n) of Figures 4 and 5. Thus, the set of Y (k) of Figure 6 is the DFT of y (n) and Parseval's theorm yields,

$$\sum_{n=0}^{N-1} \left| y(n) \right|^2 \quad , \quad \sum_{k=0}^{N-1} \left| Y(k) \right|^2$$

In practice, a given velocity filter covers a range of frequencies narrow compared to the total DFT spectral range. Thus, H (k) would be zero for most values of k. For a given doppler filter with H (k) the (augmented) DFT of h (n), the computation corresponding to q (L) in Figure 4 is.

$$q(L) = \sum_{K} \left| H(k) \;\; X(k) \right|^2$$

where K refers to the range of non-zero H (k)'s.

It is usually true that the response to a burst, as shown in Figure 5, is more useful after transients caused by switching the signal on and off have died down. For this reason, it may be desirable to weight the data. In effect, this weighting modifies the matched filter of Figure 4 so that the transient portion of q (n) (beginning and end) have the least effect on the subsequent measurements. Two ways of weighting are; (a) the augmented x (n) is multiplied by the weighting function w (n) prior to DFT or (b) the set of X (k) is convolved with W (k) (the DFT of w (n)).

V. GENERAL PURPOSE DIGITAL RADAR SIGNAL PROCESSOR.

The parameters for an air traffic control radar, a weather measurement radar and a ground mapping radar are very different and each system will require unique operations in addition to the common function of range-gated pulse doppler. For this reason, it is useful to develop a computer architecture which can be applied to a variety of radar signal processing systems and which is of sufficient generality to be adapted to the peculiarities of individual radars. The remainder of this paper, represents a preliminary attempt to develop such an architecture. Figure 1 shows the overall radar system for which this architecture is to be developed. In this paper we will not discuss signal design problems (such as chirp waveforms, Barker codes, etc.) but will assume that the preprocessor receives in-phase and quadrature signals which are sampled for inclusion into the range bins chosen by the processor program. Thus, samples for successive range bins can be made to enter the main memory at a uniform rate and, indeed, can be sent to the processor via the memory at the same uniform rate. In what follows, we restrict our discussion to the flow of data into and out of the main memory and into and out of the processor. A timing picture which fulfills the real time requirements is shown in Figure 7.

During the "Enter" phase,

(a) The pre-processor feeds samples from contiguous range gates into the main memory.
(b) The main memory feeds successive samples from a given range gate in the main memory to set up the processor data base.
(c) The main memory also sends results of past processing on the same range gate to the processor to be used for post-detection integration and thresholding.

During the "Do" phase the processor performs filtering and detection on a single range gate. These operations may include,

(a) Weighting the data.
(b) Discrete Fourier transform.
(c) Recursive or non-recursive digital filtering.
(d) Matched filtering via the FFT.
(e) Post-detection integration
(f) Centroid measurements on the FFT to estimate frequency.
(g) Thresholding for yes/no decisions on targets.

The return phase consists of the return of all necessary post-detection, threshold and display information for storage by the main memory until the next turn of the just processed range gate information.

In order to perform these tasks in real time, the 3 phases must be overlapped as shown in Figure 7. In the present concept, control resides exclusively in the processor and it is only the processor which physically consists of high speed circuits capable of both processing and controlling the addressing of the main memory. Before describing the processor-main memory complex in more detail, let us describe an addressing algorithm which allows for the unimpeded flow of contiguous range data from the pre-processor to the main memory and (simultaneously) the flow of data from a single range gate in the main memory to the processor for the DFT. This can be accomplished by an addressing scheme illustrated in Figure 8 for the case of a 5 point DFT and 7 range gates. After the memory is filled by storing incoming samples sequentially (column 1) the address is then incremented by 7 (modulo 34) in column 2. For each new address a main memory word is first transferred to the processor and then that vacant location is filled by the incoming datum from the pre-processor. If M is the number of range gates and N the length of a DFT, then the addressing sequence obeys the rule,

$$A_n(k) = A_n(k-1) + M^n$$

Where both the addition and the term M^n are understood to be taken module (MN-1) and n refers to the nth iteration through the entire memory (a single column in Figure 8).

VI. MORE DETAILS OF THE PROCESSOR AND MAIN MEMORY.

Figure 9 shows a structural diagram of the processor and intermediate memory. Since, in general, this memory will be large and does not have unduly high speed requirements, we can imagine it to be a core memory. The memory and arithmetic associated with the processor we imagine to consist of the fastest commercially available circuits. Since processing speed is a prime consideration, the computer has certain structural features for greater speed. These are:

(a) Separate program and data memories so that instruction fetch and data fetch take place in parallel.
(b) Overlap of data memory and arithmetic operations.
(c) An arithmetic element especially designed for high speed signal processing, in particular the FFT.
(d) Double length words so that two operations can be performed in parallel.
(e) Double instructions performed simultaneously.

The details of the large FFT array are shown in Figure 10. Maximum speed and minimum control of the butterfly algorithm is obtained at the cost of added hardware by making the array box purely combinational. Using emitter coupled logic gates and similar full adder integrated circuit packages results in an array speed of about 200 nsec. Processor memory cycle time is about 100 nsec. Thus, the timing diagram of Figure 11 where the array propagation time is taken to be twice the memory time, shows how a butterfly is accomplished in the time to perform two reads and two writes. For such a memory and for a 4 multiplier array, the structure does an FFT butterfly in minimum time (about 400 nsecs.)

When the array is not used to do a butterfly, it can be controlled to perform two simultaneous multiplications. This means that the weighting operation is very efficient, as is the matched filtering. The preliminary order code we have devised (Figure 12) allows completely general purpose programming. The F register, which communicates between program and data memory is also used to address the intermediate memory and the input-output instruction initiates the intermediate memory read or write pulse.

VII. PERFORMANCE OF SIGNAL PROCESSOR.

The time to do a butterfly is the basic figure of merit of the radar signal processor. This time can be translated into the maximum allowable number of range gates which can be processed in real time, given the sampling rate and the size of the FFT. For simplicity, let us assume that if the processing time of a butterfly is 400 nsec., then the expanded time due to other operations for a typical radar is 1 microsecond. Using this number, Table I shows how many FFT's (in round numbers) per second can be processed for various FFT sizes.

FFT Length	FFT's Per Second
16	31,000
32	12,000
64	5,200
128	2,200
256	1,000

We have thus demonstrated that it is possible to build digital signal processing equipment for range-gated pulse doppler radars with capability of measuring velocity or detecting moving targets in hundreds or thousands of gates. Such a capability is not economically feasible using analog signal processing. Furthermore, the structure we have defined, since it is general purpose, could be used with slight changes for different radar systems.

REFERENCES

1. Linder, R. A., Kutz, G. H., "Digital Moving Target Indicators" Aerospace and Electronics Systems Technical Convention 1967 Record, pp. 374-385.

2. Hall, E. L., Lynch, D. D., Young, R. E., "A Digital Modified Discrete Fourier Transform Doppler Radar Processor", Electronic and Aerospace Convention Record (1968), pp. 150-159.

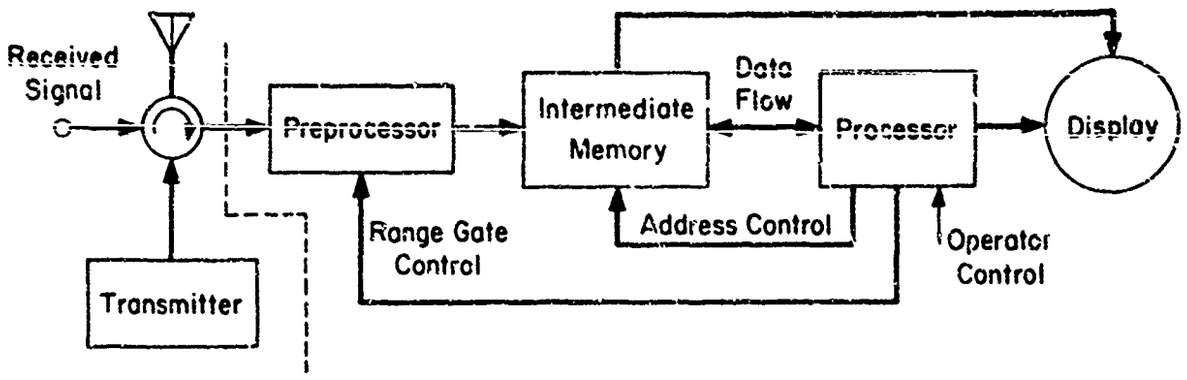


Figure 1. Block Diagram of a Range-Gated Pulse Doppler Radar

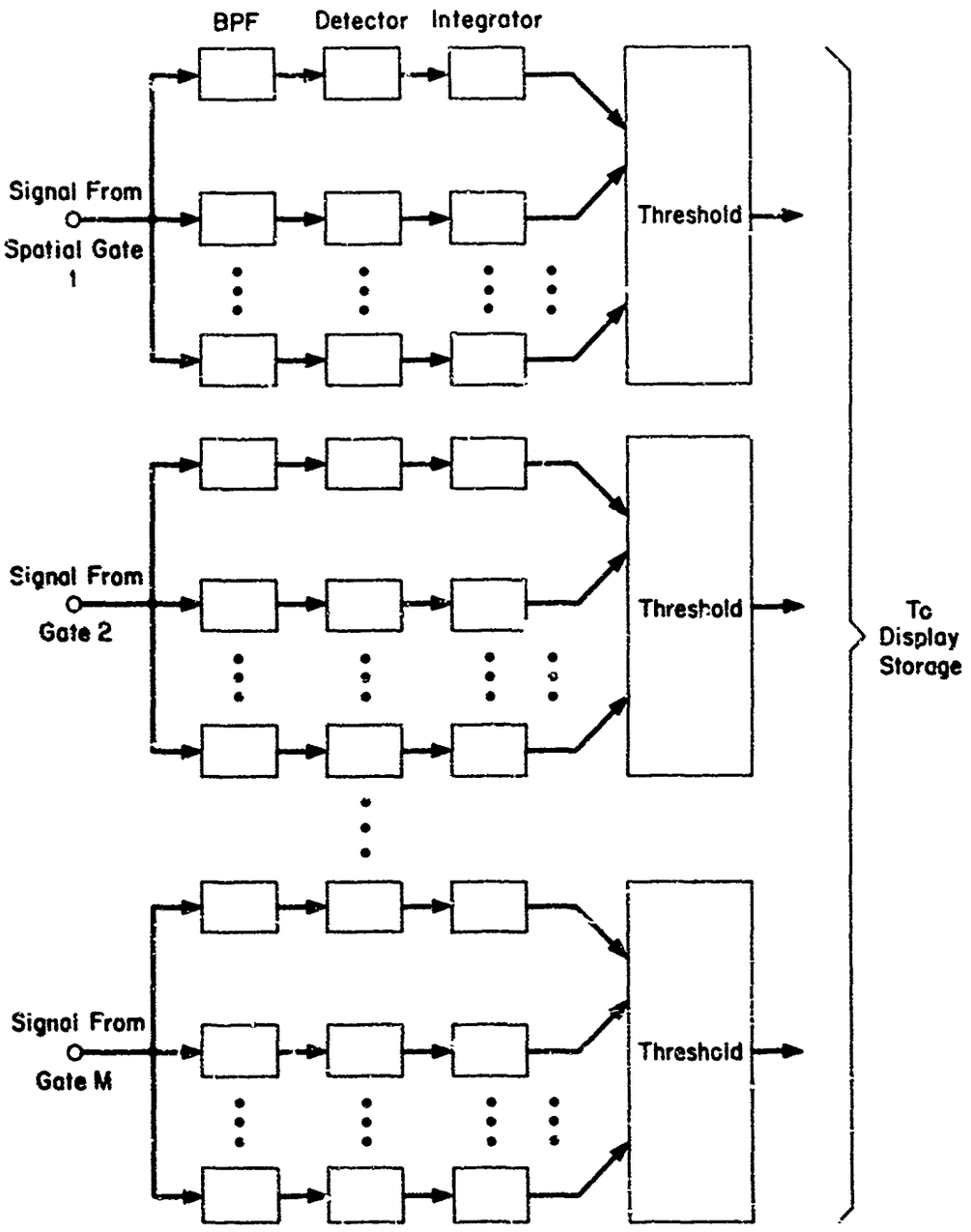


Figure 2. n Banks of Doppler Filters, Each Bank Processing the Gated Signal from a Specific Region of Space as Determined by the Radar Timing

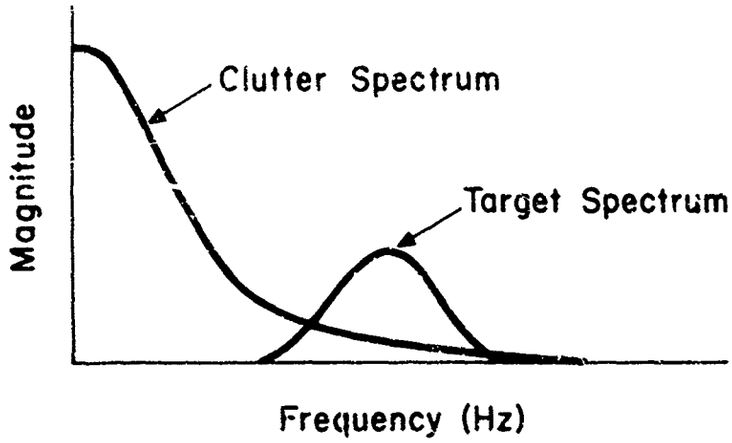


Figure 3. Spectrum of a Moving Target in a Large Clutter Background

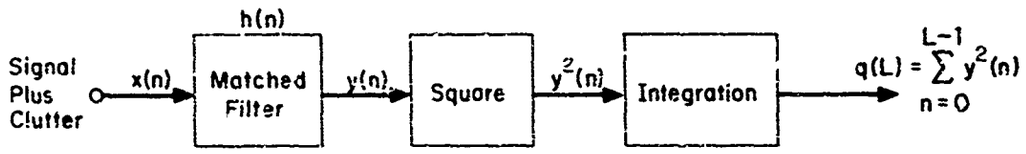


Figure 4. Matched Filter Detection of a Target in Clutter

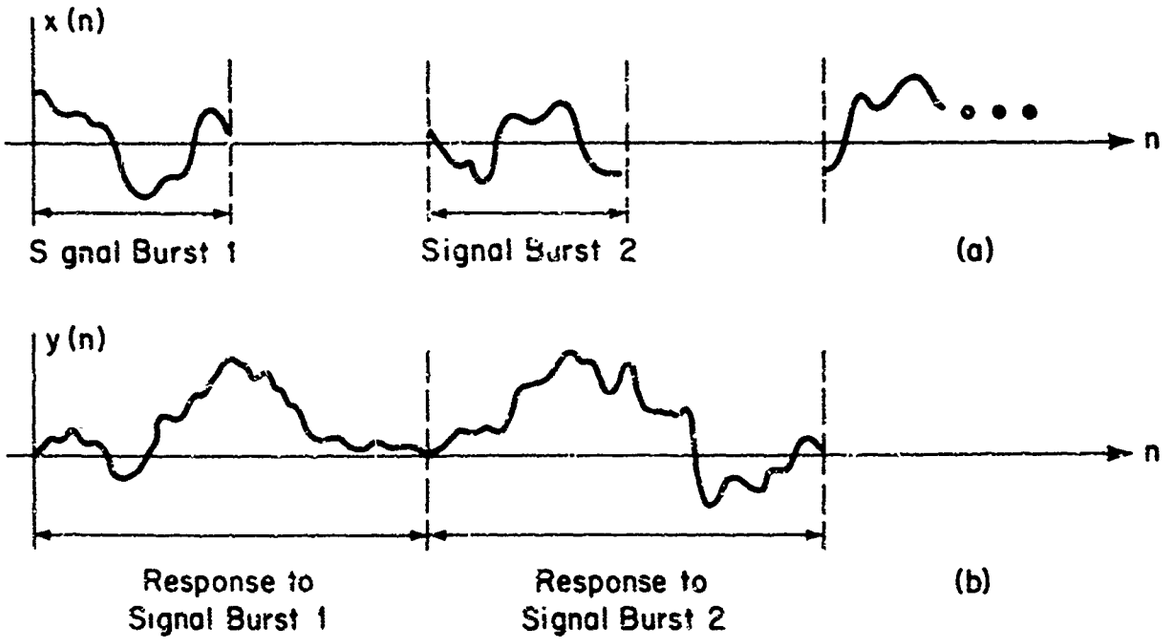


Figure 5. Input and Output Signals of Matched Filters

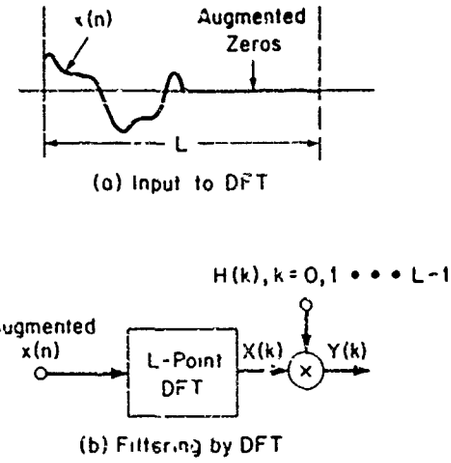


Figure 6. Fourier Transform which is Equivalent to Matched Filter of Fig. 4 when Excited by Signal Bursts

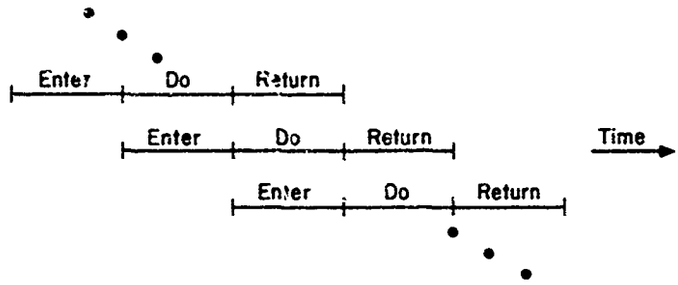


Figure 7. Overall Timing Diagram of the Radar Signal Processing

n	1	2	3	4	5	6
0	0	0	0	0	0	0
1	7	15	3	21	11	
2	14	30	6	8	22	
3	21	11	9	29	33	
4	28	26	12	16	10	
5	1	7	15	3	21	
6	8	22	18	24	32	
7	15	3	21	11	9	
8	22	18	24	32	20	
9	29	33	27	19	31	
10	2	14	30	6	8	
11	9	29	33	27	19	
12	16	10	2	14	30	
13	23	25	5	1	7	
14	30	6	8	27	18	
15	3	21	11	9	29	
16	10	2	14	30	6	
17	17	17	17	17	17	
18	24	32	20	4	28	
19	31	13	23	25	5	
20	4	28	26	12	16	
21	11	9	29	33	27	
22	18	24	32	20	4	
23	25	5	1	7	15	
24	32	20	4	28	26	
25	5	1	7	15	3	
26	12	16	10	2	18	
27	19	31	13	23	25	
28	26	12	16	10	2	
29	33	27	19	31	13	
30	6	8	22	18	24	
31	13	23	25	5	1	
32	20	4	28	26	12	
33	27	19	31	13	23	
34	34	31	34	34	34	

Figure 8. Example Showing Address Sequence for Intermediate Memory for a 5-Point FFT and 7 Range Gates

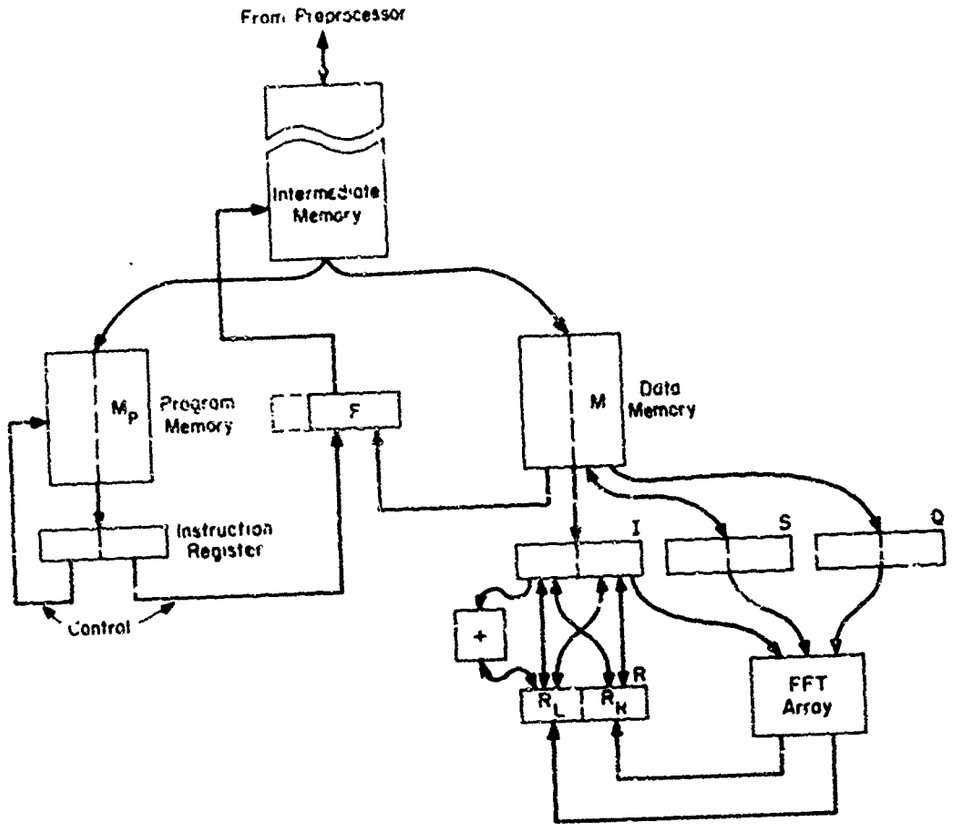


Figure 9. Structure of Processor and Intermediate Memory

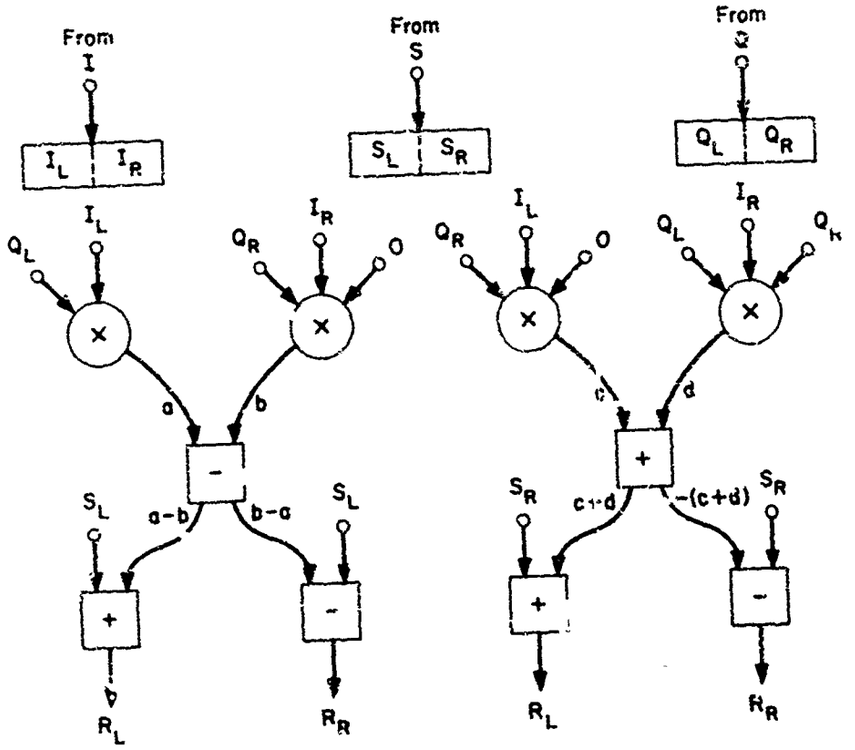


Figure 10. FFT Array for Performing Butterfly

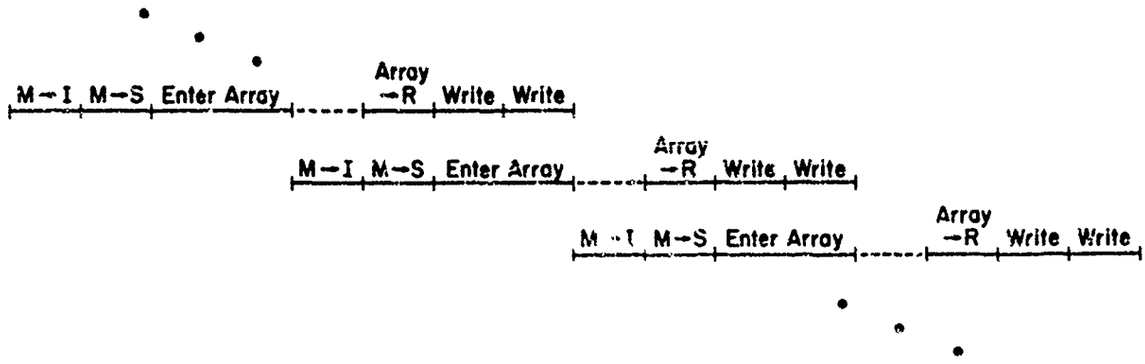
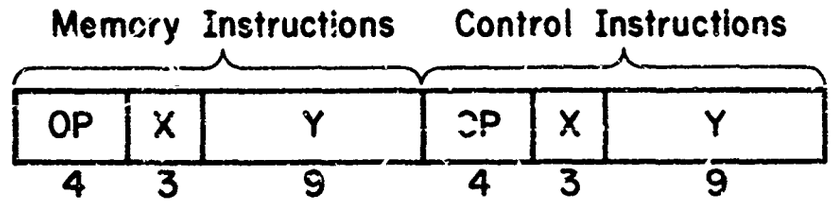


Figure 11. Timing of Executed Instructions Needed to Perform an Efficient Butterfly



Memory Instructions

(1) M → F
(2) F → M
(3) M → I
(4) M → Q
(5) M → S
(6) R → M
(7) E → M
(8) M → E
(9) Arithmetic Group

Control Instructions

(1) Y + X → X
(2) X - Y → X
(3) Y - X → X
(4) Y → X
(5) Jumps
(6) Input-Output Control
(7) F-Group Instructions

Figure 12. Tentative Partial Instruction List for Structure of Fig. 10

DATA PROCESSING CONSIDERATIONS IN ADVANCED PHASED ARRAY RADAR SYSTEMS

by

G. Cokas and J. H. Parsons

Hughes Aircraft Company, Ground Systems Group
P.O. Box 3310, Fullerton, California 92634, USA

DATA PROCESSING CONSIDERATIONS IN ADVANCED PHASED ARRAY RADAR SYSTEMS

S. Cohen
J. H. Parsons
HUGHES AIRCRAFT COMPANY, GROUND SYSTEMS GROUP
P. O. Box 3310, Fullerton, California 92634

SUMMARY

This paper has been prepared to present a broad overview of data processing considerations for advanced phased array radars. Data representing radar processing experience of the last decade has been collected from a number of organizations engaged in radar design and computer programming. This data has been correlated with industrial experience obtained from other data processing applications. The necessary growth of radar data processing is traced together with the consequences of this growth. Examples of particular solutions to data processing problems are cited to demonstrate how technology has changed. Residual problems are evaluated and some possible solutions are discussed.

1. INTRODUCTION

During the past decade, phased-array radars have increased in complexity because of the increased requirements of radar applications in defense and air traffic control. These phased-array radars have developed from simple fan beam radars to high-resolution pencil-beam radars. The operational modes of these radars have progressed from simple search to automatic track and automatic acquisition in adverse environments. Today, radars are programmed in an adaptive manner so that their power and measurement capability are used efficiently to cope with varying tracking, environmental, and load situations. In addition, extensive signal and data processing is now used to extract the information from the radar returns. All of this added capability has increased the amount of digital data-processing equipment built into the radar, the size and speed of the tracking and control computers, and the number of instructions in the operational program.

In this presentation we will trace the growth of the digital equipment, the computers, and the computer program and show the effect this growth has had on the cost and complexity of radar processing systems. We will also show how technical advances in digital circuits, integrated circuits, large scale integration, and computer design have made processing techniques available that were previously either not technically possible or economically feasible. Older processing implementations based on clever, special-purpose digital devices are being discarded for more complex but more cost-effective designs. The results of advances in the characteristics of logical elements, circuits and processing techniques have not only extended radar capability but have made possible these extensions without an accompanying increase in hardware cost. During the time that the data processing requirements have increased, the processing capability available per dollar has increased at a greater rate. We will see that while hardware costs have remained stable, the rising costs of designing, writing, testing and documenting the operational computer programs has now become significant. We will indicate how future advances in such items as compilers for operational programs to be performed on multiple processors, and preprogrammed standard operational program modules should tend to alleviate this problem.

2. INCREASES IN DIGITAL PROCESSING WITHIN THE RADAR

The requirements for digital processing equipment designed into the radar has increased by a factor greater than 100 in the last decade (see Figure 1) due to:

- Increased phase control computations within the radar.
- Increased signal processing requirements.
- Increased number of digital implementations of signal processing formerly performed by analog circuits.
- Increased system bandwidth (range resolution).
- Increased buffering of the radar returns to permit multiple target handling capability.
- Close control of the radar on a beam by beam basis.

Figure 2 shows how this increase in processing requirements has resulted in a corresponding increase in the number of digital elements (flip-flops, gates, or memory cells) in the radar in the last decade. This amount of digital equipment is contrasted with the amount of similar equipment in "small" computers. The particular radars used to establish data points in this study were of advanced design and capability; hence they represent extremes in complexity. While this increase in complexity has continued through the years, the cost of logical elements has decreased by a factor of one thousand. (It should be noted that the fastest flip-flops are not available at the lowest price.) The result has been a, more or less, constant cost of digital hardware in the radar despite the increases in processing and control capability. The cost and speed of available logical elements in the last decade are contrasted in Figure 3.[1]

To take advantage of the cost savings permitted by medium and large-scale integration, it was necessary to abandon earlier design concepts that stressed clever design for reducing equipment costs through minimizing the number of logical elements used to perform specific operations. For example, some time ago, to compute the phase shifts for each transmitting element using the incremental phase shifts supplied by the

control computer, a single adder was used repeatedly, under the control of a wired program. The phase-shift words were held in a core memory prior to gating to the individual phase shifter controls. The individual phase shift words were computed in approximately 1 millisecond, limiting the PRF of the radar. Later, the phase shift computations were performed in a number of adders and the phase shift words were directly gated to the individual phase shifter controls. No core memory was used. This implementation was not only faster but more economical to build. Future phase shift computers may be implemented with an adder at each phase shift element and the phase shift addendeas will be according to the rows and columns of the phase shift elements. Counters (performing the function of adding) at each element controlled by row and column timing pulses is another possible future implementation. The transition from the earlier solutions to the future implementations is dictated solely by the economics of integrated circuits and medium-scale integration. In the future, each adder and control driver would be implemented as a special purpose integrated circuit for each transmit element. Even though the number of logical elements to do the specific task increased, the cost of design and production is drastically reduced.

Signal processing for pulse compression, clutter processing, and other data-extraction operations have previously been performed by analog circuitry. The advantages of greater processing capability (such as non-linear processing and digital coding) and higher reliability (no critical adjustments) inherent with digital processing were realized long before the economics of logical elements made the use of digital implementations competitive. Logic elements are now available with transition times in the order of a few nanoseconds; single-chip shift registers and other multiple-circuit units are available for signal processing digital implementations. Ultimately digital implementations may be replaced by optical and acoustical analog processing methods which are now under development.

The return-data buffer (radar-computer interface equipment) is becoming increasingly important as the number of simultaneous tracks increases. The buffer not only prevents overloading the computer interface during peak data-return rate loads but also serves as the data distribution device to gate the return data to the appropriate processor or memory of a multiprocessor control and tracking computer complex. Associative processing using non-destructive ferrite memory elements was an early approach for this application. Correlation processor memories using flip-flops have been developed for this application. Write speeds have increased by a factor of a thousand and the correlation processing time have reduced by a factor of one hundred in these devices when integrated circuit logic devices were used rather than ferrite devices. This implementation also reduced the power requirement and eliminated associated heating problem. For this type of application, the use of flip-flops as memory elements have become economically competitive to the use of non-destructive ferrite memory elements or to thin film memories.

3. INCREASES IN THE COMPUTER REQUIREMENTS

The number of instructions in the operational program that controls the radar and processes the radar returns has increased by a factor of one hundred in the last decade. In addition, the number of instructions that must be performed for each radar pulse transmission and for each return has increased (in one case) from about 250 instructions to 8,000 instructions. The pulse-repetition frequencies of the radars are dictated only by range requirements and, hence, have not changed through the years. The result is that computers must perform a larger number of instructions per second. The increased number of instructions per radar cycle is due to:

- The inclusion of programs to provide adaptive control of the radar power, signal coding and processing to match the requirement of the varying operational environment.
- The inclusion of automatic track-correlation that is necessary for automatic acquisition.
- Advanced tracking algorithms that permit more accurate track measurement, smoothing and extrapolation.

In addition, the operational program now includes:

- Built in operational test programs.
- Computer and radar maintenance diagnostics.

The quantitative results of these increases is shown in Figure 4 where the number of instructions per second required by radar systems is contrasted with the available computer speed. The phased-array radars used to establish data points for this comparison were all large, advanced systems and hence represent extremes (as previously stated). Figure 5 shows the number of instructions per second that can be performed by the most advanced computers (years correspond to dates of the introduction of computers and to the dates when required in radar systems). Early radar systems taxed available computers capability; for this reason, special-purpose computers, possessing an instruction repertoire designed for radar operation, were used to meet the real time computational requirements. Since then, ample computer speed has become available and the use of computers not specifically designed for radar operation has been much more cost-effective for satisfying the majority of the data processing requirements for radar control and tracking.

The above discussion does not mean that special-purpose devices are not worthwhile in specific applications. Special-purpose coordinate converters, correlation processors, and display-data generators have all been found to be cost effective. Each application of these devices, used to relieve computer load, must be considered on its own merit.

In Figure 5 the curve depicts the computational speed or capacity available. (References 2 and 3 and many other sources.) The units of the ordinate are millions of instructions performed per second (MIPS).

The computer examples chosen represent the fastest computers introduced within the year designated. Earlier machines are single processor while later machines are parallel processors. The highest-speed machine is a newly completed computer using sixty-four parallel processors. The increase in speed indicated in the figure will probably not continue through more than one more order of magnitude. Physical limitations of length of wire, propagation time, and heat dissipation are beginning to limit speed.

The cost of the computer can be discussed in terms of instructions that can be performed per purchase or rental dollar throughout the life of the machine or system. Figure 6 shows the results (basic data from references 2 and 3 and many other sources) of such an analysis again using large scale computers. The conclusion is obvious; the instructions performed per dollar have increased dramatically over the last decade. This increase is due to:

- o Lower cost logical elements.
- o Faster logical elements.
- o Better computer design.

It should be emphasized that this curve is for the largest machines only. It is a consequence of Grosch's Law* that smaller computers (less computational capacity) cannot be as cost effective as large computers. For comparison, the two squares shown represent examples of these smaller computers recently available. Because of the probable limit of computer speed(4), the trend of stable computer cost in spite of drastically increased requirements will not continue unless (as it is hoped and believed) the radar computational requirements will also naturally reach a limit.

4. COMPUTER PROGRAMMING

We have already seen that the number of instructions that must be performed for each radar transmission has increased dramatically. If we look at the length of the computer program that must be written for the radar computer, we would naturally expect a similar increase in size. We find, however, that the number of instructions in the total operating program has increased even more. The reason for this excess is that the radars operate in a number of alternate modes selectable in an adaptive manner according to environment; each one of these modes requires a separate program with a large number of instructions. In addition, multimodal tracking algorithms are used that are selectable according to accuracy and smoothing desired on individual tracks. Further, the number of test, maintenance, and fault isolation programs included with the operational programs have increased because they are needed to cope with the increased radar and computer complexity. The results of these increases are shown in Figure 7.

In the previously cited increases in demands on the hardware elements, logical element and computer technology have more than compensated resulting in no appreciable cost increases in system production costs. In programming there are no comparable compensating factors. In fact, the cost of programming per instruction (including analysis, coding, key punch, assembly and documentation) has increased by almost a factor of ten in the last decade as shown in Figure 8. The reason for this increase is not all in the increased salaries of the programmers. The lower line represents the early experience in cost of programming per instruction. Note that the squares, shown almost on this curve represent the cost per instruction (national average) for scientific programming.(5) The upper line represents recent experience in radar operational programming. The slope is the same but fewer instructions are written per month per person due to increased documentation (flow diagrams, operational descriptions, etc.) and increased program complexity required by the user. The circles represent costs that are being experienced by various programming organizations. The spread is more a function of the variation in the definition of what is included in programming rather than programming efficiency or salary variation.

The trends in digital equipment costs, computer costs, and programming costs are contrasted in Figure 9. Simply stated the cost of digital equipment and computers is more or less stable. The cost of programming is becoming the dominant factor in radar control and track processing. In a recent system involving several netted radars the cost of programming was estimated to be about 15 million dollars while the cost of the computer was approximately 3 million dollars. The national experience for general-purpose computer centers indicates a similar trend in that equipment costs are now only twenty-five percent of operating costs. A direct comparison of programming costs to computer costs for radar systems is unfair because the same program is used in all systems of the same types. Thus, if ten identical systems are built, the cost per system for programming is one-tenth the original programming costs, if no site program adaptation is required. But even in this case the programming cost per system has become significant.

A number of techniques are being studied to help reduce the cost of programming. Compilers specially designed to meet the needs of radar data processing are being developed to reduce operational programming difficulty. Such compilers should also simplify the documentation problem. Since every advanced radar performs the same set of basic operation, pre-designed modular subprograms are being considered. These modules would be designed for ease of understanding, simple module interfacing, simple documentation and multisystem adaptability. Clever design of programs that are very efficient because of the exact tailoring to the system would be abandoned. Computer requirements would increase and programming costs would go down. This may be the most cost effective solution to the problem.

---

*Computer power is proportional to the square of the cost.

5. CONCLUSION

At the start of the decade the accent of radar data processing engineering was directed toward hardware. Technology has solved these problems. Today the important problem is how to reduce the cost of programming and to this problem the radar community must now direct its attention.

REFERENCES

1. Various industrial sources.
2. Knight, K., "Changes in Computer Performance", Datamation, Sept. 1966, pp. 40-54.
3. Knight, K., "Evolving Computer Performance 1963-1967", Datamation, Jan. 1968, pp. 31-35.
4. Dinnerstein, L. J., "The CDC-7600 - A Giant in Our Time", Data Processing Magazine, May 1969, pp. 48-59.
5. Brandon, R., as quoted in Electronic News, April 15, 1968, p. 5.
6. Rudenberg, H. Gunther, "LSI - Past Promises and Present Accomplishment", Fall Joint Computer Conference 1969 (Later in Proceedings of Conference).

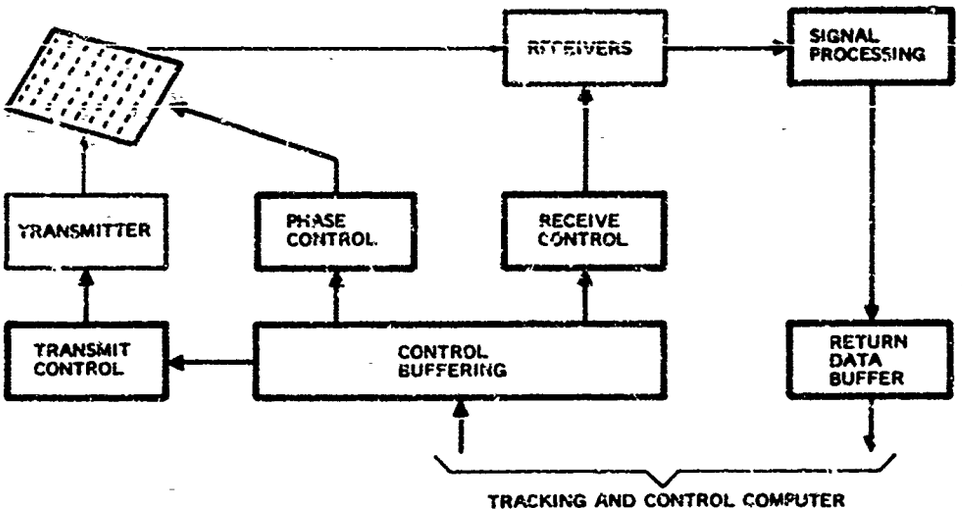


Figure 1 THE AREAS WITHIN RADAR WHERE DIGITAL PROCESSING AND CONTROL EQUIPMENT HAVE INCREASED BECAUSE OF INCREASED SYSTEM CAPABILITY.

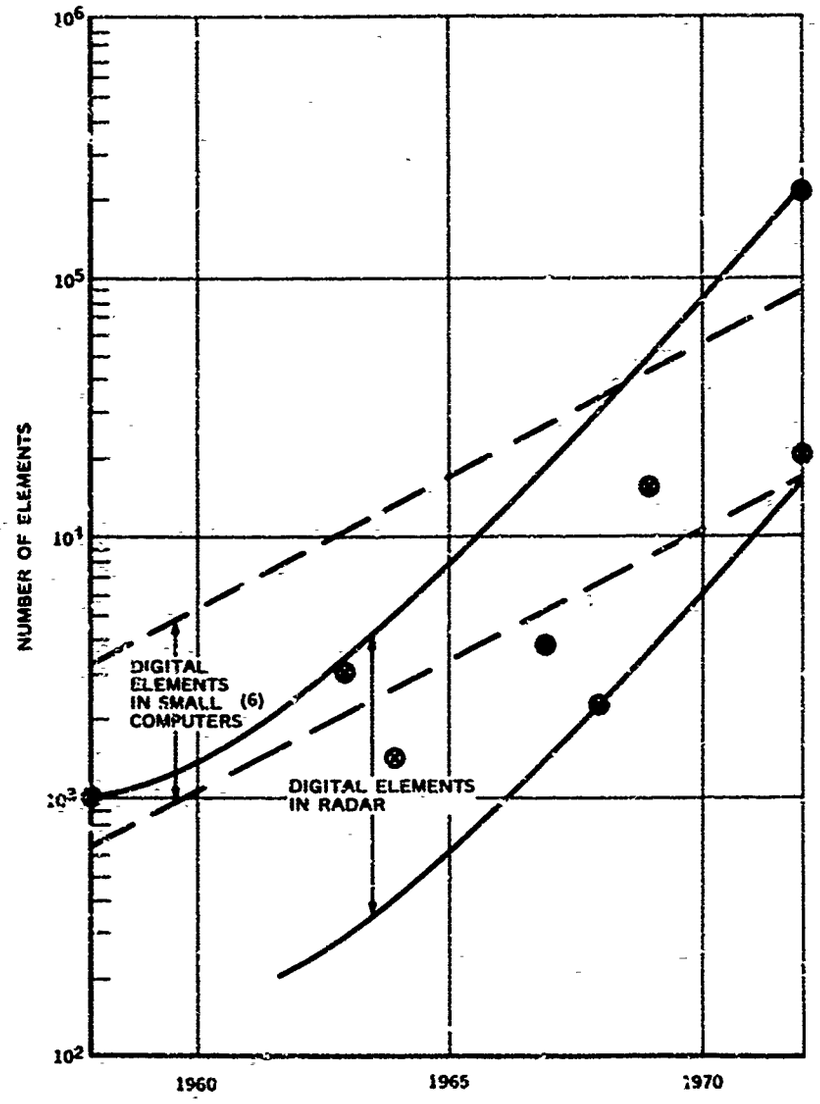


Figure 2 THE NUMBER OF LOGICAL ELEMENTS THAT ARE AN INTEGRAL PART OF THE RADAR. THE PAST DECADE HAS SEEN A FACTOR OF ONE HUNDRED INCREASE IN DIGITAL EQUIPMENT WITHIN RADAR.

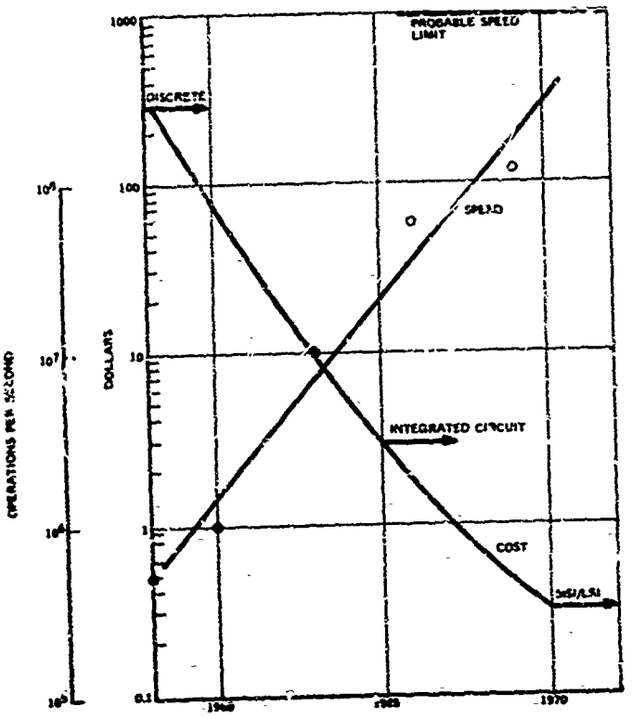


Figure 3 COST AND SPEED OF LOGICAL ELEMENTS (NOT NECESSARILY AVAILABLE IN SAME PACKAGE). OPERATIONS PER SECOND PER DOLLAR HAVE INCREASED BY MANY ORDERS OF MAGNITUDE OVER THE LAST DECADE.

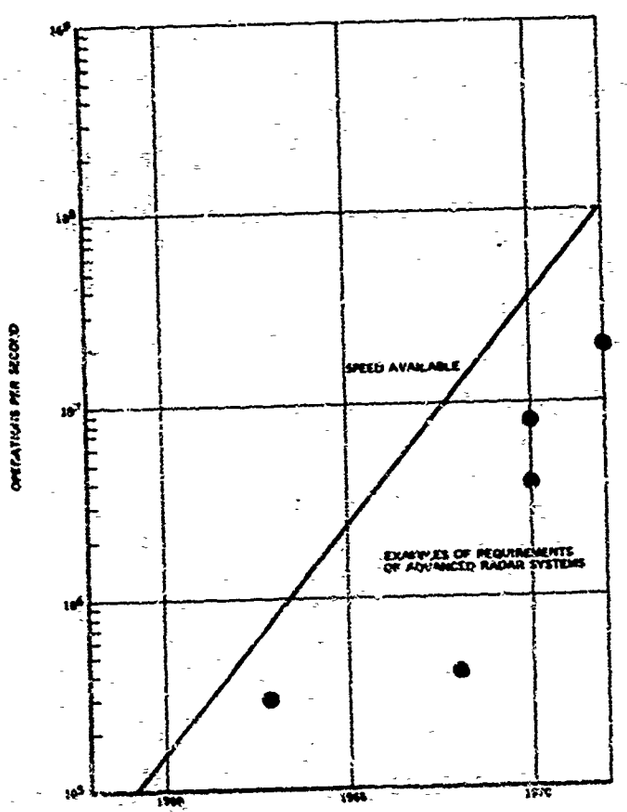


Figure 4 SPEED REQUIRED OF CONTROL AND TRACKING RADAR COMPUTERS COMPARED TO SPEED AVAILABLE IN ADVANCED COMPUTERS. CURRENT AVAILABLE CAPABILITY EXCEEDS REQUIREMENTS.

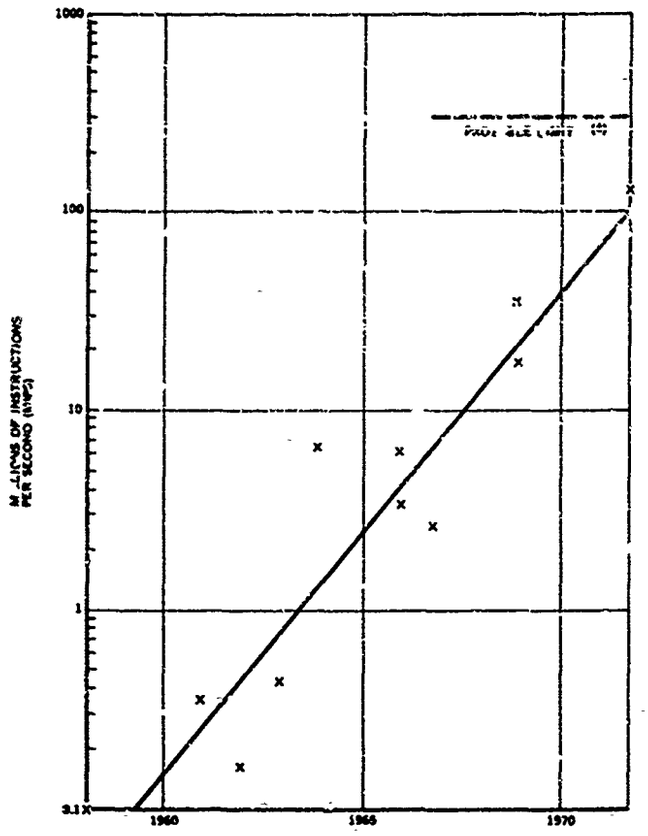


Figure 5 COMPUTER SPEED IN MILLIONS OF INSTRUCTIONS PERFORMED PER SECOND. COMPUTER SPEED HAS INCREASED OVER THE LAST DECADE AS A RESULT OF BETTER COMPONENTS AND BETTER COMPUTER DESIGN.

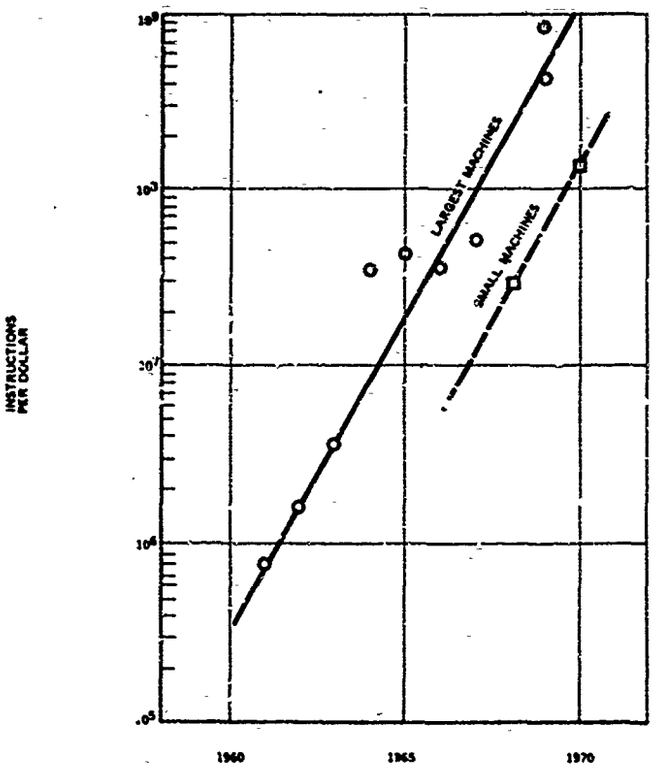


Figure 6 INSTRUCTIONS PERFORMED PER DOLLAR OF PURCHASE OR RENTAL PRICE[2] [3]. COST OF COMPUTATIONS HAVE BEEN REDUCED BY LARGE FACTORS.

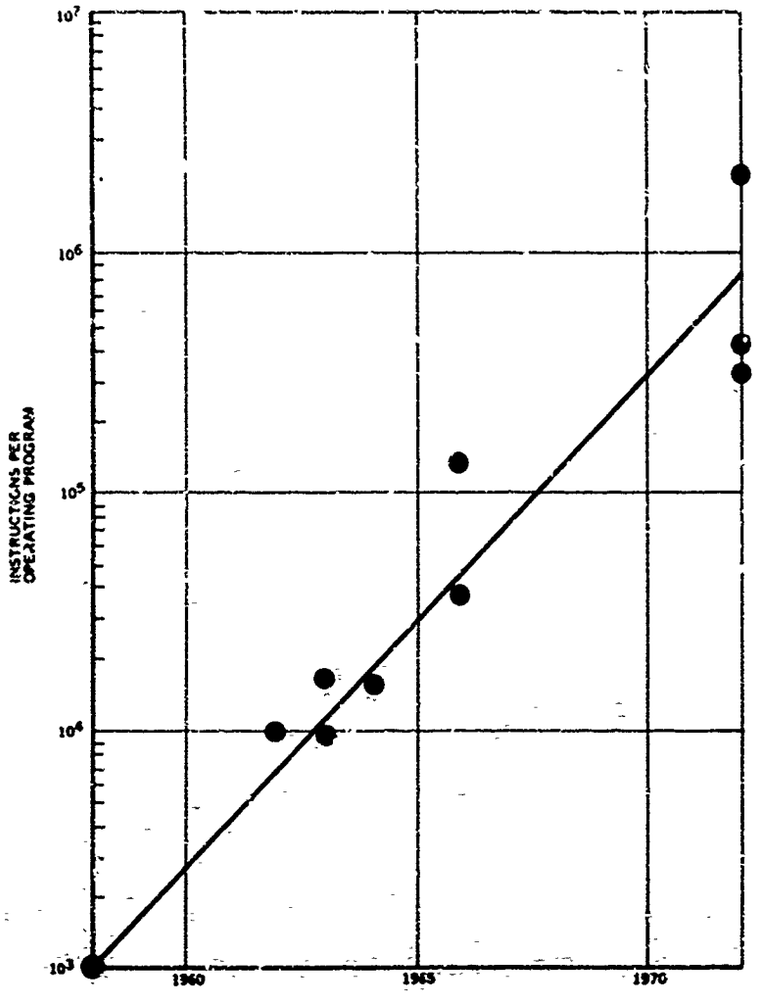


Figure 7 TOTAL INSTRUCTIONS PER OPERATING PROGRAM. THE NUMBER OF INSTRUCTIONS PER OPERATING PROGRAM HAVE INCREASED BY A VERY LARGE FACTOR.

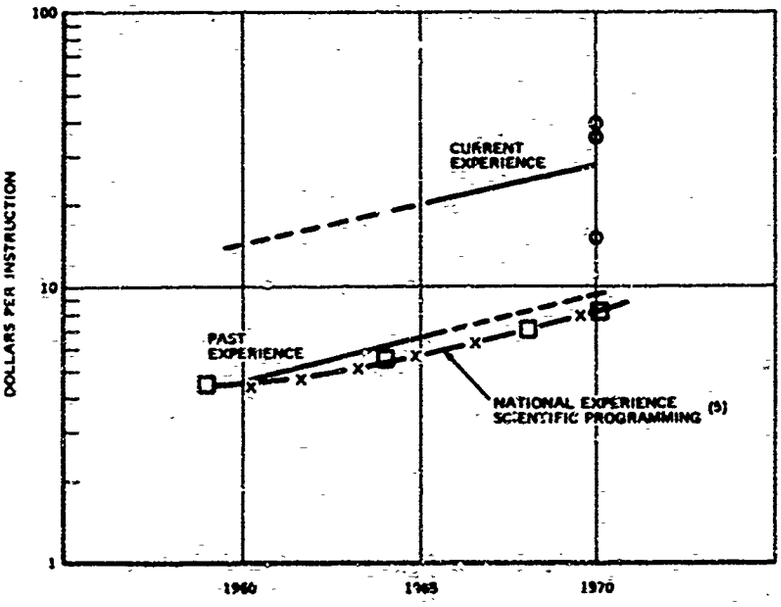


Figure 8 COST OF PROGRAMMING IN TERMS OF DOLLARS PER INSTRUCTION (INCLUDING ALL COSTS). PROGRAMMING COSTS HAVE INCREASED BECAUSE OF COMPLEXITY AND DOCUMENTATION.

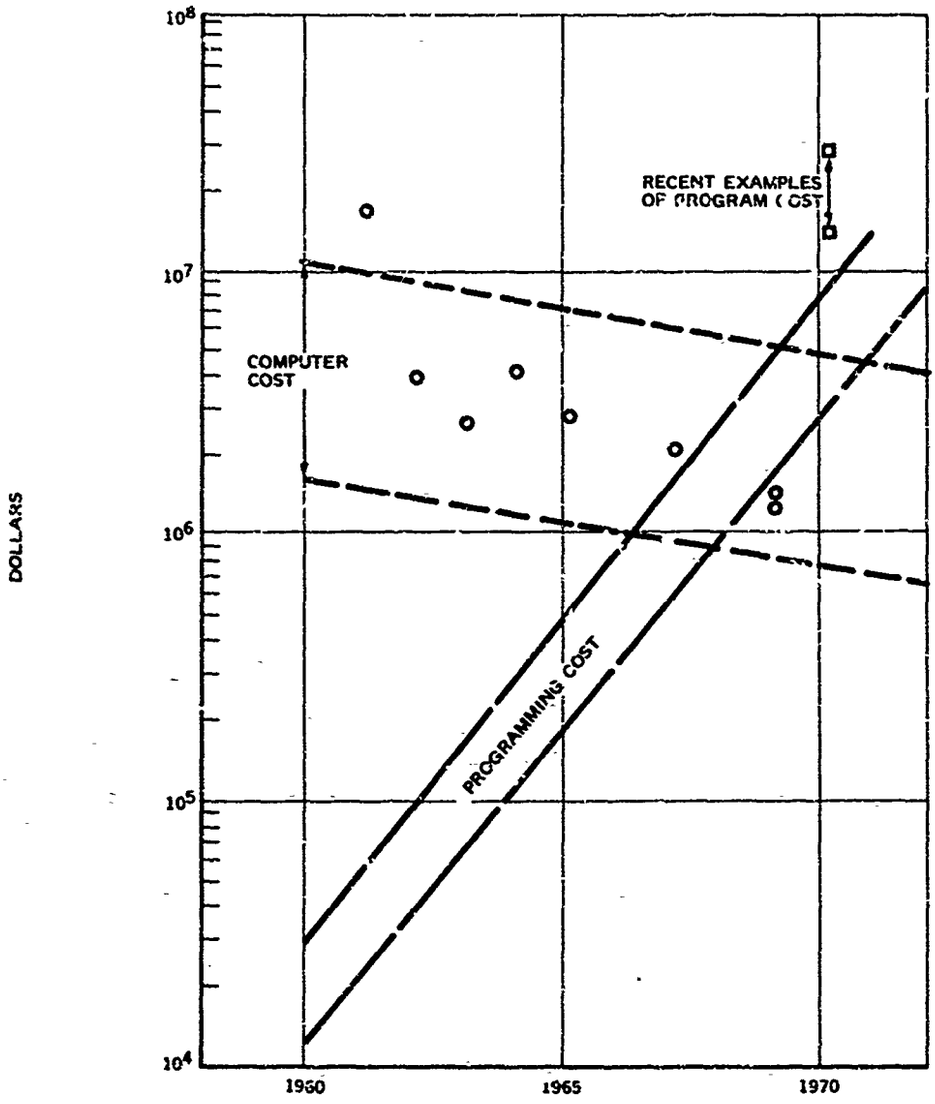


Figure 9 COMPUTER COSTS CONTRASTED WITH PROGRAMMING COST. COMPUTER COSTS ARE UNDER CONTROL. WAYS MUST BE FOUND TO REDUCE PROGRAMMING COSTS.

OPERATIONAL EMPLOYMENT AND DETECTION PERFORMANCE OF ADVANCED COMPUTER-CONTROLLED RADAR SYSTEMS

by

Ir J.A. Bijvoet

SHAPE TECHNICAL CENTRE

Systems Research Division

The Hague, Netherlands

SUMMARY.

To be competitive from a cost/effectiveness standpoint, it is essential that the inherent flexibilities of the computer-controlled Multi-Target Tracking radar system (M.T.T.R.) are exploited to the fullest possible extent.

The purpose of the present paper is to provide an introduction to the study of a suitable operational control of full or partial electronically-steered radars.

"Horizon" and "grazing" search modes are discussed which provide the minimum search time, which leave the least possible freedom for target concealment in clutter and which provide a maximum of time left for use of anti-clutter measures during sub-horizon searching.

Design curves are presented relating the number of obtainable beam positions, in non-adaptive volume search modes, to required transmitter peak power and numbers of hits on target.

Design curves are presented relating volume search time to system range and required volume coverage, and taking into account pulse repetition rate adaptation to beam elevation, off-broadside beam widening and adaptation of the number of hits to beam elevation for providing a constant volumetric detection probability.

A constant spatial detection probability, optimum use of the time and power, and a constant tangential target location accuracy can be achieved by the use of an "Inversely-Proportional" antenna, of which both the azimuth and elevation beamwidths are inversely proportional to range.

The importance of the use of very narrow beams is stressed, for reasons of reducing ground illumination, improving performance against clutter, and improving target tracking performance.

The results of an investigation are presented into optimum phased-array beam separation.

For the evaluation of phased-array detection performance a "Radar Beam Disposition Factor, $L_d^{(2)}$" is introduced. Values of $L_d^{(2)}$ for different radar operating conditions are given.

Design curves are presented enabling selection of the optimum beam separation as a function of the detectability conditions.

A very small transmitter power is shown to be sufficient for the type of system operation described. Interesting system construction aspects may arise from this.

Several of the arguments given apply also to systems having phase-frequency steering or using one-dimensional inertialess steering only.

A limited discussion is included concerning adaptive multi-target tracking procedures.

OPERATIONAL EMPLOYMENT AND DETECTION PERFORMANCE OF ADVANCED COMPUTER-CONTROLLED RADAR SYSTEMS

by

Ir J.A. Bijvoet
SHAPE Technical Centre, The Hague, Netherlands

1. INTRODUCTION.

Various advanced radar systems, which use full or partial electronic beam steering and operate under computer control, have come to be feasible.

For this second generation of radar systems to be competitive from a cost/effective standpoint, however, it is essential that their inherent flexibilities be exploited, as part of the overall systems control, to the fullest possible extent.

The classic, continuously-rotating, search radar spreads its energy indiscriminately throughout space and generates information irrespective of the momentary needs. This not only results in the proportion of useful information being small but involves both wanted information and large quantities of unwanted data being fed to the signal and data processing system, thereby creating enlarged data processing requirements.

The flexibility of the computer-controlled radar can be exploited in respect of many of the detection, data extraction and tracking functions. The combination of the improvement factors, some of which may be small in themselves, will lead to greatly improved performances.

The purpose of the present paper is to provide an introduction to the study of a suitable operational control of full or partial electronically-steered radars.

To dissociate the terminology from any particular system design, such as "phased array" and "frequency-scanned" radar, we have introduced the term "Multi-Target Tracking Radar", or "MTTR", so relating the radar to its principal operational function and at the same time indicating that, in general, selective target tracking only is sufficient.

Finally, the information which follows investigates the possible use of very narrow beams.

2. HORIZON SEARCH.

With air-surveillance radars there is no need to perform a full-volume search. It is sufficient to carry out an "Horizon Search" over a large azimuth sector but with limited elevation coverage, followed by the selective, multi-target tracking of targets detected by the horizon search.

For example, in the case of a radar with a detection range of, say, 200 km, an horizon search with an elevation coverage of 2° is sufficient to detect all targets which enter the radar coverage at heights up to 50,000 ft. (See Figure 1)

Selected targets are tracked individually by selective search of only those small volumes containing the predicted possible next positions of the targets. The number of beam positions needed per data interval for the updating of tracks is small compared to the number of beam positions which would be required to search the full volume, even when a large number of targets has to be tracked. For example, for 200 targets to be tracked, and assuming that, on average, per position update, two beam positions are needed for searching and position extraction, only 400 beam positions will be required.

For the horizon search example shown in Figure 1, with an azimuth sector coverage of 120° and a narrow pencil beam radar with 0.5° beamwidths in azimuth and elevation, 1400 beam positions are needed, giving a total number of beam positions for searching and multi-target tracking of 1800.

The number of beam positions needed to search a full solid angle of 120° in azimuth by 20° in elevation, assuming the same narrow pencil beam antenna, amounts to 9600. Therefore, for this example, the number of beam positions needed for horizon search and multi-target tracking is only 20% of that needed in the case of full-volume search.
A further large reduction in the time spent in each beam position can be achieved as discussed below, and, since the full antenna gain is maintained, there will be a substantial reduction in required transmitter power.

These considerations also apply, to a large extent, to half-electronically steered systems, which employ electronic scanning, e.g., in elevation only.

3. GRAZING SEARCH.

A profitable mode of operation, called "Grazing Search", is one in which the beam elevation follows the horizon contour. A slight additional elevation is applied so that most of the beam remains above terrain. This reduces the magnitude of the ground clutter return.

The small volume left uncovered below the elevated beam only needs a single low-elevation scan or "Terrain Search". The need for MTI (Moving Target Indication) operation is limited in this way to a single, or at most two, azimuth scans.

Horizon contour-following causes a reduction of the total number of beam positions needed since there is no need to transmit in screened directions. In the terrain screening example shown in Figure 2 this reduction is 17%.

With the Grazing Search mode, the smallest number of beam positions requiring MTI operation is obtained. This enables the number of hits for MTI operation to be increased, so providing improved anti-clutter performance.

If, for example, an average of three hits is available per beam position for the example in Figure 2, the system can be operated so as to have two hits on target in non-cluttered directions and about six hits for searches with MTI.

Programming the system control function and storing of data for the Grazing Search is not a complicated matter since the amount of data needed is small. The subsequent clutter-reduced elevation positions can be obtained by initial manual beam control.

4. AVAILABLE NUMBER OF BEAM POSITIONS

The procedure of using the available number of beam positions for deriving the relations given in the following sections conveniently provides the fundamental system parameters.

The number of beam positions obtainable is a function of only the maximum range, the data rate required, and the number of hits on target, n_h, needed for the specified detection probability.

For an unambiguous range equal to R_{max}, the number of beam positions, N_b, obtainable within a data interval or search time, T, is:

$$N_b = \frac{1.5 \times 10^5}{n_h} \times \frac{T(sec)}{R_{max}(km)}$$

Considering different types of system, the maximum radar range and the required data interval can be seen to be approximately related, making the number of beams a function of the number of hits only:

For example, if R_{max} = 300 km and T = 6 seconds, $N_b = \frac{3000}{n_h}$.

The absolute maximum number of beam positions obtainable is 3000 (which would occur in the case of one hit per beam position only).

5. HITS ON TARGET, TRANSMITTER POWER, AND NUMBER OF BEAMS

The required transmitter power can be directly related to the available number of beam positions and the number of hits required on target, as follows:

From the radar equation it follows that the transmitter peak power, P_t, is proportional to D_s, the signal-to-noise ratio required at the receiver; all other radar parameters are assumed to be constant (Ref. 1). For a given target signal fluctuation characteristic and a particular false alarm probability, D_s is a (non-linear) function of the probability of detection and the number of hits, n_h, on target (Ref. 1 and 2). For a given required probability of detection, P_{d_r}, this "detectability relation" is: $D_s = f(n_h)_{P_{d_r}}$.

Therefore, for a linear dependence of maximum range and required data interval:

$N_b = 3000/n_h$ and

$P_t' \propto f(n_h)_{P_{d_r}}$ where P_t' is the relative peak power, i.e. as related to the peak power required for the case of one hit on target.

The dependence of N_b, P_t' and the relative transmitted energy per beam position on n_h is shown in Fig. 3, for P_{d_r} = 80%.

Without taking account of cost/effectiveness considerations it can be expected that the use of about three hits per beam position will provide best integration efficiency.

For n_h = 3, the relation between search time, maximum range and number of beams is

$$N_b = 0.5 \times 10^5 \times \frac{T(sec)}{R_{max}(km)} \qquad \text{(Fig. 4)}$$

For this optimum number of hits and the earlier assumed linear dependence of R_{max} and T, the number of obtainable beam positions is a constant:

$$N_b = 1000.$$

This number of beams can therefore be considered as an approximate system design constant.

If more transmitter power is applied than that corresponding to the best integration efficiency, N_b can be increased (Fig. 5).

For our example, with n_h = 3, R_{max} = 300 km and $\theta_A = \theta_E = 0.5^\circ$ and taking as an example the following set of radar parameters (Ref. 1),

$P_f = 10^{-6}$ m = 1 dB

Fast scintillating signal C_b = 0 dB

Non-coherent integration	L_x = 2 dB
$f_c \approx 5000$ MHz ($\lambda \approx 6$ cm)	$\overline{NF}$ = 4 dB
$G_t = G_r$ = 51 dB	L_r = 2 dB
$L_d^{(2)}$ = 2 dB (corresponding to a minimum Energy Factor at n_h = 2 and P_d = 80% Ref. 8).)	L_t = 2 dB
	τ_t = 4 μs

a pulse peak power of only 9 kW or a mean transmitter power of 18 W is required for 80% probability of detection on a 10 m^2 target (Ref. 1).

This low value of transmitter power is primarily due to the high antenna gain corresponding to the $0.5^\circ \times 0.5^\circ$ pencil beam assumed. Interesting possibilities from the point of view of system construction can result from such a low transmitter power requirement.

6. THE RADAR BEAM DISPOSITION FACTOR, $L_d^{(2)}$.

The solid angle that can be covered is determined by the number of beam positions available, the type of scan pattern, and the spacing between beam centre positions.

There exists an optimum value of the beam centre spacing when considering the total energy needed to search a given volume. Early investigations by Bradley (Ref. 4) optimized the beam spacing on the basis of maximum average power in space. However, average signal-to-noise ratio has no operational meaning; optimization of operational performance must in all cases be performed on the basis of average probability of detection. This has been pointed out also by Barton and Hall (Ref. 5).

Improved procedures taking into account the effect of beam overlap were reported by Evans and Kanyuck (Ref. 6) and by Hahn and Gross (Ref. 7). Their results, however, were valid only for a single hit on target and for 50% detection probability.

When optimizing on the basis of average probability of detection, the optimum spacing becomes dependent on the detectability conditions, such as the number of hits on target, the type of signal scintillation and the probability of detection, because of the non-linear interdependence of these parameters.

An optimization has been carried out at SHAPE Technical Centre with respect to the Energy Factor, E, which represents the product of transmitter power and the time required to search a given volume while providing the required average detection probability. Values of E as a function of detectability conditions, and for both optimum and fixed beam spacings, are shown in Figure 6.

It can be seen that only for 50% detection probability can a fixed spacing be used without losses.

A full set of data for the optimum beam spacing for a range of detectability conditions and for all Swerling type signal scintillation models is provided in Ref. 8.

In all cases the staggered beam arrangement, or triangular lattice, is optimum because of the reduced spatial power variation. For this arrangement of beam positions the dependence of optimum beam spacing on n_h and P_d is shown in Figure 7. The beam centre spacing S_p' is expressed relative to the half-power antenna beamwidth.

To enable calculation of the required transmitter power, etc., for each setting of the beam spacing, a "Radar Beam Disposition Factor, $L_d^{(2)}$" is introduced, which factor replaces the well-known "beam shape" or "pattern loss" factor of the scanning type radar. These are not applicable to the stepped-scan, phased-array radar because there is no modulation by the beam shape of the received train of pulses. Also, the effect of an increase in detectability due to beam overlap has to be accounted for, which overlap can cause this factor to be greater than 1.

Figure 8 gives a number of values of the radar beam disposition factor for different detectability conditions (Ref. 8).

The solid angle of ψ_s° in azimuth x ϕ_s° in elevation that can be covered, assuming a staggered beam arrangement, is

$$\psi_s \times \phi_s = \tfrac{1}{2}\sqrt{3} \times S_p'^2 \times \theta_A \theta_E N_b$$

The time required to search this volume is

$$T = \frac{2n_h}{3 \times 10^5} N_b R_{max} \text{ (km)}.$$

For our example with $\theta_A = \theta_E = 0.5^\circ$, $N_b = 1000$, $n_h = 3$, and $R_{max} = 300$ km, and for a relative beam spacing $S_p' = 0.9$ (corresponding to a minimum energy factor at $n_h = 3$ and $P_d = 80\%$) :

$$\psi_s \times \phi_s = 175 \text{ degr.}^2$$

That is, for instance, 3.5° in elevation x 87° in azimuth with a data interval of T = 6 seconds.

For $S'_p = 0.9$, the relative solid angle Ω' covered, is

$$\Omega' = \frac{\theta_s \times \phi_s}{\theta_A \times \theta_E} = 0.7\ N_b$$

The dependence of Ω' on P_t is shown in Figure 9.

7. VOLUME COVERAGE HORIZON SEARCH.

In the above example, providing 87° azimuth coverage, no account was taken of the natural beam broadening occurring at large angles away from the direction normal to the antenna aperture.

The antenna beamwidth at an angle ψ away from broadside increases due to the reduction of antenna aspect and is approximately :

$\theta_A(\psi) = \theta_A(0)/\cos\psi$, where $\theta_A(0)$ is the beamwidth at broadside, the relation being valid for angles up to about 85° off broadside.

Due to the natural beambroadening the number of beam positions, $N_{b\ az}$, required in order to cover an azimuth sector ψ_s reduces to

$$N_{b\ az} = \frac{2 \sin \frac{1}{2} \psi_s}{S'_p\ \theta_A(0)(\text{rad})}$$

The number of beams required in order to cover the volume $\psi_s \times \phi_s$ (when ψ_s is large), using a staggered beam arrangement, and the time to search this volume are, respectively :

$$N_b(\psi_s \text{ large}) = \frac{1}{\frac{1}{2}\sqrt{3}\ S'^2_p} \times \frac{\phi_s}{\theta_E} \times \frac{2 \sin \frac{1}{2} \psi_s}{\theta_A(0)(\text{rad})}$$

$$T(\psi_s \text{ large}) = \frac{2n_h}{3 \times 10^5} \times N_b(\psi_s \text{ large}) \times R_{max}(\text{km})$$

For our example it follows that an azimuth sector of 81° can be covered in 6 seconds or, alternatively, an azimuth sector of 120° in 7.4 seconds.

The azimuth coverage obtained in this way is not uniform due to the reduction of antenna gain as a function of off-broadside angle, ψ. It is possible, by adjusting the mutual coupling between antenna elements, to reduce the variation of antenna gain with angle at the expense of reducing the maximum antenna gain at broadside, G(0).

For the purpose of the present general discussion we will assume that no such frequency-sensitive measures are taken, and that the antenna gain, $G(\psi)$, is proportional to the antenna aspect, so that :

$$G(\psi) = G(0) \cos\psi.$$

As a result, the detection range drops proportional to $\sqrt{\cos\psi}$ and a plan position coverage is obtained as shown in Figure 10.

The loss in coverage towards the sides can be compensated for by increasing the number of hits transmitted in these directions. Because the antenna gain only drops rapidly towards the edges, only a small increase in the total number of hits, and consequently in the total search time, is needed.

Examples of plan position coverages with and without compensation are shown in Figure 10. The search times are, respectively, 9.7 and 7.4 seconds for the parameters as assumed above.

If horizon-contour-following is used, these search times are reduced. For the terrain-screening example shown earlier, giving a reduction of 17%, the search times are 8.3 and 6.3, respectively.

8. REDUCING SEARCH TIMES.

These search times, or data intervals, can be drastically reduced further, if required, by either

- increasing the beamwidth in azimuth or elevation,
- reducing the number of hits,
- increasing the beamspacing above the optimum,
- using multiple simultaneous beams,

or a combination of these. All of these measures will involve increased transmitter power.

It should also be noted that not all azimuth angles have to be searched.

The above considerations did not include the possible use of the sequential detection technique. A substantial literature on this technique already exists. It has been shown that, by using this technique,

a reduction in the required number of hits or in the required transmitter power by a factor of about two can be achieved for a low number of resolution cells.

The use of this technique is particularly well suited to the horizon search mode since the range interval of interest can be restricted to a small value near maximum range. Only for the lowest beam elevation position does a large range interval have to be searched.

It is concluded that a large azimuth sector can be searched in a few seconds even when a narrow beam is used.

All of the time left within the required minimum data interval will be used for increasing the time spent in the lowest elevation directions for improving MTI performance.

9. FULL-VOLUME SEARCH.

For the sake of completeness, this short discussion on full-volume searching is included although, as stated earlier, there is no real need for a search of all elevation angles.

When searching for aircraft-type targets at the higher elevation angles the computer-controlled pencil beam radar can conveniently take advantage of the reduced target ranges resulting from the aircraft height ceiling. Firstly, the time inverval between pulses can be reduced proportionally to the maximum unambiguous range, R_{max}, required. Secondly, for a given constant detection probability the required signal-to-noise ratio, D_s, is reduced.

The change of R_{max} and D_s as a function of elevation angle, ϕ, is shown in Figure 11, taking into account earth curvature and atmospheric diffraction, for a target at a constant height of 53,000 ft. (Figure 1) The required signal-to-noise ratio is reduced by 30 dB at 20° elevation, assuming no antenna gain variation. Consequently, the probability of detection rapidly increases to 100% (Figure 12), all parameters being assumed constant.

If the complexity arising from the use of antenna gain variation has to be avoided, the extra energy available will enable a reduction to be made in the number of hits transmitted as a function of elevation angle. An example is shown in Figure 12.

Due to reducing both the number of hits and the pulse travel time a significant reduction is obtained in the time to search the higher elevation angles. Figure 12 shows the incremental search times, $n_\theta T_1$, per unit of elevation angle for fixed and for reducing numbers of hits. It follows that the entire volume above 2° elevation can be searched in a time roughly equal to that required for the horizon search.

The probability-of-detection curve resulting from the use of this procedure (Figure 12) still has a large region of very high detection probability, indicative of the non-optimum use of the available energy.

To achieve optimum conditions, basically the objective in controlling the system should be to provide a constant volumetric detection probability. In this case a minimum search time is also obtained.

This objective could be met elegantly if forced beam broadening by antenna illumination control is introduced in such a way that both the beamwidths of the pencil beam antenna in azimuth and elevation are made to vary in inverse proportion to target range; that is,

$$\theta_A(\phi)/\theta_A(0) = \theta_E(\phi)/\theta_E(0) \propto R(0)/R(\phi).$$

This differs from the usual cosecant-squared type of gain variation in that both the azimuth and elevation beamwidths are changing.

A particularly noticeable effect of such an "Inversely-Proportional" antenna is that the tangental plot accuracy, which is proportional to $R.\theta$, remains constant in both azimuth and elevation.

The incremental search times of the "Inversely-Proportional" antenna rapidly reduce with elevation angle and become very small at elevation angles above 10° (Figure 12).
In our example, the beamwidths will increase from 0.5° at the horizon to 3° at 20° elevation.

Since about three hits provide optimum integration efficiency there is no need also to change the number of hits though, if necessary, this could be done to reduce the maximum beam broadening.

It is concluded that, by the use of p.r.f. adaptation and either hit reduction or inversely-proportional beam widening the full-volume search time can be held to acceptable values even when very narrow beams are used.

The inversely-proportional beamwidth adaptation could be approximated by changing the beamwidths in steps, in particular between 2° and 10° of elevation.

The conventional cosecant-squared antenna is inefficient from the time management point of view since neither the pulse repetition rate nor the number of hits on target can be varied.

10. NARROW BEAMS.

The present paper assumes the use of narrow antenna beams because of their inherent advantages from the operational point of view.

The increased antenna gain itself is not the primary advantage as the resulting reduced transmitter cost has to be balanced against the increased cost of the antenna, although it should be remembered that, for the same radar range, the required transmitter power is inversely-proportional to the square of the number of antenna elements.

The main virtues of the narrow beam system from the operational point of view are :

- the high target location accuracy, which also allows a reduction in the volumes to be searched for target position updating in the multi-target tracking phase;
- the reduced ground illumination, giving reduced ground clutter return and the possibility of minimising the volume to be searched under clutter conditions;
- the reduction in size of the radar resolution cell which represents the intercepted clutter volume and therefore reduces rain clutter return;
- the reduced sensitivity to external interference.

11. FREQUENCY AGILITY.

Also assumed in the example is the use of frequency agility (pulse-to-pulse frequency jumping). The reasons are that it has the following advantages:

- increased detection probability, due to the fast target signal scintillation : 3-dB gain for three hits and 80% detection probability (Ref. 1,2,3)
- reduced target glint error, due to decorrelation of returns
- reduced multi-pass error and reduced "lobing" effect, due to filling-in of interference nulls caused by ground reflection
- improved interference resistance.

12. "FIRST-GRADE"- AND "SECOND-GRADE" CONTROL.

The discussion above on optimising radar operational performance has been restricted mainly to what could be called "First-Grade Control", that is, simple system control procedures, optimising performance primarily by management of the spatial power distribution and the distribution of search and tracking times and not requiring multiplication of hardware or extensive computer processing.

The purpose in limiting the discussion to first-grade control is to show that, even with simple system control procedures a large increase in efficiency of the time and power management can be obtained, and that acceptable data rates can be achieved even when narrow beams are used.

Included under first-grade control are horizon-search and grazing-search procedures, and the use of a single, multiple-channel receiver.

"Second-Grade Control" would imply control procedures requiring a multiplication of hardware or extensive data processing. Examples are sequential detection methods requiring detection channels for each resolution cell and extensive adaptive control procedures.

Forced beam-widening would also be classified as second-grade control since its complexity does not result from switching off antenna sections, for example, but from the need for power redistribution or resetting of element power levels to restore the proper antenna illumination function for the maintenance of sufficiently low side-lobe levels.

13. MULTI-TARGET TRACKING.

Tracking is discussed here only in the context of beam management and time distribution.

As mentioned above, the tracking of targets detected with the horizon search would require a number of beam positions which is small compared to that required for searching. If the system is operated in such a way that target-position updates are performed before the targets' next positions have an uncertainty corresponding to one beamwidth, and assuming two-dimensional monopulse position extraction, only one beam position per target is needed. Also, in general, not all targets in radar cover need to be tracked, since certain parts of the air space may not be of interest, and some tracks could be terminated when they are no longer of interest.

Consequently, the data interval, or the interval between looks on the same target for track updating will not be much larger than the search times derived for the examples discussed above. This additional time required for multi-target tracking is further drastically reduced because, in most cases, the elevation angles are relatively high and ranges short, which means proportionally smaller pulse intervals (Figure 12).

Furthermore, it should be realised that the achievable data intervals mentioned earlier constitute the average of the track update intervals of the different targets. These intervals can be different, depending, for example, on the track quality, the particular manoeuvring status of each track, and the collision chance. In "track-while-scan" systems, using continuously-rotating search-type radars, the available data rate is necessarily related to the worst target manoeuvring possibility. In the steered system , adaptive control of data rates can be simply accomplished. Since the chance of all targets turning at the same moment or otherwise requiring close attention is small, the steered system can employ increased data rates on a few selected targets and reduced data rates on others. Consequently, the effective data rate is much higher than the average data rate.

It is possible in principle to obtain a further substantial reduction in total tracking time by updating several target positions within a single pulse travel period, corresponding to the target at the longest range. In this mode pulses are transmitted sequentially in the subsequent directions of increasing target range, all pulses being sent within the maximum pulse travel time. Returns are listened to in the same way. The total number of targets that can be illuminated in this way within one pulse repetition period is limited by the range of the nearest target.

This mode of operation requires, however, a very rapid recalculation and repositioning of the beam. The tracking of several targets within one pulse repetition period also creates an increased demand on the transmitter capabilities. Consequently, interpulse-period track updating is classified as second-grade control.

When however updating targets in subsequent data intervals, the computation of the next beam direction can be performed in the preceding data interval.

Using the narrow pencil beam, three-dimensional target tracking is achieved. By providing a fixed tilt of the antenna, of say 30°, targets can be tracked to very high elevation angles, so that the system will not have a "zenithal null". No significant reduction in the horizon search performance will occur as a result of this tilting due to the slow change of beamwidth and antenna gain for small off-broadside angles. Separation of the tracking from the searching function permits optimization of the system for both functions independently. For instance, the most suitable transmitter waveforms can be selected for use in each of the two phases.

14. CONCLUSIONS.

To be competitive from a cost/effectiveness standpoint, it is essential that the inherent flexibilities of the computer-controlled radar system are exploited to the fullest possible extent. With proper control, the large flexibility of operation of this second generation of radar systems leads to a greatly improved performance as compared with the non-intelligently rotating, full-volume search, radar.

Even when the systems management is limited to "First-Grade Control", not requiring multiplication of hardware or extensive computer processing, a large increase in efficiency is achieved, both in the use of the time and power available, in generating the wanted information.

For air space surveillance, including the detection and tracking of airborne targets, it is sufficient to perform an "Horizon Search", with a small elevation sector, covered by a few elevation beam positions, followed by the individual and selective tracking of the targets detected (that is, "Multi-Target Tracking"). The substantial power wastage of the large-volume search and tracking system is in this way avoided while a better performance is obtained.

Air-Traffic Control is a good example of a civil system not requiring a full-volume search. Flight plans, target handover procedures and prescribed corridors limit the volume to be searched; consequently, the use of Multi-Target Tracking alone would often suffice.

Inclusion of the horizon-contour-following and "Grazing Search" procedures as part of the horizon search mode limits the total volume to be searched and the number of ground illuminating beam positions required, and leaves the maximum time for anti-clutter operation. In this way anti-clutter performance is improved, particularly if narrow antenna beams are used.

With regard to the volume that can be covered within a given data interval, 1000 beam positions appears to be an approximate system design constant.

Optimum integration efficiency occurs when the number of hits on target is approximately three.

To maintain an optimum "Energy Factor" (the product of search time and transmitter power) it is necessary to adjust the spatial beam disposition to match the radar mode of operation, particularly for detection probabilities higher than 50%. Values for the optimum setting of beam spacing for different numbers of hits on target, detection probabilities, and target scintillation models, have been provided.

For the evaluation of phased-array detection performance a "Radar Beam Disposition Factor, $L_d^{(2)}$" is introduced, which factor replaces the beam-shape-loss or pattern-loss factor of the continuously-rotating search-type system, which is not applicable to phased-array radars. Values of $L_d^{(2)}$ for different radar operating conditions are given.

It has been shown that, even when using very narrow pencil beam antennas, a large azimuth coverage can be obtained within acceptable data intervals.

A very small transmitter power is shown to be sufficient for the type of system operation described. Interesting system construction aspects may arise from this.

An example is given of compensating the loss in detection probability at large off-broadside angles of fixed arrays by increasing the number of hits on target.

Coverage of the full volume above the horizon-search is not considered necessary. It is shown that, if nevertheless such a search is desired, this can be performed within a time comparable to that required for the horizon search, by exploiting the reduced target ranges for reduction of the pulse repetition period and reducing the number of hits required to meet a given detection probability.

A constant spatial detection probability and a constant tangential target location accuracy can be achieved by the use of an "Inversely-Proportional" antenna, of which both the azimuth and elevation beamwidths are inversely proportional to range.

The importance of the use of very narrow antenna beams is stressed, for reasons of reducing ground illumination, improving performance against clutter, and improving target tracking performance.

Several of the arguments given apply also to systems having phase-frequency steering or using one-dimensional inertialess steering only.

It is concluded that, with an electronically-steered radar system, the time and energy required for multi-target tracking is small compared to the volume search time. Further reduction of the total target tracking time is obtained by exploiting the reduced target ranges.

The effective tracking data rate is much larger than the average system data rate, because of the possibility of adapting target data rates to match track conditions.

A small tilt of the antenna permits target tracking for all elevation angles, in this way avoiding the usual "zenithal null".

In the computer-controlled system, using the mode of operation described, it is possible to separately optimize the searching and tracking functions.

A new approach to radar systems design and evaluation is necessary. For optimizing the information generating capabilities, a reversed approach should be taken : In the controlled multi-target tracking radar system, the data user will continuously determine target and search area priorities and information requirements. By systems control the data processing system is instructed accordingly and determines track update and search area requirements, and in turn instructs the radar and signal processing system to generate the required information. The volumes left to be searched are continuously being minimized as part of the systems control process.

It is to be noted that this procedure is the reverse of that customary for conventional search and tracking radar systems which generate information continuously, irrespective of the momentary need, and which leave it to the signal and data processing system to filter out the wanted information from the large amounts of unwanted data.

The fact that little has been done so far towards such an approach is exemplified by this Symposium which does not include a "System Control" category, and where this paper has been listed under "Data Processing", in line with the classic sequence of events in radar and data processing.

John Allen has said, at the Long Island IEEE Symposium in 1969, that "Phased-Arrays cannot compete with mechanically scanned radars". It seems to me, however, that the classic search radar, as such, will disappear and that the future of radar (for civil application too) lies in systems-controlled narrow-pencil-beam, multi-target tracking radars.

It may be disappointing to the radar engineer to realize that radar will become subordinate to and under continuous control of the system, rather than being an autonomous data generating source; but we can no longer afford to go on spreading all our pulses inefficiently into the space around us in an indiscriminate manner.

The radar engineer, however, may find some consolation in the fact that the advent of adaptive control procedures mark the end of the radar as a classic example of a compromise system.

REFERENCES.

1. J.A. Bijvoet, "Standard Methods for Predicting and Specifying Performance of Air Surveillance Radar Systems",
SHAPE Technical Centre, Technical Report TR-50-U, April 1969.

2. J.I. Marcum and P. Swerling, "Studies of Target Detection by Pulse Radar",
IRE Transactions on Information Theory, Vol. IT-6, no. 2, April 1960.

3. J. Eendebak, "Detectability Curves for Swerling Cases III and IV Target Models",
SHAPE Technical Centre, Technical Report TR-50-U, Supplement 1, April 1970.

4. G.A. Bradley, "How to determine Antenna Pattern Loss",
Microwaves, Vol. 6, no. 1, p. 38-41, January 1967.

5. Correspondance between D.R. Barton, W.M. Hall and G.A. Bradley,
Microwaves, Vol. 6, no. 6, p. 138-142, June 1967.

6. W.T. Evans and A.J. Kanyuck, "Statistical Analysis of Antenna Pattern Loss Effects",
IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-5, no. 1, p. 77-82, January 1969.

7. P.M. Hahn and S.D. Gross, "Beam Shape Loss and Surveillance Optimization for Pencil Beam Arrays",
Correspondance in IEEE Transactions on Aerospace and Electronic Systems,
Volume AES-5, no. 4, p. 674, 675, July 1969.

8. J.A. Bijvoet and J. Eendebak, "The Radar Beam Disposition Factor, $L_d^{(2)}$",
SHAPE Technical Centre, Professional Paper PP ..... , ... 1970.

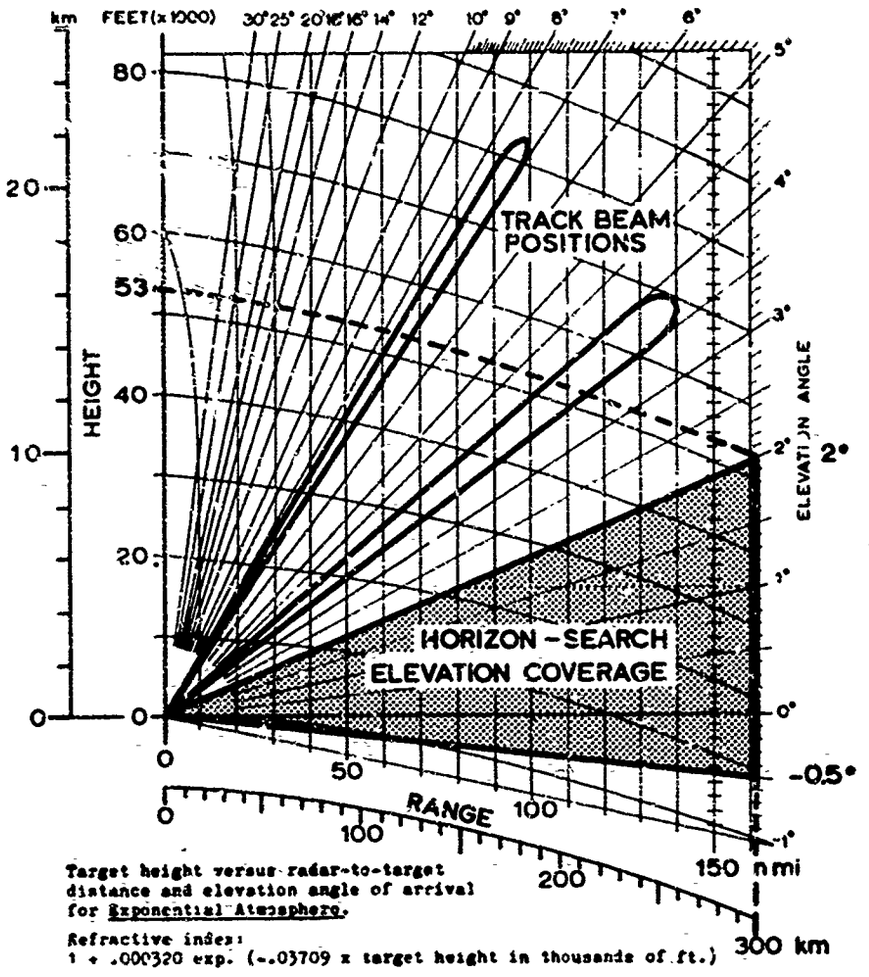


Fig. 1. "Horizon search", with limited elevation coverage, followed by selective Multi-Target Tracking.

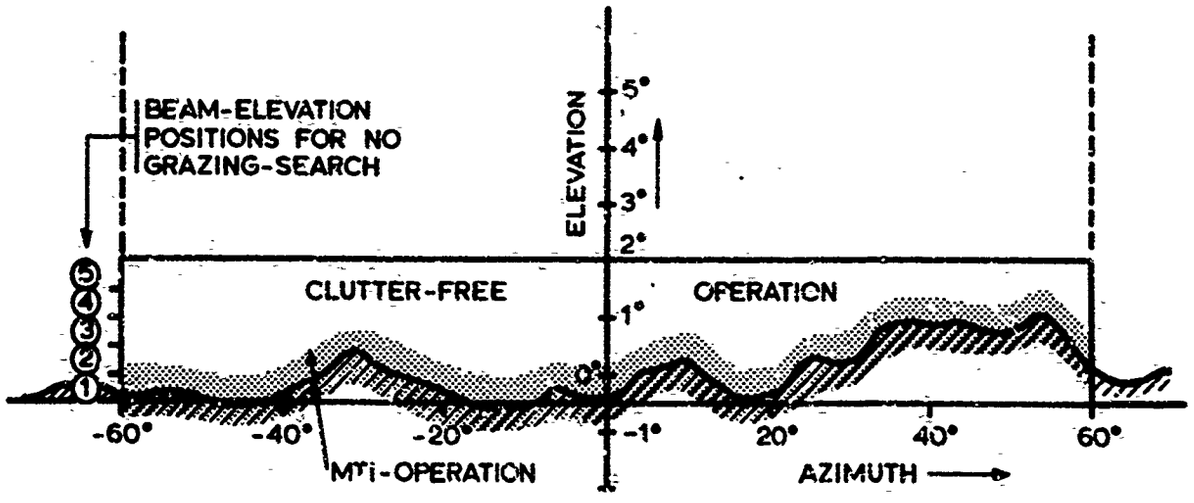


Fig. 2. Elevated grazing-search, reducing search volume and MTI-operation requirements.

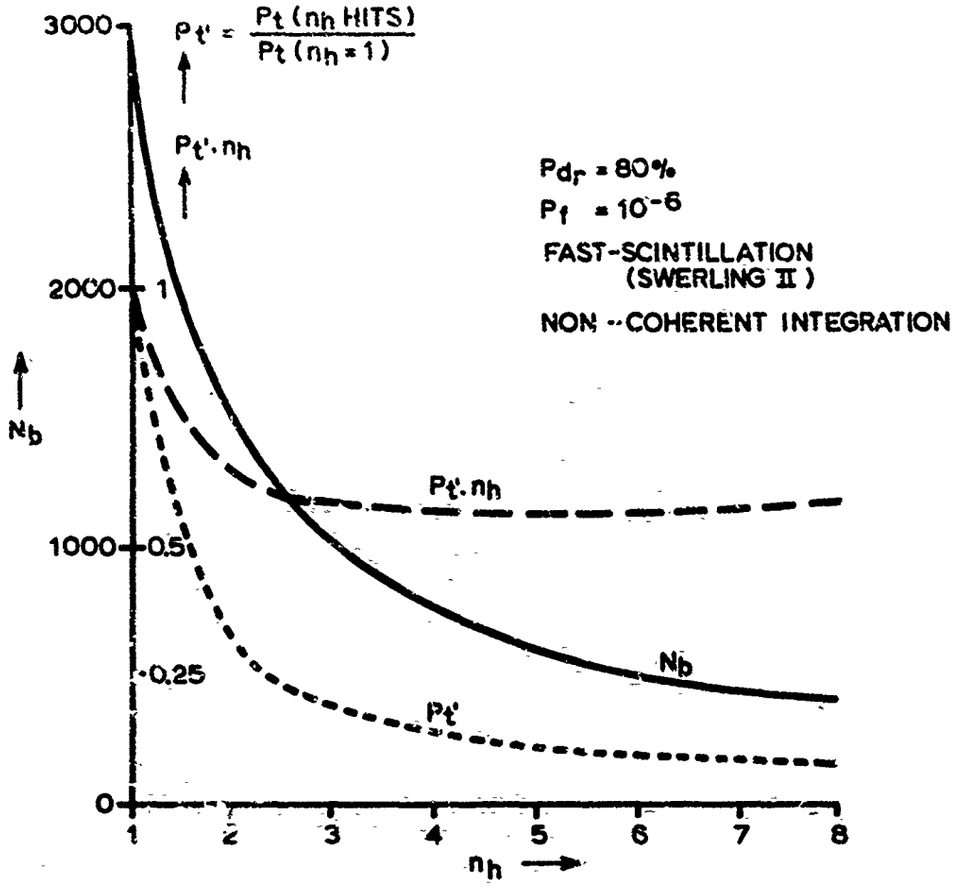


Fig. 3. Effect of choice of number of hits of target, n_h, on number of obtainable beam positions, N_b, required relative transmitter peakpower, P_t', and relative transmitted energy per beam position $P_t' \cdot n_h$.

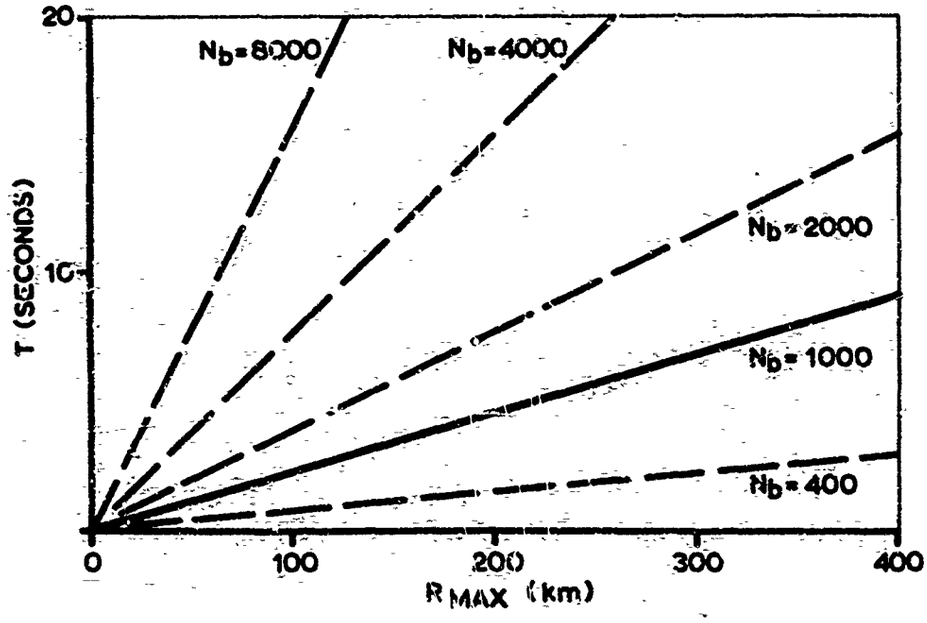


Fig. 4. Relation between data interval, T, maximum unambiguous range, R_{max}, and number of obtainable beam positions, N_b, for three hits on target.

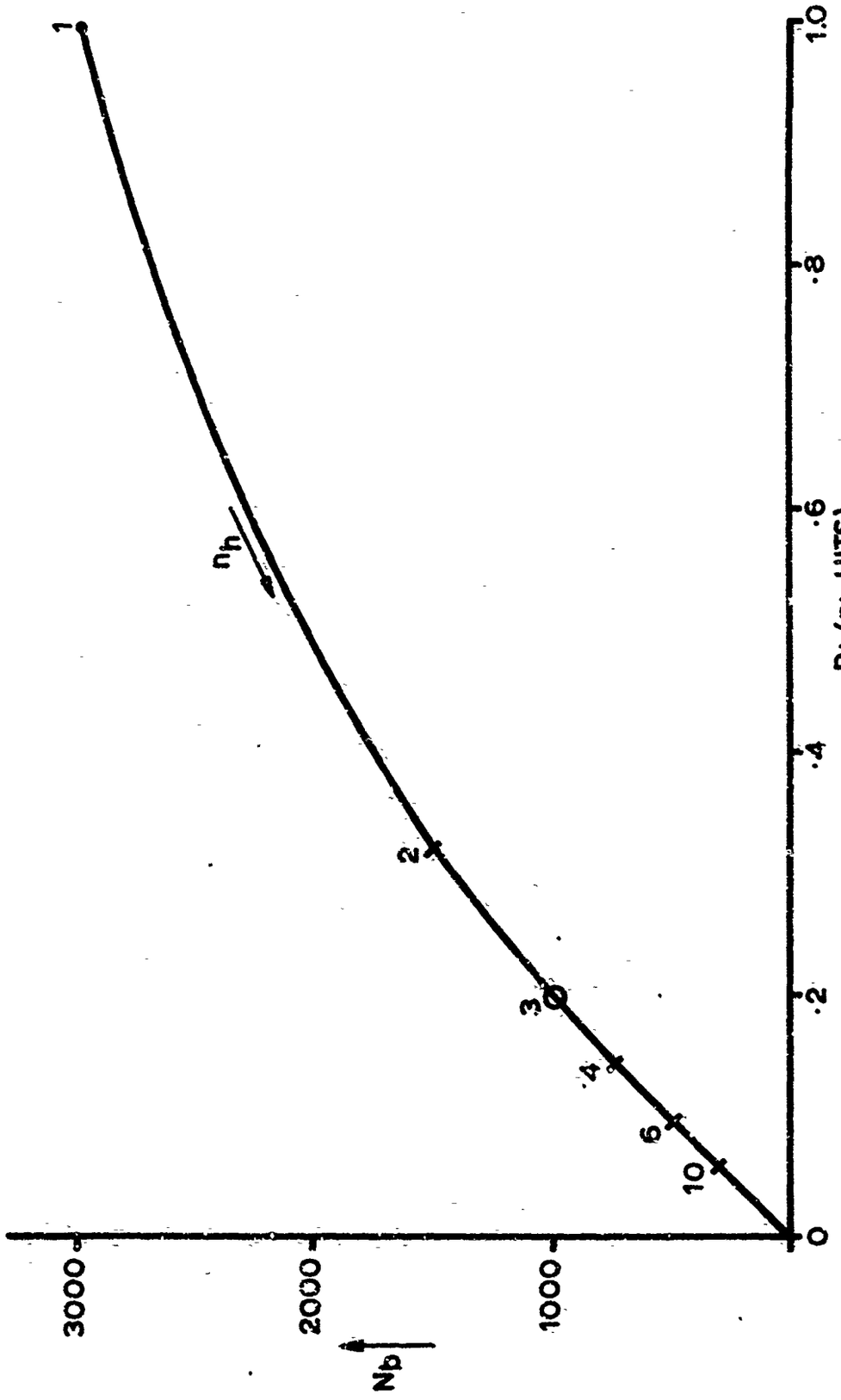


Fig. 5. Trade-off between number of beam positions, N_b, obtainable and required relative transmitter peakpower, P_t'.

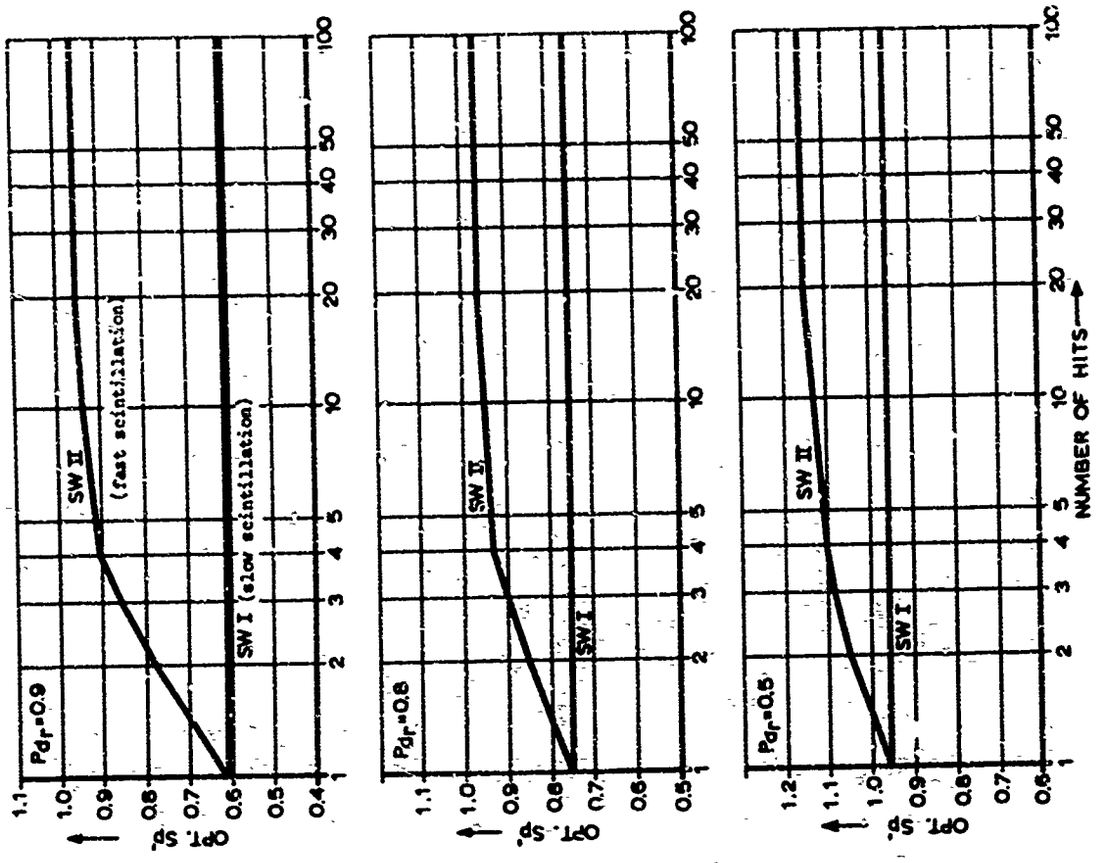


Fig. 7. Optimum beam spacing as a function of detectability conditions (staggered beam arrangement). $P_f = 10^{-6}$ Non-coherent integration.

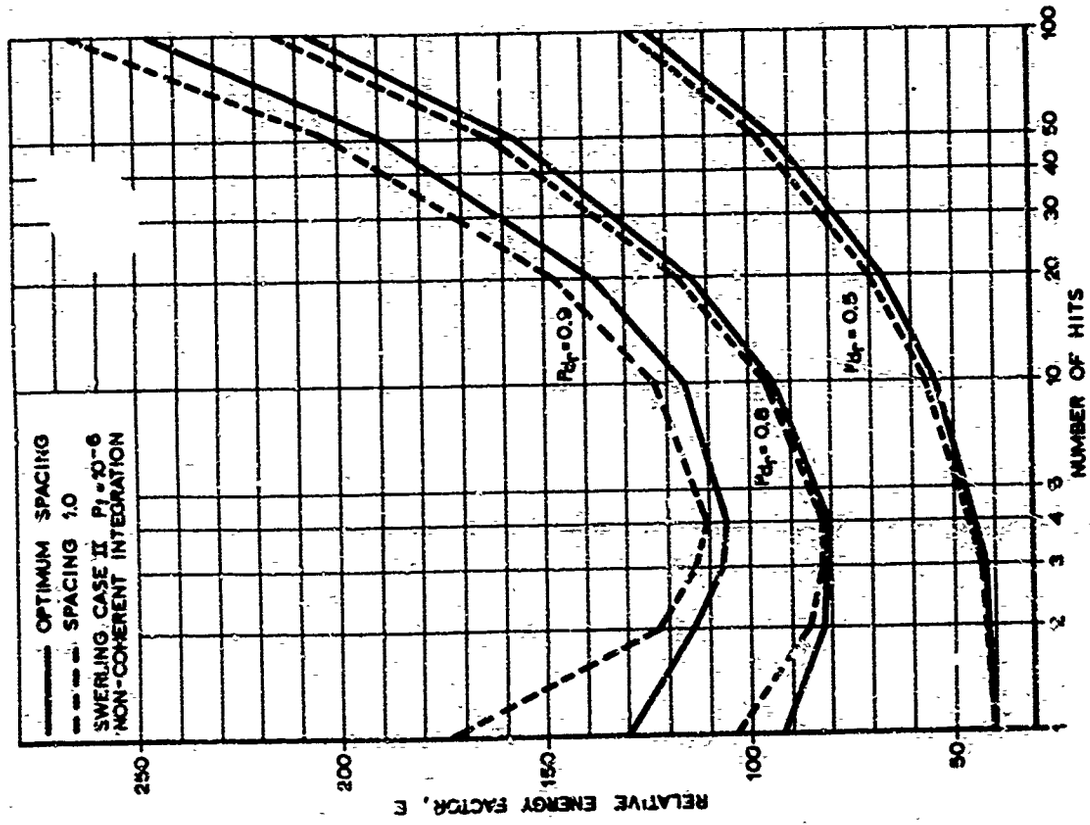


Fig. 6. Relative energy factor, E, as a $f(n_h, P_d)$ for optimum and fixed beam spacing.

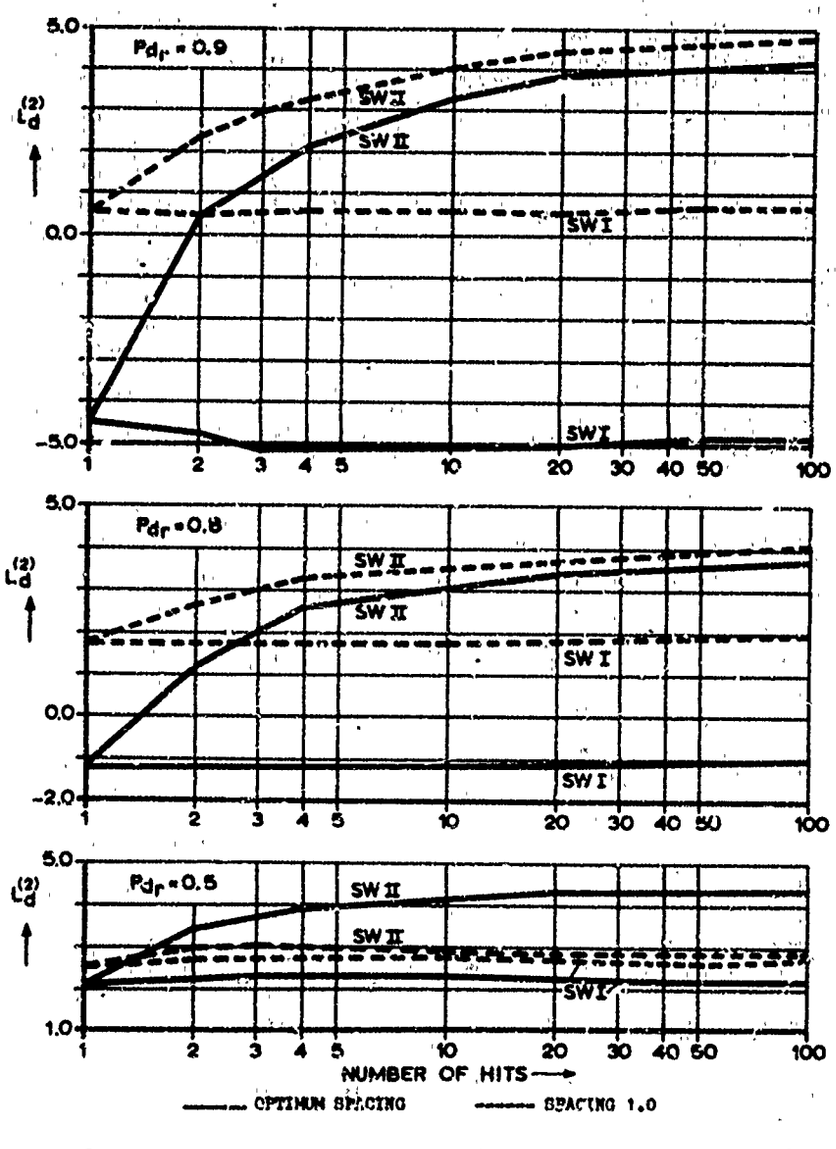


Fig. 8. Radar beam disposition factors (staggered beam arrangement). $P_f = 10^{-6}$ Non-coherent integration.

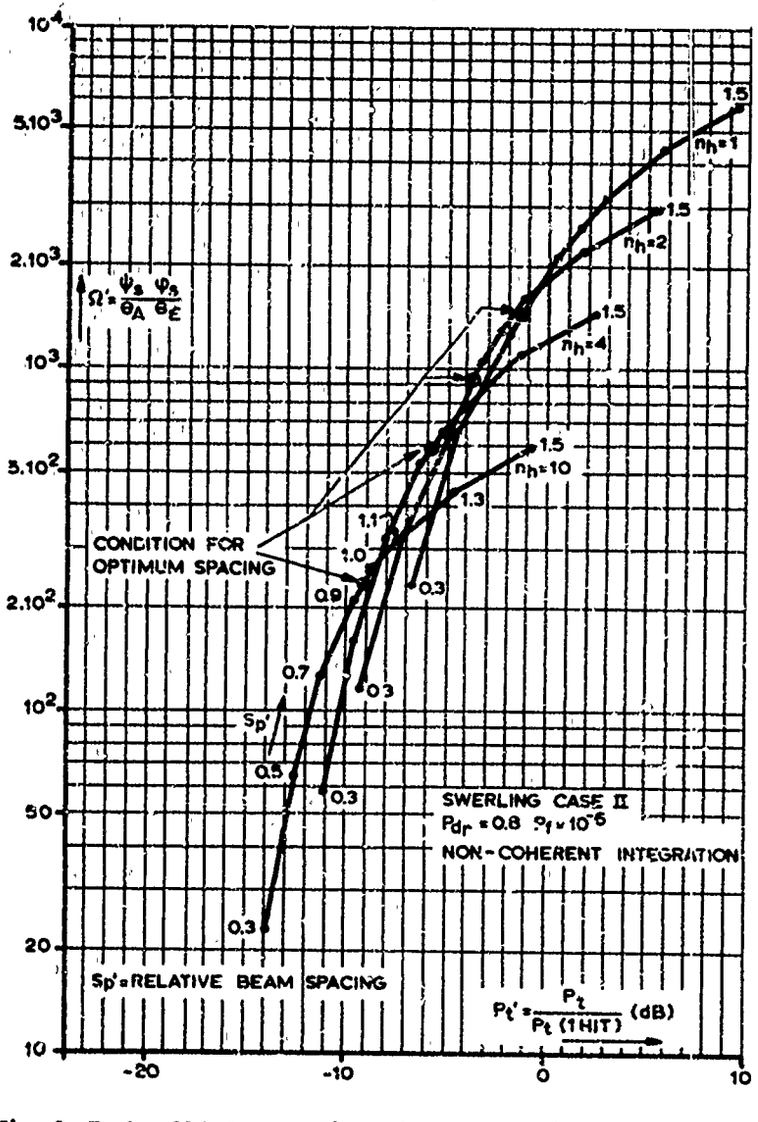


Fig. 9. Trade-off between solid angle coverage, Ω', and transmitter power, P_t'.

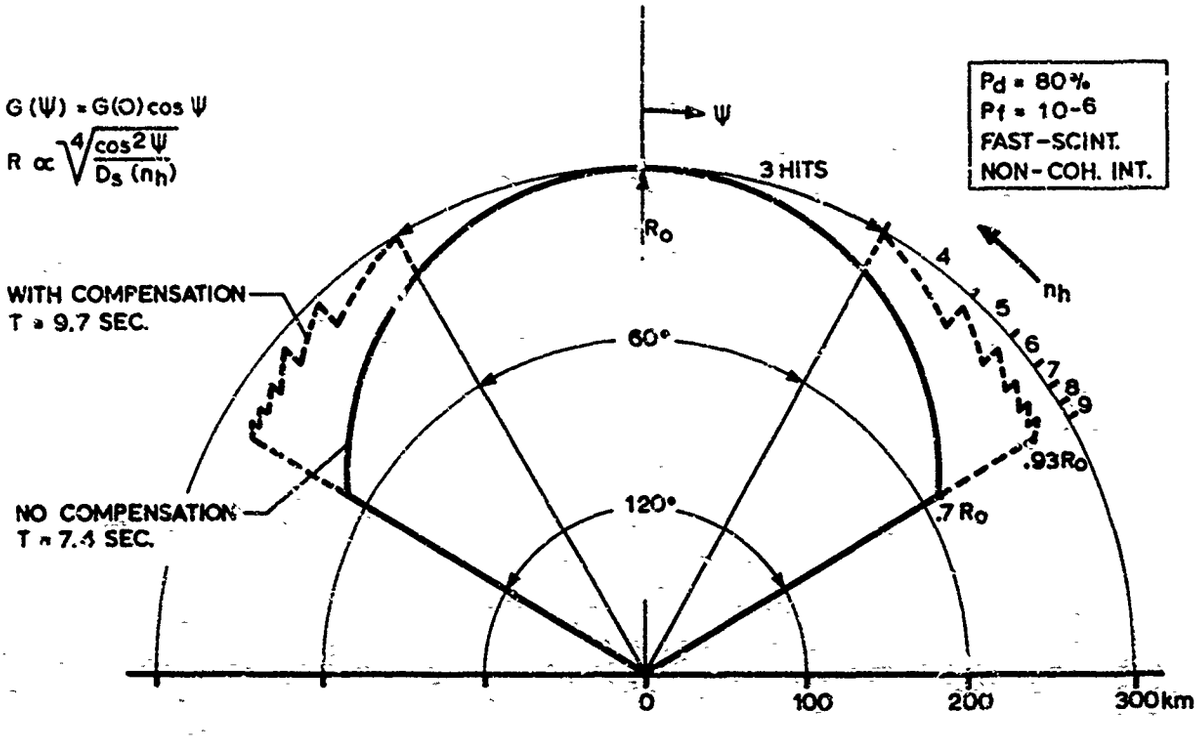


Fig. 10. Plan position coverages with and without compensation of antenna gain reduction at off-broadside angles.

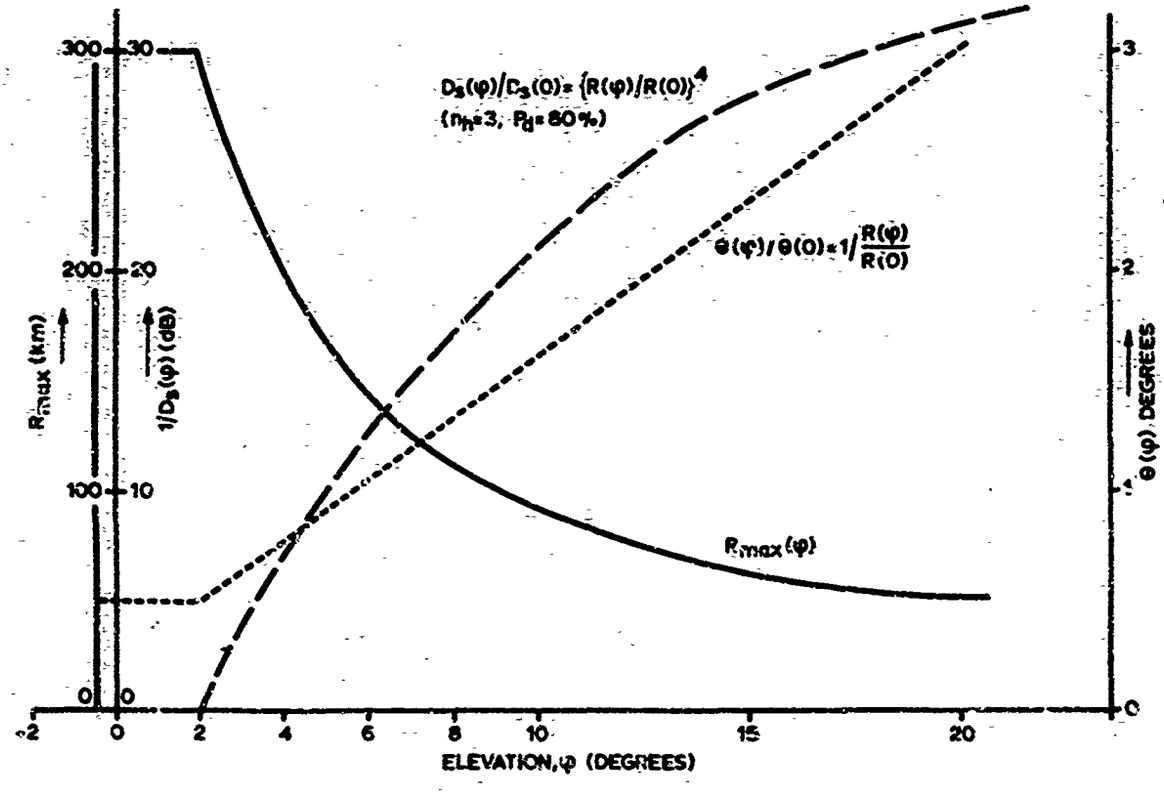


Fig. 11. Dependence on target elevation of maximum range, R_{max}, required signal-to-noise ratio, D_s, and change of azimuth and elevation beamwidth, θ, for "inversely-proportional" antenna. (Target height = 53,000 ft.)

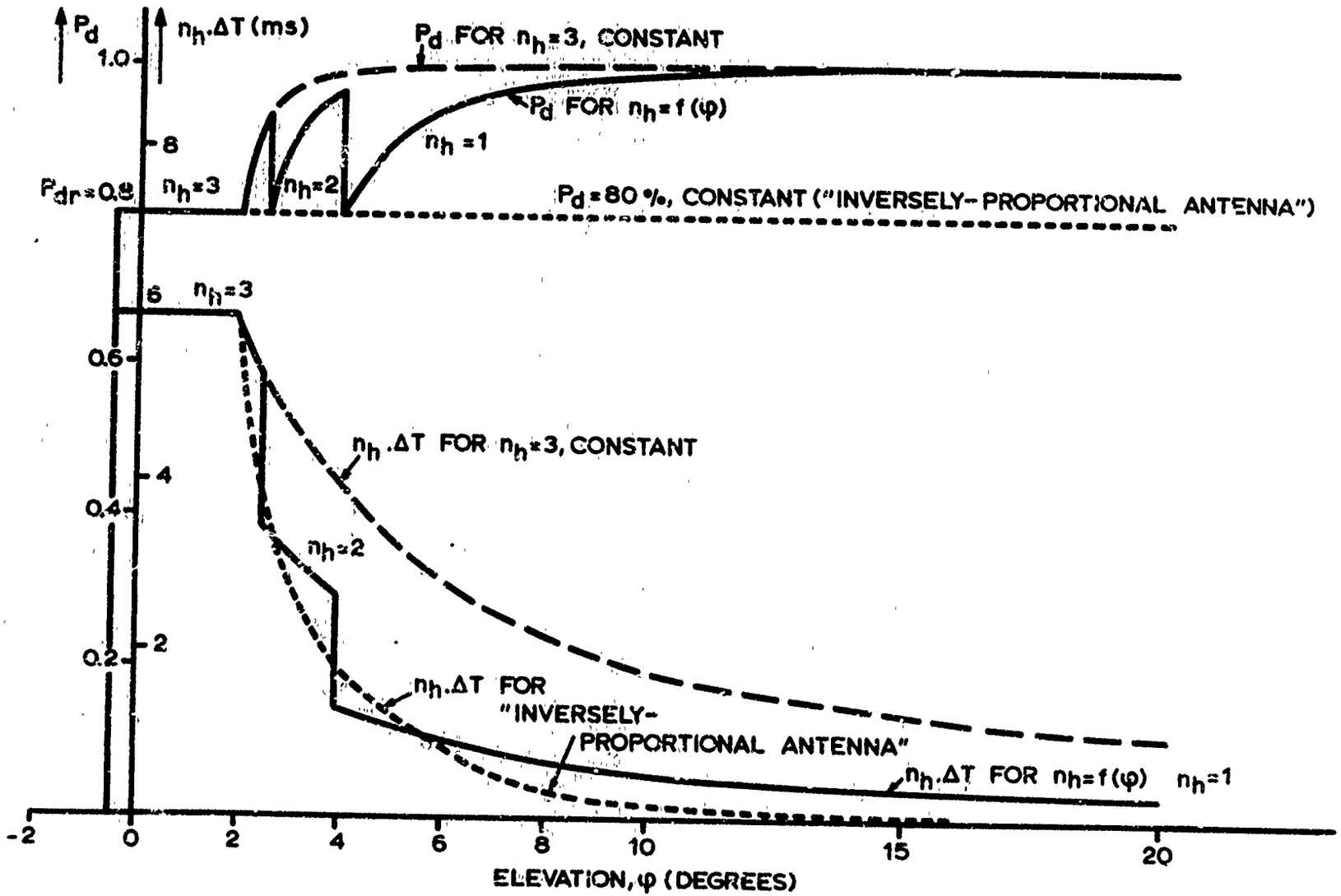


Fig. 12. Distribution in elevation of incremental search times, $n_h.\Delta T$, and detection probability, assuming pulse repetition rate adaptation to target range, and for additional adaptation of the number of hits on target or for "inversely-proportional antenna".

DIGITAL PLOT EXTRACTION OF PRIMARY RADAR

by

N. H. A. Smith and I. R. Jackson

Plessey Radar Ltd.,
Addlestone,
Surrey, U.K.

DIGITAL PLOT EXTRACTION OF PRIMARY RADAR

N.H.A. Smith B.Sc. M.I.E.E.
I.R. Jackson B Sc. Eng.

Plessey Radar Ltd.,
Addlestone,
Surrey, U.K.

SUMMARY

This paper covers the digitisation or plot extraction of primary radar data which can be used to fulfil two purposes. Firstly, to permit narrow band transmission of radar data from a remote radar site to a centre at which displays are located, and secondly, to provide a bright display of primary radar data using conventional fluoride displays without involving scan conversion. The paper will cover the techniques involved in both hit detection and plot detection of aircraft echoes derived from a primary radar together with the various techniques employed to reduce the incidence of false plots due to clutter and weather. Reference to secondary radar plot extraction will be made and the system requirements for processing and displaying plot extracted radar data will also be discussed.

1. INTRODUCTION

1.1 Plot extraction of primary radar may be defined as the process of correlating the returns from successive radiated pulses so as to determine, once per antenna rotation, the positions of illuminated targets. The process is essentially probabilistic and a position report is generated whenever the likelihood of a target being present, at the specified position, exceeds some threshold level. Thus, reports may be generated when targets are not present (probability of false plot), or reports may not be generated for targets which are present (probability of detection).

1.2 On the conventional radar cathode ray tube display, the human operator can very effectively make a visual correlation of the returns to determine the presence of targets. Using phosphors of moderate persistence, his correlation field, in practice, extends to cover several antenna rotations. He not only correlates the returns to detect plots, but also correlates the plots from successive antenna scans so as to observe the movement of the target. Thus, he can track the target. His ability to carry out this tracking process, especially for any length of time, is limited. He can really only manage to deal with a few targets at a time. The number he can manage reduces dramatically as the speeds of the targets increase and their movements become more erratic.

1.3 The shortcomings of the human operator led to a requirement in military intercept systems and in missile systems for the tracking to be performed by firstly analogue and subsequently digital computers. This in turn, required the process of plot extraction to be carried out automatically. Thus, the development of plot extraction techniques owed much originally to the needs of military systems.

1.4 More recently, however, the rapid expansion of civil aviation and the use of secondary radar has led to the requirement in air traffic control systems, for the direct labelling on the radar position display of significant information, relating to the targets displayed. This has caused considerable development effort to be put into extraction systems for both primary and secondary radar, and systems are now operational using tracking computers connected to primary and secondary plot extractors. As well as providing outputs in a form suitable for computer processing, plot extractors also considerably reduce the bandwidth required for the transmission of radar data. Thus compared with the several MHz video bandwidth of a radar, after plot extraction the data may typically be carried over circuits of a few kHz bandwidth. This means that where a single centre is to be used to provide control over a large area, the whole area can be covered by a series of remote radars, sited so as to give the best radar coverage. The radars can then be connected to plot extractors and the extracted data economically transmitted into the centre.

1.5 After suitable processing the output of a plot extractor may be written on a display with each plot being re-freshed at a rate of typically 20 times per second. In this way, a bright flicker-free display is produced enabling radar and procedural controllers to sit side by side in an air traffic control centre, thus easing the communication problem between them.

1.6 Returning to the analogy of the human operator carrying out plot extraction and tracking, it should be observed that in such a system the human operator will subconsciously lower his acceptance threshold in the area in which he expects, on the basis of the previously observed track, to see a target return. Thus he will accept a weak response as confirming the continued presence of the target he was already tracking. In just the same way a tracking computer can be used to control the acceptance thresholds in a plot extractor so as to increase the probability of maintaining a continuous track.

2. PLOT EXTRACTION

2.1 Over the past ten to fifteen years, the basic techniques of digital plot extraction have become relatively well known. Although many different machines have been, and are being made, whose performance varies from one to the next, they are all derived from the same principles. One may perhaps, divide primary extraction into three main areas:-

(a) Hit detection.

(b) Plot detection.

(c) Clutter elimination.

2.2 Hit Detection

From a conventional incoming video containing an abundance of wanted and unwanted signals, a digital representation must be produced. Hopefully, this will contain information about all the wanted signals and nothing of the unwanted ones. The process of quantization is essential for the plot detection following it and may be performed in a variety of ways.

2.3 Perhaps the most common method adopted is to operate with a single amplitude threshold and to accept all levels above this threshold as a logical 'one' and to reject all below as logical 'zero'. Such a threshold would usually be controlled by some type of noise-counting feed-back loop, in order to preserve a constant false alarm rate (CFAR) in the presence of varying thermal noise.

2.4 The video would be better represented by quantizing into more than one state, using several thresholds at different voltages; multi-level quantization. With signals in thermal noise alone, it can be shown that the improvement using conventional plot extraction techniques is limited to around 2dB and since a large increase in hardware is needed to implement it, multi-level quantizing is usually dismissed as not cost effective. However, in clutter areas it is possible that the improvement in detection may be much greater, and with a continuation of the present downward trend of storage cost, multi-level quantizing may yet come into use.

2.5 The amplitude quantized signal so far produced will in most systems also be time quantized by a convenient range clock. At this stage it is convenient to assess the width of each hit, in order to reject those which are much shorter or much longer than the original transmitted pulse.

2.6 In some systems a video may be received which has already been quantized; for example, in certain types of digital MTI. This is, however, merely a shift of hardware within the overall system.

2.7 The information content of the signal emerging from hit detection will have a certain probability of true signal (P_S), and a certain probability of false signal (P_N). Whatever the quantizing system employed, these two quantities will be related to the signal-to-noise ratio of the original video and, in a non-clutter situation, the relationship will usually be readily calculable.

2.8 Plot Detection

Having produced a digital representation of the incoming video, a plot extraction device must store that data for a period corresponding to several single illuminations of a target. This is necessary in order that successive returns may be range correlated to produce the required detection probability and decrease the false alarm rate. There are many ways in which this may be done, each method having slightly different characteristics of its own. Three such methods are conventionally referred to as 'moving window', 'staircase integrator' and 'maximum likelihood estimator'. Some of the methods require data storage for only a few p.r.i.s. (p.r.i. = pulse repetition interval), whereas the 'maximum likelihood estimator' needs somewhat over twice the radar beamwidth of storage. The 'moving window' detector requires the storage of data for one beamwidth and has, for a given false alarm rate, the highest probability of detection of any of these methods.

2.9 The azimuth resolution varies from one detection method to another. In general terms, those detectors which attempt to resolve targets separated by less than the radar beamwidth also produce a high proportion of target splits, due to missing hits. Azimuth accuracy also varies between the methods, those which retain the complete hit pattern, like the 'moving window', providing the best performance.

2.10 At this point, the two common methods of organising storage of hit data should be mentioned. One way is to allow a separate range bin for every possible hit position during a p.r.i., over the complete beamwidth, arranged so that they are addressed by a real-time range word. This is called range-oriented storage. Alternatively, one may retain only the actual hits received by storing their range in any chosen location. When storing a new hit in this fashion, its range must first be compared with others in storage. If the ranges are within some selected tolerance of each other the new hit enters the detector storage at the original address; if they are not within the tolerance, a new address is selected for the new hit. This is frequently referred to as address-oriented storage.

2.11 Address-oriented storage is used in order to save storage but can become saturated if a large number of hits are received. Range-oriented storage cannot saturate and is essential for most types of clutter-elinination.

2.12 Range oriented storage yields range resolution which is limited by the chosen granularity of the range bin (or the radar pulse width, whichever is greater). The range resolution of address-oriented storage is limited by the correlation tolerance. In both cases, the reporting accuracy of range may be much higher than the resolution; typically $\frac{1}{4}$ mile resolution but 1/16 mile accuracy.

2.13 For each application it would be possible to assess the storage and detector costs necessary to achieve the desired extractor performance. However, even with the present cost of storage, a range-oriented 'moving window' system may often provide the most effective extractor. It combines high probability of detection with high range and azimuth accuracy, has adequate range resolution, is non-saturable and has azimuth resolution realistically related to the radar beamwidth.

2.14 After the plot extraction process, there will be a certain probability of real plot detection (P_D) and a certain probability of false plots (P_{FA}). When consideration is taken of targets in termal noise alone, P_D and P_{FA} may be optimised by adjusting the parameters of the hit and plot detectors. For a particular extractor there will be mathematically expressible functions connecting them. Figure 1 shows a graph of (final) P_D against signal-to-noise ratio for a fixed P_{FA} of 10^{-6}, corresponding typically to 2 or 3 false plots per antenna rotation. This figure is for a 'moving window' extractor optimized for a radar producing a theoretical 16 hits per 3dB beamwidth.

2.15 Clutter Elimination

When false signals are present of comparable amplitude to wanted signals and which are not distributed according to any well defined probability function, extra processing must be introduced. Such processing may be applied within the radar receiver with great effect against certain types of clutter. Alternatively or additionally, a further statistically derived function may be introduced into the plot extractor. Such a function would attempt to retain a constant false alarm rate, at the expense of detection probability, by assessing the clutter distribution surrounding the current detection window.

2.16 As mentioned earlier, range-oriented storage permits this assessment by providing a complete area of hit pattern which can be analysed. Systems incorporating this method have achieved considerable success. Figure 2 shows the type of function used to maintain a 'moving window' detector at a constant false plot rate of 10^{-5}. The independent variable is the sum of the hit pattern over a carefully chosen area surrounding the extraction window, and the dependent variable is the leading edge threshold to be applied to that window.

2.17 The limitation of false plots due to clutter is especially important where the extracted data is to be transmitted via a narrow band link, which, by definition, will have a relatively restricted message handling capability. With the display system at the same site the problem may not be so acute although the display system itself will have a finite capacity.

2.18 Secondary Radar Extraction

With secondary radar, the problems are somewhat different. To start with, the hit detector is not concerned to any great extent with thermal noise, but on the other hand, it must recognise time intervals to a high accuracy, so as to extract valid code pulses, whilst taking account of such problems as garble and interleaving. In fact, the 'hit detector' expands to become a system element in its own right, which is common to both automatic and manual secondary radar systems.

2.19 The plot detector also has different tasks to perform. Fruit, i.e. non-synchronous secondary returns, must be eliminated, and the extractor must compare and validate the 12 bit codes associated with each target. Since three mode interlace interrogation is now common practice, storage must be available for a very minimum of 36 extra bits of data with each plot as compared with a primary extractor.

2.20 In general terms, secondary extractors must handle much more real and wanted data than primary extractors for a given target density, but the elimination of unwanted data is simplified. The actual plot detector type is much less important and clutter elimination is unnecessary.

2.21 Combined Extraction

When extracting data from both primary and secondary radars which are co-located and synchronized in rotation and pulse repetition frequency, an extractor may operate on both simultaneously, and correlate in real time the primary and secondary returns from the same aircraft. This will reduce the data to be transmitted to the display site; typically this saving could be as high as 25% of line time.

2.22 There does exist, however, a finite possibility that two returns will not correlate, and this possibility rapidly increases with any separation between the two antennae. With a separation of say 240m, the probability of non-correlation in range is approximately as shown in Figure 3. The law relating the probability of range splitting to azimuth is very nearly sinusoidal, and splits may be reduced by arranging for the trigger of one radar to be delayed by a simple function of the azimuth. Figure 4 shows the resultant decrease in split probability.

2.23 Non correlation in azimuth is not of such great concern, due to the relatively low azimuth resolution of the radars, although it must be borne in mind that even in the perfectly co-located condition, azimuth splits may arise due to the different hit patterns which could be received from the two radars.

2.24 Where the two radars are so separated that azimuth splits become troublesome, or alternatively where the radars are not synchronized, it is necessary to use independent extractors and to perform correlation as a separate function. To retain the saving in data transmission time it is preferable to do this at the radar site, possibly using a hardware correlator. If however, any processing is required at the radar site of a type which is best done using a general purpose computer, then the computer itself may additionally perform the correlation. Alternatively, at the expense of data transmission time, the correlation can be carried out in the display processing computer.

2.25 Track Feedback

If tracking is introduced, there exists a very attractive further extension to the system; dynamic detection criteria control, (DDCC). This process closes the loop between plot detection and tracking and begins to resemble the human operator.

2.26 When a track has been established, the detection criteria in an area surrounding the targets next predicted position are deliberately lowered from their normal setting. This results in a small increase in false plot rate, but yields a large increase in track confidence.

2.27 If tracking and extraction are carried out at different locations the need for reverse data transmission may introduce organizational and storage problems. For this reason it appears likely that future plot extractors will have some limited in-built hardware tracking facilities, even if final tracking is performed by a remote software processor.

3. DATA TRANSMISSION REQUIREMENTS

3.1 Plot extractors are frequently located at radar sites remote from the centres at which the data is required. In systems for air traffic control, now operational or in the planning stage, data is being sent over distances of several hundreds of kilometres. Use is being made of narrow band channels within existing microwave communications links, or in some cases of leased telephone circuits.

3.2 In typical systems data rates of some 2400 bits per second are required to cover the transmission of primary radar extracted targets. With s.s.r. data included, data rates of about twice this value are required. To guard against transmission errors it is usual to provide simple, single dimension parity checking systems. Whilst this affords no possibility of error correction there is little requirement for this, since a revised data package on each target is normally available one antenna rotation later.

3.3 The data transmission rate required can be reduced by carrying out tracking at the radar site. In this way full track data on each target need only be sent occasionally, with track correction data being sent more frequently, at a maximum rate corresponding to once per antenna rotation. However, such a system relies more heavily on the integrity of the data transmission system, so that more sophisticated error detection and error correction techniques must of necessity be introduced. Further, the demands on a tracking computer located remotely from the display centre are much more severe if advantage is to be taken of the possibilities of lower data transmission rates. In such a case fully automatic track initiation is essential, otherwise data, possibly of importance, will be prevented from reaching the display system.

4. DISPLAY PROCESSING REQUIREMENTS

4.1 When we come to the question of displaying plot extracted radar data for air traffic control systems several possibilities exist. In all cases some processing is required. In the simpler cases this can readily be carried out using special purpose processors. For more sophisticated applications stored program computers are more applicable.

4.2 Invariably, there is a requirement to carry out conversion of the position data from polar to cartesian co-ordinates. In the simplest display system this is really the only processing required, in which case the co-ordinate conversion may most conveniently be carried out using special purpose hardware. This can either be implemented within the plot extractor or within the display system.

4.3 The simplest display system may in fact involve no more than receiving plots from the extractor and painting a position symbol on the display at the appropriate position. Using a long persistence phosphor the display is then similar to a conventional radar p.p.i. display. Although there is no longer a rotating spoke on the display, the plots are nevertheless painted more or less in synchronism with the antenna rotation so that a rotating effect is observable. If a background map is required, this may most conveniently be produced synthetically using digitally controlled vector generation techniques.

4.4 In such a system there is no particular problem of display writing speed. Using displays of only moderate performance the time required to paint the arriving plots is quite insignificant, so that a great deal of time can be devoted to writing a synthetic background map. Maps of moderate complexity, comprising typically some 500 vectors can readily be written under these circumstances at a refresh rate of 20 times per second.

4.5 To provide a bright display the plots must also be stored and then written on to the display, preferably at refresh rates of 20 times per second or more. Arrangements can be made for retaining plots from earlier antenna rotations and painting them at lower intensity so as to produce a radar 'trail'. The plots can be labelled with identity and height information using secondary radar data.

4.6 Such a system requires displays having higher writing and recovery speeds. Even so, a compromise has to be reached depending on the parameters of the particular system, so as to apportion the available time between writing plots and writing synthetic maps. In a typical system currently being implemented, a plot capacity of some 1000 plots or trail positions is provided, together with a synthetic map capability of some 300 vectors.

4.7 At the next level of sophistication, tracking can be introduced. In civil air traffic control this is not likely to be implemented except in systems which include plot extraction of secondary radar. The track labels may now be used to provide pertinent flight information for which controllers would otherwise need to refer to flight strips or some other data display.

4.8 In such systems the use of 4096 discrete codes provides a powerful means of automatic track initiation and tracking can be used very effectively on secondary radar only. However, it is well known that the aircraft antenna designer has considerable difficulty in producing an antenna with an isotropic radiation pattern. In consequence secondary radar responses are apt to fade at crucial points in an aircrafts flight. For this, and other reasons, the ability of the system to continue tracking on primary radar data is of the utmost value.

4.9 Systems are in operation using the processes described and more are currently being planned. We may now look forward to a decade in which considerable effort will be put into refining the basic techniques and improving their implementation.

ACKNOWLEDGEMENTS

The authors wish to thank the Directors of the Plessey Company for permission to publish this paper.

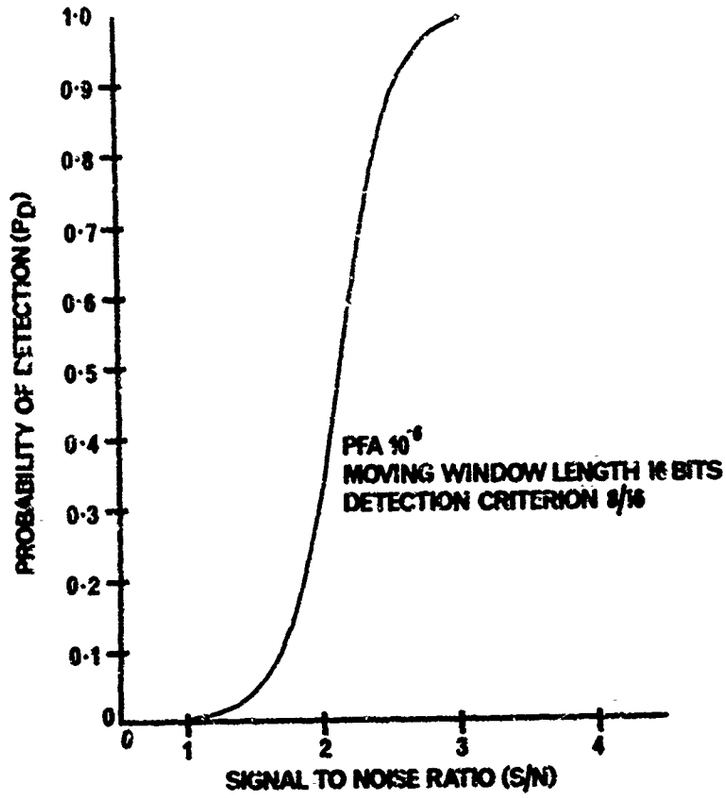


Figure 34.1 Probability of detection versus signal/noise ratio.

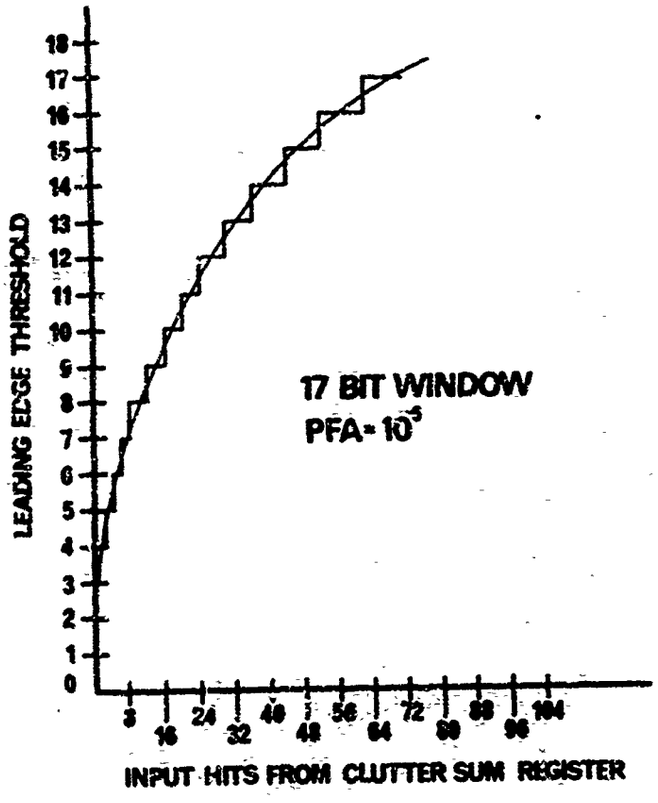


Figure 34.2 Digital function generator curve. 17-bit window. PFA = 10^{-5}

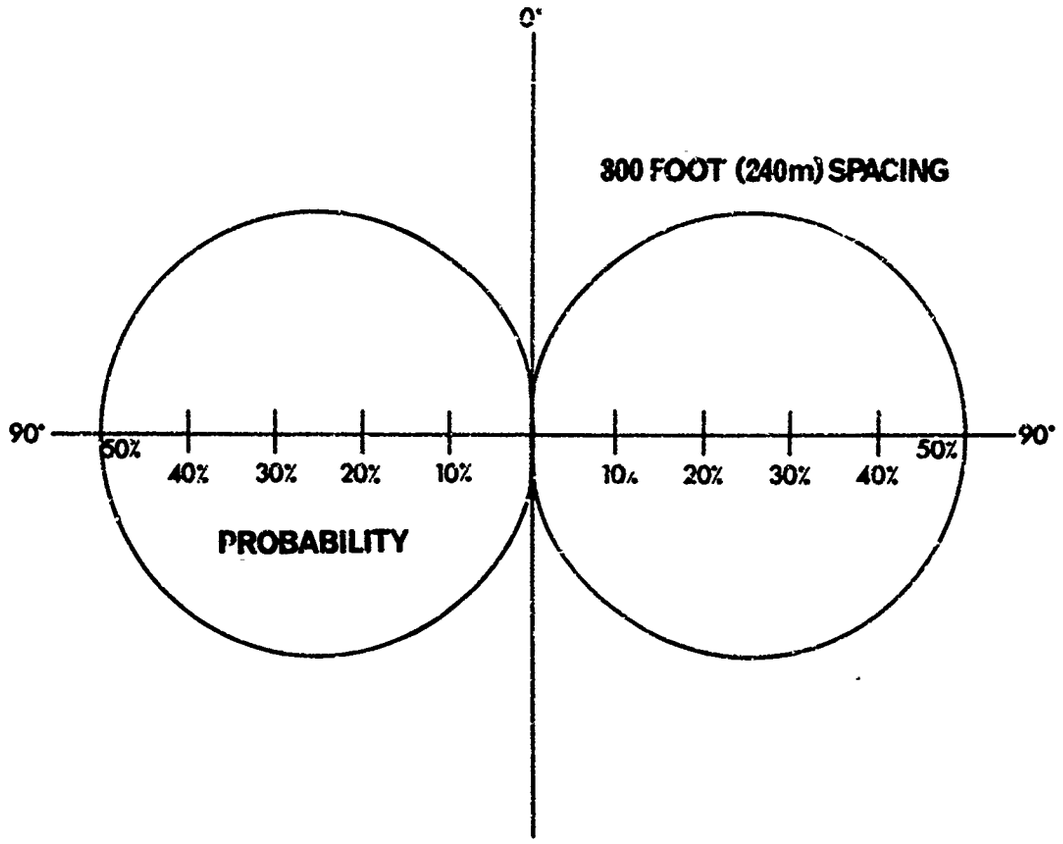


Figure 34.3 Probability of range split at certain azimuth.

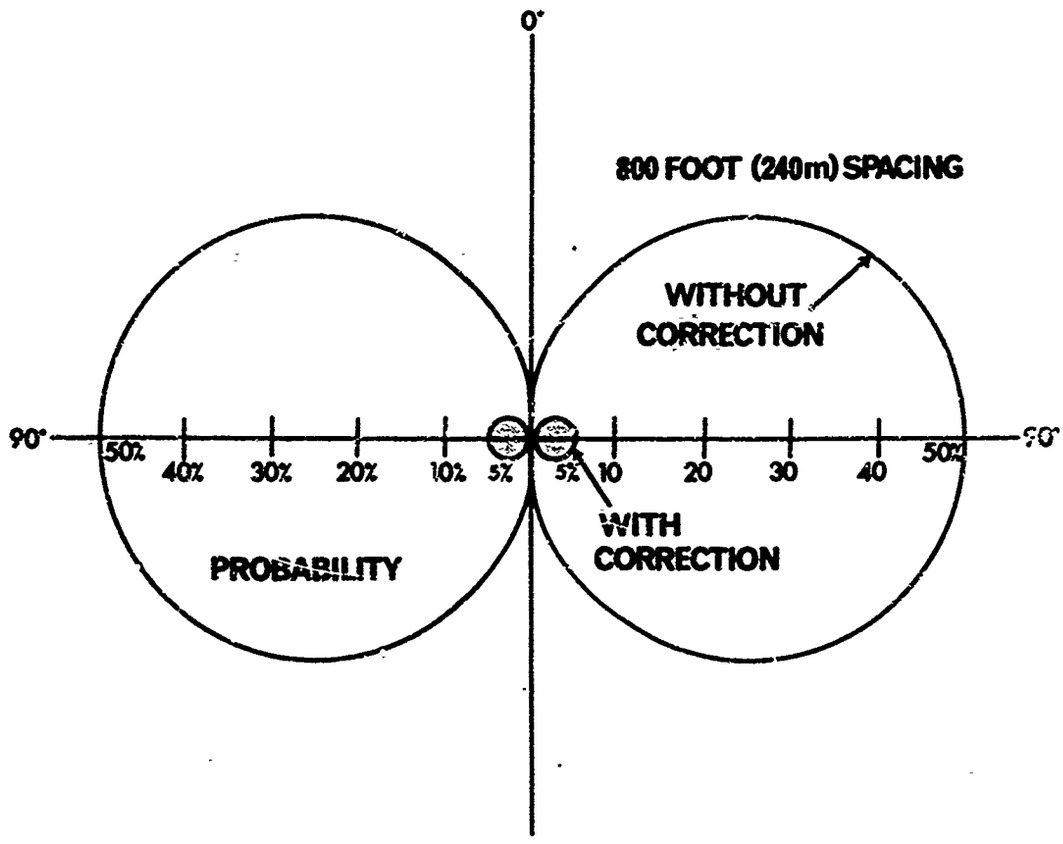


Figure 34.4 Probability of range split at certain azimuth with trigger corrections.

DESCRIPTION OF A TRACKING STRATEGY FOR PHASED ARRAY RADAR

by

G. van Keuk

Forschungsinstitut Werthhoven
Germany

DESCRIPTION OF A TRACKING STRATEGY FOR PHASED ARRAY RADAR

G. VAN KEUK
FORSCHUNGSINSTITUT WERTHHOVEN
DEUTSCHLAND

SUMMARY

In this paper a strategy for simultaneous tracking of many targets by a phased-array-radar system is described. The accuracy of each track is defined by a limiting volume, so that the probability of the target to lie outside vanishes exactly. Errors in location and limiting values for the cinematical behaviour of the targets are taken into account. The expectation value of the total costs for tracking a fixed number of equidistributed objects of one kind is calculated. It depends strongly on the accuracy of the track and the number of stored echoes. Because no smoothing methods are used, we easily can evaluate the density of the distance between the target and the center of the expectation volume. There is a simple relationship between the implied cinematics and the tracking costs, so that the advantage of adaption during the tracking of a target can be derived approximately.

INTRODUCTION

For the control of a phased-array-radar (PAR) system a computer is neccessary, which especially has to start and serve the different phases of processing. So the antenna resp. digitalextractor is linked up with the computer much more, than this is the case with common surveillance radar sets. From this it can be concluded, that the quality of operation depends not only on the hardware, but is strongly influenced by the program-systems brought into action.
By using the stored information from the past, the advantage of a PAR system moreover lies in the possibility, to send the concentrated energy in such regions of space, where observations probably will give much information. For using this reality, suitable strategies and algorithms are to be searched for. While with common surveillance radar all space ranges are scanned for equidistant times, with PAR we can vary the sampling-periods and adapt the decision thresholds to optimize the whole system.
Roughly simplifying we can distinguish three phases of processing

1. The searchmode: In this mode the PAR system is controlled by a special processing-algorithm (see for instance 1), which aims at a most rapid and sufficient secure acception of new flying objects at the border of the range. For optimizing the searchstrategies there are many different criteria depending on the operation of the whole system.
Especially at the border of the range the received signals in this mode consist of true but also of lots of false ones, which are transmitted into storage buffers. A discrimination between true and false signals or a separation of the built up false traces can without additional information only be realized by observations of correlations in space and time. The implied strategies influence the probability of escape, because only a portion of limited depth of the range can be processed during this mode.

2. The mode of trackinitiation: In this mode of processing the discrimination between true and false signals is accomplished and the trackphase is started. In special cases of interest identifications are undertaken and cinematical parameters and intentions are estimated.

3. The trackmode: It is the responsibility of this mode to process all traces, which are controlled in the range of surveillance. Because each additional scan of a target will give rise to costs (time resp. energy), adaptive strategies are to be set up, which minimize the whole tracking costs. We here only consider this phase. A trackphilosophy has to imply a variety of à priori unknown or only partly known facts: manoevres of the tracked objects, to escape for instance,

decision errors, which give finite detection and falsealarm probability, clutter and location errors of statistical or systematic kind. The consequence of these facts can be untimely cancellation of true traces, branching and instability induced by processing false traces.

So, generally, there is no uniform best trackphilosophy, because actually the special assumptions on the above mentioned facts and further requirements are to be involved. If you know, that the possible manoevres of the objects of interest (described by maximum values of accelerations across and parallel to the velocity, Q resp. B) are small, smoothing methods are used. But if you have to expect a high mobility, it may be better to study the cinematical restrictions more exactly . The limits between these two aspects are not sharp, so adaptive branches are part of a more sophisticated tracking logic.

By definition, a trace in the computer consists of a cyclic list of space-time-points of passed locations of one object. For continuing a trace, new positions are required for definite times. To this, estimated positions are to be transmitted to the antenna-system, which essentially define the center of a spatial expanded expectation volume (EV). In this area, possibly, a search with more than one antenna position is started for localizatio. of the expected object. The membership of datas (r, ϑ, φ, t) to a given trace is defined by belonging to the joined EV.

We now come to assumptions defining our model: Let the characteristic of the search-pencilbeam be nearly square with a width of $\phi = \pm$ 0.5 dgr. in ϑ and φ. In the case of a detection a measuring of ϑ and φ up to $\varepsilon = \pm$ 0.1 dgr. in ϑ and φ takes place. We assume a quantization of the distance r to the radar-center of $\Delta r = 150$ m and take a range of 150 km . For numerical results we take a meanvalue of the decision length (target or no target) at one antenna position of $\tau = 5$ ms. γ denotes the time spent for position measuring. We set $\gamma = \tau$.

Further we assume, that only one accuracy for all tracked targets is required, although this is immaterial to our algorithm. For the maximum values B, Q we take 2.5 resp. 10 m/sec^2. We started from tracking-logics for Monoradar[2] and Multiradar[3] as they are developed and studied theoretically[4].

EXPECTATION VOLUMES

Let $t_1 > t_2 > \ldots > t_N$ be a sequence of scan-times for one target and let $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N$ denote the joined received positions. Each localization $\vec{r}_i$ will have errors $\vec{e}_i$ with densities, that are independent of i and, as an assumption, vanish outside a sphere of a certain radius ρ. Consider the set of all cinematical possible curves $\vec{r}(t)$ with accelerations $\vec{q}(t)$ and $\vec{b}(t)$ across resp. parallel to the actual velocity and $q(t) \leq Q$ resp. $b(t) \leq B$ for all t, so that the locations $\vec{r}_i$ in spacetime could be possible. If we call these curves permissible, $\bar{V}(t; t_1, \ldots, t_N)$ is defined as the pointset of all possible endpoints of permissible curves $\vec{r}(t)$ at time t. By listing the times t_i behind the symbol $\bar{V}$ we express the fact, that these locations are taken into account constructing the expectation area $\bar{V}$. We therefore have

$$\bar{V}(t; t_1, \ldots, t_N) \subset \bigcap_{i \neq k} \bar{V}(t; t_i, t_k)$$

Let V be the volume of the pointset $\bar{V}$, then it follows

$$V(t; t_1, \ldots, t_N) \leq \min_{i \neq k} V(t; t_i, t_k)$$

So, for an estimation, we are only to calculate the EV basing on two locations in past. By the minimum condition the EV depends on all locations in the course of time.

In fact, the pure cinematical EV basing on two locations is difficult to calculate exactly[5] and is distorted by the position errors $\vec{e}$ drastically. But if the time distances of scanning are short we get approximately a cylindrical area with radius L and depth l by expansion for small angles.

$$L = \tfrac{1}{2} \cdot Q \cdot (t - t_1) \cdot (t - t_2)$$

$$l = B \cdot (t - t_1) \cdot (t - t_2)$$

So we get for the pure cinematical portion of the EV

$$V_{cin} = \frac{\pi}{4} BQ^2 (t - t_1)^3 \cdot (t - t_2)^3 \quad .$$

We now consider the errors of location by a worst-case estimation. First we state, that the volume of location is cylindrical too. To get rid of direction depending errors, we circumscribe the volume of location by a sphere of radius d in the distance r

$$d = \left[(r \cdot \hat{\varepsilon})^2 + \left(\frac{\Delta r}{2}\right)^2 \right]^{1/2}$$

Let us now assume that the error-vector $\hat{\varepsilon}$ is equidistributed in this sphere, as this would be the case for devictions that stem from quantization. To consider the error-propagation we have to put a layer around the cinematical EV of a depth

$$d \; \frac{2t - t_1 - t_2}{t_1 - t_2}$$

and therefore get

$$\begin{aligned} V(t;t_1,t_2) = {} & \frac{\pi}{4} BQ^2 (t - t_1)^3 \cdot (t - t_2)^3 \\ & + \pi Q (B + Q/2) d (t - t_1)^2 \cdot (t - t_2)^2 \cdot \frac{2t - t_1 - t_2}{t_1 - t_2} \\ & + \Delta V \cdot \left(\frac{2t - t_1 - t_2}{t_1 - t_2} \right)^3 \end{aligned}$$

with $\Delta V = \frac{4\pi}{3} d^3$.

In the limiting case of small accelerations, V reduces to the stretched volume of location because we don't use any smoothing methods. For small d or large B,Q-values the cinematical portion of the EV dominates. Because we started from an expansion for small angles, our EV is independent of velocities.

Let us go over to equidistant scans in time and let be $t = 0;\ t_1 = -mT;\ t_2 = -nT$, then we get

$$V(T;m,n) = An^3m^3T^6 + Bn^2m^2 \; \frac{n+m}{n-m} \; T^4 + \Delta V \left(\frac{n+m}{n-m}\right)^3$$

$$A = \frac{\pi}{4} BQ^2 \; ; \; B = \pi Q (B + Q/2) d \; ; \; n > m \geq 1 \; .$$

For given values of T,B,Q and r we get m,n from the minimum condition

$$V(T;m,n) \longrightarrow \min$$

The result is unique: $m = 1$ and $n = n^o(T, B, Q, r)$. For small sampling periods T and especially for great distances r we have $n^o > 2$. Really n is limited by the length of our list of stored plots.

By selecting an optimal n resp. choosing an appropriate set of locations from the list, we balance the error-propagation against the increase of the EV by manoevres. For determining n^o see the result on Fig. 36-1.

The so calculated EV is expanded around the center

$$\vec{r}_t = \vec{r}_1 + \frac{\vec{r}_1 - \vec{r}_n}{t_1 - t_n} \; (t - t_1)$$

and contains all endpoints of permissible curves.

DENSITIES IN THE EXPECTATION VOLUME

We consider a permissible curve with B = 0 and fixed $\vec{q}(t)$. Let A denote the distance between the true target position and the center of the EV. Then we are interested in the normalised density G(A).

If $(\vec{r}_1, t_1)$; $(\vec{r}_2, t_2)$ for $t_1 > t_2$ are the true positions of the object, the measured positions will have the form $(\vec{r}_i + \vec{e}_i, t)$; i = 1,2 with independent error-vectors which are equidistributed in a sphere of radius d.

In our model we get

$$\vec{A} = \vec{r}_1 + \vec{e}_1 + \frac{t - t_1}{t_1 - t_2} \; (\vec{r}_1 + \vec{e}_1 - \vec{r}_2 - \vec{e}_2) - \vec{r}(t; t_1, t_2)$$

$$\vec{r}\,(t; t_1, t_2) = \vec{r}_1 + \frac{t - t_1}{t_1 - t_2} \; (\vec{r}_1 - \vec{r}_2) + \frac{1}{2} \, q \, (t - t_1) \cdot (t - t_2) \cdot \vec{f}$$

$$|\vec{f}| = 1 \, .$$

So the vector

$$\vec{A} = \vec{e}_1 + D \, (\vec{e}_1 + \vec{e}_2) + \Sigma \vec{f}$$

with $$D = \frac{t - t_1}{t_1 - t_2} \; ; \; \Sigma = \frac{1}{2} \, q \, (t - t_1) \cdot (t - t_2) \, .$$

is a sum of independent random vectors. So we can approximately write down a normal distribution following the central limit theorem. After calculating the covariances and transforming to the absolutevalue A, we get:

$$G(A) = \frac{1}{\sqrt{2\pi}} \cdot \frac{A}{d \, \sigma \, \Sigma} \; \exp\left[- \frac{1}{d^2 \sigma^2} (A - \Sigma)^2 \right] \cdot \left(1 - \exp\left[- \frac{2 A \Sigma}{d^2 \sigma^2} \right]\right)$$

where σ^2 is equal to $\frac{2}{5} \cdot (\frac{1}{2} + D^2 + D)$.

Let us denote the second time of scan by t = -nT, then we get in good approximation

$$\sigma \cong \sqrt{\frac{1}{5}} \; \frac{n}{n-1}$$

$$\Sigma \cong \frac{1}{2} \cdot Q T^2 n$$

The distance from the center of the EV to the border is

$$\rho = \frac{1}{2} \cdot Q T^2 n + d \, \frac{n+1}{n-1}$$

Generally the density depends on a variety of parameters. In fact, the density of A/ρ varies only slightly, which is a consequence of balancing the error-portion and the cinematical part of the EV. We get approximately

$$\rho \cong d \cdot \frac{n^o + 5}{n^o - 1}$$

if n^o denotes the optimal chosen n. In the limiting case of no acceleration (q = 0), the density G (A) takes on a maximum at

$$A_o \cong d \sqrt{\frac{2}{5}} \cdot \frac{n^o}{n^o - 1}$$

So we see, that A_o/ρ does not depend strongly on different parameters and A_o/ρ takes on a value of .3 in moderate distances r. As is to be assumed, the error-portion of the EV is after balancing strongly apparent. If we calculate the maximum of the density for $q = Q$, which was the value to generate the EV, we get nearly $A = \sum$. If this value is compared with A_o it follows

$$\sum / A_o \approx 7/n^o$$

for large $n^o > 2$, for instance at the border of the range or for small distances r the influence of the cinematics is diffuse. In moderate ranges we have $\sum / A_o \approx 2$.

This behaviour is shown by simulations. Fig. 36-2 demonstrates some typical densities G.

We conclude, that without smoothing no relevant cutoff of the EV can be obtained because the densities reach at the border of the EV.

FORMULATION OF THE ALGORITHM

The relation of received plots to processed traces is defined by the EV. Being comparatively large, these EV may contain more than one signal, so that branching cannot be excluded (creation of parasites). False branches either diverge from the true trace or - in most cases - will converge and be absorbed after some time. A branch or a trace is cancelled, if no signal can be found after defining appropriate decision thresholds. We now study the scan-period.
Higher scan-rates/trace generally will give a better track-accuracy. We introduce a free parameter V_o, which describes the maximum allowed inaccuracy of a track in the course of time. Our expectation volumes are upper limits of inaccuracy and increase with the time t. Not to cross the limiting volume V_o, the next scan is required, if the actual volume of the joined EV is equal to - or larger than - V_o. This algorithm leads to local equidistant scans in time with a period $T_o(n, V_o, r, B, Q)$ independent of angles.
At the scan-times the EV possibly decomposits in more than one of elementary cells because of the limited width of the pencil-beam, whose processing will give additional costs. For calculating the number of cells we circumscribe the EV by a sphere of radius $\rho(T,n,r)$. For all cases of interest we have Q definitely larger than B, so we get

$$\rho^2(T,n,r) = \frac{Q^2+B^2}{4} \cdot n^2T^4 + QT^2dn\,\frac{n+1}{n-1} + d^2\left(\frac{n+1}{n-1}\right)^2$$

where n indicates the second used data set in the list of the considered trace. If we call δ the local width of the pencil-beam, $\delta = \bar{\phi}\, r$, the number of cells approximately can be expressed as

$$F(T,n,r) = \max\left(1,\ 1+\frac{\pi}{2}\left(\rho/\delta\right)^2\right)$$

It is not very effective to use directional information here.

Now we change over to the local track-costs for one target in the distance r. Let <Tr.local> be the fraction of the unit of time, that-in the sense of a timeaverage-is used for tracking a target in fixed distance r. We have

$$\langle Tr.local\rangle = \frac{V + \tau \cdot F(T,n,r)}{T}$$

where $T = T_o(n, V_o, r, B, Q)$ is the really tuned scan-period. We get n and T from the following minimum condition

$$\langle Tr.local\rangle \longrightarrow \min$$

It follows, that for large distances r from the radar center, T will be equal to T_o, while for moderate and especially small r, or for big volumes V_o, T comes out smaller than T_o, to minimize the local track-costs. Differently expressed,

V_o is not reached in all parts of the range. $\langle$ Tr./local $\rangle \approx 1$ indicates an instabil mode of operation induced by too high an accuracy. V_o can be chosen differently for each trace and can be varied in the course of time. Let us concede, that too pessimistic an assumption of the costs for searching the elementary cells is implied, because there exist strategies to find a known number of targets in a limited volume of space, that manage with smaller costs - work in this direction is not completed. So we get here as an result, that nearly all over the range we have only one cell: $l = 1$.

At the end we list some consequences of the above algorithm

1. For given values of V_o, r the optimal chosen n does in very good approximation not depend on B and Q.
2. For given values of V_o, r the optimal scan-period T as a function of the accelerations B,Q varies as

$$T(\lambda B, \lambda Q) = \frac{1}{\sqrt{\lambda}} T(B,Q)$$

3. The local track-costs f in good approximation follow the relation

$$f(\lambda B, \lambda Q) = \sqrt{\lambda} f(B,Q).$$

4. The local scan-algorithm is stabil against fluctuations in the scan-times and converges in a few seconds to equidistant scans in time after changing the V_o or B,Q parameters.
5. For the time-average of the track-inaccuracy we get with

$$\langle V \rangle_{n,T,V_o} = \frac{1}{T} \int_0^T dt\ V_{n,V_o}(t)$$

$$\langle V \rangle_{n,T,V_o} \leq \frac{1}{5} V_o + \frac{4}{5} \cdot \Delta V \left(\frac{n+1}{n-1}\right)^3 .$$

GLOBAL TRACK-COSTS

For the sake of simplicity we consider a surveillance area of range R, heigh h (20 km for numerical results) and of points with $r \geq h$. Let p (r) denote the density of objects to be tracked in a distance r

$$\int_h^R r dr \cdot p(r) = 1$$

For a given number N of independently tracked objects the integrated track-costs are

$$\langle \text{Tr./sec} \rangle = N \cdot \int_h^R r dr p(r) \min_{\substack{T \leq T_o \\ n \geq 2}} \frac{\gamma + \tau F}{T}$$

with the above defined expression T and F, where T obeys the equation

$$V_o = An^3 T_o^6 + Bn^2 \left(\frac{n+1}{n-1}\right) T_o^4 + \Delta V \left(\frac{n+1}{n-1}\right)^3 .$$

Here we have implied, that the whole system operates in the stability region : $\langle$ Tr./sec $\rangle$ definitely smaller than 1, so that tracking of different objects works independently. If we only state, that N traces are to be continued, the time-average of the costs can be evaluated by assuming that these objects fly randomly. So we get $p(r) = 2/(R^2 - h^2)$ by phase-space arguments :

$$\langle \text{Tr./sec} \rangle = \frac{2N}{R^2 - h^2} \cdot \int_h^R r dr \min_{\substack{T \leq T_o \\ n \geq 2}} \frac{\gamma + \tau F}{T}$$

For given values of B,Q and known behaviour of the position-noise the global track-costs as a linear function of N depend on the limiting volume V_o and the maximum number of stored plots. Figure 36-3 shows this set of curves for a fixed number of N = 50. From the asymptotic behaviour of the costs for large V_o we get the smallest possible costs for 50 equidistributed objects. Because V_o has to be larger than

$$\Delta V \left(\frac{n+1}{n-1} \right)^3$$

the curves consist of several branches.

ADAPTION

After finished classification resp. for other reasons an adaption can be obtained by associating appropriate B,Q-values to different traces. By distributing special V_o parameters to different targets according to their importance an adaption can be reached too. In these cases our algorithm needs no modification, because V_o and B,Q can be changed. Following the above result, we see that by reducing the B,Q-values by a factor λ the costs for tracking will decrease by at most $\sqrt{\lambda}$.

Without discrimination between classes of flying objects, more interesting are adaptive strategies for objects with large mobility. We have described the motion of a target by means of a Markov-process and the construction of the joined trace in the computer by methods of Markov-chains. So we got estimations for the advantage of adaption by studying the ergodic properties of the underlying chain. For discrimination between several manoevres, sequential decision procedures of high speed are available, whose OC-function partly defines the costfunction.

For the special case of only two classes of mobility G,K with Q = 5 resp. Q = 20 m/sec^2 (B = Q/4) Tab. 36-6 shows approximative results for the save of costs for one target in the course of time. The results depend on the average duration of state $\langle G \rangle$ and $\langle K \rangle$.

REFERENCES

1. Posner, E.C., Rumsey, H. Continuous sequential decision in the presence of finite number of hypotheses IEE Trans. of Inform. Th. Vol. IT-12,2 , 1966
2. V.Issendorff, H. Organisation und Strategie eines automatischen Monoradar-Beobachtungssystems. Elektronische Datenverarbeitung 8, 1968
3. Pandikow, W. Strategie und Organisation der Multiradar-Beobachtung Elektronische Datenverarbeitung 9, 1968
4. Hoenerloh, H. Vergleich zwischen Prozeßmodellen und simulierten Verfolgungsprozessen mit echten und falschen Spuren. Interner Bericht Nr. 120
5. Fleckner, H., Hoenerloh, H. Erwartungsgebiete für die automatische Flugobjekt-Beobachtung. Elektronische Datenverarbeitung 7, 1968

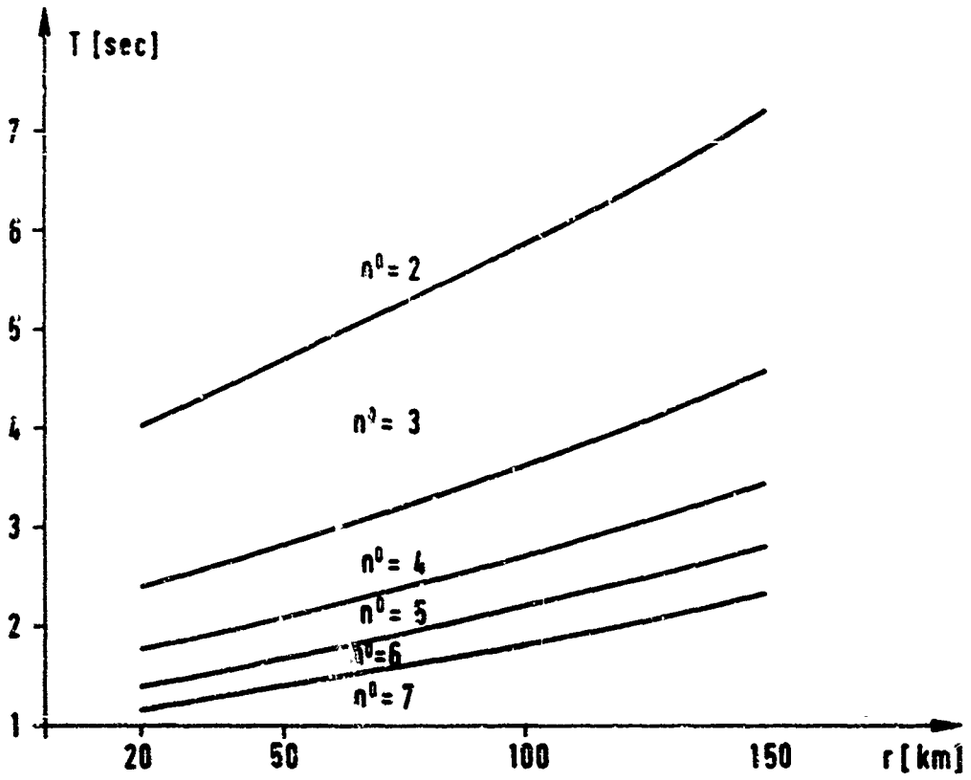


Fig. 1 Regions of optimal chosen n^o depending on T and r.

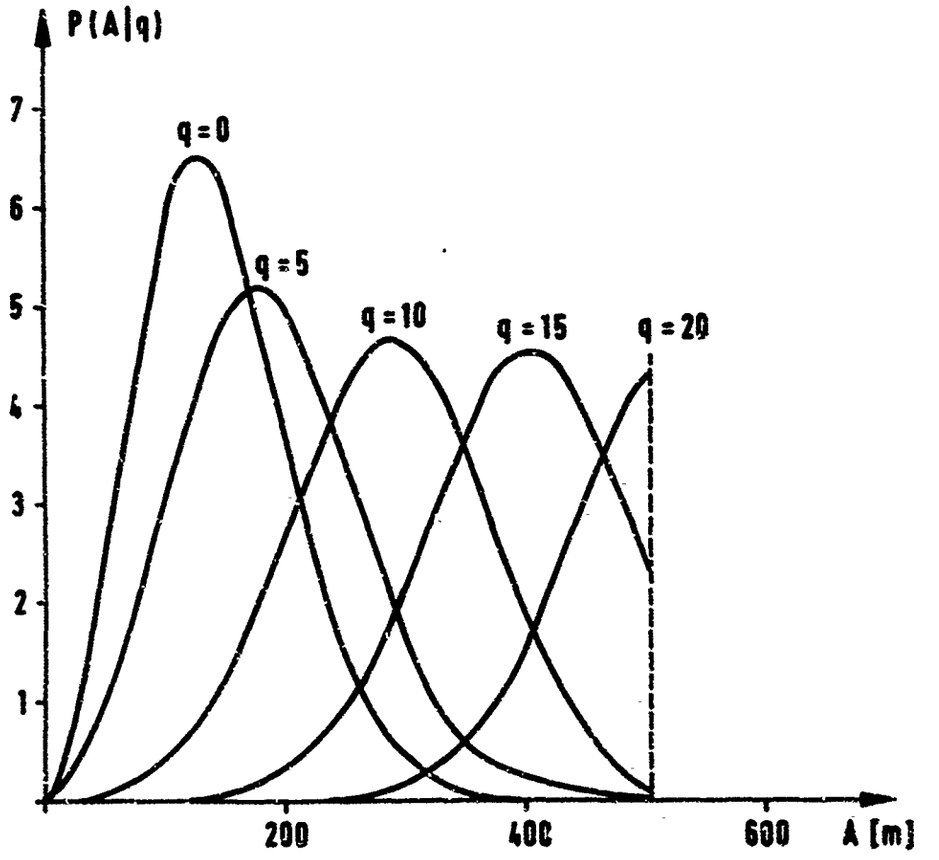


Fig. 2 Densities in the EV for different accelerations q in m/sec^2 for r = 60 km, T = 4.1 sec, $n^o = 3$.

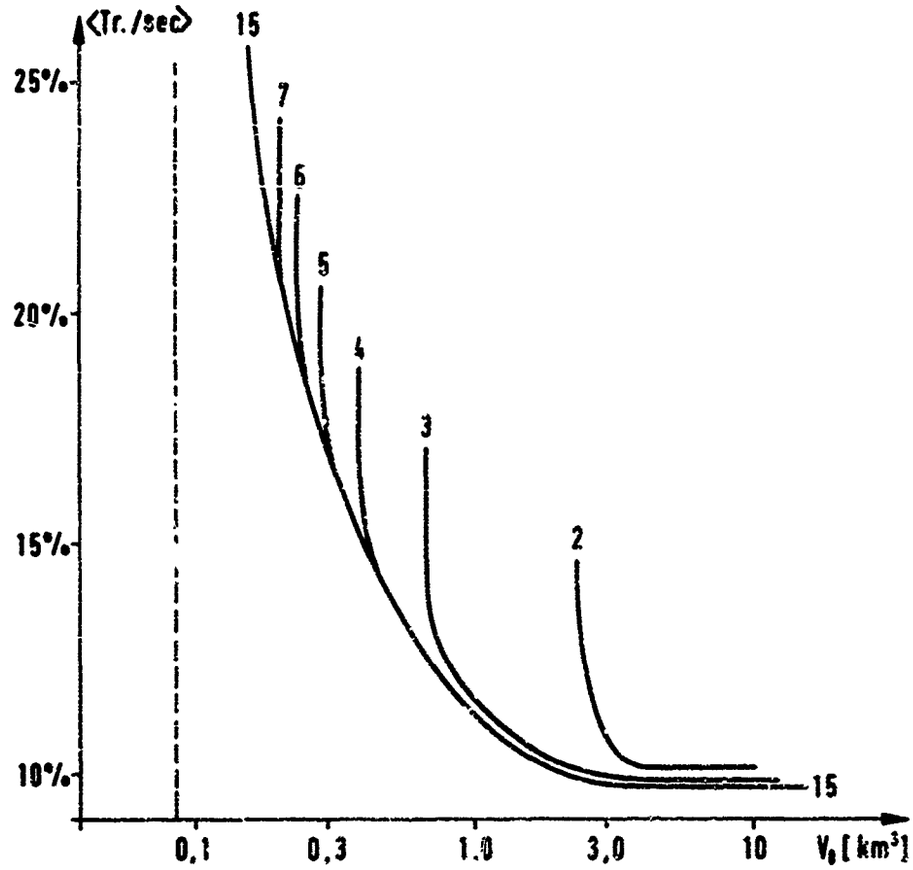


Fig. 3 Track-costs for 50 objects depending on the number of stored echoes and V_o.

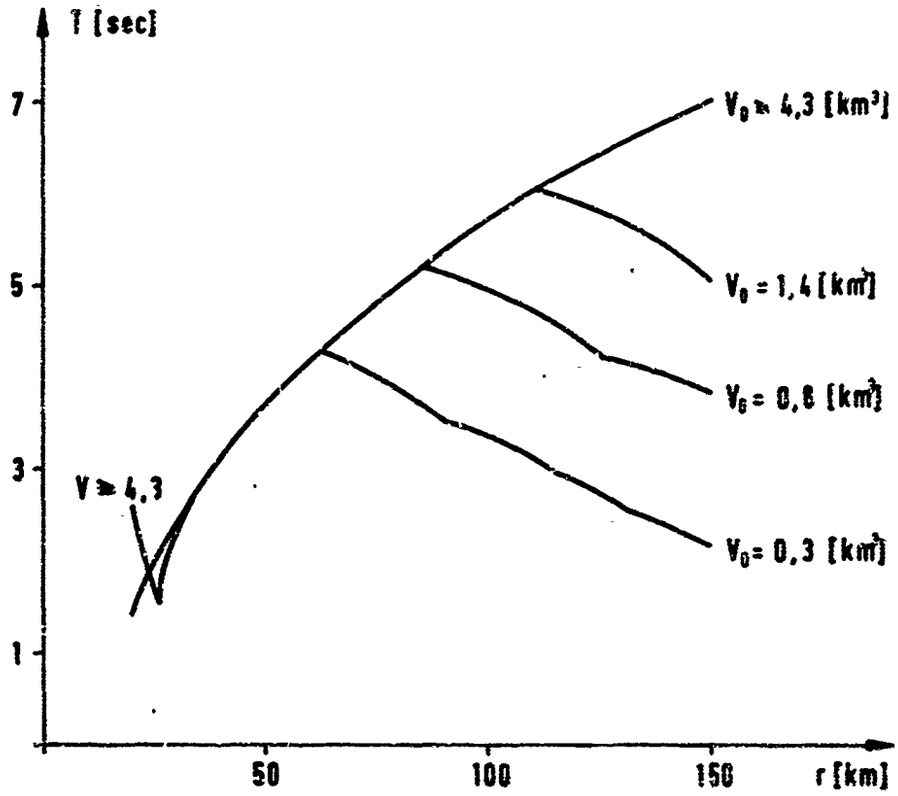


Fig. 4 Scan-period for different limiting volumes V_o.

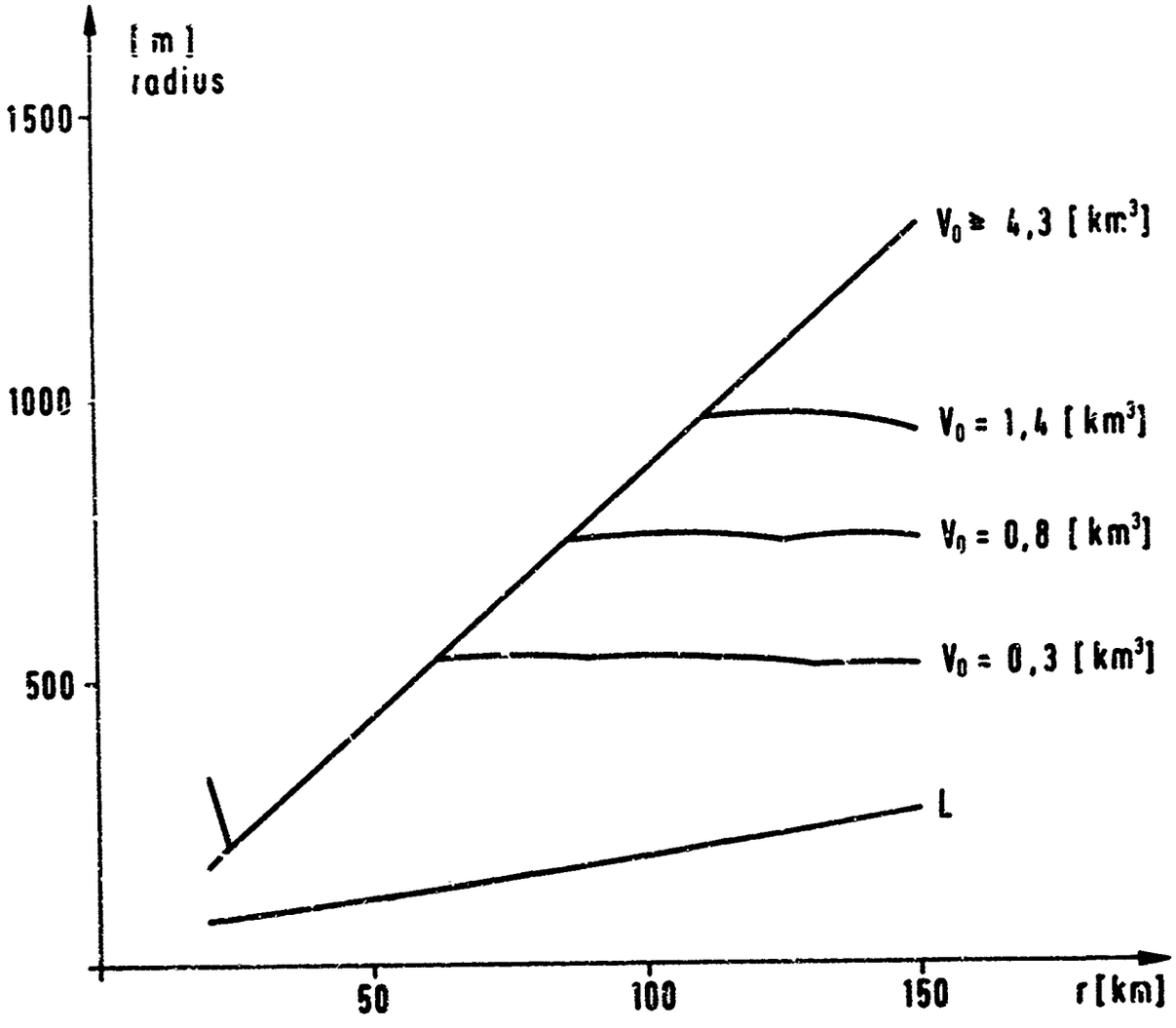


Fig. 5 Radius of location (L) and radius of circumscribed sphere for different V_0.

G \ K	30 s	60 s	90 s	120 s	150 s	180 s	210 s	250 s	300 s
30 s	- 2 %	- 2 %	- 1 %	0 %	0 %	0 %	0 %	0 %	0 %
60 s	13 %	10 %	8 %	6 %	5 %	5 %	4 %	4 %	3 %
90 s	21 %	17 %	14 %	11 %	10 %	10 %	8 %	7 %	6 %
120 s	27 %	22 %	19 %	17 %	15 %	13 %	12 %	11 %	9 %
150 s	30 %	26 %	23 %	20 %	18 %	17 %	15 %	14 %	12 %
180 s	33 %	29 %	26 %	23 %	21 %	19 %	18 %	16 %	14 %
210 s	35 %	31 %	28 %	25 %	23 %	22 %	20 %	18 %	16 %
250 s	37 %	34 %	31 %	28 %	26 %	24 %	23 %	21 %	19 %
300 s	39 %	36 %	33 %	31 %	29 %	27 %	25 %	23 %	21 %

Tab. 6 Estimated save of costs by adaption depending on the average time of state in sec.

SOME STATISTICAL RESULTS COMPARING DIFFERENT TRACKING LOGICS

by

W.A.Both and H.Fleckner

Forschungsinstitut für Funk und Mathematik
5307 Wachtberg-Werthhoven, Germany

SOME STATISTICAL RESULTS COMPARING DIFFERENT TRACKING LOGICS

W.A. Both and H. Fleckner
FORSCHUNGSINSTITUT FÜR FUNK UND MATHEMATIK
5307 Wachtberg-Werthhoven, Germany

SUMMARY

The idea of a radar-digitalextractor-computer controlled air space resulted until now in a variety of track-operators,which all have the common objective to approximate the real air traffic situation by an adequate track-net. In the sequel two sensible functions are given which in a certain sense measure the quality of a track-operator's approximation of a real net. Furthermore, some knowledge about real- and false-track-processes concerning a special track-operator is gained. As an example the functions are applied to some track-operators based on primary radar information, which have been developed in the "FFM, Werthhoven".

INTRODUCTION.

The automatic air space control works with complex scanning- and computer evaluation-systems, that are generally based on the following principles:

A radar digital extractor system (RDS),/7/, scans in discrete time points the momentaneous air traffic situation and gives this information as coordinates in a space-time point-grid to a computer processing system. This system successively evaluates information, thus constructing a more or less good track-net as a mapping of the actual air traffic situation. This net, stored in the computer, is steadily updated and forms the base for a simultaneously working automatic flight-control and -guidance. The quality of the track-net depends on the RDS's possibilities as well as on the programmed track logic (/1/, /2/, /5/, /6/, /8/). In this paper we will give a sensitive measure for its quality, this measure being realized by two functions. Therefore we first define in 1. the assumptions for the basic structure of an ideally working track operator by the properties of the mapping given by the operator. These are generalized in 2. in a natural way to the real case (RDS-errors!). The properties of the mappings constitute ideal cases, which are only to be approximated in realiter by track logics (program systems). The need for measuring the approximation's errors leads to the definition of adequate measures in 3. For other measures see /4/, /5/. In the following section we develope our ideas about the creation of a mastertape-track-net, gained from real radar data of a digital extractor (primary data), see again /4/, /5/. Finally we give some evaluation results, which are shortly interpreted. An appendix contains some characte-

ristics of the evaluated mastertapes as well as a summary of the used shortenings.

1. THE IDEAL CASE. Assumptions and definitions.

We assume, that there exists an ideal scanning system ATS being able to give the three dimensional coordinates of all flying objects, moving in the ATS-controlled area ST, at an arbitrary time during a finite time interval T. By subdividing the time interval into discreet parts together with the assumptions (V1), (V2), (V3), we formalize the above mentioned facts.

(V1) A given time interval T is subdivided into aquidistant points $t_1, t_2, \ldots, t_m$, $m \in \mathbb{N}$.

(V2) Over T there exists a by n, $n \in \mathbb{N}$, flying objects FO_j generated air traffic situation S(t), changing with time, $j \in \{1, 2, \ldots, n\}$, $t \in T$.

(V3) For all t and all j the coordinates a_{tj} of FO_j are known, the a_{tj} being zero if and only if FO_j is not in the controlled air space of ST; $t \in T$, $j \in \{1, 2, \ldots, n\}$, $a_{tj} \in \mathbb{R}^3$.

Hence, the curves $FB_j = FB_j(t)$ along which the flying objects move are known, i.e. each curve can by interpolated by a polygon. Thus the vector FS_j,

$$FS_j = (a_{j1}, \ldots, a_{jm}),$$

constitutes a special interpolation polygon of the curve FB_j with respect to the support points (V1). Let $S_i := S(t_i) \subset \mathbb{R}^3$. An aquivalent definition is:

$$S_i := \{P_i \mid \forall j,\ j \in \{1, 2, \ldots, n\}: P_i = FB_j(t_i)\},\ i = 1, 2, \ldots, m,$$

i.e. for each element $P \in P_i$ there exists one and only one flying object, (V2), so that P is its scanning coordinate at the time t_i (ideal scanning). To identify the individual FO_j or its scanning coordinates, we introduce the identification K of FO_j in def. 1.

Def. 1. Let $P_i \in S_i$.

$K(P_i) := j :\Leftrightarrow: P_i = FB_j(t_i)$

A procedure, being able to correlate the t-scans S_i pointwise in such a way, that in accordance with the interpolation FS_j the track-net (set of all FS_j) over T is reproduced, is called a (semi-real) track operator. Is this procedure interpreted as an unequivocal mapping f, well defined on S_i, then we have:

$$S_i \ni P_i \xrightarrow{f} P_{i+1} \in S_{i+1}.$$

Def. 2. A polygon $PG = (P_{i_1}, \ldots, P_{i_r})$ is called a "track", if (S1) and (S2) are valid for PG.

(S1) $f(P_{i_r}) = \phi$, $f^{-1}(P_{i_1}) = \phi$

(S2) $f(P_{i_j}) = P_{i_{j+1}}$, $j = 1, 2, \ldots, r-1$.

P_{i_1} is called "track-germ", P_{i_r} "track-end" and $L(PG) := r$ length of the polygon PG.

Def. 3. f is called "track-operator".

Def. 4. f is called "semi-real", if (I1) and (I2) are valid.

(I1) $PG = (P_{i_1}, \ldots, P_{i_r})$ track (see def. 2.)

$\Rightarrow \dot{\forall} j,\ j \in \{1, 2, \ldots, n\} \ /\ FS_j = (o, \ldots, o, P_{i_1}, \ldots, P_{i_r}, o, \ldots, o)$

(I2) $j \in \{1, 2, \ldots, n\}$

$\Rightarrow \dot{\forall}$ PG, PG track / PG $\equiv FS_j$ (see (I1)).

(I1), (I2) guarantee for each flying object the existence of exactly one track PG=FS_j and vice versa. Consider the imbedding of PG into FS_j being implied (norming of the dimension).

2. GENERALIZATION.

In the following we will investigate the loss of scanning information compared with ST by an inexact ATS: RDS (f.e. radar detection system). Consider the set system

$$\{\tilde{S}(t_1), \ldots, \tilde{S}(t_m)\}.$$

Again let $\tilde{S}_i := \tilde{S}(t_i) \subset \mathbb{R}^3,\ i = 1, 2, \ldots, m.$

Let $\tilde{S}_i$ be the traffic situation as reproduced by an inexact technical device at the time t_i. In general we have

$$\tilde{S}_i \neq S_i.$$

The probability for $P_i \in S_i \cap \tilde{S}_i$ is called PE (detection prob.).

$$PE = W\ (P_i \in S_i \cap \tilde{S}_i).$$

The probability for $\tilde{P}_i \notin S_i \cap \tilde{S}_i$ is called PF (error prob., i.e. the probability for a false signal).

$$PF = W\ (\tilde{P}_i \notin S_i \cap \tilde{S}_i).$$

The assumptions for the existence of a semi-real track-operator (No.1) are no longer valid in the generalized system, no complete information about the air traffic situation to the adequate interpolation points being given and false signals coming up. This neccessarily leads to a certain weighting of the $\tilde{P}_i \in \tilde{S}_i$.

Def. 5. (a) $\tilde{P}_i \in \tilde{S}_i$ is called "real" (real-signal) if and only if $\tilde{P}_i \in S_i \cap \tilde{S}_i$.

(b) $\tilde{P}_i \in \tilde{S}_i$ is called "false" (false-signal) if and only if $\tilde{P}_i \notin S_i \cap \tilde{S}_i$.

In a generalized analogy to the definition of a track operator (semi-real, No. 1) we have to consider a pair of mappings $\theta = (f, \tilde{f})$, where $\tilde{f}$ is to reproduce the lost information in such a way, that $\tilde{f}$, linked to f, is able to give the above mentioned correlation of points. In this sense let us now consider the pair of mappings $\theta = (f, \tilde{f})$ with

$$\tilde{f} : \tilde{S}_i \to S_i,\ i = 1, 2, \ldots, m.$$

$$f : \tilde{f}\ (\tilde{S}_{i-1}) \to S_i,\ i = 2, 3, \ldots, m.$$

For completeness we give in Def. 6., according to Def. 2. and Def. 4., the analogue definition for r-trucks (real-tracks).

Def. 6. (a) A polygon $\tilde{PG} = (\tilde{P}_{i_1}, \ldots, \tilde{P}_{i_r})$ is called "r-track" (real-track), if (S11), (S22), (S33).

(S11) $f^{-1}\ (\tilde{P}_{i_1}) = \phi,\ f(P_{i_r}) = \phi$ (see (S1)).

(S22) $f\ (\tilde{P}_{i_j}) = \tilde{P}_{i_{j+1}},\ j = 1, 2, \ldots, r-1.$

(S33) $\tilde{P}_{i_j} \in \tilde{f}\ (\tilde{S}_{i_j}).\ i_j \in \{i_1, \ldots, i_r\} \subset \{1, 2, \ldots, m\}.$

Def. 6. (b) θ is called "real track-operator" if and only if (R1), (R2) are valid for θ.

(R1) $\tilde{f}(\tilde{S}_i) = S_i$, $i = 1, 2, \ldots, m$.

(R2) f semi-real.

Hence, a real track-operator in the sense of the definition has the following properties:

(R11) It suppresses the false signals, occuring in $\tilde{S}_i$ because of PF>0.

(R12) It compensates the in $\tilde{S}_i$ because of PE<1 missing signals.

(R21) It gives the real tracks of ST.

The problem is to approximate the operator θ under the given assumptions. In its usual meaning a track-operator (track logic) always constitutes a more or less qualified approximation of θ by means of a real time processed program system. Any approximation of θ must of course make use of information of the curve's history, if it exists. This is done in the so called prediction part, which is of eminent importance for any approximation, what so ever. The inner structure of prediction parts (/1/, /2/, /3/, /6/) is not investigated, since their parameters don't figure explicitly in the quality measuring functions yet to be developed. Furthermore, in the evaluation with the aid of these functions, we restrict ourselves to the for prediction logics essential parameters Π_A, Π_L, Π_Z, because they are of eminent importance for the very track-creation and track-pursuit (see Def. 8). Π_A is the minimal number of signals from a flying object for its track to be initiated, Π_L is the number of the last points of a track being used for its prediction and Π_Z the maximal number of ineffective successive predictions, the track then being dropped.

3. ON THE APPROXIMATION.

Let an approximation operator APO for the real operator θ be given. To get a statement on the APO's quality as required above, a suitable measure is necessary. Since the process to be approximated is essentially a stochastic one, any approximation will be more or less bad.

(1) The filter for false signals will never operate perfectly.

(2) The compensation of missing signals will never be accomplished perfectly.

I.e. the curve interpolation of a single FO_j will be considerably falsified. In fact there will be linking of real signals to false signals, as well as linking of real signals from different flying objects. Some flying objects will never be traced, false tracks are looked upon as real ones. The confusion degree (approximation error) of an APO can be easily determined in a natural way, comparing the ideal track-net of an ST with the one created by the APO. Therefore we need a function measuring the approximation error with respect to the ideal track-net, and another function measuring the tendency of the ideal net to be disturbed by false tracks during its approximation. These functions are given in what follows.

To define them, we need another assumption, since in general we have:

$$S_i \cap \tilde{S}_i = \phi,$$

i.e. no APO at all will be able to meet the problems to be solved. The distance coordinates of a flying object are known only up to an error ellipsoid, the half axes of which depend on the radial error, the azimutal error and finally the elevation error of the RDS. Therefore the additional assumption (V4) is necessary:

(V4) $\tilde{P}_i \in \tilde{S}_i \Rightarrow (\tilde{P}_i \in S_i \cap \tilde{S}_i :\Leftrightarrow: \tilde{P}_i$ is in the ellipsoid of $P_i)$, i.e. each point $\tilde{P}_i \in \tilde{S}_i$ is interpreted as real, if and only if it is in the error ellipsoid of $P_i \in S_i$.

In this context there is to be considered, that any RDS gives its scanning coordinates in a discrete subspace of a space-time-continuum, i.e. the systematic error of the mapping into the discrete subspace is blended with the special measuring error of the RDS. Let now an APO $\tilde{\theta}$ for θ be given.

Def. 7. $S(\tilde{\theta}) := \{S \,/\, S$ r-track resp. $\tilde{\theta}$ over ST$\}$, i.e. the set off all r-tracks on ST created by the APO $\tilde{\theta}$.

We now consider an r-track from $S(\tilde{\theta})$ with the representation $S = (\overset{s}{\beta}_1, \ldots, \overset{s}{\beta}_m)$ and define in definition 8. the conditiones for a track part $S_{ir} = (\overset{s}{\beta}_i, \ldots, \overset{s}{\beta}_r) \subset S$, $1 \leq i \leq r$, to be a "true"-track.

Def. 8. S_{ir} is called "true" (t-track), if (E1), (E2), (E3) are valid.

(E1) $\bigvee j,\ j \in \{1, 2, \ldots, n\} \,/\, \overset{s}{\beta}_k = \alpha_{ji+k},\ k = 0, 1, 2, \ldots, r-i.$

(E2) $\overset{s}{\beta}_{i-1} \neq \alpha_{ji-1},\ \overset{s}{\beta}_{r+1} \neq \alpha_{jr+1},\ 1 < i < r < m.$

(E3) $L(S_{ir}) = r \geq n_A.$

We will use the following shortenings:

$$S' \wedge S(\tilde{\theta}) :\Leftrightarrow: (S, S \in S(\tilde{\theta}), S' \subset S, S' \text{ true}).$$

$$S' \in \{j\} :\Leftrightarrow: (S' \wedge S(\tilde{\theta}), \tau(P) = j \wedge P, P \in S').$$

Hence, $S' \wedge S(\tilde{\theta})$ means, that S' is a true-track in the sense of the Def. 8, $S' \in \{j\}$, however, is the identification (see Def. 1.) of S'. The definition 8 is to interpreted as follows: it requires, POST 1, only such tracks to be looked upon as true-tracks, which are part of real-tracks in the favourable case, where an ST has been scanned with PF=0. Thus an uncontrolled "positive" influence of the stochastic false track process on the real track process is avoided. This is the reason, why (E3) implicitly links the on n_A depending part of the prediction to the following quality functions.

Def. 9. $\tilde{\theta}(j) := \{\beta \,/ \bigvee S,\ S \wedge S(\tilde{\theta}),\ S \in \{j\}: \beta \in S\}.$

$\tilde{\theta}(j)$ is the set of real signals of FO_j satisfying Def. 8.

$M(j) := \{\beta / Vi,\ i \in \{1, 2, \ldots, m\},\ \beta \in S_i \cap \tilde{S}_i,\ K(\beta) = j\}.$

$M(j)$ contains exactly those signals given by the radar from real FO_j.

$L(j) :=$ number of positive coordinates of FO_j (FS_j) (:= length of the curve of FO_j).

We regard $\tilde{\theta}$ as to meet the demand POST 2 with respect to the curve FS_j, if

$$\text{Anz}\,(\tilde{\theta}(j)) \geq \text{Anz}\,(M(j)).$$

The weak requirement POST 2 is nothing but a natural considera tion of the individual detecting probability u_j of FO_j:

$$u_j = \text{Anz } (M(j)) \;/\; L(j),\quad j = 1, 2, \ldots, n.$$

This is easily seen, f.e. in the case where $(1-u_j) \cdot \ddot{u}_A > 1$. u_j therefore is a required value to be approximated by $\tilde{\theta}$ using

$$v_j = \text{Anz } (M(S)) \;/\; \text{Anz } (\tilde{\theta}(j)).$$

The value of the functional

$$\lambda_j^* := |u_j - v_j|$$

gives the error of $\tilde{\theta}$ with respect to θ on ST. POST 2 suggests to especially weighten the case in which we have

$$\text{Anz } (\tilde{\theta}(j)) < \text{Anz } (M(j)):$$

Thus we get

$$\lambda_j := \begin{cases} \lambda_j^*, \text{ if Anz } (\tilde{\theta}(j)) \geq \text{Anz } (M(j)) \\ 1, \text{ otherwise.} \end{cases}$$

Hence, we have the function $Q(\tilde{\theta}, \theta, ST)$ to measure the approximation error of $\tilde{\theta}$ against θ on ST.

$$Q(\tilde{\theta}, \theta, ST) := \frac{1}{n} \sum_{i=1}^{n} \lambda_i^2$$

with

$$0 \leq Q(\tilde{\theta}, \theta, ST) \leq 1, \text{ because } 0 \leq \lambda_j \leq 1,\ j = 1, 2, \ldots, n.$$

Def. 10. $\quad L(S(\tilde{\theta})) := \sum_{S \in S(\tilde{\theta})} L(S)$

$$LF(S(\tilde{\theta})) := L(S(\tilde{\theta})) - \sum_{S' \Delta S(\tilde{\theta})} L(S')$$

With the aid of the definition 10 the false track process is easily described:

$$FQ(\tilde{\theta}, \theta, ST) := LF(S(\tilde{\theta})) \;/\; L(S(\tilde{\theta}));$$

its value is proportional to the disturbance influence of the false track-on the real track-process, which could not be eliminated by $\tilde{\theta}$.

Remark: $\quad Q = 0,\ FQ = 0,\ \Delta ST \Rightarrow \tilde{\theta} = \theta$ (ideal case).

We give in Def. 11a our conditions for two APO's to be of semi-aquivalent quality with respect to the functions Q, FQ and in 11b we introduce the saying for one APO to be better than an other one.

Def. 11. a) $\tilde{\theta}_1 \overset{\varepsilon}{\underset{\varepsilon'}{\sim}} \tilde{\theta}_2 :\Leftrightarrow: (A_\varepsilon),\ (A_{\varepsilon'})$ (i.e. $(\varepsilon,\varepsilon')$-S-aquivalent), $\varepsilon, \varepsilon' \in \mathbb{R}^+$

(A_ε) $\quad |FQ(\tilde{\theta}_1, \theta, ST) - FQ(\tilde{\theta}_2, \theta, ST)| < \varepsilon$

$(A_{\varepsilon'})$ $\quad |Q(\tilde{\theta}_1, \theta, ST) - Q(\tilde{\theta}_2, \theta, ST)| < \varepsilon'$, i.e. $\tilde{\theta}_1$ is $(\varepsilon, \varepsilon')$-superior to $\tilde{\theta}_2$, if the absolute difference of (A_ε), $(A_{\varepsilon'})$ is less ε or less ε'.

Def. 11. b) $\tilde{\theta}_1 \overset{\varepsilon}{<} \tilde{\theta}_2 :\Leftrightarrow: (A_\varepsilon),\ (A_1)$, (i.e. $\tilde{\theta}_1$ ε-superior $\tilde{\theta}_2$)

(A_1) $\quad Q(\tilde{\theta}_1, \theta, ST) < Q(\tilde{\theta}_2, \theta, ST)$, i.e. $\tilde{\theta}_1$ is ε'-superior to $\tilde{\theta}_2$ if (A_ε) of

(a) holds and the approximation degree of $\tilde{\theta}_1$ is greater than that of $\tilde{\theta}_2$.

The value of $(\varepsilon, \varepsilon')$ results from the requirements a special user has concerning the real-and false track process.

4. THE MASTERTAPE AS REALISATION OF THE ASSUMPTIONS (V1), (V2), (V3), (V4).

Def. 12. A mastertape is regarded as a real air traffic situation ST, scanned by a radar-detection-system and stored on a tape, every signal carrying a label K = 0 or K$\in$ {1, 2, ..., n} showing the signal to come from a flying object or to be false (see Def. 1); the curves of the flying object (n flying objects), which have contributed to this traffic situation are known.

On the production and information security we make the following remarks:
The primary RDS scans during the time interval T_T (length of T) the air space situation ST. Technical reasons caused the scanning to be only two dimensional (no elevation). At the same time the PPI of a simultaneously working radar is photographically integrated. The so produced PPI-sum allows up to small errors identification of the respective curves (see fig. 1). These are, after correct identification, assigned natural numbers as labels. The label carrying PPI is added to the detector data on a display, and to these data the adequate labels can be given over a computer, this procedure being rather time consuming, but very simple. The thorough investigation of the PPI (see fig. 1) leads to a rather perfect decision security of man's eye concerning flying manoeuvers, starting and landing. Already after a short time the characteristics of a detector picture are controlled (comparing the detector data with the PPI). Some testmasters were created without the help of any PPI-information and later on compared with the resp. PPI-information. This test resulted in the fact, that the adaptive man's eye recognizes the PPI tracks up to a small error (d % ~ 5 %). It must be mentioned, that in this case, one depends up to a certain degree on the detector's decision-logic, see /9/ and POST 2.

Until now two representative mastertapes have been created. Each mastertape carries a heading, containing the radar-detector characteristics (GRS-radar, (L-Band), der BFS in Neunkirchen, Odenwald, pulse length 2 µs, distance = 120 n.m.), the place and the time of the data acquisition, the range of the controlled space (minimal, maximal distance), length of T, the number of flying objects creating ST, length of the respective curves, the number of marked signals from ST. For the above mentioned mastertapes we have:
acquisition place: air port Frankfurt, for T there were mainly chosen time periods with high traffic density. $T_T = 35 \pm 5$ min. $\hat{=}$ m = 200, number of flying objects: $280 \leq n \leq 320$, length of the curves: $8 \leq L(FS_j) \leq 200$, marked signals: $5 \leq M(j) \leq 195$, total number of signals 35000-45000, marked total 10000-15000. Hitherto there have been processed about $8 \cdot 10^4$ data that are now available.
Remark:
There has not yet been made any statement about the PPI's security. We regard the PPI to be at this moment one of the most important and reliable devices for air traffic control in Germany. That is why we believe to have a minimal decision security -in general- of

about 85 % - 95 %. The special detection probabilities μ_j are of a decision uncertainty not greater than that involved in the air trafric control as operated nowadays.

5. SOME RESULTS.

Some APOs $\tilde{\theta}$ with respect to Q and FQ have been evaluated. The inner structure of $\tilde{\theta}$ in it's prediction parts is based on ideas of W. Hauck /2/ and of H. Springer /1/, and was realized at the "FFM", Werthhoven /6/, /8/, where it is available under the program identification "Monoradar". "Monoradar" is a real-time Algol operator. The different logics of these APOs were realized by choice of $\Pi_A \in \{2, 3\}$, $\Pi_L \in \{2\}$, $\Pi_Z \in \{1, 2, \ldots, 6\}$. The evaluation of $\tilde{\theta}$ was done with (a) PF = 0 (PF = 0 means an à priory total suppression of unmarked signals.), (b) PF > 0. The results are shown in fig. 2, 2.1, 3, 3.1 and shall be shortly interpreted for fig. 2.1. 3.1.

INTERPRETATION:

The Q-curves in fig. 2.1 show clearly the influence of greater Π_Z (maximal number of successive predictions up to dropping). In the Π_Z-interval [3, 5] the values of FQ change but slightly, whereas in the interval [4, 5] the approximation degree increases up to about 70 %. Considering Π_Z, the operators $\tilde{\theta}_1$ with $\Pi_Z = 4$ and $\tilde{\theta}_2$ with $\Pi_Z = 5$ are optimal; in addition they are (0.02, 0.05)-aquivalent. The Q-curve in fig. 3.1 shows the unavoidable minimal error of $\tilde{\theta}$. The difference between the Q-curves of fig. 2.1 and fig. 3.1 gives an impression of the uncerntainty of $\tilde{\theta}$'s predicition part. FQ in fig. 3.1 gives the confusion degree of $\tilde{\theta}$, responsible for the false linkage of real signals coming from different flying objects. Added there to is in fig. 2.1 the very false track process and the disturbance caused by false signals, hindering the exact real track process, this disturbance alone being responsible for not reaching the maximal approximation degree in curve fig. 2.1 given by the curve of fig. 3.1. The pairs (Q_1, FQ_1), (Q_2, FQ_2) shown in fig. 2.1 represent the test result of a "logic" having been evaluated for $\Pi_Z = 4, 5$ (so called normal track start, track begin, i.e. for this logic $\Pi_A = 2$); these logics are seen to be far from optimality.

Fig. 4 and fig. 5 show the μ_j distributions of the according master-tapes (I and II). Pay special attention to the interval $30 \leq \mu_j \cdot 100 \leq 60$.

ABBREVIATIONS.

"$\dot{\vee}$"	"there exists one and only one"
"$\vee$"	"there exists"
"$\wedge$"	"for all"
":<=>:"	"then and only then per def."
$\mathbb{N}$	set of natural numbers
$\mathbb{R}^+$	set of positive real numbers
$\mathbb{R}^3$	three dimensional space
ϕ	empty set
{...}	explicit set representation
Anz{...}	number of elements of a set

ACKNOWLEDGEMENT

We are obliged to the "Bundesanstalt für Flugsicherung, BFS" for their friendly help during the data acquisition.
We are indebted to Mr. W.J. Grünewald, who designed the whole program-system (man-machine) for the mastertape production, and in addition coded and checked it.
We particulary thank our young information-assistants, Miss J.Dannebaum, Miss G.Jansen, Miss E.Luitjens and Miss H.Reinsberg and as well Mr. M.Butter for their valuable and unselfish help in the mastertape production through many nights.

LITERATURE

/1/ Springer, H.
Modell eines Datenverarbeitungssystems für die Flugsicherung
Elektron. Datenverarbeitung, Heft 3, 1965

/2/ Haack, W.
Beobachtung des Luftraumes durch automatische Verarbeitung der Information von Rundsichtgeräten mittels digitaler Rechenautomaten
Arbeitsgemeinschaft für Forschung des Landes Nord-Rhein-Westf., Heft 27, S.39-64

/3/ Kamm, K., Heijnen, W.J.B.
A Detailed Description Of The STC Tracking Evaluation Methode, TM-125, Oct.1965

/4/ Steeadman, R.H.
An Off-line Tracking Evaluation S.A.T.A.N.
Extraction Digitale et Truitment des Informations radar. TOME II, Colleques des 7. 8. et 9 juin, 1966, Bruxelles (Eurocontrol)

/5/ L. Leth - Espensen
Evaluation Of Track-While-Scan Computer Logic
Radar Techniques For Detection Tracking and Navigation, of The Eighter Symposium Of The Agard Avionics Panel. London 21.-25. Sept. 1965

/6/ v. Issendorf, H.
Organisation und Strategie eines automatischen Monoradar-Beobachtungssystems
Elektron. Datenverarbeitung, Heft 8, 1968

/7/ Wirth, W.D.
Beitrag zur Vereinfachung der Radarauswertung durch Digitalisation
Dissertation TU Berlin, 1962

/8/ Pieckner, H., Hönerioh, H.
Erwartungsgebiete für die automatische Flugobjektbeobachtung
Elektron. Datenverarbeitung, Heft 7, 1968

/9/ Both, W.A.
Zielerkennung in Radarsignalfolgen mittels binärer Digitalisierung und Filterung
Diplom-Arbeit, Saarbrücken 1968

Fig.1 PPI sum

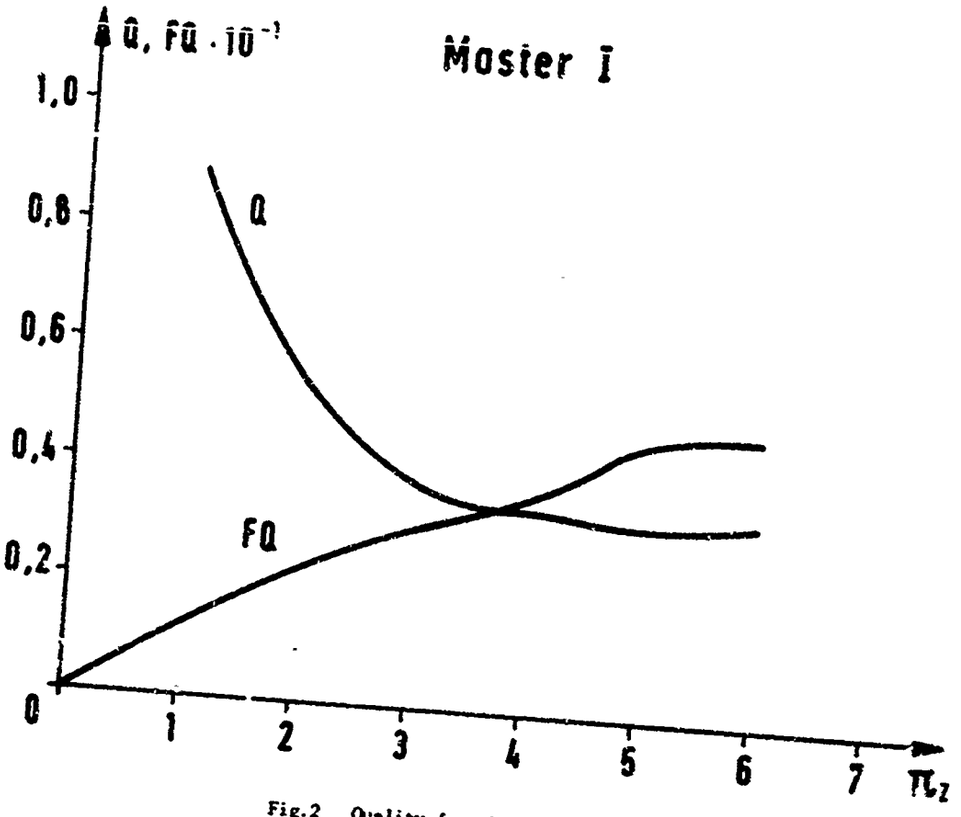


Fig.2 Quality functions of master I,
$\Pi_L = 2$, $\Pi_A = 3$, PF $\neq 0$

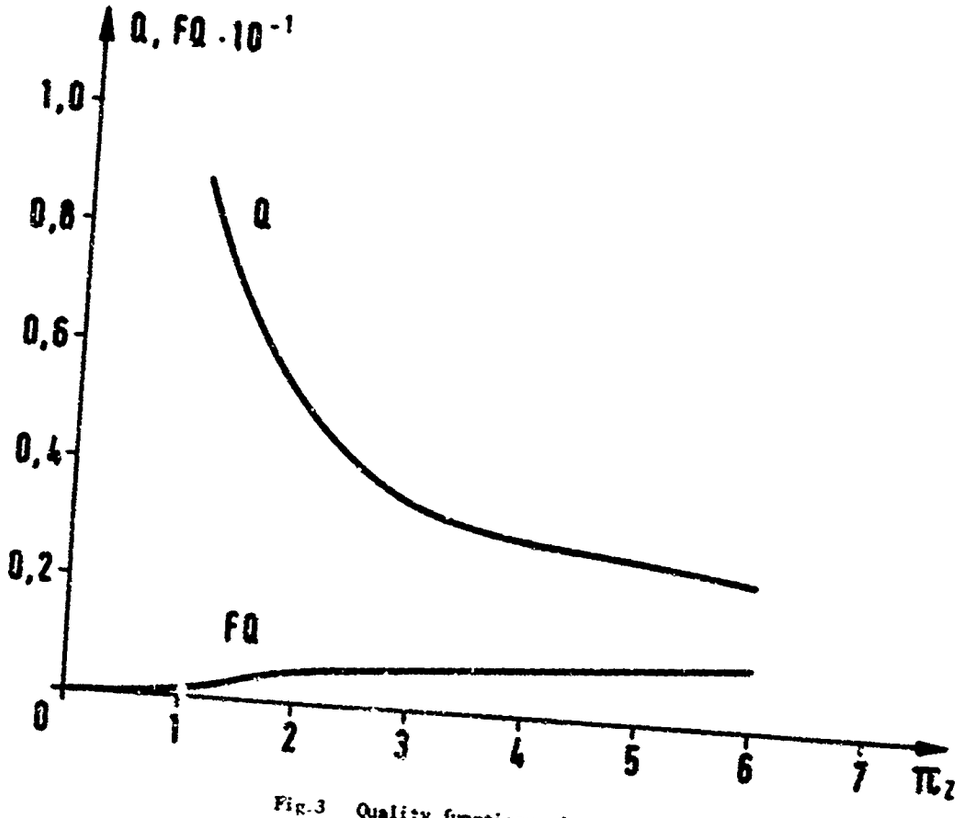


Fig.3 Quality functions of master I.
$\Pi_L = 2$, $\Pi_A = 3$, PF = 0

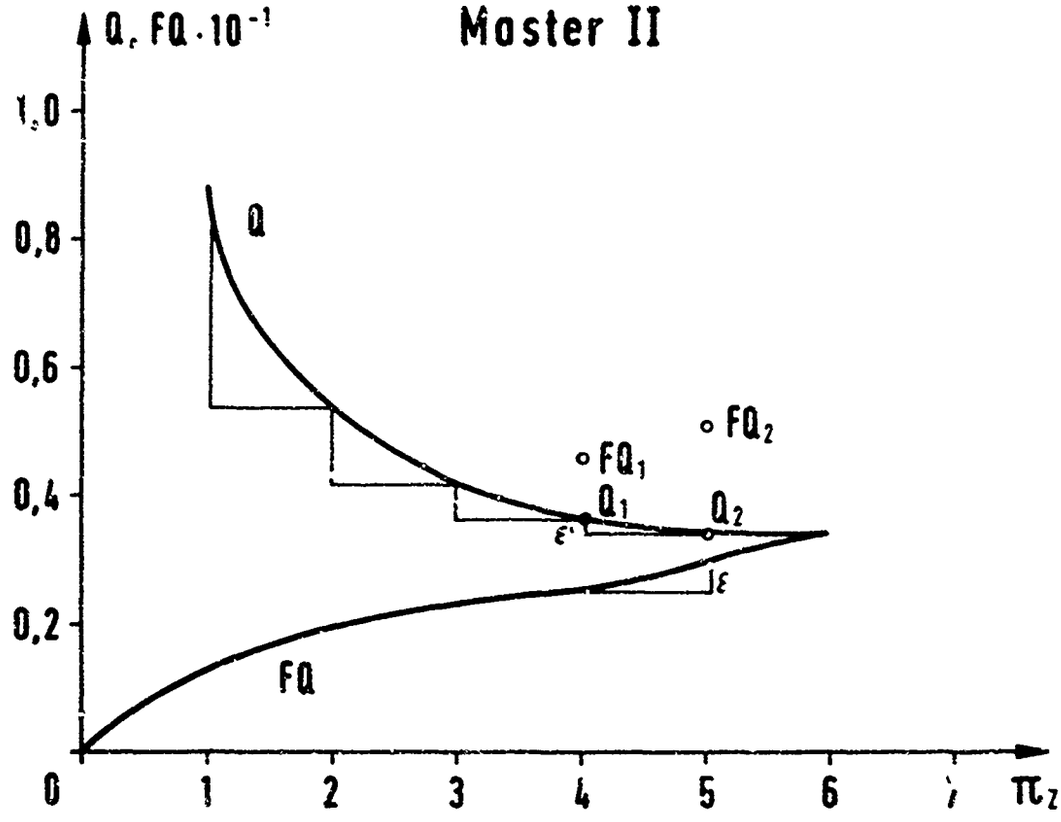


Fig.2.1 Quality functions of master II.
$\Pi_L = 2$, $\Pi_A = 3$, PF $\neq 0$

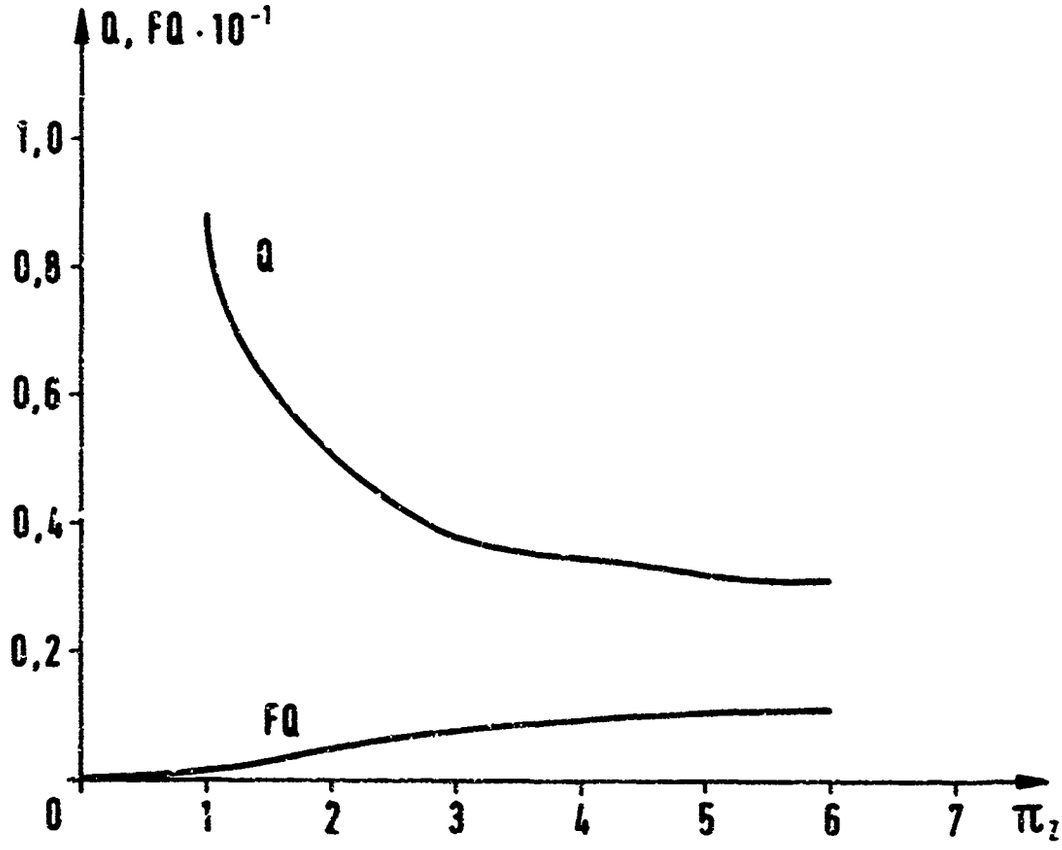


Fig.3.1 Quality functions of master II.
$\Pi_L = 2$, $\Pi_A = 3$, PF = 0

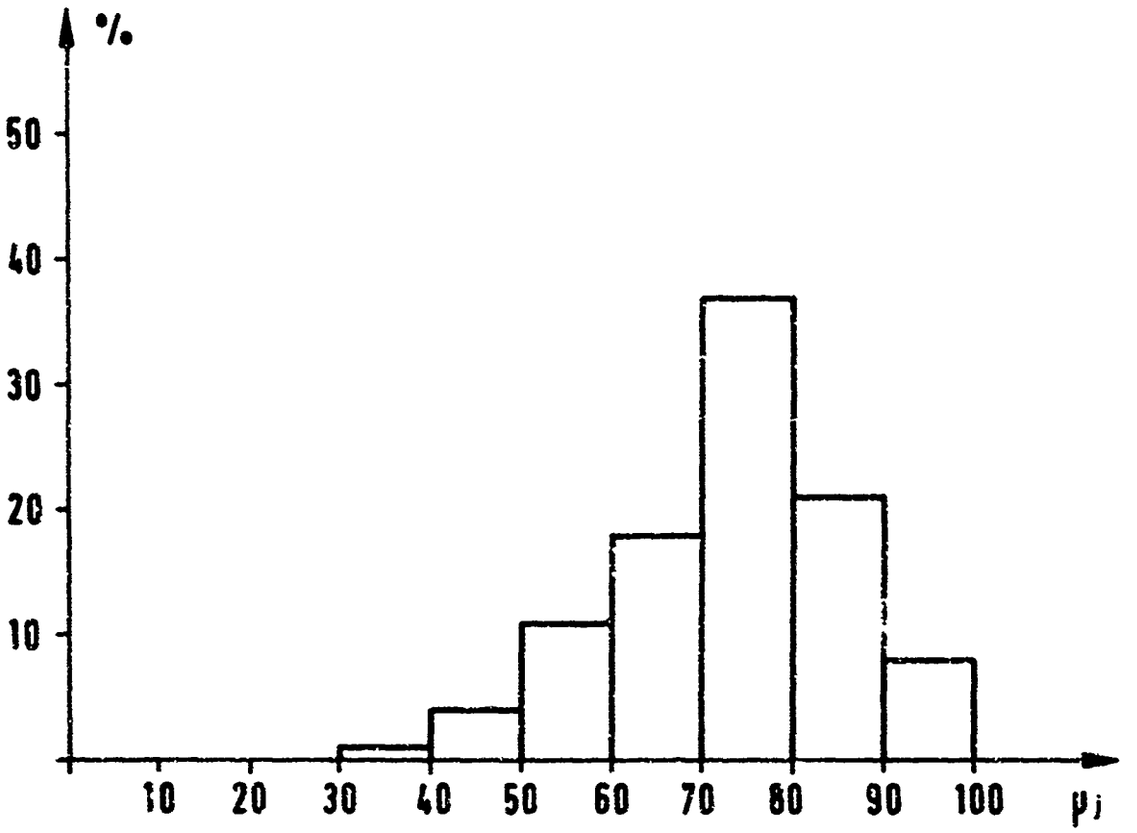


Fig.4 μ_j-distribution of master I

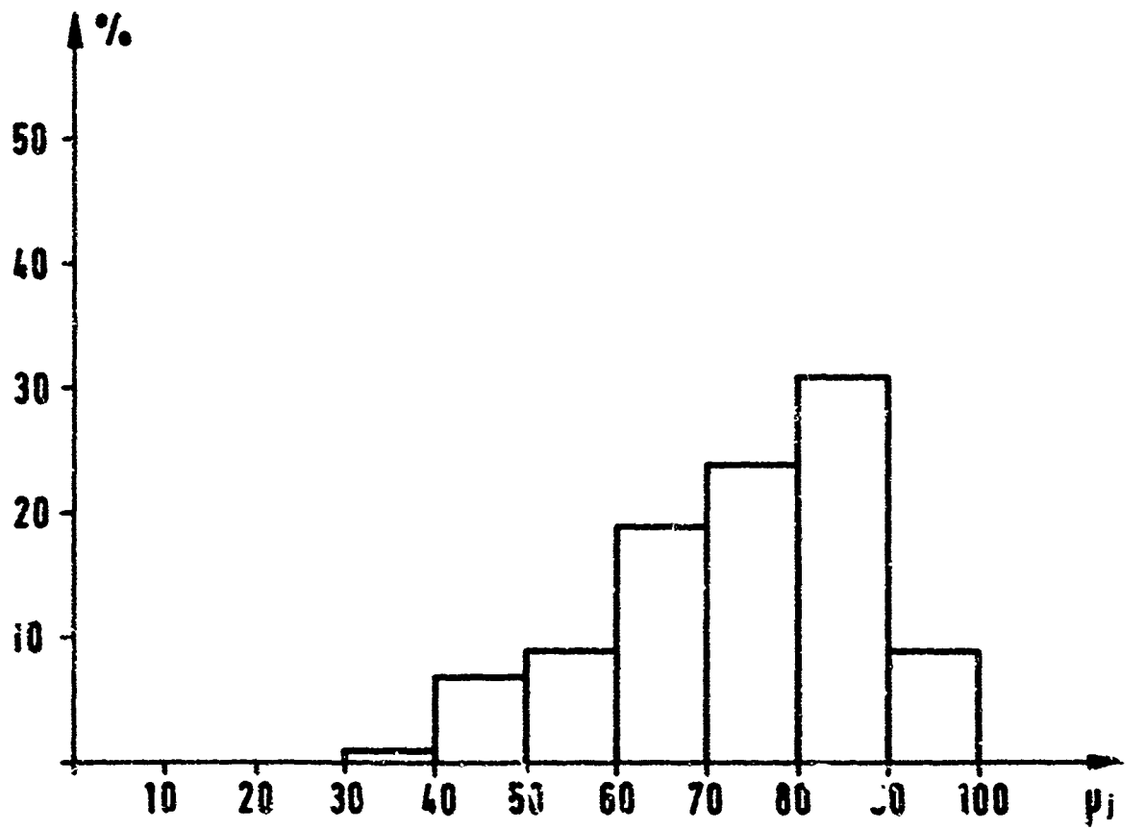


Fig.5 μ_j-distribution of master II

DIGITAL RADAR PLOT EXTRACTOR

A SYSTEM MODULE OF AN AUTOMATIC AIR TRAFFIC CONTROL SYSTEM

by

H. Ebert and D. Siegenthaler

AEG-TELEFUNKEN
Ulm and Konstanz, Germany

DIGITAL RADAR PLOT EXTRACTOR

A SYSTEM MODULE OF AN AUTOMATIC AIR TRAFFIC CONTROL SYSTEM

by

Dr. Ing. H. Ebert and
Dipl. Ing. U. Siegenthaler,
AEG-TELEFUNKEN, Ulm and
Konstanz, Germany

1. Basic Functions of Radar Plot Extractors

Improved methods for air traffic control are increasingly based on the information obtained by radar on air traffic movements. The fail safe capability of the radar control therefore becomes of increased importance. This can in the first place be achieved by multiple coverage of the controlled airspace with the aid of an increased number of radar installations. In order to be able to utilize the informations from these radars at a central point, it is necessary to construct a closely integrated telecommunication network which will enable the radar information to be transmitted efficiently, even over long distances.

This still further increased input of radar information at the air traffic control centres is in no way equalled by an increase in human capability to cope with the processing of this data. Large control centres must, in the future, be equipped with digital computers which will be the only possible way to handle the processing of the high rate of input information. An important task foreseen for computers is that of organizing the display of air traffic data on screens of a new technical concept, which are designed for computer control and on which alphanumeric as well as geographical information can be displayed. To enable the presentation of the air traffic situation to be continually brought up to date, the radar installations (primary and secondary radar) must be directly connected with the computer at the control centre. The radar installations were not originally designed for such a direct connection. Their signals must therefore first be put into digital form. An interfacing installation is thus necessary between radar and computer, this being known as a radar plot extractor.

The equipment and processing techniques that will be used in the future are shown in Fig. 1. New equipment at the radar site is the digital radar plot extractor. This item is first of all an analog-digital converter, that is, it converts the analog radar signals into digital form. Besides this, it evaluates the signals from the primary and secondary radars in the sense of an automatic target recognition, that is, it takes over a function that up to now has been carried out by controllers at the display tube. The data flow is thereby reduced to such an extent that the transmission capacity of two to three normal telephone cables (bandwidth each about 3 kHz) serves to transmit a radar display. The information that is transmitted contains basically the position coordinates of the detected target and the additional information from the secondary radar. By this method, which allows the normal telephone network to be used for the transmission of radar display information, the required solution to the problem of efficient transmission of radar information has been found

The most extensive new equipment in the control centre is the air traffic control computer system, which takes over a large amount of data processing and coordinating functions. In the survey given here on future methods of radar processing, we shall limit ourselves to the description of the computer system's function in target tracking and presentation of processed radar data. Tracking, i.e. automatic recognition of aircraft movements, is done in a first digital computer (Fig. 1) based on the transmitted information of several radar installations whose coverage areas overlap (Multiradar Tracking System). A second computer is used for processing the information that is presented at the individual controller displays.

As shown in Fig. 1, the radar plot extractor consists of two parts, the plot detector and the plot processing unit. The plot detector is a system specifically designed for automatic target recognition and determination of target coordinates. For further processing of target information, which will be described in more detail later, again a digital computer is used. The computers used within the plot extractor system for automatic target tracking and for the display system are all of the same type and size (TR-86 computer), thus simplifying problems of maintenance and software and hardware handling.

2. Principles of Automatic Target Detection Technique

Before the description of special items of the radar plot extractor presented here is continued, a short introduction to the techniques of automatic target detection should be given:

The same basic problems occurring in manual control and caused by radar signal characteristics also have to be overcome with automatic target detection. The main characteristic points of the radar echo signal are, on the one hand, its weakness owing to the reduction in reflected power as a fourth power of range, and on the other hand, that it is always mixed with clutter caused by thermal noise as well as by cloud and ground reflections. The usual radars are so arranged that a target is covered by several pulses, on average between nine and twenty. Target recognition is based on the detection of a pulse sequence correlated by the target in which the single pulses come from the same range. Noise pulses appear, on the contrary, to be more or less uncorrelated and come from different ranges. By adding several target echoes on the screen the observer can determine the useful target, which will be distinguished from a single noise echo. Target detection is therefore the result of an integration of several target echoes. For automatic target detection, the controller is replaced by circuits with similar integrating capabilities. The operation is divided into the two steps of echo detection and echo integration. The echo detection clarifies the question of whether a single received echo should be evaluated as coming from a target or as noise. For this purpose a voltage comparison switch, known as the first threshold, is used (Fig. 2). If the voltage of an echo pulse exceeds a preset threshold voltage, this pulse will coordinate a digital output. If it is smaller than the threshold voltage, then it will be suppressed (evaluated as digital zero). Once an echo pulse exceeds the first threshold voltage, then it will be processed as coming from a target, even though it has possibly originated from noise. Efficiency of target detection is therefore considerably dependent on the value of the threshold voltage. If it is set too low, all target echoes will be detected but it also increases the possibility that noise pulses can exceed the low threshold level. If, on the other hand, it is set too high, only a few noise pulses will exceed it, but then low voltage target echoes will be lost.

Echo integration, that is, the actual target detection, is achieved in the following way: Within an antenna beamwidth and a definite range interval, quantized echoes are examined for correlation, i.e., their number is compared with a further threshold criterion, known as the second threshold. This procedure is carried out by means of a switch, which by virtue of its manner of working is known as a Sliding Window Detector (SWD). Fig. 3 clarifies the method of operation of the sliding window detector. A target is assumed to be at a range of 58.25 nautical miles and that it will be detected within the rotating antenna field. Within the field, a total of nine radar periods should be able to acquire the target. During period P1 an echo is received, which after exceeding the first threshold, is recorded as "1" in the first stage of a nine stage shift register. The total contents of the register are then serially shifted up from each stage of the register to the adjacent one. Due to this the contents of the ninth stage are then erased and the first stage in which the "1" received during period P1 was recorded will again be vacant since this "1" is now occupying the second stage. During period P2 an echo is received which does not exceed the first threshold level. A "0" will therefore be recorded in the first stage of the shift register. During the third to the sixth periods, a "1" will also be recorded in the register, from which up to this point five quantized echoes have been received from the target. This value of five echoes out of a possible total of nine is defined as the criterion for the second threshold, that is, the target detection threshold. If the number of alarms received from a target reaches this minimum value, then a "start of target" is recorded. Further turning of the antenna means that the target will be outside the coverage area. If the register contents are less than a definite value, for example four as in the diagram, then "end of target" is recorded. According to the discrimination capability and range coverage of the radar, some five hundred to one thousand sliding windows are necessary in order to obtain optimum results for target detection within the coverage area. It is then possible to divide the range coverage area into rings, whereby each of the sliding windows corresponds to a definite range. The range covered by a ring is suited to the pulse length and in practice is from two to four microseconds. The number of positions of the windows is dependent on the maximum alarm number during an antenna pass. In general this is between nine and twenty and is calculated from the antenna directivity, the number of antenna rotations, and the pulse rate. From the azimuth values of the "start of target" and "end of target" information given by the SWD, the azimuth value of the target centre can be calculated. The position number of the window in which a target is detected already corresponds to a definite range. Therefore the coordinates of a detected target are known and the basis of an automatic target detection system is obtained.

The procedure that has been described will also be used in a similar manner for automatic recognition of secondary radar targets. For reasons of expediency, the automatic target recognition switching circuits for both radar systems are combined in a single unit, the combined digital radar plot extractor for primary and secondary radar.

Information on automatically detected targets is transmitted to a computer for further processing. The data flow from the plot detector to the computer is statistical because each information is transmitted immediately after detection. Therefore the transmitted data has to be buffered in the memory of the computer before further processing.

3. Clutter Suppression

In most of the theoretical considerations it is assumed that the background against which detections of primary radar targets have to be made is thermal noise, whereas in fact the background may consist of meteorological or land clutter in addition to receiver noise. Consequently, an automatic detection equipment must be designed so that it will not generate too many false alarms in clutter areas, but will continue to detect any targets of sufficient amplitude to show above the clutter.

Circuits for clutter suppression may be used within the system for hit detection (first threshold) as well as combined with the switching circuits for plot detection (second threshold). In both cases the basic concept is to raise the threshold for unwanted signals.

As a system of the first kind we use a circuit ahead of the first threshold, which is sometimes used in the conventional radar technique for rain clutter suppression (Fig. 4). It is known that the resultant amplitude distribution of such clutter closely approaches Gaussian distribution. When such signals have been detected by a logarithmic receiver such as we use, the d.c. deviation component is independent of the signal amplitude, while the d.c. component increases with increasing signal power, i.e. clutter power. In order to subtract the d.c. component of the signal, a delay line circuit is used. The video is applied to a delay line of twenty radar pulse lengths. The mean d.c. level of the background is obtained by two adding circuits, one connected to a series of line taps ahead of the centre top, the other connected in the same manner behind the centre top. The d.c. levels obtained at the output of the adding circuits are compared and the greater is subtracted from the signal sample taken from the centre top.

This circuit was tested with an AEG-TELEFUNKEN GRS-Radar. A remarkable suppression of rain clutter was observed and the measured loss of sensitivity was about 0.6 db.

A further clutter elimination system is used in connection with the sliding window detector (Fig. 5). The hit history of eight sliding windows before and behind the sliding window "in process" is examined for the false hit probability and the amount of hit correlation.

The probability of false hits p is

$$p = \frac{f_1}{n} \qquad (1)$$

f_1 is the number of measured hits in the examined area, n the total number of possible hits or hit cells, which is in our case 16 x 15 = 240 for a sliding window length of 15 cells.

The correlation can be derived from the probability p_{11} of "double hits", i.e. the probability that a second hit immediately follows a first hit in azimuth direction (Fig. 5). If the number of such double events f_{11} is known, p_{11} can be calculated by

$$p_{11} = \frac{f_{11}}{f_1} \qquad (2)$$

p_1 and p_{11} are calculated for each sliding window process and their values are used to adjust the second threshold automatically in such a way that preset values of false alarm probabilities are kept. This means that the second threshold is raised from a normal value of eight out of fifteen to higher values, when the number of hits and double hits increases, indicating a clutter area. The preset probabilities of false alarm can be individually adjusted. The complex hardware system, consisting of TTL integrated circuits, is fully digitized.

The first tests of this clutter eliminator done in the laboratory with tape-recorded radar data have shown that clutter can be eliminated to a great extent, depending on the preset false alarm probability, which is kept constant. The great advantage of this system is that the systems following the target detector, i.e. the TR-86 computer for target processing and the data transmission over telephone lines can be automatically prevented from overflow caused by unwanted data. This technique has the further important advantage of being quiescent in areas not containing clutter and yet acting instantaneously in clutter areas.

Photographs showing the effectiveness of both clutter eliminating circuits are given later in chapter 10.

4. Block Diagram of a Combined PR/SSR Plot Extractor

Figure 6 is a block diagram of our combined PR/SSR plot extractor equipped with a TR-86 computer. The group denoted "plot detector" contains the circuits for target detection, which is done in separate channels for PR and SSR targets. The tasks of the computer are given in the block diagram and described in more detail in chapter 6.

Since in our installations the antennas of both radar systems are mechanically linked, a common timing unit in the plot detector is sufficient for the central generation of clock pulses and digital range and azimuth information. If a start or end of a target is declared, the associated azimuth and range data are sent to the computer by a transfer circuit, consisting of a shift register buffer. In addition to these data the received code information is connected through to the computer for further processing (code validation) if the start of an SSR target has been declared. The SSR decoder unit has similar functions for SSR-video handling as the first threshold unit for processing of PR-video. Additionally it includes capability for emergency code alarm and special degarbling functions to reject erroneous code data due to SSR code overlap.

The total number of range increments of each channel is 1024. Each range increment has a minimum width of 2.5 µsec, but other values may be chosen by changing the crystal of the master clock. For fine range measurements each range increment is subdivided into four further increments, each at least 0.625 µsec wide.

In the PR-channel a unique pulse splitter circuit operates on the quantized video to determine final acceptance and range of each signal crossing the first threshold. The circuit determines the center position of each radar hit rather than using the conventional leading edge of the return signal. This enables deciding for only one hit also in cases of overlap between adjacent sliding windows and reduces the sweep-to-sweep variations in the position reports of a target. Rejection of signals too narrow to be live returns is also included. For processing of normal and MTI video two separate first threshold units are used. They are followed by a switch which is controlled either automatically by the computer or manually, and connects one of the two video types to the sliding window detector. A quality factor is generated according to the hit number reached by a PR-target and sent to the computer.

SSR-plot detection is normally evaluated for one identity code plus altitude code detection in any interlace ration. In that case, a SWD with a run length of ten bits is used. However, plot and code detection for any two identity codes (plus detection of altitude code) may also be used. Two SWD's each five bits wide are then used for plot detection.

The plot extractor includes complicated built-in test features that permit checking the system operation. A major feature is the internal PR/SSR test target, which provides a means for continuous on-line check (for more details, see chapter 5). Further built-in tests permit a rapid check of SWD memory and switching circuit operation, clutter eliminator operation, data transfer to the computer and proper SSR-code processing.

For interstage control of the whole extractor system, a 16" monitor display is included, which permits presentation of unprocessed PR- and SSR-video, quantized video, acceptance criteria of PR- and SSR-SWD's, several map videos (MTI-map, low and high density clutter areas) and synthetic plot representation by several types of symbols (PR-, SSR-, combined plots etc.).

A special PR/SSR-plot extractor system had to be developed for use in connection with a complicated back-to-back or "Janus" radar system. Figure 7 is a block diagram of the whole installation, consisting of an AEG-TELEFUNKEN SRE-LL radar and the special extractor for this radar installation. As can be seen from the diagram, the radar installation contains one SSR-beam and two PR-beams on each of its halves, thus providing a total of two SSR- and four PR-beams or "radars". To cope with this system, one would at first expect that two SSR- and four PR-plot detectors are needed. However, several aspects allow diminishing the additional need to about 30% for a Janus plot extractor system with 100% standby at the computer group and 50% standby at the detector group, compared with a normal detector installation with 100% standby equipment:

1) The computer is also able to handle the larger data flow caused by the back-to-back radar. Therefore the computer group remains unchanged.

2) In the PR-part of the system a pulse length of 4.5 µsec is used. Therefore it is possible to split the PR-part of the SWD-memory in two halves of 512 range cells, thus preparing two independent SWD's for the upper and lower PR-antenna diagram. In that case, the PR-range increment is 5 µsec, so that the same coverage area as before is reached. For SSR, a range increment of 2.5 µsec is used as before. Therefore only one first threshold unit and a clutter eliminator unit for the second PR-beam are needed additionally in the plot detector group. As can be seen from the diagram, a complete PR/SSR-plot detector (altered and enlarged as just described) is used on each side of the two radar halves, while a third unit is provided for standby purposes.

5. Data Link between Plot Detector and TR-86 Computer

5.1 Structure of the Plot Part Messages

Figure 8 shows the structure of the plot part messages which are transmitted from the plot detector to the TR-86 computer. For PR and SSR, several types of plot part messages are used. In the case of PR, the sliding window detector generates a PR plot start message when the second threshold level is reached. Also a plot end message is generated when the plot-end-threshold is reached. Both messages have the same range number. In addition the PR plot start message contains the fine range and the PR plot end message contains the quality factor. Azimuth information is transferred by separate azimuth messages which are generated at the beginning of every sweep. In the case of SSR, the sliding window detector generates SSR acceptance and SSR end messages when the associated thresholds are reached. SSR code information is transmitted to the computer when the SSR plot-start-threshold is reached. This threshold is adjusted to a value below the SSR-acceptance-threshold.

5.2 Radar Data Input Channel

The transmission of the plot part messages from the plot detector is carried out by the Radar Input Channel, a special hardware unit. It has to feed the plot part messages into two separate input buffers of the TR-86 memory, one for PR and one for SSR. For this purpose the radar data input channel is equipped with two separate address generators. Azimuth messages are fed into both buffers. The messages are transported directly to the buffers, i.e. without touching the central processing unit.

The maximum input rate which can be handled by the radar data input channel is one TR-86 word of 24 bits per 1.25 µsec.

6. Tasks of the TR-86 Computer within the System

The tasks of the TR-86 computer within the system are given in Fig. 6. These tasks are:

1. Validation of the SSR code messages
2. Determination of the centre azimuth of PR and SSR plots
3. Correlation of PR and SSR plots which are generated by the same target
4. Discrimination of false targets (if possible)
5. Generation of message formats for the extracted plot messages which are transmitted via telephone lines.
6. Storage of plot messages of up to 1 antenna revolution (This task may be necessary because the transmission rate of the telephone lines is limited to 3 times 2400 bits per second for each linked station).
7. Organization of the data transfer process, i.e. providing the different telephone lines with the extractor output messages
8. Supplying the monitor display of the extractor with synthetic radar data
9. Overflow control
10. Generation of test plots which are transmitted back to the plot detector, processed there and again transmitted to the computer. Thus, by comparing transmitted and received test plots an examination of the proper working of the system is possible.
11. Control of plot detector

The TR-86 computer system used for the plot extractor includes papertape punch and reader as well as control teleprinter. These units and a special software operating system enable the operator to correspond with the system. Supervisor functions are carried out by special hardware and software systems, which announce failure situations.

7. Amount of Real and False Plots within Coverage Area

Depending on weather conditions and air traffic situation, the plot detector will transmit to the computer the plot part messages of up to 350 real and 650 false plots during one antenna revolution. That means that during one antenna revolution the computer receives up to 2000 PR plot part messages and about 7000 SSR part messages (including the SSR code messages). In the case of back-to-back radar the plot part messages of about 500 real targets have to be processed during one information repeating period of about 5.5 seconds.

7.1 Amount of Plot Accumulation

Moreover, a considerable portion of the plot part messages of the real targets plots may appear within a sector of less than 10 degrees. Thus the computer must be able either to store or to process several thousand plot part messages within a time period of about one half second. This corresponds to a value of plot part messages of about 60 SSR plots within one radar beam.

7.2 Maximum Short Time Input Rate

Finally we must consider that within a period of about 100 µsec up to 65 plot part messages may be generated by the plot detector. To avoid saturation of the link between plot detector and computer, the plot detector is fitted with a flipflop buffer unit while the radar data input channel is able to handle one plot part message per 1.25 µsec.

8. Special Problems

8.1 Code Validation

As mentioned previously, the computer has to validate the SSR code replies received from the plot detector. In order to inform the computer which SSR codes belong to the same SSR target, the plot detector adds to the code information an address number of 7 bits, which is generated in a cyclic manner. The computer then checks that at least two subsequent code replies of the same modus and address number are identical.

These tasks are carried out by the code validation program of the computer. This program has the highest priority within the program system for radar data processing because it has to operate on-line in order to suppress the large number of redundant code replies. The suppressing factor caused by the on-line processing of code replies is about 5 to 1, thus reducing the required memory for SSR plot part message storage from 10 to 2.5 k TR-86 words.

8.2 Scan-to-Scan MTI

Another special task the computer has to execute is known as Scan-to-Scan MTI, which means that all plots are suppressed which do not change their position until the next antenna revolution. All pure PR plots which are not correlated to a corresponding SSR reply and which are not farther apart than 50 nautical miles will be processed by the scan MTI program. The scan MTI area is divided into elements of about 0.2 nautical miles edge length. A plot will be recognized as coming from a fixed target if it does not leave this area during one scan.

At the present time the TR-86 computer is programmed to examine up to 800 plots which may be fixed targets, a number which can be changed if necessary.

8.3 Providing Air Traffic Control Centres with Synthetic Radar Information

The capacity of the data link we use between plot extractor and ATC station enables transporting the corresponding plot messages at about 35 plots per second. One plot message consists of three to six fields of 16 bits each. That means two 2400 bits/sec telephone lines are sufficient. In order to provide the necessary meantime-between-failure, an additional telephone line of 2400 bits/sec is installed for each ATC station for standby purposes. The computer system enables providing up to three ATC-stations with different processed radar information.

In order to avoid an uncontrolled overflow if the plot detector delivers too many messages, special provisions are foreseen to reduce the data flow in a step-by-step manner. A similar process to reduce the data flow is used if two of the three telephone lines to one ATC-station stop working properly. The reduction of data flow is done in the following manner: At first only messages from privileged areas are transmitted. If necessary the SSR code replies are suppressed also.

If this is not sufficient, the solo PR and finally the solo SSR plots are suppressed so that only combined PR/SSR messages are transferred.

On the other hand, the processing program enables transmitting individual data to the different ATC stations. The stations may, for instance, require only plots which are recognized by SSR, they may receive plots coming from especially interesting regions and they may, in addition, claim plot positions either in x, y or ϑ, ϱ -co-ordinates.

8.4 Control of the Plot Detector

As mentioned before, the computer controls the Normal/MTI video switch in the plot detector. A further switch permits stopping plot detection within wanted areas. For this purpose the switching information is stored within two files, the "MTI" file and the "Blank Out" file. The coverage area is divided into area elements of 2.8 degrees and 6.4 nautical miles. Within each file one area element is represented by 1 bit, which gives the information for the required switching position.

While the Blank Out file must be changed by papertape reader or control teleprinter input, the MTI file can also be learned by the computer. The information stored within this file must be dependent on the actual weather situation, which is known by the clutter elimination system of the plot detector and given to the computer.

The map information concerned will be transmitted to the plot detector every 2.8 degrees. The serial channel, which is another special hardware unit, serves as a data link from the computer to the plot detector. It transports several message types, including monitoring and test plot information.

8.5 Standby Procedures

In order to secure the required meantime-between-failure the computer system is outfitted with a 100% standby unit in which two identical TR-86 computers are used. The active computer is connected to the data link but both units, active and standby, process radar data. Switching over from one computer to the other is controlled by the software and watched by the switchover control unit of the computer system: Every 100 msec each computer checks its own status and sets a status flipflop, which can be interrogated by its partner. On the other hand, every 100 msec each computer interrogates the status flipflop of its partner. If the standby computer recognizes that the active unit is not properly working it will then call for switchover. This information is transferred to the switchover control unit. Switchover, however, will only be carried out if the switchover control unit recognizes that the status flipflop of the active computer announces "not okay".

9. Software Organization

Figure 9 is a block diagram of the software structure. The programs called "interrupting programs" interrupt the radar data processing program if necessary. They belong to the operating system. The interrupting programs: 1) announce special failure situations, 2) serve the control printer, papertape units, monitoring display, and a clock pulse.

The radar data processing program is divided into two main parts. One is the code validation program as mentioned before and the other is the radar data extraction program, which carries out all other tasks of radar data processing. As can be seen from the flow diagram in the centre of Fig. 9 , radar data extraction can only be carried out if code validation rests temporarily, thus providing the necessary priority of code validation with respect to data extraction. There are two cases in which radar data processing will be in a waiting situation:

1) The SSR- and PR-input buffers are empty

2) The output buffer for the data link is completely filled.

During both situations for a period of up to 1 msec the radar data processing program delivers program management to the so-called waiting program, which is the program part of the operating system with the lowest priority. The tasks of the waiting program are such as preparing output messages or processing input orders, which are handled by the control teleprinter or papertape units. Another task of the waiting program is to check all the TR-86 statements.

The arrangement of the different files and the flow of radar data may be seen in Fig. 10. The radar data input channel fills the input buffers for the PR- and SSR-plot messages. Code validation is done by the code validation file and causes the clearing of the SSR input buffer as well as the filling of the SSR intermediate buffer. Radar data extraction will be carried out by means of a range orientated file and the plot correlation file.

Solo PR plots are stored within the fixed plot file for one antenna revolution and are transmitted to the ATC centres only if they have escaped the scan MTI process, as mentioned earlier. Radar data extraction causes the clearing of the PR input buffer and SSR intermediate buffer. On the other hand, the output buffers for the monitoring display and the data link to the ATC centres are filled up by extracted data. Finally it should be mentioned that the radar data processing program and the operating system use about 5 k TR-86 words (word length = 24 b ts) while the buffers and files need about 9 k, thus leaving 2 k as spare room in the total 16 k memory with which the TR-86 computer used is equipped.

10. Results of Practical Tests

Fig. 11 is a picture of the plot detector group

Fig. 12 shows the computer TP-86

Until now the whole extractor system has been tested with real PR video recorded with a special radar video recorder by Ampex, Great Britain, simulated SSR-plot data (multiple SSR-plot data prepared with the aid of a Data Pulse 203 Program Generator), and simulated data transfer. Simultaneous running of recorded and simulated input data and data output showed that the computer system is able to handle the specified data flow and that spare computer capacity is still available.

The effectiveness of the rain suppression unit and the clutter eliminator in the PR-channel is shown by several photographs taken from the monitor PPI. Another photograph shows the output data of the extractor system, presented at the monitor PPI as different types of symbols.

11. Conclusion

This abstract was to demonstrate what great efforts have been made during recent years in radar plot extraction technique. One of the main tasks was to develop efficient plot extractor systems for practical use in ATC. Among the work done for years, particularly in USA, Great Britain, France, Sweden, and also in West Germany, the system just described may be one of the most modern and efficient ones, especially because of its use of the process computer.

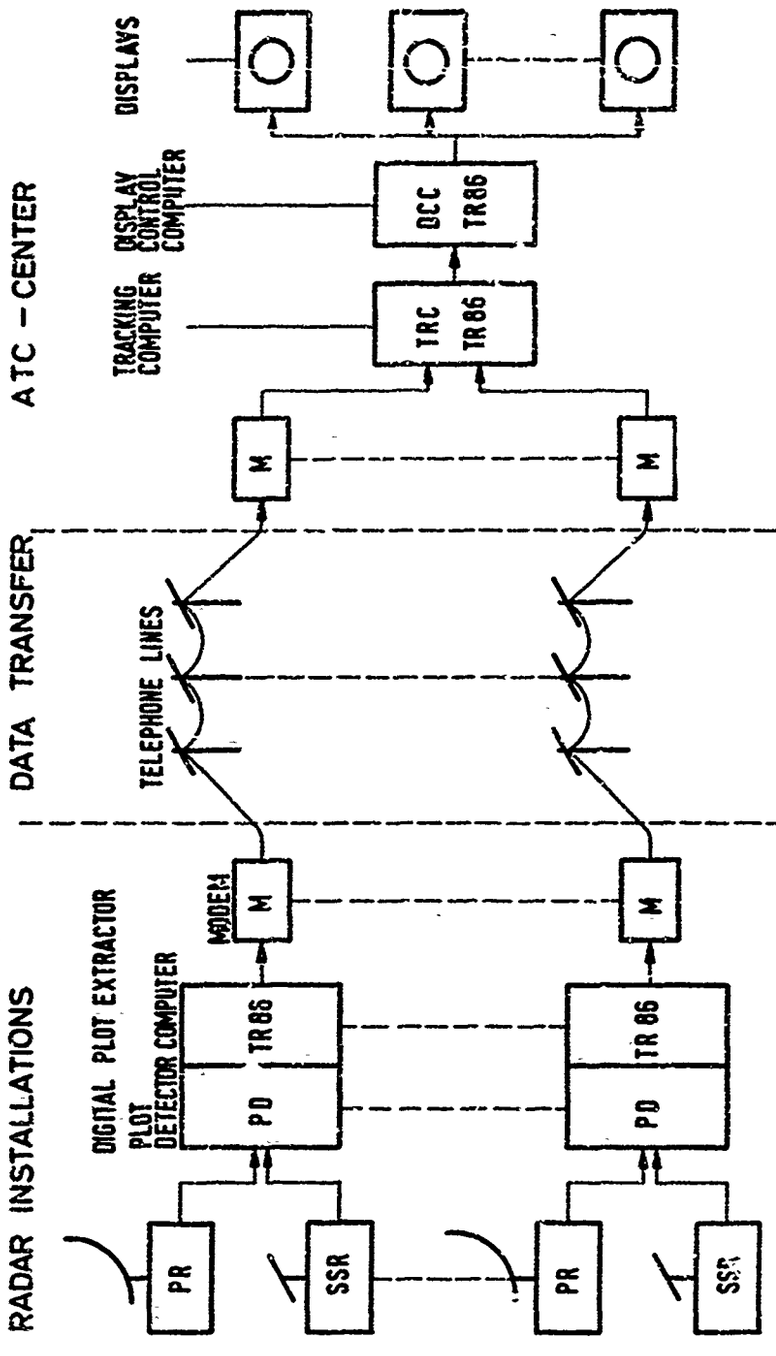


FIG. 1 BLOCK DIAGRAMM OF AN ATC-SYSTEM WITH AUTOMATIC RADAR DATA PROCESSING

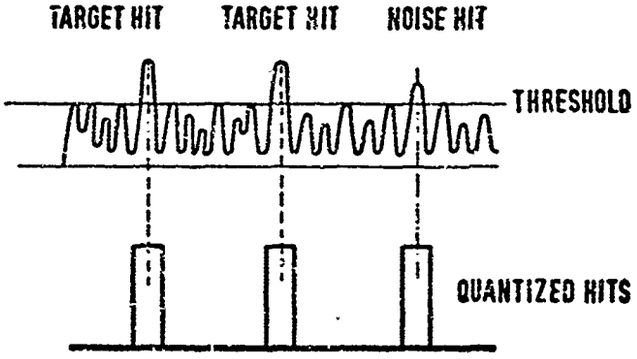


FIG.2 PRINCIPLE OF HIT DETECTION

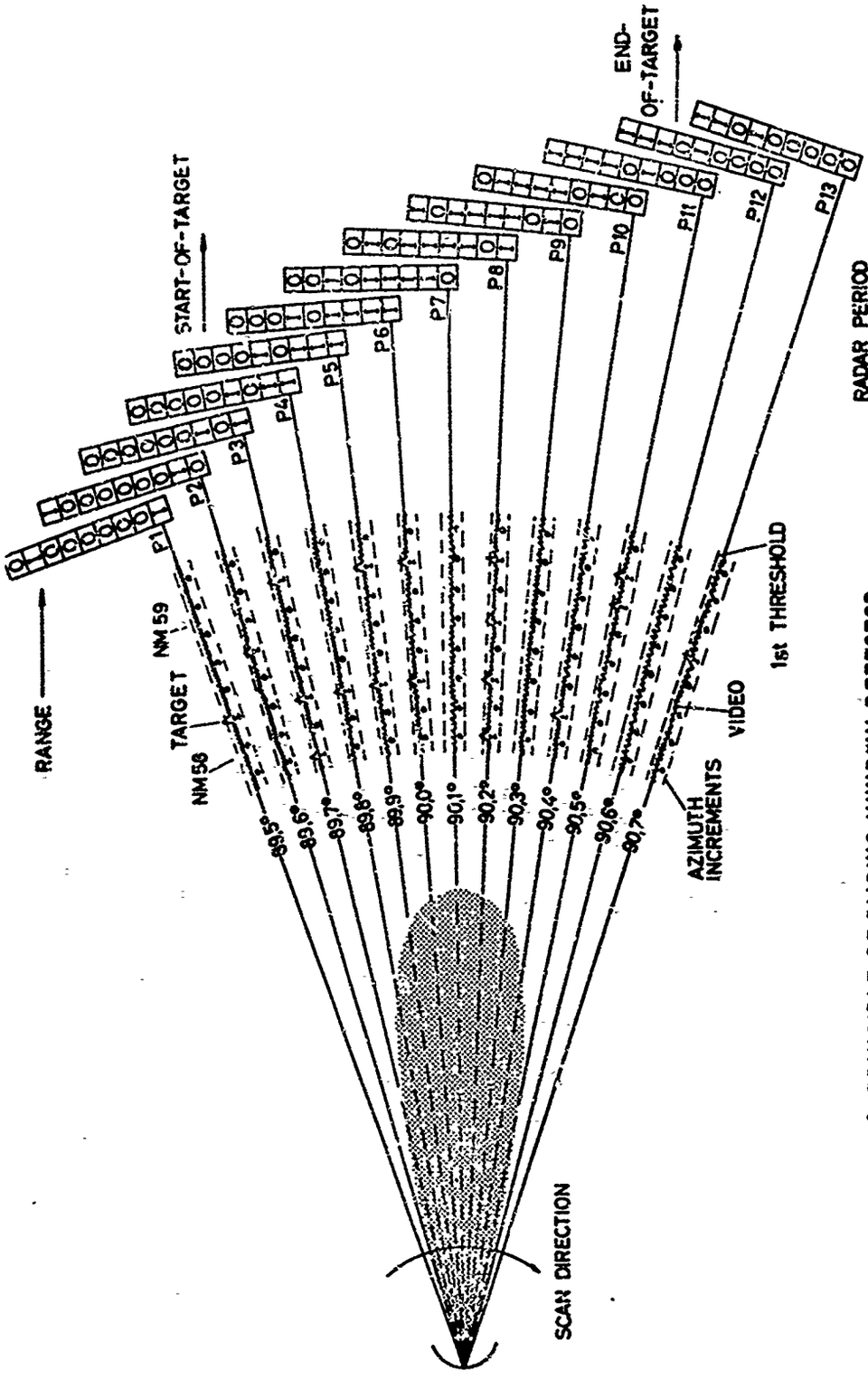


FIG.3 PRINCIPLE OF SLIDING WINDOW DETECTOR

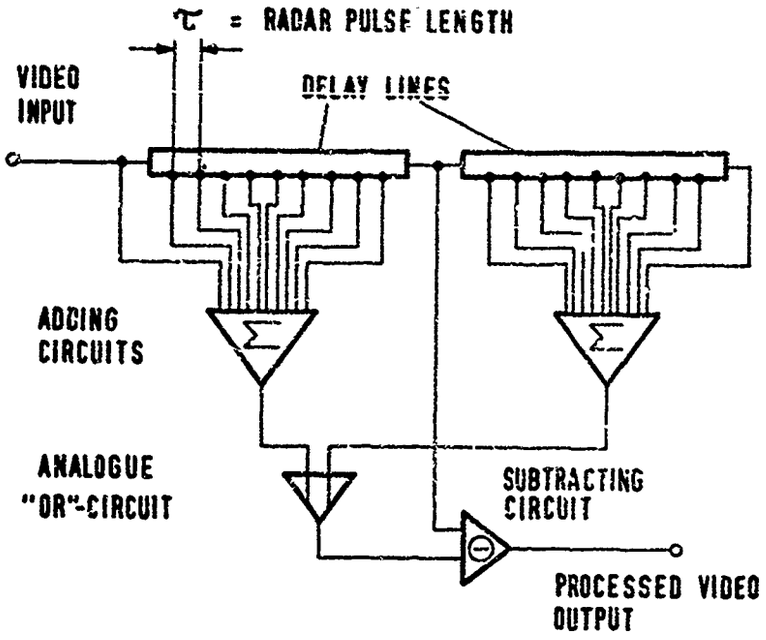


FIG. 4 RAIN CLUTTER SUPPRESSION CIRCUIT

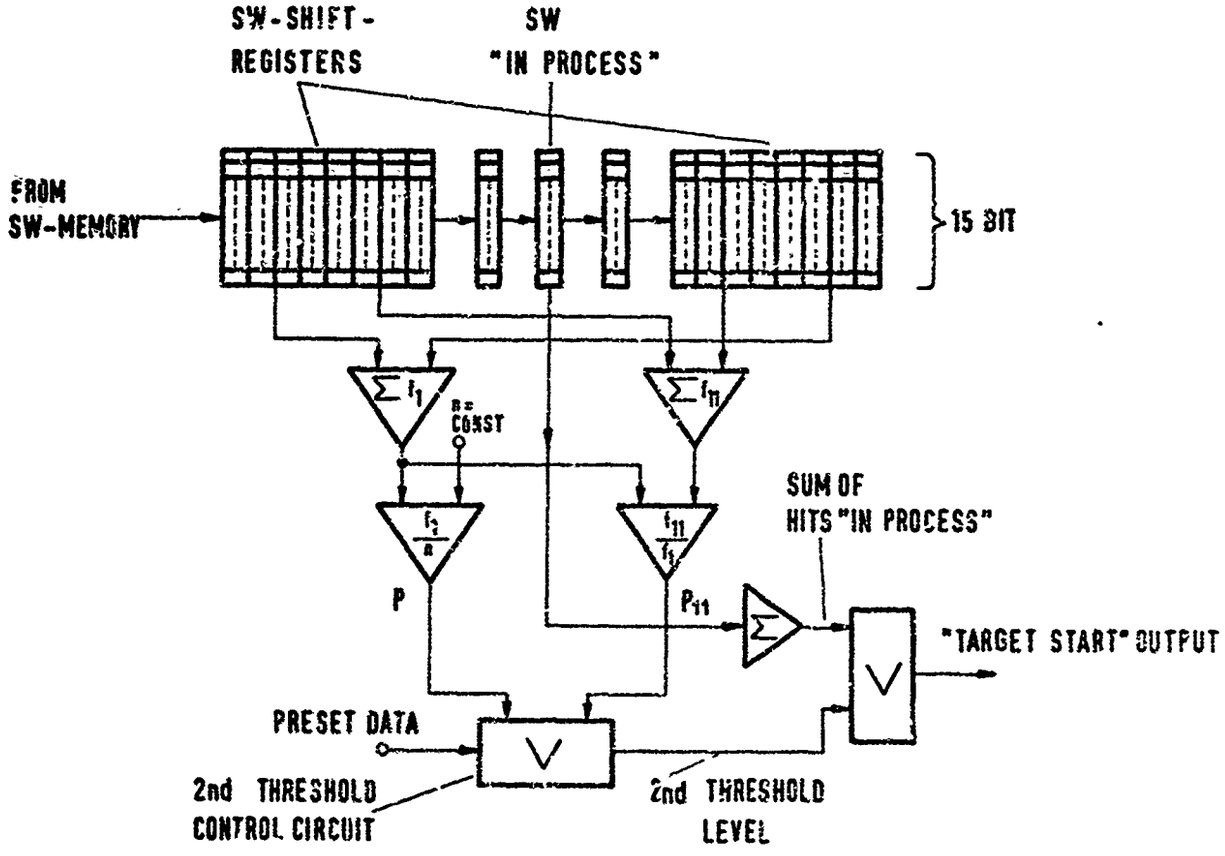


FIG.5 PRINCIPLE OF CLUTTER ELIMINATION CIRCUIT

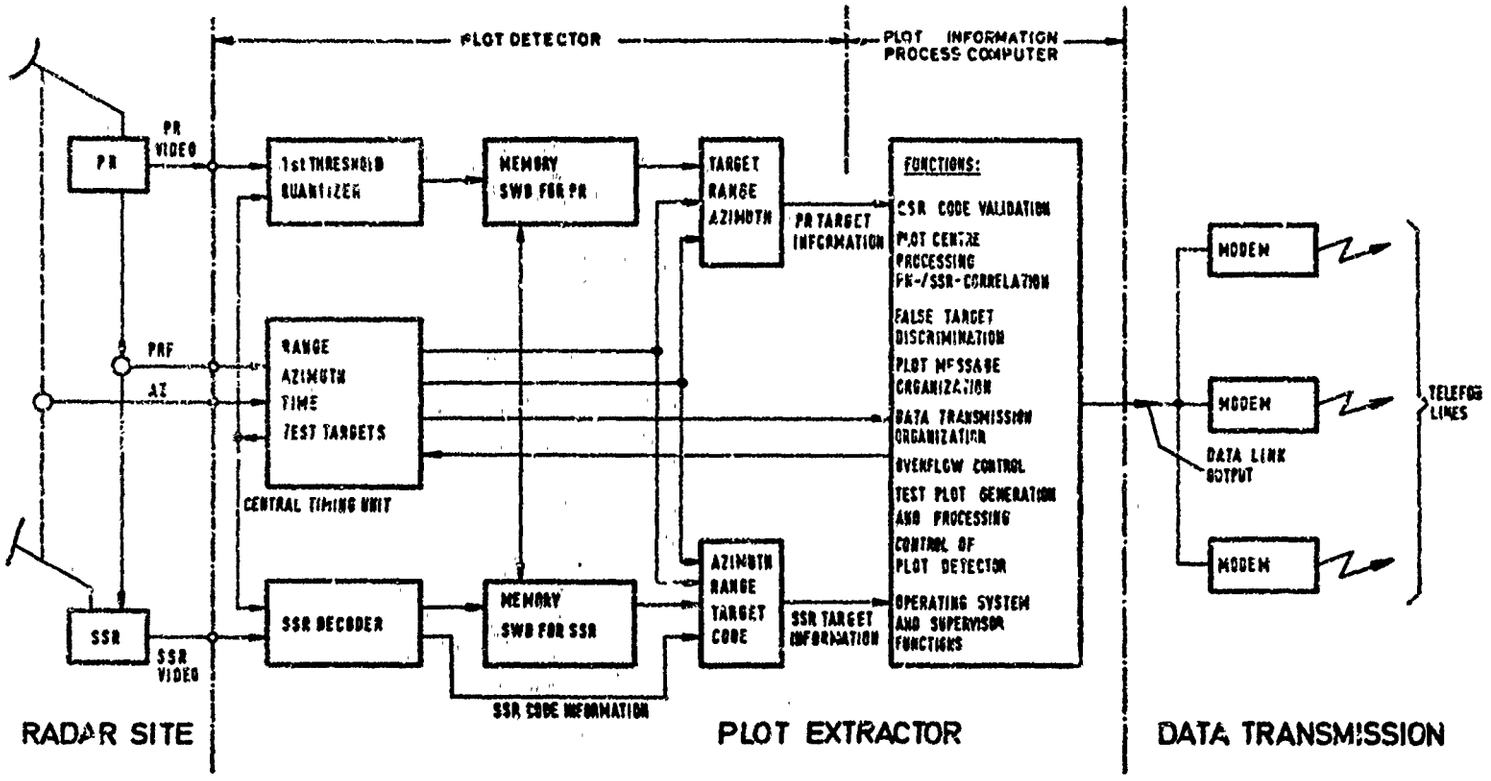


FIG. 6 COMBINED PR/SSR PLOT EXTRACTOR FUNCTION DIAGRAM

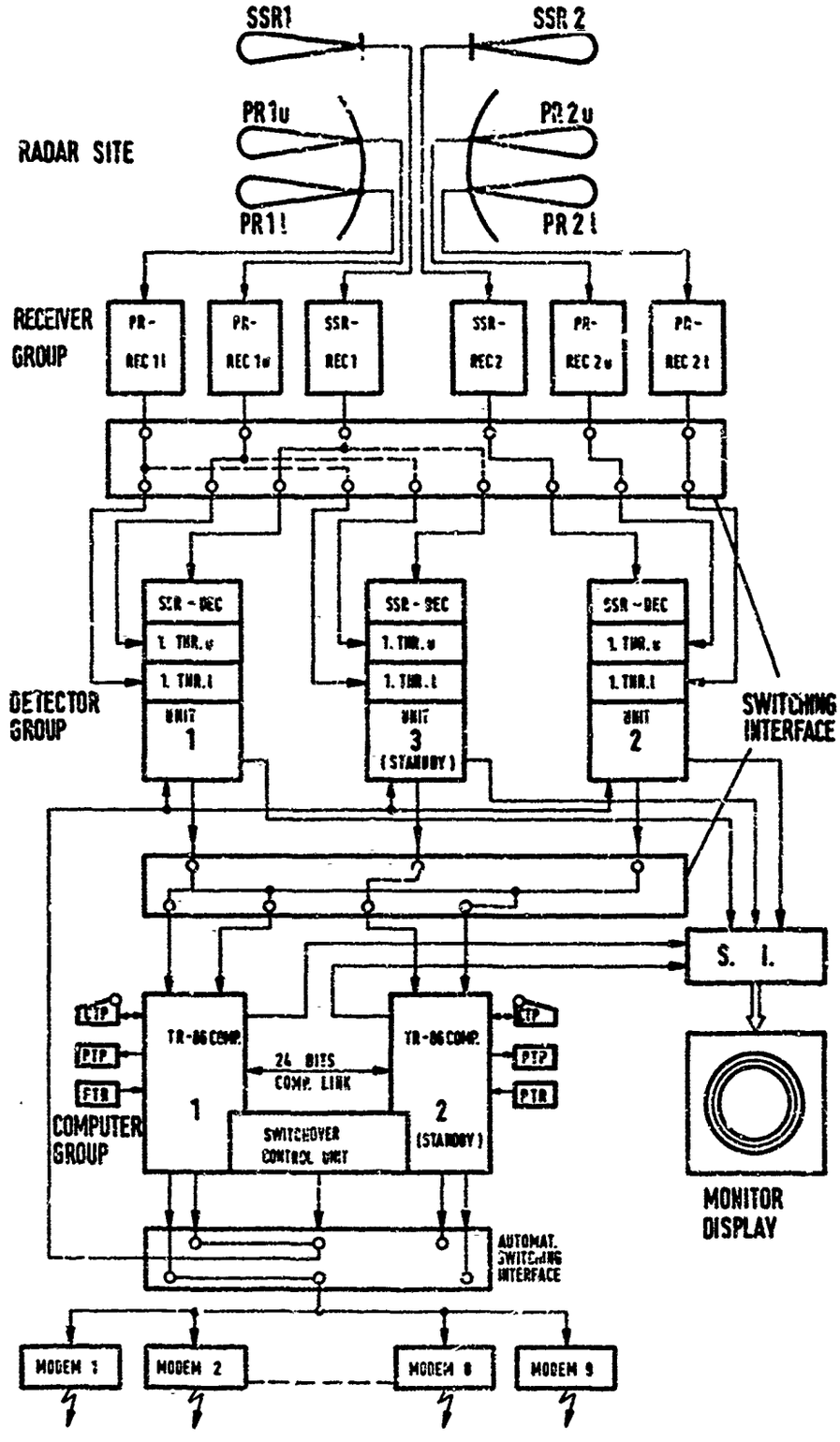


FIG. 7 BLOCK DIAGRAM OF PR/SSR PLOT EXTRACTOR FOR SRE-LL BACK TO BACK RADAR

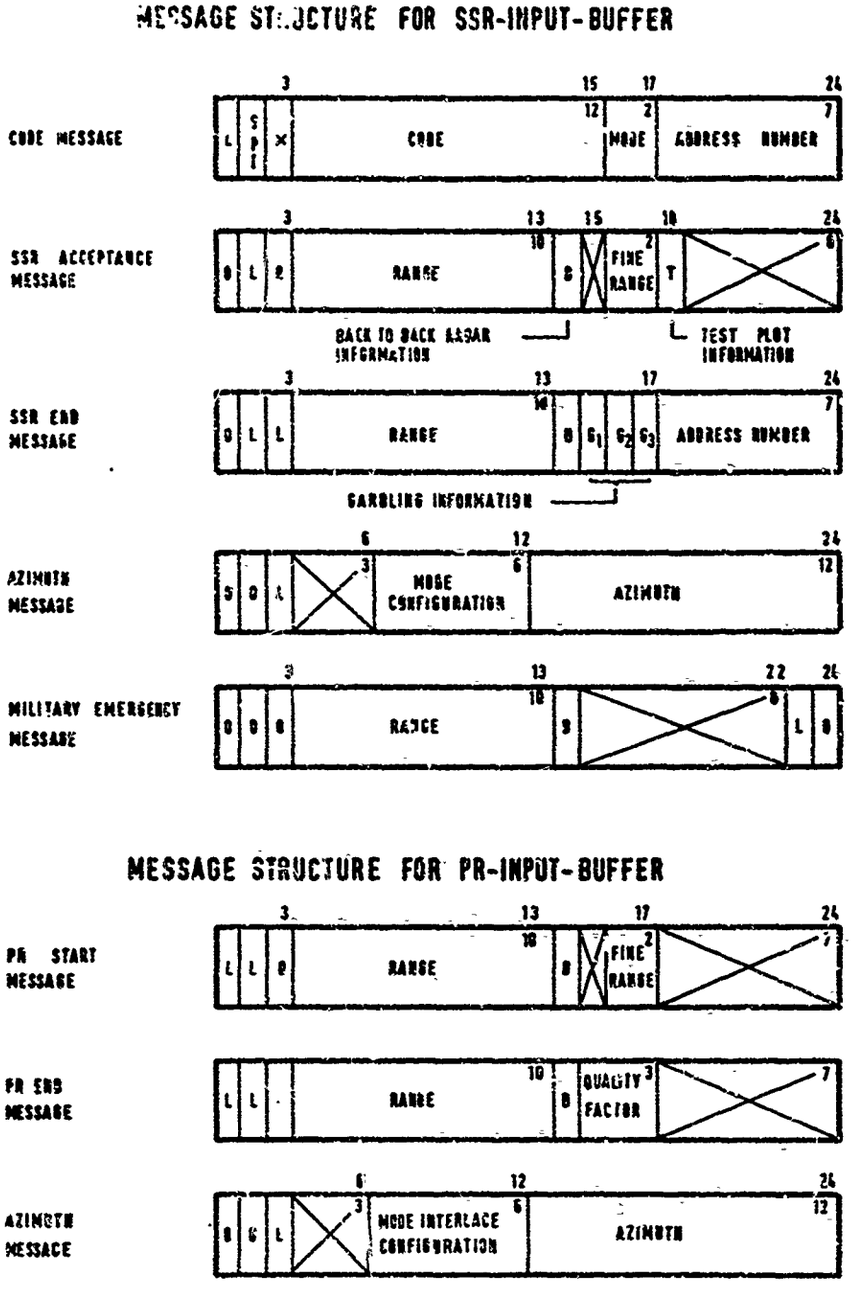


FIG. 8 STRUCTURE OF PLOT PART MESSAGES

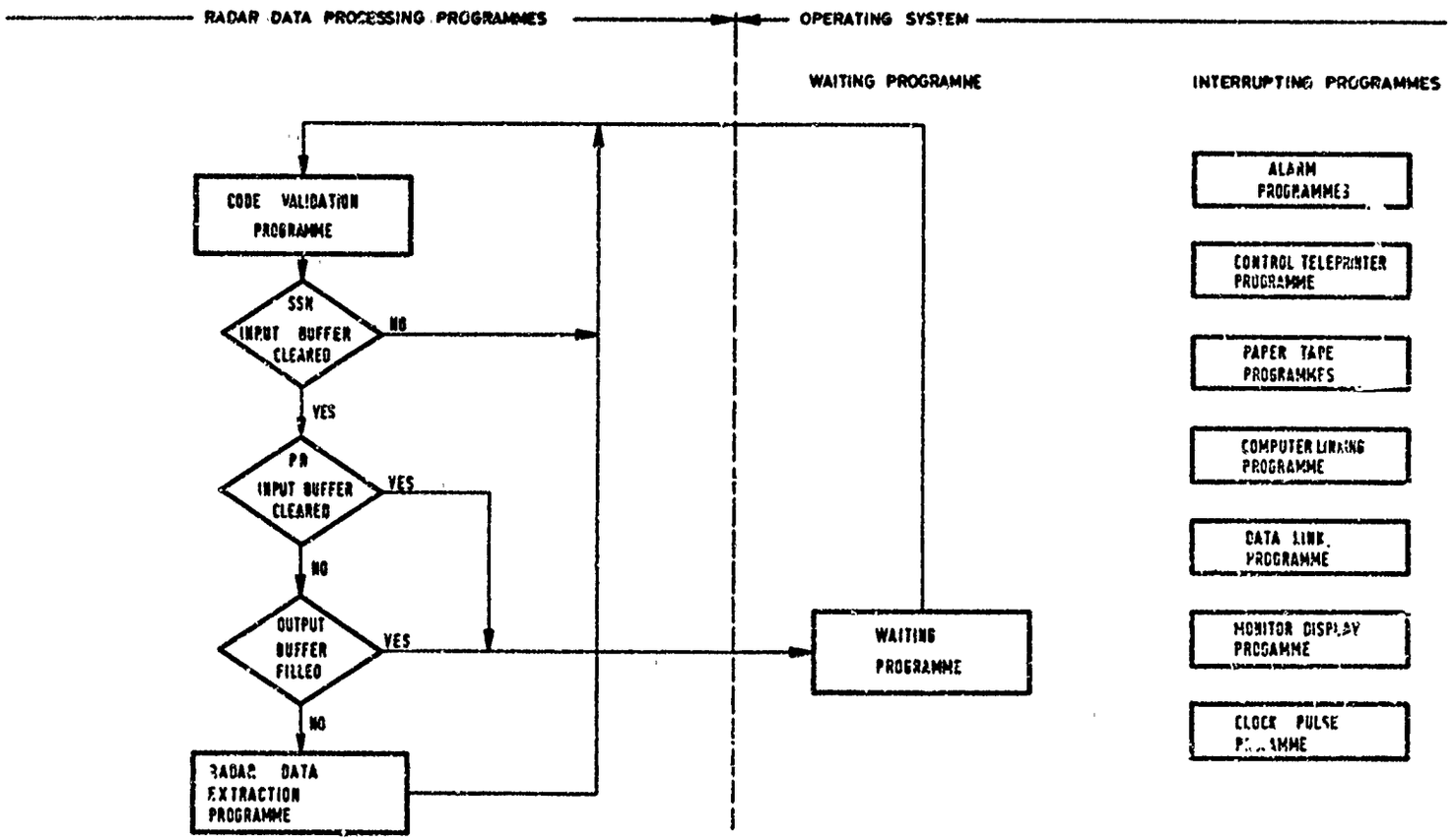


FIG. 9 BLOCK DIAGRAM OF SOFTWARE STRUCTURE

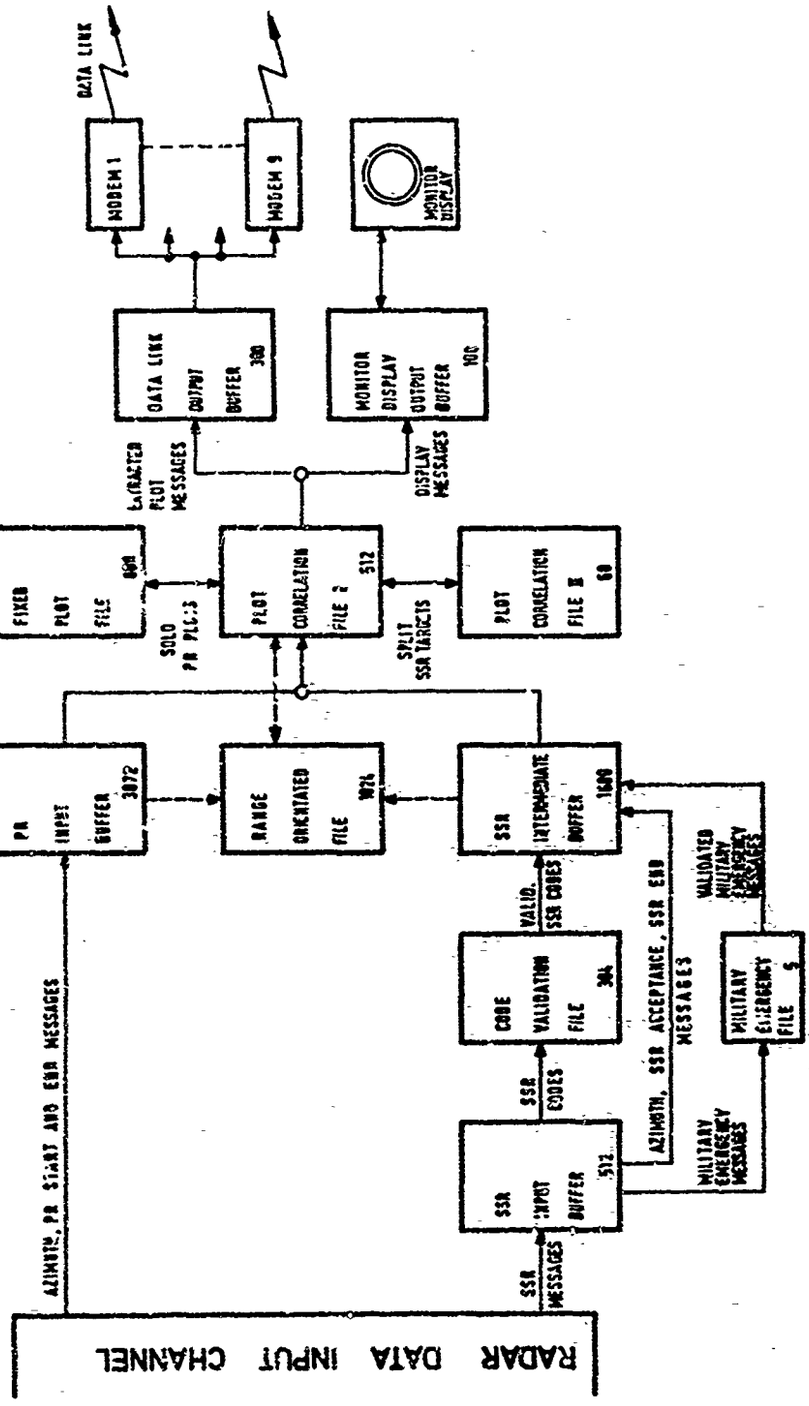


FIG.10 ARRANGEMENT OF FILES AND RADAR DATA FLOW

RUDAR - AN EXPERIMENTAL NOISE RADAR SYSTEM

by

J. A. Smit

Christiaan Huygenslaboratory
Noordwijk, Netherlands

RUDAR - AN EXPERIMENTAL NOISE RADAR SYSTEM

C.A. Swit
CHRISTIAAN HUYGENSLABORATORY
Noordwijk - Netherlands

SUMMARY

The results of experiments with a noise radar, performed in 1964, are given. Attention has been focussed on a flexible cross-section of a possible complete system. As a consequence of this we were able to check in a straightforward manner the system performance with relatively simple hardware. Several features of the system have been shown, viz.:
- the antenna system can be used as a space filter;
- there is an option for azimuth measurement by mechanical- or frequency-scanning.

A discussion will be given of the advantages and disadvantages of such a system.

1. INTRODUCTION

In the literature several references can be found quoting noise as a possible radar signal source (1, 2, 3, 4, 5,). However application of this type of signal in a system has, as far as we know, only been mentioned a few times (6, 7, 8). For this reason it may be of interest to give the results of our experiment with a noise radar performed in 1964. Prior to the present experiment we did some work on microwave correlators for measuring purposes. As an example a noise-reflectometer-system could be mentioned. In fact it is a radarsystem for short distances but with a high resolution (about 30 cm) and a high accuracy in target-distances (about 1 cm) and target strength (about 5 %). In fig. 39.1 an example of a measurement is given. Our further research was directed to the use of frequency dispersive antennas as a means to spread the available frequency spectrum in space. For when a broadband signal (e.g. a noise signal) is transmitted by an end fed linear array, the spectrum of this signal will be spread out in space owing to the antenna "squint" (fig. 39.2). In a direction θ, in the far field, the spectrum received, will have a center frequency specified by the direction θ and a form specified by the antenna pattern. The bandwidth of this spectrum with an antenna beamwidth of θ_{3dB} and a squintfactor of K °/MHz, will be: $B = \theta_{3dB}/K$(MHz). For detection of a signal that has been reflected by an object in a direction θ, the receiver should be matched to the signal:

a. in the frequency domain,
with a filter, with center frequency f_θ, and a form dependent on the beam pattern
b. in the time domain,
with a correlator, because we deal with a noise signal.

The block diagram of the noise radar with a dispersive system will be as given in fig. 39.3. In the ideal case a precise distance measurement can theoretically be done by means of the maximum of the fine structure. But as result of fluctuations in electromagnetic propagation and target motion, the delay time τ is subject to variations. Therefore an envelope correlator detector is required. Three possible configurations for envelope correlators are given in figures 39.4 a to c. In the appendix 7.1 more details of these circuits appear.

2. THE SCOPE OF THE EXPERIMENT

The object of the experiment was an investigation into a complete noise-radar system, making a cross-section of the system. In fig. 39.5 a block diagram of a complete system is given. Target bearing is measured by means of the filters F_1 to F_m, target range by means of the delays τ_1 to τ_n, while target velocity is measured by means of the doppler shift f_d, in filters f_{d_1} to f_{d_4}. The observed space is divided in n x m space-cells, and each cell will be divided in 1 doppler slots. A cross-section as mentioned can be given by:

a. one direction, given by F_k, and n delay steps, with n correlators and n x 1 doppler filters;
b. one distance, given by τ_j, and m bearing filters, with m correlators and m x 1 doppler filters;
c. one direction, given by F_k, and one distance, given by τ_j, with one correlator and one doppler slot, f_{d_i}.

If in the third alternative F_k, τ_j and f_{d_i} are made variable, a good impression of the behaviour of a complete system may be obtained. The scope of an experiment with such a flexible set-up is:

a. to test the principle of noise radar
  a.1. is it in fact possible to detect a target in very poor S/N conditions ?
  a.2. can a dispersive linear array be used as a space filter ?
  a.3. is it possible to resolve targets with different speeds in the same space-cell?
  a.4. what is the optimum bandwidth of the bearing filters?
  a.5. is it allowed to extrapolate the results of one space-cell to a system with n x m x 1 facets ?
b. to discover the drawbacks
  b.1. the detrimental effect of the cross-coupling between transmitter- and receiver antenna;
  b.2. the influence of nearby targets;
  b.3. the effects of spurious signals in the delay line.
c. measurements
  c.1. target position and velocity;
  c.2. cross correlation function;
  c.3. antenna squint characteristics.

3. THE SYSTEM

3.1 System requirements.

Another object of the experiment was to verify the expected theoretical relations of the various parameters. Therefore it is of little use to specify in advance center frequencies, bandwidth, etc. As a consequence of this, it was decided:
- that the maximum bandwidth of the receivers should not be limited by the choice of the delay line;
- to construct one filter-correlator combination, with the requirement that the center frequency

and bandwidth of the filter must be variable:
- to obtain the speed information from the fluctuation rate of the fine-structure of the correlation function.

From the above the principle of the system can be given. The method given in fig. 39.4-c is most suitable. The reason for this are the following:
- there are no image frequencies;
- the videobandwidth can be half the r.f.-bandwidth.

The video amplifiers (filters F_1 and F_2) are not made as real low-pass filters, but a cut-off frequency > f_1 is given for the lower frequencies. Then signals with a frequency f_1 are suppressed, they originate at:
- the S.S.B.-modulator, by crosstalk;
- the mixer, as a consequence of unbalanced amplitude modulation.

In fig. 39.6 the block diagram of the realized system is given.

3.2. Transmitter

The transmitter consists of a noise source and of two travelling wave tube amplifiers. As a noise source a fixed-frequency magnetron type JP 9-7A was used. It is possible to bring a magnetron into a condition of high noise output without stable or unstable oscillations. The best results were obtained by shunting the magnetron magnet. This shunt has a great influence on the spectrum form, while both the anode voltage and heater current will effect the output power. With a noise figure of 60 dB, a bandwidth of 200 MHz, and an amplifier gain of about 57 dB, the available output power was 350 mW.

3.3. Antenna-system

The antenna-system consists of two identical end-fed slotted wave guide aerials, with a length of 5,5 m. They are mounted with a special frame on a gearbox with servo-drive (fig. 39.7). Each antenna has a beamwidth of 0,4° between the -3 dB points, a sidelobe level of -26 dB, and a gain of 35 dB. The squintfactor of this type of antenna is 8.83×10^{-3} °/MHz. With a uniform noise spectrum, the transmitted spectrum in a direction ϑ, will be:

$$B_\vartheta = \frac{\alpha_4}{g\vartheta_3}\,10^{-3} = 45\ \text{MHz}$$

After reception with an identical antenna, the bandwidth is reduced to 22 MHz. The cross-coupling between antennas was about -70 dB, mainly due to the two-channel rotating joint. It could be improved to approximate -110 dB.

3.4. Delay line

To meet the specification for the delay line such as: a broad bandwidth of 200 MHz and a delay of at least 20 μ sec., a microwave delay line was used. This consists of a non-dispersive transmitting and receiving antenna system, a propagation path reasonably free of secondary obstructions, a reference target and a travelling wave tube amplifier. In this delay line spurious responses were maximally 18 dB below the wanted signal. For this reason there was some inaccuracy in the measurements of correlation functions.
A variable delay line was used in the video amplifiers; with a maximum delay line of about 0.5 μsec in steps of 6.6. μsec (= 1 m).

3.5. Correlator performance

The transfer function of the receivers are given in fig. 39.8. With a uniform spectrum on the input of the mixers, the measured envelope of the cross correlation function is as given in fig. 39.9. The lack of symmetry indicates that the phase-transfer functions of the two amplifiers are not identical. Computer calculations give the phase difference function arg H(f) of fig. 39.10.

There was no time to correct this difference. The sensitivity of the correlator receiver is dependent on the two input signal levels. In fig. 39.11 this dependence is given for an output signal-to-noise ratio of 10. From this figure it can be seen that there is an optimum level for the reference channel of about -50 dBm. This can be explained by the fact, that for small reference signals the conversionloss increases, while for greater reference signals the self-noise increases.

The output signal-to-noise ratio is dependent on the system function. If the power of the output signal- and noise-components are measured behind the low-frequency bandpass filter, the $(S/N)_{out}$ is (9):

$$\left(\frac{S}{N}\right)_{out} = \frac{2B}{B_1}\,\frac{\frac{1}{2}\left(\frac{S}{N}\right)_1\left(\frac{S}{N}\right)_2}{\left(\frac{S}{N}\right)_1+\left(\frac{S}{N}\right)_2+\left[1+\left(\frac{S}{N}\right)_1\right]\left[1+\left(\frac{S}{N}\right)_2\right]}$$

where:

B = the effective bandwidth of the video receiver; (= 15,7 MHz);
B_1 = the effective bandwidth of the intergration bandpass filter; (= 60 Hz);
$(S/N)_1$ = correlator input signal-to-noise ratio in signalling channel;
$(S/N)_2$ = correlator input signal-to-noise ratio in reference channel.

The input noise is partly of the thermal origin and partly due to the cross-coupling effect:

$$N_{in} = N_t + \alpha \overline{P_z}$$

where:

N_t = input thermal noise power
α = the cross-coupling factor
$\overline{P_z}$ = average-transmitted signal power.

In view of this the following signal-to-noise values were available:

$$10 \log \left(\frac{S}{N}\right)_1 = -35\ \text{dB}$$

$$10 \log \left(\frac{S}{N}\right)_2 = -10\ \text{dB}$$

The output signal-to-noise ratio in this case is:

$$10 \log \left(\frac{S}{N}\right)_{out} = +12\ \text{dB}$$

In the correlator circuit itself we found a maximum disagreement between theory and practice of about -4 dB, due to the 1/f noise from the diodes in the multiplier circuit. Therefore a minimum output signal-to-noise ratio of +8 dB is to be expected.

3.6. System possibilities

The realized system possibilities are:

a. Target detection under very low signal-to-noise conditions;
b. The antenna-system is used as a space filter;
c. As a result of (b), there is an option for azimuth measurement by antenna rotation or by local oscillator frequency tuning, or a combination of both;
d. For distance measurement there is a choice between either microwave-, i.f.-, or video delay lines, or a combination of the same.
e. By proper selection of signal-bandwidth the accuracy of distance- and bearing measurement can be exchanged;
f. Radial target velocity is obtained by measuring the doppler-shift of the microwave signal. A reference doppler-shift can be introduced to measure different target velocities with a fixed filter.

4. RESULTS

4.1. Target detection

For this measurement we used a conveniently situated fixed target. Detection can be ac-

complished by scanning in distance as well as in azimuth. Figure 39.12 gives the result for azimuth scanning and a fixed delay. Figure 39.13 the same for distance scanning and fixed bearing. From these figures it can be seen that a signal-to-noise ratio of at least + 8 dB has been realized. It has been therefore proved possible to detect a target with a noise radar.

4.2. Frequency scanning

For a linear array the relation between the direction of the mean beam (with respect to the normal) and the frequency (or wavelength) of a C. W. - signal is given by:

$$\sin\theta = \frac{\lambda}{\lambda_g} - \frac{\lambda}{\lambda_{g_o}} \qquad \text{(fig. 39.3)}$$

An approximation for small angles is:

$$\theta \approx K\,\Delta f$$

with:

$\Delta f = f - f_{oo}$

f_{oo} = frequency for beam direction $\theta = 0$

K = squint factor : in this case: $8.83\ 10^{-3}$ °/Mhz

In fig. 39.14 the results of the measurements are given where in stead of a C.W.-signal, a noise signal with a bandwidth of 14 Mhz id used. The measured points of the mean beam direction are given along with the theoretical curve. The agreement is quite good. The measured diagrams of the antenna system as a space filter are shown in fig. 39.15. The signal bandwidth affects the accuracy of the bearing measurements. As long as the bandwidth of the receiver filter is smaller than the equivalent bandwidth of the antenna system, there is little loss in directivity. In the other cases a loss must be expected as can be seen in fig. 39.16. It will be obvious from the above that there is a possibility to exchange accuracy in distance and direction. More details will be given in appendix 7.2. To check the correspondence of frequency- and mechanical azimuth scan we did two measurements:

a. with the antenna aimed at the target, the local-oscillator was swept over a broad frequency band, and the correlator output was measured;
b. with the local-oscillator at a fixed frequency the antenna was rotated.

Figure 39.17 shows the results obtained.

4.3. Doppler measurements

To get an insight in the behaviour of the system on moving targets we used a reference target that causes a doppler shift of 230 Hz. The same doppler filter can be used, only the audio-oscillator must be tuned to $(f_1 \pm 230)$Hz. In fig. 39.18 the results are illustrated. With the method used it is possible to detect and measure the speed of moving targets against a background of high level fixed objects. The crosstalk between the doppler filters is given by the doppler filter characteristics.

5. DISCUSSION OF THE RESULTS

On the base of these results it now becomes possible to check in how far the objects of the experiment have been met:

5.1. Test of the principle of noise radar

5.1.1. It is indeed possible to detect a target in very poor S/N conditions, and that even with non-optimum hardware with a $(S/N)_{in} = -35$ dB the $(S/N)_{out} \approx +12$ dB.

5.1.2. A dispersive linear array can be used as a space filter; even for a broadband noise signal. The measured relation between direction and frequency is in good agreement with the theoretical curve, provided the bandwidth of the transmitted spectrum is much greater than the receiver bandwidth. Mechanical- and electrical-scanning gave almost the same results. However the electrical-scanning method is very sensitive to interference, originating from spurious signals in the delay-path.

5.1.3. Targets in the same space-cell but with different speeds can be resolved by measuring the doppler shift. The crosstalk between the different doppler channels is given by the filter characteristics.

5.1.4. In appendix 7.2 a relation is given for the optimum bandwidth of the bearing filter. There will be little loss in bearing accuracy as long as the equivalent bandwidth of the videofilter is smaller than or equal to the antenna "bandwidth": $B \leq \theta_{-3dB}/2K\sqrt{2}$. This agrees quite well with the measurements. The antenna pattern broadens as soon as the bandwidth of the bearing filter is greater than the bandwidth of the received signals. This results in fewer independent directions.

5.1.5. Bearing in mind, that we would like to extrapolate the present results on one space cell with one doppler-slot (one facet) to a system with n x m x l facets, we can remark that:

- it is possible to get an insight in the behaviour of a complete system by comparing the different measurements for one facet;
- a specification can be given of filter- and amplifier-bandwidth, of amplification factors, of minimum noise factors, of maximum allowable spurious signals in delay lines, etc.
- however, the dynamic behaviour of such a system of n x m cells can not be predicted from these experiments;
- for the design of a more complex system the know-how of one facet is in fact essential; but multiplying the hardware of one to n x m x l - facets will not necessarily provide an optimum system.

5.2. The system drawbacks

The crosscoupling between transmitter- and receiver-antenna was about - 70 dB. Because the transmitted power was not so high, the self-noise was below the 1/f-noise of the correlator. This was also true for the nearby targets. However the spurious signals in the delay line on the one hand and the crosscoupling together with reflections from nearby targets on the other hand gave correlated output signals. Especially with frequency scanning this effect was rather severe. In fig. 39.19 the correlator output signal is given for frequency- resp. mechanical scanning. The interference effect of the target signal with the spurious signal is obvious from fig. 39.19-a. The correlator output signal in this case is (9):

$$S_o(\Delta\tau_1, \Delta\tau_2, f_d) = \frac{1}{2}\left[|J_1(f_d)|^2 + |J_2(\Delta\tau_2, f_d)|^2 + 2|J_1(f_d)||J_2(\Delta\tau_2, f_d)|\cos x\,(\phi_{f_d} + \phi_{f_d} \cdot (\Delta\tau_2 - \Delta\tau_1)\right]$$

with: $|J_1(f_d)|^2$ = envelope of correlation function for the target

$|J_2(\Delta\tau_2, f_\Delta)|^2$ = envelope of correlation function for the spurious signal;
$\Delta\tau_1$ = delay difference for target signal;
$\Delta\tau_2$ = delay difference for spurious signal;
f_Δ = $f_o - f_c$
f_c = local oscillator frequency
f_o = center frequency of antenna.

A 2 π phase variation in the interference term is obtained for a local oscillator variation of:

$$f_{c_1} - f_{c_2} = \frac{2}{\Delta\tau_1 - \Delta\tau_2}$$

In our case there was a delay difference of $\Delta\tau_2 - \Delta\tau_1 = 0{,}33$ μsec, that would give a 2π phase variation for every 6 MHz of local oscillator variation. This agrees quite well with the measured results. For a 1 % output fluctuation and a $(S/N)_{in} = 10^{-3}$, the spurious signal in the delay line must at least be 50 dB below the wanted signal.

5.3. The measurement accuracy

The accuracy of the measurement depends on the output signal-to-noise ratio and on the presence of interfering objects. In view of 5.2., both were not too good unfortunately. Differences of 1 meter in the distance to one target could easily be measured. The resolution between two targets was about 30 meter. The measured antenna squint characteristic agrees well with the theoretical curve. Deviations were caused by the non-uniform transmitter spectrum.

6. CONCLUSION

From the discussion of the measurements it has been shown that a noise radar can be realized. Also is has been proved that the spectrum of a noise signal can be spread out in space with a dispersive array and that position (distance, bearing) and velocity of a target can be measured.

The difficulties that appear in the realisation of such a system are many. This experiment has been very useful to recognize these difficulties, to learn their importance, and if possible to deal with them.

The results obtained with a dispersive linear array have a more general importance in that they can easily be extrapolated to the case of a two-dimensional dispersive array.

In such a manner the broadband noise signal can be dispersed in two independent co-ordinates viz. azimuth (θ) and elevation (ψ). Every direction in space, defined by θ and ψ is uniquely determined by a specific center frequency of a spectrum, characterized by the squint factors K_θ and K_ψ. The form of the spectrum as before will depend on the antenna pattern. So it should be possible to obtain in one measurement target position in three dimensions (τ, θ, ψ).

Attractive possibilities arise when combining mechanical- and frequency-scanning. For instance θ mechanical scan and ψ frequency scan.

Alternatively when scanning in one co-ordinate the combination in mechanical- and frequency-scan offers the possibility to improve the dwell-time on a target.

A further possibility is to make the squint factor K, which was so far supposed to be a constant, a system variable. Every type of signal has its own limitations.

As is known from the literature the usefulness of a random noise signal, as every continuous signal, is limited by the target environment.

Methodr to overcome this limitation have been studied (4, 12).

The systems hardware could be another limitation.

The decision whether to use a random noise signal for a particular application can only be arrived at by careful consideration of its weighted advantages and disadvantages.

It was the object of this paper to demonstrate the feasibility of the use of noise as a radar signal, and to provide sufficient insight for visualized applications.

7. APPENDIX

7.1. Microwave envelope-correlators

In the following a discussion will be given of three methods to construct an envelope-correlator. These methods are shown in figures 39.4 - a to c. We shall use the complex-signal notation (10, 11).

7.1.1. Single-sideband modulator and bandpass correlator

In fig. 39.4a the block diagram is given. In this relative simple structure the filters F_1 and F_2 are supposed to be microwave filters. Filter F_1 selects a frequencyband from the reference signal. The combination of the dispersive antenna and filter F_2 provides the bearing of the detected signals. It will be shown that this circuit yields the envelope of the correlation function. After the filters F_1 and F_2 we have the signals

$$s_1(t+\tau_1) = Re\left[u_1(t+\tau_1)e^{j2\pi f_o(t+\tau_1)}\right]$$

$$s_2(t+\tau_2) = Re\left[u_2(t+\tau_2)e^{j2\pi f_o(t+\tau_2)}\right]$$

$u_1(t)$ and $u_2(t)$ are low frequency functions with respect to the carrier f_o. The two signals originated from the same source and have different delay paths τ_1 and τ_2.

$x_A(t)$ is the output signal of the single-sideband mixer:

$$x_A(t) = \left[a(t)\cdot s_1(t+\tau_1)\right]_{f_o+f_m}$$

with:

$$a(t) = A\,Re\left[e^{j2\pi f_m}\right]$$

$$x_A(t) = \frac{A}{2}\,Re\left[u_1(t+\tau_1)e^{j2\pi(f_o+f_m)(t+\tau_1)}\right] + \gamma\frac{A}{2}\,Re\left[u_1(t+\tau_1)e^{j2\pi(f_o-f_m)(t+\tau_1)}\right]$$

where $\gamma \leq 1$, and gives the suppression of the unwanted sideband. The multiplier circuit gives:

$$y(t) = \left[s_2(t)\cdot x_A(t)\right] = \frac{A}{4}\,Re\left[u_1(t+\tau_1)u_2^*(t+\tau_2)e^{j2\pi f_m(t+\tau_1)}e^{j2\pi f_o(\tau_1-\tau_2)}\right] + \gamma\frac{A}{4}\,Re\left[u_1^*(t+\tau_1)u_2(t+\tau_2)e^{j2\pi f_m(t+\tau_1)}e^{-j2\pi f_o(\tau_1-\tau_2)}\right]$$

The integration bandpassfilter has an impulse reponse

$$h(t) = Re\left[h_o(t)\,e^{j2\pi f_m t}\right]$$

with $h_o(t)$ a real function.
After this filter we get:

$$g(t) = Re\,\frac{1}{2}\left[\int_{-\infty}^{\infty} h_o(\xi)\left\{x_A(t-\xi)\cdot s_2(t-\xi)\right\}_- e^{j2\pi f_m \xi}\,d\xi\right]$$

The following relations are used:

$$u_1(t+\tau_1) = \int_{-\infty}^{\infty} M_1(f)\,e^{j2\pi f(t+\tau_1)}\,df$$

$$u_2(t+\tau_2) = \int_{-\infty}^{\infty} M_2(f')\,e^{j2\pi f'(t+\tau_2)}\,df'$$

$$Z(f) = \int_{-\infty}^{\infty} h_o(\xi)e^{-j2\pi f\xi}\,d\xi = 2H(f), \text{ if } f>0 \quad = 0, \text{ if } f<0$$

For the "signal" part of g(t) we could write, with: $f = f$; $Z(o) = 2$, and $\tau_1 = \tau_2 = \tau$:

$$g_{sign}(t) = R_e\left[\frac{A}{4}\int_{-\infty}^{+\infty} M_1(f)M_2^*(f)e^{j2\pi f\tau}\,df\cdot e^{j2\pi\{f_m(t+\tau_1)+f_o\tau\}} + \right.$$
$$\left. + \gamma\frac{A}{4}\int_{-\infty}^{+\infty} M_1^*(f)M_2(f)e^{-j2\pi f\tau}\,df\cdot e^{j2\pi\{f_m(t+\tau_1)+f_o\tau\}}\right]$$

The signals $s_1(t)$ and $s_2(t)$ originate from the same source, but have traversed different paths with frequency amplitude characteristics $H_1(f)$ and $H_2(f)$:

$$M_1(f) = W_s(f)\cdot H_1(f)$$

$$M_2(f) = W_s(f)\cdot H_2(f)$$

where: $W_s(f) = \psi_s(f+f_o)$, is the spectrum of the complex envelope of the source signal:

$$\Psi_s(f) = 2S(f) \qquad f>0$$
$$= 0 \qquad f<0$$

Substitution of these functions will give for the envelope of $g_{sign}(t)$:

$$|g_{sign}(t)| = A\sqrt{1+\gamma^2+2\gamma\cos\alpha}\cdot\sqrt{R_e^2 J(\tau)+I_m^2 J(\tau)}$$

with: $\operatorname{tg}\alpha = \dfrac{I_m J(\tau)}{R_e J(\tau)}$

With a detector after the bandpass filter the envelope of the correlation function can be measured. To scan in azimuth the filters F_1 and F_2 must be varied synchronous.

The S.S.B.-modulator must be suitable for a bandwidth of 200 MHz.

The multiplier circuit is a balanced mixer with microwave mixer-diodes.

For a frequency azimuth-scan this circuit also must have a bandwidth of 200 MHz.

An acceptable conversion loss can be obtained with 1 mW of reference signal power on the output of the S.S.B.-modulator. The multiplier circuit must be well balanced, especially so to eliminate the mix-products of the spectrum of the strong reference signal. An unbalance better than 25 dB below the input signal is not to be expected for signals with bandwidths of 20 MHz lying within a band of 200 MHz.

Therefore it is recommended to choose the center frequency of the narrow bandpass filter outside this spectrum.

A suppression of the unwanted sideband in the S.S.B.-modulator, better than 20 dB is not to be expected, so the measured envelope of the correlation function has a ripple of about ± 10 %.

A disadvantage of this S.S.B.-modulator is the small permissable signal level, the own noise contribution, and the conversion loss of 8 à 10 dB.

From the results of experiments with a set-up according to this method, we found that for an output signal-to-noise ratio of 10 dB, an interference free input signal level of about - 70 dB, (with a bandwidth of 20 MHz) is needed.
This method is not suitable for a radar system because:
- the sensitivity is to low;
- the filters F_1 and F_2 must be tuned synchronously and with high precision.

7.1.2. Combination of a normal correlator with a "90°"-correlator.

In fig. 39.4-b an example is given. A mixer circuit in front of the multiplier circuit shifts the spectra of the signals to a band where it is possible to use conventional amplifiers.

These amplifiers could also be used as selective directional filters. When choosing the intermediate frequency care must be taken to eliminate the image frequency band.

Therefore pre-selection filters will be necessary. In principle, azimuth scanning is possible by tuning of the local-oscillator. It is possible to use an i.f.-delay line. A special case, that approaches the third method mentioned in 7.1.3., arises when f_o is in the middle of the spectrum of s(t).

The filters F_1, F_2 and F_3, and the amplifiers become low pass circuits with bandwidths that are about half the bandwidth of the incoming signals. It is now even possible to use video delay lines. Moreover the problem of the rejection of the image frequencies is thereby avoided.

It is very important to equalize the conversion loss in the multiplier circuit and the amplification in channels 2 and 3.

The construction of a circuit for squaring correlation functions with sufficient stability and dynamic range is not easy. The output signal is the square of the envelope of the correlation function. A derivation will be given below.

The input signals are again $s_1(t+\tau_1)$ and $s_2(t+\tau_2)$ (see 7.1.1.).
These signals are mixed with:

$$a(t) = A R_e\left[e^{j2\pi f_c t}\right]$$

$$b(t) = A\, I_m\left[e^{j2\pi f_c t}\right]$$

The output signals of the mixers are filtered in the i.f.-amplifiers, where upon the multiplications $x_A(t).x_B(t)$ and $x_A(t).x_B(t)$ are made. The derivation is analogous to the method given in 7.1.1. With an integration time large with respect to the correlation time of the signals and $H_{12}(f) = H_{22}(f)$, it is easy to show that the output signal becomes:

$$u_{out}(\tau) = \left(\frac{A^2}{4}\right)^2\left[R_e^2 J(\tau) + I_m^2 J(\tau)\right]$$

with:

$$J(\tau) = \int_{-\infty}^{+\infty} M_1(f)M_2^*(f)H_{12}(f)H_{12}^*(f)e^{j2\pi f\tau}\,df$$

7.1.3. Bandpass correlator with video amplifiers

The correlator receiver element for the present radar system has been build according to this principle (fig. 39.4-c).
Again the input signals are $s_1(t+\tau_1)$ and $s_2(t+\tau_2)$.

The local oscillator signals are:

$$a(t) = A\,R_e\left[e^{j2\pi f_c t}\right]$$

and a signal that is shifted in frequency by a S.S.B.-modulator over f_1Hz ($f_1 < 1$ MHz).

$$b(t) = B\,R_e\left[e^{j2\pi(f_c+f_1)t}\right]$$

The frequency-difference terms in the mixer output signals will be passed by the video filters. After the video amplifiers, with frequency amplitude characteristics $H_{12}(f)$ and $H_{22}(f)$ we get the signals $x_A(t)$, resp. $x_B(t)$. The signal $(x_A(t).\ x_B(t)$ is integrated in a bandpass filter tuned to f_1. The envelope of the signal term of the output signal is:

$$|g_{sign}(t)| = \frac{AB}{2}\sqrt{R_e^2 J(\tau)+I_m^2 J(\tau)}$$

with:

$$J(\tau) = \int_{-\infty}^{+\infty}|S(f+f_c)|^2 H_1(f)H_{12}(f)H_{12}^*(f)H_{22}^*(f)e^{j2\pi f\tau}\,df$$

In the video amplifiers it is possible to delay the signals by a video delay line. The bandwidth of this delay line may be half the bandwidth of the received signals. By choosing f_c somewhere inside the spectrum of the signals $w_1(t)$ and $w_2(t)$, $s_1(t)$ and $s_2(t)$ are selected and band limited by the low

pass filters F_1 and F_2. The correlator input signals are selected rather with filters symmetric about f_c and with a pass band given by F_1 and F_2. Azimuth scanning is obtained by tuning of the local oscillator.

The frequency shift f_1 is essential for the determination of the envelope of the correlation function.

Without this frequency shift, the correlation function with fine structure would be measured as a D.C.-signal. Choosing f_1 <1 MHz has various advantages:

- the filters F_1 and F_2 can be identical; when f_1 << bandwidth of the filters;
- if F_1 and F_2 are not real low-pass filters, but have a cut-off frequency >f_1 signals with frequencies f_1, resulting from the S.S.B.-modulator, as a result of cross-talk, are suppressed;
- the 1/f-noise from the microwave mixer-diodes is also suppressed;
- integration filters with a bandwidth of about 100 Hz are easily realized.

7.2. Influence of receiver bandwidth on bearing resolution

The amplitude transfer characteristics of a linear array in a direction θ is:

$$H_a(f,\theta) = \frac{\lambda_c}{a}\int_{-\frac{a}{2k}}^{+\frac{a}{2k}} A(x)\, e^{j2kx\,u(f,\theta)}\,dx$$

with:

a = length of the array

$A(x)$ = the illumination function

$$u(f,\theta) = v - \frac{f_{oo}}{f'}v_{oo} - \sin\theta$$

$$v = \frac{\lambda}{\lambda_G}$$

θ = 0, for $f' = f_{oo}$

If the antenna is aimed at a target in a direction θ_o, and the receiver is tuned to a center frequency f_c, (fig. 39.20), we have:

for $f' = f_o$: $U(f_o,\theta_o) = 0$ and $\sin\theta_o = K(f_o - f_{oo})$

for $f' = f_o + \delta f = f_c + f$ $\quad U(f_c+f) = \sin\theta_c - \sin\theta_o + Kf$

with $\quad \sin\theta_c - \sin\theta_o = K(f_c - f_o)$

$K = v_{oo}/f_{oo}$, and is approximately constant for small frequency deviations.

In case of a constant local oscillator frequency f_c and a rotating antenna, θ_c is constant and θ_o is the variable azimuth angle with respect to the normal to the array.

For a constant delay difference τ, the output of the correlator will give an "antenna pattern" with center θ_c.

This antenna pattern is measured as the envelope of the correlator output signal:

$$\left|J(\tau,\Delta\theta)\right| = \left|\int_{-\infty}^{+\infty} \left|S(f+f_c)\right|^2 H_{at}(f)\,H_{ar}(f)\,H_{12}(f)\,H_{21}^{*}(f)\,H_{22}^{*}(f)\,e^{j2\pi f\tau}\,df\right|$$

with: $S(f+f_c)^2$ = noise source power spectrum

$H_{at}(f)$ = amplitude transfer characteristic of transmitting antenna

$H_{ar}(f)$ = the same for the receiving antenna

$\Delta\theta = \theta_c - \theta_o$

For a symmetrical illumination function $A(x)$, $H_a(f)$ is a real function.

If the transmitting- and receiving-antenna are identical:

$$H_{at}(f) = H_{ar}(f) = H_a(f)$$

$H_{12}(f)$ and $H_{22}(f)$ are the receiver filter amplitude transfer characteristics.

$H_{21}(f)$ is the amplitude transfer characteristics for some other filter.

The matching condition for maximum output signal and constant $\Delta\theta$, is:

$$H_a^2(f)H_{12}(f) = H_{21}(f)H_{22}(f)$$

The two receiver filters are made identical to ensure equal phase functions, while $H_{21}(f)$ could be used as a correction filter in case of an asymmetric $A(x)$. With these assumptions we get:

$$H_{12}(f) = H_{22}(f) = H(f)$$

$$H_{21}(f) = 1$$

With a broad uniform source spectrum, the normalized antenna bearing characteristic will be:

$$\frac{\left|J(o,\Delta\theta)\right|}{\left|J(o,o)\right|} = \frac{\left|\int_{-\infty}^{+\infty} |H(f)|^2 \left|H_a(\Delta\theta\, kf)\right|^2 df\right|}{\left|\int_{-\infty}^{+\infty} |H(f)|^2 \left|H_a(kf)\right|^2 df\right|}$$

Insight in the influence of the receiver bandwidth on the bearing resolution can be obtained by using some approximations:

For instance if the filter transfer function is a gaussian function and the standard deviation is defined as: $\sigma_H = \frac{B}{D}$

then: D= 1,179

if: B= the bandwidth between f = o and f = f_{-3dB}

The same is done for the antenna, for which the standard deviation is:

$$\sigma_a = \frac{\theta_{-3\,dB}}{2D}$$

with: θ_{-3dB} = the antenna beamwidth between the - 3 dB points.

The normalized antenna characteristic becomes in this case:

$$\frac{\left|J(o,\Delta\theta)\right|}{\left|J(o,o)\right|} = \exp\left[-\frac{1}{2}\,\frac{\Delta\theta^2}{\frac{k^2}{D^2}\left\{B^2 + \left(\frac{\theta_{-3dB}}{2k\sqrt{2}}\right)^2\right\}}\right]$$

The "beamwidth" of this pattern will be:

$$\theta'_{-3dB} = 2k\sqrt{B^2 + \left(\frac{\theta_{-3dB}}{2k\sqrt{2}}\right)^2}$$

A measure of the loss in the angular resolution could be defined as the ratio :

$$\eta = \frac{\theta_{-3dB}}{\theta'_{-3dB}} = \frac{1}{\sqrt{1 + \left(\frac{2B\,k\sqrt{2}}{\theta_{-3dB}}\right)^2}}$$

In the experiment we used two video bandwidths which where equivalent resp. to a "broad" r.f. bandwidth of about 30 MHz and a "narrow" r.f. bandwidth of about 14 MHz.

The antenna beamwidth was 0,4 °

For the two cases we obtain resp:

with 2B = 30 MHz: η = 0,73

with 2B = 14 MHz: η = 0,91

From the measurements (fig. 39.16) follows resp. : η = 0,71 and η = 1. This agrees quite well.

It can be shown that the receiver bandwidth B, which would give joint optimum resolution in distance as well as in angle is given by the relation:

$$B = \frac{\theta_{-3dB}}{2k\sqrt{2}}$$

8. REFERENCES

1. Siebert, W.M. "A radar detection philosophy" I.R.E. Trans. I.T.- 2, Sept. 1966, p. 204.

2. Grant, M.P., e.a., "A class of noise Radar systems".Proc. I.E.E.E., July 1963. p. 1060.

3. Carpentier, M. "Radars, Concepts Nouveaux" Dunod, Paris 1966.

4. Woodward, P.M. "Radar ambiguity analysis". R.R.E. Technical note no 731, Feb. 1967.

5. Berkowitz, R.S. "Modern Radar" J. Wiley. 1965.

6. McGillem, C.D., e.a. "An experimental random signal radar". Proc. Nat. Electr. Conf. vol. 23 Oct. 1967.

7. Poirier, J.L. "Quasi-monochromatic scattering and some possible radar applications". Radio Science, vol. 3, Sept. 1968, p. 881.

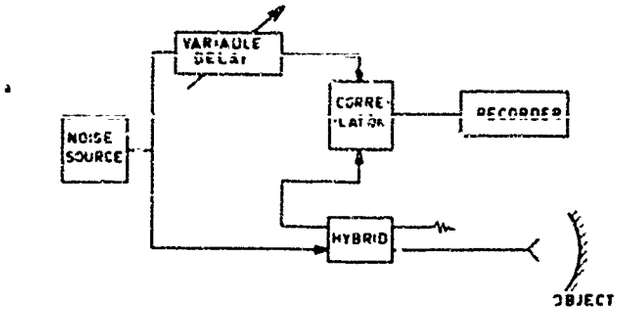
8. S.E. Craig, W. Fishbein, O.E. Rittenbach, Continuous-Wave radar with high Range Resolution and Unambiguous Velocity Determination I.R.E. Trans. MIL - 6, April 1962, p. 153

9. Smit, J.A. "An analogue microwave envelope correlator", Christ. Huygenslab. Rep. Dec. 1964

10 Gabor, D., "Theory of communication" J. Inst. Elec-Engrs. (London) Pt III, vol. 93 1946. p. 429.

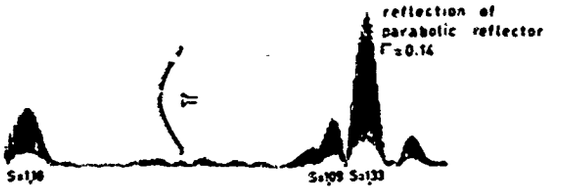
11 Rihaczek, A.W. "Principles of high-resolution radar", McGraw-Hill 1969.

12 Smit, J.A. "Signalform and resolution" Christ. Huygenslab. Rep., Dec. 1965.

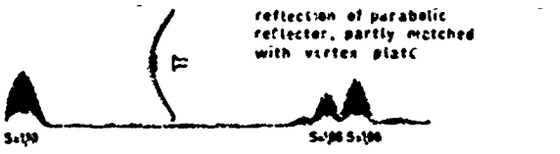


a BLOCK DIAGRAM OF REFLECTOMETER

b. CORRELATOR OUTPUT SIGNAL - FOR CROSS-COUPLING OF HYBRID (REFERENCE „REFLECTION"), AND REFLECTION OF HORN. (S = VSWR)



c CORRELATOR OUTPUT SIGNAL, AS IN b, AND HORN PLACED IN FOCUS POINT OF A PARABOLIC REFLECTOR.



d. CORRELATOR OUTPUT SIGNAL, AS IN c, PARABOLIC REFLECTOR PROVIDED WITH A „VERTEX PLATE"

Fig. 1 Example of measurement with a microwave noise-reflectometer system. Matching of a Parabolic-reflector at 9000 MHz. Spectrum bandwidth: 1000 MHz

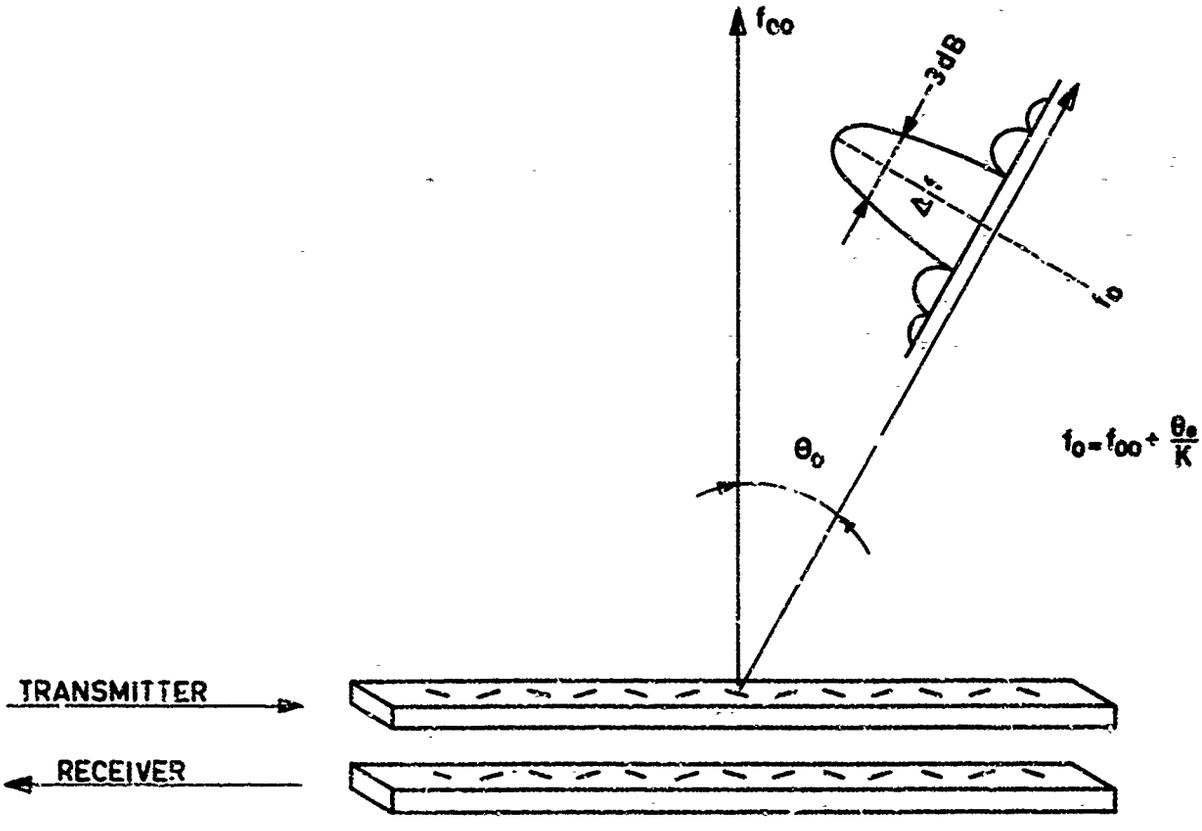


Fig. 2 Linear array as a space filter

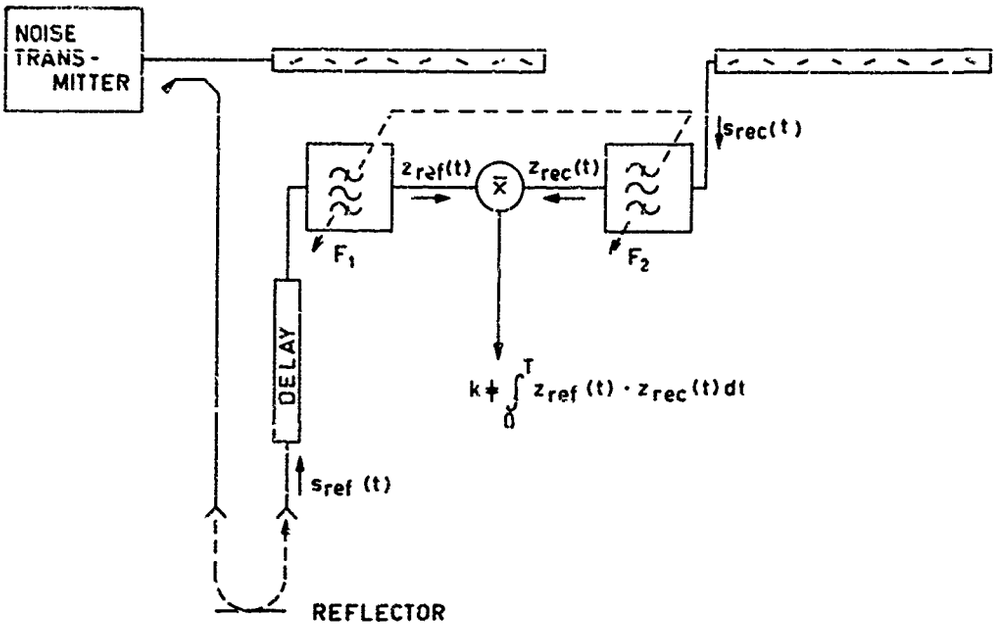
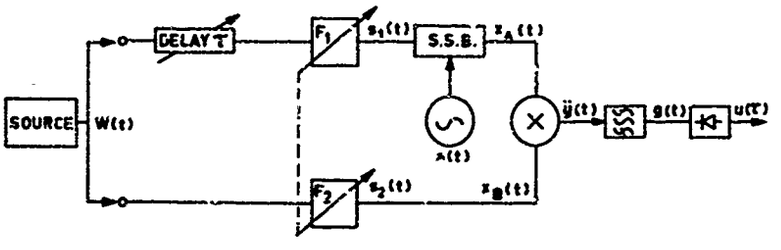
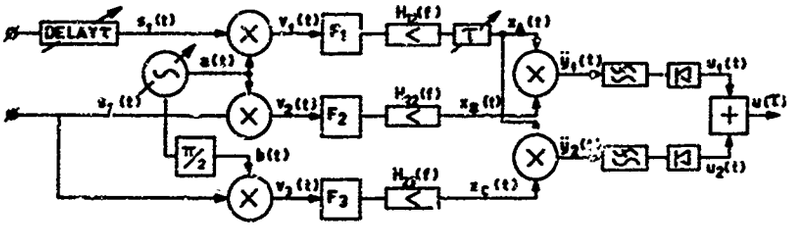


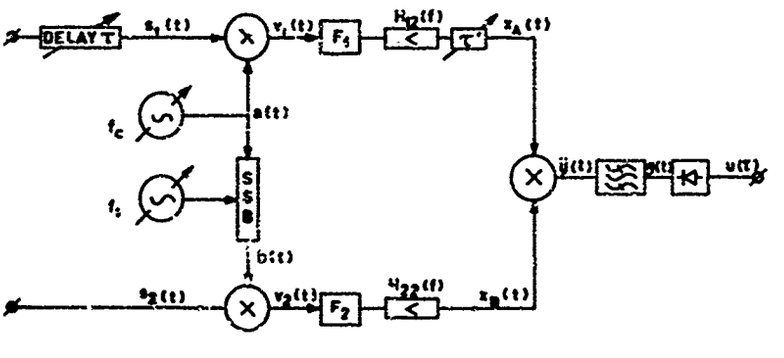
Fig. 3 Block diagram of a noise radar



—a—



—b—



—c—

Fig. 4 Microwave envelope-correlator configurations

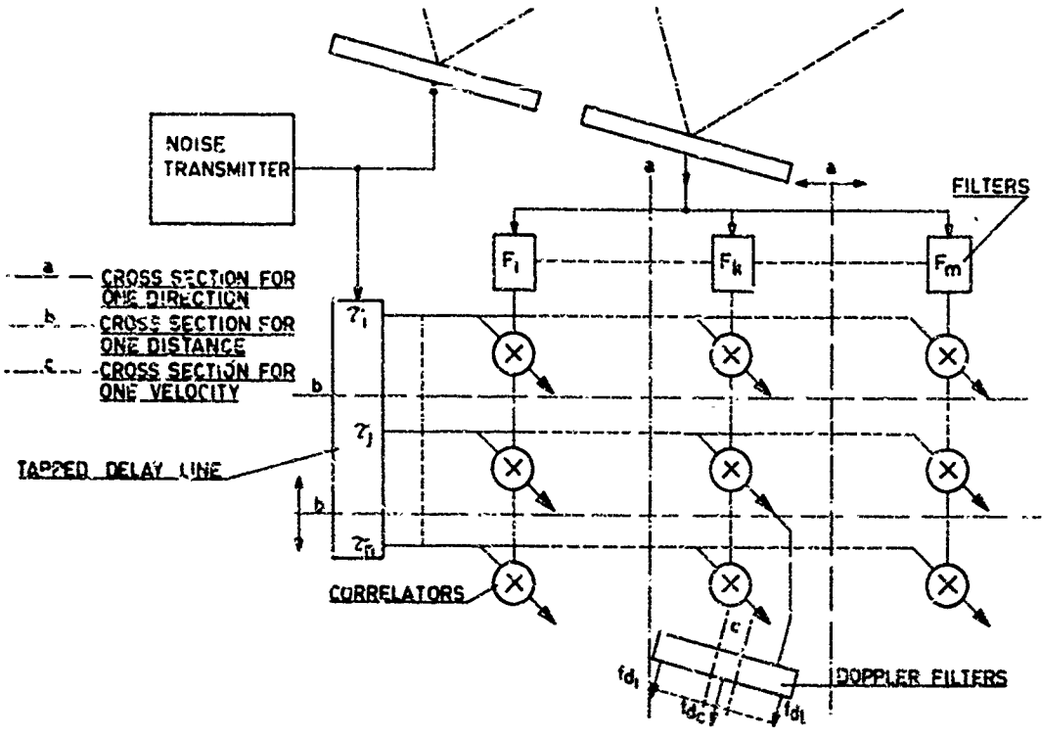


Fig. 5 Block diagram of a complete noise radar system

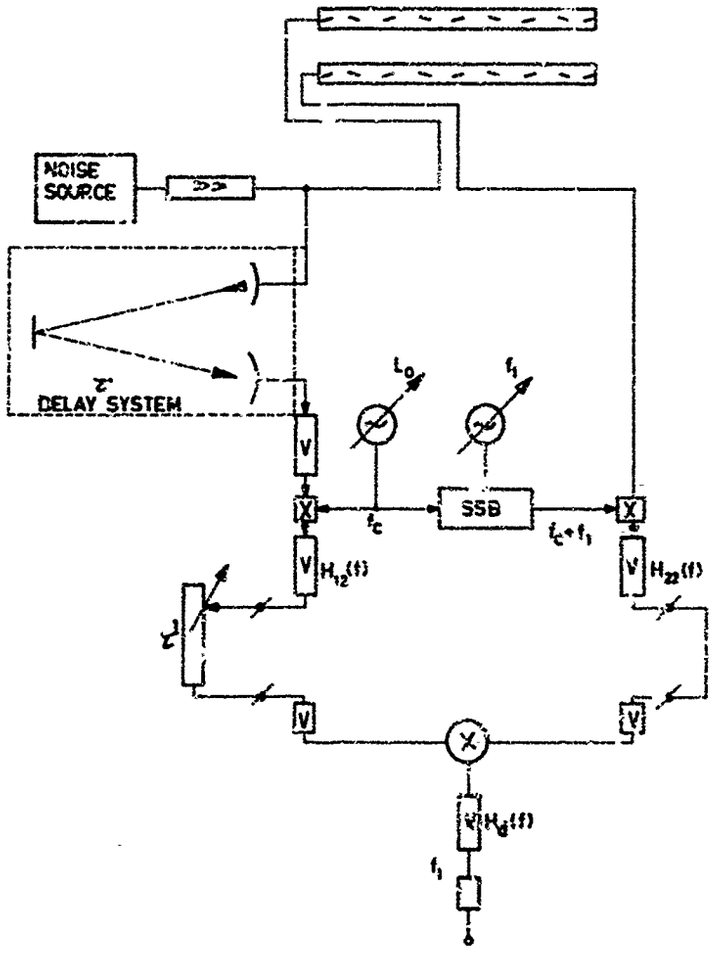


Fig. 6(a) Block diagram of realized RUDAR-system

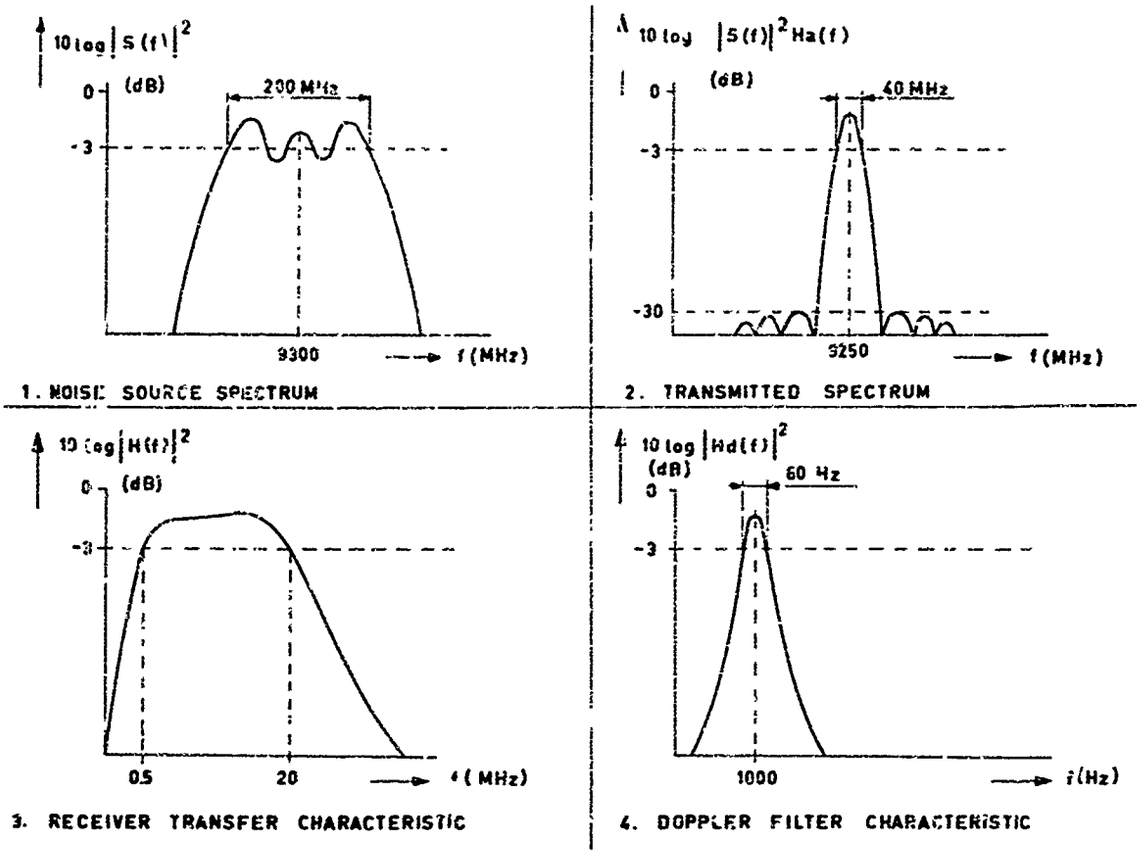


Fig.6(b) Spectra at some points in the system

Fig.7 The RUDAR antenna system

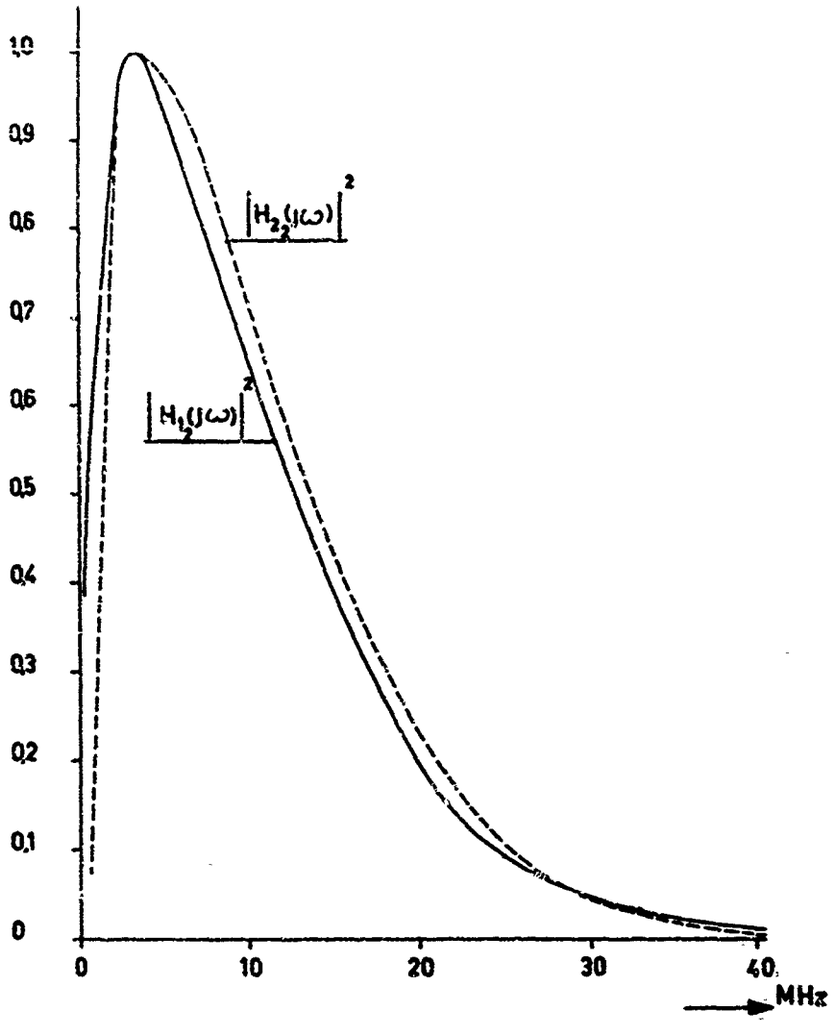


Fig.8 Transfer function of corralator receivers

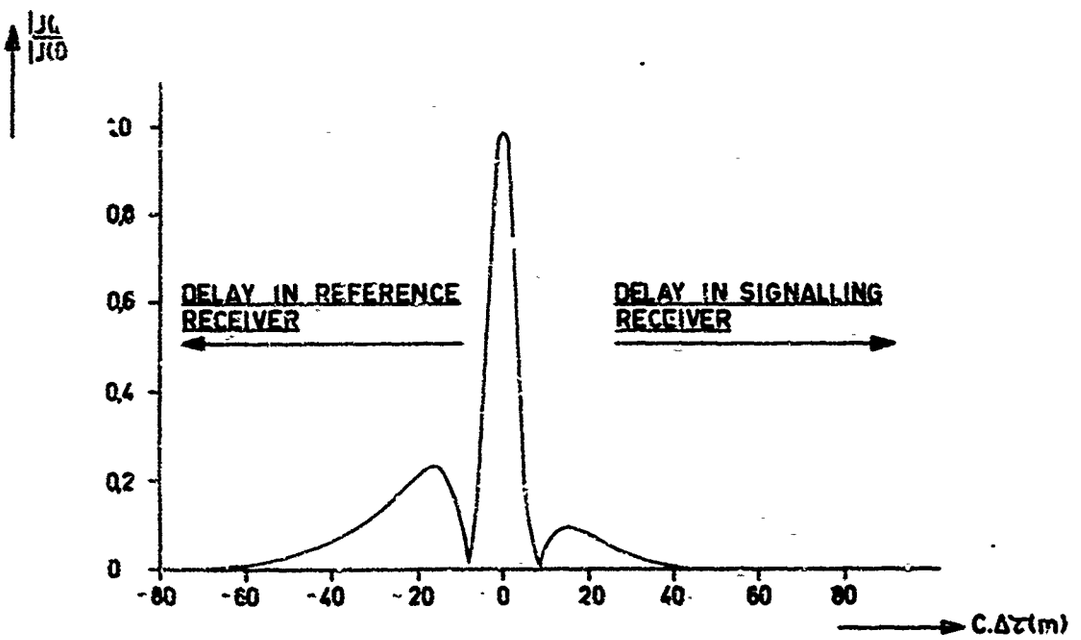


Fig.9 Measured normalized envelope of the cross-correlation function of RUDAR-receiver

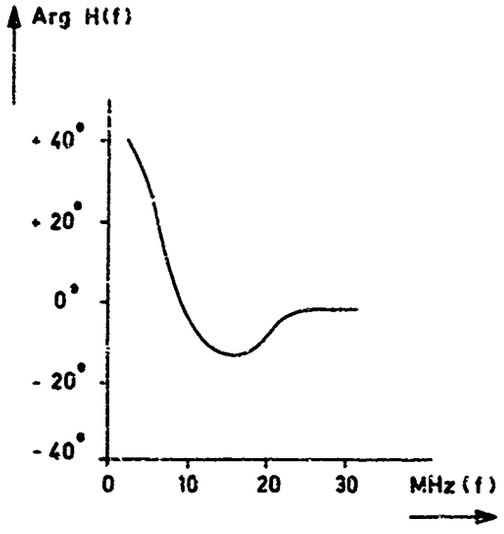


Fig. 10 Calculated phase-difference function of the correlator receiver

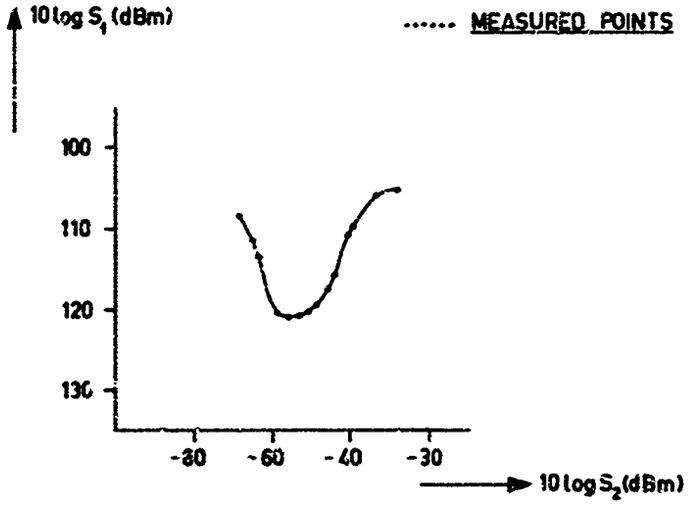


Fig. 11 Input signal level S_1 as a function of the reference signal level S_2 for $(S/N)_{out} = 10$

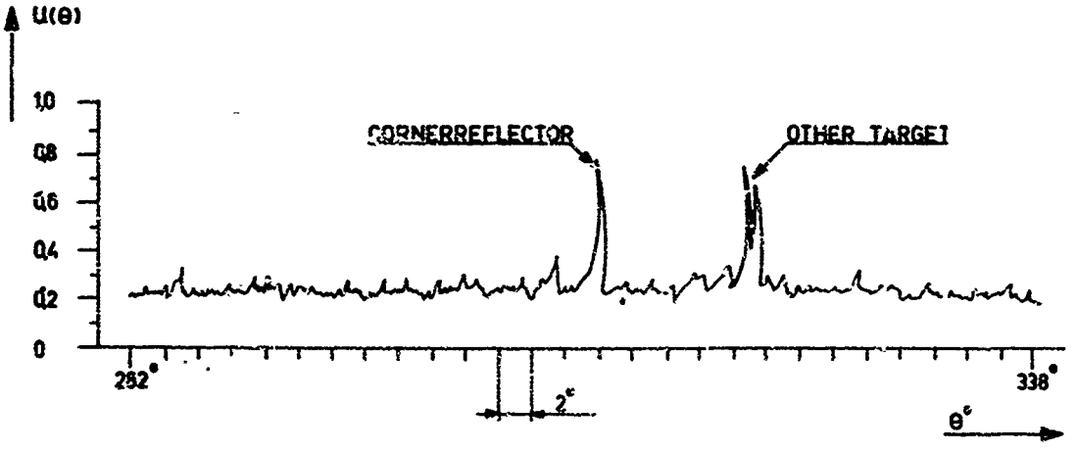


Fig. 12 Correlator output signal, for mechanical azimuth scanning over 50°

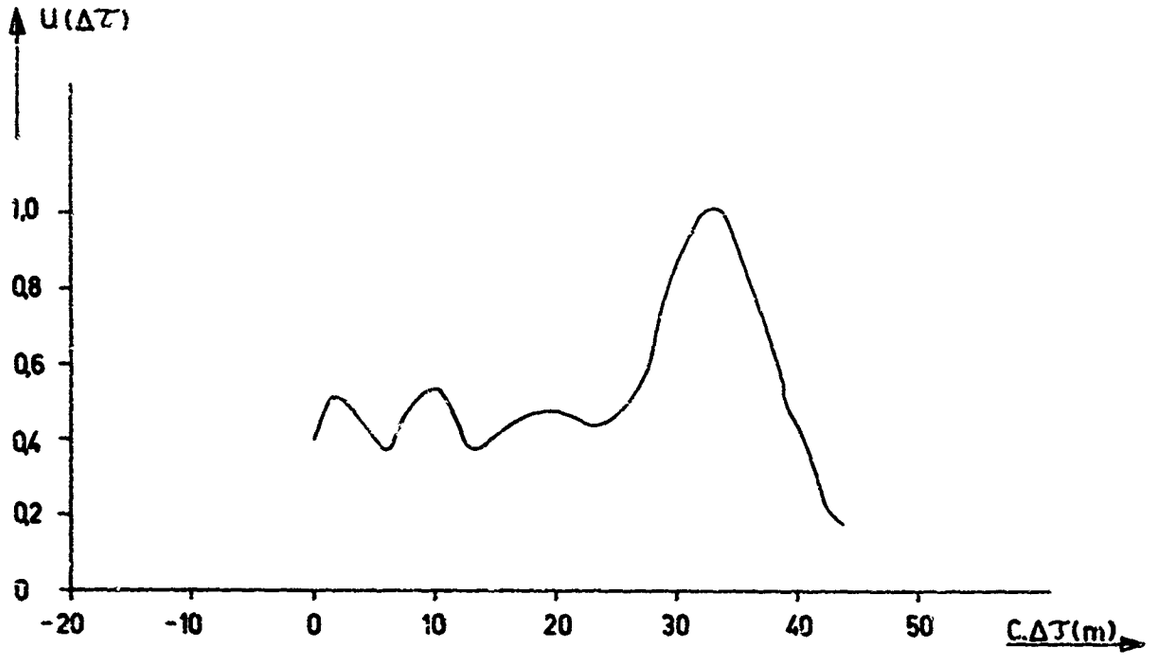


Fig. 13 Correlator output signal, for distance scanning

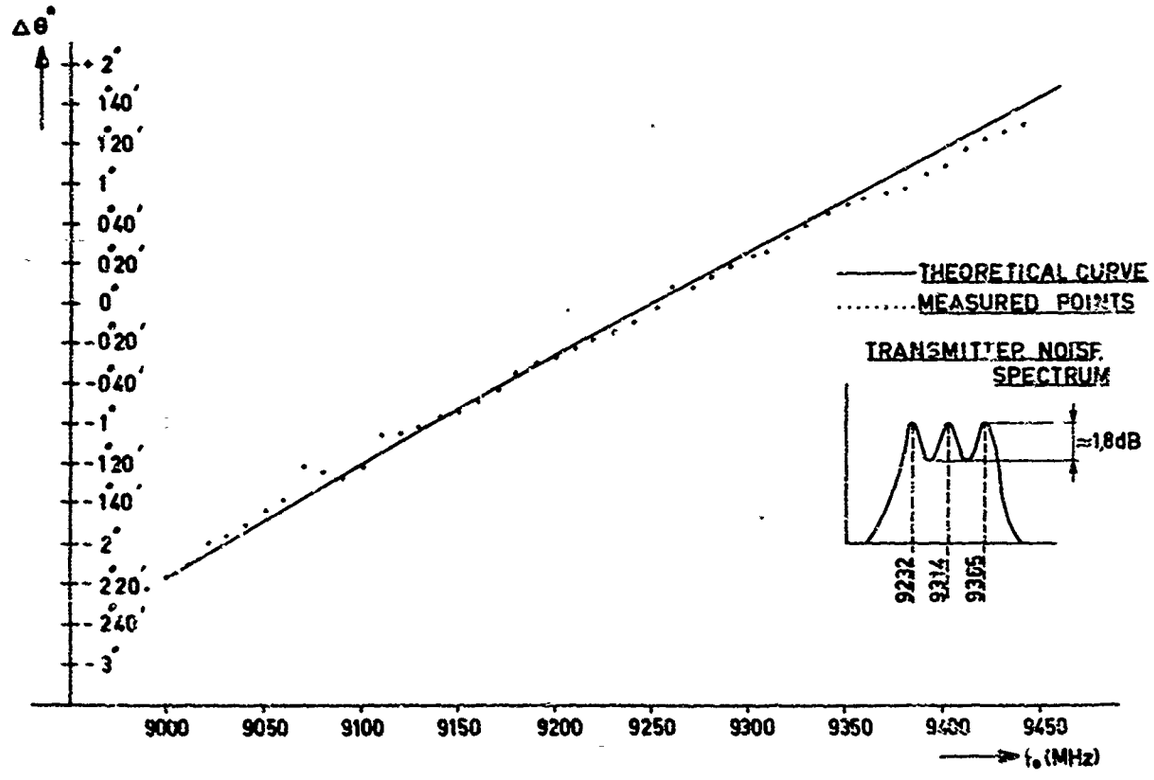


Fig. 14 Antenna-squint measurement with RUDAR (narrow video bandwidth)

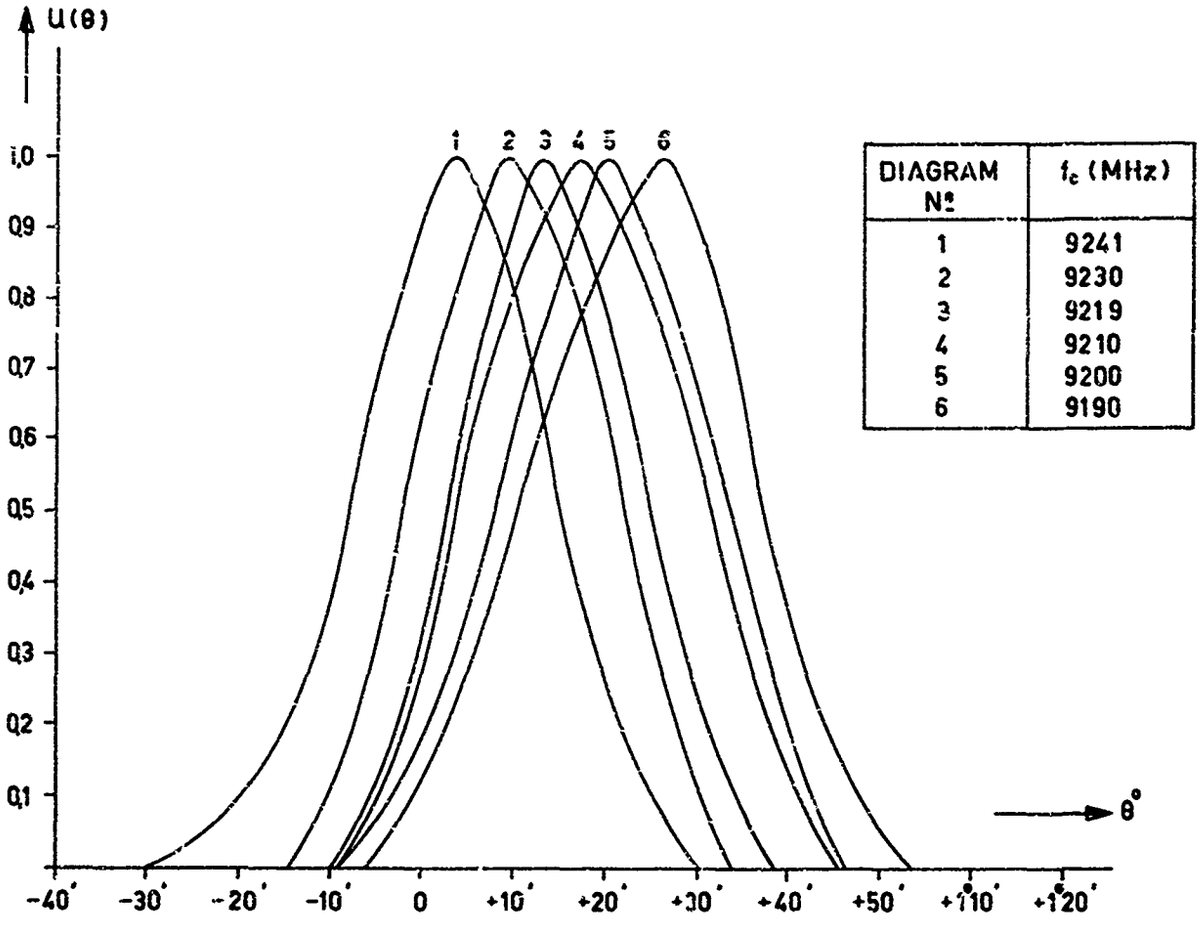


Fig. 15 Antenne system as a space-filter (narrow video bandwidth)

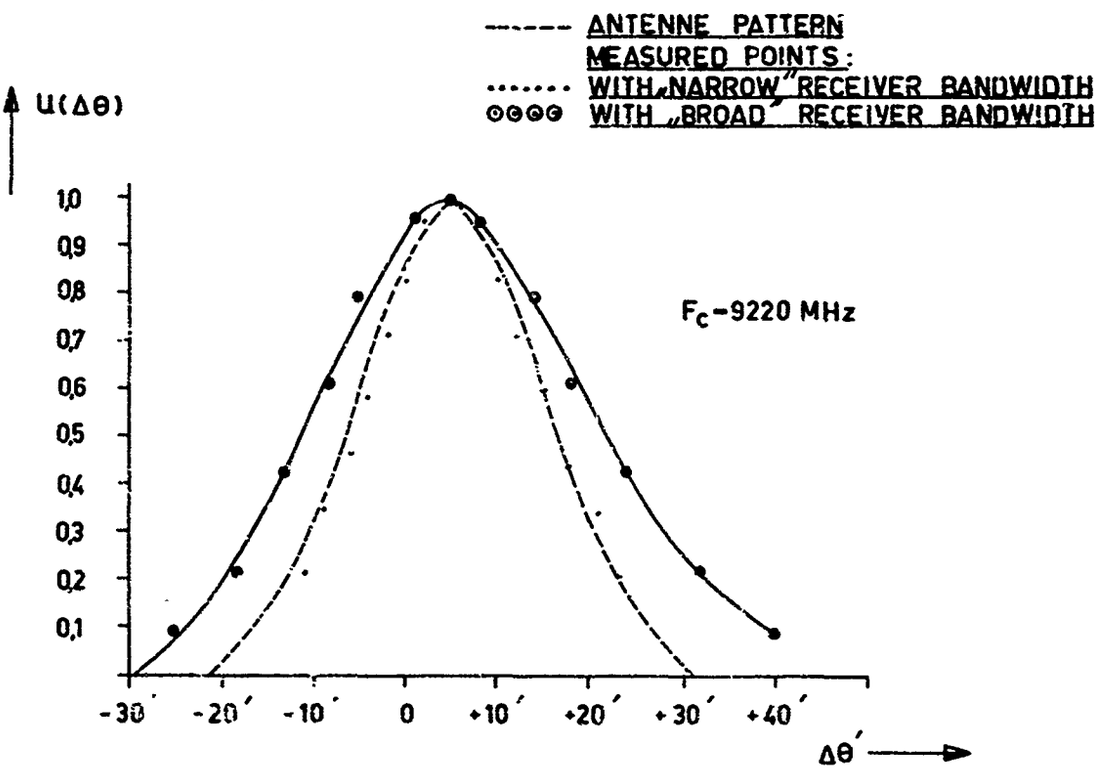


Fig. 16 Influence of receiver bandwidth on measured antenna main beam

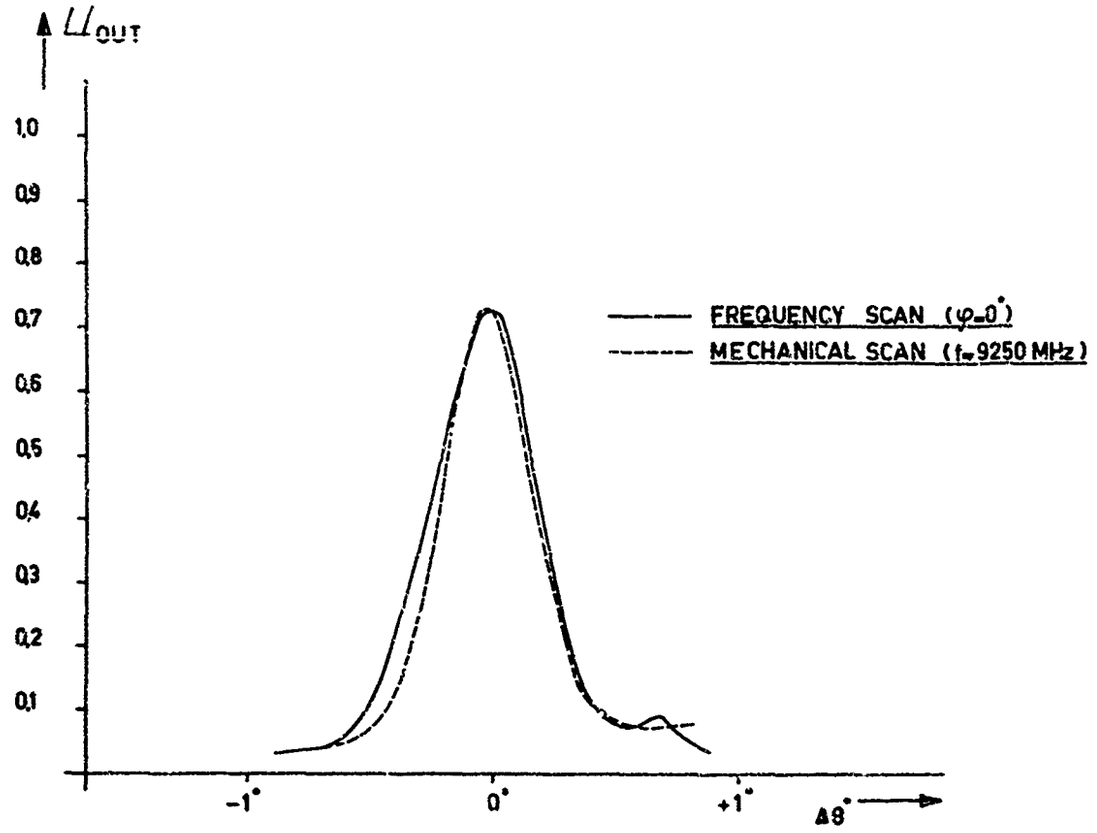


Fig. 17 Comparison of mechanical- and frequency-azimuth scanning

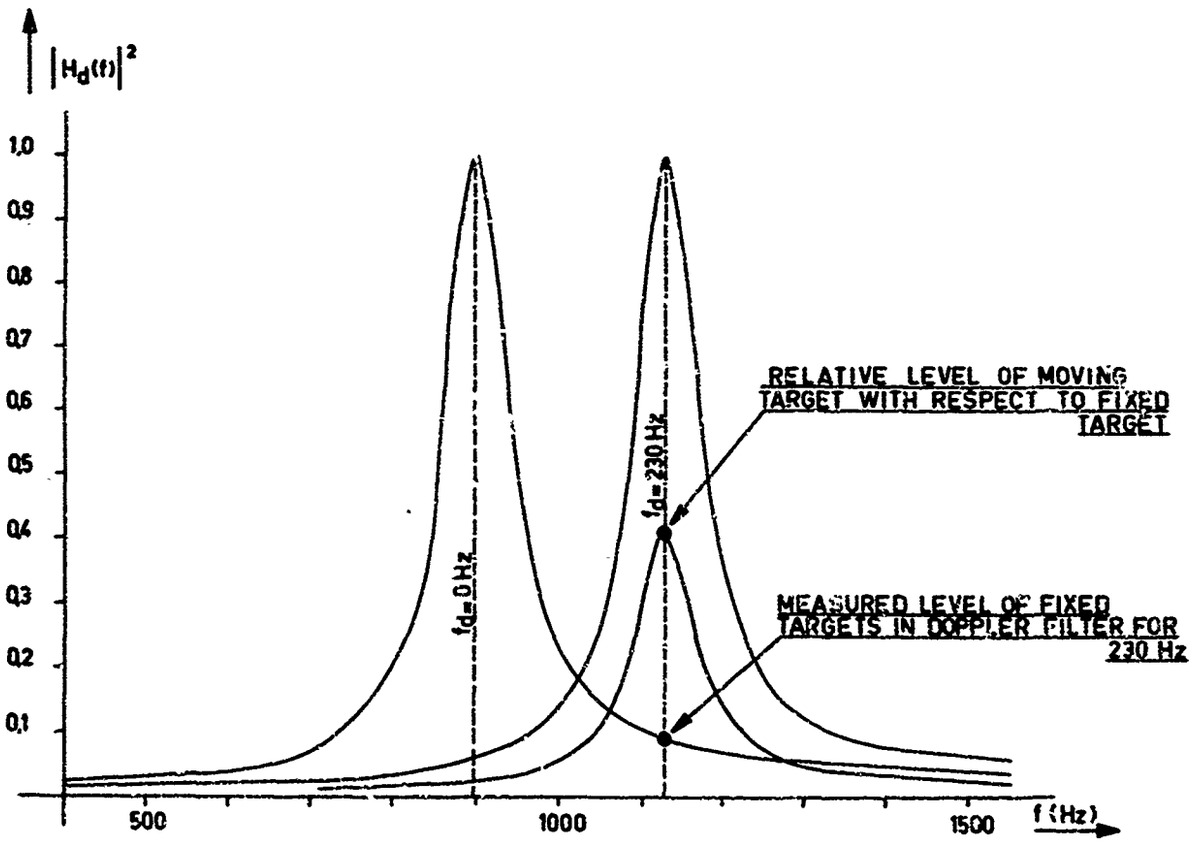


Fig. 18 Doppler filter characteristics

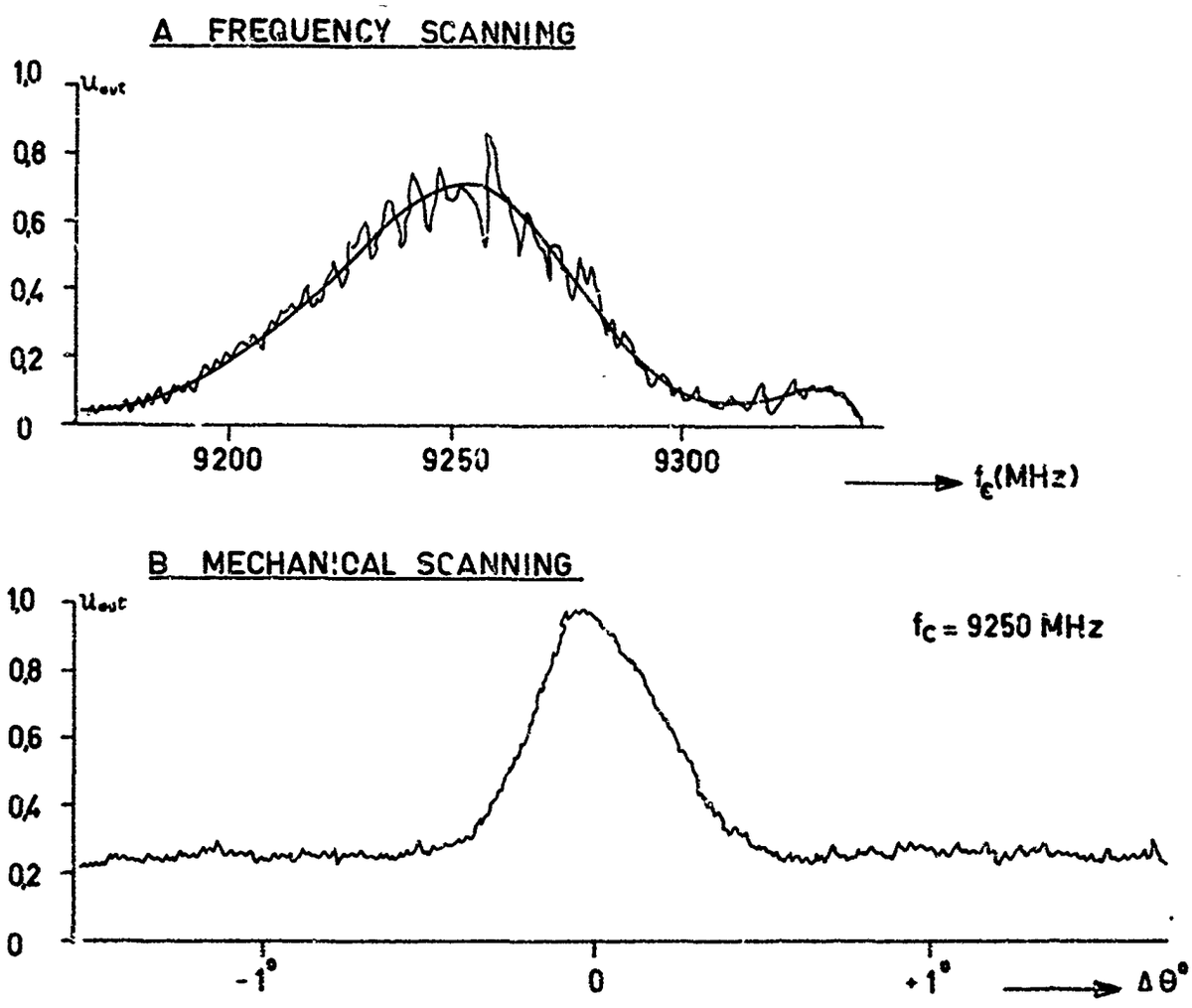


Fig. 19 Influence of spurious signals in delay line on frequency scanning

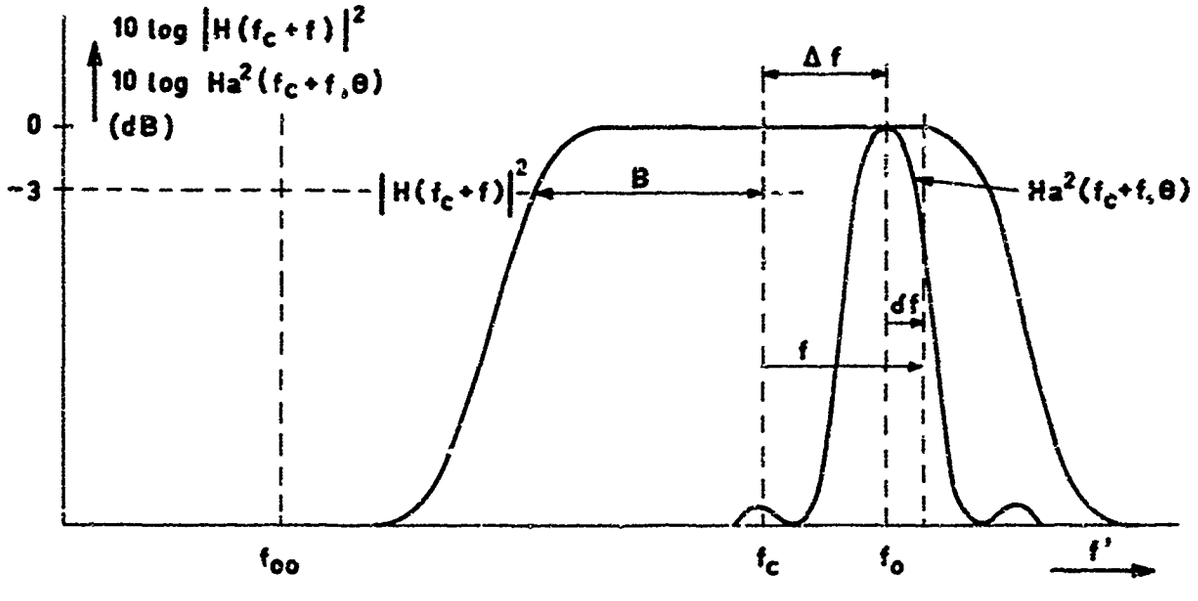


Fig. 20 Frequency relationship between antenna and receiver filter

RADAR D'EVITEMENT D'OBSTACLES POUR HELICOPTERE

par

Gérard Collot

Electronique Marcel Dassault
92 - Saint-Cloud - France

RADAR D'EVITEMENT D'OBSTACLES POUR HELICOPTERE

Gérard COLLOT
ELECTRONIQUE MARCEL DASSAULT
92 - Saint-Cloud - FRANCE

SOMMAIRE

Le radar SAIGA a été développé par l'ELECTRONIQUE MARCEL DASSAULT sous contrat du Service Technique des Télécommunications de l'Air en 1966 et 1967. Les essais en vol ont commencé en 1968 et se sont poursuivis jusqu'à la mi 1969. A la fin de ces essais en vol, il y a eu plusieurs évaluations tant françaises qu'étrangères pour vérifier la validité du matériel dans différentes conditions de vol.

CARACTERISTIQUES PRINCIPALES

fréquence d'émission	:	35 000 MHz (bande Ka)
puissance crête d'émission	:	de l'ordre de 7 kW
largeur d'impulsion	:	0,2 µs
antenne : diamètre approximatif	:	450 mm
largeur du faisceau	:	1,3°
facteur de bruit global	:	inférieur à 13 dB
stabilisation de l'antenne	:	en roulis et tangage
volume exploré		
en gisement par rapport à l'axe hélicoptère	:	supérieur à $\pm$ 60°
en site par rapport au vecteur vitesse	:	supérieur à $\pm$ 10°
Figuration des échos dangereux	:	par rapport à une ligne de référence
pour une vitesse comprise entre	:	0 et 500 km/h
pour une altitude de garde minimale de	:	30 m
Présentation des informations	:	type C ou "télévision" (limitée à 5 km)
	:	type B : échelle 2 ou 5 km
Tests intégrés	:	permanents
Poids		
radar avec radome	:	50 kg environ
indicateur	:	10 kg environ
Consommation électrique		
réseau 200 V - 400 Hz	:	inférieure à 700 VA
réseau 28 V continu	:	inférieure à 2 A
réseau 26 V - 400 Hz	:	inférieure à 30 VA

PRINCIPES DE FONCTIONNEMENT

Le radar SAIGA est un matériel héliporté, il permet le vol tous temps à très basse altitude. Les principes mis en oeuvre dans le radar SAIGA sont adaptés aux conditions particulières du vol d'un hélicoptère, appareil qui peut s'enfoncer profondément dans le relief.

L'équipement se compose essentiellement de deux unités :

- le radar proprement dit, monté dans le nez de l'hélicoptère et dont le rôle est d'explorer l'espace à l'intérieur duquel le vol va se poursuivre (figure 40-1),

- l'indicateur, installé dans le poste pilote, dont la face avant groupe toutes les commandes d'exploitation (figure 40-2).

L'ensemble constitue un matériel rustique, compact et léger, qui allie la facilité d'installation à une maintenance aisée. Il répond aux normes concernant les matériels aéroportés et a été conçu en tenant compte des conditions d'environnement rencontrées à bord d'un hélicoptère.

La réalisation à haute densité, fait appel à des éléments d'une technologie avancée : transistorisation très poussée, circuits intégrés, circuits imprimés.

C'est pourquoi, le poids du radar, y compris le radome et l'indicateur, est relativement faible : environ 60 kg.

Utilisation opérationnelle

Le radar SAIGA permet les fonctions opérationnelles d'évitement d'obstacles et de visualisation du sol (sur option).

L'évitement d'obstacles peut s'effectuer dans deux plans, et comme le pilotage est manuel, il est possible de choisir la manoeuvre de dégagement la plus appropriée, à savoir :

- franchissement,
- contournement,
- freinage.

A cet effet, la situation est clairement présentée au pilote sur un indicateur cathodique qui affiche tous les obstacles dangereux ainsi que des repères facilitant l'exploitation.

Détection des câbles

Le radar SAIGA, grâce à l'utilisation de la bande Ka qui permet d'obtenir un faisceau d'antenne très étroit, détecte les obstacles isolés tels que les pylônes ainsi que les câbles des lignes à haute tension.

Cette possibilité permet le vol à très basse altitude par tous temps.

Visualisation du sol

Cette fonction peut être ajoutée, sur option, au radar SAIGA.

Dans ces conditions, la représentation utilisée est du type B. Cette fonction permet :

- la navigation et la recherche d'un site dégagé pour des manoeuvres telles que :
  - la navigation ou le largage du matériel ou de personnel,
  - la récupération du personnel en mer.

L'exploration de l'espace s'effectue à l'aide de deux antennes couplées à l'intérieur d'une large portion de sphère située à l'avant de l'hélicoptère (figure 40-3).

- en gisement

balayage sur	:	+ 80°
résolution	:	1,3°
temps d'examen (une ligne)	:	40 ms environ

- en site

balayage sur	:	20° à 40° avec les 2 antennes couplées
résolution	:	1,3°
temps d'examen	:	1 seconde environ

- distance

zone utile	:	70 m à 5000 m
résolution	:	30 m

L'enfoncement maximal de l'hélicoptère dans le relief est lié à la finesse de l'analyse qui s'obtient par la mise en oeuvre d'un pouvoir séparateur élevé en site, gisement et distance.

Compte tenu des faibles dimensions de l'ensemble et de celles des antennes en particulier, il a fallu rechercher un faisceau relativement fin, par l'adoption d'une longueur d'onde aussi courte que possible. Le choix de la bande Ka permet d'obtenir une ouverture de faisceau de 1,3° dans les deux plans il reste compatible avec les portées nécessaires quelles que soient les conditions météorologiques.

Le pouvoir séparateur en distance est de 30 m, valeur homogène avec les définitions angulaires données par le faisceau à l'intérieur du domaine de recherche.

Cette valeur, qui correspond à une impulsion relativement brève (0,2 µs) permet en outre de réduire les perturbations dues aux réflexions sur un rideau de pluie.

La navigation en aveugle, à très basse altitude n'est possible que si le pilote dispose d'un moyen capable d'évaluer le relief qui va être survolé, afin de déterminer les obstacles dangereux à la poursuite du vol.

Ce moyen est réalisé pratiquement grâce à la "surface de garde" lieu des obstacles nécessitant une manoeuvre d'évitement de l'hélicoptère (figure 40-4).

Les obstacles situés au-dessous de la courbe de garde ne sont pas dangereux et ceux situés au-dessus sont dangereux. Cette surface de garde est située sous l'hélicoptère à une hauteur appelée "hauteur de garde", dont la valeur peut être choisie entre 30 et 300 m au gré du pilote.

L'expérience de l'ELECTRONIQUE MARCEL DASSAULT dans le domaine des radars et des calculateurs appliqués à la détection d'obstacles a permis de définir une solution simplifiée de méthode de calcul pour l'application au vol basse altitude d'un hélicoptère.

Le principe général consiste à comparer en permanence la position et la variation de position des échos reçus par rapport à un gabarit défini à l'avance.

Le gabarit choisi pour SAIGA consiste en deux segments de droite dont la position du point A d'articulation dépend de la vitesse du porteur (figure 40-4). La pente du deuxième segment de droite a été définie une fois pour toutes après expérimentation.

La position du gabarit (courbe de garde) par rapport au porteur dépend de la hauteur affichée et de la vitesse de descente, les équations des points caractéristiques de la courbe sont les suivantes :

- position verticale $h = h_o + T_1 V_z$ seulement pour $V_z < 0$
- point d'articulation $D = D_o + T_2 V_x$
- pente du deuxième segment de droite = 2 à 6°

avec :

h_o = hauteur affichée par le pilote
V_z = vitesse de descente
T_1 = constante de temps qui dépend des caractéristiques aérodynamiques du porteur
V_x = vitesse sol horizontale
T_2 = constante qui dépend des caractéristiques aérodynamiques du porteur

Deux types de présentation sont réalisés :

- une présentation gisement-site type "télévision" ou "type C" particulièrement bien adaptée à la situation puisqu'elle restitue sensiblement l'image du paysage situé devant l'hélicoptère;
- une présentation gisement-distance de "type B" donnant une image très fine du relief découpé par la surface de garde. La distance maximale est au choix de 2 ou 5 km.

Dans les deux cas, et afin d'être facilement interprétable, l'image est stabilisée en gisement et les échos dangereux sont présentés en surbrillance sur fond noir.

A ces images sont incorporés :

- une ligne de référence permettant le pilotage dans le plan vertical, et qui figure la surface de garde. Les obstacles dangereux sont situés au-dessus de cette ligne de référence et les obstacles non dangereux sont situés au-dessous,
- un marqueur correspondant à la direction du vecteur vitesse de l'hélicoptère, le pilotage consistant essentiellement à maintenir ce repère dans une zone dégagée. Ce vecteur vitesse est symbolisé par un cercle en type C et par une ligne pointillée en type B,
- un marqueur figurant la route à suivre ou le cap obtenu à partir de l'affichage des données du système de navigation de bord.

Sur la présentation "type B", en plus des marqueurs vecteur vitesse et route à suivre, il y a :

- des marqueurs de distance, espacés de 500 m sur l'échelle 2 km et de 1000 m sur l'échelle 5 km,
- le marqueur "distance d'arrêt" symbolisé par une ligne pointillée. Cette distance d'arrêt varie en fonction de la valeur de la vitesse horizontale de l'hélicoptère.

La présentation des obstacles et des marqueurs associés permet au pilote :

- d'avoir une vue d'ensemble du relief au voisinage de la route suivie,
- de pouvoir se dérouter, afin d'éviter le survol d'un obstacle important se trouvant devant lui,
- de pouvoir freiner à temps si le déroutement n'est pas possible.

Pilotage en saut d'obstacle

Le problème consiste à donner à l'hélicoptère une pente de montée correcte. La figure 40-5 montre le processus.

La surface de découpe liée à l'hélicoptère doit être normalement en contact avec le sol. Lorsqu'il se présente un obstacle, le pilote doit donner à son hélicoptère une pente de montée α de façon à passer à la verticale de l'obstacle à une hauteur h_0. L'angle de pente α est précisément l'écart de position par rapport à la ligne de référence horizontale indiqué par le marqueur cercle.

La seule opération à effectuer par le pilote consiste simplement à maintenir le centre du cercle du vecteur vitesse sur le sommet de l'obstacle.

Les trois figures 40-5a, b, c, indiquent :

a) le vol est trop haut parce que la ligne de référence symbolisant la surface de découpe est au-dessus du sol ; le pilote doit maintenir son vecteur vitesse (cercle) en contact avec le sol,

b) le vol est trop bas, le sol est au-dessus de la ligne de référence ; le pilote doit faire cabrer son hélicoptère,

c) le vol est correct : - sol, référence et cercle sont confondus.

Le radar SAIGA est équipé d'un système de tests intégrés qui vérifie, en permanence le bon fonctionnement de l'équipement.

Cette vérification s'effectue grâce au test simultané de l'ensemble Emission-Réception et des circuits de calcul.

L'ensemble Emission-Réception est reconnu en bon état lorsque la mesure du niveau vidéo en fin de chaîne excède un seuil prédéterminé. Dans le cas contraire, le pilote est averti de la défaillance du matériel par l'allumage du voyant "Alarme".

La validité des circuits de calcul est signalée par la présence d'une zone lumineuse à la partie gauche de l'indicateur. Ce test est réalisé grâce à un générateur d'échos fictifs incorporé à la chaîne de réception et constitué par un oscillateur déclenché périodiquement.

On notera de plus, que l'allumage du voyant "Alarme" peut également être provoqué si la hauteur de vol donnée par l'altimètre est inférieure à la hauteur de vol affichée.

Stabilité des images

Afin d'éviter un glissement de l'image présentée sur l'écran de l'indicateur, celle-ci est gelée par rapport à des références fixes situées dans les plans site et gisement.

Dans le plan gisement, l'image est gelée par rapport au marqueur de route (ou de cap). Pour des variations de l'écart de route de $\pm 60°$, l'image est recentrée automatiquement sur l'écran et le marqueur de route est décalé de la même quantité.

Pour des écarts de route dépassant l'ouverture en gisement du radar, l'image reste gelée et le marqueur de route est visible en butée à droite ou à gauche de l'écran.

Dans le plan site, l'image est gelée par rapport à l'horizontale. Le recentrage de l'image est effectué par quantités fixes de 10° pendant les variations de l'angle de pente. Lorsque l'angle de pente est supérieur à $\pm 20°$, une flèche verticale s'éclaire.

DESCRIPTION (Planche 40-6)

Le radar groupe les éléments suivants :

- l'antenne et les mécanismes associés,
- l'émetteur-récepteur,
- les circuits de traitement de l'information.

L'indicateur comporte :

- le tube cathodique et les circuits associés,
- les commandes d'exploitation.

Antenne et mécanismes

L'ensemble mobile est constitué par une sphère comportant un anneau support d'antenne et deux demi-radomes permettant d'éviter l'effet de ventilation.

L'antenne est constituée par deux paraboles placées dos à dos (figure 40-7) animées dans le plan horizontal d'un mouvement de rotation continue, à la vitesse de 660 tr/mn. Au cours de la rotation, un commutateur à ferrite aiguille successivement l'énergie hyperfréquence sur chaque parabole, afin de réaliser le balayage en gisement du domaine d'exploration.

L'ensemble est stabilisé en roulis et en site, à partir des informations de la centrale de verticale.

Le mouvement en site effectué en une seconde environ, s'obtient grâce à une oscillation des deux antennes de $\pm$ 10° autour de l'axe de stabilisation.

La liaison hyperfréquence, à travers les axes d'articulation est assurée par trois joints tournants.

Emetteur-récepteur

L'énergie hyperfréquence est délivrée par un magnétron dont la puissance crête est de l'ordre de 7 kW dans la bande Ka. La modulation est réalisée grâce à un ensemble thyratron ligne à retard à charge résonnante.

L'énergie reçue est amplifiée par un récepteur du type supernétérodyne à contrôle automatique de fréquence, dont la loi de réception permet d'éviter toute saturation.

A cet effet, l'amplificateur moyenne fréquence présente une caractéristique linéaire-logarithmique, particulièrement adaptée aux variations importantes du niveau. Le gain est variable dans le temps afin de s'affranchir des perturbations que pourraient apporter les échos proches reçus par les lobes secondaires de l'antenne.

Le signal vidéo recueilli en sortie est dirigé vers les circuits de traitement de l'information.

Les circuits de traitement de l'information ont principalement pour rôle l'élaboration de la surface de garde, en vue de la discrimination des obstacles dangereux.

Le calcul de la surface de garde s'obtient à partir des informations suivantes :

- composantes de la vitesse/sol (données par le radar doppler ou tout autre appareil),
- hauteur de garde (affichée par le pilote),
- position de l'antenne.

En plus de cette fonction, les circuits de traitement de l'information ont également pour but :

- le calcul de la distance d'arrêt à partir de l'information vitesse,
- la génération des différents marqueurs, y compris l'horizon artificiel, à partir des données de la centrale ou de tout autre appareil : horizon recopié, par exemple.

Indicateur

L'indicateur réalisé sous forme d'un coffret parallélépipédique de faibles dimensions, englobe le tube cathodique et les circuits qui lui sont propres : amplificateurs, alimentations.

Le tube cathodique est du type à entretien d'image, à deux canons d'inscription ce qui permet de réaliser simplement l'affichage des différentes informations.

La première voie est réservée à la présentation de l'horizon artificiel et la seconde à la visualisation des signaux radar et des marqueurs.

Sur la face antérieure, sont disposés les diverses commandes d'exploitation, ainsi que les voyants d'état du radar.

Les principales commandes d'exploitation sont les suivantes (figure 40-2) :

- mise en oeuvre,
- choix de la hauteur de garde,
- présentation de l'image,
- gammes de recherche,
- tests "voyants",
- choix de la polarisation de l'onde émise,
- commande de virage,
- déblocage de l'antenne.

Maintenance

La maintenance du radar SAIGA s'effectue soit sur l'hélicoptère soit sur le banc de maintenance.

RESULTATS DE L'EXPERIMENTATION EN VOL

Portée

Le diagramme (figure 40-8) représente en abscisse la distance sur échelle logarithmique et en ordonnées la puissance reçue en dB/W. Le seuil de visualisation est situé à - 105 dB/W (+ 16 dE/KTBF). On peut voir que les arbres sont détectés à une distance de l'ordre de 2 à 3 km (mesures effectuées en été).

Les diagrammes (figures 40-9 et 40-10) donnent les portées obtenues sur pylônes de lignes HT et câbles HT. Dans les deux cas, on peut voir que ces obstacles sont aisément détectés à partir de 1500 mètres.

Détection des lignes HT

Le radar SAIGA permet de détecter les câbles HT en trois points :

- un point selon la normale au câble,
- un point de chaque côté, à un angle d'environ 20° qui dépend du type de câble.

Sur la figure 40-11 dans le cas de la position (a) la direction de l'hélicoptère est perpendiculaire à la ligne haute tension, nous avons un point de détection exactement dans l'axe et un point de chaque côté avec un angle de 20°. Dans le cas de la position (b) il y a toujours une détection selon la normale aux câbles, et également un point de chaque côté de ce point central. C'est particulièrement intéressant car cela permet à l'hélicoptère de sauter par-dessus les lignes haute tension. La figure 40-11c donne l'explication du phénomène.

Lorsqu'une onde radioélectrique de longueur d'onde λ vient frapper un obstacle, il y a une réflexion en sens inverse selon la normale, ce qui est tout à fait classique et lorsque l'on s'écarte de cette position, il y a une perte due au pouvoir réflecteur de l'obstacle. Mais dans le cas où nous avons affaire à des obstacles à structure périodique comme un câble torsadé (à notre connaissance, il n'y a parmi les lignes haute tension que des câbles torsadés) on constate lorsque l'on s'écarte d'un certain angle particulier θ, représenté sur la figure, une concordance des phases de réflexion de l'onde émise de chaque élément torsadé. Pour cet angle particulier, il y a un maximum d'énergie réfléchie, ce maximum est obtenu quand :

$$d \sin \theta = n \frac{\lambda}{2}$$

avec :

- λ = longueur d'onde : 8,6 mm
- n = nombre entier positif
- θ = angle d'incidence du faisceau radioélectrique
- d = pas entre spires

Installation d'essais

La photo 40-12 nous montre l'hélicoptère ALOUETTE III avec le radar équipé de son radome et la caméra d'axe qui nous permettait de prendre la photo du panorama en même temps que le radar.

Les photos 40-13, 14, 15, 16, montrent les détections obtenues sur colline, pylône isolé et ligne HT.

Un grand nombre de diapositives en couleurs seront présentées au cours de la conférence et permettront de mieux se rendre compte de l'intérêt des résultats obtenus.

Mais déjà on peut avoir une idée des possibilités de détection du radar SAIGA sur les quelques photos présentées ici.

Photo 40-13 : l'hélicoptère évolue en cabré le long de la pente d'une colline (vecteur vitesse au-dessus de la ligne de référence). La colonne lumineuse à gauche représente le test intégré.

Photo 40-14 : l'hélicoptère saute par-dessus le pylône isolé ; pour ce faire, il maintient en permanence le cercle (vecteur vitesse), sur le sommet de l'obstacle.

Photo 40-15 : détection d'une ligne HT de 90 kV en type C.

Photo 40-16 : la même ligne HT en présentation type B. Les marqueurs distance sont espacés de 500 mètres. La ligne est visible entre 600 et 1500 mètres.

Conclusion

Le radar SAIGA a démontré qu'il était particulièrement bien adapté aux conditions de vol d'un hélicoptère, appareil qui peut s'enfoncer beaucoup plus profondément dans le relief qu'un avion et qui doit éviter les obstacles aussi bien dans le plan vertical que dans le plan horizontal. Les résultats obtenus au cours d'essais en vol sur tous les terrains avec des obstacles naturels ou artificiels tels que pylônes et câbles haute tension, vérifient la validité des solutions originales choisies pour ce matériel.

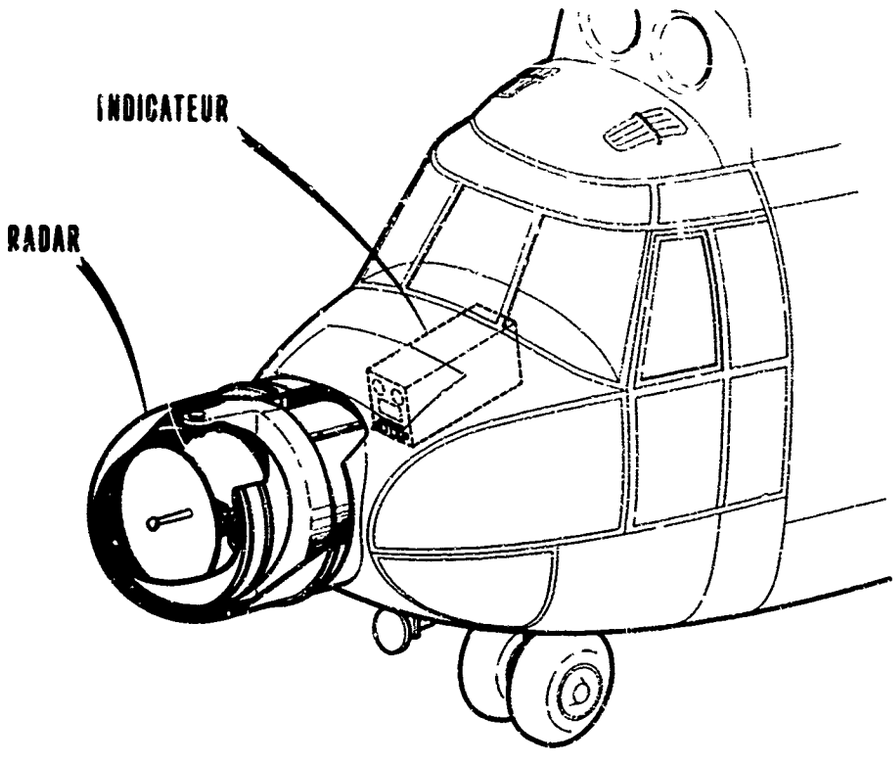


Fig. 1 SAÏGA radar d'evitement d'obstacles pour hélicoptère

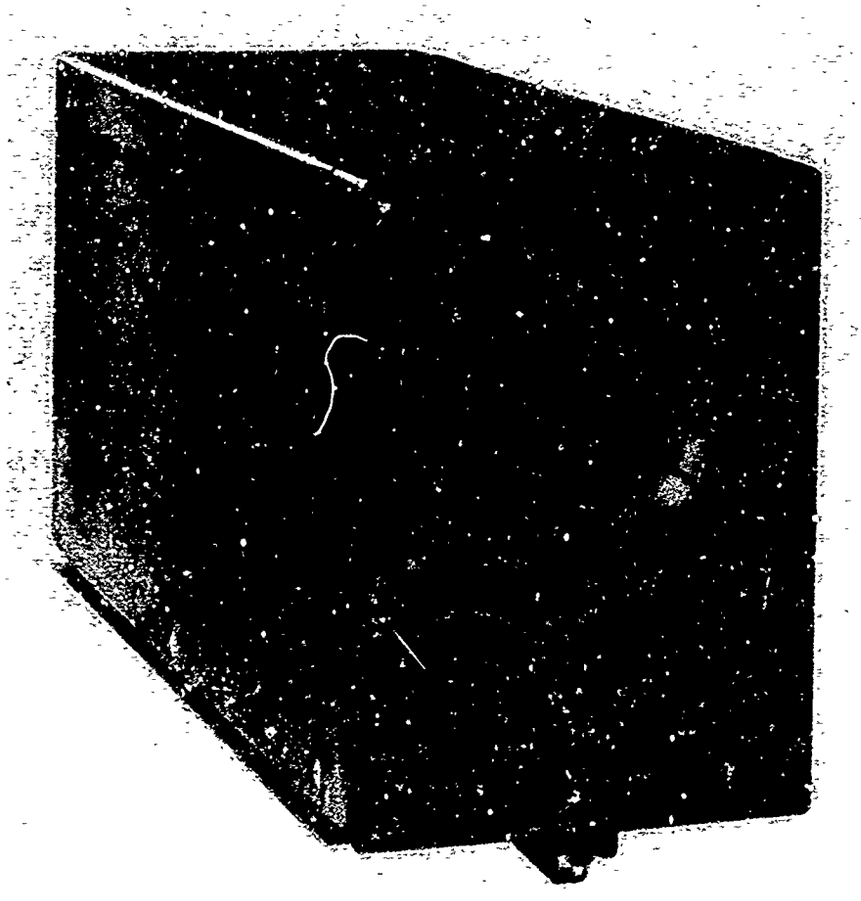


Fig. 2 Indicateur

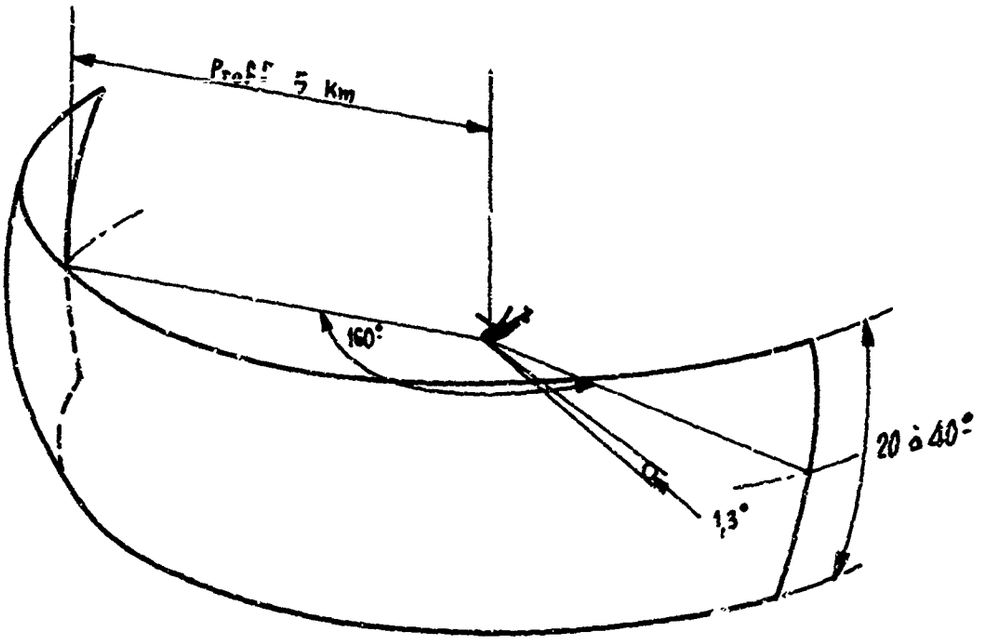


Fig. 3 Domaine d'exploration

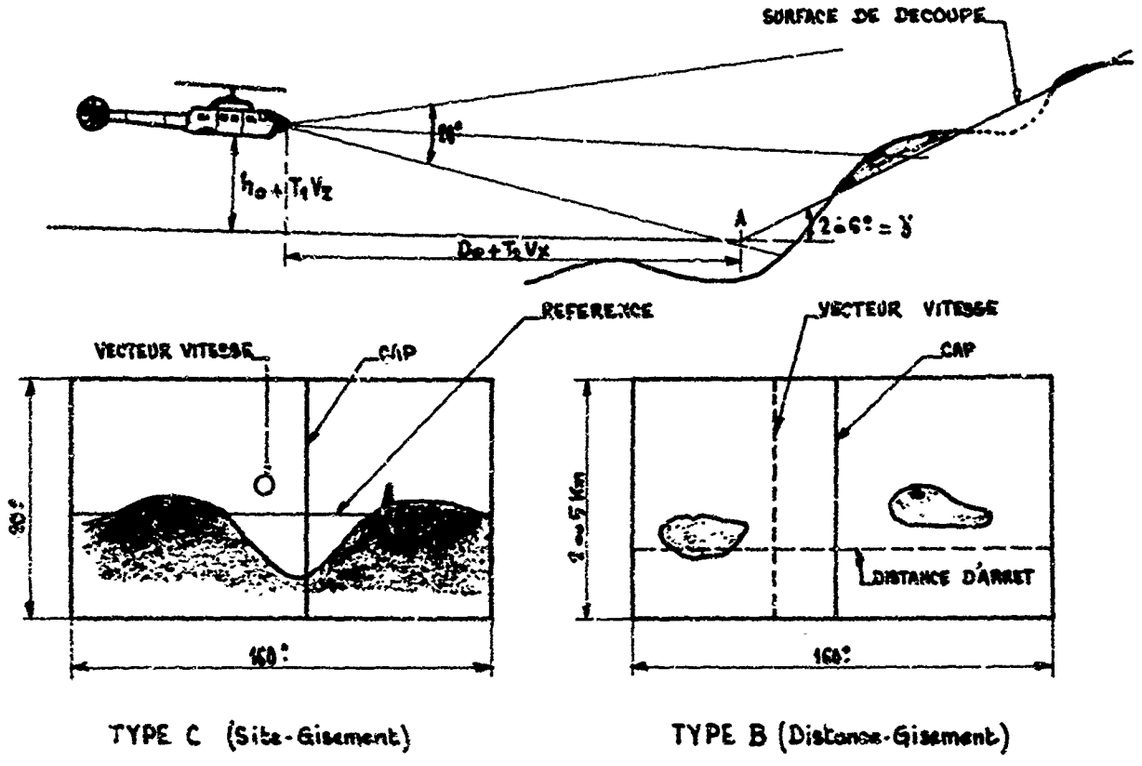


Fig. 4 Principe de localisation des obstacles

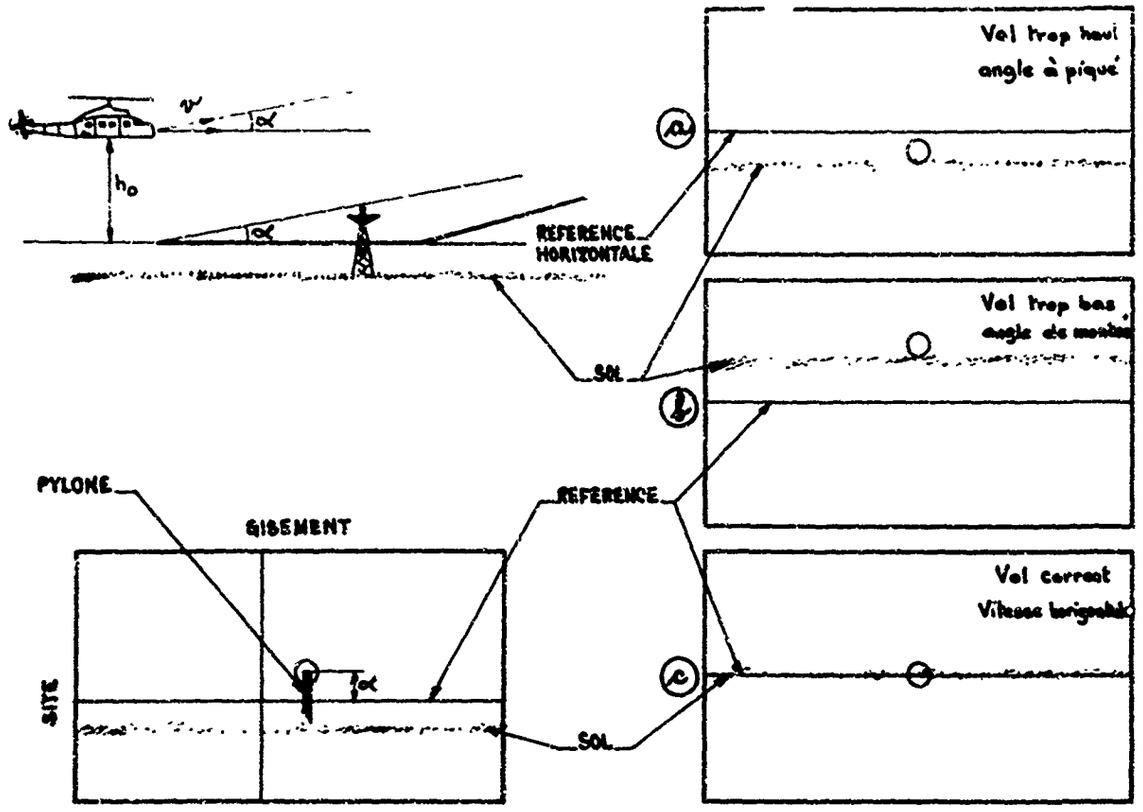


Fig. 5 Pilotage en type C

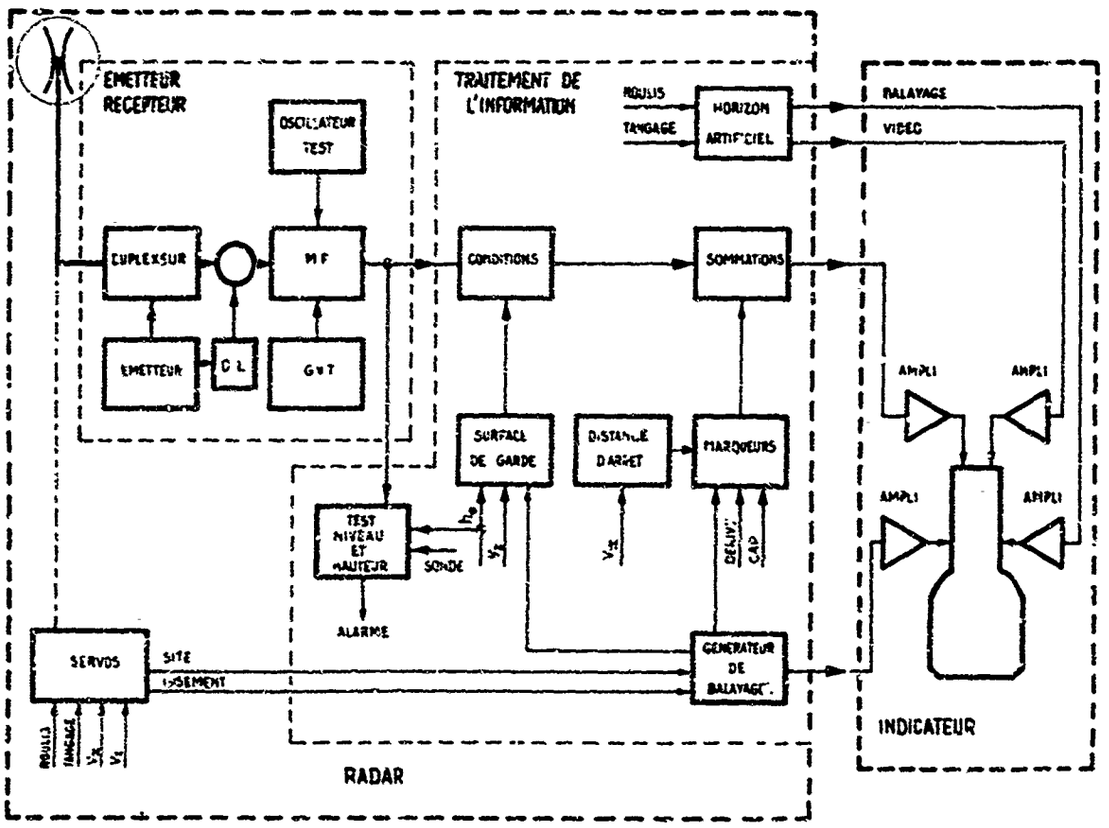


Fig. 6 Schéma synoptique général

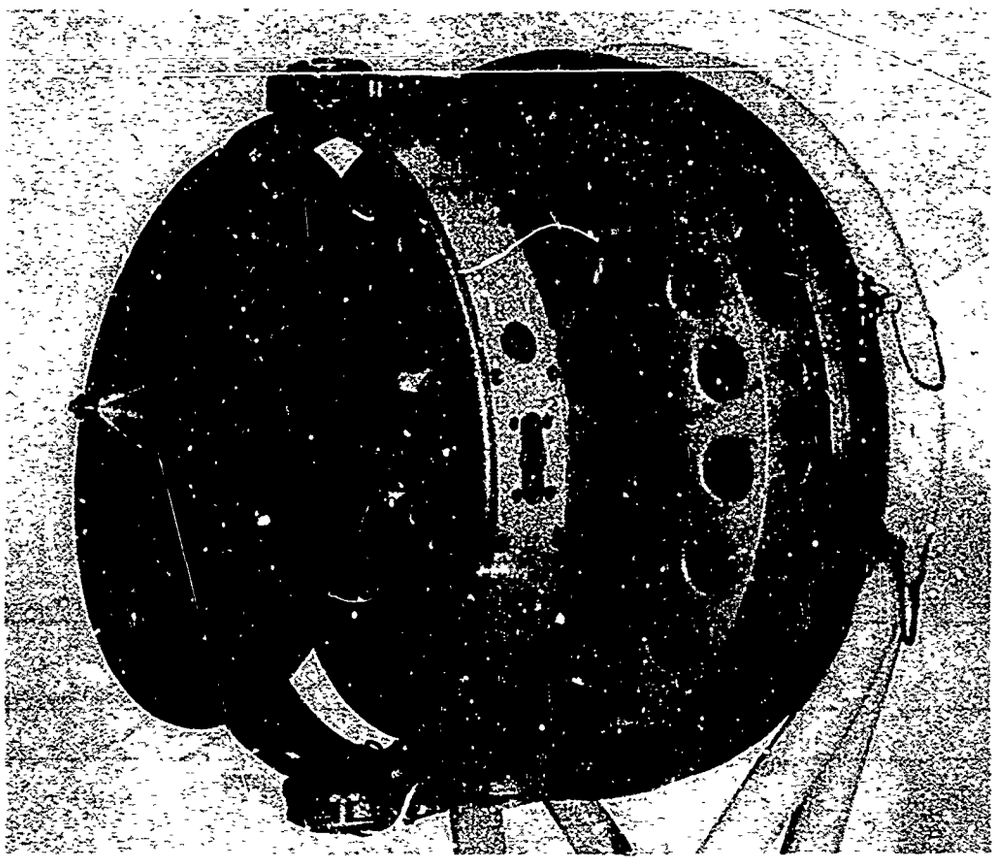


Fig. 7 Ensemble mobile

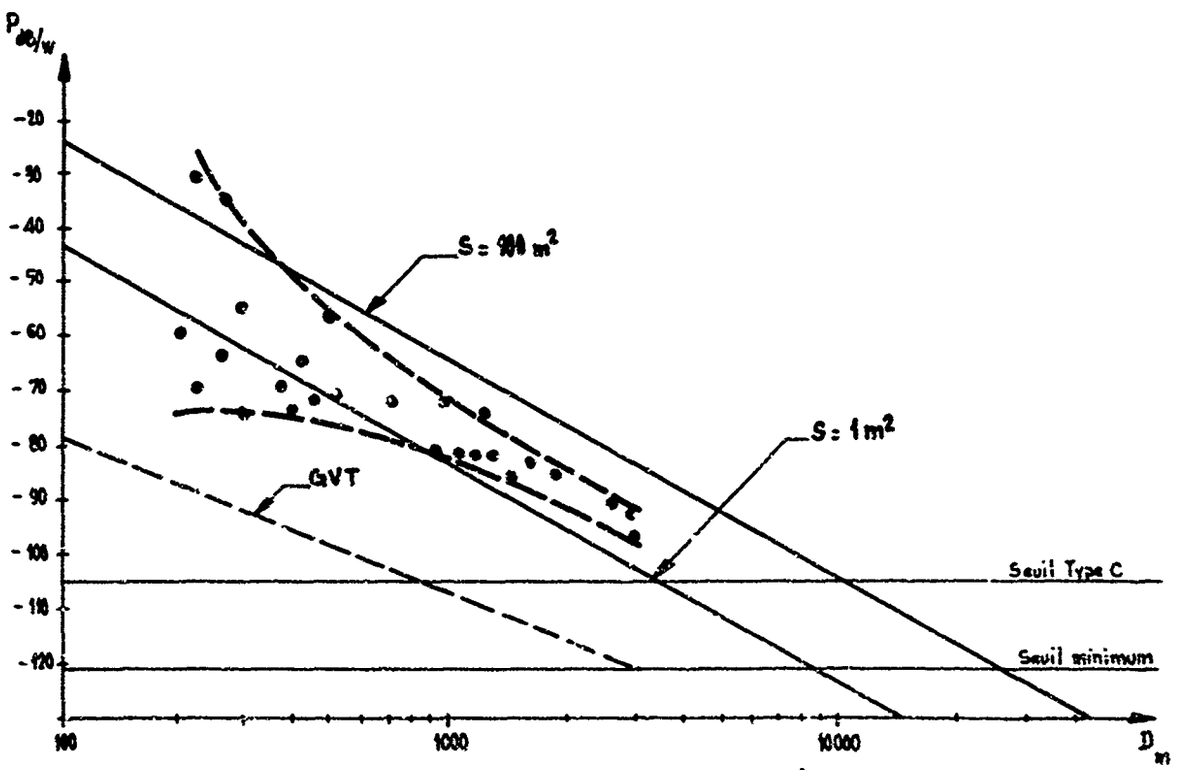


Fig. 8 Portée sur forêt

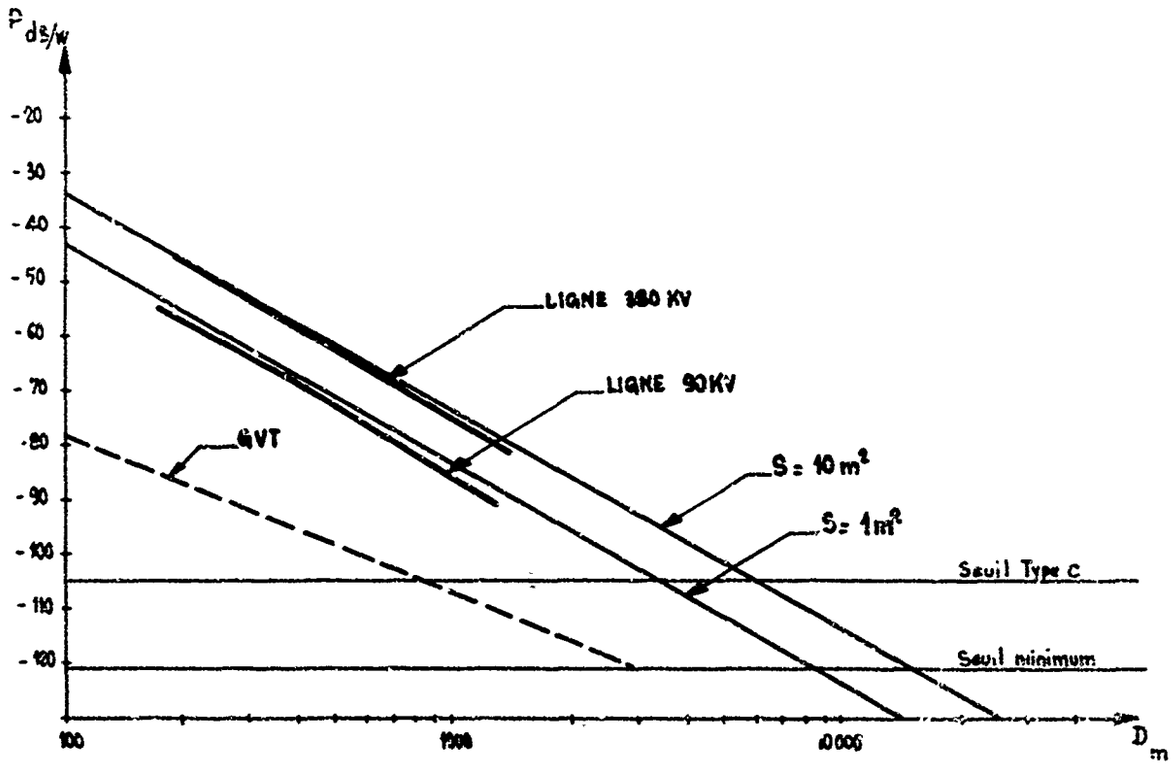


Fig. 9 Portée sur pylône HT

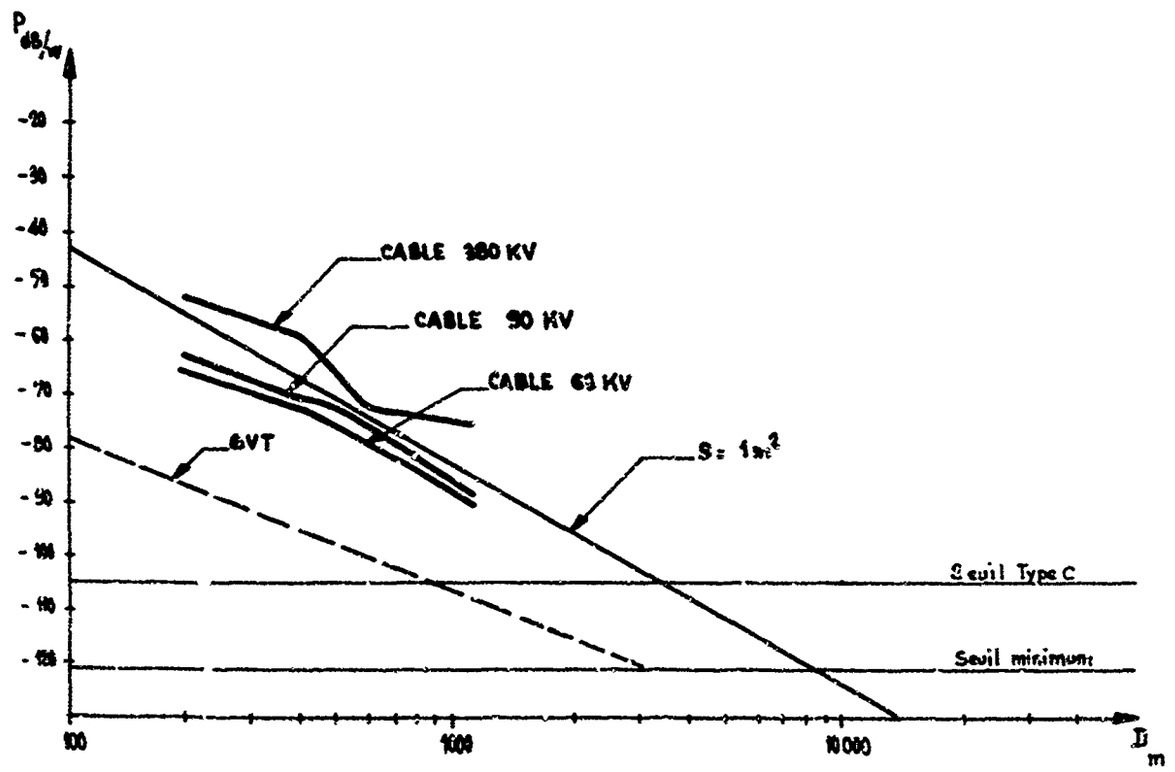


Fig. 10 Portée sur câbles HT

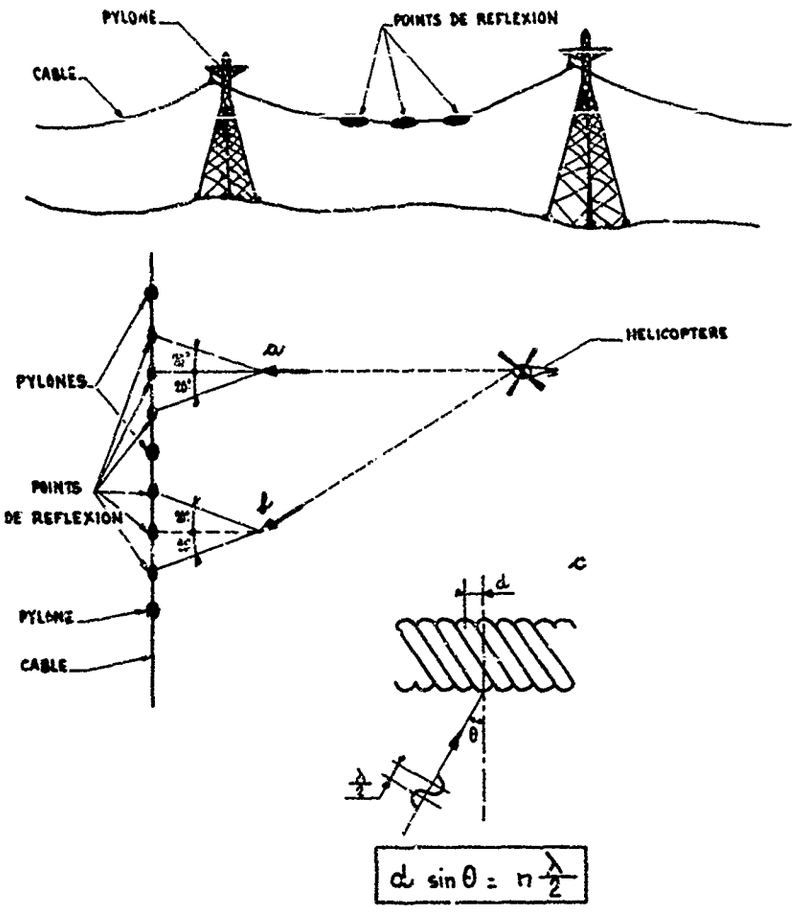


Fig. 11 Détection des câbles HT

Fig. 12 Radar installé sur hélicoptère ALOUETTE III

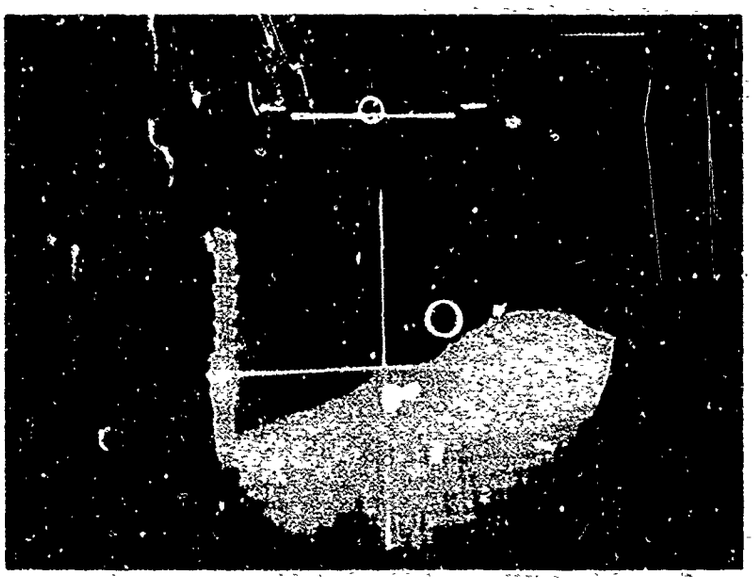


Fig. 13 Image type C sur colline

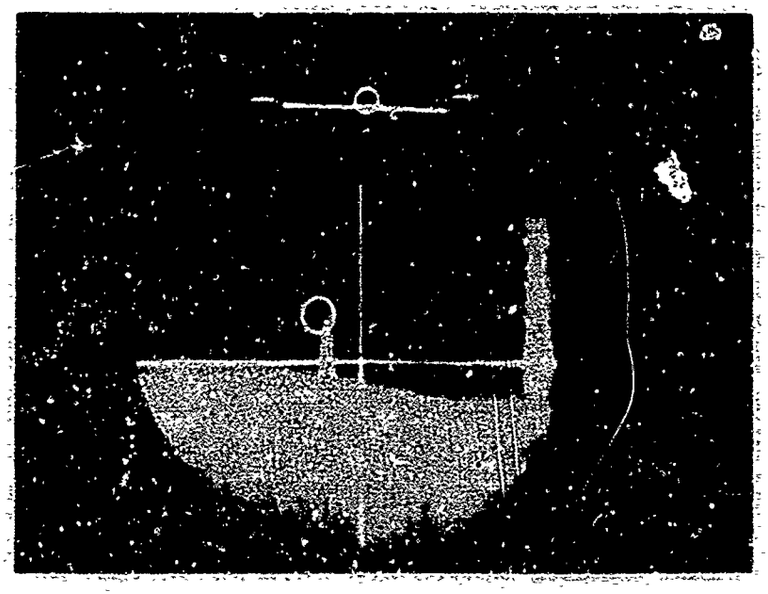


Fig. 14 Image type C sur pylône isolé

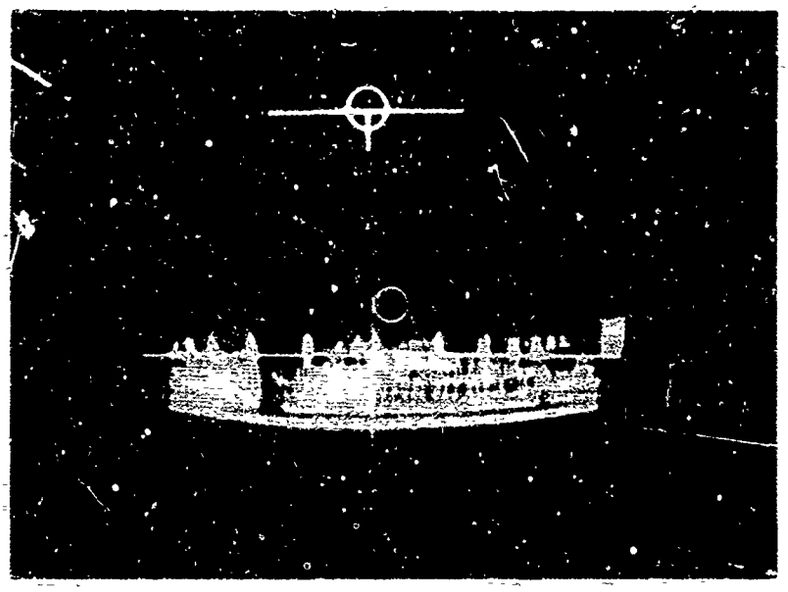


Fig. 15 Image type C sur ligne HT

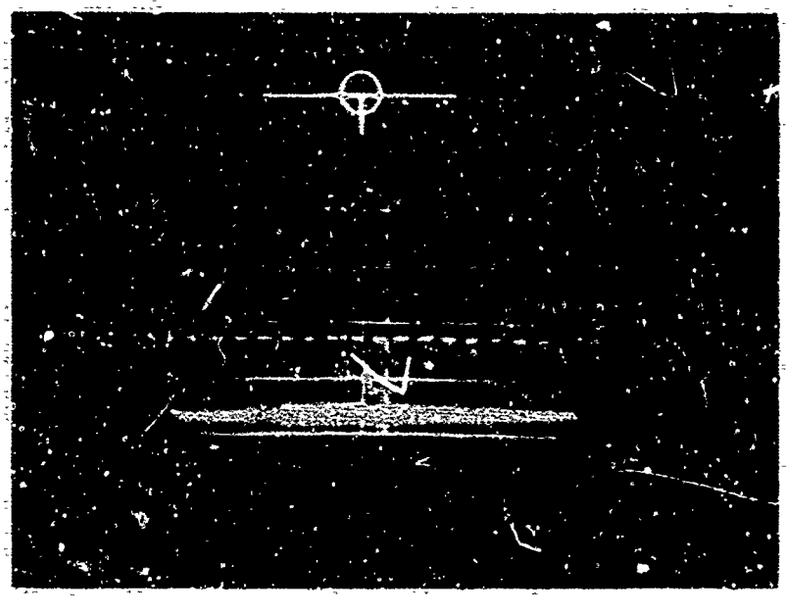


Fig. 16 Image type B sur ligne HT

A COST EFFECTIVENESS STUDY OF THE AIRBORNE EARLY WARNING AIRCRAFT AS AN INTEGRATED PART OF AIR DEFENCE SYSTEMS

by

T.H.Kerr

Ministry of Defence, UK

A COST EFFECTIVENESS STUDY THE AIRBORNE EARLY WARNING AIRCRAFT AS AN INTEGRATED PART OF AIR DEFENCE SYSTEMS

T.H. Kerr, BSc., AFRAeS
MINISTRY OF DEFENCE U.K.

SUMMARY

Advances in radar technology now make it possible to detect low flying targets against the ground clutter. Using this type of radar (*Pulse Doppler*) installed in Airborne Early Warning (AEW) aircraft as part of a mix with fighters and Surface-to-Air Missile (SAM) systems, it is possible to obtain an all-altitude defence in depth and, in particular, considerably improve the effectiveness of the defences against attack by low flying aircraft. The effectiveness of a constant cost mix of fighters with AEW as compared with that of fighters alone and the improvement in effectiveness is shown. In addition, the variation of overall effectiveness of the defences of a task force with AEW radar range is examined.

Using a mathematical model to simulate the operation of AEW aircraft and radar systems, the number of aircraft required to man one station is presented as a function of the aircraft and radar size, power etc. and the probability of being on patrol.

1. INTRODUCTION

The amount of early warning of an enemy attack that can be obtained from surface-based radar is severely limited by the earth's curvature and to a certain extent by local terrain and vegetation screening. Whenever possible these limitations have been partly overcome by locating the sensor on high ground; for example, on the edge of a cliff when looking over the sea, or on a hill-top. This has considerably increased the radar horizon, but it is still severely limited unless the radars can be sited well forward of the area to be defended. A greatly increased radar horizon can be obtained by locating the radar in an aircraft flying well above the earth's surface. Developments in radar now make it possible to detect low altitude attackers against the ground clutter by virtue of their differing radial velocities. If a radar of this type is installed in an aircraft, it can be integrated with the overall defence systems and provide the valuable early warning necessary for alerting all the systems in the defence environment. The concept of the airborne early warning system and the coverage that might be obtained from the radar are shown in Figure 1-42; any attacker, flying at low altitude, can be detected, the defences alerted and the fighters scrambled and directed against the raid. Without such a system, the attackers can underfly the coverage of the ground radars, achieve surprise against the defence systems and more readily saturate them. The airborne early warning system, therefore, appears to offer a solution to the problem of defending against the low altitude attacker and in this paper an attempt is made to take some very simple situations and demonstrate the advantages on a cost effectiveness basis of including the airborne early warning system in a defence mix.

2. THE ASSUMPTIONS

2.1. Basic Concept

When examining alternative policy options for resource allocation in the future, the work is most easily presented and understood if force mixes of constant cost and varying effectiveness or constant effectiveness and varying cost are compared. In this paper the former presentation is used. In order to establish the costs of the alternative aircraft mixes to be purchased, it is necessary to consider the research and development costs and the total purchase and operating costs for the life of the aircraft. The life for aircraft is normally assumed to be 10 years. Research and development costs are not used in this paper. A separate treatment of R&D costs is essential because of the large variations caused by different assumptions on the method of costing, and whether it is a fixed initial cost, a cost shared with other nations, or amortised over the total buy. In this paper it will be assumed that it is an initial fixed cost not included in the comparisons. This means that the comparisons will assume that we are dealing with alternative ways of increasing the force effectiveness in the airfield and task force scenarios at a constant incremental cost.

2.2. The Fighter

The fighter is assumed to be capable of flying at a Mach number of 2 at heights up to 60,000 ft and to have a maximum endurance of two hours. It is assumed to carry four medium range air to-air missiles which can be fired at targets flying at heights above or below the fighter. The airborne intercept (AI) radar is assumed to be of the *pulse doppler* type which can discriminate targets against the clutter from ground returns. The AI is assumed to have a *lock on* range of 40 nm against 5 sq. metre targets. The combined weapon system is assumed to have 0.5 expected kills per pair of missiles when fired after being positioned against the raid by broadcast type of ground controlled interception, and 0.25 expected kills per pair of missiles fired when operating autonomously. Such a fighter is assumed to cost £3.1M including the initial purchase price and 10-year running costs.

2.3. The Airborne Early Warning Aircraft and System

In a typical deployment (Figure 1.42) the airborne early warning aircraft would fly above the target to be defended. The coverage of the ground radar is shown and it can easily be seen that because of the ground radar horizon limitations, an attacking aircraft flying at low altitude can underfly this coverage and remain undetected until it is very close to the target. With the airborne system a low

flying attacker would be detected at a much longer range. Typical layouts of such an aircraft system are shown in Figure 2.42. Type A has aerials in the nose and tail of the aircraft, and Type P has a mushroom-type aerial above the fuselage. Each type is assumed to be capable of detecting 5 sq. metre targets (typically large fighter ground attack aircraft) at a range of 150 to 250 nm, depending upon the details of the aerial size and the power transmitted. The main advantage of the fore and aft layout over the mushroom arrangement is derived from the increased aerial depth for a given aircraft size and, therefore, the improved height-finding capability that can be obtained.

For this paper it has been assumed that the all-up-weight of the aircraft is about 60,000 lb, an endurance of five hours on station at 30,000 to 50,000 ft, and a detection range of 160 nm against 5 sq. metre targets. The purchase price and 10-year running costs are assumed to be £6.1M per aircraft and five aircraft are required to ensure that one will be on station for more than 90% of the time. (The investigation of the numbers required to man a single station will be given in detail later in the paper.)

2.4. Surface-to-Air Missile Systems

Surface-to-air missile systems are usually deployed near the defended area. For the defence of a task force at sea (discussed later) it is assumed that two types of surface-to-air missiles are available:

(a) A short range system with a maximum range of 5 km for point defence.

(b) Two long range systems of 30-km range at low altitude and 70 km at high altitude.

The longer range systems provide area defence for the task force. The costs are assumed to be sunk costs and the variations of defence effectiveness as a function of AEW range are investigated.

3. THE DEFENCE OF AN AIRFIELD

3.1. Introduction

The fighters on the airfield are assumed to be deployed to protect attack aircraft based on the same airfield and to maintain facilities, etc. They can also provide area defence for other facilities or armies within their area of operation. The mode of operation depends upon the availability of warning and direction of attack. With warning, the fighters can be scrambled from ground alert: without it, they must be operated partly or completely on continuous airborne patrol (CAP). These points are discussed in more detail below.

It has been assumed that the total purchase and 10-year running costs of the fighters and fighter/AEW mix will be the same. Taking the estimates of cost given earlier of £3.1M per fighter and £6.1M per AEW aircraft, the constant cost mix at £93M would give:

20 fighters plus 5 AEW

or

30 fighters.

3.2. Operations With AEW

It is assumed that the fighters are held at ground readiness and when alerted, scramble at 30-sec intervals after a 3-minute delay. Against a raid of low altitude attackers penetrating at 500 knots, it is assumed that they cruise out at 20,000 ft at high subsonic speeds, descending to 5000 to 10,000 ft to make a first attack head-on against the raid. After this attack, the fighters turn into a tail chase and accelerate to supersonic speeds, attacking again with a pair of missiles. The following table shows the expected number of kills, assuming that the last intercept must take place at least 10 nm from the airfield.

KILLS TO BE EXPECTED FROM FIGHTERS SCRAMBLED BY AEW FROM GROUND ALERT

Number of Fighters Available	Air Intercepts per Aircraft	Total Intercepts	Expected Kills
20	1.5	30	15 to 30

3.3. Operations Without AEW

Two modes of operation without AEW will be considered. The first assumes that the direction of attack is unknown and a continuous airborne alert is therefore required covering attack from any direction and using all possible fighters. The second mode assumes that the situation allows patrols to be operated over a limited sector and the airborne fighter on patrol can alert other fighters at readiness on the ground.

All Fighters on CAP. Experience has shown that some five fighters are required to maintain one on CAP. Thirty fighters would therefore provide six airborne at any time. It is also assumed that these fighters fly a circular track some 40 miles radius from base (Figure 3.42). The separations between the fighters and their AI search patterns will give a minimum detection range of some 50 nm from base. The number of intercepts possible in the autonomous mode will probably be 0.5 to 1.0, and using the missile load in a single attack the kills possible would also be 0.5 to 1.0.

Mixed CAP and Ground Alert. Although the risks are higher, if it is thought that there is a preferred attack direction (for reasons of terrain, limited range, etc.) it is possible to use some of

the fighters as AEW aircraft and scramble from ground readiness those fighters not involved in the CAP operation. One possible type of patrol pattern is shown in Figure 4.42 providing warning over a 180° sector. In this operation it has been assumed that three aircraft would be maintained on CAP, utilizing 15 of the total available. In the most favourable assumptions, the remaining 15 could be scrambled against the incoming raid. In this autonomous mode of operation one of the three on CAP and the 15 scrambled might achieve 0.5 to 1.0 intercepts and each intercept has 0.25 chance of a kill; resulting in 2 to 4 kills.

3.4. Overall Comparison of Force Effectiveness

At a constant cost of £53M the effectiveness of the fighter defences for an all-fighter or fighter/AEW mix is shown in Figure 5.42. This indicates an improvement in effectiveness of some 15 to 30 times when the force includes AEW.

4. PROTECTION OF A TASK FORCE AT SEA

4.1. Concept of Operation

The previous section considered the additional effectiveness achieved by fighters when an AEW aircraft is included in the force mix. This section examines the variation in effectiveness of the AEW with range when included in a fighter and missile environment.

Let us imagine the defence of a task force at sea in a limited war environment in which conventional HE weapons are used. The systems that might be used in the attack are illustrated in Figure 6.42. They consist of direct attack at low, medium and high altitudes, and stand-off attack by short and long range weapons. These have to be countered by the defence systems (Figure 7.42) consisting of fighters supported by AEW and surface-to-air (SAM) systems. For the purpose of the study, only low altitude direct and stand-off attacks will be considered, since the medium and high altitude attacks can be countered by the defence systems supported by the radars in the ships. The success of the defence systems, and in particular the AEW, will be measured by the probability of damage to the most important target within the task force.

4.2. The Defence Effectiveness

Defence effectiveness is calculated in two stages; first the fighters and then the SAM systems which cover an area of radius 20 nm from the main elements of the task force.

Let us assume that seven fighters are available. Of these, two are on continuous airborne patrol and the remaining five are scrambled at 3,3½,4,4½,5 min. The climb and speed profile of the fighter, the detection and tracking of the AI radar, the missile launch, flight path and probability of kill are modelled. The model can then be used to estimate the expected kills per sortie. Assuming that the low altitude attackers are flying at 600 knots and that the AEW radar can detect them at 200 nm, then the two fighters on CAP can make the first intercept in head-on attack at 110 nm (Figure 8-42) and the scrambled fighters at 90 to 80 nm. By this time the CAP fighters have turned and overtaken the raid, making the second intercept at 70 nm, followed by the other aircraft at 50 to 40 nm. In this way the expected intercepts and kills can be calculated. This is a very simple description of a detailed and complex tactical model being used to assess the capability of detection, conversion to attack and kill.

The calculations can be repeated assuming that the AEW radar detection range is 150 and 100 nm and the fighters' capability re-assessed. It has been assumed that the fighters do not enter the 20-nm defended zone. The attackers can, of course, be re-engaged when they leave the SAM-defended zone.

A mathematical model of the SAM defences (Figure 9-42) is used to assess the kill capability of the SAM systems deployed, taking into account their position, range, single shot kill probability, etc. Unless there is an obviously preferred attack direction, the SAM effectiveness is generally assessed over all possible directions of attack. The surviving raiders are assumed to deliver their weapons and the effectiveness of the attack in terms of damage is calculated. This will depend, of course, on the weapon type and load, delivery accuracy and target vulnerability.

In order to illustrate the effectiveness of the defences for different AEW ranges, a number of different strengths of the attacking force have been considered, the largest force being a little over 20 aircraft. Because of mutual safety considerations, this is about the largest force that can be co-ordinated within a single raid. Larger numbers of attackers may be assumed to be organized in two raids separated in time by more than the reload time of the SAM systems.

For a low altitude attack of a given strength, it is possible to show how the damage in the task force varies with the range of detection of the attackers (Figure 10-42). If the numbers of aircraft in the low altitude raid is less than five, the SAM systems alone can provide adequate protection and little damage is achieved. If the number of attackers is increased to 10, the SAM systems plus two fighters on CAP cannot prevent considerable damage. If, however, an AEW system of 100-nm range is included, fighters can be scrambled and the damage caused by 10 raiders reduced to a minimum. Fifteen raiders can be countered in a similar way by including a 150-nm range AEW, and with a 200-nm range AEW a further small improvement can be obtained. Twenty aircraft in a single raid saturate the defences and a further increase in AEW range is of no value unless the fighter defences are strengthened. The effectiveness of the low altitude attack, therefore, varies with AEW radar detection range and the manner of this variation is illustrated in Figure 11-42. As the fighters' AI radar and CAP station can provide about 50-nm warning, there is no advantage in deploying an AEW of this range. When the range exceeds 50 nm, the effectiveness of the defences increases rapidly to about 150 nm and then the rate of increase falls until at about 200 nm there is no further increase in effectiveness with range. The curve is of the classic *S* shape and indicates a choice of radar range of 150 to 200 nm.

The middle diagram indicates the change in effectiveness as a function of range when stand-off weapons are considered. When this and the direct attack effectiveness are combined, the overall attack effectiveness is as shown in the lower diagram. The same pronounced *S* shape of the curve is maintained, but at ranges greater than 200 nm the effectiveness of the AEW continues to increase but at a low rate. When the defence decision planner chooses the AEW system, the relative magnitude of the direct attack and stand-off threat must be considered as well as the size and cost of the AEW system.

It is important to stress that in these calculations there are many factors that cannot be taken into account directly. For example, it is not possible to say how many kills by the fighters would break up the raid and thereby ease the problems for the SAM defences or cause the attackers to abandon the attack, or even what level of defence would deter the enemy from initiating the attack. These are matters of experience and judgement that depend upon the particular environment of the study, but may be very important.

5. A SIMULATION OF AEW AVAILABILITY

The AEW system consists of a medium to large aircraft combined with a medium sized radar. Such a system is complex and any country considering introducing it into its force structure must be satisfied that it will be on station and operating when required. It is possible to predict the reliability and maintainsbility of such a system and use these data in a mathematical model to simulate the operation of the system and to investigate the number of aircraft required to operate one station and the effect of different aircraft and radar designs. The most important factors that can be used as variables in the model are:

(a) The aircraft

- (i) speed and height of patrol
- (ii) endurance
- (iii) defect rate
- (iv) repair time and manpower required

(b) The radar

- (i) the power and aerial size
- (ii) defect rate
- (iii) repair time and manpower required.

It is accepted that reliability is notoriously difficult to predict and is continually changing (improving, one hopes) throughout the service life of an equipment or aircraft. Nevertheless it is a very important parameter in any model seeking to present a complete picture of the defect rates and therefore of the spares and manpower required to support any piece of equipment. The larger and more complex the equipment, the lower the availability will be and the higher the spares and manpower required. To provide a base line from which we might predict future equipment reliability after one to two years in service, a comparison was made between the defect rates predicted by a component count method proposed by the Royal Radar Establishment on current equipment and data available from records of defect rates of the aircraft radar in squadron operations. It was found that there was a factor of approximately 3 between these estimates, the actual number of defects being higher than predicted. Component counts were done for the new equipments and the same factor applied. Studies were also made of various airframe and engine combinations. It was found that the defect rate varied directly with size and this was used to predict the reliability of possible future aircraft.

Before these data could be used as inputs to a model designed to investigate the overall availability of the possible AEW systems, two further items of information were required:

(a) The distribution of total time elapsed between landing and being available for a further sortie.

(b) The distribution of the occurrence of defects in flight.

There is a tendency for maintenance organizations to adopt a somewhat different mode of operation in war, compared with that used in peace-time. In war, some of the defects of a minor nature are not repaired until an opportune time and it is assumed, for the purpose of illustration, that only the most important two-thirds of the defects are repaired. Figure 12-42 shows the cumulative probability of the aircraft being ready for the next flight as a function of the time elapsed since landing for sustained rates of operation of up to three months. This shows that there is a 50% probability of an aircraft being available after four hours and a 90% probability after 12 hours.

The distribution of occurrence of defects in flight has been measured in trials and in simulated flight conditions in the laboratory. It was established that a large number of defects were found to occur at or just after take-off. Figure 13-42 shows measurements taken at 30-minute intervals of the defects arising per half hour. On a typical aircraft and radar system, it was found that the probability of a defect occurring is approximately 10 times higher in the first 30 minutes than in any subsequent half hour. This effect is probably caused by a combination of factors such as switch-on failures, inability to test some equipments adequately on the ground, failures accumulated after the aircraft was prepared for flight, the environment during take-off, etc. The knowledge that such a distribution exists is extremely important when attempting to predict the effect of increasing the sortie duration.

This information was then used as input to a *Monte Carlo* model of the aircraft operations. Figure 14-42 shows pictorially the flow diagram used. With this model, tests can be made to show the effects of reliability, sortie duration, in-flight refuelling, changing repair philosophy, rules for allowing scheduled servicing, etc.

A typical output from the model, assuming a total of four aircraft are available to provide one on CAP, is indicated in Figure 15-42. The probability that an aircraft will be on station at height is called the *system warning probability*. It can be seen that on average, an aircraft would be on station 95% of the time with a spread of possible results from 85% to nearly 100%. The system warning probability for three, four or five aircraft available is shown in Figure 16-42 and the expected values are about 85%, 95% and 97% respectively.

With the results of these models it is possible to examine the variations in effectiveness and cost as a function of all the important parameters governing the capability of each size and power of the AEW systems considered. The variation in system warning probability with the number of aircraft to man each station can also be examined.

6. CONCLUDING REMARKS

In view of the existing weakness in low altitude air defence and the assumption that an attacker will always exploit a weakness, the AEW system offers the possibility of detecting the attackers early enough to alert the defences, scramble fighters to intercept the raid and, with a suitable airborne intercept radar (*pulse doppler*) air-to-air missile system in the fighters, provides a means of engaging them long before they reach the target area. It has been shown that, by taking reasonable assumptions of cost and effectiveness, within a given budget a fighter/AEW mix can be much more effective than fighters alone, and by considering the fighter plus AEW/SAM mix, it is possible to indicate the variation in effectiveness of the defences as a function of AEW range. There is obviously a wide spectrum of possible solutions to the requirement for the AEW system in terms of both the aircraft and radar. The relative advantages and disadvantages of each solution can be investigated by means of the defence models and the availability studied by means of the simulation of the operation of the various AEW solutions.

If the addition to the defences of an AEW system gives such a large increase in the defence effectiveness, will not the enemy make it a priority target and, therefore, can it be of any real value? This question must be asked by any country considering its introduction. The enemy may attack it in two ways; firstly by electronic warfare and secondly, by physical attack. Against the jamming threat, countermeasures such as frequency diversity, etc. can be built into the AEW radar and the capability of the enemy can be degraded by changes in physical deployment such as position or height when these are possible. With reference to physical attack, the AEW is not an easy target. It is deployed within a defence system and is in the best position to know when and if it is likely to come under attack. It can, of course, use AI or missile jammers to defend itself. Every situation requires a separate examination, but an investigation of many of these possibilities leads to the conclusion that these arguments do not negate the significant advantages in favour of deploying such a system.

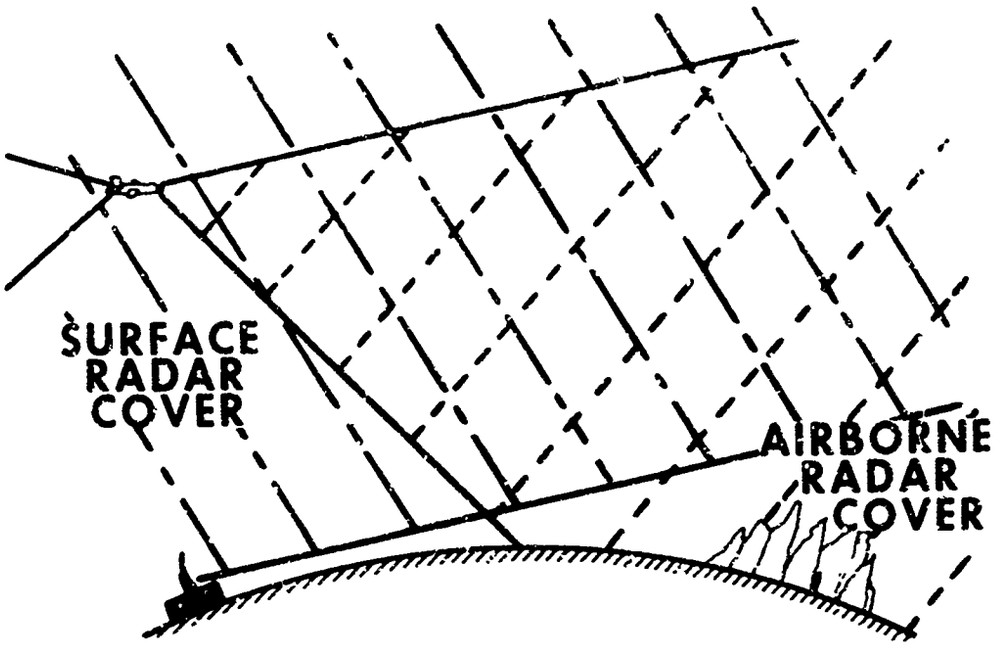
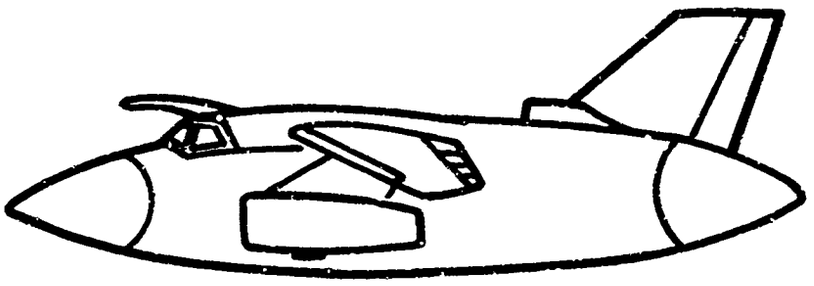
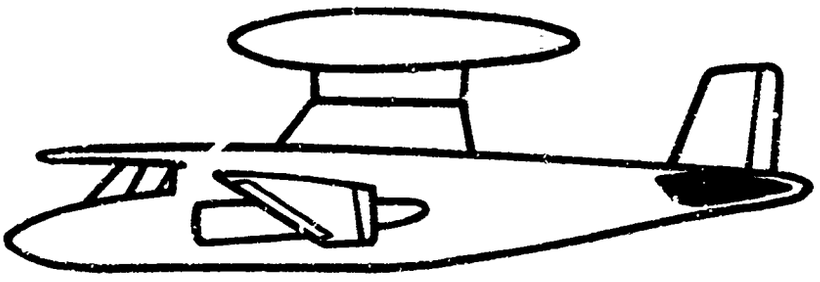


Fig.1 The radar environment



Fore and Aft Layout



Mushroom Layout

Fig.2 Typical AEW aircraft

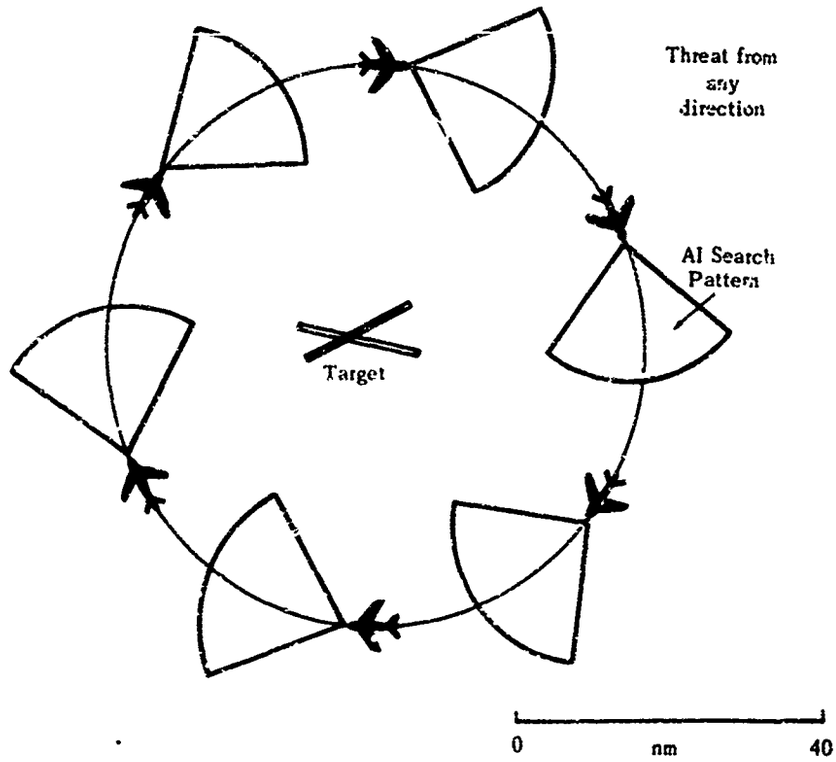


Fig.3 CAP operation of six fighters

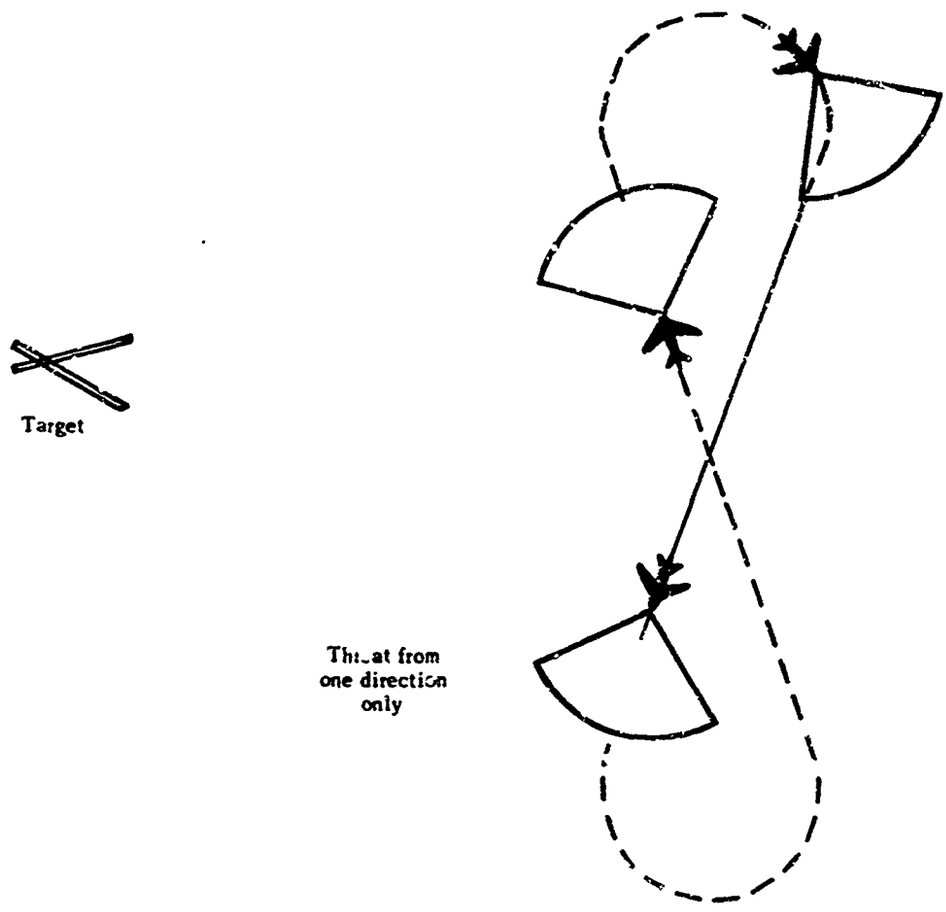


Fig.4 Limited CAP operations by two fighters

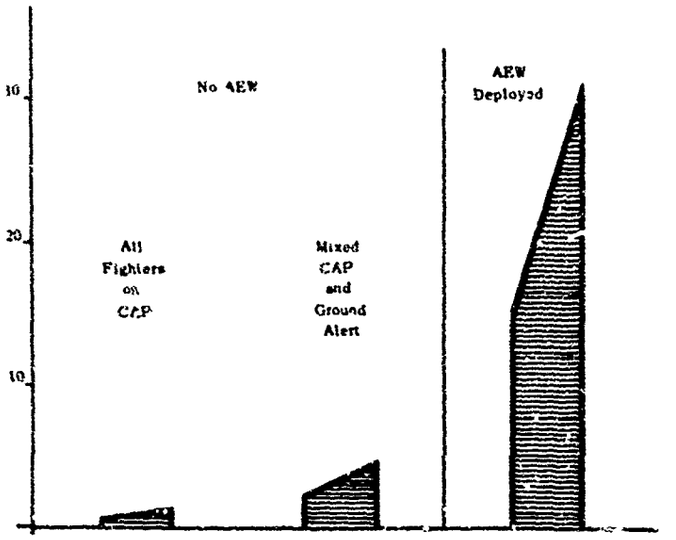


Fig.5 Comparison of the defence effectiveness with and without AEW

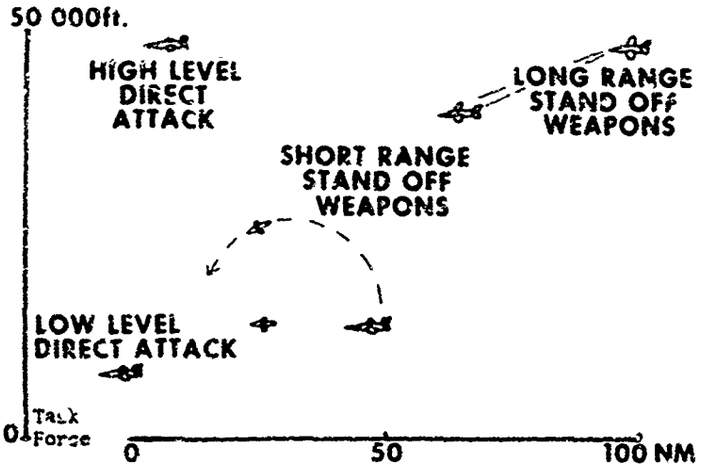


Fig.6 Attack systems

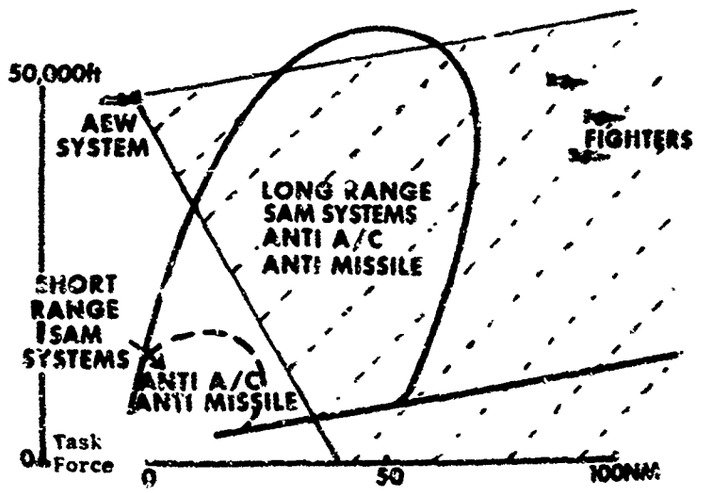


Fig.7 Defence systems

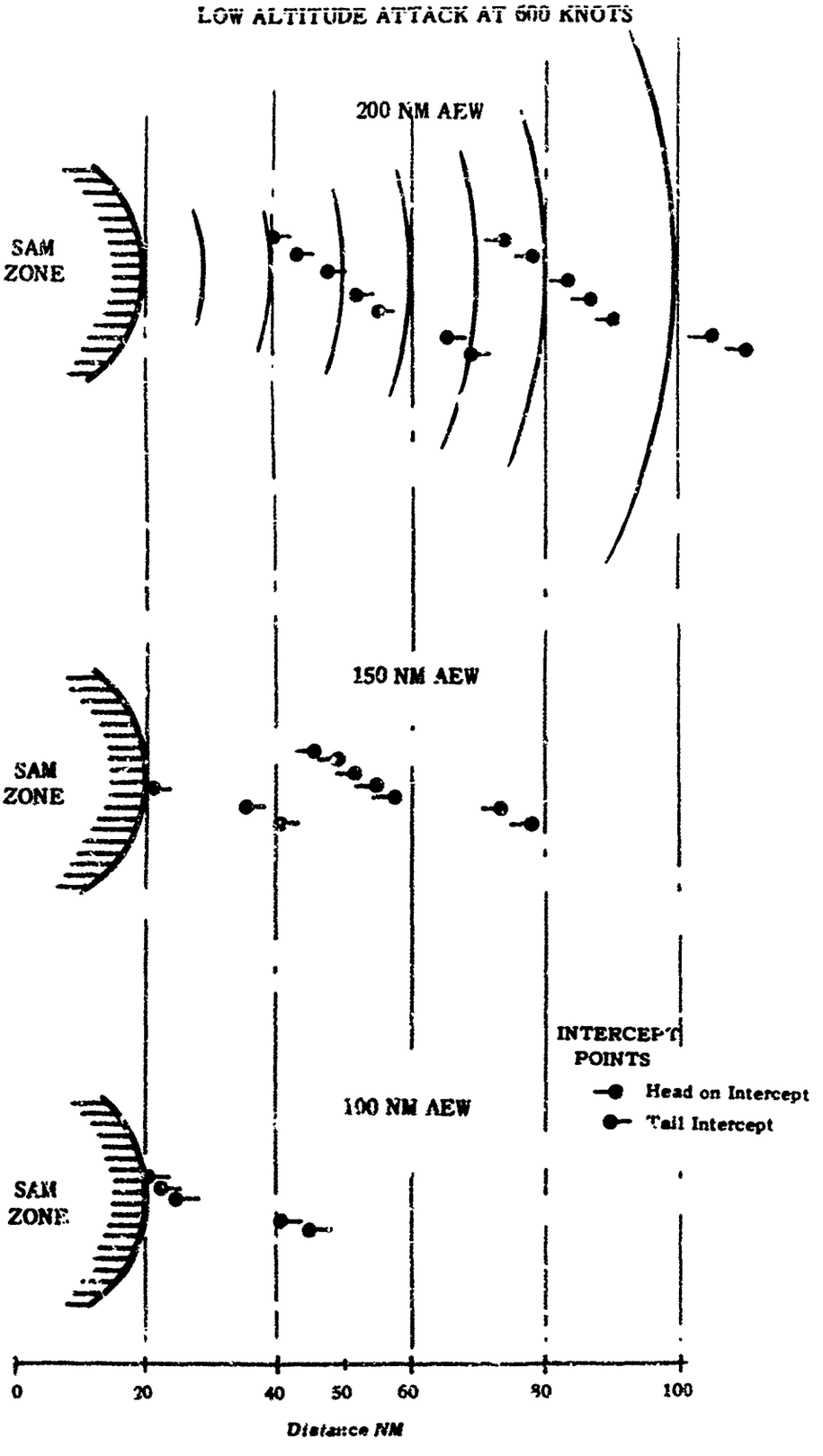


Fig.8 Range of intercept from the task force as a function of AEW radar range

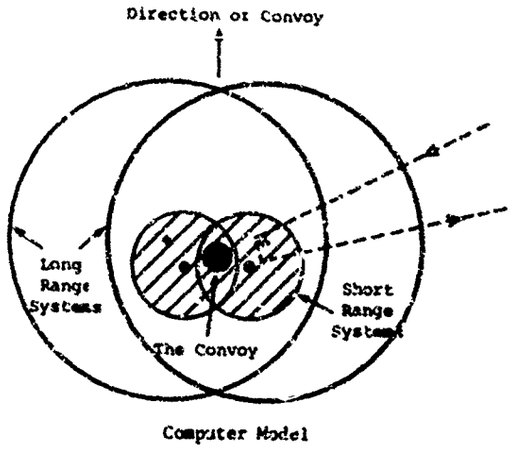


Fig. 9 Computer model of SAM systems

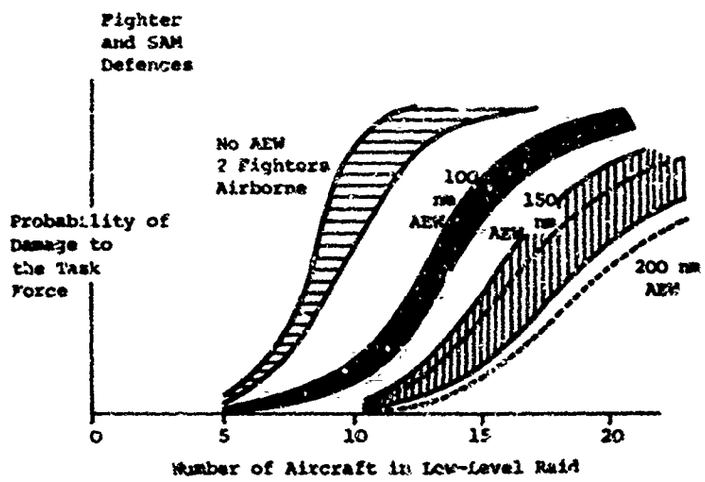


Fig. 10 Probability of damage at the task force as a function of AEW range and raid size

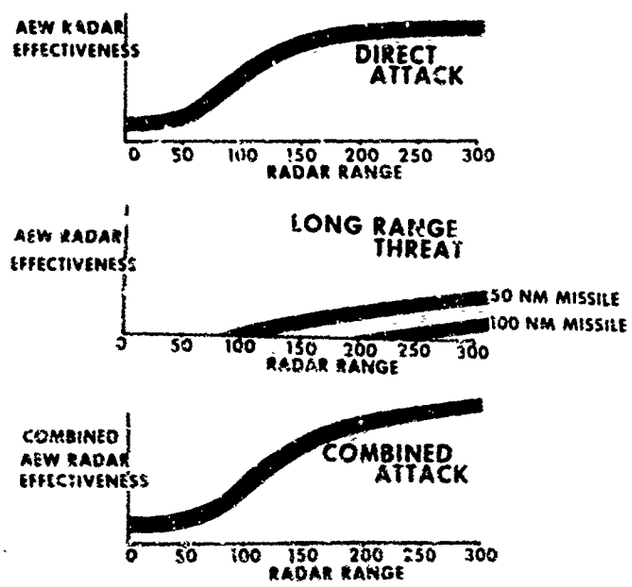


Fig. 11 Variation in defence effectiveness with AEW range

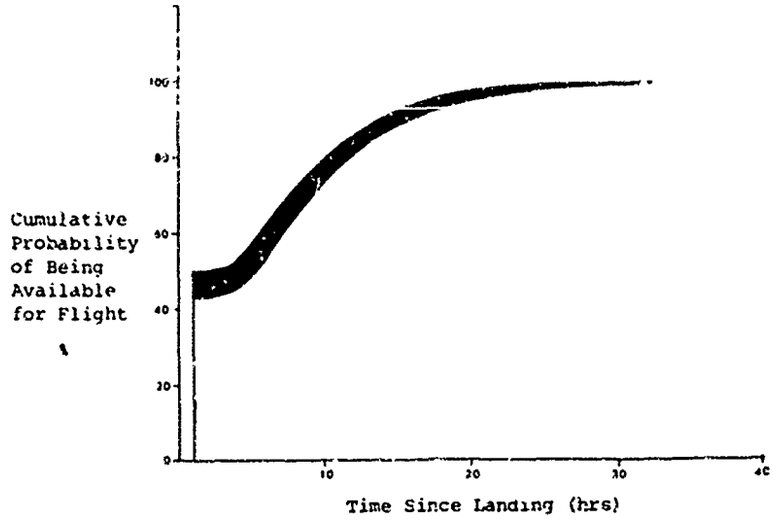


Fig.12 Cumulative probability of being available for flight. As a function of time since landing

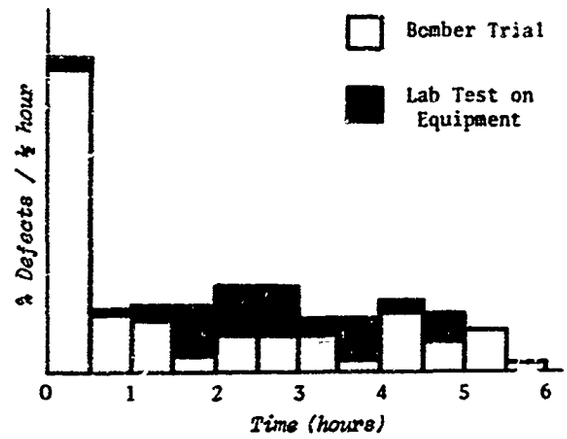


Fig.13 Distribution of defects in flight

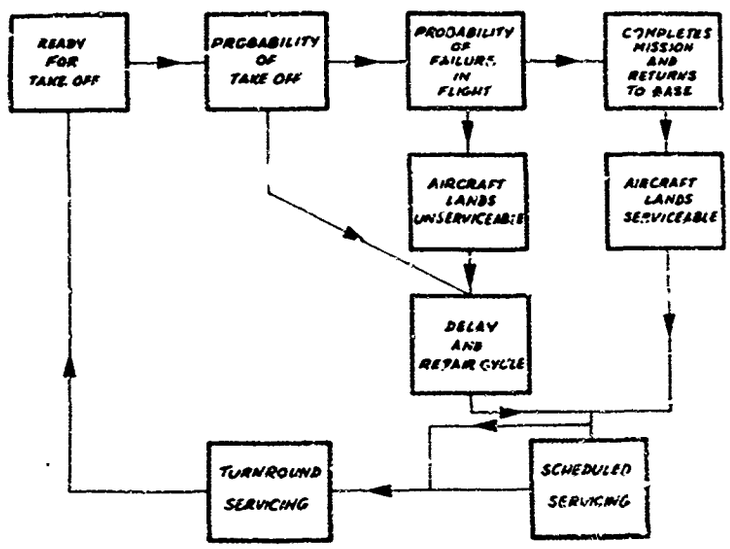


Fig.14 Typical aircraft cycle

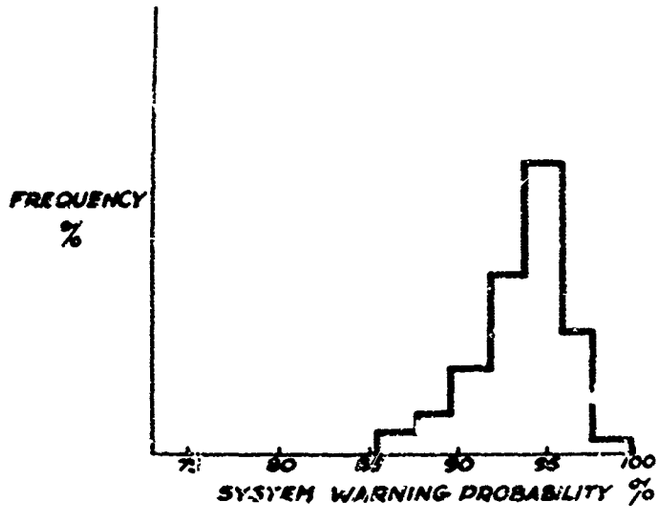


Fig. 15 Probability of being on station - 4 aircraft available

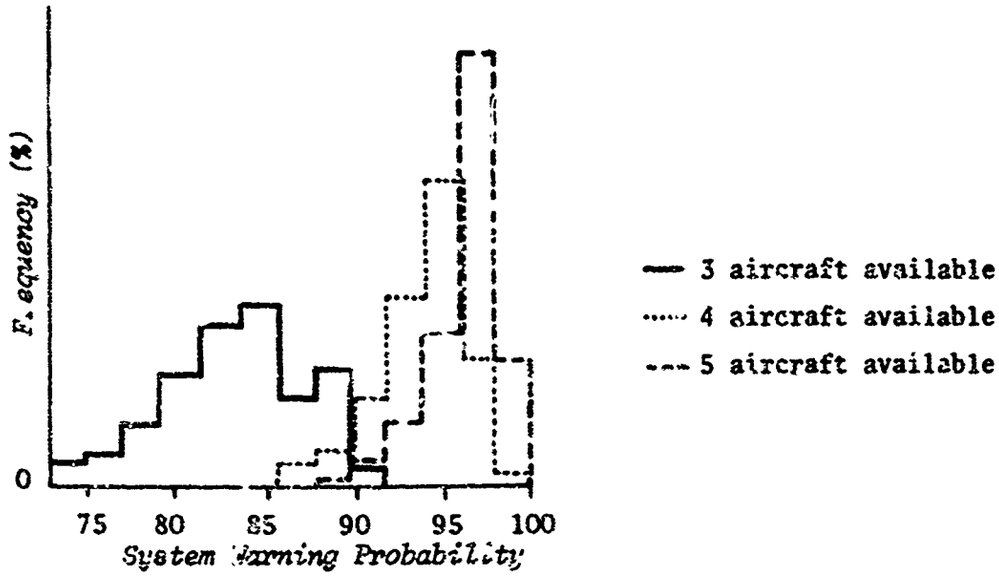


Fig. 16 Distribution of system warning probability with different numbers of aircraft

DISCUSSION ON PAPERS

DISCUSSION ON PAPERS

Paper No. 1. Radar Sea Return in High Wind/Wave Conditions
by K. W. Guinard and J. R. Ransone Jr.

MR. J. FREEDMAN: How well does the model proposed predict performance at low angles of incidence? Does the independence of the back scatter of wavelength still obtain at low angles of incidence?

MR. N. GUINARD: The model has not been verified for shallow grazing angles over the sea. Some laboratory experiments have been performed on sinusoidal and slightly rough planar surfaces which indicate that the model is valid for these scattering surfaces. We have very recently conducted shallow angle studies over the open ocean which should shed more light on the validity of the model at low angles. The wavelength dependence at low angles has to my knowledge never been measured. We hope to increase the power output of the 4FR system. If this is done we may be able to determine it over our range of wavelengths.

MR. P. BRADSELL: Does the clutter model lead to an asymptotic form of grazing incidence? This is important in assessing performance of sea-search radars.

MR. N. GUINARD: If the model is weighted with the probability distribution of "A new model for sea clutter" referenced in our paper, such an asymptotic form may be found. The use of a maximum slope condition should of course be used.

Paper No. 2 An Experimental Study of Some Clutter Characteristics
by M. P. Warden.

MR. G. HANSON: The purpose of experimental research is to provide data from which target signal to clutter ratio can be improved, I presume. Can the interesting ground data results described be extrapolated to apply to airborne radar performance at distances for which the incident angles are very small?

MR. M. WARDEN: The unprocessed clutter results cannot be readily extrapolated to give the information you require of the non-stationarity of the time series of signals. It would be interesting to see the results of subtracting the time varying mean from the signal, as discussed in the paper, compared to measurements made with an airborne radar.

Paper No. 4 Some Measurements of Radar Angels (False Radar Targets)
by W. Fishbein, E. Frost and W. E. Vander Meer.

DR. T. KESTER: I would like to comment on your remarks concerning Fig. 13 of your paper showing two doppler peaks for wind blown rain during a period of 20 seconds. We have tried to show in our paper (paper 3 of this conference) that the line shape of the doppler spectrum of rain clouds need not be gaussian and depends on the wind profile prevailing at the time as well as on the radar parameters, the height of the cloud and its distance from the radar. In my opinion it is conceivable that wind profiles are stationary for a period of 20 seconds.

Paper No. 6 More Measurements of the Effects of Frequency Agility on Aircraft Radar Returns
by W. J. Whitlock, A. M. Shepherd and A. L. C. Quigley.

DR. A. RIHACZEK: Dr. Whitlock stated that his measurements do not seem to agree with theory, which predicts that the frequency jump necessary for decorrelation depends on target size. Actually, a more precise statement would be that the frequency jump depends on the range spread for the essential scattering centers of the target. I am sure that re-examination of his results from this point of view would show agreement with theory.

DR. W. WHITLOCK: Some published theories assume a wide distribution of scatterers over the aircraft, but this is not realistic as Dr. Rihaczek implies, for in general the dominant scatterers will be grouped within the physical extremes of the aircraft. Nevertheless it was anticipated that aircraft of very different sizes should show differing behaviour to frequency agility but the experimental study failed to resolve this difference.

MR. P. BRADSELL: Does your work give a value in decibels for the improvement in detectability of an aircraft due to frequency agility?

Were any of the observations made in rain clutter? What was the effect?

DR. W. WHITLOCK: a) Insufficient data has been accumulated to make a quantative statement on improvement in detectability.

b) As the inter-pulse period of the radar was long compared with the natural decorrelation time of rain clutter it was not appropriate to carry out frequency agility measurements. However other workers have demonstrated that the reciprocal pulse length rule applies to rain clutter.

Paper No. 11 Integrated Subsystems on Composite Substrates
by M. Lemke, P. Host and H. J. Schmitt.

DR. D. PAGE: How do the mechanical properties of these substrates compare with conventional substrates? Are they more brittle?

DR. H. SCHMITT: Ferrites, and also composite ferrite substrates are more brittle than Al_2O_3 substrates. Mechanical shock tests have not yet been made, however. Mechanical processing like polishing, drilling etc., is much easier than for Al_2O_3.

MR. E. ALPANDARI: Avez-vous rencontré des difficultés pour la reproduction en grand nombre de ce type de déphaseurs à ferrite?

DR. H. SCHMITT: From experience with similar ferrite devices we would expect excellent reproducibility of material properties. Only small numbers of integrated digital phase shifters have been made so far for laboratory purposes.

MR. J. FREEDMAN: What is your estimate of the cost of such phase shifters when reproduced in quantity?

DR. H. SCHMITT: The price will strongly depend on quantity. Because of easier machining and processing it should be competitive with all other ferrite phasing schemes.

Paper No. 14 Pattern Compression by Space-Time Binary Coding of an Array Antenna
by S. Drabowitch and C. Aubry.

MR. R. GRAHAM: With a conventional phased array which uses single bit (0° or 180°) phase shifters, if the residual phase errors are randomly spread between $\pm 90^\circ$, there is a loss of gain of 3.9 dB.

With the binary coding techniques the l[illegible] gain is only 3 dB.

Would Dr. Drabowitch care to comment?

DR. J. DRABOWITCH: Dans un reseau classique à 1 bit avec dispersion erratique des lobes parasites, la perte de gain est 3.9 dB mais la perte aller-retour est doublée: elle atteint 4.8 dB.

Dans un reseau à codage periodique, la perte en gain est environ 3 dB mais l'existence de deux lobes principaux permet de doubler le temps de mesure sur un écho, ce qui fait que la perte aller-retour n'est que de 3 dB environ.

Paper No. 15 Flood-lighting with Nyquist Rate Scanning
by A. K. Edgar and T. L. Jones.

MR. R. VOLER: I should first of all like to take the opportunity to congratulate yourself and Mr. G. F. Clarke on inventing such a novel and intriguing radar principle. I am sure that it will find many applications.

I wonder whether, if it does not involve going into too much detail, you could say a few words about the signal processing you would employ when the system is made to incorporate pulse compression.

MR. A. EDGAR: The beam scanning process may be considered as the Fourier transformer of the modulated signals from the array. With pulse compression the scanning modulation must of course be matched to the signal bandwidth and the subsequent addition in a dispersive filter designed for the final output bandwidth will provide a signal time code giving range and bearing. I would like to emphasize that we have not made any experimental equipment for handling this type of signal and also to note that the final signal bandwidth will be N times greater than the input waveform because of the high speed scan - as in the system I described in the paper.

Paper No. 17 The Linear Array for Beam Steering
by R. Reitzig.

MR. N. GUINARD: Is the speaker aware of the extensive work being done by Drs. Marston and Brown at USNRL. They have been working for a number of years in the design of feeding networks for cylindrical array antennas with minimum mutual coupling.

DR. R. REITZIG: I am afraid, no, I don't know the work the gentlemen you mentioned have been performing on minimising the element coupling through the feeding network. However, I should very much appreciate if you could trigger off a mutual exchange of more detailed information on this matter.

Paper No. 18 Microwave S.S. Power Sources
by F. Sterzer.

MR. P BRADSELL: What is the capability of these transistor and transferred electron devices in the pulsed regime?

MR. F. STERZER: The maximum power output of the pulsed transistor amplifiers is in some cases 2 or 3 times as high as that of CW amplifiers. The difference between the maximum power output of pulsed and CW transferred electron devices can be as high as 1 or 2 orders of magnitude. At C-band frequencies, for example, the highest CW power output from transferred electron oscillators are currently approximately 1 watt, while the highest pulsed outputs (at a few per cent duty cycle) are of the order of 40 watts. The highest power outputs obtained to date with pulsed transferred electron amplifiers are several watts at C-band frequencies with an efficiency of 6%.

Paper No. 20 Lightweight Phased Array Systems for Airborne and Ground Based Applications
by M. Simpson.

MR. R. WILES: I see from the introduction to your paper that you were aiming at a beam pointing accuracy of 1/30 of a beamwidth. Could you say whether you had to take any special precautions to allow for the differential heating of the array and over what temperature range you achieved the required result?

And, secondly, could you kindly say why an airborne weather radar is a candidate for electronic beam steering?

MR. M. SIMPSON: The major problem with temperature is provided by the variation in phase vs temperature. Differential phase compensation is provided very effectively by the use of the flux transfer drive principle in the ferrite phase shifters. Insertion phase control is best obtained by maintaining the antenna phase shifter array at a uniform temperature by the use of circulating air or liquid. Satisfactory antenna performance can be obtained over a temperature range of $-40^{\circ}F$ to $+158^{\circ}F$.

The use of the particular antenna with a weather radar was due to the desire for a rapid airborne evaluation with an operational radar, namely, the AN/APN-158 System. The antenna was intended to be expanded at a later date to accomplish more functions such as Terrain Clearance and Terrain Avoidance.

MR. A. LEENHOUTS: a) The application of space-feed in the linear arrays which have been described introduces phase-difference between radiating elements due to the unequal pathlengths for the electromagnetic waves. Is compensation for this undesired phase-shift obtained by hardware or by including numerical corrections in the computed phase-settings?

b) Could the application of sub-arrays be considered in the described linear antenna systems?

MR. M. SIMPSON: a) Compensation for the cylindrical phase front is provided by prestored phase compensation as a function of frequency.

b) Sub-arrays could be employed where there is some advantage to be gained. Examples of this are very long arrays which use sub-arrays with time delay control in order to increase instantaneous bandwidth. Another example is the use of sub-arrays fed by separate power amplifiers in order to increase overall system power and reliability.

MR. E. BOLDBOHM: 1. Could you elaborate somewhat on the way tracking was organised?

2. How did you approach the problem of mutual coupling? Were the elements measured in the environments of similar elements or did you calculate the effect?

MR. M. SIMPSON: 1. A dual primary feed is inserted in the apex of the parallel plate space feed. The dual feed is fed by a hybrid to provide in-phase sum pattern and out-of-phase difference pattern illumination of the rear face of the aperture.

2. The element patterns were measured in an array environment as well as in a waveguide simulator to evaluate mutual coupling effect.

Paper No. 22 Radar de Poursuite a Grand Pouvoir Separateur en Distance
par E. G. Forestier et M. Chevalier.

DR. A. RIHACZEK: The speaker stated that the processing principle under discussion applies for tracking radar. It applies more generally to search radar, but it is necessary to provide parallel channels to the extent that the time-bandwidth product is reduced by correlation (as with any correlator). Such a system is entirely practical.

MR. E. FORESTIER: I agree with your statement but from the economical point of view I think that the processing principle exposed applies pretty well for tracking radar where only one channel is necessary in correlation or more generally three channels in case of monopulse application.

Paper No. 23 Low Distortion Dispersive Networks for Wideband Signal Processing Systems
by F. G. Herring, P. M. Krencik and A. J. Axel.

DR. A. RIHACZEK: Would the speaker please comment on the problem of maintaining the sidelobes at -40 dB under operational conditions, and on the possibility of further reducing the sidelobe level for this particular approach.

MR. G. HERRING: It is difficult to obtain a 40 dB sidelobe level and a simple analogy is that of retaining a 40 dB carrier balance in a balanced modulator. Whether a 40 dB carrier level can be maintained under operational conditions depends on the type of radar. For tactical or field radars which must operate over a wide range of environmental conditions with a minimum of maintenance, it would be, in most cases, impossible. However, for large, ground-based, instrumentation type radars which operate in a controlled environment and have maintenance on a daily or weekly basis, a 40 dB operation level is entirely practical. This is particularly true if the receiver processor contains an adjustable transversal equalizer which can be used to compensate for slow changes in the electrical characteristics due to time (aging), temperature, or variations in operating conditions. To reduce the sidelobes further than 40 dB would result in a substantial increase in equipment cost and complexity.

MR. P. BRADSELL: What are the important parasitic quantities in the type 3 network?

MR. G. HERRING: Two important parasitic quantities are: (1) the discontinuity which exists at the shorted end of the line and (2) the discontinuity which exists at the input and output due to the transition from a (constant-impedance) coupled line to a (constant-impedance) single unbalanced line to ground. The first discontinuity can be partially compensated by the addition of a capacitive tab extending past the short. The second can be partially corrected by tapering the transition from a single to a coupled line.

Paper No. 25 Switched Capacitor Storage Arrays for A.M.T.I. and Bandwidth Compression
by D. S. Harris.

MR. J. FREEDMAN: 1. What was the tolerance required in selection of the capacitors?

2. How much additional weight and volume would be required to introduce more conventional bandpass filters following the gating rather than the capacitor subtraction method used?

MR. D. HARRIS: 1. Cancellation ratio is not affected by capacity as the subtraction of returns is carried out on a single capacitor within each slot. However slot-to-slot ripple is affected by capacity tolerance and a tolerance of $\pm$ 2% was selected. Generally slot-to-slot ripple is better than 30 dB below signal level stored in the capacitors.

2. This question is a little difficult to give a factual answer to as I have very little experience in the packaging of range gate filters. However the capacitor storage technique packs twelve stores in a volume of 0.75 cubic inch. Modern techniques and the use of active filters may well give a comparative system in weight and volume.

MR. R. VOLES: The module that you have described is most versatile and is likely to find many applications. I wonder, therefore, whether you propose to translate the form of construction from Cordwood to MSI or LSI?

MR. D. HARRIS: The Cordwood technique was used to enable the experimental system to be built as simply as possible and to enable component tolerances and characteristics to be selected as required. Future production may well use thick film technology rather than MSI or LSI where cost could be fairly high.

Paper No. 26 Methode de Traitement Digital de Signaux Monopulse en Presence de Clutter
par H. G. P. Forestier.

DR. T. KESTER: Vous venez de decrire un système de suppression de l'echo fixe basé sur une triangulation. Comment ce système se comporte-t-il en presence de clutter ayant un

spectre avec une raie Doppler d'une largeur non nul? Je pense au spectre des échos météorologiques de pluie par exemple.

MR. H. [illegible]: Le système se comporte exactement comme un MTI classique. Une compensation de vitesse moyenne de pluie est aisément réalisable et peut être introduite en décalant la fréquence de MTI.

Paper No. 29 Système de Traitement associé à un Radar Cohérent pour la Distance et la Vitesse Radiale des Cibles
par M. J. Sirven.

MR. J. COLIN: 1. Fonctionnement en présence de deux cibles mobiles.

2. Rapport signal sur bruit minimum pour lever l'ambiguité de vitesse.

MR. J. SIRVEN. 1. Le système fait une erreur dans l'estimation de la vitesse radiale si deux cibles situées à la même distance ont des fréquences Doppler qui diffèrent d'un multiple de la fréquence de répétition de probabilité de cet événement est très faible.

2. L'erreur sur la mesure de phase dépend du rapport signal sur bruit et du nombre total d'ambiguités qu'il est possible de résoudre, donc du rapport T/τ Une probabilité de mesure exacte suffisante, peut être obtenue, avec un rapport signal sur bruit de 15 dB après integration et pour un rapport T/τ de 20, correspondant à 20 ambituités.

Paper No. 33 Operational Employment and Detection Performance of Advanced Computer Controlled Radar Systems (Abstract only)
by J. A. Bijvoet.

MR. C. MUEHE: What differences would it make in your analysis of search and track using phased arrays if multiple beam antennas were available which could provide several beams in space at the same time?

MR. J. BIJVOET: The use of multiple beam antennas would proportionately reduce the search time. However this does not change the basic argument since such a system can be considered as a number of parallel systems requiring a multiplication of transmitters and receivers and a more complex antenna.

Paper No. 34 Digital Plot Extraction of Primary Radar
by N. H. A. Smith and I. H. Jackson.

DR. E. HANLE: From what reference did you get your Figure 2? I think that the threshold must be proportional to the mean value of the clutter amplitude and not logarithmic as shown in your Figure. Do you have any practice on this subject?

MR. N. SMITH: The questioner is perhaps mistaken as to the implications of Figure 2. It refers to the second detection threshold, rather than to an initial quantizing level, and is obtained as follows:- assume the clutter hit density over the chosen area to be randomly distributed. With the window size and a desired probability of false plot rate as constants TL is then given via the Binomial Probability Distribution function given below where p = clutter sum achieved divided by the maximum possible clutter sum. Although this relationship assumes random clutter hit distribution extensive field trials have shown that the calculated functions preserve the chosen PFA to better than an order of magnitude, irrespective of actual

clutter distribution.

$$P_{FA} = \sum_{s=T_L}^{W} {}^{W}C_s \, p^s (1-p)^{W-s}$$

Paper No. 30 Digital Radar Plot Extractor - a System Module of an Automatic Air Traffic Control System
by H. Ebert and U. Siegenthaler.

MR. P. BRADSELL: 1. Present-day systems have circuits giving CFAR at the first threshold for, say, rain clutter. They also have adaptive control of the second threshold to give constant false alarm rate in plot detection. In your opinion is there an optimum balance between the effectiveness of these actions?

2. With regard to the Scan-scan MTI,

a) is there a threshold of velocity?
b) what type of clutter is it intended to remove?

3. Could you make a comparison between the method of associating primary and secondary radar plots and associating primary and secondary track data in the computer?

DR. H. EBERT: 1. In my opinion, CFAR circuits at the 1st threshold and present day systems of control of the 2nd threshold (such as the clutter eliminator described in the abstracts) cannot be compared directly. A CFAR is principally suited to reduce false alarm probability without a remarkable loss of sensitivity (for instance, the system described in the abstracts has a measured loss of sensitivity of about 0.8 db). But on the other hand CFAR is not suited against clutter of high correlation in azimuth direction (for instance angels) because it does not use the information of several subsequent radar periods. The clutter eliminator, on the other hand, uses information of several radar periods and is therefore able to cope with correlated clutter, but by setting the 2nd threshold to higher values, sensitivity is lost. The best way would be to have a system combined with the 1st threshold, which is able to cope with correlated noise without remarkable loss of sensitivity. One solution (as proposed by D7) could be to digitize pulse amplitude and to analyze the neighbourhood of every "hit in process" for correlation - of course a very expensive solution.

2. The Scan MTI is an attempt to cope with angle echoes. A fixed area (as described in the paper, and whose edge length can be adjusted) is defined and a target is declared as a fixed target, if it does not escape this area during the next antenna revolution. This may be in some way a threshold of velocity. Praxis must prove whether the intended effect is reached.

3. In the plot extractor the association of PR and SSR is done merely on the basis of target location. (If a PR- and an SSR- target are in adjacent range increments and overlap in azimuth direction). The advantage is in reducing the data flow to the tracking computer in the ATC-center via the telephone lines. (It is expected that most of the real targets can be correlated by this method under normal radar conditions.) As for the tracking process, tracking of "combined" (i.e. correlated) targets can be done on an SSR tracking basis, which needs a less sophisticated program than PR tracking normally does. Therefore some amount of computer work can be spared.

MR. N. SMITH: The functions carried out within the 'plot information process computer' are in other systems carried out either by special hardware at the radar site or in a computer at the centre. Could Mr. Ebert please give the reasons, which in his view lead to the use of a computer at the radar site?

DR. H. EBERT: The main reasons for a computer for plot information processing rather than a special hardware are:

1. Special wishes of the customer for plot information processing: connection to 3 different ATC-centers which get different data, Scan MTI, sophisticated on-line test system, map control of plot detector, etc.

2. Flexibility of system with respect to subsequent wishes and growing experience.

3. Lower development costs for software instead of hardware for an equipment of which only a small number will be built.

A computer in the ATC center cannot be used because the data flow from the plot

detector group to the computer is random, i.e. on-line with incoming radar information (a small delay within the plot detector neglected).

Paper No. 39 Rudar - an Experimental Noise-Radar System
by J. A. Smit.

MR. H. VOM STEIN: Would you be so kind to give some detailed information concerning your correlator (analog, digital, upper frequency limit)?

MR. J. SMIT: Information about the correlator as used in our experiment is given in the appendix and in fig. 4, 6, 8 and 10 of the paper. As you will see it is an analogue envelope correlator. The maximum bandwidth of the correlator is 40 MHz. This bandwidth is centered around the local-oscillator frequency f_c (fig. 4). In our case this frequency could be shifted $\pm$ 400 MHz. For other r.f. circuitry this will be different.

Paper No. 40. Obstacle Warning Radar for Helicopters by Mr. G. Collot.

MR. C. M. STEWART: 1) What was used to measure the velocity vector of the helicopter in elevation plane.

2) What was the operational experience with this radar and particularly the pilot's feelings about flying for long periods on a display which needs considerable interpretation.

MR. G. COLLOT: 1) A doppler Radar.

2) The longer period was 1½ hours with window curtains down, and with a sortie previously prepared - the pilot said that the orders were easy to follow and the course wasn't too tiring.

Paper No. 42 A Cost Effectiveness Study of the Airborne Early Warning Aircraft as an Integrated Part of Air Defence Systems
by T. H. Kerr.

MR. J. SIRVEN: Une portée de l'ordre de 150 NM semble être nécessaire pour un système AEW. Cette portée peut-elle être obtenue si le terain est accidenté et si l'ennemi utilise le relief éviter d'être detecté.

MR. T. KERR: We have investigated this possibility in a number of areas of the world. In only a small proportion of these areas, is it possible for the low flying aircraft to take advantage of the cover provided by terrain. Generally the penetration achieved is small and as far as can be seen it does not unduly degrade the warning provided by the early warning system.