

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
RADIATION LABORATORY SERIES

Board of Editors

LOUIS N. RIDENOUR, *Editor-in-Chief*

GEORGE B. COLLINS, *Deputy Editor-in-Chief*

BRITTON CHANCE, S. A. GOUDSMIT, R. G. HERB, HUBERT M. JAMES, JULIAN K. KNIPP,
JAMES L. LAWSON, LEON B. LINFORD, CAROL G. MONTGOMERY, C. NEWTON, ALBERT
M. STONE, LOUIS A. TURNER, GEORGE E. VALLEY, JR., HERBERT H. WHEATON

1. RADAR SYSTEM ENGINEERING—*Ridenour*
2. RADAR AIDS TO NAVIGATION—*Hall*
3. RADAR BEACONS—*Roberts*
4. LORAN—*Pierce, McKenzie, and Woodward*
5. PULSE GENERATORS—*Glasoe and Lebacqz*
6. MICROWAVE MAGNETRONS—*Collins*
7. KLYSTRONS AND MICROWAVE TRIODES—*Hamilton, Knipp, and Kuper*
8. PRINCIPLES OF MICROWAVE CIRCUITS—*Montgomery, Dicke, and Purcell*
9. MICROWAVE TRANSMISSION CIRCUITS—*Ragan*
10. WAVEGUIDE HANDBOOK—*Marcwitz*
11. TECHNIQUE OF MICROWAVE MEASUREMENTS—*Montgomery*
12. MICROWAVE ANTENNA THEORY AND DESIGN—*Silver*
13. PROPAGATION OF SHORT RADIO WAVES—*Kerr*
14. MICROWAVE DUPLEXERS—*Smullin and Montgomery*
15. CRYSTAL RECTIFIERS—*Torrey and Whitmer*
16. MICROWAVE MIXERS—*Pound*
17. COMPONENTS HANDBOOK—*Blackburn*
18. VACUUM TUBE AMPLIFIERS—*Valley and Wallman*
19. WAVEFORMS—*Chance, Hughes, MacNichol, Sayre, and Williams*
20. ELECTRONIC TIME MEASUREMENTS—*Chance, Hulsizer, MacNichol, and Williams*
21. ELECTRONIC INSTRUMENTS—*Greenwood, Holdam, and MacRae*
22. CATHODE RAY TUBE DISPLAYS—*Soller, Starr, and Valley*
23. MICROWAVE RECEIVERS—*Van Voorhis*
24. THRESHOLD SIGNALS—*Lawson and Uhlenbeck*
25. THEORY OF SERVOMECHANISMS—*James, Nichols, and Phillips*
26. RADAR SCANNERS AND RADOMES—*Cady, Karelitz, and Turner*
27. COMPUTING MECHANISMS AND LINKAGES—*Svoboda*
28. INDEX—*Henney*

RADAR SYSTEM ENGINEERING

SCIENCE LIBR.

Edited by

LOUIS N. RIDENOUR

PROFESSOR OF PHYSICS
UNIVERSITY OF PENNSYLVANIA



FIRST EDITION



NEW YORK AND LONDON
MCGRAW-HILL BOOK COMPANY, INC.

1947

7 11 22 7 3

1947

v. 1

2/6

RADAR SYSTEM ENGINEERING

COPYRIGHT, 1947, BY THE
MCGRAW-HILL BOOK COMPANY, INC.
PRINTED IN THE UNITED STATES OF AMERICA

*All rights reserved. This book, or
parts thereof, may not be reproduced
in any form without permission of
the publishers.*

THE MAPLE PRESS COMPANY, YORK, PA.

RADAR SYSTEM ENGINEERING

EDITORIAL STAFF

LOUIS N. RIDENOUR

AVIS M. CLARKE

CONTRIBUTING AUTHORS

L. Y. BEERS	F. J. MEHRINGER
B. V. BOWDEN	R. D. O'NEAL
W. M. CADY	C. F. J. OVERHAGE
R. E. CLAPP	E. C. POLLARD
G. B. COLLINS	E. M. PURCELL
A. G. EMSLIE	L. N. RIDENOUR
W. W. HANSEN	C. V. ROBINSON
L. J. HAWORTH	A. J. F. SIEGERT
R. G. HERB	R. L. SINSHEIMER
M. M. HUBBARD	D. C. SOPER
P. C. JACOBS	G. F. TAPE
M. H. JOHNSON	L. A. TURNER
W. H. JORDAN	M. G. WHITE
J. V. LEBACQZ	A. E. WHITFORD
F. B. LINCOLN	J. M. WOLF
R. A. MCCONNELL	C. L. ZIMMERMAN

Foreword

THE tremendous research and development effort that went into the development of radar and related techniques during World War II resulted not only in hundreds of radar sets for military (and some for possible peacetime) use but also in a great body of information and new techniques in the electronics and high-frequency fields. Because this basic material may be of great value to science and engineering, it seemed most important to publish it as soon as security permitted.

The Radiation Laboratory of MIT, which operated under the supervision of the National Defense Research Committee, undertook the great task of preparing these volumes. The work described herein, however, is the collective result of work done at many laboratories, Army, Navy, university, and industrial, both in this country and in England, Canada, and other Dominions.

The Radiation Laboratory, once its proposals were approved and finances provided by the Office of Scientific Research and Development, chose Louis N. Ridenour as Editor-in-Chief to lead and direct the entire project. An editorial staff was then selected of those best qualified for this type of task. Finally the authors for the various volumes or chapters or sections were chosen from among those experts who were intimately familiar with the various fields, and who were able and willing to write the summaries of them. This entire staff agreed to remain at work at MIT for six months or more after the work of the Radiation Laboratory was complete. These volumes stand as a monument to this group.

These volumes serve as a memorial to the unnamed hundreds and thousands of other scientists, engineers, and others who actually carried on the research, development, and engineering work the results of which are herein described. There were so many involved in this work and they worked so closely together even though often in widely separated laboratories that it is impossible to name or even to know those who contributed to a particular idea or development. Only certain ones who wrote reports or articles have even been mentioned. But to all those who contributed in any way to this great cooperative development enterprise, both in this country and in England, these volumes are dedicated.

L. A. DUBRIDGE.

Preface

THE earliest plans for the Radiation Laboratory Series, made in the fall of 1944, envisaged only books concerned with the basic microwave and electronic theory and techniques that had been so thoroughly developed during the wartime work on radar. These plans were laid aside for a time when it became clear in this country that several months of fighting remained in the European war.

When work on the Series was resumed in the early summer of 1945, the books planned, as before, dealt with basic matters and with techniques. Every effort was made to point out the general applicability of the work reported and to avoid special emphasis on its application to radar, since radar itself was thought to have only a limited importance.

The end of the Pacific war made it possible to put more effort on the job of preparing the Series than had been available earlier. The books on theory and techniques having been planned as comprehensively as appeared to be worth while, the work was extended by the addition of five books concerned with radar and allied systems.

Of those five books, this is the only one that deals with radar itself. One book takes up the use of radar in navigation, one concerns the design of radar scanners and radomes, one treats the design and construction of beacons, and one describes hyperbolic navigational systems—in particular Loran.

This book is intended to serve as a general treatise and reference book on the design of radar systems. No apology seems to be needed for the fact that it deals primarily—though by no means altogether—with microwave pulse radar. Thousands of times as much work has gone into pulse radar as into any other kind, and the overwhelming majority of this work has been concerned with microwave pulse radar. The superiority of microwaves for almost all radar purposes is now clear.

The first eight chapters of this book are intended to provide an introduction to the field of radar and a general approach to the problems of system design. Chapters 9 through 14 take up the leading design considerations for the various important components that make up a radar set. These chapters are so thorough in their treatment that Chap. 15, which gives two fairly detailed examples of actual system design, can be quite brief. Chapters 16 and 17 take up two new and important ancillary

techniques that are not dealt with fully elsewhere in the Series: moving-target indication and the transmission of radar displays to a remote indicator by radio means.

For fuller information than can be found in this book on any detailed point of design, the reader is referred to one of the other books of the Series. In a sense, this book specializes to radar the techniques reported more fully elsewhere in the Series.

Radar is a very simple subject, and no special mathematical, physical, or engineering background is needed to read and understand this book.

Because the book covers the entire field of effort of the Radiation Laboratory and the other wartime radar establishments, its contributing authors are more numerous than those listed for most other volumes of this series. I am especially grateful to L. J. Haworth and to E. M. Purcell, whose contributions have been more extensive than those of other authors, and whose advice on editorial problems has often been extremely helpful. In addition to the authors already listed, whose names appear in the book in connection with the material they have written, I wish to thank the following men for their work in providing essential background material that did not eventually find its way into the book: R. M. Alexander, A. H. Brown, J. F. Carlson, M. A. Chaffee, L. M. Hollingsworth, E. L. Hudspeth, R. C. Spencer, and I. G. Swope. Changing plans for the book also reduced the acknowledged contribution of E. C. Pollard far below the very considerable quantity of material he prepared.

I owe an apology to all the authors for the liberty I have often taken in altering their original text to fit the final framework of the book and my own ideas of style. Because most authors left the Laboratory immediately on finishing their writing, and much of the editorial work had to be deferred until the book was substantially complete, it has not always been possible to adjust with the authors the alterations in their manuscripts that have seemed desirable to me.

The general acknowledgments I owe as Editor-in-Chief of the Series are set forth in the Series Index. In connection with the preparation of this book, however, it is a pleasure to thank Dr. B. V. Bowden, of the British Air Commission, not only for his assistance as an author but also for his general comments on the book as a whole. I am grateful to Lois Capen for her work in following the preparation of illustrations, and to Phyllis Brown for general secretarial assistance.

LOUIS N. RIDENOUR.

CAMBRIDGE, MASS.

June, 1946.

Contents

FOREWORD BY L. A. DUBRIDGE	vii
PREFACE	ix
CHAP. 1. INTRODUCTION	1
1-1 What Radar Does	1
1-2 How Radar Works	3
1-3 Components of a Radar System	6
1-4 The Performance of Radar	8
1-5 Radar Systems	12
1-6 The Early History of Radar	13
1-7 Wartime Radar Development in the U.S.	15
CHAP. 2. THE RADAR EQUATION	18
THE RADAR EQUATION FOR FREE-SPACE PROPAGATION	18
2-1 The Meaning of Free-space Propagation	18
2-2 Antenna Gain and Receiving Cross Section	19
2-3 Scattering Cross Section of the Target	21
2-4 The Radar Equation	21
2-5 Beams of Special Shapes	22
2-6 The Beacon Equation	27
THE MINIMUM DETECTABLE SIGNAL	28
2-7 Noise	28
2-8 Receivers, Ideal and Real	30
2-9 Receiver Bandwidth and Pulse Energy	33
2-10 The Statistical Problem	35
2-11 Effect of Storage on Radar Performance	41
MICROWAVE PROPAGATION	47
2-12 Propagation over a Reflecting Surface	47
2-13 The Round Earth	53
2-14 Superrefraction	55
2-15 Attenuation of Microwaves in the Atmosphere	58
CHAP. 3. PROPERTIES OF RADAR TARGETS	63
SIMPLE TARGETS	63
3-1 Cross Section Expressed in Terms of the Field Quantities	63
3-2 Rayleigh Scattering from a Small Sphere	63
3-3 Scattering of a Plane Wave by a Sphere	64
3-4 Approximations for Large Metal Targets	65
3-5 The Corner Reflector	67
3-6 Target Shaping to Diminish Cross Section	68
3-7 Use of Absorbent Materials	69

COMPLICATED TARGETS	73
3-8 Return from Two Isotropic Targets	73
3-9 Actual Complex Targets	75
3-10 Compound Targets Extended through Space	81
3-11 Extended Surface Targets	85
GROUND PAINTING BY AIRBORNE RADAR	88
3-12 Specular and Diffuse Reflection	89
3-13 Sea Return and Ground Return	92
3-14 Mountain Relief	96
3-15 Structures	99
3-16 Cities	101
3-17 Navigation	108
CHAP. 4. LIMITATIONS OF PULSE RADAR	116
4-1 Range, Pulse-repetition Frequency, and Speed of Scan	116
4-2 Bandwidth, Power, and Information-rate	121
4-3 Pulse Radar and C-w Radar	123
4-4 Clutter	124
CHAP. 5. C-W RADAR SYSTEMS	127
5-1 General Considerations	127
5-2 Transmitted Spectra	129
5-3 Effect of Target	130
5-4 Class of Systems Considered	131
5-5 Utility of C-w Systems	132
SPECIFIC SYSTEMS	132
5-6 Simple Doppler System	132
5-7 Range-measuring Doppler System	139
5-8 F-m Range-measuring System	143
5-9 Multiple Target F-m Range Measurement	147
5-10 Alternative F-m Ranging System	149
5-11 Pulse-modulated Doppler System	150
5-12 Summary	157
CHAP. 6. THE GATHERING AND PRESENTATION OF RADAR DATA	160
6-1 Influence of Operational Requirements	160
TYPES OF RADAR INDICATORS	161
6-2 Definitions	161
6-3 Summary of Indicator Types	163
6-4 One-dimensional Deflection-modulated Displays	164
6-5 Representation of the Horizontal Plane	167
6-6 Plane Displays Involving Elevation	171
6-7 Three-dimensional Displays	174
6-8 Error Indicators	175

EXAMPLES OF THE MAJOR OPERATIONAL REQUIREMENTS	175
6-9 Early Aircraft Warning Radar	175
6-10 PPI Radar for Search, Control, and Pilotage	182
6-11 Height-finding Involving Ground Reflection	184
6-12 Height-finding with a Free-space Beam	187
6-13 Homing	196
6-14 Precision Tracking of a Single Target	203
6-15 Precision Tracking during Rapid Scan	210
CHAP. 7. THE EMPLOYMENT OF RADAR DATA	213
7-1 The Signal and Its Use.	213
EXTERNAL AIDS TO RADAR USE.	214
7-2 Aids to Individual Navigation.	214
7-3 Aids to Plotting and Control	218
7-4 The Relay of Radar Displays.	225
EXAMPLES OF RADAR ORGANIZATIONS	226
7-5 Radar in the RAF Fighter Command	226
7-6 The U.S. Tactical Air Commands	229
7-7 Close Control with SCR-584	238
7-8 Teleran.	240
CHAP. 8. RADAR BEACONS.	243
RADAR-BEACON SYSTEMS.	246
8-1 Types of Radar-beacon Systems.	246
8-2 Systems Planning	250
8-3 General Identification System—IFF.	251
8-4 Radar Interrogation vs. Special Interrogators.	252
8-5 Independence of Interrogation and Reply	254
8-6 Frequency Considerations	260
CODING	263
8-7 Interrogation Codes	263
8-8 Reply Codes	264
STATISTICAL CONSIDERATIONS.	265
8-9 Traffic Capacity.	265
8-10 Unsynchronized Replies	268
CHAP. 9. ANTENNAS, SCANNERS, AND STABILIZATION.	271
9-1 The Antenna Equation.	271
9-2 Round and Cut Paraboloid Antennas	272
9-3 Fan Beams.	274
9-4 Nonscanning Antennas.	277
9-5 Construction of Radar Antennas.	279
RADAR SCANNING PATTERNS	280
9-6 Simple Scans	281
9-7 Complex Scans	281

MECHANICAL SCANNERS	282
9-8 The Kinematics of Mechanical Scanners	282
9-9 The Weight of Mechanical Scanners.	283
9-10 R-f Transmission Lines.	283
9-11 Data Transmission.	284
9-12 Examples of Mechanical Scanners.	284
ELECTRICAL SCANNERS.	291
9-13 The AN/APQ-7 (Eagle) Scanner	291
9-14 Schwarzschild Antenna.	295
9-15 SCI Height Finder	298
9-16 Other Types of Electrical Scanners.	302
THE STABILIZATION PROBLEM.	304
9-17 Stabilization of the Beam.	305
9-18 Data Stabilization.	311
9-19 Installation of Airborne Scanners	312
9-20 Installation of Surface-based Scanners	313
9-21 Radomes.	314
9-22 Streamlining	315
9-23 Electrical Transmission.	316
9-24 Structural Design of Radomes.	316
9-25 Examples of Radomes	317
CHAP. 10. THE MAGNETRON AND THE PULSER	320
THE MAGNETRON.	320
10-1 Construction	321
10-2 The Resonant System	325
10-3 Electron Orbits and the Space Charge	330
10-4 Performance Charts and Rieke Diagrams	336
10-5 Magnetron Characteristics Affecting Over-all Systems Design	340
10-6 Magnetron Characteristics Affecting Pulsar Design	352
THE PULSER	355
10-7 Pulsar Circuits	356
10-8 Load Requirements	362
10-9 The Hard-tube Pulsar	367
10-10 Line-type Pulsars	373
10-11 Miscellaneous Components	383
CHAP. 11. R-F COMPONENTS.	391
11-1 The R-f Transmission Problem	391
11-2 Coaxial Lines.	393
11-3 Waveguide	398
11-4 Resonant Cavities.	405
11-5 Duplexing and TR Tubes.	407

MICROWAVE COMPONENTS OF THE RECEIVER	411
11-6 The Mixer Crystal	412
11-7 The Local Oscillator	414
11-8 The Mixer	416
11-9 Automatic Frequency Control.	418
MOUNTING THE R-F PARTS.	419
11-10 Reasons for an R-f Package.	419
11-11 Design Considerations for the R-f Head	421
11-12 Illustrative Examples of R-f Heads	425
CHAP. 12. THE RECEIVING SYSTEM—RADAR RECEIVERS	433
INTRODUCTION	433
12-1 The Role of the Receiving System.	433
12-2 A Typical Receiving System	435
THE RECEIVER	441
12-3 Special Problems in Radar Receivers.	441
12-4 I-f Amplifier Design	442
12-5 Second Detector.	449
12-6 Video Amplifiers.	450
12-7 Automatic Frequency Control.	453
12-8 Protection against Extraneous Radiation.	457
TYPICAL RECEIVERS.	460
12-9 A General-purpose Receiver.	462
12-10 Lightweight Airborne Receiver	464
12-11 An Extremely Wide-band Receiver	470
CHAP. 13. THE RECEIVING SYSTEM—INDICATORS	475
THE CATHODE-RAY TUBE.	475
13-1 Electrical Properties of Cathode-ray Tubes.	475
13-2 Cathode-ray Tube Screens	479
13-3 The Selection of the Cathode-ray Tube	483
COORDINATION WITH THE SCANNER	486
13-4 Angle-data Transmitters	486
13-5 Electromechanical Repeaters	490
BASIC ELECTRICAL CIRCUITS	492
13-6 Amplifiers	492
13-7 The Generation of Rectangular Waveforms.	496
13-8 The Generation of Sharp Pulses.	501
13-9 Electronic Switches	503
13-10 Sawtooth Generators.	510
INDICES	513
13-11 Angle Indices	514
13-12 Range and Height Indices; Synchronization.	518

DISPLAY SYNTHESIS	524
13-13 The Design of A-scopes	524
13-14 B-Scope Design	528
13-15 Plan-position Indicator	532
13-16 The "Resolved Time Base" Method of PPI Synthesis	534
13-17 Resolved-current PPI	538
13-18 The Method of Pre-time-base Resolution	544
13-19 The Range-height Indicator	545
SIGNAL DISCRIMINATION, RESOLUTION, AND CONTRAST	548
13-20 Resolution and Contrast	548
13-21 Special Receiving Techniques for Air-to-land Observation	550
CHAP. 14. PRIME POWER SUPPLIES FOR RADAR	555
AIRCRAFT SYSTEMS	555
14-1 Choice of Frequency	555
14-2 Wave Shape	557
14-3 Direct-driven Generators	557
14-4 Motor-alternator Sets	561
14-5 Voltage Regulators	563
14-6 Speed Regulators	571
14-7 Dynamotors	579
14-8 Vibrator Power Supplies	581
14-9 Summary of Recommendations for Aircraft Radar Power	582
GROUND AND SHIPBOARD SYSTEMS	583
14-10 Fixed Locations	583
14-11 Large Systems Where No Commercial Power Is Available	584
14-12 Smaller Mobile Units	585
14-13 Ultraportable Units	585
14-14 Ship Radar Systems	586
CHAP. 15. EXAMPLES OF RADAR SYSTEM DESIGN	588
15-1 Introduction	588
15-2 The Need for System Testing	590
DESIGN OF A HIGH-PERFORMANCE RADAR FOR AIR SURVEILLANCE AND CONTROL	592
15-3 Initial Planning and Objectives	592
15-4 The Range Equation	595
15-5 Choice of Pulse Length	596
15-6 Pulse Recurrence Frequency	598
15-7 Azimuth Scan Rate	599
15-8 Choice of Beam Shape	600
15-9 Choice of Wavelength	604
15-10 Components Design	606
15-11 Modifications and Additions	609

DESIGN OF A LIGHTWEIGHT AIRBORNE RADAR FOR NAVIGATION	611
15-12 Design Objectives and Limitations	611
15-13 General Design of the AN/APS-10	614
15-14 Detailed Design of the AN/APS-10	616
CHAP. 16. MOVING-TARGET INDICATION	626
INTRODUCTION	626
16-1 The Role of Moving-Target Indication	626
16-2 Basic Principles of MTI	626
16-3 A Practical MTI System	632
16-4 Alternative Methods for Obtaining Coherence	635
PERFORMANCE CRITERIA AND CHOICE OF SYSTEM CONSTANTS	638
16-5 Stability Requirements	638
16-6 Internal Clutter Fluctuations	642
16-7 Fluctuations Due to Scanning	644
16-8 Receiver Characteristics	646
16-9 Target Visibility	649
16-10 Choice of System Constants	653
MOVING-TARGET INDICATION ON A MOVING SYSTEM	655
16-11 Compensation for Velocity of System	655
16-12 The Noncoherent Method	656
16-13 Beating Due to Finite Pulse Packet	657
COMPONENT DESIGN	658
16-14 The Transmitter and Its Modulator	658
16-15 The Stable Local Oscillator	659
16-16 The Coherent Oscillator	662
16-17 The Receiver	665
16-18 The Supersonic Delay Line	667
16-19 Delay-line Signal Circuits	672
16-20 Delay-line Trigger Circuits	675
16-21 Special Test Equipment	677
CHAP. 17. RADAR RELAY	680
INTRODUCTION	680
17-1 The Uses of Radar Relay	680
17-2 The Elements of Radar Relay	681
METHODS OF SCANNER DATA TRANSMISSION	682
17-3 General Methods of Scanner Data Transmission	683
17-4 Methods of Combating Interference	685
17-5 The Method of Incremental Angle	689
17-6 The Phase-shift Method	695
17-7 Methods of Relaying Sine and Cosine	701
17-8 Pulse Method for Relaying Sine and Cosine	705
17-9 Comparison of Synchronization Methods	711

THE RADIO-FREQUENCY EQUIPMENT	713
17-10 Antennas, Frequencies, and the Radiation Path.	713
17-11 General Transmitter and Receiver Considerations	717
17-12 A 300-Mc/sec Amplitude-modulated Equipment.	719
17-13 A 100-Mc/sec Frequency-modulated Equipment.	721
17-14 Microwave System for Point-to-point Service.	723
RADAR RELAY SYSTEMS	726
17-15 A Ground-to-ground Relay System	726
17-16 Relay System for Airborne Radar	732
INDEX.	737

CHAPTER 1

INTRODUCTION

BY LOUIS N. RIDENOUR

1-1. What Radar Does.—Radar is an addition to man's sensory equipment which affords genuinely new facilities. It enables a certain class of objects to be "seen"—that is, detected and located—at distances far beyond those at which they could be distinguished by the unaided eye. This "seeing" is unimpaired by night, fog, cloud, smoke, and most other obstacles to ordinary vision. Radar further permits the measurement of the range of the objects it "sees" (this verb will hereafter be used without apologetic quotation marks) with a convenience and precision entirely unknown in the past. It can also measure the instantaneous speed of such an object toward or away from the observing station in a simple and natural way.

The superiority of radar to ordinary vision lies, then, in the greater distances at which seeing is possible with radar, in the ability of radar to work regardless of light condition and of obscuration of the object being seen, and in the unparalleled ease with which target range and its rate of change can be measured. In certain other respects radar is definitely inferior to the eye. The detailed definition of the picture it offers is very much poorer than that afforded by the eye. Even the most advanced radar equipment can only show the gross outlines of a large object, such as a ship; the eye can—if it can see the ship at all—pick out fine details such as the rails on the deck and the number or character of the flags at the masthead. Because of this grossness of radar vision, the objects that can usefully be seen by radar are not as numerous as the objects that can be distinguished by the eye. Radar is at its best in dealing with isolated targets located in a relatively featureless background, such as aircraft in the air, ships on the open sea, islands and coastlines, cities in a plain, and the like. Though modern high-definition radar does afford a fairly detailed presentation of such a complex target as a city viewed from the air (see, for example, Fig. 3-35), the radar picture of such a target is incomparably poorer in detail than a vertical photograph taken under favorable conditions would be.

One further property of radar is worth remarking: its freedom from difficulties of perspective. By suitable design of the equipment, the picture obtained from a radar set can be presented as a true plan view,

regardless of the obliquity of the angle from which the target is seen. This is shown clearly in Fig. 1-1, which compares with a chart of the tip of Cape Cod a radar picture and a direct photograph taken simultaneously

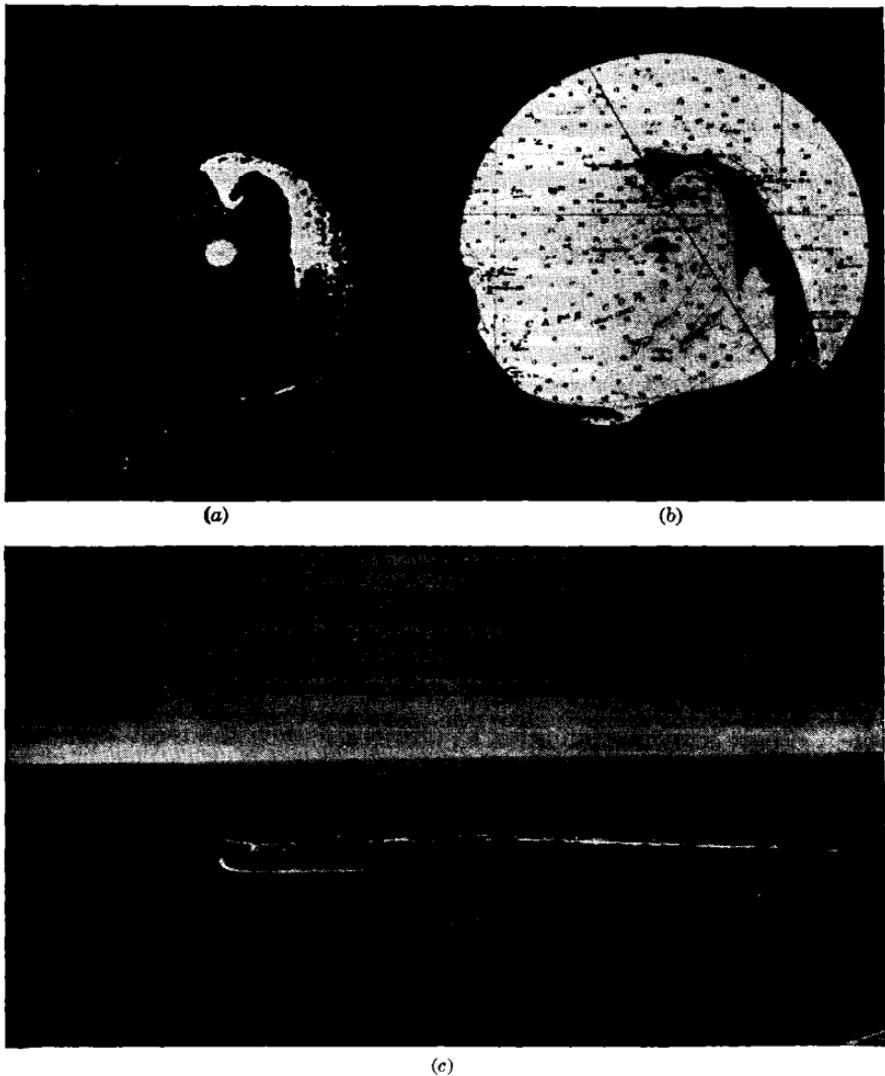


FIG. 1-1.—Cape Cod: (a) radar photograph, (b) map, (c) optical photograph from aircraft carrying the radar set. The radar photograph shown in (a) is one of the earliest pictures taken with 3-cm airborne equipment (summer 1942).

from an airplane flying over Cape Cod Bay. There can be little question of the superiority of the radar picture for most purposes, especially in view of the fact that had it been night, or had the weather been foggy,

the radar picture would have been unaffected while photography or ordinary vision would have been useless.

1-2. How Radar Works.—The coined word *radar* is derived from the descriptive phrase “radio detection and ranging.” Radar works by sending out radio waves from a transmitter powerful enough so that measurable amounts of radio energy will be reflected from the objects to be seen by the radar to a radio receiver usually located, for convenience, at the same site as the transmitter. The properties of the received echoes are used to form a picture or to determine certain properties of the objects that cause the echoes. The radar transmitter may send out c-w signals, or frequency-modulated c-w signals, or signals modulated in other ways. Many schemes based on transmissions of various sorts have been proposed and some of them have been used. Chapter 5 of this book treats the general radar problem, in which any scheme of transmitter modulation may be used, in a very fundamental and elegant way.

Despite the great number of ways in which a radar system can in principle be designed, one of these ways has been used to such an overwhelming degree that the whole of this book, with the exception of Chap. 5, is devoted to it. When radar is mentioned without qualification in this book, pulse radar will be meant. No apology for this specialization is needed. Thousands of times as much effort as that expended on all other forms of radar put together has gone into the remarkably swift development of pulse radar since its origin in the years just before World War II.

In pulse radar, the transmitter is modulated in such a way that it sends out very intense, very brief pulses of radio energy at intervals that are spaced rather far apart in terms of the duration of each pulse. During the waiting time of the transmitter between pulses, the receiver is active. Echoes are received from the nearest objects soon after the transmission of the pulse, from objects farther away at a slightly later time, and so on. When sufficient time has elapsed to allow for the reception of echoes from the most distant objects of interest, the transmitter is keyed again to send another very short pulse, and the cycle repeats. Since the radio waves used in radar are propagated with the speed of light, c , the delay between the transmission of a pulse and the reception of the echo from an object at range R will be

$$T = \frac{2R}{c}, \quad (1)$$

the factor 2 entering because the distance to the target has to be traversed twice, once out and once back. Figure 1-2 shows schematically the principle of pulse radar.

The linear relation between delay time and range shown in Eq. (1) is

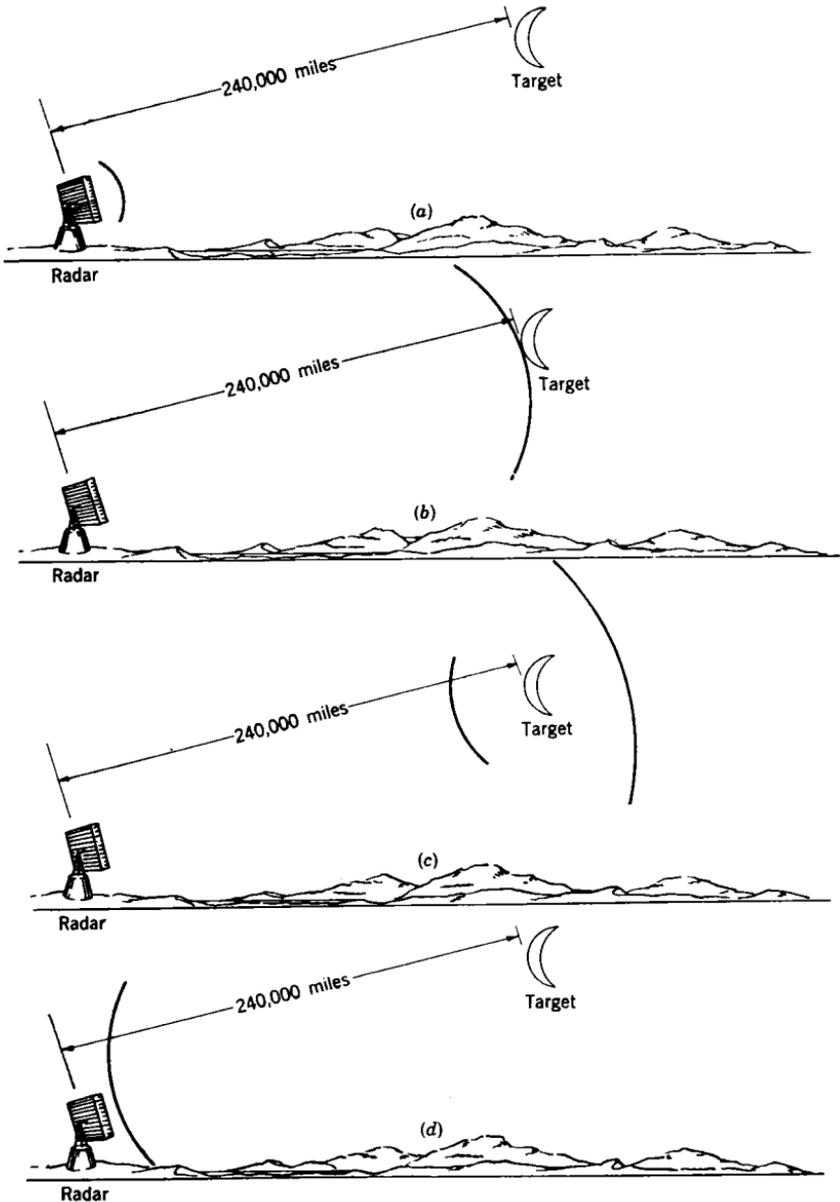


FIG. 1-2.—The principle of pulse radar. (a) Pulse has just been emitted from radar set. (b) Pulse reaches target. (c) Scattered energy returns from target; transmitted pulse carries on. (d) Echo pulse reaches radar.

the clue to the ease with which range can be measured by radar. Range measurement is reduced to a measurement of time, and time can be measured perhaps more accurately than any other basic physical quantity. Because the velocity of light is high, the intervals of time that must be measured in radar are short. Numerically, the range corresponding to a given delay time is 164 yd for each microsecond elapsing between the transmission of the pulse and the reception of the echo. If it is desired to measure range to a precision of 5 yd, which is necessary in some applications of radar, time intervals must be measured with a



FIG. 1-3.—The simplest radar display, the A-scope. The echo at the right is from the moon.

precision better than $\frac{1}{30}$ μ sec. Modern electronic timing and display techniques have been developed to such a point that this can readily be done.

One of the simplest ways in which radar echo signals can be displayed is shown in Fig. 1-3. The beam of a cathode-ray tube is caused to begin a sweep from left to right across the face of the tube at the instant a pulse is sent from the transmitter. The beam is swept to the right at a uniform rate by means of a sawtooth waveform applied to the horizontal deflection plates of the CRT. The output signals of the radar receiver are applied to the vertical deflection plates. To ensure that the weakest signals that are at all detectable are not missed, the over-all gain of the receiver is high enough so that thermal noise originating in the receiver (Sec. 2-7) is perceptible on the display. The two signals that rise significantly above this noise in Fig. 1-3 are, on the left, the "tail" of the transmitted

pulse leaking into the receiver, and on the right, the echo signal from a radar target. The target in the particular case of Fig. 1-3 is the earth's moon.

The measurement of range by means of radar is thus a straightforward problem of time measurement. It is also desirable to be able to measure the *direction* in which a target lies as viewed from a radar station. In principle, this can be done on the basis of triangulation, using range information on the same target from two or more separate radar locations. Although this method permits of great accuracy and has occasionally been used for special purposes, it is far more desirable from the standpoint of simplicity and flexibility to measure direction, as well as range, from a single radar station. Measurement of target bearing was made possible by the development of radio techniques on wavelengths short enough to permit the use of highly directional antennas, so that a more or less sharp beam of radiation could be produced by an antenna of reasonable physical size.

When the pulses are sent out in such a beam, echoes will be received only from targets that lie in the direction the beam is pointing. If the antenna, and hence the radar beam, is swept or scanned around the horizon, the strongest echo will be received from each target when the beam is pointing directly toward the target, weaker echoes when the beam is pointed a little to one side or the other of the target, and no echo at all when it is pointing in other directions. Thus, the bearing of a target can be determined by noting the bearing of the radar antenna when that target gives the strongest echo signal. This can be done in a variety of ways, and more precise and convenient means for determining target bearing by means of radar have been developed (Chap. 6), but the method described here illustrates the basic principle.

It is convenient to arrange the radar display so that, instead of showing target range only, as in Fig. 1-3, it shows the range and angular disposition of all targets at all azimuths. The plan-position indicator, or PPI, is the most common and convenient display of this type. Figure 1-1 is a photograph of a PPI-scope. The direction of each echo signal from the center of the PPI shows its direction from the radar; its distance from the center is proportional to target range. Many other forms of indication are convenient for special purposes; the various types of indicator are catalogued in Chap. 6.

1-3. Components of a Radar System.—A radar set can be considered as separable, for the purposes of design and description, into several major components concerned with different functions. Figure 1-4 is a block diagram of a simple radar set broken up into the components ordinarily distinguished from one another.

In the set illustrated in Fig. 1-4, a cycle of operation is begun by the

firing of the modulator. This sends a high-power, high-voltage pulse to the magnetron, which is the type of transmitting tube almost universally used in modern radar. For the brief duration of the modulator pulse, which may typically be $1 \mu\text{sec}$, the magnetron oscillates at the radio frequency for which it is designed, usually some thousands of megacycles per second. The r-f pulse thus produced travels down the r-f transmission line shown by double lines in Fig. 1-4, and passes through the two switches designated as TR and ATR. These are gas-discharge devices of a very special sort. The gas discharge is started by the high-power

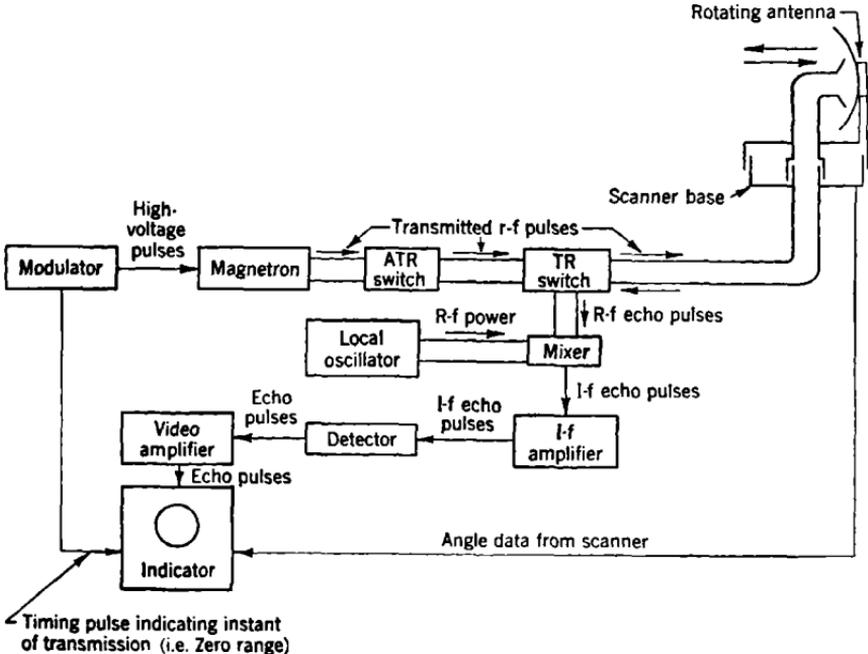


FIG. 1-4.—Block diagram of a simple radar.

r-f pulse from the transmitter, and maintained for the duration of that pulse; during this time the TR (for transmit-receive) switch connects the transmitter r-f line to the antenna, and disconnects the mixer and the rest of the radar receiver shown below the TR switch. The ATR (for anti-TR) switch, when fired, simply permits the r-f pulse from the transmitter to pass through it with negligible loss. Between pulses, when these gas-discharge switches are in an unfired state, the TR switch connects the mixer to the antenna, and the ATR disconnects the magnetron to prevent loss of any part of the feeble received signal.

After passing through these two switches, the transmitter pulse travels down the r-f line to the antenna, where it is radiated. The

antenna is designed in such a way that the beam shape it produces is suitable for the requirements the radar set must meet. It is mounted on a scanner which is arranged to sweep the beam through space in the manner desired; simple azimuth rotation is indicated in Fig. 1-4.

After the transmission of the pulse, the discharges in the TR and ATR switches cease and the system is ready to receive echoes. Echoes are picked up by the antenna and sent down the r-f line to the mixer. The mixer is a nonlinear device which, in addition to receiving the signals from the antenna, is supplied c-w power from a local oscillator operating at a frequency only a few tens of megacycles per second away from the magnetron frequency. The difference frequency that results from mixing these two signals contains the same intelligence as did the original r-f echoes, but it is at a sufficiently low frequency (typically, 30 Mc/sec) to be amplified by more or less conventional techniques in the intermediate-frequency amplifier shown. Output signals from the i-f amplifier are demodulated by a detector, and the resulting unipolar signals are further amplified by a video-frequency amplifier similar to those familiar in television technique.

The output signals of the video amplifier are passed to the indicator, which displays them, let us say for definiteness, in plan-position form. In order to do this, it must receive a timing pulse from the modulator, to indicate the instant at which each of the uniform range sweeps out from the center of the PPI tube should begin. It must also receive from the scanner information on the direction in which the antenna is pointing, in order that the range sweep be executed in the proper direction from the center of the tube. Connections for accomplishing this are indicated in the Fig. 1-4.

In Chaps. 9 to 14, inclusive, the detailed design of each of the components shown in Fig. 1-4 is treated. In addition, consideration is given to the problem of supplying primary power in a form suitable for use with a radar set; this is especially difficult and important in the case of airborne radar.

1-4. The Performance of Radar.—In discussing the performance of radar, one usually refers to its *range performance*—that is, the maximum distance at which some target of interest will return a sufficiently strong signal to be detected. The factors that determine range performance are numerous and they interact in a rather complicated way. Chapter 2 is devoted to a discussion of them, and Chap. 3 deals with the important matter of the properties of radar targets.

The usual inverse-square law which governs the intensity of radiation from a point source acts to determine the range dependence of the fraction of the total transmitted energy that falls on a target. So far as the echo is concerned, the target can also be thought of as a point source of radia-

tion, so that the inverse-square law must be applied again to determine the range dependence of the amount of echo energy reaching the receiver. In consequence, the echo energy received from a target varies with the inverse fourth power of the range from the radar set to the target, other factors being constant.

To be detectable, a signal must have a certain minimum power; let us call the minimum detectable signal S_{\min} . Then the maximum range of a radar set on a target of a given type will be determined by S_{\min} , according to the expression

$$S_{\min} = \frac{KP_t}{R_{\max}^4},$$

where K is a constant and P_t is the power in the transmitted pulse, to which the received signal power will clearly be proportional. Rearranging,

$$R_{\max} = \left(\frac{KP_t}{S_{\min}} \right)^{1/4}. \quad (2)$$

Equation (2) displays the difficulty of increasing the range performance of a radar set by raising its pulse power. A 16-fold increase in power is required to double the range.

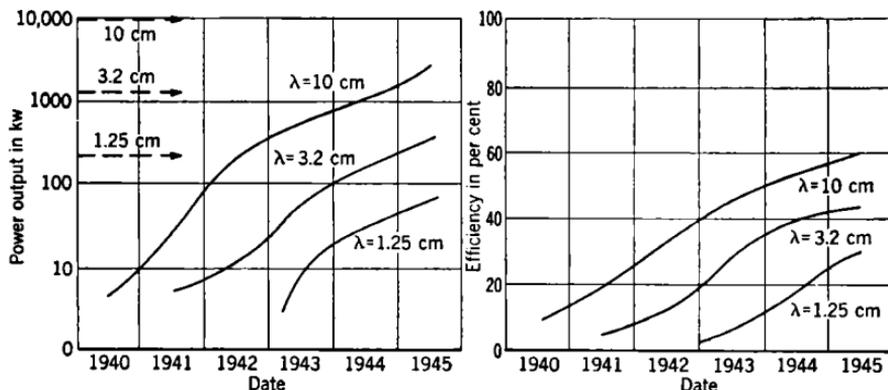


FIG. 1-5.—Historical development of microwave magnetrons.

However formidable this requirement appears, one of the most remarkable facts of the wartime years of development of radar is that practicable pulse powers in the microwave frequency range (about 1000 Mc/sec and above) have increased by a factor of *hundreds* in a relatively short time. This stupendous advance resulted from the invention and rapid improvement of the multicavity magnetron, which is described in Chap. 10. Figure 1-5 shows the history of magnetron development, with respect to pulse power and efficiency, at the three most important microwave bands exploited during the war. The curves are rather arbitrarily drawn, and only their general trend is significant. Not every upward

step in output power was due to an improvement in the magnetron itself. The increase at 10-cm wavelength in the early part of 1941 was brought about by the development of modulators of higher power.

It is important to realize that the curves of Fig. 1-5 lie above one another in the order of increasing wavelength not because development was begun earlier at 10 cm than at 3 cm, and earlier at 3 cm than at 1 cm, but because magnetrons of the type used in radar are subject to inherent limitations on maximum power which are more severe the shorter the wavelength. The same is true of the r-f transmission lines used at microwave frequencies. The horizontal dashed lines shown in Fig. 1-5a show the maximum power that can be handled in the standard sizes of "waveguide" used for r-f transmission at the three bands.

A similarly spectacular decrease in the minimum detectable signal, due to the improvement of microwave radar receivers, has marked the war years. In the wavelength bands above about 10 m, natural "static" and man-made interference set a rather high noise level above which signals must be detected, so that there is little necessity for pursuing the best possible receiver performance. This is not true at microwave frequencies. Natural and man-made interference can be neglected at these frequencies in comparison with the unavoidable inherent noise of the receiver. This has put a premium on the development of the most sensitive receivers possible; at the end of 1945 microwave receivers were within a factor of 10 of theoretically perfect performance. Improvement by this factor of 10 would increase the range of a radar set only by the factor 1.8; and further receiver improvement can today be won only by the most painstaking and difficult attention to details of design.

Why Microwaves?—The reader will have observed that when radar is discussed in what has gone before, microwave radar is assumed. This is true of the balance of this book as well. So far as the authors of this book are concerned, the word *radar* implies not only *pulse* radar, as has already been remarked, but *microwave* pulse radar. Though it is true that the efforts of the Radiation Laboratory were devoted exclusively to microwave pulse radar, this attitude is not entirely parochialism. The fact is that for nearly every purpose served by radar, microwave radar is preferable. There are a few applications in which longer-wave radar is equally good, and a very few where long waves are definitely preferable, but for the overwhelming majority of radar applications microwave radar is demonstrably far more desirable than radar operating at longer wavelengths.

The superiority of microwave radar arises largely because of the desirability of focusing radar energy into sharp beams, so that the direction as well as the range of targets can be determined. In conformity with the well-known laws of physical optics, by which the sharpness of

the beam passing through an aperture of given size depends on the ratio of the diameter of the aperture to the wavelength of the radiation in the beam, the sharpness of the beam produced by a radar antenna (which can be thought of as a sort of aperture for the radio energy) depends on the ratio of the antenna dimensions to the wavelength used. For an antenna of given size, the breadth of the beam produced is proportional to the wavelength. These statements are made precise in Sec. 9.1.

Particularly in the case of airborne radar, where a large antenna cannot be tolerated for aerodynamic reasons, it is important to produce sharp radar beams with an antenna structure of modest size. This demands the use of microwaves. Roughly speaking, microwaves are radio waves whose wavelength is less than 30 cm.

Radar definition, its ability to discriminate between targets close together in space, improves as the beamwidth is narrowed. Targets at the same range can be distinguished by radar as being separate if they are separated in azimuth by an angle larger than one beamwidth; thus the quality of the picture afforded by radar improves as the beamwidth is reduced. For an antenna of given size, the beamwidth can be decreased only by lowering the wavelength.

The finite velocity of light sets a limit to the desirable beamwidth if a region of finite size is to be scanned at a given speed by a radar set. Chapter 4 considers this and other limitations of pulse radar in some detail.

The Propagation of Microwaves.—Further limitations on the performance of radar arise from the propagation properties of radio waves in the microwave region of the electromagnetic spectrum. Like light, microwaves are propagated in straight lines. Unlike radio waves at frequencies lower than about 30 Mc/sec, microwaves are not reflected from the ionosphere. This means that the maximum range of a radar set whose performance is not otherwise limited will be set by the optical horizon which occurs because the earth is round. This is in fact the limitation on the performance of the best radar sets developed during the war. Under certain conditions, bending of the microwave beam around the earth is produced by meteorological conditions (Sec. 2.14). This can increase the range of a radar set beyond the optical horizon, but such phenomena are relatively rare and essentially unpredictable.

A lower limit on the wavelengths which can be used for practical radar systems is fixed by the onset of atmospheric absorption of microwave energy. Below a wavelength of about 1.9 cm, serious absorption occurs in moist atmosphere, because of a molecular transition in water vapor which can be excited by the radiation (Sec. 2.15). For this reason, 2 cm is about the shortest wavelength at which radar systems of good range performance can be built. For certain very special applications

where high absorption can be tolerated or is even welcome, shorter wavelengths can be used, but 2 cm is a good practical limit. The wartime development of radar components and systems at 1.25 cm antedated the discovery of the strong water-vapor absorption at this wavelength. A wavelength of 1.25 cm is, fortuitously, very nearly the most unfortunate choice that could have been made in the development of a new short-wavelength band.

1.5. Radar Systems.—The uses made of radar were so various under wartime conditions that many different systems were developed to fill different needs. These systems usually differed more in regard to beam shape, scanning means, and mode of indication than in regard to any other properties. Chapter 6 gives a brief conspectus of the principal varieties of radar, with especial emphasis on those types that promise to have an important peacetime use. Two examples of the detailed design of radar systems are given in Chap. 15, after the components of radar systems have been discussed.

Considerable use has been made of radar beacons. These are devices which, on receiving a pulse or a series of properly coded pulses from a radar set, will send back in reply a pulse or a series of coded pulses. A great increase in the flexibility and convenience of the use of radar under certain conditions can be obtained by the use of such beacons. A brief account of their properties and uses, though not of their design, will be found in Chap. 8.

Toward the end of the war, two major developments occurred which promised to extend greatly the applicability of pulse radar under unfavorable conditions. Means were developed for reproducing radar indications at a point distant from the set that gathered the original data; the intelligence necessary was transmitted from the radar to the distant indicator by radio means. This radar relay, as it has come to be called, is described in some detail in Chap. 17.

Chapter 16 deals with another important development—namely, the modification of pulse-radar equipment so that it will display only targets that are in motion relative to the radar. Such moving-target indication is potentially of great importance in freeing radar from the limitations of site. At the present, a radar site must be chosen with careful attention to the surrounding terrain; hills or buildings within the line of sight can return strong “permanent echoes” which mask target signals over a large part of the desirable coverage of the set. In mountainous terrain, this problem is very serious. An arrangement that gives signals only from targets that are moving appears to be the best solution to the permanent-echo problem.

A fact that has been too little recognized when radar systems are discussed is that the organization which is to make use of the positional

information afforded by radar is usually at least as important as is the radar itself. A good organization can make excellent use even of inferior radar information, as was proved by the success of the British Home Chain of radar stations, the first large-scale radar installation to be made. An inadequate organizational set-up can do a poor job, even though provided with splendid radar from the technical standpoint. The many problems that enter into the creation of an adequate organization for the use of radar data have not received the study that they should. Despite this fact, Chap. 7 attempts to provide an introduction to this sort of planning, and to raise some of the important problems, even though they may not yet be satisfactorily solved.

1-6. The Early History of Radar.—Though the complete history of the origins and the growth of modern radar is a long and complicated one,¹ it will be of some interest to sketch here its main lines, with especial reference to Allied developments.

Successful pulse radar systems were developed independently in America, England, France, and Germany during the latter 1930's. Back of their development lay half a century of radio development for communication purposes, and a handful of early suggestions that, since radio waves are known to be reflected by objects whose size is of the order of a wavelength, they might be used to detect objects in fog or darkness.

The fact that radio waves have optical properties identical with those associated with ordinary visible light was established by Heinrich Hertz in 1886, in the famous series of experiments in which he first discovered radio waves. Hertz showed, among other things, that radio waves were reflected from solid objects. In 1904 a German engineer, Hulsmeyer, was granted a patent in several countries on a proposed way of using this property in an obstacle detector and navigational aid for ships. In June 1922, Marconi strongly urged the use of short waves for radio detection.

The principle of pulse ranging which characterizes modern radar was first used in 1925 by Breit and Tuve, of the Carnegie Institution of Washington, for measuring the height of the ionosphere.² After the successful experiments of Breit and Tuve, the radio-pulse echo technique became the established method for ionospheric investigation in all countries. The step from this technique to the notion of using it for the detection of aircraft and ships is, in retrospect, not such a great one; and various individuals took it independently and almost simultaneously in America,

¹ For the fullest treatment of radar history available, the reader is referred to the official history of Div. 14, NDRC, "Radar" by H. E. Guerlac, to be published by Little, Brown, & Co., Boston.

² M. A. Tuve and G. Breit, "Terrestrial Magnetism and Atmospheric Electricity," Vol. 30, March-December 1925, pp. 15-16. Also *Phys Rev*, 28, 554 (1926).

England, France, and Germany, about ten years after the original work of Breit and Tuve.

The research agencies of the American Army and Navy have a long and complicated history of early experiment, total failure, and qualified success in the field of radio detection. The interested reader will find this dealt with at length in Dr. Guerlac's history.¹ Here it will be sufficient to report the earliest full successes. In early 1939, a radar set designed and built at the Naval Research Laboratory was given exhaustive tests at sea during battle maneuvers, installed on the *U.S.S. New York*. The first contract for the commercial manufacture of radar equipment was let as a result of these tests, for the construction of six sets, designated as CXAM (Sec. 6-9), duplicating that used in the trials. In November 1938, a radar position-finding equipment intended for the control of anti-aircraft guns and searchlights, designed and built by the Signal Corps Laboratories of the Army, was given extensive tests by the Coast Artillery Board, representing the using arm. This set also went into quantity manufacture, as the SCR-268 (Sec. 6-14). An Army long-range aircraft-detection set whose development had been requested earlier by the Air Corps was demonstrated to the Secretary of War by the Signal Corps Laboratories in November 1939. A contract for the production of this equipment, the SCR-270 (and SCR-271; see Sec. 6-9) was let in August 1940.

British radar was developed at about the same time but its application proceeded at a somewhat faster pace under the immediate threat to England and with considerably greater realism during the early years of the war. During the winter of 1934-1935, the Air Ministry set up a Committee for the Scientific Survey of Air Defense. Among the suggestions it received was a carefully worked out plan for the detection of aircraft by a pulse method, submitted by a Scottish physicist, now Sir Robert Watson-Watt, then at the head of the Radio Department of the National Physical Laboratory.

The first experimental radar system of the type suggested by Watson-Watt was set up in the late spring of 1935 on a small island off the east coast of England. Development work during the summer led to the blocking-out of the main features of the British Home Chain of early-warning stations (Sec. 6-9) by fall. Work began in 1936 toward setting up five stations, about 25 miles apart, to protect the Thames estuary. By March 1938, all these stations—the nucleus of the final Chain—were complete and in operation under the charge of RAF personnel.

British radar development effort was then brought to bear on airborne radar equipment. Two types were envisaged: a set for the detection of surface vessels by patrol aircraft (called ASV, for *air to surface vessel*),

¹ *Op. cit.*

and an equipment for enabling night fighters to home on enemy aircraft (called AI, for aircraft interception). Work was concentrated on ASV first, and an experimental equipment was successfully demonstrated during fleet maneuvers in September 1938. Experimental AI equipment was working by June, 1939, and it was demonstrated to the chief of RAF Fighter Command in August of that year. The Air Ministry asked that 30 such systems be installed in aircraft in the next 30 days. Before the end of September all these systems had been installed, four having been ready on the day war broke out.

Emphasis on airborne radar underlined the point that, if sharp radar beams were ever to be produced by antennas small enough to carry in an airplane, wavelengths shorter than the $1\frac{1}{2}$ m used in early British airborne equipment would have to be employed. This led to the effort that the British put into developing a generator of microwaves which could give pulse power adequate for radar use. By early 1940, a British version of the multicavity magnetron had been developed to the point where it was an entirely practicable source of pulsed microwave energy, and the history of modern radar had begun.

1-7. Wartime Radar Development in the United States.—Before the end of 1940, the work on radar of American and British laboratories had been combined as a result of an agreement between the two governments for exchange of technical information of a military nature. A British Technical Mission arrived in Washington in September 1940 and mutual disclosures were made of British and American accomplishments in radar up to that time. Members of the British mission visited the Naval Research Laboratory, the Army Signal Corps Laboratories at Fort Monmouth, and the Aircraft Radio Laboratory at Wright Field, as well as manufacturing establishments engaged in radar work. They demonstrated their version of the cavity magnetron and furnished design information that enabled U. S. manufacturers to duplicate it promptly.

In discussions with the Microwave Committee of the National Defense Research Committee, which had been set up a few months before, members of the British Mission proposed two specific projects which they suggested that the United States undertake: a microwave aircraft-interception equipment, and a microwave position finder for anti-aircraft fire control.

To implement their decision to follow these suggestions, the Microwave Committee of the NDRC decided to set up a development laboratory staffed primarily by physicists from a number of universities. They were encouraged in this step by the success that the British had already experienced with civilian wartime radar development agencies staffed with physicists having no special radio experience but good general scientific training. After exploring several possibilities, the

Microwave Committee persuaded the Massachusetts Institute of Technology to accept the responsibility of administering the new laboratory. The Radiation Laboratory, as it was named, opened its doors early in November 1940. The director of the laboratory throughout its 62 months of life was Dr. L. A. DuBridge.

The Army and Navy development laboratories were glad to depend on the new Radiation Laboratory for an investigation of the usefulness for radar of the new microwave region of the radio spectrum. They were fully occupied with the urgent engineering, training, and installation problems involved in getting radar equipment that had already been developed out into actual military and naval service. At the end of 1940, the use of microwaves for radar purposes seemed highly speculative, and the Service laboratories quite properly felt it their duty to concentrate on radar techniques that had already been worked out successfully.

During 1941, while the Navy was installing long-wave search radar and medium-wavelength fire-control radar on ships of the fleet, and the Army was sending out Signal Aircraft Warning Battalions equipped with the SCR-270 and anti-aircraft batteries with the SCR-268, not a single item of radar equipment based on the new microwave techniques was delivered for operational use. However, development work at the Radiation Laboratory had broadened far beyond the two specific projects suggested by the British Technical Mission, and microwave equipment was showing great promise for many wartime uses.

A few important dates will indicate the way in which this development was proceeding. On Jan. 4, 1941, the Radiation Laboratory's first microwave radar echoes were obtained. A successful flight test of a working "breadboard" model of an airborne radar intended for AI use was made on March 10, in a B-18A furnished by the Army Air Corps. In this flight it was found that the equipment was extremely effective in searching for ships and surfaced submarines at sea.

In the late spring of 1941, an experimental microwave sea-search radar equipped with a PPI was installed on the old destroyer *U.S.S. Semmes*. On June 30, the Navy let the first production contract for microwave radar equipment based on the work of the Radiation Laboratory. This was for a production version of the set that had been demonstrated on the *Semmes*.

At the end of May, a prototype of the microwave anti-aircraft position finder developed at the Radiation Laboratory was in operation. It accomplished the then-astonishing feat of tracking a target plane in azimuth and elevation wholly automatically.

These and other early successes led to an increasing Service interest in microwave radar, which had seemed so speculative a venture in 1940. The tremendous expansion of the development program can be measured

by the fact that the personnel of the Radiation Laboratory, which had been about 40 at the beginning of 1941, rose to nearly 4000 by mid-1945. Similarly, the Radar Section of the Naval Research Laboratory increased its personnel to 600. The Radio Position Finding section of the Signal Corps Laboratories grew into the Evans Signal Laboratory, with a peak personnel of more than 3000. A similar growth took place at the Aircraft Radio Laboratory at Wright Field.

A tremendous amount of work was carried out during the war by the research and engineering staffs of many industrial concerns, both large and small. In some cases, these firms, working either independently or on development or production contracts with the armed forces or with NDRC, engineered certain types of radar sets all the way through from the basic idea to the finished product. To a larger extent, the contribution of industry was to take the prototype equipment produced in government laboratories and make the design suitable for quantity manufacture and for service use under combat conditions. The art advanced so rapidly in the early years that manufacturers were often called upon to make major changes during the course of production in order to take account of new lessons from both the laboratories and the battlefields.

The growth of the radar industry, which scarcely existed before 1940, is indicated by the fact that by the end of June 1945, approximately \$2,700,000,000 worth of radar equipment had been delivered to the Army and the Navy. At the end of the war, radar equipment was being produced at a rate of more than \$100,000,000 worth per month.

The enormous wartime investment of money, skill, and productive facilities in radar paid the Allies handsome dividends with the fleet, in the air, and on the battlefield.¹ The uses of radar in a peaceful world were just beginning to be worked out in 1946. Some of these are dealt with in Vol. 2 of this series. But the technical achievement represented by the wartime development of radar seems very nearly unparalleled. In terms of the time intervening between the reception of the first radar signals and the large-scale use of radar in the war, it is as if, seven years after the first faltering flight of the Wright brothers at Kitty Hawk, the airplane had been developed into a powerful weapon of which thousands were in constant use.

¹ The story of radar's operational use in the war is told, in a way that is somewhat blurred about the edges by the censorship obtaining just before the end of the war with Japan, in a pamphlet entitled "Radar: A Report on Science at War," released by the Joint Board on Scientific Information Policy on Aug. 15, 1945. It is obtainable from the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.

CHAPTER 2

THE RADAR EQUATION

BY E. M. PURCELL

The operation of a radar set depends on the detection of a weak signal returned from a distant reflecting object. The factors which control the strength of the signal so received are clearly of first importance in determining the maximum range of detection of a given target by a specified radar set. In Secs. 2·1 to 2·6 we shall formulate and examine the basic relation between these quantities, which is commonly known as the "radar equation." Specifically, we want to derive an expression for the peak radio-frequency signal power S , available at the terminals of the radar antenna, which will involve measurable properties of the transmitting and receiving antenna system, the transmission path through space, and the target itself. Now this relation will not suffice to fix the maximum range of detection unless the minimum power required for detection, S_{\min} , is known. This important quantity S_{\min} we prefer to discuss separately, beginning in Sec. 2·7 below. It will be found to depend on many other factors, not all readily accessible to measurement, ranging from thermal noise in a resistor to the integrating property of the eye of the radar observer. Thus we choose to divide the problem into two parts, by a fictitious boundary, as it were, between the radar antenna and the rest of the set. The relations which we shall develop in Secs. 2·1 to 2·6 are wholly *geometrical* ones in the sense that the factors upon which the received power S depends are all lengths, apart from the transmitted power P , to which, of course, S is always proportional.

THE RADAR EQUATION FOR FREE-SPACE PROPAGATION

2·1. The Meaning of Free-space Propagation.—Fortunately, the quasi-optical nature of microwave propagation permits us to concentrate our attention at the outset on a very simple case, which we shall call "free-space propagation." The circumstances implied would be realized exactly if radar set and target were isolated in unbounded empty space. They are realized well enough for practical purposes if the following conditions are fulfilled:

1. No large obstacles intervene between antenna and target, along an optical line of sight.

2. No alternate transmission path, via a reflecting surface, can be followed by a substantial fraction of the radiated energy.
3. The intervening atmosphere is homogeneous with respect to index of refraction, at the frequency used.
4. The intervening atmosphere is transparent, i.e., does not absorb energy from the wave, at the frequency used.

Condition 1 restricts our attention to targets within the horizon. Condition 2 bars, for the present, consideration of radar search at low angles over water, although later we shall include this case by a suitable modification of the radar equation. Microwave radar over land appears to be relatively free, even at low angles, from the reflection effects which are so pronounced at longer wavelengths. In any case, if the directivity of the antenna pattern is such that very little energy strikes the reflecting surface, Condition 2 is fulfilled. Any implications of Conditions 3 and 4 which are not self-evident will be clarified in the last part of this chapter, where other types of propagation will be discussed. If, now, these conditions of free-space propagation apply, the result is very simple: The transmitted wave, at any considerable distance from the antenna,¹ has spherical wavefronts—limited in extent, of course, by the radiation pattern of the antenna—which spread so that the intensity of the disturbance falls off with the inverse square of the distance.

2-2. Antenna Gain and Receiving Cross Section.—If the transmitting antenna were to radiate energy isotropically—that is, uniformly in all directions—the power flow through unit area at a distance R from the antenna could be found by dividing P , the total radiated power, by $4\pi R^2$. A *directive* antenna, however, will concentrate the energy in certain directions. The power flow observed at some distant point will differ by some factor G from that which would be produced by an antenna radiating isotropically the same total power. This factor G is called the “gain” of the antenna in the direction in question. By our definition, the gain of the hypothetical isotropic radiator is 1 in every direction. For any other antenna G will be greater than 1 in some directions and less than 1 in others. It is clear that G could not be greater than 1 in every direction, and in fact the average of G taken over the whole sphere must be just 1.

Usually we are interested in antennas for which G has a very pronounced maximum in one direction, that is to say, antennas which

¹ The limitation implied is to distances greater than d^2/λ , where d is the width of the antenna aperture and λ the wavelength. At distances R less than this (less than 360 ft, for example, for $\lambda = 3$ cm, $d = 6$ ft), the intensity does not fall off as $1/R^2$. Although this region has been until now of no interest for radar applications, one can anticipate the development of short-range, very-high-resolution radar for which the near zone, so defined, will be of primary importance.

radiate a well-defined *beam*. This maximum value of G we shall denote by G_0 . The narrow, concentrated beams which are characteristic of microwave radar require, for their formation, antennas large compared to a wavelength. In nearly every case the radiating system amounts to an aperture of large area over which a substantially plane wave is excited. For such a system, a fundamental relation connects the maximum gain G_0 , the area of the aperture A , and the wavelength:

$$G_0 = \frac{4\pi A f}{\lambda^2}. \quad (1)$$

The dimensionless factor f is equal to 1 if the excitation is uniform in phase and intensity over the whole aperture; in actual antennas f is often as large as 0.6 or 0.7 and is rarely less than 0.5. An antenna formed by a paraboloidal mirror 100 cm in diameter, for a wavelength of 10 cm, would have a gain of 986 according to Eq. (1) with $f = 1$, and in practice might be designed to attain $G_0 = 640$.

The connection between gain and beamwidth is easily seen. Using an aperture of dimensions d in both directions, a beam may be formed whose angular width,¹ determined by diffraction, is about λ/d radians. The radiated power is then mainly concentrated in a solid angle of λ^2/d^2 . An isotropic radiator would spread the same power over a solid angle of 4π . Therefore, we expect the gain to be approximately $4\pi d^2/\lambda^2$, which is consistent with Eq. (1), since the area of the aperture is about d^2 . For a more rigorous discussion of these questions the reader is referred to Vol. 12, Chap. 5.

A complementary property of an antenna which is of importance equal to that of the gain is the *effective receiving cross section*. This quantity has the dimensions of an area, and when multiplied by the power density (power per unit area) of an incident plane wave yields the total signal power available at the terminals of the antenna. The effective receiving cross section A_r is related to the gain as follows:

$$A_r = \frac{G\lambda^2}{4\pi}. \quad (2)$$

Note that G , not G_0 , has been written in Eq. (2), the applicability of which is not restricted to the direction of maximum gain or to beams of any special shape. Once the gain of the antenna in a particular direction is specified, its effective receiving cross section for plane waves incident *from* that direction is fixed. Equation (2) can be based rigorously on the Reciprocity Theorem (see Vol. 12, Chap. 1). Comparing Eqs. (2) and (1) we observe that, if the factor f is unity, the effective receiving cross

¹ Wherever a precise definition of beamwidth is intended, we shall mean the angular interval between two directions for which $G = G_0/2$.

section of an antenna in the principal direction is precisely the area of the aperture; in other words, all the energy incident on the aperture is absorbed. Quite generally A_r will depend on the area of the antenna aperture and not on λ , whereas G_0 will depend on A/λ^2 .

2.3. Scattering Cross Section of the Target.—We have to consider how the target itself enters the radar problem. Evidently we need some measure of the amount of power reflected by the target. For this purpose we define the scattering cross section of the target σ as follows: σ (dimensions of an area) is to be 4π times the ratio of the power per unit solid angle scattered back toward the transmitter, to the power density (power per unit area) in the wave incident on the target. In other words, if at the target the power incident on an area σ placed normal to the beam were to be scattered uniformly in all directions, the intensity of the signal received back at the radar set would be just what it is in the case of the actual target. In some respects "radar cross section" is a more appropriate name for σ in so far as it indicates that we are concerned only with the power scattered directly back toward the transmitter.

It is essential to realize that the cross section of a given target will depend not only upon the wavelength but upon the angle from which the target is viewed by the radar. The fluctuation of σ with "target aspect," as it is called, is due to the interference of reflected waves from various parts of the target (see Chap. 3). Only for certain special cases can σ be calculated rigorously; for most targets σ has to be inferred from the radar data themselves. Usually this cannot be done in any uniform way because of the fluctuation referred to, and it may be well to assert at this point that intelligent use of the formulas which we shall derive, in all of which σ appears, requires an appreciation of these limitations.

2.4. The Radar Equation.—With the pertinent quantities defined it is now a simple matter to formulate the radar equation. If S is the signal power received, P the transmitted power, G the gain of the antenna, λ the wavelength, σ the radar cross section of the target, and R the distance to the target or range, this relation must hold:

$$S = \left(\frac{PG}{4\pi R^2} \right) \left(\frac{\sigma}{4\pi R^2} \right) \left(\frac{G\lambda^2}{4\pi} \right). \quad (3a)$$

The quantity in the first parenthesis is the power density in the incident wave at the target. The first two terms in parentheses together give the power density in the returning wave at the radar antenna, and the last factor will be recognized as the receiving cross section of the radar antenna, from Eq. (2). Rearranging terms, for compactness,

$$S = P \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4}. \quad (3b)$$

Again we call attention to the fact that Eq. (3b), like Eq. (2), contains G rather than G_0 , and is not restricted to any particular direction or to beams of any special shape. The sole restriction which has not yet been made explicit is that G should not vary significantly over the angle which the target subtends at the radar antenna.

Usually we shall be interested in the signal that is returned when the target lies somewhere along the maximum of the radar beam, and we should then replace G by G_0 . It is instructive to proceed then to eliminate G_0 by means of Eq. (1), obtaining

$$S = \frac{P\sigma A^2 f^2}{4\pi R^4 \lambda^2}. \quad (4a)$$

Note that λ^2 now appears in the denominator, while the numerator contains the square of the area of the antenna aperture. A further manipulation of Eq. (4a) is of interest. Suppose the minimum power required for satisfactory detection, S_{\min} , is known; we may solve Eq. (4a) for the maximum range of detection, R_{\max} :

$$R_{\max} = \sqrt[4]{\frac{P\sigma A^2 f^2}{4\pi S_{\min} \lambda^2}}. \quad (4b)$$

At this point it may be well to get an idea of the order of magnitude of the quantities involved by inserting numbers not unusual in wartime pulse-radar practice. If we choose $\lambda = 0.10$ ft ($= 3.0$ cm), $P = 10^5$ watts, $A = 10$ ft², $f = 0.6$, $\sigma = 100$ ft² (typical for small aircraft), $S_{\min} = 5 \times 10^{-12}$ watts, we obtain $R_{\max} = 155,000$ ft or 29 statute miles.

We observe that a 16-fold increase in transmitted power is required to double the maximum range; on the other hand, it would appear that R_{\max} could be doubled by doubling the linear dimensions of the antenna. But the latter step would at the same time reduce the beamwidth by a factor of 2 and as we shall see in Sec. 2-11 that this indirectly affects S_{\min} . A change in wavelength is even more difficult to discuss as it entails changes in S_{\min} , P , and possibly σ as well.

2-5. Beams of Special Shapes.—In several applications of radar, use is made of an antenna designed to spread the radiated energy out over a considerable range in angle in one plane. The object usually is to increase the angular region covered at one time. An example of such a radiation pattern is the simple "fan beam" sketched in Fig. 2-1. It is easy to produce such a beam by means of an antenna whose effective aperture is wide in the direction in which the beam is to be narrow and narrow in the direction in which the beam is to be broad. The connection given by Eq. (1) between aperture and gain still holds, implying that by reducing the vertical dimension of the antenna aperture (in Fig. 2-1), gain has necessarily been sacrificed; of course this is inevitable,

for the radiated energy has been spread out over a large solid angle. Consider the problem of designing a radar set with the requirement imposed that the vertical beamwidth be β . Let us recast Eq. (4) in a form appropriate for this case, introducing explicitly the vertical and horizontal dimensions of the antenna aperture, which we denote by b and a respectively. To a good enough approximation we can set $b = \lambda/\beta$. Then $A^2 = a^2b^2 = a^2\lambda^2/\beta^2$, and inserting this in Eq. (4b) we have

$$R_{\max} = \sqrt[4]{\frac{P\sigma f^2 a^2}{4\pi S_{\min}\beta^2}} \quad (5)$$

The range no longer depends *explicitly* on the wavelength.

From what has been said it should be clear that an excess of antenna gain, in any direction, over what is needed to meet any requirement for

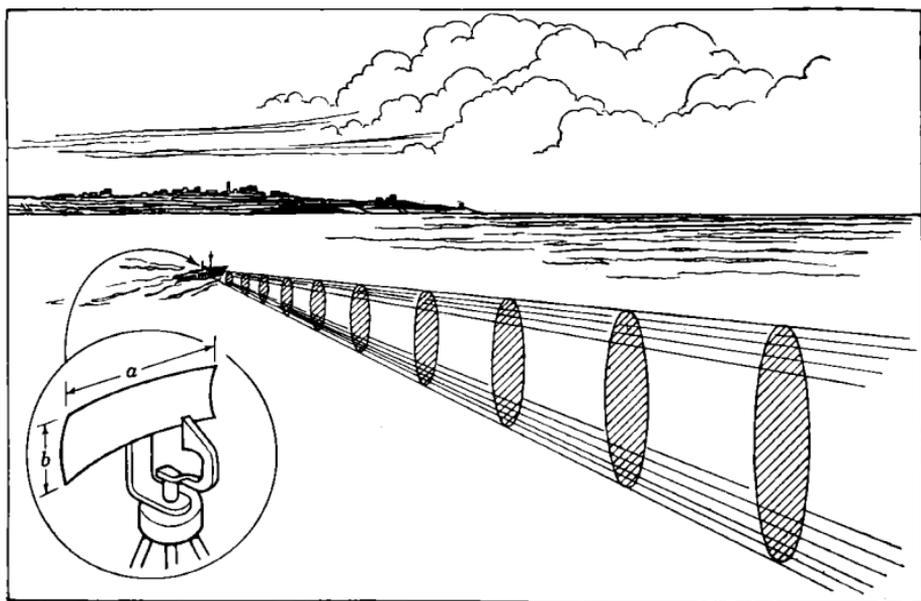


FIG. 2-1.—A radiation pattern of this shape is called a "fan beam."

detecting targets in that direction is wasteful. It would be most desirable to adjust the directional pattern of the antenna so that just the desired angular coverage, and no more, would be obtained. A prescription for such an antenna pattern can be obtained from the basic radar equation, Eq. (3b), in any given case. A particularly important and instructive example is that of airborne ground-mapping radar, which we shall now examine in detail.

The object here is to obtain a radar picture, from above, of a circular area on the ground stretching out in all directions from the aircraft to some maximum range R_0 . This is done by illuminating, at any one

time, a narrow radial strip extending outward from, for example, the point directly beneath the aircraft and rotating or "scanning" this strip about a vertical axis (see Fig. 2-2). Evidently some sort of fan beam, as narrow as possible in the horizontal direction but spread out vertically, is demanded. To find the shape which the beam should have, in a vertical plane, the properties of the target must be taken into account. Unlike the targets previously discussed, which were assumed to be small compared to the cross section of the beam and to be characterized by a fixed radar cross section σ , the ground is an *extended*, or *compound* target. It consists of a multitude of small scattering or reflecting objects, many of

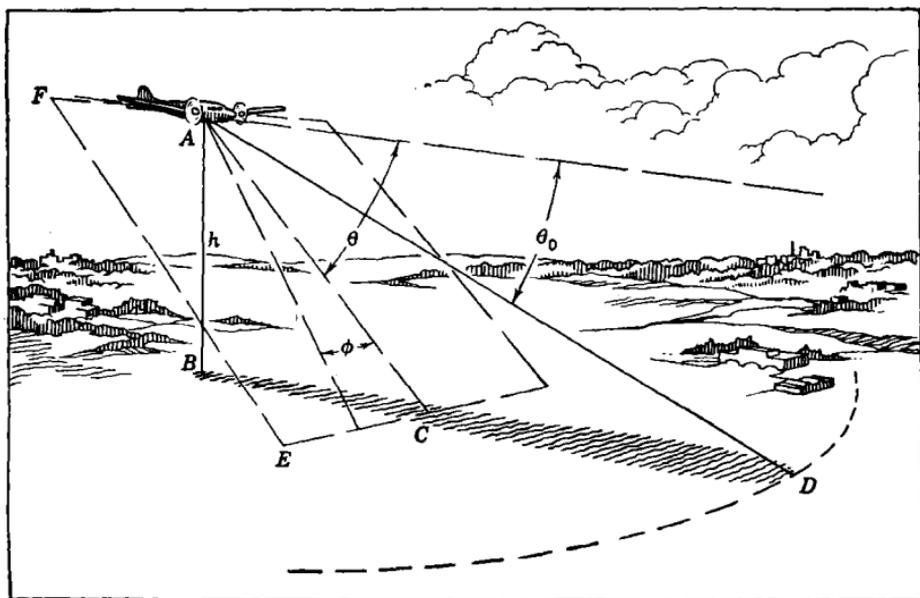


FIG. 2-2.—The problem of the ground-mapping antenna.

which contribute to the echo received at any instant. It is easy to see that the area of the patch on the ground, such as C in Fig. 2-2, from which echoes can be received at one instant is proportional to the width of the radar beam at the range in question and to the pulse duration. Also, the effectiveness of such a patch undoubtedly depends on the angle from which it is viewed. A detailed discussion of this matter must be reserved for Chap. 3, but the foregoing remarks should make plausible the following hypothesis:

$$\sigma = LR\alpha \sin \theta, \quad (6a)$$

in which α is the beamwidth in azimuth and L is a factor having the dimensions of a length that contains the pulse length and otherwise depends only on the characteristics of the terrain. The factor $\sin \theta$

expresses well enough for our purposes the effect of viewing the echoing area obliquely, at least when θ is small. (In Fig. 2-2 the vertical angles have been exaggerated, for clarity.)

Substitution of this expression for σ in the radar equation, and the replacement of R by $h/\sin \theta$, h being the height of the aircraft, leads to

$$S = P \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} = \frac{P \lambda^2 L \alpha G^2 \sin^4 \theta}{(4\pi)^3 h^3}. \quad (6b)$$

If S is to be independent of θ between $\theta = \theta_0$ and $\theta = \pi/2$, we must require that $G(\theta)$ vary as $\csc^2 \theta$ through this angular interval. The ideal antenna pattern then would be described by

$$\left. \begin{aligned} G &= G_0 \frac{\csc^2 \theta}{\csc^2 \theta_0}, & \text{for } \theta_0 < \theta < \pi/2, \\ G &= 0, & \text{at all other angles.} \end{aligned} \right\} \quad (7)$$

The requirement of Eq. (7) imposes a restriction on the maximum gain G_0 which can be achieved and hence on the maximum range. To see this, let us compute the average of the gain of the antenna over all directions, or $1/4\pi \int \int G d\omega$ extended over the whole sphere, where $d\omega$ is the element of solid angle. By the definition of gain this integral must be equal to 1. But first something must be said about the shape of the antenna pattern in the other plane—that is, in a plane normal to the fan beam. Suppose that the horizontal width of the antenna aperture has been fixed by other design considerations at the value a . The maximum gain possible will be obtained if the illumination of the antenna aperture is uniform horizontally, and in this case it can be shown that in any plane such as *ACEF*, Fig. 2-2, the gain as a function of the angle ϕ will be¹

$$G_\phi = G_{(\phi=0)} \frac{\sin^2 \left(\frac{\pi a \phi}{\lambda} \right)}{\left(\frac{\pi a \phi}{\lambda} \right)^2}. \quad (8)$$

As a function of θ and ϕ then, the pattern is completely described by

$$G = G_0 \frac{\csc^2 \theta}{\csc^2 \theta_0} \frac{\sin^2 \left(\frac{\pi a \phi}{\lambda} \right)}{\left(\frac{\pi a \phi}{\lambda} \right)^2}, \quad (9)$$

and we have to require $\frac{1}{4\pi} \int_0^{2\pi} \int_{-\pi/2}^{+\pi/2} G \cos \phi d\phi d\theta = 1$. Since G is very

¹ This expression will be recognized as the diffraction pattern of a rectangular aperture. See Vol. 12, Sec. 4-9.

small except for small values of ϕ , and since G vanishes for values of θ outside the interval θ_0 to $\pi/2$, it is permissible to write and evaluate the integral as follows;

$$\frac{1}{4\pi} \frac{G_0}{\csc^2 \theta_0} \int_{\theta_0}^{\pi/2} \int_0^{\infty} \csc^2 \theta \frac{\sin^2 \left(\frac{\pi a \phi}{\lambda} \right)}{\left(\frac{\pi a \phi}{\lambda} \right)^2} d\phi d\theta = \frac{G_0 \lambda}{4\pi a} \cos \theta_0 \sin \theta_0 = 1. \quad (10)$$

Hence

$$G_0 = \frac{4\pi a}{\lambda} \sec \theta_0 \csc \theta_0,$$

or from Eq. (7)

$$G = \frac{4\pi a}{\lambda} \tan \theta_0 \csc^2 \theta. \quad (11)$$

Returning now to Eq. (6b), if G is replaced by the right-hand side of Eq. (11), and α replaced by λ/a , we obtain, for S ,

$$S = \frac{PLa\lambda \tan^2 \theta_0}{4\pi h^3} \quad (12)$$

When θ_0 is small, as is usually the case, it is permissible to replace $h/\tan \theta_0$ by R_{\max} , the maximum range, which leads to the final relation

$$S = \frac{PLa\lambda}{4\pi h R_{\max}^2}. \quad (13)$$

The appearance of λ in this formula, which is to be contrasted with the result obtained for the simple fan beam and point target, Eq. (5), can be traced to the influence of the horizontal beamwidth upon the effective cross section of the extended target. It will be observed that once the other system parameters, P , S_{\min} , and a , are specified, the quantity hR^2 is fixed. That is to say, the maximum range obtainable is inversely proportional to the square root of the height of the aircraft, keeping everything about the radar set constant but the vertical radiation pattern of the antenna, which we assume to be adjusted to optimum shape for each height.

A problem related to the preceding one is met in the design of ground-based air-search radar, which may be required to provide uniform coverage at all ranges for point targets (aircraft) flying at some limiting altitude h . Here, however, σ is generally assumed to be constant: the reader will easily verify that this assumption leads again to the requirement that the gain vary as $\csc^2 \theta$, but with a *different* final result for the dependence of S upon h , R , and a . It turns out in fact that for the point target with σ independent of angle, the quantity hR_{\max} is constant, rather than hR_{\max}^2 , and S is proportional to a^2 rather than to $a\lambda$.

In practice it is not feasible to produce a pattern which exactly meets the specification of Eq. (7) but a reasonable approximation has been achieved in several instances. A comprehensive discussion of the problem of the "cosecant-squared" antenna, as this type has been called, will be found in Chap. 14 of Vol. 12.

2.6. The Beacon Equation.—So far we have confined our attention to the "two-way" radar problem, in which the route from transmitter to receiver is a complete round trip involving scattering of the energy by some remote object. Radar beacons operate on a different principle. A receiving antenna at a remote point picks up directly energy sent out from the radar transmitter. The signal is amplified and enabled to initiate the transmission of another signal, in reply, by an associated transmitter. This signal, received back at the radar, provides an artificial echo which can be utilized in various ways (Chap. 8). Here the analogue of the radar equation for radar-beacon operation will be discussed.

It is clear that we have to do with two entirely independent processes, each of which consists simply of one-way transmission and reception. Consider the first process, usually called "interrogation." Let P_r be the power transmitted by the radar, S_b the signal power received by the beacon antenna. These must be related by

$$S_b = P_r \left(\frac{G_r G_b \lambda^2}{16\pi^2 R^2} \right), \quad (14)$$

where G_r and G_b are the gains of the radar and beacon antennas respectively. The subsequent process of beacon reply is described by a similar equation with the subscripts b and r interchanged throughout. If the same antenna or similar antennas are employed for transmission and reception at the beacon and likewise at the radar, as is nearly always the case, the quantity in the parenthesis has the same value for interrogation and reply, and we can infer the corollary relation

$$\frac{S_b}{S_r} = \frac{P_r}{P_b}. \quad (15)$$

In practice, the gain G_b of the beacon antenna is fixed at a relatively low number by the requirement of something like omnidirectional coverage. The remaining factor in the parenthesis in Eq. (14) is the quantity $G_r \lambda^2$ which is proportional to the aperture of the radar antenna. This leads one to suspect that long beacon ranges should be, in general, more difficult to achieve at shorter wavelengths. Antenna apertures are rarely increased in area when shorter wavelengths are employed: on the other hand, available r-f power generally decreases markedly with decreasing wavelength. Actually this has not proved to be a serious

limitation down to wavelengths of the order of 3 cm, because ranges that can be achieved are already very great and are limited usually by the horizon rather than by the relation expressed in Eq. (14). That is to say, the condition of free-space propagation, which we have assumed in this section, often does not apply at the extremity of the microwave-beacon range.

Long beacon ranges are, of course, a result of the enormous advantage of one-way over two-way transmission. If we compare Eq. (14) with Eq. (4), we see that in signal strength for given transmitted power the beacon process enjoys the advantage of a factor of $4\pi G_b R^2 / G_r \sigma$. It is interesting to compute the limiting free-space range for a beacon operating in conjunction with a radar set of the characteristics assumed in our earlier radar example. Let us suppose that the gain of the beacon antenna is 10, and that the transmitted power and minimum required signal power for the beacon are the same as those assumed for the radar set. Using Eq. (14) we obtain for the maximum range, either for interrogation or reply, 60,000 statute miles.

In conclusion, we may point out that Eq. (14), although it has been written in notation appropriate to the radar-beacon problem, applies, of course, to any one-way transmission problem where free-space propagation can be assumed. Applications of Eq. (14) are to be found in the fields of microwave radar relay, radar jamming, and microwave communication.

THE MINIMUM DETECTABLE SIGNAL

2-7. Noise.—It is well known that despite our ability to amplify a feeble electrical signal by practically any desired factor, it is still not possible to discern an arbitrarily weak signal because of the presence of random electrical fluctuations, or "noise." If the true signal entering any receiver is made weaker and weaker, it subsides eventually into the fluctuating background of noise and is lost. What is the origin of these fluctuations, and what factors determine precisely the level at which the radar signal is hopelessly obscured by them?

Before we attempt to answer these decisive questions, it is worth while to consider briefly the limit of useful sensitivity of an ordinary low-frequency radio receiver. This limit is also set by random disturbances, but in this case the largest random disturbances with which the signal must compete originate generally not in the receiver itself but elsewhere in space. Whatever their source—and this may range from a passing trolley car to the mysterious reaches of interstellar space—these disturbances enter the receiver by way of the antenna. The crucial quantity is therefore the ratio of the field strength of the signal in the neighborhood of the antenna to that of noise or interference.

The absolute magnitudes of signal and interference power available at the antenna terminals are of little importance; only their ratio, which, for example, might be favorably altered by the use of a directional antenna pattern, determines the ultimate performance. This circumstance is perhaps one reason why the effective receiving cross section of an ordinary radio antenna is a number of little interest and is indeed rarely discussed in radio engineering texts. More significantly, it explains why the emphasis in the development of radio receivers has been mainly on improving discrimination against some of the external noise (for example, by greater frequency selectivity) rather than on reduction of noise inherent in the receiver.

The situation is different in the microwave region. Substantially all the noise originates in the receiver itself, not because microwave receivers are noisier or more imperfect receivers than low-frequency receivers—for they are not, as a rule¹—but because noise and interference originating outside the receiver are enormously greater at the lower frequencies. In fact, such noise in the microwave region is almost wholly negligible in existing receivers.² It is the noise that originates in the receiver itself against which the signal power determined by Eq. (3) must compete for recognition.

The level of fluctuations in an otherwise quiescent electrical circuit can be described in many equivalent ways. One description of which we shall make use is the following. Let R be the real part of the impedance measured, at frequency f , between two points in a passive electrical network, all dissipative parts of which are in thermal equilibrium at absolute temperature T . Then let the voltage e between these two points be measured by an ideal voltmeter which is capable of indicating, without disturbing the circuit in any way, the time average of the square of e , denoted by \bar{e}^2 , over a small frequency range \mathfrak{B} ,³ in the neighborhood of the frequency f . It will be found that \bar{e}^2 so defined is given by

$$\bar{e}^2 = 4RkT\mathfrak{B}, \quad (16)$$

where k is Boltzmann's constant (1.38×10^{-23} joules/degree).

¹ Very few low-frequency receivers approach the best modern microwave receivers in respect to signal-to-noise ratio, despite the fact that it is easier to attain a low noise figure at low frequencies, other things being equal.

² Future large-scale exploitation of the microwave bands may change the picture, although not, perhaps, as much as one might at first suppose. Both the optical nature of microwave propagation and the vast frequency band available in the microwave region of the spectrum mitigate mutual interference.

³ We use \mathfrak{B} for bandwidth throughout, for uniformity, although it is common practice, and in some respects preferable, to write Eq. (16) in differential form $d\bar{e}^2 = 4RkT df$. Where a finite bandwidth is implied, it will suffice for our purposes in this chapter to suppose that the voltmeter or amplifier in question has a rectangular pass band of width \mathfrak{B} cps.

Another description that is often useful involves the notion of "available power." Again, let some passive circuit, which we shall call C , be in thermal equilibrium at the temperature T . If now we connect across the terminals of C an impedance element Z , itself at absolute zero, the maximum rate of transfer of energy from C to Z within the frequency band \mathfrak{B} is just $kT\mathfrak{B}$. This rate, which is actually attained when Z is adjusted to be the complex conjugate of the impedance observed between the terminals of C , we call the "available noise power" from C . If Z is not at absolute zero, the transfer is not wholly one-sided; and if Z is at the same temperature as C , the *net* transfer vanishes, as of course it must.

This law, which has been derived in various forms by Nyquist¹ and others, is based on the equipartition law of statistical mechanics and not on any special assumption as to the details of the mechanism responsible for the observed fluctuations.

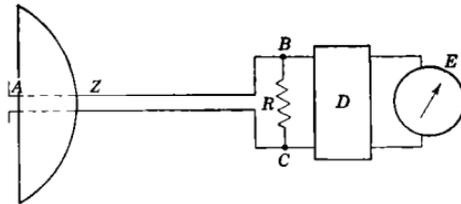


Fig. 2·3.—Schematic diagram of an ideal receiving system.

Other sources of noise are to be found in receivers—notably shot noise in vacuum tubes—which do depend on a special mechanism. These we do not intend to examine here since our purposes are first to demonstrate the fundamental limitations of an ideal receiver, no matter how constituted, and second, to study the factors that influence the distinguishability of the signal in noise of a specified relative level. In the latter problem it is allowable to lump together thermal noise and noise from other sources within the receiver only because these other types of noise share with thermal noise the bandwidth factor \mathfrak{B} , and give rise, at all frequencies of interest to us, to signals of precisely the same character.

2·8. Receivers, Ideal and Real.—Returning to the original problem, let us consider the rudimentary receiver shown in Fig. 2·3.

For simplicity, we shall assume that the antenna A constitutes a matched load on the transmission line of characteristic impedance Z , the latter being terminated at the other end in the resistor BC , whose resistance R need not be equal to Z . The box D is an ideal amplifier of bandwidth \mathfrak{B} , equipped with an output meter E , whose reading is proportional to the mean square voltage $\overline{e^2}$ across BC . The bar as before denotes a time average, or, more explicitly, an average over a time long compared with $1/\mathfrak{B}$.

¹ H. Nyquist, *Phys. Rev.* **34**, 110 (1928). See also Vol. 24 of this series, Chap. 4.

We are interested in the magnitude of \bar{e}^2 in the absence of signal, which we denote by \bar{e}_n^2 . In order to apply Eq. (16) we must know the resistance R_{BC} between B and C . This consists of R in parallel with the impedance of the line, Z , or what amounts to the same thing in this case, in parallel with the radiation resistance of the antenna. That is,

$$R_{BC} = \frac{RZ}{R + Z} \quad (17)$$

What temperature is to be associated with the resistance Z ? Or to put a more restricted question, under what circumstances can we say that R and Z are at the same temperature? Clearly this is the case if all surroundings of the antenna with which the antenna is capable of exchanging energy are at the same temperature T as is R . Let us assume for the moment that this is so. Then

$$\bar{e}_n^2 = 4kT\mathcal{B}. \quad (18)$$

Now let a signal be received by the antenna, and let S be the available signal power—that is, the power which would be absorbed by a matched load ($R = Z$) at BC . In general, the mean-square signal voltage \bar{e}_s^2 developed across R will be

$$\bar{e}_s^2 = S \frac{4R^2Z}{(R + Z)^2}, \quad (19)$$

from which we have

$$\frac{\bar{e}_s^2}{\bar{e}_n^2} = \frac{S}{kT\mathcal{B}} \cdot \frac{R}{R + Z}. \quad (20)$$

From Eq. (20) we should infer that the most favorable condition possible is attained by making R infinite, that is, by terminating the antenna line in an open circuit. In this special case, $\bar{e}_s^2/\bar{e}_n^2 = S/kT\mathcal{B}$. The physical interpretation is at once evident; in this case the noise originates entirely outside the system, arriving at the antenna as radiation.¹ The available noise power from the external region is $kT\mathcal{B}$, just as S is the available signal power.

In practice, however, the antenna is usually *not* terminated in an open circuit, for reasons which were ignored when we made our earlier assumptions about the ideal nature of the amplifier D . Without trying to do justice to this complicated and important problem we shall say only that when other sources of noise—in particular, noise in the output of the

¹ It is interesting to note that this result can be derived directly from the theory of black-body radiation by assuming (as we have tacitly done already) the antenna to be enclosed in a region all at temperature T , and making use of the receiving cross section of the antenna given by Eq. (2).

first amplifier stage—are taken into account, together with the fact that the input impedance of microwave receivers is not unlimited, the optimum value of R turns out to be finite. The condition $R = Z$ is, in fact, not unusual. This condition, in our ideal system of Fig. 2-3, leads to

$$\frac{\bar{e}_s^2}{e_n^2} = \frac{S}{2kT\mathcal{B}}. \quad (21)$$

A question of fundamental interest, though not as yet of much practical moment, is the one raised earlier about the temperature to be assigned to the antenna radiation resistance. This is certainly *not* in general the temperature of the metal parts of the antenna. The reader may correctly surmise, from what has been said, that the effective temperature of this element of the circuit is that of any surroundings of the antenna with which the antenna can exchange energy by radiation—that is to say, it is the temperature (or a suitable average of the temperatures) of whatever would absorb energy radiated by the antenna as a transmitter. This has indeed been demonstrated experimentally. It has been shown¹ that a microwave antenna pointed at the sky receives only a very small amount of radiation, corresponding to an absolute temperature of at most a few degrees. If our receivers were very nearly ideal this would have the practical result of making it much easier to detect aircraft appearing at high angles of elevation. In the best existing receivers the reduction in noise output which could be obtained by pointing the antenna upward would amount to some 10 per cent.

The foregoing somewhat academic discussion of thermal noise would be inappropriate in this place were it not for two facts. First, microwave receivers, even as this is written, have been brought so close to the pinnacle of ideal performance that it is well for the radar engineer to appreciate the nearness, and the finality, of the goal. Second, thermal noise, although it is not wholly to blame for the noise background in microwave receivers, provides a very convenient standard in terms of which the performance of an actual receiver can be specified.

We shall define, as the *over-all noise figure* N of a receiver, the ratio of signal power available from the antenna to $kT\mathcal{B}$,² when the mean noise power and the signal power are equal as observed at some stage in the receiver where both have been amplified so highly as to override completely any noise introduced by succeeding stages. In framing the definition so broadly we have in effect included, under “receiver,” not only the mixer but all associated r-f circuits. It is well to do so at this stage, for the analysis of the contribution of each part of the input system to the over-

¹ R. H. Dicke, “The Measurement of Thermal Radiation at Microwave Frequencies,” RL Report No. 787. See also Dicke *et al.*, *Phys. Rev.*, **70**, 340 (1946).

² In this definition T is customarily, and arbitrarily, taken to be 291°K.

all noise figure is complicated by many subtle questions which should not now concern us. By our definition the noise figure N of the ideal system of Fig. 2-3 is just 2, if the *antenna temperature* and the temperature of R are both 291°K. For the open circuit termination, $N = 1$. Actually, over-all noise figures of 10 or lower are not now uncommon in the best microwave receivers. A noise figure of 10 and a bandwidth of 3 Mc/sec, for example, imply that a signal of 1.2×10^{-13} watts will be sufficient to increase the receiver output by an amount equal to the average noise output.

2-9. Receiver Bandwidth and Pulse Energy.—In all the expressions for noise power there appears the width of the amplifier pass band, \mathcal{B} . Apart from certain technical limitations, this factor is completely at our disposal, and we must now decide how its value shall be chosen. To do this we must for the first time in this chapter admit that we are concerned with the detection of *pulses*. Suppose that the r-f signal pulse consists of a wave train with a rectangular envelope, of duration τ seconds. The response of an amplifier (or filter) to such a pulse depends on the width of its pass band, which is in effect an inverse measure of the time of response of the amplifier. If this response time, in order of magnitude $1/\mathcal{B}$, is very much longer than τ , the peak output signal power will be proportional to the square of the bandwidth \mathcal{B}^2 . On the other hand, if the response of the amplifier is relatively very fast, the output will quickly reach a level determined solely by the input power and the gain, and independent of \mathcal{B} . Thus if we plot peak output signal power against \mathcal{B} , for constant (during the pulse) input power a curve like a of Fig. 2-4 will be obtained.

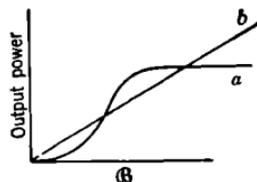


FIG. 2-4.—Variation of output signal power a and output noise power b with bandwidth at constant pulse duration.

The average noise power, however, is just proportional to \mathcal{B} according to Eq. (16) and is represented by the line b of Fig. 2-4. We must therefore expect a *minimum* in the *ratio* of noise output to peak signal output, at some value of \mathcal{B} closely related to $1/\tau$. It is not obvious that such a minimum should represent precisely the best condition for the detection of the signal. Indeed it would be surprising if any one criterion, appropriate to all of the various methods of final detection that can be used, could be found. This need not discourage us as it turns out that all such criteria lead to similar results, and that moreover the minimum is fairly flat.

Some results obtained by the group headed by J. L. Lawson in their investigation of signal discernibility are displayed in Fig. 2-5. A simple range time base, or A-scope, was used; the video bandwidth was so wide (10 Mc/sec) as to have a negligible influence on the outcome, at least in

the neighborhood of the minimum. The signal threshold power plotted in Fig. 2-5 is essentially the signal strength, relative to average noise power, for which an observer would identify the signal correctly nine out of ten times. For a more precise definition of the experimental criterion for threshold signal in these tests, the reader is referred to Vol. 24, Sec. 8-2. Our interest just now is in the position of the minimum of each curve that is seen to lie near $\mathfrak{B} = 1.2/\tau$. However, departure from this

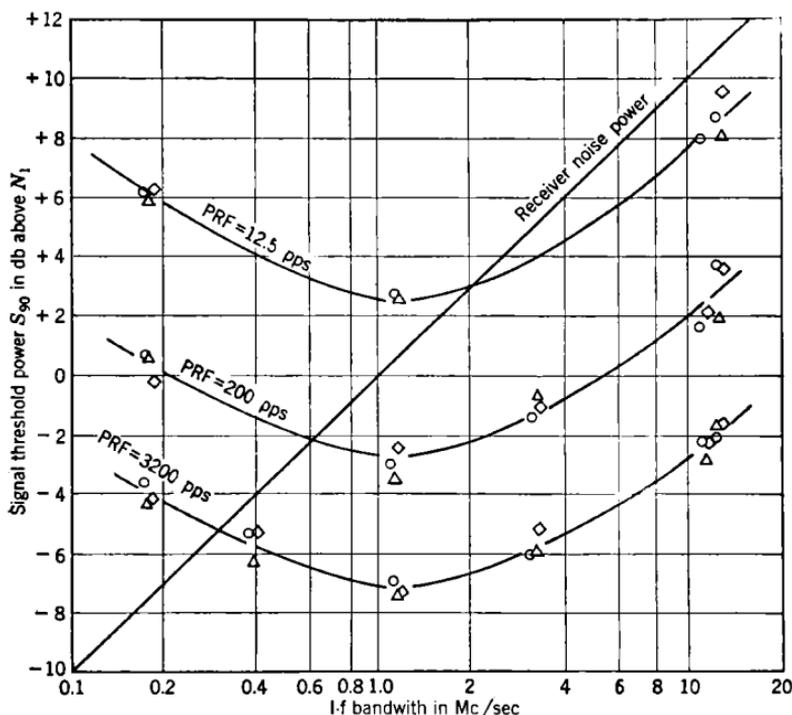


FIG. 2-5.—Signal threshold vs. i-f bandwidth for a pulse duration of $1 \mu\text{sec}$. The signal power is measured in units of the noise power, N_1 , within a band 1 Mc/sec wide. In these experiments the video bandwidth was 10 Mc/sec, the signal presentation time was 3 sec, and the length of the pulse on the screen of the A-scope was 1.7 mm.

value by a factor of 2 in either direction increases the minimum discernible signal power by less than 1 db. The radar designer is inclined to take advantage of this latitude to set \mathfrak{B} somewhat greater than $1.2/\tau$ as this eases the requirement of accurate frequency control. A value of \mathfrak{B} in the neighborhood of $2/\tau$ is typical of present practice; there is some evidence also from controlled experiments on intensity-modulated indicators (Vol. 24, Chap. 9) which favors this higher value.

The quantity $S/kT\mathfrak{B}$, which has been our primary concern in this section, can now be rewritten by expressing \mathfrak{B} in terms of τ . If, for simplicity, we require $\mathfrak{B} = 1/\tau$ then $S/kT\mathfrak{B}$ becomes $S\tau/kT$. The prod-

uct of the received signal power S and the pulse duration τ is just the energy contained in the received signal pulse, while kT is the average stored energy in a quiescent system with one degree of freedom. That the ultimate criterion of detectability should involve a comparison of these two quantities is certainly not surprising. Thus it is always possible to rewrite the radar equation, Eq. (4b), in which S_{\min} will always involve $kT\mathcal{B}$, in a form containing, explicitly, not the transmitted power P , but the energy emitted per pulse, $P\tau$. For example, if the pulse duration τ is increased by a factor of 4, peak pulse power being maintained constant, a 40 per cent increase in maximum range would be expected. This can be regarded as the result of increasing the total energy in the transmitted pulse. The pulse energy is a more significant measure of the usefulness of a transmitter in long-range detection than is the pulse power, in so far as the choice of pulse length and bandwidth is not restricted by other considerations.

2-10. The Statistical Problem.—It might seem that we now have within our grasp a specification for S_{\min} which can be introduced into the radar equation to yield the maximum range of detection, R_{\max} . Unfortunately, it is not so simple as that. Let us summarize what we *do* know, once we are provided with the over-all noise figure and bandwidth of the receiver, the transmitted power, and the geometrical factors in the radar equation which concern the antenna and the target. We know the ratio of the amplified signal power to the average value of the amplified noise power. We are *not* yet able to say how large this ratio must be before the signal can be identified with reasonable certainty. The root of the difficulty is that we have to do with a statistical problem, a game of chance. The answers will be given as probabilities, and will depend upon many features of the system by which the signal is presented to the observer, as well as upon the precise description of the “reasonable certainty” mentioned above. We approach the problem by examining first the character of the noise fluctuations.

The output noise power will fluctuate continually and irregularly above and below its *average* value, the latter being understood as an average taken over a time long enough to include very many fluctuations. An example that is easy to discuss quantitatively is the following. Suppose that the intermediate frequency amplifier is followed by a square-law detector and by further (video) amplification in stages whose pass band is so broad as to introduce no distortion of the signal. The final output voltage is then a direct measure of the instantaneous power output of the i-f amplifier. Let this voltage be displayed, as a function of time, on an oscilloscope. Figure 2-6a shows a typical single trace which might be obtained in this way. Now in spite of the wildly irregular nature of the noise voltage, the theory of random processes applied to this case

(Vol. 24, Chap. 8) leads to certain positive statements. First, the rapidity of the fluctuations is determined by the bandwidth \mathfrak{B} of the amplifier. That is to say, in a time short compared with $1/\mathfrak{B}$, it is extremely unlikely that the output power will change noticeably.¹ On the other hand, during a time long compared with $1/\mathfrak{B}$, fluctuations will almost certainly occur, and the value of the power P at the end of such a long time will bear no systematic relation at all to the value it happened to have at the beginning. This is a roundabout way of saying that values of P determined at times differing by much more than $1/\mathfrak{B}$ are statistically independent, or, in other words, $1/\mathfrak{B}$ is a measure of the correlation time of the fluctuations.

The second statement which can be made concerns the probability that at some arbitrarily selected instant the output power will be found to lie between some specified level, P , and $P + dP$. This probability, which we shall label $W_1(P) dP$, is given by

$$W_1(P) dP = \frac{1}{P_0} e^{-\frac{P}{P_0}} dP. \quad (22)$$

P_0 is the average power, determined over a long time.

A statement which is easily seen to be equivalent to this is: The probability that the power *exceeds* some specified level P , at an arbitrarily selected instant, is just $e^{-\frac{P}{P_0}}$. Thus there is always a finite chance of getting a high noise peak. For example, the probability that at a given time P is greater than $5P_0$ is e^{-5} or 0.0068. For the discussion that follows it is convenient to simplify the problem somewhat by dividing the time base into discrete intervals each $1/\mathfrak{B}$ long. In Fig. 2-6a we would have 50 such intervals. The essential features of the noise background can then be described by regarding these intervals as independent² and associating some one value of P with each. Again Eq. (22) correctly expresses the probability that the power, in an arbitrary interval, will lie between P and $P + dP$.

The task of detecting a signal amid the noise in this simplified case amounts to selecting an interval which displays so large a value of P that one is justified in *betting* that the peak was due to a combination of signal and noise, and not to noise alone. This should dispose of any hope that we shall be able to define once and for all the minimum detectable

¹ The intermediate frequency itself is assumed to be high compared with \mathfrak{B} , so that it is permissible to speak of the instantaneous power while actually meaning the power averaged over one cycle of the intermediate frequency.

² The intervals so defined are actually not entirely independent, for the output power is, after all, a continuous function of the time. We are here substituting for a continuous random process a discrete random process, which is easier to discuss in elementary terms.

signal S_{\min} . One can never be *absolutely* sure that any observed peak is not due to a chance noise fluctuation, and one cannot even say how *probable* it is that the peak is not due to noise, unless one knows how probable it is, a priori, that the peak is due to something else—namely, signal plus noise. Knowledge of the a priori probability of the presence of signal is possible in controlled experiments such as those described in Vol. 24, Chap. 8. For example, it might be arranged to present a signal at a given point on the oscilloscope sweep about half the time, the observer being required to judge whether or not the signal was there. A shrewd observer would call "Signal" whenever the scope deflection at this point was only a little greater than average, and could score a pretty good "batting average" even for rather weak signals. If, however, he expected the signal to occur only very rarely, he would naturally set his standard higher, in order to avoid frequent false identification of what were, in fact, noise fluctuations.

For obvious reasons, no such clear delineation of the statistical problem is possible in the normal operation of a radar set. However, we can roughly distinguish different orders of what was called above the a priori probability of the signal. Compare, for example, the use of a radar set for early warning (detection of enemy aircraft at a great distance) and for ground control of aircraft, in which a selected plane, already picked up on the radar, is followed, its position being noted in each scan. In the latter case, we know about where to look for the echo in a given scan, having observed the position in the previous scan and knowing approximately the course and speed from still earlier information. Hence, we accept even a weak indication of the presence of signal if it comes near the expected place, with comparatively small chance of error. In the former case, however, a large area of the radar scope must be kept under surveillance for a long time. The likelihood that a high noise peak will be observed is much greater: only a very pronounced peak can now be confidently identified as a new signal.

We have divided our time base into intervals $1/\mathcal{B}$ long. In one second then, \mathcal{B} intervals come under observation. If \mathcal{B} is 2 Mc/sec, for example, this amounts to 2,000,000 intervals per sec. According to Eq. (22) we expect to find among these some hundred that represent noise peaks greater than $10P_0$. In one hour there is better than an even chance of observing one noise peak greater than $20P_0$. Must we therefore require that to be detectable a single signal pulse be at least greater than $20P_0$, or at the terminals of the antenna, greater than $20NkT\mathcal{B}$?

We should indeed, if we are speaking of a *single* signal pulse. But anyone who has had anything to do with the design or operation of a radar set will perceive at once that an essential part of the problem has so far been ignored. The fact is that detection of a target by radar

practically never depends on the detection of a lone signal pulse, but on reception, from the same range and bearing, of repeated signal pulses. This has a profound influence on the statistical problem that we have been discussing because, whereas the noise fluctuations during successive sweeps¹ are completely independent, the signal can be made to appear at the same position in each sweep. To take advantage of this essential difference between signal and noise it is evidently necessary to resort to some method of information storage, or integration, which would make use of the information contained in several sweeps. It might be possible,

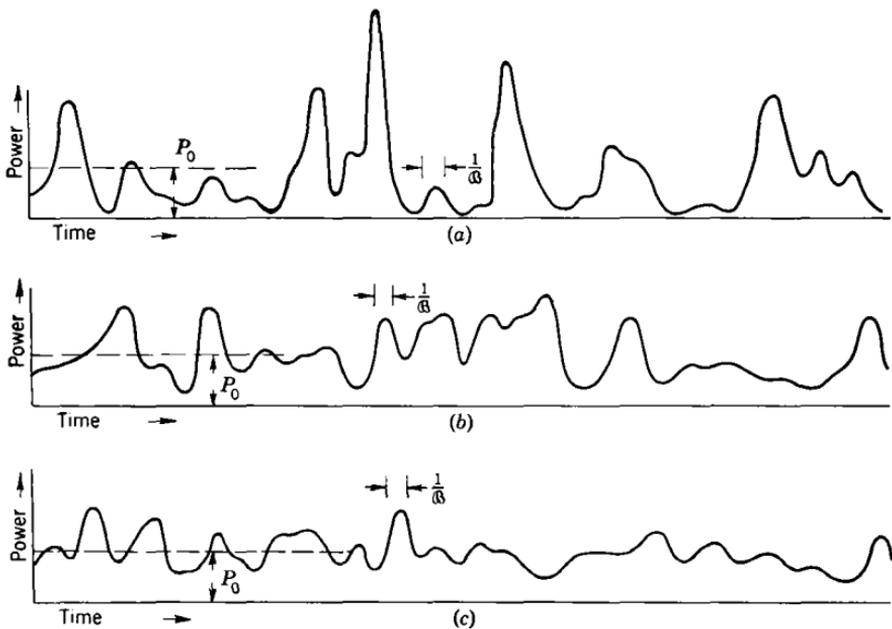


FIG. 2-6.—Output noise power: (a) single sweep; (b) average of two sweeps; (c) average of four sweeps.

for example, to average over a number of sweeps—that is, to present to the observer a single trace whose height at any time would represent the average of the power output which was obtained in each of the several sweeps at that same time (measured from the transmitted pulse). Now this is substantially what many radar indicators do; it is so nearly what they do, in fact, that quantitative conclusions drawn from a study of the ideal process outlined are of practical significance and are surprisingly well borne out by actual observations. Therefore, we want to inquire more carefully into the simple averaging or integration process.

¹ We use the term “sweep” to mean one entire interval between successive pulse transmissions. Such a term unavoidably suggests a particular type of display, but a more general and formal terminology seems too cumbersome.

Let us turn to Fig. 2-6*a*, which represents a single oscilloscope trace starting at the time of the transmitted pulse or at some known time thereafter. This is a typical single trace showing noise only, without signal.¹ Clearly only a rather strong signal could stand out conspicuously among these fluctuations. Now examine Fig. 2-6*b* in which the average of two noise traces is plotted, each point in one trace being averaged with the corresponding point in the other. The fluctuations are here less violent. We note a scarcity of excessively high peaks: this is not surprising as it is unlikely that very high peaks, being in any case infrequent events, should occur at the same place on two original traces. In Fig. 2-6*c*, four sweeps with noise have been averaged, and we notice a further reduction in the magnitude of the fluctuations. As the size of a *signal peak* (which means, really, *signal-plus-noise*) will not be essentially altered by the averaging provided the signal occurs at the same place on successive sweeps, it will be easier to detect a small signal on the trace *c* than on *a*. More than this we cannot say without elaborate analysis of the statistical problem. Such analysis can be carried through by well-known methods (see Chap. 8, Vol. 24) and one can obtain, for example, the results displayed in Table 2-1 below.

TABLE 2-1.—PROBABILITY OF NOISE EXCEEDING SIGNAL-PLUS-NOISE

n	$w = 10^{-1}$	$w = 10^{-2}$	$w = 10^{-3}$	$w = 10^{-4}$	$w = 10^{-5}$	$w = 10^{-6}$
1	5.0	8.9	10.9	12.3	13.3	14.2
2	3.0	6.7	8.6	9.9	10.8	11.6
4	1.0	4.6	6.5	7.7	8.6	9.3
8	-0.8	2.7	4.5	5.7	6.5	7.3
16	-2.5	0.8	2.5	3.7	4.5	5.2
32	-4.2	-1.1	0.5	1.6	2.5	3.2
64	-5.9	-2.9	-1.4	-0.4	0.4	1.1

The entries in Table 2-1 are the strength of a signal, in decibels, relative to the average noise power P_0 . The number n , at the left, is the number of sweeps averaged—for example, 4 in the case of Fig. 2-6*c*. The quantity w at the head of each column is the probability that an arbitrary interval along the final averaged sweep will display a larger deflection than the interval that contains the signal. For example, if the signal is 4.6 db above P_0 (third row, second column) and the integration has been carried over four sweeps, there is one chance in a hundred that a selected interval will actually exceed in power the interval con-

¹ This trace was constructed, as were the others in Fig. 2-6, by random sampling of a number field prepared according to the appropriate distribution function, the sweep having been divided into intervals for this purpose. Again we are overlooking the difference between a discrete and a continuous random process.

taining the signal. If the sweep length were equivalent to, for example, 500 intervals (remember that what we call an "interval" is about $1/\beta$ long), one would expect to find several (roughly, five) noise peaks in the sweep which were higher than the signal peak. In this case we should certainly need either a stronger signal or more integration.

Now it is the *trend* of the numbers in Table 2-1 which is of interest to us, since any real radar problem will differ in many particulars from the ideal process to which Table 2-1 applies exactly. If we select any column ($w = \text{constant}$) we observe that the variation of signal power with n is something intermediate between $1/n$ and $1/\sqrt{n}$. For large n and particularly for w not too small, the variation is not far from $1/\sqrt{n}$. This implies that doubling the number of sweeps integrated, other things being equal, allows the signal power to be reduced by a factor $1/\sqrt{2}$. This relation has been strikingly verified by the experiments of J. L. Lawson and others (Vol. 24) in the detection of signals on the A-scope, under conditions where n was large. On the other hand, for small n , and especially for very small values of w , Table 2-1 would require the signal power to vary more nearly as $1/n$. This is not too important in practical radar design, for so many factors are involved in the real problem that we cannot hope for, and do not need, a very precise answer to such questions. We shall most frequently assume, in later chapters, that the required signal power varies as $1/\sqrt{n}$.

It must be observed, to put the above considerations in proper perspective, that the benefits of integration are not confined to the smoothing out of thermal or purely random noise. A very important requirement, in practice, is discrimination against isolated but powerful disturbances such as transients from nearby electrical apparatus, or, very commonly, pulse interference from other radar sets. Indeed, the frequency of such disturbances, in most locations, renders academic our earlier remarks about the likelihood of getting an abnormally high noise peak once an hour or so. The fact that the desired signal occurs repeatedly allows such isolated disturbances to be discarded easily, or disregarded. Strictly it is not *integration*, but a sort of *coincidence selection*, that is most potent against these scattered flashes of interference. Such selection is inherent to a greater or lesser degree in nearly all radar systems and as a rule very little selection suffices. In practice then, it is still the noise that we have to combat, not the interference, if we want to reduce the minimum detectable signal power.¹

¹ Because this is true, an essential part of any test of the condition or quality of a radar set includes a measurement of the noise level, or a quantity proportional thereto. The determination of minimum discernible signal power under any reproducible conditions of observation constitutes such a measurement.

2-11. Effect of Storage on Radar Performance.—How is the integration process, or its equivalent, actually carried out in a radar set? The earliest and simplest radar indicator is the A-scope, the horizontal range sweep with vertical deflection by the signal. Usually the screen of the cathode-ray tube does not give a persistent glow but the picture decays immediately. It is difficult to see any integrating or storage mechanism here, and indeed there is none. The storage is accomplished in the eye and mind of the observer and is rather astonishingly effective.¹ Just *how* it is done we do not know in detail, but we may see the results in Fig. 2-5 (page 35). Note now the significance of the three curves in that figure which correspond to different *pulse repetition frequencies* with the same total time of observation. An increase in PRF by a factor of 16, from one curve to the one below, reduces the signal power required by very nearly a factor of 4. It is natural to assume that, whatever the effective number of integrations, it is proportional to the PRF, for constant observation time, and we may regard the experiments described by Fig. 2-5 as evidence favoring the relation

$$S_{\min} \propto \frac{1}{\sqrt{n}} \quad (23)$$

In Fig. 2-7, which has been taken from Vol. 24 of this series, an attempt has been made to show directly the effect of integration in the A-scope. Here the camera has been substituted for the eye, and the number of sweeps overlapping in one picture has been varied. Signals of different strength are present at two positions in each sweep and by referring to the explanation accompanying the figure, the reader will be able to trace the improvement in discrimination from one picture to the next. The noise and signal shown, incidentally, were derived from a *linear* detector.

In the case of an intensity-modulated indicator with a persistent screen, what we have called "integration" is very nearly exactly that—namely, the repeated addition, and hence averaging, of a number of successive sweeps. The difference is that the *number* of sweeps so averaged is not sharply defined, since the luminosity of the screen decays gradually with time. However there will be some time which is characteristic of the decay, and we can say simply that if the electron beam traces repeatedly over the same strip of the screen, the number of sweeps averaged will be of the order of the pulse repetition frequency times this decay time. In practice, the radar is usually *scanning*. That is, before the decay time elapses, the radar beam moves off the target and the cathode-ray-tube trace moves to a new position on the screen. The number of

¹ Failure to appreciate the power of the observer to integrate was responsible for some disappointment with early electrical integrating devices, which, through no fault of their own, did not bring the expected improvement over the simple A-scope.

sweeps integrated then is limited by the number of pulses striking the target during the scan and not by the screen persistence. If, however, the latter is so long that storage is effective from one scan to the next, we

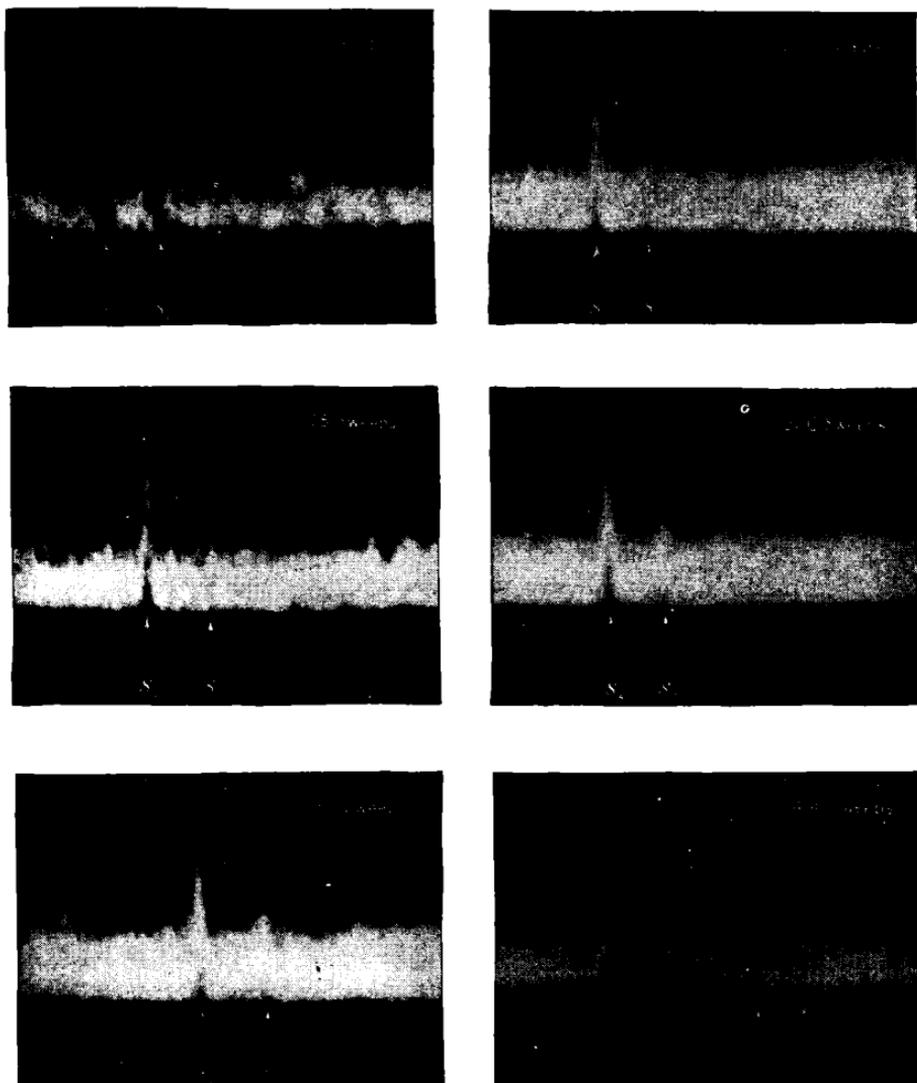


FIG. 2-7.—In this series of photographs, the number of A-scope sweeps recorded during each exposure was varied from 12 to 400. Artificial signals were injected at the locations marked, with the following strength, relative to average noise power: S_1 , +10 db; S_2 , +5 db; S_3 , 0 db; S_4 , -5 db.

must include more than one scan in estimating the amount of the integration, and the screen persistence again enters.

Let us try to elucidate this rather complicated state of affairs by

tracing through an example of the familiar plan-position indicator, PPI. Suppose that the effective storage time of the screen is 4 sec, and that the angular width of the radar beam is 6° : let the rate of rotation be variable. Assume a constant pulse repetition frequency of 500 pps. Let us begin with the antenna stationary, pointing at a single target. The number of sweeps available for integration, n_i , will be limited by the screen persistence and will be $4 \text{ sec} \times 500 \text{ sec}^{-1}$, or 2000. This will remain true, even though the antenna rotates, until the rate of rotation becomes so great that the beam dwells on the target less than 4 sec, in other words, up to $\frac{1}{4}$ rpm. With increasing speed of rotation, the number of sweeps integrated will decrease as $1/(\text{rpm})$ until the beam rotates faster than one revolution in 4 sec, at which time we begin to include more than one scan in the integration. Then n_i will begin to level off, its asymptotic value, for high speed, being given by

$$(\text{screen storage time}) \times (\text{PRF}) \times \frac{(\text{beam width})}{360^\circ}$$

or, in this case, $n_i = 33$ at speeds much greater than 15 rpm. Altogether n_i has changed by a factor of $\frac{1}{50}$ between low and high speed, or by just the ratio of the beamwidth to 360° .

If now we assume that the signal power required for detection, S_{\min} , is proportional to $1/\sqrt{n_i}$, a plot of S_{\min} against scanning speed will look like Fig. 2-8. To be sure, we have so drastically oversimplified the problems that no significance can be claimed for the exact shape of the curve. Nevertheless, the main features of the curve—especially the existence of the plateau regions I and III, whose vertical separation is determined by the ratio of the beamwidth to the total scan angle—are worthy of attention, and have been confirmed by controlled tests (Vol. 24, Sec. 9-1).

The effect of scan speed on threshold-signal level is often called "scanning loss," the condition of scan-speed zero being implied as the standard of reference. This condition merits no more than any other the distinction of being a standard. On the contrary, the factors determining S_{\min} are in this case especially elusive, if one wants to take them all into account. We prefer not to use the term "scanning loss," and if a basic standard of reference is deemed desirable we would choose, on logical grounds, the single A-scope trace, which contains all the information normally available from a single pulse transmission. With

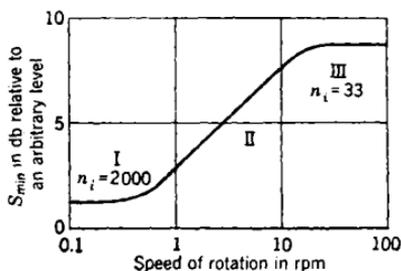


FIG. 2-8.—Effect of scanning speed on storage gain, assuming 6° beamwidth, and 4-sec storage time, with PRF = 500 pps.

respect to this standard all other modes of indication provide varying amounts of *storage gain*, by making use, more or less imperfectly to be sure, of information gathered on repeated pulse transmissions. Thus we say that Fig. 2-8 describes the effect of scanning on storage gain.

The number of other independent variables, yet unmentioned, which affect the amount of gain realized through storage is appallingly large. The more prominent of these must be discussed briefly, if only to suggest the scope of the general problem treated comprehensively in Vol. 24.

The *video bandwidth* b , as distinguished from the i-f bandwidth \mathcal{B} , has hitherto been assumed to be so great as to allow reproduction of the detected i-f signal without distortion. This is seldom true in practice and the question arises, how does the video bandwidth influence the signal discernibility, and how narrow can it be without harm? Making the video bandwidth b considerably less than $1/\tau$ amounts to putting the rectified signals and noise through a long-time-constant filter. This is a sort of *averaging*, or integration process, in effect, the final output at any instant being an average over *several* adjacent intervals, each of the order of τ in duration. It is clearly an integration process operating at a *loss* rather than at a *gain*, in contrast to the sweep-to-sweep integration discussed earlier, for it includes with the signal an unnecessarily large amount of noise. The result is an increase in signal power required for detection: for $b \ll 1/\tau$, S_{\min} is roughly proportional to $1/\sqrt{b}$. Of course such a decrease in video bandwidth is accompanied by a marked deterioration in the time discrimination, or *resolution in range*, of the system unless the range resolution is already limited by other factors. Because of this, and for other reasons as well, designers have usually aimed at adequately wide video bands.

Actually, the effect of video bandwidth narrowing would not deserve even this much discussion here, were it not that, in a less obvious way, a very similar effect arises from certain other influences, having to do with the *spot size* and the *sweep speed*. Let us consider an intensity-modulated indicator tube, such as the PPI. Let d be the diameter of the light spot due to the beam, and let v be the speed with which the beam is swept across the face of the tube. If $d/v < \tau$, the pulse duration, the *intervals* of which we spoke earlier are spread apart on the tube; if $d/v > \tau$, the spot overlaps several such intervals, or better, several such intervals contribute to the same spot of light on the tube. Now this is precisely what happened when the video bandwidth was made too narrow, and we must expect the same consequences. In effect, the intensity-modulated cathode-ray tube is a low-pass filter, whose bandwidth is of the order v/d . For example, a radial sweep covering 50 miles of range on a 7-in. PPI tube would be written in with a velocity of about 1.5×10^4 cm/sec. If d is 0.1 cm, the bandwidth of the system is about 150 kc/sec.

To use such a display for a 1- μ sec pulse would entail a penalty of about 2 db in S_{\min} .

An extreme, indeed a limiting, case of the spot-size sweep-speed effect is met in the type C indicator (Sec. 6-6) which presents only elevation and azimuth data, not range. Here the sweep speed v is zero or very nearly zero, the spot moving only as the antenna scans. All the noise and the signal, if any, from one range sweep are piled up on one spot. If T is the duration of such a sweep,¹ we should expect an increase in S_{\min} due to this superposition of noise, of the order of $\sqrt{T/\tau}$.

Video mixing, which is practiced in certain radar applications, consists in superimposing two radar pictures on the same screen. Because this results in adding noise from one picture onto a signal from the other, we would expect, and we find, an increase in S_{\min} . If two similar but independent² radar pictures are superimposed on one cathode-ray-tube screen, a signal on one of these pictures will have to be about 1.5 db, or $\sqrt{2}$ stronger to be detected than would have been necessary if the other picture were absent. A related problem is this: Suppose that on a single radar indicator we simply remove the signal from alternate sweeps (by holding off the transmitter pulse, for example). The resulting increase in S_{\min} will amount to about 3 db in this case, for what we have done is equivalent to the following *two* changes; (1) reduction in PRF by factor of $\frac{1}{2}$ with corresponding reduction in n_t ; (2) mixing in an equal number of foreign noise-bearing sweeps, as in the previous example.

Whether the *detector* is a square-law detector, a linear detector, or something else, will no doubt influence S_{\min} in any given case.³ However, we may as well lump together the effect of the detector, the cathode-ray-tube modulation characteristics, the screen characteristics, and those of the eye, and admit that no brief discussion of these factors is possible. In most cases one has to rely on experience; that is to say, one has to design by making relatively short extrapolations from previous practice, or by making preliminary tests under conditions approximating those selected.

¹ Because the C-scope was found useful only for targets at short range, it was the practice to apply a *range gate* covering only the useful range interval and suppressing the electron beam during the remainder of the period between pulses. In this case the time T represents the duration of the "on" part of the sweep.

² By "independent," it is meant that the signal from a radar target will appear on one of the pictures only, or at any rate, not at the same position on each. The superposition is usually accomplished by interlacing the sweeps.

³ In so far as the question of discernibility can be stated in terms of the probability of signal-plus-noise being larger than noise, the law of the detector, and indeed of any subsequent elements, can be eliminated from the *theoretical* problem. Table 2-1, for example, holds for any detector whose output is a function of the absolute value of the i-f voltage.

Although we have not arrived at a universal prescription for the minimum detectable signal power, we have studied the influence of various factors on S_{min} , and we are now in a position to predict the *relative* change in S_{min} which will accompany some proposed alteration in the radar system. One way to establish some absolute basis is to give the observed value of the ratio of minimum detectable signal power to average noise power, for one particular system. This is done in Table 2-2, in which are displayed the relevant constants of the system selected.

TABLE 2-2.—CHARACTERISTICS OF A SAMPLE SYSTEM

Type of indication.....	PPI; P7 screen; decay time about 7 sec	
I-f bandwidth.....	$B = 1.2$ Mc/sec	
Video bandwidth.....	$b = 5$ Mc/sec	
Pulse duration.....	$\tau = 1.0$ μ sec	
Pulse repetition frequency.....	320 pps	} Hence about 50 pulses on one point target per scan, with negligible scan-to-scan storage.
Scan rate.....	6 rpm	
Beamwidth.....	$\theta = 6^\circ$	} Note that $d/v > \tau$, which is not the best condition, but is a typical one.
Sweep speed.....	$v = 2 \times 10^4$ cm/sec	
Spot diameter.....	$d = 0.1$ cm	
Conditions of observation.....	Signal occurs on one scan, at known azimuth, in one of six range positions. Observer must locate it with 90% certainty. This corresponds to a value of w , in Table 2-1, of the order of 10^{-2} .	

Ratio of minimum detectable signal power to average noise power, under above conditions.. 1.25 (+1 db)

Working from Table 2-2, we can infer that if the system there described had an over-all noise figure N of, for example, 15, the signal power required at the terminals of the antenna would be $S_{min} = 1.25NkT(B) = (1.25)(15)(1.37 \times 10^{-23})(291)(1.2 \times 10^6) = 9 \times 10^{-14}$ watts. Now if for some other system N is given, together with the new values of the parameters listed in Table 2-2, the reader should be able to estimate S_{min} for that system. The estimate will be better, of course, the less extreme the departure from the conditions of Table 2-2. We must emphasize, however, the limited utility of a value of S_{min} obtained in this way. The experiments upon which Table 2-2 is based were performed in the laboratory under ideal conditions of observation; moreover, the value of w (the parameter of Table 2-1) is here certainly much too large for any radar search operation, although it might not be inappropriate for the continuous tracking of a target already detected.

The experimental determination of the maximum range of an actual radar set is as uncertain a matter as is the calculation of the maximum range by the method outlined in this chapter. Some arbitrary procedure for observation must be specified, and the limit of range must be defined

in a manner consistent with the inherently *statistical* nature of the problem. That R_{\max} is not rendered so uncertain by these statistical effects as to lose any usefulness is largely due to the inverse-fourth-power law expressed by the radar equation, which makes the range relatively insensitive to moderate changes in the system parameters.

It should not be forgotten that in these latter pages we have been wholly concerned with that obscure marginal region between strong signals and utterly undetectable signals. This margin is not, after all, very broad. A glance at Table 2-1, for instance, shows how very rapidly w diminishes as the signal becomes strong. Most of the useful signals in any radar set are strong enough to be practically unmistakable.

MICROWAVE PROPAGATION

In the first part of this chapter the radar equation was derived under the assumption of free-space propagation. The assumption is frequently not justified, and we must now turn to some of the important cases that fail to fulfill one or more of the requirements laid down in Sec. 2-1. We shall try, where possible, to modify the radar equation to suit the new circumstances, but it will be our broader purpose to describe, if only qualitatively, certain propagation phenomena peculiar to the microwave region. This is a vast subject. It includes some exceedingly difficult problems in mathematical physics, not yet completely solved; it includes topics in meteorology; above all, since it involves the weather and the variegated features of landscape and seascape never susceptible of exact mathematical description, it includes a large collection of observations and experience, rarely easy to interpret. We could not here treat such a subject comprehensively, but certain aspects with which the radar engineer should be familiar are not hard to explain. Their influence on radar planning and design is felt, or should be, at a very early stage.

2-12. Propagation over a Reflecting Surface.—If the transmission path lies near a reflecting surface it may be possible for energy to reach the target, and hence also for scattered energy to return to the radar antenna, by way of the surface as well as directly. The result of combining the direct and the reflected wave at the target will depend on the relative intensity and phase of the reflected wave, which in turn will depend not only on the difference in the length of the two paths but upon changes of phase or intensity introduced in the process of reflection. The analysis is very easy in the case of a *flat, perfectly reflecting*, surface.

Let us consider a nondirective transmitting antenna A located at a height h_1 above a flat reflecting surface S , as in Fig. 2-9. The field strength at some other point B of height h_2 can be described by giving the *ratio* of the field strength at that point to the field strength which would have been observed in the absence of the reflecting surface—in

other words, under free-space conditions. It is this ratio, which we shall call F , that we want now to compute.

The difference in length between the paths AB and AMB , if the angle θ is small, is given quite closely by

$$\overline{AMB} - \overline{AB} = \frac{2h_1h_2}{R}. \quad (24)$$

This is responsible for a difference in electrical phase of $4\pi h_1h_2/R\lambda$ radians between the two waves arriving at B , to which we must add any phase

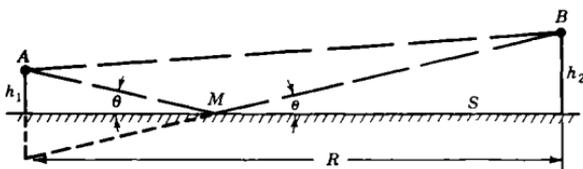


FIG. 2-9.—Propagation over a flat reflecting surface.

shift ψ resulting from the reflection of the one wave at M . The total phase difference ϕ is then

$$\phi = \psi + \frac{4\pi h_1h_2}{R\lambda}. \quad (25)$$

If the reflection at M is total, both waves arriving at B will have substantially the same intensity, namely, that corresponding to free-space propagation.¹ Taking the phase difference into account, then

$$F^2 = 2(1 + \cos \phi). \quad (26)$$

The angle ψ , contained in ϕ , depends upon the nature of the reflecting medium, upon the angle θ between the surface and the direction of wave travel, and upon the polarization. If the surface were a perfect *conductor*, ψ would be π radians for horizontal polarization (electric vector parallel to surface) and zero for vertical polarization² (magnetic vector parallel to surface). We shall be chiefly concerned, however, with the surface of the sea, and to a lesser extent with land, and these mediums behave more like *dielectrics* than *conductors* at microwave frequencies. For horizontal polarization it is still true that $\psi = \pi$, but for vertical polarization the situation is more complicated. If θ is considerably

¹ The fact that the waves arrive at B from slightly different directions, as well as the effect of the slightly unequal path lengths upon the relative intensity, need not be taken into account so long as $R \gg h_1 \gg \lambda$.

² This use of the terms "horizontal" and "vertical" polarization, though not meticulous, is common radar practice, and causes no confusion when the directions of propagation make small angles with the horizontal. Of course the electric vector of a vertically polarized wave traveling from A to M is not precisely vertical.

smaller than 6° , in the case of water the phase shift ψ is nearly π ; for nearly normal incidence, on the other hand, ψ is nearly zero. Also, the *magnitude* of the reflection coefficient, contrary to what we assumed in writing Eq. (26), varies widely as θ is changed, being in fact very nearly zero for $\theta = 6^\circ$ (Brewster's angle). This is as far as we care to pursue this complicated matter, apart from displaying, in Fig. 2-10, curves showing the dependence upon θ of the reflection coefficient and the phase shift, for reflection of 10-cm radiation at a water surface (there is very little difference between salt and fresh water at microwave frequencies).

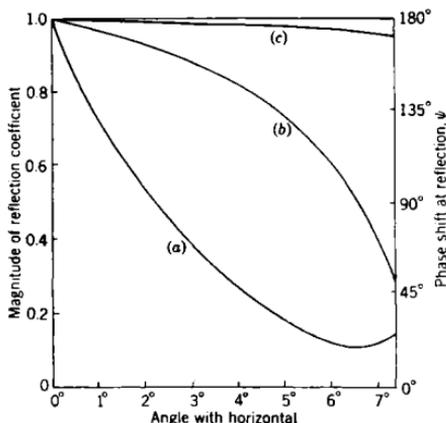


FIG. 2-10.—Reflection by water at 3000 Mc/sec. Curve a: amplitude of reflection coefficient, vertical polarization; Curve b: phase shift at reflection, vertical polarization; Curve c: amplitude of reflection coefficient, horizontal polarization. Phase shift for horizontal polarization is 180° throughout range shown.

Fortunately, the case which is of the greatest practical importance, that of nearly grazing incidence, is peculiarly easy to discuss, for if θ is less than, for example, 2° there is not much error in taking the phase shift to be π and the reflection coefficient, unity, for both polarizations. In this region, then, we can write

$$F^2 = 2 + 2 \cos \left(\frac{4\pi h_1 h_2}{\lambda R} + \pi \right) = 4 \sin^2 \left(\frac{2\pi h_1 h_2}{\lambda R} \right). \quad (27)$$

Now F^2 measures directly the ratio of the *power* incident on a target at B to that which would strike the target under free-space conditions. This ratio, according to Eq. (27), varies between 0 and 4, the latter value being attained when the direct and reflected waves arrive precisely in phase, or when the geometry is such as to satisfy the condition

$$\frac{4h_1 h_2}{\lambda R} = n, \text{ an odd integer.} \quad (28)$$

The effect of the interference described by Eq. (28) is to break up the original radiation pattern of the source at A into a lobe structure, as sketched in Fig. 2-11; at least this is an appropriate description of the effect at a considerable distance from the source. The lowest maximum in Fig. 2-11 occurs at an elevation angle of $\lambda/4h_1$ radians. Below this angle the field strength diminishes until, at the reflecting surface itself, it vanishes. No modification of the above argument is required for a *directive* antenna at A , unless the directivity is high enough to affect the relative intensity of the waves traveling along AB and AM (Fig. 2-9)

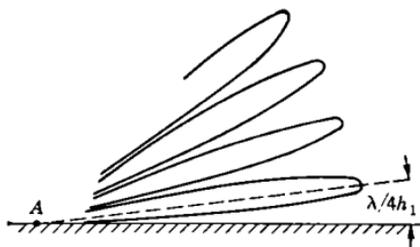


Fig. 2-11.—Lobes caused by interference of direct and reflected waves.

respectively. In any case, the width of an interference lobe will be less than the width of the primary radiation pattern since $\lambda/4h_1$ will certainly be greater than the vertical aperture of the antenna.

As the geometry of the problem is not altered if we interchange transmitter and target, the interference must affect the return of the radar echo in the same degree as it affects the pulse transmission. Therefore the required modification of the radar equation is obtained by multiplying by F^4 the expression given in Eq. (3) for the received signal power S ; that is,

$$S = P \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \cdot 16 \sin^4 \left(\frac{2\pi h_1 h_2}{\lambda R} \right). \quad (29)$$

To discuss the practical consequences of this we first distinguish between the problems of "high coverage" and "low coverage." In the former case we are concerned with targets so high or so close that $2\pi h_1 h_2 / \lambda R > 1$; that is, with targets lying well up in, or above, the lowest lobe of Fig. 2-11. Evidently the detection of such a target will depend critically on its range *and* altitude. If an airplane flies in toward the radar at constant altitude h_2 we must expect the signal to vanish and reappear repeatedly, as the lobe structure is traversed. The smaller λ and the larger h_1 , the more finely divided is the lobe pattern. A welcome change, on the other hand, is the additional factor of 16, effective in Eq. (29) for directions of maximum constructive interference. This factor should double the maximum range of detection, compared to free-space conditions. Something like this increase is actually observed, although it should not in every case be taken for granted, in view of the idealized model upon which it is based.

"Low coverage" refers to a situation in which the target lies well below the lowest maximum in Fig. 2-11, in the region where

$$\frac{2\pi h_1 h_2}{\lambda R} < 1. \tag{30}$$

For this case the sine in Eq. (29) can be replaced by its argument, leading to

$$S = 4\pi P \frac{G^2 \sigma (h_1 h_2)^4}{\lambda^2 R^8}. \tag{31}$$

We see that in this region the signal strength falls off as $1/R^8$ rather than $1/R^4$, if the other factors are held constant. This region is often referred to loosely as the “eighth-power region,” or occasionally as the “far zone.”

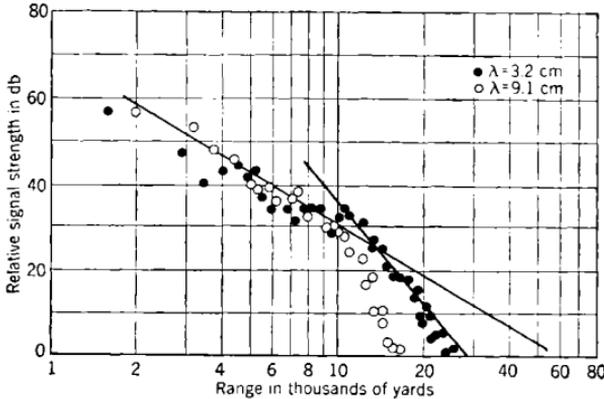


FIG. 2-12.—Radar performance over water. Target: 5000-ton freighter. Antenna height: 21 ft.

As radar and target are separated, keeping h_1 and h_2 constant, the range at which this inverse eighth-power dependence begins varies as the reciprocal of the wavelength; in other words, the longer the wavelength, the higher the lowest lobe in Fig. 2-11 is tilted, and the sooner an outgoing target recedes underneath it. This gives a heavy advantage to shorter wavelengths where the primary task of the radar set is to search the sea for low-lying targets, the radar antenna itself not being mounted at a great height. The advantage of the shorter wavelength is maintained throughout the inverse eighth-power region. Comparison of Eq. (3) with Eq. (31) shows that if two radar sets, operating at wavelengths λ_1 and λ_2 , receive equal signals from a target in the free-space region, the ratio of respective signal strengths in the region to which Eq. (31) applies will be $S_1/S_2 = (\lambda_2/\lambda_1)^4$.

Figure 2-12 records some radar observations made at two wavelengths with the same ship as the target. The signal strength in decibels is plotted against the logarithm of the range. Clearly distinguishable are two regions in which the slope corresponds respectively to an R^{-4} and to an R^{-8} dependence, and the relative advantage of the shorter wavelength in the latter region is apparent.

The maxima and minima which one might have expected in the nearer region are not conspicuous in Fig. 2-12. The reason for this is that the target was not a point with a unique height h_2 but a complicated object extending from the surface up to some maximum height. In this "near zone" some parts of the ship lie on maxima and some lie in the nulls, and the net result at any instant is some sort of average, in the fluctuations of which one could hardly expect to discern traces of the regular interference pattern predicted by Eq. (29). The absence of a sharply defined break in the curve between the two regions is readily justified on the same grounds. The location of the break, ill-defined as it is, can however be used to compute some "effective height" h_2 .

Although the relation in Eq. (30) fixes roughly the inner boundary of the " R^{-8} region," were we to cling to our flat-earth hypothesis there would be no outer limit. Actually, of course, the region is ultimately

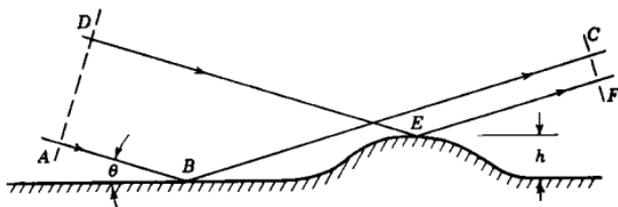


Fig. 2-13.—Reflection from an irregular surface.

bounded by the radar horizon, where a new and even more drastic falling-off in signal strength sets in. In many instances, especially at microwave frequencies, this limit occurs so soon after the beginning of the eighth-power region as to reduce the latter to rather inconsequential size.

Before we discuss the effect of the earth's curvature, we should comment on certain other, less important, shortcomings of the preceding simplified argument. For one thing, we have ignored the fact that the reflecting surface is not smooth; even in the case of the sea, the irregularities (waves) are both wide and deep compared to a microwavelength; in the case of a land surface, we may be confronted with any imaginable irregularity. Nevertheless, at least for reflection on the sea at nearly grazing incidence, the observed reflection coefficient is rather close to what a glassily smooth sea would give.¹ This is, perhaps, no more surprising than the fact that an ordinary piece of paper displays, at nearly grazing incidence, specular reflection of light, and it can be made plausible by some argument such as this: In Fig. 2-13, two parallel rays, AB and DE , strike a "rough" surface, the roughness consisting of a single bump of height h . The reader will find with little trouble that the net path difference between the two rays $\overline{EF} - \overline{AB}$ is just $2h \sin \theta$. If

¹ The reflection coefficient would have to be substantially less than 1 to change the result significantly, so far as the radar problem is concerned.

only this difference is small compared with λ , the surface should reflect as though it were optically smooth. But $h\theta < \lambda$ implies $h < \lambda/\theta$, and if θ is small, bumps which are actually large compared to λ can be tolerated.

Something like this may occur in some circumstances over land. Indeed, reflection has been observed on unusually flat terrain, such as an airfield, at microwave frequencies. Except in such unusual circumstances there is very little evidence of ground reflection at wavelengths of 10 cm or less. For microwaves, then, the results of this section are almost wholly restricted to transmission over water at nearly grazing incidence.

2.13. The Round Earth.—The distance R_h to the optical horizon, from an observer situated h feet above the surface of a spherical earth of diameter D_0 ft would be given by the formula, $R_h = \sqrt{D_0 \cdot h}$ if the atmosphere did not bend the rays of light. Actually, the earth's atmosphere decreases in density with height, introducing a downward curvature in all rays, which allows a ray to reach somewhat beyond the distance given by the above formula. For rays of small inclination to the horizontal, and for heights small compared to the thickness of the atmosphere, this effect can be taken into account by replacing the true diameter of the earth by a somewhat larger number D_{eff} . How much larger it is depends on the rate of change with height of the index of refraction of the atmosphere. This effect may be expected to show local variations. A reasonable choice of a "standard" condition leads to a value $R_{\text{eff}} = 1.33R_0$, and thence, thanks to a fortuitous numerical relation between units, to an easily remembered formula for the distance to the radar horizon,

$$R_h \text{ (statute miles)} = \sqrt{2h \text{ (feet)}}. \quad (32)$$

The formula Eq. (32) predicts a somewhat greater horizon distance than does the corresponding formula for the *optical* horizon, because the conditions assumed as "standard" include a moderate gradient of water-vapor concentration. Water vapor, although it has but a minor influence on the atmospheric refraction of visible light, displays a very pronounced refractive effect at all radio frequencies, including microwave frequencies. This effect, caused by the permanent electric moment of the water molecule, can, in some circumstances, drastically affect the propagation of microwaves in a manner to be described in the next section.

Confining our attention for the moment to the atmosphere of "standard" refractive properties, to which Eq. (32) applies, let us see why it makes sense to speak of a microwave horizon when radiation of much lower frequency, as is well known, travels far beyond any such horizon. In the first place, the *ionosphere* does not, to any appreciable degree, reflect or refract microwaves. In the second place, the spreading of waves *around* the curved surface of the earth, essentially by *diffraction*, is much reduced at microwave frequencies because the wavelength is so

small compared with the size of the obstacle, to put it very crudely. There is some spreading of this sort, of course, and it is a phenomenon upon which a good deal of theoretical effort has been expended. Methods have been developed for calculating the intensity of the radiation in the "diffraction region," that is, beyond the horizon (Vol. 13, Chap. 2). In this region, however, the field strength normally diminishes so rapidly with increasing range that any additional radar coverage thus obtained is of little value. From the point of view of the radar designer, targets over the horizon might as well be regarded as totally inaccessible under "standard" conditions of propagation.

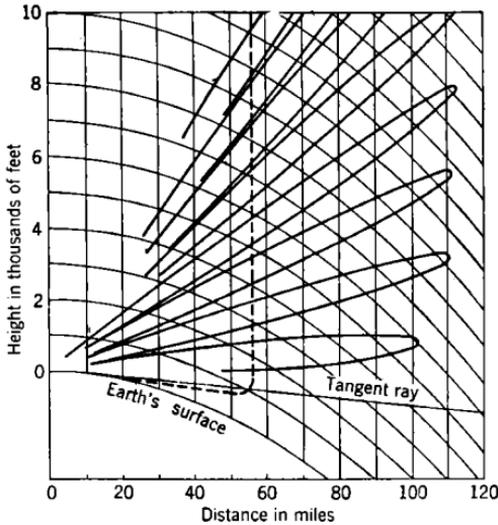


FIG. 2-14.—Coverage diagram for 2600 Mc/sec, transmitter height 120 ft. Solid curve for totally reflecting earth. Dotted curve for nonreflecting earth.

As for regions well within the horizon, the curvature of the earth at most complicates the geometry of the interference problem discussed in the preceding section. Naturally, we have no right to apply Eq. (29), as it stands, to targets near the horizon. We need not concern ourselves here with these complications, which are adequately treated in Vol. 13. Methods have been worked out for rapidly calculating the field strength over a curved reflecting earth. The radar designer usually prefers to display the results in the form of a *coverage diagram*, which shows contours of constant field strength plotted in coordinates contrived to show directly the effect of the curvature of the earth. One such contour, shown in Fig. 2-14, is calculated for an omnidirectional antenna transmitting at 2600 Mc/sec from a height of 120 ft above a totally reflecting earth. The dotted curve is the corresponding contour with a nonreflecting earth. Both contours would usually be modified by the directional pattern of the

radar antenna and it is not difficult to take this into account in the calculation. Note the exaggeration of the vertical scale.

A contour of constant field strength is a contour of constant power intensity as well; moreover, the intensity of the signal received from a radar target of given cross section, located anywhere along such a contour, is the same.¹ Thus, if the proper contour is chosen it will represent, as nearly as any curve can, the boundary of the region in which a given target can be detected. The reader's earlier introduction to some of the statistical factors involved in radar should prepare him for the warning that such coverage diagrams are not to be taken too literally. They are, nevertheless, useful in the planning and design of long-range search radar.

2-14. Superrefraction.—As we have seen, the effect of the normal vertical gradient of refractive index in the atmosphere is to introduce a slight downward curvature in the path of light and of microwaves. Were this curvature only a few times greater, it would equal the curvature of the earth itself, and it would be possible for a ray to bend around the earth without leaving the surface; in other words, there would be no *horizon*. Whatever misgivings we may have about the use of the word "ray" in this connection, it would not be surprising if some interesting departure from standard microwave propagation were to manifest itself under such conditions.

Refractive index gradients of the requisite strength (5 parts in 10^8 per ft) can be produced under some conditions by temperature gradients alone. For example, if land heated by the sun cools by radiation at night, a fairly thin layer of cold (therefore *dense*) air may be formed just above the ground, which results in an unusually rapid decrease of refractive index with height, the index of the lowest layer being abnormally great.

A more widespread cause of strong vertical gradients in refractive index, and therefore of excessive bending of rays, is the refractive effect of water vapor mentioned earlier. Over most of the surface of the ocean the region above the water is not saturated with water vapor, whereas the layer directly in contact with the water must be very nearly saturated. There is, in other words, a continual evaporation of water from the sea

¹ One must be careful not to confuse the directional pattern of an antenna with the plot of contours of constant intensity in the field of the antenna, which it occasionally superficially resembles. If we plot the *gain* of an antenna as a function of *angle*, in polar coordinates, we have an antenna pattern that has meaning, strictly, only if the antenna is isolated in space. On the other hand, contours of constant intensity, the coordinates of which refer directly to positions in space, can be used to describe the radiation field no matter what the surroundings or type of propagation involved. The fact that a signal of the same intensity is received from a target at any position on one such contour does not depend on the inverse-square law or any other law of propagation but only on the Reciprocity Theorem.

and a diffusion of the vapor upward into the overlying air mass. This implies the existence of a vertical gradient in the concentration of water vapor with, normally, the highest concentration at the surface and decreasing upward.

A typical condition to which such an effect may lead is that of a relatively shallow layer just above the surface within which the vertical gradient of refractive index is negative and exceeds the critical value of 5 parts in 10^8 per ft. Such a region is called a "duct" for reasons that will appear shortly, and the level at which the gradient has just the critical value is called the "top of the duct." Were we to trace the path of an initially horizontal ray at this level, we would find it curving downward just enough to keep up with the curvature of the earth, and therefore maintaining constant height. Farther down in the duct, where the gradient is stronger, it would be possible for a ray launched at a slight upward inclination to be bent back to the surface again and thus to proceed by a series of bounces, trapped as it were, within the duct.

From this temptingly graphic picture of propagation within a duct it is easy to draw false conclusions. Since the description of the process in terms of rays, traced by the rules of geometrical optics, nowhere involves the *wavelength* of the radiation, we should be led to expect similar effects at all frequencies for which the index of refraction has the same value, namely for all radio frequencies. But, actually, ducts such as we have described have *no observable effect* on the propagation of low-frequency radio waves. The reason for this is that the duct is effective in "trapping" and guiding radiation only if the wavelength is less than some critical value determined by the height of the duct and the steepness of the gradient of refractive index within the duct. The ducts which we have described have no great vertical extent, their heights being of the order of a few tens, or at the most a few hundreds, of feet; their influence on propagation is usually confined to frequencies in the 1000-Mc/sec range and above.¹ For gradients in refractive index which would not be unusual in these surface ducts over water, the relation between height of duct and the longest wavelength strongly affected by the duct is suggested by the figures in Table 2-3.

TABLE 2-3

Height to top of duct, ft.....	25	50	100	200	400
Longest wavelength trapped, cm*.....	1.8	5	15	40	110

* These numbers are based on an arbitrary, although reasonable, criterion for trapping, and upon a simplified model in which the refractive index decreases upward through the duct at the constant rate of 8 parts in 10^8 per foot. They are intended only to be illustrative.

¹ In certain parts of the world, the effect of trapping has been observed for frequencies as low as 200 Mc/sec.

That the height of the duct and the wavelength should be related, and in such a manner that the heights involved are hundreds of times greater than the corresponding wavelengths, may perhaps be made plausible to the reader acquainted with propagation through waveguides. The duct is, in a sense, a waveguide. Let us consider an oversimplified model of a duct in which the index of refraction, n_1 , is constant from the surface up to some height a , where it abruptly changes to n_2 (Fig. 2-15). If n_1 exceeds n_2 by some very small amount δ , both n_1 and n_2 being very nearly unity, a wave incident on the boundary CD at a grazing angle α smaller than $\sqrt{2\delta}$ will experience total internal reflection. Under such conditions the region between AB and CD can be regarded as the interior of a waveguide bounded by two reflecting surfaces. But in an ordinary

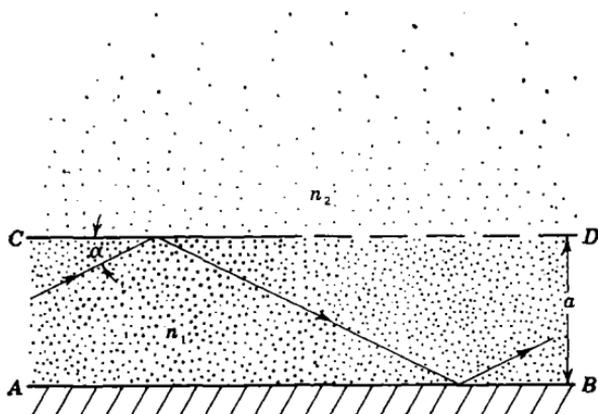


FIG. 2-15.—Propagation within a duct (oversimplified).

waveguide there is a basic relation between the width of the guide and the angle of incidence of the plane waves into which the simpler waveguide modes can be resolved. The longer the wavelength and the narrower the guide, the larger the angle α according to a relation which for very wide guides reduces to $\alpha = \lambda/2a$. But if, in our model sketched in Fig. 2-15, α becomes larger than $\sqrt{2\delta}$, total internal reflection no longer occurs, and the energy leaks rapidly out of the guide. We might therefore anticipate some such restriction as $\lambda/2a < \sqrt{2\delta}$ or $\lambda^2 < 8a^2\delta$ for an effective duct. The form of this result is not inconsistent with the figures quoted above which were based on a constant gradient of n_1 in the duct, although it must be admitted that we have brutally oversimplified a problem that abounds in mathematical difficulties and subtleties.

Actually, no sharp distinction between trapping and standard propagation can be drawn. Even our naïve model suggests this; for we need

not require *total* reflection at *CD* to get appreciable guiding for a considerable distance. Nor does the source of radiation have to lie within the duct to permit a portion of the energy to be partially trapped in the duct, although it may not be too far above it. The variation of field strength with distance from the transmitter and height above the surface is very complicated, and to cover this type of propagation by a mere modification of the radar equation is entirely out of the question, as the reader who pursues this subject into Vol. 13, where it is treated at length, will learn. It is perhaps best here to summarize the aspects of super-refraction which have a significant bearing on radar planning and design.

1. The guiding of microwaves by refractive anomalies of the duct type appears to be the sole means by which coverage beyond the horizon can be obtained. Extensions of range up to several times the horizon distance have often been observed.

2. The most prevalent, the best understood, and probably the most important example of the type is associated with the surface evaporation duct, which seems to exist most of the time over large areas of the oceans of the world.

3. Short wavelengths are required to take advantage of the guiding effects of ducts, and, for surface ducts, low antenna (and target) heights. This is especially true of the evaporation duct.

4. *Substandard* as well as *superstandard* radar ranges can be caused by refractive anomalies, if the transmission path is nearly horizontal.

5. Propagation at angles steeper than a few degrees with respect to the horizontal is not affected by the refractive anomalies here discussed.

2-15. Attenuation of Microwaves in the Atmosphere.—The earth's atmosphere, excluding the ionosphere, is for all practical purposes transparent to radio waves of frequency lower than 1000 Mc/sec. Even over a transmission path hundreds of miles long no appreciable fraction of the energy in the radio wave is lost by absorption or scattering in the atmosphere. With the extension of the useful range of radio frequencies into the microwave region we have at last entered a part of the electromagnetic spectrum to which the atmosphere is not wholly transparent. Indeed an upper limit to frequencies useful for radar, imposed by the properties of the atmosphere, is now within sight. The radar engineer must therefore acquaint himself with certain phenomena falling heretofore within the exclusive province of the molecular spectroscopist.

Broadly speaking, there are two ways in which energy can be dissipated from a radar beam: (1) by direct absorption of energy in the gases of the atmosphere; (2) through absorption or scattering of energy by condensed matter such as water drops. All such processes lead to an exponential decrease of intensity with distance from the source, superimposed on, and eventually dominating, the inverse-square dependence.

The effect is therefore truly an *attenuation* in the sense in which the word is applied to transmission lines, and is properly measured in decibels per kilometer. We shall discuss first the absorption by the gases of the atmosphere.

Of the three abundant gases of the atmosphere—nitrogen, oxygen, and water vapor—the latter two are intrinsically capable of interacting with and absorbing energy from a radio wave by virtue of the permanent *electric* dipole moment of the water molecule and the permanent *magnetic* dipole moment of the oxygen molecule. We know, however, that molecules absorb radiation in more or less well-defined absorption lines, or bands, and we have to inquire whether either of these molecules exhibits absorption lines in the microwave range—that is, at frequencies much lower than those usually associated with molecular absorption spectra. It has been found that both oxygen and water vapor do in fact display such absorption. Although the effects observed would be classed as very weak by a spectroscopist, radar involves transmission over such long paths that very serious attenuation is encountered in certain parts of the spectrum. In Fig. 2-16 are plotted curves showing the course of the water vapor absorption and oxygen absorption, as a function of wavelength. The absorption is measured by the rate of attenuation in decibels per kilometer.

The most prominent feature of the water-vapor absorption is a single “line” which appears as a broad maximum centered about 1.3-cm wavelength, superimposed on the residual effect of a multitude of far stronger lines located at much shorter wavelengths. The solid part of the curve is based on extensive direct measurements. These confirmed the main features of the theoretical predictions,¹ on the basis of which the remainder of the curve has been sketched in; one cannot, however, rely on the quantitative accuracy of the dotted part of the curve. The curve is plotted for an atmosphere containing 10 g of water vapor per cubic meter. This corresponds to a relative humidity of 66 per cent at a temperature of 18°C, for example. Over the range of absolute humidities normally encountered in the atmosphere one may assume that the attenuation is simply proportional to the absolute humidity. The rapid rise of the curve below 3 mm is evidence of the powerful absorption displayed by water vapor throughout the far infrared. No further transparent regions are to be found until we reach a wavelength of the order of 15 microns (0.0015 cm).

The oxygen absorption rises to a high peak at 5-mm wavelength; this has been quantitatively verified by direct measurement. At longer

¹J. H. Van Vleck, “Further Theoretical Investigations of the Atmospheric Absorption of Microwaves,” RL Report No. 664, March 1, 1945. See also Vol. 13 of this series.

wavelengths, in contrast to the behavior of the water-vapor curve, a residual absorption persists up to wavelengths of the order of 30 cm. This effect is small enough to have escaped experimental detection, but it is clearly predicted by the theory.¹ Except near the center of the absorption maximum, the attenuation due to oxygen should vary about as the *square* of the pressure; hence the effect rapidly diminishes at high altitudes. It is perhaps unnecessary to remark that attenuations

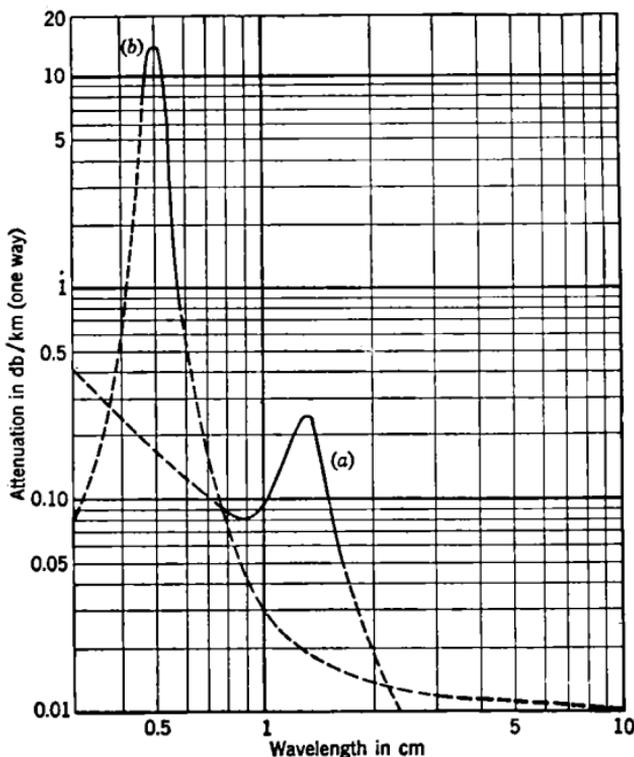


FIG. 2-16.—Attenuation caused by water vapor (Curve *a*) and oxygen (Curve *b*). Curve *a* applies to an atmosphere containing 10 g of water vapor per m³. Curve *b* applies to an atmosphere which is one fifth oxygen, at a total pressure of 76 cm Hg.

due to independent causes, such as water vapor and oxygen, are directly additive.

Water drops in the atmosphere can affect the passage of microwave radiation in two ways. In the first place liquid water is a very imperfect dielectric at microwave frequencies, and absorbs energy from an oscillating electric field just as any lossy dielectric would. For extremely small drops, such as those in fog or clouds, this is the only important effect, and in this limiting case the attenuation at a given wavelength is simply proportional to the aggregate liquid water content of the atmosphere,

¹ J. H. Van Vleck, *op. cit.*; see also Vol. 13.

measured, for instance, in grams per cubic meter.¹ The effect of larger drops is more complicated, depending not only upon the total mass of water per unit volume, but upon the diameter of the drops as well. The absorption process itself is no longer simple, and *scattering* of energy by the drops, which depends very strongly on the ratio of wavelength to drop diameter, begins to play a role. Energy scattered out of the directed beam must, of course, be counted as lost.

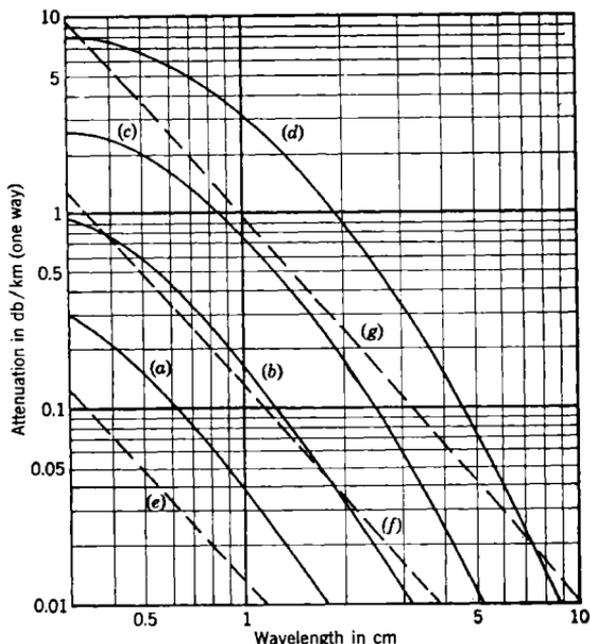


FIG. 2-17.—Solid curves show attenuation in rain of intensity (a), 0.25 mm/hr (drizzle); (b), 1 mm/hr (light rain); (c), 4 mm/hr (moderate rain); (d), 16 mm/hr (heavy rain). Dashed curves show attenuation in fog or cloud: (e), 0.032 g/m³ (visibility about 2000 ft); (f), 0.32 g/m³ (visibility about 400 ft); (g), 2.3 g/m³ (visibility about 100 ft).

The attenuation resulting from these effects can be calculated for drops of any given diameter. Reliable and accurate though they may be, such results are not in themselves of much use, for it is neither customary nor convenient to describe a rain in terms of the drop diameter and the number of drops per cubic meter. In any case, the drops are never all of one size and it is not sufficient to know merely the average diameter. Instead, one has to make use of empirical meteorological data correlating drop size (really distribution-in-size) with precipitation rate to arrive finally at a relation connecting attenuation in decibels

¹ This conclusion holds so long as the diameter of the drops is very much less than λ/n , where n is the index of refraction of water at the frequency in question.

per kilometer, with precipitation rate in millimeters per hour. The most extensive analysis of this sort has been carried out by J. W. and D. Ryde, upon whose work¹ the curves of Fig. 2-17 are based. The direct measurements of attenuation and rainfall which have been made confirm these predictions satisfactorily. The chief difficulty in such experiments is connected with the measurement of the rainfall, which is homogeneous in neither time nor space.

The dashed curves of Fig. 2-17 show the attenuation in fog or clouds which, as we have said, does not depend on the drop diameter. According to Ryde a certain limit of optical visibility can be, at least loosely, associated with each of the dashed curves. For the conditions to which Curve *f* applies, for example, the optical range is limited to about 400 ft. At a wavelength of 3 mm the radar range could be 50 to 100 times as long.

It would be easy, but not very instructive, to introduce an exponential factor into the radar equation to take account of the attenuation that we have been discussing. We leave this task to the reader, who will find no difficulty in calculating, for any given case, the reduction in range caused by a specified strength of attenuation, which is effective, of course, on both the outgoing and return path. One general observation should be made, however, which is that the effect of an exponential term in the radar equation is insignificant at very short ranges but overwhelming at very long ranges. What we mean by *short* and *long* is determined by the rate of attenuation. An entirely arbitrary criterion, which will serve as well as any other for discussion, is the range for which the presence of the atmospheric attenuation just *doubles* the normal rate of decrease of signal intensity with range. If α is the rate of attenuation in db/km, the range R_0 so defined is given by $R_0 = 8.68/\alpha$ km. At shorter ranges than this the inverse-square law is the more important factor; at ranges greater than R_0 the exponential factor controls the situation and any slight improvement in range must be bought at enormous price. In other words, once attenuation takes hold, it is of little avail to struggle against it.

¹ J. W. Ryde and D. Ryde, Report 8670 of The Research Laboratory of General Electric Company, Ltd. This is a British publication.

CHAPTER 3
PROPERTIES OF RADAR TARGETS
SIMPLE TARGETS

BY A. J. F. SIEGERT, L. N. RIDENOUR, AND M. H. JOHNSON¹

3.1. Cross Section in Terms of Field Quantities.—In the preceding chapter the quantity “cross section of a target” was introduced phenomenologically. Theoretical considerations which in certain cases will allow the prediction of the value of these quantities from the known properties (shape, material) of the target will be presented in this chapter. The following considerations will be restricted to cases where the individual target is sufficiently small, compared with the distance from the transmitter, to permit the incident electromagnetic field at the target to be approximated by a plane wave propagating in the direction of the target away from the transmitter; this is chosen as the z direction. The problem of finding the cross section then reduces to the mathematical problem of finding that solution of Maxwell’s equations which at large distances from the target reduces to the incident plane wave and at the target fulfills the proper boundary conditions.

Suppose a solution of this problem has been found. At large distances from the target the component of the electric field parallel to the receiving dipole can be written in the form $E_0 e^{\frac{2\pi i}{\lambda}(z-ct)} + S \frac{e^{\frac{2\pi i}{\lambda}(r-ct)}}{r} +$ terms decreasing faster than r^{-1} , where E_0 is the amplitude of the incident plane wave, S/r is the amplitude of the only important part of the scattered wave, λ the radar wavelength, c the velocity of light, and t the time. Usually E_0 and S contain complex phase factors; S is in general a function of the scattering angles. In the following discussion the symbol S_R is used to denote the value of S in the direction of the receiver. In terms of E_0 and S_R , the cross section σ defined in Sec. 2.3 is given by

$$\sigma = 4\pi \frac{|S_R|^2}{|E_0|^2}. \quad (1)$$

3.2. Rayleigh Scattering from a Small Sphere.—As an illustration of the use of this equation, we shall derive the Rayleigh law for the case

¹ Sections 3-1-3-4 and 3-6 by A. J. F. Siegert, 3-5 by L. N. Ridenour, and 3-7 by M. H. Johnson.

of back scattering from a sphere of radius a small compared to the radar wavelength. In this limiting case the incident field is considered homogeneous over the extension of the target. In the homogeneous field the sphere becomes an electric dipole with a dipole moment $pe^{-\frac{2\pi i}{\lambda}ct}$, where

$$p = \frac{\epsilon - 1}{\epsilon + 2} a^3 E_0 \quad (2)$$

and ϵ is the dielectric constant. The field of this dipole, observed at a distance r in the wave zone and in a direction perpendicular to the polarization, is

$$S_R \frac{e^{\frac{2\pi i}{\lambda}(r-ct)}}{r} = \left(\frac{2\pi}{\lambda}\right)^2 p \frac{e^{\frac{2\pi i}{\lambda}(r-ct)}}{r}. \quad (3)$$

The cross section is therefore

$$\sigma = 4\pi \left(\frac{2\pi}{\lambda}\right)^4 \left|\frac{p}{E_0}\right|^2 = 4\pi \left(\frac{2\pi}{\lambda}\right)^4 \left|\frac{\epsilon - 1}{\epsilon + 2}\right|^2 a^6. \quad (4)$$

If $|\epsilon| \gg 1$, which is true, for instance, for a raindrop, we obtain the simpler expression

$$\sigma = 4 \left(\frac{2\pi a}{\lambda}\right)^4 \pi a^2, \quad (5)$$

which in this form serves to compare the radar cross section with the "geometric cross section" πa^2 .

For the case of a metal sphere, one would be tempted to use the same formula with $|\epsilon| = \infty$. This, however, is not correct, because, in the case of a conducting sphere, surface currents which have a magnetic dipole moment are induced by the field. The radiation of both the electric and the magnetic dipole must therefore be considered. We then obtain

$$\sigma = 9 \left(\frac{2\pi a}{\lambda}\right)^4 \pi a^2. \quad (6)$$

3-3. Scattering of a Plane Wave by a Sphere.—The general problem of scattering of a plane wave by a sphere is solved in detail in J. A. Stratton's book, *Electromagnetic Theory* (pages 563 ff), and references to the original papers can be found there. The cross section for back scattering from a metal sphere¹ divided by πa^2 is plotted vs. a/λ in Fig.

¹ P. J. Rubenstein, RL Report No. 42, Apr. 3, 1943. Numerical values for the problem of scattering from a dielectric sphere can be found for certain values of the dielectric constant in the references given by Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 563 ff. See also Robert Weinstock, RRL Report

3-1, where a is the radius of the sphere and λ the radar wavelength. For wavelengths large compared to the radius, the cross section is given by Rayleigh's law. In the opposite limiting case ($\lambda \ll a$) the cross section approaches the geometrical cross section πa^2 . Between these limits are the resonance maxima.

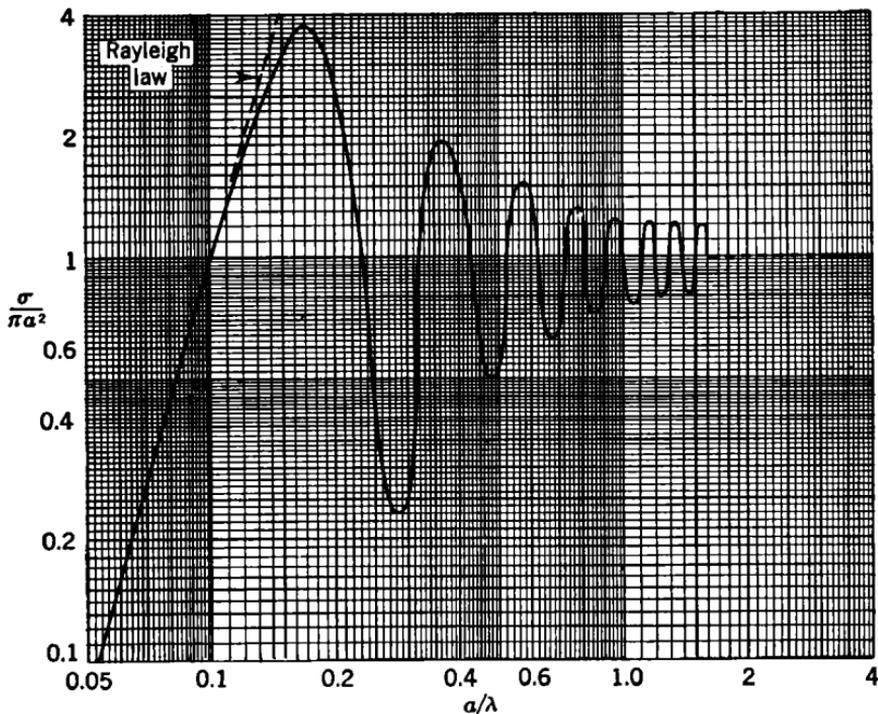


FIG. 3-1.—Back scattering from a metallic sphere. In the region where the line is shown dotted, no calculations have been made.

3-4. Approximations for Large Metal Targets.¹—For metal targets whose dimensions as well as radii of curvature are large compared with the radar wavelength, the following approximation yields good results. For every surface element the current is used which would be caused by the incident field if the surface element were a part of an infinite plane sheet. The radiation from all the surface elements is added in the proper phase relations, and the resultant field is considered as the scattered field from the target. For a flat metal sheet of area A perpendicular to the

No. 411-125, Nov. 14, 1944 (for cylinder and sphere); Morse and Rubenstein, *Phys. Rev.*, **54**, 895 (1938) (for elliptical cylinder); L. J. Chu, RRL Report No. 4, Oct. 22, 1942; Marion C. Gray, "Reflection of Plane Waves from Spheres and Cylinders," BTL Report MM 42-130-95.

¹J. F. Carlson and S. A. Goudsmit, "Microwave Radar Reflections," RL Report No. 43-23, Feb. 2, 1943.

incident radar beam we thus obtain the cross section

$$\sigma = \frac{4\pi A^2}{\lambda^2}. \quad (7)$$

For general angles of incidence the cross section is a function of the angle of incidence and it varies very rapidly when, as is assumed here, the wavelength is small compared with the linear dimensions of the plate. In a diagram of return power vs. angle of observation these variations show up as the "lobes." The strong main lobe is normal to the plate and the side lobes decrease rapidly with increasing angle. For small angles θ , but excluding the main lobe, the average cross section $\bar{\sigma}$ (averaged over several lobes) is given approximately by

$$\bar{\sigma} \approx \frac{4\pi\lambda^2}{(2\pi\theta)^4}. \quad (8)$$

This result is independent of the size of the target, subject of course to the limitation that the linear dimensions of the target are large compared to the wavelength.

For a cylinder of radius R and length l , both large compared to λ , we obtain

$$\sigma = 2\pi \frac{Rl^2}{\lambda} \quad (9)$$

for incidence perpendicular to the axis. For the same cylinder, if the beam forms an angle θ with the normal to the axis the average cross section (averaged over several lobes) is approximated by

$$\bar{\sigma} = \frac{R\lambda}{2\pi\theta^2}, \quad (10)$$

valid for small angles θ excluding the main lobe.

For curved surfaces, the formulas of geometrical optics can ordinarily be used.¹ For a segment of spherical surface of radius R we find

$$\sigma = \pi R^2 \quad (11)$$

independently of the size of the segment and of the wavelength λ . This formula is valid as long as the diameter d of the segment (perpendicular to the incident beam) is larger than $\sqrt{2\lambda R}$, and provided that the edge of the segment deviates sufficiently ($>\lambda/4$) from a plane perpendicular to the incident beam, since otherwise edge effects may become important. If the surface is not spherical the formula is still valid if we take for R the geometric mean of the two principal radii of curvature. The result is the same whether the concave or convex side is turned toward the radar transmitter.

¹ For more detailed discussion see Vol. 13 of this series, Chap. 6.

3-5. The Corner Reflector.—It is often desirable to make a compact radar target with a large cross section. A flat plate of dimensions large compared to a wavelength exhibits a large cross section when viewed along its normal, because of specular reflection, but the cross section falls off sharply in other directions [see Eqs. (7) and (8), Sec. 3-4]. The problem of designing a target that will give strong specular reflection for almost any direction of illumination has been solved by taking over into microwave radar practice the corner reflector familiar in optics. The small glass reflectors used in highway markers work on this principle.

A corner reflector consists of three mutually perpendicular intersecting planes (Fig. 3-2*a*). If a beam is directed into the corner formed by the planes, triple reflections occur which send it back in the direction from

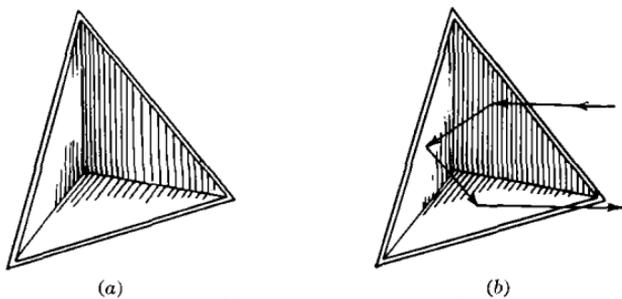


FIG. 3-2.—The triangular corner reflector.

which it came (Fig. 3-2*b*). The effective area for triple reflection depends on the direction in which the corner is viewed, but it is large over most of the octant in which a single corner is effective. When the area for triple reflection grows small, double reflection (from two planes whose line of intersection is nearly normal to the line of sight) and single reflection (from a plane nearly normal to the line of sight) begin to make important contributions to the radar cross section. A single corner will be effective only for directions of illumination that cover one octant of a sphere centered at the reflector, as has been remarked; but all directions can be covered by making a cluster of eight such corners (Fig. 3-3).

We can find the cross section for a corner by considering it equivalent to a flat reflecting plate whose area is the effective area of the corner for triple reflection. Equation (7) gives, for area A and wavelength λ , the cross section

$$\sigma = \frac{4\pi A^2}{\lambda^2}.$$

The maximum area for triple reflection will be that afforded by the corner when it is viewed along its axis of symmetry. This maximum area is that of the regular hexagon formed by cutting off the corners of the projection of the corner on its axis of symmetry; it is given by

$$A_{\max} = \frac{a^2}{\sqrt{3}}$$

where a is the edge of the corner. This yields for the maximum cross section

$$\sigma_{\max} = \frac{4}{3} \frac{\pi a^4}{\lambda^2},$$

showing that the free-space range at which a corner reflector can be seen by a given radar is proportional to its linear dimensions. Figure 3-4 shows the maximum cross section at various radar frequencies as a function of edge length.

As a/λ (and consequently the gain of a corner reflector) is increased, greater attention must be paid to the mechanical construction of the reflector. When the angle between one pair of planes differs from 90° , the triply reflected return from the corner splits into two divergent beams, with a con-

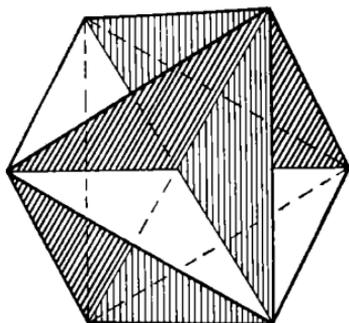


FIG. 3-3.—A cluster of corner reflectors.

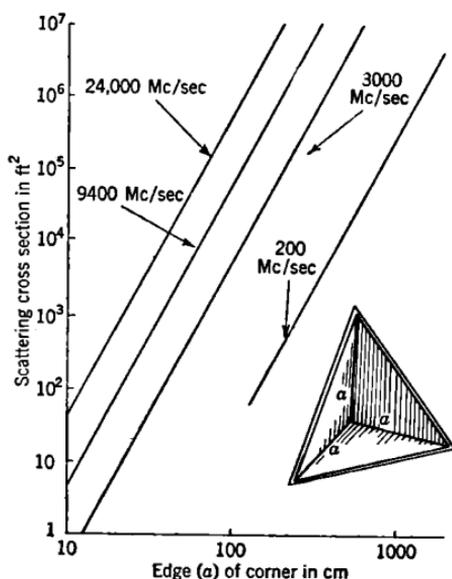


FIG. 3-4.—Maximum scattering cross section of triangular corner reflector.

sequent reduction in signal return. If two angles are in error, there are four beams; if all three angles are wrong, six beams result. The theory of the effect of errors¹ yields the result that a loss in returned signal of 3 db is caused by a displacement of the outer edge of the corner of about $\lambda/2$. This limitation requires extremely close attention to the mechanical construction of corner reflectors of high gain. It also presents the possibility of applying modulation to the reflection from a corner by slight motion of one side.

3-6. Target Shaping to Diminish Cross Section.—A corner reflector gives a radar return by specular reflection for nearly all directions of

¹ R. C. Spencer, "Optical Theory of the Corner Reflector," RL Report No. 433, Mar. 2, 1944.

approach. A target consisting of two planes intersecting at right angles gives specular reflection for all directions of observation perpendicular to the line of intersection. A cylinder standing perpendicular on a plane is also strongly visible for most directions of approach. If the problem is, on the contrary, to construct a target with a low radar cross section for most directions of approach, such configurations must be avoided. Furthermore, to eliminate ordinary specular reflection, surfaces perpendicular to the probable directions of observation must be eliminated. Partial camouflage of targets has been achieved by proper shaping.

3-7. Use of Absorbent Materials.—The possibility of reducing the radar cross section by the use of materials absorbent at radar frequencies has been the subject of considerable investigation. Contributions to the cross section can be separated roughly into two parts. The first arises from flat or gently curved surfaces of considerable area (in square wavelengths) which are normal to the line joining the target and transmitter. Radiation returned from such surfaces may be properly classed as specular reflection; cross sections arising therefrom will be referred to as “specular cross sections.” If any dimension of the target is smaller than a wavelength, or if the backward radiation arises from the secondary maxima in the diffraction pattern, the cross section will be referred to as a “diffraction cross section.” When both types of backward scattering are present, the specular cross section is usually much greater. In considering the use of absorbent materials, this distinction must always be borne in mind. The theory that follows is applicable only to specularly reflected radiation, and therefore to the specular cross section.

Absorbers in general are of two types. In the first kind, reflections occurring at the front surface of the absorber are canceled by destructive interference with the wave that enters the layer and subsequently reemerges. This type, analogous to the antireflection coatings applied to optical lenses, will be referred to as “interference absorbers.” In the second kind, the material of the absorber is so designed that no reflection takes place at the front surface and the attenuation in the layer extinguishes the entering wave. A continuous gradation from one kind of absorber to the other exists.

Consider the reflection from an infinite plane sheet of material that is characterized by a complex dielectric constant ϵ and a complex magnetic permeability μ . The imaginary part of ϵ arises from dielectric loss and electrical conduction in the medium. The physical cause for the imaginary part of μ in the ultrahigh and microwave frequency range is not known. Let the bounding plane between the medium and air be at $z = 0$. The solution of Maxwell's equations which represents a plane wave incident on the bounding surface, a wave reflected from the surface, and a wave transmitted into the medium is given by

$$\left. \begin{aligned} E_y &= e^{ik_0z} - \alpha e^{-ik_0z} \\ H_x &= e^{ik_0z} + \alpha e^{-ik_0z} \end{aligned} \right\} \quad z \geq 0 \quad (12)$$

$$\left. \begin{aligned} E_y &= A e^{ikz} \\ H_x &= \sqrt{\frac{\epsilon}{\mu}} A e^{ikz} \end{aligned} \right\} \quad z \leq 0. \quad (13)$$

In these equations α is the amplitude reflection coefficient.

The boundary conditions that the tangential components of E and of H be continuous at $z = 0$ immediately yield

$$\alpha = \frac{\sqrt{\frac{\epsilon}{\mu}} - 1}{\sqrt{\frac{\epsilon}{\mu}} + 1}. \quad (14)$$

It is clear that α will be 0 if $\epsilon = \mu$. Material with such a property will serve as an absorber of the second kind, providing ϵ has a considerable imaginary part. If the refractive index n and the absorption index κ are introduced by the familiar relation

$$n + i\kappa = \sqrt{\epsilon\mu}, \quad (15)$$

Eq. (14) can be rewritten

$$\alpha = \frac{n + i\kappa - \mu}{n + i\kappa + \mu}. \quad (16)$$

Let us now examine the behavior of an absorber of the first kind in which the internal reflection occurs from a metal surface at the plane $z = -d$. It will be assumed that κ^2 is small compared to n^2 . The calculation of the resultant reflection can be made by summing the emergent rays and adding this sum to the wave reflected from the plane $z = 0$. Let $g = 4\pi d\kappa/\lambda$ be the damping of a wave for one passage from the front to back surface and return. Let $\phi = 4\pi nd/\lambda$ be the change in phase for the same passage. Finally let β be the transmission coefficient of the front surface. When the index of refraction is high, Eq. (16) gives

$$\alpha = 1 - \frac{2\mu}{n}. \quad (17)$$

If it is remembered that the coefficient for the internal reflection at the front surface is $-\alpha$, the following table can be constructed.

TABLE 3-1.—EMERGENT RAYS AFTER MULTIPLE REFLECTION

No. of passages through the layer	Amplitude of emergent wave
0	α
1	$\beta e^{-\sigma+i\phi} \cdot \beta$
2	$-\alpha\beta e^{2(-\sigma+i\phi)} \cdot \beta$
3	$\alpha^2\beta e^{3(-\sigma+i\phi)} \cdot \beta$
...	...
m	$(-\alpha)^{m-1}\beta e^{m(-\sigma+i\phi)} \cdot \beta$

The resultant amplitude of the wave is then

$$\begin{aligned} R &= \alpha + \beta^2 e^{-\sigma+i\phi} [1 - \alpha e^{-\sigma+i\phi} + \alpha^2 e^{2(-\sigma+i\phi)} - \alpha^3 e^{3(-\sigma+i\phi)} + \dots] \\ &= \alpha + \frac{\beta^2 e^{-\sigma+i\phi}}{1 - \alpha e^{-\sigma+i\phi}} = \frac{\alpha + e^{-\sigma+i\phi}}{1 + \alpha e^{-\sigma+i\phi}} \quad (18) \\ &= \frac{\alpha + e^{-\sigma}(\cos \phi + i \sin \phi)}{1 + \alpha e^{-\sigma}(\cos \phi - i \sin \phi)}. \end{aligned}$$

The minimum reflection clearly occurs when $\sin \phi = 0$ and $\cos \phi = -1$. Hence

$$\begin{aligned} \phi &= (2p - 1)\pi \quad p = 0, 1, 2, \dots \quad (19) \\ d &= (2p - 1) \frac{\lambda}{4n}. \end{aligned}$$

Thus for cancellation the phase change must be π and the layer an odd multiple of a quarter wavelength thick. The minimum reflection will actually be 0 if

$$g = \ln \frac{1}{\alpha} \approx \frac{2\mu}{n},$$

and since

$$\begin{aligned} \kappa &= \frac{gn}{\phi} = \frac{gn}{\pi} \quad (20) \\ \kappa &= \frac{2\mu}{\pi}. \end{aligned}$$

This condition determines the necessary attenuation in the layer in order that the emergent wave will produce complete cancellation. It is qualitatively obvious that the attenuation for each passage must become smaller as the reflection at the front surface becomes greater.

When Eq. (20) is satisfied the power reflection coefficient is given by the relation

$$R^2 = \frac{4\alpha^2 \cos^2 \frac{\phi}{2}}{(\alpha^2 - 1)^2 + 4\alpha^2 \cos^2 \frac{\phi}{2}} \quad (21)$$

If the bandwidth $\Delta\lambda$ of the absorber be defined as the range of wavelengths in which more than half the incident power is absorbed, it can be shown from Eq. (21) that

$$\frac{\Delta\lambda}{\lambda} = \frac{8\mu}{\pi n} \quad (22)$$

Therefore the bandwidth of an absorber is proportional to μ and inversely proportional to n . The qualitative nature of this behavior follows at once from the way in which α depends upon n and μ . For a given refractive index, a continuous transition from absorbers of the first kind to those of the second kind may be effected by allowing the value of μ to range from 1 to ϵ .

Absorbent materials have been produced in Germany for the

radar camouflage of U-boats. The type of absorber that was actually put into service was of the interference kind. The dielectric constant and permeability were produced by a high concentration of spheroidal metal particles (carbonyl iron). The concentration of metal was 80 per

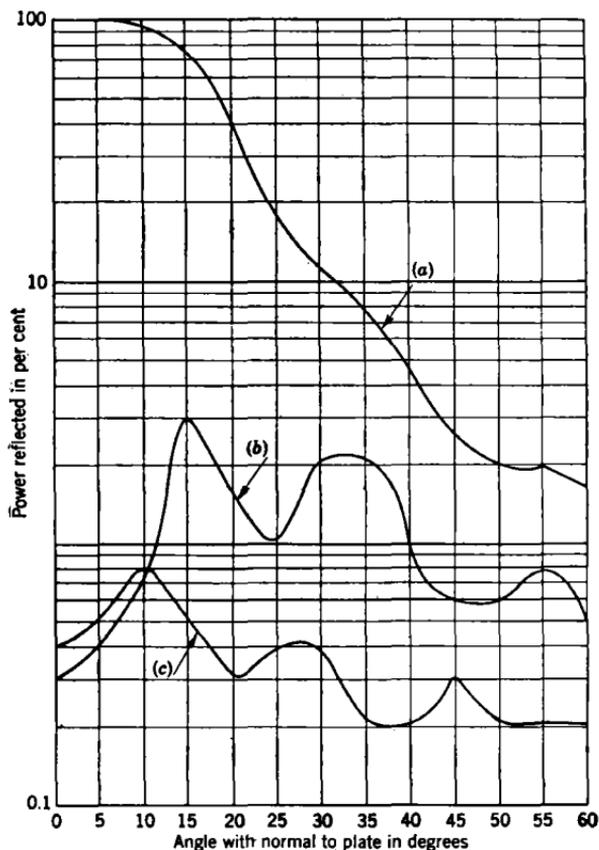


FIG. 3-5.—Effect of absorbing material on energy returned from a flat plate at various angles of incidence. Curve *a*: diffraction pattern of a metal plate, electric vector of radiation normal to plane of incidence. Curve *b*: plate covered with absorbing material of high dielectric constant. Curve *c*: plate covered with magnetic absorbing material of low dielectric constant.

cent by weight, and values of dielectric constant and permeability were $\epsilon = 7$, $\mu = 3.5$.

An absorber of the second kind was also developed in Germany. It consisted of a series of layers whose conductivity regularly increased with depth. The layers were separated by a foam-type plastic whose dielectric constant was close to 1. The absorption was excellent from 4 to 13 cm. However, the complete absorber was a rigid structure $2\frac{1}{2}$ in. thick, and it was never actually used.

When absorbent materials are applied to targets whose cross section is mainly specular, the cross section is reduced by the factor $|R|^2$ —a result that is confirmed by field tests.

The manner in which absorbent materials influence the diffraction cross section is shown in Fig. 3·5. The diffraction pattern of a small plate exhibits a strong maximum at an angle corresponding to that for specular reflection, and at other angles falls off with the oscillations shown. If the plate is covered with absorbing material of high refractive index, the principal diffraction maximum is greatly reduced whereas the secondary maximum is in general only slightly changed. Covering the plate with magnetic material of low refractive index not only reduces the principal diffraction maximum, but also makes the secondary maxima considerably lower. Generally speaking, material of high refractive index can be expected to reduce the specular cross section while it leaves the diffraction cross section substantially unaltered. Magnetic material of low refractive index may, however, also effect a reduction in the diffraction cross section.

COMPLICATED TARGETS

BY A. J. F. SIEGERT AND E. M. PURCELL

Most targets are much more complicated than those dealt with in the previous sections. A distinction will be made between “complex targets” and “compound targets” (Sec. 3·10). The latter term denotes targets consisting of many independent elements (rain, vegetation) which generally fill the volume illuminated by a “pulse packet” (Sec. 4·2) completely. The former denotes complicated targets (such as ships, aircraft, and structures) which are large, but still smaller than the illuminated region. The power received from compound targets is thus dependent on beamwidth and pulse length, since these govern the size of the pulse packet; the signal from a complex target is not.

3·8. Return from Two Isotropic Targets.—The outstanding features of the signal returned from complex targets—its fluctuations and its wavelength dependence—can be studied by considering a simple model consisting of two equal isotropic targets a distance l apart. This distance is assumed to be smaller than $c\tau/2$, where τ is the pulse duration, so that the signals overlap—at least partly. The ratio of the received power from the two targets to that which would be received from one of the targets alone (at a distance large compared with l) is given by

$$\begin{aligned} \frac{P}{P_0} &= \left| e^{\frac{4\pi i}{\lambda} \left(r + \frac{l}{2} \cos \theta \right)} + e^{\frac{4\pi i}{\lambda} \left(r - \frac{l}{2} \cos \theta \right)} \right|^2 \\ &= 4 \cos^2 \left(\frac{2\pi l}{\lambda} \cos \theta \right) = 2 \left[1 + \cos \left(\frac{4\pi l}{\lambda} \cos \theta \right) \right]. \end{aligned} \quad (23)$$

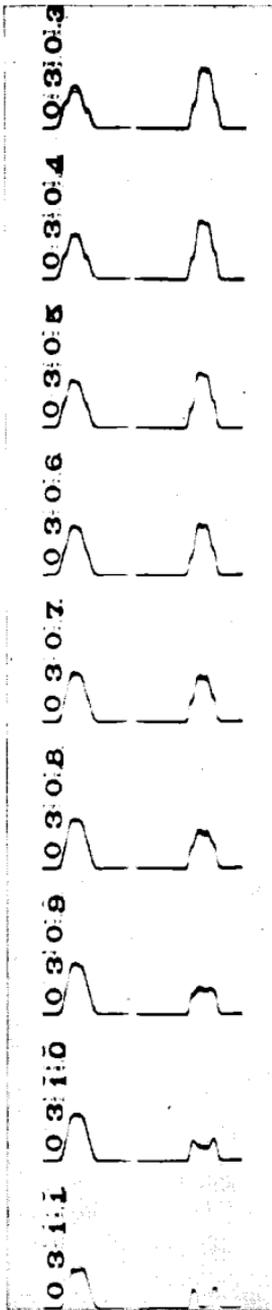


FIG. 3-6.—Successive frames of a photographic recording of an A-scope trace at $\frac{1}{16}$ -sec intervals.

Because of interference, the received power varies between 0 and 4 times the power received from one of the targets.

A beautiful example of the change of signal power with l , as predicted by this formula, is given in Fig. 3-6, which shows successive frames of a photographic recording of an A-scope trace at $\frac{1}{16}$ -sec intervals. These recordings were obtained on Deer Island in Boston Harbor. The signal on the right-hand side is received from the antenna towers of Radio Station WBZ. The signals received from these two towers overlap in the center and there show the effects of destructive and constructive interference. In the last frame (No. 0311) and in the enlarged frame (No. 0479, Fig. 3-7) the parts of the signal where there is no overlapping are separated by a gap of zero power due to destructive interference in the overlapping region. In Frame No. 0303 and in Frame No. 0538, Fig. 3-7, we have full constructive interference. The deflection of the A-scope used was not linear in the power, so that the highest deflection with constructive interference is only 3 times that of the individual signals, whereas the power itself is 4 times the power received from the towers individually. The change in power is due to changes in the distance l arising from a swaying motion of the towers. The other possible cause, change of wavelength due to change in frequency or refractive index, has been ruled out.¹ Correlation was actually observed between the wind velocity and the rate of fluctuation. The towers would have to move only one inch relative to each other to change from constructive to destructive interference.

In Eq. (23) we note further that the return power from a complex target consisting of two equal isotropic scatterers a fixed distance apart depends upon λ , and that the change in signal for a given wavelength change increases with the value of l . This fact has been applied in a

¹ See *Propagation of Short Radio Waves*, Vol. 13, Sec. 6-17, Radiation Laboratory Series.

device to distinguish small targets of large radar cross section from large complex targets such as ships by observing the change in signal intensity caused by a change in wavelength.

3-9. Actual Complex Targets.—The actual targets encountered in the practical use of radar are of a much higher degree of complexity than the simple model just considered. Only a rough estimate of the cross

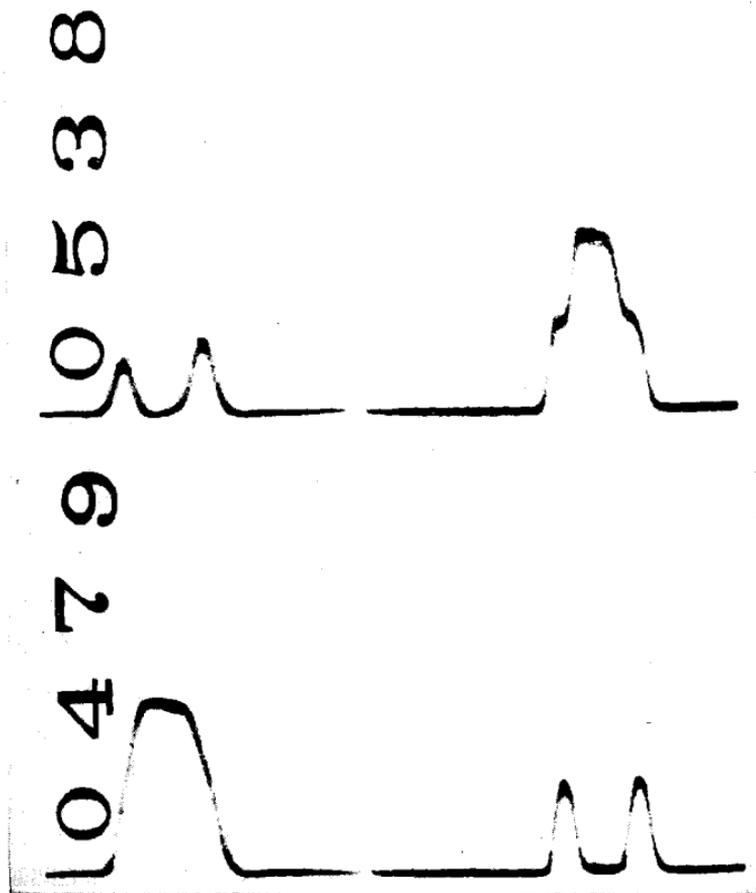


FIG. 3-7.—Enlarged frames of a photographic recording of an A-scope trace.

section of such targets as aircraft or ships can be obtained by calculation. Even if one could carry through the calculation for the actual target (usually one has to be content with considering an oversimplified model) the comparison of calculated and observed cross section would be extremely difficult because of the strong dependence of the cross section on aspect. In Figs. 3-8 and 3-9 the observed return power¹ from a B-26

¹ Ashby *et al.*, RL Reports No. 931, Apr. 8, 1946, and No. 914, Mar. 28, 1946.

and an AT-11 aircraft, respectively, is shown as a function of azimuth angle.

The airplane in each case was mounted on a turntable in surroundings free from other reflecting objects, and was observed with a near-by radar set. The signal strength was automatically recorded as the airplane was slowly revolved at a uniform rate.

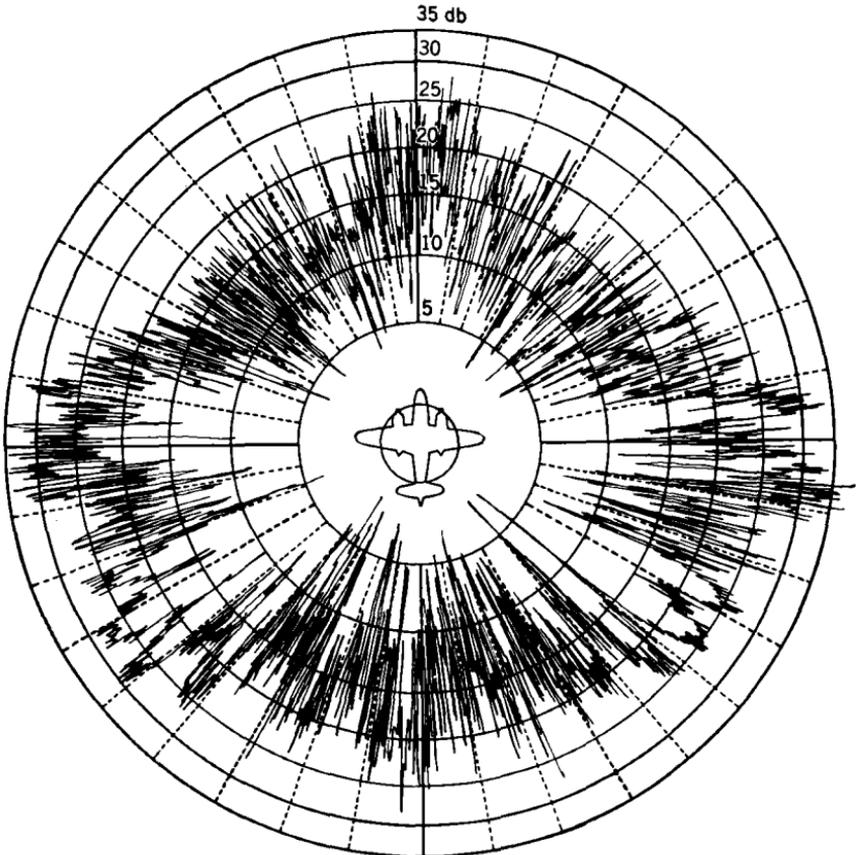


FIG. 3-8.—Return power from a B-26 aircraft at 10-cm wavelength as a function of azimuth angle.

In many positions, the power received changes by as much as 15 db for a change of only $\frac{1}{3}^\circ$ in aspect angle. To a lesser, but still very noticeable, extent the cross section is a function of the position of the propeller, so that modulation¹ of the received power is produced when the propeller is rotating (Figs. 3-10 and 3-11). The modulation is far from sinusoidal,

¹ Ashby, *loc. cit.*

and can be described better as a series of flashes. The power in the higher harmonics decreases only slowly with increasing harmonic number. Because of these variations of cross section, it has been necessary to modify the definition of cross section, in order to make at least the experimental definition unique.

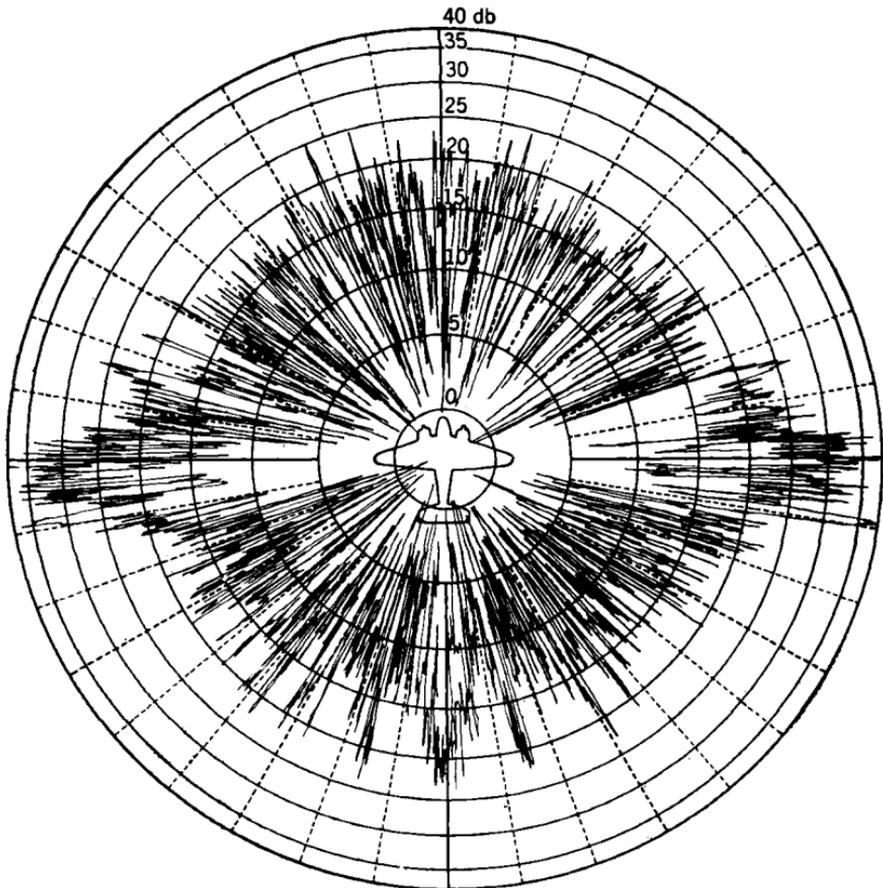


FIG. 3-9.—Return power from an AT-11 aircraft at 10-cm wavelength as a function of azimuth angle.

In RL Report No. 64-10,¹ where a number of measured cross sections are tabulated, the convention is made that the cross section attains the tabulated value during one half of a series of time intervals. The signal is considered as seen in a time interval Δt if at least once during this interval it was distinguishable in the noise. The interval length is

¹L. B. Linford, D. Williams, V. Josephson, W. Woodcock, and supplement by L. B. Linford, RL Report No. 64-10, Nov. 12, 1942.

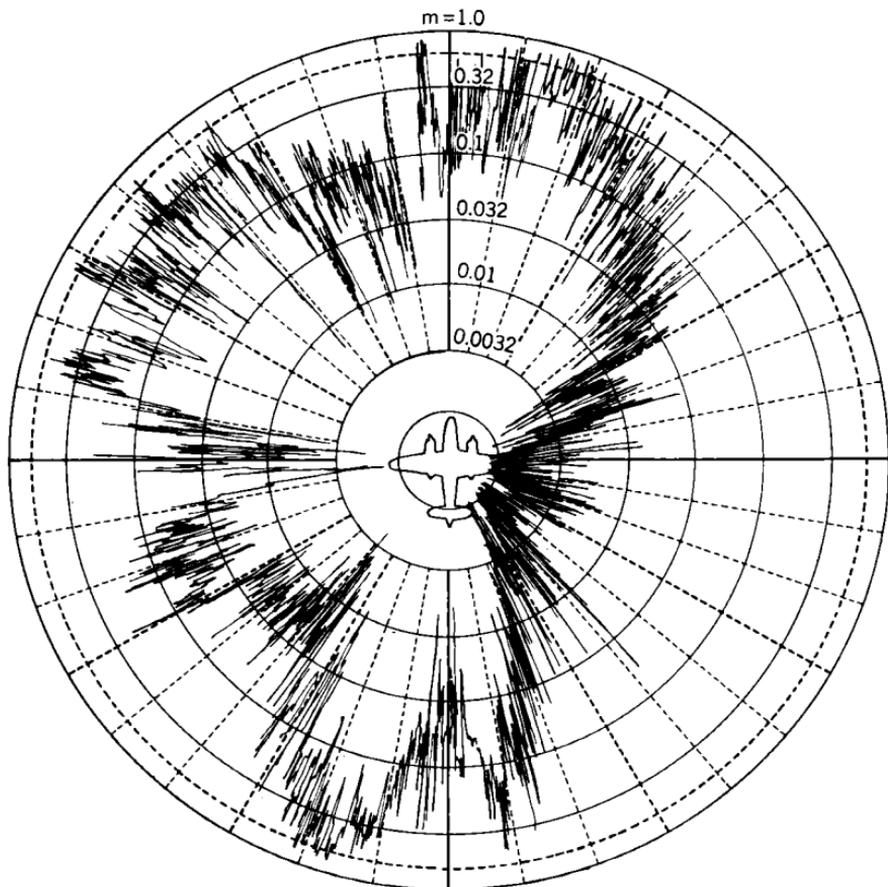


FIG. 3-10.—Fractional modulation of returned signal from a B-26 at blade frequency; left propeller rotating, right propeller stationary.

chosen as 5 sec. The cross section thus defined exceeds the average cross section by an amount as yet unknown.

The values for cross sections in Table 3-2 have been obtained using this definition of cross section.

TABLE 3-2.—RADAR CROSS SECTION OF AIRCRAFT

Aircraft Type	Cross Section, ft ²
OS2U.....	170
Curtiss-Wright 15D.....	410
J2F Grumman Amph.....	440
B-18.....	640
B-17.....	800
SNB.....	230
AT-11.....	200
PBY.....	560
Taylorcraft.....	170

A more serious modification of the definition of cross section is made necessary by the reflection of radar waves when the target is observed at an elevation angle smaller than half the beamwidth and over a water surface. In such cases, the received power is modified by the interference of the direct wave with the reflected wave, both for the incoming and for the outgoing wave (Sec. 2-12). The concept of cross section in the

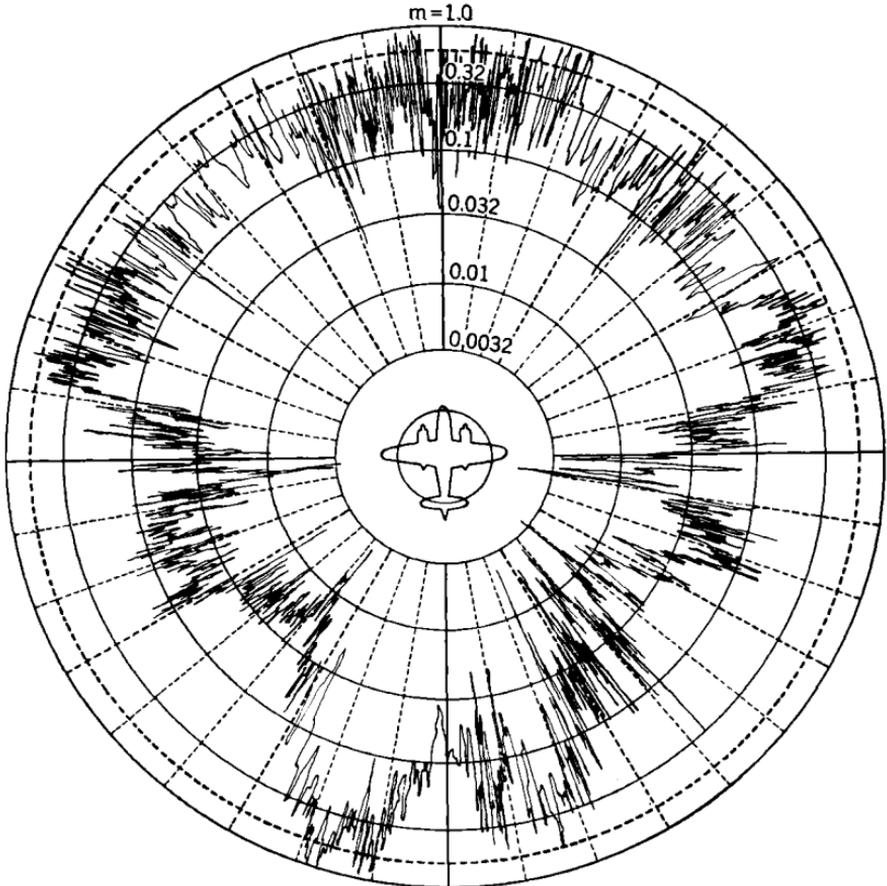


FIG. 3-11.—Fractional modulation of returned signal from a B-26, both propellers rotating.

original sense breaks down in such cases, because it is then no longer possible to define a single quantity characterizing the target—namely the cross section—in such a manner that the received power depends only on the properties of the radar set (P , G , and A), on the range, and on the cross section.

As long as the target does not extend in elevation over more than one lobe (an assumption that is usually correct for airplanes) a modification

of the definition of cross section can be made in a natural and simple fashion if the target is not too near to the radar horizon. Assuming a flat earth with reflection coefficient -1 , the power per unit area at the target is modified by the factor

$$4 \sin^2 \left(2\pi \frac{h_1 h_2}{\lambda R} \right),$$

where h_1 and h_2 are the heights of the antenna and the target, respectively, λ is the radar wavelength, and R the range (see Sec. 2-12). The same factor has to be applied again in computing the return signal to take account of the interference along the return path. The assumption of a flat earth is a good approximation¹ up to ranges somewhat greater than $R_0 = \frac{4h_1 h_2}{\lambda}$. The cross section is thus redefined by the equation

$$S = \frac{PG}{4\pi R^2} \cdot \frac{A}{4\pi R^2} \cdot \sigma \left[4 \sin^2 \left(2\pi \frac{h_1 h_2}{\lambda R} \right) \right]^2.$$

An attempt has been made in RL Report No. 401 to extend the validity of this definition to include ship targets by determining experimentally an "effective value" for R_0 . This is done by observing the range at which the attenuation of the return signal becomes greater than that predicted by the inverse fourth-power law. In this way the cross sections in Table 3-3 were obtained. The procedure amounts to a replace-

TABLE 3-3.—RADAR CROSS SECTION OF SHIPS

Type of ship	Cross section, ft ²	
	$\lambda = 10$ cm	$\lambda = 3$ cm
Tanker.....	24×10^3	24×10^3
Cruiser.....	150	150
Small freighter.....	1.5	1.5
Medium freighter.....	80	80
Large freighter.....	160	160
Small submarine (surfaced).....	0.4	1.5

ment of the real target, which extends over an elevation range from zero to its actual height and whose illumination varies with elevation, by a target at an "effective height" determined experimentally.

A different attempt has been made by M. Katzin,² who computes the

¹ O. J. Baltzer, V. A. Counter, W. M. Fairbank, W. O. Gordy, E. L. Hudspeth, "Overwater Observations at X and S Frequencies," RL Report No. 401, June 26, 1943.

² Navy Report RA3A213A.

effect on the received power of the variation of illumination for a vertical rectangular sheet and uses the formula thus obtained as a definition of cross section.

3-10. Compound Targets Extended through Space.—Targets such as rain, vegetation, "window,"¹ or the surface of the sea are much more complicated than the targets considered in the preceding sections. It is, however, easier to predict and verify certain features of the return from such targets because the compound targets here considered are composed of large numbers of independent individual scatterers, and for that reason statistical considerations can be used in their treatment. The aim of this statistical treatment is to predict the probability distribution of the returned power and of the correlation of successive measurements of the return power.

It will be useful to distinguish between compound targets distributed throughout a volume (rain, "window") and those distributed over a surface (vegetation, waves). In this section the first class is considered, the second being deferred to Sec. 3-11.

The Rain-echo Problem.—Echoes from rain are frequently observed on microwave radar (Fig. 3-12). Under some circumstances, the ability to map out storm areas by radar may be put to good use. In many applications, however, the presence on the radar screen of storm echoes (that is, echoes from rain drops within a storm area) is objectionable for either of two reasons. (1) A small isolated storm cloud at long range may be mistaken for a legitimate target; airborne sea-search radar is especially vulnerable to this kind of confusion. (2) Echoes from a storm area may mask or confuse the echoes from targets at the same range and azimuth.

The nuisance of "rain clutter," which the latter effect is sometimes called, is most severe when the radar cross section of the desired target is small. Other factors that determine the intensity of the rain echo relative to that of the target echo are the beamwidth, pulse length, and wavelength of the radar, the distance to the target, and the number and size of the water drops in the neighborhood of the target. In the case of a radar set using a pencil beam or a simple fan beam, these factors enter the problem as follows:

$$\frac{\text{Average rain-echo intensity}}{\text{Target-echo intensity}} = \left(\frac{R^2 \lambda^2}{A} \cdot \frac{c\tau}{2} \right) \frac{N\sigma_0}{\sigma_t} \quad (24)$$

¹"Window" is the British and most commonly used code name for conducting foil or sheet cut into pieces of such a size that each piece resonates as a dipole at enemy radar frequency. When this material is dispensed from aircraft, large volumes of space can be filled with it. It falls at a speed of only a few miles per hour. The strong signals it returns so effectively mask the radar signals from aircraft that are in the midst of a cloud of window that several tons of aluminum used to be dispensed

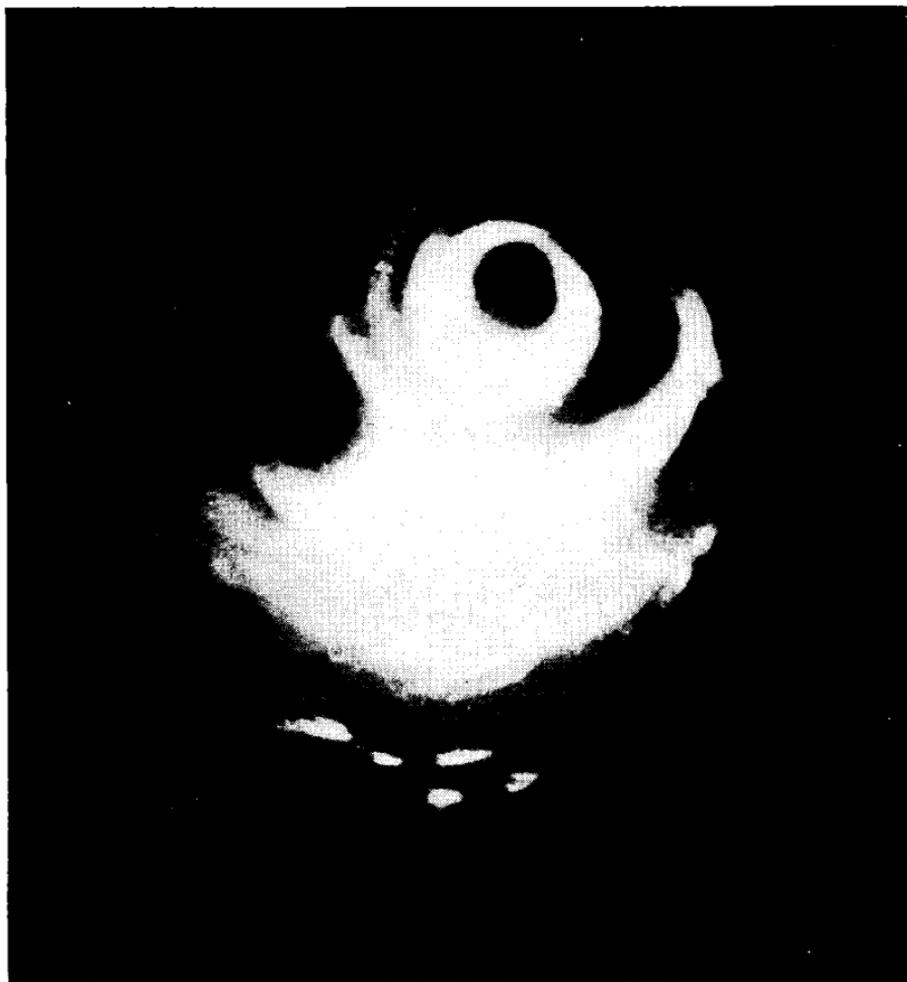


FIG. 3-12.—Echoes from a typhoon on the scope of a 10-cm shipborne radar system. "Eye" of storm is clearly visible above center of scope.

where

R = range

λ = wavelength

c = velocity of light

τ = pulse duration

A = area of antenna aperture

N = number of drops per unit volume

σ_0 = radar cross section of an average drop

σ_t = radar cross section of target.

in this form on each European heavy bomber raid. The U.S. Army referred to this material as "chaff"; the Germans called it "Dueppel."

The quantity in parenthesis in Eq. (24) will be recognized as the volume of the region in space from which, at a given instant, reflected signals can be received. One must assume, to obtain Eq. (24), that the rain is distributed over a region larger than this pulse packet and that the target is smaller than the pulse packet.

The average total power received from all the raindrops that contribute to the return at a given range is the sum of the return powers of the individual drops. The return powers, not the return fields, must be added¹ because the random distribution of raindrops in space results in random phases of the individual contributions.

The cross section σ_0 for raindrops is that of a small sphere of large dielectric constant (Sec. 3-2) and is given by

$$\sigma_0 = 4\pi \left(\frac{2\pi}{\lambda} \right)^4 \bar{a}^6 \approx 300 \frac{\bar{d}^6}{\lambda^4} \quad (25)$$

where a is the radius of the raindrops, d their diameter, and the bar denotes averaging over all drops that contribute to the return. Since neither N nor the distribution in drop size is very well known, it will not be possible to test this formula experimentally with great precision, but the existing measurements of average intensity can be explained, assuming reasonable values of N and \bar{a}^6 .

Using Eq. (25) we can now rewrite Eq. (24) as

$$\frac{\text{Average rain-echo intensity}}{\text{Target-echo intensity}} = 150 \frac{R^2 c \tau \bar{d}^6 N}{A \lambda^2 \sigma_t}, \quad (26)$$

which displays the strong dependence of rain echo upon wavelength and drop diameter.

What has been said above pertains to the *average* intensity of the rain echo. Actually, the signal received from a given region, being the vector sum of the waves reflected from the individual drops, fluctuates continually in amplitude as these drops shift in position relative to one another.² This fluctuation obeys a simple statistical law, which, for our purpose here, can be stated as follows: the probability of receiving, at any time, an echo of intensity (power) P or greater is just $e^{-\frac{P}{P_0}}$, where P_0 is the *average* intensity over a time long compared to the fluctuation time. Precisely the same law describes the distribution in intensity of thermal noise power as amplified by the i-f amplifier of the radar set, and this is one reason for the striking similarity between rain clutter and receiver noise as seen on a radar oscilloscope. There is, however, one

¹ This point is discussed in greater detail in Chap. 7, Vol. 13 of this series.

² The changes of intensity caused by statistical fluctuations of the number of raindrops within a pulse packet are small compared with these changes.

essential point of difference:¹ the rate of fluctuation of the receiver output is determined by the bandwidth of the receiver, whereas the rate of fluctuation, in time, of the echo from a given element of volume in the rain is fixed by the time required for the drops in the volume to assume a new configuration. This is ordinarily so long that several successive radar pulses find the drops disposed in nearly the same way, that is, with changes in relative positions amounting to a small fraction of a wavelength only. A quantitative formulation of this statement is derived in Vol. 24, Sec. 6-2.

The latter effect makes the detection of a desired target echo within the rain clutter even more difficult, for the very persistence of the target echo on successive pulses helps greatly to distinguish it from noise (Sec. 2-10). Such help is of no avail against rain unless the target echo is received over a time long compared to the rain fluctuation period discussed above.

To distinguish a target in the midst of rain clutter, we must make use of some peculiar feature of a raindrop as a radar target. One such feature, perhaps the only one unique to rain, is that the raindrops are round; thus the intensity and phase of the reflection from a single drop do not depend on the direction in which the incident beam is polarized. This cannot be said of most radar targets, which, being complicated objects usually not rotationally symmetrical about the line of sight, show very large variations, with polarization, of the total (complex) reflection coefficient. An experimental verification of the theory that this property of symmetry which is peculiar to raindrops can be used to distinguish a target signal in the midst of rain clutter has been carried out.

Very briefly, the principle is this: if a sphere is struck by a circularly polarized plane wave, formed by passing a linearly polarized plane wave through a quarter-wave plate,² the scattered wave observed in the back-

¹ There are other minor differences: the length in range of a "noise spot" is determined by the receiver bandwidth and the cathode-ray-tube spot size. The length in range of a "rain spot" depends on the pulse length as well. The width, in azimuth, of a noise spot depends only on cathode-ray-tube spot size, or sweep interval, whichever is larger, whereas the azimuthal width of a rain spot depends on the antenna beamwidth.

² Quarter-wave plates and half-wave plates for radar, entirely analogous to those familiar in optics, can be made. What is necessary is to make the phase velocity of the wave as it passes through the plate depend upon its direction of polarization with respect to some direction in the face of the plate. This is done by making the plate of a stack of parallel metal sheets spaced an appropriate distance from one another. When a linearly polarized radar wave encounters the edges of the sheets forming the stack, the component whose electric vector is normal to the edge passes through unaffected. The component whose electric vector is parallel to the edge, however, finds itself in a waveguide of great height but of finite width. Its wavelength

ward direction will be circularly polarized. However, the *sense* of rotation of the vector which represents the field of the scattered wave is such that if the scattered wave passes back through the original quarter-wave plate it will emerge as a wave whose polarization is perpendicular to that of the initial linearly polarized wave (Fig. 3-13). It will therefore not enter the antenna which was the source of the original linearly polarized wave.

Experimental results are shown in Fig. 3-14. The average intensity of the rain echo at a given range was reduced approximately 26 db, while the ground targets (buildings) which were being observed at the same time suffered a reduction of 4 to 8 db. Imperfections of the quarter-wave plate, as well as the slight ellipticity of falling raindrops, prevent complete cancellation of the rain echo.

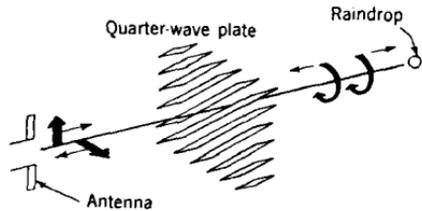


FIG. 3-13.—Scheme for reduction of rain echo.

slight ellipticity of falling raindrops, prevent complete cancellation of the rain echo.

3-11. Extended Surface Targets.¹—In many cases the individual scatterers in a compound target are confined to a relatively thin layer, which can be treated as an extended surface target. Nearly all of the signals received by airborne radar are from targets that fall into this classification. One example is vegetation, which covers most of the land over which an airplane is likely to fly. Another is the diffuse return from irregularities on the surface of the sea.

Let us examine briefly the processes involved in the reception of a radar signal from a layer of scatterers, such as the layer of vegetation on the ground. The radar set transmits a pulse of duration τ , which travels from the airplane toward the target at the velocity c . A particular scatterer will be illuminated for a time equal to the pulse length, as the advancing pulse goes by, and will send a reflected pulse of the same duration back to the receiver. The signal received at a time t , measured from the moment the transmitter begins to radiate its pulse, will consist of contributions from all those scatterers which lie within the antenna beamwidth α and within a range interval ΔR , where

$$\begin{aligned} R' - \Delta R &= \frac{1}{2}c(t - \tau), \\ R' &= \frac{1}{2}ct. \end{aligned} \quad (27)$$

in this guide is greater than the free-space wavelength (Sec. 11-3), and the phase of this component can be advanced over that of the other component by any desired amount, simply by choosing the width of the sheets along the direction of propagation appropriately. See W. E. Kock, "Metal Plate Lenses for Microwaves," BTL Report MM-45-160-23.

¹ By R. E. Clapp.

As we can see from Fig. 3-15, this group of scatterers will lie within a patch of surface of nearly rectangular shape, with its area approximately given by

$$A = (R\alpha)(\frac{1}{2}c\tau \sec \theta). \quad (28)$$

Here we have used

$$R = R' - \frac{1}{2} \Delta R, \quad (29)$$

which is the range to a point in the center of the "target area," the area shaded in Fig. 3-15.

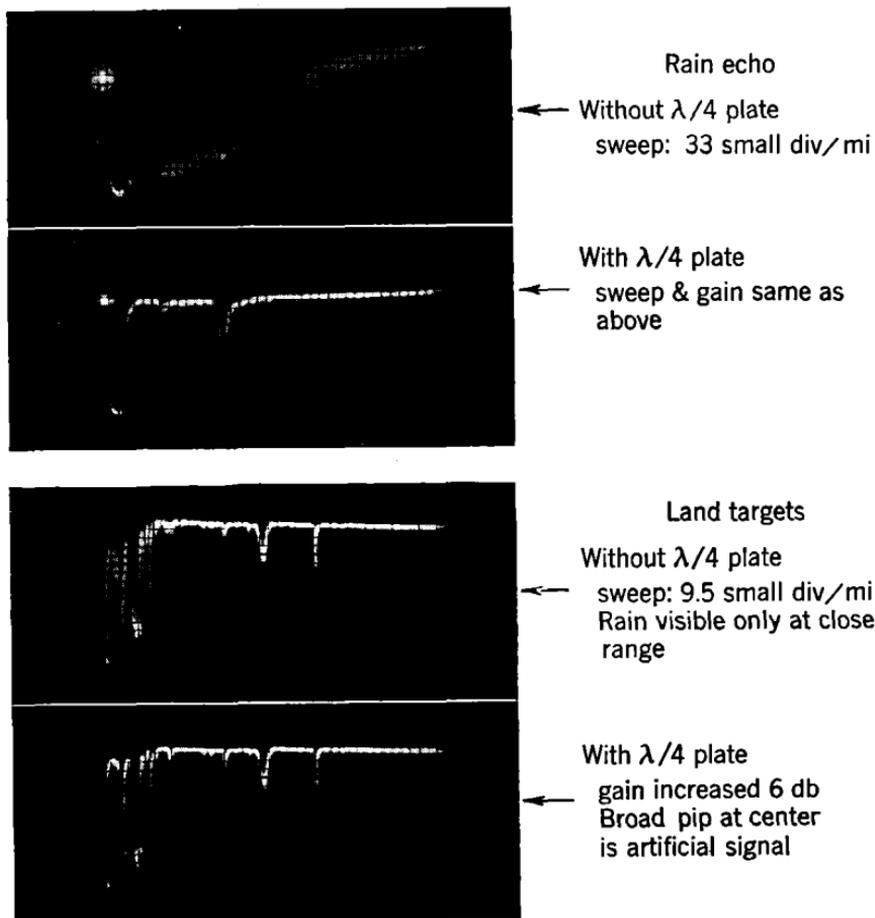


FIG. 3-14.—A-scope observations of rain and land targets.

The radar equation was derived in Chap. 2 for the case of a discrete target with cross section σ , Eq. (2-3b). This equation, repeated here, is

$$S = \frac{P\lambda^2 G^2}{(4\pi)^3 R^4} \sigma. \quad (30)$$

We can generalize this formula to the case of an extended surface target, if we replace σ by an expression whose form will be deduced through statistical arguments similar to those used in Sec. 3-10. As long as the relative positions of the objects in the target layer are random,¹ we add the power received from each object in the area to find the average

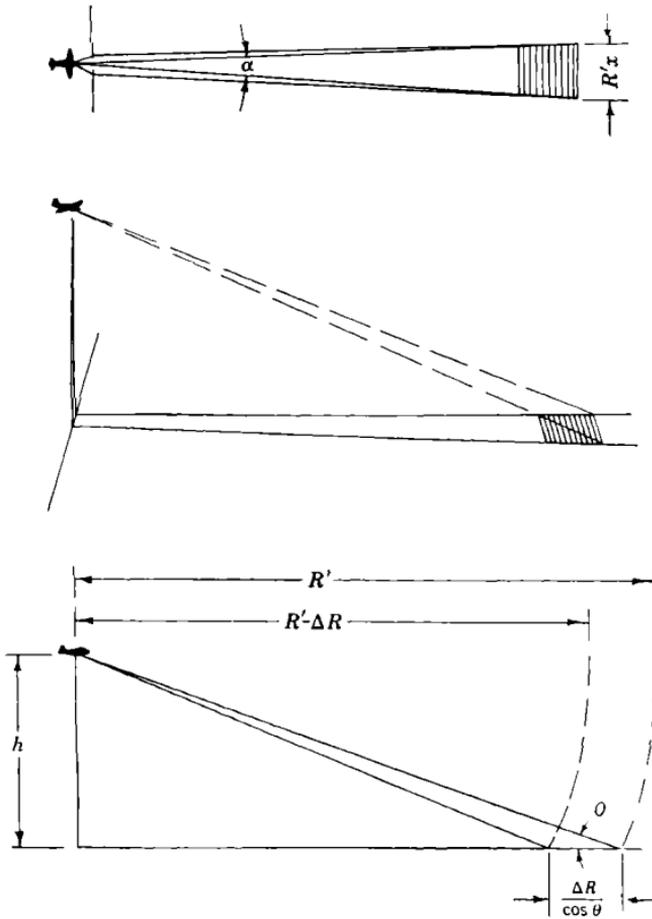


FIG. 3-15.—The area contributing to the instantaneous power received from an extended surface target is limited in range and azimuth by the resolution of the radar system.

level of the probability distribution. The average signal S will thus be proportional to the area A , but it can also be expected to depend on the aspect angle θ , in a way which need not be specified here beyond

¹ Specifically, we require that the wave trains received from the individual scatterers within the target area should combine in random phase, which will be the case if the ranges R_i to the scattering objects are distributed with a randomness or deviation much greater than one wavelength, over a range interval ΔR itself much greater than one wavelength.

denoting it by a function $F(\theta)$. The aspect function $F(\theta)$ will thus include three effects: (1) the amount of power intercepted by the area A will depend upon the orientation of the surface with respect to the incident wave; (2) only a part of the incident radiation will be reradiated diffusely, the rest being absorbed (or, as in the case of sea return, reflected specularly); (3) the scattered power may be reradiated preferentially in certain directions, depending on the properties of the scatterers. If we let $F(\theta)$ include these three factors, the expression for the effective cross section of an extended surface target becomes

$$\sigma = (R\alpha) \cdot (\frac{1}{2}cr \sec \theta) \cdot F(\theta). \quad (31)$$

Substitution of Eq. (31) into Eq. (30) will give the received signal S .

At medium and long ranges (compared to the altitude of the airplane), the factor $\sec \theta$ is approximately unity and can be neglected. On the other hand, for the computation of the altitude signal,¹ Eq. (31) is not sufficiently precise, because the aspect angle θ varies considerably over the large area, directly below the airplane, which lies within the range interval ΔR when R is approximately equal to h , the altitude. In that case, the received signal must be obtained from an integration over the large area contributing to the instantaneous power level in the radar receiver.²

GROUND-PAINTING BY AIRBORNE RADAR

BY C. F. J. OVERHAGE AND R. E. CLAPP

Airborne radar equipment has been extensively used in military aircraft for navigation by pilotage under conditions of restricted visibility. The performance of these radar sets in displaying topographic features below the aircraft depends on point-to-point variations in the radar-reflection properties of the earth's surface. The information contained in the received echo signals is generally presented to the observer as a brightness pattern on an intensity-modulated persistent-screen cathode-ray tube in which radial distance from the center corresponds to slant range or ground range, and azimuth to relative or true bearing. While the coordinates of this plan-position indicator (PPI) presentation thus lend themselves to comparison with maps, the correlation between the brightness pattern and the topographic features of the ground is a matter of varying difficulty, depending on the nature of the terrain, the experience and skill of the operator, and the particular

¹ The first signal to arrive is the reflection from the ground directly beneath the aircraft; it is called the "altitude signal" because its range is equal to the altitude of the aircraft.

² See the chapter on "The Altitude Signal," R. E. Clapp, "A Theoretical and Experimental Study of Radar Ground Return." RI Report No. 1024, 1946.

radar system that is used. The identification and correlation are based upon intensity contrasts which fall into several categories, such as the contrasts between land and water, between hill and valley, and between built-up areas and open countryside. As each type of contrast is discussed, it will be illustrated by radar scope photographs.

3-12. Specular and Diffuse Reflection.—Of the several ways in which airborne radar gives its information, the most important is through the contrast between rough and smooth surfaces. A smooth surface on the ground appears as a black area in the radar picture, while a rough surface appears bright in contrast. In the case of the smooth surface, the incident radiation is deflected away, as in Fig. 3-16. Where the ground is rough, the incident radiation is scattered in all directions, as

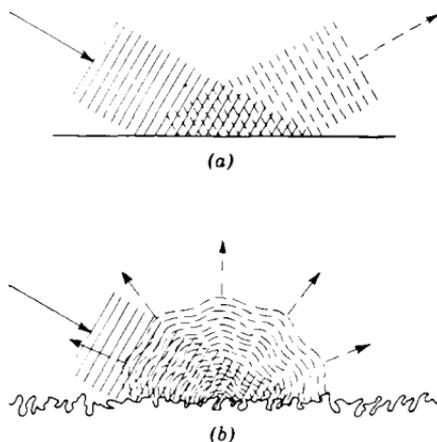


FIG. 3-16.—(a) Specular reflection. (b) Diffuse reflection.

in Fig. 3-16, a part of it returning to the receiver to be amplified and shown on the radar scope.

The signal strength from land areas is so much greater than that returned by water surfaces that the interpretation of land-water boundaries is the simplest of all recognition problems. In a region of highly indented coastline the presentation is so strikingly similar to ordinary maps that navigation by pilotage can be performed even by inexperienced operators. Situations of this type are illustrated in Figs. 3-17, 3-18, 3-19, and 3-20. In Fig. 3-17, for example, the position of the aircraft (given by the spot at the center of the picture) can immediately be identified by reference to the Oakland waterfront, Treasure Island, and the Bay Bridge. Figure 3-20 illustrates the aid in recognition afforded by a broad river valley.

The fineness of the detail which can be shown depends on the range and azimuth resolution of the radar system and on the sharpness of

focus (the "spot size") of the oscilloscope. Figures 3-17 and 3-18 show the resolution of 3-cm systems with an azimuth beamwidth of about 3° , as compared with slightly less than 1° for the 1.25-cm systems of Figs. 3-19 and 3-20. With many radar systems the azimuth resolution is not as fine as the range resolution, so that individual signals appear

N

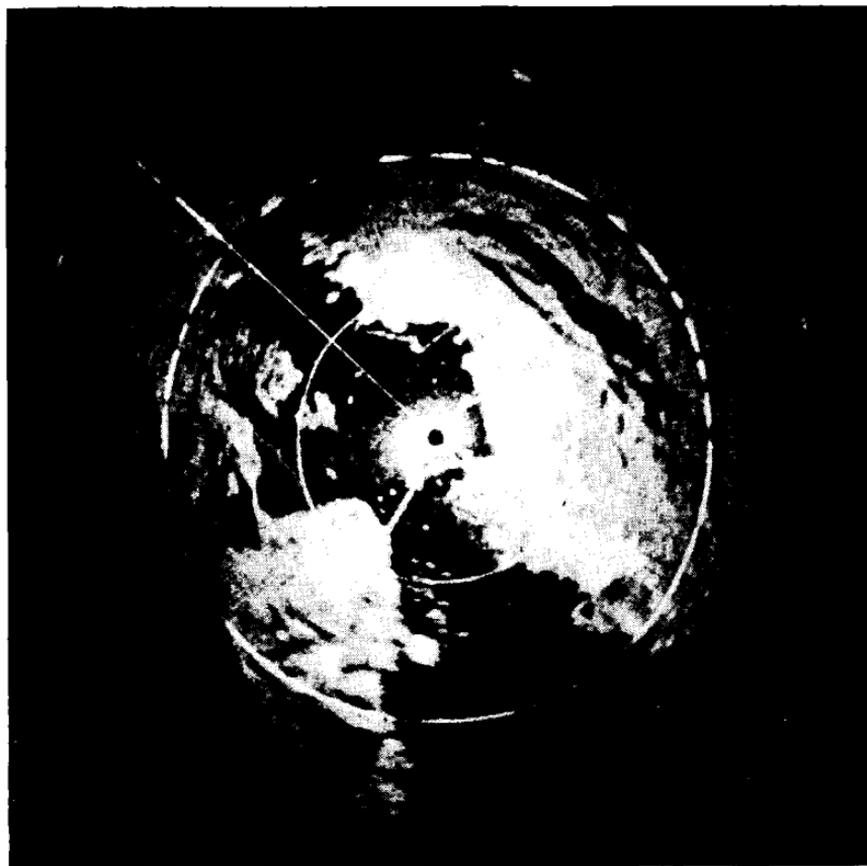


FIG. 3-17.—San Francisco Bay, with Golden Gate Bridge and city of San Francisco at left, Oakland shore at right, and Bay Bridge and Treasure Island near the center of the picture. Wavelength = 3.2 cm, altitude 8000 ft, radius 15 nautical mi. $37^\circ 51' N$, $122^\circ 21' W$.

as narrow circular arcs as in Fig. 3-18. In general, a radar picture may be considered as being painted with brush strokes whose size and shape depend on the beamwidth, pulse length, and oscilloscope spot size. Narrow rivers and inlets can best be resolved when they are seen from the direction which sets them parallel to these brush strokes.

Whether a particular surface appears rough or smooth depends on the wavelength of the radar system. For example, in Fig. 3-21 the paved

airport runways stand out clearly against the grass-covered ground between the runways. At the wavelength of 1.25 cm the grass is thoroughly rough and the runways quite smooth, resulting in strong contrast. The contrast is reduced at much shorter wavelengths (ordinary light), for which the runways as well as the grass appears rough. At much

N

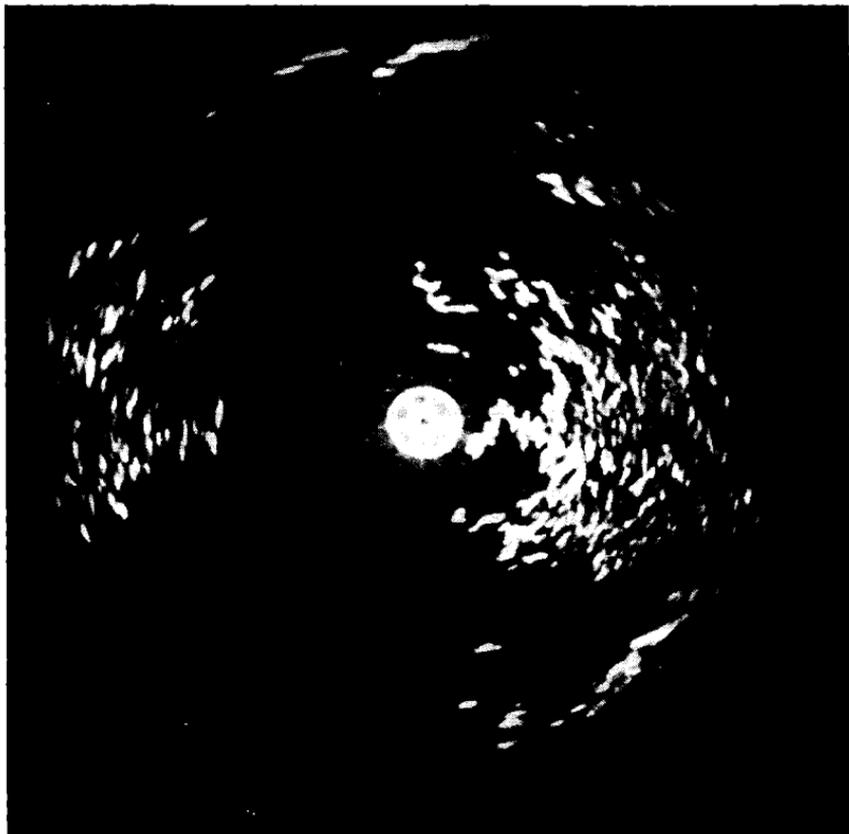


FIG. 3-18.—Coastal areas of Kyushu and Shikoku in the vicinity of Uwajima, Japan. Wavelength = 3.2 cm, 3° beam, altitude 10,000 ft, radius 24 nautical mi. $33^\circ 03'N$, $132^\circ 20'E$.

longer wavelengths both the grass and the runways appear smooth in comparison with wooded areas around the airport. There is still a small signal from the grass and a smaller signal from the paved runway, (because any irregularity scatters a certain amount of power), but on the radar picture a surface that is flat in comparison with the wavelength appears dark in comparison with a rough surface, except when the flat surface is viewed at normal incidence.

At normal incidence the full strength of the reflected wave from a specularly reflecting surface hits the receiver with an intensity much larger than the signal from a diffusely reflecting surface. For this reason the altitude signal is much stronger over water areas than over land.

N



FIG. 3-19.—Tip of Cape Cod, Mass., with Provincetown Harbor. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft. radius 3 nautical mi. 42°03'N. 70°06'W.

3-13. Sea Return and Ground Return.—Many surfaces are neither perfectly smooth nor thoroughly rough. Although a quiet water surface is as near to a specular reflector for microwaves as nature provides, it has been observed that a water surface agitated by wind and tide reflects a strong signal known as “sea return.” Considerable sea return is shown in Figs. 3-17 and 3-22, but it does not interfere with identification of the land areas. Since sea return is usually much weaker than ground return and lacks the sharp outlines of land areas, there is little chance of

confusion between sea and land. Sea return is strongest in the direction from which the wind is blowing. This effect is shown clearly in Fig. 3-22, which not only shows substantial sea return in the upwind and downwind directions (north and south), with no visible return in directions crosswind from the radar, but also shows a much weaker sea return in the protected inner harbor than in the outer harbor.

N

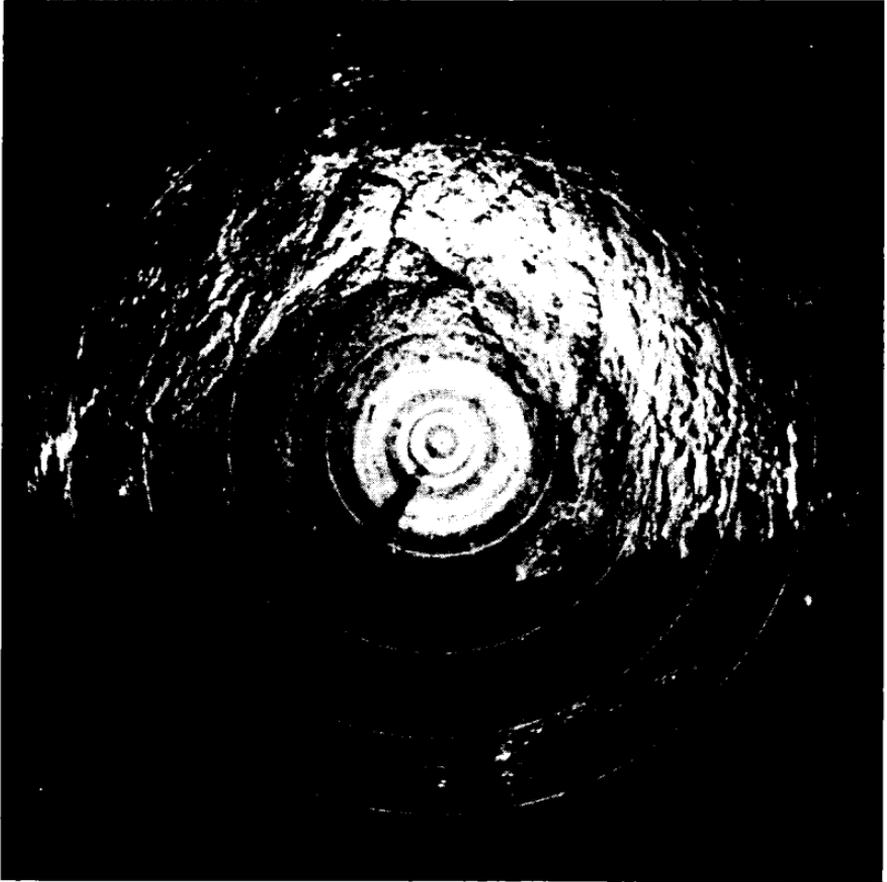


FIG. 3-20.—Connecticut coast between New Haven and New London. Wavelength = 1.25 cm, 0.8° beam, altitude 7000 ft, radius 20 nautical mi. The north shore of Long Island is faintly visible 17 mi. southeast of the aircraft. $41^\circ 23' N$, $72^\circ 30' W$.

For microwave radar most land areas are wholly on the rough side of the division between rough and smooth surfaces—primarily because of the presence of vegetation. Because the ground is thoroughly rough, the aspect function $F(\theta)$ of Sec. 3-11 takes, for ground return, the following simple form:

$$F(\theta) = K \sin \theta, \quad (32)$$

where K is a numerical constant which can be interpreted as the reflection coefficient of the ground. Equation (32) is based on the assumption that every bit of the incident energy that is headed for an area on the ground will strike some object in that area, where a fraction, $(1 - K)$, will be absorbed, and a fraction, K , will be reradiated, all directions of reradiation being assumed equally probable. Some experimental

N

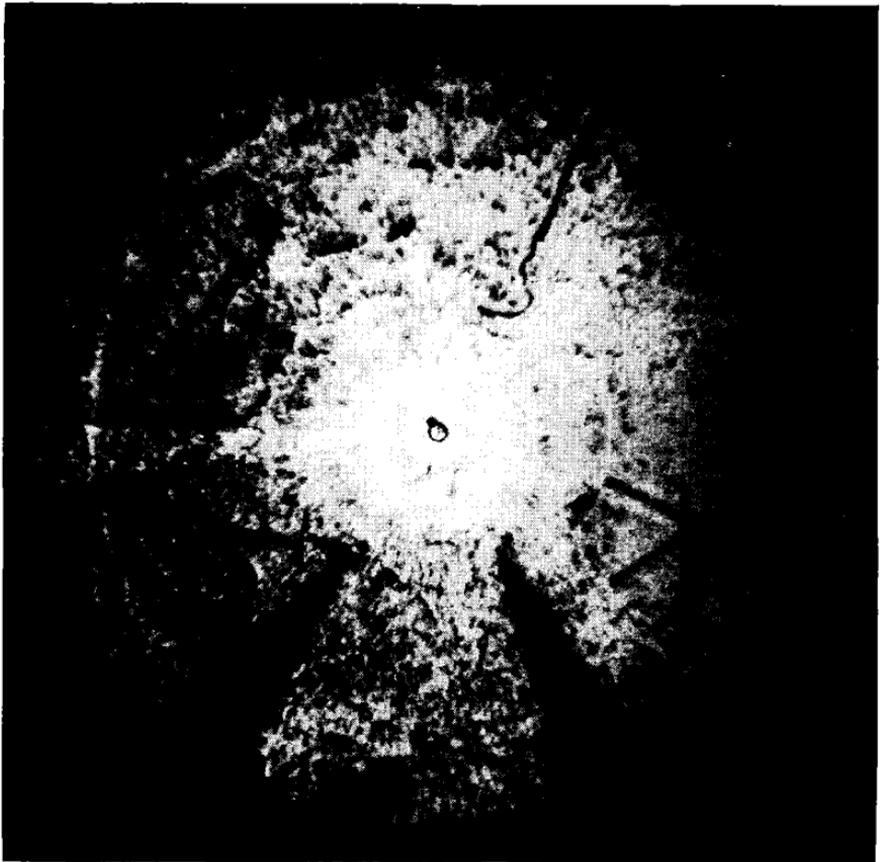


FIG. 3-21.—Airport runways at Bedford, Mass. Wavelength = 1.25 cm, 0.8° beam, altitude 1500 ft, radius 3 nautical mi. The blank sectors to the south are shadows cast by the wheels of the aircraft. $42^\circ 29' N$, $71^\circ 19' W$.

measurements have been made of the radar signals received by an airborne radar system from level, vegetation-covered ground, and of the variation with aspect angle of the intensity of the signals. These measurements are in substantial agreement with Eq. (32).

Introduction of Eq. (32) into Eq. (31) gives the radar cross section of the ground:

$$\sigma = (R\alpha)\left(\frac{1}{2}c\tau \sec \theta\right)K \sin \theta. \quad (33)$$

Since the angle θ that appears here is the angle between the plane of the ground and the line of sight, σ depends on the slope of the ground, being larger for ground that slopes upward (as viewed from the aircraft) and smaller for ground which slopes downward. Two areas of ground which are each level in comparison with the dimensions of σ and rough

N

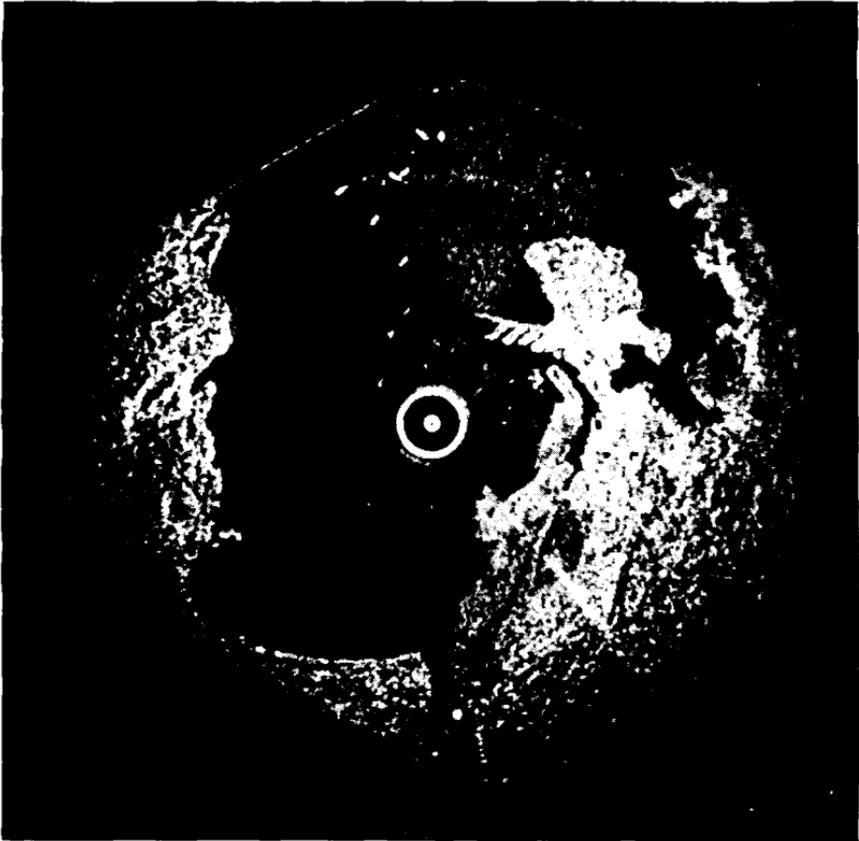


FIG. 3-22.—Atlantic entrance of the Panama Canal, with the docking facilities at Cristobal. Wavelength = 1.25 cm, 0.8° beam, altitude 1500 ft, radius 3 nautical mi. $9^\circ 21' N$, $79^\circ 59' W$.

in comparison with the wavelength, and which have the same reflection coefficient K , reflect signals of equal intensity and must be distinguished from each other through large-scale irregularities such as hills, structures, and bodies of water. To make these large-scale irregularities most easy to detect, it is desirable that the antenna pattern be smooth and properly shaped to produce on the radar picture an even background against which irregularities will stand out by contrast. The shaping of the

antenna pattern is discussed in Sec. 2-5, where an expression for the ideal antenna pattern is derived from an approximate expression for σ :

$$\sigma = (R\alpha)(L) \sin \theta. \quad (34)$$

In Sec. 2-5 it was assumed that L was a constant, and a comparison of Eq. (34) with Eq. (33) shows this to be approximately true, since $\sec \theta$ differs appreciably from unity only for very short ranges. Based on Eq. (34), the ideal antenna pattern is

$$G(\theta) = G_0 \csc^2 \theta. \quad (35)$$

A more exact expression, based on Eq. (33), is

$$G(\theta) = G_0 \csc^2 \theta \sqrt{\cos \theta}. \quad (36)$$

The effect of using an antenna whose pattern fits Eq. (35) rather than Eq. (36) is to increase the strength of the signals at close ranges above those at medium and long ranges. This increase is not large, amounting only to 3 db at 60°; its main effect is to make the altitude signal stronger than the succeeding ground signals. Radar experience indicates that the specification of what curve the pattern should follow is less important than the requirement that the pattern be smooth.

In Fig. 3-20 the diffuse bright rings at the center of the picture are the result of an imperfectly shaped antenna pattern. These intensification rings¹ make it difficult to distinguish irregularities on the ground, but they are not as objectionable as the black rings or "holes" which would have appeared if the antenna pattern, instead of being too strong, had been too weak in those regions.

Equations (30) and (33), combined, give only the average level S of ground return. Particular signals vary widely. On the PPI, with a medium gain setting, ground return can be seen to consist of many bright signals. Among these bright signals are weaker signals, some of them too faint to show on the screen. The resulting stippled or "beaded" texture of ground return is clearly visible when fast sweeps are used, as in Fig. 3-25. If the receiver gain control is set high, most ground signals rise to saturation and ground return takes on the more flattened texture seen in Fig. 3-17.

3-14. Mountain Relief.—The presence of hills and mountains in the area covered by the radar presentation is indicated by the bright returns from the mountain sides facing toward the aircraft and by the shadow regions on the far side of the crests. These conditions follow directly from the geometry of the illumination, and produce a very realistic

¹ The rather diffuse rings of intensification should not be confused with the range markers, which are narrower and appear at equal radial intervals out to the edge of the picture.

effect of observing land forms in relief, although the radial illumination from the center differs from the unidirectional illumination from northwest that is conventionally used in relief maps. Two illustrations of this effect are shown in Figs. 3-23 and 3-24, the former obtained with a 3° beam at 3 cm, the latter with a 0.8° beam at 1.25 cm. Figure 3-23

N

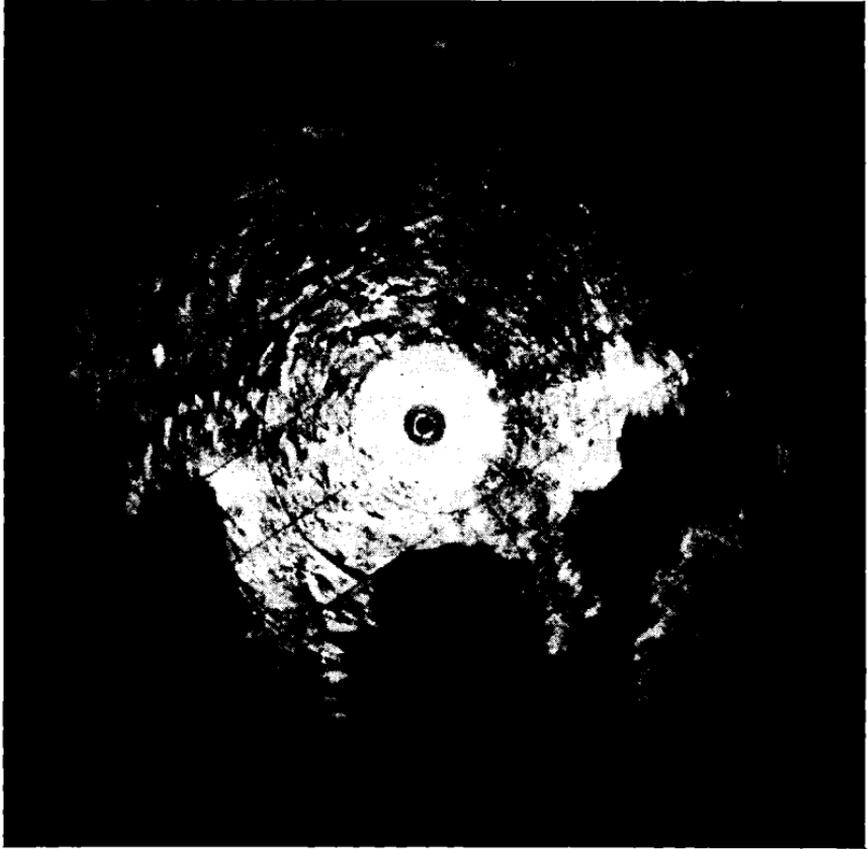


Fig. 3-23.—Tokyo Bay, with Tokyo 21 mi east of the aircraft, at the head of the bay, and Fujiyama to the southwest of the aircraft. Wavelength = 3.2 cm, 3° beam, altitude 30,000 ft, radius 50 nautical mi. $35^\circ 33' \text{N}$. $139^\circ 14' \text{E}$.

was observed roughly midway between Tokyo and Mt. Fuji, and shows the characteristic shadow of the cone. Slightly to the southeast, Mt. Echizen, less than half the altitude of Fuji, throws a semicircular shadow. Many of the crests and valleys north of Fuji are identifiable by reference to a map. In Fig. 3-24 the valleys formed by tributaries of the Susquehanna can be traced clearly in the relief.

It is often useful to visualize the radar presentation of the earth's

surface in terms of a small-scale relief model illuminated by a rotating fan-shaped beam of light, with the source of light held above the model at a height equivalent to the altitude of the aircraft. The direction and shape of terrain shadows and the brightness of the near sides of the mountains are properties of the radar picture which can readily be understood through such scale models.

N

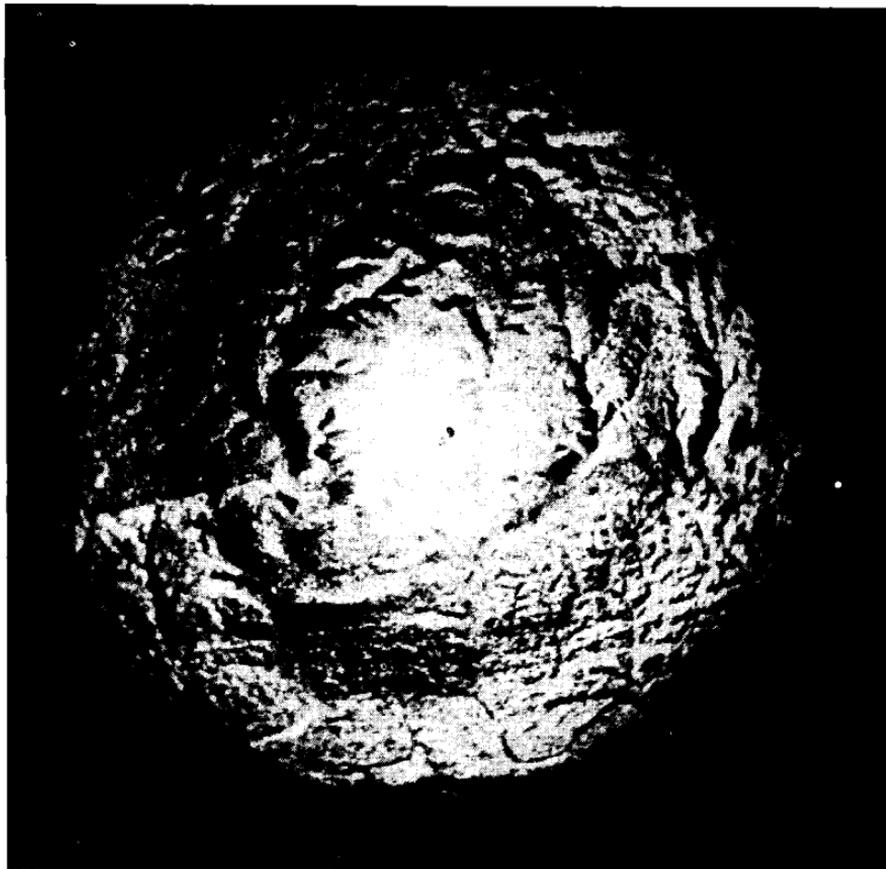


FIG. 3-24.—The Susquehanna River and its tributaries flowing through the Allegheny Mountains near Jersey Shore, Pa. Wavelength = 1.25 cm, 0.8° beam, altitude 9500 ft, radius 10 nautical mi. $41^\circ 19' N$. $77^\circ 16' W$.

As an aircraft approaches a hill or mountain, it will eventually (if its altitude is greater than the height of the mountain) reach a position where the farther slope is no longer hidden behind the crest. Although the shadow is gone, it is still possible to distinguish the mountain on the radar picture through the contrast between the brighter signal from the near slope and the weaker signal from the far slope. In Fig. 3-24, for

example, the contours of the land are indicated partly by shading, partly by true shadows. The identification of hills through intensity shading depends largely upon the radar operator. There is a natural tendency on the part of the operator to set the gain too high, raising all ground signals to saturation and making small variations in intensity indistinguishable.

N

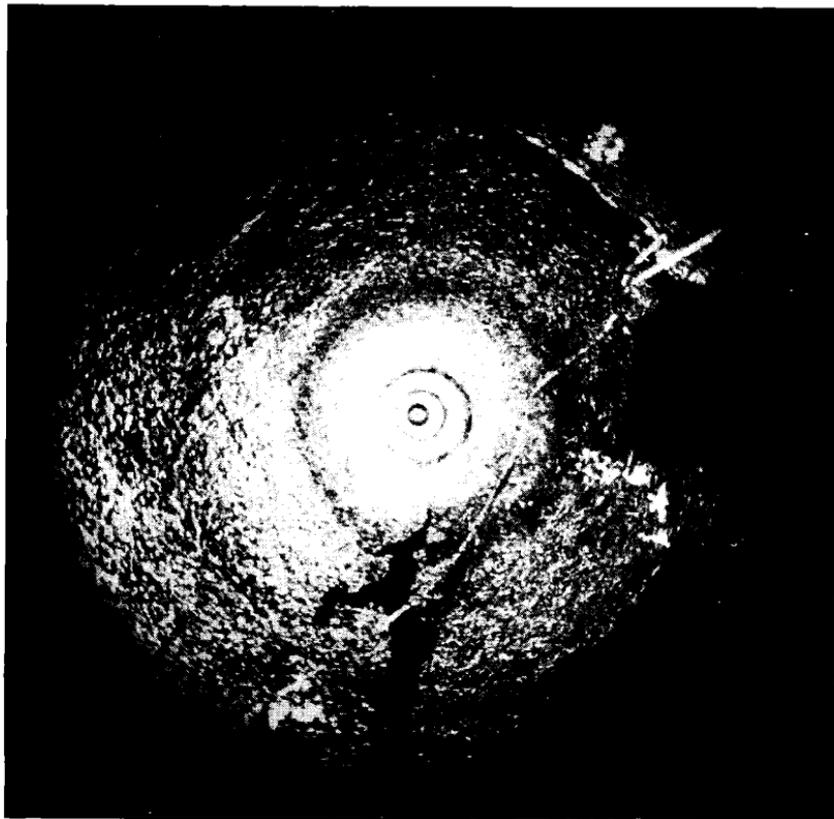


FIG. 3-25.—Mouth of the Susquehanna River at Havre de Grace, Md. Buildings of Edgewood Arsenal and Aberdeen Proving Ground are visible as bright patches south of the electrified railway line. Wavelength 1.25 cm, 0.8° beam, altitude 4000 ft, radius 10 nautical mi. 39°32'N. 76°13'W.

3-15. Structures.—The general appearance of structures and clusters of buildings is illustrated in Fig. 3-25, which was obtained with a high-resolution system in the region between Baltimore and Havre de Grace, Md. The most prominent single feature is the curved line of the Pennsylvania Railroad. Bridges over two branches stand out in sharp contrast with the water. Southwest of the position of the aircraft, near the

periphery of the picture, the buildings comprising Edgewood Arsenal are clearly visible; a similar group southeast of the aircraft corresponds to Aberdeen Proving Ground. A third bridge over the Susquehanna River, just north of the two visible in the picture, is barely suggested by two bright dots at the upper end of the small island in the river. In general, the three bridges appear as equally strong signals; in the particular sweep corresponding to this photograph the third bridge returned a poor signal.

Many structures stand out against ground return because they project vertically above the surrounding level ground and intercept energy intended for target areas behind them. The central portion of the railroad in Fig. 3-25 is an example of a structure which casts a shadow. Seen broadside from a relatively low angle, the wires and supports of the overhead electrification system and the embankment upon which the tracks are laid intercept energy and cast their shadows in the same

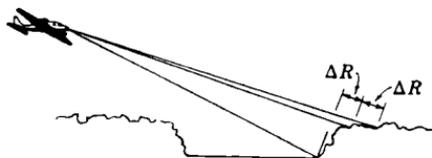


FIG. 3-26.—Increased angle intercepted by inclined target area. The length in range of a pulse packet, ΔR , is $\frac{1}{2}cr$.

way as would a small hill or ridge. This type of contrast is less effective when the target is viewed from a higher angle, for the same reasons that make mountain contrast less effective at high angles of incidence.

Natural structures also show radar contrast. Examples are the rows of trees that line canals and streams in otherwise treeless regions, the hedgerows of Normandy and the cliffs of Dover, or the edges of forests¹ and the banks of rivers. As an illustration, Fig. 3-26 shows the increased angle intercepted by a target area which spans a river bank. In Fig. 3-25, the bright line marking the east shore of the Susquehanna River is characteristic of the strong reflections returned by sharply inclined river embankments.

The radar signals received from man-made structures are often too strong to be fully explained in terms of the solid angle intercepted by the target. Figure 3-25 shows several examples of the bright signals from groups of buildings, without accompanying shadows. In order to account for contrasts as strong as those in Fig. 3-25 a certain amount of retrodirectivity in the target objects themselves must be present.

This retrodirectivity can arise in several ways. Strong specular

¹ The near edge of a forest or group of trees gives a bright signal; the far edge gives a weakened signal or casts a shadow.

reflection will result whenever a flat surface happens to be oriented normal to the line of sight; yet the mere presence of flat surfaces is not enough to guarantee a strong reflection. If these surfaces were oriented in random directions, the probability of finding one at just the right orientation would be so low that the average signal from such a group of flat surfaces would be no stronger than the average signal from a collection of isotropic scatterers filling about the same volume. Therefore the flat surfaces must be so oriented that the reflection is concentrated in the direction of the radar receiver. In a group of buildings, a large proportion of the flat surfaces will be vertical walls, while many others are smooth pavements or flat roofs. There are many opportunities for combinations of three flat surfaces at right angles to form corner reflectors (Sec. 3-5), which are highly retrodirective targets. The full potentialities of these tremendous corner reflectors are never realized in practice because of the strict tolerances imposed by the short wavelengths of microwave radar. Insufficient flatness in the walls makes a huge, imperfect corner reflector behave like one which is smaller but perfect. Inadequate perpendicularity results in several return beams in the vicinity of the aircraft instead of a single return beam pointed directly at the aircraft. Nevertheless, the average effect of many triple corners is to provide retrodirectivity in the radar target.

Vertical and horizontal surfaces can combine into double (rather than triple) corners, giving directivity in elevation under certain conditions. It was shown in Sec. 2-12 that rough surfaces like the ground can serve as satisfactory mirrors for more distant targets when the angle of incidence is sufficiently low. Because double-corner directivity depends on the mirror-like properties of a horizontal surface in front of vertical structures, we should expect strongest signals from these structures when they are seen from low angles. Many buildings or groups of buildings return strong signals at long ranges but tend to fade at shorter ranges when the higher angle of incidence reduces their retrodirectivity. If the line of sight is nearly horizontal, strong signals are sometimes observed by direct specular reflection from vertical surfaces without the benefit of mirror reflection from an intermediate horizontal surface.

Special cases of target directivity in azimuth arise when, for instance, large groups of buildings have parallel walls. The signals in directions perpendicular to these walls are often intensified, as can be seen in Fig. 3-35.

3-16. Cities.—The brightest signals within a built-up area (Boston, Mass.) are presented in Fig. 3-27. This particular photograph was obtained with a so-called “three-tone” presentation (Sec. 13-21), in which gain and limit level are electronically switched back and forth from levels most suitable for land-water contrast to levels giving the

best overland contrast for the brightest target highlights. In this manner, the points yielding the brightest returns are presented over a "base map" which shows the location of the aircraft and which participates in any geometrical distortions of the radar presentation. Studies of this type of presentation show that (1) the location of the brightest

N



FIG. 3-27.—Boston, Mass., with "three-tone" presentation (Sec. 13-21). Wavelength = 3.2 cm, 3° beam, altitude 4000 ft, radius 5 nautical mi. 42°20'N. 71°05'W.

signals changes from instant to instant, and (2) the great majority of these highlights cannot be identified with any particular prominent structures in the city. These bright signals, therefore, must be accidental strong reflections produced at random by favorable illumination of particular surfaces and by constructive interference of reflections from different surfaces within a signal pulse packet or target area. "City return" is similar to ground return and sea return in that each

signal represents the superposition of reflections from the surface elements located within a region bounded by beamwidth and pulse length. The individual targets, however, because of their low absorption and high retrodirectivity, are more effective than the scatterers and irregularities responsible for ground return and sea return.

N



FIG. 3-28.—The Kanto plain north of Tokyo. Wavelength = 3.2 cm, 3° beam, altitude 10,000 ft, radius 28 nautical mi. 35°59'N. 139°59'E.

In spite of the fluctuations of individual signals the average intensity of the return from a city is sufficiently high to form a relatively stable bright area on the screen. The shape of this area and the brightness of particular sections remain sensitive to altitude and direction of approach, but the whole group of signals is strong enough to give a reliable indication under most circumstances.

The appearance of city signals of various sizes is illustrated in Fig. 3-28, a photograph obtained over the Kanto plain north of Tokyo. The

urban area of Tokyo is a large group of signals near the southern limit of the picture, with the industrial suburbs of Urawa and Omiya, northwest of the city, showing as a strong elongated signal. East of Tokyo is Funabashi, a strong signal at the head of Tokyo Bay. Numerous small towns in the Kanto plain are shown as individual arcs, particularly to the west of the aircraft. The Tone River is visible just below the

N

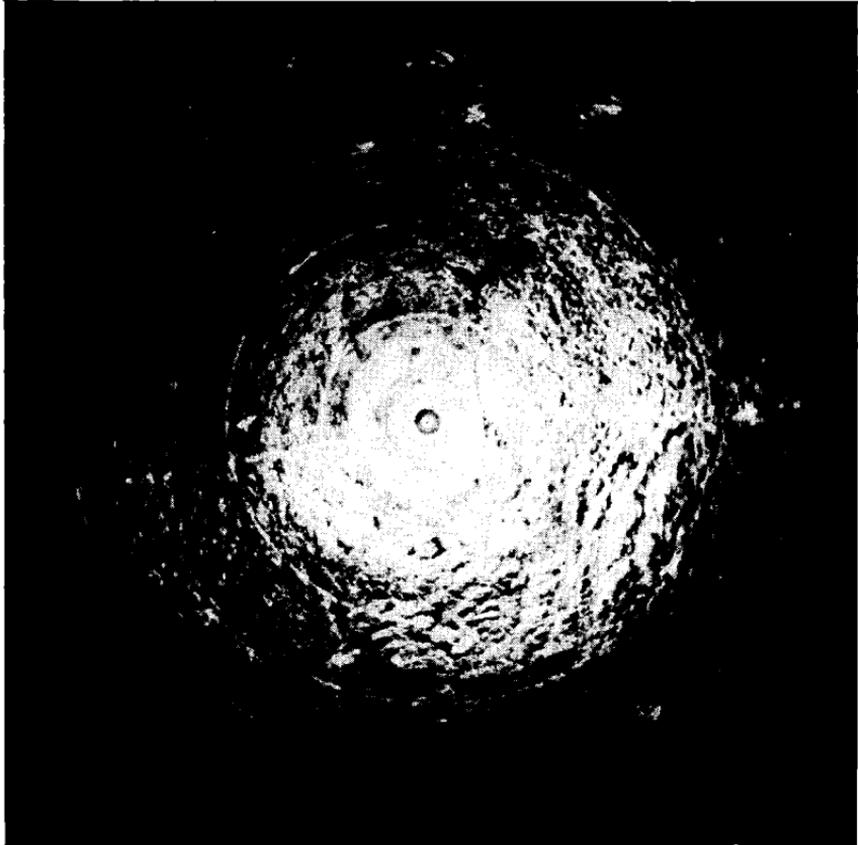


FIG. 3-29.—Small towns near Worcester, Mass. Wavelength = 1.25 cm, 0.8° beam, altitude 7000 ft, radius 24 nautical mi. $42^\circ 16' N$, $71^\circ 48' W$.

center of the picture, with the Toride railroad bridge appearing as a strong signal just south of the aircraft.

Figure 3-29, taken at somewhat closer range with a system of higher resolution, shows a number of small towns in the vicinity of Worcester, Mass. A characteristic group of five bright signals appears near the top of the photograph. The signal nearest the aircraft is a mountain signal, readily identified as such by its triangular shadow. The remaining

four signals represent the towns of Gardner, Fitchburg, Leominster, and Ayer.

Figures 3-28 and 3-29 illustrate situations in which identification depends on the recognition of particular groups or "constellations" of towns; at long ranges the returns from individual towns lack distinctive

N

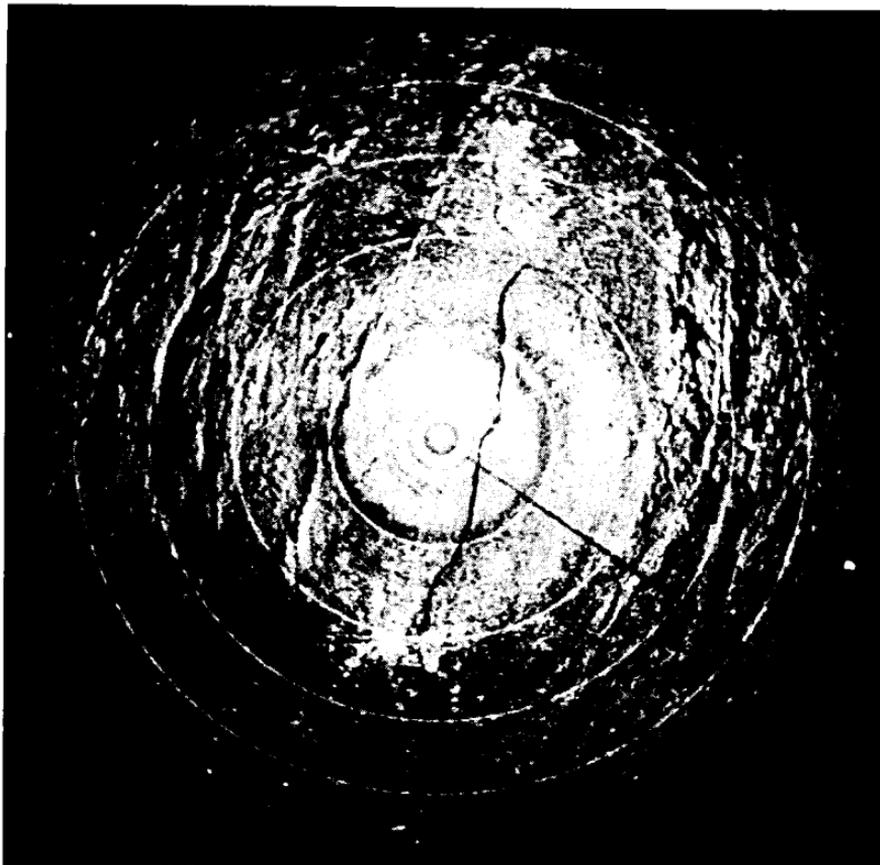


FIG. 3-30.—Part of the Connecticut River valley, with Springfield, Mass., to the north and Hartford, Conn., to the south. Their shapes correspond roughly to the principal built-up areas. Wavelength 1.25 cm, 0.8° beam, altitude 7000 ft, radius 20 nautical mi. $41^\circ 57' N$, $72^\circ 39' W$.

characteristics of their own, although at shorter ranges and with systems of higher resolution such characteristics do appear. Fig. 3-30 shows the cities of Springfield, Mass., and Hartford, Conn., observed at ranges of about 10 miles with a 1.25-cm system. Here the bright returns form characteristic shapes roughly corresponding to the densely built-up parts of the cities. Intermediate conditions are illustrated in Figs. 3-31 and 3-32, which show two Japanese towns as seen with medium-resolution

systems at short ranges. These towns had to be identified by reference to other targets on sweeps of longer range, but the signals begin to show characteristic shapes.

In almost all built-up regions additional information is contained in the detailed brightness variations which become apparent at close range. Many cities contain important and characteristic water areas.

N

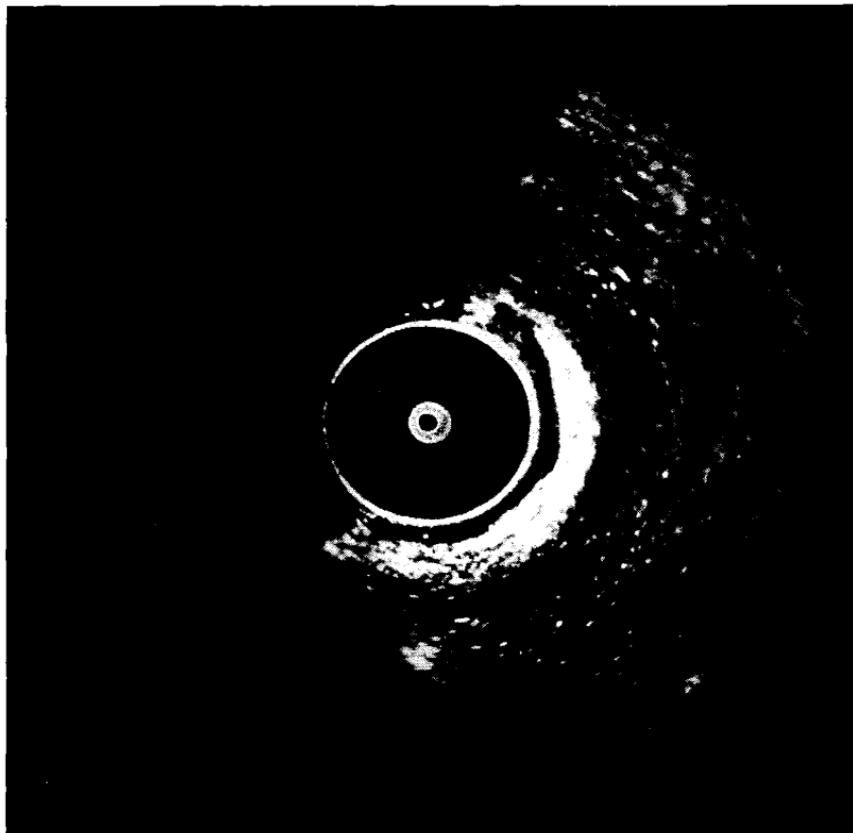


FIG. 3-31.—Kanazawa, Honshu. Wavelength = 3.2 cm, 3° beam, altitude 10,000 ft, radius 10 nautical mi. 36°40'N. 136°40'E.

In Fig. 3-33 the bright region northwest of the center represents the densely built-up area of Boston, Mass. The shape of the Charles River Basin with its two vehicular bridges provides unmistakable identification. A small dark patch slightly closer to the center corresponds to Boston Common, and various water courses together with the shoreline provide identification of various parts of the city. Figure 3-34 shows the further detail which becomes visible at closer range. Figure 3-35 shows a

portion of New York City in which immediate identification is possible by reference to the Hackensack, Hudson, Harlem, and East Rivers, together with Central Park and the various bridges.

It is sometimes possible to discern a few prominent features of the street pattern of a city. A complete presentation of the street pattern would be ideal, but present radar equipment lacks the resolution which

N

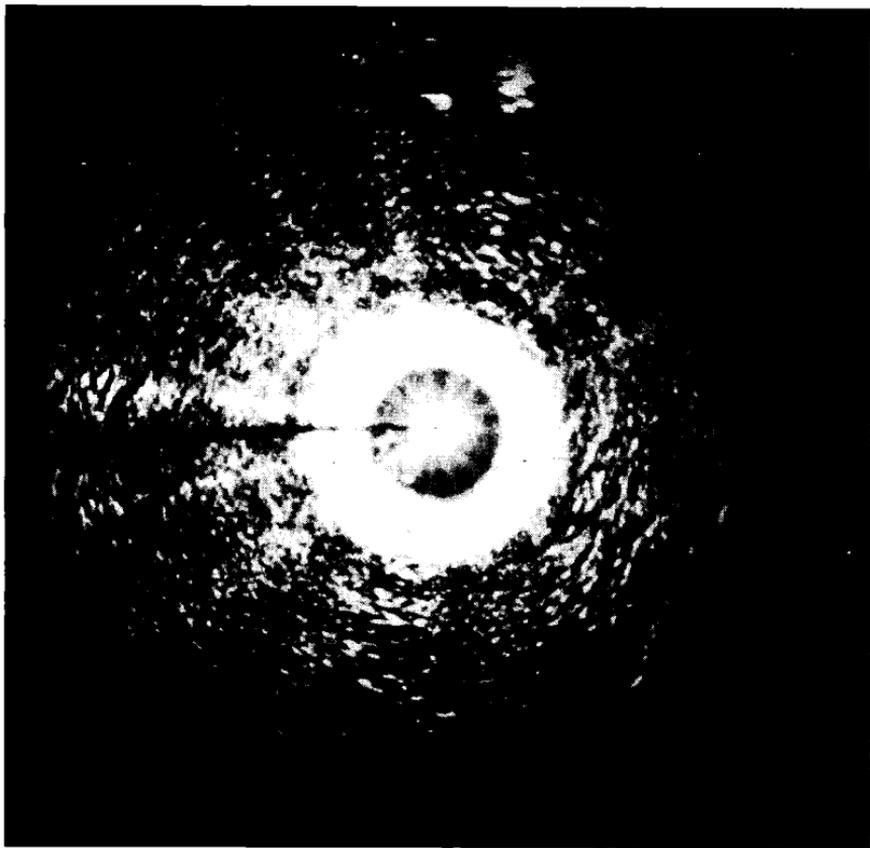


FIG. 3-32.—Obihiro, Hokkaido. Wavelength = 3.2 cm, 3° beam, altitude 12,000 ft, radius 15 nautical mi. 42°41'N. 143°12'E.

would make this possible. At very low altitudes and very short ranges, major thoroughfares may occasionally be seen as dark lines in the bright mass of city return. There is some evidence of the Manhattan street pattern near the center of Fig. 3-35. Airport runways, which are normally much wider than streets, can often be seen clearly at low altitudes, as in Fig. 3-21. On the other hand, the concentration of large buildings along major streets and the presence of elevated railways or

overhead trolley systems often results in a concentration of particularly bright signals along such streets. The street patterns of Chicago and Detroit, partially visible in Figs. 3-36 and 3-37, are of this type. Some caution is necessary in the interpretation of such displays; bright radial lines are occasionally caused by directionally selective reflection from a mass of buildings with parallel surfaces. Such lines appear to move along

N



FIG. 3-33.—Outer harbor, Boston, Mass. Wavelength = 1.25 cm, 0.8° beam, altitude 8000 ft, radius 10 nautical mi. $42^\circ 18' N$, $70^\circ 58' W$.

with the aircraft and can thus be distinguished from streets, which are stable with respect to other signals.

3-17. Navigation.—Sections 3-12 to 3-16 have been devoted to discussions of the kinds of targets and target contrasts encountered with airborne radar. It remains to describe their integration into the radar picture as a whole and the use of this picture as a navigational aid. A fuller discussion of this subject will be found in Vol. 2 of this series.

Airborne radar is best used as a supplement to standard navigational methods, rather than as a substitute for them. Overland navigation by the traditional dead-reckoning procedure can be based upon determinations of ground speed and drift angle made by the radar operator through measurements of the motion of ground targets, while the position of the

N



FIG. 3-34.—Boston, Mass. Detailed resolution of land-water boundaries permits direct comparison with conventional maps. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft, radius 5 nautical mi. $42^\circ 21' N$, $71^\circ 03' W$.

aircraft is checked at intervals against a map by the identification of radar landmarks along the route.

The characteristics of the radar system must be kept in mind when flights are planned. Each radar set has its limitations in range and resolution. The maximum range is limited by many factors,¹ including the transmitted power, the receiver sensitivity, and the antenna gain.

¹ See Chap. 2.

The maximum range also depends upon the type of target and often upon the altitude of the aircraft. Ground return, for example, reflects a stronger signal at higher altitudes for the same reasons that make the near side of a hill appear bright: a given area of ground intercepts a larger solid angle and receives more of the incident radiation. Thus the range

N



FIG. 3-35.—Lower Manhattan, New York, and vicinity of Jersey City, N.J. Rivers, bridges, and railway lines, together with the dark rectangle of Central Park, form an abundance of good reference marks. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft, radius 5 nautical mi. $40^\circ 46' N$, $74^\circ 01' W$.

at which ground return and land-water boundaries can be distinguished increases with altitude, provided the gain of the antenna is unchanged.

Large or highly retrodirective structures are usually visible to considerably longer ranges than is ground return. The difference in ranges is well illustrated in Fig. 3-38, which was obtained with low gain adjustment at an altitude of 4000 ft over the Pennsylvania Railroad main line in New Jersey. The land-painting range in this case extends approxi-

mately to the 5-mile range circle, while the strong reflections from the overhead structure of the electrified railway line can be seen out to 10 miles. A higher gain setting would have resulted in increased ranges both for structures and for ground return, with the maximum ranges limited by the presence of noise on the picture.

Specification of maximum useful ranges for radar systems is difficult.

N

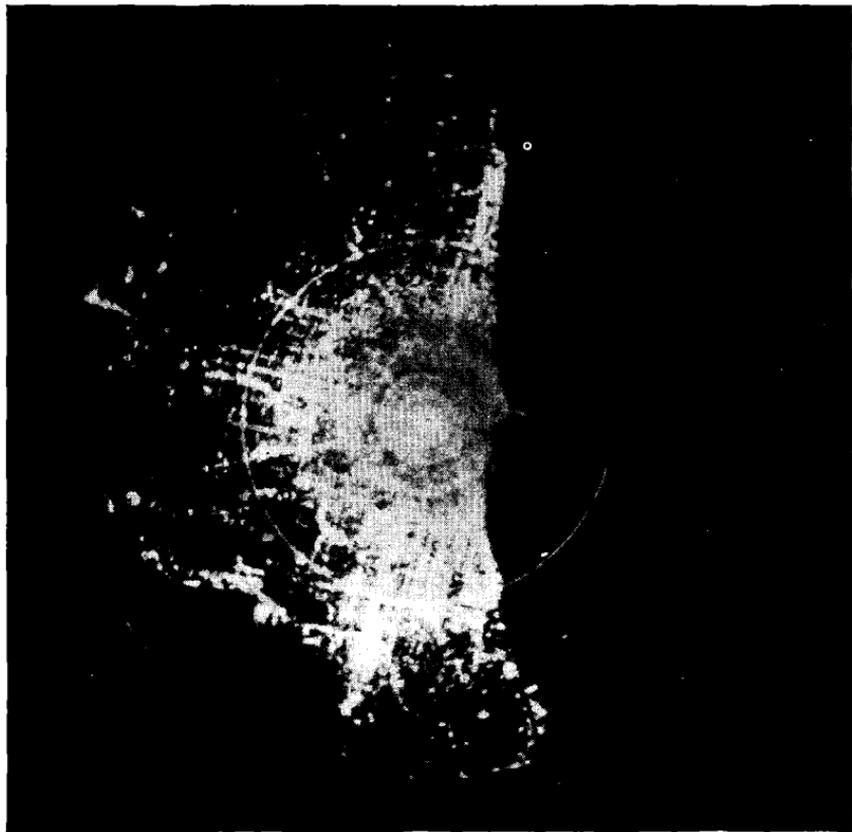


FIG. 3-36.—Chicago, Ill. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft, radius 10 nautical mi. 41°52'N. 87°39'W.

The conditions for rigorous tests are so hard to establish, and the results in general are accompanied by so many cumbersome qualifications, that maximum-range figures have been controversial. Nevertheless, because of the importance of range performance to radar navigation, a few rough figures will be mentioned. At medium altitudes (10,000 ft), with airborne radar systems designed for overland navigation and bombing, ranges of about 40 (nautical) miles on ground return and about 100

miles on large targets are observed with 3-cm systems such as the AN/APQ-13 and AN/APS-15, while with 1.25-cm systems like the AN/APS-22 the ranges available at present are about 15 miles on ground return and 30 miles on large targets, for average atmospheric conditions in the Eastern United States. Under conditions of high atmospheric humidity,

N



FIG. 3-37.—Detroit, Mich. The Ford Rouge River plant shows as a bright patch 7 mi southeast of the aircraft. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft, radius 10 nautical mi. $42^\circ 21' N$. $83^\circ 01' W$.

the absorption of energy by water vapor becomes an important factor limiting the maximum range of a radar set at 1.25 cm (see Sec. 2-15). Even strong signals from large cities are attenuated below noise level if the cities are much beyond the maximum range for ground return. Figure 3-39, which illustrates this uniform range cutoff, was obtained with a 1.25 cm system at an altitude of 20,000 ft near the Pacific entrance of the Panama Canal, with an average water-vapor concentration of

6.9 g/m³. Land-painting extends slightly beyond the 15-mile range circle.

Some radar landmarks are more satisfactory than others. Routes which run parallel to a coastline, a large river, or a range of hills are more easily followed by radar than by contact flying with good visibility, because radar eliminates effects of perspective. Other routes are easily

N

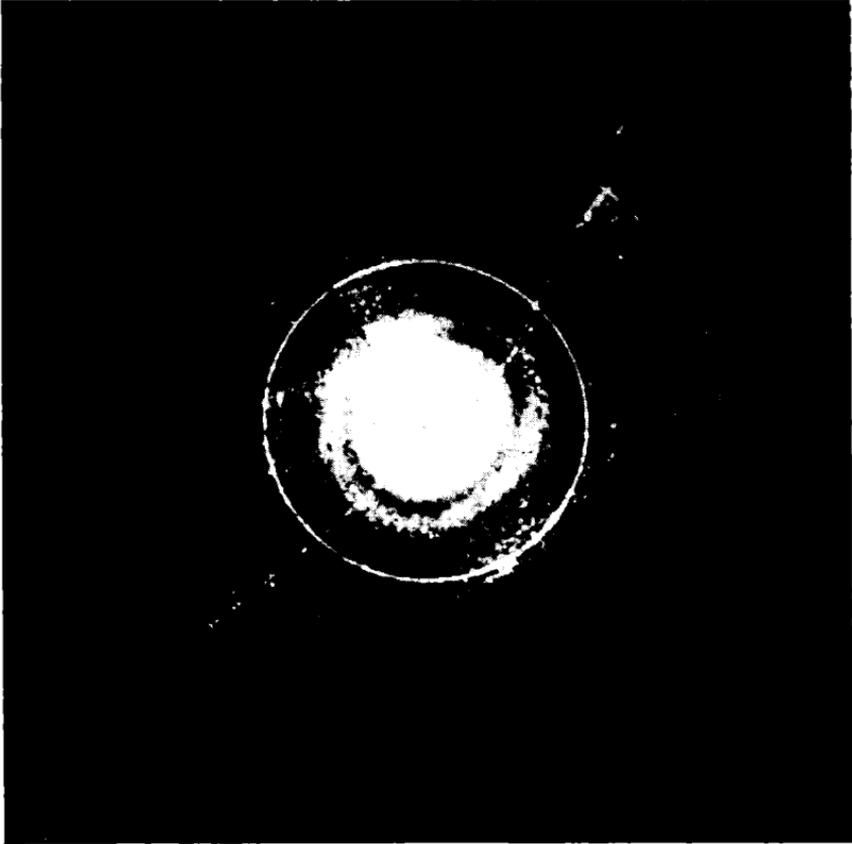


FIG. 3-38.—Main line of Pennsylvania Railroad near New Brunswick, N.J. Wavelength = 1.25 cm, 0.8° beam, altitude 4000 ft, radius 15 nautical mi. 40°22'N. 74°33'W.

followed if they are laid out with the selection of a few distinctive topographical features along the way. A good radar target like a lake may nevertheless be an unsatisfactory radar landmark if there are other lakes in the same region, with which it might be confused. The choice of the best landmarks varies with the radar system, as well as with the terrain. When the radar system has high resolution but a short range or a small field of view, the detailed character of a target is relied on for identifica-

tion, while with systems of less resolution but greater range it is often the spatial relationship of a group of targets which results in their identification.

In mountainous regions the presentation of a given area varies considerably with the altitude and position of the observing aircraft, and skill is required in using radar pilotage over this sort of terrain. How-

N



FIG. 3-39.—Pacific Coast near Balboa, C.Z. Average water-vapor concentration over flight path 6.9 g/m^3 . Wavelength = 1.25 cm, 0.8° beam, altitude 20,000 ft, radius 23 nautical mi. $9^\circ 01' \text{N}$. $79^\circ 35' \text{W}$.

ever, the Fifteenth Air Force of the U.S. Army has used the method extensively in navigating across the Alps. One difficulty inherent in the use of radar pilotage over mountainous terrain is illustrated in Fig. 3-40. West of the aircraft a mountain chain rising to 6000 ft throws long shadows into the Mogami Valley beyond. The town of Yamagata is seen as a bright signal beyond a spur of Mt. Taki extending northwest

from the main range. If the aircraft had been slightly south of the position shown here, this town would have been completely obscured by the long shadow visible just south of it in the picture. Such situations can be foreseen by intelligent map analysis and routes chosen to minimize the difficulty.

Airborne radar is particularly valuable when navigational emergencies

N

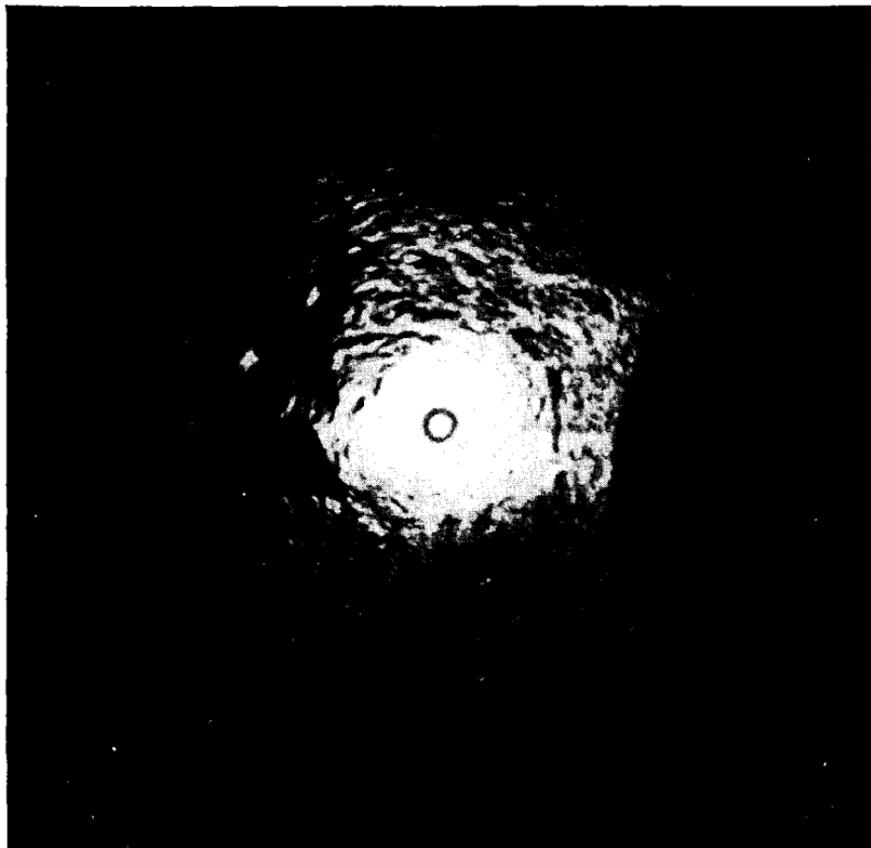


FIG. 3-40.—Sao Mountains, northern Honshu. Wavelength = 3.2 cm, 3° beam, altitude 10,000 ft, radius 30 nautical mi. 38°09'N. 140°38'E.

occur. The reorientation of an aircraft that is off course or the location of an airport for an emergency landing may depend upon the radar system and upon the ability of the radar operator to interpret adequately the brightness pattern on his oscilloscope. Timely notice of the location of thunderstorms is given on the radar screen. The most graphic warning of all is the black shadow extending to the edge of the radar picture, cast by a mountain whose summit rises high enough to threaten a crash.

CHAPTER 4

LIMITATIONS OF PULSE RADAR

BY E. M. PURCELL

To a thoughtful observer, one of the most striking features of a microwave radar screen is the quantity of information that is available at a glance and continually being renewed. In some installations the map of an area of many thousands of square miles is drawn every few seconds. Such a map may comprise in effect some 10^5 to 10^6 separate "elements" of information similar to the elements from which a television picture or a half-tone cut is constructed. This in itself is no cause for complacency; an ordinary photograph, recorded in a fraction of a second, usually contains much more information. Indeed, the uninitiated, comparing the rather fuzzy radar picture with the pin-point detail of the photograph, may conclude that the obvious deficiencies of the former merely betray the primitive state of the art, and that vast improvement in distinctness of detail is to be expected in the normal course of development. These conclusions are only partly true. In the first place many of the unique capabilities of radar, such as direct range measurement or detection of very remote objects despite cloud and darkness, often deserve more emphasis in radar design than does the ability of the set to produce a lifelike picture. In the second place, the pulse radar process is subject to certain inherent limitations. These limitations are of obvious origin. A few have been mentioned already in preceding chapters, but since the interest there was merely in the detection of energy reflected from a single target and not in over-all radar system design, their implications were not pursued.

4.1. Range, Pulse-repetition Frequency, and Speed of Scan.—It is the function of most radar sets to search continually through some region in space by *scanning*. Naturally the radar designer strives always to enlarge the region which can thus be searched, to increase the rapidity with which it can be completely explored, and to improve the ability to distinguish detail within the region. In this endeavor he is made acutely aware of two of the fundamental limitations of pulse radar which can be blamed respectively on the finite velocity of light and the necessity of funneling all information in sequence through a single electrical channel. The effect of these two limitations and the close connection between them can be seen in a simple example.

After each pulse transmission enough time must be allowed for energy to travel to, and return from, the most distant targets, the time required for range R being $2R/c$, or nearly $11 \mu\text{sec}$ for each statute mile of range. If it is supposed that a radar set operates with a *pulse-recurrence frequency* of 500 pps, during the 2000- μsec interval following each pulse echoes will be received from objects within a range of 186 miles. A target 250 miles away may, however, return a signal strong enough to be detected. If it does, this signal will arrive 690 μsec after the *next* transmitted pulse, in exact imitation of an echo from a target at 64 miles. To decide how serious a complication the possibility of a "second time around echo," as it is called, presents, two cases must be carefully distinguished.

1. Targets beyond 186 miles are not of interest. The second-time-around echo is a nuisance only. It would be well to get rid of it if there were some way to do so; if there is not, it would be desirable to identify it as an interloper.
2. The extension of coverage beyond 186 miles is for some reason important. Echoes from targets beyond that range constitute valuable information that ought to be sorted out and presented unambiguously.

If Case 1 applies, the difficulty can hardly be regarded as fundamental. For one thing, there are several ways in which the echo in question can be identified as originating from the preceding pulse. The use of a slightly irregular pulse-recurrence rate, for instance, will prevent overlapping of successive echoes of this type from the same target, without affecting the superposition of echoes from a target lying within the range limit defined above. If a more powerful remedy is needed, the unwanted echoes can be removed altogether by some such scheme as the following one. The frequency of the transmitter may be changed for each pulse, with a corresponding shift in the frequency of the local oscillator of the receiver. For example, the transmitter frequency—that is, the *carrier* frequency—might alternate between two values, f_1 and f_2 , separated by a frequency interval greater than the receiver pass band. Echoes originating from the first transmitted pulse, at frequency f_1 , would not be amplified if received during the interval following the *second* pulse, for during this interval the receiver would be in tune only for echoes of frequency f_2 , and so on. This rather clumsy expedient, although it is actually feasible, would scarcely be justified solely as a means of avoiding a reduction in pulse-repetition frequency. It is mentioned only to show that objectionable second time around, or even "nth time around," echoes could be eliminated if necessary without reduction in PRF and without a drastic change in the radar process.

The situation is different in Case 2, for if it is required to receive both first- and second-time echoes, to separate out the latter, and to present them in proper relation to a range scale, some method must be used which is the equivalent of operation on two or more frequency channels *simultaneously*. Various schemes to circumvent such operation may occur to the ingenious mind, but close examination will show that each is either tantamount to multichannel operation¹ or involves some sacrifice in performance through an abnormally wide pass band, incomplete use of available time, or the like.

Supposing that the reader accepts, with more or less reservation, the force of the range-PRF restriction, we proceed with the example, in which the PRF is 500 pps and the range limit is accordingly 186 miles. Suppose that the width of the radar beam in azimuth is effectively 1° , and that the operation required is a search through 360° in azimuth by rotation of the antenna. Clearly, if the antenna turns at a rate exceeding 500° per sec, during one revolution some narrow sectors in the region to be covered will fail to be illuminated. We therefore conclude that the shortest possible time in which the area can be searched is $\frac{360}{500}$ or 0.72 sec. If the rate of rotation were much faster than $1/0.72$ sec or 1.4 rps, a retentive screen could be used to accumulate information over several revolutions until all gaps were filled, but the time for collecting a complete picture would again be 0.72 sec or longer.

As a matter of fact, rotation at a rate higher than 1.4 rps is ruled out by quite another consideration. In order to receive an echo from the target, the antenna must point toward the target with an accuracy measured by the beamwidth. At a speed of 1.4 rps, in the present example, this requirement is already jeopardized since the antenna will have rotated through just 1° between the transmission of a pulse and the arrival of the corresponding echo from a target at the range limit.

The extreme limiting conditions which have been assumed allow any target to be struck by one pulse of energy, at most, within the time of one complete searching cycle, or scan. In Chap. 2 it was pointed out that the minimum detectable signal power depends sensitively on the number of echoes from the same target which can be accumulated and integrated. The reduction in S_{\min} purchased at the price of increased scanning time by allowing the beam to dwell in the target for several pulse intervals, is almost always worth the cost.

The number of pulses striking the target during one scan is an important parameter in radar design which will reappear frequently in later

¹ The use of different directions of *polarization* as a means of distinguishing one arriving signal from another suggests itself immediately. This would be an effective and elegant method for operation on two channels with common antennas were it not that radar echoes, in general, are substantially *depolarized*.

chapters. This number will be denoted by N_{sc} and defined, where exact specification is required, by the relation

$$N_{sc} = \frac{\nu_r \cdot \Theta}{\omega_{sc}}. \quad (1)$$

In Eq. (1), ν_r is the pulse-repetition frequency, Θ is the width in radians of the radar beam, between half-gain points as usual, and ω_{sc} is the angular velocity of scanning in radians per second. Of course N_{sc} is closely related to the more general and more loosely defined quantity n_s introduced in Sec. 2-11. If N_{sc} is required to be, for example, 10, the rate of rotation of the antenna in the previous example must be restricted to 0.14 rps or about 9 rpm.¹

The numbers arrived at in this example are typical of long-range ground-based microwave radar. Any considerable improvement in all three characteristics—range, scanning speed, and angular resolution—will be blocked by irreconcilable requirements as long as a single radar set is relied upon for the entire coverage. But the restriction falls even more severely on radar systems designed to search rapidly in *two* angular coordinates rather than in one.

Suppose that it is our ambition to design a radar system which will locate in azimuth, elevation, and range any aircraft within 20 miles with the angular accuracy that can be achieved through the use of a beam 2° wide in both azimuth and elevation. The solid angle which the beam itself includes is roughly 0.001 steradians whereas the hemisphere to be searched represents a solid angle of 2π . Some 6000 “patches” in the sky must therefore be covered. The number of radar pulses needed during a complete scanning operation, regardless of the order in which it is carried out, cannot be less than 6000 N_{sc} . The maximum pulse-recurrence frequency consistent with a 20-mile range² is about 4500 pps. Hence the minimum time for completion of the scanning operation is $6000N_{sc}/4500$ or $1.3N_{sc}$ sec. Even if N_{sc} is permitted to be as small as 3, for a complete picture a scanning time of 4 sec is required, which for many purposes is uncomfortably long.

The only physical constant which has been invoked in the foregoing discussion is the velocity of light. The only assumption which has been made about the radar system amounts to this: There is only one channel, through which elements of information in the form of echoes

¹ It should be noted that imposing this restriction incidentally insures that the antenna will not turn too far away from the direction of the target before the echo returns.

² Certain practical considerations which prevent the use of the entire interval between pulses for reception of signals and which therefore set the limit on ν_r , even lower than $c/2R$ (see Chap. 12) are here ignored.

move in single file. It must be concluded that the restriction which thus arises can be overcome only by the use of multiple channels. The simplest—at least in conception—and the most direct way to do this is to use two or more radar systems, apportioning to each a part of the region to be covered.

This method is not the only practical scheme for multichannel operation. In fact, in one of the most important existing applications of multichannel operation, the "V-beam" radar, an increase in the amount

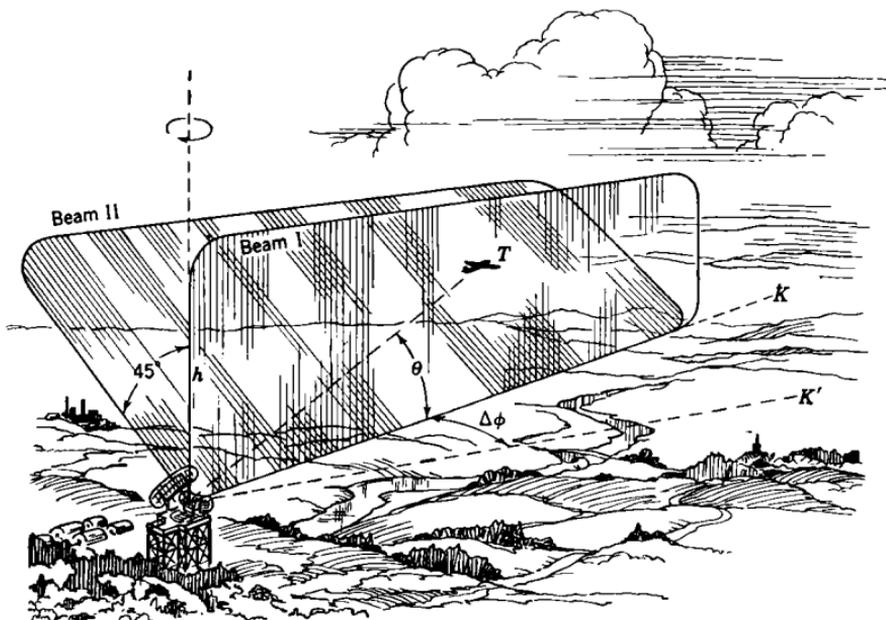


FIG. 4-1.—Principle of V-beam height measurement.

1. Plane of Beam I is vertical. Plane of Beam II falls back 45° from Beam I. Ground edges of both beams are common at K .
2. As both beams rotate together, target T is picked up in Beam I when ground edges of both beams are at K .
3. Target T is later picked up in Beam II when ground edges of both beams are at K' .
4. The angle formed by the advancement of the ground edges from K to K' ($\Delta\phi$) depends on the angle θ of elevation of the target T .

of information provided is achieved in a more subtle and, under the circumstances, a more effective way. The V-beam principle permits a single scanning radar to give *height* as well as range and azimuth of aircraft. In addition to the vertical fan beam which is characteristic of microwave search radar (cf. Chap. 15), the V-beam set provides an additional fan beam that is rotating at the same speed and the plane of which is tilted out of the vertical. If the azimuth and range of a target are known from the first beam, the height can be inferred from the time of appearance of the same target in the second, or "slant" beam (see Fig. 4-1). This latter beam is associated with a completely

separate transmitter and receiver. The V-beam radar is thus an authentic example of the use of a separate radar system to overcome a scanning speed limitation.

In a certain sense, however, the use of the second channel in this way more than doubles the number of angular elements searched in a given time. That is to say, an isolated target can be located in azimuth and in elevation with an accuracy that would ordinarily require a sharp "pencil" beam. But to scan the whole region with such a pencil beam would take, according to the earlier discussion, a very much longer time. It might therefore appear that the V-beam system eludes the very restrictions which have been claimed to be fundamental. This is to be explained by the fact that the V-beam height-finding method works only if relatively few targets appear on the screen at one time; otherwise there is essential ambiguity in the interpretation of the picture. The system is, therefore, not fully equivalent to a pencil-beam scan of the same angular region.

Actually more than two separate radar sets are involved in the V-beam system, for both the vertical and the slant beams are themselves composite. This, however, has nothing to do with the scanning speed limitation, but is required merely to get adequate range and vertical coverage. It is a consequence, in other words, of the radar equation for fan beams discussed in Sec. 2-5.

4-2. Bandwidth, Power, and Information Rate.—If a radar system were used to find only the direction of a target, range information being suppressed, it would be operating very much like a television camera. The picture so obtained would, in effect, be divided into a number of elements equal to the ratio of the total solid angle scanned to the solid angle included in the beam itself. Thus the number of pieces of information that the system can collect per second is simply the number of picture elements multiplied by the number of complete scans per second, exactly as in television. In the radar system giving hemispherical coverage which was used as an example in the preceding section, this number comes to 1500 elements per second, which is not very impressive by television standards. The type C indicator already mentioned in Sec. 2-11 and described in more detail in Sec. 6-6 gives information of just this sort. It has not found wide use for two reasons: (1) the accompanying increases in minimum detectable signal power, explained in Sec. 2-11, and (2) the fact that the method discards a large fraction of the information available in the radar system, the range information.

An even closer approach to the television method has been visualized, in which the proposed radar receiving system is built something like a television camera. A mosaic of microwave-sensitive elements located at the focal surface of a receiving antenna would be scanned by an

electron beam. The system would be handicapped by the limitation on angular resolution imposed by wavelength and antenna aperture, but it should be said that the possibilities of *storage* inherent in the mosaic method might eventually offset to some degree the other apparent disadvantage of the system—its low sensitivity. As usual in c-w radar, knowledge of target range would not be afforded by such a system without further elaboration.

If range is taken into account, the rate at which the radar collects information can be computed in a very direct way. Under ideal circumstances (use of entire interval between pulses, PRF consistent with

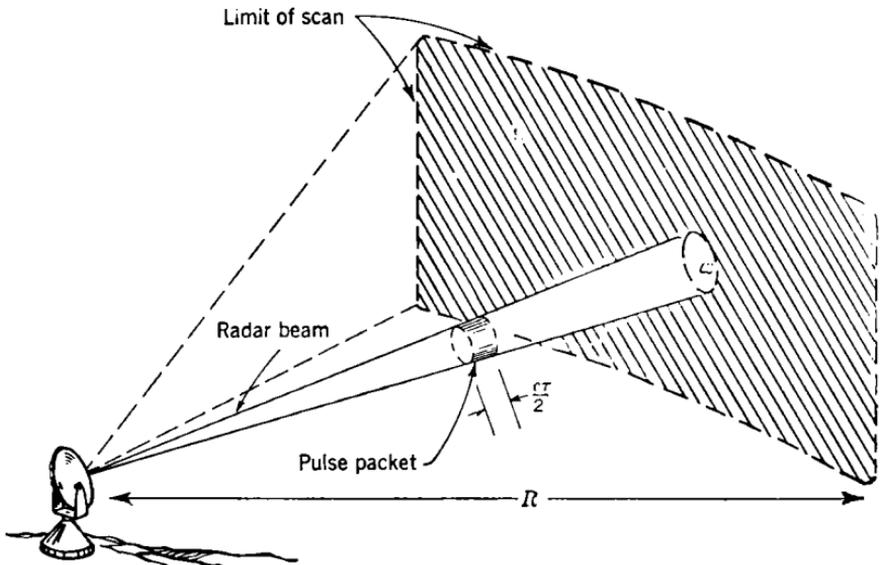


FIG. 4-2.—Analysis into pulse packets of the region scanned. Note that the pulse-packet, which defines a volume from which echoes are received simultaneously, is only half as long as the outgoing wave train.

range limit, and adequate bandwidth in both video amplifier and indicator system), the rate at which information of all types is collected is simply the rate at which separate echo pulses can be received. This rate would normally be about $1/\tau$, diminished by the factor $1/N_{sc}$ which takes account of repetition of pulses from the same target.¹ This number, $1/\tau N_{sc}$, is correct no matter how the information is divided between angular coordinates and range coordinates. The same result can be obtained by observing that the radar system analyzes a region in space by dissecting it into elements of volume, so-called "pulse

¹ To be sure, repetition of pulses from the same target does give extra information in that it helps to distinguish a weak echo from noise. In this section, however, sensitivity is not the primary concern, and it has been tacitly assumed that the echoes appearing on the radar screen are all well above the noise background.

packets," whose dimensions are determined by the angular size of the beam together with the radial distance $c\tau/2$. If Ω (see Fig. 4-2) is the total solid angle scanned in a particular case, ω the solid angle included in the beam, and T the total scanning time, it is necessary to apply the restrictions discussed in the previous section which require that

$$T = \frac{\Omega N_{sc}}{\omega \nu \tau} = \frac{N_{sc} \Omega 2R}{\omega c} \quad (2)$$

On the other hand the number of pulse packets in the volume scanned is $(\Omega/\omega)(2R/c\tau)$. Hence the number of separate volume elements examined per second is

$$\frac{\Omega 2R}{\omega c \tau} \frac{1}{T} = \frac{1}{\tau N_{sc}} \quad (3)$$

Since the pulse duration τ and the bandwidth \mathcal{B} of the radar receiver are ordinarily related approximately by $\mathcal{B} = 1/\tau$, it can also be stated that the rate at which elements of information are collected is of the order of magnitude of the bandwidth \mathcal{B} . This conclusion, which has been approached by a roundabout way, is familiar to the communication engineer, whether he is concerned with voice, facsimile, or television transmission.

Let us see how closely a typical radar system approaches this fundamental limit. Consider an airborne ground-mapping radar with a 1.5° beamwidth and PRF of 1800 pps, scanning at 15 rpm. Suppose that the pulse duration is 1 μ sec and that a region 30 miles in radius is mapped on the indicator screen. By a little arithmetic it is found that N_{sc} is 30 in this case and that the number of separate patches on the ground that are examined in 1 sec is 18,000. The product of these numbers, 540,000, is to be compared with the bandwidth of the system described, which would probably be about 2 Mc/sec. The important point is that the numbers do not differ by orders of magnitude. It must be said, however, that it is not always easy to make full use on the indicator of the information available in the radar system (cf. Secs. 13-20 and 13-21).

4.3. Pulse Radar and C-w Radar.—The last section has a bearing on the relative capabilities of pulse radar and so-called "c-w radar," by which is meant a system operating at relatively low peak power, with a very narrow receiver pass band, and making use of the doppler principle, or of frequency modulation. As was explained in Sec. 2-9, a reduction in pulse power, accompanied by a corresponding increase in pulse length and decrease in receiver bandwidth, leaves the maximum range of a radar system unaffected. Proceeding to a limit in this direction, imagine the pulse to be made so long that it fills the whole interval,

$1/\nu_r$ long, which previously separated successive pulses. At the same time the average power is kept constant by a reduction in pulse power by the factor $\tau\nu_r$. The receiver bandwidth, meanwhile, is reduced to about ν_r cps. The result is essentially a c-w system in so far as power and sensitivity are concerned, with a bandwidth of ν_r . It can see as far as the previous radar system, but it cannot see as much in the same time. Unless it is provided with multiple channels it can collect information only at the rate ν_r .

The reader must be cautioned against taking too literally the result of the above argument, for the hypothetical variation of system parameters cannot be duplicated in practice. A change in pulse power, by a large factor, entails a change in the type of tube used as a transmitter, and perhaps in other components as well. The radar designer, in other words, cannot adjust pulse duration arbitrarily, keeping average power constant; nor can the receiver pass band be made arbitrarily narrow as τ is increased, for effects such as fluctuations of the echo eventually impose a limit. This problem will be discussed in detail in the following chapter which is concerned specifically with c-w radar. The general conclusion, however, remains valid—an essential advantage of high pulse power, as used in pulse radar, is that it permits information to be obtained rapidly.

For certain tasks, no very high information rate is required, and for such applications the various c-w methods (Chap. 5) deserve consideration. In the problem of the radio altimeter, for example, what is required is merely the range of a single, large, ever-present target—the earth below. It is significant that the most important practical application of c-w radar has been made in this field.

4-4. Clutter.—One of the most formidable limitations to the usefulness of pulse radar until recently was that imposed by “clutter.” Often a radar system sees too much, rather than too little; the picture is confused by unwanted echoes, or clutter. This can be made up of echoes from surrounding objects on the ground (ground clutter), of echoes from the irregular surface of the sea (sea clutter), or even of echoes from storm clouds. The problem is to find the desired echo in the midst of the clutter. Although various palliatives have been invented (Chap. 12) which accentuate a fairly strong echo relative to a diffuse background of clutter, this difficulty is plainly fundamental as long as there is no essential difference between the echoes that make up the clutter and the echoes from what we may choose to call the true target. How can the radar system distinguish between the echoes from boulders, ridges, trees, and a multitude of irregularities on the side of a hill, and the echo from a tank moving down the hill? How can the periscope of a submarine be seen against the background of echoes from a considerable area of rough water?

How can a low-flying plane be tracked across a radar map filled with "permanent" echoes?

The only basis that is known for a distinction between the echoes from a particular target and equally strong echoes from its immediate surroundings is the motion, if any, of the target relative to its surroundings.¹ The reflection of a wave by a moving object gives rise to a signal as received in the neighborhood of the (stationary) transmitter, which differs slightly in frequency from the outgoing wave. This is the familiar doppler effect. The frequency difference Δf is but a small fraction of the original frequency f_0 , in the case of electromagnetic waves, being in fact $(2v/c)f_0$, where v is the velocity of the target along the line of sight and c the velocity of light.

Doppler detection systems, or "c-w radar" as they are sometimes loosely called, utilize this principle. In fact the doppler effect and c-w radar were from the beginning so closely identified as to create a rather widespread impression that pulse radar was inherently incapable of capitalizing on this essential difference between fixed and moving targets. This is not true, fortunately, as has been vividly demonstrated by the development of methods, described in Chap. 16, which make possible a distinction between moving targets and their surroundings. In some cases the echoes from the latter are automatically rejected.

This is a spectacular advance in pulse-radar technique but it does not entirely eliminate the clutter problem. It is applicable only to moving targets, and then only when the unwanted echoes from the surroundings do not fluctuate so rapidly as to defeat the pulse-to-pulse comparison which is an essential part of the scheme.

Another direction from which the clutter problem can be attacked is that of resolution. To just this quality is due the great advantage of microwave radar over long-wave radar in respect to clutter. As the radar beam is made narrower and the pulse shorter, the amount of clutter signal superimposed on the target signal decreases. Also, improved resolution often allows recognizable forms to be distinguished in the radar picture. A road or a river may be plainly discernible on a radar map even though any *one* of the many traces making up the picture would be undecipherable. How far we can hope to go in improving resolution depends on factors already discussed, namely the beamwidth-aperture-wavelength relation, considerations of scanning speed, and the practical limits on pulse length and range resolution. Within these limitations, however, there is still much room for development.

¹ One exception to this statement must be made. Echoes from spherical water drops (rain clutter) accurately preserve the polarization of the incident wave, and it has been demonstrated that rain clutter can be almost entirely eliminated by making use of this fact (Sec. 3-10).

The reader may well ask whether a phenomenon has been overlooked which could be used to distinguish some targets from others. There appears to be no possibility for such a phenomenon in the elementary process of reflection of electromagnetic waves from inhomogeneities in the medium through which they travel. A returning wave is characterized by *frequency* (including *phase*), intensity, and polarization. If two targets within the radar beam—for example, a telephone pole and a stationary man—produce echoes similar in the respects listed, they are utterly indistinguishable, as much as we might prefer to label one clutter and the other the true target. Such echoes may very well be identical in the respects listed since no significant difference exists at these frequencies between the electromagnetic properties of a man and those of a piece of wood. To put it another way, the dimension of “color” is not available because the radar cross section of most objects varies in no systematic way with frequency. Distinction by shape, on the other hand, is possible only when the radar beam is considerably smaller in cross section than the object viewed.

CHAPTER 5

C-W RADAR SYSTEMS

BY W. W. HANSEN

5-1. General Considerations.—We shall begin with a discussion that is fundamentally useful in understanding what follows, but has as its immediate object a definition of the scope of this chapter more precise than is given by the chapter title.

The simplest conceivable radar system consists of a source of r-f power, an antenna, and a scattering object or target. The antenna emits waves with time factor $e^{i\omega t}$, which, on striking the target at a distance r , are reflected and returned to the antenna, where their time factor will be $e^{i(\omega t - 2kr)}$ with $k = 2\pi/\lambda$. The returned waves will be reduced in amplitude by a factor α which depends on the target and various geometrical factors, as discussed in Chaps. 2 and 3.

This returned wave will give rise to currents and voltages in the antenna, which add to those produced by the power source and so give rise to a voltage-current ratio, or impedance, which is changed from the value existing when no target is present.

In principle, this impedance, or some similar quantity, might be measured on an absolute scale, and the deviations from the normal, or no-target value, ascribed to the presence of a target.

Practically, this procedure is impossible because the returned signals are often 10^{-9} times as large, on a voltage basis, as the outgoing signals, with the result that the variations in impedance might be of the order of a part in 10^9 . Such variations can hardly be measured in the laboratory at low frequencies, let alone in the field and at microwave frequencies.

It is therefore necessary to cause the returned signals to vary in some manner so that the variation can be measured, rather than to attempt such an absolute measurement as that described above.

This desired variation is caused by a modulation process of some sort. The word "modulation" is used here in a general sense to include changes induced by the target as well as changes in signal introduced at the transmitter. All possible radar systems can then be classified by describing the type of modulation and the use made of the resulting information.

As a specific example of these general remarks, let us take a very simple radar system which consists, as shown in Fig. 5-1, of an r-f power source, an antenna, a rectifier, a high-pass filter, an indicator, and a

target. It is assumed that the target is moving radially, thus supplying the needed modulation.

In consequence of the radial motion of the target, which we suppose to be at constant velocity v_r , the phase of the returned signal, relative

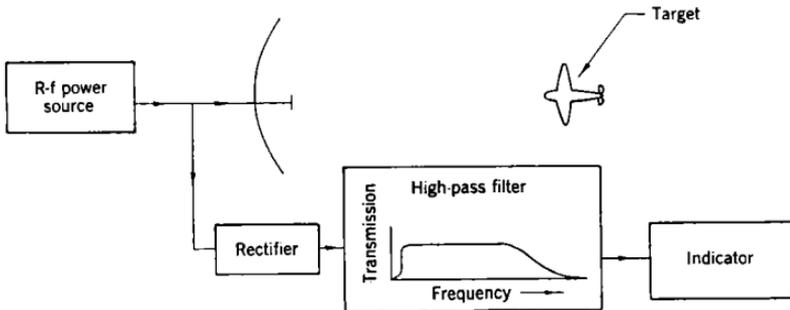


FIG. 5-1.—Simplest possible radar system, using single transmitted frequency, with modulation derived from target motion.

to the outgoing one, shifts continuously. In other words, the returned frequency is different from the outgoing frequency, the difference being $(2v_r/c)f$ with v_r the radial velocity and c the velocity of light.

Numerically, with v_r in mph and f defined by a wavelength λ in cm, the frequency difference, which is called the "doppler frequency," is given by

$$f_D = 89.4 \frac{v_r}{\lambda}, \quad (1)$$

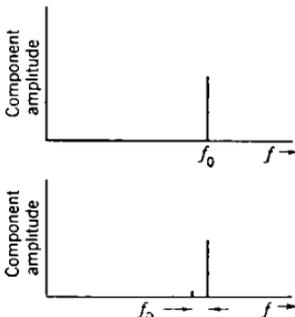


FIG. 5-2.—The upper diagram shows the single Fourier component f_0 radiated by the transmitter; the lower shows the components present in the receiver, these being f_0 from the transmitter and $f_0 \pm f_D$ from the target.

or, at 10-cm wavelength, f_D amounts roughly to 9 cps/mph. The doppler-shifted return signal, when added to the transmitter voltage and rectified, gives rise to a voltage with small pulsations recurring at the doppler frequency. The steady component is removed by the highpass filter, which may be simply a transformer or a series condenser. The fluctuations are amplified and used to actuate the indicator. Thus the presence of a (moving) target is detected.

Although the system just described is hardly practical, apparatus working on exactly this principle but with slight technical modifications can be useful.

In Fig. 5-2 such a system is described diagrammatically in terms of type of modulation and use made of the returned signal. The upper part of the diagram shows the amplitude of the Fourier components radiated by the transmitter; the lower part shows the amplitude of the

components present in the receiver. The modulation is produced by target motion, and the system makes use of this modulation by determining the presence or absence of a signal of doppler-shifted frequency, and usually also, by giving some indication of the magnitude and possibly the doppler shift of this signal.

5.2. Transmitted Spectra.—Consider now a radar system of a general type, as described from the Fourier-analysis point of view, either by a diagram such as Fig 5.2 or, more precisely, by a Fourier series (or integral) for some transmitter voltage or current. Such a series might be written as

$$E = \sum a_n e^{i\omega_n t}, \quad (2)$$

and we know that if $\omega_n = n\omega$, such a series can represent any periodic¹ function. This series is sufficiently general—in fact, that is the difficulty. The series contains more information about the transmitted signal than is commonly available.

To make use of Eq. (1) we must therefore relate the different a_n with various known quantities. A precise and complete connection for all possible forms of modulation might well occupy a book, but sufficient information for present purposes is contained in the brief series of statements that follow.

In the amplitude-modulation case the signal generally consists of a carrier of angular frequency ω_0 , which is modulated at a frequency associated with the repetition rate. Thus $\omega_n = \omega_0 + n\omega_r$ and we have a carrier ω_0 and a number of evenly spaced sidebands, as shown diagrammatically in Fig. 5.3, which shows a spectral distribution resulting from a sequence of pulses. Of course the signal is not strictly periodic unless ω_0 chances to be an integral multiple of ω_r .

The sidebands have significant amplitude over a range that depends on the pulse width, being approximately $1/\tau$ on a frequency scale. Thus the number of important sidebands is roughly $1/f_r\tau$ or T_r/τ . As is

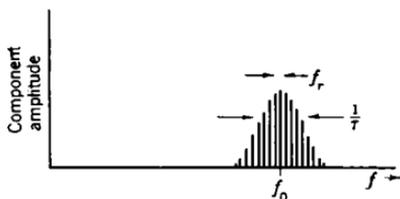


FIG. 5.3.—Approximate frequency spectrum of a sequence of pulses of carrier frequency f_0 , repetition frequency f_r , and duration τ .

¹ The voltages in radar systems are often, and even usually, not really periodic—there is no phase relation between the various pulses. We choose to simplify the situation by ignoring this fact. Ordinarily no harm is done, for in f-m and doppler systems, voltages are periodic, and in normal pulse systems the relative phases from pulse to pulse are ignored. In the special form of pulse system known as the “MTI type” (Chap. 16) the voltages are not actually periodic, but special relations between receiver and transmitter combine to give the same effect.

pointed out in Chap. 4. $1/\tau$ is also the maximum rate at which information can be received.

When frequency modulation is used, the variation of a with ω depends on the form of modulation just as it depends, in the amplitude-modulation case, on the pulse shape. Usually the frequency is varied either (a) sinusoidally or (b) in a sawtooth manner. The results are shown qualitatively in Figs. 5.4a and 5.4b, which are drawn for the normal case in which the frequency deviation is large compared to the modulation frequency. Also, only the envelope of the $|a_n|$ has been drawn. The actual $|a_n|$ fluctuate irregularly below this value.

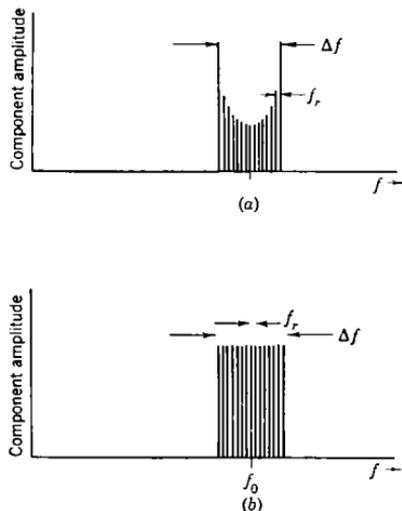


FIG. 5.4.—Approximate frequency spectra for frequency modulation, (a) with sinusoidal and (b) with linear sawtooth variation of instantaneous frequency. More precise figures would show a few sidebands outside Δf , the maximum frequency deviation, and the intensities of the various sidebands as fluctuating irregularly about the mean value here shown.

amplitudes are therefore not equally reduced. This effect can be expected to be notable only when the target extent is comparable to c divided by the bandwidth. Third, there is a phase change, linear in ω , with proportionality constant depending on the distance. Fourth, all frequencies are shifted by the doppler frequency. Finally the scattering cross section varies with time; this introduces an additional modulation which broadens each of the returned sidebands by an amount depending on the rate of the fluctuation. (All this is for a single target. Usually there are many targets, in which case the above specification is further complicated by a summation over all targets.)

For pulse modulation, the phases are such as to make the component waves add to a maximum periodically. For example, if one pulse is centered on time $t = 0$, the phases are all zero. In the frequency-modulation case, the phases are such that the various components add to give a result that varies more or less sinusoidally with time, but with constant amplitude.

Of course there are many possible types of modulation besides simple AM and FM but these will not be treated here.

5.3. Effect of Target.—The effect of reflection from the target on this spectrum will now be considered. First, all amplitudes are greatly reduced, as is discussed in detail in Chap. 2. Second, the scattering cross section of the target is, in general, a function of ω ; the various

Such, then, is the information potentially available to the receiver of a radar set. The amount of this information, which varies with set type, may be enormous, and use is seldom made of all of it. We may complete our classification of possible radar types by specifying what use the receiver makes of this information. For example, in a radar of the usual pulse type, the linear phase change in going from component to component of the returned waves may be just compensated by an increased time. As a result, the returned waves add to a maximum at a time later than do those in the outgoing spectrum. This time delay, which is the observed quantity, may then be regarded as a measure of the rate of change of phase with frequency of the returned components.

5-4. Class of Systems Considered.—In this chapter we will consider that class of systems in which the modulation is such that energy is emitted all, or nearly all, the time. Quite arbitrarily, we take “nearly all” to mean at least 10 per cent of the time—a large percentage in comparison with that usual in pulsed radar, where energy is generally emitted during less than 0.1 per cent of the time. In Fourier terms this may mean, in the case of frequency modulation, almost any bandwidth but phases such as to give constant amplitude. Or, when amplitude modulation is employed, it means that there are, at most, 10 significant sidebands. There may exist only the carrier, as in the simple system described in Sec. 5-1, where the modulation is provided by the target.

In considering systems of this type, we will first describe a number of specific systems and for each system the relevant theory, leaving any generalizations until the end of the chapter. This appears to be the only feasible course since of many possible systems few, if any, will be familiar to the average reader, and general theory would therefore be scarcely comprehensible.

The reader should be warned that in the descriptions of these specific systems quantitative information will be lacking on various points, even on important ones. For example, in the first system discussed, the aural detection of doppler frequencies in the presence of hiss noise is employed. Various questions immediately arise that are similar to those treated in the last part of Chap. 2. Here they are not treated—adequate information simply is not available. This situation and others like it are the result of the fact that very little research has been done on c-w systems in comparison with that devoted to pulse systems.

There are various reasons for this comparative neglect of c-w problems, one of which is certainly valid. All the c-w systems to be described have only a small effective receiver bandwidth (as explained in Chap. 4) and, therefore, a limited rate of information transmission. Obviously this is true for amplitude-modulation systems with 10 sidebands or less and with sideband spacing fixed (see Chap. 4) by the maximum unam-

biguous range, and it will appear later to be true also of frequency-modulation systems.

5-5. Utility of C-w Systems.—In spite of the fact that c-w systems are limited in their rate of information transmission they are of value for a number of reasons. First, there are instances in which a rapid rate of transmission is of no advantage. For example, in the case of an altimeter, there is one target, the earth. Its general direction is known and fresh information as to its distance is hardly needed more than a few times a second. Observation 100 times a second is actually employed in radar altimeters, to make possible the reduction of certain errors by averaging, but even so the rate of information transmission is small.

Second, there are the situations in which, though a little information may be obtained with ease, a lot is impossible to obtain. Such a situation arises in the presence of very severe clutter, where pulse systems, even with MTI equipment, may fail to give any information. Certain doppler systems, on the other hand, will provide useful and even adequate information.

Third, the price of rapid transmission of data is a certain degree of apparatus complexity. In some cases, the gain is not worth the price and the simpler c-w system is adequate.

Finally, there are some things which pulse-type systems simply cannot do—work down to zero range, for example.

In the descriptions of c-w systems to follow, they will be presented roughly in the order of their complexity as to objects, conception, and apparatus. Most of the systems described have seen some use, although a few that are included have been tried only briefly.

SPECIFIC SYSTEMS

5-6. Simple Doppler System.—We describe first a system capable of detecting one or more moving objects in the presence of large amounts of ground clutter. In detail, the specifications called for the detection of an airplane 50 ft above the ground at 10 miles range and aircraft at higher altitudes at 15 miles range. Also, the system later proved most useful for measuring velocities of projectiles.

It will be observed that, in principle, the simple device shown in Fig. 5-1 can do all that is required. Practically, however, two important modifications must be made.

First, steps must be taken to keep as much transmitter power as possible out of the rectifier. There are two reasons for this: (1) the only practical rectifiers at microwave frequencies are crystal detectors and these would burn out if connected, as in Fig. 5-1, to even a low-power transmitter; (2) if the rectifier did survive, it would respond to amplitude modulation of the transmitter and, since it can be found from Eq. (2-4a)

that the returned signal will be 10^{-7} or 10^{-8} times the transmitted signal in voltage, it follows that amplitude modulation of the transmitter would have to be held below this value in order not to be obtrusive. Though perhaps not impossible this would certainly be difficult.

The easiest way to keep transmitter power out of the receiver is to use separate transmitting and receiving antennas, and this was the course adopted. With a moderate amount of work on the antennas, the leakage from transmitter to receiver can be made to be of the order 10^{-5} or less in power. Another power of 10 is canceled by an adjustable leakage path from transmitter to receiver. Further than this it does not pay to go since reflection from nearby ground objects contributes a leakage of this same order.

It has often been suggested that a single antenna would be satisfactory if a bridge-like system were used similar to that used in two-way telephone repeaters. Ordinarily however the single antenna is not satisfactory. For one thing, the increased antenna gain resulting from greater available dish area is lost because of the power used by the "artificial" antenna which balances the real one. More important, since very slight mechanical changes will spoil a 60-db balance between two equal voltages, such bridge systems tend to be highly microphonic.

The second modification relates to the intermediate frequency of the crystal mixer. The reader will have observed that the system is well described as a superheterodyne with zero intermediate frequency, the leakage from the transmitter constituting the local oscillator power and the modulation frequency being the doppler frequency. The only unconventional feature is that the signal is single sideband.¹ Of course, the absolute sensitivity limit of such a system depends on kT and the bandwidth of the amplifier. Experimentally, however, this limit is not even remotely approached because a crystal detector, when passing current, generates a noise analogous to carbon microphone hiss. This noise increases with decreasing frequency and is enormous compared to thermal noise for audio frequencies. To avoid this excess noise a local oscillator is introduced and amplification done at some normal intermediate frequency, 30 Mc/sec for example. At this frequency the excess noise is negligible. In the i-f amplifier a strong component is found due to the beat between leakage from the transmitter and power from the local oscillator, and a much weaker component, due to the target, displaced by the doppler frequency. After suitable amplification these two frequencies are passed into a second detector whose output signal consists of a d-c component associated with leakage from the

¹ It is not difficult to make systems that determine on which side of the carrier the sideband lies, and are thus able to discriminate between approaching and receding targets.

transmitter, and the doppler frequency. The d-c component is removed by passage through a transformer, or otherwise, and the doppler frequency is used to actuate the indicator.

An incidental but important advantage of this modification is that it enables the use of a suitable and easily adjustable crystal current.

These modifications are introduced in a manner indicated in Fig. 5-5 which shows a block diagram of the system. Everything is straightforward, except for the local-oscillator power which is obtained by modulating the transmitter and selecting a suitable sideband by means of a bandpass filter. This insures a constant intermediate frequency and avoids any tuning problems. Frequency-doubling is employed in the modulator to avoid difficulties in keeping leakage from the crystal oscillator out of the i-f amplifier.

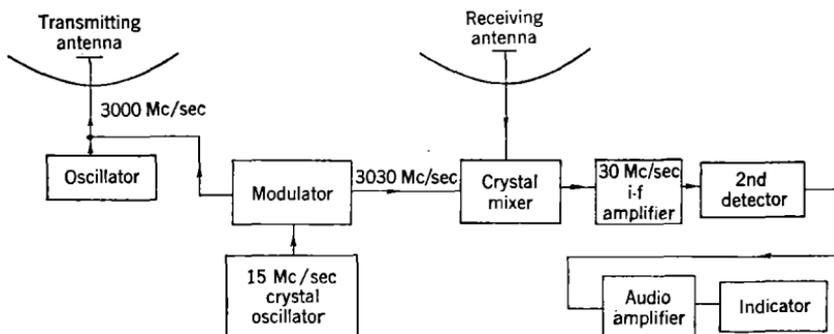


FIG. 5-5.—Block diagram of simple doppler-type radar system.

Design Procedure.—Having blocked out a proposed system, we may now sketch the design procedure that fixed the various apparatus constants and dimensions.

Design commenced with the choice of an indicator. A voltmeter will certainly work, but experience has shown that a speaker or pair of earphones, in conjunction with the operator's hearing sense, is much more effective provided the frequencies to be detected lie in the range between a few hundred and a few thousand cycles per second.¹ Within this range, the ear easily recognizes pure tones in the presence of hiss even when the pure note is much too weak to be seen on an oscillograph. This phenomenon is analogous to the averaging performed by the eye when looking at an A-scope, as discussed in Chap. 2. Rough experiments show that the ear will detect a pure tone in the presence of hiss when the tone is stronger than the noise in a 200-cps bandwidth. If the ear can do this over the range 200 to 3000 cps, then by using the

¹ A detailed study of the aural detection problem can be found in RRL Report No. 411-86, May 5, 1944.

simplest possible indication method, a power gain of roughly 15 to 1 over what might be had by using voltmeter or oscillograph can be achieved. Aural indication is therefore the logical choice.

This decision, together with the maximum target speed, almost determines the wavelength, for the doppler frequency must be chosen to lie in a range for which the ear is sensitive. Thus $\lambda = 10$ cm gives, by Eq. (1), 8.9 cps/mph or about 3000 cycles for 300 mph. At the time this system was designed, faster planes were not common. Even so, a somewhat longer wavelength might have been desirable, but 10 cm was chosen because good tubes were available at this wavelength.

This wavelength proved satisfactory, but it sometimes gives doppler frequencies rather below the frequency region in which the ear is sensitive. To overcome this difficulty, provision was made to modulate the audio signal with a 500-cps tone. The modulator was a balanced one, so that the 500-cps carrier output was zero when the signal was zero. The resulting variations in amplitude of the tone were quite distinctive, even in the presence of the noise modulation, and carried the effectiveness of the system down to frequencies as low as desired.

The wavelength being determined, the antenna size was chosen. To get as much range as possible, this was taken as large as was feasible without either (a) making the device impractically large or (b) getting the beam so sharp that at the specified angular rate of scan the beam would be on the target too short a time for the listener to hear it. Both requirements led to a diameter in the neighborhood of 40 in., the value finally chosen.

It now remained only to estimate the power required. This could be done either on the basis of experience with similar systems or by calculations of the sort outlined in Chap. 2. As a result of such considerations, a power of 10 to 15 watts was chosen.

This completes the discussion of the basic design of the system. A few of the apparatus details will be given later but first various points will be discussed which are not mentioned in Chap. 2 but which are of importance in computing the range. These all relate to the effective bandwidth, the geometrical factors that determine the transmission attenuation being the same for pulse and c-w systems.

Effective Bandwidth.—What then, in principle, determines the minimum bandwidth of a doppler system?

If we use a simple indicator, such as a voltmeter, the bandwidth must be sufficient to include the doppler frequencies of all targets of interest. This bandwidth is then determined simply by the wavelength and the range of target velocities.

If, however, we consider only targets of one radial velocity, or use some more complex indicator that divides the possible doppler range

into small bands and looks at each band separately, there are certain limits that determine how small the band may be—limits other than the obvious one of the response time of the indicator. Three such limits are discussed in the following paragraphs.

The first is due to the modulation arising from scanning. If the system is scanned, the beam will be on the target for only a finite time, with the result that even a single-frequency doppler signal will be spread over a band whose extent will be roughly the reciprocal of the time during which the beam is on the target. This is obvious enough, and would not be worth further discussion, were it not for the fact that the same consideration arises in connection with the system's ability to reject clutter. The presence of scattering objects on the ground gives rise, as we have seen, to a d-c component in the output signal of the second detector. This direct-current component will be a function of azimuth angle and so, as the system scans, we find at the second detector output "varying direct current"—that is, direct current plus various low-frequency components extending up to a maximum frequency roughly given by the reciprocal of the time the beam is on the target. But the ground returns are so enormous compared to target signals that one might fear that even the tails of the ground-return frequency spectrum would be large compared with the desired signals. Actually, this fear is not realized, as will be shown by a brief calculation.

The general idea on which this calculation is based is as follows: If the system scans at a uniform rate, the system output signal as a function of time as it scans across a fixed target depends on the directivity as a function of angle. The frequency spectrum is then the Fourier transform of the directivity function. But this directivity depends on the illumination of the dish and is in fact the Fourier transform of the illumination as a function of distance across the dish. But the Fourier transform of a Fourier transform is the function itself, and so we find the interesting theorem that the frequency spectrum due to scanning has the same form as the function representing the illumination of the dish. Actually, the above statement represents a slight oversimplification because we have assumed that the antenna directivity enters only once whereas actually it enters twice, once in the sending process and once in the receiving. Taking this into account and carrying out the calculation, we find that the spectrum due to scanning, which we shall call $g(\omega)$, is given by

$$g(\omega) \approx \int E(y)E(\omega - y) dy,$$

with

$$y = kx \frac{d\theta}{dt}, \quad (3)$$

where $E(x)$ is the dish illumination as a function of x , the transverse distance from the axis of rotation; $k = 2\pi/\lambda$; and $d\theta/dt$ is the angular rate of rotation. Now E is zero outside the range $x = \pm d/2$ for a dish with diameter d and so $g(\omega)$ is zero for $kd \frac{d\theta}{dt} < \omega$. This corresponds, incidentally, to the doppler frequency associated with the motion of the edge of the dish or, what is usually about the same thing, the number of beamwidths scanned per second. Although this calculation is more complicated for separate transmitter and receiver dishes, etc., the important general conclusion stands—namely that, in so far as the approximations used are good, the spectrum produced by scanning is definitely confined to a finite frequency range.

Another lower limit on bandwidth is set by the modulation due to fluctuation in target cross sections. If the target is an airplane, for example, inspection of Fig. 3-8 shows that even moderate yawing will introduce large fluctuations in returned signal. The frequency spread so introduced depends on the rate of yaw and on the ratio of the target dimensions to the wavelength, this ratio determining the number of pattern lobes per radian. Calculations are difficult because of lack of data, but experimentally a value of about 30 cps at $\lambda = 10$ cm is found. Another and equivalent point of view is that when the plane is turning different parts have different doppler frequencies.

Finally there is, in principle, a limit set by target acceleration. Thus if we have an accelerating target it may happen that the doppler frequency will transit the pass range of the band-determining filter before the filter has time to build up, in which case the signal may be missed. If the acceleration is a , and the band is $\Delta\nu$, so that we require a buildup time of $1/\Delta\nu$, we then find

$$\Delta\nu > \left(\frac{60a}{\lambda}\right)^{1/2} \quad (4)$$

with a in ft/sec². and λ in cm.

Apparatus Considerations.—Having blocked out the system and determined the leading design constants, we can now proceed with the detailed engineering. Figures 5-6 and 5-7 show the final result; there follow a few remarks as to the more important ways in which the engineering technique for this system differs from that used in pulse systems.

There are no high pulse powers and both intermediate- and audio-frequency amplifiers are narrow-band. Gains per stage are limited by stability rather than by bandwidth; in other words, the technique is like that of ordinary radio rather than like that of television. Because of the large size of the leakage signal relative to the target signal, only

relatively small i-f amplifications can be used, the remaining amplification being done after the second detector.

The most important consideration in doppler work is keeping the transmitter frequency modulation down. This will be discussed in

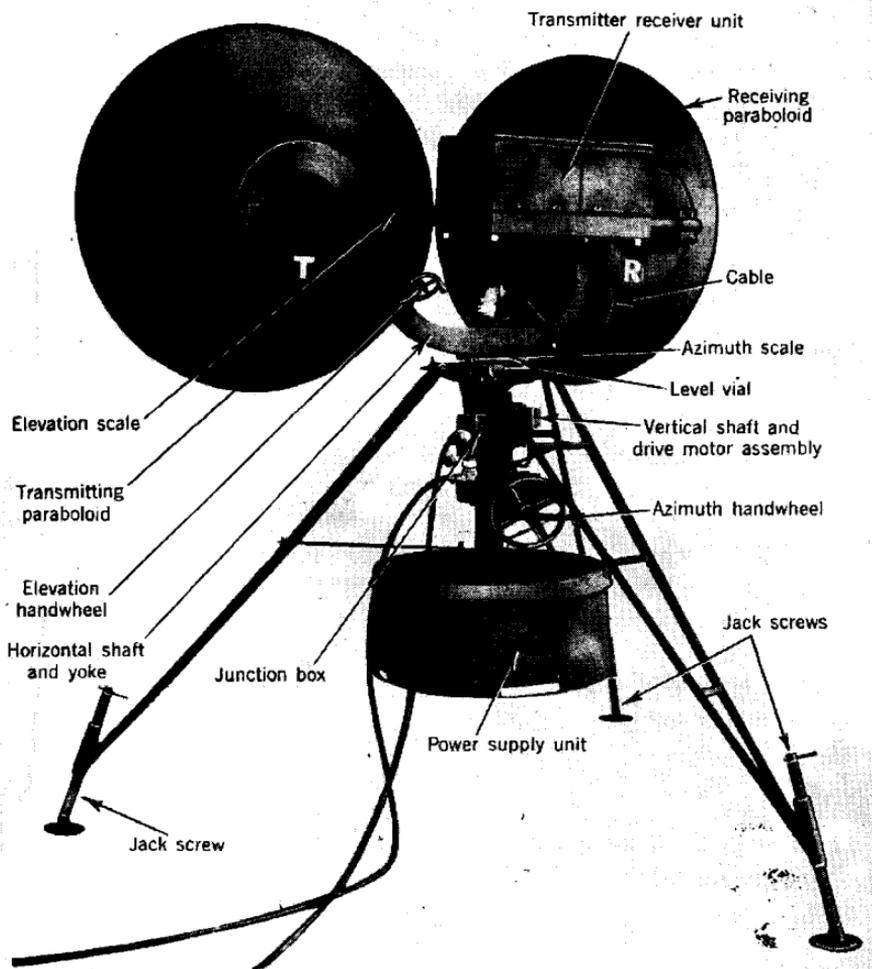


FIG. 5-6.—Rear view of 10-cm doppler system. (Courtesy of Sperry Gyroscope Company, Inc.)

detail in connection with another system but it should be noted here that short-time frequency stabilities of the order of a part in 10^{10} must be attained if the system is to work with full sensitivity in the presence of ground clutter. This requires careful attention to microphonics and to power-supply filtering. Also, the transmitter filament must be sup-

plied with direct current or conceivably with alternating current of frequency above the pass frequency of the audio amplifier.

5-7. Range-measuring Doppler System.—In order to measure the range of one or more targets, the outgoing wave must be marked, or modulated, in some way and the time required for the marks on the wave train to return must be measured. The modulation may be of either the amplitude or the frequency type, the techniques for the two methods being quite different.

The most familiar method for measurement involving amplitude modulation is to emit a pulse of waves and determine the time delay in

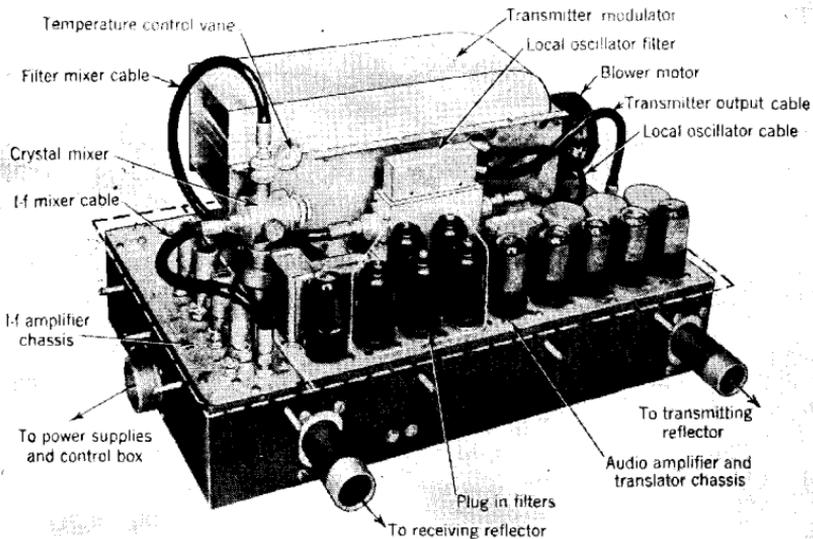


FIG. 5-7.—Transmitter-receiver chassis of 10-cm doppler system. (Courtesy of Sperry Gyroscope Company, Inc.)

the arrival of the reflected pulse. This modulation gives rise to a large number of sidebands, and the distance of the target may be regarded as determining the relative phase shifts of the various sidebands on their return path.

The fundamental point here is the determination of the relative phase shifts, not the multiplicity of sidebands. Actually only two frequencies are needed, as we will now show by considering the simplest possible amplitude-modulated c-w system capable of measuring range.

The system consists, in principle, of two separate systems like that just described, the two systems having transmitter frequencies differing by an amount f_r . The transmitted and received spectra then appear as in Fig. 5-8 with the two receivers receiving f_0 and $f_0 + f_r$ by leakage

and $f_0 \pm f_D$, $f_0 + f_r \pm f_{D'}$, from the target, with f_D and $f_{D'}$ very slightly different because of the difference in transmitter frequencies. This slight difference in doppler frequencies leads to a phase difference between the two doppler-frequency outputs which is a linear function of time, just as the target range is a linear function of time for constant radial velocity. This suggests that the phase difference is a measure of the range.

This is in fact the case, as may be shown analytically, or qualitatively by the following argument. If the target is very close to the system,

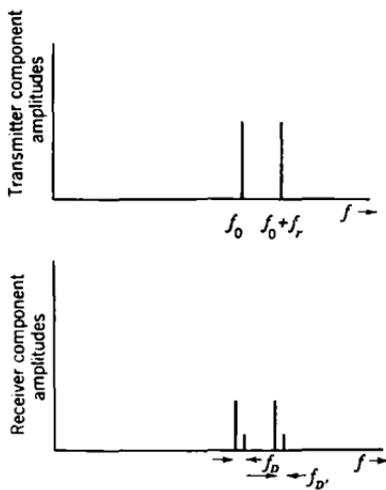


FIG. 5-8.—The upper figure shows the frequencies present in the transmitter spectrum. The lower gives the same information for the receiver. Frequencies f_0 and $f_0 + f_r$ in the receiver are due to leakage from the transmitter; the same frequencies $\pm f_D$ come from the target.

is not nearly so complicated as appears from the description since the two transmitters can be combined, as also can most parts of the receivers. The apparatus is essentially like that of the system described in Sec. 5-6 with the addition of a modulator for the transmitter, another audio channel, and a phase meter. No detailed description will be given because, as yet, no great practical application of the method has been made. The principal difficulty at present is the lack of a well-developed phase meter that will work over a range of both frequency and amplitude.

Nevertheless, some further discussion is in order both because of the principles involved and because some future use may be made of the idea.

Four points of principle should be noted. First, to measure the

number of wavelengths from transmitter to target and back will be the same, even though the two transmitter wavelengths differ slightly. Thus the leakage and target signals will be in phase or out of phase at the two receivers simultaneously and the doppler outputs will be in phase for this range. As the target gets further away, this phase difference increases, finally becoming 2π when the number of wavelengths to the target and back is one greater for $f_0 + f_r$ than for f_0 . Analytically one easily finds that the range is given by

$$r = \frac{1}{4\pi} \frac{c}{f_r} \phi \quad (5)$$

with ϕ the phase shift between the two doppler frequencies.

This is a perfectly feasible system and has worked in the field. It is

distance to a single target, only one additional sideband is needed. Second, the presence of the doppler shift, applying as it does to both sidebands, does not interfere with distance measurement. Third, the two receivers are separate and their bandwidths are determined by the doppler frequency, not by the modulation or repetition frequency. Fourth, we now have two design frequencies, f_D and f_r , and it is a matter of great importance which is the greater. If $f_D \ll f_r$, as the diagram above implies, the apparatus works as described. If $f_r < f_D$, the operation is the same in principle but the beat f_r between the two leakage signals will be within the doppler band and unless special measures are taken this beat will completely dominate the target signals.

Since the sense of this inequality is an important factor in all the systems to be described from here on, this section closes with a few further remarks concerning it.

The doppler frequency is, by Eq. (1), dependent only on the wavelength and the radial velocity. Values as a function of wavelength and for various speeds are given in the curves of Fig. 5-9.

The modulation frequency is ordinarily made as large as possible, in order to get the maximum possible change of phase with change of range. The limit is reached when, at extreme range, the phase shift is 2π radians; ranges beyond this are ambiguous. This condition leads to the inequality

$$f_r \leq \frac{c}{2r_{\max}} \quad (6)$$

a result that is entirely analogous to the similar relation discussed in Chap. 4. A plot of Eq. (6) using the equality sign is given in Fig. 5-10.

Inspection of the last two figures will show that, for radar systems in general, there is no "usual" case. One may perfectly well have either sign of the inequality. For example with $\lambda = 3.3$ cm and velocities greater than 100 mph, f_D is greater than 2950 cps and a maximum range of 93 miles would give $f_r \leq 1000$ cps. On the other hand a lower velocity and shorter range might, in conjunction with a longer wavelength, lead to numbers like $v_r = 10$ mph, $\lambda = 10$ cm, $r = 9.3$ miles, values which would give $f_D = 89.4$ cps, $f_r \leq 10,000$ cps.

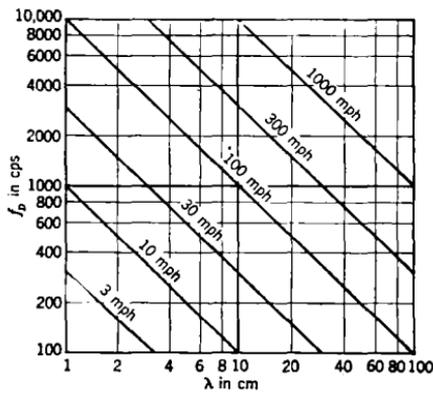


FIG. 5-9.—Doppler frequency as a function of wavelength, for various radial velocities.

When $f_r < f_D$, the system as outlined must be modified by placing infinite-attenuation filters after the second detector. The filters are designed to remove f_r completely, while leaving all but closely adjacent frequencies undisturbed. In principle, this is no different from the infinite-attenuation filters already present to remove the d-c signals due to clutter. Practically, somewhat more trouble is involved, since zero frequency filters are stable because neither the frequency nor the filter tuning can change, while both possibilities are present for frequencies other than zero. Usually also, either by accident or design, various harmonics of f_r will be present and these also must be filtered out.

This is the first case we have encountered of a general and almost exact¹ theorem to the effect that any periodic modulation can produce

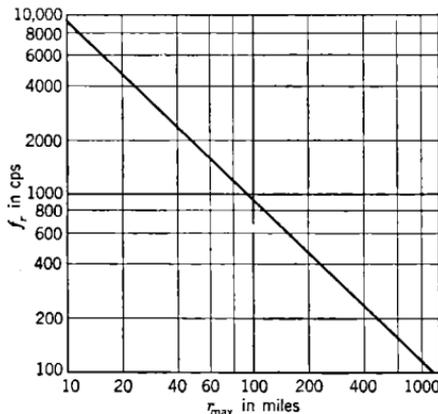


FIG. 5-10.—Maximum repetition or modulation frequency as a function of maximum unambiguous range.

in the receiver, whether by direct leakage or reflection from stationary objects, only output signals that have the same periodicity; these output signals can therefore be removed by a sequence of infinite-attenuation filters tuned to frequency zero and to all harmonics of the modulation frequency.

Finally, we note that this system will handle only one target. If more are present the precise behavior depends on the type of phase meter used and the best that can be done is to choose a type that measures range to that target which gives the most intense reflection.

If the system is required to handle more than one target at a time, discrimination may be made either on the basis of range or of doppler frequency. If range is used, more sidebands are called for, and the

¹ The theorem would be exact if the carrier frequency were an integral multiple of the modulation frequency. When this is not so, the deviations are of the order f_r/f_0 or less and are usually negligible.

number of range intervals that can be distinguished is essentially the same as the number of sidebands added. A system of this type will be described later. If discrimination of targets is to be made by doppler frequency, the various doppler tones can be separated with filters.

The more sidebands used, the more information is obtained per second because of the wider band. In discriminating by frequency, either duplicate audio systems may be used to increase the information flow, or a sweeper of some sort may be used—with resulting loss of speed.

5-8. F-m Range-measuring System.—The next system to be described is one that illustrates the f-m technique of range measurement. The specific problem is that of a radio altimeter. In this case, we wish to measure the distance to a single target, no clutter is present, and the target radial velocity may be zero. Also, the system must work down to zero range. The system described below was very successful and

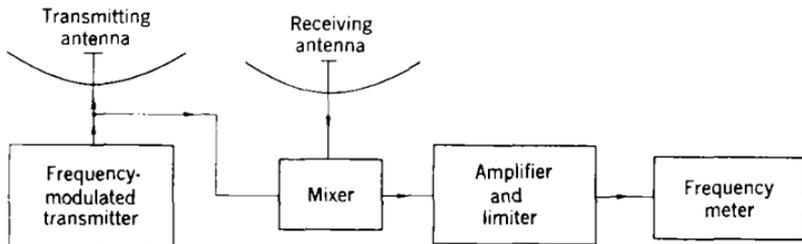


FIG. 5-11.—Simplified block diagram of f-m system for measuring range.

it and the proximity fuze are the two c-w systems that have been most widely used.

The general method may be explained by the block diagram of Fig. 5-11 and the graphs of Fig. 5-12. Here the transmitter emits waves of a frequency¹ that varies linearly with time, oscillating above and below the mean frequency f_0 , as shown in Fig. 5-12. These waves arrive at the receiver both by a direct connection and by reflection from the target. Since the trip to the target and return takes time, the received frequency curve, indicated by the dotted line, is displaced along the time axis relative to the transmitted frequency. Also there might be a displacement along the frequency axis due to doppler effect. This we disregard for the moment. The two frequencies, when combined in the mixer,

¹ We note that two types of "frequency analysis" are useful in analyzing this and other systems. One is the Fourier method, in which some curve is decomposed into a sum of sine waves, each of constant amplitude and frequency and extending in time from $-\infty$ to $+\infty$. According to the other method, frequently useful in discussing f-m systems, we say that any function of time can be represented by a function of the form $a \cos \omega t$ where a and ω are functions of t and are so chosen as to get the best fit at any value of t . In cases where the latter procedure is useful, a and ω are slow functions of t .

give rise to a beat¹ f_a , as shown in the bottom graph of Fig. 5-12. Plainly enough, the greater the target distance, the greater this beat frequency; its magnitude is then a direct measure of the range. The signal of this frequency is therefore amplified and limited, and the frequency measured, usually by a cycle-counting device of some sort. The frequency meter is then calibrated in terms of range.

Such a system would work, as described, and would also work on multiple targets. But linear frequency modulation is not easy, and for single targets the operation is not greatly affected by a change from linear to sinusoidal frequency modulation. The difference frequency then varies sinusoidally with time, but the important point remains, namely that the mean magnitude of the difference frequency depends on the range. Sinusoidal frequency modulation is therefore adopted since it requires simpler apparatus and accomplishes the same result.

There is one subtlety that is worth some discussion. If the mean or carrier frequency were an integral multiple of the modulator frequency, it is obvious that the output of the mixer would be periodic with periodicity corresponding to the modulating frequency. This would mean that each modulation period would contain the same number of cycles. This number is then integral and we see that the frequency meter can read only 1, 2, . . . , n times the modulating frequency. Actually, the same conclusion holds even with no special relation between modulation and carrier frequencies, if, as is usual, we use a cycle-counting type of frequency meter. Naturally, it is of interest to translate this step-wise behavior of the frequency meter into altitude readings. For the linear-modulation case we easily find $f_a = 4f_m \Delta f h / c$ with f_m the modulating frequency, Δf the total frequency swing, f_a the beat frequency, and h the height. Solving for the height and introducing the fact that f_a is quantized in steps of f_m , we find for the error δh

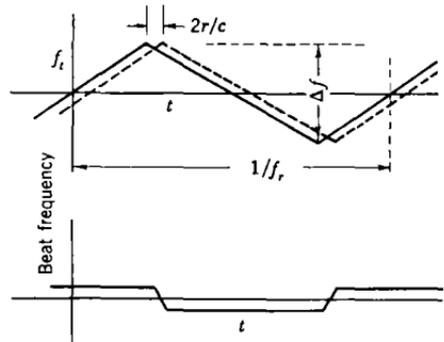


FIG 5-12.—The upper figure shows, in the full line curve, the instantaneous transmitter frequency as a function of time. The dotted curve is the received frequency. The lower figure shows the difference, or beat, between transmitted and received frequencies, as a function of time. The horizontal axis in the upper figure corresponds to the mean transmitter frequency f_c .

¹ It is beats of this sort, which may be described as due to time-delay demodulation of FM, that are regarded as spurious signals in the system described in Sec. 5-6 and that must be avoided by reducing FM of the transmitter to the lowest possible value

$$\delta h = \frac{c}{4\Delta f}. \quad (7)$$

This is of the nature of a maximum error and can be much reduced by averaging over a number of modulation cycles.

Design Procedure.—Design of such an altimeter is begun by using the above equation, together with an allowable altitude error, to pick a suitable Δf . In the present case Δf was chosen as 40 Mc/sec, which corresponds to $\delta h = 6$ ft, and so, because of averaging, to a rather smaller operational error.

The mean transmitter frequency is now chosen. This must be fairly high in order to keep the frequency variation from being an impossibly large fraction of the mean frequency. On the other hand there is no gain in very high frequencies and there may be some loss in intensity over such terrain as forests. These considerations, plus a consideration of the tubes available, led to the choice of a mean frequency of 440 Mc/sec, which can be obtained from acorn tubes; the frequency modulation is accomplished mechanically.

The modulation frequency is next chosen to give a convenient range of beat frequencies f_a , subject to the restriction that the time of a modulating cycle shall be long compared to the maximum signal transit time. A value of 120 cps was picked. This gives $f_a = 8000$ cps at an altitude of 400 ft. A second range of 0 to 4000 ft is obtained by reducing Δf to 4 Mc/sec.

This completes the major specifications, except for the power. Practically speaking, power must be decided on the basis of experience, calculations of available power and thermal noise power being quite useless since, in practice, the limitation is not thermal noise but microphonics, etc. Thus if we consider the earth as a diffuse reflector we easily find the received power is $(1/2\pi)(A/r^2)$ times the earth's reflection coefficient and times the transmitted power. Even for very small transmitter powers and reflection coefficients, this power is large compared to thermal noise. On the basis of experience, then, it was decided that the 0.3 watt available from an acorn tube would be sufficient for altitudes up to 5000 ft, provided certain points of apparatus detail were correct, as explained below.

Apparatus Considerations.—The leading parameters having been specified, one can proceed with the detailed design. This is straightforward, except for certain steps necessary to reduce various spurious signals in the receiver which might limit the range.

These signals, for the most part, are due to amplitude modulation in the transmitter. This amplitude modulation comes from microphonics and from slight variation of transmitter amplitude with frequency.

Modulation from the latter source is at the modulating frequency and low harmonics thereof. These spurious signals may be greatly reduced by two devices. First, one makes the mixer of Fig. 5-11 a balanced one

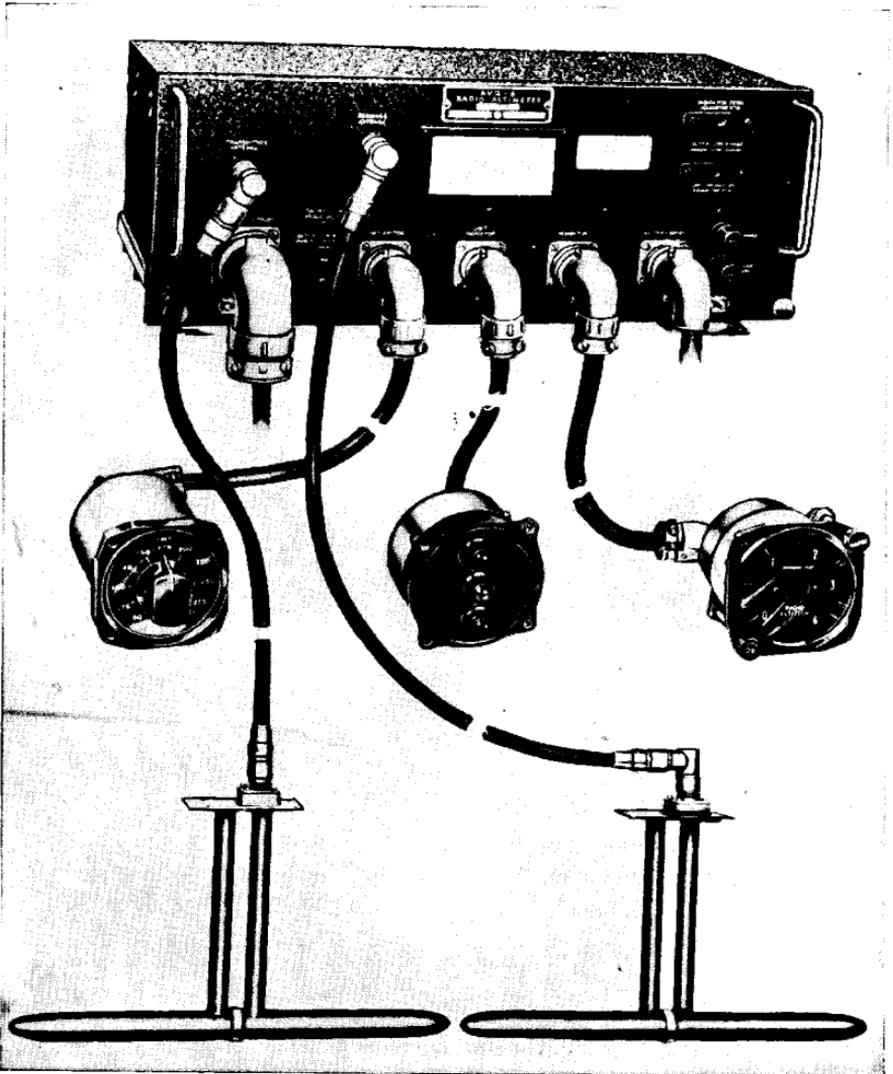


FIG. 5-13.—The AN/APN-1 frequency-modulated radar altimeter. (Reprinted from *Electronics*.)

(see Vol. 24) with the result that, if the balance is good, amplitude modulation from the transmitter balances out in the detector output. Second, the amplifier that precedes the limiter is given a response that is a rising function of frequency. The spurious signals, being largely of

frequency lower than the desired signals, are thus discriminated against. Roughly speaking, the response should rise linearly with frequency, since, on a voltage basis, the incoming signals go down at about this rate in range and thus in the equivalent frequency. This frequency distortion serves a second purpose in that it makes the output signal less dependent on altitude so that the limiter does not have to work over such a large range of amplitude. The complete device is illustrated in Fig. 5-13.

The effect of doppler shift is to raise or lower the dotted curve (Fig. 5-12) corresponding to the returned signal. With the proportions here used and with the aircraft in reasonably level flight, the doppler frequency is almost always less than the beat frequency due to altitude; consequently the end result is that there are somewhat fewer cycles of beat frequency in one half of the modulating cycle, and somewhat more in the other, the total number per cycle remaining the same. The difference in numbers of cycles in the two halves of the modulation period is then a measure of the doppler frequency. This difference can be measured in various ways, and has been used, along with the altitude information, in various developments of the device here described.

5-9. Multiple-Target F-m Range Measurement.—In the absence of clutter and doppler shift almost the same methods can be used with a plurality of targets as with a single target. In this case linear frequency variation with time is almost essential, and the triangular form of Fig. 5-12 is probably most convenient, though a saw-tooth variation might be used. The detector output then contains a number of frequencies, one corresponding to each target range present. Preferably these frequencies are detected by some device such as a Frahm vibrating-reed frequency meter, which indicates all frequencies simultaneously. If, however, time is no object, a device that scans the frequency range may be used. A variable frequency can be added to the signal frequency and observations made when the sum frequency falls in the pass band of a resonant circuit. Variants of this idea sweep either f_r or Δf and observe when the target frequency falls in the pass band of a resonant circuit.

Any of these scanning devices greatly increases the time required to obtain the desired information. Thus if we have a frequency band f which is to be split into n pieces, the time required for such a device as a Frahm meter to respond is of the order n/f , whereas if the n frequency intervals are observed in sequence the time is n^2/f .

The general design procedure is essentially the same as that described in Sec. 5-8. First we decide on the allowable range error δr and determine from this the total frequency swing Δf by means of Eq. (7):

$$\delta r = \frac{c}{4\Delta f}.$$

We may note that $1/4\Delta f$ is analogous to the pulse time; it determines the width of waveband transmitted and the range accuracy.

Next, one determines the modulation or repetition frequency f_r . As in pulse systems, this depends on the maximum range. But whereas in pulse systems f_r may actually be as much as $c/2r$, in the present systems one must have f_r markedly less than $c/2r$ in order that not too large a fraction of the time will be wasted while the sign of the beat frequency is changing. Subject to this upper limit, the choice of f_r is determined by considerations of bandwidth, response time, and apparatus convenience. The smaller f_r , the narrower the band and the slower the response. Also, the beat frequency, which varies with f_r , should be kept in a range suitable for the frequency meter to be used.

The maximum beat frequency is then determined as $4f_r\Delta f(r/c)$, and an indicator working up to this frequency by steps of f_r is designed. The number of steps is thus about $r/\delta r$.

This method of ranging, like the pulse method, makes no special requirements on the carrier frequency, which is then chosen on the basis of other considerations.

The power required depends on the usual things, including the bandwidth, which in this case is the bandwidth of one element of the frequency meter—which should be about f_r .

This scheme, although it has not been highly developed, will undoubtedly work much as described. High range accuracy and small bandwidths are possible—in fact one may have both at once. Thus one might have a repetition rate, and so a noise bandwidth, of, for example, 10 cps while having a frequency swing of 40 Mc/sec, which would give range accuracy corresponding roughly to a $\frac{1}{40}$ - μ sec pulse. And this latter would be achieved without a wideband i-f or video.

If, however, we attempt to modify the system so as to allow the presence of clutter, and therefore also a doppler shift of the signal coming from the target, two difficulties arise.

First, although the clutter may, in principle, be eliminated by infinite attenuation filters tuned to the repetition frequency and multiples thereof, the number of filters required is of the order of $r/\delta r$, and if this number is large the system may be impractical. This limitation appears to be fundamental and arises in similar form in all other systems. But even if we neglect the multiplicity of the filters, there is considerable doubt whether practical means of frequency modulation can be devised that will make successive modulation cycles as nearly identical as is needed for filtering out really serious clutter.

Second, as a result of the doppler shift, each target gives two output frequencies, and so two range indications. The seriousness of this depends on the ratio of the doppler frequency to mean beat frequency.

This ratio depends on target speed, wavelength, repetition rate, and range accuracy, and might conceivably have almost any value. But in many cases the value will be too large for this system to be useful. Specifically, the maximum beat frequency may be written as $f_r(r/\delta r)$ so that, for example, 1 per cent range accuracy and $f_r = 10$ cps gives $f_a = 1000$ cps, which may be compared with a doppler frequency of 894 cps for 10 cm and 100 mph.

5-10. Alternative F-m Ranging System.—Another scheme of the f-m type which is designed to work on multiple targets and in the presence of clutter may be understood by reference to Fig. 5-14. This shows, in the full curve, a sinusoidal dependence of the transmitted frequency on time. The total frequency swing is made large compared to the doppler frequency. The received frequency, indicated by the dotted line, is a similar curve but (a) displaced to the right by time delay resulting from transmission to the target and back and (b) displaced vertically by doppler shift. If now the difference between the transmitted and received frequencies is passed through a low-pass filter whose pass band includes frequencies as high as the maximum doppler frequency, there will, in general, be little signal output because the difference frequency will be greater than the doppler frequency over the majority of the modulation cycle. If, however, the modulation frequency is adjusted so that one cycle corresponds to the transmission delay time, the difference frequency will always just equal the doppler frequency and so a large output signal will result. Thus, as the modulation frequency is varied, output occurs whenever $n/f_r = 2r/c$ with n any integer and r the distance to a target. Ground clutter, having no doppler shift, gives an output that is periodic with period corresponding to the modulation frequency. The d-c component is easily removed and, if $f_D < f_r$, the other components are removed by the low-pass filter, which passes only up to f_D . If, therefore, $f_D < f_r$ ground clutter may, in principle, be removed.

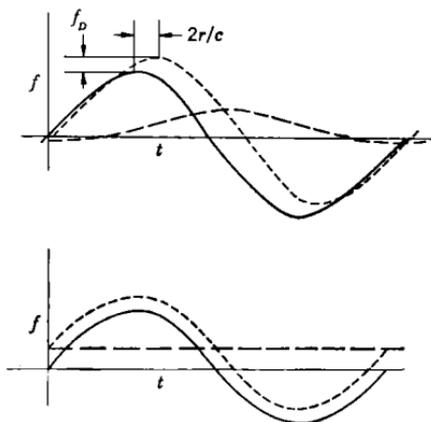


FIG. 5-14.—The full curve in the upper graph shows the instantaneous transmitter frequency as a function of time; the dotted curve shows the received signal; and the dashed curve, the beat between the two. The lower curve shows the same quantities when $2r/c = n f_r$, in which case the beat frequency is constant at the doppler value.

A variant of this scheme turns the receiver off during one modulation

cycle, and turns the transmitter off and the receiver on during the following cycle. This permits the use of a single antenna and also reduces the clutter intensity by making it possible to gate out the strong returns from near-by objects.

Difficulties with systems of this type are as follows: (1) determination of range by variation of f_r takes much more time than is justified by the receiver bandwidth; (2) the distance measurement is unambiguous only over a 2-to-1 range; (3) proportions must be such that $f_D < f_r$ —this means long wavelengths and/or slow targets; (4) it is difficult to get the modulation cycles to repeat well enough for suppression of really serious clutter.

5-11. Pulse-modulated Doppler System.—The last system to be described works against the heaviest ground clutter, and, in principle, on a plurality of targets and with no restriction on the relative values of doppler and recurrence frequencies. Practically, there are limits on the last two factors, the limits being set by questions of apparatus complexity.

A block diagram is shown in Fig. 5-15 and the operation may be described as follows. If, as we shall assume for purposes of explanation, only a single target is to be observed, the transmitter is turned on for half the keying cycle and off during the other half by means of the modulator, square-wave generator, TR and ATR tubes. If a multiplicity of targets must be handled, the transmitter pulse is made correspondingly shorter. The transmitter power then follows the Curve *a* of Fig. 5-16, and the receiver input follows Curve *b* with the time delay depending on the range and with the dotted part of the received signal rejected because the receiver is off.¹ This received signal is amplified by a conventional superheterodyne receiver whose only special feature is the derivation of the local-oscillator frequency from the transmitter frequency by the addition of 30 Mc/sec. This 30 Mc/sec is then subtracted in a second detector so that the received signal is translated down to zero frequency.

Consider the voltage at point 1 in Fig. 5-15 just after the second detector. If the returns are from a stationary target the voltage is of the form shown in Graph *a* of Fig. 5-17. This voltage is periodic and has magnitude depending on the target cross section and on the exact phase of the returned signal—a change of range of $\lambda/4$ resulting in a reversal of sign. If the target moves, the change of range causes a periodic oscillation of amplitude, the resulting signal being like that of Graph *b* Fig. 5-17 and having an envelope which is a sine wave of doppler frequency.

¹ Actually, the receiver gate is not opened until several microseconds after the transmitter is off. In this way ground clutter due to nearby objects is eliminated.

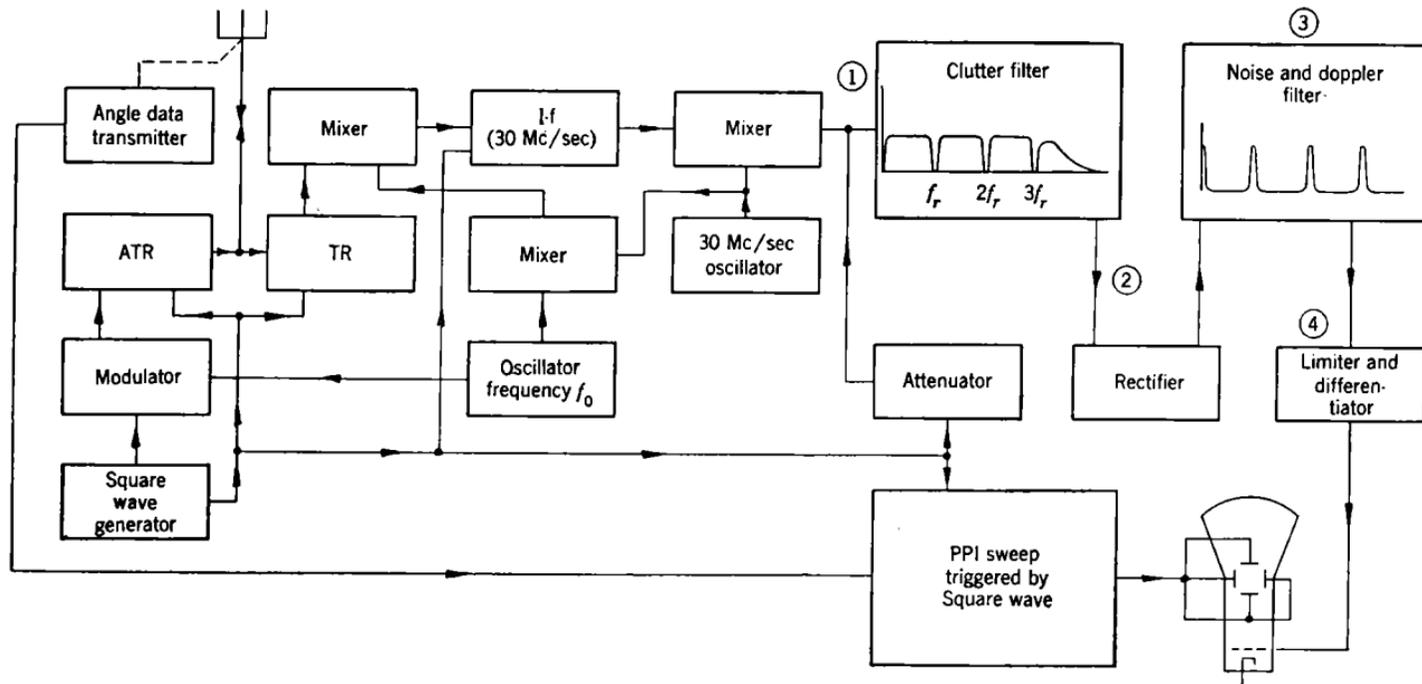


FIG. 5-15.—Block diagram of pulse-modulated doppler system.

Now the Fourier resolution of Curve *a* will contain the repetition frequency and multiples thereof, whereas Curve *b* contains $f_D \pm nf_r$, but no multiples of f_r (unless f_D is a multiple of f_r). Thus the first or "clutter" filter that attenuates f_r and multiples thereof, while leaving all else, will remove all signals due to stationary objects, while leaving those due to moving objects undisturbed. At point 2 in Fig. 5-15, therefore, the signals are like those of Fig. 5-17*b*, and do not contain the frequency f_r .

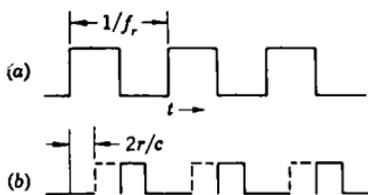


FIG. 5-16.—Transmitted and received powers as a function of time.

The signals are now passed through a full-wave rectifier, the voltage at Point 3 in Fig. 5-15 being like Fig. 5-17*c*. This curve now contains harmonics $nf_r \pm mf_D$.

The various doppler sidebands and most of the noise are now removed by passage through a filter that passes only f_r and its multiples, whereupon the voltage at Point 4 has a form like that shown in Fig. 5-17*d*. Thus the clutter, the doppler frequency, and much noise have been removed, while the position of the trailing edge of the pulse, which is a measure of the range, is undisturbed.

The time delay of the trailing edge of the wave at d , relative to the modulating pulse, is now a measure of target range. This time, or range, may be displayed on the linear radial sweep of a PPI tube by differentiating the voltage and using the result to intensity-modulate the beam.

A few minor points remain to be considered. First, the explanation has assumed that the bandwidth is large enough to pass square waves without distortion. Actually, the operation is substantially the same when only a few harmonics of f_r are passed. Second, the function of the attenuated square-wave voltage introduced after Point 1 is to balance

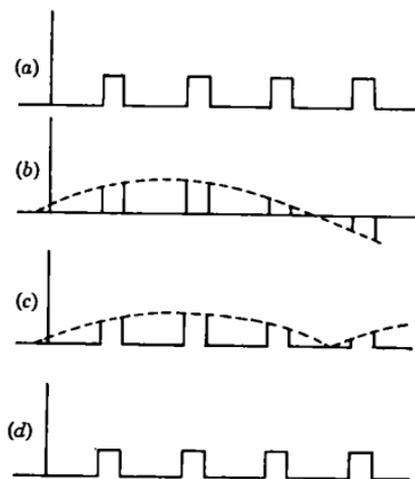


FIG. 5-17.—Curve *a* shows the voltage at Point 1 (cf. Fig. 5-15) due to a stationary target and *b* shows that due to a moving target; the dotted envelope curve corresponds to the doppler frequency. Curve *c* is the result of passing *b* through a full-wave rectifier, and *d* is the result of filtering *c*.

out the square-wave voltage due to the d-c component of receiver noise, as modulated by the gating. Finally, it should be noted that the last filter can have fairly narrow pass bands, thereby considerably reducing the noise.

Design Procedure.—The leading requirement in the design of a system such as that described above was a range of 75 miles on single targets moving at speeds up to 400 mph, even in the presence of extreme ground clutter.

First, since ability to deal with multiple targets was not required, the modulation cycle was taken to be half on and half off. Next, the repetition frequency was chosen to be as high as possible, consistent with getting a reasonable fraction of the extreme range signals to return during the time the receiver is on. Thus the value of $f_r = 1000$ cps, which was selected, lets the entire return signal into the receiver at $r = 46.5$ miles and none at 93 miles, so that at 75 miles about half is lost. The advantage of a high value of f_r is the reduction in number of stop bands required in the filter.

The dish size and wavelength were chosen—on the basis of range, propagation factors, and beamwidth—to be 12 ft and 40 cm, respectively. This latter value also determines the doppler band, which extends from zero to about 1000 cps for 400-mph targets. Thus we need carry only one harmonic of the repetition frequency.

The transmitter power depends on geometrical factors in the usual way, and on the effective bandwidth of the receiver. It might be thought that this would be the total width of the pass bands in the second filter of Fig. 5-15, but actually, because of the rectifier between the two filters, the noise depends on the bandwidths, both before and after this detector. In fact, the effective bandwidth is the geometric mean of the two bandwidths. The last bandwidth was chosen at about 4 cps, this being the smallest consistent with the time during which the beam remains on the target; and the first bandwidth was about twice the maximum doppler frequency. Thus the effective bandwidth, for noise-computation purposes, was about 100 cps. With this bandwidth, a transmitter power of 100 watts should give the required range.

Apparatus Considerations.—The leading design parameters having been chosen, the next subject of discussion is the technical difficulties involved. There are two chief ones, both having to do with elimination of ground clutter: (1) frequency, and also amplitude, modulation in the transmitter must be held to very low values; and (2) the filters, which reject frequencies 0, f_r , $2f_r$, etc., present a considerable problem.

As to the first point, the entire operation of the system is predicated on the assumption that the ground returns are periodic. This assumption would be *completely* falsified, if the range of a ground target, measured in wavelengths, were to change by as little as $\lambda/2$ in one repetition cycle. Since, in round numbers, there are 10^6 wavelengths contained in twice the maximum range, it is apparent that the operation will not even approximate the above description unless the short-time frequency stability

exceeds one part in 10^6 . Actually, much better than this must be done since the ground returns may be 90 db or thereabouts above the target returns. Thus, one might expect stabilities of the order of 10^{10} to be needed. Much calculation can be done on this point but, though such calculation was useful in showing that a workable system was possible, it will suffice to say here that by very careful attention to detail it was found possible completely to eliminate all trouble due to frequency (and amplitude) modulation of the transmitter. The main points were: crystal control of both transmitter and modulator, very careful regulation and filtering of power supplies, and operation of important filament supplies with alternating current (3000 cps) obtained by multiplication from the repetition frequency.

We next consider the "infinite rejection" filters that are to remove the ground clutter.

The maximum attenuation required is the quantity of most importance. Naturally, this is indefinite since it depends on the terrain, and pertinent measurements are very scanty. But the attenuation needed may be estimated in several ways, which agree moderately well, and finally it was measured on the finished system. Data from all sources indicate that ground returns in mountainous terrain may be much larger than is commonly realized—in the present case, of the order of 90 db above noise.

The first method of estimation is to calculate the largest possible ground return on the assumption that the return is from a mountain so large as to fill the beam completely. Such a mountain is certainly possible, and assuming it to be hemispherical its returns will be larger than those from a target of cross section σ in the ratio $r^2 A / \lambda^2 \sigma$ with A the dish area. In the present case, this comes out at about 110 db. This is certainly an overestimate because the mountain will generally not have unity reflection coefficient, and because the part of the model that does the reflecting—namely, the part of the hemisphere normal to the line of sight—will usually be missing in actuality. Nevertheless, the calculation is interesting as indicating that *very* large ground returns are possible.

A second estimate can be derived from measurements made at Ellenville, N.Y. with a more or less normal (1- μ sec, 100-kw) pulse system. These measurements showed a ground return 75 db above noise. To compare these measurements with our present problem, a number of factors must be taken into account. These are: the difference in pulse lengths (\sim 27 db), the different receiver bandwidths (\sim 30 db), and the different transmitter powers (\sim 30 db). Taking these factors into account, the clutter returns at Point 1, Fig. 5-15, may be very crudely

estimated as 100 db above the noise in a 10^3 -cycle band, or 130 db above that in a 1-cycle band.

A third estimate can be made from measurements on the completed system at this same location, which showed a clutter 100 db above the noise in a 1-cycle band. Part of the discrepancy between this and the above 130-db estimate is no doubt due to crudities in the estimation, the rest to the fact that the clutter in a 500- μ sec pulse will certainly be less than 500 times that in a 1- μ sec pulse because the clutter is not spread uniformly in range. In any case, whatever the reason, the clutter is experimentally found to be about 70 db above the noise in the 1000-cycle audio band, or about 80 db above the noise in the effective band of 100 cycles.

The system, as actually made and tested at this same location, had 92-db attenuation at the "infinite attenuation" points of the filters and experience showed that this was just comfortably adequate.

From the above results we conclude that the filters must have about 90 db of rejection for satisfactory operation at this particular site. Measurements at other sites would be valuable in answering the question as to how representative this site is. Lacking such data, we can only venture the opinion, based on personal observation, that the location did not appear unusual in any way—in fact it appears likely that mountain areas with even higher ground returns may be common.

The width of the rejection bands depends on the variability of the ground returns. Until now, the ground returns have been assumed constant, so that the voltage at Point 1 in Fig. 5-15 is actually periodic and could therefore be removed by filters with infinitesimal bandwidth. Actually the ground returns vary, with the result that the filter rejection bands must have a finite width. Using some data of H. Goldstein,¹ a width of 4 cycles 12 db up from the 92-db bottom of the curve was chosen. Even if the ground returns had been constant, much the same bandwidth would have been needed because of modulation due to scanning. The rest of the curve was then made as narrow as possible. This turned out to be about 200 cycles at the 3-db point.

Additional filters for eliminating "window" signals or rain clutter were also provided and could be switched in when desired. These filters had much less attenuation than the clutter filters, but the attenuation extended over a wider frequency range, designed to exclude doppler frequencies due to motion of "window."

Although the audio amplifier began to cut off rapidly above 1000 cps, and the strength of the harmonics of f_r decreased fairly rapidly, it was found necessary to have infinite attenuation filters at 2000 and 3000 cps,

¹ See Sec. 6-20, Vol. 13 of this series.

in addition to those at 0 and 1000 cps. Naturally, these latter filters were not nearly so critical as the one at 1000 cps.

This concludes the discussion of the design and the leading difficulties of a system of this type.

Performance was much as expected. In particular, the strong ground clutter mentioned above was completely eliminated, targets being tracked regardless of range or position. All other known systems that have been developed to the field-trial stage were tried at this same site; none of them eliminated the ground clutter.

The only deficiency of this system, as has been pointed out, is its inability to function properly when there are a number of targets in the beam. When this is the case, the range indicated is that of the strongest target. This difficulty may be reduced as desired by reducing the length of the transmitted pulse, by correspondingly increasing the number of filter bands, and so increasing the complexity.

Since the general objects of this system are much the same as those of the MTI system described in Chap. 16, some comparison is in order, even though this comparison is difficult and possibly dependent on personal viewpoint.

First, the two systems regarded from an abstract point of view work on similar principles. Both have, therefore, much the same fundamental limitations. Practically, however, the two systems use entirely different apparatus, and consequently when, as is usually the case, apparatus limitations dominate, the systems may be expected to be entirely different. Furthermore, because of this apparatus difference, one may expect future developments to improve the two in different ways and by different amounts.

Second, the two systems differ in their ability to handle a multiplicity of targets. As actually constructed, MTI systems use 1- or 2- μ sec pulses and so are capable of handling targets a thousand feet or so apart in range. On the other hand, the system described in this section works on only one target, and although a moderate number of additional filters might be considered as possible whenever anyone wishes to deal with multiple targets there is certainly a practical limit to their number. Possibly dividing the total range into 10 pieces would be reasonable. For traffic control around a busy airport, however, the c-w system would be almost useless whereas the MTI system would be very good. For long-range detection of aircraft, an ability to handle 10 targets might well suffice and the c-w system therefore be adequate.

When, on the other hand, one considers severe ground clutter the system described above has the advantage over the MTI systems. MTI systems at present reduce the clutter by perhaps 30 db—which in many locations is adequate. But mountainous terrain that does not appear

in any way unusual has been found to have ground clutter 75 db above noise. Under these conditions, the MTI system would not be expected to work; the above-described system does work.

5-12. Summary.—In an effort to help the reader who wishes to compare the various systems described in this chapter, Table 5-1 has been prepared giving, in tabular form, a summary of the more important systems of characteristics, both qualitative and quantitative. This table is worthy of careful study but two limitations must be remembered. First, the quantities given for numbers of sidebands, noise band, etc., are only qualitatively correct; really precise definitions would require so many qualifications as to render the quantities either useless or confusing. Second, neither the table nor this chapter pretends to describe *all* c-w systems; at least a dozen more are known and no doubt still more could be invented. It is believed, however, that representatives of all important types have been included.

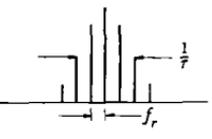
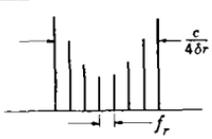
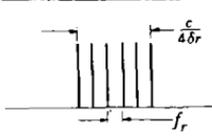
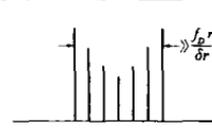
From this table, or what has gone before, the reader will perceive that there have been grouped together, in this chapter, some rather diverse systems. Perhaps a better, if longer, chapter title would have been "Systems That Are Not Pulse Systems."

If we attempt a classification, the two best starting points appear to be the questions—does it utilize the doppler shift or not, and does it use delay-time demodulation of f-m signals for range determination? Further questions are—will it work on multiple targets, and will it work down to zero range?

It is hoped that the latter part of this chapter will have answered the question asked at the beginning as to whether there are any uses for the systems here described. Naturally, a general answer cannot be given as to when c-w systems are indicated—we can only suggest a careful study of the requirements and a comparison of the potentialities of the various types of systems. Perhaps the only useful general remark is the obvious one that c-w systems are most useful on problems where pulse systems fail! For example, if one wants to measure a range of 10 ft a pulse system would hardly be suggested; that described in Sec. 5-8 would have no trouble. Also this c-w system will measure such short ranges much more accurately than pulse systems. Or suppose the velocity of a bullet is wanted—the system described in Sec. 5-6 measures such quantities easily and directly. Or again, one may want simplicity—this is provided by such a system which, when carried to the ultimate in this direction, becomes the proximity fuze.

TABLE 5.1.—SUMMARY OF

 τ = pulse width r_{\max} = maximum range δr = range discrimination

System type	Repetition or modulation frequency	Transmitter spectrum	Number of sidebands	Noise band
Pulse	$f_r < \frac{c}{2r_{\max}}$		$\frac{1}{f_r \tau}$	$\frac{1}{\tau}$
MTI	$f_r < \frac{c}{2r_{\max}}$		$\frac{1}{f_r \tau}$	$\frac{1}{\tau}$
Simple doppler with audio split into n bands (Sec. 5-6)			1	$\frac{f_D}{n}$
Range-measuring doppler system, two transmitted frequencies (Sec. 5-7)	$f_r < \frac{c}{2r_{\max}}$		2	f_D
F-m range measuring system (Sec. 5-8)	$f_r \ll \frac{c}{2r_{\max}}$		$\frac{c}{4f_r \delta r}$	$\frac{r}{\delta r} f_r$
Multiple target f-m system (Sec. 5-9)	$f_r \ll \frac{c}{2r_{\max}}$		$\frac{c}{4f_r \delta r}$	f_r
F-m system (Sec. 5-10) doppler band split into n ranges	$f_r = \frac{c}{2r}$		$\gg \frac{f_D r^2}{c \delta r}$	$\frac{f_D}{n}$
A-m doppler system (Sec. 5-11)	$f_r < \frac{c}{2r}$		$\frac{1}{f_r \tau}$	$\left(\frac{f_D}{\tau}\right)^{1/2}$

SYSTEM CHARACTERISTICS

 f_r = modulation frequency or PRF f_D = doppler frequency

Re- sponse time	Informa- tion rate	Number of targets possible	Ground clutter rejected?	Zero range possible?	Frequency restrictions	Remarks
$\frac{1}{f_r}$	$\frac{1}{\tau}$	$\frac{1}{f_r \tau}$	No	No		
$\frac{1}{f_r}$	$\frac{1}{\tau}$	$\frac{1}{f_r \tau}$	Yes	No		In 1946, maximum clutter reduction about 30 db.
$\frac{n}{f_D}$	f_D	n	Yes	Yes		If band is not split $n = 1$. Ground clutter filters are very simple.
$\frac{1}{f_D}$	f_D	1	Yes	Yes		
$\frac{1}{f_r}$	f_r	1	No	Yes	$f_D < \frac{r}{\delta r} f_r$	In practice, noise limit usually set by other factors than noise band.
$\frac{1}{f_r}$	$\frac{r}{\delta r} f_r$	$\frac{r}{\delta r}$	Yes	Yes	$f_D \ll \frac{r}{\delta r} f_r$	Clutter rejection not known, probably not extremely high.
$> \frac{r n}{\delta r f_D}$	$\frac{f_D}{n}$	$\frac{r}{\delta r}$	Yes	No	$f_D < f_r$	Clutter rejection not known, but probably not extremely high.
$\frac{1}{f_n}$	$\frac{f_n}{f_r \tau}$	$\frac{1}{f_r \tau}$	Yes	No		Rejects extreme clutter; f_n is bandwidth of noise rejection filter.

CHAPTER 6

THE GATHERING AND PRESENTATION OF RADAR DATA

BY L. N. RIDENOUR, L. J. HAWORTH, B. V. BOWDEN, E. C. POLLARD

6-1. Influence of Operational Requirements.¹—The purpose of any radar system is to present information on the positions of targets within the volume of space it surveys. The nature of such targets, their number, and the character of the information required about them will depend profoundly on the function which the radar is called upon to perform. For example, the radar altimeter used in aircraft deals with only one target—the surface of the earth beneath the plane—and offers only one datum—the minimum distance of the altimeter from that target. A radar system used for control of air traffic near a landing field, on the other hand, must display the range, bearing, and altitude of all aircraft within or near the boundaries of the control zone. At a given moment, there may be more than a hundred individual targets, for each of which the radar must provide sufficiently accurate positional information to enable controllers to issue prompt instructions to pilots.

Radar equipments suitable for such different purposes will present wide differences in design. This chapter and the next deal with the principal fundamental differences that appear in various conventional pulse radar designs. In this chapter the effect of operational requirements on the design of the radar itself is considered; in the next is discussed the more difficult question of organizing the facilities and services required to interpret and make use of the data provided by radar.

The most important and fundamental of the radar design differences that arise from different functional requirements are those which concern the beam pattern produced by the antenna, the arrangements for scanning a certain volume of space with that beam, and the indicator or indicators necessary to display the positions of targets detected. Because of the extremely close interrelationship among these three factors, and because all three are principally determined by the functional aim of the radar equipment, it would appear desirable to begin with a discussion of the various radar functions, showing for each the choice of these three factors which experience has so far recommended. Without an appropriate introduction, however, such a treatment might be hopelessly confusing. Accordingly, there follows a catalogue of existing indicator

¹ By L. N. Ridenour.

types, without attention to the functional requirement which has called each into being, and without a description of the technical means of realizing such indications. The latter subject is undertaken in Chap. 13. The postponed discussion of radar functions, and of the choice of important radar properties to fit these functions, is found after this catalogue of indicator types.

TYPES OF RADAR INDICATORS

BY L. J. HAWORTH

6-2. Definitions.—The device that presents radar data in observable form is called the *indicator*. It is usually a cathode-ray tube; alternatively it may be a loudspeaker or telephone headset, a flashing light, a moving-coil meter, or a pen-and-ink recorder.

The cathode-ray tube (CRT) permits an interpretation of electrical phenomena in terms of a picture painted on a phosphorescent screen by a sharply focused beam of electrons controlled in position and intensity by electrical signals. It is capable of using and displaying many millions of separate data per second. The geometrical expression which the CRT gives to electrical phenomena is peculiarly appropriate to radar, because a geometrical situation involving the various radar targets is precisely what must usually be represented. Of the other devices mentioned as possible indicators, only the pen-and-ink recorder is capable of giving such a geometrical interpretation, and it is slow and cumbersome by comparison.

Thus, in spite of certain inadequacies, a cathode-ray tube is used as the radar indicator in all situations that involve any appreciable target complexity. The picture presented on the tube face is called the *indication*, the *display*, or the *presentation*. The tube itself is referred to as the *indicator*, *indicator tube*, *display tube*, *CRT*, or *scope*. When a tube presenting a particular form of display is to be identified, a descriptive adjective or code designation is prefixed to "scope." The word *indicator* is often extended to include devices and circuits auxiliary to the cathode-ray tube.

Both magnetic and electrostatic cathode-ray tubes are used (Sec. 13-1). In their design every effort has been made to achieve the optimum in resolution, light intensity, deflection sensitivity, and compactness. The materials (*phosphors*) used in the screen are of particular importance. If, as is usually the case, the scanning interrupts the picture for longer than the retentivity time of the eye, it is necessary to introduce persistence into the screen. Screens of various rates of decay are available for this purpose. The detailed characteristics of the tubes, including the screens, will be discussed in Secs. 13-1 to 13-3.

In forming the displays, radar echo signals may be used either to displace (*deflection-modulate*) the electron beam, as in the ordinary oscilloscope, or to intensify (*intensity-modulate*) it, as is done in television. Deflection modulation affords precise information about the strength and character of the signals delivered by the receiver, but leaves only one dimension of the tube face free to represent a geometrical quantity. On the other hand, intensity modulation permits the presentation of a two-dimensional figure on which the signals appear as bright spots or patches, but offers only qualitative information about signal intensity. The CRT screen cannot present any true three-dimensional picture, of course; this is the most fundamental limitation of the cathode-ray tube as a radar indicator. Under very simple conditions, it is possible to present a third dimension in an understandable though unnatural way, but in any complicated three-dimensional situation one is forced to use more than one display.

The fundamental geometrical quantities involved in radar displays are the spherical coordinates of the target measured from the origin at the radar antenna: range, azimuth angle (or bearing), and elevation angle. Practically every radar display uses one or two of these quantities directly as coordinates on the tube face, or is a simple modification of a display which does.

The vast majority of displays use as one coordinate the value of slant range, its horizontal projection (ground range), or its vertical projection (altitude). Since slant range is involved in every radar situation, it inevitably appears in at least one display on every set. It is the coordinate most frequently duplicated when more than one type of display is used by a given radar set, partly because displays presenting range have the greatest signal-to-noise discrimination, and partly for geometrical reasons.

Range is displayed by causing the electron beam of the CRT to move (*sweep*) across the tube at a uniform rate¹ starting from a given point or line at a definite time in each pulse cycle. Thus, distances on the tube face along this *range sweep* or *time base* are proportional to increments of slant range. The *scale factor* relates distances on the tube to actual range, and the *sweep length* is the distance represented; both are determined by speed of the sweep. The origin of range may be on or off the tube face; sometimes the start of the sweep is delayed for some time after the instant of transmission of the outgoing pulse.

The angle in which the scanner is pointing, either in azimuth or elevation, may enter into a display (1) directly as a polar angle, (2)

¹ The range sweep must be nonuniform when it is desired to project range in a plane not containing the radar set; as, for example, in the case of true ground mapping on the indicator of an airborne radar.

directly as a cartesian coordinate, or (3) as the basis for resolving a range sweep in a particular direction. This will be further discussed in connection with specific types of display.

Indices or Scale Markers.—Both mechanical and electrical means are used for making geometrical measurements on the displays.

In measuring range, "electronic" range markers are practically always used. These consist of artificial video signals introduced into the display by a precision timing circuit; thus, inaccuracies in the display do not enter as sources of error in the measurements. Almost every display entails a set of discrete, regularly spaced markers derived from an oscillator properly phased with respect to the firing of the modulator. Fairly accurate ranges can be read from these markers at a glance. In many cases, however, interpolation errors are larger than can be tolerated; the fixed markers are then supplemented by a manually controlled, continuously movable, calibrated index. This index has the advantage of extremely high precision, especially on an expanded sweep, but it requires appreciable time in its use. Mechanical indices which move in front of the CRT face are seldom used in measuring range.

For the determination of angle, it has been most usual in the past to use mechanical indices: either a set of fixed indices engraved on a transparent overlay, or a movable mechanical cursor. These can give errors due to parallax and to inaccuracies in the display. More recently, electronic indices have begun to replace mechanical range markers. A discrete set of fixed indices, or a continuously variable index, may be provided; for some techniques of display synthesis, the latter is not easy to achieve.

The controls of movable markers, both of range and of angle, are often connected to devices providing remote data transmission.

6-3. Summary of Indicator Types.—Many considerations enter into the choice of the display geometry. In a three-dimensional problem, the designer must decide how to divide the coordinates between two displays, or how to present all the information on a single display, if this is feasible. Even a two-dimensional display is complicated, and often involves conflicting requirements, such as the need for high resolution in angle and range without sacrifice in the total field of view. In some cases, the needs can best be met by deliberately deforming the picture; in others it may be necessary to use more than one display, either alternately on a single tube or simultaneously on different ones.

Many different display schemes have been invented to deal with these problems. The following summary, which classifies them according to the spatial geometry represented, includes the important geometries actually used. Although many others are possible, they have had little or no practical application to date.

- I. One-dimensional deflection-modulated displays. These invariably present range. Two such displays on a single tube are sometimes used to obtain directional information by comparing signal intensities from two different beam lobes.
- II. Two-dimensional intensity-modulated displays.
 - A. The representation of a horizontal or vertical plane.
 1. Undeformed displays. Because of the nature of the radar data these are usually in the form of polar plots of range and angle (PPI).
 2. Deformed displays.
 - a. Radial deformation of a polar plot created by a shift in the range origin.
 - b. Linear deformation produced by "stretching" a polar plot along one rectangular axis.
 - c. Rectangular plots of range and angle.
 - B. Rectangular plots of azimuth and elevation.
 1. True displays which follow the antenna orientation.
 2. Error indicators. (Such displays are not always intensity-modulated.)
- III. Three-dimensional intensity-modulated displays. These are all modifications of two-dimensional displays which make use of one or both of the coordinates on the tube face to present, in a formalized way, information about the third dimension being displayed.

6-4. One-dimensional Deflection-modulated Displays.—Since one-dimensional displays yield little geometrical information, their only justification is that they permit the use of deflection modulation, which gives a maximum of information about the intensity and form of the echo signals. For this purpose, it is best to display the signals as a function of time or range, and thus the only displays using deflection modulation are those in which the deflections are applied perpendicular to a range sweep (Fig. 6-1). This sweep may represent either a small part or nearly all of the period between successive pulses. In the former case, the particular range interval that appears on the display is determined by the delay elapsing between the transmission of the outgoing pulse and the starting of the range sweep.

The general classification "type A" is used to describe one-dimensional displays. An A-scope is universally used for observing radar signals and various circuit waveforms in a radar set during test and alignment. It may be either part of the permanent installation or a piece of portable test equipment. The laboratory analogue of the A-scope has come to be known as a "synchroscope"; it is an indispensable tool in the design and testing of electronic circuits.

When the antenna of a modern narrow-beam radar is scanning, the beam passes over each target in a comparatively short interval of time. The resulting rapid appearance and disappearance of signals on an A-scope makes this type of display confusing and of little use as a primary indicator except for sets with very broad beams, extremely slow scanning rates, or the operational function of "searchlighting"—following a chosen target constantly. Even then, the chief utility of the A-scope

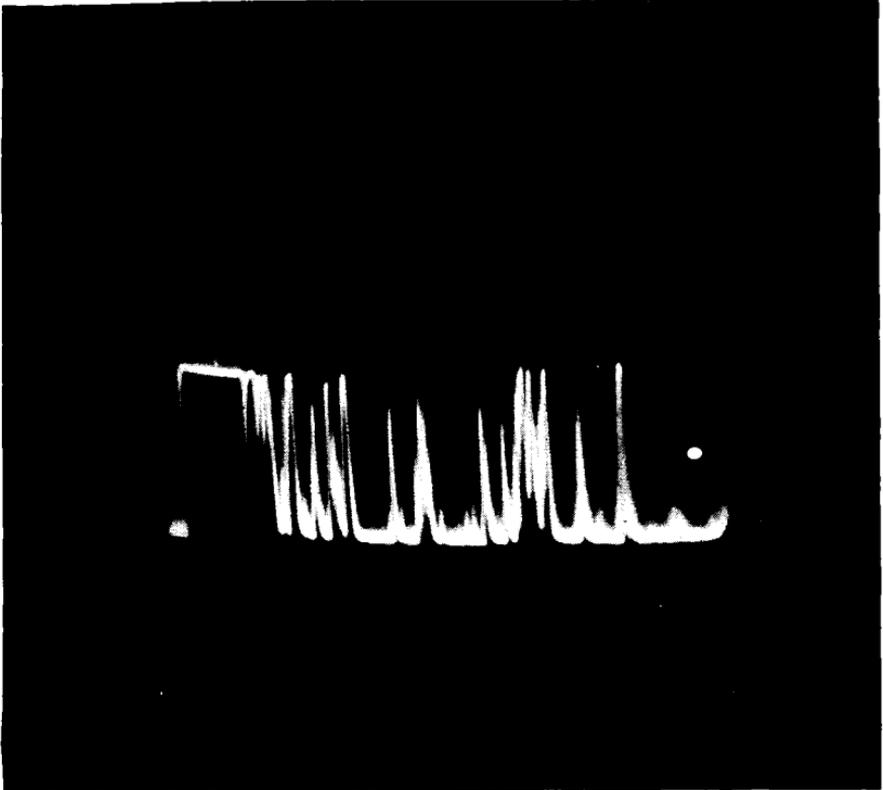


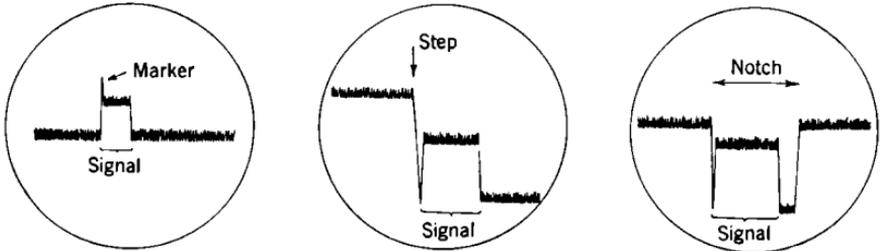
FIG. 6-1.—A-scope. Range increases from left to right.

is as an instrument for accurate range determination, or for reading certain types of coding of beacon signals (see Chap. 8).

Range Scopes.—The term "R-scope" (for range) is used to designate several forms of modified A-scope that are used for accurate range-finding. In all of them, a greatly expanded sweep is combined with a precision timing device. In some cases, the sweep delay is calibrated only crudely, the entire precision being embodied in a timing circuit which puts accurate range markers on the display itself. In others, the delay circuit itself is a precision device, and forms a part of the complete

timing equipment. The displays are sometimes subclassified in terms of the particular type of electronic marker used (see Fig. 6-2).

Sometimes an R-sweep and an A-sweep are shown simultaneously on the same tube by switching between them on alternate pulses. The A-sweep is used for general surveillance of targets and for determining

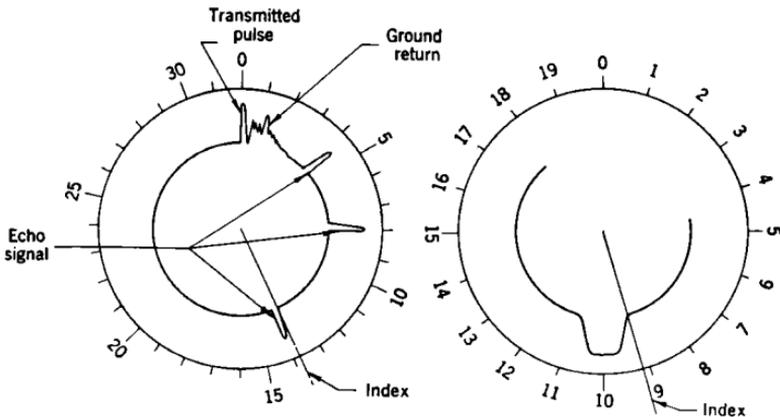


(a) R-scope with marker (b) R-scope with step marker (M-scope). (c) R-scope with notch.

FIG. 6-2.—Range scopes. Each displays a delayed, expanded, range interval.

the proper delay for the R-sweep, either by inspecting the range scale or by displaying on the A-sweep an electronic marker which indicates the setting of the R-sweep delay. Such an indicator is called an “A-and-R-scope.”

One form of deflection-modulated indicator which deserves special mention is the J-scope, in which the sweep is circular and the deflections



(a) Reading in thousands of yards. (b) Reading in hundreds of yards.

FIG. 6-3.—Coarse and fine J-scopes. The pair of scopes shown is that used in the SCR-584 (Sec. 6-14).

are applied radially by means of a central electrode (Fig. 6-3). The circular sweep is derived by applying a properly phased precision sine voltage in quadrature to two pairs of deflecting plates at right angles to one another. The frequency of the sweep sinusoid may be many times the pulse repetition frequency, in which case the intensity grid of the

CRT is used to blank the trace on all but one sweep cycle; a delayed expanded sweep is thus produced (see Fig. 6-3b).

"Pip-matching" Displays.—It is frequently useful to compare the strength of the radar echoes received from a single target by means of antennas whose patterns differ in direction (see, for example, Sec. 6-13). For convenience of comparison, the two echoes are usually presented on the same scope; they can be identified with the corresponding antenna by either of two arrangements—the K-scope or the L-scope. The K-scope (Fig. 6-4a) is so arranged that the range sweeps corresponding to the two antennas start from different origins, with the result that the echoes to be compared are side-by-side. In the L-scope, the signals from the two antennas produce deflections of opposite sign, the range



(a) Side-by-side presentation (K-scope).
Signal return from the right lobe is the stronger.

(b) Back-to-back presentation (L-scope).
Signal return from the right lobe is the stronger.

FIG. 6-4.—Pip-matching displays.

origin being common. The signals to be compared thus occur back-to-back (Fig. 6-4b).

6-5. Representation of the Horizontal Plane.—The large radius of curvature of the earth and the shallowness of the layer of air above it in which conventional aircraft are accustomed to fly make it useful to project all signals from radar targets in a horizontal plane. Most radar sets use a simple azimuth scan (with a beam that is fanned somewhat in the vertical direction if aircraft are to be observed), and present their data on a two-dimensional intensity-modulated display. Even when target height is important, some sort of range-azimuth display is usually basic to the indication system. The third coordinate usually appears in a separate presentation on another tube, often as a projection of all targets on a vertical plane. Intensity-modulated displays of plane surfaces constitute the most important class of radar indicators.

The Plan-position Indicator.—If the slant range and azimuth coordinates of the various targets in the field of view are represented, respectively, by distance from the center of a CRT and by azimuth on the tube face, the result is a map. Its only defect is the error introduced by using slant range instead of ground range (that is, slant range times cosine

of elevation angle). Except when viewing aircraft at very substantial elevation angles, or when mapping the ground from a high-flying aircraft, this error is imperceptible and can be neglected under practically all circumstances.

The indicator used for the display just described is called the *plan-position indicator* or PPI.¹ In the PPI, an outward range sweep is rotated about the range origin in synchronism with the azimuthal scanning of the antenna. Targets then appear in the proper directions as well as at the proper relative distances from this origin and the result is a map.

The origin of the PPI may be at the center of the cathode-ray tube, giving an equal field of view in all directions. Frequently, however,

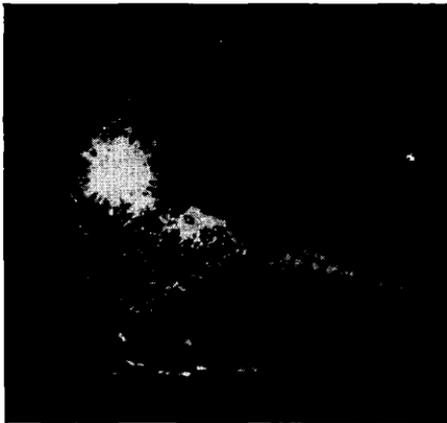


FIG. 6-5.—Off-center PPI.

it is displaced, sometimes far off the tube face, in order to give a maximum expansion to a given region; such a display is called an *off-center PPI* (Fig. 6-5). The expression *sector display* is often used when the displacement is extreme.

The PPI is the most widely used and versatile of all displays. In presenting the information with which it deals, its only fundamental shortcoming is that, in common with all maps or charts, it cannot simultaneously possess a highly expanded scale and a large field of view.

The vertical analogue of the PPI may be formed by substituting elevation for azimuth angle. However, since only a restricted range of elevation angle usually has interest, it is more customary to use distorted displays for this purpose.

¹ See Secs. 13-15 to 13-18. The abbreviation "PPI" is used interchangeably to represent both the display and the equipment (plan-position indicator). Abbreviations for other types of indicators are used in the same way.

In spite of the usefulness of a true map display such as the PPI, occasions arise in which it is not ideal. These situations usually involve the need for providing high resolution or dispersion in some particular coordinate without restricting the field of view too severely in other dimensions. A number of types of deformed displays have been devised for such purposes, and are particularly useful in dealing with point targets.

Radial Deformation.—A polar plot such as the PPI can be deformed radially by shifting the range origin, by using a nonlinear range scale, or both. The first means has been widely used.

On a normal PPI it is difficult to determine accurately the direction to a target whose range is a small fraction of that covered by the display. This difficulty is often partly overcome by expanding the zero-range origin of the PPI into a circle so that the radius vector to the echo from a nearby target is greatly increased. Such an arrangement is known as an *open-center PPI*. Range and bearing retain their identities and their linear scales. For the degree of center-opening ordinarily used, the deformation introduced into the sizes, shapes, and relative positions of the targets is serious for only a fractional part of the range portrayed. An open-center PPI has been widely employed for station-keeping and for homing on such objects as ships, airplanes, runways, and the like. The open center is switched in only a fraction of the time; for example, during the closing stages of a homing operation.

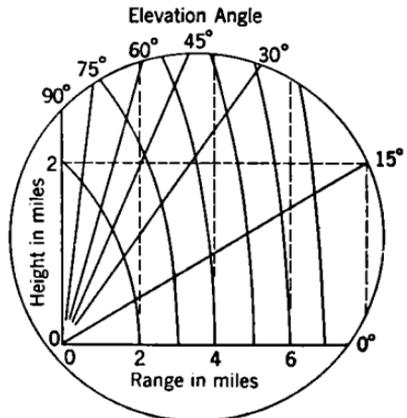


FIG. 6-6.—Stretched PPI. Two-mile grid is shown by dashed lines.

It is sometimes desirable to provide an expanded range scale over an interval at a distance from the radar site without sacrificing an all-round view as an off-center PPI would do. This may be accomplished by delaying and expanding the range sweep in an otherwise normal PPI so that a ring-shaped area is collapsed into a solid circle. Because of the technique used in producing it, this display is called a *delayed PPI*. As in the open-center PPI, range and bearing retain their identities and their linearity. The deformation is very great except at the extreme edge of the display. The delayed PPI is used to increase the resolution on closely spaced targets and on coded beacon signals, to facilitate accurate range measurements, and to increase the signal-to-noise or signal-to-clutter discernibility under some circumstances.

Linear Deformation.—A deformed display of considerable utility, especially in the vertical plane, can be formed by “stretching” a polar plot in one rectangular dimension as though it were on a sheet of rubber. On such a display, the cartesian coordinates parallel and perpendicular to the axis of stretch retain their original meaning and their linearity, but they have different scale factors. Straight lines remain straight but, except for those parallel to the coordinate axes, their directions are changed. Circles of equal range appear as ellipses with their major axes in the direction of the stretch.

In the horizontal plane this geometry, called the *stretched PPI*¹ (Fig.



FIG. 6-7.—B-scope. The azimuth interval displayed is 180°; the range marks are 2 miles apart.

6-6), is useful principally in connection with the control from a remote point of aircraft approaching a landing, or ships navigating a channel. The stretching, done in a direction perpendicular to the desired course, aids greatly in detecting slight deviations therefrom.

This technique finds its greatest utility in the vertical plane where vertical stretching is used to enhance the accuracy of aircraft height determination (see discussion of RHI, Sec. 6-6).

Rectangular Presentation of Range and Angle.—A plane surface is often represented in a deformed manner by combining range and angle in cartesian rather than in polar coordinates. This

is accomplished by moving a range sweep laterally across the tube face in synchronism with the antenna motion so that the origin is stretched out into a line.

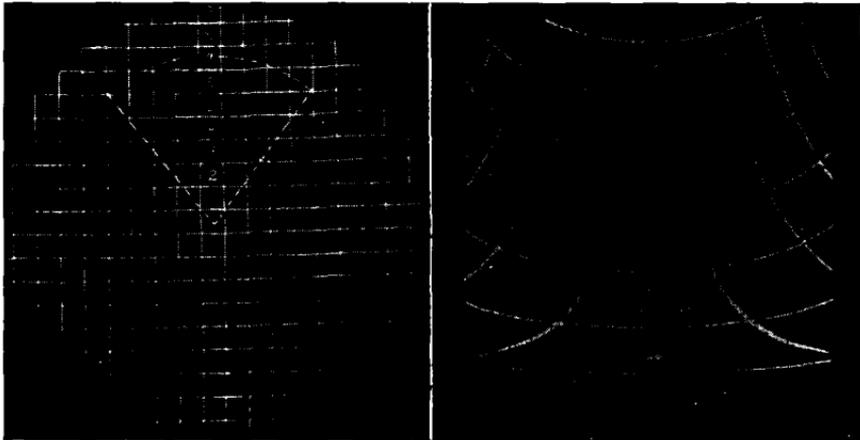
In range and azimuth these rectangular displays are of two different sorts:

1. Displays in which no attempt is made to minimize the deformation, either because it is unimportant in the particular circumstances, or because certain advantages can be gained by neglecting it. Any desired range interval and azimuthal sector may be covered, although in practice more than 180° is rarely used. The display is always normalized to make optimum use of the tube face.

¹ This has sometimes been called, erroneously, an “expanded azimuth” display.

Range is always plotted vertically, and at short range the resolution in angle is far greater than that afforded by the PPI. This display, called "type B," (shown in Figs. 6-7 and 6-8) is used in situations where the chief considerations are the range and bearing of point targets or groups of targets, with little or no importance attached to the shapes of extended targets or the relative locations of widely separated targets. It is especially useful in homing, and in telling plots of range and bearing for recording elsewhere in polar form.

2. Displays in which the desired angular field of view is so small that the distortion can be made negligible by proper normalization of



(a) 150-mile sweep, normal PPI with grid squares. (b) 100-mile sweep, B-scan, zero delay.

FIG. 6-8.—Distortion caused by B-scope display.

the range and angle scales. Such a display, known as a *micro-B*, is simply a substitute for a PPI sector; it is used either because it is technically easier to attain, or because the same indicator is alternately used for a regular type B display. Proper normalization requires that the angular dispersion be kept proportional to the range to the center of the display.

The vertical analogue of the type B display, known as "type E," will be described in the next section.

6-6. Plane Displays Involving Elevation.—One rarely desires to know the elevation angle or altitude of an aircraft target without also requiring its range and bearing. Simultaneous azimuth and elevation information can be obtained by two-dimensional scanning with a single antenna, by the V-beam principle (Sec. 6-12), or by the use of separate radar systems scanning respectively in azimuth and in elevation. In the last case, the

elevation scanner must be adjustable in azimuth, since otherwise its beam must be so wide that range performance is poor or target identification is difficult.

Two-dimensional scanning limits range performance. The average power reaching a point target is reduced below that corresponding to a simple azimuth scan by a factor approximately equal to the number of

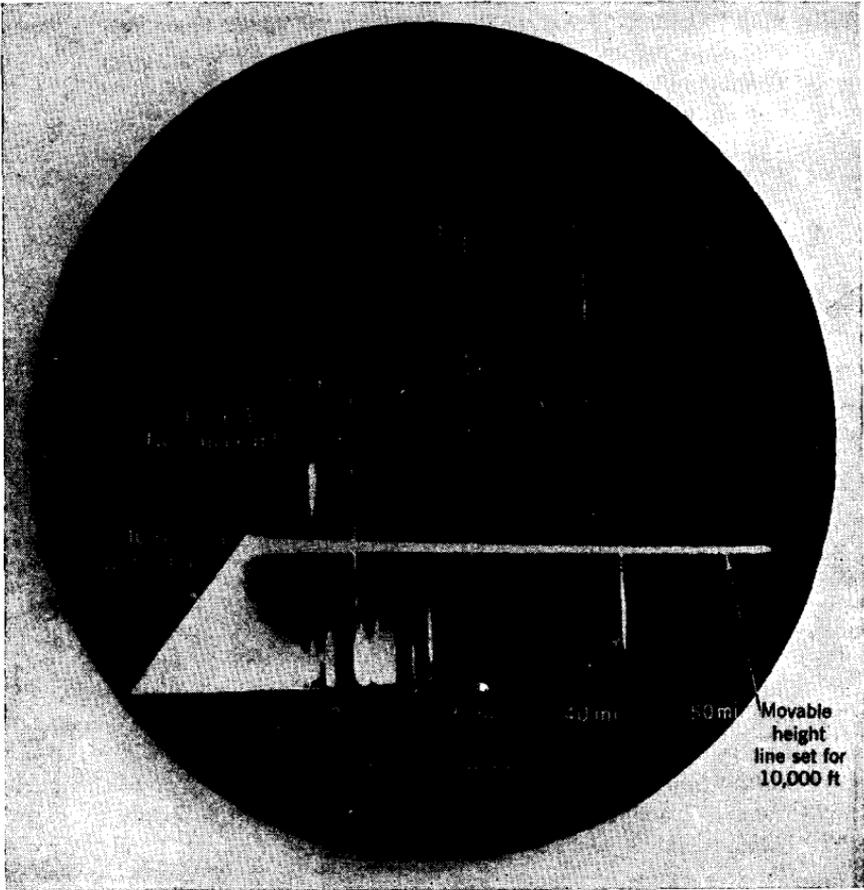


FIG. 6-9.—Range-height indicator (RHI).

beamwidths scanned vertically. The signal-to-noise discernibility on a PPI or a B-scope is reduced by this same factor, if the total scanning times are equal in the two cases.

The Range-height Indicator (RHI).—The RHI, shown in Fig. 6-9, is the elevation-angle analogue of the stretched PPI, as indicated in the last section. Target range is displayed as a horizontal coordinate and

the display is expanded in the vertical dimension, so that the height interval to be covered occupies approximately the same distance on the tube face. Lines of constant height are horizontal and equally spaced. Arrangements are sometimes made to move the range origin off the tube face to allow expansion of a region of interest.

If the height-finding antenna is required to scan over an appreciable range of azimuth angles, the RHI is usually blanked except during a relatively narrow azimuth sector, in order to improve the ratio of signal to noise and to avoid confusion between targets at different azimuths. This sector can be chosen by adjustment of a calibrated control.

The RHI is always used in conjunction with a PPI or other display of the horizontal plane; usually this auxiliary indicator obtains its data from another radar set. The signals from a given target are correlated in the two displays on the basis of range and azimuth position. In some cases the height-finding antenna searchlights in azimuth and is manually aimed in the proper direction. If it scans in azimuth, the center of the sector which is shown on the RHI is indicated by a mechanical cursor on the PPI of the search set.

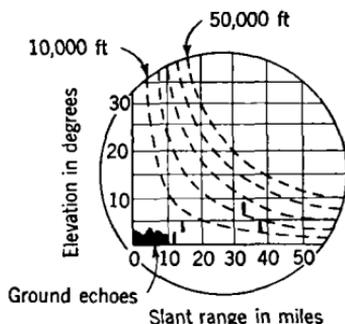


FIG. 6-10.—E-scope. Lines of equal height are shown dotted.

The E-scope.—The E-scope is a rectangular display in which range is the x-coordinate and elevation angle the y-coordinate. Lines of constant height are hyperbolas (Fig. 6-10). As in the case of the B-scope, a delayed range sweep is often used to allow range expansion. The elevation analogue of the micro-B (Sec. 6-5) has found no application, since the distortion of the E-scope is not particularly harmful and, by allowing it to occur, the dispersion in the elevation coordinate can be normalized to make the most effective use of the tube face.

In some older sets, the E-scope is used for height indication, but its distortion and its poor height dispersion at large ranges make it inferior to the RHI for this purpose.

The C-scope.—The type C display presents the azimuth and elevation angle coordinates of the scanner as rectangular coordinates on the tube

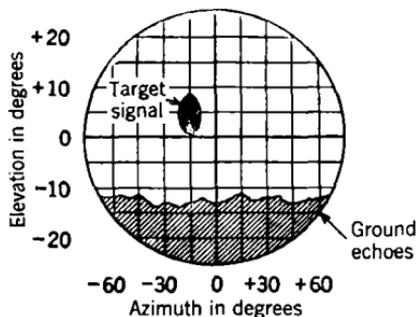


FIG. 6-11.—C-scope.

face (Fig. 6-11). It can be used in conjunction with a PPI or a type B display as an aid to homing (Sec. 6-13). This display has the great disadvantage that its signal-to-noise discrimination is very poor if signals are displayed for an appreciable fraction of the pulse cycle (see Chap. 2). For this reason, the C-scope is usually blanked except for a short range interval including the signal of interest. Initial recognition of the target signal is accomplished on the other display, and the proper range setting for the C-scope display interval is determined either by inspection or by displaying on the search tube an electronic marker derived from the C-scope brightening signal.

The *V-beam height indicator* is an example of the type of display considered in this section, but its description is deferred to Sec. 6-12.

6-7. Three-dimensional Displays.—In spite of the fact that the face of the cathode-ray is two-dimensional, two conventionalized three-

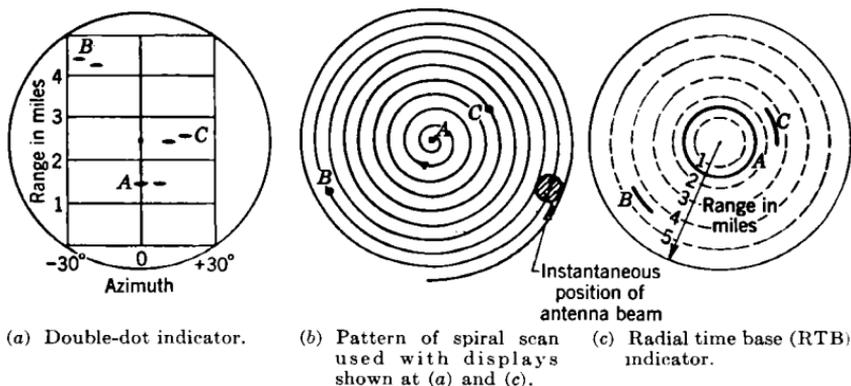


FIG. 6-12.—Three-dimensional displays. A, B, and C are targets.

dimensional displays [as distinguished from error indicators (Sec. 6-8)] have been successfully used in situations involving relatively few targets.

The "Double-dot" Display.—A modified type B presentation on which elevation angle is roughly indicated is known as the "double-dot" display. On alternate range sweeps the origin of the pattern is moved to the right and left respectively by a fixed amount (Fig. 6-12a). On the sweeps corresponding to the right-hand position the origin is simultaneously shifted vertically by an amount proportional to the sine of the elevation angle. Any single echo therefore appears in two neighboring positions, and the slope of the line joining the two "dots" is a rough measure of elevation angle which is accurate to two or three degrees under the usual circumstances of use. This display has been used in air-to-air homing (AN/APS-6).

The "Radial Time Base" Indicator.—The radial-time-base indicator is a three-dimensional display which, like the double-dot display, is best

suiting for use with a spiral scan. A range sweep moves from the center of the tube in a direction corresponding to the projection of the antenna beam on a plane perpendicular to the axis of scanning (Fig. 6-12c). The echoes therefore appear at radial distances corresponding to their range. If the target is on the symmetry axis of the scan, it is equally illuminated at all "spin" angles and the echo appears as a full circle. If it is far off axis it is illuminated only through a narrow spin angle (target B) and the arc on the display is short. Target C indicates an intermediate case. The display is surprisingly easy to interpret after a short period of observation.

6-8. Error Indicators.—A cathode-ray tube may be used as a radar indicator in a quite different way from those so far described; that is, as a meter on which to display various forms of intelligence. One common use of an error indicator is to indicate accuracy of pointing in connection with a conical scan. The signal intensity is combined electrically with the scanning information to provide voltages proportional to (or at least increasing with) the pointing error in both azimuth and elevation. These voltages are used to deflect the spot. A departure of the spot from the tube center indicates the direction and, to a qualitative degree at least, the magnitude of the pointing error. The radar signal is sometimes used to intensify the spot in order to distinguish between perfect pointing and no target. Range can be indicated by causing the spot to grow "wings" whose length is some rough inverse measure of the range. In this form (Fig. 6-13) the indication gives to a surprising degree the illusion of an actual aircraft which apparently grows larger as it approaches.

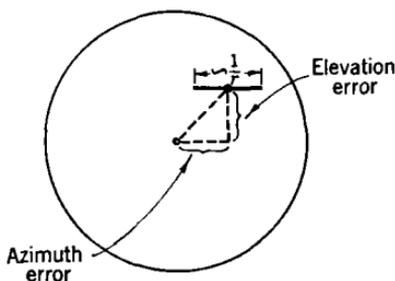


FIG. 6-13.—Spot error indicator with wings to indicate target range.

EXAMPLES OF THE MAJOR OPERATIONAL REQUIREMENTS

6-9. Early Aircraft-warning Radar.¹—*The British Home Chain.*—The first radar to have actual combat use was the British CH (Chain, Home) equipment. Work was begun in 1936 toward setting up five stations, about 25 miles apart, to protect London and the Thames estuary, and by the time war broke out in 1939 this Chain had been extended to cover the greater part of the south and east coasts of England and Scotland. During the next few years the Chain was extended and its equipment and performance were improved. In spite of the introduction of more modern and efficient equipment, the Chain provided the basis of the British

¹ By B. V. Bowden and L. N. Ridenour.

reporting system for the greater part of the war. At the very end of the war the system was used to plot V-2 rockets. The great success of the equipment, which saved the country from defeat during the Battle of Britain, had an important effect on the course of radar development. It called attention to the revolutionary usefulness of radar in modern air warfare, but at the same time it tended to freeze operational thinking concerning radar along lines of the static defense which was all that the CH stations could provide.

The main Chain operated at frequencies between 22 and 28 Mc/sec. In spite of the disadvantages inevitable in any system operating on so low a frequency, the choice was made deliberately after consideration of the tubes and techniques available at the time.

“Throughout the whole of the development period the small team who were responsible ruthlessly sacrificed all refinements, elegancies and versatilities in the desperate need for something to be going on with. They never turned aside from their cult of the third best—“The best never comes, the second best comes too late.”¹

At the frequency used, strong reflections from the ionosphere occur, and an extremely low repetition rate of 25 pps was used in order to ensure that ionospheric echoes were not confused with echoes from aircraft targets. The repetition rate could be reduced to 12.5 pps when ionospheric disturbances were unusually troublesome.

The pulse width could be varied from 6 to about 25 μ sec and the receiver bandwidth could be changed from 50 to 500 kc/sec. In normal operation the pulse length was about 12 μ sec and the bandwidth 150 kc/sec. The final design made use of continuously evacuated transmitting tubes with uncoated tungsten filaments, whose c-w power output was about 80 kw. Peak pulse power using these tubes was about 150 kw; later, it was increased to about one megawatt. Both transmitters and receivers were massive and elaborate. They were housed in separate bombproof buildings, usually about half a mile apart.

Separate antennas were used for transmitting and receiving. The main transmitting antenna was an array, usually of six dipoles with suitable tuned reflectors, hung between two gantries on a 350-ft steel tower. These towers were usually installed in sets of three in order to provide spare antennas and standby frequencies. Radiation from the transmitting antenna floodlit a quarter-sphere in front of the station, the greater part of the energy being confined within 20° of the horizontal and within approximately $\pm 50^\circ$ of the main “line of shoot.” Enough

¹ Sir R. A. Watson-Watt, “Radar in War and in Peace,” *Nature*, **156**, 319–324, (1945).

energy was radiated backwards to make it uncertain, without special arrangements, whether a target was in front of or behind a station.

The main receiving antenna consisted of a pair (or a stack of pairs) of crossed dipoles, mounted on 240-ft wooden towers. Balanced pairs of concentric feeders connected the dipoles to the field coils of a goniometer whose rotor was connected to the receiver. The final display was on a 12-in. A-scope. The range of a target was determined from a scale on the oscilloscope, and its azimuth by rotating the goniometer until the echo disappeared. Ambiguities of 180° were resolved by the use of

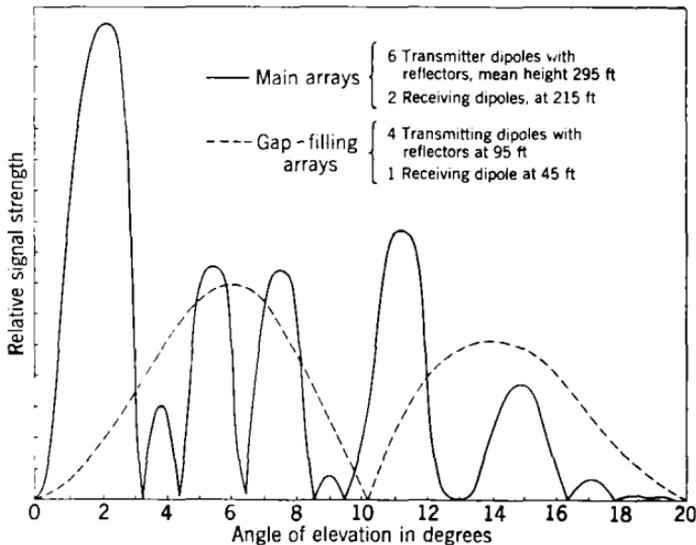


FIG. 6-14 —Typical elevation pattern of CH radar, east coast of England, 1939. Wavelength = 10.1 meters.

reflectors mounted near the receiving antennas. These reflectors could be switched in or out of circuit by remote control from the ground. From their effect on the echo, the operator could deduce whether the target was in front of or behind the station.

The performance of CH stations was controlled by the reflections from the ground of both the transmitted and the received signals. The antenna patterns of the transmitting and receiving arrays were not identical and there were large gaps in the vertical coverage of the stations. In order to minimize the importance of the gaps, separate auxiliary "gap-filling" transmitting and receiving arrays were installed to allow aircraft to be followed through the gaps produced by the main arrays (see Figs. 6-14 and 6-15). It was desirable that the ground surrounding a station should be flat for several miles in every direction; few stations had really ideal sites. Even when all possible precautions had been taken,

station performance was never predictable from theory. It was always necessary to "calibrate" a station by an elaborate series of test flights.

The maximum reliable range of a CH station was usually about 120 to 150 miles on single aircraft, but the range depended very much on the skill of the operator and of the maintenance crew as well as on meteorological

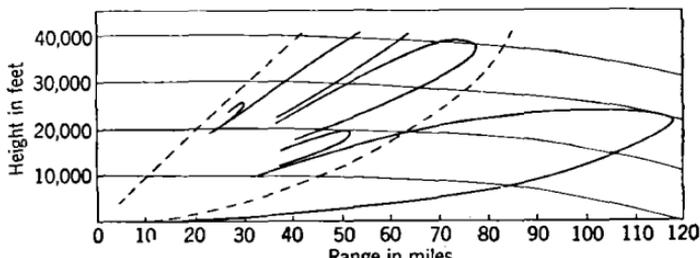


FIG. 6-15.—Polar diagram of typical CH station for reliable pickup of single fighter aircraft along the line of sight. Solid line represents main arrays; broken line represents gap-filling arrays.

logical conditions. As early as 1941 a Photographic Reconnaissance Spitfire, flying at about 40,000 ft, was plotted by the CH station at Canewdon, just north of the Thames estuary, as it flew almost all the way to Genoa at a maximum range of over 450 miles.

Height-finding was performed by making use of the lobe pattern of the receiving antennas.

Figure 6-16 shows the vertical pattern produced by two receiving antennas, respectively at 220 and 90 ft above ground. If one compares the signals received in these two antennas as a function of angle of elevation, one obtains a curve shown dotted in Fig. 6-16. It follows that, by measuring the signal ratio, it is sometimes possible to deduce the angle of elevation of the target. The comparison was performed by means of the goniometer already described.

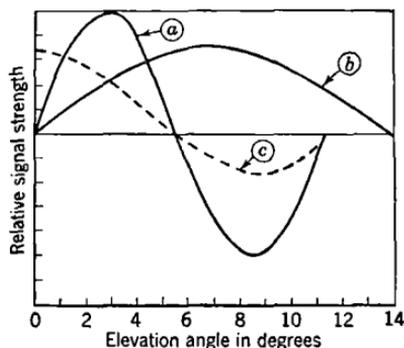


FIG. 6-16.—Relative signal strengths from CH height-finding antennas. Wavelength = 13.22 meters. (a) Two dipoles, mean height 220 ft. (b) One dipole at 90 ft. (c) Ratio of (a) to (b) on an arbitrary scale.

Under normal circumstances, the field coils of the goniometer were connected to two dipoles about 220 ft above the ground. By means of a switch, one goniometer coil was connected to a dipole at 220 ft, the other to a dipole at about 90 ft. By turning the goniometer to give minimum signal, the angle of elevation of the target could be determined and hence its height could be deduced (see Fig. 6-17).

Measurements of angles of elevation above about 7° were ambiguous on this so-called "A" system, and measurements of angles below about 1° were inaccurate owing to the "flatness" of the height curve for these small angles. The maximum usable angle of elevation was extended by using two receiving antennas 90 ft and 45 ft high respectively (the "B" system). Figure 6-17 shows the goniometer readings for the A and B systems as a function of angle of elevation. The curves are plotted for a flat site and for antennas perfectly matched to their feeders. In practice, the height curves for all stations were verified by an elaborate series of test flights. Owing to irregularities of sites the height curves varied considerably with azimuth, with the result that it was sometimes neces-

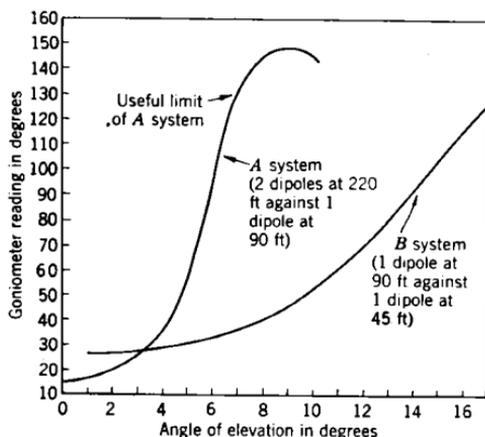


FIG. 6-17.—Height curves for CH station. Wavelength = 13.22 meters.

sary to use half a dozen different height calibrations. Some stations could measure height only in one or two favorable azimuths.

The operations involved in applying azimuth errors and interpreting height measurements were performed automatically by a system of standard telephone selector switches and relays. The operator measured range by a pointer on the tube, azimuth by turning the goniometer to give zero response from the echo. She¹ then pressed a button and the grid reference (Chap. 7) of the target appeared on a board. The operator then switched to height-finding, turned the goniometer again and pressed a button, whereupon the height was displayed in front of the teller. If the height measurement was ambiguous or unreliable, that fact was shown by the machine, and the operator could try again on the "B" height system. Operational use of radar information is discussed in Chap. 7.

The shortcomings of the CH system are fairly clear. Its long wavelength necessitated the use of large and expensive towers. There

¹ After 1940 almost all operators were women.

were large gaps in its vertical coverage (Figs. 6-14 and 6-15). The range on low-flying aircraft was poor. The system was supplemented in 1940 by 200-Mc/sec equipment, the so-called "CHL" (Chain, Home, Low) system. It was often very difficult to find sites that were flat enough for stations. Since the station at Dover was on a 400-ft cliff and Ventnor was on a cliff 1000 ft above the sea, it was impossible for these stations to measure height at all. The processes involved in deducing the plan position and the height of an aircraft were complicated and not very reliable. The maximum traffic-handling capacity was low (a good operator could pass about 6 plots per minute), and the system could not keep track of large numbers of independently routed aircraft. Nevertheless, formations of aircraft were plotted reliably, and the number of aircraft in a formation could be estimated by an experienced operator. With all its imperfections, this system was the basis of the British radar defense for the greater part of the war.

The CXAM.—While the British Home Chain was being designed and installed, the Army and Navy development agencies in the United States were independently developing pulse radar equipment. The earliest service equipment to be commercially manufactured was the CXAM radar, designed at the Naval Research Laboratory. A laboratory-built prototype of this set was tested at sea during battle maneuvers in early 1939, aboard the *U.S.S. New York*, and its performance was so promising that a contract was let in October 1939 for the manufacture of six sets of similar equipment.

The CXAM was operationally quite different from the CH. Instead of using separate, fixed, broad-beam transmitting and receiving arrays, it employed a common antenna for transmitting and receiving. To produce as narrow a beam as possible, it operated at the then ultrahigh frequency of 195 Mc/sec, and employed a "mattress" array of dipoles with reflectors, giving a gain of 40 and a beam 14° wide in azimuth by about 70° in elevation. The antenna could be rotated in azimuth at a speed of 5 rpm, or manually trained to follow a particular target. The peak pulse power was 15 kw, the pulsewidth 3 μ sec, and the repetition rate 1640 pps. Range against bombers was about 70 miles, against fighters about 50 miles.

The display was an A-scope in which the trace was lengthened by causing the sweep to take place from left to right across the tube, then drop down and return from right to left. Range was estimated with the help of marks on the face of the tube; bearing was determined as the direction of antenna-pointing which yielded maximum signal. The height of targets could be estimated with the help of nulls in the vertical antenna pattern (Sec. 6-11).

Despite its early design and its lack of adequate coverage against low-

flying aircraft, this simple, rugged equipment proved highly satisfactory in service use. It was the direct forerunner of the later shipborne long-range air-search radar equipments (SA, SC and its various redesigns, SK and its redesigns) used on large ships until the end of the war.

SCR-270 and SCR-271.—The early efforts of the U.S. Army Signal Corps to design service radar equipment were in the direction of produc-

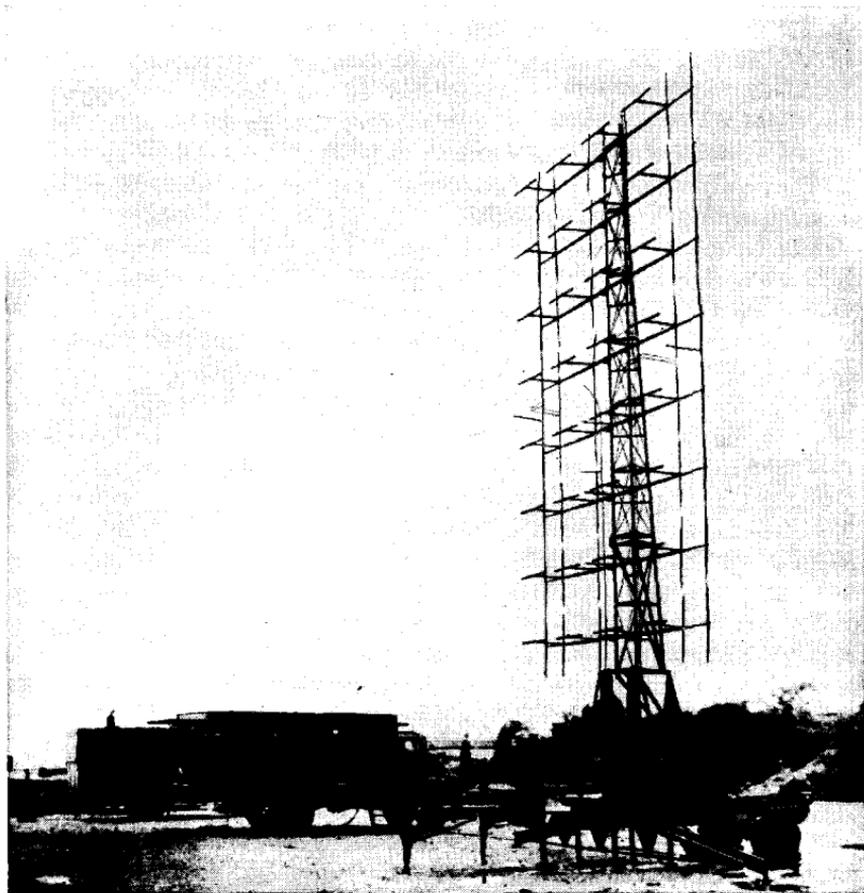


FIG. 6-18.—SCR-270D.

ing an equipment for antiaircraft fire control and searchlight control (SCR-268; see Sec. 6-14). In May 1937 the Air Corps requested the development of a "long-range detector and tracker." One laboratory version of the resulting equipment was demonstrated in November 1939 and showed a range of more than 100 miles on bombers. In August 1940 a contract for quantity production of this equipment was let.

The set was built both for mobile use (SCR-270) and for fixed-station installation (SCR-271). During the general advance in radar technique, the set went through many detailed changes and improvements and was still in active use at the close of the Japanese war.

The SCR-270¹ and the CXAM are quite similar in their operating features. The SCR-270 operates at a frequency of 106 Mc/sec, has a peak power of 100 kw, a pulsewidth from 10 to 25 μ sec, and a reliable range against single bombing aircraft of more than 100 miles. The repetition rate is 621 pps. The same antenna is used for both transmission and reception. This antenna (Fig. 6-18) consists of a dipole array originally four dipoles wide and nine high, with a reflector; the antenna gain is 140 and the beamwidths 28° in azimuth and 10° in elevation. Provision is made either for continuous rotation of the antenna at 1 rpm or for manual training to maximize target signals.

The SCR-270 has an A-scope display, with a control for so varying the start of the sweep that a selected target signal can be brought under a marker on the face of the tube; a scale on the phasing control then indicates target range directly. Target bearing is determined by reading off the antenna azimuth when the echo signal from the target in question appears to be strongest. The azimuth accuracy depends on proper siting, as in the case of all radar making use of ground reflections, and attainable accuracy in practice is about $\pm 4^\circ$. Height-finding can be performed by the method of nulls (Sec. 6-11).

Later models of this set include, among other improvements, an antenna giving a beam narrower in azimuth, and a PPI indicator. The simplicity, reliability, and ruggedness of the SCR-270 made it a very useful and very widely used equipment despite its great bulk and weight, its poor low-angle coverage, its vulnerability to ground clutter in anything less than a very carefully chosen site, its lack of range resolution and azimuth discrimination, and its low traffic-handling capacity.

6-10. PPI Radar for Search, Control, and Pilotage.²—In the last section, an important historical development has been traced. The first radar sets used fixed antennas and floodlighting technique, and required that a manipulation be performed to find the azimuth of each target. Next were designed relatively narrow-beam, continuously scanning radars, in which each target in the field of view is displayed periodically, so that its azimuth and range can be read off at regular intervals. A further step was made with the development of the plan-position indicator, or PPI, which, although the radar beam was still narrow and still

¹ Since the radar design of the SCR-270 does not differ from that of the SCR-271, the designation SCR-270 will be used, as a matter of convenience, to stand for either equipment.

² By L. N. Ridenour.

showed instantaneously only those targets located in a narrow azimuth sector, made use of the persistent property of CRT screen phosphors to preserve from one scan to another the target signals returned on the most recent scan. The PPI thus presents a continuous map-like display of all targets in the field of view of its radar, and permits attention to be concentrated on particular targets without the penalty of losing sight of the general situation.

It is interesting to notice that, before the advent of the PPI, users of such equipment as the CXAM and the SCR-270 frequently constructed a plan display with pencil and paper, plotting in polar coordinates the targets whose range and azimuth were called off by the radar operator. Further, skilled operators of such sets developed an ability to form, by watching the A-scope, a sort of mental PPI picture of the radar targets, remembering from scan to scan their ranges and bearing angles. When the PPI became available, the operator was released from this necessity, and could apply his skills to a more sophisticated mental appreciation of the target situation; for example, a good PPI operator will often keep in his mind the directions in which important targets are moving.

With the exception of special-purpose sets all radar equipment designed in 1941 or later contained one or more PPI scopes. In many types of equipment the use of the PPI is fundamental to the purpose of the radar. PPI radar sets can be divided into three classes: land-based, shipborne, or airborne.

Ground-based PPI sets are usually employed for the detection and plotting of aircraft, either to give early warning of enemy air activity, or to control friendly aircraft by radio instructions, or both. Mounted on shore, such sets can also be used for plotting ship traffic to permit warning and control, but this use is simply a two-dimensional case of the same problem as that involved in plotting aircraft.

In wartime, radar sets on shipboard must provide air-search and control facilities, and in addition must give a display of surface targets, to permit station-keeping in formation and pilotage in narrow waters under conditions of poor visibility. In peacetime, usually only the surface-search function is of importance. Because of the different beam-shape requirements, separate radar systems are ordinarily used for air search and for surface search.

Airborne radar equipments employing a PPI are used for pilotage in the vicinity of shorelines or over land. This pilotage may be of a general navigational character, as it always is in peacetime and usually is in war; or it may be as precise as the limitations of the radar will permit, for the purpose of blind bombing of a radar target.

The narrow beam provided by a microwave radar is important in all these cases. In ground radar, a narrow beam permits the detailed resolu-

tion of a complicated air situation, and also is beneficial in reducing the effects of ground clutter. Ship radar for displaying surface targets will show greater range, other things being equal, the shorter the wavelength used. This is a result of interference between the direct beam and the beam reflected from the water surface (Sec. 2-12). The greater azimuth resolution provided by a sharper beam is also useful in resolving complicated target situations. The only drawback of microwave radar for ship search is the resultant narrowing of the beam in elevation; in order to keep the rolling of the ship from tilting the beam away from the surface of the sea, either the antenna mount must be mechanically stabilized (Chap. 9), or the vertical dimensions of the antenna must be reduced to fan the beam sufficiently in elevation to take account of roll. In airborne PPI radar, the attainment of sufficiently good display detail to permit navigation over land, away from distinctive shorelines, demands the use of microwave frequencies. In fact, a considerable premium is placed on attainment of the narrowest azimuth beamwidth possible. Antenna stabilization is also desirable in aircraft, to correct for the effects of changes in attitude; but it is so costly in weight and complication that it has been seldom used.

Typical examples of PPI radar systems intended for ground and for airborne use are analyzed in detail in Chap. 15.

6-11. Height-finding Involving Ground Reflection.¹—The height of an aircraft target is usually determined by finding separately its range and its angle of elevation, and then solving the equation $H = R \sin e$. In the wartime use of radar, height was such an important datum that considerable ingenuity was expended on the problem of its measurement.

The resulting methods can be roughly separated into two classes depending on whether the reflection of energy from the level surface surrounding the radar antenna is important to the scheme of height-finding used. In this section are considered methods which do involve ground (or sea) reflection; in the next, free-space beam methods which do not.

If the antenna of a radar of medium wavelength, for example, about one meter, is less than about fifty wavelengths above the ground, the reflected energy from the part of the beam that strikes the ground will produce maxima and minima in the elevation pattern. These are shown for a representative case in Fig. 6-19.

Null Readings.—In a pattern such as that shown in Fig. 6-19, the range at which an aircraft first appears on the screen depends upon the height of the aircraft. If the antenna pattern is known, height may be estimated from range of first appearance. If the course of the aircraft

¹ By E. C. Pollard and L. N. Ridenour.

is approximately directly toward the station, the range at which the first fade is seen is also determined by the target height. If several fade points are determined, the height can be inferred with better accuracy. This method has been used extensively with the 200-Mc/sec aircraft-search radar used on naval vessels. It is especially useful with this type of radar because the elevation pattern can be well known (the ocean is a good reflector, and the antenna height is determined by the class of ship and not by the shape of the terrain), and because a radial approach to a task force is generally chosen by either enemy or friend.

Height-finding by nulls, however, has inherent drawbacks which seriously reduce its accuracy in operational use. The elevation pattern of a ground-based set must be carefully calibrated experimentally, because the reflection coefficient of the surrounding ground will determine the pattern, and because uneven ground gives different elevation patterns at different azimuths. Further, changes in the aspect presented to the radar by the aircraft may cause great changes in the echo strength. Finally, variations in the atmospheric refractive index can cause large variations in the amount of radiation falling on the ground or water surface surrounding the radar and thus appreciably change the pattern.¹ Despite these difficulties, if there is plenty of friendly

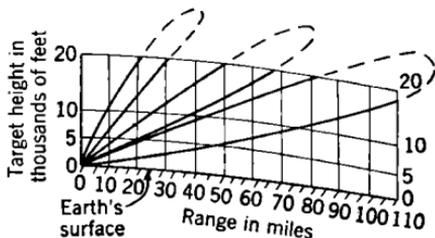


FIG. 6-19.—Typical elevation lobe pattern.

air traffic and if radio contact can be used to enable a continual check on height accuracy, adept radar plotters can give heights accurate to within one thousand feet more than half the time by the observation of signal fades. The method is slow; a reading cannot be made in less time than that required for a plane to fly through several nulls of the pattern. This drawback, and the elaborate calibration necessary at an overland site, have led to the development of other height-finding means, even for long-wave radar sets.

Signal Comparison.—A more rapid and convenient means of height-finding which is also based on the existence of maxima and minima in the elevation pattern of a long-wave radar employing ground reflection is signal comparison. Two or more antenna systems with different elevation patterns are provided at a single radar station, and the intensity of the echo received on one antenna is compared with that received on another. The comparison may be made either by measuring the vector

¹ Since refraction of one-half degree is possible, and since variations of echo strength due to aspect can be 12 db, altitude errors of several thousand feet can occur even in the favorable case of sea reflection.

resultant of simultaneously received signals (as is done in the CH by means of a goniometer; see Sec. 6-9), or by presenting alternately received signals side-by-side in a type K display and estimating visually their ratio in height.

The British Type 7 equipment is the most widely used radar exploiting

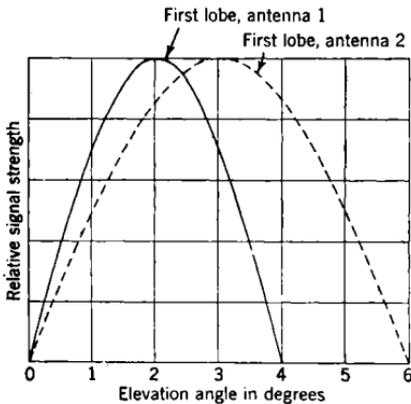


FIG. 6-20.—Elevation coverage diagram.

this signal-comparison principle. A copy of this equipment built in Canada for the United States is called the SCR-588; American-built redesigns of the same basic equipment are the SCR-527 and the SCR-627. The antenna of this set, is placed very close to the ground in such a way that the beam is tilted up in elevation and low-angle coverage deliberately sacrificed. A lobe pattern such as that shown in Fig. 6-20 is thus produced.

The lobes are few in number and very wide. A second antenna, placed at a different height, produces a second series of lobes which overlap those of the first antenna. The echo strength for one lobe system is compared with that for the lobe system from the other antenna. A type K display system is used which shows the two responses side-by-side on a 12-in. scope. From the relative intensities and the range of the aircraft, height can be deduced.

For either lobe pattern, the electric field strength at the target depends on the phase difference between the direct and reflected beams, and so will vary nearly sinusoidally with the altitude of the target. Figure 6-21 is drawn for antenna heights of 35 ft and 23 ft, on the assumption of sinusoidal variation. It shows the ratio between the signal strength in the lower beam and that in the upper as a function of elevation angle.

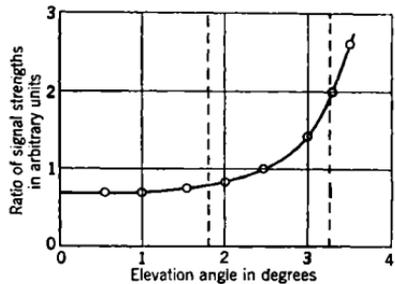


FIG. 6-21.—Ratio of signals in lower and upper beams.

The ratio varies slowly for small elevation angles, then rapidly as the lower lobe just touches the aircraft—in this case at an elevation angle of about 4°. By estimating the signal-strength ratio correct to 10 per cent, the angle of elevation can be told with an average accuracy of about one-third degree, corresponding to a height error at 50 miles of ± 1000 ft.

The angular range through which this height-finding scheme is useful is shown in Fig. 6-21 by dotted lines. At low angles the variation of ratio with angle is far too slow. At angles near 4° the variation is fast, but the illumination of the plane by the lower beam may be so weak that no echo is obtained, and therefore no height can be inferred. If both the antennas are nearer the ground, the effective angle is increased, but the beams are both tipped up so much that coverage on aircraft at low and medium altitude is sacrificed. Even in the case we have considered, an aircraft 10,000 ft high and 50 miles away, or one 5000 ft high at a range of 25 miles, would fall in the unmeasurable class.

To overcome these limitations, a variety of overlapping elevation lobe patterns are made available by providing a variety of antennas with phase and antiphase feeding. This additional complexity increases the time required to find height, since if the operator cannot guess which combination of antennas is suitable, a wrong choice makes it necessary for him to wait for a full 360° rotation of the antenna mount before he can try a better pair.

In addition to the elaboration necessary in equipment for height-finding by this method, siting is also a problem. A suitable site should be level within a few feet for a half-mile radius around the radar. Even sea reflection can be troublesome, because tides cause a profound change in the calibration of the station as a height finder. For these reasons, a free-space-beam height finder, not depending on ground reflection, is usually preferred.

6-12. Height-finding with a Free-space Beam.¹ *Searchlighting.*—The simplest form of free-space-beam height finder is one which measures elevation angle of the desired target directly, by pointing an antenna producing a narrow beam directly at the target and measuring the elevation angle at the antenna mount. This technique is called "searchlighting." Even with centimeter wavelengths, beamwidths attainable with antenna reflectors of practicable size (a few degrees) are large in comparison with the accuracy desired in the measurement of elevation angle (about a tenth of a degree) for height-finding purposes. Therefore, target height is not usually determined by directing the beam at the target simply by maximizing the echo.

Instead, arrangements are made for shifting the beam rapidly in elevation angle by an amount which may be, typically, a third of a beamwidth, and comparing the echo signal received in one position with that in the other. When the two signals are equal, the bisector of the angle between the two positions of the beam is pointed directly at the target. The signals from a given target may be compared for the two positions of the beam by presenting them side-by-side on a type K indicator or

¹ By E. C. Pollard, L. N. Ridenour, and D. C. Soper.

back-to-back on a type L indicator, or by using integrating methods of signal comparison and error display.

A simple means for moving the beam in elevation angle is to rotate the beam rapidly about an axis which makes a small angle with the center line of the beam; this is called "conical scanning." It will be seen that a conical scan permits the simultaneous determination not only of the elevation error in pointing the axis of the scan at the target, but also of the azimuth error in such pointing. Full consideration of this method for angular tracking of a single target is deferred to Sec. 6-14, but it will be useful here to describe a few principal features of a radar that employs conical scan solely for the purpose of height-finding.

This principle was used in the British CMH, a mobile 10-cm equipment designed specifically for height-finding. It used a 6-ft dish, a power of 500 kw, and was capable of measuring heights to within ± 500 ft.

The aircraft to be measured was followed manually in range with the aid of an electronic marker on the range trace. Two other operators followed manually in azimuth and elevation. The range and elevation settings were used to operate an automatic height indicator which could be placed in a remote position adjacent to a PPI fed by an auxiliary search radar.

Only a few of these equipments were built owing to progress in the development of a centimetric scanning system, and to the inherent inability of the CMH to find height on more than one target at a time. The SCR-615 was a similar U.S. equipment, except that it was designed as a fixed installation and employed an 8-ft dish.

The SM equipment used by the U.S. Navy was perhaps the most widely used of allied radar sets in this category. It was employed in conjunction with suitable PPI radar for ship control of aircraft. Its wavelength is 10.7 cm, and it employs an antenna reflector 8 ft in diameter. The waveguide feed used is a little off the axis of the paraboloid, so that the beam, whose width is 3° , is off axis by 1° . The feed is rotated to produce the conical scan. Pulses from the upper half of the antenna rotation are compared with those from the lower half by means of integrating circuits, and the output difference is used to displace the spot of an error indicator (Sec. 6-8). A PPI is provided for general reporting, and two range scopes, one covering the full range of the set and one an expanded A-scope, make it possible to determine target range and to eliminate from the signal-comparison circuits all echoes except those from the target whose altitude is being determined.

In operation, either a helical scan (Sec. 9-7) or a continuous rotation in azimuth with manually controlled elevation can be used in searching for targets. The helical scan can be set to cover any 12° in elevation angle in the range from 3° below to 75° above the horizon. When it is desired

to find height on a target, the antenna is stopped and manually kept pointed in the target azimuth. The target is located on the full-range A-scope, and the range interval covered by the expanded A-scope adjusted to include the target. A portion of the sweep of the expanded A-scope is deflected downward to form a "ditch" (Fig. 6-2c). Only signals received during the range interval represented by this "ditch" are fed to the integrating circuits which control the spot error indicator. The setting of the expanded A-scope is adjusted until the desired signal is in the "ditch," and the antenna pointing is adjusted to center the error indicator. When this has been done, the elevation angle of the line of sight to the target can be read off repeater dials which indicate the elevation angle of the mount. The altitude of the target, which is computed electrically as the product of range and sine of elevation angle, is shown directly on a meter calibrated in feet.

When SM equipment is used on shipboard, angles between the antenna and the structure of the ship must be properly corrected with respect to a stable vertical, maintained by a gyro, in order to yield the true angles which would have been measured with respect to a level structure. Considerable attention has been paid to the stabilization of radar antennas on ships and aircraft, and to the data correction that can sometimes be used as a substitute (see Chap. 9).

Elevation Scanning.—The height-finding methods so far described are all limited in accuracy, in speed of height-reading, or in the number of targets which can be dealt with at once by a single height-finding radar. The advantages of a system in which a narrow beam scans in elevation and the signals are displayed directly on an intensity-modulated tube whose coordinates are derived from range and elevation angle are apparent. These very important advantages were clear before the microwave radar art had advanced to the point where it was possible to obtain narrow beams and adequate range performance with reasonably small antenna structures.

As a result, the British developed a system operated at 200 Mc/sec which provided elevation scanning with a fixed antenna. It was called Variable Elevation Beam (VEB). Two versions were developed, one using a 240-ft mast and the other a 120-ft mast, the latter being intended for rapid installation. Such tall structures were necessary in order to obtain a narrow beamwidth in elevation at this low frequency.

The antenna system on a 240-ft mast consists of nine groups, each of eight horizontal dipoles, spaced vertically one above another. This version scans over a range of elevation angle from $\frac{3}{4}^{\circ}$ to $7\frac{1}{2}^{\circ}$. The beam-tilting is achieved by altering the relative phasing of the groups of dipoles; each group is fed from a mechanical phase shifter mounted in the center of the mast. The banks of dipoles can be tilted mechanically

to move the scanning range to the elevation interval from $7\frac{1}{2}^{\circ}$ to 15° . Elevation can be read to an accuracy of $\pm 0.15^{\circ}$ and readings can be obtained to a maximum range of about 70 miles.

The 120-ft system consists of nine groups of four dipoles each, stacked vertically. Its scan covers the range of elevation angles from $1\frac{3}{4}^{\circ}$ to 15° . Elevation angle readings can be obtained to an accuracy of $\pm 0.3^{\circ}$ out to about 55 miles range.

The horizontal antenna pattern of both sets is a fixed wide beam covering approximately 60° on either side of a predetermined "line of shoot."

The indicator is an E-scope in which a horizontal intensity-modulated range sweep moves vertically up the tube in synchronism with the move-

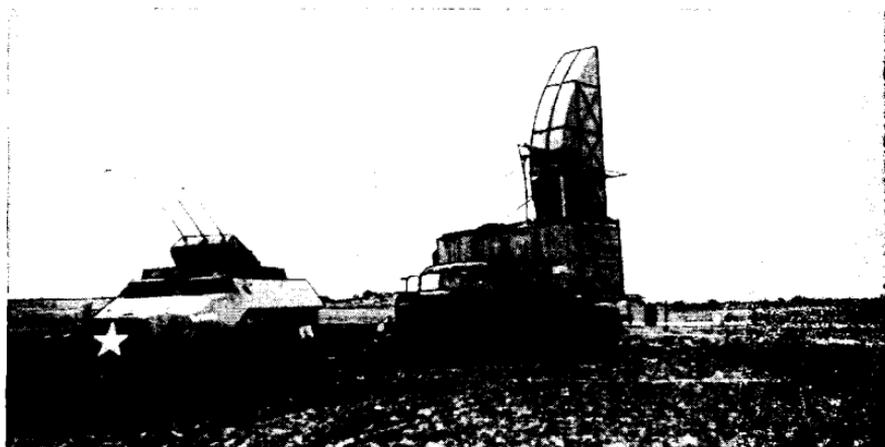


FIG. 6-22.—British A.M.E.S. Type 13 height finder.

ment of the radar beam. Signals appear as vertical lines at a certain range along the trace centered about the elevation angle of the target. Calculation of height from range and elevation angle is performed automatically by the selector switch and relay equipment used in CH stations (Sec. 6-9). The range and elevation of the target are set into the computer by the operator, who adjusts range and angle markers to cross at the center of the target signal on the scope. The computer indicates target height in numerals on a lamp display.

Development of the microwave art permitted the design of elevation-scanning height finders which, unlike the VEB, could be made mobile. The first such set to be used operationally was the British Type 13. The reason for the late appearance of this very successful height finder is the technical difficulty of illuminating an antenna reflector of large aperture ratio. This is accomplished in the Type 13 by illuminating a double "cheese" (Fig. 6-22) by means of a horn feed (Chap. 9). Although side

lobes are not objectionable in this antenna, it displays considerable frequency sensitivity. A standing-wave voltage ratio of 1.33 (Chap. 11) is encountered for a frequency change of 2 per cent. The beamwidths are 1° in the vertical dimension and 6° in the horizontal dimension. The pulse width is $2 \mu\text{sec}$, the output power 500 kw.

In operation, the antenna is moved up and down in the expected azimuth of search at a rate of one oscillation in ten seconds. An RHI display is used (Fig. 6-23). Because of the slow rate of scan, it is imperative to use a second radar to determine the approximate direction of the target. The Type 7 equipment was widely used for this function. The azimuth position of the Type 13 radar is shown on the PPI of the Type 7 by means of a beam of light projected from behind the phosphorescent screen of the PPI. This light is yellowish and does not affect the persistence of the radar signals on the PPI. In addition there is a color contrast between it and the echoes. By rotating the antenna of the Type 13 until this "azicator" line cuts the center of the signal on the PPI, the Type 13 can be correctly enough pointed to display the desired target. The beamwidth of 6° is chosen to be wide enough to permit Type 13 operation with a PPI radar having very poor resolution.

A similar American set (AN/TPS-10) is lighter and simpler since it operates at 3 cm. This enables the antenna size to be reduced to 10 by 3 ft with the very high gain of 18,000. The range is over 50 miles on a four-motored aircraft, although the radar set uses only 60-kw output power at $1\text{-}\mu\text{sec}$ pulse width and 1000 pps. The antenna can be oscillated once per second. The beamwidths are 0.7° in elevation and 2° in azimuth. This last is much smaller than was at first thought feasible; it was argued that so small an azimuth beamwidth would render search for the target laborious. It has proved, however, to be no operational limitation.

Height-finding by such a set is reasonably accurate. The results of a calibration flight with an experimental model of the AN/TPS-10 are shown in Fig. 6-24. Relative height is accurate to about ± 300 ft under favorable conditions. Absolute height is more difficult to find accurately by radar, because of the bending of the beam by atmospheric refraction, which varies from time to time. Fortunately, relative height is usually sufficient for aircraft control. Further, it is generally possible to keep

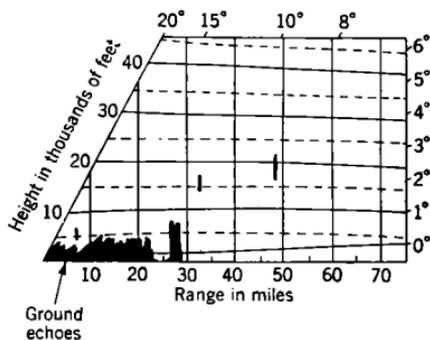


FIG. 6-23.-Range-height indicator of Type 13.

track of refraction conditions by talking to pilots of aircraft seen on the radar, and checking radar height readings with their altimeters.

Combined Plan and Height Systems.—Despite the speed, convenience, and accuracy of the elevation-scanning height finders just described, these sets are merely auxiliary to a search radar. Equipments that provide plan and height information continuously on all targets in the field of view, using a single antenna system, are obviously preferable. The limitations discussed in Chap. 4 on the range out to which a substantial area can be rapidly scanned preclude the design of a single radar system which will give both plan and height information with the necessary accuracy, as often as is desirable, and out to a range limited only by the optical horizon.

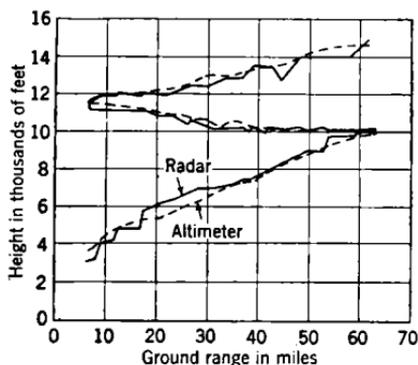


FIG. 6-24.—Calibration of height readings of AN/TPS-10.

Two or more radar systems can be combined to allow continuous and rapid indication of both plan position and height of all targets in the field of view. One example of such a radar is a ship-based set which combines a conventional PPI radar with an elevation-scan height finder of advanced design. Another, working on quite a different principle, is the V-beam radar, which reduces the measurement of height to two measurements of target azimuth.

The ship radar uses a search set and a height finder whose antennas occupy the same pedestal. The search set has a beamwidth in azimuth of 1.5° and elevation coverage up to 15° . The height finder antenna is 15 ft high by 5 ft wide, fed at 8.9-cm wavelength by a special feed (Sec. 9-16) which can cause the beam to scan at ten oscillations per second. The beamwidth in elevation is 1° and in azimuth 4° . The height finder antenna feed is so arranged that the pulses are distributed uniformly throughout the angle of scan, and not concentrated at the upper and lower angles as they are by the mechanical oscillation used in the sets just described. As a result, the antenna can be rotated at 4 rpm and height-finding to a range of 50 statute miles is quite possible. A 12° range of elevation angle can be covered.

The antenna is shown in Fig. 9-31. The display console is shown in Fig. 6-25. A three-fold indication scheme is used. One indicator gives the general PPI picture and a second gives an expansion in range and azimuth angle which enables detailed control to be accomplished. The third indicator is a range-height indicator. The RHI is brightened for a

selected range of azimuth angle which corresponds to that shown on the expanded range-azimuth indicator, so that the aircraft under detailed observation can be recognized on the RHI. Four or more display consoles, each with a full complement of indicators, can be operated independently and at the same time from a single radar.

The fundamental principle of the V-beam radar is to reduce the observation of elevation angle to a double observation of azimuth angle. The

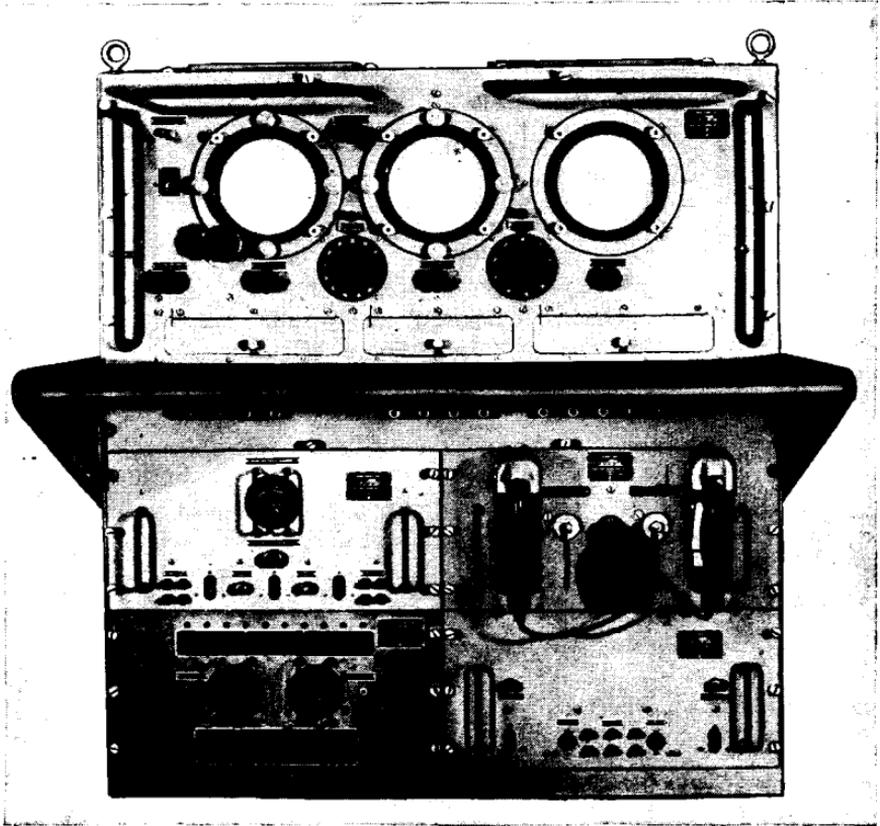


FIG. 6-25.—Console of ship-based height finding radar.

second observation of azimuth is made by a beam slanted at 45° to the ordinary search beam of a ground PPI radar. The two beams form a V-shaped trough.

Figure 6-26 (see also Fig. 4-1) shows the two beams and the direction of rotation of the antenna mount. The 45° V reproduces the value of the height as the horizontal distance between the vertical and slant beams at the appropriate range and height. This is expressible in terms of D , the ground range to the target. If $\Delta\phi$ is the angle turned by the mount,

$h = D \sin \Delta\phi$. But $D^2 = R^2 - h^2$, so that

$$h = \frac{R \sin \Delta\phi}{(1 + \sin^2 \Delta\phi)^{1/2}}$$

This simple relation permits height to be found if the target range and the difference angle can be determined. There is a danger that at

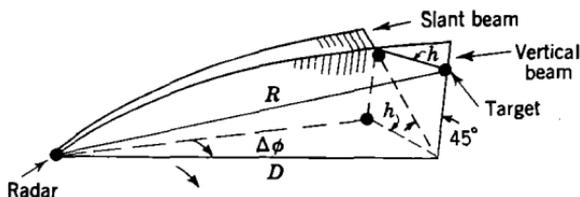


FIG. 6-26.—Schematic view of V-beam principle. For clarity, vertical and slant beams are shown without the 10° separation in azimuth which is used in practice.

low elevations the two signals from the vertical and slant beams might be displayed so close together that no height can be found. This problem is eliminated by introducing a fixed separation of 10° between the beams, so that there is a 10° separation between the signals even from a target at zero altitude. The indicator is designed around the general principle of a B-scope in that range is presented against angle. The simplest form of indicator has range along the horizontal axis and $\Delta\phi$ on the vertical axis (Fig. 6-27).

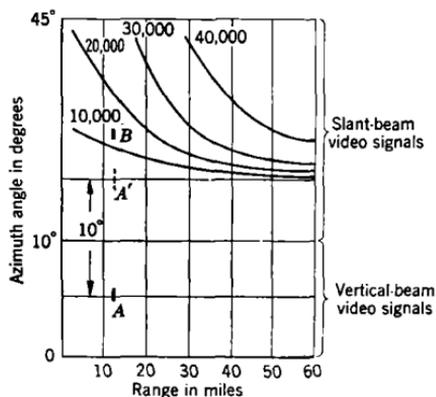


FIG. 6-27.—Simple V-beam indicator. Lines of constant height are shown.

Suppose a target appears at *A* on the vertical beam at an azimuth selected by the operator. Video output from the vertical beam appears on the scope for 10° of azimuth rotation centered on the selected azimuth. It is then blanked out and slant video substituted. If the target is at zero elevation, the slant beam will illuminate it 10° later, at *A'*. If the target has a definite height, the slant beam will rotate further before illuminating it and the signal will actually appear at *B*, as shown. A movable overlay superimposed on the scope face is set so that its baseline bisects the first signal at *A*. The 10° delay position is marked by a second line, which serves as the zero-height line for a succession of constant-height lines which appear as drawn. Height is estimated directly from the face of the scope. The two signals from a single target (e.g., *A* and *B*) always appear at the same range.

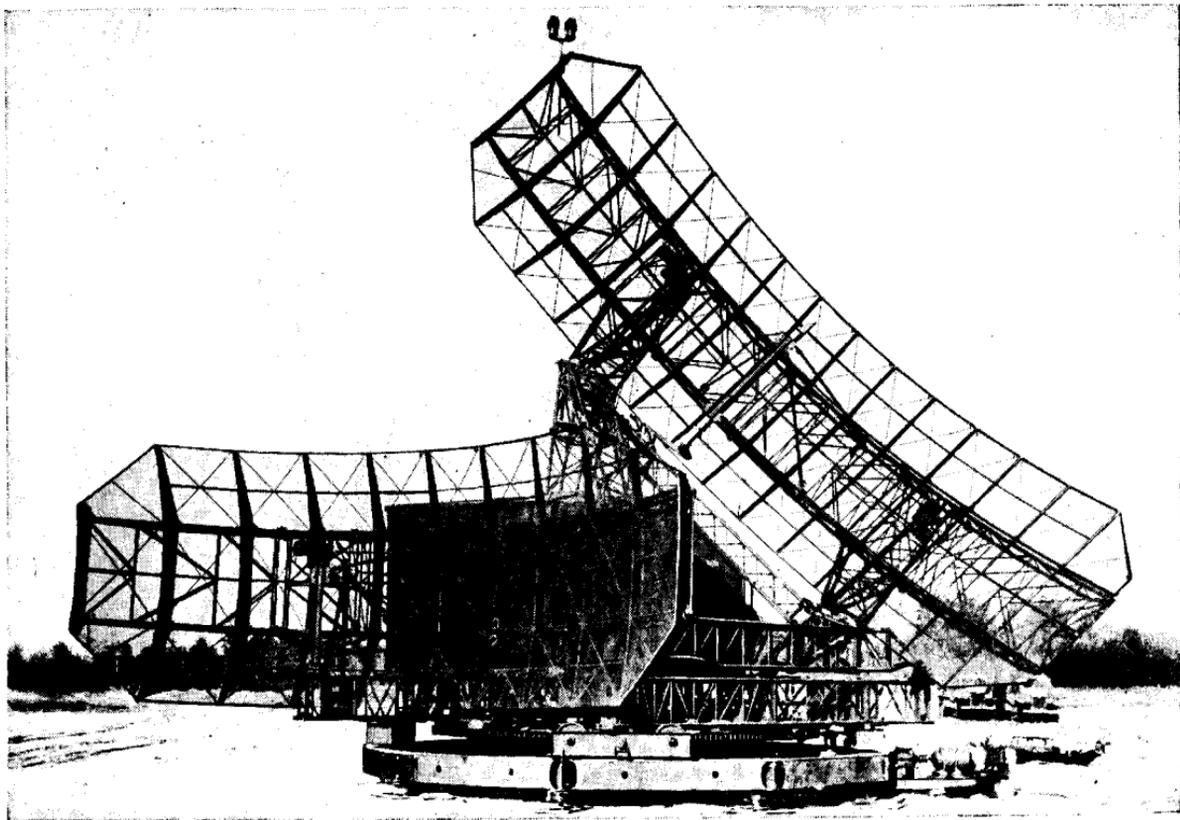


FIG. 6-28.—Antenna and turning gear for V-beam.

This type of height finder is free from the range limitation which arises from the scanning loss of a rapidly scanning system. The V-beam can, in principle, find height to any range where a search set of the same antenna size can detect aircraft. Heights on single aircraft out to 140 miles have been recorded.

The set is shown in Fig. 6-28. The two antenna reflectors are 25 by 10 ft. The vertical sheet of radiation is obtained by three separate transmitters feeding a horn and a series of dipoles. The slant beam uses only two transmitters, since early-warning range is ordinarily not needed on the slant beam which is intended for height-finding.

The results of height calibrations show that heights can be called rapidly to an accuracy of ± 1000 ft and, with care, to ± 500 ft. If an average is taken, better results can be obtained.

6-13. Homing.¹—The use of airborne radar to guide an aircraft to its target was extremely important during the war. Long-wave radar installed in aircraft is scarcely capable of performing any function more complicated than homing, since the enormous size of an antenna giving a beam sharp enough to produce a map-type display precludes its use in aircraft. Fortunately for the wartime development of airborne radar, homing is a tactically important function both in the attack of shipping by aircraft and in the interception of hostile aircraft by defending fighters. Sets for both purposes operating near 200 Mc/sec were extensively used during the war.

Homing on a Surface Target.—The simpler of the two problems is that of the detection and interception of shipping by patrol aircraft. Since targets are known to be on the surface of the sea, homing information need be supplied only in azimuth. The first operationally successful aircraft-to-surface-vessel, ASV, radar was the British ASV Mark II. A counterpart of this equipment was manufactured in the United States, being called SCR-521 by the Army and ASE by the Navy.

The SCR-521 operates at a frequency of 176 Mc/sec. Two sets of antennas are provided giving different beam patterns (Fig. 6-29); one pattern is used for search and the other for homing. Each pattern consists of a left and a right lobe. In the search position these lobes extend broadside to the airplane; in the homing position they extend nearly straight ahead, so that they overlap to a considerable extent. The exact design of these antennas underwent a complicated series of changes. In some equipments separate antennas were used for transmission and for reception, and in others both functions were carried out with a common antenna. In all cases, however, the general coverage and arrangement of the beams were those shown in Fig. 6-29.

The radar receiver is rapidly switched from the left to the right mem-

¹ By L. N. Ridenour.

ber of the pair of antennas in use, and the radar echoes are displayed on an L-scope in which signals coming from the left antenna cause a displacement of the trace toward the left, and signals from the right antenna displace the trace to the right (Fig. 6-4). The range sweep is linear, upward from the bottom of the scope, and the range is estimated by an engraved scale in front of the tube.

On the search antennas, a given target shows either on the left antenna or on the right, depending on its location with respect to the aircraft, never on both. On the homing antennas, however, because of the overlapping coverage of the two beams, a target will show a signal at the same range both to the left and to the right of the center line. A comparison of the strength of the two returns shows the radar observer in which direction course must be altered to home on the target. Care must be taken in antenna installation to make sure that the axis of equal signal coincides with the direction of flight of the aircraft. "Squint," which

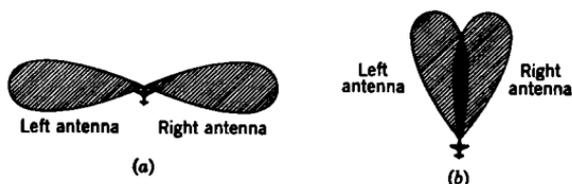


FIG. 6-29.—Beam patterns of ASV Mark II (SCR-521). (a) Search antenna pattern. (b) Homing antenna pattern.

results from improper installation or trimming of antennas, has the same operational effect as crabbing of the aircraft in a cross-wind; in either case, the operator notices that the relative signal strength of the left and right echoes changes when the aircraft is steady on a compass course initially chosen to give equal signals. A skillful operator can allow for this effect and choose a course on which the signals, though not equal, do not change relative to one another when the course is held. This, regardless of squint or crabbing, is a true interception course.

The first radar beacons (Chap. 8) were designed for use with this equipment. The type L display is just as well suited to homing on a beacon as it is to homing on a radar target, and the navigational aid provided by beacons was very useful in bringing aircraft home after long sea patrols.

A very similar equipment operationally, but of greatly reduced total weight, has been widely used by the U.S. Navy. Referred to as the ASB, it operates on 515 Mc/sec, at which frequency a dipole becomes small enough to allow the Yagi arrays used as antennas to be mounted on a rotating mechanism that permits them to be pointed broadside for search, or turned forward for homing, at the will of the operator (Fig.

6-30). A single Yagi on each side is used as a common transmitting and receiving antenna. The indication and the method of use of this equipment are the same as those of the ASV Mark II.

With the introduction of microwave airborne radar, sea-search equipment giving a map-type indication became practicable. Pilotage within 50 to 100 miles of islands or coastlines is very greatly facilitated by such

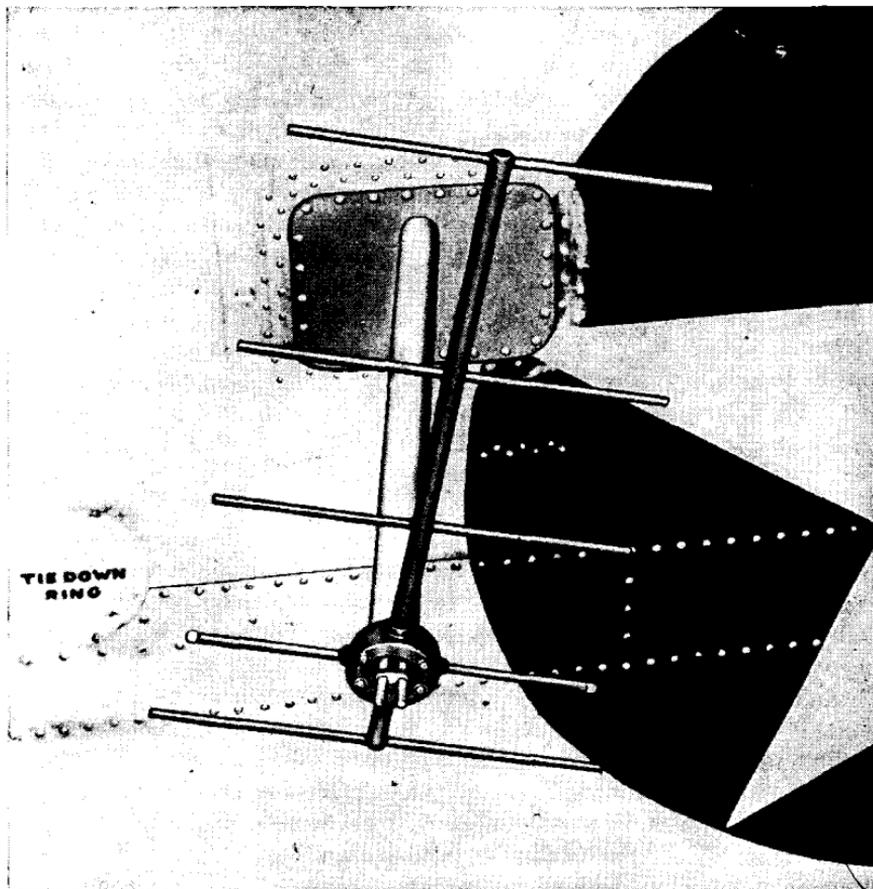


FIG. 6-30.—ASB antenna.

radar equipment, and the advantage of coverage of the complete field of view by such scanning radar (instead of coverage of the two sectors viewed by the equipments just described) is important in sizing up a complicated target situation. Only the bulk and weight of early microwave airborne radar militated against its use for sea search; these difficulties have now been overcome.

Since homing, as well as search, is important, special provision is

usually made to facilitate homing in a sea-search radar. This usually takes the form of a provision for "sector" scan and a modified indicator. Sector scan is, as its name implies, an arrangement for causing the antenna to scan back and forth across a desired sector, sometimes one centered dead ahead of the aircraft, and sometimes one that is adjustable at the will of the operator. The indicator provisions for homing are of either of two types: the set may have a type B display, which facilitates homing because the condition for an interception course is that the target signal approach down a vertical line of constant azimuth, or it may have an open-center PPI. In the latter, the target follows along a radius as the range is closed on an interception course; the open center reduces the crowding of signals at short ranges and makes homing easier. A movable engraved marker or electronic index showing constant azimuth is usually provided with either type of display to aid in homing.

The principal wartime American microwave ASV sets were:

1. The SCR-717, a 10-cm equipment whose antenna is a 30-in. paraboloid arranged either for continuous rotation or for sector scan, at the will of the operator, and whose indicator is, in different models, either a type B or a PPI arranged to permit center-opening.
2. The ASG, a very similar equipment made for the Navy by another manufacturer, which offers only PPI display.
3. The AN/APS-15, a 3-cm radar with optional sector scan and with PPI display, designed for overland bombing but used for sea search when its "cosecant-squared" antenna (Sec. 2·5) providing high-altitude coverage is replaced by one designed for about 5000 ft, the optimum altitude for sea search.
4. A series of Navy equipments operating at 3 cm and designed for the primary purpose of homing.

These latter sets are the ASD, an improved redesign of the ASD called the ASD-1 or the AN/APS-3, and the ASH or AN/APS-4, a set similar to the other two, but representing a very complete redesign with the principal object of reducing bulk and weight. The 18-in. paraboloid antennas of the ASD and the AN/APS-3 are mounted in a nacelle faired into the leading edge of the wing of a single-engined aircraft. (Fig. 6·31.) All of the AN/APS-4 except the indicator and the controls is mounted in a "bomb" hung under the plane on the conventional bomb shackles (Fig. 6·32). The paraboloid antenna reflector of the AN/APS-4 is 14 in. in diameter. In all these equipments, the 360° azimuth scan of the Army equipments is replaced by a wide sector scan centered along the line-of-flight of the aircraft, and covering about 160° in all. At the will of the operator, this scan can be replaced by a narrow sector scan executed more rapidly. Type B indication only is provided; provision is

made for expansion of the sector dead ahead when homing is being carried out. In these sets, the general mapping functions of microwave radar have been sacrificed in favor of the utmost simplicity and ease in carrying out the homing function.



FIG. 6-31.—Scanner housing of AN/APS-6 in wing of F6F. The installation of AN/APS-3 in single-engine carrier-based torpedo bombers is identical in external appearance.

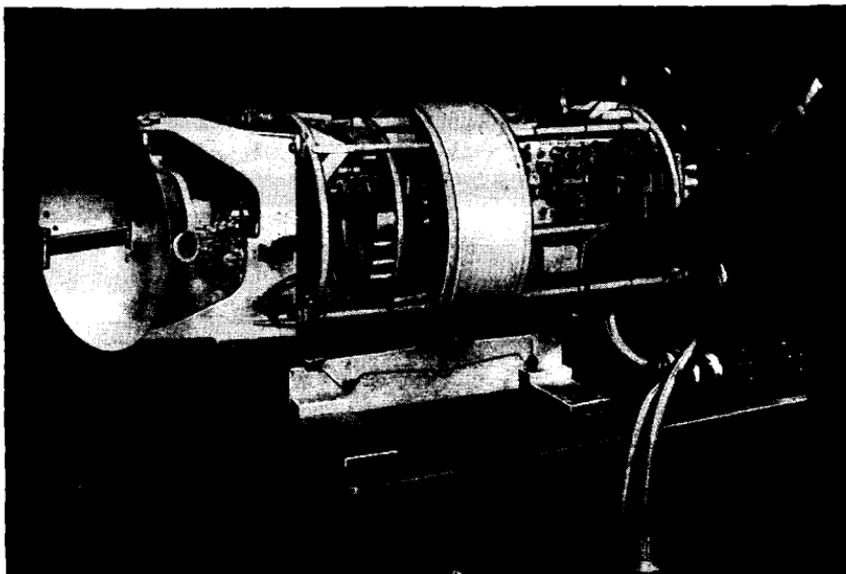


FIG. 6-32.—Units of AN/APS-4. Covers, radome, and fairing of main unit are removed. In use, main unit is suspended from bracket shown at bottom in this view.

Homing on Aircraft.—During the war, defensive nightfighters equipped with radar were developed as a reply to night bombing attacks. The homing problem faced by the Aircraft Interception, AI, radar used for this purpose is more difficult than that presented in the sea-search

application, since the target speed is much higher (and much more nearly that of the intercepting aircraft), and since homing must be carried out in two dimensions instead of one.

An early AI equipment was designed by the British as an extension to AI of the design principles of ASV Mark II. It was designated AI Mark IV by the British; a similar system made for the U.S. Army by an American manufacturer was called SCR-540. In this set, which operated near 200 Mc/sec, a single transmitting antenna sent a broad lobe in the forward direction from the aircraft. Two pairs of receiving antennas were provided; one pair produced overlapping lobes like the "homing" lobes of the ASV Mark II, to give an indication (also identical with that of the ASV Mark II) of the homing error in azimuth. The other pair of lobes was used to measure homing error in elevation angle; they overlapped exactly when viewed from above the airplane, but one was displaced slightly above the center line of the airplane and the other slightly below the center line. Elevation homing error was shown on a second indicator exactly like that of the ASV Mark II, turned through 90° so that the range sweep occurred horizontally from left to right. Signals from the upper lobe displaced the trace upward; those from the lower lobe displaced it downward. The receiver was rapidly switched to each of the four receiving antennas in turn, and the display correspondingly switched to the proper deflection plate of the appropriate indicator tube. Range could be read either on the azimuth tube or on the elevation tube. The arrangement of the receiving antenna patterns and of the indications is shown in Fig. 6-33.

The principal operational limitation of this equipment was a result of the fact that very broad beams were produced by the single-dipole-and-reflector antennas used. These broad beams gave strong reflections from the ground beneath the aircraft, restricting the maximum range at which aircraft echoes could be seen to less than the altitude at which the AI-equipped fighter was flying. In the case of ASV operating on 200 Mc/sec, similar returns from the sea were experienced, but the returns from ship targets were so strong that signals were sought and tracked at ranges *beyond* the sea return. Aircraft echoes were many times weaker, so that aircraft could be seen only at ranges shorter than the ground return.

It was clear that the best hope of escaping this limitation was to make use of a sharper beam in AI equipment, and since the maximum antenna size was limited by the necessity of aircraft installation, considerable effort was exerted to develop a microwave AI. The American equipment which resulted is the SCR-720. In this equipment, a 29-in. paraboloid reflector rotates continuously in azimuth and is slowly tilted in elevation. The helical scan thus produced covers an elevation interval of 25° . The

9.1-cm wavelength used gives a beamwidth of 10° , which is sharp enough to allow a target on the same level as that of the AI-equipped aircraft to be followed out to several times the altitude common to the two aircraft without undue difficulty from ground return.

Two types of indicator are provided. The radar operator has a B-scope which shows targets on a range-azimuth display; this is provided with a movable marker that enables the operator to feed to the other indicator—a C-scope showing azimuth and elevation—only the video signals in a range interval centered on the range at which the marker is

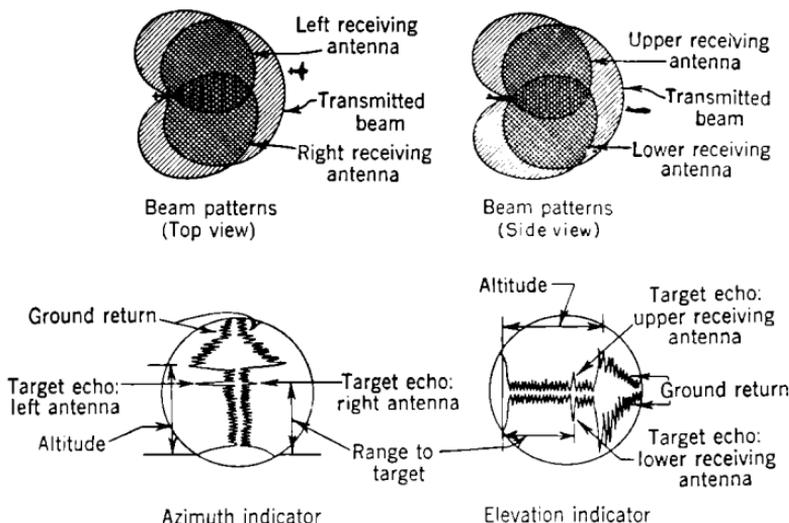


Fig. 6-33.—Beam patterns and indication of British AI Mark IV (SCR-540).

set. The unfavorable signal-to-noise characteristics of the C-scope are thus somewhat mitigated, and the C-scope, which is repeated for the benefit of the pilot, displays only the desired target. A range meter operated by the position of the radar operator's movable marker is also provided to enable the pilot to gauge the progress of the interception.

The SCR-720 has been used with success both by the RAF and the AAF.

An airborne automatic tracking radar—essentially a light version of the SCR-584 described in the next section—was designed for AI use and manufactured in small quantity. Though it was operationally very satisfactory and was highly regarded by nightfighter pilots who flew with it, it was never used in the war, principally because of the low priority put by the AAF on night air defense after 1942. This set was designated SCR-702, AN/APG-1, and (made by another manufacturer) AN/APG-2.

The difficult problem of providing AI equipment for single-engine, single-seat, carrier-based Navy nightfighters was solved in a marginally

satisfactory way by the AIA equipment and its redesigned successor, the AN/APS-6. This 3-cm equipment has an 18-in. reflector which is housed in a wing nacelle in the same fashion as that of the AN/APS-3. As this reflector is rotated at 1200 rpm about an axis passing through its feed and parallel to the line of flight of the aircraft, the paraboloid is slowly tilted from a position in which its axis coincides with the axis of rotation to a position in which there is a 60° angle between these axes. Because of the pattern thus traced out in space by the beam (Fig. 6-12*b*), this is called a spiral scan; a complete scan takes 4 sec. The display used for search is the "double dot" indicator described in Sec. 6-7 (Fig. 6-12*a*).

In the last stages of a homing operation, the assumption is made that the pilot can safely concentrate all his attention on his target. Turning a switch then changes both the scan and the indication. The tilt of the antenna with respect to the axis of its rotation is fixed at 3° , resulting in a conical scan. Two voltages, respectively linear with the azimuth error and with the elevation error in pointing the axis of the conical scan toward the target, are applied to a spot error indicator of the type described in Sec. 6-8. The pilot completes the interception by keeping the spot centered on his scope, and closing the range until the wings of the spot have grown to the desired size.

6-14. Precision Tracking of a Single Target.¹—During the war, considerable effort was expended in the design of radar for highly precise position-finding on one target at a time. This tracking was ordinarily intended to permit the control of fire against such a target, but as the war drew to a close, it was also used for detailed control of the maneuvers of the target by radio instructions from a controller at a ground station. This was done to permit blind bombing by fight-bombers too small to carry a radar set and its operator, or to enable an aircraft not equipped with radar to make a blind landing approach on instructions from a controller at a ground radar. The latter use has, and will probably continue to have, a considerable peacetime importance.

Lobe-switching and Pip-matching.—The first radar set intended for precision tracking of a single target was the SCR-268, a laboratory prototype of which was formally demonstrated to the Secretary of War in May 1937. The problem of accurately tracking a moving target with an antenna that can be elevated and trained is similar to the problem of homing with the help of fixed antennas on a movable vehicle, and the same techniques are useful. Lobe switching, which has been described in the last Section, is employed by the SCR-268 for angle-error determination; in fact, the SCR-268 was the earliest production radar to use this technique. The general appearance of the set is shown in Fig. 6-34.

¹ By E. C. Pollard and L. N. Ridenour.

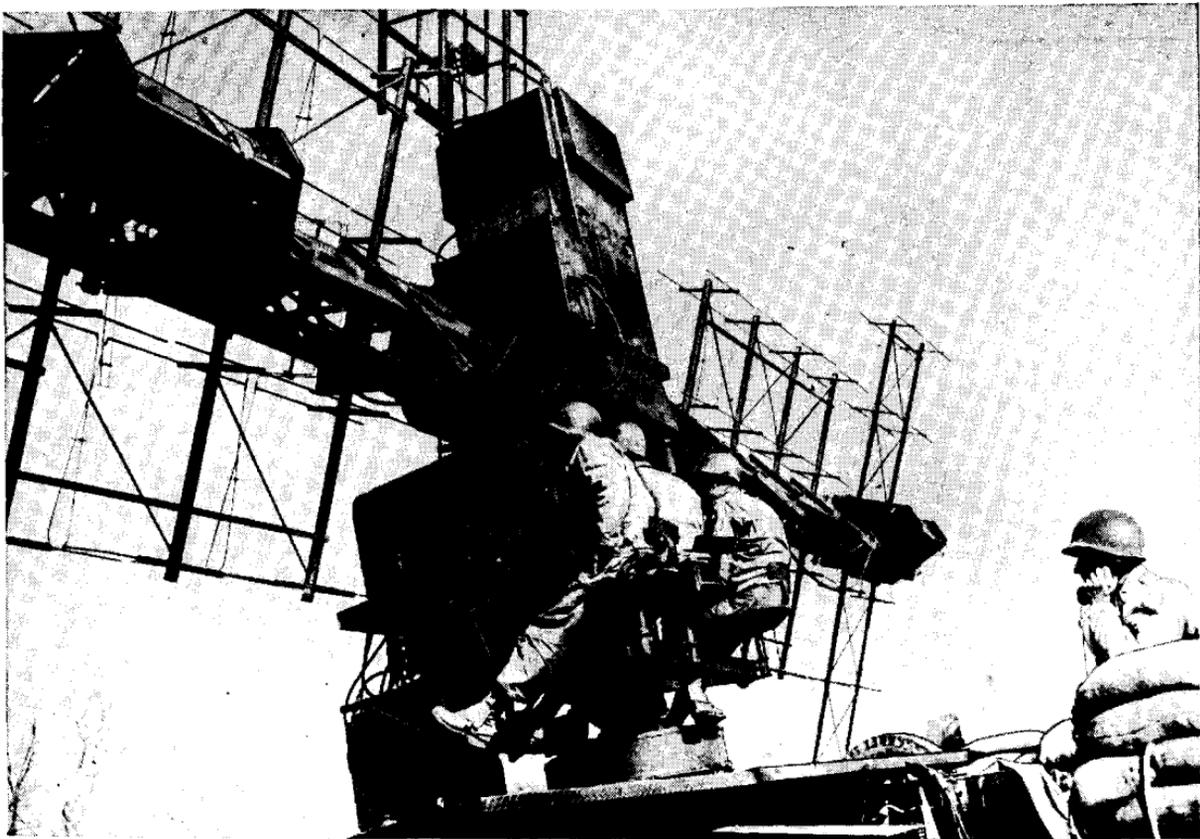


FIG. 6-34.—SCR-268 in operation.

At the operating frequency of 205 Mc/sec, a 4 by 4 array of dipoles with reflectors gives a beam about 21° wide in azimuth and in elevation. Return echoes are received on two separate antenna arrays, each with its own receiver. One array, six dipoles wide by four high, gives information to an azimuth scope; the other, two dipoles wide and six high, provides signals to an elevation scope. The receiving arrays have two separate feeds arranged to produce different phase relations between the elements of the array. Thus, by switching the azimuth receiver from one feed to another, the beam pattern of the receiving antenna can be switched from one to the other of two overlapping lobes equally displaced in azimuth on opposite sides of the normal to the antenna array. Similar switching arrangements are provided for the elevation antenna and its receiving channel.

The antennas are mounted on a single cross arm, which rotates in azimuth with a central pedestal enclosing the radar circuits and can be turned about its own axis to elevate the arrays. Three operators, each with a scope, ride with the cross arm as it turns. A range operator is provided with an A-scope having a movable marker which he keeps on the signal being tracked. His rotation of the range handwheel feeds range information to a "height converter," a computing mechanism employing a three-dimensional cam to combine slant range and elevation angle in such a way that continuous target altitude information is produced. Rotation of the range handwheel also brings the signal being tracked into the center of the azimuth and elevation display tubes. Each of these tubes is manned by an operator provided with a handwheel which moves the antenna in the appropriate angular coordinate. The display is type K (Fig. 6-4). It is the duty of the azimuth operator to keep the two signal "pips" in the center of his scope matched in height by turning his handwheel; the elevation operator has a similar task.

The maximum design range of the set is 40,000 yards, and the repetition rate is 4098 pps. With a power output greater than 50 kw and a pulse width of 7 to 15 μ sec, the set is very conservatively designed, and can track targets much beyond its rated range. Angular accuracies attainable in practice with this equipment, about $\pm 1^\circ$, are too poor to permit good blind antiaircraft fire, and because of ground reflections the equipment suffers from severe errors in reading elevation at angles smaller than about 10° . However, this versatile, rugged, and readily mobile equipment was available in quantity early in the war, and served many useful roles, being used for fire control, short-range search and warning, and searchlight control.

Conical Scan and Error Indication.—In a situation where not more than one target at a time is expected in the radar field of view, the equipment necessary to track a target can be considerably simplified. Such a case

arose in the past war in connection with the defense of bombing aircraft against nightfighters. The limitations of AI radar were such that a coordinated attack on a single bomber by more than one nightfighter was not possible, and the only practicable attack was the stern chase. Under

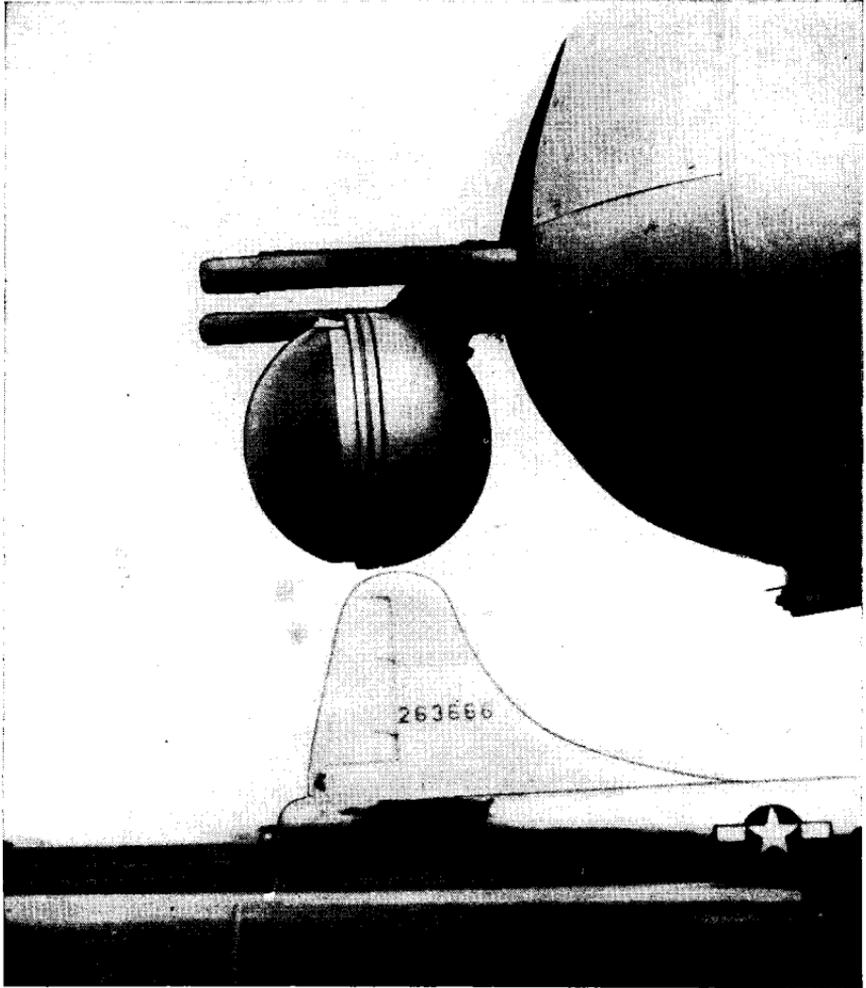


FIG. 6-35.—Scanner of AN/APG-15 in tail turret of B-29.

these circumstances, the bomber can be protected by a simple, short-range radar equipment mounted in the tail turret. The B-29 was equipped with a radar of this sort, the AN/APG-15. Figure 6-35 shows the antenna installation.

The equipment operated at 12 cm, and had a 12-in. dish giving a beamwidth of about 25°. This beamwidth was sufficient to give warn-

ing coverage of the tail cone without the necessity for scanning. The beam was conically scanned by rotating the reflector about an axis which made an angle of 2.5° with its own. Commutators were mounted on the shaft of the antenna rotation mechanism so that signals from the upper half of the scan could be integrated and compared with those from the lower half; similarly, signals from the left half of the scan were compared with those from the right. The resulting error voltages were used to deflect the spot of a CRT used as an error indicator. An automatic circuit searched in range in the absence of a target, and locked on and tracked a target when an echo was encountered within the 2000-yd coverage of the set. This range circuit was used to brighten the spot only when a target was in range, in order to distinguish between correct pointing and the absence of a target. It also caused the target spot to grow horizontal "wings" whose spread increased as the range grew shorter, in order to provide an indication of range to the gunner (see Fig. 6-13).

The axis of the conical scan was aligned with the guns so that only point-blank fire with no lead was possible. This was done because the lead required in countering the usual nightfighter approach is negligible. The indicator was so arranged that its display was projected into the reflector sight, enabling the gunner to use the same technique for either optical or radar tracking. This equipment became available so late in the war that it had substantially no operational use, but its proving-ground tests indicated that its performance would have been quite satisfactory. The attainable angular accuracy was about $\pm 0.5^\circ$.

The AN/APG-15 was unusual in that a triode, the "lighthouse" tube, was used as a pulsed oscillator. At the low pulse power used (about 500 watts), this enabled transmitter, receiver, and the necessary rectifier-filter power supplies to be housed in one compact unit, shown in Fig. 6-36.

Automatic Angle Tracking.—Shortcomings of the SCR-268 as a fire-control position finder arose mainly from the use of a relatively long wavelength which resulted in broad beams from antennas of any reasonable size. When microwave radar became practicable with the development of the cavity magnetron, one of the immediate applications of the new technique was an anti-aircraft position finder. The most widely used and generally successful of the resulting equipments was the SCR-584.

Before beginning precision tracking of a single target, a radar that is to serve as the only equipment of an anti-aircraft battery must execute a general search in order to locate targets that are to be tracked and engaged. No auxiliary long-wave radar search equipment was provided with the SCR-584. Instead, the 6-ft antenna reflector was helically scanned, and targets presented on a PPI whose maximum range was 60,000 yd. A range of elevation angles up to 10° was covered by the

helical scan. Alternatively, the antenna could be set to rotate uniformly in azimuth, the elevation being manually controlled, or both azimuth and elevation could be manually controlled. The performance at 10-cm wavelength was more than adequate to give the 60,000-yd range desired on single aircraft, and the convenience of having a common radar perform both the search and the tracking functions is very great. Figure

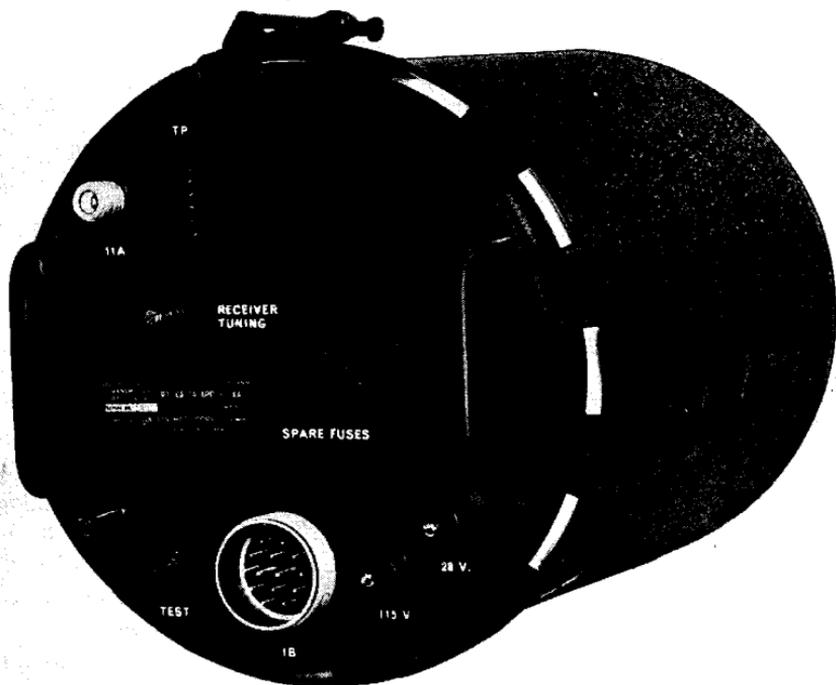


FIG. 6-36.—LHTR transmitter-receiver-power-supply unit used in AN/APG-15 and other sets.

6-37 shows an “x-ray” view of the SCR-584 with the antenna elevated into operating position. For traveling, it is lowered within the trailer.

A conical scan is executed by the antenna of the SCR-584, the dipole feed of the 6-ft paraboloid rotating rapidly about an axis which is that of its mechanical, though not its electrical, symmetry. For precision tracking, the azimuth and elevation error signals derived from this conical scan are used, not to give an indication of pointing error, but actually to drive servomechanisms which position the antenna mount. An error signal in either coordinate will cause the antenna to move in the direction necessary to reduce the error. This so-called “automatic angle tracking” was

embodied in the design because its precision of following is, unlike the accuracy of manual tracking, not subject to human errors arising from combat stresses and fatigue, *and precision.*

In order to ensure that the error signals are measured with respect to the desired radar echo only, and also to provide a continuous measurement of range, two type J displays are presented to the range operator of the radar. These are circular range sweeps (Sec. 6-4) with radial signal deflection. The coarse range scope shows all ranges out to a maximum of 32,000 yd. and the fine scope shows a magnified trace of a 2000-yd

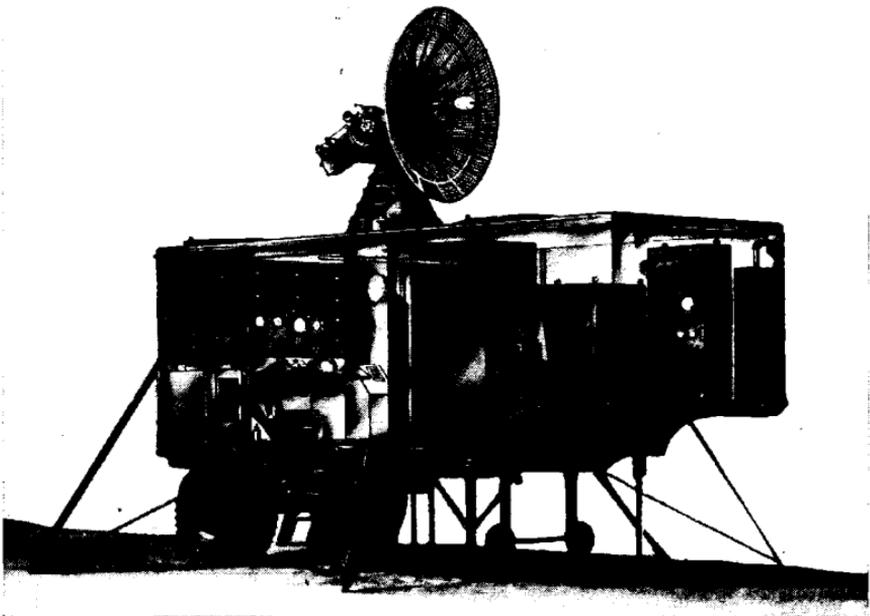


FIG. 6-37.—SCR-584 in operating position.

interval which can be chosen anywhere within the 32,000 yd. Markers geared together at 1 to 16 and driven by the range handwheel are in front of these tubes. When the marker on the 32,000-yd tube is put over the signal echo to be tracked, the 2000-yd tube displays a range interval 1000 yd on either side of the target range. The marker on the 2000-yd tube can then be set exactly on the leading edge of the target signal. When this adjustment is made the marker is kept in register with the desired signal, and continuous target-range information is sent to the computer. Further, only those signals which come in a range interval of about 50 yd immediately following the range for which the marker is set are sent to the input of the angle-error measuring circuits.

Circuits have been devised and tested which provide automatic tracking in range, as well as in angle, but it appears that the judgment of the range operator is of definite value in the operation of this equipment. The operator is especially useful in following a desired target in the presence of other near-by signals, which may be due to ground objects, other aircraft, or "window" dropped to make radar tracking difficult (Sec. 3-10).

The proving-ground accuracy of this set is about ± 25 yd in range at all ranges, and about ± 1 mil¹ or better in angle. With good radar maintenance, this accuracy is achieved under field conditions. This was one of the most widely used, versatile, and generally successful of all allied wartime radar sets.

6-15. Precision Tracking During Rapid Scan.²—At the end of Sec. 6-12 we considered the design of systems for general air surveillance and control which offered fairly accurate information, renewed every few seconds, on all the positional coordinates of all targets in the radar field of view. This was claimed to be the operational ideal; and so it is, for a radar set whose purpose is to permit the general control of aircraft. For fire control, or for the close control of aircraft which is needed if a ground controller is to coach a pilot through a blind landing approach, much more precise and much more nearly continuous positional information than that supplied by the V-beam is demanded.

The requirement of greater precision implies that narrower radar beams must be used, and the requirement of more frequent information implies an increase in the speed of scanning. These design changes, as we have seen, seriously restrict the volume of space which can be covered by the resulting radar set. However, the goal of maintaining at least a partial situation picture by scanning, while simultaneously providing highly accurate positional information on a particular target being tracked, is so attractive that several equipments have been designed to attain it. The requirement of high scanning speed has led the designers of all such sets to use "electrical" scanners (Chap. 9).

The design limitation on the volume of space that can be covered frequently enough to be useful is least troublesome in the case of a radar equipment designed to deal with surface targets. In this case, one is interested only in coverage of a plane, not in searching the volume of space that may contain aircraft. The most advanced equipment of the rapid-scan precision-tracking type that was in field use at the end of World War II was the AN/TPG-1, which had been designed for the control of shore-battery fire against ships. For the purpose of fire control against ships, a rapid-scan precision radar is most desirable, because the

¹ A mil is a thousandth of a radian; thus 17.4 mils = 1°.

² By L. N. Ridenour.

splashes of shells that miss the target can be seen on the scope of the radar. Accurate "spotting-in" of fire both in range and in deflection is possible with this type of set.

The antenna and scanning principle of the AN/TPG-1 (which was also designated, in various modifications, as SCR-598, AN/MPG-1, AN/FPG-1) are described in Sec. 9-14, and a perspective drawing of the antenna is shown as Fig. 9-25. The radar beam produced was about 0.55° wide in azimuth; this narrow beam, coupled with the $\frac{1}{4}$ μ sec pulse length used, gave the set high resolution both in range and in azimuth. A scope photograph showing the ability of the equipment to resolve closely-spaced targets is shown as Fig. 6-38.

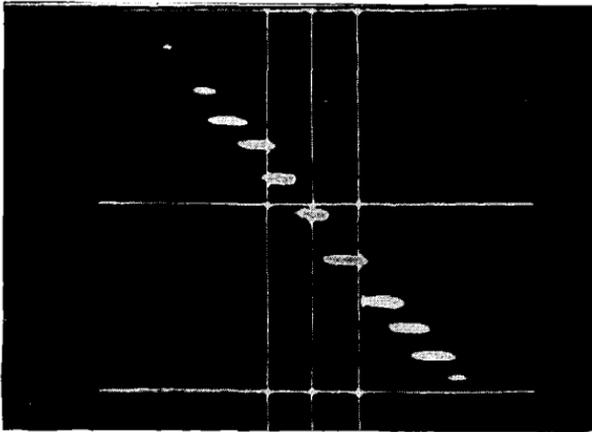


FIG. 6-38.—Row of small ships (LCI's in formation off Honolulu) shown on micro-B of AN/TPG-1. Ships are less than 300 yd apart. (Reprinted from *Electronics*, December 1945.)

The AN/TPG-1 rapid scan covered a sector 10° wide centered on the target of interest. The antenna could be mechanically rotated as a whole to swing the center of this sector to any desired azimuth position, or could be rotated continuously in azimuth during search. The set incorporated a PPI which was useful when the antenna was rotating mechanically for search. Its other indicators were a conventional type B presentation which presented the 10° sector covered by the rapid scan, and a micro-B indicator. The latter is normalized at 400 yd/in. At the ranges important in the operational use of this set, the distortion of the micro-B display is very small. Electronic range and azimuth markers are provided on the micro-B to aid in tracking. The azimuth error obtained in trials of this set was less than 0.05 degrees.

The AN/MPN-1 equipment, often called GCA (for ground control of approach), is the only radar designed during the war to provide accurate

tracking on *aircraft* targets by means of rapid scanning. The inherent difficulty of the problem is displayed by the fact that the GCA really consists of three separate radar equipments. One is a PPI set which presents a general picture of the air traffic situation; the other two, which comprise the precision system, have "beavertail" beams narrow respectively in azimuth and in elevation. Each of these is rapidly scanned in its narrow dimension, and separate indicators of the "stretched PPI" type present the two angular coordinates. The set is fully described in Sec. 8-13 of Vol. 2 of this series.

CHAPTER 7

THE EMPLOYMENT OF RADAR DATA

BY B. V. BOWDEN, L. J. HAWORTH, L. N. RIDENOUR,
AND C. L. ZIMMERMANN¹

7-1. The Signal and Its Use.—The presentation of echo signals on an indicator by no means completes the problem of designing an operationally useful radar system. It is necessary that action of some sort be taken on the basis of the information afforded by the radar. To enable this action to be taken promptly, intelligently, and correctly, an organization must be created. This organization begins with the radar indication and extends to the execution of commands that arise from the situation as displayed by the radar. The nature of these commands and the nature of the organization that assimilates the radar data and gives the commands differ widely from one functional use of radar to another.

The organization that employs radar data may be almost entirely mechanized and automatic, as is, for example, the Army system for radar antiaircraft fire control. In this system, the azimuth, elevation, and range of a target are transmitted directly and continuously from the radar to an electrical computer. The computer solves the fire-control problem, determines the future position of the target, and transmits azimuth, elevation, and range (fuze time) information to the guns of the battery. The guns are positioned automatically by means of servomechanisms, and a fuze-setting mechanism is automatically adjusted to cut fuzes to the time-setting indicated by the computer.² The duties of the gun crew are the purely mechanical ones of supplying ammunition and loading.

The Army system of antiaircraft fire control just sketched was outstandingly successful in the past war. This success can be viewed as demonstrating the principle that, when a complicated task must be carried out quickly and accurately under trying conditions, extreme mechanization is well worth while, if it removes from human operators the necessity for employing judgment and for performing complicated operations. Such mechanization substitutes for the skill of operators the design, manufacturing, and maintenance skills that are necessary to

¹ Sections 7.1, 7.2, 7.4, 7.7, and 7.8 by L. N. Ridenour, Sec. 7.3 by L. J. Haworth, Sec. 7.5 by B. V. Bowden, and Sec. 7.6 by C. L. Zimmermann.

² The use of proximity fuzes eliminates this step.

supply and maintain the equipment required by mechanization. The advantage of making this substitution is two-fold: the latter skills can be applied under favorable conditions, while operator skill is required in the field on the occasion of each crisis; further, there are usually many more operators of a given equipment than there are designers, production engineers, and maintenance men, combined. The total skill required is thus more economically used under more favorable conditions if mechanization is employed.

Most problems involving radar are far simpler than that of anti-aircraft fire control. A simple set may be used on a ship, for example, for navigation and collision avoidance. Under these circumstances, all that is required is a good PPI display on the bridge, and perhaps an alarm that gives a signal if a radar target approaches within a mile. Very good navigation can be performed by comparing the PPI with a chart and taking the range and bearing of sufficient identifiable points shown on the radar to determine the location of the ship. The organization required beyond the radar indicator consists in this case of very little more than the ship's navigator.

The problem of creating the most efficient organization for the use of radar data, in each functional situation involving radar, is a very complicated one. During the war, it lacked the systematic study that its importance and its complexity deserved. Since any treatment of system design would be incomplete without some reference to this important topic, this chapter will deal briefly with some of the devices and some of the methods which have been worked out to translate into commands the decisions taken on the basis of radar information.¹

It is to be emphasized that this subject is at least as difficult and as important as that of the technical design of the radar itself, and far less well understood. We can now build reliable radar equipment whose principal performance limitation arises because the earth is round; the major improvements to be looked for in the use of radar over the next few years will lie, for the most part, not in the category of technical radar design, but in the field of fitting the entire radar system, including its operational organization, to the detailed needs of the use and the user.

EXTERNAL AIDS TO RADAR USE

7-2. Aids to Individual Navigation.—The most frequent and important use made of radar is as an aid to air and sea navigation. The resemblance of the PPI display of a modern microwave radar set to a chart is striking, and suggested very early that navigation would be assisted by a device that enabled a map of the proper scale to be super-

¹ The problem of devising an organization for using radar information in navigation is discussed in Vol. 2 of the Series.

posed on the radar indicator. Figure 7-1 is a schematic diagram of one method of accomplishing this. The indicator screen is viewed through an inclined, partially reflecting piece of glass which reflects to the eye of the observer the image of a screen on which a chart is projected.

Special coatings for glass have been developed which give a high reflectivity in the blue region of the spectrum, while transmitting yellow light substantially without loss. If such a coating is used on the inclined mirror, the yellow persistent signals on the PPI tube can be seen with nearly their full intensity, while the chart image can be projected on the mirror in blue, so that it is reflected with little loss in intensity.

If a radar fitted with such a device is being used for the navigation of a moving ship or aircraft, the charts used for projection must be adjustable and easily changed. Methods have been worked out for storing a large amount of map information in a small space by the use of microfilm, and for providing motion of the chart in the projector. The direction and speed of this chart motion can be adjusted by the operator so that the chart and the radar display stay in register as the vessel moves; when this has been accomplished the direction and speed of the ship, as well as its instantaneous position, are known.

Few such chart-projection devices have been built. They are complicated, bulky, and expensive, and the function they perform can rather simply be done by other means.

Dead-reckoning Computers.—The direct comparison of the PPI with an appropriate chart is an unnecessary elegance if it is desired merely to take occasional fixes and to proceed between fixes by the usual methods of dead reckoning. This is, in any event, the only safe means of navigation. When the first microwave PPI radar was installed in aircraft, careless navigators neglected their dead reckoning in the belief that the radar would enable them to determine their position whenever necessary. This procedure resulted in getting lost with such regularity that radar was soon established as an adjunct to dead reckoning, rather than a substitute for it.

To assist in dead reckoning, particularly in aircraft navigation where high speed and frequent course changes make it difficult, various forms

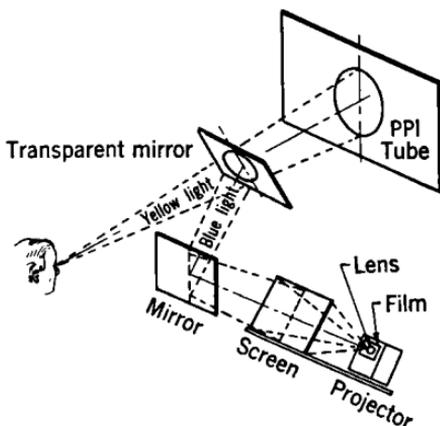


FIG. 7-1.—Schematic diagram of chart projector.

of dead-reckoning computers, or odographs, have been developed. The control box of one such computer is shown in Fig. 7-2. This device, developed in the Division 17 program of the NDRC, receives information from a true-airspeed meter and a repeater compass. These data are resolved into cartesian velocity components along the north-south and east-west directions, and these velocity values are integrated to provide continuous readings of air position with reference to any chosen point of departure.

A mechanism is also provided to indicate ground position by taking account of the effect of wind. The direction and speed of the wind (as determined from double drift readings, meteorological information, comparison of earlier wind settings with the actual course made good, radar fixes, or otherwise) can be set into the device by means of knobs shown in Fig. 7-2. Wind velocity is resolved into components along the north-south and east-west directions, integrated, and the result added to the air position to give the ground position with respect to the point of departure.



FIG. 7-2.—Control box of aircraft odograph or dead-reckoning computer.

Air position and actual ground position both appear on the counter dials shown at the lower left. The output signals of the unit can also be used to actuate a plotting arm which moves over a map table not shown in Fig. 7-2. With this addition, the device draws on a map the actual

ground track made good by the aircraft carrying the instrument. Various map scales can be selected for the plotting table by means of a knob on the control box. The control box can also be provided with a gear train that actuates a mechanism showing latitude and longitude directly, in addition to showing the departure from the last fix. In a carefully made installation, the cumulative error of this instrument amounts to only $\frac{1}{2}$ per cent of the total distance traveled from the point of departure.

Ground-position Indicator, GPI.—Although the device just described aids in performing dead reckoning, and to this extent assists in radar navigation, it is convenient to present the results of the dead reckoning directly on the radar scope. This has been done in a device called the “Ground-position Indicator,” or GPI.

The GPI provides an electronic index for the PPI tube of the radar in the form of the intersection of a circle of constant range and an azimuth

marker. Once this index has been set on a radar echo by means of adjusting knobs provided on the device, the index will move across the face of the tube with the echo, provided that the wind setting has been made properly, regardless of maneuvers of the aircraft.

A simplified schematic diagram of the GPI is shown in Fig. 7-3. To the left of the box marked "Rectangular to polar coordinate resolver," the device is identical with the dead-reckoning computer just described. Since the basic coordinates of the PPI are range and azimuth, it is necessary to convert the cartesian position information to polar coordinates; this is done in the resolver shown. The azimuth resulting from this operation is compared with the instantaneous azimuth of the radar

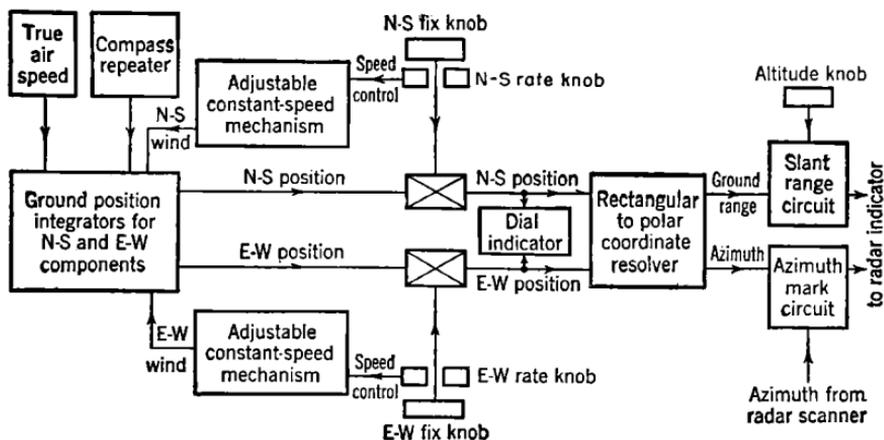


FIG. 7-3.—Simplified schematic diagram of Ground Position Indicator.

scanner. When the two are equal, an azimuth mark is put on the radar indicator electronically.

The range information provided by the resolver is measured horizontally along the ground; it must be corrected for the altitude of the aircraft if, as is usual, slant range is displayed by the radar indicator. This is done by means of the electrical triangle-solver labeled "Slant range circuit," into which the value of aircraft altitude is set by the operator. The value of slant range computed by this device is used to put a range marker on the PPI at the appropriate distance from the origin of the sweep.

Taking a fix with the GPI involves turning the N-S and E-W fix knobs until the intersection of the range and azimuth marks appears on top of an identified target signal. The "fix" dials will then read the ground range of the target from the aircraft, resolved into north-south and east-west components. As the aircraft moves, these dials will continue to read the correct position of the target relative to the aircraft and the

index will continue to stay on the target echo, provided that the correct value of wind has been set into the device by means of the two "wind" knobs shown. Even after the index and the target echo have moved off the face of the radar indicator, the dials will still show the target position relative to the aircraft.

The determination of wind requires that two fix operations like that just described be made on the same echo (which need not correspond to a target whose identity is known). The major correction which will have to be made at the second fix is that arising from the effect of wind. The wind and fix knobs of the GPI are arranged so that they can be gripped together and turned together by the same amount. A proper choice of scale factors will enable the operation of taking the second fix to correct the wind velocity at the same time. If the wind error is to be taken out exactly by this process of double-gripping the wind and fix knobs, account must be taken of the fact that the proper ratio between the scale factors of these two knobs is a function of the time elapsed since the first fix was made. The GPI is arranged with a ratio of scale factors which changes with time in the proper way, so that double-gripping will remove wind errors entirely at the second fix, providing this second fix is made at any

time up to six minutes after the first.

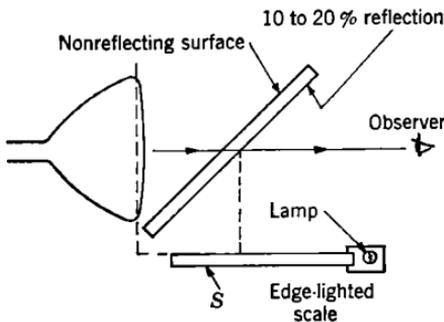


FIG. 7-4.—Method of superimposing an edge-lighted scale on a CRT pattern.

target on successive sweeps, to determine the direction and approximate speed of its motion.

Optical Superposition.—Optical devices of the same general character as the chart projector mentioned in the last section are helpful for these purposes. Two schemes involving the use of a partially reflecting mirror to place a virtual image of a screen in optical superposition with the display are illustrated in Figs. 7-4 and 7-5.

The device of Fig. 7-4 is identical in principle with the chart projector shown as Fig. 7-1. The virtual image formed in this case, however, is that of an edge-lighted screen which can be engraved with a scale or a set of indices. Multiple sets of indices with different scale factors can be

7-3. Aids to Plotting and Control.

The problem of controlling aircraft on the basis of radar information involves careful plotting of the signals seen on an indicator. The controller desires to know the position of a given signal with respect to a map or grid—which must therefore be somehow superposed on the display—and he wishes to record the position of a

conveniently provided by ruling each individual set on a thin transparent sheet. The sheets are then stacked and arranged to be edge-lighted individually; thus only the chosen scale appears. Such multiple scales are used on the indicator of the V-beam height-finding display described in Sec. 6-12, to permit the use of range sweeps of various lengths.

The arrangement of Fig. 7-5 permits the use of only one scale, but affords, as shown, the considerable advantage that a wax pencil can be used for plotting the position of each radar echo of interest, with a minimum of difficulty from parallax. As the pencil point touches the edge-lighted plotting surface, it glows brightly over the small area of contact, and small corrections in the position of the point can be made before a mark is actually applied to the screen. Although the operator must look through the screen itself at its virtual image below, the real and virtual images can be distinguished, providing the scales are not too complex, by their positions and usually by a difference in intensity or appearance.

Display Projection.—An outstanding problem in the use of radar for control is that of transferring radar plots from the face of a small indicator tube to a large board where plotting can be done, other information entered, and a display visible to many people in a large room presented. This problem has been approached in

several different ways. One—that of telling grid coordinates of a target to plotters who enter the target position on a large board—will be described in Secs. 7-5 and 7-6. Another, which seems obvious but presents considerable technical difficulty, is that of direct optical projection of an enlarged real image of the display on an appropriate screen.

The light intensity from a persistent cathode-ray-tube screen of the usual type is too low for satisfactory enlarged projection. Several alternative methods of accomplishing the same result have been used or proposed, including:

1. Splitting the cascade screen (Sec. 13-2) into two parts, the blue component being in the tube and the persistent component on the projection screen. Although this scheme is better than direct projection of a cascade screen, results are far from satisfactory.

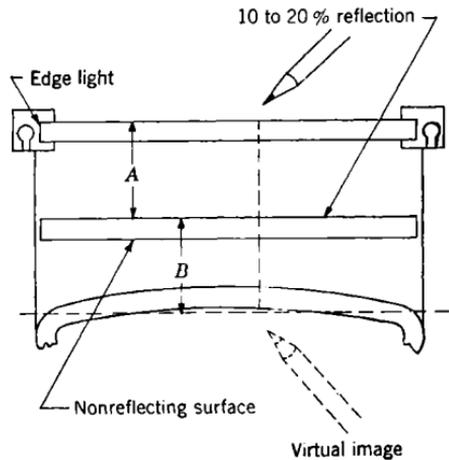


FIG. 7-5.—Method of plotting with the aid of optical superposition.

Further, the screen must be well shielded from stray light if this method is to be used, and this somewhat limits its usefulness.

2. Televising the persistent screen and using the video signals so obtained to modulate a high-intensity short-persistence tube suitable for projection. The results obtained with this rather complicated scheme up to the end of the war were mediocre.
3. Storing the radar picture on an image orthicon or other storage device which can be rapidly scanned electrically to produce television signals. This method, although not well developed at the end of the war, holds great future promise.
4. Use of a dark-trace tube, or "skiatron."
5. Rapid photographic projection.

The last two methods mentioned are the only ones developed fully enough during the war to warrant further description.

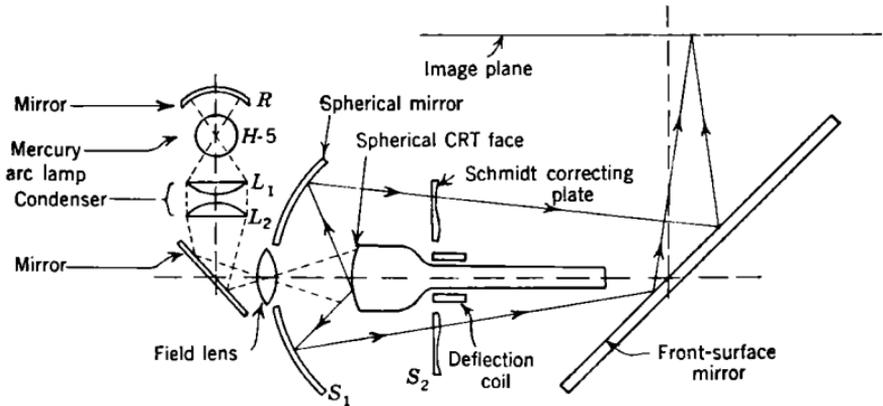


FIG. 7-6.—Schmidt projection system for skiatron.

The Skiatron.—The dark-trace tube, described briefly in Sec. 13-2, has been used as the basis for a reasonably successful projection system both in England and in America, the principal differences being in the optical systems used. In England, an extremely flat-faced tube serves as the source for a wide-aperture lens system; in this country a tube has been designed to fit a Schmidt optical system.

Figure 7-6 illustrates the latter. The cathode-ray tube, whose face is precisely spherical, is mounted with this face concentric with the spherical mirror of the optical system. Light from an intense mercury-arc or tungsten source is concentrated on the tube face by the lens-and-mirror arrangement shown. The geometry is so chosen that the Schmidt correcting plate comes roughly at the neck of the cathode-ray tube. A 45° mirror reflects the light to a horizontal projection screen which forms part of the top of the cabinet housing the equipment.

The display is formed as in any other cathode-ray tube, except that it is necessary to minimize the diameter of the focusing and deflection coils to prevent them from cutting off too much light. Between 10 and 12 kv is applied to the final anode of the tube.

Signals appear as magenta-colored patches against a white background, and they can be viewed in the presence of a reasonably high level of ambient light. This property, together with the large, flat, parallax-free image, makes the display extremely useful for measurement and plotting. However, it has serious inadequacies which arise from shortcomings of the skiatron tube. Contrast is always low, particularly at low duty ratios. Repeated signals have a very objectionable tendency to "burn in," becoming stronger than transient signals and remaining on the tube long after the radar echo that created them has moved away or disappeared. This difficulty can be considerably alleviated by occasionally raising the tube temperature and scanning the screen with a weak electron beam, both of which measures tend to bleach the screen. A fairly intense burned-in pattern can thus be removed in one or two minutes.

In its present technical state, the skiatron is most useful for observing and plotting the courses of ships, from either a shipboard or a coastal station. It is less satisfactory when used to display aircraft signals, because of the poor contrast in their more rapidly moving echoes.

Rapid Photographic Projection.—The inadequacies of other methods of providing a large projection display of a radar indicator led to the development, by the Eastman Kodak Company, of means for the photography of one full 360° sweep of a PPI, rapid photographic processing of the exposed film, and immediate projection of the developed picture. The equipment is shown in Fig. 7-7.

The camera uses 16-mm film, of which 350 ft are required for continuous 24-hr operation at two scans per minute. An instantaneous blue phosphor serves best for photographing. After exposure to a full scan, each frame is processed with metered quantities of developer and fixer, the total processing time being 13.5 sec. A total quantity of less than a gallon of the two solutions is required for 24-hr operation. The projector unit consists of a 300-watt lamp and a simple optical system. The film is cooled by an air blast during projection. The maximum diameter of the projected image is 8 ft, and images of a given point can be repeated on successive frames within a circle of $\frac{1}{4}$ -in. diameter at this magnification.

Plotting means employing this device were just being worked out at the end of the war.¹ It appeared that this was a highly effective and satisfactory method of large-scale direct presentation of a radar display.

¹L. L. Blackmer, "P⁴I (Photographic Projection PPI)," RL Report No. 725, April 26, 1945.

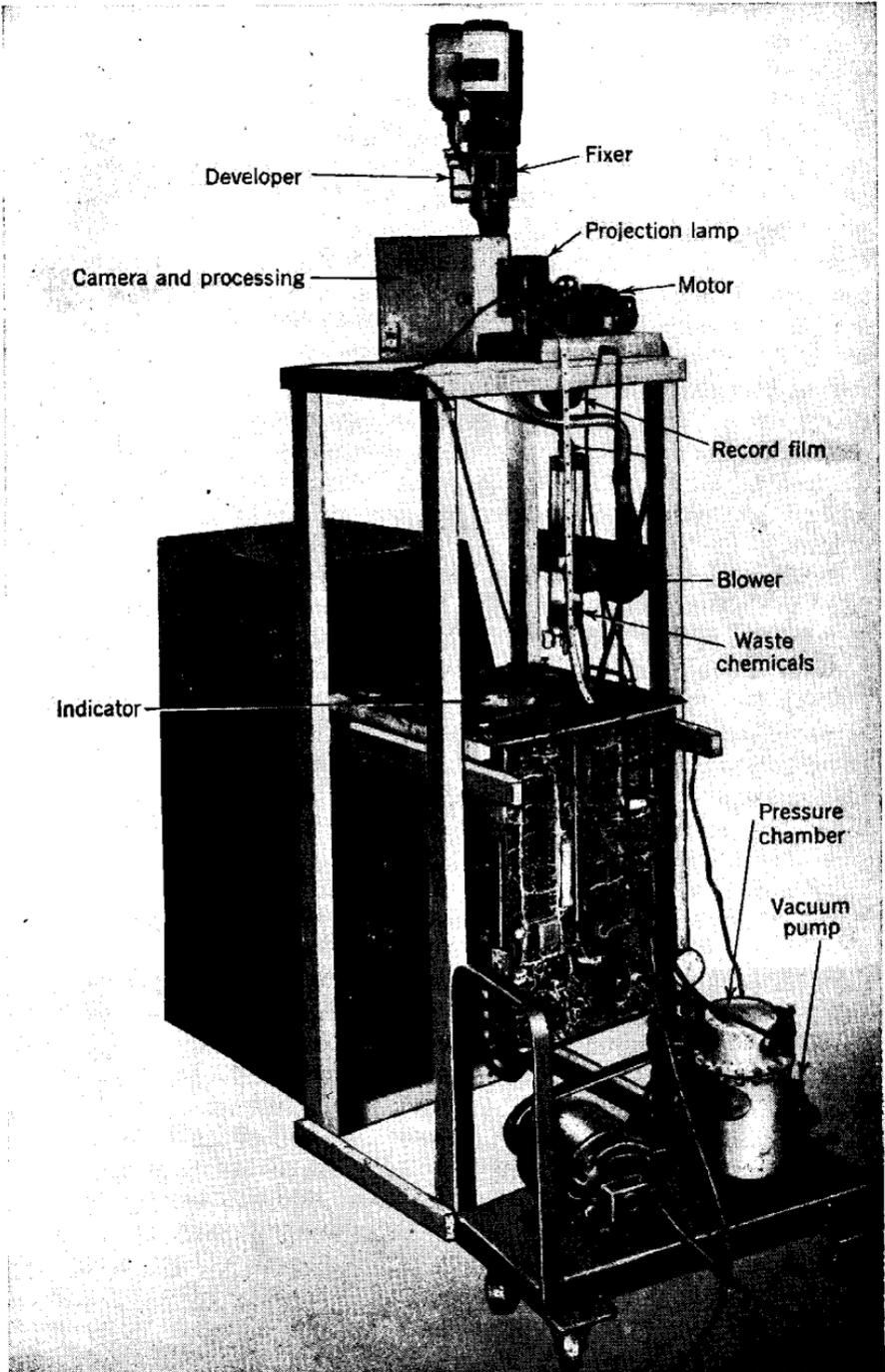


FIG. 7-7.—Rapid photographic PPI projector.

Video Mapping.—None of the methods so far described permits the superposition of an “electronic” map on the display. By this term is meant the addition to the radar signals of signals which, by appropriate intensity-modulation of the indicator, reproduce on the display itself a map, a grid, or any pattern that may be desired. The position of radar echo signals can be compared with this electronic map with complete freedom from parallax, and, in addition, any distortions or imperfections of the display affect the grid and the signals equally and thus produce no reading error. A very satisfactory method of accomplishing this has been given the name of “video mapping.”

The signals required to produce the desired modulation are derived from the device illustrated in Fig. 7-8. The desired map, in the form of

opaque lines on a transparent background or vice versa, is placed immediately in front of an auxiliary PPI of the highest precision obtainable. This PPI rotates in synchronism with the radar antenna and executes its range sweep in the proper time relation with the radar transmitter, but operates at a constant intensity, receiving no video signals. The cathode-ray tube used has a short-persistence blue screen; the intensity is adjusted to a medium or low level. As the moving spot of the auxiliary PPI moves behind the map to be reproduced, a photocell placed two

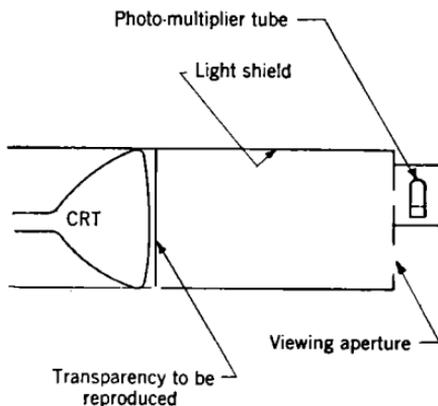


FIG. 7-8.—Video mapping transmitter. The aperture in front of the phototube is small enough so that the parallax between the spot on the CRT screen and the marks on the plotting surface is small.

or three feet from the map receives a signal whenever the rays from the spot to the photocell undergo a change in absorption. Because of the rapid motion of the spot in its range sweep, these signals are in the video-frequency range. The fast screen avoids any appreciable “tailing” of signals.

The signals from the photocell are amplified and mixed with the radar echo signals with a polarity such that on the final display the lines of the map are brighter than the background. In the absence of limiting, gradations of light and shadow can be presented. Figure 7-9 is a reproduction obtained on a second PPI tube of a photographic negative used as a subject for video-mapping transmission. Figure 6.8 was also produced by video mapping, the final display being a B-scope.

This method has the very great advantage that the map is correctly correlated with the radar display regardless of the degree of off-centering,



FIG. 7-9.—Photograph transmitted by video mapping techniques.

or of any deliberate or accidental deformation or distortion. The accuracy is determined entirely by the precision of the scanning PPI and by the geometrical relationship between the PPI and the map. Slight changes in parallax between the center and the edge of the tube can be compensated by using a suitably nonlinear range sweep.

If the radar site is moving, a compensating motion must be applied to the map in order to maintain the proper relationship on the final displays.

This method can be applied to any type of index or marker, and has obvious uses as a substitute for elaborate computing circuits under some circumstances.

7-4. The Relay of Radar Displays.—It is common for the best location of the radar station and the optimum location of a control center to be different. The radar site is chosen from the standpoint of good coverage, freedom from permanent echoes, and the like; the criteria entering into the choice of site for a control center are usually entirely different. Also, a control center should receive supplementary information from other radar installations located elsewhere, even though a single radar equipment may provide the primary data for control of operations. This will enable the coverage of the primary radar to be supplemented by information from neighboring sectors, and will provide coverage of possible "blind spots" of the primary radar.

The telling of plots by telephone land line was the technique first used (Sec. 7-5) for the transmission of radar data from one point to another. Substantial errors and delays are inherent in this procedure. Far more important, when the information to be transmitted has been gathered by a modern long-range, high-definition radar, is the low traffic-handling capacity of the system of telling and plotting. Literally hundreds of targets may show at a given time on the indicator of such a radar as the first one described in Chap. 15; an attempt to convey with adequate accuracy and speed the information provided at a rate of four sweeps per minute is hopeless under such conditions.

Considerations such as these led to the development of means for reproducing radar displays at a distant point by transmitting the radar video signals and appropriate synchronizing information by more or less conventional radio practice. The technical problems of "radar relay," as this is called, have been worked out, and systems for the purpose are discussed in detail in Chap. 17.

Radar relay is mentioned here to emphasize its usefulness as one element in the creation of an organization for the use of radar data. An example of an operational system in which this technique is important is given in Sec. 7-8.

EXAMPLES OF RADAR ORGANIZATIONS

In the following sections are given very brief sketches of a few radar organizations with which some experience has been obtained. The systems described are, for the most part, necessarily those set up in wartime for the purposes of war, and to this extent do not represent useful models for peacetime radar organizations. It will, however, be instructive to consider briefly a few typical systems.

7-5. Radar in the RAF Fighter Command.—Just as the first radar equipment to be used in operations was that installed for the air defense of the British Isles, so the first operating radar organization was that of RAF Fighter Command, the user of this equipment. When the organization was fully developed, the number of people involved in the interpretation of the radar information and in making use of it was comparable to the number required to obtain it.

Organization of the Home Chain.—It has been mentioned in Sec. 6-9 that the performance of CH stations depended very much on the nature of their sites. Moreover, owing to a shortage of trained men and to the difficulties involved in maintaining equipment of this kind in wartime, the antenna systems were often badly installed and inadequately maintained. In consequence, large errors in apparent azimuth were very common; most stations had errors of 10° or 15° , and errors as large as 30° were not unknown. The method of measuring height depended on reflection of the received waves from the ground, and almost invariably the height calibration of a station was different along different azimuths. Each station had to be checked and calibrated both for height and azimuth by an elaborate and difficult series of test flights.

The complicated nature of the corrections necessary on each aircraft plot, and the requirement for speed and accuracy in applying these corrections, led the British to design and install what is perhaps the first device intended for assistance in the use of radar data, as opposed to the gathering and display of that data. This is the celebrated "fruit machine,"¹ a complicated calculating machine made up of standard telephone selector switches and relays. The operator measured the range of the target by setting a marker to the echo on the A-scope. She then turned the goniometer until the echo disappeared, and pressed buttons transmitting range and apparent azimuth to the fruit machine. The machine automatically applied the appropriate correction and deduced the true azimuth, multiplied the target range by the sine and the cosine of this true azimuth, added in the rectangular map coordinates of the station itself, and deduced the coordinates of the target. This

¹ It was named after the English equivalent of the American slot machine used for gambling.

information was displayed in lights to a "teller," who passed the information over a telephone line to Fighter Command Headquarters. At a later stage, the plots were passed by teletype to save time. The whole computation process was completed in a second or so.

The operator then pressed a button which connected the goniometer to the height-finding antenna arrays. Again she turned the goniometer until the signal disappeared, and pressed a button. The height of the target was computed by the fruit machine and passed to headquarters.

Radar plots from all of the Chain stations were sent to Fighter Command in Stanmore, north of London. Because of the shortcomings of the early radar equipment, and the "blind spots" in the coverage of individual sets, it was felt to be essential to combine the information from all radar stations at a central point. This was done at a so-called "filter center."

In a large underground bombproof room was mounted a central table whose surface was a gridded map of England. A crew of plotters, each with a telephone connection to one of the radar stations, stood around the table at positions corresponding to the geographical locations of the Chain stations. As a plotter received plots over the telephone, she put colored disk markers in the grid positions indicated. It usually happened that two or more stations were simultaneously plotting the same aircraft, so that several girls might be putting down counters representing the same formation. Owing to the errors in the system and the variable delays in the plotting process, the interpretation of the piles of disks in terms of aircraft was not easy. Special officers known as filterers stood beside the girls and decided, for example, whether two adjacent tracks were really separate, or represented the same aircraft, erroneously plotted by one of the stations. After he had analyzed the data, the filterer put on the map a little plaque that bore his best estimate of the identity, position, height, speed, and number of aircraft in the formation.

This whole process was observed by filter officers who sat on a balcony overlooking the map. They were informed of the plans for the movements of friendly aircraft. They were able to direct the operations of the whole Chain and to decide which formations should be plotted by each station. The work of the filter officer called for a considerable understanding of the performance of CH stations, and an appreciation of the good and bad qualities of each individual installation.

Established tracks on the filter-room board were telephoned by a teller to a second plotting board maintained in an "operations room." Here the filtered radar tracks were combined with plots of the Observer Corps (visual airplane spotters), and with fixes made by means of radio direction finders. In this main control room, the aim was to display the whole of the picture of the air war; all operations were generally directed

from this point. The detailed control of fighter interceptors was carried out from group and sector headquarters, where plots were maintained on the basis of information repeated from Fighter Command Headquarters.

The operation of this system was much more difficult than might be inferred from this brief description. If the CH stations had been perfect, it would still have been difficult to interpret some hundreds of plots every minute, all subject to variable delays and to the personal errors of the observers. Quite trivial difficulties proved surprisingly hard to overcome. It was hard to find room for all the plotters around the table. They could not plot fast enough. They might disturb one set of plots when they leaned over to plot another aircraft. Such rather simple difficulties could be, and often were, the limiting factors on the use that could be made of the radar plots, and an intensive study of all the stages in plotting and filtering was made throughout the early years of the war.

Despite the large number of people necessary to this system, and its prodigal use of telephone land lines for the telling of plots, its operational limitations were severe. Under conditions of moderate aircraft density, a good filter officer with a good organization could filter plots with an accuracy of perhaps 70 per cent. When plots were sparse, the accuracy was excellent, and the only objection to the system was its unavoidable time lag in reporting. Under conditions of high aircraft density, the system broke down, and it was commonplace to cease reporting in certain areas where the density was so high that filtering was impossible.

In spite of these handicaps, very considerable success attended the use of this system in the Battle of Britain and thereafter. A tendency grew to forget that the main reason for organizing the reporting and control system in this centralized way was that a single radar could not be relied on to give a sufficiently complete or accurate picture of events in the air. The great technical improvements of 1943 and early 1944 resulted in long-range, high-definition microwave radar having good coverage if properly sited. This improved equipment made it possible to depend on a single radar installation for a substantially complete picture of the air situation.

Delays and errors unavoidable in a complicated scheme of telling and plotting are largely eliminated in a system that combines the operational organization with the radar equipment. Controllers who give instructions to aircraft are able to work directly from the radar display and therefore have a far more accurate and up-to-date appreciation of the situation than can be obtained from a plot, however well maintained it may be. This was eventually appreciated, and such systems have been put into very successful operation. Such a system is described in the following section.

7-6. The U.S. Tactical Air Commands.—Each U.S. Army that fought in Europe was accompanied by a Tactical Air Command whose mission was that of providing air defense in the forward sector occupied by the Army concerned, of conducting offensive operations against enemy ground troops and installations in the immediate tactical area, and of conducting offensive operations in close cooperation with friendly ground units. The aircraft of a Tactical Air Command, or TAC, were almost entirely fighters and fighter-bombers. Each TAC went into the fighting on the Continent with a radar organization of the general character shown

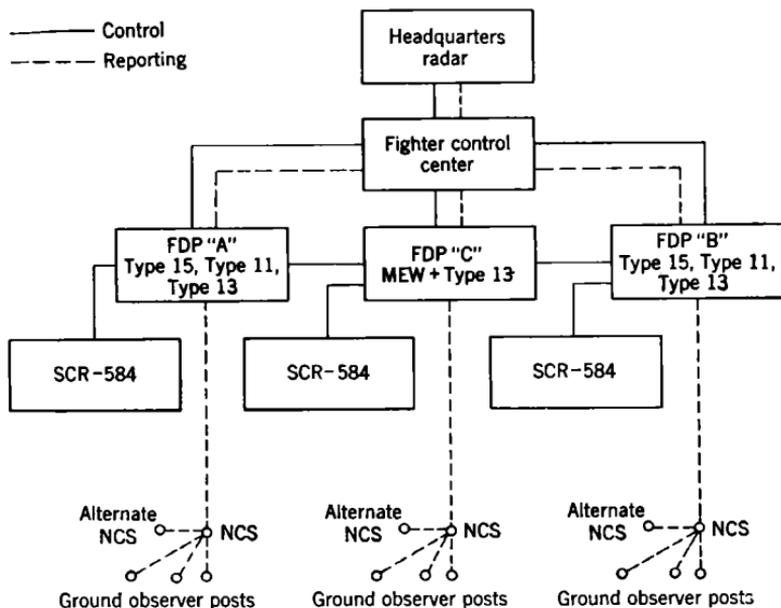


FIG. 7-10.—Fighter control system of IX Tactical Air Command.

in Fig. 7-10, which displays the setup of the 555th Signal Air Warning Battalion attached to IX TAC (which operated in support of the American First Army).

Each of the three Forward Director Posts (FDP's) had operating with it five Ground Observation Posts (GOP's). Men placed far forward observed air activity that might consist of low-flying hostile craft not covered by the radar placed farther back of the front lines. Each GOP reported by radio telephone to a Net Control Station common to the group of five. The NCS then phoned pertinent information to the FDP; an alternate NCS was maintained for use in emergencies.

The radar of the TAC was mainly deployed at the FDP's. These sites were charged with reporting air activity to the Fighter Control Center, which was located at the forward headquarters of the TAC,

usually occupied jointly with Army forward headquarters. In addition, they executed direct control of airborne fighters on certain types of missions.

Let us consider first the reporting function of the FDP, leaving until later the consideration of control directly from the FDP. The following steps must occur before radar information is in the form in which it was used for decision and action:

1. The coordinates of a target signal are read off the display at the radar station.
2. A number is assigned to this particular track, on which continuous plotting will be maintained.
3. An attempt is made to identify the track from the following data:
 - a. Known location of friendly aircraft.
 - b. Response to IFF interrogation (Chap. 8).
 - c. Intelligence information on enemy tracks.
4. A series of symbols is telephoned to the FCC, giving the following information:
 - a. Identification of track as friendly, enemy, or unknown.
 - b. Number assigned to track.
 - c. Location of plot.
 - d. Estimated number of aircraft.
 - e. Height.
 - f. Direction of flight.

At the FCC the following steps take place:

1. The filter officer must determine whether the plot is a continuation of an old track or the beginning of a new one.
2. The track is reidentified, if possible, from the fuller intelligence information available at the FCC.
3. Plots reported from more than one radar station must be recognized as belonging either to the same track or to different tracks.
4. Cards bearing the information listed under (4) of the preceding paragraph are put into a marker called a "Christmas tree."
5. The Christmas tree is moved on to a large plotting board to occupy the position reported for the track.

All of the steps just described require time, and the large plotting board at the FCC is usually between three and five minutes behind the existing air situation. Although this means that the aircraft positions displayed are wrong by some 20 or 30 miles, the delay is usually tolerable for air-warning purposes.

In the "sector control" that is executed by the FCC, aircraft controllers observe the relative positions of tracks on the large plotting board

and issue radiotelephone instructions to friendly fighters in an effort to bring them into favorable positions to counter air attacks. For defensive operations, this sort of control is usually adequate; enemy air attacks will ordinarily be directed at one of a few vital areas where substantial damage can be done. Even with his five-minute-old information, the controller can make a shrewd guess as to the target of such an attack, and can marshal defensive fighters accordingly. For many types of mission however, direct control from the FDP's is required.

Reporting Capabilities of FDP's.—Between Oct. 21 and Nov. 16, 1944, a study was made by the Operations Analysis Section of Ninth Air Force of the reporting work done by the three FDP's of IX TAC. Because of the difference in the amount and character of air activity by day and by night, daily and nightly averages were separately computed for the following data: the average number of separate plots per track reported, the average track duration in minutes, and the average number of plots per height reported. The results are shown in Table 7-1.

TABLE 7-1.—REPORTING BY FDP'S OF IX TAC, 21 OCTOBER TO 16 NOVEMBER 1944

Station	Total tracks	Avg. track duration	Avg. number plots per track	Avg. number plots per height
Day averages				
FDP "C" (Microwave)	1561	11.1 min.	5.4	8.7
FDP "A"	2416	10.0	5.6	8.7
FDP "B"	828	8.4	5.9	5.0
Night averages				
FDP "C" (Microwave)	1249	15.1 min.	7.8	10.4
FDP "A"	1805	10.5	6.3	9.0
FDP "B"	774	8.2	6.6	5.4

The FDP "C" was equipped with one of the preproduction 10-cm radar sets designed for long-range air warning and control (see Chap. 15). The other two FDP's had British long-wave equipment. The longer range of the microwave radar shows in the longer track duration for the FDP equipped with that set.

Differences in the total number of tracks reported arise from a differing density of air traffic in different areas, and from the fact that the fraction of time spent off the air for maintenance and other reasons differed from one FDP to another.

All the plots observed at the FDP are not passed to the FCC. Since control executed by the FDP makes sector control by the FCC outdated for many operations, it is unimportant to report all plots. A study made at one of the FDP's of XIX TAC from Sept. 22 to Oct. 21, 1944, a period during which air activity was limited by bad flying weather, showed that

50 per cent of the total plots obtained were passed on. However, 84 per cent of the tracks were reported, the difference being due to single-plot tracks which were not reported. The average number of plots per track was 6.2 and the average track duration 9.5 min. Hence, on the average, a plot was obtained on each track once every 1.8 min, which is considered to be fairly continuous tracking.

Close Control by the FDP.—At the turn of the year 1940–41, the Luftwaffe abandoned its daytime attacks on England and turned to night bombing, the attacking aircraft operating individually. This led to the development by the British of an elaborate system of nightfighter defense which has been described elsewhere.¹ One of the features of this system was exact control of the nightfighter, prior to his interception of the hostile bomber, by a ground controller at a GCI (for Ground-controlled Interception) radar station. The aim of this control was to put the nightfighter behind, a little below, and on the same course as the hostile aircraft, in which position he was well situated to complete the interception with the help of airborne radar equipment he himself carried.

By 1943, both this GCI technique, involving precise control of single aircraft by a controller located at a radar station, and the daytime sector-control technique described in Sec. 7-5, involving the control of formations of fighter interceptors from a central station at which radar data were assembled and assessed, were firmly established procedures. In that year two things happened which gave rise to the technique of close control as it was practiced by the FDP's of the American TAC's in the Continental phase of the European war. A young American officer, who was working with the RAF as a member of the Electronics Training Group sent to the United Kingdom soon after the United States went to war, conceived the idea of controlling formations of day fighters directly from the information available at his radar station alone. Meanwhile battle experience in Africa showed that an independent radar set could pass to airborne fighters useful information on the disposition of enemy aircraft.

The American officer's novel idea did not gain immediate acceptance since control of several squadrons of fighters in daylight requires techniques quite different from those of the GCI control of a single nightfighter. Finally, the officer was given permission by the late Air Chief Marshal Sir Trafford Leigh-Mallory to control wings of fighters directly from his radar station, entirely independently of the sector-control system.

While the new technique was being worked out and the confidence of pilots and controllers was being developed, 15 RAF planes were lost for

¹ See, for example, "Radar (A Report on Science at War)," Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., Aug. 15, 1945.

10 enemy planes downed. After this initial shakedown period, 90 enemy planes were destroyed for a loss of 13 to the RAF, before enemy jamming of the 1.5-m equipment used in this first work put an end to its usefulness.

The success of this experiment led the British to design a higher-frequency radar specifically for the purpose of close control, and resulted in the modification of the first of the American 10-cm sets, on its arrival in England, to fit it for similar close control. Because of the African experience the control setup was made mobile. Figure 7-11 shows the

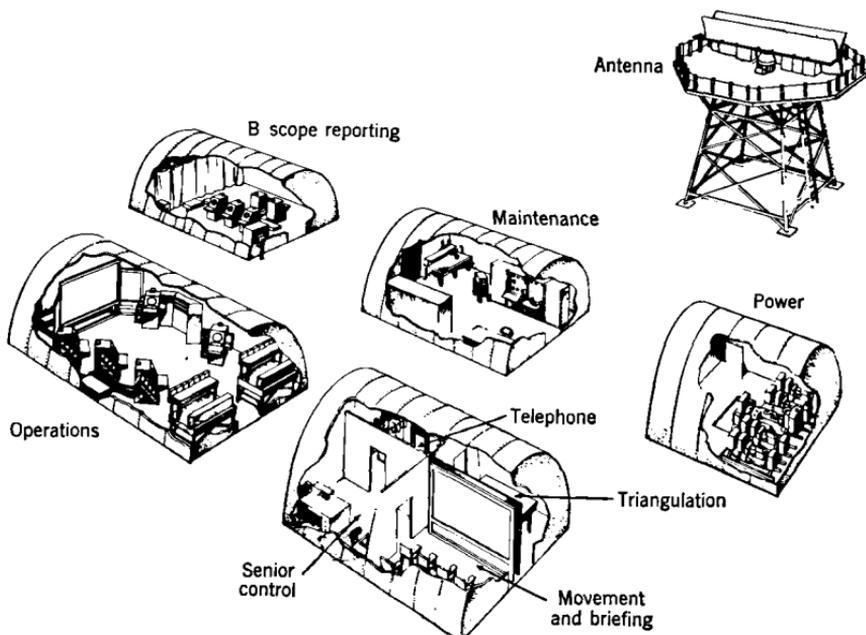


FIG. 7-11.—Layout of radar control center.

arrangement of facilities arrived at for an installation designed for use in close control.

The antenna is shown mounted on a low structural-steel tower. For mobile operation the antenna is mounted on a trailer (see Fig. 9-15). Power is supplied by diesel-electric units housed either in a Jamesway shelter, as shown, or in trailers. The maintenance shelter contains, in addition to workbenches, spares, and test equipment, the power control unit for the set. Another shelter houses a number of B-scopes used for reporting signals to the operations shelter, where they are plotted. The coverage of each B-scope can be chosen to give the best total coverage of the area important to the operations being carried on by the station. Another shelter houses the telephone switchboards, a triangulation table

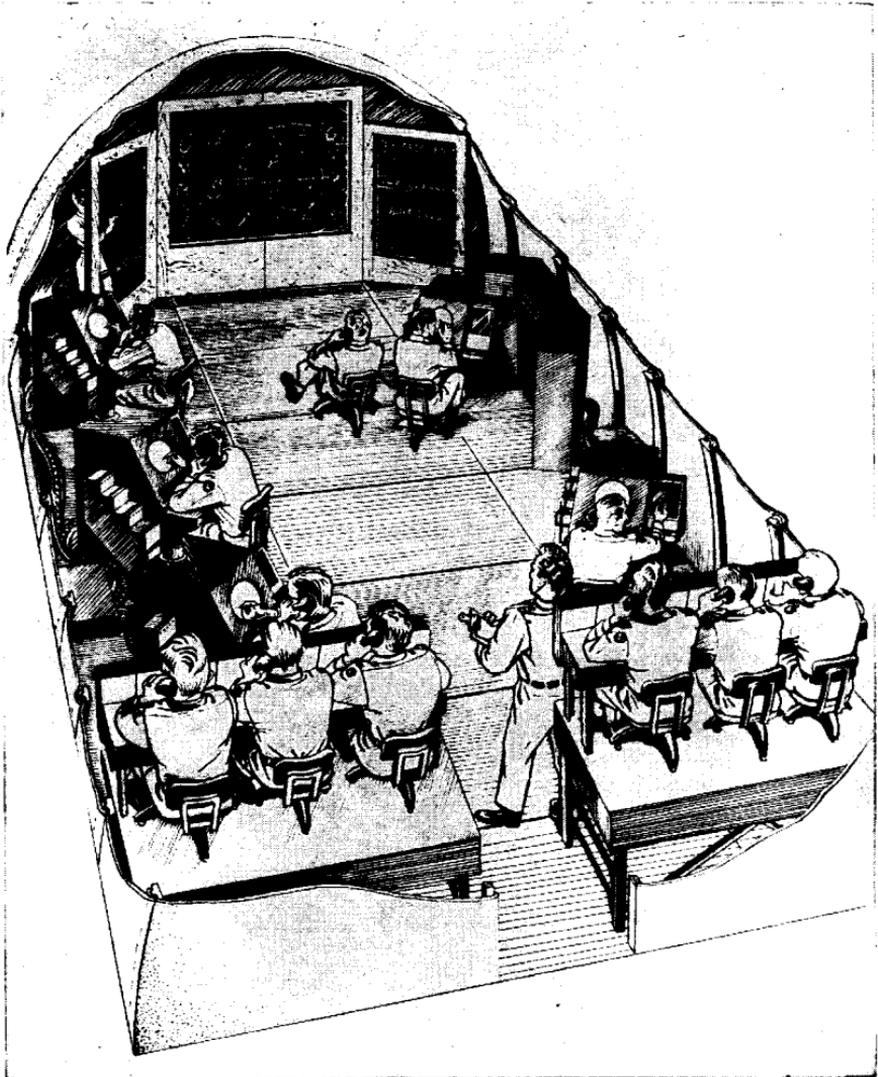


FIG. 7-12.—Interior of operations shelter.

for making radio direction-finder fixes, an office for the senior controller, and a briefing room for controllers.

The operations shelter is the heart of the system. Its interior is shown in more detail in Fig. 7-12. On a dais at the rear of the shelter sit the chief controller, who is responsible for the general operation of the station, and officers concerned with identification of tracks, liaison with antiaircraft artillery, Army, and Navy (as required), and other coordina-

tion functions. Off-center PPI scopes are provided for the chief controller and the deputy controllers, the latter sitting along the sides of the shelter. At the right of the shelter, near the plotting board, sits the supervisor, who is responsible for the technical operation of the radar set. He has a console patchboard by which he can communicate with any of the operators, maintenance men, tellers, or plotters. In front of the plotting board is a filter officer who decides which plots are to be told to the FCC.

The large vertical plotting board is shown in more detail in Fig. 7-13. It is made of transparent plastic and is edge-lighted, so that plots put on

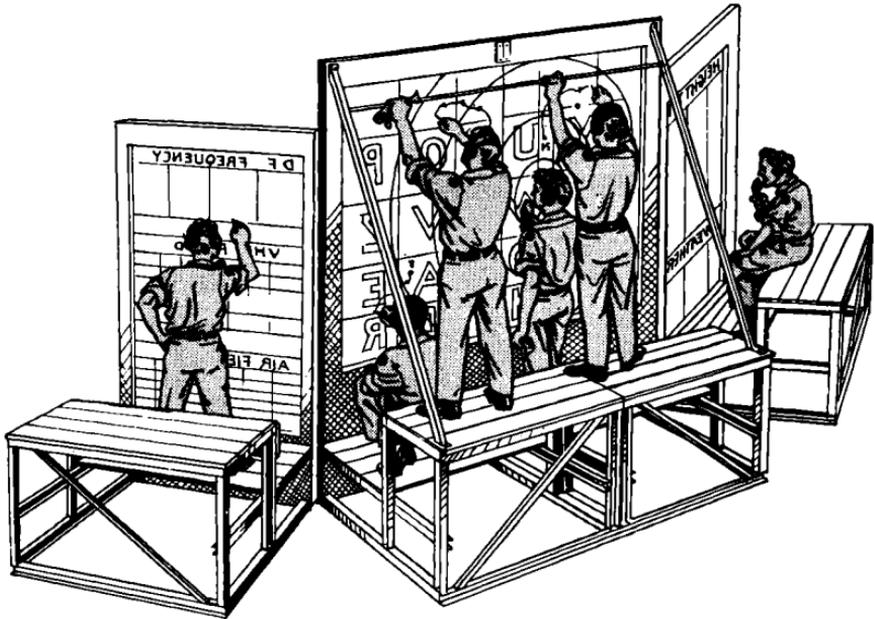


FIG. 7-13.—Rear view of vertical plotting board.

in grease pencil by plotters working behind the board appear brilliantly illuminated when viewed from the front. Each plotter is connected by telephone with a teller who is watching a B-scope in the reporting shelter. Auxiliary boards to the right and left of the main plotting board give pertinent information on such things as the heights reported (by an auxiliary radar height-finder) for various tracks, weather and winds aloft, radio-frequency channels in use for communications and for direction-finding, and the ready status of aircraft at various fields.

The deputy controllers, each of whom is charged with giving instructions to a certain formation as directed by the chief controller, depend on the plotting board for their general picture of the air situation, and work directly from their PPI scopes to determine the instructions to be given



FIG. 7-14.—Controller at off-center PPI.

aircraft under their control. Usually a gridded map of the region covered is at the elbow of each controller for reference, as shown in Fig. 7-14. A grease pencil is used on the face of the controller's PPI to enable him to determine, from sweep to sweep, the direction in which the aircraft of interest are moving.

The setup just described is an idealization of the FDP which never



FIG. 7-15.—FDP of U.S. Tactical Air Command, France, March 1945.

saw service in the European war, but it is based soundly on months of combat experience with improvised and frequently changed systems and devices. An actual FDP using one of the preproduction sets is shown in position in eastern France in Fig. 7-15. The central antenna is that of the main radar; the two that flank it are those of British Type 13 heightfinders (Sec. 6-12). The system is largely mounted in and operated from trucks, and it can all be taken down, loaded on trucks, unloaded at a new site, and erected and put back on the air in a total time of less than 24 hr.

7-7. Close Control with SCR-584.—Figure 7-10 shows, attached to the FDP's, SCR-584 radar sets. This equipment, designed for accurate tracking of a single aircraft at a time, in order to permit antiaircraft fire control, can also be used to provide to a ground controller the information necessary for highly precise control of the aircraft being tracked. It was so used by the TAC's in Europe.

Since the information on target position is normally transmitted to the antiaircraft computer by means of synchros or accurate potentiometers (Sec. 13-4), it is necessary to provide a supplementary method of displaying target position to enable the set to be used for ground control. This was first done by means of the 180° plotting board shown in use in Fig. 7-16. An arm, pivoted beneath the surface of the board on the side nearest the controller, swings in azimuth in accordance with the orientation of the radar antenna. Range information from the set is used to control the position of a range carriage that runs in and out along the azimuth arm. The carriage has an optical system which projects a small spot of light up through the glass surface of the plotting board. Over the surface of the plotting board is stretched a map of the area surrounding the radar set, carefully adjusted so that the point on the map occupied by the radar is directly above the pivot of the moving arm. The scale factor of the range-carriage mechanism is adjusted to correspond with the scale of the map being used, which may be either 1 to 50,000 or 1 to 100,000. When the radar is tracking a target, the spot of light from the range carriage shows up on the map just at the position over which the airplane is flying at that moment.

The controller can thus issue his instructions with full and constant knowledge of the exact position of the aircraft under his control. This is to be contrasted with the situation that obtains in ground control from a scanning radar; in that case the controller has only one "look" at his target per scan—that is, perhaps once in 15 sec. However, he sees all other targets in the air within range of his radar. The controller in the SCR-584 sees the position of his aircraft continuously, and in proper relation to the terrain, but the price paid for this is that he has no knowledge of the whereabouts of other aircraft.

Figure 7-16 also gives an idea of the compact operational organization of a control SCR-584. Range and PPI operators sit at the radar panels, behind the controller. The controller works alone, in full charge of the operation once an aircraft or a flight has been handed over to him by the FDP. The man at the right has communications lines that connect him with the radio truck associated with the SCR-584 station, with a radio direction-finding station, and with the FDP under which the SCR-584 is operating.

The 180° plotting table used with the SCR-584 was admittedly a



FIG. 7-16.—Interior of SCR-584 modified for close control, showing 180° plotting board. Near Cologne, December 1944.

makeshift, and there was later designed a plotting board, shown in Fig. 7-17, which embodied many improvements. It plots in rectangular rather than in polar coordinates, contains "smoothing" circuits that greatly improve the accuracy of the plotted aircraft position, and draws an ink record of the position of the aircraft being tracked, so that the controller has a knowledge of the aircraft course, as well as position, at all times. This device also plots in ground range rather than in slant range.

The range at which aircraft can be tracked by the SCR-584 can be greatly increased, and the accuracy of tracking somewhat improved, if an airborne beacon (Chap. 8) is carried in the aircraft under control. Such a

beacon also facilitates identification of the plane to be controlled at the time of initial pickup.

The operational use made of the SCR-584 as modified for aircraft control was basically any that required precise navigation under restricted visibility. Large aircraft could carry radio and radar navigational aids which often made it possible for them to carry out such missions without assistance, but the fighters and fighter-bombers of the TAC's had neither

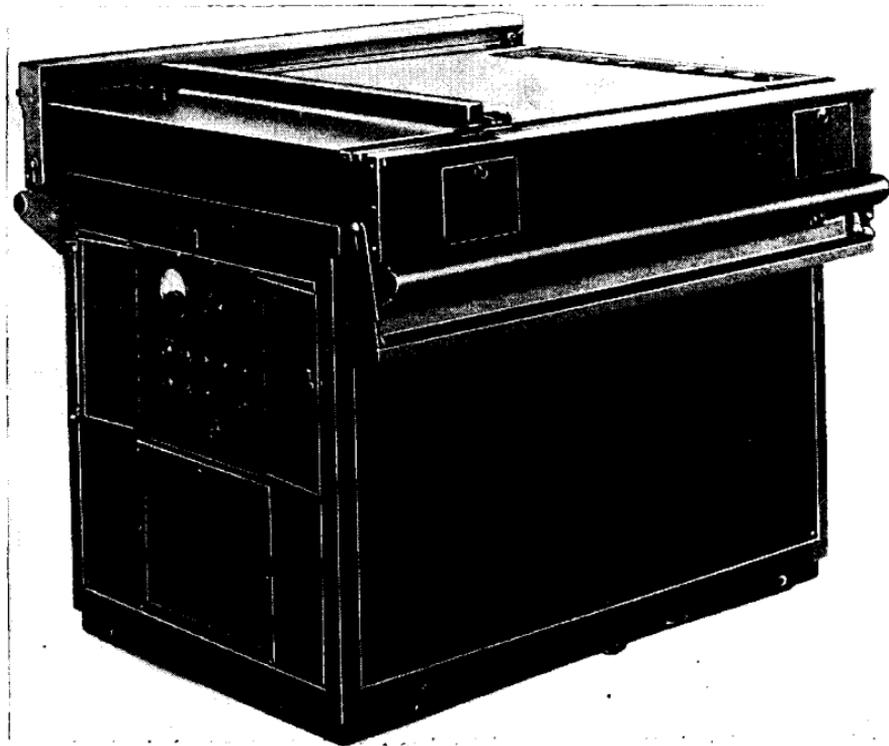


FIG. 7-17.—X-Y plotting board for SCR-584.

space nor operators for such equipment. Their navigational aids had to be external to the aircraft.

A total complement of 8 officers and 38 enlisted men was required to operate a single control SCR-584, the majority of these men belonging to the organization involved in making use of the radar data.

7-8. Teleran.—The examples of operational radar systems will be closed with a brief account of a system still in the developmental stage. It is called "Teleran" (for *Television Radar Air Navigation*) and involves the coordinated use of air and ground equipment.

The basic idea of Teleran is that high-performance ground radar, if supplemented by simple and reliable height-finding means and by

weather and route information, can provide all of the data needed by aircraft for en-route navigation and for airport approach. Further, high-precision ground radar can supply all the information that a pilot requires, in addition to his own flight instruments, for instrument approach and landing. All that is required is to get this information into the aircraft, where it can be displayed to and used by the pilot.

It is proposed to add to the display of a ground radar a chart that contains map, airways, and weather information of interest to pilots, and to send the resulting picture by television to the aircraft, where it is displayed to the pilot. Televising of the display is preferred to the use of the radar-relay methods described in Chap. 17 because television does not require the use of persistent phosphors, with their low maximum level of light intensity, in the aircraft cockpit, where the level of ambient light is likely to be high.

On the basis of the assumption that display of all aircraft in the picture offered a pilot would be useless and confusing, it is proposed to separate the radar signals according to the altitude at which the corresponding aircraft are flying. Thus, for example, a pilot at 5000 ft would be sent signals showing the positions of all aircraft in his neighborhood and in the altitude range from 3500 to 6000 ft. This can be accomplished technically by supplying each plane with a transponder beacon whose reply signal is coded (Sec. 8.8) in accordance with the altitude at which the plane is flying. This coding can be controlled automatically by an altimeter. At the ground radar station, signals in the various altitude intervals can be sorted out and displayed separately for transmission to aircraft using the system. The transponder also increases the range of the radar equipment and, if a reply frequency different from the transmitter frequency of the radar is used, eliminates difficulty from ground echoes which might interfere with seeing low-flying aircraft.

Figure 7-18 shows a Teleran display as it might appear to a pilot flying at 11,000 ft. At this altitude, since topographical features are of little interest, nothing is shown except other aircraft, towns, airports, airways, frequency channels in use by Teleran ground stations, and the direction and velocity of the wind. To assist the pilot in flying a course, the compass reading is repeated on a disk mounted over the face of the indicator and ruled with parallel lines showing the heading of the aircraft. These lines are shown dashed in Fig. 7-18.

It is essential for the pilot to know which of the beacon responses is that of his own plane. This is accomplished by having a rotating radial line, centered at the radar station (which appears in the center of the picture), form part of the display information transmitted from the ground station. This line is normally invisible, but the television receiver is controlled by a signal from the transponder in such a way that the line

appears bright when the beacon is being interrogated. Since the beacon is interrogated only when the radar is pointing at it, the bright line passes through the signal pip representing the individual plane carrying the equipment. Each pilot sees a different radial line indicating his own plane.

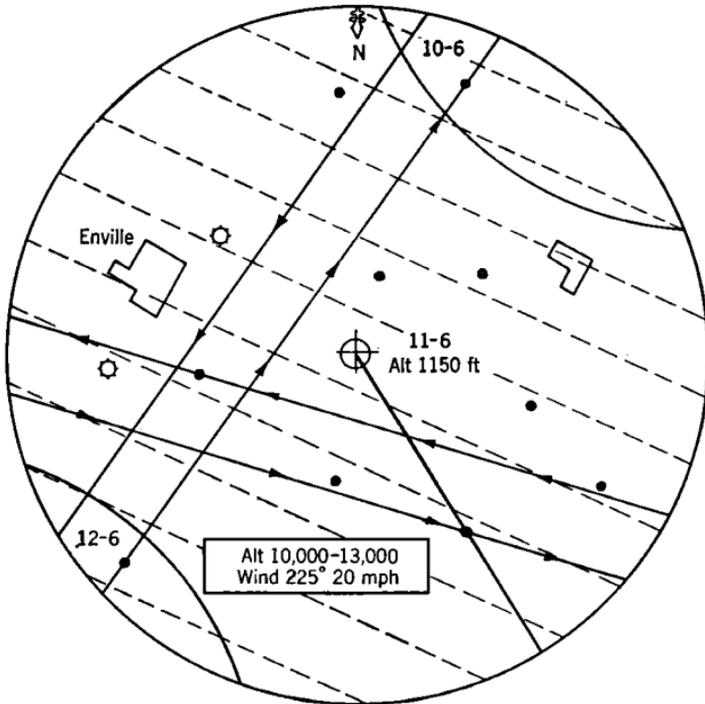


FIG. 7-18.—Schematic drawing of Teleran presentation in an aircraft flying at 11,000 ft.

The Teleran system has been worked out in much more detail than can be described here. Enough has been said to indicate the fashion in which system planning, based always on the needs of the user, leads to a functionally simple but technically elaborate final result, embodying not only radar but also beacons, television, aircraft instruments, and always—inescapably—men.

CHAPTER 8

RADAR BEACONS

BY L. A. TURNER

Introduction.—Radar waves are reflected by targets of different sizes regardless of their importance to the user of a radar set. The echo from an aircraft may be lost in much larger echoes from near-by mountains, or it may become too weak to be followed to ranges as great as desired. The echo from a friendly aircraft is like that from a hostile one. The exact location of a place on the ground may be of importance to a radar-equipped aircraft even though there is no distinguishable radar target at that point. In nearly all cases where it would be advantageous if an echo could be made much stronger or more readily distinguishable from other confusing ones, the use of a radar beacon is indicated. (An enemy aircraft obviously constitutes one difficult exception.)

The usefulness of beacons was demonstrated with the early radar sets that were operated at long wavelengths. A large proportion of the beacons used in the war operated at frequencies about 200 Mc/sec. These included the beacons used with ASV Mark II search radar, the transponders used for identification, the portable Eureka beacons that were part of the independent Rebecca-Eureka beacon system, and a much-used system for precise bombing, the Oboe Mark I. Another system for precise bombing, the Gee-H system, used beacons of even lower frequency. In this chapter, more emphasis is put on the newer beacons at higher frequencies since the trend in radar is toward microwaves.

The beacon is essentially a repeater of radar pulses. It has an antenna and receiver that convert pulses of energy, received at high frequency from a radar set or special interrogator, into triggering signals. Each such signal triggers the transmitter in the beacon and causes it to radiate one or more pulses of radio energy that may have almost any desired power, frequency, duration, number, and characteristic spacing. Figure 8-1 gives a block diagram of a beacon. Since it takes time for the beacon to react, the first reply pulse comes back to the radar set slightly delayed and indicates a range slightly greater than the true one. In many applications this delay is negligible, in others it is made to have a constant known value for which allowance can be made. The delay can be kept down to a few tenths of a microsecond when necessary. In the special case where the radiated pulse is single, has approximately the

duration and frequency of the radar pulse, and is not appreciably delayed, the beacon acts somewhat like an echo amplifier. The intensities of received and transmitted signals are, to be sure, independent, rather than in fixed proportion as in the case of a true amplifier. In general, however, the frequency of the reply will be different from that of the triggering radar set. In order to receive such replies, one needs a receiver tuned to the frequency of the pulses sent out by the beacon instead of to the frequency of the initial transmitted pulses. This may be either the receiver of the radar set tuned to the new frequency, or a second independent receiver tuned to the beacon. In either case, the receiver of the beacon signals does not receive radar echoes since it is not tuned for them. Thus,

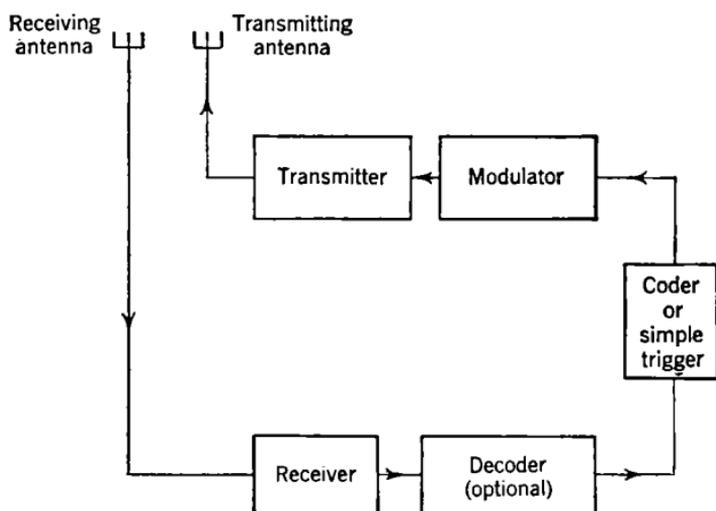


FIG. 8-1.—Block diagram of a beacon.

the beacon signals are separated from radar reflections and can be displayed without being swamped by heavy permanent echoes. Also, since the pulse power of the beacon transmitter can be made as great as desired, there is no limit to the strength of the reply. The range is limited only by the power of the radar transmitter and the sensitivity of the beacon receiver, which determine whether the beacon transmitter is triggered or not. Figure 8-2a shows the radar echoes and Fig. 8-2b the beacon replies on the indicator of the same 3-cm radar set. The pictures were taken one immediately after the other on the same flight.

By its very nature, the radar-beacon combination involves two send-receive links as does any two-way communication system. The two links are ordinarily connected automatically in a simple regular way and are uninfluenced by human reactions. Since the channels exist, however, they afford the basis for a communications system. In the past,

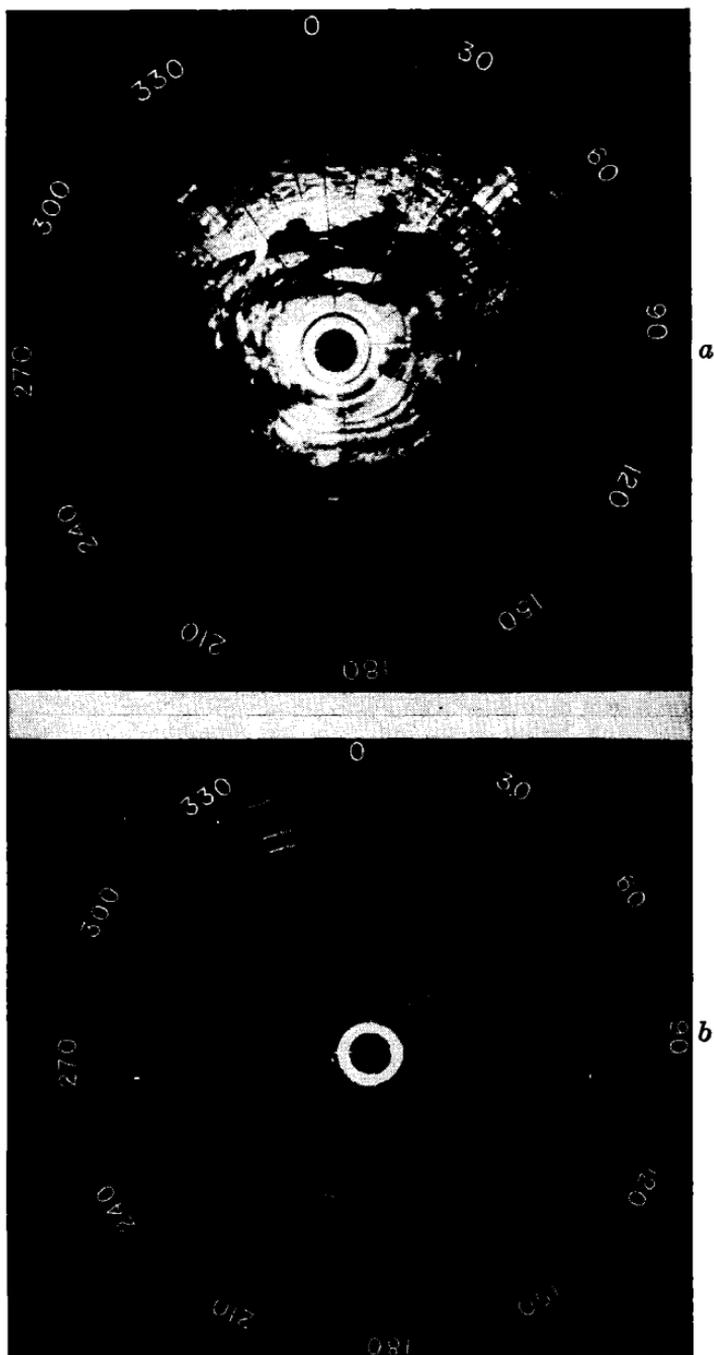


FIG. 8-2.—Radar and beacon signals of a 3-cm radar set: (a) shows only the radar echoes, (b) shows the beacon reply; both were received at nearly the same position.

beacons have sometimes been used in providing communication systems of a rudimentary sort and also for exercising remote control. Intelligence has been conveyed from the radar to a beacon-carrying vehicle by modification of the repetition rate, the length of interrogating pulses, their spacing in groups, or the duration of the intervals of interrogation. The replies of the beacons have also been modulated in such ways. One 200-Mc/sec beacon was used for two-way voice communication while still functioning for its normal use. For the most part, however, this use for communications has been somewhat incidental. Much fuller use of the channels could be made since the portion of the spectrum required for transmission of beacon signals of a simple sort is ample for conveying much more complicated intelligence.

Beacons of the synchronous sort just described have been variously called "radar beacons," "responder beacons," "racons," and "transponders," there being no essential distinctions among these terms. The discussion here is confined to such beacons since other free-running types, more like ordinary radio beacons, appear to be less useful in conjunction with radar sets. From the free-running type, only the bearing can be determined. Radar information is particularly useful because it gives accurate determinations of range; it is obviously sensible to provide, as an adjunct to radar, the sort of beacon that makes the best use of this property.

This chapter aims to give a brief résumé of the main points involved in the use of radar beacons. The design of beacons and of systems using them is treated at length in Vol. 3 of this series.

RADAR-BEACON SYSTEMS

8-1. Types of Radar-beacon Systems.—In a discussion of beacons it is convenient to classify them as fixed ground beacons, shipborne beacons, airborne beacons, and portable beacons. All of these may be used in conjunction with ground, ship, or airborne radar sets, or with special interrogator-responders. The following combinations have proved useful so far:

Ground Radar.

1. Shipborne beacons. This combination is of use principally for identification of particular ships since, in general, the radar echo from a ship is distinct enough.
2. Airborne beacons. This combination has proved to be of great usefulness for identification and for various purposes where ground surveillance and control of air traffic is desired. Figure 8-3 shows the AN/APN-19, an airborne 10-cm beacon.

A special system of precision bombing, known as the "Oboe

system," involved simultaneous determination by two ground stations of the ranges of a beacon-carrying aircraft with great accuracy. Coded signals for guiding the navigation of the plane were given by the radar sets by suitable modulation of the interrogating pulses.

3. Portable beacons. Portable beacons used with portable interrogators form the basis of a new and accurate system of surveying.

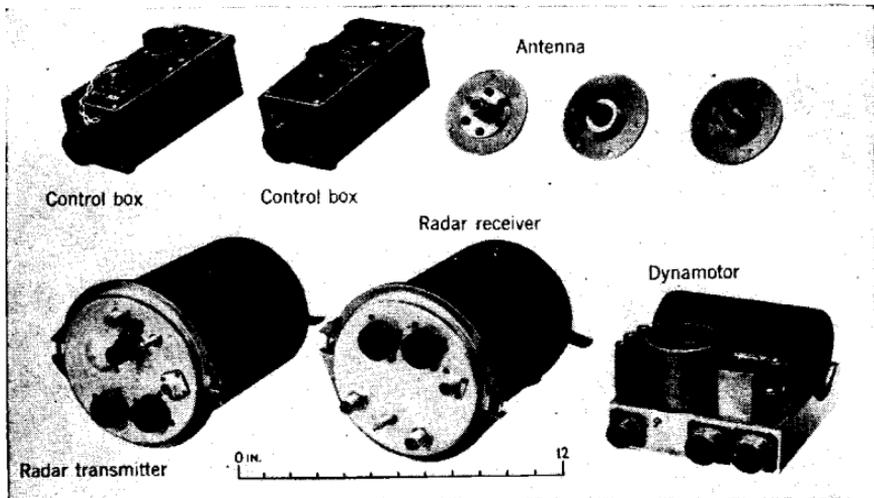


FIG. 8-3.—Component parts of the AN/APN-19 beacon. The AN/APN-19 is a 10-cm beacon designed for installation in aircraft.

The position of ground radar sets can sometimes conveniently be fixed by measurement of ranges to portable beacons placed at known points.

Ship Radar.

1. Fixed ground beacons. The beacon is here the radar analogue of the old-fashioned lighthouse. Since range to the beacon is measured as well as its bearing, observation of a single beacon gives two intersecting lines of position and a fix more accurate than is customarily obtained by other methods.
2. Shipborne beacons. This combination can be of use principally for identification.
3. Airborne beacons. Beacons have facilitated control of military aircraft from a ship, and have been useful for identification of aircraft. In the past war, they were used to enable ships to home on aircraft that were orbiting in regions where submarines had been sighted.

4. Portable beacons. Small beacons have been used for marking life rafts and small target rafts.

Airborne Radar.

1. Fixed ground beacons. Beacons used with microwave radar having PPI presentation give most satisfactory navigational fixes by graphically locating the aircraft with respect to several known points on the ground. Figure 8-4 shows one of the AN/CPN-6 beacons, a 3-cm beacon of high power designed for use in air navigation.

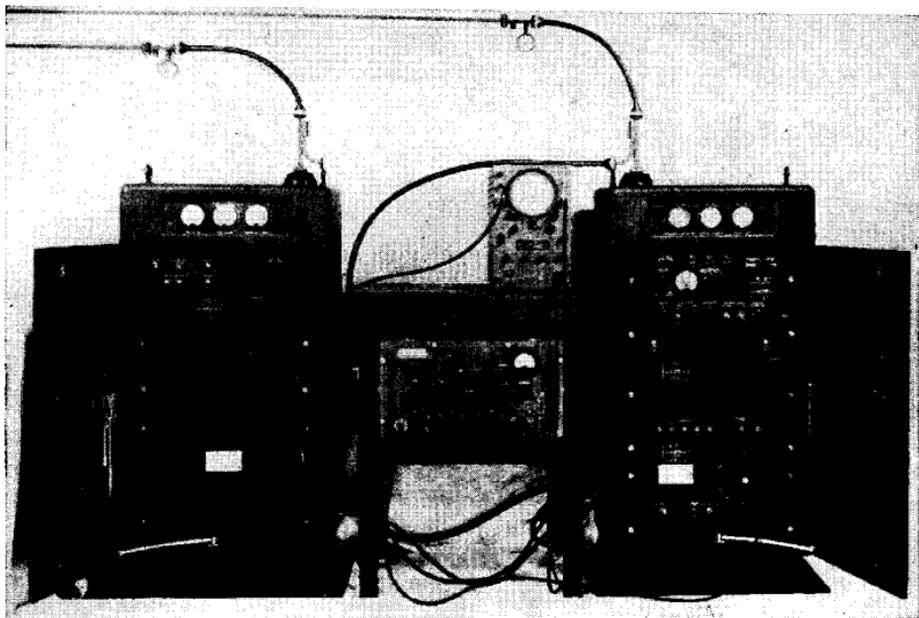


FIG. 8-4.—The AN/CPN-6 beacon. The AN/CPN-6 is a high-powered 3-cm beacon intended for use in permanent installations on the ground.

Special systems designed for measuring with great accuracy the ranges from two ground beacons at known positions have been used for mapping and blind bombing of high precision. Since in such systems there is no need for determining azimuth, lower frequencies can be employed advantageously, omnidirectional antennas can be used in the aircraft, and the beacon replies (which are then steady) can be presented on special indicators designed for ease in getting great accuracy of measurement.

Beacons at longer wavelengths have been used in blind-approach systems.

2. Shipborne beacons. Some beacons have been used for identifying ships, others for enabling aircraft to home on their carriers or on convoys.

3. Airborne beacons. This combination has been useful as an aid in effecting rendezvous and for identification.
4. Portable beacons. Such a combination has been of military use under circumstances where fixed ground beacons could not be set up readily, but small portable beacons could be used instead.



FIG. 8-5.—The AN/PPN-2 beacon. The AN/PPN-2 is a 1.5-m beacon designed for use principally by paratroops.

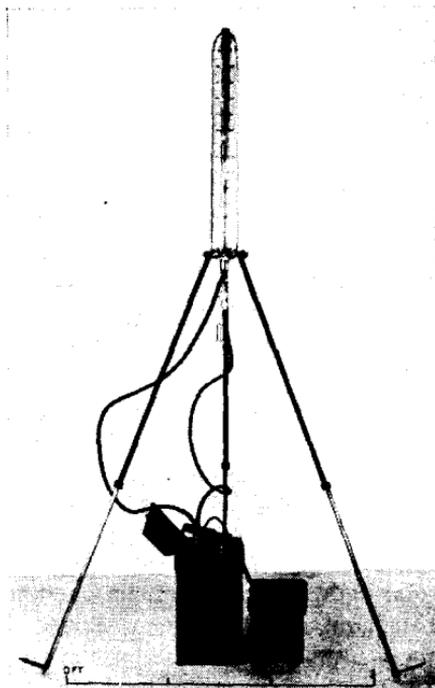


FIG. 8-6.—The AN/UPN-1, a portable battery-operated 10-cm beacon.

Such beacons have been used principally to indicate to supporting aircraft the location of isolated forward elements on the ground, such as advance parties of paratroops or secret agents. Figure 8-5 shows the AN/PPN-2, a 1.5-m beacon carried by pathfinder paratroopers; Fig. 8-6 shows the AN/UPN-1, a 10-cm beacon that was similarly used. Similar beacons have been used for marking life rafts at sea.

Table 8-1 collects and summarizes all these applications.

TABLE 8-1.—SUMMARY OF PRINCIPAL USES OF BEACONS

Radar or other interrogator	Beacons			
	Ground (fixed)	Shipborne	Airborne	Portable
Ground	Identification	Identification Ground-controlled precise navigation	Surveying
Shipborne	Pilotage	Identification	Control of aircraft Homing	Life rafts Shore bombardment
Airborne	Navigation Blind approach	Identification Homing	Identification Rendezvous	Temporary marking of points on land Life rafts

8-2. Systems Planning.—Given a particular radar set, it is simple enough to provide a beacon for any special use. The receiver of such a beacon can be of narrow bandwidth and tuned to the frequency of the radar. The reply can be either at the same frequency or at one just different enough to permit separation of radar echoes and beacon replies, but still receivable merely by minor adjustment of the radar receiver. When, however, it is desired to provide a beacon that will be useful to many different radar sets of the same type, the problem is more complicated. Radar sets of a given kind are usually operated at somewhat different frequencies in order to avoid mutual interference. Thus, the receiver of the beacon has to have sufficient bandwidth to receive the interrogating signal from any one of the radar sets. For airborne 3-cm radar sets, for example, the band from 9320 to 9430 Mc/sec was used. This band of 110 Mc/sec was needed to take care of the variations of the frequencies of the magnetrons as manufactured plus further changes to be expected in adjustment and use in the field. The reply, however, had to be made at some particular frequency; that used was 9310 ± 2 Mc/sec. Provision must be made in the radar receiver for quick and accurate tuning to a chosen beacon reply frequency if beacon signals are to be used.

Radar sets and beacons cannot be so designed without careful planning. If the full potentialities of radar beacons are to be realized, *both radar sets and beacons must be planned together as parts of a unified radar-beacon system.* This now seems trivially obvious, but it is not the way that much of the existing equipment was designed. In the development of radar, the beacons came as an afterthought. The result was that

beacon performance often was not so good as can be obtained. The designer of future radar systems, having knowledge of radar and radar-beacon possibilities, should decide on the scope of desired operational characteristics of his system and then design his radar set and beacons together in order to achieve the result most efficiently.

If the frequency of the beacon reply is to be very different from the frequency of the radar transmitter, it is usually necessary to provide a separate antenna and receiver for the beacon signals. The scanning of this second antenna must then be synchronized with that of the radar antenna. In some cases, this synchronization has been accomplished by using the same parabolic reflector with separate feeds for the two frequencies.

If the frequency of the beacon is close to that of the radar set, other problems arise. The same antenna can be used, but then the design of the duplexing system becomes more complex than it is for the simple radar set. Attenuation of the received signals at both frequencies must be held to a small amount.

In any case, it is desirable to include some device that will automatically keep the local oscillator for beacon signals in tune. When the beacon receiver is completely separate from the radar receiver, it is advantageous to provide switching arrangements so that the radar operator can have his choice of either radar or beacon signals alone, or both together. It is often useful to include separate adjustments for the saturation levels of the signals, so that stronger beacon signals will stand out when superimposed on saturated ground "clutter" from the radar. Also, if the pulses of the beacon are short, it is possible to improve the display by stretching them in the video amplifier of the receiver.

Some beacons are made to reply only to interrogating pulses of proper length, or to those having other special characteristics, as discussed below. In such cases, the corresponding changes have to be included in the modulator of the radar set, with appropriate controlling switches.

The above list of radar design features needed for best use of beacons appears somewhat formidable, and so it is when one is trying to patch up an antique radar set that does not have such features included in its original design. When starting afresh, however, the list does not involve unreasonable additional complications. Actually, the most compact airborne microwave radar set produced during the war was also by far the most simple to operate, and it incorporated nearly all of the features mentioned above. This set, designated AN/APS-10, is described in Chap. 15 of this book.

8.3. General Identification Systems, IFF.—The discussion of this chapter is confined almost entirely to the use of beacons for various navigational problems and for identification systems of somewhat

restricted scope. Mention should be made, however, of the more general problem of identification, which gave rise to a system known as IFF—Identification of Friend or Foe. This system was probably the most important single field of application of radar beacons in the war. The goal was to provide every friendly military ship and aircraft with a transponder that would give an identifying reply signal in response to proper interrogation. Since the location of the ship or aircraft in question was almost always determined by radar, the problem was actually one of determining the friendly or hostile character of radar targets. The IFF interrogating equipment was thus used in the great majority of cases as an adjunct to a radar set. The problem was enormously difficult in view of the rapid development of many types of radar sets in both Britain and America, the tremendous density of airplane traffic to be dealt with in many theaters and the necessity for having replies coded with sufficient elaborateness to prevent the effective use of captured IFF transponders by the enemy for deception. In addition to these and other inherent technical difficulties, there were still others of a political sort. Among these were the problems of getting agreement among the many branches of the British and U.S. Armed Services concerning details of desired characteristics of systems and of planning manufacture, distribution, installation, and maintenance of equipment in such a way that an IFF system could actually go into effective widespread use in a theater by some target date. It seems that problems of this sort are to be encountered in connection with any such system that is meant to have widespread use. In spite of all these difficulties, one such system was put into use in nearly all theaters, and was of great help where its potentialities were understood and its limited traffic-handling capacity was not exceeded. It is beyond the scope of this chapter to go further into the problems of IFF. A peace-time requirement that is likely to pose problems of similar character, although of lesser complexity, is that of control of air traffic in the neighborhood of airports where the traffic is likely to be heavy.

8-4. Radar Interrogation vs. Special Interrogators.—Much of the preceding discussion implies that beacons are to be used principally with radar sets and to be interrogated by them. For IFF, however, because of the great variety of radar sets at different wavelengths, it is necessary to pick particular bands of frequencies for the IFF beacons and to supply supplementary interrogating equipment to work with the radars. For every proposed radar application where beacons would be of use, the question will arise whether it is better to use radar beacons or to provide such separate equipment to work at frequencies set aside for the purpose. There seems to be no single correct answer. Separate cases must be considered separately, but it is desirable that in so doing the proposed

systems be conceived as broadly as the possibilities for unified planning, control, and operation will permit.

Let us first consider the planning of beacons for use with airborne radar, which has been the subject of some controversy in the past. If radar beacons are to be provided, a whole new set of them is required for every new band of frequencies used. British opinion has inclined to the view that the resulting multiplicity of beacons is intolerable, that it involves far too great a cost for the design, manufacture, installation, and maintenance of these many different beacons. The British policy has been to design for every radar set a supplementary synchronous beacon interrogator for interrogating beacons in the 200-Mc/sec region and receiving replies at such a frequency, the replies to be displayed whenever possible on the indicator of the radar set. These interrogators were to be replaced in due course with similar equipment operating at a new and higher region of frequency. Special separate antennas are obviously required. The contrary view, more widely held in the United States, is that the provision of suitable beacons is but a small part of the over-all complication and expense of introducing radar in a new band of frequency; that the frequencies used for radar sets for any given purpose tend to group in a small number of reasonably narrow bands; that since the radar set itself can be as good an interrogator-responder as one could desire, it is putting complication in the wrong place to add more equipment to a crowded airplane; and that the display of the beacon signals and the performance as a whole will be inferior when the beacon frequency is considerably below the radar frequency.

The decision must depend, to a considerable extent, on the relative importance of using beacons for navigation and for identification. For an identification system, it is obviously necessary to have but one relatively simple type of beacon and to require that the identifying radar sets be accommodated to it. For navigational purposes, the requirements of the airplane become relatively more important.

If all information needed for the desired use can be obtained with beacon signals alone, it is not necessary to have a proper radar set at all, and the radar set can then be replaced by an interrogator-responder. This has a transmitter and receiver like those of a radar set, but it is in general somewhat smaller and lighter since not so much transmitted power is required for triggering beacons as is required for getting adequate radar echoes. Such equipments can be especially economical in size and weight if range only is wanted, or if the sort of azimuth information obtainable by lobe-switching is adequate. In such cases the sets can be run at low frequencies, and the equipment is relatively compact. If, however, it is desired to have PPI presentation of the beacon signals, which affords azimuth information comparable to that given by modern

radar sets, the interrogator-responser becomes so nearly a complete radar set that it may as well be made to be one.

In the case of beacons carried by aircraft to extend the range and to facilitate identification by ground radar sets, it is obviously desirable to have only one type of beacon for all aircraft under the surveillance of a comprehensive system of like ground radar sets. It is also desirable that these beacons give signals of the same degree of resolution in azimuth as that of the radar signals. It seems likely that the radar sets and beacons will be designed together as parts of a system operating at 3000 Mc/sec or higher. There appears to be little to be gained by trying to have this system operate at the same frequencies as the beacons for use with airborne radar sets unless it is required that the same beacons that are to be interrogated by the ground radar sets also reply to interrogation by airborne radar sets.

8-5. Independence of Interrogation and Reply.—Although the fundamental considerations that underlie the design of beacons and beacon systems are the same as those for radar in general, there are certain consequences of the almost complete independence of the interrogation and reply links that need explicit mention.

Range.—The expression for the power received by the beacon is given by Eq. (2-14). For its maximum value, using a notation defined below, we get

$$(P_{Si})_{\max} = \frac{P_{Ti}(G_{Ti})_{\max}(G_{Si})_{\max} \cdot \lambda_i^2}{16\pi^2 R^2} \quad (1)$$

Subscript *i*—interrogation leg.

Subscript *r*—response leg.

Subscript *T*—transmitting components.

Subscript *S*—receiving components.

R = distance from interrogator to beacon } same units.
 λ = wavelength of the radiation }

P = power in watts.

G_{\max} = maximum value of effective gain.

$(P_{Si})_{\max}$ = maximum value of the available peak power received by the beacon.

P_{Ti} = peak value of the power transmitted by the interrogator (radar).

$(G_{Ti})_{\max}$ = maximum value of the gain of the transmitting antenna of the interrogator.

$(G_{Si})_{\max}$ = maximum value of the gain of the receiving antenna of the beacon.

λ_i = transmitted wavelength.

If we let P_{Si}^0 represent the peak value of available received power necessary for triggering the beacon and R_i^0 represent the corresponding

range, Eq. (1) gives

$$R_i^0 = \frac{\lambda_i}{4\pi} \left[\frac{P_{T_i}}{P_{S_i}^0} (G_{T_i})_{\max} (G_{S_i})_{\max} \right]^{1/2}. \quad (2)$$

In Eq. (2) R_i^0 is the maximum free-space range for interrogation of the beacon; when it is at greater ranges it will not be triggered.

The corresponding expression for the reply link is

$$R_r^0 = \frac{\lambda_r}{4\pi} \left[\frac{P_{T_r}}{P_{S_r}^0} (G_{T_r})_{\max} (G_{S_r})_{\max} \right]^{1/2}. \quad (3)$$

In Eq. (3), $P_{S_r}^0$ is the peak value of the available received power required to give a satisfactory beacon signal on the indicator of the interrogating system, due account being taken of scanning losses and other circumstances of use. The range R_r^0 is the greatest one for which reply signals from triggered beacons will be usable. This cannot be defined as clearly as is R_i^0 since having a signal "satisfactory" is less definite than having it either present or absent. The other quantities in Eq. (3) are of obvious significance if it is kept in mind that the subscript r refers to the reply link.

In the preceding paragraphs it was assumed that the triggering of the beacon and the intensity of the displayed reply depend on the pulse powers of the two pulses. This implies that both the receiver in the beacon and the one in the interrogating system have a bandwidth sufficient to insure that both received signals rise to a value that depends on the pulse power and is independent of the length of the pulse within wide limits. This condition is usually the desired one and the receivers are designed accordingly. If a receiver of very narrow bandwidth should be used, however, the peak value of its output would then depend upon the energy of the pulse, rather than upon the pulse power. For such cases Eqs. (2) and (3) would be modified by substituting the transmitted and received values of the energies per pulse for the respective pulse-power values.

It is sensible to design the system so that $R_r^0 \sim R_i^0$ since there is no use in having replies that are too weak to be observed even though the beacon is being interrogated, as can be the case if $R_r^0 < R_i^0$. Likewise nothing is gained by making provision for strong replies if they are absent because of the failure of the interrogation. This situation can arise if $R_r^0 > R_i^0$. The common value of $R_i^0 = R_r^0$ must be made equal to the desired maximum range, and indeed somewhat larger if a reasonable factor of safety is desired. It is sometimes useful to make R_r^0 somewhat larger than R_i^0 in order to facilitate recognition by the operator of beacon signals in the noise.

In the case of beacons interrogated and received by a radar set, λ_r is nearly equal to λ_i ; $(G_{Ti})_{\max}$ usually equals $(G_{Rr})_{\max}$ since the same antenna is likely to be involved; and $(G_{Ri})_{\max}$ usually equals $(G_{Tr})_{\max}$ since beacon antennas for receiving and transmitting need to have the same radiation pattern and are therefore alike. The condition for balance of the two links, namely that R_i^o equal R_r^o , then becomes

$$P_{Ti}/P_{Si}^o = P_{Tr}/P_{Sr}^o \quad \text{or} \quad P_{Tr}/P_{Ti} = P_{Sr}^o/P_{Si}^o. \quad (4)$$

The receiver of the beacon is almost always less sensitive than that of the radar, either by necessity (because of the greater bandwidth), or by choice to avoid undesired triggering (see below); hence $P_{Sr}^o < P_{Si}^o$. In the balanced system, then, the transmitter of the beacon may as well be less powerful than that of the radar in the same proportion.

In actual cases, Eq. (1) and a corresponding equation for the received power in the reply link involve two additional factors. One of these is a geometrical factor that expresses the effects of interference and diffraction. The other takes account of the attenuation of the radiation in the atmosphere. These matters as they pertain to radar echoes are discussed in Secs. 2-12 and 2-15. It should be noted, however, that the numerical factors appropriate here are the square roots of those for radar signals, when for the latter the effects on both the transmitted and reflected pulses are lumped together. These factors drop out of the expression (4) just as the antenna gains do, since they have the same value for both links. In any beacon system in which the frequencies of interrogation and reply are substantially different, however, all of the factors in Eqs. (2) and (3) must be retained and these additional ones must be added. The two links have to be designed separately to get proper operation at all ranges up to some fixed value or to meet other particular requirements.

It is apparent from Eq. (1) that another consequence of the independence of the two links is a different law for the relation between signal strength and range. Radar echoes vary in pulse power as the inverse fourth power of the free-space range, whereas beacon replies vary only as the inverse square. In general, beacon signals do not vary as much between particular values of range, do not cut off as abruptly with increasing range, and do not give as deep interference minima as radar signals do.

Azimuth.—One further consequence of the separation of the two links is that the scanning sector through which beacon replies are obtained is limited either in the interrogation link or in the reply link according to circumstances. Equation (1) becomes more generally

$$P_{Ri} = \frac{P_{Ti}G_{Ti}G_{Ri}\lambda_i^2}{16\pi^2R^2}. \quad (5)$$

Here P_{Ri} is the available pulse power at the beacon for any relative position of the interrogator and the beacon. The gain of the transmitting antenna, G_{Ti} , is a variable that depends upon the angular position of the antenna with respect to the line from it to the beacon. The received power is a maximum only when the interrogator is really looking at the beacon. Similarly G_{Ri} , the gain of the beacon receiving antenna, is a variable depending on the orientation of the antenna with respect to the line from the beacon to the interrogator.

Let us consider, as an example, a ground radar set scanning around through the complete azimuth circle and interrogating an airborne beacon. Curve *a* of Fig. 8-7 represents schematically a polar diagram of the logarithm of the power received at the beacon as a function of the angular position of the radar antenna.

The line to the beacon is assumed to be the upward vertical. In the ideal case the curve is the same as the antenna pattern of the radar; in actual cases it is often modified by reflections from near-by buildings, hills, etc. If circle *b* represents the threshold power for triggering, the beacon will be triggered only through the sector *AOA'*. If, however, the values of received power should all be increased by 12 db or so (by decreasing the range to the beacon), the beacon would also be interrogated by side lobes of the antenna.

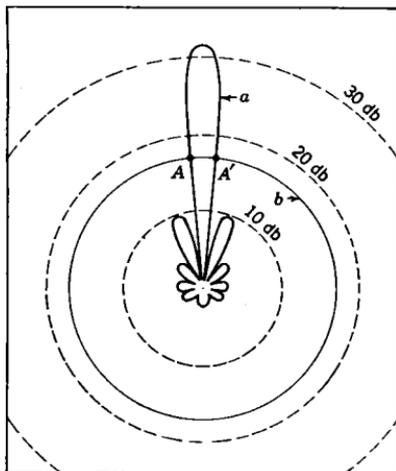


Figure 8-8 shows this effect. It is clear that if we consider interrogation that begins when the separation is first large and then decreases, the beacon

is first interrogated at maximum range through a very narrow angular region. As the range decreases, the sector of interrogation increases, but remains of the order of magnitude of the half-power beamwidth of the antenna until a range about one tenth of maximum is reached. There, assuming the side lobes of the antenna pattern to be about 20 db down, side-lobe interrogation begins and grows with decreasing range. Finally, in practical cases, a range is reached for which the beacon is interrogated no matter which way the antenna is pointed. It is also clear that reduction of the radiated power would produce a narrowing of the sector of interrogation of any particular beacon, but would lead to failure to interrogate those at greatest range.

FIG. 8-7.—Schematic logarithmic polar diagram of the antenna pattern of a microwave radar. The triggering of a beacon occurs when the energy radiated in a given direction is in excess of some necessary minimum amount.

Similar considerations apply to the reply link. If we assume the beacon to be interrogated for all angular positions of the antenna of the interrogator and the sensitivity of the receiver to be kept constant, the angular width of the beacon arc on the PPI tube will vary as did the sector of interrogation in the previous case. Over a large distance the reply arc will be of the order of magnitude of the half-power beamwidth of the receiving antenna. For close-in beacons, extensions attributable to side lobes appear, and for very close beacons a complete circle is obtained. The complete circle or side-lobe pattern can be reduced to the narrow arc by suitable reduction of the gain of the radar receiver, again with the loss of more distant beacons. Here, however, it is possible to use a

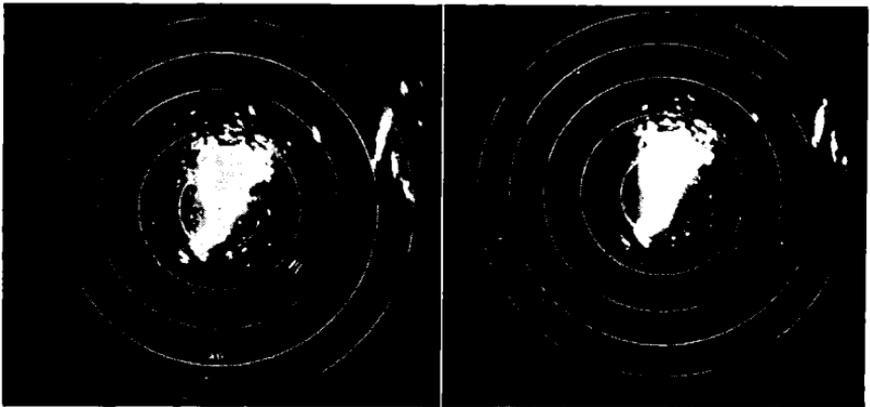


FIG. 8-8.—Interrogation of a beacon by side lobes. Interrogation of a shipborne 10-cm beacon by a radar set on Mt. Cadillac on Mt. Desert Island, Me. (a) shows the extended pattern at about 4 o'clock in the picture; (b) shows how it can be reduced to the reply in the main beam by manipulation of the gain control of the receiver of the beacon signals.

sensitivity-time-control (STC) circuit that gives automatic variation of the sensitivity of the receiver as a function of the elapsed time after emission of the interrogation pulse. The gain is thus automatically adjusted to be correct for beacons at all ranges with the result that they all appear as approximate half-power beamwidth arcs. This circuit must be carefully designed if it is to give good results with ground radars interrogating beacons on aircraft flying in the maximum of the pattern of the receiving antenna and still not unduly attenuate weaker replies from beacons in low-flying aircraft.

It is clear from the foregoing that the width of the reply arc is always limited in either one link or the other, whichever is the narrower, and that it can be controlled by adjusting either transmitted power or receiver gain. Where manual adjustment for good reception of a particular beacon is to be used, cutting down the transmitted power would be

preferable since overinterrogation is thereby minimized. Manual adjustment would be helpful in systems where a large number of airborne radar sets are likely to be interrogating some particular beacon on the ground or on a ship. This procedure has not been used so far, however, since the reduction of the transmitted power usually requires more complica-

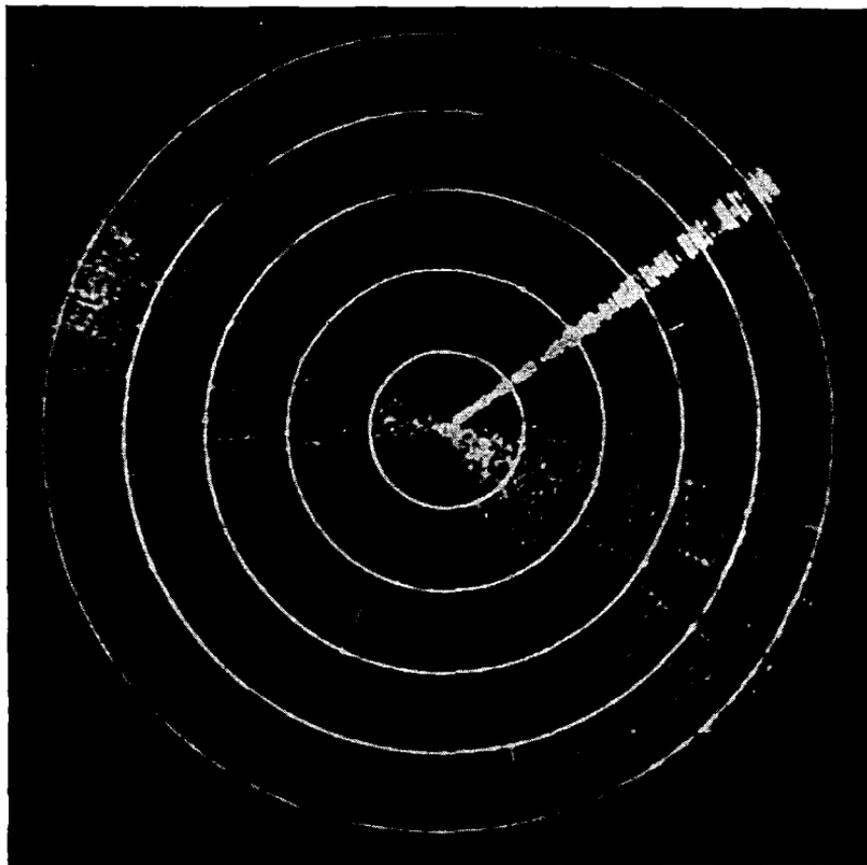


FIG. 8-9.—Replies from an experimental beacon interrogated at 10 cm and replying at 1.5 m. The exposure was continuous; the strong radial line in the picture is therefore made up of the successive beacon replies. The sweep was delayed, so that the replies extend from 50 to 100 miles.

tion of the radar set than does reduction of the receiver gain. Automatic adjustment of the sensitivity of the receiver is more appropriate for ground radar sets interrogating airborne beacons at many different ranges, or wherever the operational simplicity of automatic adjustment of gain is more important than the avoidance of possible overinterrogation.

It has proved convenient for some purposes to provide beacons which are interrogated by microwave radar sets and which reply at a frequency

lower than that of the radar by a factor of 20 or so. Either the beamwidth of the antenna for receiving replies from the beacon is made very broad or this antenna is made omnidirectional. Microwave angular discrimination is still provided by the sharpness of the interrogating beam—again only for beacons at ranges beyond that for which side lobes begin to cause broadening of the reply. Figure 8-9 shows a whole set of replies from such a beacon, exposures having been made every few sweeps of the PPI. The interrogations were at 10 cm, the replies at 1.5 m. The principal reason for using such a beacon rather than one that gives a microwave reply is the relative simplicity of the low-frequency transmitter and antenna of the beacon. Also, by placing several simple antennas and receivers for the interrogation at different places on an aircraft, it is possible to achieve a good approximation to an effectively omnidirectional pattern for the interrogation link. This is almost impossible to achieve with single simple microwave antennas because of interference effects. It is necessary to put the outputs of such multiple receivers in parallel after detection so that there will not be a combination of r-f signals in varying phases to give rise to interference maxima and minima.

The foregoing discussion implies that the broadening of the arcs for beacon replies is greater than that for radar signals. This is not the case, however, if the maximum range for both is the same. If G_{\max} and $G_{(\text{lobe})}$ are the respective gains of the antenna at the center of the pattern and at the maximum of the first side lobe and R_i^0 and $R_{i(\text{lobe})}$ the corresponding maximum ranges for interrogation of the beacon (to give minimum triggering power), from Eq. (5) we get $G_{\max}/(R_i^0)^2 = G_{(\text{lobe})}/R_{i(\text{lobe})}^2$.

The same line of argument followed through with the radar equation gives $G_{\max}^2/(R_i^0)^4 = G_{(\text{lobe})}^2/R_{i(\text{lobe})}^4$, which gives the same value of $R_{i(\text{lobe})}$ when R_i^0 is the same. In practice, however, the broadening of the beacon signals is likely to be greater since it is customary and usually desirable to set the sensitivity of the beacon receiver so that the maximum free-space range is considerably greater than the radar range to an average target. This is done partly to ensure increased range and partly to allow for the considerable unavoidable variations between the receivers of different beacons as installed and operated.

8-6. Frequency Considerations. *General.*—Since beacon replies are much like radar echoes, the considerations involved in choosing the frequency region for a system in order to get the desired resolution are much the same. In general, the higher the frequency, the better the azimuth discrimination for a given tolerable size of antenna system. This advantage of improved angular separation with increased frequency is somewhat offset by a tendency toward increased size and weight of transmitters needed for a given range performance.

The most satisfactory presentation of beacon replies for purposes of ordinary navigation is undoubtedly the PPI since it gives the whole situation at a glance in a direct way. The angular widths of beacon replies should be limited to a few degrees, if possible. For an airborne interrogator this suggests the use of 3-cm or shorter wavelength; for systems having ground interrogators with large antenna arrays a considerably lower frequency can be used. Azimuth determination by lobe-switching, involving comparison between the strengths of signals received on two antennas pointed in different directions, is adequate for many purposes but relatively clumsy in use. It permits the use of lower frequencies and much broader antenna patterns, with corresponding reduction in size of the antennas, but it results in a great increase in the difficulty of identifying beacon replies with particular radar echoes. It is, however, satisfactory for aircraft homing on the beacon. In the wartime systems which provided means of precise navigation by accurate measurement of the ranges to two ground beacons at known positions, special indicators were used. They displayed and measured the positions of the steady beacon signals obtained by using omnidirectional antennas rather than scanners.

There is an upper limit on the frequency to which it is desirable to go. At about 15,000 to 20,000 Mc/sec the attenuation by water vapor in the atmosphere of the earth begins to become of consequence. In conditions of bad weather, when navigational beacons are most needed, the range of beacons at such frequencies would be lowered so much that they would be of little use.

In nearly all cases, the frequency of the beacon reply should be different enough from that of the interrogating pulses to obviate simultaneous reception of beacon replies and radar echoes by one receiver. In this way the swamping of beacon replies by strong echoes is eliminated. Even when it is desirable that radar echoes and beacon replies be presented simultaneously, the saturation video levels for the two kinds of signals can be made different by using separate receivers; thus beacon replies can be made to stand out even when superimposed on saturated ground clutter. A very striking differentiation between radar and beacon signals from aircraft has been achieved by putting them on separate PPI tubes that give signals in two different colors, the signals being effectively superimposed by optical means.

If any interrogator or radar set of a given type is to be able to locate a beacon readily, the beacon must transmit at some known frequency and the receiver for the signals must be readily tunable to that precise frequency, whether the beacon replies are received or not. When the location of the beacon is unknown (as is usually the case) and the interrogator is scanning, it is not tolerable to increase the general uncertainty about

the beacon by having to search over a band of frequencies. Thus, adequate stabilization of the beacon transmitter is required, and the receiver for the signals must have automatic frequency control, or a wavemeter must be provided, or it must be possible to set the frequency accurately to preset values by having mechanical parts of sufficient precision. All of these techniques have been worked out and are now available.

This use of a spot frequency for the reply does have a disadvantage, however. Each interrogating radar will receive not only the responses of the beacon to its interrogation pulses, but also the beacon responses to interrogating pulses from other sets. The replies that a radar set receives to its own pulses are synchronous and appear at the proper range in a regular way. Replies to other radar sets are, in general, not synchronous. This is discussed further in Sec. 8-9.

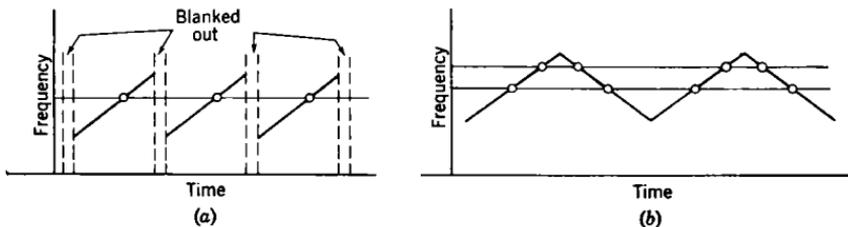


FIG. 8-10.—Variation of frequency with time in a beacon. With sweeps as in (a), equal intervals between replies will be obtained for all frequencies. For (b) the distribution depends on the frequency.

Sweeping Frequency.—In some cases it has not been feasible to use a single spot frequency of reply either because the number of replies at the frequency would be excessive or because it was not practical to modify a group of radar sets to receive the chosen frequency. In these cases, beacons have been used in which the common frequency of receiver and transmitter was made to sweep periodically over the band of frequencies of the interrogators. A given interrogator got replies when the beacon came into tune with it. This system is useful only with interrogators that are pointed continuously at the beacon. If the interrogator scans in space, it will not in general be pointed at the beacon during the interval that the beacon is tuned to its frequency band unless the scanning period is made short compared with the interval during which the beacon replies.

Some such beacons have a single tube used both as a superregenerative receiver and as a pulsed transmitting oscillator so that by mechanical changes of the tuning of the resonant circuit both frequencies can be changed together. The frequency for reception and that for transmission are nearly the same; they are not identical because of the different voltages and transit times involved at the two different levels of operation. Figure 8-10 shows two readily realizable cycles of variation of frequency

with time. The intersections of the horizontal lines with the sawtooth curves give the times at which replies are received by a radar tuned to the frequency corresponding to the horizontal line. In case *a* all interrogators at different frequency get regularly spaced pulses, in case *b* they come in pairs, the grouping depending on the frequency. This use of a single tube for receiver and transmitter is convenient especially because it tends to mitigate the effects of antenna mismatch often encountered in installations in the field. The shift of frequency for both transmitter and receiver is automatically almost exactly the same. Separate receivers and transmitters whose tuning controls have been ganged can be used if desired but this method requires more care in design and installation.

CODING

The simplest type of beacon replies with a single pulse to every pulse of sufficient strength received within a certain band of frequency. For purposes of reducing interrogation by confining it to only those interrogators that are intentionally seeking beacon replies, coding of the interrogation may be used. Likewise, the replies may be made more complicated in a variety of ways for the purpose either of identifying the beacon or of using it as a part of an auxiliary communication system.

8-7. Interrogation Codes. *Frequency.*—If the beacon is one that replies to all radars of a given type, the frequency for interrogation is not highly characteristic. If it is one to be used in a system with specially designed interrogator-respondors, certain discrete interrogation frequencies can be used as part of a code characterizing the particular beacon.

Pulse Length.—In order to avoid excess interrogation of ground beacons by airborne radars not interested in beacon replies, some beacons have been designed so that they are triggered only by pulses longer than those used for ordinary radar search. Careful design of the beacon receiver is required to prevent the stretching of strong, short pulses to give the same effects as longer ones of medium strength, but this problem has been solved satisfactorily. The beacon reply is obviously delayed since it cannot occur until after the lapse of a time longer than the duration of the search pulse. This delay must be allowed for in estimating the range to the beacon.

Multiple Pulses.—The beacon can be made to include a decoder so that a trigger is developed only upon reception of a pair of pulses having the proper separation, or it can be made so that it will reply only upon reception of a still more complicated group of properly spaced pulses. A considerable time delay may be involved, but in many cases this can be obviated by having the sweep of the interrogating radar started by the last of the group of interrogating pulses. One of the principal benefits

to be derived by the use of this more complicated system is relative freedom from random triggering by other radar sets in the same band since the probability of getting such a group of pulses by accident decreases rapidly as the complication of the group increases. This matter is discussed more fully below. All such systems, however, are subject to some difficulties resulting from the presence of reflected signals that are delayed because of the longer path traversed.

Two-frequency Interrogation.—Another method related to the use of groups of pulses is the use of coincident interrogation at two different frequencies. The two pulses may be actually coincident, or one delayed with respect to the other by any fixed amount. Both the value of the second frequency and that of the delay give additional coding possibilities.

Communication by the Interrogation Link.—So far, methods of interrogation coding have been discussed which determine whether or not the beacon will reply when interrogated. More complicated schemes for coding the interrogating beam enable the transmission of information to the vehicle carrying the beacon. Such schemes are usually applied only in cases where the interrogating antenna points continuously at the beacon. Information in the telegraph code can be conveyed by turning the interrogator on and off, by varying the width of the pulses, by changing the spacing between pairs of them, by changing the number of pulses in groups, and by varying the pulse repetition rate. When the repetition rate is made high enough, the relative spacing of pairs of pulses or the pulse width can be varied at voice frequencies to give telephonic communication.

8-8. Reply Codes. *Frequency.*—Some coding can be had by use of several spot frequencies for beacon replies.

Gap Coding.—The beacon transmitter can be turned on and off to give Morse letters. This system is useful only when the interrogator is expected to look at the beacon steadily. A variant that has been suggested for use with scanning radars is a rapid code of this kind which would break the reply arc on the PPI into dots and dashes. Only simple codes could be used because of the limitation of beamwidth and the complete uncertainty as to the part of the coding cycle for which the broken reply arc would be centered.

Width Coding.—Letters of the Morse code can be transmitted by changing the width of the pulses in the proper sequence instead of turning the pulses on and off. Beacons emitting such width-modulated replies will give a display on scanning radars and have the additional slow gap code which can be read if the scanning is stopped.

Range Coding.—Codes readily visible on the PPI, as in Figs. 8-2a and 8-2b, are obtained by having the trigger operate a coder that causes the emission of a number of reply pulses. These may be grouped in time in

different ways in order to give various groups in apparent range on the indicator. They are instantly readable by rapidly scanning radars, but have the disadvantage of being more subject to confusion by overlapping patterns, when numerous beacons are present at nearly the same bearing, than are replies that give simpler patterns.

Voice Communication.—Here also, when the repetition rate is high enough, the relative spacing of pairs of reply pulses or the width of pulses can be varied at voice frequency to give telephonic communication, usable when the interrogator-responser points steadily at the beacon.

STATISTICAL CONSIDERATIONS

8-9. Traffic Capacity.—For every beacon there is some limit to the total amount of energy it can radiate per second, with a corresponding limitation of the number of its replies per second and of the number of interrogators that can work with it simultaneously. Beacons have been made in which the limit to the rate of reply was set by the amount of interrogation necessary to overheat the beacon until it started a fire, but this informal system has drawbacks. It is usually desirable to incorporate an arrangement for limiting the average rate of reply to a safe value. There are two principal ways of doing this.

In the first method, the sensitivity of the beacon receiver is decreased as the average number of replies increases. Weaker interrogations are thus rejected and the replies are limited to stronger interrogators up to the maximum number tolerable. For navigational beacons, in general, this is not a good scheme since it tends to eliminate replies to the more distant interrogators, which are likely to be the ones most in need of getting them. For some purposes, where beacons of short range are to be used in connection with landing systems, this might prove to be a useful method.

The second method involves reducing the ratio of the number of replies to the number of interrogations to keep the total number of replies within the safe limit. The beacon display of all interrogators is somewhat impaired if there is sufficient overinterrogation to reduce this ratio markedly, but there is no discrimination against the far-off interrogators. This result is accomplished by varying the period of insensitivity of the beacon following the emission of its coded reply. There is always a finite minimum time between successive beacon replies; the maximum number of possible replies is thus the reciprocal of this time. Usually this maximum number is too great for safety unless the "dead" time has deliberately been made greater than it would naturally be. The simplest way to protect the beacon against overloading is to arrange that the receiver will always be dead for the desired time after making a reply and will then recover sensitivity so quickly that stronger interrogators are

not favored. This simple system can lead to occasional unnecessary interference of one interrogator with another. Let us assume, for example, that there are but two interrogators and that the beacon could readily reply to both steadily. If the dead time is long, there will nevertheless be times when the interrogations from one beacon will arrive in the dead time following the reply to the other. A more complicated system that avoids this defect has been used. The dead time is normally short but is arranged to increase, as needed, with increasing average rate of interrogation. Too short a dead time can be troublesome, for it may lead to multiple interrogation by the same interrogator pulse that has reached the beacon not only directly, but also by other and longer paths involving reflection.

If W represents the probability that the beacon will reply to an interrogating pulse of sufficient strength, it is easy to show by a statistical argument that

$$W > \frac{1}{1 + n\tau} \quad (6)$$

and

$$W < \frac{1}{1 + (n - 1)\tau} \quad (7)$$

In these expressions, n is the number of like interrogators having almost equal recurrence rates, and τ is the ratio of the dead time to the total interval between interrogating pulses. Thus τ is the fraction of the time that the beacon is insensitive to a second interrogator if it is replying fully to a first one. If the recurrence rates of the interrogators were exactly equal, there would be a fixed phase relationship between pulses and the statistical argument would be inapplicable. When interrogating radars have crystal-controlled repetition rates that are almost exactly equal, the radar that first emits its pulses may for a time steal the beacon completely away from a second radar that emits its pulses somewhat later. Equation (6) results from assuming complete randomness of interrogations, and gives a result that is absurd when there is but one interrogator. Equation (7) results when we assume that the probability of reply to a particular pulse from one interrogator is completely uninfluenced by the existence of the other interrogations by that interrogator. This assumption is false since there is indirectly such an influence. The exact expression for W is cumbersome but this is of little practical consequence since it need not be used. The two values obtained from Eqs. (6) and (7) bracket the true value and differ from each other unimportantly in practical cases. A useful approximation for W is given by

$$W = \frac{1}{1 + (n - W)\tau}, \quad (8)$$

which transforms to give

$$n = \frac{W^2\tau - W + 1}{W\tau}. \quad (9)$$

Figure 8-11 gives a plot of Eq. (9) for two values of τ . For existing radars and beacons τ is ordinarily of the order of 0.1.

It is clear from the foregoing that there is no simple way to estimate the traffic capacity of a beacon system since it depends on the lowest value of W tolerable for satisfactory operation, and that in turn depends to a considerable extent on the indicators of the interrogator-responder. For certain actual microwave radars and beacons, something like five radars giving steady interrogations are allowable without serious impairment of the presentation. Where necessary, this number can be increased by designing the beacon to have a shorter dead time.

So far in this discussion, the results of scanning by the interrogating radars have been ignored. If a radar scans through 360° and interrogates the beacon through only 6° , it will load the beacon only $\frac{1}{60}$ as much as if it pointed at it steadily. The traffic capacity of the beacons in such a system is increased by a factor of about 60. Actually, it is necessary to take account of the statistical fluctuations of the interrogation. Therefore, the number of interrogating radars should be substantially less than would be allowable if they remained uniformly spaced in their intervals of interrogation. Interrogating planes near the beacon tend to interrogate through much greater sectors of their scan and thus load the beacon disproportionately. For this reason, control of the width of the reply arcs by decreasing the interrogating power would be desirable.

In any case, if large traffic is to be expected, it is desirable to use some sort of interrogation coding to ensure that beacons will be triggered only by those radar sets that want to look at them, and not by ones uninterested in getting replies from the beacon.

The effectiveness of multiple-pulse interrogation codes in cutting down unwanted beacon replies can be seen as follows. Let us assume that N random pulses per second are received by a beacon, and that the beacon is triggered if any pulse is followed by a second one that arrives later by an interval of time between $D - T$ and $D + T$. D is the delay, and $\pm T$ is the tolerance. Each pulse thus produces a following interval

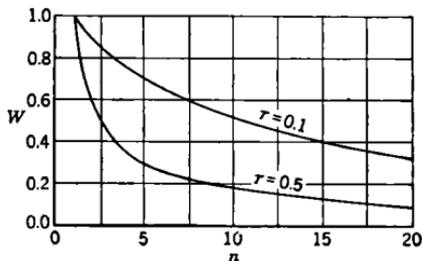


FIG. 8-11.—The fraction of interrogating pulses, W , getting replies as a function of the number of the interrogating radar sets, n .

of sensitivity to triggering by a second pulse of amount $2T$. If N pulses arrive per second on the average, the average sensitive time per second will be $2TN$. This is also the fraction of the time that the beacon is sensitive, or the probability that a random pulse will find it sensitive and trigger it. Since N random pulses arrive per second the number of triggerings per second will be N times this probability, or $2TN^2$. Thus, the ratio of the number of beacon replies to the number of random pulses per second is $2TN$. If, for example, $T = 5 \mu\text{sec}$ and $N = 500$ pps, the product $2TN$ has a value of $\frac{1}{200}$, the number of random triggerings per second being reduced from 500 to 2.5. If a three-pulse code is used, a similar argument applies for the third pulse and an additional factor of $2TN$ will come in to give $4T^2N^3$ as the number of random replies per second.

The argument also applies to two-frequency interrogation. Here the number of random interrogations will be $2TN_1N_2$ if N_1 and N_2 are the respective random rates at the two frequencies, and the beacon is made so that the two pulses have to arrive in a certain order for triggering to occur. If the order is immaterial, the rate of random interrogation will be twice as great, namely, $4TN_1N_2$.

In making a comparison between the two systems, it is necessary to keep in mind that one must use estimates of the numbers of random pulses that will be present after the system has been introduced. Changing radar sets to give double pulses increases the number at the radar frequency; adding coincident interrogating pulses at a second frequency increases the number at that frequency.

8-10. Unsynchronized Replies.—When a beacon is being interrogated by several radars at the same time it replies to all, and the receiver of each interrogator-responder will respond to all of these signals. Replies to the interrogation by a particular radar are synchronized and give a stationary pattern on its indicator as do its radar echoes; the replies to interrogations by the other radar sets appear to be unsynchronized and thus show at random on the screen and are similar to random noise from other causes. Figure 8-12 shows several beacon signals appearing on the PPI of a 3-cm radar; one of them is sufficiently overinterrogated to show unsynchronized replies. This effect is displeasing aesthetically, but experience has shown that a great amount of overinterrogation is necessary to interfere seriously with recognition of the beacon. The problem is more troublesome when numerous adjacent ground or ship interrogators are interrogating numerous airborne beacons. Unsynchronized replies coming from the multiple interrogation of a near-by beacon may interfere with detection of the signals from one farther away at the same azimuth. It is difficult to make a definite statement about the extent of the interference because it depends so much on the characteristics of the radar set. The use of

delayed fast sweeps is helpful since the density of random signals on any region of the scope is thereby reduced. Also, fewer unsynchronized replies are encountered when narrow-beam, scanning radars are used since the chance that several radars will be looking at a given beacon at the

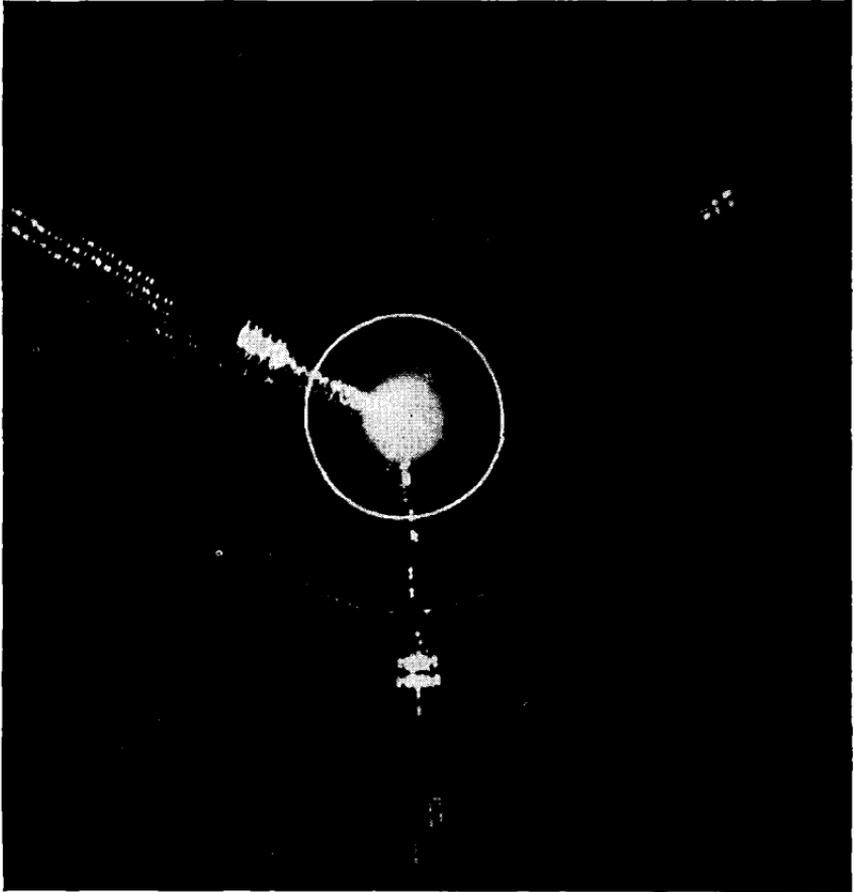


FIG. 8-12.—Four 3-cm beacons on the scope of a plane flying over England.

Beacon at 50°	(Beachy Head)
Beacon at 180°	(Defford)
Beacon at 280°	(Alconbury)
Beacon at 340°	(Halesworth)

The Alconbury beacon shows numerous unsynchronized replies resulting from interrogation by other radar sets.

same time is reduced. This trouble is mitigated (as is the one of over-interrogation of a particular beacon) by all arrangements that help in avoiding interrogation of beacons except when the beacon replies are wanted.

In systems that use crystal oscillators for accurate measurement of

range and incidental fixing of pulse repetition rate, the replies to interrogations from a given interrogator may appear as signals from a false beacon on the indicator of a second set because the two repetition rates are almost exactly equal. In two such systems used in the past war, it was found necessary to jitter the timing of the interrogation by whole numbers of periods of the crystal in somewhat random fashion in order to avoid this effect and the associated difficulty already mentioned—that of possible “stealing” of the beacon by one interrogator.

CHAPTER 9

ANTENNAS, SCANNERS, AND STABILIZATION

BY W. M. CADY, C. V. ROBINSON, F. B. LINCOLN, AND
F. J. MEHRINGER

The antenna is the sensory organ of the radar set. Its function is to accept r-f energy from the transmission line, to distribute this energy into space as desired, to gather in the radar echoes of this energy, and to direct these echoes back into the transmission line. Since the nature of electromagnetic radiation is such that a good transmitting antenna is also a good receiving antenna, we need consider only the former function. In most cases the antenna is required to form the energy into a sharp beam which may be aimed in various directions. The supporting structure is known as the "pedestal," and the entire assembly is termed the "antenna mount" or "scanner."¹ Airplanes and ships are notoriously unsteady vehicles, whose motions disturb the direction of a radar beam of energy transmitted from them. The compensation for such motions is called "stabilization."

9.1. The Antenna Equation.²—Two of the salient requirements of most radar sets are that they be able to reveal distant objects and be able to give accurately the direction of such objects. The design of the antenna has a great influence on the attainment of these requirements. We have seen (Sec. 2.1) that under certain conditions the maximum range of detection of a given target varies as the square root of the area of the antenna. This is one reason for the use of large antennas. Another advantage of large antennas, having to do with the resolution of the radar set, is that the beamwidth varies inversely as the linear dimension of the antenna. Mathematically the beamwidth θ (degrees) is usually related to the width D of the antenna and the wavelength λ of the radiation by the approximate formula

$$\theta \approx 70 \frac{\lambda}{D}, \quad (1)$$

if D and λ are measured in the same units. We see that shorter wavelengths make possible sharper beams; this accounts for the very con-

¹ For a fuller account of radar scanners the reader is referred to *Radar Scanners and Radomes*, Vol. 26 of this series.

² Secs. 9.1 to 9.12 by W. M. Cady.

spicuous trend toward shorter wavelengths which has characterized radar ever since its earliest development.

In the course of developing and testing a new antenna, the distribution of the energy in the beam may be expressed by means of a polar diagram or antenna pattern. Figure 9-1 displays such a pattern and

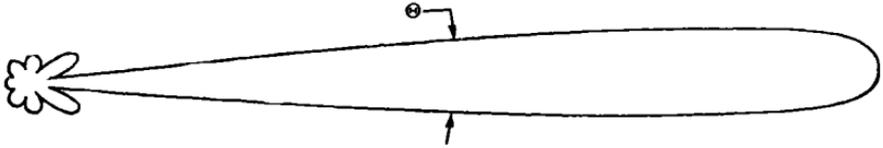


FIG. 9-1.—A typical antenna pattern showing the main beam of width Θ and the side lobes. shows that the beamwidth Θ is the full beamwidth at half power. A small amount of power is unavoidably radiated in undesired directions, forming the "side lobes" shown in this figure.

9 2. Round and Cut Paraboloid Antennas.—One very common type of microwave radar antenna takes the form of a paraboloidal reflector¹ with a source of radiation, or "antenna feed," at its focus. Since the

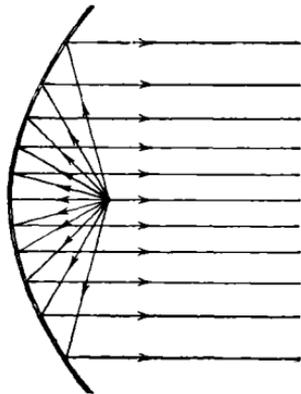


FIG. 9-2.—Ray diagram of a paraboloid antenna.

wavelength of the radiation is short in comparison with the dimensions of the reflector, it is profitable in an introductory discussion to regard the operation of the antenna as a problem in ray optics (Fig. 9-2). Figure 9-2 does not explain the side lobes or the beamwidth, but it does serve to emphasize the fact that the feed is always so designed as to be directive. Figure 9-3 represents an automatic record of the energy radiated at various angles from a well-designed feed, the directional power being displayed as a function of angle. It is obvious that nearly all of the energy from the feed will strike a properly placed paraboloid, which will then collimate the radiation. Figure 9-4

shows the radiation from an incorrectly designed feed. A correct design, used in the 3-cm band, is illustrated in Fig. 9-5, which shows two dipoles excited by the radiation field from the $\frac{1}{2}$ -by 1-in. waveguide transmission line. The dipoles are so adjusted in length and position that their combined effect is to direct the radiation back within a cone surrounding the line without reflecting radiation into the waveguide.

Equation (1) requires further comment in the case of a paraboloid reflector of which the perimeter has been so trimmed ("cut paraboloid") that its width and height are unequal. If the beam is pointed horizon-

¹ The subject of paraboloid and other types of antennas is developed in *Microwave Antenna Theory and Design*, Vol. 12 of this series.

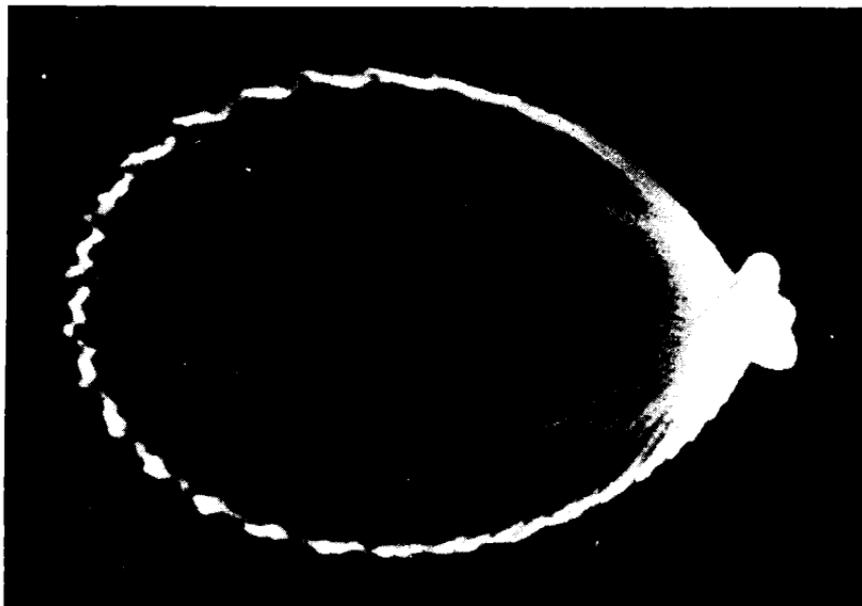


FIG. 9-3.—The angular distribution of energy from a two-dipole feed.

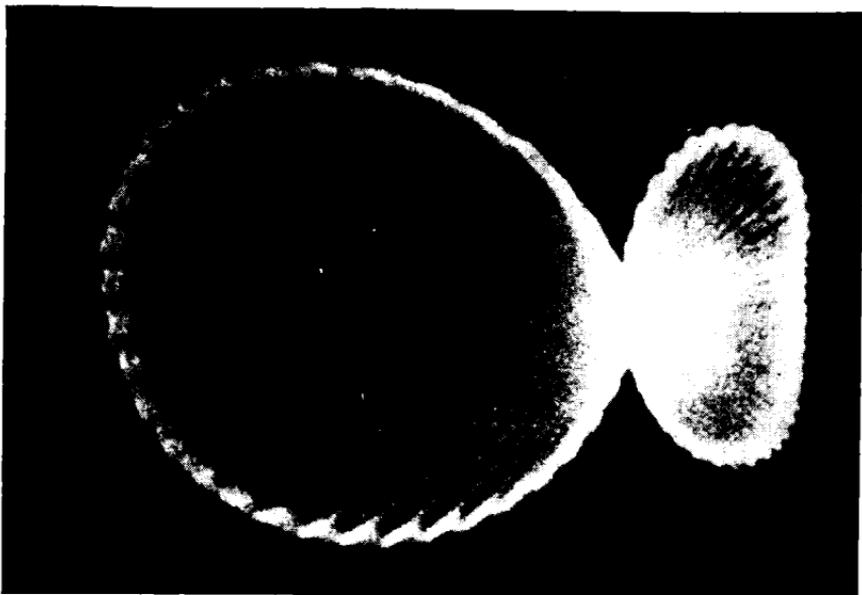


FIG. 9-4.—The angular distribution of energy from a two-dipole feed incorrectly designed.

tally, the beamwidth can be measured by exploring the intensity of the radiation to the right and left of the center of the beam or, alternatively, above and below the center of the beam. It is an important fact that the beamwidth as measured vertically depends, according to Eq. (1), upon the vertical dimension of the cut paraboloid, and the horizontal beamwidth depends upon the horizontal dimension. Thus, in order to realize high resolution in azimuth, the antenna must be wide but not necessarily tall, whereas a height-finding radar, affording accurate measurement of the angular elevation of an airplane, must have a tall antenna which need not be wide. These two types are exemplified by the scanners shown in Figs. 9-15 and 9-14 respectively.

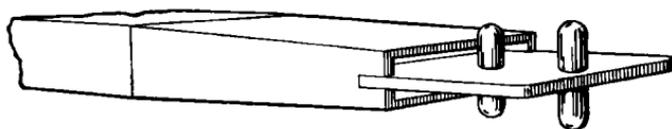


FIG. 9-5.—Two-dipole feed radiating toward the left.

We have not mentioned the "mattress" type of antenna, which is perhaps the most familiar to the public because it is commonly seen on naval ships. These antennas are rectangular arrays of radiating elements so phased as to produce a broadside beam. They are widely used in radars of wavelength over 1 m. With the trend toward shorter wavelength, the mattress antennas are giving way to the types discussed above.

9-3. Fan Beams.—One of the functions for which airborne radars are designed is aiding air navigation, and an extremely important type of airborne navigational radar presents a map of the terrain around the aircraft. In order to map the ground the transmitted energy must be directed toward the ground, and rather than having the energy beamed like a searchlight, a "fan beam" must be employed as shown in Fig. 9-6 for complete coverage. A uniformly intense map is desired, which will display ground objects that lie at depression angles between, for example, 5° and 70° below the horizontal. The energy in the fan beam must therefore be properly distributed to give adequate illumination of the most distant objects while not overilluminating those which are at greater depression angles. It can be shown, subject to certain simplifying assumptions, that the energy should be distributed as the square of the cosecant of the depression angle (see Sec. 2-5). Figure 9-7 shows how well such a pattern is realized in practice, with an antenna 12 in. in height (shown in Fig. 9-8) used at the 3-cm band. In this polar diagram the square root of the power is plotted rather than the power itself, since the desired distribution is then given by a straight line as shown. Figure 9-9 is a photograph of the PPI display, showing the performance of this antenna.

The cosecant-squared fan beam is used not only in airborne navigational radar but also in surface-based radar for the detection of airplanes. In this case the fan is inverted, the intention being that a target airplane

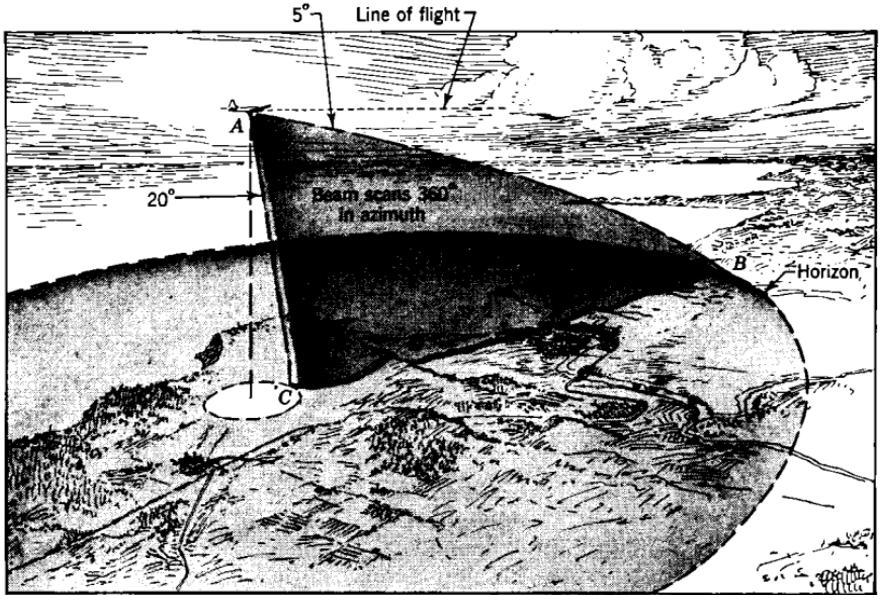


FIG. 9-6.—A fan beam which enables an airborne radar to scan the surface of the earth.

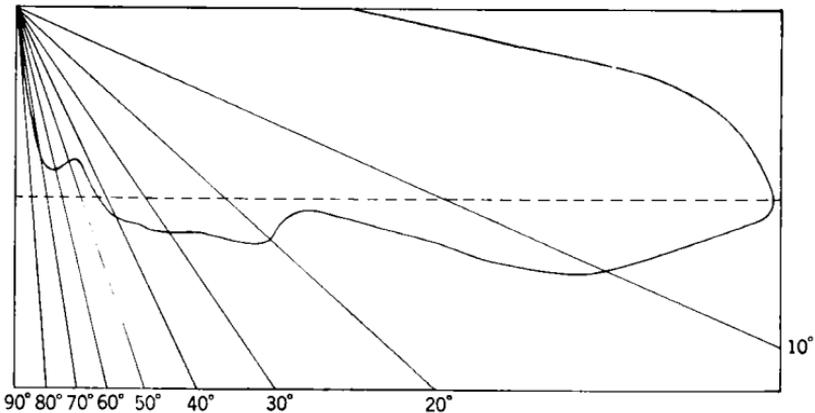


FIG. 9-7.—The energy distribution to a square-root scale as a function of the depression angle in an airborne navigational radar.

flying at any constant altitude will be indicated with a constant intensity, independent of range.

Several antenna types have been developed for producing a cosecant-squared fan beam; those most widely used can be discussed in terms of

ray optics. Figure 9-10 shows two distortions of a paraboloid used for this purpose. Another widely used method employs a *cylindrical reflector*, shaped like the blade of a snowplow or bulldozer, and having only single curvature (Fig. 9-8). Such a reflector is illuminated by the radiation from a linear feed rather than a point feed. In this case the linear feed defines the beam sharpness in regard to azimuth and the reflector distributes the energy in elevation. The ray diagram shown in Fig. 9-10a will serve again to illustrate the formation of the cosecant-squared fan.

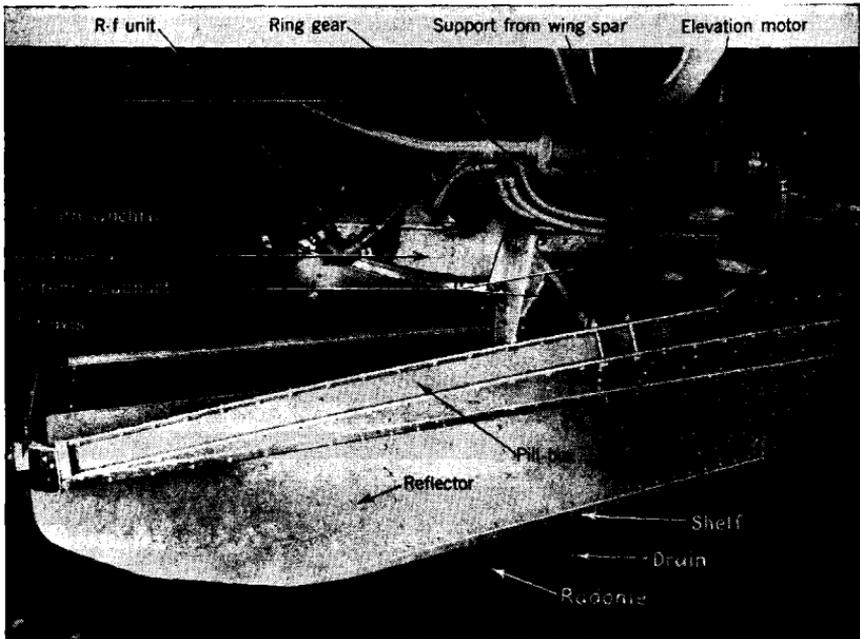


FIG. 9-8.—A scanner with a 60-in. shaped cylindrical reflector, ventrally installed in a B-29.

In this figure the linear feed is imagined perpendicular to the plane of the drawing. The feed is frequently the straight opening between two identical parallel conducting sheets which are joined along their curved backs by a parabolic reflector strip. Just inside the center of the opening, at the focus of the parabolic strip, is located the open end of a waveguide which irradiates the parabola, thereby setting up the propagation of energy between the two sheets and straight out through the opening. This assembly is called a "pillbox." Microwaves may be propagated between the sheets with the electric field polarized either parallel or perpendicular to the sheets. In parallel polarization the spacing must be held accurately in order to prevent distortions of the wavefronts; slender spacing posts are used since they have almost no effect on the waves. In

perpendicular polarization, posts scatter the radiation badly and are not generally used; but fortunately the wavefronts are not disturbed by gentle variations in the spacing of the sheets. Both polarizations are used in practice.

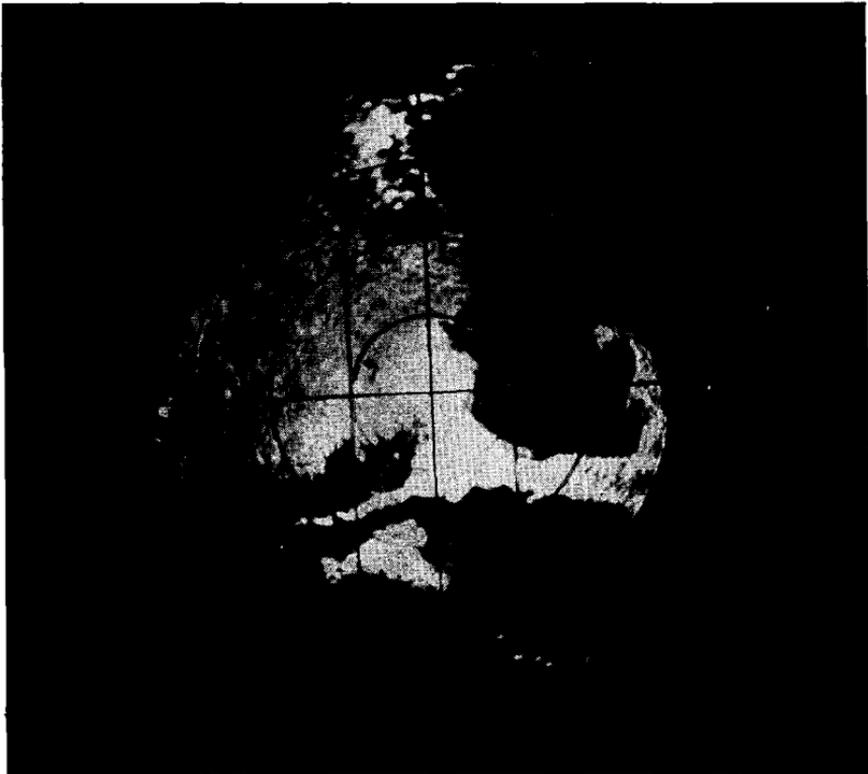


FIG. 9-9.— Photograph of PPI scope of 3-cm airborne radar: Cape Cod area.

Another type of linear source of radiation is the linear array of dipoles. If these are so excited as to emit energy in phase, the resulting radiation is a fan in the broadside direction. Further discussion of line sources of radiation will be found in Secs. 9-12 and 9-14.

9-4. Nonscanning Antennas.—There are certain types of radar antenna which are not ordinarily required to scan. In this category are the end-fire antennas which are sometimes attached to airborne guns in order to permit radar range-finding. As an example, one end-fire linear array (Fig. 9-11) is a series of dipoles excited in such a phase relationship that their energy is beamed in the direction of the array. Another end-fire array is the Yagi antenna. Commonly, this is an array of dipoles, only one of which is excited by the transmission line; behind it is a parasitic *reflector* dipole and in front are several parasitic *director* dipoles. The

Yagi array has been widely used for wavelengths of several decimeters; an example is shown in Fig. 6-30. An end-fire antenna can also consist of a *polyrod* (a rod of polystyrene) of appropriate dimensions, into one end

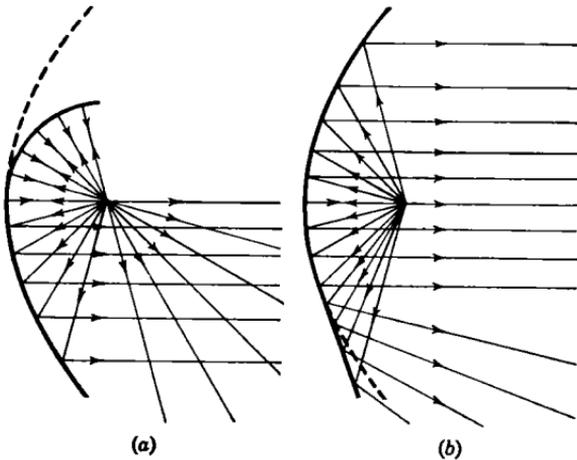


FIG. 9-10.—A paraboloid reflector may be distorted in various ways in order to produce a fan beam.

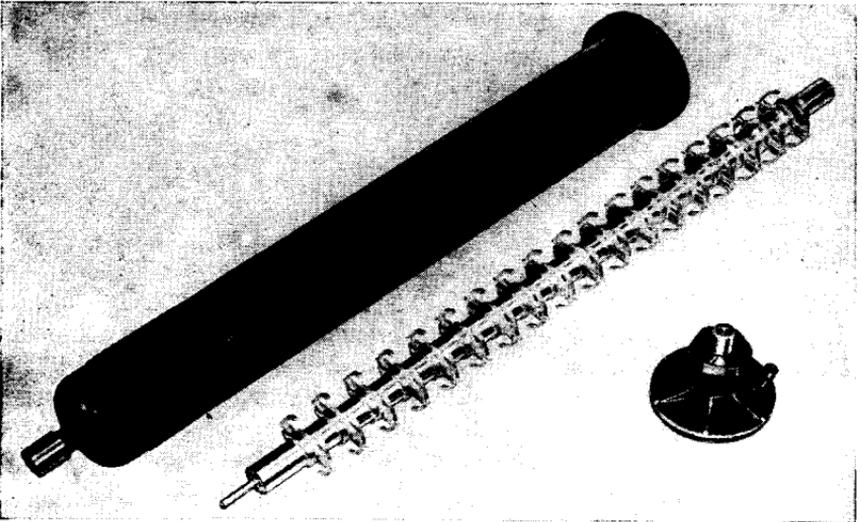


FIG. 9-11.—An end-fire array and its housing for use in the 10-cm band.

of which r-f energy is introduced. These end-fire antennas are useful in cases where a paraboloid would obstruct vision or add unduly to the aerodynamic drag of an airplane. An array of 42 polyrods, placed parallel and excited in phase, has been used as a broadside array as shown later in Fig. 9-32.

Beacon antennas are often made up of a vertical array of radiating elements excited in phase. By correct design of such an antenna the pattern can be made reasonably uniform in azimuth and at the same time well confined to the region of the horizon. Like other beacon antennas, the antennas used for IFF equipment are not strictly radar equipment and will not be discussed. The same statement applies to the antennas used in the search and jamming functions of radar countermeasures.

9-5. Construction of Radar Antennas.—The main requirements in the mechanical construction of airborne antennas are accuracy of form, the ability to withstand field service conditions, and light weight. Surface-based antennas must be no less accurate, and considerations of weight, inertia, and wind forces are paramount. Considerable advances have been made in the mechanical design of airborne, and particularly of surface-based, antennas for microwave radar during the recent war. In this section only antennas in which the feed is distinct from the reflector are considered; array antennas are not discussed.

Sheet aluminum or magnesium alloy is always used for the airborne reflectors. Small reflectors up to 30 in. in diameter are simply spun or otherwise formed, a bead around the rim being added for stiffness, as in Fig. 9-17. The tolerance allowable in the surface of the reflector is about $\pm \frac{1}{32}$ of a wavelength.

The larger sizes of airborne reflectors, whether paraboloid or shaped cylindrical surfaces, thus far installed in a streamlined airplane are considerably wider than high and are used for circular scanning. On these it is good practice to use aircraft construction methods. The selection of materials used depends upon the size of the reflector and the stresses involved. Magnesium or aluminum alloy will serve for airborne equipment. Since aluminum alloy has better forming characteristics than magnesium, it is considered the best material for the reflector face. The most accurate method of forming aluminum alloy sheet for the reflector face is by stretching the material over a metal die which has been cast and ground to the desired contour; these dies are usually made from Kirksite, a lead and zinc alloy. After forming, the sheet metal conforms to the exact contour of the die providing the draw is shallow enough, as is the case with most radar reflectors. The next problem is to support this reflector face rigidly and to maintain the contour already obtained by forming. This can be done, as in the 5-ft reflector of AN/APQ-13 shown in Fig. 9-8, by riveting to the rear surface of the reflector several stamped or machined ribs contoured thereto and boxing in by riveting another sheet of aluminum or magnesium alloy across the back; flush rivets are preferable on the front face. A 42- by 10-in. reflector with single curvature can be made within a tolerance of ± 0.005 in.

In the case of a large reflector, over 5 ft in width, weight may be saved

by using sheet-metal ribs or channels formed by the hydropress method, and boxed in at the rear as in Fig. 9-8, or by using a stamped outer frame made by the drop-hammer method, as in the 8-ft reflector shown in Fig. 9-12. These waffle-shaped frames require no ribs for stiffening because the mating section conforms to the contour of the back of the reflector face. An aluminum-alloy supporting bracket can be riveted to

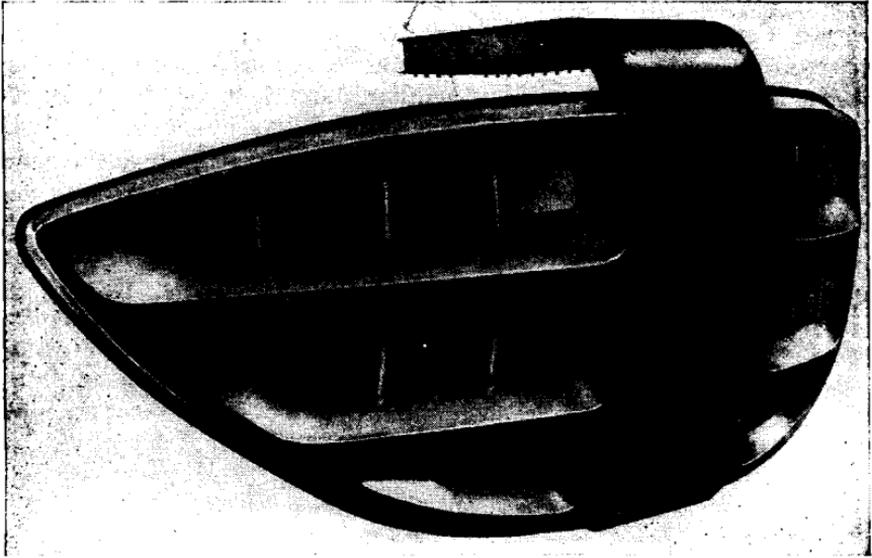


FIG. 9-12.—Rear view showing the construction of an 8-ft airborne paraboloid trimmed to a 3-ft height.

the structure; or, if the elevation angle of the beam must be adjustable, the ends of the structure may be supported by ball-bearing trunnions.

Surface-based antennas are usually not housed, and therefore are subject to wind forces. These forces can be reduced by as much as 50 per cent if the reflector is a grille (Figs. 9-14 and 9-28) or a perforated sheet (Fig. 9-13). Fortunately such perforation does not impair the electrical performance if the holes are not larger than about one-eighth wavelength in the direction of the magnetic vector.

RADAR SCANNING PATTERNS

We have seen that although a few radars have been designed with radiation beams that are fixed in direction, the great majority of radar beams are made to scan. The motion of the beam may be thought of as the motion of a point on a sphere that is centered at the scanner. The beam sweeps over a certain region on the sphere determined by the operational function of the radar; often it does so periodically, and in a set geometrical pattern termed the "scan." The great variety of scans may be divided into two categories, simple and complex. A simple scan is one

in which the beam sweeps with but one degree of freedom, that is, it covers repeatedly one and the same arc on the sphere. A radar employing a simple scan can tell the range of a target and only one of its angular coordinates, e.g., azimuth. In a complex scan the beam ranges over a certain solid angle, by virtue of possessing two degrees of freedom. A radar possessing a complex scan can tell the location of a target in space by giving the range and two angular coordinates, e.g., elevation and azimuth.

9-6. Simple Scans.—One of the simple scans in common use is the circular (or “horizon,” or “all around looking,” or “360°,” or “A”) scan; the beam travels continuously around the horizon, or may be adjusted to scan around at any constant angle above or below the horizontal. This scan is widely used in radars providing surface-based surveillance of land, vessels, and distant aircraft; and it is used (see Fig. 9-6) in airborne radar, serving, for instance, as an aid to air navigation. In certain scanners the beam can be adjusted to a position above or below the horizon at the will of the operator. The scan rate is usually in the range between 4 rpm for large scanners and 30 rpm for small. Sector (or “B”) scan is a modification of circular scan in which the beam scans to and fro on an arc. This motion is common in certain airborne radars for surface search. The sector is typically 75° wide, and 1 to-and-fro cycle may occupy 1 sec. The third simple motion of the beam is conical scan, in which the path described on the sphere is a circle of a very few degrees diameter. This scan is not used for search, but finds wide use in accurate tracking of an individual target (see Sec. 6-14). The diameter of the circle described by the center of the beam is chosen with this application in mind; it is commonly such that the intensity radiated to the center of the circle is somewhat greater than half the peak intensity. The scan rate is usually at least 1200 rpm.

9-7. Complex Scans.—A radar having spiral scan was used in single-seat nightfighter aircraft during World War II. This complex scan may be described as a conical scan in which the angular diameter of the circle described by the beam is continuously varied from 0° to a maximum of, for example, 60° and back to 0°. By this motion the solid angle covered is scanned completely in, say, 1 sec as the beam spirals outward.

A helically scanning radar was used in nightfighters that carried a radar operator in addition to the pilot. The beam revolves rapidly around a vertical axis, as in horizon scan, while the elevation angle is made to oscillate much more slowly between limits a few degrees above and below the horizon. In this way targets may be sought within a horizontal 360° zone except for whatever blanking is caused by the structure of the nightfighter itself. The time required to explore this zone is approximately 3 sec.

A marriage of the horizon scan (or, alternatively, the sector scan)

with the conical scan produces the Palmer scan, the cycloidal motion of the beam which results when a conical scanner slowly traverses the horizon. The region of search is a horizontal rectangle with semicircular ends. Palmer scan was used to accomplish an easy transition from the search function to the gunlaying function of a fire-control radar. The name derives from a familiar exercise in the Palmer system of calligraphy.

Other complex scans have been used mainly in experimental equipment. Probably the most useful of these is a scan in which the elevation (or azimuth) angle oscillates rapidly while the azimuth (or elevation) angle oscillates slowly.

MECHANICAL SCANNERS

In most cases the entire radar antenna is put through certain angular motions in order to make the beam scan. The antenna assembly, including the mount that supports it and makes it move, is called a "mechanical scanner." The term "electrical scanner" is reserved for cases where the beam is moved not by a motion of the antenna as a whole, but rather by relatively subtle motions of the feed or other parts of the antenna. The borderline between mechanical and electrical scanning is not well defined. For instance, in denial of the foregoing definitions, the mechanical category includes the conical scans that are mechanized by the circular motion of a point feed around the focus of a fixed paraboloid. An electrical scan is frequently used to produce a rapid sector scan of small angular amplitude.

9-8. The Kinematics of Mechanical Scanners.—The intended use of a radar fixes the type of scan, and the type of scan fixes in turn the kinematic aspects of the scanner design. A circular scan is mechanized by simply revolving the antenna on a horizontal turntable. If it is desired to raise or lower the beam, the turntable must carry bearings so that the antenna may be tilted about a horizontal axis at right angles to the beam; the scanner may be described as a tilting antenna on a rotating assembly on a fixed base. Airborne scanners of this type are usually mounted on the under side of the fuselage, the antenna being below the pedestal. Sector scanners are kinematically similar to circular scanners.

To enable the radar operator to control the antenna from a distance the tilting is usually actuated by a motor. This motor may be located on the turntable (in which case it must be powered through slip rings) or else on the base (in which case a special mechanism must be provided for transferring the mechanical power from the tilt motor to the revolving antenna). Both methods are used in practice although only the former is encountered in surface radars.

Conical scan may be effected by spinning the antenna about an axis not quite parallel to the beam (AN/APS-6, Fig. 9-17); or by spinning the

paraboloid about an axis passing through the feed, the paraboloid being a few degrees "drunk" in relation to this axis (AN/APG-15); or by spinning an electrically asymmetrical feed about the axis of the fixed paraboloid (SCR-584, Fig. 9-13). All these methods and others are in wide use.

Of the complex motions, the spiral scan has been mentioned as derived from conical scan, and more particularly from three varieties just listed. This is exemplified in the first two cases by the Navy's AN/APS-6 nightfighter radar and the British AI Mark VIII. Helical scan is used in the SCR-720 Army AI radar and in the SCR-584 anti-aircraft set. In the latter a conical scan is superimposed so that the result is a sort of helical Palmer scan.

9-9. The Weight of Mechanical Scanners.—The design of a mechanical scanner is largely dependent on its antenna. Although only a small percentage of the scanner weight is attributable to the antenna, a large antenna necessitates a heavy scanner. Airborne scanners show this relation very strongly. A survey of weights¹ indicates that, very roughly, the weight of scanners having simple scans is given by $0.09D^2$ pounds, where D is the paraboloid diameter in inches. The formula for complex scans is $0.13D^2$. The survey shows that among airborne radars with mechanical scanners the scanner weighs 13 to 106 lb, representing from 6 to 21 per cent of the weight of the set, with an average of about 14 per cent. Surface-based scanners are in general larger than airborne scanners. The weight runs from 75 to 5500 lb in shipborne antenna mounts, and as high as 28,000 lb for land-based. The antenna mounts of surface-based radars represent 10 to 40 per cent of the total weight of the set.

9-10. R-f Transmission Lines.—The r-f energy generated at the transmitter is radiated by the antenna, and the echo is led back from the same antenna to the mixer. The same transmission line² is used for both the transmitted and the received energy. A rotary joint must be inserted in the line whenever more than a few degrees of rotation of an antenna are required. Such a joint is always installed with its axis coinciding with the axis of the corresponding degree of freedom. Pressurization of the rotary joints, when necessary, is accomplished either by means of a composition rubber ring revolving snugly around a polished steel tube, or by means of a polished carbon ring revolving in contact with a polished steel annulus. Helical scanners have at least two rotary joints, i.e., for azimuth and elevation, as have spiral scanners.

9-11. Data Transmission.—The continuous transmission of information from the scanner to the indicator, giving the attitude of the antenna in relation to its base, is known as "data transmission." Often these

¹ W. L. Myers, USNR, "Weight Analysis of Airborne Radar Sets," RL Report No. 450, Jan. 1, 1945.

² Chap. 11 contains a discussion of the transmission line and associated components.

data are transmitted as voltages induced in synchros.¹ There is, for instance, a synchro generator built into the AN/APQ-13 scanner shown in Fig. 9-8 which revolves 10 times for each azimuth revolution of the antenna; a synchro motor in the indicator imitates the motion of the generator, and through a gear reduction rotates the deflection coils of the PPI in synchronism with the antenna. Occasionally a potentiometer or a sine-wave generator is used as the data-transmitting element. A similar type of data transmission is required of servo-controlled scanners: a voltage from the transmitting element is compared with a voltage from the knob, etc., that is to control the antenna, and the antenna is automatically driven in the direction which will bring the error signal or difference voltage to zero.

9-12. Examples of Mechanical Scanners.—The above generalities will be made more concrete in the description of a few specific scanners. These are chosen to illustrate a variety of purposes, shapes, and sizes.

The SCR-584.—The antenna mount of the SCR-584 (Fig. 9-13) has been cited as a Palmer scanner, inasmuch as it combines helical with conical scan. The prime function of this radar,² located in a trailer, is to locate enemy aircraft and to provide an antiaircraft director with data on slant range, azimuth, and elevation angle of aircraft. The beam (at the 10-cm band, about 4° wide) is formed by a 6-ft stamped paraboloid illuminated by an electrically offset dipole. This 120-lb reflector is perforated with 6400 half-inch holes to reduce weight and windage. A gear train in the base actuates the searching in azimuth at 6 rpm. A similar gear train is located behind the reflector and actuates the elevation motion; the excursion in elevation is covered once per minute. To enable tracking, the electrically asymmetrical dipole feed spins at 1800 rpm, thus scanning the beam conically; this has the incidental effect of gyrating the plane of polarization of the energy. Five units, driven by a spring-loaded gear train, are provided for azimuth data transmission: four synchros, of which one is geared to rotate at 16-speed, and one potentiometer. As is common practice in data transmission, each take-off may be individually adjusted by loosening its mounting clamps and rotating the stator case for proper angular setting in relation to the antenna. After the eventual removal of the SCR-584 radar to a new site, any necessary reorientation of the take-offs may be effected by an ingenious "group adjustment" of the data gear train. Units analogous to all but one of the azimuth synchros are geared to the elevation mechanism. The phase of the conical scan is continuously signaled by means

¹ Potentiometers and synchros are discussed in Chaps. 8 and 10 of *Components Handbook*, Vol. 17 of this series.

² War Department Technical Manual TM-1524.

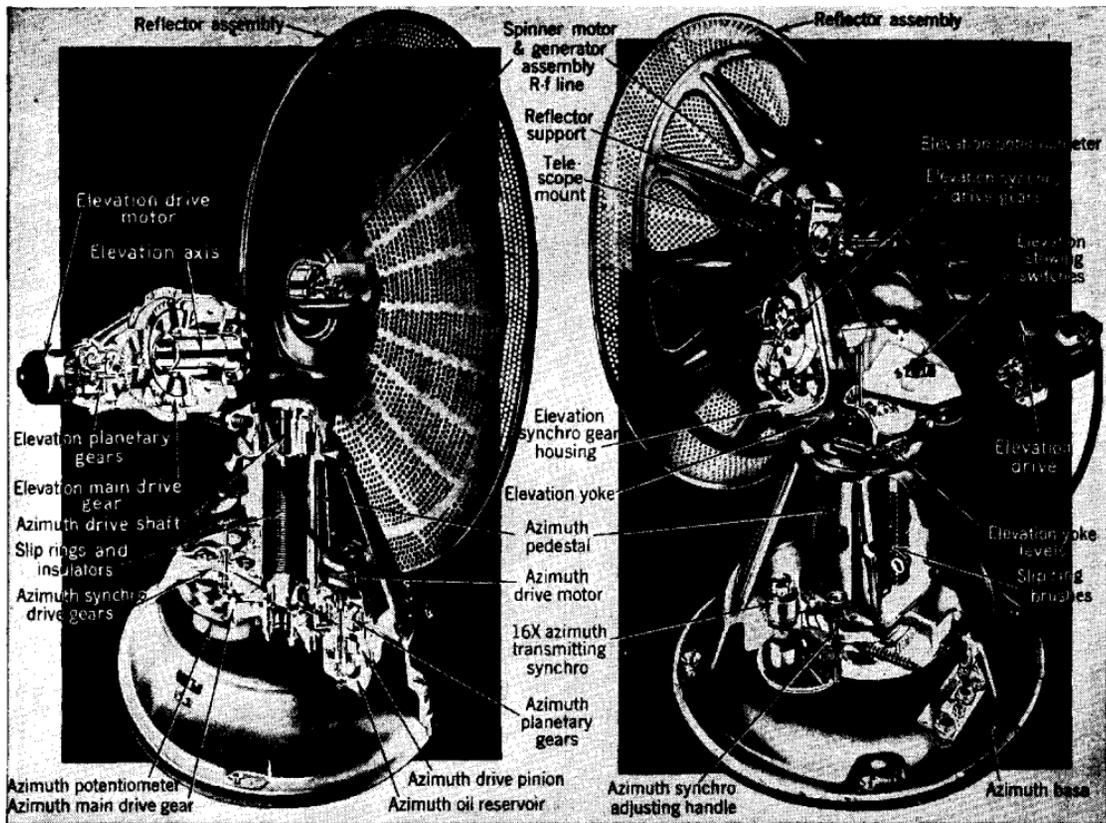


FIG. 9-13.—The SCR-584 anti-aircraft antenna mount with 6-ft reflector. (Courtesy of Chrysler Corporation.)

of a two-phase sine-wave generator whose armature rotates with the spinning antenna feed.

In order to bring power to the elevation and spin motors and to bring the elevation and spin data down to the indicator and servo circuits, 48 slip rings are provided concentric with the azimuth axis. The transmission line is coaxial. It runs from the transmitter up through the hollow azimuth shaft, in which there is an r-f rotary joint; thence into the hollow

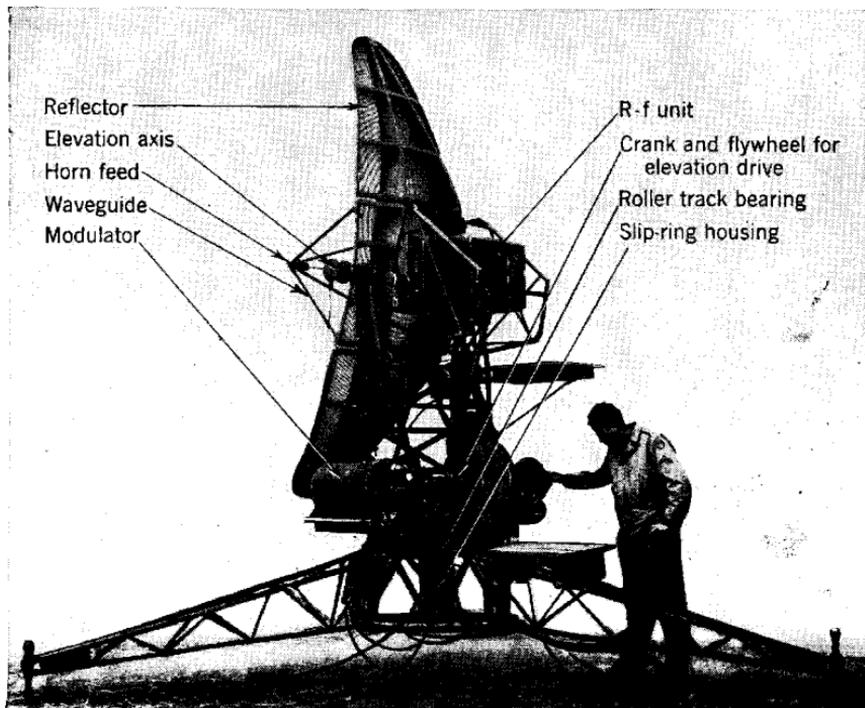


FIG. 9-14.—The AN/TPS-10 height-finding antenna mount with a reflector 10 ft high.

elevation shaft, where a second rotary joint is provided; and then into the reflector at its vertex, where it passes through the hollow shaft of the sine generator, a high-speed rotary joint, and the hollow shaft of the spin motor. In order to exclude moisture, the entire line, including the three rotary joints and the feed, is pressurized to about 5 lb/in.² The mount, weighing some 2200 lb, is installed on an elevator in the trailer and so may be raised for use or lowered into the trailer for stowage.

The AN/TPS-10.—A second antenna mount of interest is that of AN/TPS-10¹ (Fig. 9-14), a radar intended for detecting airplanes and determining their altitude, especially in mountainous terrain. Although

¹ War Department Technical Manual TM 11-1568.

it may be mounted and used on two trucks, this set is more commonly based on the ground. As mobility requires, the mount is easily assembled and disassembled in the field, and only four of its parts exceed 80 lb in weight.

The antenna, a paraboloid made of a grid of curved $\frac{1}{4}$ -in. tubes, is trimmed to an oval contour 10 ft high by 3 ft wide. The beam is therefore of a "beavertail" shape and since the radiation is at the 3-cm band, the beam is 2.3° wide in azimuth and 0.7° wide in elevation. The polarization is vertical. In the height-finding function the beam scans in elevation between 2° below and 23° above the horizontal at a manually controlled azimuth. For search purposes the azimuth motion may be motor-driven at 4° per sec, either boxing the compass or scanning a sector. This azimuth rate cannot be exceeded if successive sweeps in elevation are to cover without gaps all portions of the zone being scanned. The "oscillating beavertail" scan which results is, however, rather slow for search, and this function is sometimes relegated to another set with which the AN/TPS-10 may form a team.

The main bearing of the mount is a roller-ring bearing, defining the vertical axis of the pedestal. On the turntable are mounted the antenna and the elevation drive, consisting of motor, gear reduction, crank, and connecting rod. The turntable also supports the pressurized modulator and its controls, the pressurized r-f head, and the power supply for the receiver. The r-f head is so mounted in order to eliminate the need for an azimuth rotary joint, and the other components named are so mounted in order to reduce the number of slip rings on the azimuth axis.

Long Range Ground Radar. The system now to be described was designed for microwave early warning and surveillance of enemy aircraft and the control of friendly aircraft. Two antennas (Fig. 9-15), usually back-to-back on a single mount, characterize this set. One is for long-range low-angle coverage and the other, radiating a fan beam, covers higher elevations. The set can detect single heavy bombers to above 30,000-ft altitude and 200-mile range, provided the aircraft are above the horizon. The antennas are similar in that each has a shaped cylindrical reflector 25 ft long, in 11 sections, fed by a linear array of dipoles at the 10-cm band and forming a beam 0.9° wide in azimuth. The waveguide transmission line is weatherproof, the array being housed in a Plexiglas cover. The antennas are different in that the low-angle reflector is 8 ft high and parabolic in vertical section, its radiation vertically polarized; whereas the high-angle reflector is 5 ft high, of an empirically determined shape producing an approximately cosecant-squared beam, and its radiation horizontally polarized.

The reflectors are not perforated for the reason that the small amount

of energy which would penetrate straight through would still be focused in azimuth and so would cause spurious signals. A *solid* construction is therefore typical of shaped cylindrical reflectors. Paraboloid reflectors need not be solid, since the radiation leaking through them from the point antenna feed is not focused and cannot cause sharply defined false echoes.

The antennas rotate about a vertical axis at 2 or 4 rpm, together with their respective transmitters and common modulator. There is no ele-

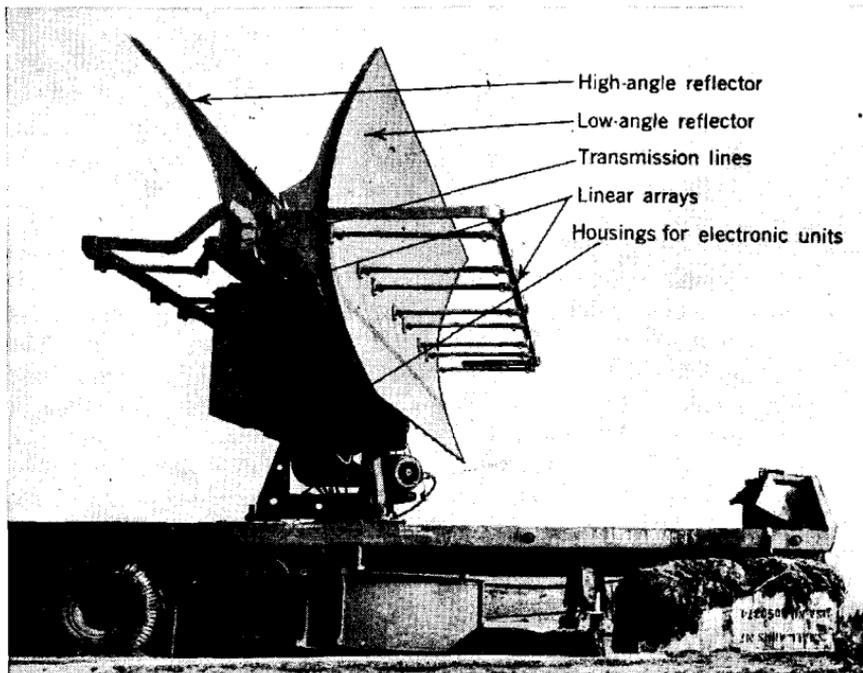


Fig. 9-15.—Long-range microwave antenna mount with two 25-ft cylindrical reflectors.

vation control of the beam except for initial adjustment at the time of assembling in the field. No r-f rotary joints are necessary in the waveguide transmission lines. Slip rings are required around the azimuth axis in order to carry 60-cps power, control voltages, and video signals.

Although weighing 6 tons, the ground-based version of this mount can be disassembled and stowed on trucks in 4 hr. The truck-mounted version is illustrated. The mount has operated in winds up to 80 mph, and the design is such that 125-mph winds should not cause damage.

The AN/APS-10.—In spectacular contrast with the foregoing is the more recent of two alternative scanner designs for the AN/APS-10 navigational radar (Fig. 9-16); it is airborne and weighs about 13 lb. The antenna is a paraboloid 18 in. in diameter, fed at the 3-cm band by a

horizontally polarized dipole feed (cf. Fig. 9-5) at its focus. The beam-width in azimuth is about 5° . The lower portion of the paraboloid is distorted as in Fig. 9-10*b*, so as to reflect a portion of the energy downward

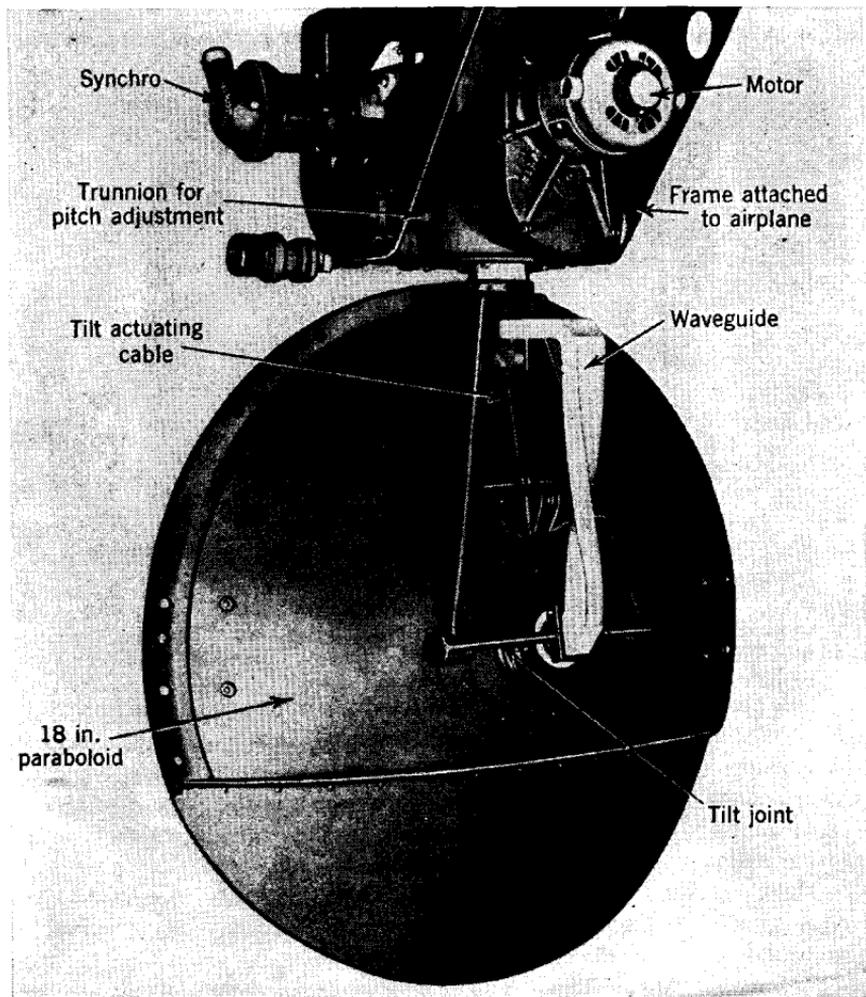


Fig. 9-16.—The AN/APS-10 airborne navigation scanner with an 18-in. reflector. (Courtesy Houston Corporation.)

to illuminate the foreground. This fan beam is approximately cosecant-squared between 5° and 30° depression angle when the antenna is at its normal attitude. In order to raise or lower the beam the radar operator may tilt the antenna up or down a few degrees at will by energizing a small motor, contained within the scanner base, which pulls a light steel cable attached to the reflector. The waveguide transmission line is

pressurized all the way from the transmitter-receiver unit through a flexible section to the scanner, the azimuth rotary joint, and the tilt joint to the feed, which is enclosed in a hollow Styrallyloy ball. Parts of the waveguide are made of nickel electrodeposited on a die-cast form which is subsequently melted out. The scan is circular, at 30 rpm.

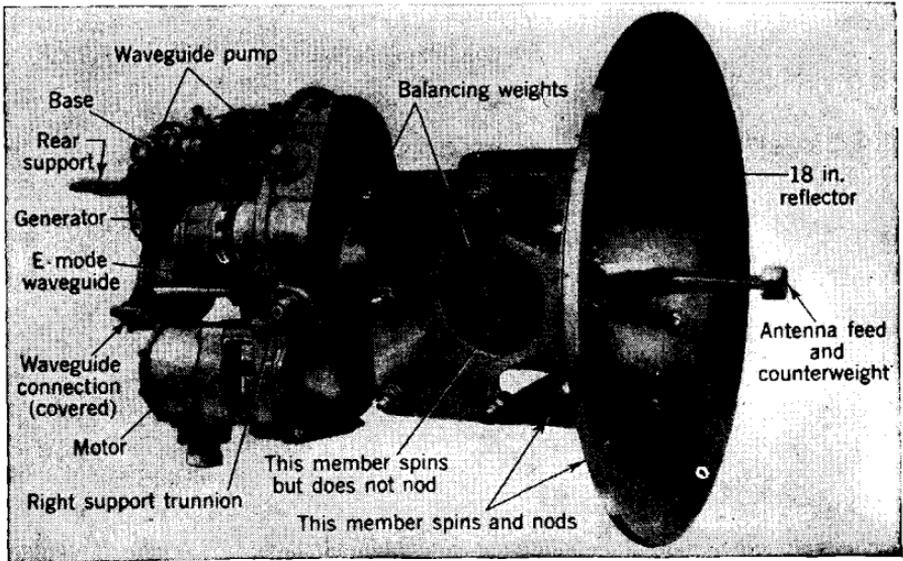


FIG. 9-17.—The AN/APS-6 spiral scanner. (Courtesy of Dalmo Victor Manufacturing Company.)

This scanner has an unusual feature which allows the main axis to be adjusted to a vertical attitude despite changes in the angle of attack of the airplane. This pitch adjustment is electrically driven and remotely controlled by hand. A second unusual feature is the nature of the r-f elevation joint, which allows the beam to be raised or depressed from its normal attitude. This is merely a choke-to-flange coupling, similar to the rigid coupling commonly employed (see Sec. 11-3), except that the flange is slightly beveled so as to allow $\pm 9^\circ$ relative motion. This "wobble joint" is enclosed in a flexible airtight tube.

The AN/APS-6.—The AN/APS-6 system imposes unusual requirements on the scanner (Fig. 9-17). This radar is used for airborne detection of aircraft under blind conditions, and therefore requires a search over a solid angle in the forward direction. The beamwidth is about 5° .

The scan is spiral, and one turn of the spiral is described in $\frac{1}{20}$ sec, which causes the plane of polarization to gyrate at this speed. The beam is made to spiral outward from 0° (straight ahead) to 60° and back again in 2 sec by the nodding of the antenna in relation to the yoke which forms the forward end of the horizontal main shaft. By throwing a switch the

operator can halt the nodding of the antenna, which then executes a conical scan to permit accurate homing. A single motor, rated at 600 watts mechanical output, provides power for the nod and spin motions. The data take-offs are a 2-phase sine-wave generator for the spin angle and a potentiometer for the nod angle, both being mounted on the main gear case to obviate the need of slip rings. The gear case is unusual in airborne practice in that it is oil-filled. Two r-f rotary joints are required, one on the spin axis and one on the nod axis.

A difficult problem in dynamic balance is presented by this scanner. The difficulty arises from the fact that the fast (1200-rpm) rotor must be in good balance for all attitudes of the antenna which is supported thereon. In order to effect balance a special procedure has been evolved whereby the correct number of standard lead weights may be attached to each of eighteen points provided on the antenna and the main shaft. A special type of balancing machine has been produced for this purpose.

ELECTRICAL SCANNERS

BY C. V. ROBINSON

No radically new principles are embodied in the mechanical scanners discussed in the previous section. By contrast, the electrical scanners to be discussed in the following sections have many novel features of fundamental design and so merit a more thorough discussion.

9-13. The AN/APQ-7 (Eagle) Scanner.—The AN/APQ-7 (Eagle) high-resolution navigation and bombing equipment was developed at Radiation Laboratory; Bell Telephone Laboratories carried through the production engineering.¹

The antenna developed for this equipment produces at the 3-cm band a horizontally polarized beam of radiation which has a width of 0.4° to 0.5° in azimuth and is shaped in elevation to give an approximately cosecant-squared coverage down to 70° angle of depression. The beam scans a 60° azimuth sector in $\frac{2}{3}$ sec, or 1 to-and-fro cycle in $1\frac{1}{3}$ sec. The range of the Eagle system for ground-mapping is about 50 miles.

The antenna is a slender linear array 16 ft long, which fits into a streamlined housing (vane) hung laterally beneath the aircraft that uses it (Fig. 9-18). The body of the array consists of a fixed 16-ft extruded aluminum channel and a movable 16-ft aluminum plate which together form a guide of variable width (Fig. 9-19). A row of 250 dipole radiators is mounted in the fixed channel in such a way as to draw power from the guide formed by the two long aluminum members. A mechanism varies

¹L. W. Alvarez, "Microwave Linear Radiators," RL Report No. 366, June 30, 1942; R. M. Robertson, "Variable Width Waveguide Scanners for Eagle and GCA," RL Report No. 840, April 30, 1946.

the guide width (and thereby the wavelength in the guide)¹ periodically; this, in turn, varies the effective dipole phasing and moves the beam in azimuth.

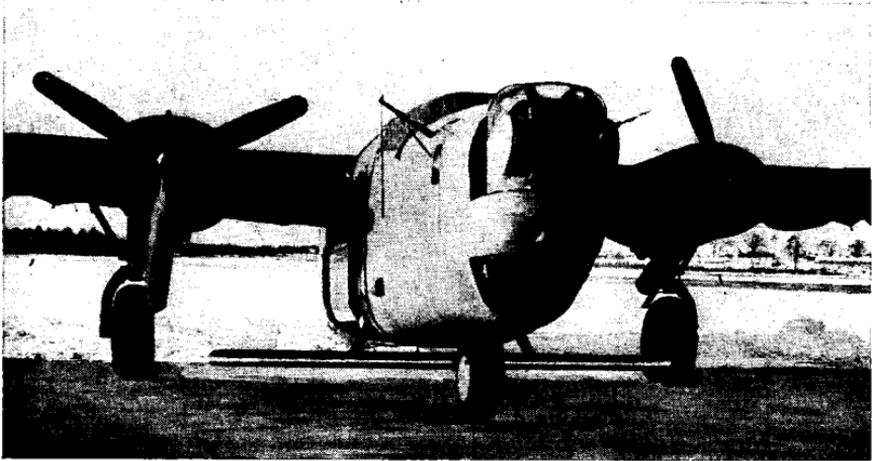


FIG. 9-18.—AN/APQ-7 antenna housing on a B-24.

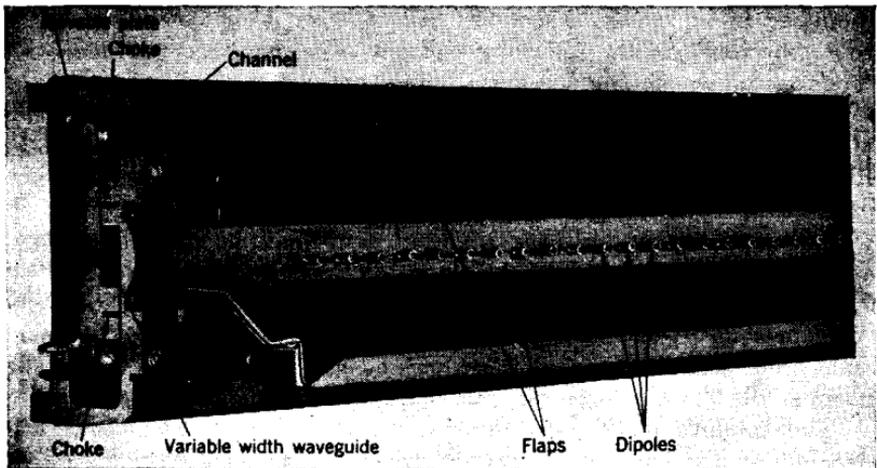


FIG. 9-19.—View of AN/APQ-7 scanner showing dipoles.

If all the dipoles are excited in the same phase, the beam is broadside to the array and directed straight ahead of the aircraft. With the dipole spacing actually used, this occurs when the inside width a of the guide is 1.200 in. Let us assume that the energy is propagated in the waveguide from left to right. Then, as the guide is narrowed to 0.660 in., the beam moves 30° toward the left, that is, in the *backfire* direction.

¹ See Sec. 11-3.

If the waveguide carries the energy from right to left, the scan extends for 30° to the right of the center line of the aircraft. By feeding r-f power into the ends of the array alternately, a total scan of 60° is realized. This alternate-end technique is always used in practice with the Eagle scanner; it is accomplished with the aid of a fast-acting r-f switch which operates synchronously with the guide-squeezing mechanism.

The azimuth beam angle θ is related to the guide width a , the dipole spacing s and the free-space wavelength λ_0 by the formula

$$\sin \theta = \sqrt{1 - \left(\frac{\lambda_0}{2a}\right)^2} - \frac{\lambda_0}{2s}$$

Here θ is positive in the *end-fire* direction.

No provision is made for tilting the beam in elevation. Fixed 16-ft aluminum flaps (Fig. 9-19) above and below the array serve to shape the beam in elevation to a roughly cosecant-squared distribution.

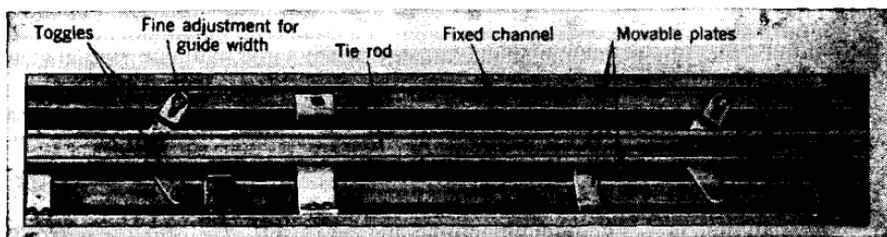


FIG. 9-20.—Back view of part of AN/APQ-7 scanner, showing toggles.

The fixed aluminum channel, which measures about $2\frac{3}{4}$ in. by $5\frac{3}{4}$ in. by 16 ft over all, is first extruded, after which the critical surfaces are machined. The plate is similarly manufactured. The plate is linked to the channel by a system of 10 pairs of toggles and a tie rod, as shown in Fig. 9-20. A cam actuates the motion of the plate, while the linkage holds the a dimension uniform along the variable waveguide. To one of the toggles is attached a cam so cut as to drive a synchro data take-off through an angle equal to θ .

The chokes shown in the channel and in the plate serve to prevent leakage from the varying guide so that the loss due both to the finite conductivity of the waveguide and to the leakage past the long choke is kept down to about $\frac{1}{2}$ db/m during most of the scan. Although the wavelength in the guide is nearly that which would be indicated by the a dimension, small corrections must be allowed for the chokes and for any variation of the clearance between the channel and the plate at the chokes. For this reason the b dimension is maintained to a close tolerance by means of rollers constraining the movable plate (Fig. 9-21). It has been found very profitable to adjust variations in the effective guide

width (or the phasing) by means of screw adjustments on the toggles of the squeeze mechanism. The a dimension is adjusted within a few

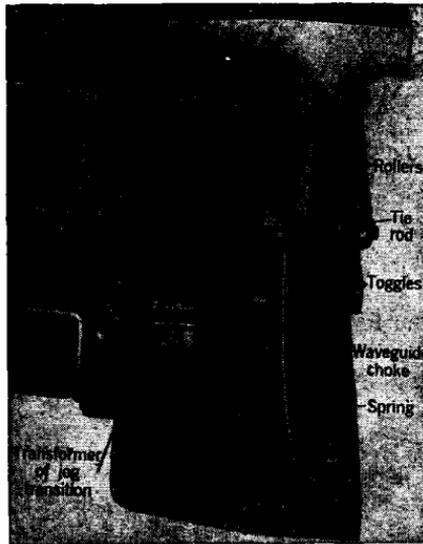


FIG. 9-21.—End view of AN/APQ-7 scanner.

thousandths of an inch as measured by an inside micrometer, after which the antenna is set up and its gain maximized by adjustment of the toggle screws.

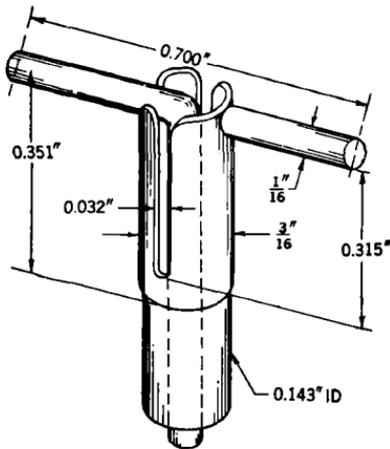


FIG. 9-22.—Construction of dipoles used in AN/APQ-7 antenna.

All the dipoles shown in Figs. 9-19 and 9-22 are horizontal and therefore radiate energy polarized horizontally; but alternate dipoles are reversed end-for-end in assembling the array. The spacing between the centers of adjacent dipoles is one-half the guide wavelength when the guide is 1.200 in. wide. Therefore in each adjacent pair the dipoles are excited in phase opposition; but because of the reversal of one of them they radiate in phase, producing a broadside beam. If alternate dipoles were absent the beam would still be formed, but would be accompanied by other strong beams comparable to undesired orders of diffraction from an optical diffraction grating. As it is, the side lobes are from 2 to 5 per cent as intense as the main beam.

The transmission line is a waveguide ending in a choke flange attached to the channel at the pad shown in Figs. 9-19 and 9-21. The transition is matched by means of the *jog transformer*¹ illustrated in Fig. 9-23. The transmission line contains the r-f switch mentioned above and sketched in Fig. 9-24. It is at the intersection of four waveguides. When the vane of the switch is at one of its 45° positions the energy is directed through the array in one direction and that 5 to 10 per cent of the energy which is not radiated is routed into a matched load. At the middle of the 60° scan the vane is flipped within 0.05 sec into its other 45° position and the energy flow in the array is reversed.

Of the many significant facts that have emerged in the study of the properties of the Eagle array, two are worth emphasizing. One is that practically all defects in the azimuth pattern of such a long array are caused by phase errors and not by departures from the desired partition of energy among the dipoles; the other is that random phase errors over lengths corresponding to a few dipoles are negligible when compared to systematic phasing errors.

The production scanner weighs 165 lb; the housing vane weighs 247 lb, including the gear box and motor. Thus the entire assembly, exclusive of the two mounting struts, weighs 412 lb.

9-14. Schwarzschild Antenna.—It is known (Secs. 9-3 and 11-1) that energy may be propagated in the space between flat parallel sheets of metal. In this and the two sections following we shall discuss electrical scanners in which curved parallel plates are used. In each scanner the plates are straight along one side; the microwaves escape from the straight gap between the plates on this side and then illuminate a parabolic reflector. The electrical scan is a periodic time variation in the direction of the energy falling upon the reflector. The antenna² about

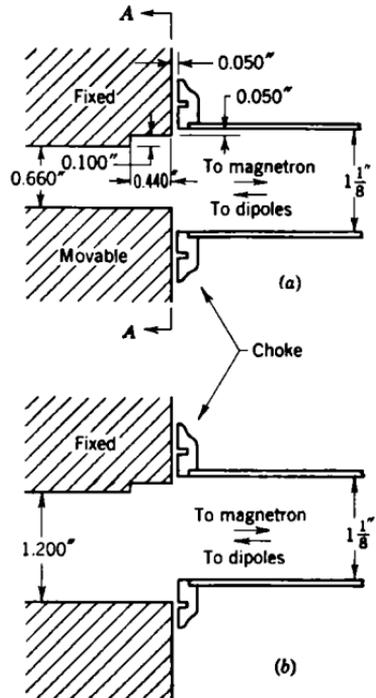


FIG. 9-23.—Jog transition from the rigid to the expandable guide. The latter is shown when narrow at *a* and wide at *b*. A section through A-A is shown in Fig. 9-21.

¹ R. M. Robertson, *loc. cit.*

² H. A. Straus, "Brief Description of AN/TPG-1, AN/FPG-1, SCR-598," RL Report No. 456, Oct. 8, 1943; also War Department Technical Manual TM 11-1566.

to be described was developed for the 3-cm band to provide coastal defense batteries with accurate target position and "splash-spotting" data for fire control. An azimuth beamwidth of 0.55° and a short pulse ($\frac{1}{4}$ μ sec) make accurate tracking possible in both azimuth and range. The sector that is rapidly scanned is 10° wide and the center of this sector may be swung to any azimuth position for search purposes.

Externally the scanner consists of a mount, which ordinarily rests on a tower platform; the antenna housing, 143 by 56 by 27.5 in., constructed

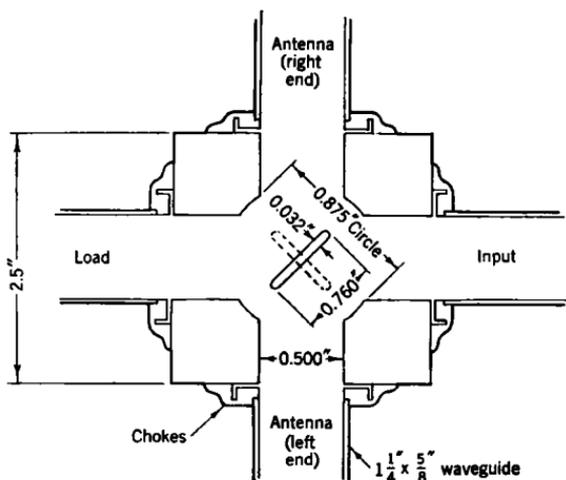


FIG. 9-24.—Detail of AN/APQ-7 r-f switch.

of plywood; and the copper-coated plywood reflector, 2 by 12 ft, hinged to the deck of the housing (Fig. 9-25). The reflector is a half parabolic cylinder of 1-ft focal length, and is fed by the 11-ft by 1.4-in. linear aperture of the enclosed portion of the antenna to produce the beam 0.55° in azimuth by 3° in elevation. Terminating in this aperture is a long "horn," a parallel-plate region of copper-coated plywood, folded as in Fig. 9-26. The conducting surfaces are spaced $\frac{1}{4}$ in. apart and increase in width from about 2 ft at the edge *AA* to about 11 ft at the edge *BB*. This horn extends over a length of about 12 ft and is folded three times. The feeding edge *AA* of the parallel plates consists of a boundary which forms a 90° arc, $14\frac{1}{2}$ in. in radius. Radiation is fed into the parallel plates between these flanges by four feed horns, each 2 in. wide, with suitable chokes meeting the flanges. The four horns are mounted on four radial waveguides which are fed at their intersection by a four-way r-f sector switch (Fig. 9-27). This switch consists mainly of a rotary joint¹ which terminates in four waveguides *F* at 90° to each other. These

¹ See Fig. 11-13.

arms rotate at 240 rpm, as do the four arms *H* which support the feed horns. They remain in the angular relation shown. Between the two sets of arms is a stationary ring with a 105° window through which energy may pass into whichever feed horn is opposite the feeding edge of the folded horn. When the horns and switch are rotated the effect is to cause the beam of the antenna to scan a 10° azimuth sector each

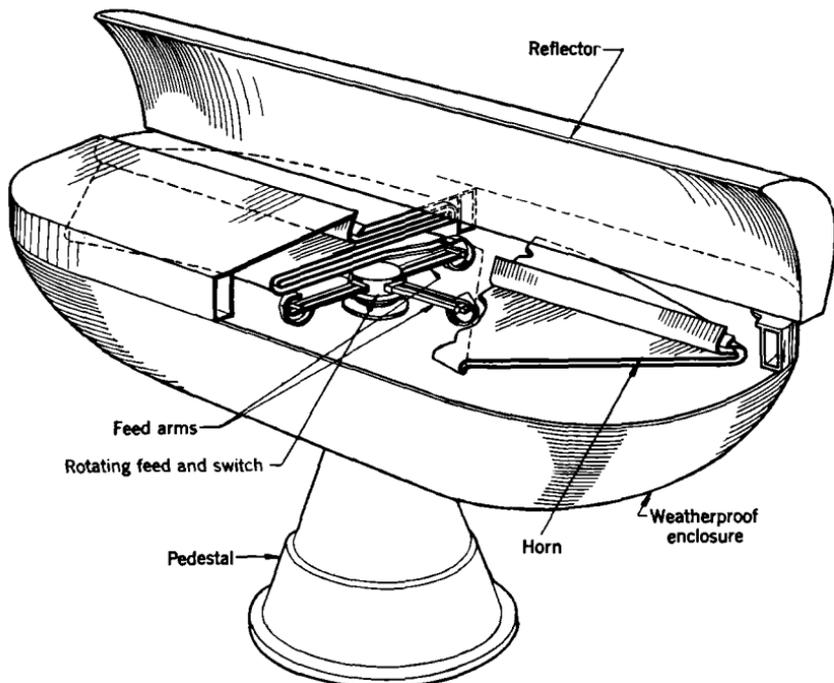


FIG. 9-25.—Perspective of horn, reflector, and feed of Schwarzschild antenna used in AN/MPG-1.

time a horn runs its 90° course. Since about 10° of the 90° are lost in switching, we can think of the parallel-plate region as effecting an 8-to-1 optical reduction from the 80° scan of the broad-beamed primary horn to the 10° scan of the very sharp antenna beam.

The transformation of the moving feed horn into the scanning beam is made by means of two double-curved bends *CC* and *DD* in the parallel plates (Fig. 9-26). Although one parabolic bend *CC* would serve to collimate the primary radiation and provide a sharp beam on axis, two bends may be so designed as to correct for coma and thus give a good beam off axis as well. The theory of two-mirror telescopes was used to calculate the bends for the Schwarzschild antenna. The cylindrical

bend EE near the aperture has no optical effect and was introduced for the sake of compactness. A plastic closure prevents moisture from entering the horn at the edge BB . Accurate wood-working makes possible the use of plywood construction which is preferable to aluminum in weight and ease of manufacture.

The drive motor runs at 1725 rpm and is connected by a 7-to-1 gear reduction to run the rotor about 4 rps. The azimuth data transmitter is a capacitive voltage divider connected to a high-frequency oscillator.

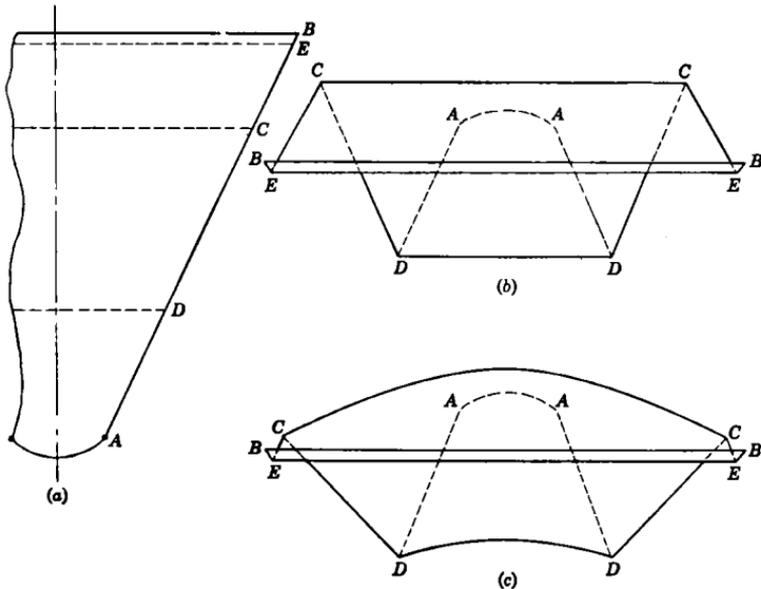


FIG. 9-26.—Successive steps imagined in folding the Schwarzschild horn.

The voltage developed across one of the condensers is detected and applied to the indicator. This condenser is of the variable rotary type and is driven at 16 rps, i.e., one revolution per scan.

The antenna is mounted on the vertical axis of the mount used for the SCR-584 antenna, with certain alterations.

Approximately 100 plywood antennas were constructed during the war. The production models that saw service showed no deterioration over a period of one year.

9-15. SCI Height Finder.—This antenna is intended for use in rapid coverage height-finding.¹ It is one of a pair of antennas (Fig. 9-28) mounted on the same pedestal and rotating together in azimuth at 4 rpm. The other antenna is used for long-range search and azimuth

¹ C. V. Robinson, "The SCI Rapid Scan Height-finding Antenna," RL Report No. 688.

determination. The height-finder beam, at the 10-cm band, is 3.5° in azimuth by 1.2° in elevation, and scans linearly one way for 10.5° in elevation from the horizon up, ten times a second. It makes continuous height-finding possible on small aircraft out to 50 miles, and farther on larger planes.

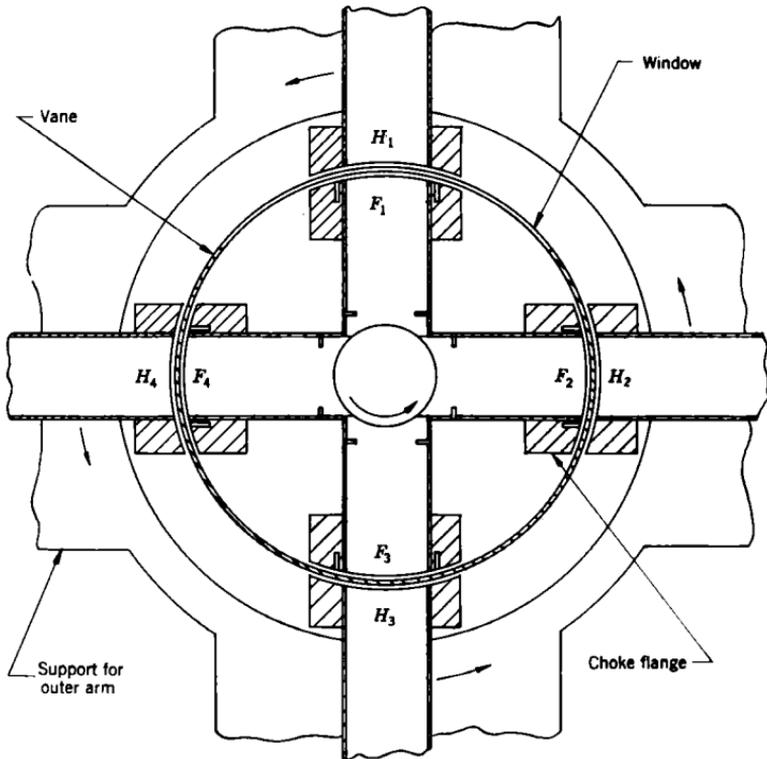


FIG. 9-27.—Four-way r-f switch used in the AN/MPG-1.

The principal elements of the antenna are a 5- by 15-ft grating reflector which stands on end, an 8- by 2- by 1-ft convoluted parallel-plate horn whose aperture is 5 ft from the reflector, and a rotating waveguide feeding the horn. The feed, which does not show in Fig. 9-28, is driven by a motor in the box attached to the horn. The parallel-plate horn serves to transform the circularly moving waveguide feed into an apparently linearly moving source. The beam scans 10.5° in elevation for each rotation of the guide.

The principle may be better understood by considering a simplified antenna which consists of a reflector and parallel trapezoidal plates with a long flared base "illuminating" the reflector; the plates are fed with

perpendicularly polarized waves by a waveguide at the small base (Fig. 9-29). As the waveguide feed horn is moved from top to bottom, the beavertail beam scans from minimum elevation angle to maximum. The

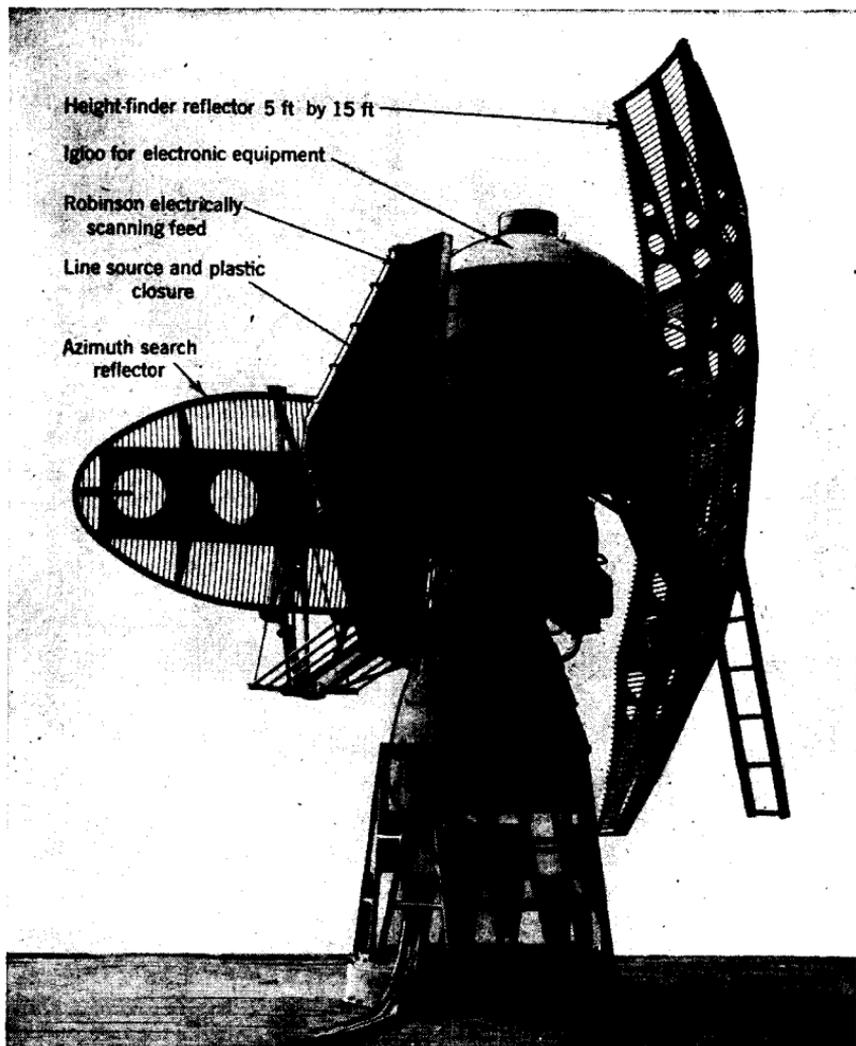


FIG. 9-28.—The three-axis antenna mount of the SCI radar.

principle is the same as moving the feed in an ordinary paraboloid reflector. The reflector for this antenna is not paraboloidal, however, but is astigmatic, having different focal lengths in the horizontal and vertical planes because the points of divergence of the rays in the feed differ in the two planes.

In actual fact, the feed horn moves, not linearly, but circularly. The rotating feature is achieved by folding and bending the trapezoidal plates, keeping them parallel with a minimum of stretching, in such a way that the feed base is rolled into a complete annular aperture while the flared base is held straight. Figure 9-30 illustrates how this can be done with a single trapezoidal sheet. Various additional mechanical and electrical complications appear when a two-sheet metal horn is made in this shape, but the convoluted parallel-plate region has substantially the same *optical*

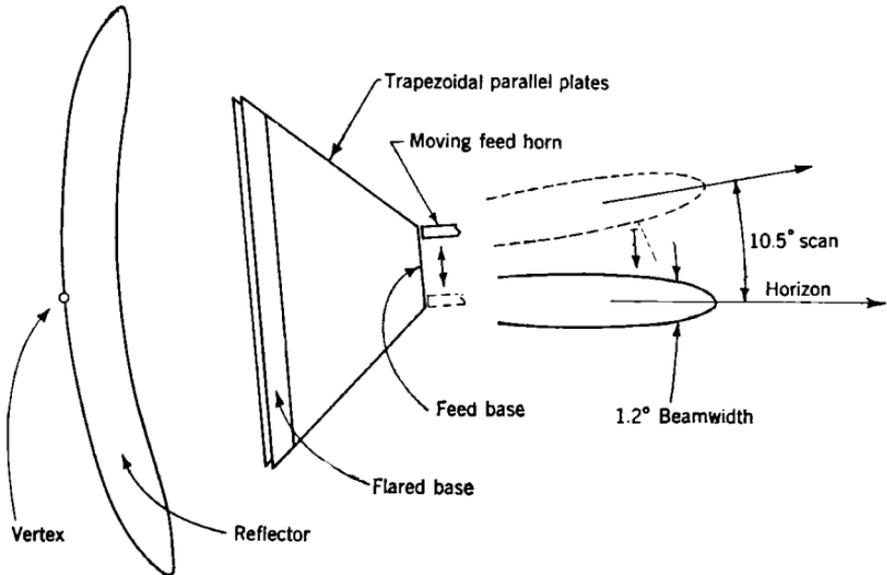


FIG. 9-29.—Simplified SCI height finder, side view.

properties as the trapezoidal plates before rolling. When the rotating guide feeds into the annular aperture of the horn, the beam describes a linear "fro" scan in elevation. The feed position marked O in Fig. 9-30 gives a beam on axis. When the feed is in the neighborhood of the cross-over point (about $\frac{1}{8}$ of the time required for one complete scan) the signal is ambiguous because there are two beams.

The parallel-plate assembly consists of two fabricated aluminum surfaces spaced $\frac{3}{4}$ in. apart, each consisting of a cast and machined center portion which supports a ribbed sheet metal "wing" section above and below. Specially designed steel posts are used to help maintain the spacing between these surfaces without ill effect on the electric waves. The rotary waveguide feed was made radial for compactness, resulting in the arrangement of the horn, chokes, and annular aperture shown in Fig. 9-31. The absorbing septum shown in this figure was introduced to reduce the amount of ambiguity caused by a double beam and to improve

the impedance match in this angular region by absorbing such radiation as may try to leak across that region as the horn approaches.

9-16. Other Types of Electrical Scanners.—Probably the first electrical scanner to be used, developed by Bell Telephone Laboratories, operates

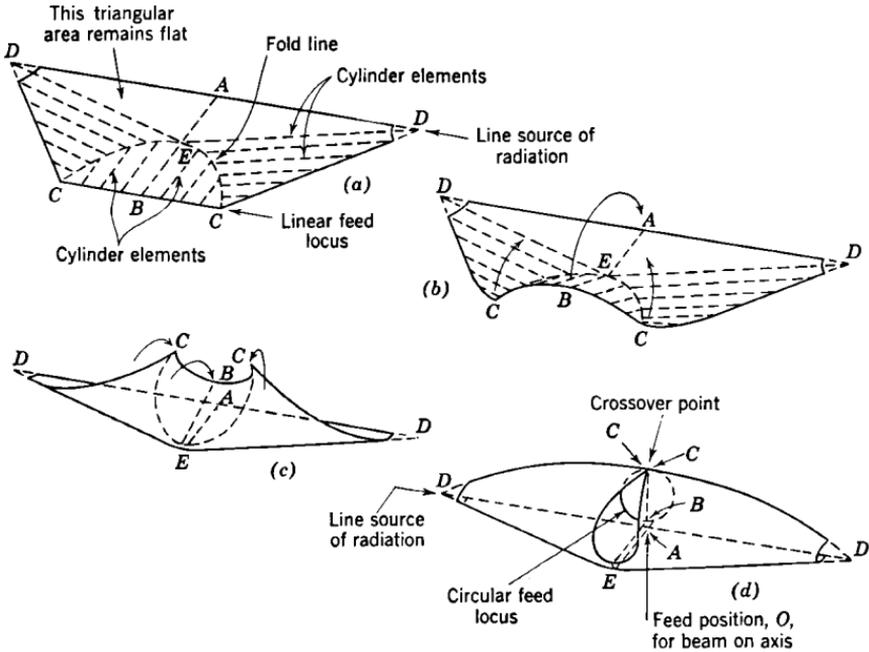


FIG. 9-30.—Illustration of trapezoid before, during, and after rolling.

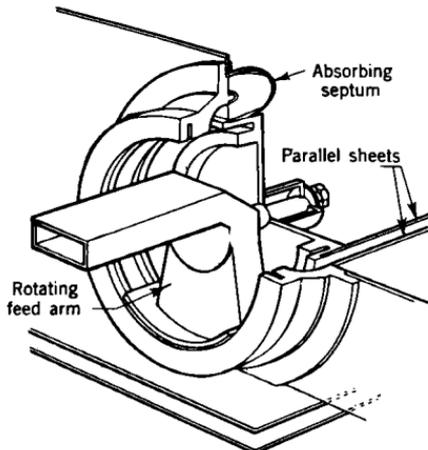


FIG. 9-31.—Detail of radial feed horn.

at about ^{39 Hz} 10 cm and produces a beam 2° in azimuth by 6.5° in elevation, which scans 29° one way in azimuth 10 times per second. The antenna radiates by means of an array of 3 by 14 polyrod radiators (Fig. 9-32). Each of the 14 vertical rows of 3 elements is phased as a unit and the

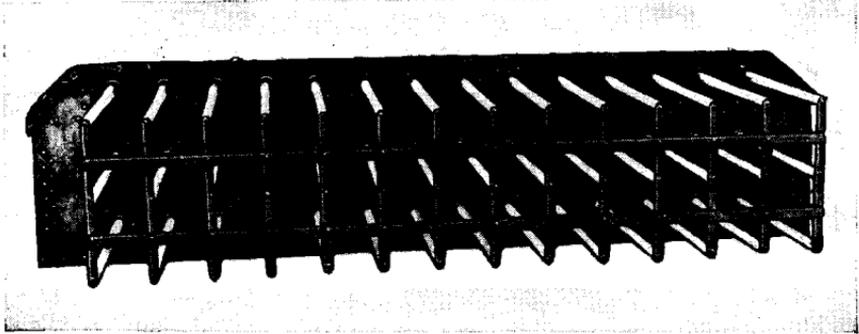


FIG. 9-32.— Array of end-fire polyrod radiators.

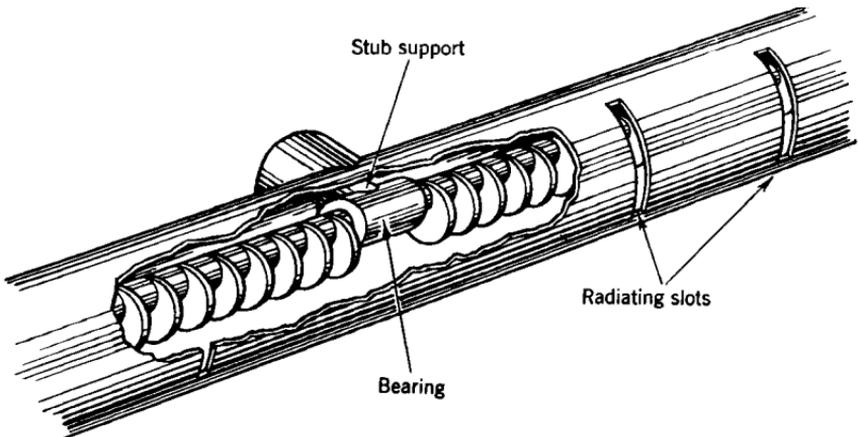


FIG. 9-33.— Cutaway section of eccentric line scanner.

units are phased linearly by a row of rotary phase shifters, each of which adds 720° of phase per revolution.

An interesting array-type antenna which reached a highly developed experimental stage is the Long Range, Aircraft to Surface Vessel, LRASV, corrugated coaxial line array. This antenna was designed to be used with a 10-cm system, mounted as a long array feeding a reflector along the side of an airplane, giving a beam 0.8° by 9° for long-range search. The array, shown in Fig. 9-33, consists of a 15-ft coaxial line whose outer conductor is pierced with radiating slots spaced at intervals of $2\frac{1}{4}$ in., which is less than the wavelength. The inner conductor consists of a $\frac{1}{4}$ -in. shaft on which are many $\frac{7}{16}$ -in. diameter eccentric disks which serve

to make the wavelength less in the line than in free space. This shaft is displaced from the axis of the outer conductor and supported by bearings at 10-in. intervals so that it can rotate. As the inner conductor turns, the wavelength in the line varies, causing the beam to scan 6.5° . Larger scans by this method are quite feasible.

Another rolled parallel-plate scanning antenna, similar to the type described in Sec. 9-15 in the respect of having a single rotary waveguide feed, and also capable of rapid scan, has been developed. The equiva-

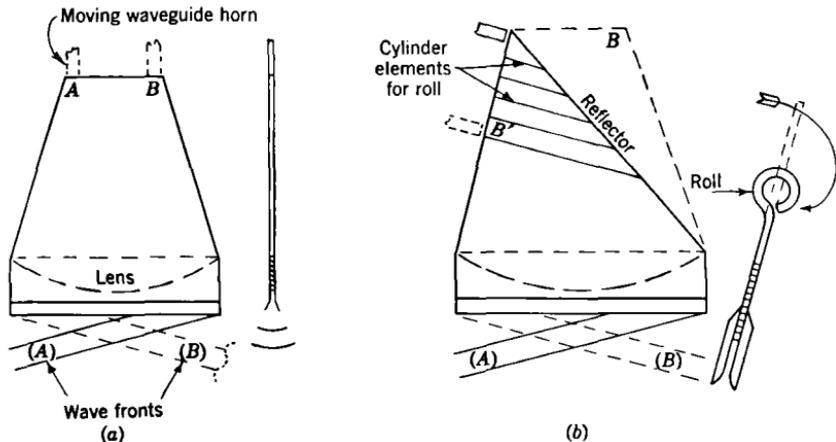


FIG. 9-34.—Reflector roll scanner.

lent simple antenna is a polystyrene lens placed between parallel trapezoidal conducting plates and fed by a waveguide at one edge of the plates (Fig. 9-34a). Moving the guide tilts the beam. A conducting reflector strip is placed between the plates in such a way that the waveguide feeds the lens from an adjacent edge of the plates (Fig. 9-34b). It is now possible to roll up plates and reflector strip in such a way that the feed path is annular. This type of roll is ordinarily simpler to build than the one described in Sec. 9-15.

In general it can be said that antennas of the array type (Eagle and LRASV) are compact but that the beam position depends on the frequency, while antennas of the parallel-plate type (for instance, Schwarzschild, SCI) have a beam position independent of frequency but have more bulk and weight. Thus, the array has found use in airborne work and mobile equipment, while the others tend to be used in ground and ship equipments.

THE STABILIZATION PROBLEM

Airplanes and ocean vessels are unsteady bases from which to make observations by radar or by any other means. It has been necessary to

devise means of compensating for the angular motions of such vehicles, and this practice is known as stabilization.¹ The technique falls into two broad divisions, depending on whether the beam of radiation is stabilized, or the data displayed on the indicator are stabilized or corrected for the distortions caused by a tilting vehicle.

In general there is no choice between the two types of stabilization. If the vehicle rolls or pitches to such an extent that its search radar no longer scans the desired part of the field of view, beam stabilization is necessary. This will prevent targets from fading from the indicator; if they are then indicated, but at a false position, data stabilization becomes a concurrent need. If the vehicle momentarily changes heading (yaw) while on a certain course, the relative azimuth of targets will momentarily alter; the consequent aberration of the display can be, and has been, corrected by either type of stabilization. Either sort of stabilization requires the vehicle to carry gyroscopes or other devices which are sensitive to direction in space and "know" the direction of north or up.

There are several ways of more or less completely mechanizing the stabilization of the beam. The most obvious and perhaps the most elegant is to provide a stable base for the scanner. This may be a platform mounted on gimbals and controlled automatically to compensate the pitching and rolling of the vehicle. If the vehicle rolls much but pitches little, the pitch-gimbal axis may be omitted, and the stable-base stabilization degenerates into roll stabilization.

A very different method of beam stabilization is possible in a scanner having its main axis always perpendicular to the floor of the vehicle. Use is made of the tilt axis of the antenna to the end that the antenna is automatically directed toward the horizon, or to whatever angle above or below the horizon is desired. The line of sight is thereby stabilized.

9-17. Stabilization of the Beam. *Airborne Antenna Stabilization.*²—Antenna stabilization for airborne radar is a scheme to preserve the same conditions of radiation illumination in nonlevel flight as in level flight. This is necessary to prevent (1) uneven illumination of the area being scanned, (2) loss of radar range, or (3) distortion of the PPI presentation. Stabilization equipment must be designed and constructed to stabilize the beam of radiation, whether it be pencil or fan, against any maneuver of the aircraft such as a climb, a glide, or any combination thereof. This must be accomplished with components that are not affected by the various accelerations, attitudes, or vibrations which might be encountered in turns, climbs, glides, or banks.

A stabilizer may generally be broken down into three main components: the gyroscope with its potentiometer or synchro take-offs, the

¹ The theory and methods of stabilization are discussed in Vol. 26 of this series.

² By F. B. Lincoln.

servoamplifier, and the follow-up system which includes the servomotor, the gear train or linkage system, and a take-off actuated by the controlled member. The gyro may be mounted on the rotating scanner or remote from the scanner, using slip rings and resolvers to take off the required error information. The error is fed into a servoamplifier whose output operates the servomotor which in turn aligns the antenna with the horizontal through a gear train or linkage system.

The gyro should preferably have either a two-speed erection mechanism or a device to disconnect the erection mechanism when the aircraft goes into a turn. In a turn the gyro rotor tends to align itself with the resultant of the true gravity and the centrifugal force caused by the acceleration in the turn, instead of with true gravity alone. This can be minimized or eliminated by using a two-speed erection mechanism, or by employing an electromagnetic clutch to throw out the erection system entirely in turns.

Continuing the general classification previously referred to, stabilizers may be further typed as stable-base, roll, or line-of-sight, depending on the stabilization compensation afforded.

The *stable-base* stabilizer, as the name indicates, provides complete stabilization of the platform upon which the antenna is mounted. This requires use of a two-axis gyro transmitter to provide alignment or error information for both roll and pitch axes. The gyro is usually mounted remotely from the antenna structure in this type. In addition two separate servoamplifiers and follow-up channels are required, one for the roll-error component and one for the pitch-error component. Although this provides complete stabilization, the weight of the components usually appears excessive so that another type is preferable.

The *roll* stabilization referred to is stable-base stabilization with the pitch channel removed. This requires a single-axis gyro transmitter for the roll axis, and but one servoamplifier and follow-up channel. Roll stabilization would be appropriate in large aircraft for which the pitch component might be negligible or tolerable. In deciding between a roll and a platform stabilizer the accuracy requirements must be balanced against the weight added by the pitch channel.

The *line-of-sight* method is one which maintains the axis of the beam of radiation in a horizontal position or a fixed angular distance above or below the horizon as the antenna rotates, scanning the horizon. This is done by automatically tilting the antenna about an axis perpendicular to the plane of the beam and parallel to the floor of the aircraft. Figure 9-35 shows the stabilizing attachment, AN/APA-15, mounted for photographic test on the scanner of the AN/APS-2 circular-scanning radar. The black tubular construction of the scanner yoke can be seen, together with the silver-plated $\frac{7}{8}$ -in. coaxial transmission line. The 29-in. paraboloid has

been removed from the gently curved support casting, which can be seen supported between the tilt bearings at the tips of the yoke. A table clamped to the yoke supports the gyro-torque unit and a servoamplifier. The gyro is mounted in gimbals and is provided with one synchro take-off

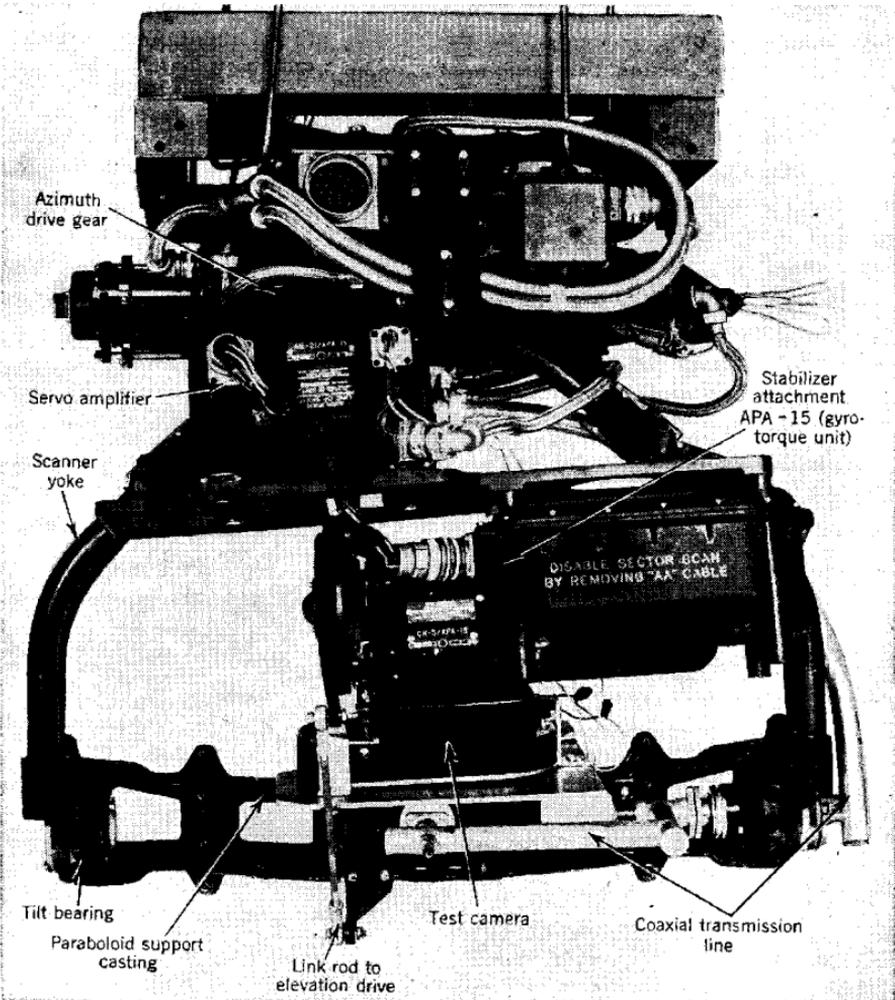


FIG. 9-35.—A stabilizing attachment mounted on an airborne scanner. The reflector has been replaced by a camera for test purposes.

at the gimbal axis parallel to the tilt axis. A synchro is also located at the output crankshaft of the torque unit. The servoamplifier essentially amplifies the algebraic difference between the voltages from these take-offs, and thus actuates the motor in the torque unit to assure alignment between the output shaft and the gyro. The output shaft is linked to the paraboloid support by a rod seen at the bottom of Fig. 9-35. When there

is a misalignment between the beam axis and horizontal, the synchros in the gyro-torque unit transmit the error to the amplifier, a phase-sensitive detector whose output controls the servomotor, which in turn causes the linkage system to tilt the antenna up or down until the beam axis is again horizontal. Because the gyro is mounted on the scanner, it is liable to the precession caused by the inertia of the supporting gimbal system. Under some conditions this precession causes the gyro to indicate a false vertical, thus producing serious errors in the stabilization system.

Although in the above example the gyro is mounted on the rotating yoke, the effects of precession are avoided in the most recent designs by mounting the gyro on the airframe and providing a take-off on each gimbal axis. A voltage indicating the correct tilt angle is fed to the servoamplifier. A special rotary inductor mounted on the scanner base provides this voltage, which it computes as a function of the antenna relative azimuth and the roll and pitch voltages from the gyro.

As previously discussed, the purpose of antenna stabilization is to preserve range performance by assuring constant illumination in non-level flight as in level flight. In considering the accuracy requirements various factors must be compared, such as beam shape and operational use contemplated for the radar. If a pencil beam moves up or down because of maneuvers of the aircraft or poor stabilization, the scope signal will fade. Quantitatively, if a beam moves up or down by one-half its beamwidth, it can be shown that the set will suffer a range reduction of 29 per cent. The range will be reduced 10 per cent if the beam varies from the stabilized position by 0.28 beamwidth and by 5 per cent if the beam varies 0.19 beamwidth. As a range reduction of 10 per cent is considered the maximum allowable, it can be seen that the beam must be stabilized so that it remains within 0.28 beamwidth of a completely stabilized position. The adequacy of the above calculations has been verified by rough observations in flight. Beam stabilization in current airborne navigation radars should be accurate within about $\pm 1.3^\circ$, depending on the set, lest range performance suffer noticeably.

In practice it is quite difficult to realize stabilization accuracies high enough to satisfy the foregoing tolerances for the various types of beams and radars. This is principally due to the early stage of airborne stabilization development and the stringent weight requirements on airborne gyroscopes. Current developments should lead to more highly accurate lightweight airborne antenna stabilization. The present static accuracy of airborne gyros ranges from $\pm \frac{1}{4}^\circ$ to $\pm \frac{1}{2}^\circ$. However, errors of from 1° to 5° may occur if the erection mechanism is not cut off in a turn. This error is a function of the duration of the turn.

Shipborne Antenna Stabilization.—In recent practice the stabilization of airborne ground-mapping antennas is by the line-of-sight method, in

sphere with the antenna imagined at its center. All lettered points are on the sphere: Z is the zenith, M is the prolongation of the mast, and H is the ship's heading. If the antenna is not stabilized, its pencil beam will at some instant be directed toward A below the horizon. The indicator will then show an angle measured by the arc HA . If the beam is stabilized by a two-axis mount, it will at that instant be at the point B where the great circle AM intersects the horizon; since HB exceeds HA the indicator will report the position of the beam falsely. If the beam is stabilized by a three-axis mount, it will be directed toward C ; the indication is again false since HC is less than HA . Thus either of these types of stabilization results in a distortion of the bearing of targets on the horizon.

A recent successful device for the stabilization of shipborne antennas includes computers and an associated gyroscope, and is located within the ship. The principal components of the stable vertical gyroscope are a gimbal system so supported that the outer and inner gimbal rings tilt respectively about the roll and pitch axes of the ship, a second gimbal system supported by and within the first, and the gyroscope wheel supported in the second gimbal system. The wheel axis seeks the vertical because of a pendulous magnet carried by a third gimbal system, and compensation is provided in the gyroscope for the effect of ship's turns or changes of speed.

The second gimbal system is held parallel to the plane of the gyro wheel, that is, horizontal, by means of a servomechanism; and from this gimbal system two push-rods extend upward and provide roll and pitch inputs to the mechanical computer located immediately over the gyro unit. The force to move these rods is provided by the servomechanism rather than by the gyro itself. In addition to roll and pitch, the desired angular position of the beam in space is fed to the computer as another mechanical input provided by a servomechanism. The computer is kinematically equivalent to a small replica of the antenna mount which is to be stabilized. The outputs are the deck-tilt correction and elevation angle for a two-axis mount, or deck-tilt correction and level and cross-level angles for a three-axis mount. They are mechanical but are converted to voltages by means of synchros, and control the servomechanisms which actuate the motors for the three axes of the mount.

More than one computer located on the gyro unit can be actuated simultaneously, making possible the simultaneous stabilization of more than one antenna. No computer is needed in stabilizing stable-base mounts, since roll and pitch data are furnished directly from the stable vertical.

Returning for the moment to generalities, two alternative methods are at hand by which the deck-tilt correction can be effected. One is to

let the antenna scan uniformly during search and to control the deflecting coil of the indicator to correct the deck-tilt error. This is a species of data stabilization; it is not commonly used. The other cure for the deck-tilt error is to drive the azimuth sweep of the radar indicator at a constant rate during search and to control the train angle of the antenna by means of a computer and servomechanism. As we have seen, it is one of the functions of the computer to provide the deck-tilt correction. This computation is easily mechanized, because the deck-tilt error is mathematically similar to the error in transmitting angular motion through a universal joint.

The gyroscopes used in shipborne antenna stabilization are accurate to within 2' to 7'; the error seems to be a function of the roll, pitch, and heading of the vessel. The large synchros used in this work are liable to err by about 10' to 30'. It follows that a good shipborne equipment, using synchros geared as high as 36 to 1 for increased accuracy, holds the beam with an average error of as little as 5', although momentary errors of more than twice this magnitude may arise.

9-18. Data Stabilization.¹—It has been remarked that even after the mechanical engineer has provided a stabilized antenna mount and thereby ensured that the radar will correctly perform its scanning function despite rolling or pitching of the vehicle, the indicator may still display the position of targets falsely. We are concerned in this section with the problem of improving the indicator in this respect.

The older PPI radars, both airborne and shipborne, were designed to display targets at their bearings relative to the heading of the vehicle. More recent design allows a display of targets in their true bearings, that is, with north at the top of the screen regardless of heading. The relative-bearing indication is bad for two reasons: (1) the entire display rotates and becomes confused if the course of the vehicle is altered; (2) the natural random changes in the heading of the vehicle cause a corresponding blurring of the display on the persistent screen. The true-bearing indication is free of these faults. It is a data stabilization in regard to yaw and changes of heading.

Yaw stabilization is implemented with the aid of a horizontal gyroscope which is kept pointing north by manual or automatic reference to the earth's magnetic field. Through a mechanical or electrical differential there is provided at the indicator a voltage signifying the true bearing of the beam, that is, the difference between the relative bearing of the beam and the relative bearing of the true north.

Deck-tilt error is serious in airborne as well as in shipborne radar. Let us imagine an airplane equipped with a bombing radar having a fan beam; the base of the scanner is not stabilized, and the antenna may or

¹ Secs. 9-18-9-20 by W. M. Cady.

may not be line-of-sight stabilized. If this airplane banks, the relative bearing of ground targets will be, as we shall see, falsely indicated. For example, let us consider the case of an airplane banking for a right turn, the relative bearing of the beam being straight ahead. It is clear that since the fan is perpendicular to the floor of the airplane it is not at the moment vertical, and that therefore the targets on the straight line on the ground now being illuminated by the beam are not at zero bearing. Since such targets are nevertheless displayed as if at zero bearing, the display is distorted; these targets are shifted by an amount equal to the altitude of the airplane multiplied by the tangent of the bank angle. Other targets are less seriously disturbed, those at 90° and 270° being correctly indicated on a banking airplane.

Two methods have been considered for correcting the airborne deck-tilt error. One method is so to stabilize the antenna that the plane of the fan beam remains indeed vertical. This can be accomplished by stable-base stabilization or by a three-axis mount. The other method is by data stabilization, and several circuits have been devised for removing the distortion if supplied with voltages depending on the angles involved. None of these circuits has been incorporated in systems now in wide use, although one radar designed for production incorporates a feature that accurately corrects the display at the one target designated as the aiming point.

9-19. Installation of Airborne Scanners.—In discussing antenna installations we will treat first the airborne case. Here, one of the pressing problems is the choice of location. The radar must have the unobstructed "vision" required for its operation, and its antenna should be so installed that it may be housed in an aerodynamically acceptable radome.

A very common requirement is circular vision of the ground well out toward the horizon, as in Fig. 9-6. The only good location is then below the fuselage. If rearward vision is not required, an antenna with circular scan can be mounted just behind and below the nose of the airplane, where the radome can be completely faired in.

Scanning of only a forward sector is possible if the antenna is mounted in the nose. Such a location is impossible in single-engine airplanes, and external housing becomes necessary. Aerodynamically, perhaps the best site in this category is at a wing tip, although a blister faired into a leading edge has been widely used. For maintenance reasons easy replacement of the set is desirable; this has been accomplished in one instance by packaging the greater part of the set in a "bomb" suspended under one wing. In the aircraft-interception radar of nightfighters the scanner is located for vision in all generally forward directions. Such scanners can be mounted in any of the locations that are good for sector scanners. Bomber-borne scanners for protective fire control are located

where they can search for and track enemy fighters approaching from the rear or other anticipated directions, or are located on the turret guns.

An important question in mounting a scanner in the belly of an airplane is "How far shall the antenna protrude below the keel line?" Too great a protrusion will add to the aerodynamic drag, whereas if the antenna is retracted too far the vision is inadequate in nearly horizontal directions because of partial blocking of the field of radiation. The ray diagram of the radiation field is not a sufficient guide in planning an antenna installation. In prototyping each installation the minimum protrusion allowable from the standpoint of the diffraction of the radiation must be determined by measurements of the antenna pattern made

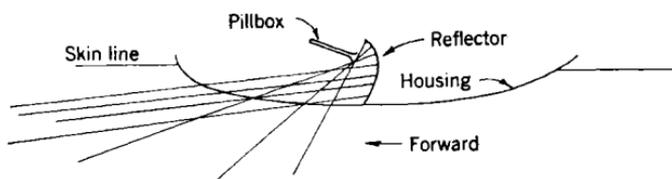


FIG. 9-37.—The ray diagram shows that fan-beam antenna may be installed partly inside the fuselage.

with the antenna installed in a mockup.¹ Such measurements have shown that some circularly scanning fan-beam antennas may be so installed that nearly half their height is above the keel line of the airplane, and this conclusion has been confirmed in observation of the indicator in flight. Figure 9-37 shows qualitatively why such an installation is possible with the type of scanner shown in Fig. 9-8.

In attaching the scanner to the airframe it is not necessary or advisable to use shock mounts unless electronic equipment is inseparably attached to the scanner. It is frequently essential, however, to provide adjustments for accurately orienting the scanner base. To allow this adjustment it is customary that the short length of transmission line between the scanner and the r-f unit be flexible: this flexibility also allows the r-f unit to be shock-mounted.

9-20. Installation of Surface-based Scanners.—The installation of scanners on shipboard presents a siting problem slightly simpler than in the airborne case. It is necessary to avoid locations where the structure of the ship will blank the radiation or produce spurious reflections. Sometimes it becomes necessary to install two scanners, each of which covers the region in which the other is blind. When a ship has more than one radar on a given wavelength, one can "jam" the other, filling its screen with spurious indications; this effect can be mitigated by proper siting of the scanners and by blanking the receiver of each at the instant

¹ A recent installation is described by W. M. Cady in RL Report No. 848, "The AN/APQ-13 (60-in.) Scanner in B-29 Airplanes." See also Figs. 9-7, 9-8, and 9-39.

the other is transmitting a pulse. On account of the curvature of the earth and for reasons of microwave propagation it is usually advantageous that the scanner be installed as high as possible. Radomes are not widely used to house shipboard scanners, the tendency being to make the antennas of such low wind resistance and such sturdy construction as to render radomes unnecessary.

Ground-based radars may be grouped, according to their mobility, as transportable, mobile, and fixed. Because the transportable scanners are disassembled for moving, their installation requirements include easy assembly. The AN/TPS-10 scanner shown in Fig. 9-15 disassembles into small parts for easy mobility in mountainous terrain. A firm and adjustable footing must be incorporated for transportable scanners. A truck or trailer on which a scanner is mounted must also be firmly braced while the radar is in use; the scanner itself must be adjustably mounted to the vehicle, if accurate position-finding is desired.

The siting of ground-based scanners is to be regarded as an aspect of the installation problem. A suitable site must be one where no mountain seriously blocks off the region beyond. A more subtle siting requirement is the avoidance of ground "clutter"; permanent echoes are bothersome at certain sites, particularly if the radar targets frequently pass at nearly the same range and direction as such echoes. Moving-target indication, discussed in Chap. 16, is now practicable to alleviate ground clutter.

9-21. Radomes.¹—It is necessary to protect all airborne and some surface-based microwave antennas from wind and weather. The protective housings have come to be known as "radomes." A fuller discussion than the following will be found in Vol. 26 of this series.

The circular cylinder shape was early used for the radome. For such a shape the radiation falls on the radome essentially at normal incidence, the angle of incidence being defined as the angle between the incident radiation and the normal to the surface at the point of incidence.

The development of microwave airborne radar systems forced the development of the science and art of radome design. It resulted in amplification of the electrical theory for radomes through which the radar beam passes at high angles of incidence. Plywood, first used for radome construction, gave way to more satisfactory synthetic materials; and aerodynamics, at first ignored, was given consideration in the design. Excellent installations are possible if the radar systems are considered in the initial design studies of aircraft.

Since the airborne radome involves all the problems of the surface-based installation in addition to many others peculiar to itself, this discussion will be limited to the airborne case. Some of the factors controlling the design of an airborne radome are the radar antenna and its

¹Secs. 9-21-9-25 by F. J. Mehringer.

function, radome materials, transmission of radar energy, structural requirements, aerodynamics, access to the radome cavity, and specification and test.

In this discussion only the three major items—*aerodynamics, electrical transmission, and structural design*—will be discussed, and these only briefly. None of these items can take preeminence over the others since a design aerodynamically or structurally poor can prevent the radar set from being useful, just as would a radome that had faulty transmission characteristics. The radome design must, therefore, be a compromise of these three major factors if accuracy and effectiveness of the airborne unit are to be secured.

9-22. Streamlining.—There is no simple foolproof answer to the aerodynamical problem. Each installation has peculiarities of its own, depending upon its location on the airplane in question, the plane itself, the speed of the plane, and the size and shape of the radome. A few simple considerations can be set forth but the real answer can only result from careful test in a wind tunnel.

Although, from an electrical viewpoint, the circular cylinder is the most desirable shape for a radome, aerodynamically it is objectionable due to its high drag. By streamlining a cylinder its air resistance can be reduced to a sixth or less. Any radome extending beyond the fuselage line of the airplane should therefore be streamlined.

The drag for a protuberance of given shape increases in approximate proportion to its frontal area. It is therefore desirable to reduce the projected area of the radome as much as possible even to the extent of having the radome surface coincide with the skin of the plane.

If a protuberance is necessary, the drag can be minimized by locating the radome at a position where there is already some disturbance of the airflow. The turbulence of the air stream tends to be a maximum toward the tail of the plane. While practically no protuberance is permissible near the nose of a high-speed airplane, one extending several inches amidship or further aft may cause a hardly noticeable increase of drag.

The truly satisfactory design, however, can be verified only by tests in the wind tunnel where very slight differences in aircraft performance are readily detectable. The wind-tunnel test can also be used to furnish pressure distributions on the radome surface, which are necessary for the structural design computations.

The location of the antenna on the airplane must also be the result of operational requirements and compromises with the other components. For navigation and bombing it is desirable to have the antenna scan 360° in azimuth. This can best be realized in a location on the belly of the plane about midship, but generally only by adding a large protuberance. A location in the "chin" of the airplane offers a more favorable location

so far as the aerodynamics is concerned, but limits the rearward view. This space, too, is desirable for the location of certain crew members and of the bombsight. Interference with other installations is usually less in an aft location, but the space is often limited by the size of the fuselage at this point and the forward view may be cut off by the forward portion of the plane.

9-23. Electrical Transmission.—The ultimate shape of the radome is generally limited by electrical design considerations. The retraction of the antenna is controlled by edge interference, for an edge can cause blanking and diffraction of a radar beam. The result is a distortion and deterioration of the antenna pattern. Any discontinuity such as that caused by reinforcement of the radome or by a piece of metal may cause diffraction if it intercepts the radar beam.

The permissible range of incident angle is the other electrical factor limiting the shape of the radome. The range of satisfactory incident radiation transmission varies with different materials and radome wall constructions. The properties of the materials affecting the transmission are the dielectric constant and the loss tangent; the latter is an index of the absorption and the former is a factor determining the reflection. For optimum combinations to date, satisfactory transmission can be obtained over a range of 0° to 70° . Beyond this value the transmission decreases very rapidly, reaching zero transmission at 90° incidence.

Having determined a reasonable shape for the radome, it is then necessary to check the angles of incidence by means of an antenna ray diagram. Typical ray diagrams for cosecant-squared antennas are shown in Figs. 9-2, 9-10, and 9-37. If the antenna is stabilized, the angles will have to be checked throughout the range of stabilization. When the angles of incidence exceed the limiting value the contours of the radome must be adjusted until the upper limit of angle of incidence is not exceeded.

9-24. Structural Design of Radomes.—The thickness and construction of the radome wall can be determined on the basis of structural considerations within limits depending upon the allowable reflection. The reflection is a cyclic function depending upon the dielectric constants, thickness of the material, and angles of incidence of the radiation.

There are two basic types of construction: the single plastic or laminated wall, and the so-called "sandwich" construction involving two or more materials. The sandwich construction consists of high-strength skin surrounding a core of lower density, which results in a high-strength, low-density construction. The sandwich construction also provides a greater variety of wall construction for high strength and efficient electrical transmission.

The approach to the structural design problem will depend upon the shape of the radome and the pressure distribution. These two factors

will determine whether the problem is one of strength or stability. The stability problem is a difficult one to meet, since no ribs or other means of reinforcement can be used for the extensive transparent area because of possible distortion of the electrical pattern.

The design of the mounting of the radome should have due consideration. Servicing requirements necessitate removal of the radome. The means of mounting the radome must be quick-fastening as well as strong.

Just as the final proof of a satisfactory aerodynamical and electrical design are aerodynamical and electrical tests, so structural tests are

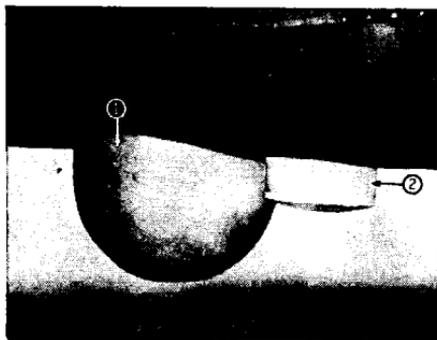


FIG. 9-38.—A retractable radome (1) partly extended on a B-29, and the bulge to accommodate the r-f unit (2) aft of the radome.

necessary to assure a satisfactory design, for in a streamlined shape the calculations can at best be only approximate.

9-25. Examples of Radomes.—Types of radomes vary greatly in size, shape, and method of construction. Attention is called to a few representative radomes.

Figure 9-38 shows an early type of retractable airborne radome mounted on a B-29. The radome is a cylinder $35\frac{1}{4}$ in. in diameter and extending about 33 in. when unretracted. While the shape is ideal electrically, its effect on the speed of the aircraft is most objectionable.

The leading edge of the Eagle vane (Fig. 9-18) is a radome made of two Fiberglas walls with occasional ribs as spacers. While the airfoil section is in itself of low drag, the attachment to the plane increases the drag until it is almost as high as that for the cylindrical radome. This increase in drag is due to the interference and stagnation points set up by mounting the airfoil to the plane. An additional feature of this radome was that it was designed to be deiced by means of hot air. The optimum wall spacings for heat transfer and for electrical transmission do not coincide, so that deicing is not very efficient.

A good bombing antenna installation for a B-29 is shown in Fig. 9-39. This radome houses the antenna shown in Fig. 9-8. The radome extends

a minimum amount (about $10\frac{1}{2}$ in. in the final installation), thus presenting a minimum of frontal area. These factors make this installation aerodynamically superior to those shown in Fig. 9-41 and 9-38. This radome is made of Fiberglas laminate in a sandwich construction. Its electrical transmission exceeds 88 per cent for both parallel and perpendicular polarizations through an angle-of-incidence range of 0° to 70° .

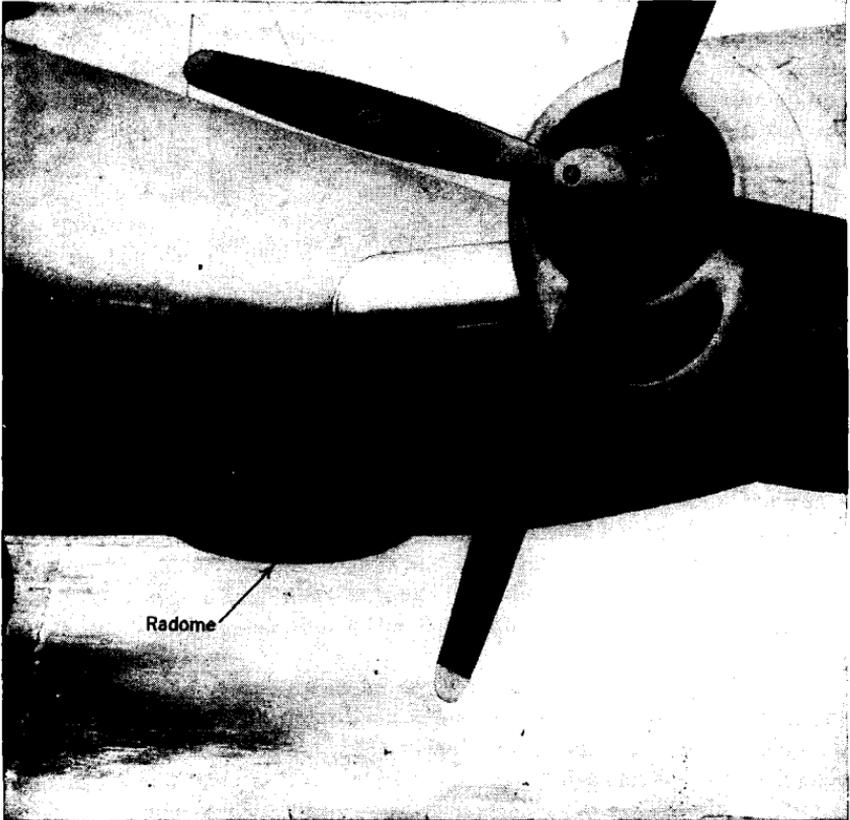


FIG. 9-39.—A recent installation on a P-29. Here it is shown extending 7 in.; the radome was later redesigned to extend $10\frac{1}{2}$ in. to avoid interference between the radiation and the fuselage.

Figure 9-11 shows a radome covering an end-fire array. This housing consists of a thin-walled tube of Laminac and cotton duck. In addition to protecting the antenna, the housing is used to pressurize the antenna.

The largest airborne installation made during the war is in a TBM airplane (Fig. 9-40). The radome illustrated, housing the antenna shown in Fig. 9-12, is almost 9 ft wide, $3\frac{1}{2}$ ft deep, and 13 ft long. Notwithstanding its large size it does not seriously alter the performance of the

aircraft; two additional stabilizers and a larger motor are all that are necessary to maintain performance despite the presence of this large protuberance. This radome depends upon a sandwich wall construction of Fiberglas laminated skins and Hycar core to give it the desirable electrical and structural characteristics.

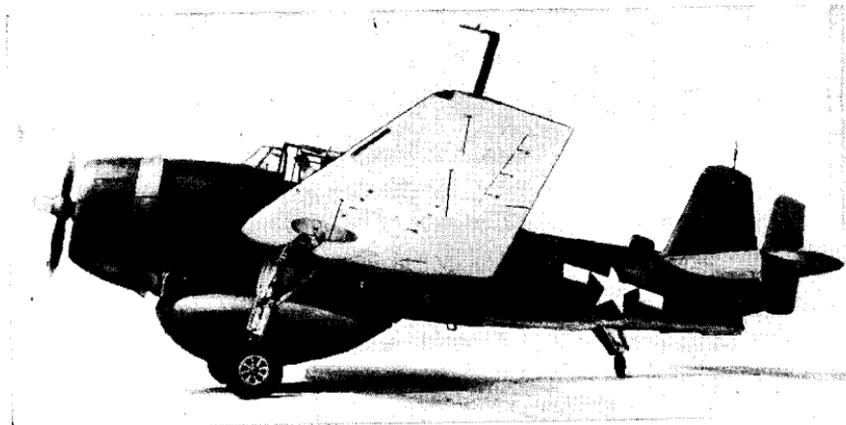


FIG. 9-40.—The largest airborne radome. It houses an 8-ft antenna under a TBM aircraft.

CHAPTER 10

THE MAGNETRON AND THE PULSER

BY G. B. COLLINS, J. V. LEBACQZ, AND M. G. WHITE

THE MAGNETRON

BY G. B. COLLINS

Strenuous efforts were made by the British, beginning about 1938, to develop high-power pulsed sources of radiation at very high frequencies, because of the operational need for microwave radar. Two lines of attack were followed. One consisted of attempts to improve and modify conventional types of transmitting tubes, and the other of efforts to devise entirely new forms of transmitting tubes.

Modifying the conventional type of transmitting tube to satisfy the requirements of radar met with difficulties. To a first approximation the electronic characteristics of a low-frequency oscillator or amplifier are preserved at high frequency only if all dimensions of the tube are scaled in proportion to the wavelength λ ; as a result, it is necessary to reduce the size of tubes as the frequency is increased. The practical consequences of this become serious as frequencies of 1000 Mc/sec are approached. The reduction of cathode and plate areas, which under these conditions vary as λ^2 , rapidly reduces the available peak emission and plate dissipation. Electrode clearances become so small that they are difficult to maintain accurately. Although many improvements in the design of triodes or tetrodes for use at microwave frequencies have been made, the general limitations just outlined have resulted in greatly reduced efficiency and power output in the microwave region.

Fortunately, a new type of pulsed microwave generator was invented whose performance was indeed spectacular. These generators are now known as "microwave" or "cavity" magnetrons, and they constitute the most important single contribution to microwave radar. The description, method of operation, and characteristics of this type of magnetron are the subject of the first part of this chapter.

As a source of high-power microwaves the multicavity magnetron represents a very great advance over both the conventional space-charge and the velocity-modulated or klystron-tube types. A few numerical comparisons will emphasize this superiority. Above frequencies of 3000 Mc/sec, space-charge tubes such as triodes cease entirely to be practical sources of r-f power, while magnetrons produce pulse powers

of the order of hundreds of kilowatts at frequencies as high as 24,000 Mc/sec. The average power output of magnetrons at 3000 Mc/sec is in the neighborhood of hundreds of watts, or one hundred times that of a triode operating at the same frequency. Klystrons are very useful sources of low c-w power at frequencies as high as 24,000 Mc/sec, but cannot be considered high-power pulsed sources. Their pulse-power output ranges from a few hundred watts at 3000 Mc/sec to a few milliwatts at 24,000 Mc/sec.

Magnetrons are self-excited oscillators and their output does not have the frequency stability possible at frequencies where power amplifiers are available and the output frequency is established by crystal-controlled oscillators. When properly designed and used, they exhibit stability adequate to the demands of pulse radar.

10-1. Construction.—Microwave magnetrons, or cavity magnetrons as they are frequently called, are basically self-excited oscillators whose purpose is to convert the d-c input power into r-f output power. Figures 10-1 and 10-2 show a particular design of a 10-cm magnetron which is typical of this class of transmitting tubes. Between the cylindrical cathode *C* and anode block *A* is an interaction space *I* in which the conversion of d-c to r-f power takes place. A constant and nearly uniform magnetic field is maintained in this interaction space in a direction parallel to the axis of the tube. In operation, the cathode is maintained at a negative potential while the anode block is usually at ground potential. The anode block is pierced in a direction parallel to the axis by a number of *side cavities R* which open into the interaction space so that the anode surface consists of alternate segments and gaps. The ends of the resonating cavities open into chambers which are called *end spaces*, through which the lines of flux extending from one resonator to the next pass. The coupling between the resonators is increased (in the design shown in Figs. 10-1 and 10-2) by conducting bars called *straps S* which connect alternate segments. Power is extracted from one resonator, one method being the use of a *coupling loop L* which forms a part of the output circuit. The combination of resonant cavities, end spaces, straps, and output circuit is called the *resonant system*.

A more detailed discussion of these parts of a magnetron follows. For pulsed operation, the cathode *C* is usually oxide-coated and heated indirectly by an internal heating coil of tungsten or molybdenum. The cathode structure is attached mechanically to two cathode stems supported by glass to provide anode-to-cathode insulation. At each end of the cathode there is an end shield, or hat *H*, whose purpose is to prevent electrons from leaving the cathode structure in a direction parallel to the axis of the magnetron. These end shields must be kept at a temperature too low to cause the emission of electrons.

The dimensions of the interaction space depend upon the wavelength and voltage at which the magnetron is to operate and also upon the number of resonant cavities, or oscillators. The ratio of cathode diameter to anode diameter is determined principally by the number of oscillators, no matter what the wavelength or operating voltage is to be. For a 12-oscillator magnetron the diameter of the cathode is about one-

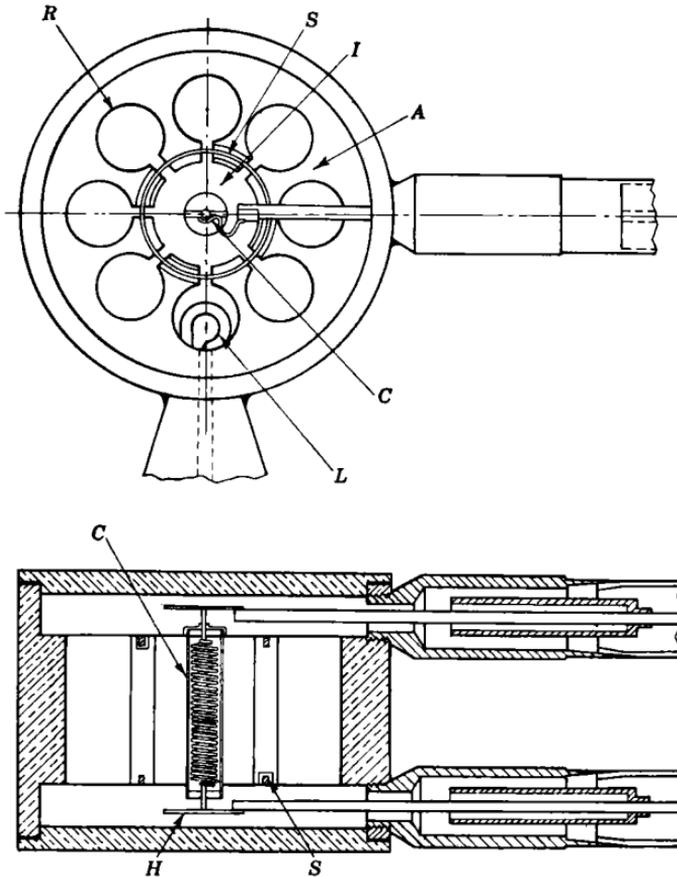


FIG. 10-1.—A typical cavity magnetron.

half the anode diameter. For fewer oscillators, it is somewhat smaller than one-half, and for more than twelve oscillators, somewhat larger. The anode and cathode diameters for any given type are proportional to the wavelength and to the square root of the anode voltage.

The particular resonant system shown consists of eight side cavities. Each of these cavities is similar to a simple oscillating circuit consisting of a lumped L and C . Although the inductance and capacity of the magnetron cavity are not strictly lumped, the inductance of the oscil-

lator resides mainly in the circular hole, and the capacitance mainly between the parallel plane surfaces of the slot. The dimensions of these

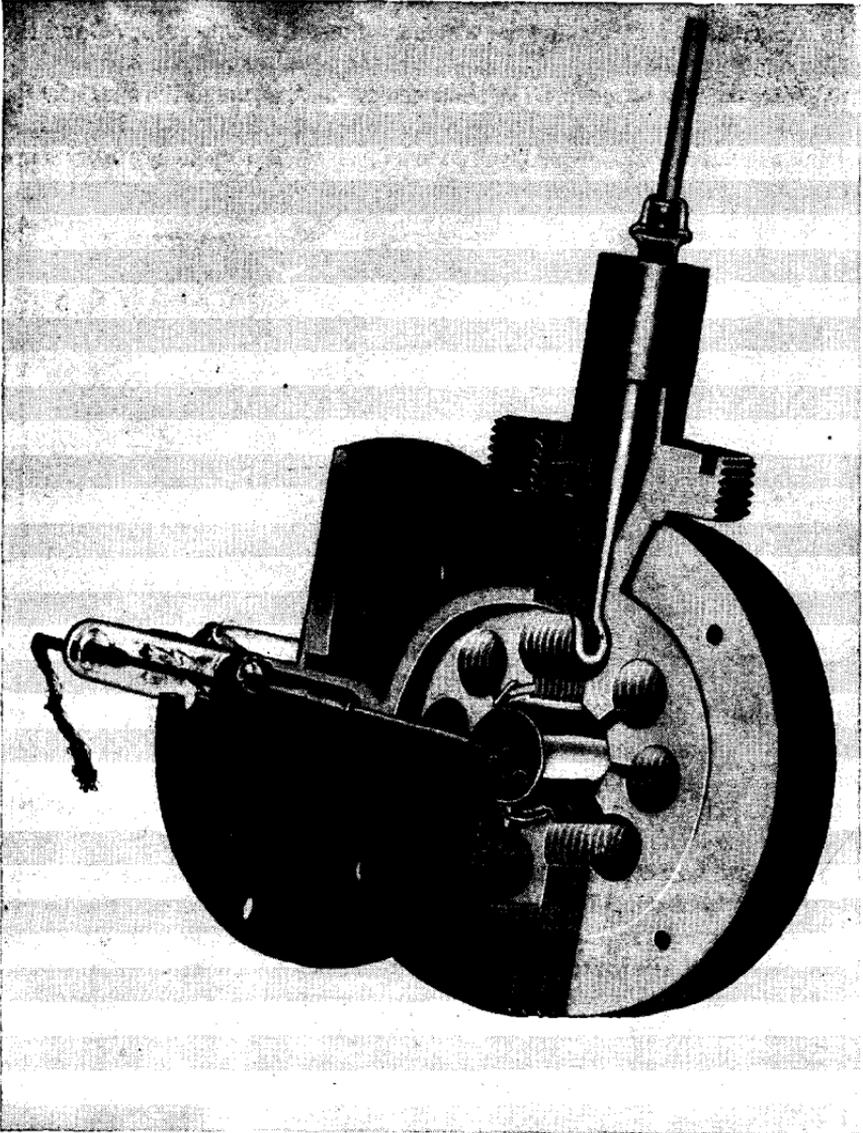


FIG. 10-2.—Cutaway view of type of magnetron shown in Fig. 10-1.

cavities determine to a large degree the oscillation frequency of the magnetron. Since the frequency is a function of the product LC , the same frequency can be obtained from a variety of shapes which have the

same LC but different L/C . It should be noted that the arrangement of these cavities is such that for the desired mode of operation their individual C 's are connected in parallel, and so are their individual L 's. Thus the effective capacitance for the whole magnetron oscillator is NC and the effective inductance is L/N , where N is the number of resonators. The frequency of the magnetron is thus nearly that of an individual cavity. The performance of magnetrons is not very sensitive to the impedance of the oscillators, L/C , and a wide variety of oscillator configurations are possible. Three types are shown in Fig. 10-3.

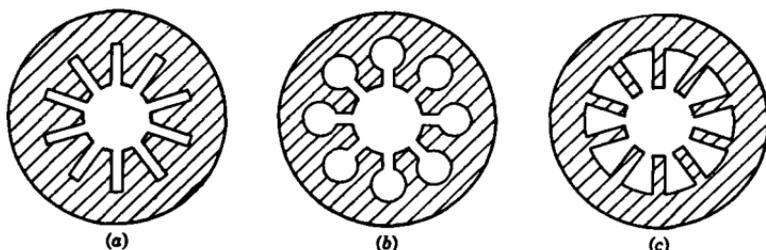


FIG. 10-3.—Magnetron anode blocks. (a) Slot-type magnetron. (b) Hole-and-slot-type magnetron. (c) Vane-type magnetron.

The simplest is perhaps the slot type (Fig. 10-3a), each of whose cavities may be thought of as a quarter-wave line (see also Fig. 10-22). It is a low-impedance oscillator having a large C and small L . The vane type (Fig. 10-3c), is obviously a high-impedance oscillator, whereas the hole-and-slot type (Figs. 10-2 and 10-3b) usually has an impedance intermediate between the two. These oscillator configurations were selected to illustrate some shapes that have been found convenient; the number of possibilities is unlimited.

Returning to Fig. 10-1 it is seen that the side cavities open at both ends into chambers, or end spaces. These end spaces assist in coupling the separate oscillators by allowing magnetic flux to pass from one oscillator into the next. The straps S consist of rings which are connected only to alternate segments. The connections are made in such a way that one set of alternate segments is connected by a strap at one end of the anode, and the other set of segments is connected by a corresponding strap at the other end of the anode. Magnetrons may be either single- or double- (or even quadruple-) strapped, meaning they have either one or two (or four) straps at each end of the anode. Figures 10-3a, b, and c show magnetron anode blocks, without straps; Figs. 10-1 and 10-2, single-ring strapping; and Fig. 10-20, double-ring strapping.

The magnetic field parallel to the axis of the cathode is often produced by an electromagnet or permanent magnet with pole faces external to the magnetron. Figure 10-4a shows a typical permanent magnet and a

magnetron with radial cathode supports. Another type of magnetron construction, favored for the higher-frequency magnetrons where magnet weight is of importance, is shown in Fig. 10-4b. The construction of such a magnetron-magnet combination, or "packaged magnetron" as it is frequently called, is shown in Fig. 10.5. The cathode is usually supported axially through iron pole pieces which extend quite close to the

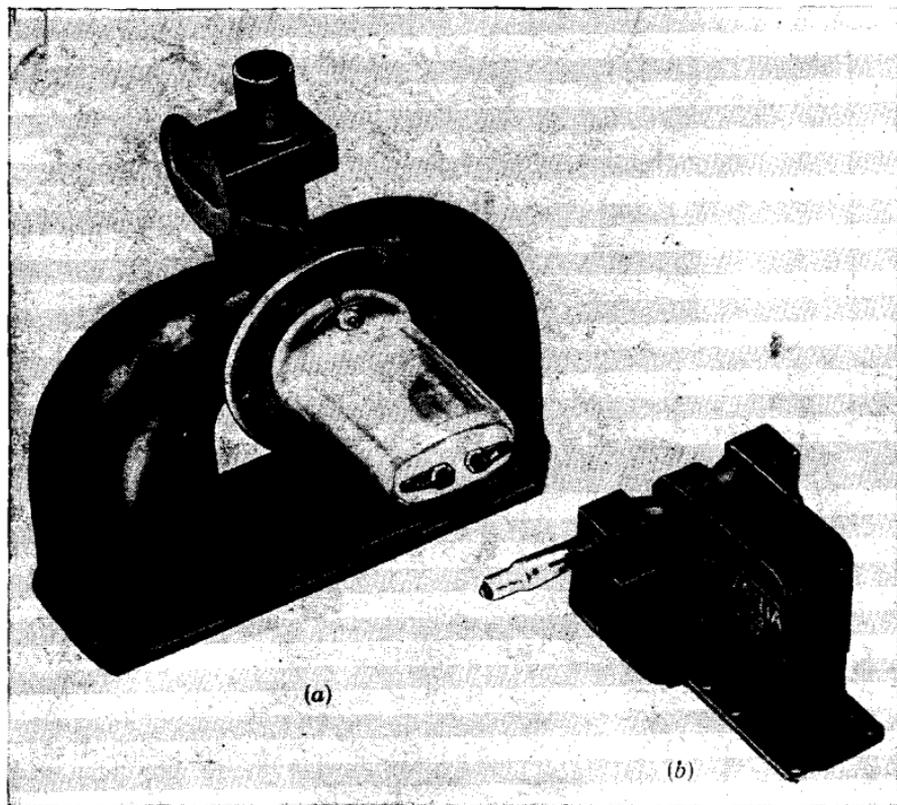


FIG. 10-4.—Two types of magnetron construction: (a) radial cathode and separate magnet; (b) axial cathode with attached magnet.

anode and thus reduce the magnetic field gap. Since the weight of a magnet which will produce a given magnetic field strength over a given cross-sectional area increases very rapidly with the length of the gap, considerable magnet weight can be saved in this manner. It is customary to supply this type of magnetron permanently attached to its magnet.

10.2. The Resonant System.—In an operating magnetron the charge distribution in the resonant system produces electric fields which interact with the space charge in such a way as to sustain the oscillations. Figure 10.6a illustrates such a disposition of charge and electric field at an instant when the concentration of charge on the ends of the anode seg-

ments is a maximum. One quarter of a period later the electric field and charges have disappeared and currents are flowing around the inside of the cavities, producing a magnetic field along the hole portion of the cavities. Figure 10-6*b* depicts the currents and fields at this instant.

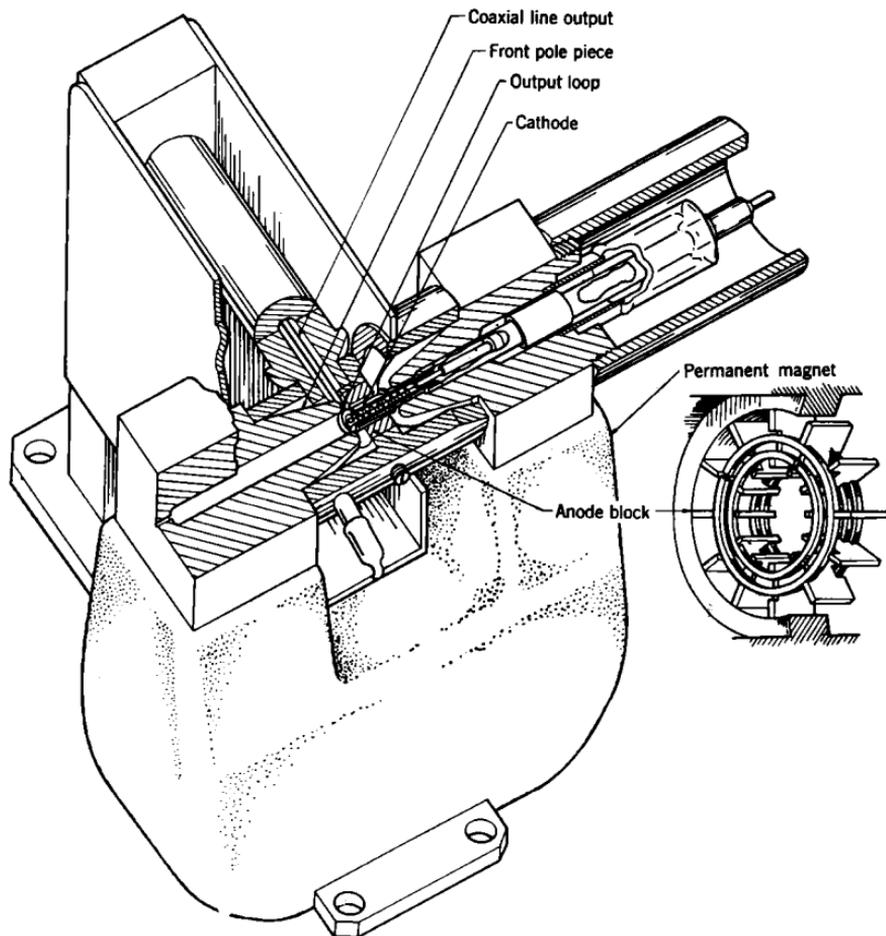


FIG. 10-5.—Magnetron with axial cathode and inserted pole pieces.

Figure 10-6*c* shows the charges and electric fields another quarter period later.

Figure 10-7 is a view of the magnetron opened out in a plane parallel to the cathode so that the anode faces lie on a plane. The broken lines show the paths of magnetic flux at the instant shown in Fig. 10-6*b*.

Figures 10-6 and 10-7 illustrate the so-called " π -mode" in which magnetrons normally operate; (π refers to the phase difference between adjacent anode segments). Actually, the number of possible modes of

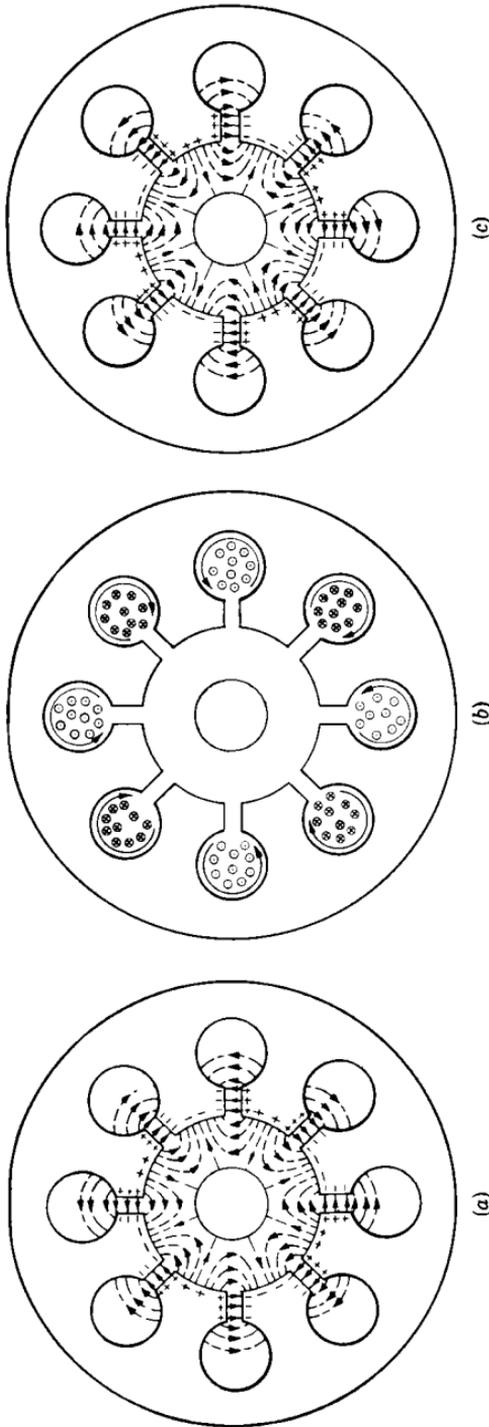


FIG. 10-6.—Distribution of r-f fields, current, charges, and magnetic fields in an oscillating hole-and-slot anode block; (a) r-f fields and charge distribution when charge is at maximum; (b) currents and magnetic fields one-quarter period later; (c) same as (a) one-half period later.

oscillation is at least one less than the number of oscillators. Figure 10-8

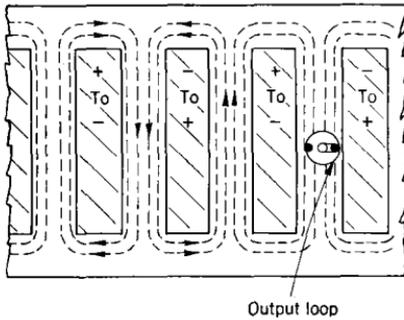


FIG. 10-7.—View from cathode of anode block opened out to show longitudinal distribution of magnetic field.

shows the distribution of charge and electric field for a magnetron oscillating in the $\pi/2$ -mode. A comparison of the electric fields with those of the π -mode shown in Figs. 10-6a, b, and c indicates a fourfold, instead of eightfold, symmetry, and shows that the field falls off less rapidly toward the cathode. Experience has shown that magnetron operation in any but the π -mode is unsatisfactory.

Possible modes for an 8-oscillator magnetron total seven: two $\pi/4$ -modes, two $\pi/2$ -modes, two $3\pi/4$ -modes, and one π -mode. In general, because of differences in the effective inductance and capacity of the

oscillating circuits, each of these modes has a different frequency of oscillation.

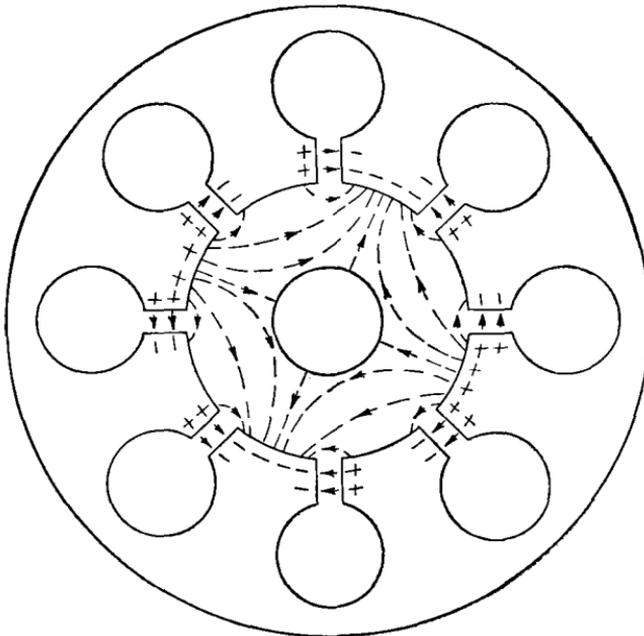


FIG. 10-8.—Distribution of r-f fields and charges in a hole-and-slot anode block oscillating in the $\pi/2$ anode.

oscillating circuits, each of these modes has a different frequency of oscillation. One of the objectives of magnetron design is to arrange

matters so that the electrons excite only the π -mode. Adequate frequency differences between the desired mode and all others favor stable operation; in most magnetrons, the frequency of the mode nearest in frequency is at least 10 per cent greater or less than that of the π -mode. Unfortunately it has been impossible to design magnetrons that do not show multiple frequencies under some conditions and this phenomenon remains as one of the most troublesome encountered in pulsed magnetron operation.

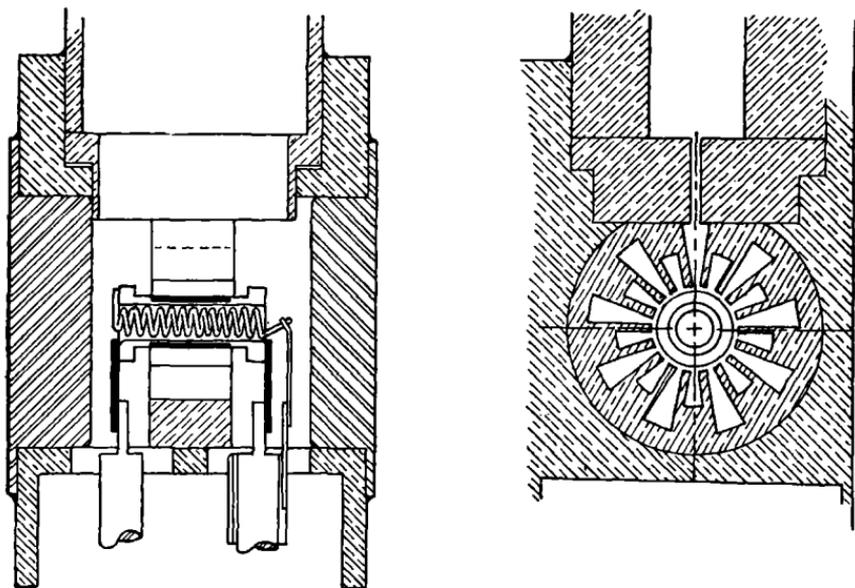


FIG. 10-9.—Rising-sun magnetron with waveguide output and radial cathode leads.

Power is extracted from only one of the oscillating cavities. This may be done by means of either a coaxial line output or a waveguide output. For the coaxial line output, it is usual to insert a loop L (Fig. 10-1) into the inductive portion of one of the oscillators in such a way that it links the magnetic flux. An r-f potential difference is thus induced between the ends of the loop, which are connected to the inside and outside conductors of the coaxial line. This type of output is satisfactory for magnetrons whose frequency is 3000 Mc/sec or less. At frequencies above 3000 Mc/sec a waveguide output is to be preferred as small coaxial lines arc over and are unduly lossy. Power is extracted from the back of the cavity by means of a slit which expands, usually with discontinuities, until its dimensions correspond to those of the waveguide (Figs. 10-9 and 10-10).

The magnetrons discussed above are referred to as "strapped"

magnetrons. The majority of existing magnetrons are of this type. A variety of other designs have been constructed, most of which have not proved to be as satisfactory as the strapped type. One type, however, has several marked advantages over strapped tubes. It is known as the "rising sun" magnetron (Figs. 10-9 and 10-10). Mode separation is accomplished in this design by making alternate resonators of a different

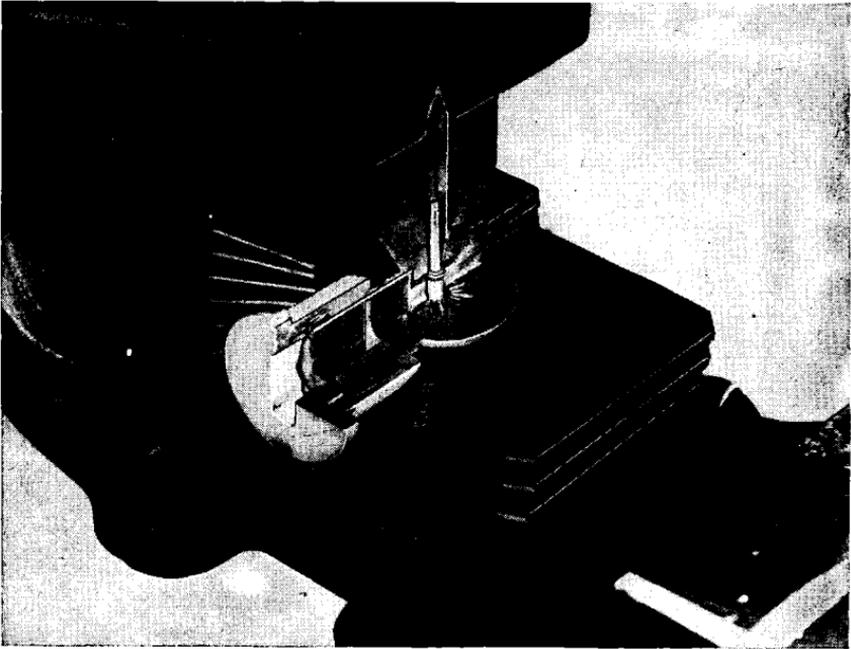


FIG. 10-10.—Cutaway view of rising-sun magnetron with axial cathode. The wavelength is 1.25 cm. (Courtesy of Bell Telephone Laboratories.)

frequency. It is interesting to note that this resonant system results in an operating or π -mode which lies intermediate in frequency between other modes while, for strapped tubes, the π -mode has the lowest frequency. For very high powers, the rising-sun system has considerable promise. Unlike strapped systems, the rising-sun system permits the anode to be made with a circumference and a length which are not small compared to a wavelength without producing mode instabilities. Thus even at high frequencies large cathodes and large emissions are possible.

10-3. Electron Orbits and the Space Charge.—An electron in the interaction space of a magnetron is acted on by a constant magnetic field parallel to the axis of the cathode, a constant radial electric field resulting from the applied d-c potential, and the varying electric field extending into the interaction space from charges concentrated near the ends of the

anode segments. In addition, the electron is part of a space charge with extreme variations in density. The resulting electron motion presents a problem of extreme complexity, and a detailed theory has not been developed. A qualitative understanding in simple terms of the processes responsible for the excitation of the magnetron is, however, possible.

Consider the simple case of a single electron in the interaction space of a magnetron in the absence of any perturbing r-f fields. In crossed

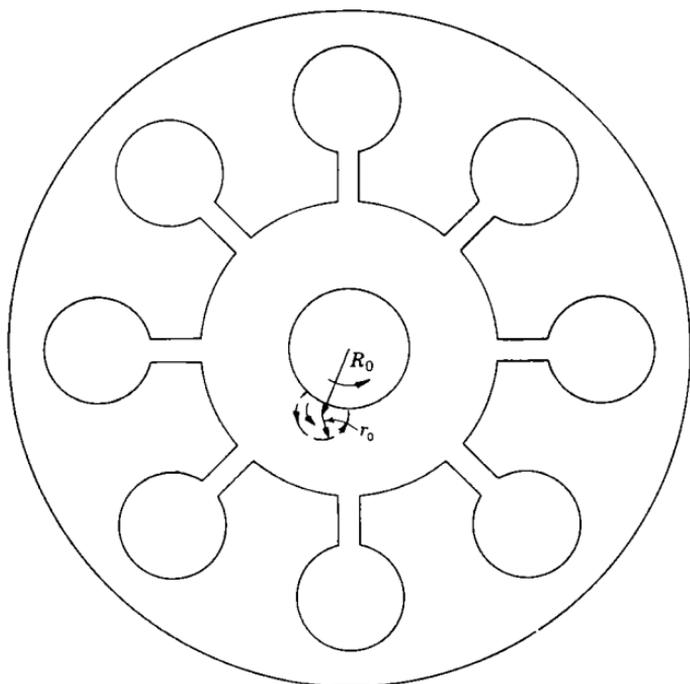


FIG. 10-11.—Path of electron in nonoscillating magnetron.

magnetic and electric fields, there is a force $-e\mathbf{E}$, due to the electric field, and another, $\frac{e}{c} \mathbf{v} \times \mathbf{H}$, due to the magnetic field, where \mathbf{E} and \mathbf{H} are the electric and magnetic field strengths, e and \mathbf{v} the charge and velocity of the electron, and c the velocity of light. The resulting motion is approximately represented by superposing a slow rotation around the cathode at nearly constant radius R_0 (the R_0 rotation), and a faster circular motion with a smaller radius r_0 (the r_0 rotation). The resultant of these two motions corresponds roughly to the motion of a point on the circumference of a wheel as it rolls around a circle somewhat smaller than the cathode in such a way that its center moves in a circle of radius R_0 (Fig. 10-11). The speed of the slow R_0 rotation is given approximately by the ratio E/H . The angular velocity of the fast r_0 rotation is determined by H

alone; it is $\omega_0 = eH/mv$. Although the angular velocity of this r_0 rotation is constant, the magnitude of r_0 depends on the initial kinetic energy of the electron and may vary for different electrons. The maximum distance any electron can proceed toward the anode ($R_0 + r_0$) in the absence of r-f oscillations is fixed by the ratio E/H , and for good operating conditions is made to be about half of the way from cathode to anode.

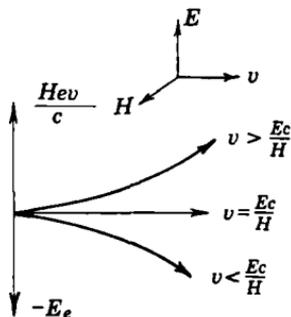


FIG. 10-12.—Paths of electrons in uniform crossed electric and magnetic fields.

The above description is exact only for the case of small r_0 (and a particular form of radial field). The picture is certainly correct in a qualitative way and this description of the electron motion may be used in an explanation of the interaction of the electrons with the alternating electric fields.

In an oscillating magnetron, the electrons pass through the r-f fields shown in Figs. 10-6a and 10-6c, and a change in their velocity results. A somewhat unusual fact is that those electrons which are speeded up have their radius of curvature reduced and return to the cathode, while those which are slowed down have their radius of curvature increased and move out toward the anode.

To make this appear reasonable, let us consider a greatly simplified case shown in Fig. 10-12. An electron moves through crossed, uniform electric and magnetic fields with a velocity \mathbf{v} that is normal to \mathbf{E} and \mathbf{H} . The directions of \mathbf{v} , \mathbf{E} , and \mathbf{H} are made to agree with those acting on the electron in Fig. 10-11. The force equation is

$$\frac{Hev}{c} = eE + \frac{mv^2}{R}, \quad (1)$$

where R is the radius of the orbit of the electron (R is positive for orbits curving down). Where the path of the electron is a straight line, the condition is obtained by letting $R = \infty$. Equation (1) then reduces to $v = Ec/H$. Inspection of Eq. (1) also shows that for $v < Ec/H$, the electromagnetic force will be reduced and the electrons will be deflected in the direction of the electric force. For $v > Ec/H$, the deflection will be in the direction of the magnetic force. The deflection that an electron suffers in this example when speeded up or slowed down thus corresponds to what happens in a magnetron and it is significant that the operating conditions are ones for which $v \sim Ec/H$. This example is given only to illustrate the complex electronics of these tubes and is obviously inac-

charge. The fields in a magnetron, however, are certainly such as to produce this separation of fast and slow electrons.

The problem is shown in more detail in Fig. 10-13. Consider an electron at point *A* at the instant for which the fields are as shown. The r-f field at this point tends to speed up the electron. As it speeds up, the radius of curvature of its path is decreased, and it will move along a path corresponding to the solid line and strike the cathode with appreciable

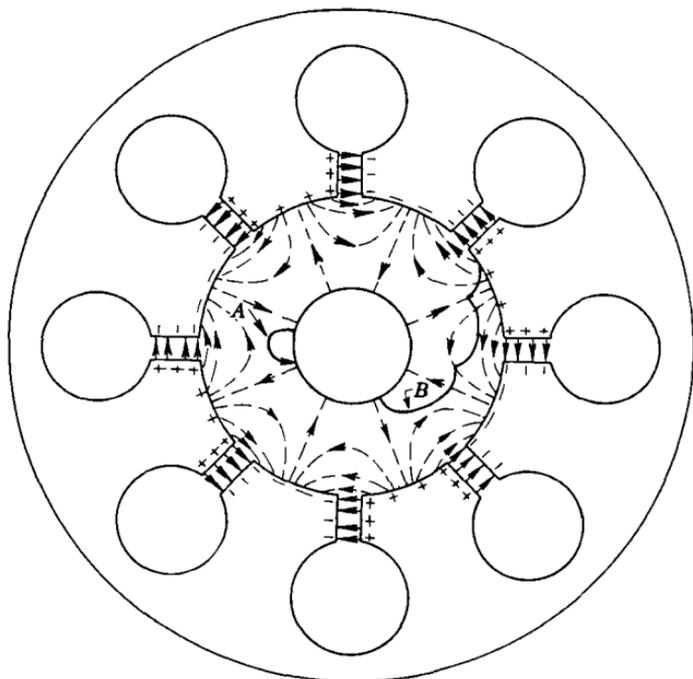


FIG. 10-13.—Paths of electrons in the presence of an r-f field.

energy. This electron is thus removed from the space charge, and plays no further role in the process except perhaps to produce a few secondary electrons from the cathode. An electron at point *B*, however, is in a decelerating r-f electric field. As a result of the reduction in electron velocity, the radius of curvature of its path is increased. If the frequency of oscillation is appropriate, this electron will always be in a decelerating field as it passes before successive anode segments. The result is that the electron, following a path of the type shown, eventually strikes the anode. Because of retardation by the r-f field, the electron gives up to the r-f field the energy gained in its fall through the d-c field to the anode.

Since the electron moves from the cathode to the anode in a very small number of oscillations, the condition that the electron keep step,

in its course around the cathode, with the variations of r-f oscillations need not be satisfied exactly. Electrons once in step with the r-f field remain in this state long enough to get to the anode, even if their angular velocity is not exactly correct. This explains why the operating conditions of magnetrons are not very critical with respect to the magnetic field, anode voltage, or other quantities which might affect the velocity of the electrons.

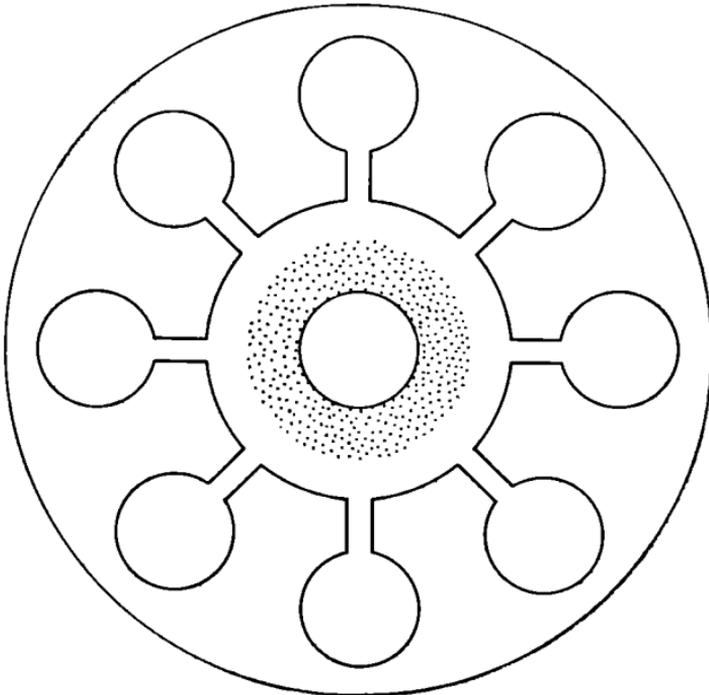


FIG. 10-14.—Space charge in nonoscillating magnetron.

Appreciable energy is associated with the r_0 rotation. This motion takes place, however, in a substantially constant r-f field, since the R_0 rotation keeps the electron in step with the variations of the r-f field. As a result, the r-f field has little or no effect on the energy associated with the r_0 rotation.

This qualitative picture shows how those electrons whose initial phase relationship is such that they absorb energy from the r-f field are eliminated at once from the space charge. Such electrons strike the cathode in the course of the first r_0 period. On the other hand, electrons which leave the cathode at such a time and place that they transfer energy to the r-f field continue around the cathode in a cycloidal path which expands toward the anode, transferring to the r-f field the energy they gain from the d-c field.

It must again be emphasized that the above description is correct in terms of general principles, but the detailed picture is doubtless more complicated. One experimental fact difficult to explain is the extraordinarily high electronic efficiencies of some magnetrons: up to 85 per cent.

In addition to describing the paths taken by individual electrons in the interaction space, it is helpful to consider the behavior of the space

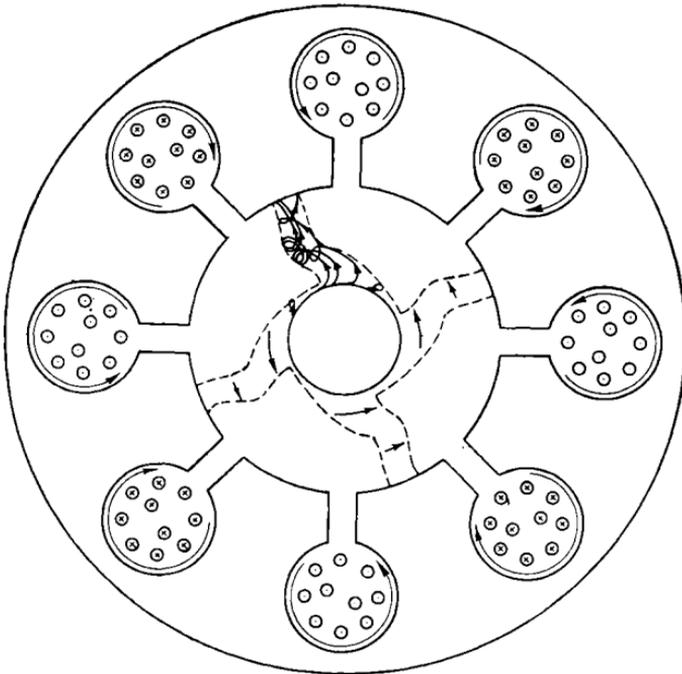


FIG. 10-15.—Space charge in oscillating magnetron.

charge as a whole. In the absence of r - f fields, the space charge forms a rotating cylindrical sheath around the cathode extending out about half-way to the anode (Fig. 10-14). The above reasoning suggests that, under the influence of the r - f fields, the electrons in this space charge which are in an accelerating r - f field travel back toward the cathode, while those in a decelerating r - f field travel toward the anode.

The result may be seen from Fig. 10-15. The rotating cylindrical sheath is distorted (for an eight-oscillator magnetron) into a smaller cylinder with four spoke-like ridges running parallel to its axis. This space charge configuration rotates with an angular velocity which keeps it in step with the alternating r - f charges on the anode segments, and the ends of these spokes may be thought of as brushing by the ends of the

anode segments and thus transferring charge from the cathode to the anode.

The r-f current set up in the oscillators is principally a displacement current produced by this rotating space charge. As the spokes of the negative space charge pass in front of an anode segment, a positive charge is induced on its surface. Half a period later, this positive charge has flowed around the back of the two adjacent oscillators to the two adjacent anode segments and the spoke of the space charge has rotated to a position in front of the next anode segment.

In addition to these displacement currents, conduction currents are produced by the flow of electrons from space charge to the anode. Electrons arrive at the anode at such a time as to constitute a conduction current approximately 90° out of phase with the r-f voltage, and thus have little effect on the oscillations.

10-4. Performance Charts and Rieke Diagrams.—Four parameters determine the operation of the magnetron; two are associated with the input circuit, and two with the output circuit. A typical set is H , I , G , and B . H is the magnetic field, I is the anode current, and G and B are the real and the imaginary parts of the r-f load on the magnetron measured at some arbitrary point in the output line. The observed quantities are three in number, usually power P , wavelength λ , and voltage V . The problem of presenting these observed quantities in terms of the four parameters is greatly simplified by the fact that the input and output parameters operate nearly independently of one another. Thus, it is possible to keep G and B (the load) fixed, and study the effect of H and I on P , λ , and V with the assurance that the nature of the results will not be greatly altered by changes in G and B . Conversely, H and I may be fixed and the effect of G and B on P , λ , and V observed. As a result of this situation, it is customary to present the operating data on two graphs. One is called the "performance chart," and shows the relationship between H , I , V , P , and λ for constant load; the other is called a "Rieke diagram," and shows the relationships between G , B , P , λ , and V for constant I and H .

Figure 10-16 is a performance chart for a typical magnetron with a frequency of about 2800 Mc/sec.

It has been customary to plot anode voltage V as ordinate, and current I as abscissa. On such a graph the lines of constant H are more or less parallel and slope upward to the right. Thus (referring to Fig. 10-16), if the magnetron is operated at a constant magnetic field, say 2100 gauss, the relations of voltage and current are given by points on the $H = 2100$ gauss line. (At 20 kv, the current drawn will be 48 amp.)

On the same chart are plotted the lines of constant power output. These are the solid lines whose form suggests hyperbolas; they show the

pulse r-f power which is obtained under varying input conditions. Thus at 20 kv and 48 amp, the power output is 470 kw. This same power can also be obtained at 25 kv and 30 amp with a magnetic field a little less than 2700 gauss.

In addition, it is customary to add curves of constant efficiency obtained directly from the above data. These lines of constant efficiency are the dotted lines looping up and to the right on the diagram.

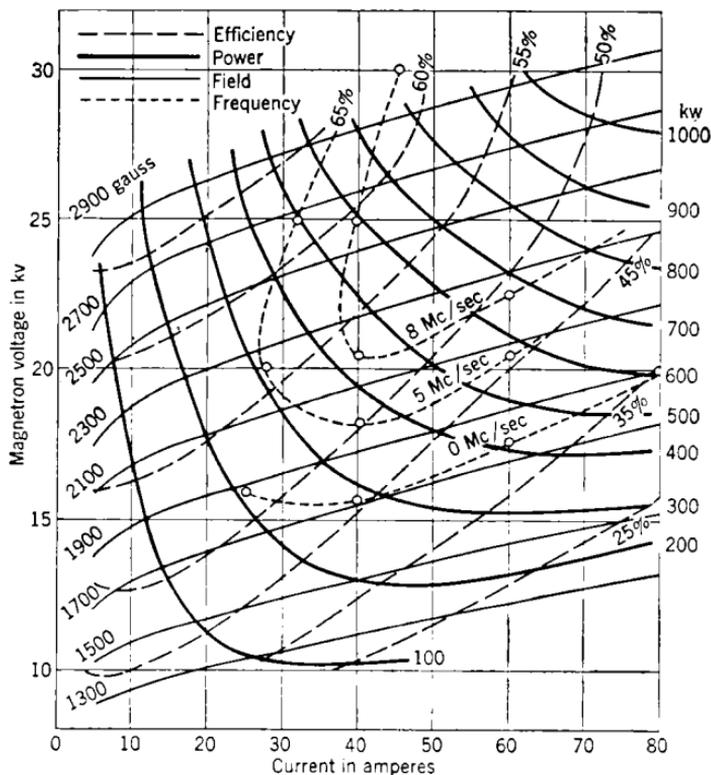


FIG. 10-16.—Typical performance chart of a magnetron (4J31). Frequency is expressed in terms of difference from a base frequency.

It is possible to add to this chart lines of constant frequency, so that the variation of frequency with input parameters may be studied. This information is useful in establishing limits on the variation of current during a pulse. (See the treatment of pulse length limitations in Sec. 10-5.) The dashed lines in Fig. 10-16 are contours of constant frequency. In this case, for currents greater than 40 amp, they are nearly parallel to the lines of constant magnetic field; this is an ideal condition, since changes in current produce no change in frequency.

Performance charts like that shown in Fig. 10-16 are indispensable

in the selection of proper operating conditions for a magnetron. The input voltage, input current, and magnetic field are chosen by making a compromise among such factors as peak power output, efficiency, allowable magnet weight, input impedance, and stability of operation. All the necessary information can be obtained from a good performance chart.

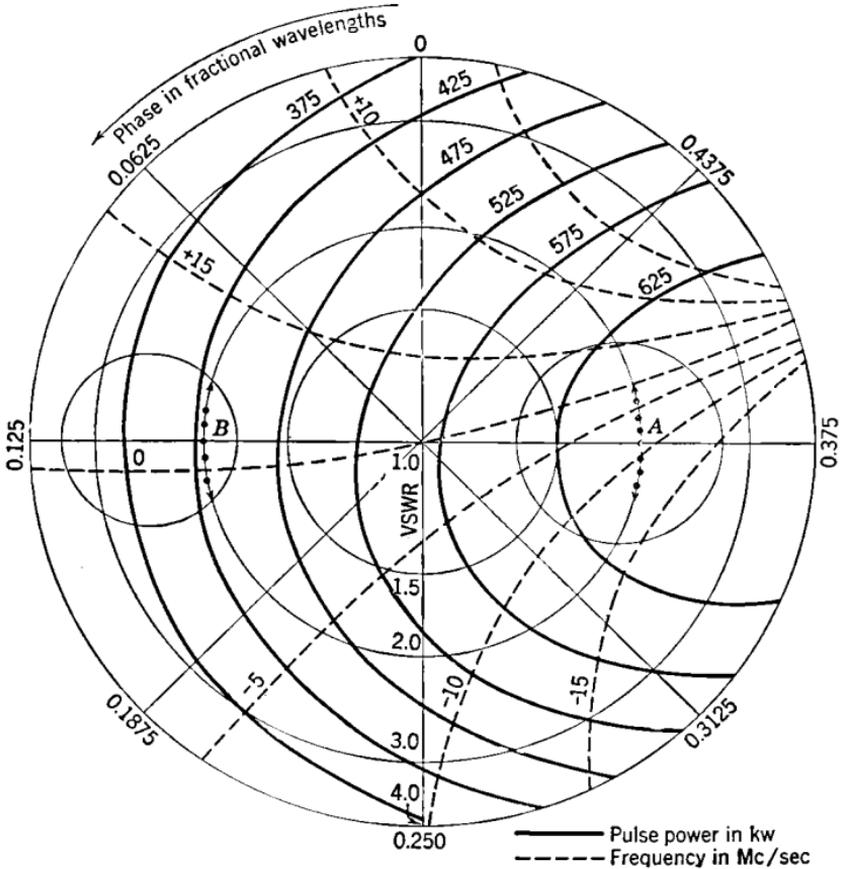


FIG. 10-17.—Rieke diagram of a magnetron. The circles to the right and left of center are transformed VSWR = 1.5 circles.

It is advisable to select an operating point that is not near regions of anomalous behavior. At low currents (below 15 amp, Fig. 10-16), or at excessive powers (above 800 kw, Fig. 10-16), instability is likely to occur. Below 1700 gauss the efficiency is obviously unsatisfactory. Considering Fig. 10-16 as an example, the operating point should be between 30 and 60 amp and 17 and 25 kv.

For magnetron-magnet combinations (Fig. 10-4), the magnetic field is fixed by the tube manufacturer, and the performance chart is reduced to a single magnetic-field line on a general performance chart. In such a case, it is only necessary to establish an operating voltage or current.

The performance of a magnetron in terms of its output parameters, or r-f loading, is presented on a Rieke diagram (Fig. 10-17). It would appear useful to express the r-f loading in terms of the resistance and reactance presented to the magnetron at the output loop. Since these quantities are difficult to determine experimentally, the Rieke diagram is a compromise between these quantities and others which can be obtained with ease experimentally. At microwave frequencies, it is customary to determine the constants of a load by observing the phase and magnitude of the standing waves set up by it, and the Rieke diagram is designed to use these experimental data directly. Since it is impractical to determine the r-f loading at the output of the magnetron, an arbitrary point in the output circuit is selected to which all measurements can be referred. The desired range of r-f loading is obtained by adjusting a tuner until the desired phase and magnitude of standing waves is indicated by a sliding pickup probe.¹ The independent variables VSWR ρ and phase ϕ are used as coordinates of a polar diagram. Power output and frequency are measured for various points on the diagram, and the data used to construct contours of constant frequency and constant power.

Magnetron manufacturers usually furnish diagrams such as this for every tube type, and these give information of considerable importance to system designers. Consider the effect of presenting such a mismatch to the magnetron that a VSWR of 2 results, and suppose further that the phase of this mismatch is such that the voltage minimum corresponds to point *A* (Fig. 10-17). The result will be a power output for the specified input conditions of about 650 kw. If the phase of the VSWR is changed, say by increasing the line length between the magnetron and the mismatch until point *B* is reached, the power output falls to 425 kw. The efficiency of the magnetron at point *A* is thus 50 per cent greater, but operation at this loading is unsatisfactory for another reason. As representative of a general class of load instability, consider a change in phase angle about the loading *A* of $\pm 7.5^\circ$ (dotted arrows) which might result from the turning of an imperfect rotary joint. The power output will be essentially unaltered, but a maximum change in frequency of 12 Mc/sec occurs. If a heavy loading corresponding to point *A* is used, frequency shifts of this magnitude, and greater, are frequently encountered and may seriously affect the system operation. At point *B*, however, corresponding to light loading, a phase shift of $\pm 7.5^\circ$ results in only a 3 Mc/sec frequency shift. This illustrates the compromise that must be made between efficiency and frequency stability. Magnetrons are usually designed with an output coupling such that the center of the Rieke

¹See *Technique of Microwave Measurements*, Vol. 11, Radiation Laboratory Series, Sec. 2-5; see also Sec. 11-1 of this volume.

diagram, which corresponds to a matched load, represents a reasonable compromise between efficiency and frequency stability.

It is possible to adjust the loading on the magnetron to any reasonable value by the suitable use of r-f transformers (Sec. 11-1) in the output line. As an example, suppose it is desirable to operate the magnetron represented by Fig. 10-17 at a point of high efficiency and low frequency stability corresponding to point *A*. This can be accomplished by introducing a transformer which sets up a 2-to-1 VSWR and making its distance from the magnetron such that the phase of this VSWR corresponds to point *A*. By moving this transformer along the line in either direction one-quarter wavelength, operation corresponding to point *B* can be obtained. In comparing the effect of different loads corresponding to various points on the Rieke diagram, it should be realized that these points represent transformations that reduce the size of a circle of constant VSWR as its center is moved away from the center of the diagram. In Fig. 10-17 the circles about points *A* and *B* represent the $VSWR = 1.5$ circle when displaced different distances from the center of the diagram.

The load points *A* and *B*, with their associated variations in load in the above example, are especially simple cases. In general, the load variations correspond to a very irregular path on the Rieke diagram whose behavior is unpredictable. A safe policy in design is to estimate the maximum variation in VSWR to be expected from the r-f circuits, and to employ a loading of the magnetron which does not produce a frequency change too large to be accommodated by the radar receiver even when this variation in VSWR is of such a character as to produce the maximum possible frequency shift.

Appreciation of the effects of the r-f load on the performance of magnetrons has contributed more than any other single factor to magnetron reliability. As a corollary, it is also true that many troubles attributed to magnetrons result from a failure to use properly the information provided by a Rieke diagram.

10-5. Magnetron Characteristics Affecting Over-all Systems Design.

One of the shortcomings of microwave magnetrons is their limited adaptability to different requirements. This circumstance has forced the design and production of an extremely large number of tube types. Although the development of microwave magnetrons in this country began only in late 1940, there now exist over 100 distinct types of magnetrons, despite early and continuing attempts at standardization. During the past war it was true, almost without exception, that each new radar system made new demands on the magnetron and required the development and production of a new type. This has not been necessary in the case of conventional types of tubes, since the associated circuit elements which are largely responsible for over-all performance lie exter-

nal to the tube and are accessible to change. In microwave magnetrons, the circuit elements are an integral part of the tube and must be incorporated with great care into every new design.

Thus any radar system designed to meet a new set of conditions or to operate on a new frequency will require the development of a new magnetron type, or at least a critical evaluation of the characteristics of existing types. Since the over-all characteristics of the radar system are so closely related to, and restricted by, the performance of the magnetron, a general knowledge of the important characteristics of magnetrons is essential.

Listed below with a discussion of each are the characteristics of magnetrons of particular importance to system design. These characteristics are not usually independent of one another and their relationships are also considered.

Wavelength Scaling.—Since the wavelength of the radiation from a magnetron is fixed or at best variable over a limited range, operation on different wavelengths requires different tubes. To a first approximation, magnetrons of different wavelength are derived from one another by a simple over-all scaling process. All essential dimensions of the tube are altered by the scaling factor $\alpha = \lambda/\lambda_0$, where λ is the new wavelength desired and λ_0 is the wavelength associated with the original dimensions.

If this is done, the new tube at wavelength λ will operate at the original voltage and current, and at a magnetic field $H = \frac{1}{\alpha} H_0$, where H_0 is the operating magnetic field of the original magnetron. The power input, and thus the power output, increases with increasing wavelength. A rough rule is: *The pulse power output (or input) of scaled magnetrons varies as the square of their wavelength.* The change in the size of the tube with wavelength is the basis of this rule. The pulse power input is often limited by cathode emission and, since the cathode area is proportional to λ^2 , pulse power input is also proportional to λ^2 . Similar reasoning shows that if the pulse power output is limited by r-f voltage breakdown within the tube, the same variation of power with wavelength is to be expected.

This rule is an important one from the standpoint of system design. At any given time it may not be exact, because special emphasis may have been given to obtaining high peak powers at a particular wavelength and a better design evolved as a result. In the long run, however, the validity of the rule is reestablished, because any new design can, within limits, be used to advantage at other wavelengths.

Pulse Power.—The most outstanding characteristic of pulsed microwave magnetrons is their extremely high pulse power output, made possible by the very large emission yielded by oxide cathodes when pulsed,

and the high efficiency of magnetrons even at very short wavelengths. The pulse powers available extend over a range of 10^5 (0.02 kw to 2000 kw).

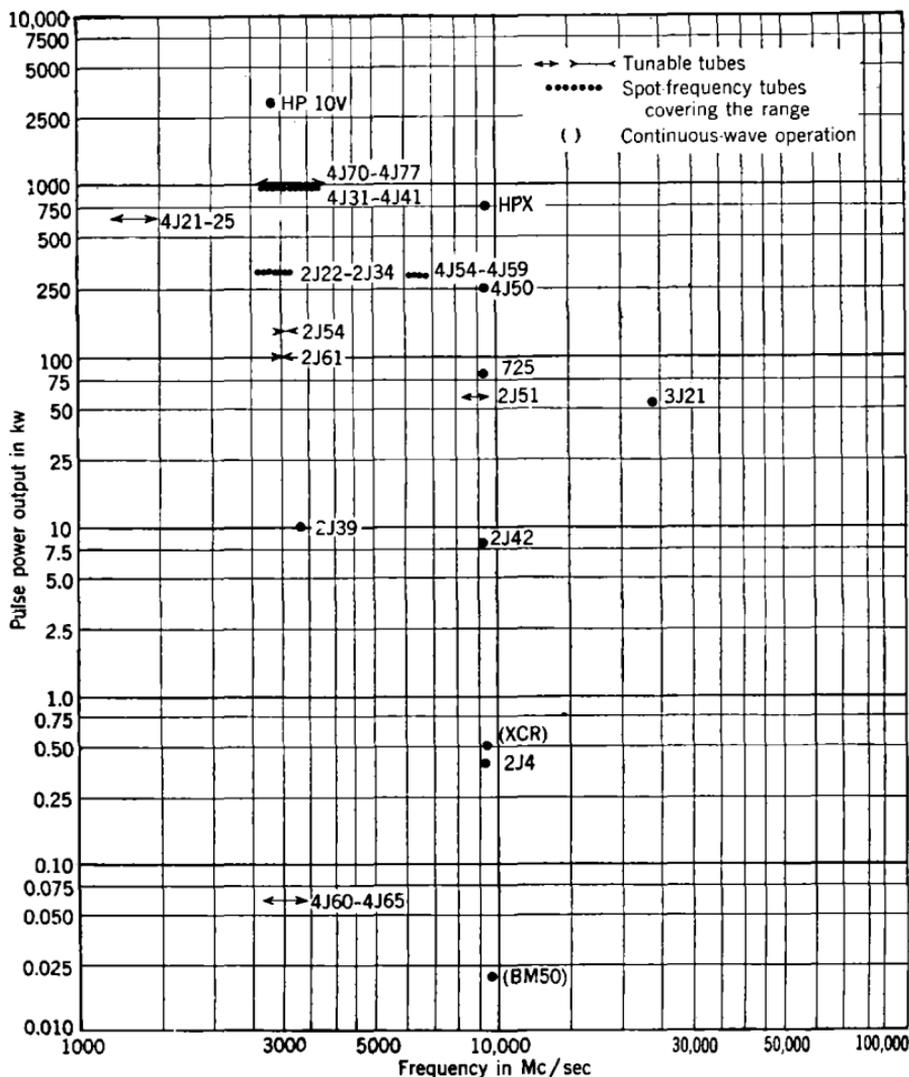


FIG. 10-18.—Diagram showing power and frequency distribution of representative microwave magnetrons developed up to 1945.

The demands of microwave radar resulted in a rather extensive development of magnetrons whose frequencies are concentrated more or less into two bands. Figure 10-18 shows on a logarithmic chart the frequency and pulse power of magnetrons that have been produced in

appreciable numbers and thus constitute well-tested designs. Fixed-frequency tubes are indicated by a dot and tunable tubes by a line whose length shows the tuning range. Whenever possible the magnetrons are identified by their RMA type numbers.

The values of pulse power given in this chart are conservative ones corresponding to reliable operating conditions. Pulse powers considerably in excess of those given for the highest power tubes have been

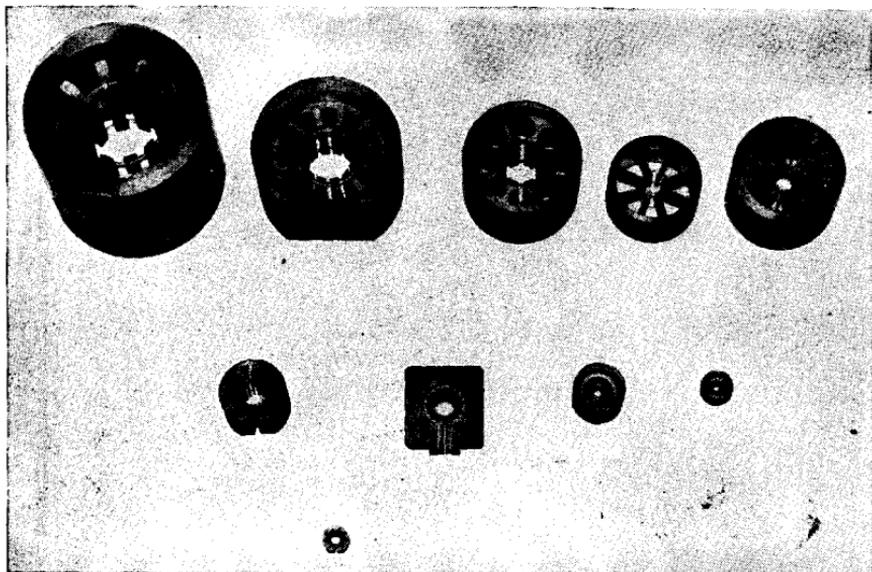


FIG. 10-19.—An array of anode blocks showing the effect of wavelength and pulse power output. Top row—10 cm, pulse power 2500 to 0.1 kw; second row—3 cm, pulse power 600 to 0.025 kw; single block—1.25 cm, 80 kw.

observed. At 3000 Mc/sec, for example, 4500 kw have been obtained from the HP10V, and the British report 3700 kw peak at an average power level of 5 kw from a similar tube, the BM735.

The general rule that maximum pulse power is proportional to wavelength squared is illustrated by the falling-off of power at high frequencies. Exceptions to this rule are the magnetrons in the 20- to 30-cm range, whose pulse power output is not as high as might be expected. This particular situation resulted from a lack of need for very high powers in this wavelength range.

Figure 10-19 illustrates the effects of wavelength and power output on the design of anode blocks. In the top row, from left to right, are 10-cm anode blocks for magnetrons with pulse power outputs of 2500, 1000, 250, 5, and 0.1 kw. The reduction in anode diameter with pulse power is evident. In the second row are 3-cm anode blocks with pulse power

outputs of 600, 200, 80, and 0.025 kw. The single anode block at the bottom is from a 1.2-cm 80-kw rising-sun magnetron. Figure 10·20 shows one variety of a high-power (1000-kw output) magnetron. This design has a rather long anode with heavy double-ring straps and a coaxial output.

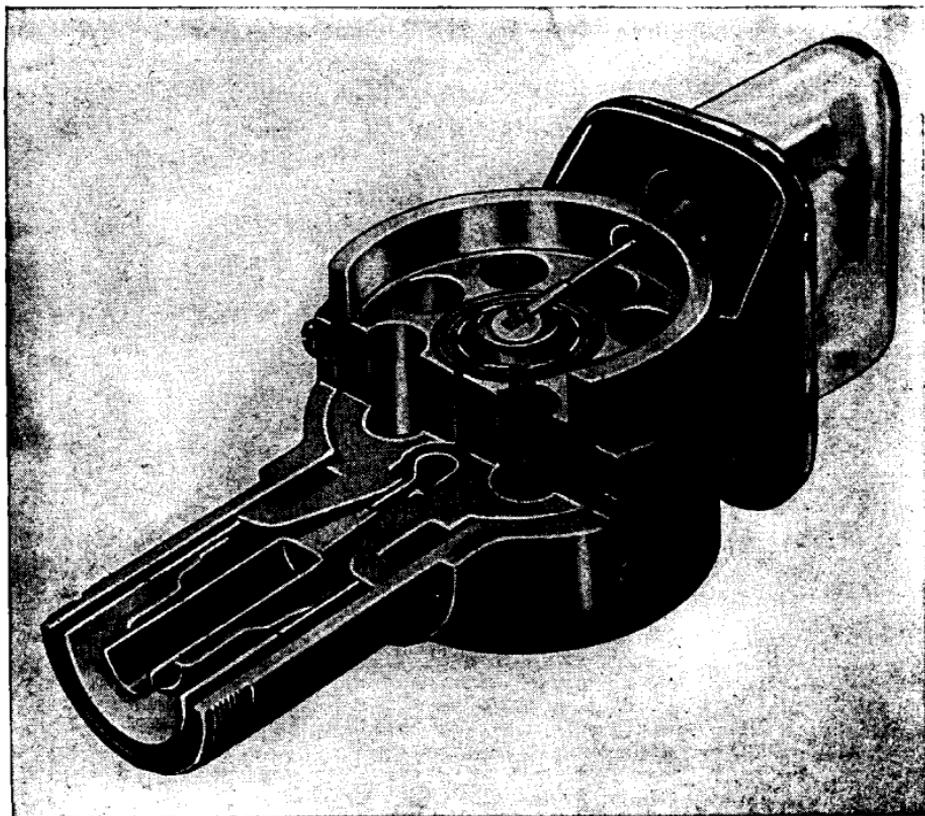


FIG. 10·20.—A high-power magnetron (type 720); 10 cm, 1000-kw pulse power output. (Courtesy of Bell Telephone Laboratories.)

Average Power.—In Sec. 10·2 it was stated that the cathodes of oscillating magnetrons are subjected to back bombardment by some of the electrons in the surrounding space charge. This bombardment produces heating of the cathode surface which amounts to about 5 per cent of the average power input to the magnetron. If this back-bombardment power exceeds the normal heater power of the cathode (which is frequently turned off while the magnetron is oscillating) abnormally short tube life results. Back bombardment usually limits the average power output. This limit can be raised by designing cathodes with very large thermal dissipation, but the average power of magnetrons designed for radar use rarely exceeds 0.2 per cent of the pulse power output. Table

10-1 shows the average and pulse power output of selected magnetron types representing the microwave range of frequencies.

TABLE 10-1.—AVERAGE AND PULSE POWER OUTPUTS OF MICROWAVE MAGNETRONS

RMA type No.	Frequency, Mc/sec	Maximum average power output, w	Pulse power output, kw	Maximum pulse length, μ sec
4J21	1,180	800	800	6.0
4J73	3,100	600	1000	2.5
725	9,400	80	80	2.5
3J21	24,000	25	55	0.5

Efficiency.—The purpose of a magnetron is to convert d-c power into a-c power at very high frequencies. Magnetrons can perform this conversion with an efficiency as high as 85 per cent, which compares favorably with the efficiency of d-c to 60-cps a-c converters. Magnetron efficiencies customarily lie between 30 and 50 per cent.

The output efficiency is the product of what is called the "electronic efficiency" and the "circuit efficiency." The electronic efficiency is the fraction of the d-c energy input which the electronic space charge converts into r-f energy. The circuit efficiency is the fraction of the r-f power going into the resonant system which appears as output power; the remainder is wasted as heat because of copper losses. Electronic efficiencies depend on the magnetic field, on r-f loading, and on the cathode-to-anode radius ratio. Electronic efficiencies of about 70 per cent are realized over a very large range of frequencies. Circuit efficiencies depend on the frequency, since this determines the conducting skin depth which in turn governs the copper losses. The C/L ratio of the resonators and the r-f loading also affect the circuit efficiency. Circuit efficiencies for the usual r-f loading vary from close to 100 per cent for the 1000-Mc/sec range to 50 per cent for the 25,000-Mc/sec range.

Realization of high magnetron efficiency affects system design in two ways. High efficiency accompanies high magnetic fields which require heavy magnets. Higher efficiencies can also be obtained by closer coupling (Sec. 10-4), which at the same time decreases the frequency stability. Experience has shown that it is expedient to provide sufficient magnetic field for obtaining good efficiencies even at the expense of some increase in weight. The variation of efficiency with magnetic field can be found from the performance chart. As discussed in Sec. 10-4, considerable efficiency should be sacrificed for the sake of frequency stability. The relationship of these two quantities can be obtained from a Rieke diagram.

Pulse Voltage.—The input impedance¹ of most pulsed magnetrons lies between 700 and 1200 ohms. For example, magnetrons with an input of 260 kw (100-kw output) require a pulse voltage of about 15 kv; inputs of 2500 kw (1000-kw output) require 30 kv, and one magnetron designed for 6000-kw input (2500-kw output) requires a pulse voltage of 50 kv.

Serious attempts have been made to design magnetrons that would operate at lower input impedances because such tubes would simplify the design of line-type pulsers. These attempts have been unsuccessful, and 350 ohms was the lower limit in mid-1946. The design of magnetrons with high impedances is, however, easily achieved, and tubes with input impedances as high as 10,000 ohms have been produced. Table 10-2 shows the input pulse power, pulse voltage, and input impedance of existing 10-cm magnetrons with input pulse powers ranging from 2 to 6000 kw. Power outputs can be estimated by assuming an efficiency of 40 per cent.

TABLE 10-2.—INPUT CHARACTERISTICS OF MICROWAVE MAGNETRONS

Tube No.	Input pulse power, kw	Pulse voltage, kv	Input impedance, ohms
4J60	2	1.5	1125
2J38	25	5	1000
2J32	250	15	900
4J31	2500	30	360
HP10V	6000	50	415

By varying the magnetic field in which the magnetron operates, any given design can be made to operate satisfactorily over a range of pulse voltage of roughly 50 per cent. For satisfactory operation, however, the tube will exhibit about the same input impedance values. Exact figures can be obtained from performance charts.

Pulse-length Limitations.—The early experiments of the British with high-power pulse techniques revealed a characteristic of oxide cathodes which is responsible in large measure for the high pulse power of magnetrons. It was found that under pulsed conditions oxide cathodes can emit as much as 20 amp/cm² as compared to about 0.2 amp/cm² for d-c emission.

More recently, currents as high as 100 amp/cm² have been obtained. This current is carried partly by primary electrons and partly by secondary electrons liberated from the oxide cathode by back bombardment.

¹ The word *impedance*, as used in this section, means the voltage-current ratio of a magnetron at the operating point. The *dynamic impedance*, the slope of the voltage-current curve near the operating point, is very much lower, usually around 100 ohms. See Sec. 10-8.

Whatever the source of the emission, if too large a current is drawn for too long a time, sparking and other instabilities result. The exact relationship between peak cathode emission and pulse length depends on the type, temperature, and age of the cathode. To a fair degree of approximation, the maximum permissible peak emission varies inversely as the square root of the pulse length: $I_{\max} \sim 1/\sqrt{\tau}$. Thus a cathode which will emit 20 amp for a pulse duration of 2 μsec will probably emit 40 amp during a 0.5- μsec pulse. In consequence, greater energy per pulse can safely be obtained for the longer pulses.

Pulse durations greater than 5 μsec are rarely employed when magnetrons are used as transmitter tubes. Frequency modulation during the pulse becomes a serious problem for longer pulses, even if sparking troubles are overcome. Pulse durations as short as 0.25 μsec have been used successfully, particularly with high-frequency magnetrons whose starting times are short.

Tuning of Magnetrons.—To change the frequency of a magnetron more than a few megacycles per second requires that a change be made in the resonant circuits of the anode. Either the effective capacity or effective inductance must be varied, and, since the resonant circuits are within the evacuated portion of the tube, variation of either of them is a troublesome problem.

For this reason early magnetrons were not tunable, and only a later need for increased flexibility of radar systems forced the design of tunable tubes. The practical advantage of tunable over fixed-tuned magnetrons is obvious. If operation on a number of frequencies is contemplated, a single tunable magnetron can replace a whole set of fixed-frequency magnetrons, and only with a tunable magnetron is it possible in general to obtain r-f power at a specified frequency. The performance characteristics of tunable magnetrons are equivalent to those of the corresponding fixed-frequency tubes, and there is thus no reason, except availability, for not using them.

Tuning of the higher-frequency microwave magnetrons is accomplished by inserting conducting cylinders into the inductive portion of each resonant cavity, thus decreasing the effective inductance. This construction, shown in Fig. 10-21, provides a tuning range as high as 12 per cent.

At frequencies lower than about 5000 Mc/sec, the magnitude of the longitudinal displacement required in inductive tuning becomes inconvenient, and other tuning methods are adopted. Figure 10-22 shows a "C-ring" type of tunable magnetron, in which a conducting surface can be moved toward or away from the straps and capacitive portion of the resonant cavities, thus changing their effective capacity. The disad-

vantage of this method is that sparking may occur within the tuning arrangement at high pulse powers.

An unsymmetrical type of tuning which has advantages for high-power operation is shown in Fig. 10-23. The frequency of the single tuning cavity is changed by distorting the diaphragm. Since the tuning cavity is tightly coupled to one of the resonant cavities, and since all the

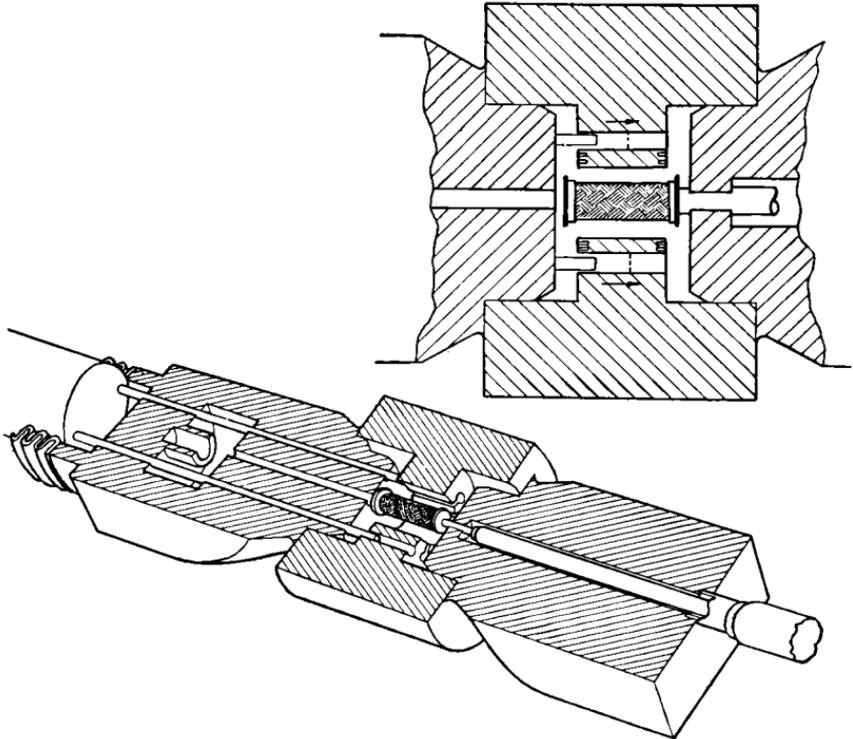


FIG. 10-21.—Magnetron with sprocket tuning. The change in frequency is accomplished by changing the inductance of the resonators.

resonant cavities are very closely coupled to one another by the straps, the oscillating frequency of the entire magnetron is altered. This unsymmetrical type of tuning distorts the electric field patterns within the magnetron and therefore limits its effective tuning range to about 6 per cent. Advantages of the method are its mechanical simplicity and its ability to handle very high pulse powers. Figure 10-24 shows output power as a function of tuning adjustment.

Various other tuning schemes have been tried, and still others will be devised, but the three types described above represent the basic methods.

Wartime experience with radar demonstrated that tunability was very desirable but that extreme tuning range was not.

Frequency Pulling by the R-f Load.—Magnetrons are self-excited oscillators whose frequency depends on output loading. In the process of scanning by the radar antenna, it is inevitable that variations in the magnetron loading will occur as a result of changing reflections from rotary joints, antenna housings, and large near-by reflecting objects. Such changes in loading produce changes in frequency which may be

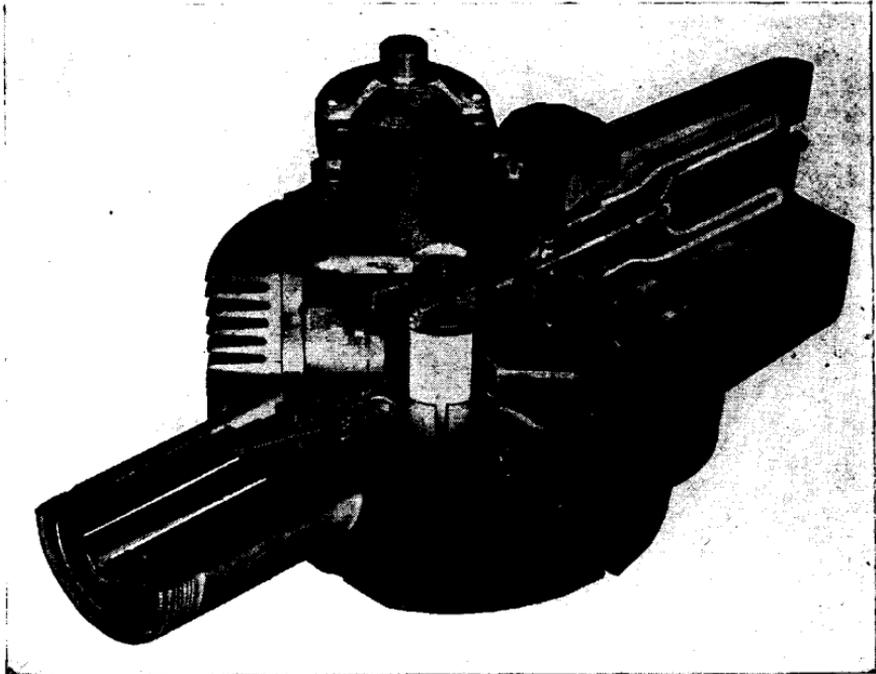


FIG. 10-22.—Cutaway view of a 1000-Mc/sec tunable magnetron showing the C-ring tuning mechanism and the slot-type resonators. (Courtesy of Bell Telephone Laboratories.)

large, and this phenomenon is a source of considerable trouble. It necessitates careful engineering of the entire r-f system from magnetron to antenna, and also affects receiver design.

The problem is attacked in three ways: (1) the magnetron is designed so that its frequency change with changing r-f loading is small; (2) the variations in r-f loading are reduced by careful design and construction of the r-f components; (3) the radar receiver is made insensitive to changes in frequency by making its pass band broad or by incorporating automatic frequency control. It is the purpose of this subsection to discuss only the first of these methods of attack, that is, magnetron design.

The effect of loading on the frequency of a magnetron is usually expressed by what is called the *pulling figure*, which is defined as *the total frequency excursion which results when a standing-wave ratio of 1.5 in voltage*

is presented to the magnetron and varied in phase over at least a full half cycle (see Fig. 10-17). The VSWR of 1.5 was selected for this definition since it is in the range usually encountered in radar systems. Pulling figures for magnetrons with comparable loading increase with frequency, since the same fractional change in frequency corresponds to a larger number

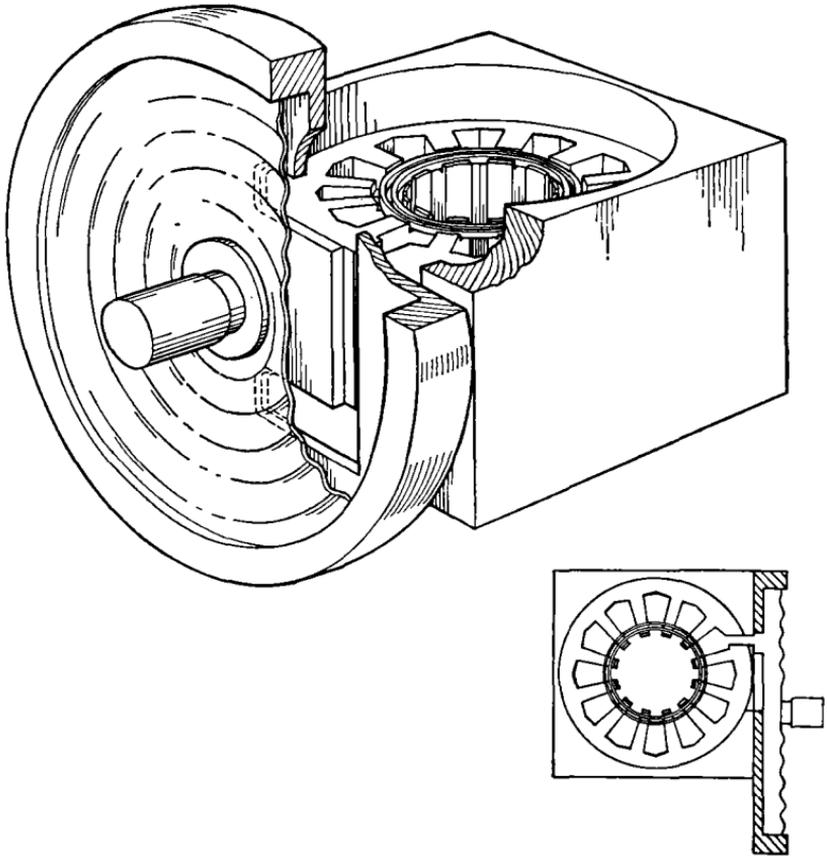


FIG. 10-23.—Magnetron with unsymmetrical tuning cavity.

of megacycles at higher frequencies. Thus, similar magnetrons at 3000, 10,000, and 30,000 Mc/sec would have pulling figures of, say, 5, 16, and 50 Mc/sec. Since receiver bandwidth is not changed substantially as the radar frequency increases, frequency stability becomes a more important consideration at shorter wavelengths.

The effect of loading on frequency stability and methods of obtaining various loadings have been considered in Sec. 10-4. As stated there, the requirements of any system usually make necessary some sort of com-

promise between efficiency and frequency stability and this compromise is arrived at from a study of the Rieke diagram.

It is possible in some cases to combine with the magnetron proper a high- Q cavity coupled to it in such a way as to reduce the pulling figure by a large amount with little or no loss in efficiency. The ratio of the pulling figures with and without the high- Q cavity for equivalent operating conditions is called the "stabilization factor," which may be as large as 10 but is usually between 2 and 5. The addition of this stabilizing

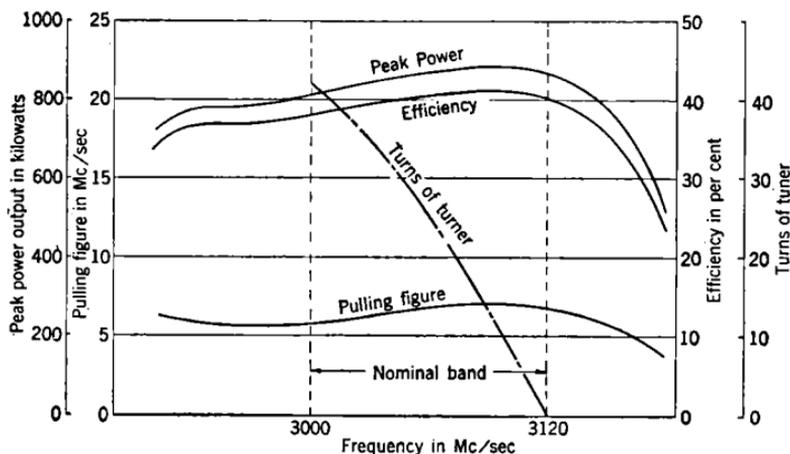


FIG. 10-24.—Characteristics of the 4J70 series.

Magnetic field = 2700 gauss

Pulse current = 70 amp

Pulse duration = 0.8 μ sec

PRF = 400 pps.

cavity has the disadvantage, however, of reducing the tuning range of the magnetron; further, it may aggravate mode instabilities. If constant-frequency operation is required, stabilized magnetrons should be seriously considered.

Frequency stabilization has been incorporated into only a few magnetrons. A good example is the 2J41, which has a pulse power output of 500 watts at 9310 Mc/sec and a pulling figure of 1.5 Mc/sec with a tuning range of 0.7 per cent. The stabilization factor is 10. The 10-cm series 4J70 to 4J77 is stabilized by a factor of 1.5 and the 1.25-cm 3J21 by a factor of 2. High-power magnetrons with high stabilization have not been developed.

Weight.—The development of airborne radar placed great emphasis on the reduction in weight of all components. As a result, magnetron-magnet combinations were produced in which every effort was made to reduce the total weight. This development has been so successful, particularly for the higher-frequency magnetrons, that the weight of a

magnetron-magnet combination is now a small fraction of the weight of the pulser required to drive it. The weight of magnetrons and their magnets has thus ceased to be a critical design consideration.

10-6. Magnetron Characteristics Affecting Pulser Design.—The magnetron, standing as it does between the pulser and the r-f system, imposes restrictions on the design of both these components. Pulser design in particular has been complicated by some very strict requirements which the magnetron places on pulse shape and voltage regulation. These requirements arise from undesirable magnetron characteristics which it has so far been impossible to remove in the design of the tubes. They may perhaps be eliminated from magnetrons in the future.

The stringency of the conditions that the magnetron places on the pulser or on the r-f components increases rapidly as the maximum operating conditions are approached. Thus the surest way to obtain reliable performance is to operate the tubes at conservative ratings.

Little need be said here about the interaction between the magnetron and the r-f components. The relationship between frequency stability, VSWR of the r-f system, and the magnetron pulling figure has been covered in Sec. 10-4. The related subject of long-line effect is discussed in Sec. 11-1.¹

Because the interaction between the pulser and the magnetron has only recently been understood, many serious difficulties have been alleviated in the past by the unsatisfactory process of cut-and-try. Four characteristics of magnetrons are largely responsible for these troubles.

Change of Frequency with Current.—The input impedance of a magnetron varies with voltage as shown in Fig. 10-29. The average input impedance is usually in the range of 400 to 1000 ohms. Very small changes in anode potentials produce large changes in anode current. For example, the performance chart of a 4J31 magnetron (Fig. 10-16) shows that, at a magnetic field of 2300 gauss, a change of voltage from 20 kv to 22.5 kv causes a change in current from 20 to 50 amp. Variations in voltage during a pulse must be kept within small limits to prevent large current variations and consequent distortion of the r-f pulse shape.

The rate of change of frequency with anode current, $d\nu/dI$, determines the amount of frequency modulation and undesirable distortion in the energy spectrum of the pulse which will be caused by current variation during the pulse. In the usual operating range for most magnetrons, $d\nu/dI$ is about 0.1 Mc/sec per amp, but it can be as large as ± 1 Mc/sec per amp. The exact value depends to a considerable degree on the operating point selected. To illustrate this problem, consider a magnetron with $d\nu/dI = 0.4$ Mc/sec per amp driven by a 1- μ sec pulse which "droops" from 55 to 50 amp. This 5-amp change in current will

¹ See also *Microwave Magnetrons*, Vol. 6, Sec. 7-2, Radiation Laboratory Series.

produce a frequency modulation of $5 \text{ amp} \times 0.4 \text{ Mc/sec}$ per $\text{amp} = 2 \text{ Mc/sec}$. Since the bandwidth of the receiver for a $1\text{-}\mu\text{sec}$ pulse would be about 2 Mc/sec , frequency modulation of this magnitude will result in a serious loss of received energy. When longer pulses and correspondingly narrower bandpass receivers are used, this problem becomes much more critical and may place very severe requirements on the flatness of the current pulse.

Instabilities.—It has so far been found impossible to construct magnetrons that do not occasionally present to the pulser either a very low impedance as a result of a gas discharge (sparking) within the tube, or a very high impedance due to a failure of the magnetron to oscillate in the proper manner (mode-changing). Either of these events may occur only once or twice per million pulses, but when such an event does take place, voltage and current surges are frequently produced in the pulser which may cause failure of some component. The pulser designer must therefore over-design components and provide special protective circuits to guard against events that may happen only once in a million pulses. All magnetrons change mode or spark occasionally, but the frequency of sparking or mode-changing can be reduced by operating at moderate peak anode currents and short pulse durations. Magnetron performance and life are materially increased if the pulser design is such that, in the event of sparking or mode-shifting, an excessive discharge does not take place through the magnetron.

These two types of instability, mode-changing and sparking, are difficult to distinguish in practice, since mode, changing usually produces sparking and vice versa. In spite of this, it is advantageous to consider them separately since the cure for each is quite specific and distinct.

Sparking is an internal discharge in the magnetron which arises as a consequence of the generation of bursts of gas within the tube. The gas may be liberated from the anode or from the cathode; in either event the frequency of the phenomenon is multiplied by an increase in anode voltage, anode current, or pulse length. Operation of a magnetron under conditions which exceed specifications for any of these quantities results in a very rapid increase in sparking rate. Sparking limits the maximum pulse length at which magnetrons can be operated.

Of all the modes of oscillation possible in magnetrons, only the π -mode is ordinarily used but all magnetrons will sometimes oscillate in an unwanted mode, or alternate erratically between two modes. This tendency is responsible for some of the most troublesome problems in magnetron and pulser design. Considerable progress toward an understanding of the phenomenon has now been made,¹ and when mode-

¹ F. F. Rieke and R. Fletcher, "Mode Selection in Magnetrons," RL Report No. 809, Sept. 28, 1945. *Microwave Magnetrons*, Vol. 6, Chap. 8, Radiation Laboratory Series.

changing difficulties arise a systematic attack is possible. The important conclusions of this study are given here.

Mode changes are of two types. One type is caused by anode currents so high that they exceed the conditions for oscillation and cause transition into a state of nonoscillation or into another mode of oscillation. The second type arises from a failure of the oscillations in the desired mode to build up rapidly enough with respect to the voltage rise at the start of the pulse. The first type is encountered usually in lower-power c-w magnetrons, rarely in pulsed tubes. The second type, called "mode-skipping," is common to nearly all high-power pulsed magnetrons and thus is of importance here.

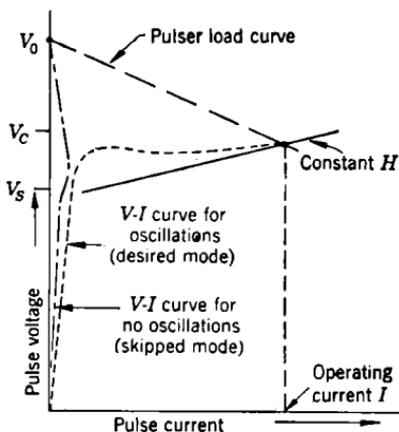


Fig. 10-25.—V-I plot of a magnetron illustrating the problem of mode skipping.

V_0 = magnetron starting voltage
 V_c = critical voltage above which oscillations will not start
 V_s = pulser voltage with no load.

(V_s) must be maintained at large currents. This calls either for a very low-impedance pulser or for a high pulser voltage V_0 at zero load current. Consider the case illustrated in Fig. 10-25. To achieve the final pulse current I , the pulser must be adjusted to a value of V_0 considerably higher than the starting voltage V_s . Under proper conditions, when the pulse is applied to the magnetron, the V-I curve is that shown dotted in Fig. 10-25. The pulse voltage never rises above the critical value V_c , because, as it reaches V_s , the magnetron draws current and loads the pulser. If oscillation fails to start promptly after the pulse voltage has reached V_s , the voltage continues to rise and may exceed V_c before current is drawn. Then, since the conditions for oscillation no longer exist, the voltage rises to V_0 and oscillations in the desired mode cannot take place. This is what occurs when a magnetron skips modes. The V-I trace for such an event is also shown in Fig. 10-25. The condition for oscillation in the desired mode occurs when the magnetron starts to oscillate and to draw current in the time interval taken for the pulse voltage to pass from V_s

to V_0 . The second type, called "mode-skipping," is common to nearly all high-power pulsed magnetrons and thus is of importance here.

Mode-skipping is dependent not only on the characteristics of the magnetron but also on the characteristics of the pulser. In practice, it has been necessary to consider the magnetron and pulser as a unit.

For given magnetic field and r-f loading, there is only a limited range of anode voltage over which the magnetron will build up oscillations in the desired mode. On the other hand, if high-power oscillations are to be maintained, a plate voltage at least as high as this starting voltage

to V_c . For this particular case the probability that the magnetron will start is increased if (1) the magnetron has a short starting time and has a large voltage interval $V_c - V_s$ and (2) the pulser delivers a slowly increasing voltage between V_s and V_c . The internal impedance of the pulser, which determines the distance that V_0 must be above V_c , also plays an important role in mode stability, but the optimum impedance depends on the characteristics of the magnetron and thus cannot be stated in general. Particular combinations of magnetron and pulser frequently present situations much more complex than those shown in Fig. 10-25. As the phenomenon of mode-changing is becoming better understood, improvements which alleviate the troubles are being incorporated into both magnetrons and pulsers. The radar designer should become fully familiar with the requirements that a magnetron places on pulse shape and pulser impedance. The reader who needs more detailed information about magnetrons than is given here is referred to Vol. 6 of this series, *Microwave Magnetrons*.

THE PULSER

BY J. V. LEBACQZ AND M. G. WHITE

Modern radar equipments are usually based upon the generation of short pulses of electromagnetic radiation. The cavity magnetron must generate electromagnetic oscillation of suitable frequency and power, and the function of the pulser or modulator is to deliver power to the magnetron in a suitable way. Cavity magnetrons have ordinarily been employed in the generation of microwaves, but a few radar equipments have been designed around triode oscillators, notably of the lighthouse type. Although the pulse techniques described here were largely worked out with the magnetron in mind, the information presented is intended to be generally applicable to any oscillator or power-consuming load. The discussion will be limited to methods peculiarly well adapted to pulse powers in the range of a few kilowatts up to several megawatts, and to pulse durations in the range from one-tenth to several microseconds. The primary aim of the balance of this chapter is to give the designer a feeling for the over-all problem, and to assist him in deciding among the compromises required to achieve a well-balanced design.

Numerous considerations enter in the design of a radar pulser; of primary concern are the nature of the load, the output pulse voltage and current, the pulse duration, and the repetition rate. The operational problem faced by the complete equipment will impose some specific requirements on the pulser, with regard to size, weight, supply voltage, etc.; such requirements may make it necessary to use an otherwise less desirable design.

It is common practice to begin the design of a pulser around a resistance load equal to the static resistance of the actual load to be used at its required operating point. Although final evaluation of the pulser performance requires consideration of the whole system and particularly of the oscillator used, much information about pulser behavior can be obtained by considering a pure resistance load. The following discussion of the types of pulsers in use at present and their characteristics is based on the assumption of a resistance load. The general characteristics of the magnetron load will be considered later, and their effect on pulser behavior and design will be studied in greater detail for the two main types of pulsers.

10-7. Pulser Circuits.—To obtain substantially rectangular pulses of short duration and high pulse power requires that energy stored in some circuit element be released quickly upon demand and be replenished from an external source during the interpulse interval. Either electrostatic or electromagnetic means of energy storage can be used.

In the latter case, energy stored in the magnetic field of an inductance through which current is flowing is released to the load by suddenly interrupting the current. This can be achieved by biasing to cutoff a high-vacuum tube in series with the inductance, the resulting inductive voltage rise being applied to the load. To restore the energy to the magnetic field, it is necessary either to pass large steady currents through the inductance between pulses, or to use a fairly complicated grid-modulating circuit whose function is to start the current flowing through the inductance a short time before the moment of interruption.

In either case, power losses in the switch and auxiliary equipment are much larger than those in pulsers of comparable output using electrostatic storage of energy. As a result, the use of pulsers employing electromagnetic energy storage has been limited to special applications (such as trigger circuits) where it is necessary to obtain a very high ratio between the pulse voltage and the supply voltage or where the load is mostly capacitive. The following discussions in this chapter will refer to pulsers whose energy is stored in an electrostatic field.

The basic circuit of most practical pulsers so far designed is given in Fig. 10-26. Assume for the present that the energy-storage element is a condenser, C_0 , charged to a potential V_c so that the energy stored is $\frac{1}{2} C_0 V_c^2$. When the switch S is closed at $t = 0$, the condenser will begin discharging exponentially through the load resistance R_L . If the switch S can now be opened suddenly at a time t_0 very small compared to the time constant $R_L C_0$ of the circuit, the voltage appearing across the load, for $0 < t < t_0$, will be given by

$$V_L = V_c \left(1 - \frac{t}{R_L C_0} \right)$$

and will be constant within a few per cent, from the assumption $t_0 \ll R_L C_0$. The voltage left on the condenser is

$$V'_c = V_c \left(1 - \frac{t_0}{R_L C_0} \right).$$

Hence, the energy dissipated in the circuit during the pulse is

$$W_0 = \frac{1}{2} C_0 V_c^2 \left[1 - \left(1 - \frac{t_0}{R_L C_0} \right)^2 \right] \approx \frac{V_c^2}{R_L} t_0.$$

In practice, the only available switches that can be used to open suddenly the large load currents are vacuum tubes—either triodes or tetrodes.

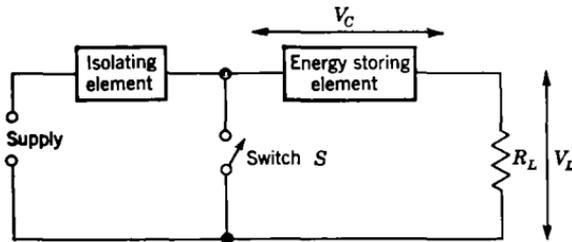


FIG. 10-26.—Basic circuit for pulsers using electrostatic energy storage.

Unfortunately, the plate resistance R_p of vacuum tubes is always rather large, resulting in a high voltage drop V_T across the switch during the pulse. The load voltage will then be given by

$$\begin{aligned} V_L &= (V_c - V_T) \left[1 - \frac{t}{(R_p + R_L)C_0} \right], \\ &= V_c \frac{R_L}{(R_p + R_L)} \left[1 - \frac{t}{(R_p + R_L)C_0} \right]. \end{aligned}$$

The energy dissipated during the pulse is again given approximately by

$$\frac{V_c^2}{R_p + R_L} t_0.$$

It must be supplied to the condenser during the interpulse interval through the isolating element. This element thus plays two very important functions in the pulser operation: it must prevent excessive power being drawn from the power supply when the switch is closed, and it must allow sufficient energy to flow during the interpulse interval to recharge the condenser C_0 . Either a high resistance, an inductance, or a series combination of resistance and inductance can be used for the purpose.

In order to avoid the large voltage drop across the high-vacuum tube, it would be desirable to use a very low-resistance switch, such as a spark

or gaseous discharge, in this circuit. When this is done, however, another difficulty presents itself immediately: the current flow through a gaseous discharge switch cannot be interrupted at will, so that all the energy stored must be dissipated in the load. If the energy-storage element is still a condenser, then its capacity C_0 could be made of such a magnitude

that $\frac{1}{2} C_0 V_0^2$ is equal to the required energy for the pulse, $(V_L^2/R_L) t_0$. This is unsatisfactory, since the wave shape of the output voltage pulse would be an exponential, instead of being substantially rectangular, as desired.

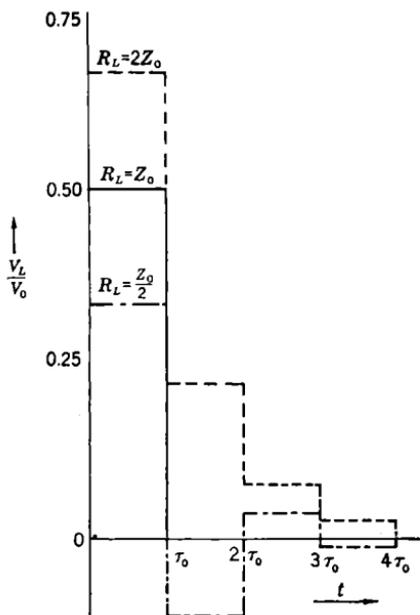


FIG. 10-27.—Output of an ideal transmission line with a resistance load.

current of magnitude $\frac{V_c}{Z_0 + R_L}$ flowing through the circuit for a time $t_0 = 2\delta$, and the voltage appearing across the load (neglecting the small voltage drop across the switch) will be given by $V_L = V_c \frac{R_L}{Z_0 + R_L}$. Successive reflections will occur, the amplitude of the n th reflection being given by

$$V_{L_n} = V_c \frac{R_L}{Z_0 + R_L} \left(\frac{R_L - Z_0}{R_L + Z_0} \right)^n.$$

The general wave shape obtained is shown in Fig. 10-27 for three values of R_L/Z_0 . Of course, if $R_L = Z_0$, no reflections occur, and a perfect rectangular pulse is obtained. Even if perfect “match” ($R_L = Z_0$) between the load and the network is not realized, little trouble will usually occur. Assuming, for example, $R_L/Z_0 = 1.15$, the voltage ampli-

A satisfactory wave shape can be obtained with a gaseous discharge switch if a transmission line, either parallel wires or coaxial cable, is used instead of the condenser. Consider an ideal (lossless) line of impedance $Z_0 = \sqrt{L/C}$, where L and C are the inductance and capacity per unit length, having a one-way transmission time δ . Elementary transmission-line theory shows that if such a line, charged to a potential V_c , is suddenly connected across a resistance R_L , there will result a discontinuous

tude of the first reflection would be only 7 per cent of that of the main pulse, and the power involved in that reflection is only 0.5 per cent of that of the main pulse. For all practical designs, a mismatch of 20 to 30 per cent is acceptable from the standpoint of energy loss. Design considerations for the best use of available components usually make it preferable to keep the impedance Z_0 small, generally about 50 ohms. The load resistances encountered in practice are approximately 500 ohms or higher. In such a case, the load and the line are matched through a pulse transformer. The voltage step-up ratio of the pulse transformer, n , is chosen to make $n^2 = R_L/Z_0$.

If $R_L = Z_0$, it can be shown easily that all the energy stored in the transmission line is dissipated in the load. Again, from elementary theory,

$$Z_0 = \sqrt{L/C} = \sqrt{L_0/C_0};$$

$$\delta = l\sqrt{LC} = \sqrt{L_0C_0},$$

where l is the length of the line, L and C its inductance and capacity per unit length, $L_0 = lL$ the total inductance of the line, and $C_0 = lC$ its total capacity. The energy stored is given by

$$W = \frac{1}{2} C_0 V_C^2,$$

but

$$C_0 = \frac{\delta}{Z_0} = \frac{t_0}{2Z_0}, \quad \text{since} \quad t_0 = 2\delta,$$

and

$$W = \frac{t_0 V_C^2}{4Z_0}.$$

On the other hand, the energy dissipated in the load is given by

$$W_0 = \frac{V_L^2}{R_L} t_0,$$

or, since

$$V_L = \frac{V_C}{2} \quad \text{and} \quad R_L = Z_0,$$

$$W_0 = \frac{V_C^2}{4Z_0} t_0 = W;$$

therefore all the energy stored in the line is dissipated in the load.

This method thus provides a very efficient way of obtaining a rectangular pulse of energy in a resistance load. It is usually impractical to use actual transmission lines or cables in actual pulsers; a cable to supply a 1- μ sec pulse would be approximately 500 ft long. This difficulty can be easily circumvented by the use of artificial transmission lines or of a pulse-forming network.

Now that a practical and highly efficient discharging circuit has been shown to be feasible, there remains the problem of replenishing the energy of the pulse-forming network. This could be done by using a resistance as isolating element. The very low efficiency of this scheme¹ makes it unsuitable for all but a very few special applications, such as systems requiring variable interpulse intervals. Accordingly, inductance charging is almost always used. It will be shown later that, if a d-c supply voltage is available, the network voltage at the time of discharge is double the supply voltage, except for losses in the inductance; it will also be shown that the network can be recharged from a source of a-c voltage, provided the repetition frequency is a multiple of one-half the supply frequency.

Comparison of the Two Types of Pulsers.—The basic circuit of Fig. 10-26 applies, as has been shown, to two types of pulse generators. In one type a small amount of the energy stored in a condenser is allowed to be dissipated in the load during each pulse. The switch, which must be able to interrupt the pulse current, is always a vacuum tube, and pulsers of this type are commonly called "hard-tube pulsers." Pulsers of the other type, where exactly the correct amount of energy is stored before the switch is closed, and the pulse is shaped by the discharge circuit itself, are referred to as "line-type pulsers," since the pulse-shaping elements or pulse-forming networks have been derived from the electrical characteristics of transmission lines.

The two types of pulsers have different characteristics, and it is of interest to analyze briefly some of the considerations involved in the design.

For instance, it is easier to change the pulse duration in a hard-tube than in a line-type pulser, since it is sufficient to change the time during which the switch is conducting. This can be done easily at low voltage in the driver stage, instead of using a high-voltage switch to change pulse-forming networks in a line-type pulser.

Methods for turning on the switch in a hard-tube pulser, discussed more fully later, generally involve a small regenerative pulser which applies a positive pulse "drive" to the control grid of the vacuum-tube switch. This small pulser nearly always requires auxiliary voltage supplies and, in addition, the switch-tube control grid must be maintained beyond cutoff during the interpulse interval. In line-type pulsers, the "triggering" of the switch (or initiation of the discharge) is usually accomplished with much less power than is necessary to drive the grid of the hard tube; in most cases, the driver power output is only a few per

¹Simple considerations show that, when charging a condenser from zero to the power supply voltage through a resistance, as much energy is dissipated in the resistance as is stored in the condenser.

cent of the pulser power output, but it runs as high as 10 per cent in some high-power pulsers. The output pulse shape is usually more nearly rectangular from a hard-tube than from a line-type pulser. Except for special cases, hard-tube pulsers are almost always built for direct output, and thus avoid the inductance and capacity added by the impedance-matching pulse transformer between the line-type pulser and the load, with the resulting oscillations in the pulse shape.

Power regulation as a function of input voltage, that is, the ratio of the fractional increase in output power to the fractional increase in input voltage which produces it, is always equal to 2 for a line-type pulser operating near matched conditions, regardless of the type of load. For a hard-tube pulser operating a load having the general characteristics of a biased diode—such as the magnetron—the regulation may be as high as 5 or 6 if the pulser is designed to operate with minimum tube drop. It is possible,

however, by choice of switch tubes with proper characteristics, to operate on such a portion of their characteristic that the power regulation near the operating point will be only about 0.5.

The pulser load line, or variation in output voltage and current as a function of load, is a plot of load voltage vs. load current (see Fig. 10-28). An equation for the line-type pulser is easily obtained, by eliminating R_L between the expressions for load current and voltage, as $V_L = V_C - I_L Z_0$. It is thus a straight line. For hard-tube pulsers, the same expression is applicable if Z_0 is replaced by the tube resistance R_p . In this case, however, R_p is not a constant, but depends on the tube characteristics.

When variable pulse spacing, very high pulse repetition rates (greater than 4000 pps), or "coded pulses" (groups of closely spaced pulses) are required, the hard-tube pulser is almost exclusively used. Although the line-type circuit can be adapted to such an application,¹ the advantages of flexibility inherent in the hard-tube pulser have restricted the use of the line-type pulser to very special cases. Such a case occurs if the supply voltage is so low that the required power output cannot be obtained from a hard tube.

"Time jitter," or the difference between the time at which a pulse is

¹ See Sec. 10-7, Vol. 5, Radiation Laboratory Series.

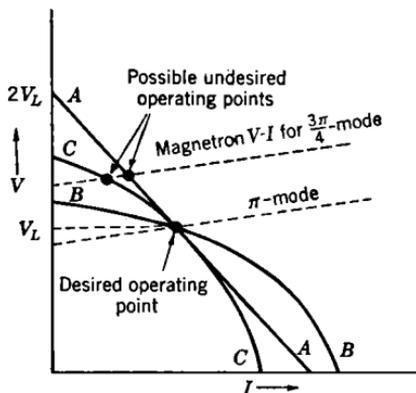


FIG. 10-28.—Pulser and magnetron voltage-current characteristics. (a) line-type pulser load line, (b) hard-tube pulser (adequate emission available), (c) hard-tube pulser near saturation.

supposed to appear and that at which it does appear, is sometimes a very important factor to be considered in the design of the pulser (Chap. 16). For hard-tube pulsers, the time jitter is practically always of the order of 0.01 μ sec or less. Some specially designed gaseous discharge switches (hydrogen thyratrons and mercury-sponge series gaps) can also give time jitters of the same order of magnitude, if the proper precautions are taken in the trigger circuit. Others have time jitters ranging from $\frac{1}{2}$ to 3 μ sec for cylindrical series gaps, to 20 to 50 μ sec for rotary gaps.

The simplest type of pulser is one of the line type, with a-c charging and a rotary-gap switch. In this case, the pulser components are reduced to a pulse-forming network, a charging transformer (which combines the functions of input step-up transformer and charging inductance), and a rotary gap mounted on the shaft of the alternator which supplies power to the charging transformer. In this way, the proper relationship between supply and repetition frequencies is easily maintained, and a very compact pulser for high power output can be made, although such a pulser has no flexibility.

In general, line-type pulser circuits are simpler and therefore easier to service. They also lend themselves to a greater efficiency, not only because the circuit in itself is more efficient, but because the overhead—auxiliary circuits, cathode power, etc.—is less. In size and weight, for a given set of output conditions, the advantage is decidedly with the line-type pulser. For airborne pulsers, these advantages in efficiency, size, and weight outweigh any advantage of flexibility or pulse shape which the hard-tube pulser may offer. As an example, the Model 3 airborne pulser (used in the AN/APS-15), rated at 144 kw pulse power output with pulse lengths of 0.5, 1, and 2 μ sec, weighs 55 lb and occupies a space 15 by 15 by 16 in.; with the techniques now available little improvement could be expected, were this pulser redesigned. On the other hand, a line-type pulser with hydrogen-thyratron switch, designed for 600 kw pulse output, for pulse widths of 0.5 and 2.5 μ sec, weighs only 98 lb and has a volume of about 17 by 17 by 24 in.

Table 10-3 summarizes the advantages and disadvantages of the two types of pulsers.

10-8. Load Requirements.—It is not enough to say that a pulser shall produce a pulse of a certain duration and magnitude. The nature of the load imposed upon the pulser, the operational problem faced by the complete equipment, and certain other practical factors usually require consideration.

Although the most usual pulser load is the magnetron, the load problem in general will be considered briefly. The eventual load will differ from a pure resistance by having a certain amount of capacity and inductance associated with it, and it will certainly be nonlinear. Further, the

TABLE 10-3.—COMPARISON OF THE TWO PULSER TYPES

Characteristics	Hard-tube pulser	Line-type pulser
Efficiency	Lower; more overhead power required for driver, cathode heating, and for dissipation in switch tube	High, particularly when pulse-power output is high
Pulse shape	Better rectangular pulses	Poorer rectangular pulse, particularly through pulse transformer
Impedance-matching	Wide range of mismatch permissible	Smaller range of mismatch permissible (± 20 -30 per cent). Pulse transformer will match any load, but power input to nonlinear load cannot be varied over a wide range
Interpulse interval	May be very short, as for coding beacons (i.e., $< 1 \mu\text{sec}$)	Must be several times the deionization time of discharge tube (i.e., $> 100 \mu\text{sec}$)
Voltage supply	High-voltage supply usually necessary	Low-voltage supply, particularly with inductance charging
Change of pulse duration	Easy; switching in low-voltage circuit	Requires high-voltage switching of network
Time jitter	Somewhat easier to obtain negligible time jitter, i.e., $< 0.02 \mu\text{sec}$, than with line-type pulser	High-power line-type pulsers with rotary-gap switch have inherently large time jitter. With care in design and use of hydrogen thyratron or mercury-sponge type of enclosed gaps, time jitter of $0.02 \mu\text{sec}$ or less obtainable
Circuit complexity	Greater, leading to greater difficulty in servicing	Less, permitting smaller size and weight
Effects of change in voltage	For design having maximum efficiency, $(\Delta P/P)$ output $\approx 6(\Delta V/V)$ input. By sacrificing efficiency in the design, $(\Delta P/P)$ output $\approx 0.5(\Delta V/V)$ input can be obtained	Better than a hard-tube pulser designed for maximum efficiency since $(\Delta P/P)$ output $\approx 2(\Delta V/V)$ input for line-type pulser, independent of design

load may also display occasional short-circuit and open-circuit conditions which must be allowed for in the pulser design.

Confining our attention to the magnetron, we note in Fig. 10-29 that it displays a dynamic impedance at the operating point of only 430 ohms, even though the V - I impedance ratio is 1480 ohms there. Both the operating and the dynamic or incremental impedance are important; the former determines the rate at which power will be absorbed from the

pulser, and the latter determines in some cases the variation in power output with small incremental changes in applied voltage. For example, a magnetron operating at 20 kv and 13.5 amp absorbs power at the rate of 270 kw, yet a decrease in applied voltage of only 3 kv will drop the current by 7 amp and virtually stop the oscillation. Since $\Delta V/\Delta I$ is small, a satisfactory pulser must produce a fairly flat-topped pulse free from voltage changes larger than 5 per cent of the peak voltage.

The V - I characteristics of most magnetrons show irregularities which are associated with various modes of oscillation whose number, magnitude, and position vary with magnetron type, r-f load applied to the magnetron, magnetic field, and cathode condition.

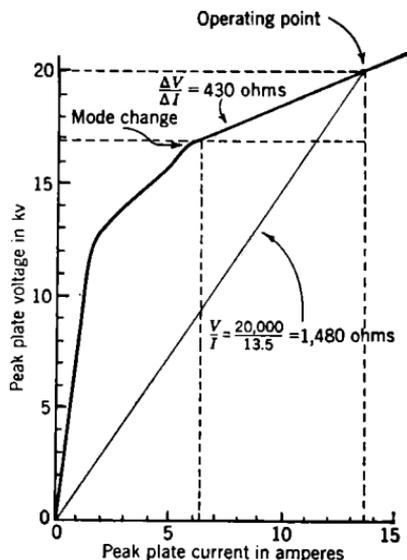


FIG. 10-29.— V - I characteristics of 4J77 magnetron. Field, 2700 gauss; $\lambda = 11.140$ cm.

Short-circuit and open-circuit conditions must be allowed for if the pulser is not to be damaged. A magnetron, or any other load, may develop an infinite impedance because of mechanical breakage of some part of the circuit, or through failure of the cathode to emit. This is not serious in the case of the hard-tube pulser, but the pulse-forming network variety generally discharges its unconsumed energy in some abnormal and undesirable way. No pulsing system should be considered which can be destroyed by a single discharge of a high-power pulse in some unexpected part of the circuit. An average-current device intended to turn off the primary

power supply if the load fails to draw power could hardly do so in less than a few pulse cycles, and thus cannot be relied upon to protect against this.

A short-circuited load can prove disastrous to a hard-tube pulser in a few pulse cycles unless provision is made to turn off the primary power when the load draws too much current, since most of the energy which the hard-tube pulser normally delivers to the load will then appear as anode heating of the switch tube. Since it is hardly practicable to design switching tubes heavy enough to withstand this sort of abuse for long, overload protection must be provided. A relatively fast-acting average-current device is suitable for this purpose.

The electrostatic capacity of the load should nearly always be kept to a minimum. Not only is the energy spent in charging this capacity

totally wasted, but stray load capacity seriously affects time of fall and back-swing. Short, well-spaced leads and a low-capacity magnetron filament transformer are helpful in reducing load capacity.

The definition of pulse duration is intimately bound up with magnetron behavior. From the radar system point of view it is the duration of the r-f pulse which is of interest and not necessarily the duration of the voltage pulse applied to the magnetron. Figure 10-30 shows the typical appearance of the r-f pulse, of the current pulse through the magnetron, and of the voltage pulse applied to the magnetron. Usually, the shape

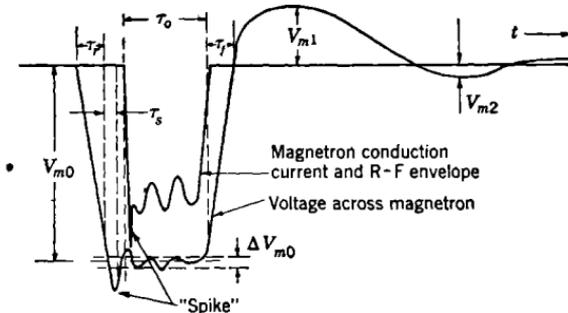


FIG. 10-30.—Time variation of voltage and current of a pulsed cavity magnetron.

- τ_r = time of rise of voltage wave
- τ_s = delay time for onset of oscillations
- τ_0 = width of current and r-f pulse
- τ_f = time of fall of voltage wave.

and duration of the r-f pulse closely approximate those of the current pulse; the voltage pulse is less steep and of longer duration.

The pulse length, τ_0 , may be defined as the interval of time during which the load current wave is within 50 per cent of peak value. In so defining the pulse length, it is necessary to ignore a rather short-duration, high-current spike often found on the leading edge of the current wave. So defined, τ_0 can be used satisfactorily in calculations of average current and average r-f power. Successful pulsers have been designed and put into production for τ_0 as short as 0.1 μ sec, and as long as 5 μ sec or a little more. Practical design difficulties increase at pulse lengths outside this range.

The voltage across the magnetron when it is oscillating vigorously in a steady state is designated as V_{m0} . In general, V_{m0} will lie between a few kilovolts and 30 to 40 kv, although it may be higher for magnetrons in the 10-Mw region. Most practical pulsers do not produce a perfectly uniform flat-top wave because this is usually costly and unrewarding. A drop of 2 to 5 per cent in voltage during the pulse, ΔV_{m0} , can usually be tolerated by the magnetron without harmful frequency modulation or mode instability.

The rate of rise and the overshoot on the leading edge of the voltage wave are important in determining the ability of the magnetron to operate at high power levels. A multicavity oscillator has several possible modes of oscillation, in only one of which it operates efficiently and smoothly. Great effort has been devoted to understanding the conditions favorable to stable oscillations in a given mode. We now know that stability depends on both pulser and magnetron. Usually the pulser can be more easily modified than the magnetron if mode-shifting or mode-jumping occurs. In any event, many interacting adjustments must be made before full power output can be assured. Sometimes a change in the r-f loading of the magnetron, in its heater current, or in the shape of the magnetic field will succeed where altering the pulse shape has failed. Almost always, too short a rise time, t_r , of the voltage pulse leads to mode instability, any rate of rise over 100 kv/ μ sec being considered fast.¹ Magnetron sparking can also be caused by high rates of voltage rise.

A "spike" frequently appears on the leading edge of the voltage pulse. Since a magnetron operates at a magnetic field far above the nominal cutoff value, it is impossible for electrons to reach the anode without the aid of the r-f field. Therefore, until the field is established, little power is drawn from the pulser and a momentary overvolting of the magnetron results. This phenomenon is particularly prominent in the case of line-type pulsers, where the voltage across the magnetron can rise to twice normal if the magnetron fails to draw power. This high-voltage "spike" can give rise to all manner of sparking and mode troubles. Its cure lies not alone in alteration of the pulser design, but also in attention to magnetron design.²

The internal impedance of the pulser plays an important role in affecting mode stability. There are significant differences between the operation of hard-tube and of line-type pulsers.³ The hard-tube pulser has an internal impedance equal to that of the output tube, which varies from 90 to 150 ohms until saturation is reached. The line-type pulser, being essentially a constant-voltage device in series with an impedance equal to the load resistance, has a resistance of 400 to 1200 ohms. Consequently, the conditions of stable operation for a given magnetron will vary with load and pulser type. Stable operation occurs for those values of V and I which simultaneously satisfy the load and generator characteristics. This would not concern us except that the various modes have different V - I characteristics, thus giving cause for instability.

¹ The 4J52 magnetron is an exception to the statement made here, being more stable with a rapid rate of rise.

² More detailed discussions will be found in *Pulse Generators*, Vol. 5 of this series, and in *Microwave Magnetrons*, Vol. 6.

³ RL Report No. 809, Sept. 28, 1945.

Figure 10-28 shows how differing internal impedances can affect mode stability: operation in the unwanted $\frac{3}{4}\pi$ mode is impossible in the case of a pulser load line such as *B*, but may occur if the load line has the form of *A* or *C*.

We must also pay attention to the reverse voltage V_{m1} which most pulsers put across the magnetron after the main-power pulse has passed, for in some cases V_{m1} may be larger than V_{m0} and lead to breakdown of the magnetron insulation. This same reverse voltage may also appear across various parts of the pulser circuit itself unless steps are taken to suppress it. Most pulsers keep the reverse voltage within safe limits by means of a resistance or, better still, a nonlinear element such as a diode. The power wasted in backswing damping is rarely more than a few per cent of the total useful power output of the pulser, and the duration of the backswing is seldom greater than 2 to 10 μsec . Some pulsing circuits produce a secondary forward pulse, V_{m2} , which can cause trouble by feebly exciting the magnetron to give a few microwatts of power. In many applications this would cause no concern because of the much greater amplitude of the main power pulse. In most radar systems, however, a few microwatts, directly from the magnetron, would swamp any but the strongest echoes which might happen to return coincidentally with the appearance of V_{m2} . As a matter of practical experience, V_{m2} should be kept below 10 per cent of V_{m0} if all possibility of trouble is to be avoided.

10-9. The Hard-tube Pulser.—A simplified diagram of the power-output stage of the hard-tube pulser is given in Fig. 10-31. The energy-storage element C_0 is recharged through the isolation inductance L_1 . In the quiescent state, the point A' is at ground potential while the high-voltage side of C_0 is at the full d-c supply voltage along with the anode of vacuum tube T_1 . The control grid of T_1 is biased to cutoff and the screen grid is held at a normal positive value. If a sufficiently large positive pulse be applied to the control grid of T_1 , the cathode will emit electrons vigorously and the anode will drop from its high positive potential to just that potential required to pass the plate current demanded by the load. Since the voltage across C_0 cannot change instantaneously, the point A' will assume a high negative voltage, which is applied to the load. Current will continue to flow out of C_0 , and around the load circuit, until the driving pulse on the grid of T_1 is removed. The system will return to its

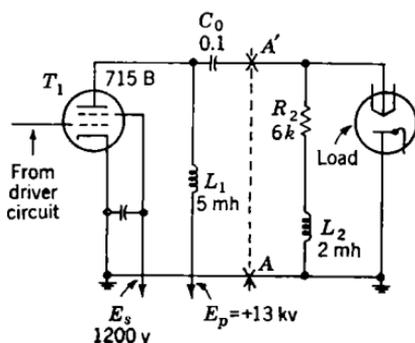


FIG. 10-31.—Schematic diagram of output stage of hard-tube pulser.

previously described quiescent state as soon as current can flow through L_2R_2 to bring the point A' back to ground, and as soon as recharging current can flow slowly through L_1 to return C_0 to its fully charged condition. During the pulse, the potential of C_0 (and hence the voltage across the load) slowly declines unless C_0 is made very large. In any practical design it is usually desirable to make C_0 no bigger than absolutely necessary, the criterion being the permissible voltage drop while carrying a current pulse of definite magnitude for the duration of one pulse. Expressed more quantitatively, $C_0 = It_0/\Delta V_{m0}$. If $I = 10$ amp, $t_0 = 1$ μ sec, and the permissible voltage drop $\Delta V_{m0} = 100$ volts, we find that C_0 must be 0.1 μ f. The sizes of L_1 , L_2 , and R_2 depend upon how quickly it is desired to bring the tail of the pulse down to zero once the grid drive has been removed, and also on how much energy one is willing to waste in these elements during the pulse.

There exists an upper limit to the size of the coupling condenser, because a very large capacity would take longer to recharge fully after a pulse. Given a fixed pulse rate and recharging impedance, an increase in coupling capacity will be accompanied by an increase in the no-load to full-load voltage ratio. In effect, a large capacity gives poor regulation. Should the designer attempt to recover this loss by decreasing the size of the recharging impedance, there will result an increased current drain from the power supply through the switch tube during the pulse, thereby causing an increased switch-tube drop. Careful balancing of these factors is required to achieve most efficient design.

Occasionally, the most efficient operating voltage of the magnetron or load does not correspond to the optimum plate voltage for T_1 , so that there must be some impedance-matching device between pulser and load. The primary of a pulse transformer may be inserted at AA' and the load connected across the secondary with or without the damping element L_2R_2 . Generally speaking, the insertion of a pulse transformer changes so much the character of the pulse tail that it is necessary to alter radically the so-called "tail damping" circuit. The most flexible and satisfactory damping system is a high-voltage diode connected across the magnetron to draw power only on the backswing.

The Switch Tube.—Since the vacuum-tube switch is the heart of the hard-tube pulser, the characteristics desirable in a switch tube will be briefly mentioned. It must be able to carry easily the current demanded by the load; in the interest of efficiency, the anode potential required should be as low as possible. Since a major source of power waste in the circuit is in the cathode-heating required to sustain the desired emission, it is customary to use oxide-cathode emitters wherever possible. Unfortunately, oxide-cathode tubes are not easily adaptable to high anode voltages because of the phenomenon of cathode sparking. With proper

care in processing and use it is possible, nevertheless, to arrive at a satisfactory and highly flexible pulser design using the oxide-cathode pulse tube up to 20 kv. Where extreme reliability and relative immunity to overload are required, the designer will prefer the thoriated-tungsten cathode switch tube in spite of its greater cathode power requirements. Switching tubes of the oxide-cathode variety usually can be counted upon to emit 300 to 600 ma/watt of heating power while the emission of the best thoriated-tungsten cathodes is only about 100 ma/watt.

The oxide-cathode tube is limited in power output by the onset of grid emission. It is impossible to prevent thermionically active cathode material from contaminating the grid structure. Great care must be taken, therefore, to cool the grid by fins, heavy rods, and other means, for it can emit a few milliamperes even at a sub-visible temperature. A few milliamperes may sound small in comparison with several amperes pulse current, but a few milliamperes flowing all the time can cause serious heating of the anode and screen, which are at relatively high potentials. Usually the result is a runaway condition in which grid control is lost. The most successful oxide-cathode high-power switch tube is the 715B developed at Bell Telephone Laboratories. The 715B has a gold-clad grid which poisons and "absorbs" active material, thereby lowering its efficiency as a thermionic emitter.

TABLE 10-4.—TYPICAL OPERATING CHARACTERISTICS

Tube type	Cathode type	Cathode power, watts	Max. plate voltage, kv	Pulse current, amp	Emission, amp/watt	Cutoff grid voltage*	Plate drop, volts	Screen voltage	Positive grid drive, volts
3D21	Oxide	10	3.5	5	0.5	- 70	400	800	150
3F29	Oxide	14	5.0	8	0.57	- 100	600	800	150
715B	Oxide	56	15.0	15	0.27	- 500	1500	1200	200
5D21	Oxide	56	20.0	15	0.27	- 500	1500	1200	200
304TH	Th-W	125	15.0	6	0.05	- 900	2000	200
6C21	Th-W	140	30.0	15	0.11	-1000	1500	1500
6D21	Th-W	150	37.5	15	0.10	- 500	2000	2000	1500
527	Th-W	770	30.0	60	0.08	-1200	1500	1500

* 0.2-ma cutoff.

A good output pulse tube must possess a sharp grid cutoff. For efficient operation, the plate current must be one milliampere or less when the grid is biased to some reasonable value. Because the anode is at very high voltage between pulses, any "leakage current" delivers excessive power to the anode. As an example, consider a pulse tube normally giving 20-amp pulses with an anode drop of 3 kv at a duty ratio of $\frac{1}{1000}$. Such a tube will display a peak power anode loss of 60 kw, but an average power loss of only 60 watts. If the leakage current at "cutoff" amounts

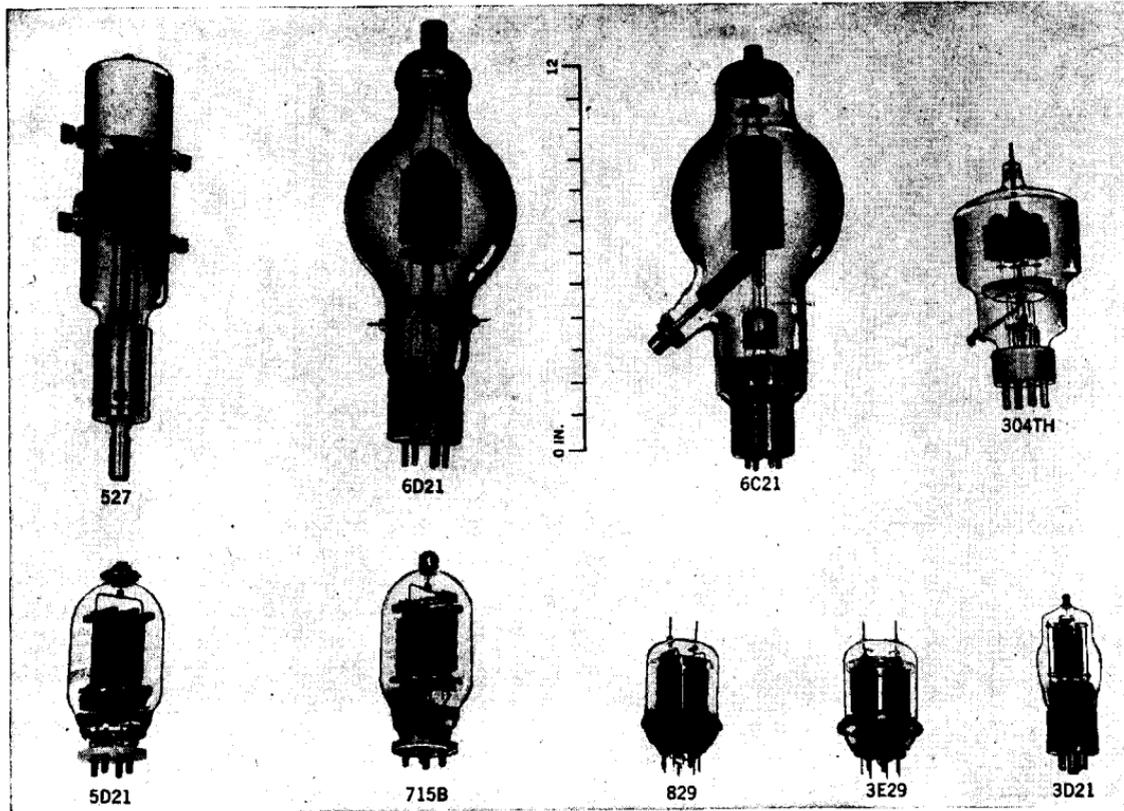


FIG. 10-32.—Group of typical high-vacuum switch tubes.

to only 1/10,000 of the peak current, or 2 ma, the power loss at 35 kv anode voltage is 70 watts. To reduce leakage current, the grid structure of a switch tube must surround the emitting area of the cathode completely, and must give a uniform field over the entire cathode. Figure 10-32 shows a group of high-vacuum switch tubes; Table 10-4 gives a set of typical operating characteristics.

Except for the screen-grid and control-grid bias requirements of power output and driver stages, the hard-tube pulser is rather simple, for (as is shown in Fig. 10-33) the grid-driving circuit is not inherently complicated. A single 829 or 3E29 tube at T_2 will satisfactorily drive two Eimac 304TH's or two Western Electric 715B tubes; either complement of tubes is adequate to deliver over 200 kw to the load. Circuit constants given in Figs. 10-31 and 10-33 are representative and have been used in a pulser produced in large quantities. With the constants shown, the output voltage pulse has a nominal duration of 1 μ sec at 12 kv and 12 amp. The maximum duty-ratio limit, set by the 715B tube, is 1/1000.

If higher powers are required, it is quite practical to use several 715B's in parallel. High-voltage tubes such as the 6C21 can be operated either singly or in parallel. In the case of parallel operation, it is sometimes necessary to insert 25 to 50 ohms into the plate and screen leads in order to prevent parasitic oscillations.

The Driver Circuit.—In the grid-driving circuit of a hard-tube pulser, all the requirements for pulse shape, accuracy of pulse spacing, and flexibility of pulse rates must be met. The important parts of the regenerative pulser circuit shown in Fig. 10-33 are the pulse-forming network at the left and the three-winding pulse transformer which provides the necessary plate-grid feedback to make the circuit self-driving.

Normally, all tubes are biased to cutoff; therefore, when a positive pulse of 100 volts is applied to the "trigger in" point (Fig. 10-33), plate current starts to flow in T_2 , inducing a further positive voltage on the grid of T_2 . Once regenerative action has commenced, the input trigger has no further effect; the plate current increases rapidly until limited by the e_p - i_p characteristic of T_2 . As i_p increases, e_p decreases until it nears the screen potential; from there on, i_p increases only slowly, thus producing a relatively flat-topped wave.

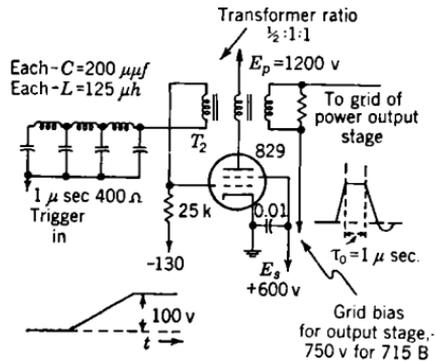


FIG. 10-33.—Schematic diagram of regenerative pulse generator used as driver of hard-tube pulser.

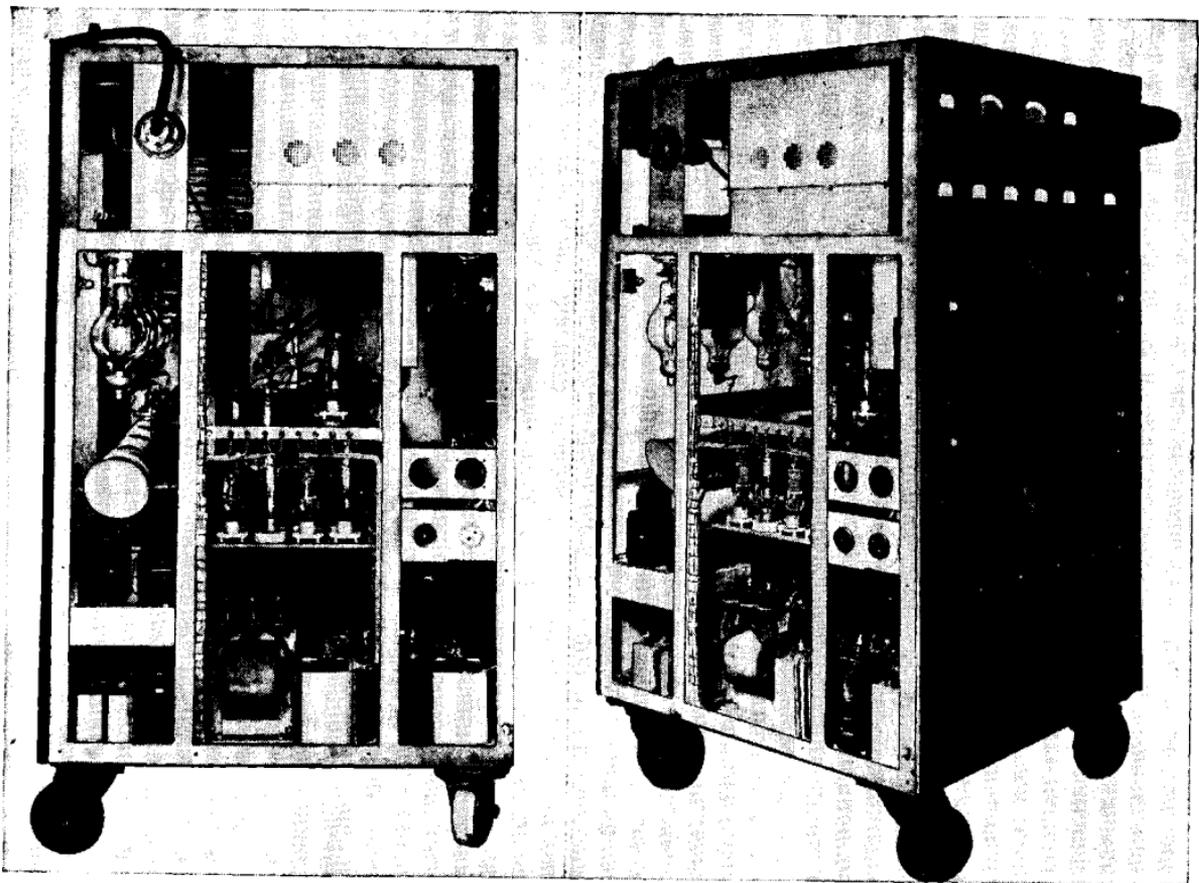


FIG. 10-34.—High-power hard-tube pulser.

During the rising part of the plate current curve, the pulse transformer induces a *negative* voltage wave on the delay line equal to the induced positive voltage on the grid. The pulse length is then fixed by the time required for this negative wave to travel down the line and back to the grid. Upon arrival at the grid, the negative wave drives the tube toward cutoff, a process helped again by the regenerative action of the pulse transformer.

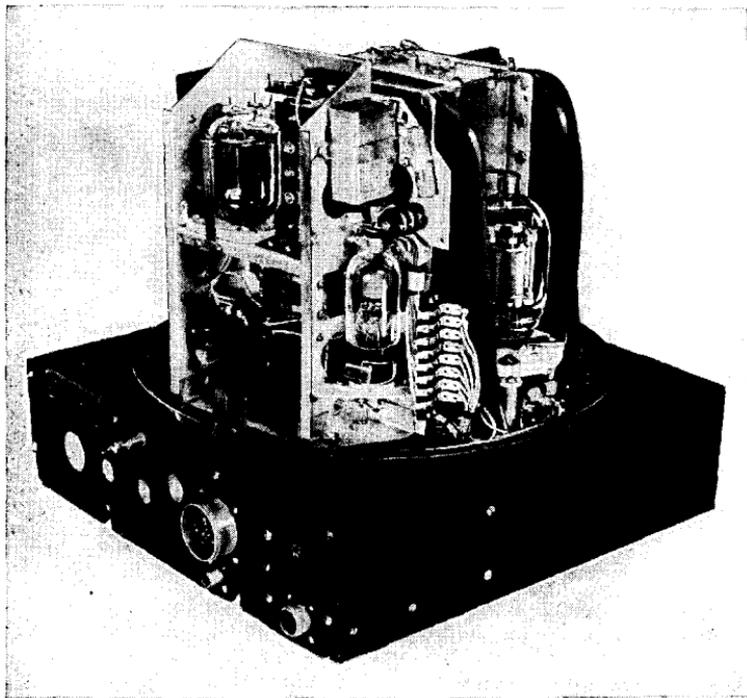


FIG. 10-35.—Hard-tube airborne pulser.

Practical Considerations.—The most economical pulser would have power flowing only during the pulse. Of course, energy must flow back to recharge those elements which store and deliver pulse power, but the design should not include vacuum tubes that conduct during the time between pulses, and are turned off during the pulse. Judicious use of plate-coupled and cathode-coupled circuits usually will permit all tubes to be biased to cutoff except during the pulse. Pulse transformers can also be used for phase reversal, impedance-matching, and circuit isolation. Means are now available for computing the important design parameters of such pulse transformers.¹

Successful hard-tube pulsers have been made with power outputs up

¹ *Pulse Generators*, Vol. 5, Chaps. 12 to 15, Radiation Laboratory Series.

to 3 or 4 Mw by operating half a dozen high-voltage 6D21 thoriated-cathode tetrodes in parallel. Figure 10-34 shows two views of a high-power hard-tube pulser rated at 3 Mw and Fig. 10-35 shows a lightweight hard-tube pulser built for airborne use. The latter is rated at 144 kw output at pulse durations of 0.5, 1, and 2 μ sec; it employs a circuit very similar to one given by combining Figs. 10-31 and 10-33.

10-10. Line-type Pulsers.—A schematic diagram applicable to most typical line type pulsers is shown in Fig. 10-36. Like the hard-tube pulser, the line-type pulser is best analyzed by separate consideration of

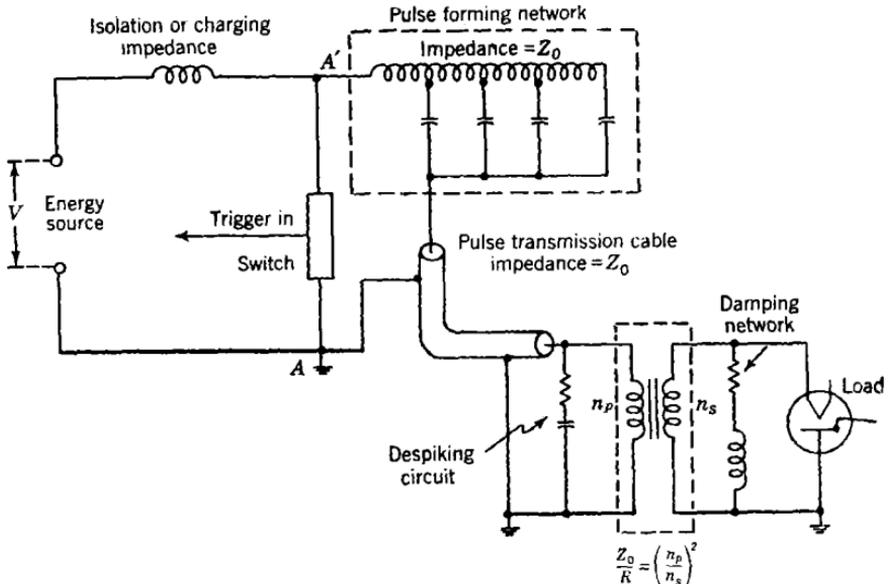


FIG. 10-36.—Basic circuit for a line-type pulser. Voltage may be either d-c or some periodic function of the time, such as $V = V_s \sin 2\pi ft$.

the discharging and the recharging circuit. The discharging circuit is very simple, and its individual components are discussed in some detail hereafter. The pulse transformer will be considered in the following section, since it is applicable to both types of pulsers. In general, it can be said that the effect of both the switch and the pulse transformer is to decrease the rate of rise of voltage at the load from that which would be produced by the network alone. The switch usually introduces an appreciable resistance during its ionizing time, and the pulse transformer introduces an additional inductance in series with the load. However, a spike voltage is often encountered in practice, as mentioned previously, because of the time required for the magnetron to draw current after the normal voltage is applied to it. When it is necessary to prevent the appearance of a spike, a "despiking" circuit is used as indicated in the diagram. The resistance of this circuit is chosen equal to the network

impedance, and the capacity is chosen small enough to be almost completely charged a very short time after the oscillator draws full load current. A damping network may also be provided to help bring down the trailing edge of the voltage pulse and prevent post-pulse oscillations.

The Pulse-forming Network.—One of the possible lumped-constant networks having electrical properties essentially equal to those of a transmission line is given in Fig. 10-37, with the values of L and C necessary for a line of given impedance and pulse length. If the inductance is constructed by winding a uniform helix on an insulating cylinder, the ratio of coil length to coil diameter should be that shown in Fig. 10-37, in order to provide the right amount of coupling between coils. The appropriate value of inductance at each end of the line differs slightly from that in the middle sections. It has been found practically that the optimum number of meshes is related to the desired pulse length as shown in the following table.

TABLE 10-5.—RELATION OF PULSE LENGTH TO NUMBER OF MESHES

Pulse length, μsec	Number of meshes
0.1–0.5	1–3
0.5–2.5	2–5
2.5–5	3–8

The greatest obstacle to design of a compact high-power network is the bulk of the necessary condensers. High-voltage condensers are inordinately large and expensive in comparison with the job they have to perform; all electronic equipment would profit from a major improvement in condenser design. Though a few new dielectrics have shown promise of increasing the energy stored per unit volume, mica, paper, and oil are still the primary dielectrics for high-voltage condensers. One new material of interest is Alsifilm or Diaplex, an aluminum-silicate clay in an organic impregnate. This material can be formed into thin, homogeneous sheets which have dielectric properties superior to mica or paper and oil. The dielectric constant is 5 to 6 and the safe operating dielectric strength is 500 to 700 volts per thousandth of an inch. Greater uniformity would raise the latter to 800 volts per thousandth. Most problems of fabrication have

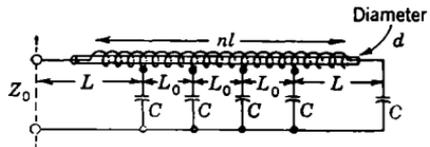


FIG. 10-37.—Type E pulse-forming network.

- $Z_0 = \sqrt{\frac{L_0}{C}}$
- $\tau_0 = 2n \sqrt{L_0 C}$
- $\frac{l}{d} = \frac{4}{3}$
- $\frac{L}{L_0} = 1.1 \text{ to } 1.2$
- $L_0 =$ inductance per mesh
- $L =$ end inductance
- $C =$ capacity per mesh
- $n =$ number of meshes
- $Z_0 =$ surge impedance
- $\tau_0 =$ pulse length
- $l =$ coil length of one mesh
- $d =$ coil diameter.

been overcome. Perhaps TiO_2 , or some similar material of high dielectric constant, will find wide use outside the low-voltage field.

Another promising approach to the condenser design problem is the electrochemical formation of very thin insulating films on metal (such as those employed in the electrolytic condenser) on which a layer of metal is then deposited. High capacity per unit area results from the very small spacing, but the insulating film must be able to withstand the extremely high field strengths thus imposed.

Further attention to the details of mechanical design of condenser fittings—containers, bushings, connectors, and the like—is also required. The over-all bulk even of condensers using conventional dielectrics can still be considerably reduced.

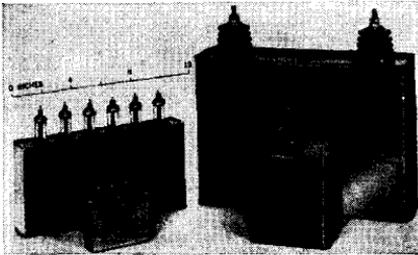


FIG. 10-38a.—Pulse-forming networks (see Table 10.6 for properties).

pulse power output; the other one is electrically equivalent to the network (1) of Fig. 10-38a, but uses Diaplex insulation.

The network designer has control of pulse length and impedance level independently of one another. The pulse length is ordinarily fixed by the nature of the application, while the impedance level is chosen to fit the characteristics of the load, the switching tube, and the power supply. A pulse transformer can be inserted between load and pulse-forming

network so that the network can be designed to use the available switching device most efficiently. Once the pulse power output is settled and the appropriate switching tube chosen, Z_0 can be determined from one of the relations $P = I_0^2 \cdot Z_0$ or $P = V_0^2/Z_0$. Usually either I_0 or V_0 is definitely limited for a given switch; this indicates the appropriate relation to use in calculating Z_0 . It must be remembered that the voltage across the pulse network (and hence across the switch) is twice the voltage delivered to a matched load.

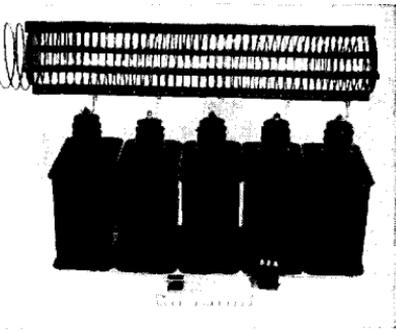


FIG. 10-38b.—Experimental pulse-forming network.

In large radar equipments, it is often desirable to separate pulser and r-f units by considerable distances. The pulse transmission cable necessary in such cases has been standardized at an impedance level of 50 ohms. Many pulse networks designed for high-power radar have an impedance of 50 ohms in order to avoid the use of a matching pulse transformer between the network and the cable. As a result, many of the more desirable switching devices have been designed to give maximum power output when used with a 50-ohm network.

TABLE 10-6.—PULSE-FORMING NETWORK CHARACTERISTICS

Network (Fig. 10-38)	Pulse length, μsec	PRF, cps	Power
No. 1	0.8	840	25 kw
	2.2	420	25 kw
No. 2	0.25	1600	250 kw
	0.5	800	250 kw
	2.6	400	250 kw
	5.2	200	250 kw
No. 3	1	1000	250 kw
No. 4	1	800	3 Mw

An important design consideration in pulse network applications is the average power to be handled. A pulse network designed to have adequate life at one pulse rate would overheat and perhaps be ruined by operation at a higher repetition frequency. Since overheating is a function of both applied voltage and repetition rate, little flexibility remains in a line-type pulser designed to achieve maximum economy of weight, space, and power. Provision can be made for shortening pulse length by bringing out taps on the pulse line. However, this is rather difficult at high power, because of the problem of designing suitable line switches.

The Switch.—The possible advantages of the line-type pulser greatly stimulated the design of low-impedance spark switches and thyratrons, since it became possible to secure flat-topped pulses without the necessity of opening the switch. Rotary spark gaps, "trigatrons," series gaps, hydrogen thyratrons, and mercury thyratrons have all been used as network switching devices. Each has its field of application. Each switch is limited in one or more of the following respects: (1) poor precision of firing, (2) low maximum pulse rate, (3) short life on long pulses, (4) narrow operating range of voltage, (5) occasional erratic firing, (6) inefficient cathode, and (7) unnecessary complication. The ideal switch has not yet been designed.

Figure 10-39 shows schematically the simple "series gap" switch. The Western Electric 1B22 is a good example of this class of switch; it consists of a cathode cylinder of aluminum surrounding an anode rod of

the same material. The tube is filled to a pressure of a little less than one atmosphere with a mixture of 75 per cent hydrogen and 25 per cent argon. Two or three tubes are used in series, with a voltage divider across them to ensure that the network voltage is divided equally among all the gaps before the pulse takes place. Triggering is accomplished as shown in Fig. 10-39 by depressing or raising quickly the voltage of the two middle electrodes. This breaks down one of the gaps and throws full voltage across the second gap. The overvolted second gap thereupon also breaks down, forming a low-impedance path between points A and A' . These switches are efficient and are quite satisfactory in use. Their chief limitations are in the allowable pulse repetition frequency (about 2000 pps maximum) and in obtainable precision of pulse timing. This type

of switch can handle several megawatts, and should be useful in applications that do not require great flexibility or accuracy of firing.

A particularly successful version of the series-gap switch is the Bell Telephone Laboratories' 1B42, which substitutes a mercury-saturated iron "sponge" for the solid aluminum cathode of the 1B22. By

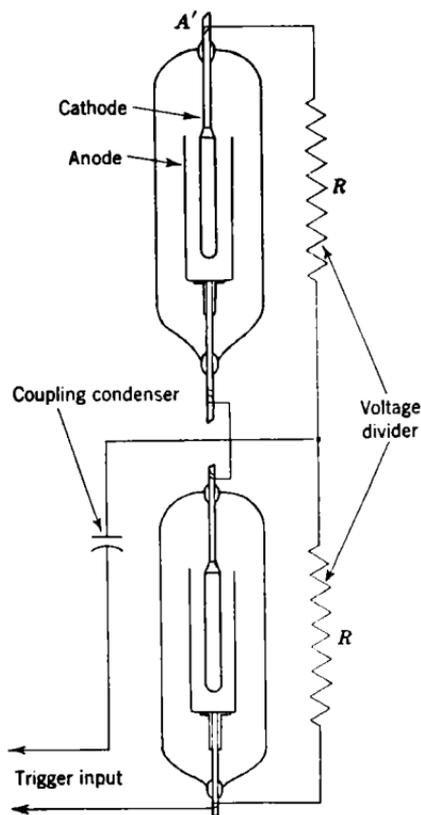


FIG. 10-39.—Series gaps.

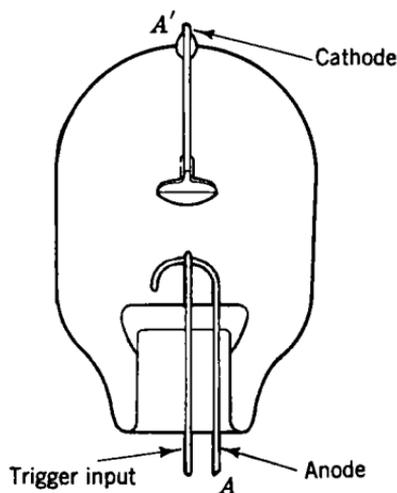


FIG. 10-40.—Trigatron.

using an atmosphere of relatively high-pressure hydrogen, the temperature variation of mercury vapor pressure is made unimportant. Since the ordinary series gap is limited in operation by pitting and spike growth, the mercury cathode provides better operation over a wide range of

operating conditions as well as longer life at higher frequencies. Apparently the mercury cathode series gap can be designed to fire very precisely.

The "trigatron" (Fig. 10-40) differs from the series gap in having but one spark gap, across which the full line voltage is applied. Conduction is initiated by applying a steep high-voltage wave to a trigger electrode; this presumably draws corona current sufficient to initiate the main discharge. A typical trigatron is the British CV85. This has three electrodes in a mixture of argon and oxygen at a pressure of approximately 3 atmospheres. A triggering voltage of approximately 4 kv is required to fire the tube when it is to deliver 125 kw of pulse power into a 70-ohm load. The precision of triggering usually is about $\pm 0.2 \mu\text{sec}$, unless special precautions are taken to insure correct trigger wave shape and amplitude.

The most versatile switch is the newly developed hydrogen thyatron (Fig. 10-41), which requires a positive trigger of only 150 volts rising at the rate of 100 volts per μsec . In contrast to spark devices, the hydrogen thyatron will operate over a very wide range of anode voltages without readjustment. This characteristic is particularly important for experimental pulsers, or for any pulser whose probable load is in doubt. In general, it is difficult to reduce the voltage across a spark switch by more than a factor of 2 from the nominal design voltage without encountering erratic operation. Voltages substantially higher than the nominal design value cause a flash-over in the switch. The hydrogen thyatron, on the other hand, is a true thyatron. Its grid has complete control of the initiation of cathode emission over a wide range of anode voltage. The anode of a hydrogen thyatron is completely shielded from the cathode by the grid. The effective grid action results in very smooth firing over a wide range of anode voltages and repetition frequencies. The hydrogen thyatron, unlike most thyatrons, has a positive grid-control characteristic, and hydrogen filling is used to reduce deionization time and make the performance of the tube independent of ambient temperature. This independence is maintained with the exception that the average dissipation rating of the tube can be lowered by excessively high air temperature.

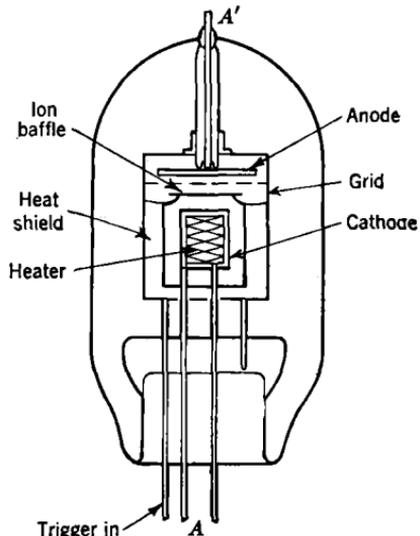


FIG. 10-41.—Hydrogen thyatron.

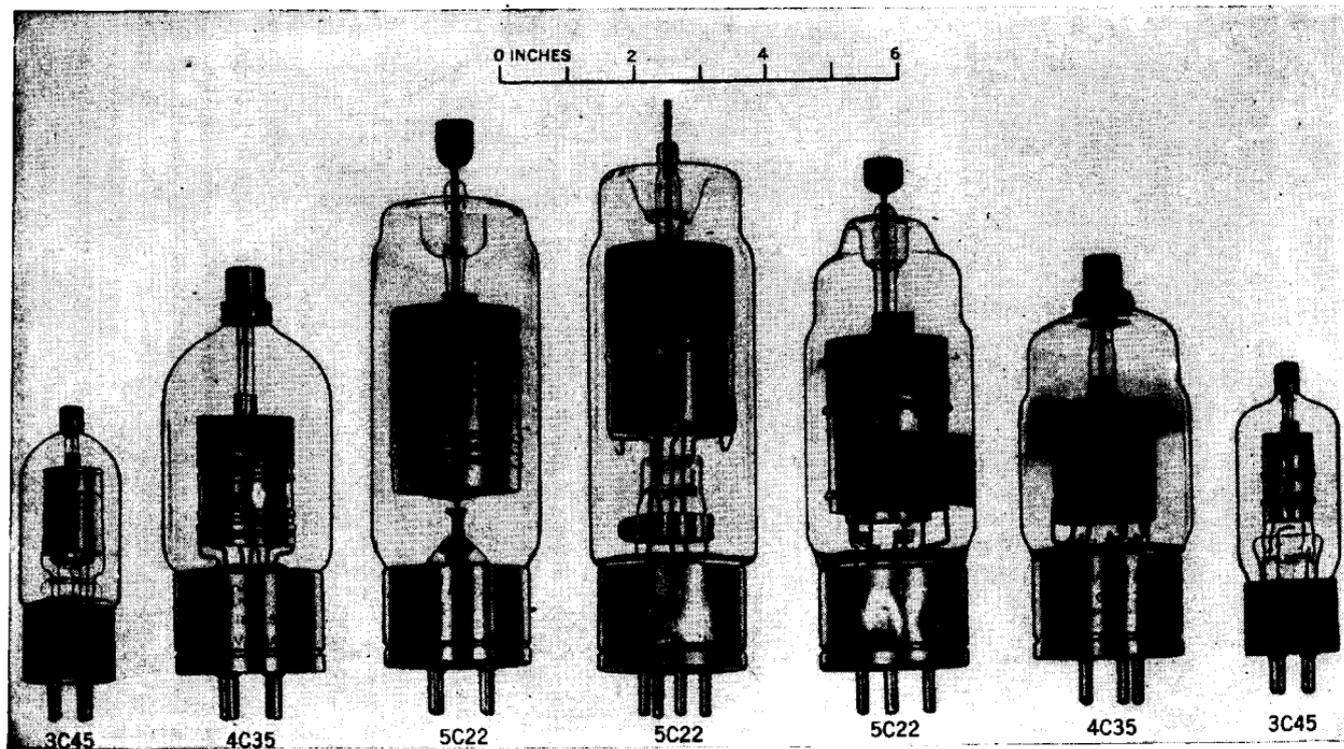


FIG. 10-42.—Production types of hydrogen thyratrons.

The hydrogen-filled thyratron (Fig. 10-41) can satisfactorily pulse at much higher frequencies than is possible with mercury or other heavy-gas fillings, because of the high ion mobility inherent in the light-gas filling. Of course, the average-power rating of the tube must not be exceeded when going to high pulse rates. In going to high pulse rates, the pulse width must be decreased faster than $1/f$, since the average anode drop at durations of a few tenths of a microsecond is higher than that during a long pulse. The hydrogen in the discharge seems to require about 10^{-7} sec to become fully ionized, so that a high tube drop is required to deliver the necessary initial current. It is therefore desirable to reduce the pulse power output as well as pulse width if tube life is to be maintained at increased repetition rates. Pulse rates up to 40,000 cps have been obtained at reduced power output.

Hydrogen thyratrons in production are the 3C45, 4C35, and 5C22, rated at 25, 250, and 1000 kw respectively; they are shown in Fig. 10-42. The development of such tubes for still higher power levels is under way.

A nonlinear inductance can also be used as a switch. An inductance with a special alloy core may be placed between points $A'A$ (Fig. 10-36).

This device has the property of possessing a high inductance when the current through it is small, and a very low inductance when the current is large. Since this type of switch has generally been used with a modulated energy source, there is initially no voltage across $A'A$ and hence the impedance is high from A' to A . As the voltage is built up the current gradually increases until suddenly, and in a regenerative way, the impedance of $A'A$ drops, thereby allowing the line capacity to discharge through the "switch," lowering its impedance still more.

A d-c bias winding is sometimes used to control the point at which the impedance suddenly drops from its high value. Virtues claimed for the nonlinear inductance switch are long life, ruggedness, and simplicity. The switch can be operated at pulse repetition frequencies up to 4000 pps and at power levels in excess of several hundred kilowatts.

The rotary spark gap, because of its great simplicity and high power handling ability, has been widely used as a switching device. Figure 10-43 shows schematically how a rotating insulating disk pierced by tungsten pins serves the function of a switch. As the disk revolves and a rotating pin approaches a fixed pin, the electric field strength becomes

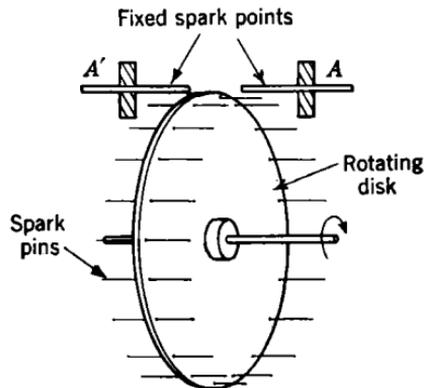


FIG. 10-43.—Rotary spark gap.

high enough for breakdown. Current flows from 4' to the disk pin and out through a second spark at point A. The simplicity of this device is partially offset by an uncertainty of up to $\pm 50 \mu\text{sec}$ in firing time, and by the narrow voltage range over which satisfactory operation occurs. The display-tube sweeps must be triggered by the appearance of the power pulse because of the great uncertainty in firing time. If f is the speed of the disk in rpm, and n is the number of pins, the resulting pulse rate is $nf/60$. For a given power output, there is an optimum spacing between rotating and fixed pins, and a minimum permissible spacing between pins in the rotating disk. These quantities vary somewhat with disk speed, so that it is not easy to design a rotary spark-gap pulser for variable pulse rate.

The Recharging Circuit.—In all forms of network pulsers it is necessary to recharge the network between pulses. This should not be done at too rapid a rate. A slow rate of charge is easily obtained by using a large inductance L_0 , which is also needed to prevent a virtual short circuit across the energy source every time the network is discharged. Inductance charging is used in practically all line-type pulsers, because it has the advantage of high efficiency and permits charging the pulse-forming network to a voltage nearly double that of the power supply, as shown below.

Consider first a d-c power supply voltage of negligible resistance, in series with an inductance, a switch, and a capacity C originally discharged. The energy supplied by the source in a time T after closing the switch is $V_s \int_0^T i dt$. If Q_c is the charge on the condenser, as long as the energy in the inductance at time T equals that at time 0,

$$V_s \int_0^T i dt = V_s Q_c = \frac{1}{2} CV_c^2.$$

But

$$Q_c = CV_c.$$

Then

$$V_s CV_c = \frac{1}{2} CV_c^2$$

or

$$V_c = 2V_s$$

and the voltage on the condenser C (or pulse-forming network) will be twice the supply voltage. It must be noted that this result is independent of the value of inductance used.

The network voltage obtainable with inductance charging in practice lies between 1.8 and 1.95 times the d-c supply voltage, because of resistance losses in the charging reactor. If L_0 and C are the values of charging inductance and network capacity, "resonance" charging is obtained when the repetition rate, $f_r = 1/(\pi \sqrt{L_0 C})$. If $f_r > 1/(\pi \sqrt{L_0 C})$, current in

the inductance does not reach zero at the time of firing, and "straight line" charging results. It is possible to operate a pulser with $f_r < 1/(\pi \sqrt{L_0 C})$ and still obtain substantially the same voltage step-up, but the network voltage at the time of firing is less than at some previous time in the charging cycle. Because of the undue stresses that are thus placed on the insulation, as well as the rapid increase in reactor losses, it is common practice to use, for this case, a "hold-off" or "charging" diode in series between the inductance and the network, to prevent the flow of reverse current.

If an a-c supply is substituted for the d-c energy source, the resulting pulser is inflexible in pulse rate and pulse length, but has the great advantage of extreme simplicity. It requires neither vacuum tubes nor auxiliary power supplies.

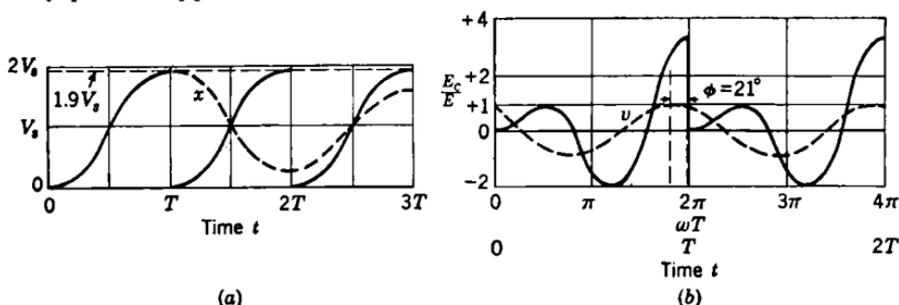


FIG. 10-44.—Line-type pulser charging waves. T = charging time; switch fires at $t = T, 2T, \dots$ (a) Network voltage for d-c resonance charging. Power supply voltage = V_0 ; pulse recurrence frequency = $\frac{1}{T_r} = \frac{2}{2\pi LC}$ cps; curve x is network voltage if switch misfires. (b) Network voltage for a-c non-resonance charging under conditions of maximum step-up ratio. Curve v is impressed a-c voltage; $v = E \cos(\omega t - \phi)$.

Figure 10-44a shows the voltage waveform across the pulse network for d-c resonance charging, and Fig. 10-44b for the condition of maximum one-cycle voltage step-up, which occurs at an impressed a-c frequency about 0.7 times the resonant frequency $1/(\pi \sqrt{L_0 C})$. A voltage step-up of about π can be obtained in a single cycle under these conditions, if the pulser is fired about 21° after the crest of the impressed a-c wave is reached.

Maximum simplicity is achieved by mounting a rotary spark gap directly on the shaft of the a-c machine exciting the network. Phasing is easily accomplished by mechanical adjustment. By proper design, the transformer used between generator and network can incorporate sufficient leakage inductance to resonate with the network capacity at the generator frequency. Pulsers of this type have been very successful; they are recommended for cases in which specifications are firmly fixed, suitable repetition rates correspond with available a-c generators, and the uncertainty in firing time of a rotary spark gap can be tolerated.

10-11. Miscellaneous Components. Pulse Transformers.—The development of pulse switching tubes and of pulse loads has proceeded so independently that usually there is little correlation between the optimum impedance levels for best operation of these two devices. Fortunately, satisfactory pulse transformers are available to match the impedance of the load to that of the generator. While the elementary theory of the pulse transformer is quite manageable, great care in design must be exercised to preserve good wave shape with fast rates of rise, a minimum of extraneous oscillation, and high efficiency. The ideal transformer is shown in Fig. 10-45a; Fig. 10-45b shows the equivalent circuit of a practical transformer (all values referred to the secondary or high-voltage

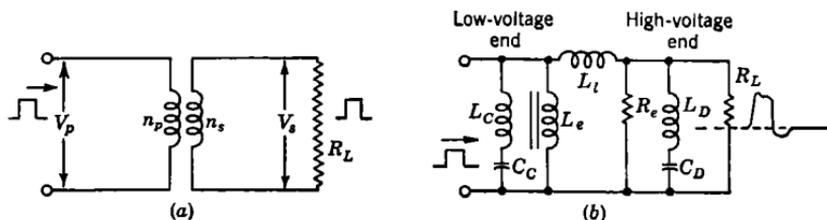


FIG. 10-45.—Pulse transformer circuit diagram.

(a) Ideal transformer; $\frac{V_s}{V_p} = \frac{n_s}{n_p}$; $\frac{I_s}{I_p} = \frac{n_p}{n_s}$; $Z_o = \left(\frac{n_s}{n_p}\right)^2$

(b) Equivalent circuit of pulse transformer.

end), exhibiting the effects of departure from the ideal case. The symbols shown in Fig. 10-45 are defined as follows:

L_l = leakage inductance due to flux from primary current which fails to link with secondary and represents effect of magnetic energy stored *between* windings by load current.

C_D = effect of electrostatic energy stored in the primary-secondary distributed capacity. If transformer is pulsed from low-voltage end C_D must be charged through L_l .

L_D = "squirted inductance" arising from nonuniform current distribution due to the charging of C_D . Current charging C_D must flow through L_D .

C_c = effect of electrostatic energy stored between primary and core.

L_c = effect of magnetic energy stored in "squirted flux" which comes from nonuniform current in primary coil arising from the charging of C_c .

L_o = effective shunt or self-inductance (input inductance with open circuit secondary).

R_e = resistance due to eddy current in iron and to hysteresis.

A good pulse transformer design attempts to maximize the shunt inductance L_o and minimize the leakage inductance L_l . In addition, undesirable oscillations arising from the series resonant circuits shown at

the input and output ends of Fig. 10-45*b* should be avoided. Leakage inductance can be kept to a reasonable value by winding the secondary and primary as close together as voltage breakdown will permit, for the space between these two windings is responsible for most of the leakage inductance. The shunt inductance and its magnetizing current constitute an additional load on the generator which must be minimized by making L_e large.

This last consideration sets a lower limit on the number of primary turns and also calls for a laminated core with high permeability at high frequencies. Laminations between 0.001 and 0.005 in. thick are necessary to maintain a core permeability of several hundred up to the frequencies of several megacycles per second present in a steep wave front. Special core materials for pulse transformers were developed during the war.

A square voltage wave of magnitude V_0 and duration t applied to the inductance L_e will build up a current of approximately $I_p = V_0 \cdot t/L_e$ amperes during the early part of its exponential rise. If this current is not to exceed a few per cent of the desired load current for pulse lengths around 1 μ sec, L_e must have a value between 10 and 20 mh. With this information, the number of turns in the primary and the required core area can be obtained from tables showing the effective permeability of the core at the flux densities and rates of rise anticipated.¹

Since the voltage drop across the leakage inductance is approximately $L_l \cdot \frac{di}{dt}$, the maximum permissible value of leakage inductance can be estimated by assuming that the current in R_L must reach 90 per cent of its final value, V_0/R_L , in a time of about $t_0/10$. The current which flows must satisfy the equation

$$V_0 = L_l \cdot \frac{di}{dt} + i \cdot R_L.$$

To a sufficient approximation, $L_l = R_L t_r \cdot \frac{I_0}{I}$, where I/I_0 is the fraction of final current built up in the rise time t_r . If $t_r = 0.1 \times 10^{-6}$ sec, $R_L = 1000$ ohms, and $I/I_0 = 0.9$, L_l must be less than 100 μ h.

Satisfactory high-power (100-kw to 5000-kw) pulse transformers have been designed which pass good wave shapes down to $t_0 = 10^{-7}$ sec and up to $t_0 = 10^{-5}$ sec, but it is quite difficult to design a pulse transformer to pass a wide range of pulse lengths. Although a long pulse calls for a large value of L_e , large L_e magnifies the difficulties of securing the small value of L_l required to pass a very short pulse. Transformers have been designed which satisfactorily pass pulse widths varying by a factor of 10; pulses of shorter or longer duration than the optimum suffer either

¹ See *Pulse Generators*, Vol. 5, Radiation Laboratory Series, Chaps. 12 to 15.

in rate of rise or in flatness of top. Wider-range transformers will be practical only when higher permeability cores and stronger insulation with lower dielectric constants are available.

Insulation between windings and adequate cooling must also be taken into account in design. High-power transformers lend themselves most readily to oil insulation because of the insulating and convection-cooling properties of the oil. In the lower-power range (200 kw and down) it is usually convenient to use solid dielectric materials, with a consequent saving in weight and size. Figure 10-46 shows the variation in size of a pulse transformer with power output. As in the case of pulse networks,

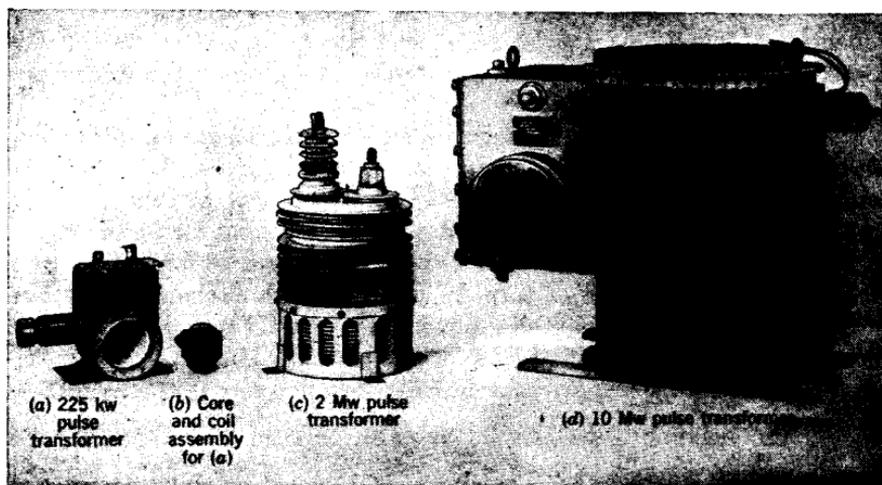


FIG. 10-46.—Pulse transformers.

the size and weight of a pulse transformer depends not only upon the pulse power to be handled, but also upon the pulse length and the repetition rate. These, together with the transformer efficiency, determine the amount of average power which must be dissipated by the transformer.

A pulse transformer contributes a certain amount of undesirable inductance and capacity to the pulser circuit. Special damping devices may be necessary to remove unwanted oscillations. Damping resistors and appropriately phased diodes are ordinarily used for this purpose.

Pulse Cables.—In handling high-power pulses with steep wave fronts, extreme care must be taken to shield the equipment sufficiently to prevent the radiation of signals which interfere with the operation of communications receivers and other electronic equipment. This shielding is especially necessary when the pulse must be transmitted several yards between pulser and load. Existing pulse transmission cables and their connectors provide satisfactory shielding and resistance to voltage breakdown, but are still rather bulky, hard to assemble, and heavy. Considerable improvement in their detailed design can be hoped for.

Energy Sources.—The energy source accounts for most of the weight and much of the complexity of any pulser except those of the a-c resonance charge type. All conventional types of rectifier circuit can be used in pulser power supplies. The only design considerations arising from the nature of the pulser load are those governing ripple and regulation. In general, the a-c supply frequency will be less than the pulse rate, and smoothing condensers capable of holding the voltage essentially constant

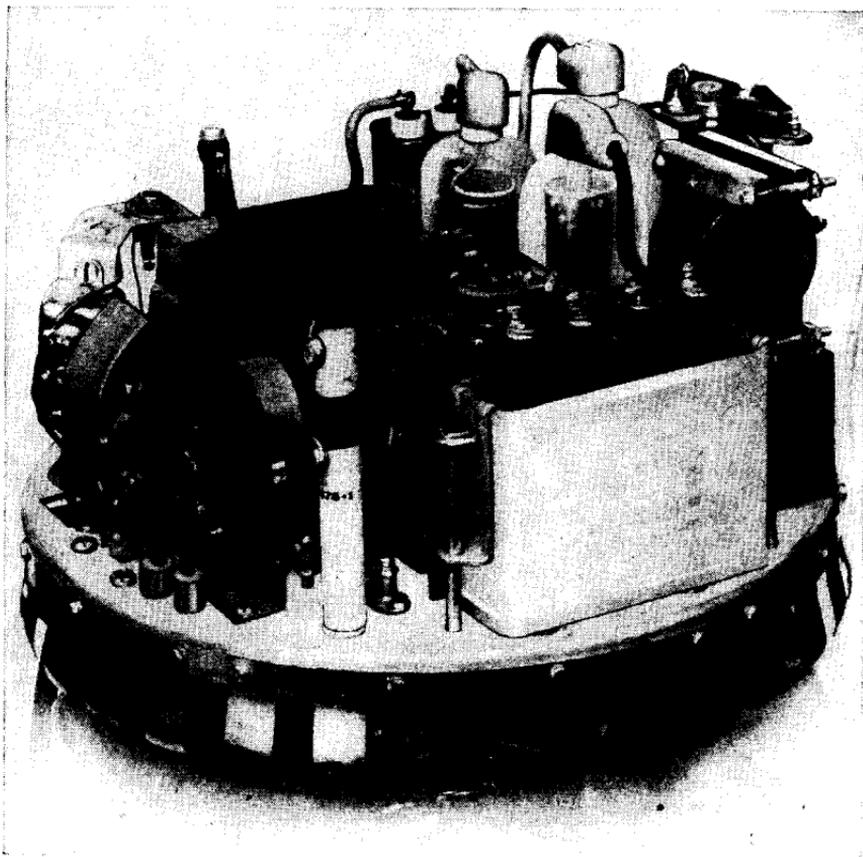


FIG. 10-47 —Low-power airborne pulser.

between pulses must be used. For example, a 60-cycle full-wave rectifier is brought to full voltage only 120 times per second, so that a pulse rate of 1200 cps would draw 10 pulses of power between successive recharge cycles. In figuring ripple it is usually satisfactory to assume a constant rectifier drain equal to the average current required by the pulser. The ripple must be kept below that which would cause a change in magnetron current sufficient to produce either a mode shift or undue frequency modulation. Apart from this consideration, amplitude modulation is

harmful only in those radars which detect moving targets (see Chap. 16) by comparing successive pulse phases.

We have already remarked that pulser internal impedance may sometimes play a deciding role in mode stability. It is also true that the power supply steady-state V-I characteristic may influence magnetron behavior.

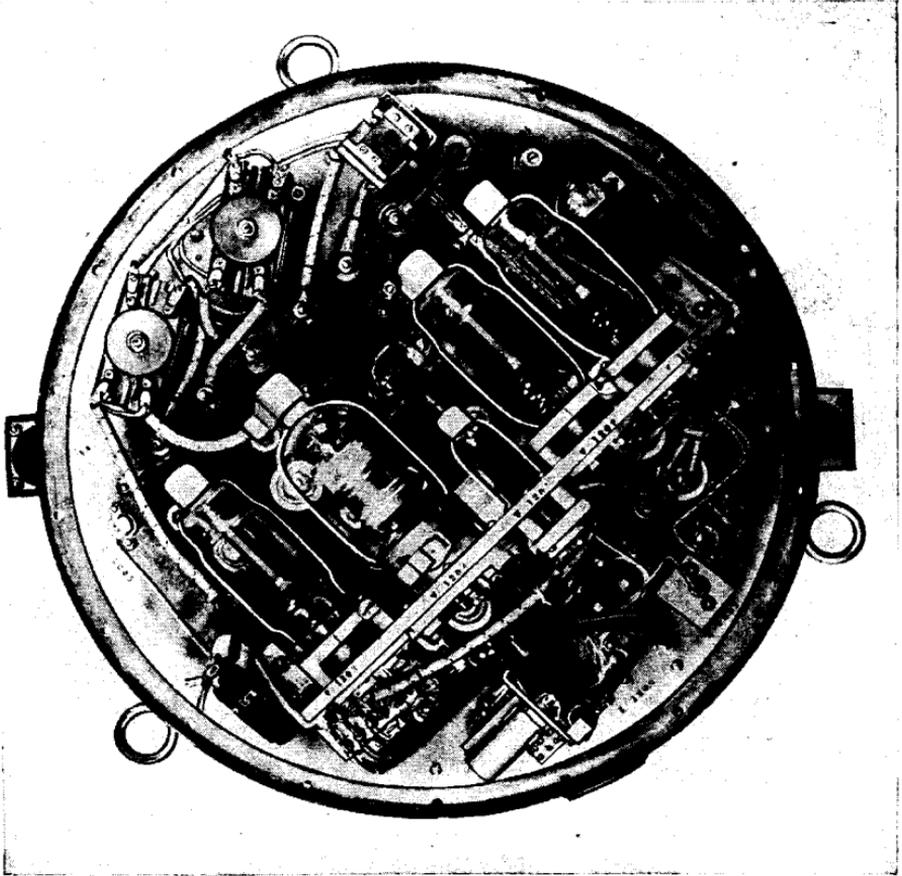


FIG. 10-48.—Medium-power airborne pulser.

A mode change in which an *increase* in voltage calls for a decrease in current can cause the instability typical of any negative-resistance load.¹ It is of considerable importance that the designer consider the problem of stability from the standpoint of the system as a whole.

The necessity of providing for a varying pulser load influences rectifier and power-supply design. It is frequently desirable to change the repetition rate of the pulser. This change is easy in principle, but in

¹ RL Report No. 809, Sept. 28, 1945.

practice many annoying points must be considered, including temperature shift, change in back-bombardment heating of the magnetron cathode, change in pulser starting delay, and change in rectifier voltage. The usual practice is to change the pulse lengths inversely with the pulse rate, thus keeping average power constant. But, because of the complications inherent in changing pulse length by switching pulse networks, the pulser designer is always inclined to stick to one pulse width. If pulse rate must be changed without a compensating change in pulse

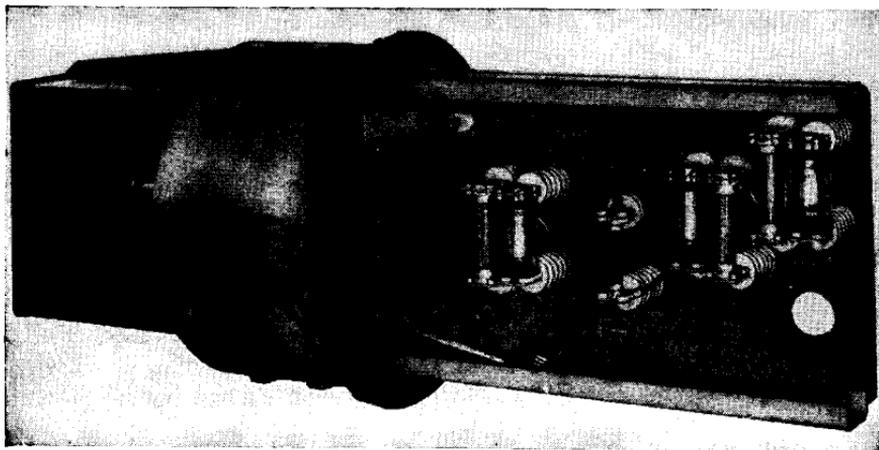


FIG. 10-49.—High-power airborne pulser.

width, it is usually necessary to control the rectifier output by switching power transformer taps or by varying a series primary impedance.

The present design of high-voltage, low-average-current rectifying diodes is quite satisfactory. There are now available diodes of reasonable dimensions, high inverse rating, and good life, covering most practically useful ratings. Thoriated tungsten cathodes are universally employed; these have satisfactory mechanical strength and are economical of filament power. Efficiency could be further increased by the development of a filamentless rectifier. This would permit operation without filament transformers which, in some circuits, must be insulated for high voltage. Perhaps a cold cathode discharge tube, or a barrier-layer rectifier, will one day replace the filamentary rectifier.

Line-type pulsers for radar applications have been built and operated successfully to cover a range of pulse power output of 1 kw to 20 Mw. The following photographs show typical pulsers designed for airborne service.

Figure 10-47 shows a hydrogen-thyratron pulser designed to supply

25 kw to the magnetron, at pulse durations of 0.8 μ sec and 2.2 μ sec. The components weigh approximately 6 lb.

Figure 10-48 shows a hydrogen-thyratron pulser designed to supply 225 kw to the magnetron, at pulse durations of 0.25, 0.5, 2.6, and 5.2 μ sec. The weight of components is about 40 lb.

Figure 10-49 shows a fixed-gap pulser designed to supply 2 Mw to the magnetron at a pulse duration of 2 μ sec. The complete pulser, including pressurized housing and cooling system, weighs around 300 lb.

MR. & MRS. JOHN W. CLINE
173 PRINCETON ROAD
AUDUBON 6, N. J.

CHAPTER 11
R-F COMPONENTS
BY A. E. WHITFORD

11.1. The R-f Transmission Problem.—In the block diagram of a basic radar system shown in Fig. 1-4, the parts shown as heavy double lines transmit the radio-frequency (r-f) energy from the magnetron to the antenna, and carry the faint echo signals into the T-branch where the receiving apparatus is located. For the types of radar treated in this book this function is performed by coaxial lines and waveguides. A considerable body of theory and a new set of techniques have grown up around this class of transmission circuits. The essential new feature is, of course, that the wavelength is of the same order as the physical size of the circuit elements; the length of the line may be many wavelengths. Although space permits only a limited treatment here,¹ sufficient introduction will be given to show the general approach, and to make understandable some of the reasons for current practices in microwave radar.

Standing Waves.—When a voltage is suddenly applied to the input terminals of a long and uniform transmission line, the current which flows in the initial interval, before reflections from the far end arrive to confuse the situation, is determined by the property of the line known as its "characteristic impedance." The characteristic impedance, Z_0 , is a function of the geometry of the conductors and insulators of the line, and for good conductors and low-loss dielectrics is almost purely resistive. For a concentric line, neglecting losses,

$$Z_0 = \frac{138}{\sqrt{k}} \log_{10} \frac{r_2}{r_1} \quad \text{ohms}$$

where k is the dielectric constant of the material in the annular space between the conductors, r_2 is the inner radius of the outer conductor, and r_1 is the outside radius of the inner conductor.

A uniform line terminated at any point in its characteristic impedance behaves as if the line were infinitely long; there is no reflection. Power introduced at the input terminals disappears into the termination with small losses in the line. However, any discontinuous change along the

¹ More extended treatment can be found in other books of this series, especially (1) *Microwave Transmission Circuits*, Vol. 9; (2) *Microwave Duplexers*, Vol. 14; (3) *Waveguide Handbook*, Vol. 10; (4) *Principles of Microwave Circuits*, Vol. 8.

line, such as might be introduced by a change in dimension of one of the conductors, or any change in geometry introduced by a sharp bend, or by a dent or an obstacle in the line, will produce a reflection. The reflected energy travels back toward the source. This results in standing waves in the line. These can be observed by sliding a small probe along a slot in the line after the manner shown for a concentric line in Fig. 11-1.

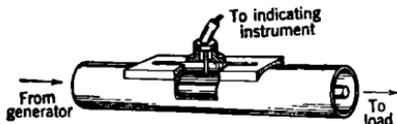


FIG. 11-1.—Slotted coaxial line and probe for observing standing waves (cutaway view).

Only a negligible fraction of the power is abstracted by the probe, but, fed into a suitable indicator, this is sufficient to register the voltage variations along the line. In general there will be voltage maxima spaced at half-wavelength intervals with minima halfway between them.

Only if the line is perfectly matched will the voltage reading be constant as the probe moves along. The ratio of the maximum to the minimum voltage is called the “voltage standing-wave ratio” (VSWR) and is the usual criterion of how well a line is matched. Complete reflection at the end of the line, such as would be expected from an open circuit or a short circuit, results in zero voltage at half-wave intervals and a VSWR of infinity. The mismatch may also be expressed as a power ratio (PSWR), or in decibels. The relations between these three measures are

$$\begin{aligned} \text{PSWR} &= (\text{VSWR})^2 \\ \text{SWR, db} &= 10 \log_{10} (\text{PSWR}). \end{aligned}$$

Quarter-wave and Half-wave Lines.—It can be shown that when a lossless transmission line of characteristic impedance Z_0 a *quarter wavelength* long is terminated in an impedance Z_i , the input impedance is

$$Z_i = Z_0^2 / Z_i. \quad (1)$$

This property is widely used. For example, two transmission lines of differing impedance can be matched to each other by joining them through a quarter-wave line whose characteristic impedance is the geometric mean of that of the two lines. This is called a matching transformer. Or, if a quarter-wave line is terminated by a short circuit, the input impedance is infinite, i.e., equivalent to an open circuit. Conversely, an open-circuited quarter-wave line appears at the input terminal to be a short circuit. For a lossless line *half a wavelength* long

$$Z_i = Z_i, \quad (2)$$

irrespective of the characteristic impedance of the line. This principle has many uses also, particularly in duplexers (Sec. 11-5) and mixers (Sec. 11-8).

Why a Matched Line?—The fraction of the incident power reflected to the source from a section of transmission line of given VSWR is

$$\text{Power reflection coefficient} = \frac{[(\text{VSWR}) - 1]^2}{[(\text{VSWR}) + 1]^2} \quad (3)$$

For the usual upper design limit of $\text{VSWR} = 1.5$, it is seen that the power reflection loss is only 4 per cent, or 0.3 db, surely not serious. In high-power systems or unpressurized airborne systems, where line breakdown is a possibility, the strain is of course higher for a high VSWR. A ratio of 1.5 means that for a given breakdown gradient, 33 per cent less power can be delivered to the load than could be delivered in a matched load. This can be a limitation.

The strongest requirement for a well-matched line arises from the properties of the magnetron. Like all self-excited oscillators, the magnetron exhibits an output frequency and a stability dependent upon the load into which it works. A mismatched line represents a resistance lower than Z_0 at voltage minimum, and higher than Z_0 at voltage maximum. At other phases it has a reactive component which may be either positive or negative. As explained in Chap. 10, magnetrons are in general designed to be stable against a VSWR of 1.5 in any phase. This is the origin of the commonly specified upper limit for mismatch.

There is a further limitation if the mismatch occurs at the end of a long line—for example 50 to 100 wavelengths from the magnetron. Then as the frequency changes the number of wavelengths in the line changes, and so also does the phase of the standing wave. The line impedance seen at the magnetron is therefore a rapidly varying function of frequency. If at a certain frequency the phase happens to be such that the variation of reactance of the line with frequency is more rapid than that of the magnetron itself, and of opposite sign, a condition results where the magnetron has no stable frequency. In another, favorable, phase the magnetron is stabilized. This is known as the "long line effect."¹ The result is that for long lines either (1) VSWR's lower than 1.5 are necessary to guarantee stable magnetron operation, or (2) a method of changing the effective line length and hence the phase of the standing wave must be included in the antenna line. The latter may be done by a "line stretcher" not unlike a trombone, or in waveguide by a squeeze section or a dielectric phase shifter.² These add an undesirable adjustment.

11-2. Coaxial Lines.—Coaxial lines consisting of concentric inner and outer conductors are not new. At lower frequencies they have usually consisted of cables with a solid dielectric and a braided outer

¹ *Microwave Magnetrons*, Vol. 6, Radiation Laboratory Series.

² *Microwave Transmission Circuits*, Vol. 9, Radiation Laboratory Series.

conductor. These are used at microwave frequencies also, especially for short, low-power interconnecting cables in the 10-cm region. However, the attenuation even in the best dielectric, as seen in Table 11-1, begins to be serious at 10 cm and gets worse at shorter wavelengths. Substitution of air for the solid dielectric eliminates dielectric losses, and only the much smaller conductor losses remain. But the center conductor must somehow be supported mechanically. Thin dielectric beads have been used as supports, but the disintegrating effect of dielectric breakdown over the surface of the bead is hard to avoid. Reflections from the beads can be largely canceled by proper spacing.¹ However, bead-supported

TABLE 11-1.—STANDARD MICROWAVE TRANSMISSION LINE

Dimensions OD, in.	Wall, in.	Wave-length, cm	Maximum power,* Mw	Attenuation,† db/m	Wave-length range, cm
A. Waveguide					
$1\frac{1}{2} \times 3$	0.080	10.0	10.5	0.039	7.6-11.8
$\frac{5}{8} \times 1\frac{1}{4}$	0.064	3.2	1.77	0.15	3.0-4.7
$\frac{1}{2} \times 1$	0.050	3.2	0.99	0.24	2.4-3.7
$\frac{1}{4} \times \frac{1}{2}$	0.040	1.25	0.22	0.5‡	1.1-1.7
B. Rigid coaxial lines, stub-supported					
Outer $1\frac{3}{8}$	0.049	10.0	4.2	0.08	9.3-11.7
Inner $\frac{5}{8}$	0.035				
Outer $\frac{7}{8}$	0.032	10.0	1.3	0.15	9.1-11.7
Inner $\frac{3}{8}$	0.032				
Outer $\frac{1}{2}$	0.032	3.2	0.36	0.49	3.1-3.5
Inner $\frac{3}{16}$	0.032				
C. Flexible coaxial cable, RG-9/U, polyethylene dielectric					
Dielectric OD 0.280 in.		10.0	0.31§	0.56	3.0-∞
Inner conductor 7 strands #21 AWG,					
Outer conductor double braid		3.2	0.31§	1.12	

* Computed for maximum gradient of 30 kv/cm. No allowance in coaxial lines for increased field around stub supports. Practical operating point $\frac{1}{2}$ to $\frac{2}{3}$ of values given.

† For brass walls. For copper or silver walls, attenuation is about half that listed, and for silver plating has an intermediate value, depending on finish.

‡ Experimental value for coin-silver tubing, generally used at this wavelength. Surface finish affects value.

§ Specification limit for cable alone. Connectors limit safe power to a few kilowatts.

¹ *Microwave Transmission Circuits*, Vol. 9.

lines have been almost completely supplanted by stub-supported lines in the 10-cm region.

The principle of the stub support is shown in Fig. 11-2. As was mentioned in Sec. 11-1, the input impedance of a quarter-wave line shorted at the far end is the same as an open circuit. When placed in parallel with the main line such a connection has no effect at all on the impedance

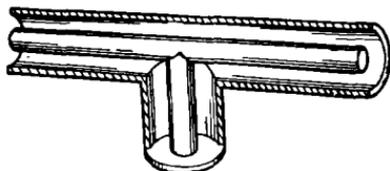


FIG. 11-2.—Simple quarter-wave stub support.

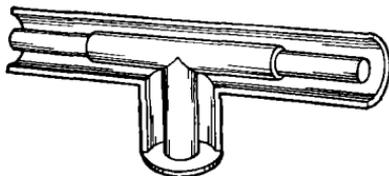


FIG. 11-3.—Broadband stub support.

and causes no reflection. The mechanical and electrical superiority of a solid piece of metal as a support and insulator is obvious, and at 10 cm the length of the stub (about 1 in.) is such as to make the projection short and unobtrusive.

Obviously a quarter-wave stub can have the desired property at only one frequency. Deviations of only 1 to 2 per cent in frequency cause the stub to have a reactance that presents an appreciable mismatch. Figure 11-3 shows a broadband stub support where the frequency sensitivity is compensated over a band of ± 15 per cent. At the center of the band the stub has an effective length of exactly a quarter wave, and the two quarter-wave sections in the main line transform to an impedance

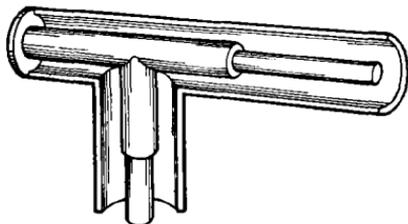


FIG. 11-4.—Stub-supported elbow.

lower than normal and then back to normal. The conditions for no reflection are satisfied. At a frequency lower, for example, than the center frequency, the stub is less than a quarter wave, but the inductive reactance thereby presented at the T-junction is made just enough to compensate for the fact that the quarter-wave transformers in the main line are also less than a quarter wave long, and would present a mismatch in the absence of the stub. Similar, but converse, conditions obtain for frequencies higher than band center. Figure 11-4 shows how a broadband stub can be used to make an elbow. There is an added complication because the sharp elbow introduces a reactance which must be compensated for in the construction of the stub.

Such supports are standard in coaxial lines used in the 10-cm region and regularly have a VSWR less than 1.03 over a band of ± 15 per cent from the center frequency.

The standard connector for joining coaxial lines is shown in Fig. 11-5. Since longitudinal currents cross the junction plane on both inner and outer conductor, good contact must be assured. For the outer conductor this is done by pulling together two mating cones of differing taper by means of the strong outer clamping rings. The fittings that solder to

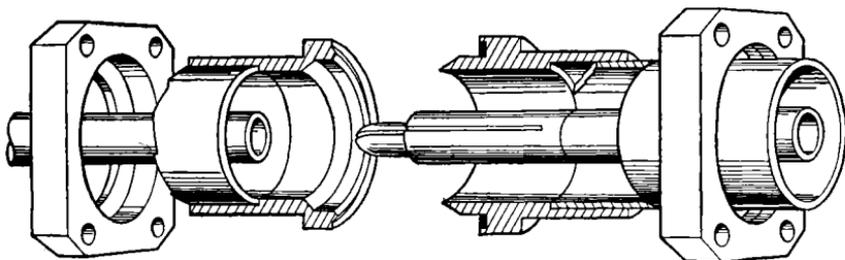


FIG. 11-5.—Coupling for coaxial line.

the outer tube also contain a gasket groove for keeping the line airtight. The inner conductor is itself a tube. The two pieces are joined by a beryllium copper "bullet" which is soldered into one piece and makes tight contact with the inner surface of the other by means of expanding prongs on the rounded tip.

In order to transfer r-f power to a rotating scanner, a rotary joint for a coaxial line is necessary. Early designs involving wiping contacts on

both the inner and outer conductors gave difficulties arising from poor contact, sparking, and wear. The superiority of the noncontact type employing choke joints has led to its universal adoption. The principle is shown in its simplest form in Fig. 11-6a. The gap between the stationary and rotating parts of both conductors is situated at the end of an open-ended quarter-wave coaxial line. As was mentioned in Sec. 11-1, the impedance at the input end of an open quarter-wave section (in this case across the gap in the line) is zero. Power flows across without loss

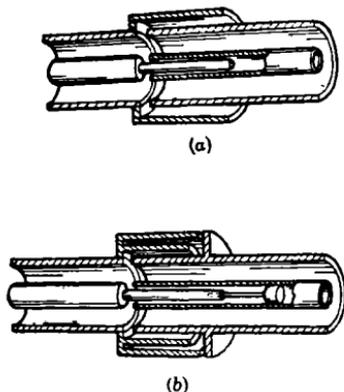


FIG. 11-6.—Choke-type rotary joints.

or sparking. Closer analysis shows that the open end of the quarter-wave section on the outer conductor is not an infinite impedance, because there is some radiation, producing a finite radiation resistance. This effect can be reduced, and the match improved, by adding a short-circuited quarter-wave line in series with the outer gap, as in Fig. 11-6b. The outermost gap can now be very small, or even a rubbing contact,

since no current flows across it. A similar improvement is made on the center conductor. Advantage is taken of the negligible current across the transition between the first and second quarter-wave sections to make a contact bearing between the innermost conductors. This is a great aid in keeping the closely spaced tubes all concentric. Quarter-wave line sections used in the manner just described to prevent loss of microwave energy into side channels are termed "chokes."

The nominal characteristic impedance of the standard rigid coaxial lines in regular use is 50 ohms, corresponding to a ratio of 2.30 for the radii of the inner and outer conductors. This is a compromise between a ratio of 3.60 (77 ohms), which, for a given outer diameter, gives the lowest attenuation due to conductor losses, and a ratio of 1.65 (30 ohms) which maximizes the power that can be carried with a given breakdown voltage gradient. Since the attenuation in a 50-ohm line is only 10 per cent greater than it is in a 77-ohm line, it is not a costly compromise. Increasing the size of both conductors to increase the air gap and thereby increase the power-carrying capacity cannot be carried on indefinitely, since a higher mode of propagation, with diametral rather than axial symmetry, can be excited when the mean circumference of the annular dielectric space exceeds one wavelength. The possibility of two modes of propagation existing simultaneously in a single line leads to serious complications. The highest powers are best transmitted by means of waveguide.

Rigid coaxial lines are not ordinarily used for wavelengths below 8 cm because the limitation on over-all size just mentioned permits too low a maximum power-carrying capacity. However, a standard $\frac{1}{2}$ -in. OD stub-supported line with appropriate couplings has been worked out for the 3-cm band. The largest stub-supported standard coaxial line is in $1\frac{1}{8}$ -in. OD tubing, with stubs designed for the 9- to 11-cm band; the higher-mode limit prohibits anything larger on this band. For moderate powers, $\frac{7}{8}$ -in. OD line with stubs designed for the 9- to 11-cm band is standard. The theoretical breakdown power for the $\frac{7}{8}$ -in. line, assuming sea-level pressure and a maximum field of 30 kv/cm, is 1.3 Mw. However, nonuniform fields around stubs and the increased gradient in the chokes on the inner conductor of rotary joints make the safe engineering design limit about 0.3 Mw.

The type of flexible coaxial cable most commonly used in the microwave region (Army-Navy designation RG-9/U) has a polyethylene

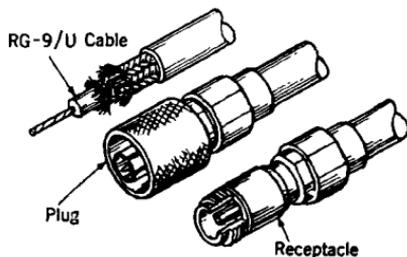


FIG. 11-7.—Coaxial cable and matched type N fittings.

dielectric of nominal outer diameter 0.280 in. A section of the cable with the standard type N connector is shown in Fig. 11-7. The connectors match the 50-ohm impedance of the cable at 10-cm and longer wavelengths. The mismatch at 3 cm is not great. Breakdown in the connectors limits the peak power to a few kilowatts; the most common use of such cable is in test equipment. Attenuation data are given in Table 11-1.

11-3. Waveguide.—Although a metallic pipe of almost any shape will transmit or guide electromagnetic waves if their wavelength in air is short enough, rectangular tubing whose internal dimensions have a ratio between 2.0 and 2.5 has been almost universally adopted where the problem is simply the transfer of microwave energy. (Use of round guide in the special case where axial symmetry is required is discussed in a later paragraph.) A detailed understanding of the propagation of waves in a region bounded by conducting walls can only be obtained from the solution of Maxwell's equations. Practically, however, the results of the mathematical analysis¹ have come to be used in a procedure which retains most of the concepts of transmission-line theory, with equivalent lumped reactances connected at suitable points to account for the effects of discontinuities.

The resemblance between a rectangular waveguide and a two-wire transmission line is shown in Fig. 11-8*a* to 11-8*d*. In Fig. 11-8*a* is shown a single quarter-wave stub support, analogous to the coaxial stub support described in Sec. 11-2. At the proper frequency the input impedance of the short-circuited stub is extremely high and there is no effect on the propagation of the wave on the line. In Fig. 11-8*b* a great many stubs, extending both ways from the two-wire line, have been added, still without affecting the propagation of the frequency in question. In Fig. 11-8*c* the stubs have coalesced into a rectangular tube which looks like a waveguide. For a single stub, a slight correction to the length is necessary to allow for the inductance of the crosspiece, but when the stubs become a solid tube, no lines of force can link the narrow side, and the quarter-wave distance becomes exact. This also implies that the length of the narrow side of the tube is not critical.

The two-wire transmission-line model explains how a waveguide can transmit all frequencies higher (wavelengths shorter) than that for which the quarter-wave stubs were designed. In such a case, as shown in Fig. 11-8*d*, the two wires become broad busbars with only as much of the wide side of the guide given over to stubs as is required by the now shorter wavelength. However, wavelengths greater than twice the broad dimension cannot be propagated because then the stubs become less

than a quarter wave and shunt the line with a rather low inductive impedance which would stop transmission.

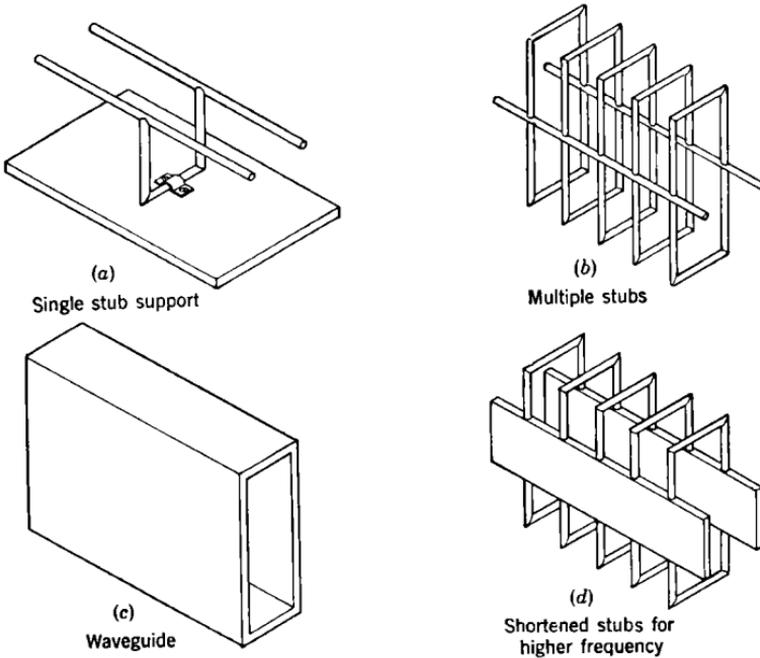


FIG. 11-8.—Waveguide derived from stub-supported two-wire transmission line.

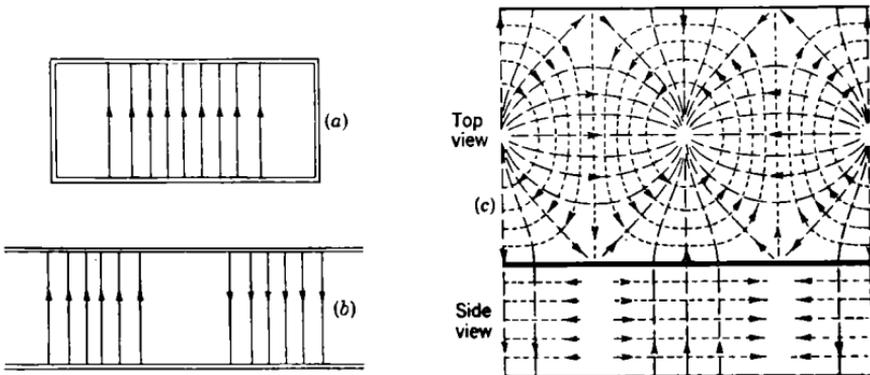


FIG. 11-9.—Fundamental mode of wave propagation in a rectangular guide; (a) and (b) show the electric field in a transverse and longitudinal cross section; (c) shows the lines of current flow in the top and side of the waveguide as long dashes; the dotted lines represent tangential magnetic field at the wall.

Figure 11-9 shows an instantaneous picture of the electric and magnetic fields in a rectangular waveguide whose wide dimension is slightly over half the free-space wavelength. The lines of current flow in the

walls are also shown. At all microwave frequencies, the skin effect confines the current to a microscopically thin layer on the inner surface. As the dimensions of the waveguide are increased, the frequency being fixed, propagation becomes possible by *modes*—that is, by particular types of “vibration” in the electromagnetic field, other than the fundamental mode illustrated in Fig. 11-9. Each of these higher modes has its own characteristic electromagnetic field configuration. Ordinarily it is advisable to avoid propagation in more than one mode, and this is most easily done by choosing the dimensions of the guide so that the lowest mode, and the lowest mode only, can propagate. However, for certain applications some of the higher modes are useful. A notable example, to which we shall return later, is the second mode in waveguide of circular cross section. This has axial symmetry and is thus useful in waveguide rotary joints.

For each type of waveguide there exists a critical, or cutoff, frequency for propagation in the lowest mode. Waves of higher than critical frequency are transmitted; those of lower frequency are rapidly attenuated.¹ Corresponding to the cutoff frequency f_c is a cutoff wavelength λ_c related to f_c by $\lambda_c = c/f_c$ where c is the velocity of light. That is, λ_c refers to the wavelength in space. For rectangular guide, as shown from the stub-supported two-wire line, the cutoff wavelength is twice the broad dimension. In other words, a guide that is to transmit a wave must have a broad dimension greater than half a free-space wavelength. If the width is more than a *whole* free-space wavelength, a higher mode of propagation with a node in the electric field down the center becomes possible, which adds most undesirable complications. Therefore, the broad dimension must lie between a half and a whole free-space wavelength. The wavelength inside the guide is longer than that in free space and is given by the relation:

$$\text{Guide wavelength} = \lambda_g = \frac{\lambda}{\sqrt{1 - \left(\frac{\lambda}{\lambda_c}\right)^2}} \quad (4)$$

where λ is the free-space wavelength and λ_c is the cutoff wavelength (here equal to twice the broad dimension). When λ_c is only slightly greater than λ , the guide-wavelength becomes very long and varies rapidly with changes in λ . This greatly increases the frequency sensitivity of quarter-wave sections of guide used in duplexers and mixers (Secs. 11-5 and 11-8) and handicaps broadband design. The other extreme of a close approach to the boundary of the higher mode, corresponding to a wide dimension of nearly a whole free-space wavelength, runs into difficulty because of

¹ In a waveguide beyond cutoff the voltage or current falls off exponentially with distance, with constants exactly calculable from the dimensions and frequency. One form of standard attenuator utilizes this fact.

too gradual an attenuation of the higher modes inevitably excited at discontinuities such as T-junctions and diaphragms. For these reasons, waveguides are ordinarily used only for frequencies where the broad dimension lies between 0.60 and 0.95 of the free-space wavelength.

Since the maximum electric field comes across the narrow dimension of the guide, it is undesirable to choose this dimension too small. In fact, the power that can be transmitted for a given breakdown field is directly proportional to the guide dimension in the direction of the field. The height is limited by the requirement that the narrow dimension be less than half a free-space wavelength in order to avoid the possibility of propagating the simplest mode having polarization at right angles to that shown in Fig. 11-9.

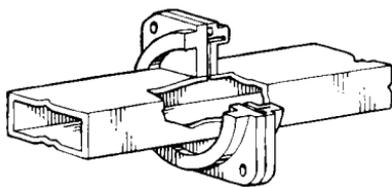


FIG. 11-10.—Waveguide choke coupling, 3-cm band.

There is no unique or generally accepted definition of the impedance of a waveguide. This might be expected from the lack of definite localized terminals at which the voltage and current could, in principle at least, be measured. However, the usual procedures of impedance-matching are carried over from transmission-line theory, and calculations are made on the basis of *normalized impedances*. The impedance of a standard waveguide is defined as unity, and resistive or reactive elements inserted in the guide are computed relative to the standard, rather than in ohms.

A typical choke joint between pieces of waveguide is shown in Fig. 11-10. The principle is identical with that discussed in Sec. 11-2 for the outer conductor of a coaxial rotary joint. The diameter of the radial section spreading out at right angles to the rectangular tube is chosen so that the average or effective distance from the inner surfaces of the waveguide is a quarter wavelength. A circular groove, likewise a quarter wavelength deep, forms the short-circuited terminating section. A rubber gasket in the outer groove serves to keep the waveguide airtight. By careful choice of dimensions, such a joint can be made to be a good match over a frequency band 12 to 15 per cent wide. Since no current flows across the gap between the choke and its mating flange, physical contact is not necessary. The power flows across a small gap with negligible loss. However, in such a case the leakage of radiation, although small compared to the transmitted power, may still overwhelm sensitive energy detectors nearby. In cases where electrical leakage must be minimized and the outer gasket groove is not needed for pressurization, an electrical gasket is substituted. Such a gasket is made by pressing a ring of woven metal gauze into the proper form.

Contact-type unions have also been used successfully. The flanges are relieved so that the bolts bring maximum local pressure to bear near the junction of the waveguide walls. Abrasion of the mating surfaces is much more serious in such a union than in the choke-type joint.

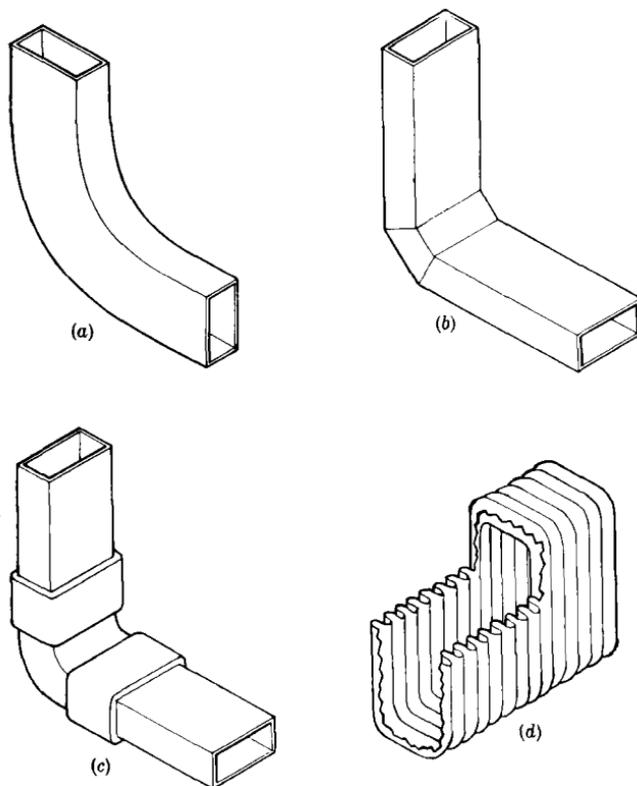


FIG. 11-11.—Waveguide elements. (a) Long-radius bend in waveguide. (b) Miter elbow. (c) Short-radius bend, electroformed fitting. (d) Flexible waveguide, corrugated wall.

Bends in waveguide such as those in Fig. 11-11a cause inappreciable mismatch if the inside radius is greater than twice the free-space wavelength. Short-radius bends are well matched if the length along the center of the guide is half a guide wavelength. Such elbows as those shown in Fig. 11-11c, are produced by electroforming—that is, plating copper or other metal on a soft metal mold which is later melted out. Two-cut miter elbows, such as those shown in Fig. 11-11b, are well matched if the distance between cuts, measured along the center of the guide, is a quarter of the guide wavelength. Flexible waveguide can be made by winding it up out of metal strip in the same way that certain

types of metal hose or conduit are produced. A molded rubber sheath pressurizes and protects the piece, as well as holding the adjacent turns in tight contact. For short lengths, convolutions small compared with the wave-length can be formed hydraulically, and thereby give flexibility to a continuous metal tube. Figure 11-11*d* shows an example.

Transitions between waveguide and coaxial lines usually take the form of a quarter-wave stub antenna on the coaxial line projecting into the waveguide a quarter guide-wavelength from an endplate, as shown in

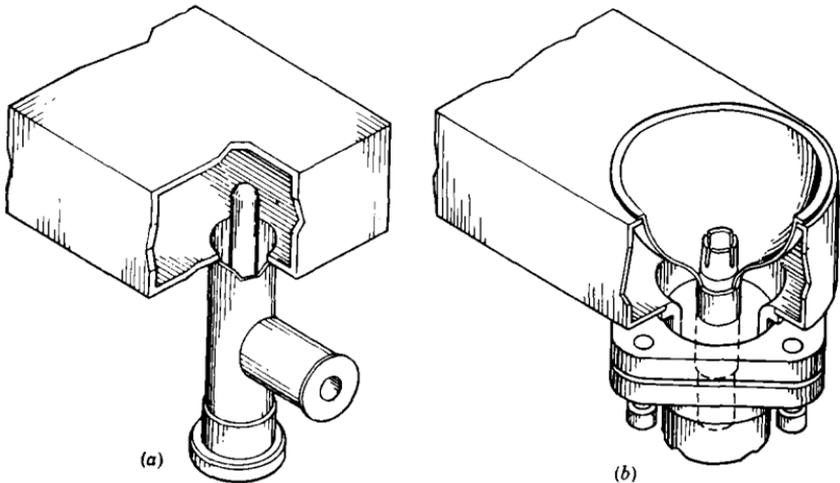


Fig. 11-12.—Coaxial line to waveguide transitions. (a) Probe transition, coaxial line to waveguide. (b) Doorknob transition.

Fig. 11-12*a*. The endplate reflects the energy going in that direction back in phase with that going down the guide. Expressed in terms of impedances, the short-circuited quarter-wave section of guide presents an open circuit at the probe. Then the load seen at the probe is only a single unit of guide impedance, rather than two units in series. Such a probe lowers the breakdown potential of the guide. The “doorknob” transition of Fig. 11-12*b* is designed to minimize breakdown. It can be thought of as a quarter-wave probe with a special form of stub support for the tip of the probe in which capacitive and inductive effect have been balanced against each other so that as a support it presents no loading of the waveguide. Rounded contours reduce the electrical gradient as much as possible.

Rotary joints between pieces of waveguide may consist of a coaxial rotary joint of the type described in Sec. 11-2 with transitions to waveguide at each end. This is common in the 10-cm region. The large $1\frac{1}{8}$ -in. OD line and the doorknob transitions just mentioned are used

where powers of the order of a megawatt are handled. In the 3-cm region transition is made to round guide, where the next-lowest mode has axial symmetry. Currents flow across the junction between the rotating and nonrotating parts of the round tube by means of the same folded choke arrangement used for the outer conductor of the coaxial rotary joint described in Sec. 11·2. Figure 11·13 shows a typical joint. The lowest mode in round guide has diametral rather than axial symmetry,

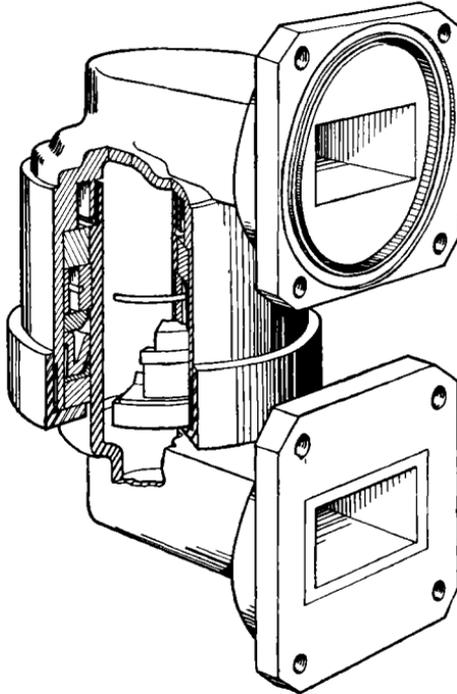


FIG. 11·13.—Waveguide rotary joint using round waveguide and axially symmetrical mode.

and if present to an appreciable extent will cause serious variation in voltage standing-wave ratio as a function of angle of rotation. The transitions from rectangular to round guide are designed to avoid exciting the undesired mode as far as possible. The guide wavelengths of the desired and undesired modes are sufficiently different so that a suitable choice of length of round guide will minimize coupling of the undesired mode from one rectangular guide to the other. Various absorbers for the undesired mode are sometimes used.

The *power-handling ability* of a waveguide, calculated from the electric fields involved, is approximately twice that of the largest coaxial line that

could be used for the wavelength carried by the waveguide. Since, however, the distortion of the normal field by the stub supports of the coaxial line makes it impossible to realize the calculated limit, the factor of two does not represent the full superiority of waveguide. In either waveguide or coaxial line, small nicks, burrs, or solder fillets can easily cause breakdown at a fifth of the calculated maximum power. Table 11.1 summarizes the properties of the waveguides widely used in microwave radar, and for comparison those of some standard coaxial lines and cable. *Attenuation* in waveguide is seen to be about half that in the largest coaxial line suitable for a given wavelength. For smaller coaxial lines of a given impedance the attenuation is inversely proportional to diameter. The conclusion to be drawn is that waveguides are superior electrically to coaxial line in nearly every respect. They are easier to fabricate because the inner conductor and its precisely machined stub supports are simply omitted. For these reasons, waveguide is almost universally used for wavelengths below 8 cm. In the 10-cm region where the size and weight of the $1\frac{1}{2}$ - by 3-in. waveguide are awkward and the power-carrying capacity is not needed, coaxial line is frequently used. For "long wave" radars (wavelengths of 50 cm and greater) waveguide is never used because of its relatively enormous size.

11.4. Resonant Cavities.—If both ends of a waveguide are closed by a short-circuiting plate, and energy is introduced by a probe so small that it does not appreciably change the properties of the enclosure, the amplitude of the standing-wave pattern in the waveguide will show a sharp maximum when the frequency is such that the length of the enclosure is an integral number of half guide-wavelengths. The reflections will then be in the proper phase to reinforce each other and cause a resonant buildup. (This is the property used in wavemeters.) For standard rectangular guide, reference to Eq. (4) shows that if the broad dimension is taken to be 0.707λ then half a guide-wavelength is also 0.707λ . The shortest resonant piece of such a waveguide is therefore square. The height does not affect the resonant wavelength, though if it is greater than $\lambda/2$, modes polarized at right angles to the desired mode become possible. Rounding off the corners of the square box shortens the resonant wavelength slightly; exact calculation shows that for a cylindrical box the resonance occurs when $\lambda = 1.30$ times the diameter, as opposed to 1.41 times the side of the square.

These round and square boxes are examples of *resonant cavities*, which play the same role in microwave transmission circuits as do resonant circuits involving lumped inductance and capacity in traditional circuit theory. Any hollow metal enclosure is capable of supporting oscillations in a large number of modes. In practice the geometry is usually chosen so only a single mode, often the lowest, is excited. For simple geometri-

cal forms the properties are completely calculable from Maxwell's equations.

In cavities of the type used in magnetrons, the inductance and capacity are fairly well separated and approximate numerical values can be calculated though they cannot be measured independently. For something like a simple cylindrical cavity, however, inductance and capacity are blended and cannot even be calculated unambiguously. The significant quantity is the resonant frequency or wavelength.

The second property of cavities that is important in microwave work is their Q . As in lumped-constant circuits, the value of Q is a measure of the sharpness of the resonance, and is determined by the dissipative elements loading the resonant circuit. If f_0 is the resonant frequency and f_1 and f_2 are the "half-power points,"—that is, the two frequencies, one above f_0 and one below f_0 , at which the voltage (or current) in the cavity is 0.707 as great as it is at resonance—then

$$Q = \frac{f_0}{f_1 - f_2}. \quad (5)$$

An equivalent, but somewhat more general, formula for Q involves the amount of electromagnetic energy stored in the oscillating field within the cavity, and the rate at which energy is dissipated in the walls or in any other way. If we denote the total stored energy by W , and the energy dissipated during one r-f cycle by w , Q is given by

$$Q = \frac{2\pi W}{w}. \quad (6)$$

If w includes only the dissipation within the cavity itself, due to the resistance of the walls and to dielectric losses in insulators within the cavity, etc., the Q defined above is called the *unloaded* Q , usually written Q_0 . If w includes, in addition, energy dissipated in external circuits coupled to the cavity, we obtain instead the *loaded* Q , or Q_L , which of course can never exceed Q_0 .

In a simple cylindrical cavity made of copper and resonant at 3000 Mc/sec in its lowest mode, the unloaded Q is about 15,000. Generally, the Q of a cavity loaded only by the resistance of its walls depends on the ratio of the *volume* of the cavity to the product of the internal surface *area* and the *skin depth*. For cavities of similar shape, the resonant frequency f_0 is inversely proportional to a linear dimension of the cavity; the skin depth varies as $1/\sqrt{f_0}$. It follows that Q_0 , for cavities of similar shape, varies as $1/\sqrt{f_0}$. On the other hand, the Q values that can be attained at microwave frequencies, typified by the example just given, are much higher than can be realized with coil and condenser combina-

tions at low frequencies. The essential reason for this is that such low-frequency circuits do not provide a correspondingly large volume for the storage of energy.

To make a cavity useful it is necessary to provide some means of introducing and removing energy, or in other words to couple it to the external circuit. This may be done by an electron stream, by a coupling loop to a coaxial line, or by an iris (hole) leading into a waveguide. Examples of these are cited in the discussion of klystrons (Sec. 11-7) and of TR switches (Sec. 11-5).

11-5. Duplexing and TR Switches.—As was explained in Sec. 1-3, the use of a common antenna for transmitting and receiving requires fast-acting switches¹ to disconnect the receiving apparatus from the antenna during the transmitted pulse, and to disconnect the magnetron during the period when echoes are being received. These two switches are called the TR (transmit-receive) switch and the anti-TR or ATR switch, respectively. The duplexer is that portion of the microwave circuit, near the T-junction of the receiving branch and the magnetron-antenna line, where the TR and ATR switches are located.

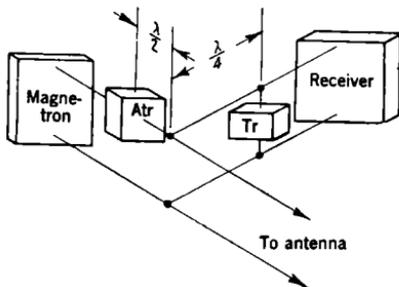


FIG. 11-14.—Duplexing system on two-wire transmission line.

The great disparity in transmitted and received powers immediately suggests that a spark gap or gas-discharge tube can be connected in the circuit in such a way as to perform the necessary switching operations. These gas-discharge tubes are referred to as TR or ATR tubes. A rudimentary system using a two-wire transmission line is shown in Fig. 11-14. The high-power pulse from the magnetron breaks down the gap in the ATR tube and the power flows out toward the antenna. The gap in the TR tube in the receiving branch likewise breaks down, and if it is designed so that the discharge takes negligible power to maintain, puts a short circuit across the line to the receiver. The delicate input circuits of the receiver are thereby protected. Since the short circuit is a quarter wavelength from the T-junction, the impedance put in parallel with the antenna line at the junction is very high and does not affect the wave traveling toward the antenna. At the end of the transmitted pulse, the discharge across the gaps goes out and the system is ready to receive echo signals. The impedance at the T-junction looking toward the magnetron is infinite because there is an open circuit half a wavelength away.

¹ *Microwave Duplexers*, Vol. 14.

Looking toward the receiver, there is a matched line. All the power goes into the receiver.

In one variation of the basic scheme just outlined, the receiver branch joins the antenna line in a series rather than in a shunt T. Then the TR switch must be a half wavelength rather than a quarter wavelength from the junction. In another variation, advantage is taken of the fact that the cold (i.e., nonoscillating) impedance of certain types of magnetrons is such as to be a bad mismatch to the line, so that nearly all the power coming toward the magnetron is reflected. Then the ATR switch can be omitted if the line length between the magnetron and T-junction is chosen correctly. This is known as "pre-plumbing." With many types of magnetron it is not feasible.

The requirements for satisfactory transmission of the outgoing pulse are rather easily met.

1. The loss in the discharge across the gaps must be a small fraction of the magnetron power.
2. The line must be matched when the gaps are fired.

From the point of view of the receiver, the requirements are much more stringent.

1. During the transmitted pulse, the power getting past the TR switch into the receiver must be less than 0.1 watt or the crystal may be damaged. This means a minimum attenuation of 60 to 70 db.
2. The TR-tube gap must fire in less than 0.01 μ sec, or the preignition "spike" of magnetron energy may burn out the crystal.
3. The gap must deionize in a few microseconds at the end of the magnetron pulse so that echoes from nearby objects will not be unduly attenuated. A typical specification would demand less than 3-db attenuation 6 μ sec after the pulse.
4. The received signal must see a reasonably good match into the receiver, and the losses must be kept to a minimum.

Some refinements in the rudimentary system of Fig. 11-14 are necessary to meet the above requirements. The fired TR-tube gap is not a perfect short circuit. If the voltage across the arc is V , the leakage power going to the receiver is V^2/Z , where Z is the impedance looking toward the receiver, measured at the gap terminals. The voltage can be made smaller by having the discharge take place in a gas at a pressure of only a few millimeters of mercury. Further reduction of leakage power is necessary, however. This may be done by a step-up transformer to the gap, and an identical step-down transformer to the receiver line. In the unfired condition, the standard line impedance is maintained on either side of the TR switch, but in the fired condition the line impedance seen at the gap appears to be very high and much less power is coupled out to the receiver.

The practical method of accomplishing this impedance transformation is by means of a resonant cavity. Figure 11-15 shows a section through a 1B27 TR tube and associated cavity, with input and output couplings. The gap across which the discharge takes place is formed by two reentrant cones on the axis of symmetry of the approximately cylindrical cavity. The cones add capacity to the resonant circuit and the cavity is smaller for the same wavelength than it would be without them. Tuning is accomplished by pushing one cone in and out on a flexible diaphragm. The unloaded Q of the cavity is lower than that of a cylinder because of

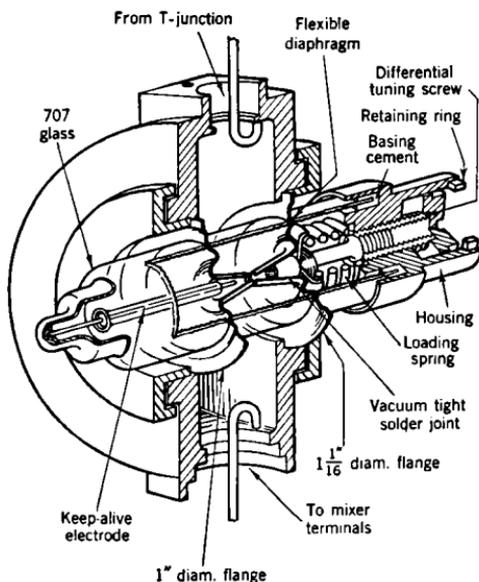


FIG. 11-15.—1B27 TR tube and cavity assembly with loop coupling.

the presence of the cones and the glass of the gas enclosure; it is about 2000. With normal input and output loading, the loaded Q is about 350. Both input and output coaxial lines end in coupling loops which play the role of the step-up and step-down transformers. They can be thought of as single-turn windings which, in proportion to their area, loop more or less of the magnetic field existing in the cavity. The smaller the loop, the higher the step-up ratio and the higher the loaded Q .

As a result of the impedance transformation, the arc coupling (i.e., the power going to the receiver as a result of the voltage across the arc discharge) is well below the danger point and is independent of input power. There is a second mechanism of coupling, called direct coupling, which gives leakage power proportional to input power. At the higher transmitter powers, direct coupling becomes more important as a source of leakage power than arc coupling. Direct-coupling power is the power

that would be coupled from loop to loop with a solid metal post in the center of the cavity, the annular space around it being a waveguide beyond cutoff.

To insure rapid breakdown at the beginning of each pulse, a supply of ions in the gap is maintained by a continuous auxiliary discharge inside one of the cones. This requires an extra electrode, known as the "keep-alive" electrode, which draws about $150 \mu\text{a}$ from an 800-volt supply. A ballast resistor drops the voltage to about 400 volts at the electrode itself. Despite this precaution, the leakage power through a TR tube shows an initial spike (Fig. 11-16) which precedes the "flat" region of constant leakage power.

For rapid deionization of the gap one constituent of the gas must have an electron affinity. After the discharge is over, in the absence

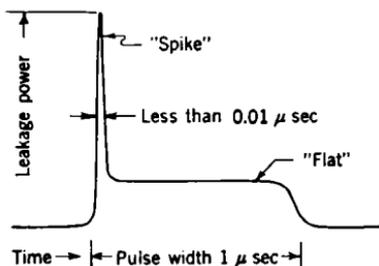


FIG. 11-16.—TR leakage power during a 1- μsec pulse.

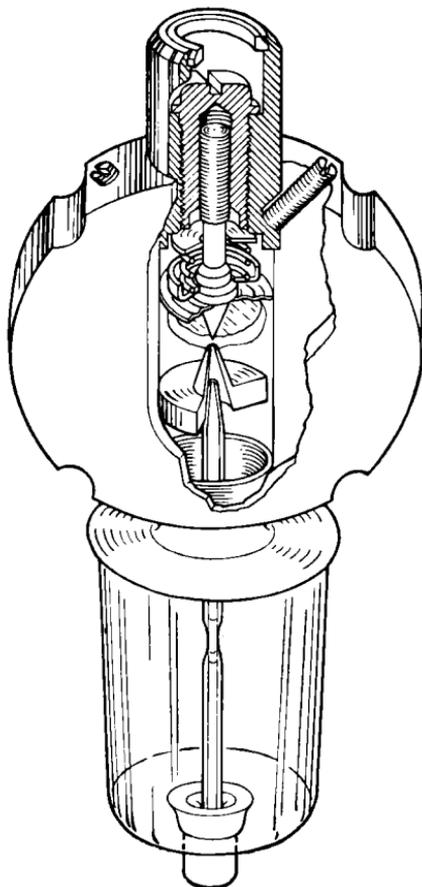


FIG. 11-17.—1B24 TR tube; 3-cm band.

of a strong field, electrons are quickly removed by attachment to molecules. Molecular ions have so much inertia that at the frequencies involved they cause negligible attenuation. Water vapor is the constituent usually introduced to hurry the electron cleanup. A typical filling would be hydrogen at a pressure of 10 mm of mercury, and water vapor likewise at 10-mm pressure. The hydrogen gives protection if the water vapor is frozen out, but deionization will then be slow. Argon is used in some tubes instead of hydrogen.

The loss to the received signal in passing through an unfired TR tube is from 1.0 to 1.5 db, occurring mainly in the walls of the cavity.

Figure 11-17 shows a cut-away view of another type of TR tube, the 1B24, which is widely used in the 3-cm band. It is of the integral-cavity type—that is, all of the cavity contains gas, not just the central portion. A gas reservoir on the side increases the total volume of gas over that contained in the relatively small 3-cm cavity. The tube is clamped between standard 3-cm rectangular choke joints, and coupling into and out of the cavity is by means of round windows or irises. Glass is soldered across them to seal the gas enclosure. Such windows play the same role in coupling cavities to waveguide as do coupling loops for coaxial line. On 10-cm systems using waveguide, the input coupling may be by means of an iris, and the output by a coupling loop.

Switches exactly like TR switches with the output coupling omitted may be used as ATR switches. However, since the double adjustment of the two switches is not always made correctly, fixed-tuned, low- Q ATR switches are preferable. Figure 11-18 shows the 1B35 tube, designed for the 3-cm band, and its mounting. One such tube will cover a frequency band of 3 per cent, and pairs of these tubes will cover a band of 6 per cent. The loaded Q of the cavity is made very low (approximately 5) by a large window in the end. The cavity is placed in series with the

magnetron line by simply substituting the face of the tube containing the window for a portion of the broad side of the waveguide. The breakdown takes place across the inner face of the elongated glass window.

A duplexer that represents in waveguide the basic scheme of Fig. 11-14 appears in Fig. 11-29 as part of an r-f transmitting and receiving system. The TR-tube input window is in contact with the narrow side of the waveguide. From the resemblance of a waveguide to a stub-supported two-wire transmission line, illustrated in Fig. 11-8, it is apparent that a T-junction on the narrow side of the guide is a parallel connection, and that the plane of the narrow side is a quarter wavelength from the effective center of the waveguide circuit.

MICROWAVE COMPONENTS OF THE RECEIVER

After emerging from the TR switch, the received signal is mixed with the local-oscillator signal and the combination applied to the crystal, in

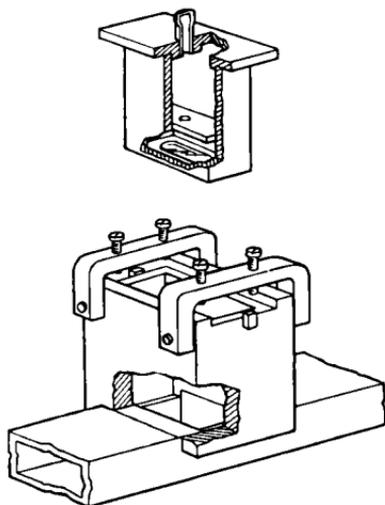


FIG. 11-18.—1B35 broadband ATR tube and mount; 3-cm band.

order to obtain a much lower beat frequency which can be amplified. A more complete and integrated discussion of microwave receivers is reserved for the next chapter, but for the sake of continuity the components of the receiver which handle microwaves are treated here in conjunction with other r-f components.

11.6. The Mixer Crystal.—Since no satisfactory amplifier for microwaves exists, the conversion to intermediate frequency must be made at

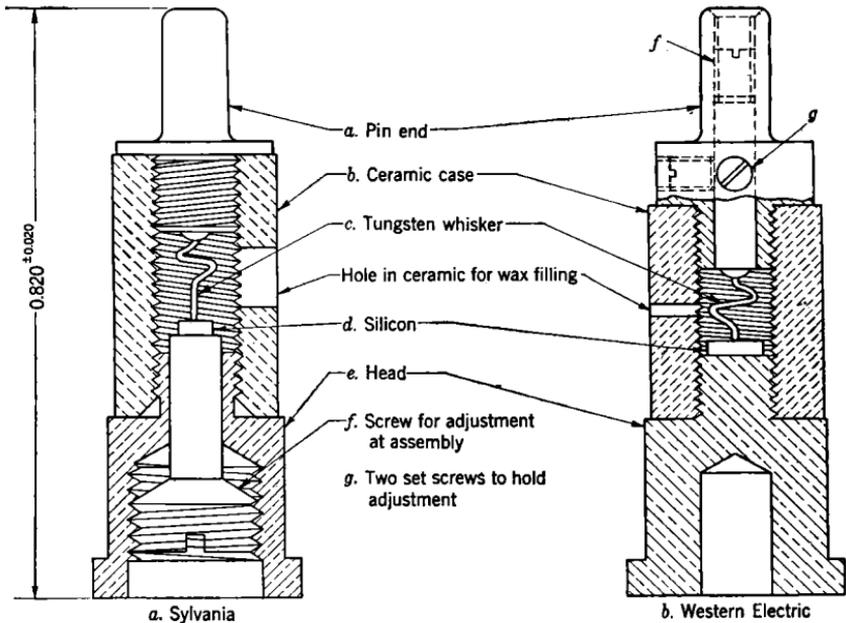


FIG. 11-19.—Microwave mixer crystals.

the low level of received signal powers. It is important, therefore, that the nonlinear element be as efficient as possible, and that a minimum amount of added noise be introduced. The most satisfactory device that has been found is the rectifying contact between a metallic point and a crystal of silicon.¹ For protection and stability the silicon and the "cat whisker" are sealed up in a cartridge. The term "crystal" ordinarily refers to the whole assembly, a cross-section view of which is shown in Fig. 11-19.

The d-c characteristic of a crystal has the form shown in Fig. 11-20. An equivalent circuit that accounts for the r-f properties is shown in Fig. 11-21. The nonlinear resistance of the rectifying contact is denoted by r_c . In parallel with it is the capacity C of the boundary layer of the

¹ *Crystal Rectifiers*, Vol. 15. Radiation Laboratory Series.

semiconductor. At sufficiently high frequencies, the capacity will short-circuit the high back-resistance of r_c and reduce the rectification efficiency. In series with this combination is R , the so-called "spreading resistance," representing the bulk resistance of the silicon. Analysis shows that RC must be small compared with the time of one r-f cycle for efficient rectification. The effects of R and C can be minimized by small contact area, and it is possible to make crystals with nearly as good conversion

efficiency as would be obtained from a simple diode at much lower frequencies. The conversion loss, defined as ratio of r-f signal power to i-f signal power, runs from 5 to 8 db for typical microwave crystals.

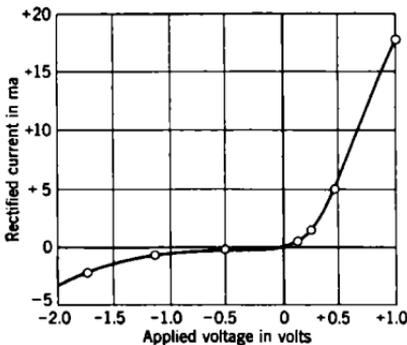


FIG. 11-20.—Typical characteristic curve of a silicon rectifier.

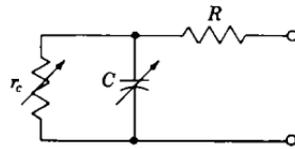


FIG. 11-21.—Equivalent circuit of a crystal rectifier.

Experimentally it is found that in the presence of local-oscillator power flowing through it, a crystal generates more noise power than would an equivalent resistor at that temperature. As a measure of this property, a crystal is assigned a "noise temperature," defined as the ratio of the absolute temperature at which an equivalent resistor would generate the observed noise to the actual temperature. Noise temperatures between 1.1 and 3.0 are typical. The noise generated by the crystal increases with local-oscillator input. There is a rather broad region of best over-all performance at 0.5 to 1.0 mw input which is a compromise between increasing noise at high inputs and greater conversion loss at low inputs. This optimum input corresponds to the widely used standard operating point of 0.5-ma d-c crystal current.

The contact area between the whisker and the silicon is of the order of 10^{-6} cm². Relatively low currents, therefore, yield high current densities, with attendant local heating and danger of burnout. For continuously applied power the danger line is of the order of a watt; this would apply to the flat part of the TR-tube leakage. The initial pre-ignition spike (Fig. 11-16) is so short (less than 0.01 μ sec) that the heat cannot be conducted away from the contact; consequently the total spike energy rather than the peak power is the important quantity. Experience has shown that burnout is far more likely to come from the spike than from the flat. Since TR-tube conditions are difficult to reproduce,

crystals are tested by sending through them a d-c pulse of duration 2.5×10^{-9} sec and total energy of the order of an erg. Crystals which pass this test are safe in a properly operating radar set.

Table 11.2 lists for comparison the rejection-limit specifications of the most widely used crystals in the three radar bands. The bearing of these figures on the over-all sensitivity limit of microwave receivers is discussed in Chap. 12.

In the manufacture of crystals to meet the specifications (which are a considerable advance over those met by the crystals used in early microwave radars) careful attention must be given to the original purity of the silicon, to "doping" with small amounts of the proper impurity, to the preparation of the crystal face, and to the sharpening and adjustment of the whisker. Crystals are stored in metal containers to avoid burnout from stray r-f fields. They can be burned out easily at the time of installation by an accidental discharge of static electricity through them.

TABLE 11.2.—SPECIFICATIONS OF CONVERTER-TYPE CRYSTALS

Type	Wavelength band (cm)	Conversion loss (db)	Noise temperature (times)	Burnout test (ergs)
1N21B	8-11	6.5	2.0	2.0
1N23B	3.1-3.5	6.5	2.7	0.3
1N26	1.25	8.5	2.5	0.1

11.7. The Local Oscillator.—Reflex klystrons are, with very few exceptions, used as local oscillators¹ in microwave receivers. Figure 11.22 shows a schematic view of a typical tube. The resonant circuit is a reentrant, doughnut-like cavity with grids across the central portion. An electron gun with a proper focusing electrode directs a stream of electrons through the grids. Upon arrival at the first grid the electrons have a velocity corresponding to 300 volts. It will be assumed that oscillations exist in the cavity. Electrons will then alternately be accelerated and decelerated by the r-f field across the grids. An electron that goes through just as the field between the grids is passing through zero on the way from acceleration to deceleration will not have its velocity changed and can be taken as a reference electron. In the space

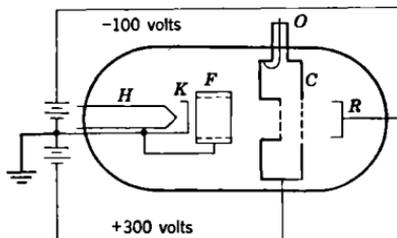


FIG. 11.22.—Schematic of reflex klystron.

- H = heater
- K = cathode
- F = focusing electrode
- C = cavity
- R = reflector
- O = coaxial output line.

¹ *Klystrons and Microwave Triodes*, Vol. 7, Radiation Laboratory Series.

just beyond the grids there is a strong retarding field produced by a reflector electrode maintained about 100 volts negative with respect to the cathode. The trajectory of the reference electron in this space is similar to that of a ball thrown into the air. It will return to the grids after a time proportional to its initial velocity and inversely proportional to the retarding field. An electron that leaves the grids earlier than the reference electron will have been accelerated by the r-f voltage across the cavity and, because of its higher velocity, will spend a greater time in the reflection space. By proper adjustment of the retarding field, the delay may be made to compensate for its earlier departure, and it may be made to arrive back at the grids at the same time as the reference electron. Similarly, an electron leaving later than the reference electron catches up by spending less time in the retarding field as a result of its lowered velocity. The net effect is that the electrons gather in a bunch.¹ At certain reflector voltages the bunch will pass through the cavity grids in such a phase that the r-f field retards the electrons. The electrons then give energy to the cavity and thereby sustain the oscillations. Oscillation is observed for more than one reflector voltage because drift times differing by a whole r-f cycle still produce satisfactory bunching.

The net energy given to the electrons during their first passage through the cavity is negligible when averaged over a whole cycle, being balanced between acceleration and deceleration. On the return passage, however, most of the electrons go through in a bunch at the most favorable phase to aid the oscillation. Half a cycle later, when returning electrons would absorb energy in being accelerated, very few electrons are passing through. Useful power is delivered to an external load through a coaxial line, loop-coupled to the cavity. The efficiency is rather low, in the neighborhood of 1 per cent. Power outputs of 20 to 50 mw are typical.

Local oscillators are tuned by mechanically changing the size of the cavity. A limited amount of electrical tuning is possible through variations of the reflector voltage. Figure 11-23 shows the frequency and power output of a 10-cm reflex klystron as a function of reflector voltage. When the reflector is made more negative, the bunch arrives at the cavity

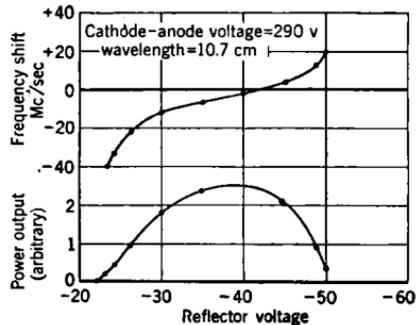


FIG. 11-23.—Reflector characteristics of the 707A.

¹ In the older two-cavity klystrons, where the bunching takes place in a field-free space, the bunch forms about the electron that passes through the first cavity as the field is changing from deceleration to acceleration.

early, hurries the oscillation and increases the frequency. Reflector-voltage tuning ranges of 30 to 50 Mc/sec between the half-power points are normal, and the tuning rate may lie between 1 and 4 Mc/sec per volt. Electrical tuning is the basis of automatic-frequency-control systems.

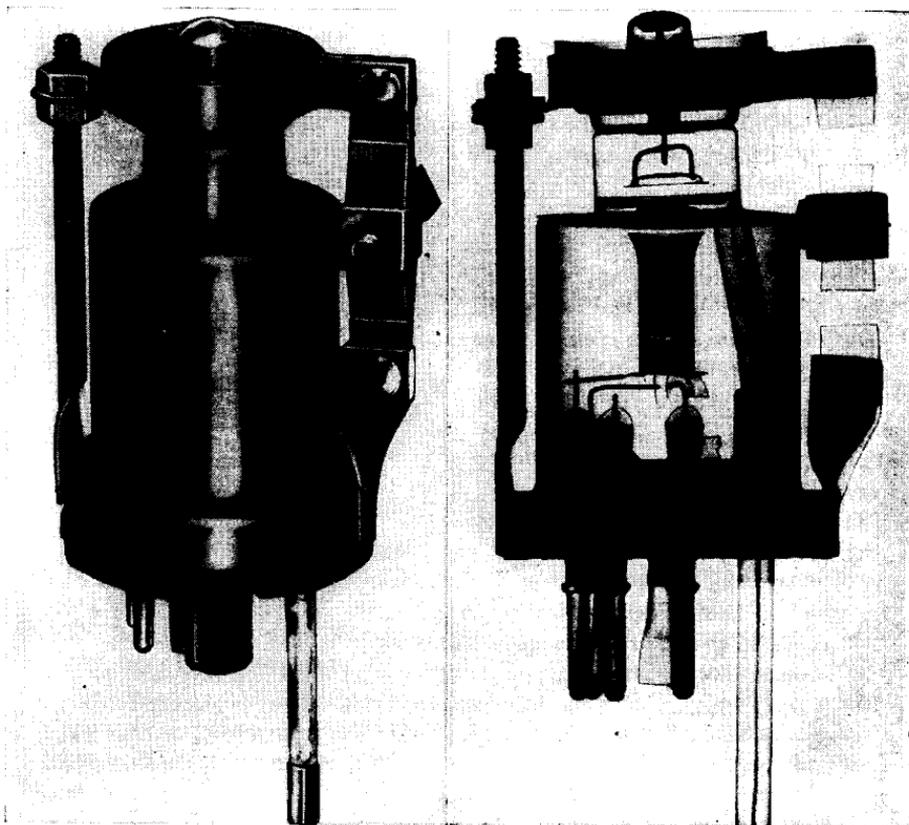


FIG. 11-24.—2K25 reflex klystron; 3-cm band: (a) exterior view; (b) x-ray view.

Figure 11-24 shows an exterior and an x-ray view of a 2K25, a widely used local oscillator for the 3-cm band. The coaxial output line ends in a probe which is inserted in the waveguide to form a matched transition like that shown in Fig. 11-12a.

11-8. The Mixer.—The mixer¹ contains the crystal. It has two sets of input terminals, one for the received signal and one for the local oscillator; the output signal goes to the first stage of the i-f amplifier. The requirements to be met are as follows: (1) the crystal must be made to appear as a matched load to the incoming signal; (2) there must be

¹ *Microwave Mixers*, Vol. 16, Radiation Laboratory Series.

minimum loss of incoming signal into the local-oscillator input; (3) the local oscillator must see a fairly good match, though the requirements are not as strict as for a magnetron.

Figure 11-25 shows a coaxial-type mixer which is widely used for the 9- to 11-cm band. It bolts directly on the TR-tube cavity and the output

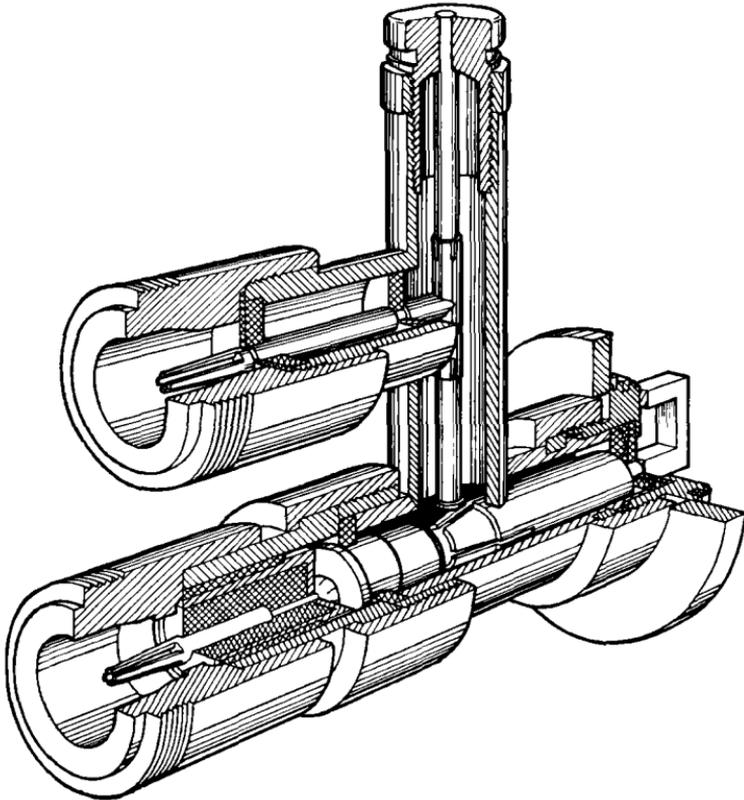


FIG. 11-25.—Coaxial-type mixer; 10-cm band.

coupling loop is integral with the mixer. The crystal is mounted as a part of the center conductor of a short coaxial line and makes a matched termination with no special transformers. Since the i-f signal must be extracted, the end of the r-f line cannot be a d-c short circuit. An effective r-f short circuit, with no metallic contact between inner and outer conductor, is provided by two concentric quarter-wave sections which form a choke not unlike that in a rotary joint. The i-f output fitting unscrews to permit replacement of crystals.

Local-oscillator power is introduced through a side arm. About 50 mw is available from reflex klystrons in the 10-cm region, which is 20 db

above the 0.5 mw needed to produce the standard 0.5 ma of crystal current. The great excess of power permits loose capacitive coupling between local oscillator and crystal. By the reciprocity principle, since there is poor transfer of power from the local-oscillator input to the signal line, there will be poor transfer from the signal line to the local-oscillator branch, thus satisfying the second condition stated above.¹ A screw on a quarter-wave stub support adjusts the proximity of the capacitive probe to the signal line and provides the method of setting the crystal current. This is the sole adjustment.

The local-oscillator input fitting contains a resistor disk which terminates the 50-ohm cable. Since the probe is half a wavelength away from the disk, the load at the probe will be in parallel with the disk (Sec. 11-1); but since the probe does not differ much from an open circuit, the load for the local oscillator is still a good match. The voltage at the probe is the same as that at the disk. It is necessary to have the probe spaced a quarter wavelength on the signal line from the effective plane of the coupling loop. With a loaded Q of 350 the TR tube is considerably off-tune at the local-oscillator frequency, which is usually 30 Mc/sec away from the signal frequency. Like all parallel resonant circuits, the TR cavity presents a very low impedance, practically a short circuit, at frequencies off resonance. If the probe were half a wavelength from the coupling loop, it would be effectively at a short-circuit point for local-oscillator frequency and no power could be transferred. At a quarter-wave point the short circuit at the coupling loop is reflected as an open circuit and the piece of line is, to the local-oscillator frequency, a stub support.

In mixers for the 3-cm region, the crystal goes directly across the center of the guide. Coupling between local-oscillator waveguide and crystal guide is by means of adjustable windows in the short face of the guide. At 10 cm coaxial cables carry the local-oscillator power to separate radar and AFC mixers, but at 3 cm double mixers are used in which the two crystals are on opposite sides of the local oscillator. Figure 11-29 shows a waveguide mixer and other r-f components for the 3-cm band with AFC and beacon features.

11-9. Automatic Frequency Control.—Radar automatic frequency control (AFC) is a scheme in which the *difference in frequency* between the magnetron and the local oscillator is compared in a discriminator cir-

¹ In an exact calculation, account must be taken of the difference in the impedance of the various branches at signal and local-oscillator frequencies. As is mentioned in the next paragraph, this is considerable for the highly resonant TR cavity. Qualitatively, however, the reciprocity argument is valid. The so-called "magic T" provides a method of decoupling the signal and local-oscillator inputs that does not depend on an excess of local-oscillator power. See *Microwave Mixers*, Vol. 16, Chap. 6.

cuit with the standard intermediate frequency being used—for example, 30 Mc/sec. An error in the difference frequency produces a voltage that is applied to the reflector of the local oscillator in such a sense as to bring the local oscillator into correct tune. Because of the close connection of the whole AFC problem with receiver design, consideration of this subject is deferred to Sec. 12-7.

MOUNTING THE R-F PARTS

Experience has demonstrated that good radar performance depends not only on a well-designed set of microwave components but also on a properly coordinated mounting of them, relative to each other and relative to the other parts of the radar. Therefore it is necessary to add to the treatment already given of the components themselves a consideration of the engineering reasons for the accepted practice of grouping certain vital components in a major unit called the "r-f head."

11-10. Reasons for an R-f Package. *Preferred Grouping of Components.*—Considerations of accessibility and convenience would often argue for putting most of the components of a radar set near the indicator rather than near the antenna, which, in order to get the proper view, must usually be at a remote or isolated point. However, the *long-line effect* discussed in Sec. 11-1 makes a long transmission line from magnetron to antenna something to be avoided if possible. In addition to the fact that in a long line even a moderately low standing-wave ratio can cause instability in the magnetron, the extra junctions, bends, and elbows needed for a long line are themselves likely to add to the mismatch. Furthermore, as seen from Table 11-1, the r-f attenuation in a long line, particularly at the higher microwave frequencies, results in the loss of several decibels when two-way transmission is considered.

The shortest possible r-f line is realized by mounting the magnetron on the back of the antenna reflector ("back-of-the-dish system"). This ideal arrangement is often not practical because of a group of components functionally associated with the magnetron which may be too large and heavy for the antenna mount to carry. The pulse transformer must be adjacent to the magnetron to avoid long high-voltage leads, which would have to be supported and properly insulated and would add unwanted capacity. The duplexer T-junction and the TR tube must, of course, go in the magnetron-antenna line. The extremely low-level echo signals coming through the TR tube must be converted to the intermediate frequency and amplified considerably before being transmitted to a remote point. The local oscillator must be close, because the AFC is based on a comparison of the frequency of the local oscillator and that of the magnetron.

Electrical and functional considerations dictate, then, that the follow-

ing minimum items be mounted in a group and that the group be as near the antenna as practical:

1. Pulse transformer.
2. Magnetron.
3. Duplexer, TR and ATR tubes.
4. Local oscillator.
5. Radar and AFC mixer.
6. AFC control circuits.
7. I-f amplifier (usually up to 1-volt video level).

This set of components, mounted as a group and usually in a closed container, is called the "R-f Head." Alternative terms are "R-f Unit" and "Transmit-receive Unit." Beacon local oscillator and beacon AFC are included where beacon reception is required. In smaller radar sets, the modulator network and switch tube and even the modulator power supply may be put with the r-f unit. An extension of this trend leads to an arrangement with all parts of the radar except the controls and the indicator tube in a single container.

Advantages of an R-f Package.—On all but the largest radar sets the group of components listed in the previous paragraph can be compactly mounted in a single container that is not too heavy for one man, or at most two men, to carry. There is considerable advantage in having the vital parts of the transmitter and receiver centralized in a demountable unit package. In the event of trouble, a spare r-f head can be substituted and connected into the rest of the radar in a few minutes. Diagnosis of trouble, repair, and readjustment, as well as assessment of performance, can all be performed on a well-equipped test bench. This is particularly helpful for airborne sets.

Good engineering practice favors pressurization of the r-f transmission lines of microwave radar systems in order to keep out water and water vapor. For high-altitude operation, pressurization may be absolutely necessary to prevent breakdown of the line. The high-voltage cathode circuit of the magnetron is a similar case. Pressurization assures dry conditions and sea-level pressure, and thus makes unnecessary the huge insulators and large air clearances that standard engineering practice would otherwise prescribe. Other parts of the r-f head can be made smaller and more compact if pressurized conditions are assumed. It is but a slight extension to consider that the r-f head should go into a single pressurized container, with the pressure common to the r-f line out to the antenna. Such a plan solves two other problems: (1) the working parts of the r-f head are protected against the most severe conditions of moisture, salt spray, and dust, not only in use, but also during shipment and storage; and (2) the shielding problem is greatly simplified, both for keeping external disturbances out of the delicate parts of the receiver, and

for confining the disturbances set up by the radar transmitter so that they do not affect other equipment.

11-11. Design Considerations for the R-f Head.—The form which the r-f head takes will depend on whether it operates on the ground, on a ship, or in an airplane; on the degree of exposure to the elements; and on the power and transmitting frequency of the magnetron. Pressurization has advantages and disadvantages. Among the latter are inconvenience resulting from inaccessibility of parts for adjustment and repair, and difficulty of transferring heat through the pressure wall. The r-f head of a small airborne set will undoubtedly be a completely pressurized unit. The r-f head of a very large ground set might operate inside a protective housing and not be pressurized at all. Between these two extreme types are designs where part is pressurized and part is not. No firm rules can be laid down. The treatment here will outline the conditions that must be met, and then give illustrative examples of two quite different designs.

*Heat Removal.*¹—The maximum safe ambient temperature for most of the r-f-head components—such as composition resistors, oil-paper condensers, and blower motors—is about 85°C. The temperature of the air around the r-f head may get as high as 50°C in desert areas inside a housing exposed to the sun's rays. If the unit is not pressurized the permitted differential of 35°C is easily met. The air from a simple blower, properly channeled, will readily carry the heat released inside the enclosure out into the surrounding air. Often the air from the magnetron cooling blower can be so directed as to do what additional cooling is needed. Subunits tightly enclosed for reasons of electrical shielding may, however, need additional circulation.

Where the free flow of external air through the r-f head is blocked off, as it must be by a pressure housing, the problem of transferring the heat to the outside air can easily be a limiting factor in design. The difficulty is not in getting the heat through the actual metal wall. A short computation shows that a differential of a fraction of a degree is sufficient for this. Almost all of the temperature drop occurs across the dead-air films on the two sides of the wall. Natural convection results in a transfer coefficient of only 0.006 to 0.010 watts/in² per °C difference in temperature between the air and the wall. If 30°C be taken as a safe figure for rise of internal air over external air, and if natural convection be assumed on both sides, then the average coefficient given above would result in a maximum thermal load of 0.12 watts/in². Forced convection from a gentle current of air along the surface raises the coefficient to 0.02 watts/in² per °C, but beyond this increased velocity results in only a slow increase. A high-velocity flow of perhaps 50 ft/sec is necessary to achieve a figure of 0.04 watts/in² per °C.

¹ *Amer. Soc. of Heating and Ventilating Engineers Handbook.*

The increase in the coefficient is due mainly to creating a turbulent layer where good mixing occurs close to the actual wall, and to reducing the thickness of the poorly conducting layer. The cooling blast for a magnetron approximates the necessary conditions, but maintenance of turbulent flow over large areas takes considerable power.

The conclusion is that with moderate blowing on the inside and natural convection outside, the loading can increase to 0.17 watts/in², and with blowing on both sides can increase to 0.3 watts/in², for an over-

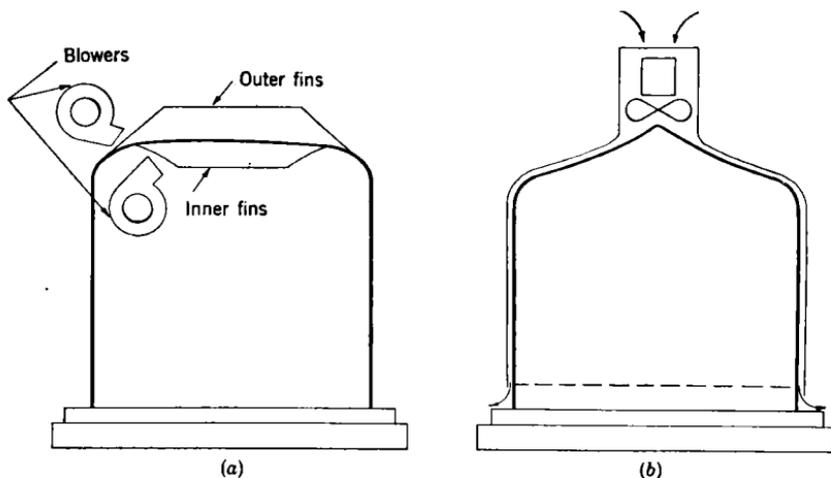


FIG. 11-26.—Methods of improving heat transfer through pressurized containers.

all rise of 30°C in air temperature. These figures are all approximate, since the geometry is never simple and uniform conditions of air flow over a whole vessel are never realized.

Radiation plays a minor role in heat transfer for the differentials here involved. Painted rather than bright surfaces are worth while, however. Fins are helpful in increasing effective area, especially with natural convection, though not in proportion to actual surface. If the forced convection is not along the fins they do not help at all, because the increased area is compensated by impediment to air flow.

Two schemes to improve heat transfer are shown in Fig. 11-26. In *a*, the internal and external blowers are directed at the same portion of the wall, which has some fins at that point at least to increase effective area. In the region of turbulent flow directly under the blower output the transfer coefficient is high. In the second case, *b*, there is an outer shell close to the wall of the pressure container. A powerful blower maintains turbulent flow over the whole interspace and a high coefficient results. This may be repeated on the inner wall of the pressure housing.

Types of Container.—If the r-f head is not pressurized, the supporting framework for the parts will be a rectangular affair consonant with the standard construction practices used on the rest of the radar. The usual rules of accessibility and convenience will dictate the placement of parts. If, however, the unit is to be pressurized, further consideration must be given to the container. It is almost essential to be able to remove the pressure cover and make tests and adjustments without disconnecting the r-f transmission line or any power or signal cable. This rules out methods of mounting the units in which the support brackets attach to the pressure cover.

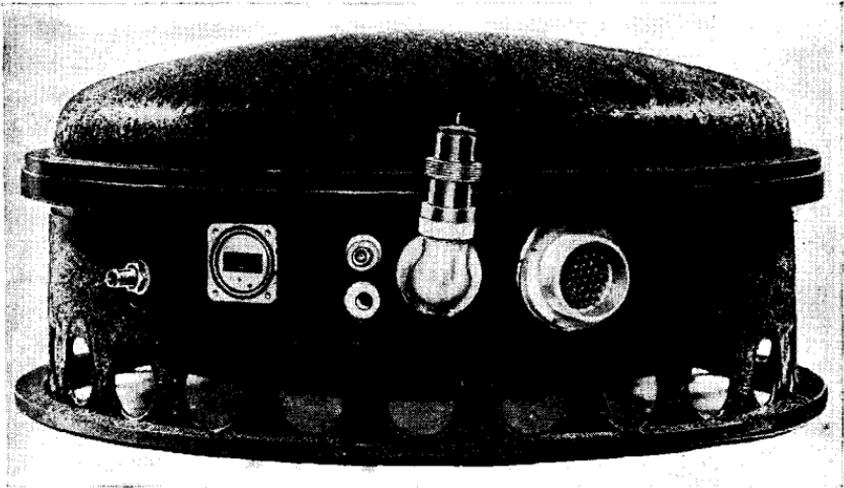


FIG. 11-27.—Exterior view of a 3-cm airborne r-f head.

Although rectangular containers are most economical of space, they are difficult to pressurize because of their tendency to become spherical under pressure, with resultant severe shear strains in the corners. In a ship set, where only small pressure differentials are expected, a rectangular cover is feasible. If it is not too large, a rectangular housing cast out of light metal can be made strong enough to stand the 30 lb/in² test pressure required for airborne r-f heads. It has been common practice, however, to use cylindrical housings with domed ends because only tensile strains are then involved and the housing can be very light.

One example of pressurized airborne container design is illustrated in Fig. 11-27, and discussed fully in Sec. 11-12.

The heat generated in the magnetron rises as the average power level is increased, but the generation of heat in the receiver components is substantially constant from one radar set to another. If the transmitter power is more than about 50 kw, the thermal load for any pressurized

container of reasonable size would be above the limit set in Sec. 11-11. Under these circumstances, it is customary to mount the magnetron anode and the pulse transformer case outside the pressurized shell, with an external blower to cool them. The cathode connection to the magnetron is made in a pressurized inner compartment, into which gasketed bushings bring the pulse transformer output and the magnetron cathode leads. The waveguide r-f line is pressurized from the magnetron output to the pressurized compartment which houses the duplexer, as well as on the antenna side. In this arrangement, high voltages appear only inside pressurized spaces, while most of the heat is liberated in the open, where it can be carried away by the air stream from the external blower.

Metering and Test Points.—Good radar performance depends on objective tests of the functioning of each part. The vital radar components are all centralized in the r-f head, and the providing of necessary test points there is a part of the design job. Testing should not require removal of any covers. A directional coupler, used to introduce the test signal, will be located on the antenna line either just outside the r-f head or just inside with cable feed to a test jack on the outside wall. At the higher microwave frequencies, cables and cable fittings introduce a variable element and the external location would be preferred. The video output jack required can be the normal output to the indicator, in which case an interchange of cable is required. If the video output tube is capable of handling parallel loads, a second jack for test purposes only would be added. If the test set requires a trigger signal at the time of the magnetron pulse, this can be obtained by inserting a few ohms in the bypass-to-ground circuit at the low end of the pulse transformer. The signal would go in a shielded cable to an appropriate external jack.

The other checks on the operation of the r-f head divide themselves into (1) minimum information continuously available to a remote operator and (2) tests to be performed with the cover removed.

The average magnetron current is the most important single item in monitoring a radar set. If it has the correct value, the magnetron is operating at the recommended point on its performance chart and is almost certainly putting r-f power into the antenna line. Sparking of the magnetron or irregular modulator operation can be observed. This meter reading is an over-all check on the operation of the radar transmitter.

The second most important item for the operator to monitor is the radar crystal current. If this maintains its original value, it is known that the local oscillator is functioning and the crystal has not changed its properties. A decrease may indicate a burned-out or damaged crystal. A rhythmical variation indicates that the AFC is searching but not locking.

A single remote meter with a selector switch, or selector relays in the r-f head, can perform these minimum monitoring functions at the operator's position. If the set has beacon AFC, the radar-beacon switching relay should transfer the remote meter to the beacon cavity crystal. Then a steady meter reading indicates not only that the beacon AFC is locked, but that it is locked at the correct frequency, since the crystal energy has come through the reference cavity.

Of the test points usually provided for cover-off checking, the most essential is a jack to measure the crystal current at every crystal, in order to permit adjustment of the local-oscillator coupling. On crystals where the same current is metered remotely as well, a cut-in jack is used. Another current worth metering is the average diode current in the second detector, for over-all sensitivity tests and checking the receiver. A jack for measuring the keep-alive current in the TR tube should be included to provide verification that the TR switch is performing normally.

In radar sets where the r-f head is accessible to an operator during operation, all of the above metering can be done on a single meter and selector switch.

The most important internal signal test point is the AFC video output. On an oscilloscope the discriminator output is seen, as well as any spurious signals. Sometimes this is made a phone jack also, so that with a pair of headphones aural indication of crossover may be compared with the frequency of best tune.

11-12. Illustrative Examples of R-f Heads. *Airborne Search Set.* Figure 11-27 shows an external view of the r-f head used on a 3-cm search and navigation set of recent design. A broad flat container was dictated by the space available in certain aircraft. A lightweight sheet-metal pressure vessel of this form factor must of necessity have a domed top and bottom. The pressure seal comes near the top of a short central cylindrical section 18 in. in diameter.

The radar set is designed to work on three pulse lengths: (1) 5 μ sec, 200 pps, for search; (2) 2.5 μ sec, 400 pps, for beacon interrogation; and (3) 0.5 μ sec, 800 pps for high definition. The long pulse, as shown in Sec. 2-9, gives greater range, particularly in mapping land-water boundaries. It must be longer than the upper discrimination limit of ground beacons. The beacon pulse is well over the minimum length of 2 μ sec necessary to trigger beacons.

The pulse is transmitted from the modulator on a 50-ohm cable, which enters the r-f head via a pressurized pulse connector fitting. The shielded cable then goes to the quadrant-shaped pulse compartment. Into this compartment project the pulse transformer bushings and the magnetron cathode bushing. The interior detail is shown in Fig. 11-28. Ordinarily the quadrant cover is on and no high voltages are exposed

when the lid of the whole r-f head is off. The pulse compartment also contains the magnetron cathode-heating transformer and the necessary bypass condensers to provide a pulse path to ground in the metering circuit. All wires that leave this well-shielded compartment go through

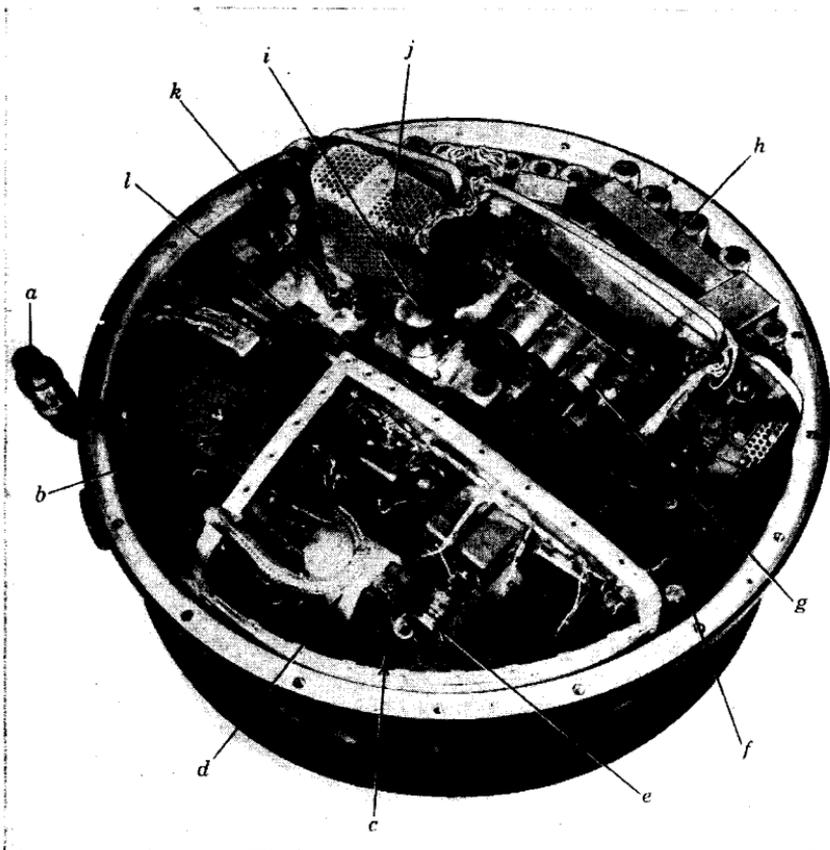


FIG. 11-28.—Interior view of a 3-cm airborne r-f head: (a) pulse input; (b) pulse transformer; (c) pulse compartment; (d) pulse transformer bushing; (e) magnetron cathode bushing; (f) blower motor; (g) AFC chassis; (h) receiver chassis; (i) beacon reference cavity; (j) shield for local oscillators; (k) double mixer; (l) TR tube.

filters which reduce pulse voltages on them to a level that will not interfere with communications. The pulse transformer case, being the source of some heat, projects outside the airtight compartment.

The r-f system of this unit is shown in Fig. 11-29. The magnetron, type 4J52, is of the "packaged" design; the magnet is an integral part of the tube. It operates at an input level of about 200 kw, with an average efficiency of about 30 per cent. Its cathode is larger than is usual for

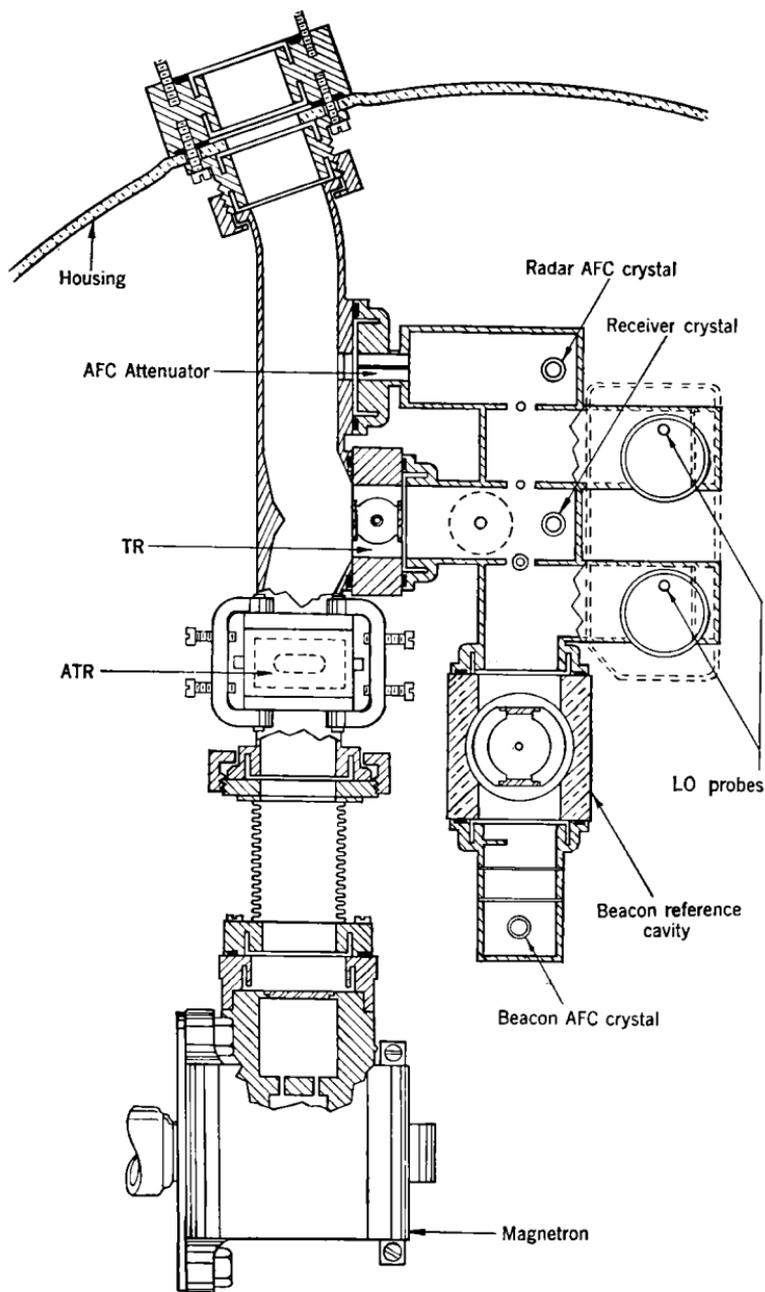


FIG. 11-29.—R-f system.

magnetrons of this power; thus the tube is better able to furnish the long 5- μ sec pulse without sparking. Its cooling blower also provides the general air circulation inside the r-f head.

The waveguide fittings inside the r-f head are not airtight; they communicate the head pressure to the antenna waveguide run. However, electrically the waveguide is kept sealed as tightly as possible in order to prevent the general level of leakage power in the closed container being high enough to interfere with AFC operation. The sealing is done by woven metal gaskets at every junction. A short section of flexible corrugated waveguide next to the magnetron absorbs the dimensional tolerances and permits all parts to be bolted together tightly without strain.

The duplexer-mixer portion of the r-f system follows closely the block diagram of Fig. 12-12. The TR tube is a 1B24. Its r-f selectivity is high enough (about 50 Mc/sec between the half-power points) so that when tuned up for radar reception the loss in receiving beacon signals may be as much as 20 db. This loss is averted by a solenoid-actuated plunger inserted in the waveguide on the low-power side of the TR tube at a distance of half a guide-wavelength from its output window. By adjusting the depth of insertion, the TR tube can be pulled into good tune for the new frequency.

The outputs of the radar and AFC crystals go in double-shielded cables to the receiver and AFC chassis respectively. The AFC control circuit is essentially of the form described in Sec. 12-7. For best use of the available space, the receiver chassis is in the form of a segment of a circle. The normal chain of i-f amplifier stages, second detector, and video output stage, is arranged around the periphery. The bandwidth of the i-f amplifier is 5 Mc/sec, but for long-pulse operation a relay switches the bandwidth to 1.0 Mc/sec by changing the loading of one of the inter-stage tuned circuits. This is necessary to realize the full gain in sensitivity to be expected from the long pulse. The bandwidth is still 5 times the reciprocal pulse length, rather than the value of 1 to 2 times established as optimum in Sec. 2-9. There are two reasons for this: (1) frequency modulation of the magnetron owing to variations of current during the long pulse can lead to a spectrum wider than theoretical; and (2) it is difficult to make a practical AFC that will hold a set in tune to a small fraction of a megacycle per second.

No d-c power supplies are provided in the r-f head, except for the TR-tube keep-alive supply. This is a small half-wave supply with resistance-capacity filter, located under the receiver chassis. All other voltages for the local oscillators, receiver, and AFC come in over wires in the large cable connector from an external centralized supply.

The various subunits in the r-f head are designed with plug connec-

tions so that they may be quickly removed for servicing or replacement with spare units. Remote metering of magnetron current, radar crystal current, and beacon cavity crystal current is provided.

Shipborne Air-search Set.—The final example is the r-f unit of a high-power shipborne set for maintaining air surveillance. The r-f head

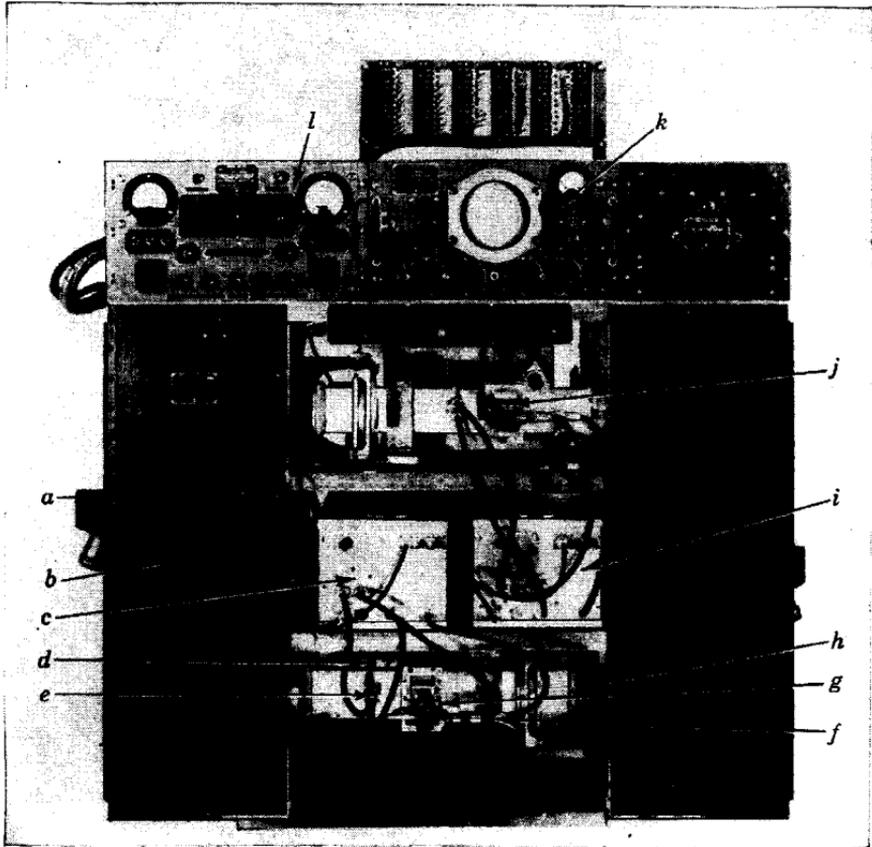


FIG. 11-30.—Front view of r-f head, shipborne radar: (a) pulse input; (b) transmitter compartment, search radar; (c) receiver, search radar; (d) ATR mount; (e) AFC mixer; (f) radar mixer; (g) TR-tube mount; (h) r-f switch for noise source; (i) receiver, height-finder radar; (j) duplexer, height-finder radar; (k, l) built-in control and test equipment.

illustrates many points of good design applicable to high-power sets in the 10-cm region whether shipborne or not. Figure 11-30 shows a front view of the r-f head.

This equipment is a dual set with separate radars on different frequencies, one for long-range search and one for height-finding. The two antennas are mounted on a turret about 5 ft in diameter which contains the r-f head and has room for an operator during tests. The

arrangement is therefore a back-of-dish mounting with short waveguide runs to the antenna, as was recommended in Sec. 11-10. The pulses from the two modulators are transmitted through rotary joints. All

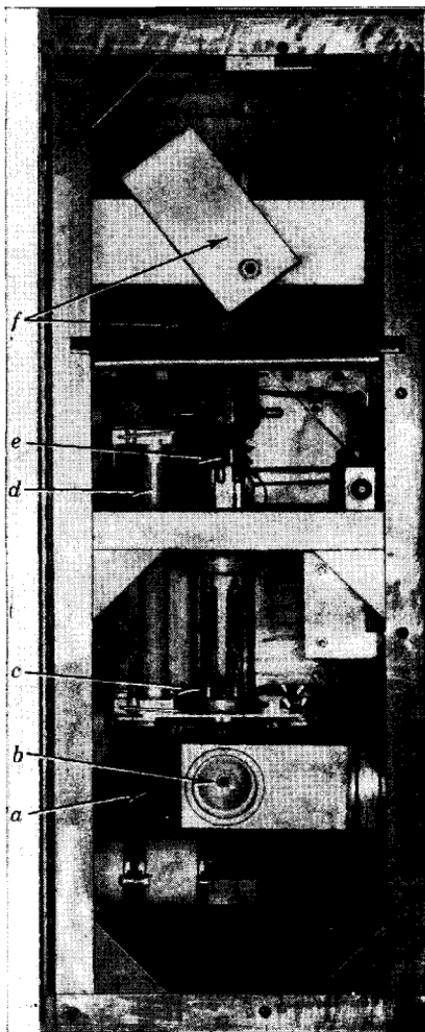


FIG. 11-31.—Transmitter compartment, shipborne radar: (a) magnet; (b) "doorknob" transition; (c) magnetron cathode bushing; (d) despiking resistor; (e) pulse-transformer output bushing; (f) pulse transformer case, directional coupler of opposite transmitter in front.

other power, signal, and control connections to the rotating turret are made via slip rings.

Each of the end compartments contains the transmitter components

of one of the radars. Figure 11-31 shows the interior detail of one of these compartments; in the other, the items are the same but are mounted the other end up. The pulse from the modulator comes in through the pulse connector on the end of the case. For the long-range search set the power from the modulator is about 2 Mw, or about 10 kv on the 50-ohm cable. This power goes in parallel to the primary of the pulse transformer and to a despiking network consisting of a 50-ohm resistor and a 400- μf condenser in series to ground. The network provides a transient load for the modulator while the voltage is rising but before the magnetron starts to draw current.

On the long-range search set, a 4J32 magnetron furnishes 0.8 Mw of r-f power. The pulse length is 1 μsec and the repetition rate is 390 pps. Since the peak voltage at the cathode is about 28 kv, good clearance must be maintained. The magnetron output is in 1 $\frac{5}{8}$ -in. OD coaxial line, but an immediate transition is made to standard waveguide 1 $\frac{1}{2}$ by 3 in. with 0.080-in. wall. Since both the coaxial line and waveguide are operating not very far from the voltage-breakdown point, a specially tapered "door-knob" transition is necessary. A small air gap in the choke-flange coupling between the transition section and the waveguide outlet absorbs the tolerances in the magnetron mounting. The whole magnetron high-voltage compartment is tightly closed in operation. The small upper compartment is mainly taken up with the fluted case of the pulse transformer. The antenna waveguide of the other radar goes through in front and the directional coupler is put in this convenient place.

In the central section are the two duplexer sections and the two receivers. The duplexer section for the search radar stretches across the lower part of the case. Nearest the magnetron is the AFC attenuator and mixer. The mixer is of the type shown in Fig. 11-25, Sec. 11-8. One coaxial cable brings r-f power from the local oscillator, which is mounted in a well-shielded compartment in the receiver box. The other carries the i-f signal to the AFC discriminator and control circuit in the same box.

Next along the guide and on the opposite side are the two ATR tubes, of the untuned low- Q type, spaced one half of a guide wavelength apart. This gives a broadband characteristic.¹ The resonant windows are made a part of the broad face of the waveguide. A half wavelength farther down the guide, and on the same side as the AFC attenuator, is found the duplexer T. The TR cavity is iris-coupled, both on the input side from the end of the waveguide and on the output to a coaxial mixer which is otherwise the same as the AFC mixer. As in the AFC case, i-f cables connect the mixer with the local oscillator and the proper parts of the receiver.

The last part of the duplexer section is an r-f switch used in testing

¹ *Microwave Duplexers*, Vol. 14.

the receiver sensitivity. In the radar operating position a straight piece of waveguide with a choke-flange joint at each end transmits the power direct to the antenna. There is, of course, appreciable r-f leakage at the air gaps of the two junctions. Accurate construction keeps this small enough not to affect the AFC mixer in the same compartment. A safety switch prevents turning on the magnetron unless the guide is in proper position. This piece slides out of the way and a matched probe-fed transition piece slides in for introducing a test signal.

The receivers are typical of ground and ship sets, except that in this case they do not contain the regulated power supplies. The usual i-f and video amplifier, the various anticlutter circuits, and the AFC circuits are included. A tight outer box provides general shielding, but the i-f amplifier and local oscillator have inner shields in addition.

One of the most noteworthy features of this radar is the built-in test equipment. An operator can go to the turret and, by transferring the usual remote controls to a duplicate set in the turret, operate the radar and view the echoes or waveforms on the scope provided. A universal meter furnishes a quick check on all the currents mentioned in Sec. 11-11. The vital check on receiver sensitivity is made with a klystron noise source, whose output signal is introduced via the r-f switch just described. The reliability and day-to-day reproducibility of the test set are excellent, and having it built in makes a quick daily check on a continuously operated radar simple to perform. Over-all noise figures of 9 to 12 db are commonly maintained.

Neither the r-f head nor the waveguide is pressurized. The moisture problem in the turret is solved by electric heaters which come on whenever the radar shuts down. Only a few degrees rise above general atmosphere temperature prevents condensation. Drain holes for water are provided at proper points in the antenna line.

CHAPTER 12

THE RECEIVING SYSTEM—RADAR RECEIVERS

BY L. J. HAWORTH AND W. H. JORDAN

INTRODUCTION

BY L. J. HAWORTH

12-1. The Role of the Receiving System.—It is the purpose of the receiving system to extract the information contained in the radio-frequency echoes, to sort it in terms of the geometrical parameters involved, and to present the results to the observer in a convenient and useful form.

The data involved are in a variety of forms. Most of them come from the radar set itself. First in importance, of course, are the echo signals, which may contain as many as several million separate pieces of information each second. The exact instant at which the radar pulse is transmitted is known by virtue of a pulse to or from the modulator. The geometrical coordinate of *range* is proportional to the return time of the echo pulses with respect to this pulse. The orientation of the antenna is available in terms of the rotational positions of shafts geared to the scanner axes. These direct data are often supplemented by others (obtained from external sources or from observation of the composite radar results), such as geographical information, the location and orientation of the radar platform, if it is a moving one, and so on.

Within the limitations of these data and the requirements of the particular situation, the equipment should present to the observer a continuous, easily understandable, geometrical picture of the radar targets under study, giving the location, size, shape, and, in so far as possible, the nature of each to any desired degree of accuracy. In many cases, the equipment should have provisions for instantaneously determining in a precise numerical way the exact position of each target with respect to the radar set or to other targets, and often it should furnish means for passing these results on to other devices in an automatic way. All of this must be done in such a manner that the echo signals have optimum sensitivity compared to internal system noise and to extraneous radiations.

The accomplishment of these objectives requires a considerable array of equipment. The ultimate link with the observer is the *indicator*. The indicator proper is usually supplemented by a considerable amount

of auxiliary equipment, the chief functions of which are to prepare the geometrical and other data, apart from the signals, for use by the indicator, to control the indicator functions, and to assist in the making of measurements. No satisfactory descriptive term exists for this auxiliary equipment, but in combination with the indicator it is known as the "indicating (or indicator) equipment." The echo signals must be amplified greatly and demodulated before being presented to the indicator, usually in the form of video signals. The path which they follow in this process will be spoken of as the "signal channel," and the equipment involved as the "receiver." On some occasions the system also contains computers or other devices which aid in the ultimate disposal of the data.

The receiver must deliver a maximum of desirable, and a minimum of undesirable, information to the indicator. It should have as little inherent noise as possible, since such noise determines the ultimate limit in signal detectability. It should afford sufficient amplification to realize this ultimate limit and sufficient dynamic range to allow wide latitude in useful signal intensities. The bandwidth must be chosen to provide sufficiently rapid transient response so that the details of the signals will be preserved; but it must not be so great as to decrease unduly the signal-to-noise discrimination. In many cases, special design characteristics enhance the discrimination of certain kinds of echo in comparison to other undesired echoes or to radiations from other transmitters.

The indicating equipment, which is almost entirely responsible for the purely geometrical aspects of the display problem, must share with the receiver the responsibility for the discernibility of the signals with respect to noise, interference, and signals from other targets. The indicator proper is almost always a cathode-ray tube whose screen presents the radar display. Synthesis of such displays involves combining the signal intelligence with antenna scanning angles and other geometrical factors to provide an intelligible picture of the dispositions and other characteristics of the radar targets. Some of the displays are, in principle, facsimile representations of the actual geometrical situation; others are deliberately deformed to improve some particular type of observation or measurement.

A wide variety of displays has been discussed in Chap. 6. The basic types are listed here for reference:

1. *Deflection-modulated displays*, in which the echo signals are used to deflect the beam laterally on the tube face. In practice, the other rectangular coordinate is invariably range (A-scope, R-scope, etc.).
2. *Intensity-modulated displays*, in which the signals serve to brighten the screen, and hence appear as bright spots or patches against a background that is partially illuminated by the receiver noise. These fall into several categories:

- a. True plots of a plane surface in which range and an angle are combined as polar coordinates (PPI).
 - b. Rectangular plots of two cartesian components of range, in general with unequal normalizations¹ (stretched PPI, RHI).
 - c. Rectangular plots of the polar coordinates, range and angle (B-scope, E-scope.)
 - d. Rectangular plots of azimuth angle and elevation angle (type C).
 - e. Modifications of the above types to indicate a third dimension in certain simple situations.
3. Presentations, not properly displays, in which the cathode-ray tube is used as a two-dimensional meter (spot error indicator).

Additional types of deliberate deformation of the basic displays and certain approximations used for technical reasons will appear later in detailed descriptions of the production of the various displays.

Observations and measurements are aided by various indices. Practically all the displays that include range are provided with a set of precisely timed "electronic markers" which occur at convenient regular intervals on the display itself. These are sometimes supplemented by a manually controlled continuously movable marker, which removes the necessity for interpolation. Measurements of angle can be made by means of similar electronic indices, but fixed lines etched on a transparent overlay plus, perhaps, a movable mechanical cursor, are often used.

12-2. A Typical Receiving System.—Figure 12-1 illustrates the principal parts and some of the subdivisions of a typical receiving system containing a cathode-ray-tube indicator.

The receiver is always of the superheterodyne type and consists of the signal channel and the local oscillator, together with frequency-control circuits for the latter. Strictly speaking, this should include all equipment concerned with the received signal, beginning with the antenna and ending with the input terminals of the indicator proper. However, in practically all radar applications the antenna, the TR tube, and the r-f line connecting them are shared with the transmission channel, and the techniques in these sections are dictated somewhat more by the transmitter requirements than by those of the receiver. We shall, therefore, consider the receiver as beginning at the point where the signal channel branches from that which is shared with the transmitter.

R-f signals from the antenna by way of the TR switch enter a crystal mixer where they are combined with the c-w output of a tuned local oscillator to form a heterodyned signal at the desired intermediate frequency, usually 30 to 60 Mc/sec. This signal is led from the mixer into an intermediate-frequency amplifier of very special characteristics, where

¹ If the normalizations are equal the display is, of course, equivalent to Category a.

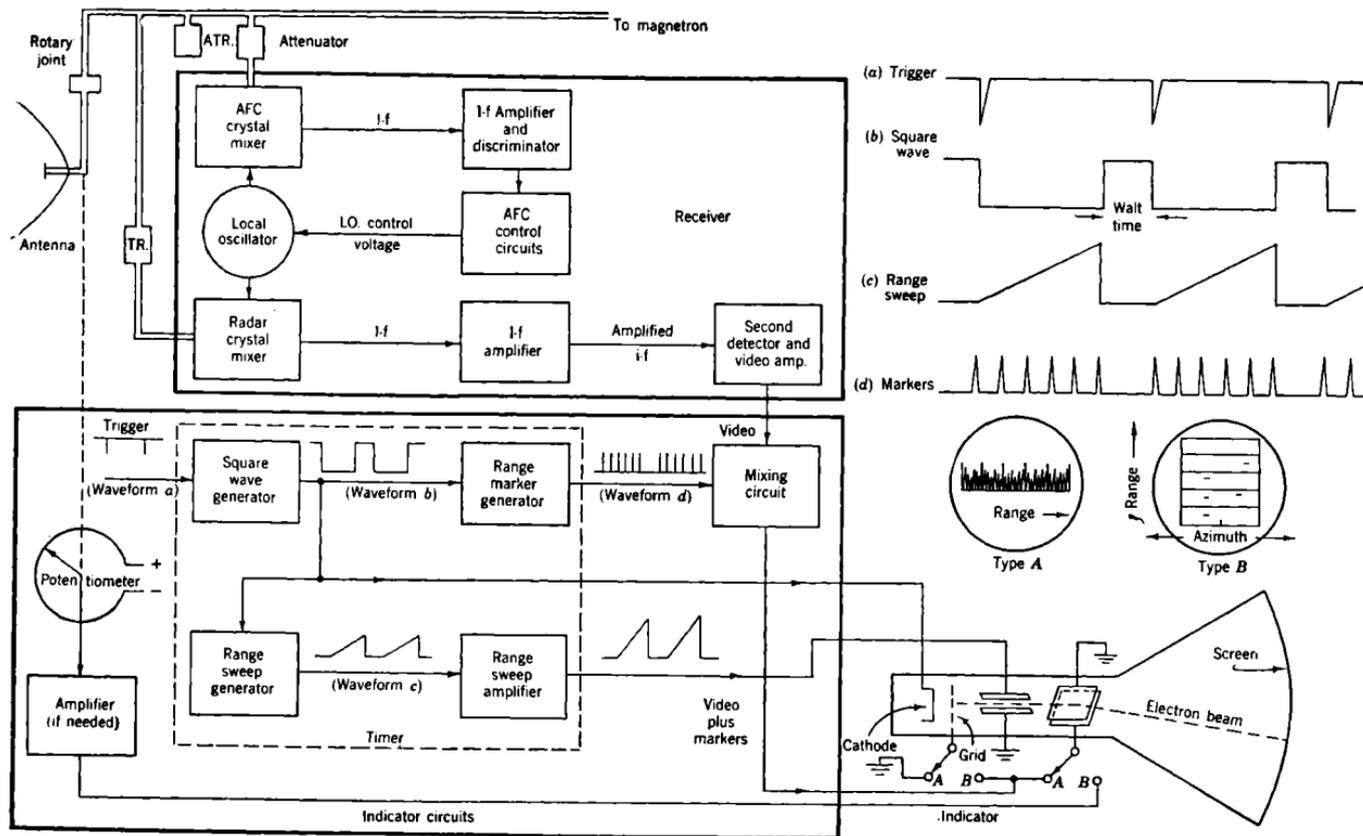


FIG. 12-1.—A typical receiving system.

it experiences an amplification of some 120 db. Following the i-f amplifier is a "second" detector and i-f filter which rectify the signals and remove the i-f components, leaving only the video signal envelope. The signal channel is completed by a video amplifier which delivers the signal at proper voltage level to the input terminal of the CRT or other indicator.

As in all superheterodynes, tuning is accomplished by controlling the frequency of the local oscillator. Since the ratio of the bandwidth to the carrier frequency is extremely small (of the order of one part in a few thousand), the tuning is critical. Some form of automatic frequency control is essential if constant manual tuning adjustments are to be avoided. Methods of AFC in pulse radar practice are quite different from those of radio or television; in radar, frequency control is exerted by passing a portion of the transmitted signal through a mixer and i-f amplifier, and maximizing its intensity by the use of a frequency discriminator and electrical circuits, which tune the local oscillator.

When the radar is to respond to beacon signals, the receiver must be tuned to the beacon transmitter rather than to the local one. In this case, an alternate local oscillator (not shown in Fig. 12-1) is forced to oscillate in resonance with a standard cavity whose frequency differs from that of the beacon by the intermediate frequency of the receiver.

The indicating equipment consists of the cathode-ray tube together with the auxiliary vacuum-tube circuits and other devices necessary to the synthesis of the display.

Those elements of the cathode-ray tube essential to a general understanding are shown in Fig. 12-1. In addition to the screen, these consist of a hot-cathode electron source, a "control grid" whose potential determines the beam intensity, and a mechanism for deflecting the beam. In the example illustrated, the deflection mechanism consists of two orthogonal pairs of parallel plates between which the beam must pass, the deflection due to each pair being proportional to the potential across it. In another type of tube, the deflection is produced by a magnetic field resulting from current in a coil or combination of coils surrounding the tube "neck." Arrangements for focusing the beam are not illustrated.

The equipment auxiliary to the cathode-ray tube varies widely with different situations but a few general statements can be made. Those parts concerned with the pulse-repetition cycle are collectively called the "timer." The timer provides synchronization with the modulator, sweeps and markers for the display and measurement of range, blanking of the cathode-ray tube during unused portions of the pulse cycle, and other related operations which may arise in special cases. The remaining equipment, apart from the necessary power supplies, is mainly concerned with the display and measurement of geometrical quantities other than range.

Since time is always measured with respect to the instant at which the modulator fires, synchronization between the modulator and the timer is of basic importance. In some cases, the timer exerts precise control of exact firing time of the modulator by sending it a trigger pulse; in others, it responds to a trigger from the modulator. In the latter case, the system is said to be "self-synchronous."¹ Although certain advantages can often be derived by giving control to the timer, this cannot always be done,² and it is sometimes inconvenient in a multiple-display system. Figure 12·1 illustrates a self-synchronous system. Some of the departures possible when the timer does the synchronizing will be pointed out in the appropriate places.

The cathode-ray tube must be turned on only during the fraction of the pulse cycle in which it is used. One portion of the timer generates a square wave (Waveform *b*, Fig. 12·1) which performs this function and usually finds other applications within the timer.

If the display includes a range sweep, as is nearly always the case, the timer generates the waveform that ultimately produces it. This waveform usually consists of a linearly increasing voltage wave which begins at the initiation of the sweep and returns to the initial condition at its end (Waveform *c*, Fig. 12·1). This involves a switching action which, as indicated in Fig. 12·1, is usually provided by the same square wave that turns on the cathode-ray tube.

The timer generates a set of discrete range markers: sharp video pulses occurring at regular, precisely known intervals. In a self-synchronous system, the markers must be recycled on each transmitter pulse; this requires a transient oscillator. In the simple example illustrated, the square wave *b* provides the necessary switching voltage. Often the discrete range markers are supplemented by a continuously variable range marker which may be generated in any of several ways. As indicated in Fig. 12·1, the markers are usually mixed with the radar video signals and the combination applied to the cathode-ray tube. In some cases, however, the signals and the markers are applied separately to different electrodes of the cathode-ray tube.

The equipment illustrated does not provide for the use of a delayed sweep. If one is to be used, separate square-wave generators are necessary to switch the range-marker circuit and to perform the other functions. The former square-wave generator must be triggered directly as

¹ A trigger generator for the modulator is sometimes housed with the indicator equipment for convenience, but the actual synchronization is accomplished by transmitting the modulator pulse to the timer proper. This case is functionally identical with that in which the trigger source is physically part of the modulator.

² Certain modulators, such as the rotary gap, cannot be triggered at all. Other types can be triggered but have so variable a response time that the modulator pulse itself must be used for synchronization in order to provide the necessary precision.

before. The latter must be provided with a delayed trigger which may come from a continuously variable timing circuit or may be one pulse of a discrete set, often the range-marker pulses themselves. The continuous type of delay is frequently associated with a continuous precision range marker (Sec. 13-12).

Many deviations from Fig. 12-1 are possible if the timer provides the synchronizing pulse. In particular, the discrete range markers can then be derived from a high-precision c-w oscillator, one of the marker pulses being selected as the modulator trigger by a scaling-down process. A continuous marker can also be provided from this oscillator by means of a continuous phase-shifting device. Other advantages of the timer-controlled synchronization entail the use of a "pre-trigger" by means of which the actions of certain circuits can be initiated somewhat in advance of the firing of the modulator. Many other functions, some of which will appear in later sections, are performed by the timer in complex situations.

The functions of those indicator circuits not included in the timer vary widely with the different display types. For illustrative purposes, Fig. 12-1 has been arranged to illustrate a simple A-scope or, alternatively, a simple B-scope. The former requires no equipment beyond that already described. The connections to the cathode-ray tube are shown as position *A* of the switch. The range-sweep voltage is applied to one pair of deflecting plates and the signal and range-mark voltages to the other. The square wave controlling the cathode-ray-tube intensity is applied to the cathode in proper polarity to brighten the tube during the range sweep.

In the B-scope display, the range sweep is applied to a deflection plate just as before. Signal modulation is applied to the control grid, and the second set of deflecting plates receives a voltage that produces the azimuthal deflection. This may be furnished, in simple cases, by a linear potentiometer geared to the scan axis. Amplification may sometimes be necessary before applying the azimuth sweep voltage to the cathode-ray tube. The circuits shown are equally applicable to a magnetic tube, the only appreciable changes being in the deflection amplifiers.

The timer of Fig. 12-1 can be used in the production of many other types of indication. For example, Fig. 12-2 shows the additional parts necessary to generate one type of PPI, and Fig. 12-3 those necessary for a (technically) different PPI or, alternatively, for one form of RHI.

The PPI of Fig. 12-2 is of the so-called "rotating coil" type. A single deflection coil driven by the range-sweep amplifier produces a radial range sweep. This is made to take the direction on the tube face that corresponds to the instantaneous antenna orientation by some form of electromechanical repeater. Except for this modification in the deflec-

tion system, the display circuits can be identical with those of the B-scope of Fig. 12-1.

Figure 12-3 shows the parts necessary to convert Fig. 12-1 to an "electronic" PPI without moving parts, or by a slight variation into an RHI. The range sweep of a PPI can be considered as made up of two orthogonal sweeps with speeds proportional to $\sin \theta$ and $\cos \theta$ (see sketch

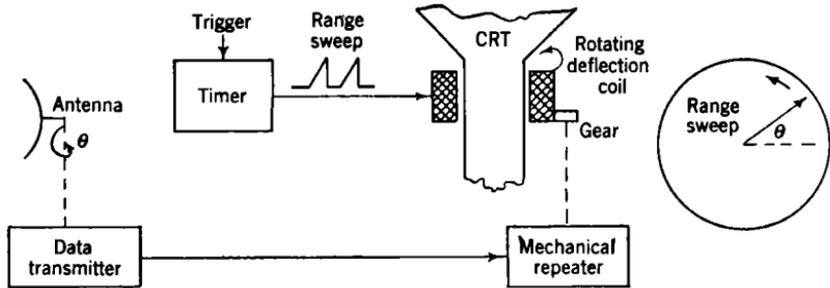


FIG. 12-2.—Rotating-coil PPI.

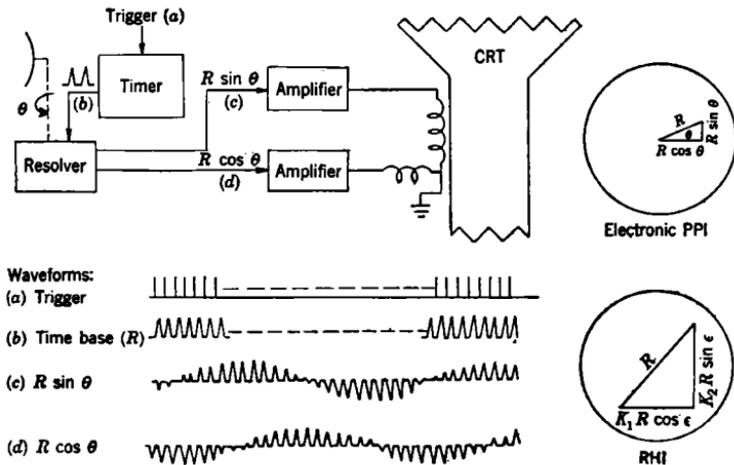


FIG. 12-3.—Electronic PPI and RHI.

in Fig. 12-3). In this particular form of PPI, these sweep "components" are formed by passing a sawtooth waveform through a sine-cosine "resolver" on the scanner. The resolver is a variocoupler or a sine-cosine potentiometer which, when excited by a given waveform, produces two "components" of the same waveform whose amplitudes are proportional respectively to the sine and the cosine of the orientation angle of the resolver. Proper amplifiers are provided for driving a pair of orthogonal coils (as indicated in Fig. 12-3) or, alternatively, the deflecting plates of an electrostatic tube.

If elevation angle is substituted for azimuth, the device becomes an RHI. Usually, in this case, the gain of the vertical amplifier is made greater than that of the horizontal, so that the display is "stretched" in the vertical direction. Since direct-coupled push-pull sweep amplifiers are usually used, they can be biased in such a way as to produce the "off-centering" shown.

THE RECEIVER

BY W. H. JORDAN

12-3. Special Problems in Radar Receivers.—Radar receivers, though similar in principle to all radio receivers, differ from them in some respects. This fact is chiefly due to a difference in emphasis on some of the functions. For example, it would be foolish to design a broadcast receiver with the ultimate in sensitivity when the weakest signal that can be detected is determined largely by man-made and natural static. In the radar portion of the spectrum, on the other hand, external sources of interference are normally negligible, and consequently the sensitivity that can be achieved in a radar receiver is normally determined by the noise produced in the receiver itself. Methods of reducing this noise are of prime importance in radar receiver design. Not only must noise be kept down, but everything possible must be done to minimize attenuation of the signal before it is amplified. How this has influenced the design of r-f components has already been seen. The effect on the design of i-f amplifiers will be developed in this section.

The lack of low-noise r-f amplifiers or converters in the microwave region has meant that most microwave receivers convert the r-f signal to an i-f signal directly in a crystal mixer and then amplify the i-f signal. This lack of previous amplification complicates the design of the i-f amplifier, as is shown by the following expression, given in Chap. 2, for the over-all receiver noise figure.

$$N_{\text{over-all}} = \frac{1}{g} (N_{i-f} + T - 1), \quad (1)$$

where

$$\begin{aligned} g &= \text{gain of the converter,} \\ T &= \text{noise temperature of the converter,} \\ N_{i-f} &= \text{noise figure of the i-f amplifier.} \end{aligned}$$

A good crystal in a well-designed mixer will have a noise temperature T that is only slightly more than 1. This means that the over-all noise figure is reduced in direct proportion to the reduction in i-f noise figure. Hence it is important to design the i-f amplifier to have a noise figure that approaches as nearly as possible the theoretically perfect value of 1.

Another very important factor in radar receiver design is that of securing a good transient response. Reception of pulses about $1 \mu\text{sec}$ in length imposes severe requirements on the pass band of the amplifier. At the same time, the amplifier must be capable of amplifying pulses several hundred microseconds in length and yet be fully sensitive to a weak signal immediately following. It must recover immediately from a signal thousands of times stronger than the minimum discernible signal.

Frequently special requirements are placed on the receiver. For example, it may be required that it respond to weak pulses even in the presence of a c-w signal of considerable strength. The problem of detecting a weak target echo in the presence of strong sea return is somewhat similar. Special circuits which go under the general title of "anti-clutter circuits" are usually required in such cases.

The unique requirements outlined in the foregoing three paragraphs have resulted in a receiver that is considerably different from anything previously existing. The design of i-f and video amplifiers to meet these requirements will be discussed in the following sections.

12-4. I-f Amplifier Design.—In an ideal i-f amplifier, all the noise would originate in the generator connected to its input terminals—that is,

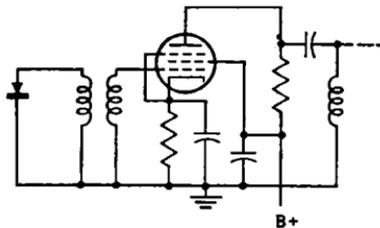


FIG. 12-4.—Pentode input circuit.

the crystal that serves as converter in the usual radar receiver. In any practical amplifier there are additional sources of noise. Thermal noise from resistances in the input circuit and shot noise due to the uneven flow of electrons to the plate of the first tube are the chief sources of excess noise, although later stages may contribute slightly. For this reason, care must

be taken in the design and tuning of the network coupling the crystal to the first grid as well as in the choice of the operating conditions and type of the first tube.

Figure 12-4 shows a typical input circuit, with the primary tuned to resonate with the crystal and mixer capacity at the intermediate frequency, and with the secondary tuned to resonate with the input capacity of the tube and socket. The coils are usually fixed-tuned or "slug-tuned"¹ in order to avoid any extra capacity. The only loading on the circuit is due to the crystal and the input resistance of the tube; thus, there are no additional resistances to contribute thermal noise. Thermal

¹ The inductance of a "slug-tuned" coil is lowered by the effect of eddy currents induced in a metallic rod (slug) inserted in one end of the coil. The degree of penetration, and therefore the inductance, can be varied by turning the slug in a threaded support.

noise due to the resistance in the coils themselves can be kept negligibly small by using coils of moderate Q . The coupling between primary and secondary is preferably magnetic, since capacitive coupling gives a somewhat poorer noise figure. The amount of coupling is usually fixed at a value that gives the best noise figure. This is necessarily greater than critical coupling, and in the case of a 6AC7 seems to be about transitional—that is, just before a double hump appears in the pass band. The bandwidth with this amount of coupling depends upon the crystal resistance and the mixer and tube capacities. Since the crystal resistance is small (around three or four hundred ohms), the bandwidth is usually adequate, being around 10 Mc/sec between half-power points. Considerably greater bandwidths can be obtained by paying attention to the mixer capacity.

A number of tube types have been tried in the first stage; however, most radar receivers today use the 6AC7, 6AK5, or 717. The 6AK5 and 717 are very similar in electrical performance and give a lower noise figure than the 6AC7; however, the latter is still widely used.

TABLE 12-1.—AVERAGE NOISE FIGURE

Circuit	Tube type		Intermedi- ate frequency, Mc/sec	Over-all receiver bandwidth, Mc/sec	Average noise, db.
	1st stage	2nd stage			
Grounded-cathode pentode to grounded-cathode pentode	6AC7	6AC7	30	1.5	3.9
	6AK5	6AK5	30	6	3.3
	6AK5	6AK5	60	16	6.5
Grounded-cathode triode to grounded-grid triode	6AK5	6AK5	30	1.5	1.5
	6AK5	6AK5	30	8	2.2
	6AK5	6J6	60	12	3.5
	6J4	6J4	180	3	5.5

The operating voltage and current for the first tube are frequently influenced by design considerations other than noise figure. In general, the plate and screen voltage should be low (75 to 120 volts are common) and the cathode current should be as high as is permitted by the tube ratings.

The noise figure that can be obtained by use of the circuit shown in Fig. 12-4 depends upon the intermediate frequency and the over-all receiver bandwidth. Since it varies considerably with individual tubes of the same type and manufacture it is desirable to quote average figures. Some representative values are given in Table 12-1.

It has long been realized that a large portion of the shot noise in a pentode is due to the interception of electrons by the screen grid. Hence,

a triode is less noisy than a pentode. The difficulty in using a triode lies in finding a circuit that is stable, uncritical in adjustment, yet with enough gain to swamp out noise originating in the following stage. A circuit that has recently been developed is shown in Fig. 12-5. The input transformer T is similar to the one used for pentodes. In order to realize the ultimate noise-figure capabilities of the circuit, the Q of the coils must be kept high.

The circuit shown in Fig. 12-5 consists of a grounded-cathode triode working into a grounded-grid triode. The input impedance of a grounded-grid amplifier stage is very low, being approximately $1/g_m$, or 200 ohms for a 6AK5. Since this impedance loads the first stage so heavily that its voltage gain is about 1, there is no tendency for it to oscillate, even without neutralization.

The loading is so heavy that the interstage bandwidth is very great. Since the value of L_2 is thus non-critical, the circuit is, in fact, fixed-tuned. Inductance L_3 is an i-f choke of almost any characteristics. Thus the circuit is stable and uncritical; it only remains to be shown that the noise contribution of the second triode is small. This is not obvious, and a rigorous proof is beyond the scope of this book (see Vol. 18 of this series). In order to minimize the second stage noise, the impedance

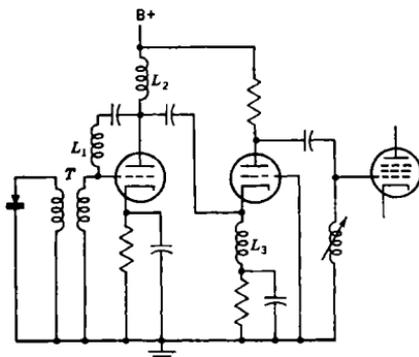


FIG. 12-5.—Triode input circuit.

seen when looking back from the cathode of the second triode must be large compared to the equivalent noise resistance at this cathode. To make this impedance as high as possible, an inductance L_1 is connected between plate and grid of the first triode. Inductance L_1 resonates with the plate-grid capacity at the intermediate frequency. This inductance is not needed for stability but does improve the noise figure about 0.25 db.

Noise figures obtainable with the double-triode circuit depend on several factors; representative values are shown in Table 12-1. Improvements of 2 db or more over the pentode circuit are usual.

Before describing the i-f amplifier, a brief discussion of some of the factors involved in choosing the intermediate frequency will be given. The over-all receiver bandwidth should be from 1 to 10 Mc/sec to pass the pulses ordinarily encountered in microwave radar sets.¹ An inter-

¹ In this chapter i-f amplifier bandwidth will be taken between the half-power points; video amplifier bandwidths will be measured between the frequencies at which the video response is 3 db down. The over-all receiver bandwidth (i.e., the

mediate frequency considerably greater than this is required in order that i-f sine waves can be filtered out of the video amplifier. Local-oscillator noise can be minimized by the use of a high intermediate frequency; however, the balanced mixer (cf. Vol. 16 of this series) provides a better solution to this problem. Many of the present AFC systems require the use of a high intermediate frequency to prevent locking on the wrong sideband. Finally, components such as condensers and coils become smaller as the frequency is raised, a distinct advantage in lightweight airborne radar sets. On the other hand, there are at least two very cogent reasons for favoring a low intermediate frequency: (1) the noise figure of the i-f amplifier is smaller at lower frequencies, and (2) manufacture and maintenance is considerably simplified because variations in tube and wiring capacitances, as well as in tuning inductances, affect the over-all receiver response much less. Thus the choice of an intermediate frequency is a compromise. Frequencies of 30 Mc/sec and 60 Mc/sec have been chosen for most of the present-day radar sets. The i-f amplifier bandwidth is not an important factor in the choice of the intermediate frequency. It is just as easy to achieve a bandwidth of, for example, 5 Mc/sec at a center frequency of 5 Mc/sec as it is at 60 Mc/sec.

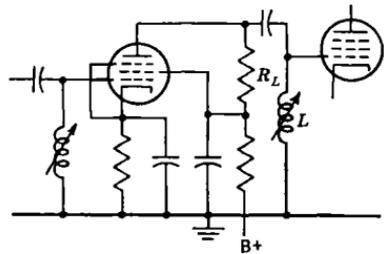


FIG. 12-6.—Single-tuned i-f amplifier.

An i-f amplifier consists of a number of cascaded stages. Figure 12-6 is a circuit diagram of a type of stage in common use.

This is known as a single-tuned stage, since there is one tuning inductance per stage. It has the advantage of being simple, easy to manufacture and align, and noncritical in adjustment. It is particularly useful in i-f amplifiers of over-all bandwidth less than 3 Mc/sec. It becomes expensive at wider bandwidths, although an improvement in tube performance would raise this figure proportionately.

The inductance L is tuned to resonate at the intermediate frequency with the combined output and input capacity plus stray capacity to ground due to sockets and wiring. It is placed in the grid circuit to provide a low-resistance path to ground. Thus, when the grid draws current during a strong signal, it does not accumulate a bias; hence the gain is not reduced and the amplifier remains sensitive to weak signals.

The gain G of the single-tuned stage shown in Fig. 12-6 is given by the expression

$$G = g_m R_L, \quad (2)$$

combined response of the video and i-f amplifier) will be taken as the equivalent i-f amplifier bandwidth unless otherwise stated.

and the bandwidth by the equation

$$\mathcal{B} = \frac{1}{2\pi} \frac{1}{R_L C'} \quad (3)$$

where C is the total capacity resonated by L . Hence the gain-bandwidth product is

$$G \times \mathcal{B} = \frac{1}{2\pi} \frac{g_m}{C}. \quad (4)$$

From Eq. (4) we can see the necessity of using tubes with as high a ratio of transconductance to input-plus-output capacitance as possible. The 6AK5 is slightly better than the 6AC7 in this respect, average tubes giving gain-bandwidth products of approximately 55 Mc/sec and 50 Mc/sec, respectively, when allowance is made for socket and wiring capacity. One can also see from Eq. (4) the necessity of keeping extraneous capacities to a minimum. It is for this reason that the inductance L is either fixed-tuned or slug-tuned, instead of being used with a tuning condenser. By the same token, the use of point-to-point wiring and the mounting of components on the tube sockets are clearly indicated.

Amplifiers with a gain of 120 db and an over-all i-f bandwidth of about 2 Mc/sec will require a stage bandwidth of over 6 Mc/sec, which permits a gain of approximately 7, or 17 db. Thus seven stages will be required. Assuming a g_m of 7000, a reasonable figure for a 6AC7, the load resistance R_L would be 1000 ohms [Eq. (2)].

Many more stages would be required to increase this bandwidth greatly, since the gain per stage must be lowered. The over-all bandwidth of an amplifier consisting of cascaded single-tuned stages is given approximately by the formula

$$\text{Over-all bandwidth} = \frac{\text{single-stage bandwidth}}{1.2 \sqrt{n}} \quad (4)$$

When n , the number of cascaded stages, is larger than 3, this formula is quite accurate.

Thus the number of single-tuned stages needed in an amplifier of given gain, even for moderately wide bandwidths, becomes prohibitively large. Lacking better tubes, the only alternative lies in the use of more effective circuits. Two things are needed: (1) a coupling circuit that will give a greater gain-bandwidth product for a given tube, and (2) a response-vs.-frequency curve that, when cascaded, does not narrow as rapidly as that given in Eq. (4). The double-tuned (transformer-coupled) circuit does very well in these respects. The elements of the usual double-tuned circuit are shown in Fig. 12-7.

The primary is tuned to resonate with the capacity in the plate circuit, the secondary with the capacity in the grid circuit. The coupling

is varied to give the desired response characteristic, usually that obtained just before the curve becomes double-humped. The loading R_p and R_s may be divided as shown, or placed on one side only. Two cases will be considered: (1) both sides are loaded so that the Q of the primary is equal to the Q of the secondary; (2) the loading is on one side only. With transitional coupling¹ the shape of the response curve is the same in either case, but the gain and bandwidth are not.

In the case of equal Q 's on both sides, the gain is given by

$$G = g_m \frac{R_p}{2} \sqrt{\frac{C_p}{C_s}} \quad (5)$$

where C_p and C_s are the circuit capacities associated with primary and secondary respectively. The bandwidth is

$$\mathcal{B} = \frac{1}{2\pi} \frac{\sqrt{2}}{R_p C_p} \quad (6)$$

Hence the gain-bandwidth product is

$$G \times \mathcal{B} = \frac{g_m}{2\pi} \frac{1}{\sqrt{2} \sqrt{C_p C_s}} \quad (7)$$

If the loading is entirely on one side, either primary or secondary, the above equations become

$$G = g_m \sqrt{2} R_p \sqrt{\frac{C_p}{C_s}} \quad (8)$$

$$\mathcal{B} = \frac{1}{2\pi} \frac{1}{\sqrt{2} R_p C_p} \quad (9)$$

$$G \times \mathcal{B} = \frac{g_m}{2\pi} \frac{1}{\sqrt{C_p C_s}} \quad (10)$$

If one compares the gain-bandwidth products given above with those previously quoted for single-tuned circuits, it will be seen that there is an improvement by a factor of 3.8 db in the equal- Q case and by a factor of 6.8 db in the case of loading on one side only.² There is the further

¹ As the coupling between two circuits of unequal Q is increased from zero, the response at resonance rises to a maximum at the "critical coupling" point and then decreases (with a flatter and flatter top to the response curve) until the point of "transitional" coupling is reached, after which the curve becomes "double-humped." If the two circuits have equal Q 's, transitional coupling and critical coupling are the same.

² Values of C_p and C_s of 7 and 14 μf , respectively, were used for this calculation.

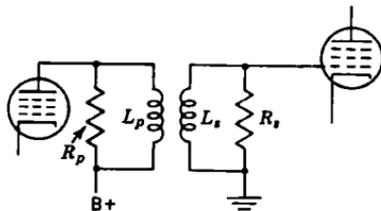


FIG. 12-7.—Double-tuned i-f amplifier.

advantage that cascaded double-tuned circuits do not narrow as rapidly as cascaded single-tuned circuits because of the rectangular form of the response curve (Fig. 12-8). For example, nine cascaded single-tuned stages, each of 6-Mc/sec bandwidth, would have an over-all bandwidth of 1.7 Mc/sec. However, nine cascaded double-tuned circuits, each 6 Mc/sec wide, would combine to give a 3.2-Mc/sec over-all bandwidth.

The over-all bandwidth of n cascaded double-tuned circuits is given by

$$\text{Over-all bandwidth} = \frac{\text{single stage bandwidth}}{1.1 \sqrt[4]{n}}$$

The usefulness of double-tuned circuits in wideband amplifiers has been demonstrated by their incorporation in a radar receiver having 120

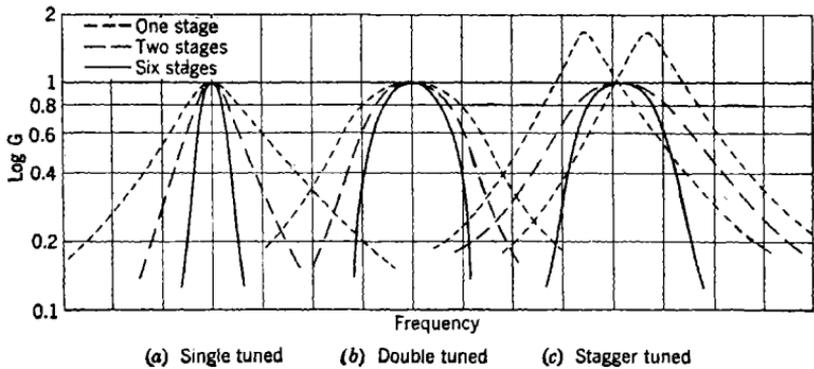


FIG. 12-8.—Amplifier response curves.

db of i-f amplifier gain and an over-all bandwidth of 12 Mc/sec. This was accomplished with only nine 6AK5 tubes in the i-f amplifier (Sec. 12-11). On the other hand, there are disadvantages in using double-tuned circuits. They are difficult to align, special equipment being required. Though this can be circumvented to some extent in wideband amplifiers by the use of fixed tuning, the sensitivity to variations in tube capacity and coil inductance is much greater than that of single-tuned circuits. This is particularly true when the loading is on one side only, so that this arrangement is ordinarily used only in very wideband amplifiers (15 Mc/sec or over).

Another means of achieving large bandwidth is by using stagger-tuned circuits.¹ Consider two cascaded single-tuned circuits of the type previously described, choosing the load resistors so that each has a bandwidth of 4 Mc/sec as shown in Fig. 12-8. The combined response curve of the two stages will be the product of the responses of the individual stages and is shown by the dashed line.

¹ H. Wallman, RL Report No. 524, Feb. 23, 1944.

This combined response curve has the same shape as that for a transitionally coupled double-tuned circuit, so that the advantage of cascading is preserved. In actual practice, a receiver that has a gain of 100 db with, for example, six stages, will be approximately twice as wide if staggered pairs are used as it would be if single-tuned circuits were used (Fig. 12-8).

The principle of stagger tuning can be carried further. Staggered triples (three different frequencies) are in fairly common use, and staggered n -uples are practicable for wideband amplifiers. The advantage of stagger-tuned amplifiers lies in the fact that simple single-tuned circuits are used throughout, a fact that makes for ease of manufacture and servicing.

There are other schemes for obtaining large bandwidths. Feedback pairs¹ and feedback chains² have both been used fairly extensively. More recently a scheme of stagger damping has been proposed.³

Before leaving the problem of i-f amplifier design, some mention of the gain required will be necessary. To detect signals barely visible above the noise it is necessary that the noise originating in the input circuit and the first stage be amplified to a point where it is easily visible on the indicator. How this gain will be split between the i-f and video amplifiers depends upon such considerations as second-detector efficiency and the complication involved in building high-gain video amplifiers. In most receivers the i-f gain is enough to amplify noise to 1 or 2 volts, thus bringing it into the linear region where diode detectors have maximum efficiency. This requires a gain of around 10^6 times at the intermediate frequency. Building an i-f amplifier with this gain while avoiding trouble with regenerative feedback has been one of the most difficult problems in radar receiver design. Ground current loops must be confined by careful bypassing and grounding; power leads must be properly filtered; coils must be wound and spaced intelligently; the shielding must be adequate. Careful attention must be paid to all of these items if the over-all response characteristic is to bear any resemblance to what is expected.

12-5. Second Detector.—The purpose of the second detector is to produce a rectified voltage that is proportional to the amplitude of the i-f waves. In most receivers it is important that this rectified voltage be proportional to the first power of the i-f amplitude (linear detector). However, in radar receivers higher powers are permissible so long as reasonable efficiency is maintained. One of the simplest and most com-

¹ Bartelink *et. al.*, "Flat Response Single Tuned I-f Amplifier," GE Report, May 8, 1943.

² A. J. Ferguson, "The Theory of I-f Amplifier with Negative Feedback," Canadian National Research Council Radio Branch Report PRA-59.

³ H. Wallman, "Stagger-damped Double-tuned Circuits," RL Report No. 538, March 23, 1944.

monly used second detectors is the diode detector shown in Fig. 12-9. The diode is usually a tube such as the 6H6 or the 6AL5, although many other tubes have been used and recently crystals have been developed for the purpose. The i-f bypass condenser C usually has around 10 μmf of capacity; it must be considerably larger than the plate-to-cathode capacity of the diode to get good detector efficiency, but it must not be so

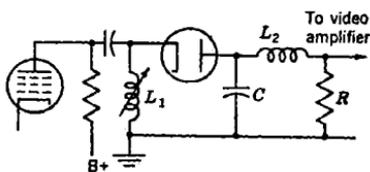


FIG. 12-9.—Diode detector.

large as to spoil the high-frequency response of the video amplifier. Theoretically, the high-frequency response would be maintained, however large C , by making R proportionally smaller. If R is too small, however, the detection efficiency is lowered and the maximum voltage that the detector can produce becomes too low. Values of R from a few hundred to a few thousand ohms are usual. Inductance L_2 is chosen to offer a high impedance at the intermediate frequency and is sometimes made resonant at this frequency.

12-6. Video Amplifiers.—One or more stages of video amplification follow the detector. The form of the video amplifier is determined largely by the number and location of the indicator tubes. In an airborne radar set there may be only a single indicator located near the receiver, in which case the video amplifier is very simple. In large ground and ship sets there will be many indicators located far from the receiver, and the video amplifiers must be more complex. In the latter case, the video amplifier following the detector drives a terminated line at a level of a few volts, and individual indicators are driven by video amplifiers bridged across this line.

The requirements placed on a video amplifier can be stated quite generally.

1. A signal of a few volts from the detector must be amplified and transmitted to the indicator, the signal level at the indicator being usually about 20 volts, although some types of indicator require a much higher voltage.
2. The amplifier must have good transient response, as defined by the following: it must pass long pulses with little "droop" on the top of the pulse; this requires good low-frequency response. The rise time¹ must be definitely shorter than the pulse length, and the overshoot on a pulse should be held to 10 per cent or less; these factors are determined by the high-frequency response. The

¹ Rise time is conveniently defined as the time required for the output voltage of the amplifier to rise from 10 per cent to 90 per cent of its final value when a square pulse is applied to the input.

amplifier must recover from large signals quickly, requiring that no control-grid current be drawn on strong signals.

3. Some form of limiting must be provided. The noise voltage at the output of the detector will be approximately 2 or 3 volts peak. In order that the noise be clearly visible on the screen, it must be amplified sufficiently to drive the indicator tube over something like half its allowable control-grid voltage swing. Since signal voltages from the detector may be 20 volts or more, the output voltage must be limited so that the indicator will not be driven outside its allowable range of voltage.

The simplest and most commonly used type of video amplifier is the resistance-capacitance-coupled amplifier shown in Fig. 12-10.

Although the gain for small signals can be simply expressed in terms of the tube and circuit constants, it is usually determined by reference to the characteristic curves of the tube. For many reasons, the high-frequency performance is most conveniently expressed by giving the frequency at which the gain is down 3 db (voltage down to 0.707). This will be referred to as the "cutoff frequency." For the amplifier shown this occurs when

$$f = \frac{1}{2\pi RC}, \quad (11)$$

where R = parallel resistance of R_p and R_o , and C is the total capacity to ground which includes output and input capacities plus stray capacities. Again the need for keeping the stray capacity to a minimum is apparent. This is sometimes difficult when large coupling condensers or long leads are involved.

There are other coupling networks that will increase the cutoff frequency and yet preserve the same gain. Perhaps the simplest of these is the "shunt peaking" circuit, which uses an inductance in series with the load resistor, R_p . Let the parameter M be defined by the equation

$$L = MR^2C. \quad (12)$$

Then M can be used as a measure of the performance of the circuit. Since a value of $M = 0.25$ corresponds to critical damping, it produces no overshoot and yet increases the cutoff frequency by a factor of 1.41. A value of $M = 0.41$ is the highest that can be used without putting a hump

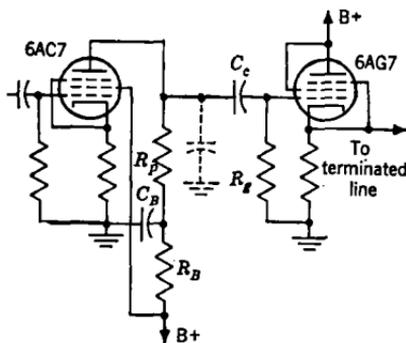


FIG. 12-10.—Video amplifier.

in the response curve. This value of M increases the cutoff frequency to a value of 1.73 times that of the simple RC circuit, but produces a 2.5 per cent overshoot on the pulse (the same amount as that produced by a transitionally coupled i-f stage).

Still greater bandwidths can be obtained by the use of series peaking¹ or a combination of series and shunt peaking. These circuits are more difficult to use, however, and the reader is referred to Vol. 18 of the Radiation Laboratory Series for details.

In order to preserve good response to long pulses, it is necessary that the time constant of the coupling circuit, $C_c(R_L + R_g)$, be kept large compared to the duration of the longest blocks of signals involved. In this event, the fractional amount of droop on the top of the long pulse is just the ratio of the pulse length to the coupling time constant. Thus if the time constant were 50,000 μsec while the pulse length was 1000 μsec , the pulse would droop 2 per cent. This is a reasonably small amount, but if there were five such circuits in cascade between the detector and the CRT, then the total amount of droop would be 10 per cent. Although it is fairly easy to maintain such long time constants provided $R_L + R_g$ is large, occasionally it happens that $R_L + R_g$ is fairly small. This means that C_c must be large, which may result in a large capacity to ground and affect the high-frequency response. For example, a 1- μf condenser may have several hundred micromicrofarads of capacity to its shell. Mounting the condenser on insulating posts is frequently necessary in such cases.

Another way of reducing the amount of droop on the pulse is by using low-frequency compensation circuits. In Fig. 12-10, if R_B and C_B are chosen properly, the beginning of the pulse will be flat and the time for the pulse to fall to 90 per cent will be greatly increased. The value of C_B must be chosen to satisfy the relation

$$C_B R_p = C_c R_g. \quad (13)$$

The value of R_B is not specified, but the amount of droop becomes less as R_B is increased. However, even for $R_B = R_p$, the time for the pulse to fall to 90 per cent is increased by a factor of 5 over the uncompensated network.

The limiting of video output signal level is usually accomplished by applying negative video pulses to the grid of a tube. When the pulses are large enough to drive the tube to cutoff, the output pulse remains constant in amplitude, no matter how much larger the input pulses may be. If the circuit shown in Fig. 12-10 were connected to the detector of Fig. 12-9, it would serve very well as a limiter-amplifier. With a sharp-cutoff tube whose screen is operated at low voltage (75 to 100 volts),

¹ In series peaking a compensating network is placed in series with the two stages.

signals of 4 or 5 volts peak reach the limit level. It is wise to have a limiter very early in the video amplifier, for then the following stages are protected from excessively large signals.

Self-bias is frequently used in a video amplifier, but, since a condenser of sufficient capacity to prevent degeneration at the lower frequencies would be prohibitively large, such a condenser is usually omitted entirely in order to assure uniform frequency response.

When the indicator is some distance from the receiver, the two are usually connected by a line terminated in its characteristic impedance. Unterminated lines have been used, but the high-frequency response is greatly impaired. Although lines of 1000-ohm characteristic impedance have been made and are fairly satisfactory for lengths of 20 ft or less, longer lines are usually standard 75- or 100-ohm coaxial cable. To drive such a line from the plate of an amplifier requires a very large coupling condenser; hence the cathode-follower circuit shown as the second stage in Fig. 12-10 is ordinarily used (see Sec. 13-6). Its chief advantage is that it can be direct-coupled. Although the gain is considerably less than 1, this is not appreciably less than would be obtained with a plate-coupled amplifier working into such a low impedance.

Following the line and bridged across it are the video amplifiers associated with each indicator. Since these amplifiers can be located very near to the control electrode of the CRT, stray capacities can be kept very small. These amplifiers can thus be either grounded-grid or grounded-cathode, depending on the polarity of signal desired.

12-7. Automatic Frequency Control.¹—The over-all bandwidth of a radar receiver may be determined by the bandwidth of either the i-f or the video circuits. However, for the purposes of this chapter we need only be concerned with the i-f amplifier, since we are interested here in questions of stability rather than in the quality of signals. As dictated by the principles outlined in Sec. 2-22, bandwidths of 1 to 4 Mc/sec are common. The narrowest r-f component, a high-*Q* TR switch, is so much wider than this that, once adjusted, no drifts large enough to cause serious detuning are likely to occur over a period of a day or so. The problem of keeping a radar in tune, then, consists essentially in maintaining the difference between the magnetron frequency and the local-oscillator frequency constant and equal to the intermediate frequency, with an accuracy of 1 Mc/sec or better. There are two reasons for making the tuning automatic. The first is that as the antenna scans, variations in standing-wave ratio arising from asymmetrical rotary joints or reflections from near-by objects can pull the magnetron several megacycles per second. Manual tuning is so slow relative to scanning that it brings in only part of the picture at a time. Second, whether or not

¹ By A. E. Whitford.

pulling exists, the magnetron and local oscillator will not hold constant frequency to one part in several thousand for several hours, nor will they vary together in some predictable manner. Variations in voltage, temperature, and pressure all produce frequency changes. Frequent tune-up by the operator, using echoes, is not dependable since serious deterioration in the sensitivity of a radar, such as that caused by drifting out of tune, is not immediately obvious on the scope. Also, in a ship on the open ocean or in an airplane over the ocean, there may often be no targets to provide echoes.

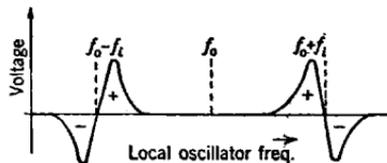


FIG. 12-11.—Height of discriminator output pulses, where f_0 is the transmitter frequency and f_i is the intermediate frequency.

For these reasons, an automatic-frequency-control circuit, AFC, has come to be a standard part of all radar sets. It is an electronic servo-mechanism that tunes the local oscillator in such a way that the proper difference frequency between it and the magnetron is maintained. The speed of tuning can usually be great enough to follow any pulling of the magnetron that may occur during the scanning cycle.

In the case of beacon reception, automatic frequency control maintains the local oscillator at a predetermined frequency that is higher or lower than the frequency of the distant beacon transmitter by just the intermediate frequency.

Radar AFC.—All forms of radar AFC work on the same basic principle. Part of the local-oscillator power is mixed, in a crystal, with a small fraction of the magnetron power drawn out during transmission of the pulse. The difference frequency is applied to a frequency-discriminator circuit whose crossover is set at the intermediate frequency. The variation of the height of discriminator output pulses as a function of local-oscillator frequency is shown in Fig. 12-11.

These pulses are integrated or otherwise converted to a voltage, which is applied to a control electrode of the local oscillator in the proper sense to push the frequency toward the crossover of the discriminator curve. This makes a degenerative feedback loop, and the frequency of the local oscillator is held very near the crossover point on the curve.

As shown in Fig. 11-22 (Sec. 11-7), the voltage on the reflector of a reflex klystron of the type used as a local oscillator tunes the tube over a range of about 30 to 70 Mc/sec. Some local oscillators are also thermally tuned by means of a control electrode which is the grid of an auxiliary triode. This provides slow tuning over a much wider frequency range (see Vol. 7 of this series).

A block diagram of a typical AFC system is shown in Fig. 12-12. The dotted portions are the additions necessary to include beacon AFC

for the beginning of the pulse and to accept a signal only from the latter part. This requires accurate timing and becomes very difficult for short pulses.

A satisfactory operating level for the AFC crystal is 0.5 mw c-w power from the local oscillator and 1 to 2 mw pulse power from the magnetron. Reduction of the magnetron power level by 65 to 90 db below the transmitted power is obtained by a combination of weak coupling out of the main r-f line and an attenuator, preferably dissipative. It is important that the r-f system be tight in order to prevent leakage power, possibly many times that desired, from reaching the crystal by stray paths. This becomes easier if the operating level of the AFC crystal is made as high as is permissible. A high operating level also requires less i-f gain ahead of the discriminator and thereby reduces the effect of gain variations.

Too high an operating level is, however, undesirable because as the crystal is driven up to 10 mw or above, it saturates. The harmonics of the difference frequency then become stronger relative to the fundamental. For example, if the discriminator crossover is set at 30 Mc/sec, corresponding to an i-f amplifier centered at that frequency, a second harmonic at 30 Mc/sec large enough to actuate the control circuit and cause locking may appear when the local oscillator is only 15 Mc/sec away from the magnetron frequency, under conditions of AFC crystal saturation.

The gain of the AFC feedback loop should be high enough to insure tight locking but not so high that the second harmonic can also cause locking. At the recommended level of 1 to 2 mw of magnetron power the harmonics are at least 20 db below the fundamental. A gain control is undesirable; therefore, care in controlling the r-f power levels and in amplifier design is required for a foolproof AFC. The most frequent sources of trouble are usually on the r-f side rather than in the electronic circuits, and arise from high power leakage into the crystal, faulty coupling to the main line, or a wrong amount of attenuation.

Beacon AFC.—To hold the beacon local oscillator at a given absolute frequency, some r-f reference standard must be provided, since the magnetron involved is in the beacon, distant from the radar. At microwave frequencies the reference standard is a resonant cavity. With proper attention to details such as temperature compensation and moisture sealing, production-line cavities can be depended upon to maintain a specified frequency to 1 or 2 parts in 10,000.

The dotted portion of Fig. 12-12 shows the beacon local oscillator attached both to the radar crystal and, through the reference cavity, to a beacon crystal. The output of the beacon crystal as a function of frequency is shown in Fig. 12-13. It is, of course, just the simple resonance

curve of the cavity. Under actual conditions of loading, the width Δf between half-power points is commonly about $\frac{1}{1500}$ of the resonant frequency itself.

Any control circuit that will lock the beacon local oscillator at the top of the resonance curve must in some manner take a derivative of the curve. The derivative, shown dotted in Fig. 12-13, has the same shape as the usual discriminator curve. In the scheme shown in the block diagram, a 1000-cycle sine wave from an a-f oscillator is added to the slowly varying voltage from the sawtooth search oscillator. This sweeps the local-oscillator frequency over a range that may be a tenth of the half-power bandwidth of the cavity. The crystal output then has an amplitude modulation of magnitude proportional to the slope of the cavity resonance curve at the frequency in question. The phase of this amplitude modulation depends on whether the slope is positive or negative. The output signal from the crystal goes to one grid of a coincidence tube, another grid of which receives the same sine wave used for modulating the local-oscillator frequency.

Only when the two coincide in phase will the coincidence tube conduct. As the frequency of the local oscillator drifts across the cavity resonance curve, the coincidence tube gives no output signal on the rising side of the curve, and passes 1000-cycle pulses increasing in amplitude as the top is crossed and the slope becomes negative. The pulse integrator uses these signals to stop the sawtooth sweep at a voltage that produces a local-oscillator frequency very near the top of the cavity resonance curve.

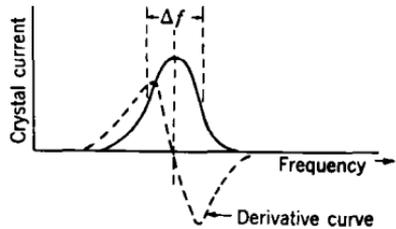


FIG. 12-13.—Beacon crystal current.

12-8. Protection against Extraneous Radiations. Antijamming.—Although the sensitivity of a radar receiver is normally limited by the noise produced in the receiver, extraneous radiation may be occasionally picked up on the radar antenna. Such radiation may, for example, be the result of the operation of other microwave equipment in the vicinity. It is good design practice to protect against such interference, and, as will be seen later, the same provisions are sometimes of value in the absence of interference. These AJ (antijamming) provisions may merely be precautions taken in the design of the receiver which do not affect its normal operation, or they may be special AJ circuits that can be switched in and out of use.

Interference may be any of several types such as CW, amplitude-modulated CW, frequency-modulated CW, r-f pulses, or noise-modulated CW. The more nearly the interference is like the echoes being received

or like the noise produced in the receiver, the more difficult it is to discriminate against.

One type of interference that must be guarded against is "hash," from rotating machinery and steep pulse wavefronts, which may leak into the i-f amplifier. Strict attention must be paid to shielding, and all leads into the receiver must be properly filtered and bypassed.

Perhaps the most fundamental AJ provisions are those that will permit the receiver to operate in the presence of considerable amounts of c-w power. If the second detector is directly coupled to the video

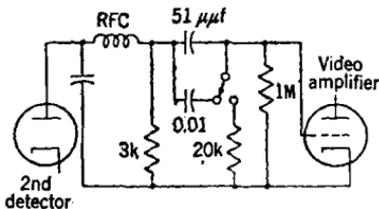


FIG. 12-14.—Fast-time-constant circuit.

amplifier, then a very small amount of c-w signal will develop enough voltage across the detector to bias the first video stage out of its operating range. This can be avoided by developing an equal amount of counter-bias, or by coupling the detector to the video amplifier through a blocking condenser. The latter course is much simpler and usually is the one

adopted. The receiver will now continue to function in the presence of a c-w signal until the i-f amplifier is overloaded. The i-f overload limit can be increased by operating the last stage or two with higher plate-supply voltage. Frequently a power tube, such as the 6AG7, is used in the last stage to increase the output capabilities of the i-f amplifier, although the 6AK5 has proven fairly satisfactory in this application. A rather wide video bandwidth (at least as wide as the total i-f bandwidth) is necessary when the frequency of the jamming signal is not the same as that of the radar transmitter. Use of a linear detector will also give the i-f amplifier a greater dynamic range. In some receivers it is necessary to apply a "gate," which sensitizes the receiver during a certain interval of time. It is desirable that such gating be done in the video amplifier rather than in the i-f amplifier, to avoid generating a pulse when a c-w signal is present. Finally, the operation of the gain control early in the i-f amplifier is of value when strong c-w signals are present. The gain control should be of a type that does not reduce the output capabilities of the controlled stages; grid gain control is satisfactory in this respect.

The precautions mentioned thus far constitute good design practice and can well be included in any radar receiver. To protect against either frequency- or amplitude-modulated c-w signals, such precautions are still necessary but not quite sufficient. In addition, there is needed between the detector and video amplifier a filter that will pass individual pulses but not the c-w modulation frequencies. Several types of filter have been tried; the simplest is the fast time constant (FTC) circuit shown in Fig. 12-14.

With the switch in the position shown, the time constant of the coupling circuit is $10,000 \mu\text{sec}$ so that moderately long blocks of signal will be passed. With the switch reversed, the time constant is $1 \mu\text{sec}$ so that individual pulses will be passed with some differentiation; long pulses and modulation frequencies below 20 kc/sec will be greatly attenuated. Even for modulation frequencies as high as 200 kc/sec , the circuit has considerable effect. It should, therefore, be switched in only when interference makes it necessary.

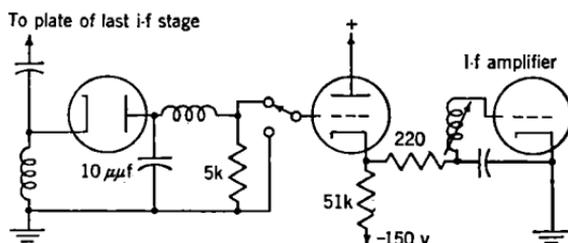


FIG. 12-15.—Instantaneous automatic gain control.

Differentiation circuits, of which the FTC circuit described above is a good example, are of real value so long as the i-f amplifier is not overloaded. When overloading occurs, the gain of the receiver must be reduced. As long as the jamming signal is constant in amplitude, manual gain control is a very satisfactory means of accomplishing the gain reduction. However, if the radar antenna is scanning or the c-w signal is modulated, manual control is much too slow to be effective. For this reason, various schemes for reducing the gain automatically and rapidly have been devised. These are variously called "instantaneous automatic gain control (IAGC)," "amplified back-bias," "back-bias," etc. The circuit shown in Fig. 12-15 shows the essential elements of an IAGC circuit.

Signals appearing at the output terminals of the last i-f stage are rectified by the detector and applied as a negative voltage to the grid of the cathode follower, which in turn controls the grid bias of one of the i-f amplifiers in such a way that the presence of a signal tends to lower the i-f gain. The time required for this reduction in gain to take place is determined by the constants of the circuits, as well as by the strength of the signal. In the particular example shown, this time is a few microseconds, so that the gain is not appreciably reduced by a single pulse but is cut down by long blocks of signals or by a c-w signal modulated at a low frequency.

The detector shown in Fig. 12-15 may be either the normal signal-detector or a separate detector. The output of the IAGC circuit may be fed back to the grid of the last stage or to that of the preceding stage. (Instability results from feeding back across too many stages.) When

it is necessary to prevent earlier stages from becoming overloaded, they can be shunted by another similar loop. The gain of the feedback loop can be increased either by using a plate detector instead of a diode, or by converting the cathode follower into a d-c amplifier. With a single feedback loop over two i-f stages, the loss in signal detectability resulting from a 60-db c-w jamming signal is only 12 db.

Although considerable thought and effort have gone into the problem of designing a receiver that can discriminate against more complex types of interference (pulses and noise), this is a much more difficult problem which will not be treated here.¹

Anticlutter Circuits.—Land, rough sea, and storm clouds reflect a considerable amount of r-f energy and can, therefore, interfere with the detection of objects in the area they occupy. If the desired signal is weaker than the "clutter" from the land, sea, or cloud, it is difficult to modify the receiver to discriminate against the clutter; for the frequencies contained in the clutter echoes are in the same range as those of the desired signals. If, on the other hand, the signal return is larger than the clutter, it should be possible to see the signal. However, the signal can be missed if the clutter is so strong as to saturate the receiver completely. This can be avoided by reducing the receiver gain to the point where the clutter is below saturation. Such a procedure would be satisfactory if the clutter were everywhere uniform, which it never is. Therefore, if the gain is reduced to the point required by strong clutter, it will be too low for regions of less or no clutter, and weak signals will still be lost. The IAGC circuit previously described is very valuable under such conditions and has been included in many receivers primarily for this reason.

In the case of sea return, the clutter is fairly constant at all azimuths. Also, the amount of sea return is a steadily decreasing function of range. It is possible to devise a circuit that lowers the receiver gain immediately following the transmitter pulse and then increases it steadily, arriving at maximum gain at the time the sea return has disappeared. This is known as a "sensitivity-time-control (STC) circuit"; it has the disadvantage that controls must be provided to adjust it for varying sea conditions. Nevertheless, STC has proven of considerable value at sea, and has also been used on some ground-based radar sets where control of aircraft close to the set is desired.

TYPICAL RECEIVERS

BY W. H. JORDAN

The practical application of the foregoing considerations will be illustrated by describing a few typical receivers chosen to cover a wide variety of purposes.

¹ See *Microwave Receivers*, Vol. 23, Chap. 10.

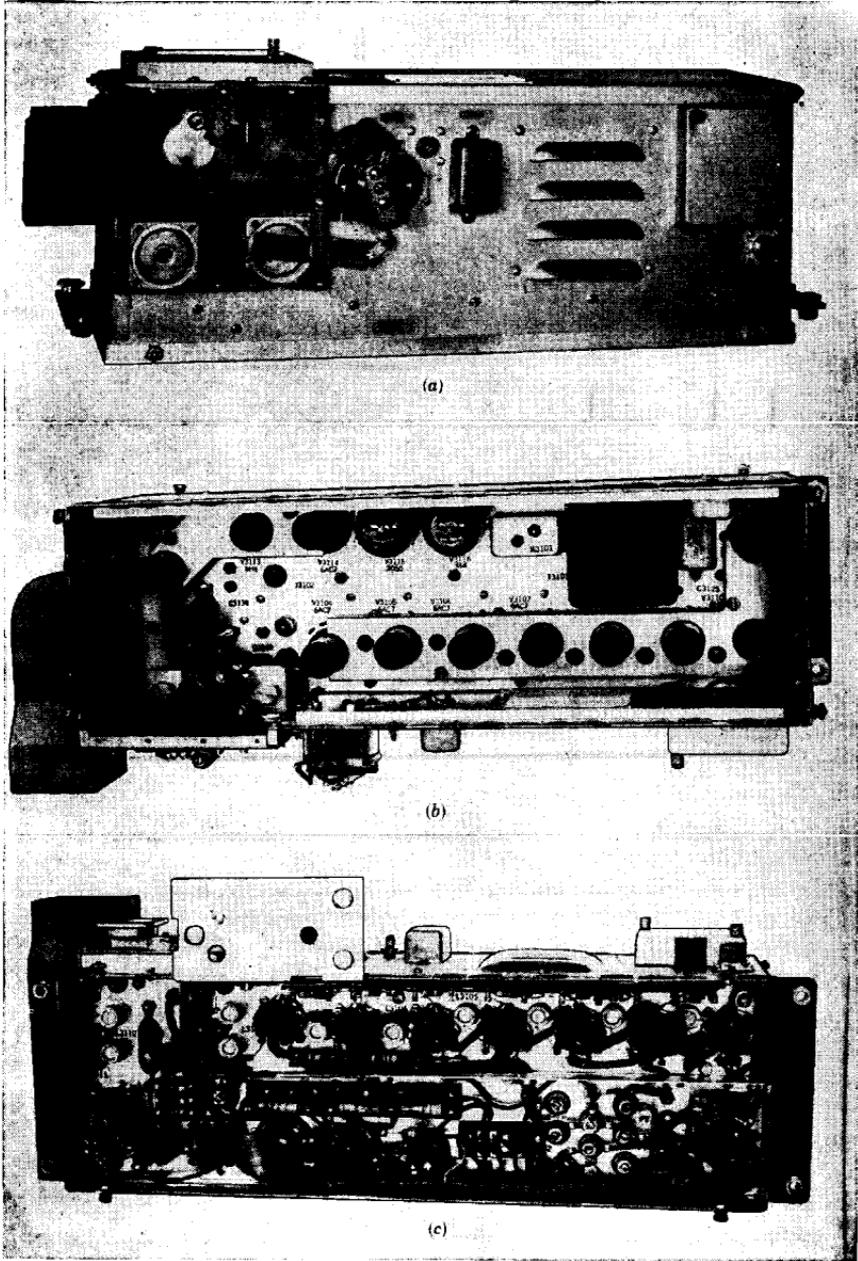


FIG. 12-16.—General-purpose receiver. (a) Side view; (b) top view; (c) bottom view.

12-9. A General-purpose Receiver.—This receiver was designed as part of a 3-cm radar set installed on small ocean-going vessels for sea-search and navigational purposes. The receiver is built into a steel box which fits into the r-f head. The general construction and layout can be seen from Fig. 12-16. Separate AFC and signal mixers are incorporated in the receiver; the choke joints for connection to the r-f system can be seen in the side view. A shutter is shown across one of the waveguides in the position it occupies when the transmitter is off. This shutter provides crystal protection against signals picked up on the antenna when the TR switch is not firing. At the other end of the receiver are a video-output jack and a terminal strip for bringing in power and certain control voltages.

The local oscillator is a 723A/B klystron. It supplies r-f power to both the AFC mixer and the signal mixer. Its frequency can be varied over a wide range by mechanical tuning and over a smaller range by varying the reflector voltage.

Except for the common local oscillator, the receiver is divided into two portions, the AFC channel and the signal channel (see Fig. 12-17). The latter consists of six i-f amplifier stages, a detector, and a video stage. Each stage of i-f amplification uses a 6AC7 pentode with single-tuned coupling circuits. All stages but the last have a 1200-ohm load resistance,¹ giving a single-stage bandwidth of a little over 6 Mc/sec and an over-all i-f bandwidth of nearly 2 Mc/sec. The nominal gain of the i-f amplifier is 120 db. A 6H6 diode second detector is capacitively coupled into a single video stage.

The video stage serves as a combined line driver and limiter. It operates as a cathode follower, the cathode resistance of 100 ohms being located at the end of the video line. The screen voltage is set at such a value that the tube draws approximately 10 ma of current with no signal; this puts 1 volt across the 100-ohm line. Negative signals from the detector may drive the grid to cutoff, thereby reducing the cathode voltage to zero. The video signals at the cathode are then negative and limited to 1 volt in amplitude.

The gain of the i-f amplifier is adjusted by varying the voltage on the grid of the second and third stages by means of an external control.

The receiver can be tuned either manually by means of an external potentiometer, or automatically by means of a 5-tube AFC circuit. The AFC circuit is of the hunt-lock type briefly described in Sec. 12-7. The reflector of the 723A/B local oscillator is swept over a range of about 30 volts, the center of the range being set by the manual frequency

¹ R-f chokes, self-resonant at 30 Mc/sec, are wound on the load resistors to reduce the power dissipation in the resistor and to keep the plate voltage up to the screen voltage.

control. The sweep voltage is generated by the 884 thyratron operating as a relaxation oscillator. When the discriminator crossover point is reached, positive pulses from the video amplifier V_{11} fire the gas tube V_{12} , thus stopping the sweep.

To obtain the control pulses a very small part of the transmitter r-f power is mixed with a portion of the local-oscillator output signal in the AFC crystal. The beat frequency generated by the crystal is then amplified by V_9 and applied to the discriminator detector. The discriminator output signal, in the form of a video pulse, is amplified by V_{11} before being applied to the control tube.

12-10. Lightweight Airborne Receiver.¹—This receiver is part of a lightweight 3-cm airborne radar (AN/APS-10) intended primarily for navigation. It was designed to require a minimum of field test equipment and a minimum number of highly trained service personnel. Accordingly, as far as possible the system was divided into small units that could be replaced, rather than repaired, in the field.

The receiver is distributed among several subunits, which, except for the video amplifiers, are enclosed (along with the r-f components, transmitter, modulator, and most of the power supplies) in a pressurized container forming one of the major units of the system. These subunits of the receiver are the following:

1. A unit containing a double crystal, mounts for 723A/B radar and beacon local oscillators, and a low- Q reference cavity for the beacon AFC.
2. An AFC unit for both radar and beacon local oscillators.
3. An i-f strip containing the entire i-f amplifier, the second detector, and a video cathode follower.
4. Power supply.

The receiver supplies low-level video signals to either one or two cathode-ray tubes, each of which is equipped with a video amplifier.

The i-f amplifier, a schematic view of which is shown in Fig. 12-18, contains eight 6AK5 pentode tubes. Its construction is shown in Fig. 12-19. The input circuit is a π -network consisting of three self-inductances. A crystal-current jack and decoupling filter are provided. The first six tubes are i-f amplifiers whose interstage coupling networks are single-tuned circuits arranged in two stagger-tuned triples. The center frequencies of these circuits are, in order, 30.0, 33.7, 26.7, 33.7, 26.7, and 30.0 Mc/sec, giving an average over-all i-f bandwidth of 5.5 Mc/sec centered at 30 ± 1.5 Mc/sec. Fixed tuning is employed in this replaceable subunit. The i-f bandwidth is made considerably wider than that

¹ By L. Y. Beers and R. L. Sinsheimer.

necessary for the reception of the 0.5- μ sec pulses employed by the system, to allow for tolerances in the adjustment of the AFC unit and in the manufacture of the i-f strip as well as for a slight spread in the frequencies of beacons. The voltage gain between the grid of the first tube and the plate of the sixth is about 30,000.

The seventh tube is a plate detector, which gives 10-db gain as well as a somewhat larger power output than could have been obtained from a diode detector of the same bandwidth. This advantage, however, is

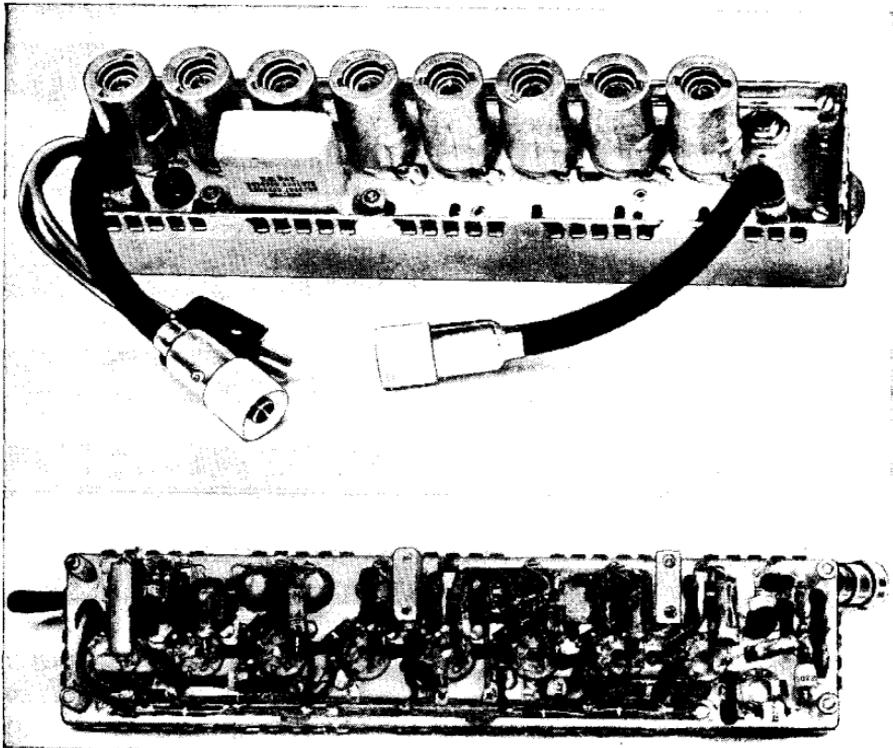


FIG. 12-19.—AN/APS-10 receiver; top and bottom views.

accompanied by three disadvantages. First, this detector is not strictly linear, a difficulty that is not too serious since this system is not intended to be operated in situations where interference is to be encountered. Second, with certain detector tubes there is some "blackout effect"—that is, the sensitivity does not recover immediately after the transmitted pulse or after other very strong signals. Third, care must be employed in the design of the power supply to reduce plate-supply ripple because such ripple is amplified by the entire video amplifier. The detector has an output test point for convenience in checking the pass band.

The last tube is a cathode follower which supplies negative video

signals at 1-volt peak, through a 75-ohm line, to the video amplifiers at the cathode-ray tubes. The half-power video bandwidth of the coupling network between the detector and the cathode follower is 6 Mc/sec.

The noise figure determined from measuring approximately 100 such amplifiers ranged from 1.7 to 4.5 db, with the average at 3.3 db. This amplifier was designed before the development of the grounded-cathode grounded-grid triode input circuit. The use of such an input circuit would require one more tube or the equivalent thereof. As an experiment, one of the standard receivers was modified to use this input circuit, the first seven tubes being in the i-f amplifier. The detector tube was replaced by a 1N34 germanium crystal, which allowed the unit to be con-

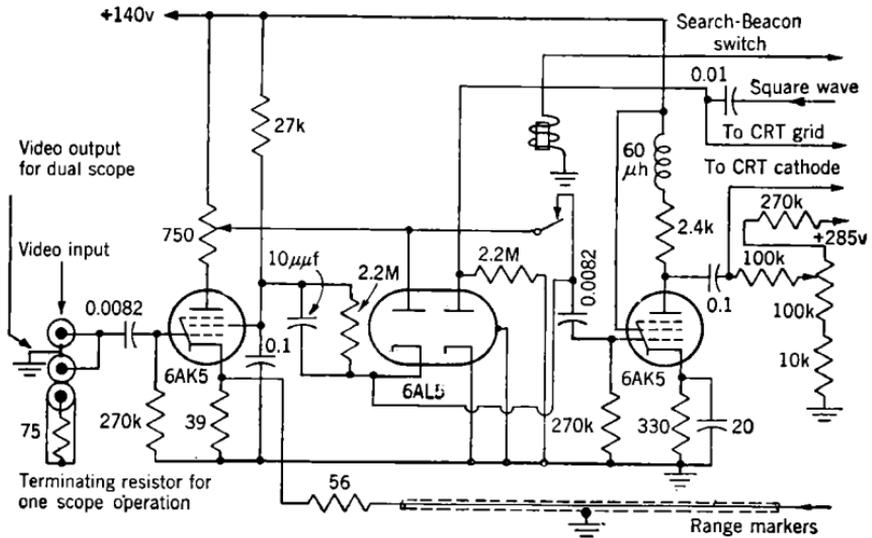


FIG. 12-20.—Video amplifier for AN/APS-10.

tained in the same chassis, though with some crowding. The noise figure of this experimental amplifier was 1.6 db.

The amplifier is mounted on a 9¼ by 2¼-in. chassis, which permits easy access to all components. The sides of the thin stainless steel cover are held by special springs at the edge of the chassis. In order to obtain stability, it was necessary to provide the chassis with two baffles to reduce feedback. These consist of small pieces of sheet metal fastened to the chassis and making contact with the bottom of the cover. Good contact between the bottom of the cover and the baffles is obtained by the use of screws which pass through the cover and are tapped into the baffles.

The i-f strip requires a power supply giving 65 ma at +105 volts unregulated for plates and screens, 1.4 amperes at 6.3 volts a-c for heaters, and a voltage variable from 0 to -10 volts (with a source

impedance not over 2000 ohms), which is applied to the control grids of the second and third i-f stages to control the gain.

The video amplifiers for the cathode-ray tubes (Fig. 12-20) have two stages and use 6AK5 tubes. The input signal is -1 volt peak; the amplifier output supplies -30 volts peak to the cathode of the indicator tube. The half-power bandwidth is 4 Mc/sec.

The chassis also contains a 6AL5 dual-diode tube. One half of this tube is used in a circuit to lengthen the duration of beacon reply pulses and so give a brighter spot on the CRT screen. On the rise of the pulse the diode conducts, and the input capacity of the second tube is charged rapidly. When the amplitude of the pulse starts to decrease, the diode no longer conducts, and the charge on the grid of the second tube leaks away slowly through a 2.2-megohm resistor.

The other half of the 6AL5 is used as a d-c restorer for the cathode-ray tube.

This amplifier is mounted on a chassis having the form of an annular ring which fits around the neck of the cathode-ray tube. The power supply requirements are 20 ma at $+140$ volts for plates and screens, and 0.5 amp at 6.3 volts for heaters.

Automatic-frequency-control Circuit.—The AFC subunit provides automatic control of the radar and beacon local-oscillator frequencies. The schematic view of the eight-tube circuit is shown in Fig. 12-21.

Both search and beacon AFC systems are of the search-and-lock type (Sec. 12-7). A sawtooth generated by the recovery of the grid circuit of a blocking oscillator, V_5 , sweeps the reflector voltage and thus the frequency of a klystron. In radar operation, i-f signals from a separate AFC crystal are amplified by V_1 and applied to the discriminator formed by V_2 and associated circuits, which is centered at 30 Mc/sec. The output pulses from the discriminator are amplified in V_3 and applied to the grid of the thyatron V_4 . When the pulses change sign and become positive at the crossover, they trigger the thyatron and stop the sweep voltage.

In beacon operation, power is transferred from the search local oscillator to the beacon local oscillator, which (as described in Sec. 12-7) is coupled to the receiver crystal and to a reference cavity tuned to a frequency 30 Mc/sec below the beacon frequency. A small 1000-cycle sinusoidal modulation, supplied by oscillator V_8 , is superimposed on the sawtooth applied to the reflector. The output of the beacon crystal, which is coupled to the local oscillator through the reference cavity, is amplified in V_6 and then applied to V_7 .

As explained in Sec. 12-7, the phase of the a-f amplitude modulation of the signal from the beacon crystal will change by 180° when the frequency of the local oscillator crosses the center frequency of the reference

cavity. This phase shift is detected in the phase-comparison tube V_7 by applying the amplified crystal output to the control grid and the audio oscillator output to the suppressor grid. Both grids are biased to cutoff, so that no current flows in the tube unless the signals are in phase. The circuit is so arranged that the 180° shift in phase of the audio signal produces phase agreement, thereby providing pulses of plate current on alternate half cycles of the audio signal. These pulses, further amplified in V_3 , then trigger thyatron V_4 to stop the sweep.

The change from search to beacon operation is performed remotely by relays which switch the grid input to V_3 , switch from the reflector voltage-control potentiometer P_1 to P_2 , change over the plate power from the search to the beacon local oscillator, and superpose the sinusoidal modulation on the sawtooth.

Switch S_1 enables a maintenance man to tune the local oscillators manually to the proper frequencies for signal reception with P_1 and P_2 (for the search or beacon local oscillators respectively), and then to revert to AFC operation centered at the correct frequency. This switch is spring-loaded so that the circuit cannot be left on manual tuning.

Jacks are provided at which the radar AFC crystal current and the beacon AFC crystal current can be measured. A pin-jack test point is provided at the grid of V_3 to aid in aligning the discriminator, and another at the plate of V_3 for checking the control signals to the thyatron.

This AFC unit requires a power source supplying 5 ma at +300 volts, 23.5 ma at +105 volts, 2 ma at -300 volts, 1.55 amp at 6.3 volts, and 0.8 amp at 6.3 volts at -225 volts from ground. None of these voltages needs to be regulated, although the plate voltages must be well filtered.

12-11. An Extremely Wideband Receiver.—The receiver shown in Figs. 12-22 and 12-23 is a good example of a design for a receiver having a very wide bandwidth. It was designed for a 3-cm radar system with a $0.1\text{-}\mu\text{sec}$ pulse width. In order to reproduce such short pulses and still maintain a good signal-to-noise ratio, an i-f bandwidth of at least 10 Mc/sec is required; the bandwidth of the receiver is 12.5 Mc/sec.

The i-f amplifier consists of 12 stages, 6AK5 tubes being used in all but the second. The intermediate frequency is 60 Mc/sec, the total gain being approximately 120 db.

An intermediate frequency of 60 Mc/sec instead of 30 Mc/sec was chosen for two reasons. First, fairly tight coupling is required between the primary and secondary of the i-f transformers to attain wide bandwidths. It is easier to make this coupling tight at 60 Mc/sec than at 30 Mc/sec. Second, the AFC operation is more certain at the higher frequency. The AFC circuit can lock the local oscillator on either side of the transmitter, provided the local oscillator is producing power at both frequencies. However, since the correct beat frequency is produced

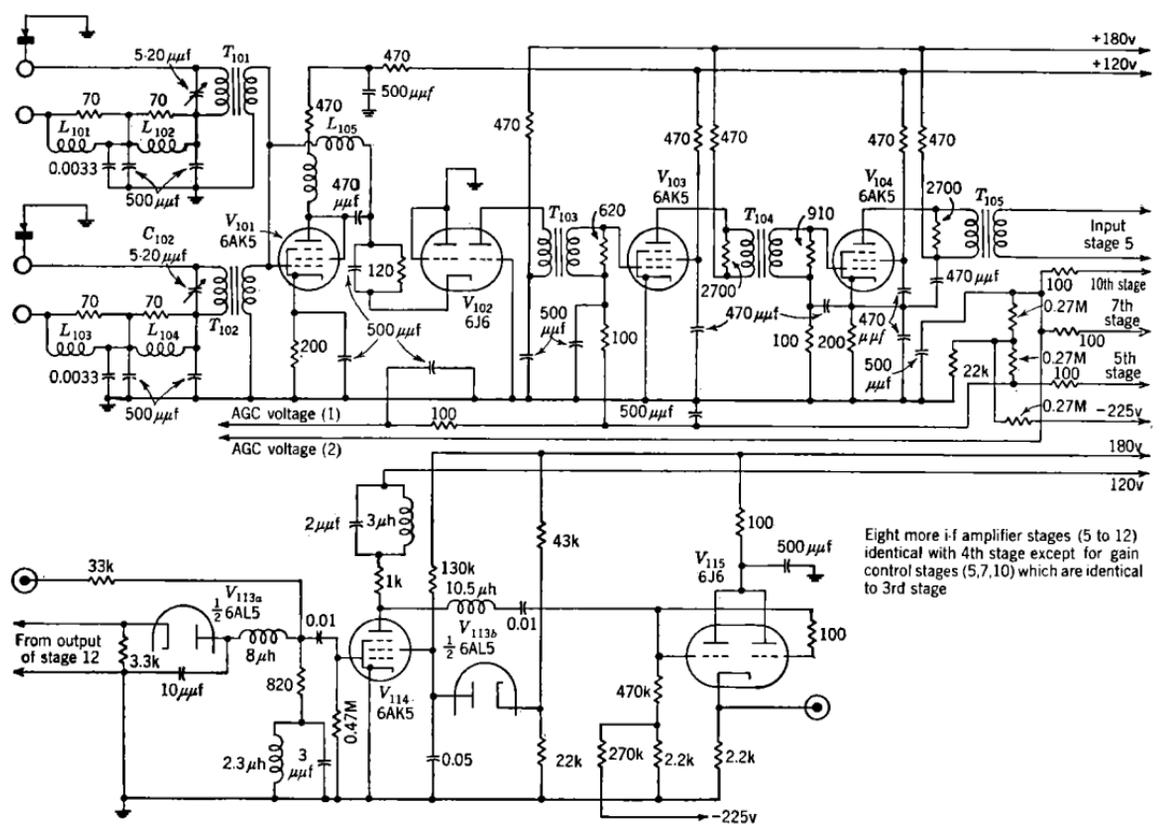


FIG. 12-22.—I-f and video amplifier for very wideband receiver.

Eight more i-f amplifier stages (5 to 12) identical with 4th stage except for gain control stages (5,7,10) which are identical to 3rd stage

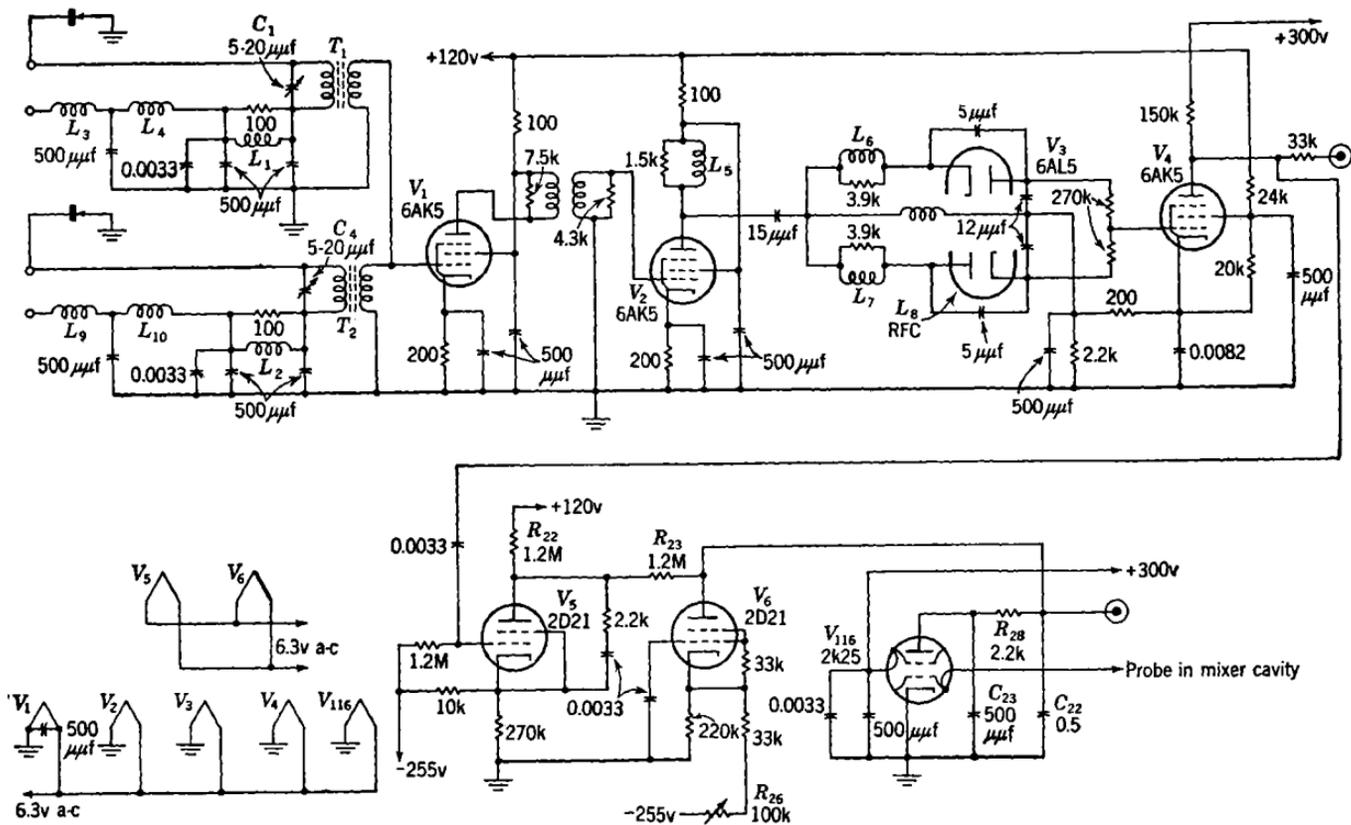


FIG. 12-23.—AFC for very wideband receiver.

in only one of these cases, it is necessary for the local oscillator to cease oscillating when it is on the "wrong" side of the transmitter frequency. This can be counted on only when the intermediate frequency is high.

Double-tuned circuits were chosen in preference to stagger-tuned, because it was felt that they would give a smaller change in the shape of the response curve for a given change in tube capacity. The circuits were loaded on both sides to reduce still further the sensitivity to tube-capacity changes. Equal values of primary and secondary Q would have been best in this respect; actually a compromise Q -ratio of 2.2 was adopted to give a little more gain than equal Q 's would provide. The transformer coils are wound adjacent to each other on a powdered-iron core to give tight coupling. A thin spacer between the two coils is of such thickness as to make the coupling just transitional. The gain per stage is approximately 11 db, and the single-stage bandwidth 25 Mc/sec.

The 6AK5 tube is chosen in preference to the 6AC7 because it is smaller, takes less power, and has a smaller variation in input and output capacity. It also gives a better receiver noise figure when used in the first stage. A 6J6 is used in the second stage because of its low plate-to-cathode capacity, an important factor in grounded-grid triode operation.

The first two stages are operated as triodes; in spite of the high intermediate frequency and the broad bandwidth, an i-f noise figure of 3.5 db can usually be attained.

Both signal and AFC mixers are of the balanced type. This effectively cancels out local-oscillator noise and provides good isolation between the two channels. There are two crystals in each mixer, provision being made for monitoring the current in each of the four crystals.

The local oscillator is a 2K25 klystron, which can be tuned mechanically over a large range and electrically over a restricted range of some 30 Mc/sec. The electrical tuning can be done manually, by means of a remote potentiometer, or automatically.

The AFC circuit operates as follows. The beat frequency obtained by mixing a small part of the transmitter power with local-oscillator power is amplified by two i-f stages and then applied to the discriminator. The video pulses from the discriminator undergo one stage of amplification before reaching the control tube, V_5 (Fig. 12-23). Tube V_5 generates a sawtooth sweep which moves the reflector voltage of the local oscillator through a range determined by the setting of R_{26} ; the sweep is stopped at the correct voltage to receive signals by the firing of the control tube.

The video amplifier consists of a limiter-amplifier stage which drives a cathode follower operating into a line terminated with 75 ohms. Limited signals of 1.5 volts amplitude appear across the line. The video amplifier is very wide, the bandwidth of each stage being about 22 Mc/sec. This bandwidth is obtained by means of a shunt-series peaking network

which gives a gain-bandwidth product 2.4 times as great as that of a simple RC circuit and still has only 0.3 per cent overshoot.

The limiter stage is operated with a low voltage (50 to 55 volts) on the screen, which means that only 2 volts of signal is required to drive the tube to cutoff. This reduces the amount of i-f gain required and increases the ratio of the maximum obtainable signal to the video limit level. During extended periods of high duty ratio the screen voltage will increase, thereby increasing the height of the limited video-output

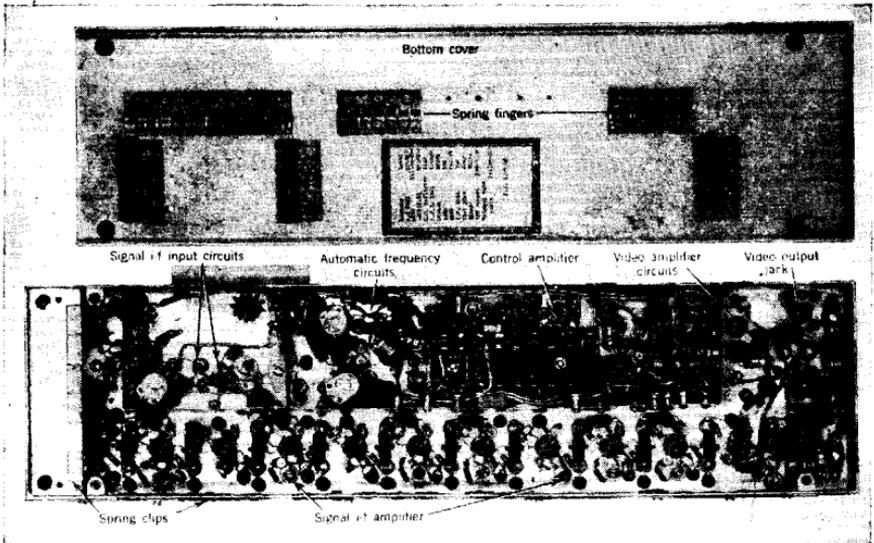


FIG. 12.24.—Receiver chassis of a very wideband receiver, bottom view.

signals. The rising screen voltage is checked at 60 volts by the diode V_{13b} .

The layout of the receiver can be seen in Fig. 12-24. The i-f amplifier is in the form of a long, narrow strip, folded at one end. The coupling transformers and associated components are laid in a channel which becomes a tube when the cover is in place. Coupling between output and input circuits is very small, since the tube acts as a waveguide attenuator at the intermediate frequency. The mixer (not shown) is attached to the top of the chassis directly behind the input coupling transformers at the upper left of the picture. The AFC circuits occupy the upper central portion of the chassis.

CHAPTER 13

THE RECEIVING SYSTEM—INDICATORS

BY L. J. HAWORTH

THE CATHODE-RAY TUBE

The effectiveness of a cathode-ray tube as a radar indicator is influenced by a number of factors, among them: (1) the size of the tube face; (2) the intensity of the emitted light, which is determined by both the strength of the electron beam and the efficiency of the screen; (3) the decay properties of the screen and the manner in which it integrates repeated signals; (4) the resolution; (5) the grid-modulation characteristics; (6) the effect of the tube on the over-all complexity, weight, size, and power dissipation of the complete indicating equipment.

These properties are not independent, and in many cases a compromise among them must be reached. As a result of the variety of operational demands a large number of types of tubes have been developed. The next few sections describe some of their salient features.

13-1. Electrical Properties of Cathode-ray Tubes.—Cathode-ray tubes are classified as “electrostatic” or “magnetic” in accordance with the method of deflection. An example of each is given in Fig. 13-1. The electron beam originates in a hot cathode at the end of the tube remote from the screen. In traversing the tube it is acted upon by a number of electrostatic, or electrostatic and magnetic, fields which serve to control and focus it.

The beam is collimated and controlled in intensity by the “control grid,” a pierced diaphragm immediately in front of the flat oxide-coated cathode. In most American tubes a “second grid” at a positive potential of a few hundred volts serves to attract the electrons from the space charge in much the same way as does the screen grid in an ordinary tetrode or pentode. The electrons are given their final high velocity by a potential difference of a few thousand volts maintained between the cathode and an anode formed by a conducting coating of carbon (Aquadag) on the inner surface of the glass at the screen end of the tube.

Focusing can be either electrostatic or magnetic. As indicated in Fig. 13-1 the same type of field is usually used for focusing and deflection in a given tube.¹

¹ Occasionally electrostatic focusing is combined with magnetic deflection; such tubes have not come into wide use in this country.

In electrostatically focused tubes a "focus electrode" (first anode) between the second grid and the high voltage anode is maintained at an intermediate potential. The various electrodes are so shaped that strong electron lens actions are produced by the fields at the two gaps. Adjustment of focus is made by controlling the potential of the focus electrode.

Magnetic focusing is accomplished by means of a longitudinal magnetic field of circular symmetry which increases in intensity from the center to the edge of the tube and thus has radial components that help provide the focusing action. External coils or permanent magnets are used to produce this field.

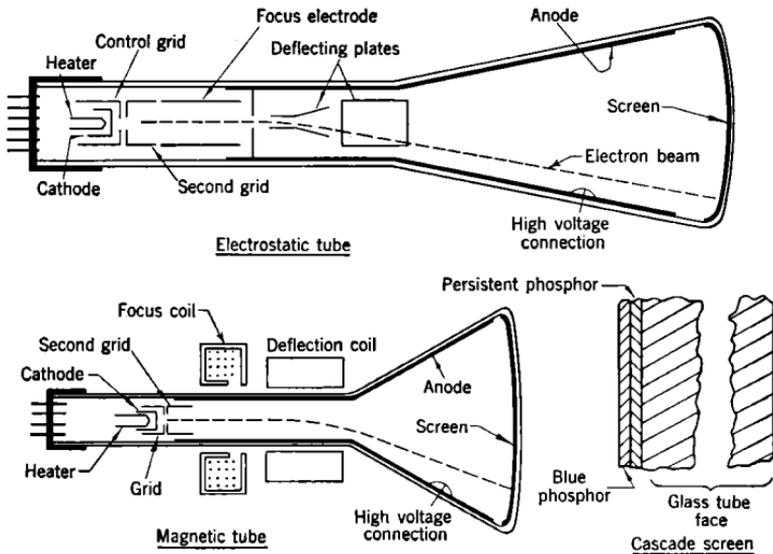


FIG. 13-1.—Elements of cathode-ray tube.

Electrostatic deflection is accomplished by passing the beam between two orthogonal pairs of deflecting electrodes or plates. The deflection due to each pair is accurately proportional to the potential difference between its members. The individual deflections due to the two pairs of plates add vectorially. The deflection sensitivity at the screen depends upon the geometry of the deflecting plates themselves, their distance from the screen, and the velocity of the electrons as they pass the plates. The two pairs are similar in geometry, but, since they are at different distances from the screen, their deflection sensitivities are unequal.

In the electrostatic tube shown, the electrons achieve their final velocity before reaching the deflecting plates. In some tubes designed for high-voltage use, deflection sensitivity is increased by applying

approximately half of the total voltage across a gap in the Aquadag at a point beyond the plates. The two parts are then spoken of as the "second anode" and "third anode."

Any variation of the mean potential between either pair of plates has a marked effect on the focus; consequently if good focus is important it is necessary to use push-pull deflection.¹ Furthermore, it is best that the mean potentials of the plates be close to that of the second anode. All

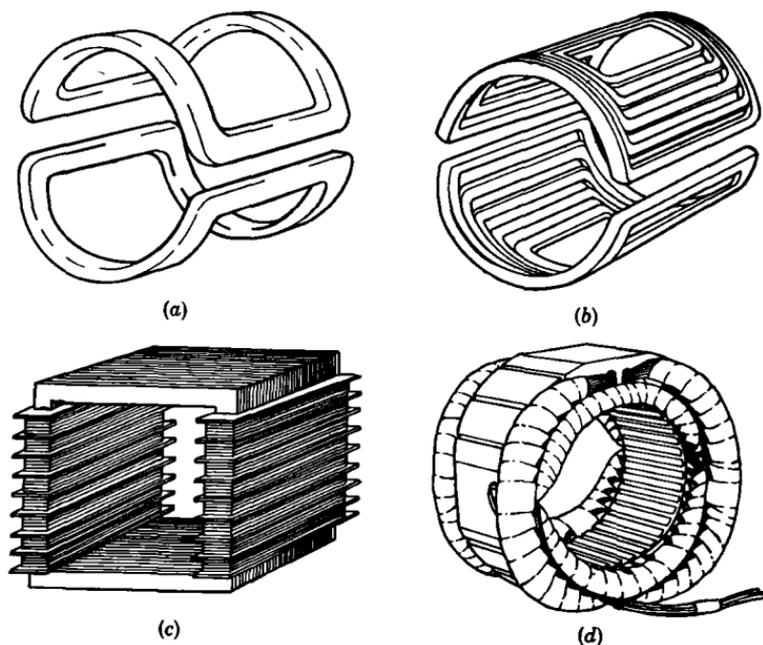


FIG. 13-2.—Deflection coils. (a) Air-core deflection coil with lumped winding; (b) air-core deflection coil with semidistributed winding; (c) iron-core deflection coil with pie windings; (d) motor-stator type of deflection coil.

these potentials are often made the same, but in many tubes it is profitable to provide an adjustment for making the first pair of plates slightly more negative.

Magnetic deflection is accomplished by passing currents of the desired waveforms through a coil or combination of coils surrounding the tube neck (Fig. 13-1). The deflection due to each coil is proportional to the current through it, and the combined deflection is the vector sum of the individual ones. A few of the coil geometries used are described below.

A single coil² is used in a rotating-coil PPI. Typical air-cored and

¹ For the sake of simplicity this is not usually indicated in block diagrams.

² A coil usually consists of at least two separate windings symmetrically placed with respect to the tube. They may be connected either in series or in parallel.

iron-cored coils for this purpose are sketched in Fig. 13-2 and pictured in Fig. 13-3. Care must be taken that surrounding fixed objects do not introduce a lack of rotational symmetry and that no magnetizable material is included in the rotating parts.

Multiple coils are used for many different purposes. Their fields must occupy the same axial region of the tube in order to avoid serious defocusing and distortion which otherwise take place when deflections

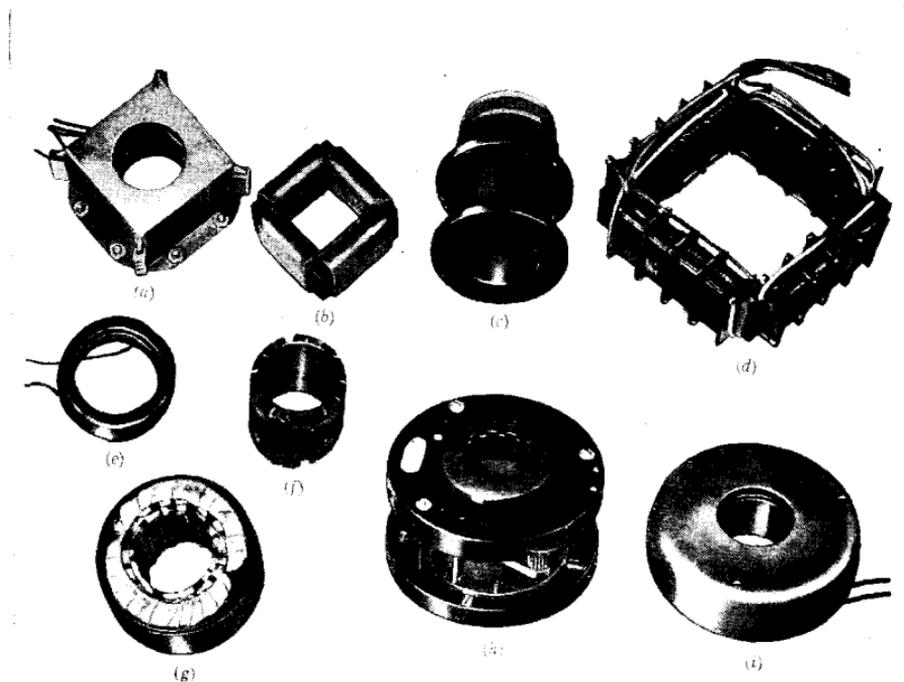


FIG. 13-3.—Deflection and focusing devices for magnetic CRT. (a) Square iron-core deflection coil in case. (b) Square iron-core deflection coil. (c) Air-core deflection coil for off-center PPI, used in conjunction with (d). (d) Off-centering coil. (e) Toroidal iron-core deflection coil. (f) Air-core deflection coil with distributed windings. (g) Motor-stator deflection coil. (h) Permanent-magnet focusing unit. (i) Focusing coil.

of appreciable magnitude are combined. In some applications requiring a fixed component of field, a permanent magnet is used in lieu of a coil.

Often coils are arranged orthogonally (Fig. 13-3) either to produce entirely independent deflections (as in a type B display), or to provide a single radial deflection (such as a range sweep) by adding its two rectangular components vectorially. The latter technique is used in the RHI, the fixed-coil PPI, and in other applications.

It is sometimes necessary to add dissimilar deflections in the same direction. Whenever practical, this is done by adding the various deflecting currents in the same coil. In some cases, however, this is

wasteful of power—for example, when a steady deflection is to be added to a high-frequency sweep that requires a coil of low inductance and therefore low sensitivity. It is then worth while to use two separate coils or a permanent magnet and a single coil. If two coils are used, they are usually placed together (on the same yoke when iron is used), but in extreme cases they must be separated to avoid interaction.

A rotating coil and a fixed coil can be used in combination—for example, to produce an off-center PPI. The fixed coil is placed outside of the rotating one, which must then contain no iron. The outer unit may contain a single coil that can be manually oriented, or it may have two orthogonal coils that can be separately excited. If the amount of off-centering is to remain fixed, a permanent-magnet arrangement can replace the outer coil.

In the specific design of the coils many factors must be considered. They should be as economical of over-all power as possible. The magnetic fields must be so shaped that they produce a linear deflection and cause no harmful effects on focus. When multiple coils are used, it must be possible to prevent harmful effects from interaction.

The current sensitivity of the coils should be made as high as practicable since this minimizes power losses. For a given geometry, the sensitivity is directly proportional to the number of turns, which should therefore be as high as practicable. However, for a given rate of deflection (i.e., a given rate of change of flux) the voltage induced in the coil is directly proportional to the number of turns. This usually sets an upper limit to the number of turns that can be used within the bounds of a reasonable power supply voltage.

These and other matters pertaining to coil design are discussed at length in Vol. 22 of the series.

13-2. Cathode-ray Tube Screens. *Phosphorescent Screens.*—The important characteristics of the screen are its decay properties, its efficiency, and the manner in which it integrates or “builds up” on repeated signals.

When scanning interrupts the picture, the screen must have sufficient persistence (“afterglow”) to permit observations and measurements on the echoes and in so far as possible to furnish a continuous picture. On the other hand, the image must not persist so long as to cause confusion on a changing picture. In terms of the scanning rate, three cases may be considered:

1. Cases in which little or no persistence is needed, either because the frame time is less than the retentivity time of the eye (about $\frac{1}{10}$ sec) or because observations are made in the absence of scanning. Screens incorporating the green willemite phosphor used in ordi-

nary oscilloscope tubes and known as type P-1 (phosphor number 1) are almost universally used in such applications. This material has an exponential decay with a time constant of a few milliseconds; it is extremely efficient in terms of light intensity.

- Applications in which persistence is needed to smooth out the effects of flicker. The best American screen for this purpose, designated P-12, utilizes a zinc-magnesium fluoride phosphor which has an exponential decay of time constant between 60 and 100 msec (Fig. 13-4). Because the response of the eye is logarithmic, this screen is satisfactory for scanning rates as slow as a few per second. On the other hand, the exponential rate of decrease prevents any long-time tailing-out and consequent blurring of a changing picture. The efficiency is somewhat less than that of the P-1 screen.

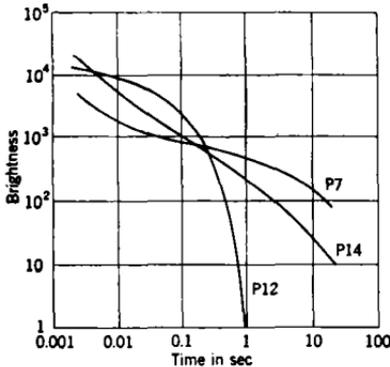


FIG. 13-4.—CRT screen characteristics.

- Applications in which the scanning is so slow that considerable persistence is needed to afford viewing time and to provide a composite picture. This classification covers scanning rates from a few per second up to one or two per minute; it includes the bulk of present-day radars with the exception of fire-control equipment.

Long-persistence phosphors have better buildup and decay properties under weak excitation than under strong excitation. Unfortunately the extremely thin layer penetrated by electrons of cathode-ray tube energies must be strongly excited to provide enough total light. On the other hand, if excitation by light could be used the excitation density could be kept very low because the screen would be excited throughout. This is accomplished by the simple but exceedingly clever expedient of covering the persistent screen with a second layer of a blue-emitting phosphor which undergoes primary excitation by the electrons. The blue light from this layer in turn excites the persistent screen. This process results in considerably lower over-all efficiencies than those of the single-layer screens.

Two varieties of such "cascade" screens are commercially available in this country. The P-14 is suitable for frame times up to a very few seconds, and the P-7 has a much longer persistence (Fig. 13-4). The persistent phosphors, which emit a predominantly yellow or orange light, are composed of copper-activated zinc-cadmium sulphides, the zinc-

cadmium ratio governing the decay characteristics. The blue layer is silver-activated zinc sulphide. Since its decay is very rapid, an orange filter is customarily used to remove its "flash," particularly at the faster scanning rates.

The decay of these cascade screens is an inverse power rather than an exponential function of the time, so that the disappearance of old signals is less clean-cut than with the P-1 and P-12 types. Unfortunately no phosphors with exponential decays of more than about 100-msec time constant have as yet been developed.

It is difficult to obtain sufficient light from intensity-modulated displays during scanning, since each point on the tube is excited only intermittently. The problem is especially acute on the slower scans, partly because of the long time between excitations and partly because the cascade screens are less efficient than those with less persistence.

In intensity-modulated displays the characteristics of the screen have important effects on the signal-to-noise discernibility. As has been pointed out in Chap. 2, the energy per pulse necessary for an echo to be just discernible is inversely proportional to the square root of the number of pulses included in the observation. From this standpoint the screen should enable the operator to integrate or average over the maximum number of pulse cycles consistent with other requirements; the intensity of each spot should represent the average of all the excitations received over a very long time in the past. The limits within which this can be accomplished are set by the achievable properties of the screen, and by the degree to which past information can be retained without causing confusion as the picture changes.

The screen properties of importance in this connection are the type of decay and the manner in which the light intensity "builds up" under successive excitations. To examine their effects, consider first a scan which is so slow that either the limitations of achievable persistence or the requirements of freedom from display confusion due to target motion prevent appreciable storage of information from one scan to the next. In such a case, the averaging must be done over a single pulse group. With modern narrow antenna beams and customary scanning rates the time occupied by this group is always short compared with the total scanning time and achievable decay times. The screen chosen should have sufficient persistence so that there is no appreciable decay¹ during the process of scanning across the target and the entire echo arc is observable at one time. In order that the average intensity of this arc shall represent all of the data, it is essential that the screen integrate the effects of all the pulses that overlap on a given focal spot—that is, it should not

¹ Except for that involved in the disappearance of the so-called "flash," which occurs for a very short time interval during and immediately after each excitation.

saturate too quickly. The number of overlapping pulses (which is often large) is given for a PPI by the expression $NS/\omega r$ at the center of the arc, where N is the PRF, S is the spot size, ω is the angular velocity of the scan, and r is the distance from the origin of the display to the echo spot. Thus for a 30-sec scan and a PRF of 400 (see Sec. 15-10) there are about ten overlapping pulses at "ranges" corresponding to the radius of a 7-in. tube, since $S/r \approx \frac{1}{200}$. The property of "building up" in intensity because of excitation by successive members of such a group of pulses is possessed to an adequate degree by all cascade screens, providing the excitation is not at too high a level.

At faster scanning rates it becomes possible to provide screens having carry-over from one scan to the next. Within the requirements set by clarity in a changing picture, the longest obtainable persistence should usually be used in order to average over the largest possible number of pulses. However, the permissible decay time is so short in the case of very rapid scans that the operator and not the screen has the longer memory. The persistence should then be determined entirely by the requirements of freedom from flicker on the one hand and freedom from blurring due to motion on the other.¹ If a set has several indicators involving different scale factors, their persistences should theoretically be graded, fast screens being used on the expanded displays where the picture changes rapidly and slower ones on those displays covering large areas. Fortunately, it is usually the latter on which the signal-to-noise discernibility is of most importance, since the expanded displays are usually confined to near-by regions.

Providing the scanning is not too slow (for example, if it is approximately one scan per second) certain types of cascade screens will, if initially unexcited, display more than twice as much intensity after two scans as after one, and so on, even though the intensity may have decayed manyfold in the time between scans. This property of "supernormal" buildup was at one time believed to be very desirable from the signal-discernibility standpoint, on the hypothesis that it would give the repeating signal an advantage over random noise. However, the property is most exaggerated on an initially unexcited screen, whereas in actual cases of successive scan integration the screen is initially excited from previous scans. Furthermore, it is not evident a priori that the

¹ When extended observations are to be made on a single target or region the display is sometimes "frozen"; that is, the target motion relative to the radar is compensated so that the picture remains stationary. It is then possible to use longer persistences without blurring. However, since observation of the frozen display is usually part of the "tracking" operation which controls the removal of the motion, the persistence must not be so long that it reduces the ease of detecting small changes in target position.

readjustment in the weighting of the successive pulse groups introduced by this phenomenon is desirable. In any case, careful observation has not detected any appreciable advantage for screens with very high supernormal buildup. The property of supernormal buildup may be connected, however, with other desirable ones. Indeed, those screens with highest buildup are in general the ones that show the longest persistence. These questions are discussed in detail in Vol. 22 of this series.

"Dark Trace" Screens.—Certain normally white salts, notably the alkali halides such as KCl, have the property of darkening for a time at a point where they have been struck by an electron beam of sufficient energy.¹ This phenomenon has been made use of in one form of cathode-ray tube, known as the "skiatron" (screen type is P-10) used for projection purposes. By means of an intense external light source, opaque projection of the tube face magnified several diameters is possible. Although of value for this purpose, this type of screen has several disadvantages. Contrast is always low, particularly if the duty ratio of the pulses is low. Furthermore, such a screen has the unhappy property that signals tend to "burn in" with time, which is a definite handicap in a changing picture and is disastrous in the presence of certain forms of interference. Except for the burning-in tendency, the normal persistence is satisfactory through about the same scan intervals as is that of the P-7 screen.

13.3. The Selection of the Cathode-ray Tube.—Both electrostatic and magnetic tubes are available in various sizes and with various screens (Table 13-1). The most widely used wartime types are pictured in Fig. 13-5. The following intercomparison of existing designs of electrostatic and magnetic tubes can be made.

1. At the excitation levels necessary for intensity-modulated displays, magnetic tubes provide much better focus than do the electrostatic types. It is possible to resolve about 175 to 200 fairly intense spots along the radius of most magnetic tubes, whereas the figure for electrostatic tubes is more like 75 to 100. On the other hand, at the lower beam currents needed for deflection modulation, the electrostatic tubes perform very well.
2. Magnetic tubes are more costly to deflect than electrostatic tubes, particularly at high frequencies. No attempt is made to deflect magnetic tubes at video frequencies.
3. Electrostatic tubes are, in general, much longer than the equivalent magnetic tubes, but the size, weight, and power dissipation of the over-all equipment is greater with magnetic tubes, partly because of the weight of the focusing and deflecting mechanisms and partly

¹ See Chap. 18, Vol. 22 of the series.

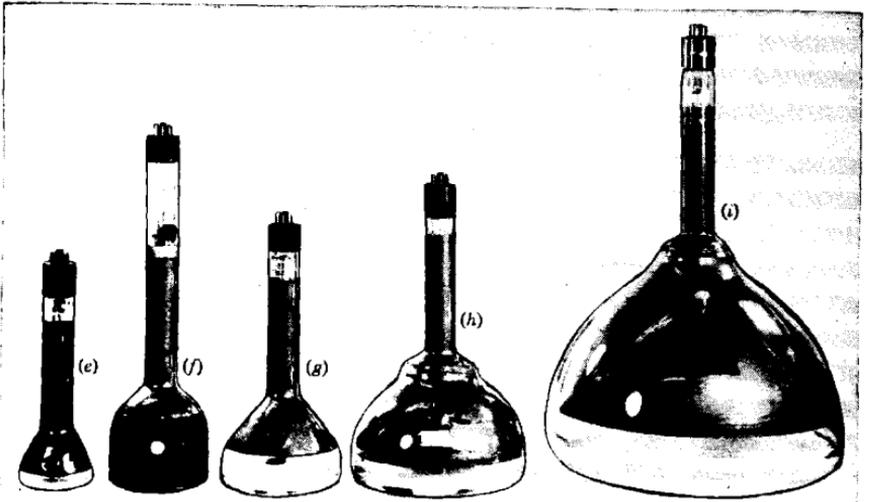
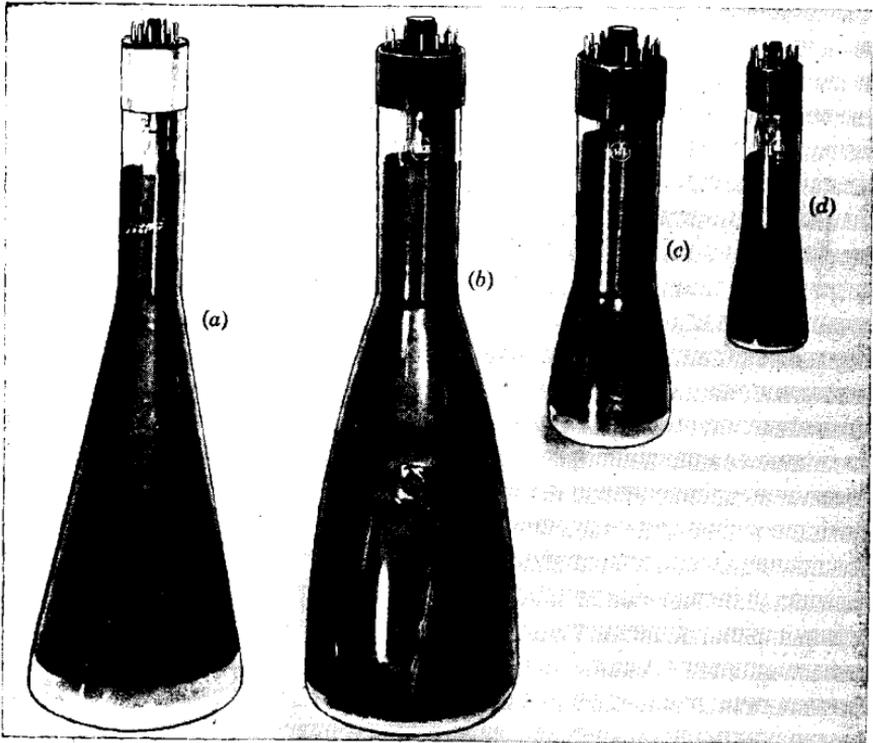


FIG. 13.5.—Representative electrostatic cathode-ray tubes: (a) 5BP1, (b) 5CP1, (c) 3JP1, (d) 2AP1. Representative magnetic cathode-ray tubes: (e) 3HP7, (f) 4AP10, (g) 5FP7, (h) 7BP7, (i) 12DP7. Most of these tubes are available with other screen types.

because of the greater power required for deflection. These differences are of significance only in airborne equipment.

4. The grid-modulation characteristic of most electrostatic tubes follows approximately a square law, whereas that of the guns used in most magnetic tubes is cubic. As compared with a linear response, these characteristics have the unfortunate property of lowering the dynamic range of usable echo intensities since they reduce the ratio between the upper limit of useful signal swing, set by the tendency to defocus, and the level below which the intensity is insufficient.

TABLE 13-1.—CATHODE-RAY TUBES COMMONLY USED FOR RADAR APPLICATIONS

Bulb number	Useful diameter, in.	Type of deflection	Max. accelerating voltage (nominal)		Specified maximum spot diameter, mm *	Commonly used screens
			Anode	Post-deflection anode		
2A	1.75	Electrostatic	1000	0.5 to 0.6	P-1
2B (New)	1.75	Electrostatic	2500	P-1
3B	2.75	Electrostatic	2000	0.55 to 0.75	P-1
3D (Central electrode)	2.75	Electrostatic	2000	0.6 to 1.0	P-1
3F	2.75	Electrostatic	2000	4000	P-7
3J	2.75	Electrostatic	2000	4000	0.75 to 0.9	P-1, P-7
3H	2.5	Magnetic	5000	0.5 to 0.6	P-7, P-12, P-14
4A	3.38	Magnetic	9000	0.3	P-10
5B	4.5	Electrostatic	2000	0.6 to 1.0	P-1
5C	4.5	Electrostatic	2000	4000	0.6 to 0.9	P-1, P-7, P-12, P-14
5F	4.25	Magnetic	7000	0.5 to 0.6	P-7, P-12, P-14
5L	4.5	Electrostatic	2000	4000	P-1, P-7
7B	6.0	Magnetic	7000	0.75 to 0.85	P-7, P-12, P-14
9G	7.62	Magnetic	7000	1.0 to 1.2	P-7, P-12, P-14
12G	10.0	Electrostatic	4000	6000	1.2 to 1.8	P-1, P-7
12D	10.0	Magnetic	7000	1.35 to 1.5	P-7

* Most magnetic tubes actually give from 0.5 to 0.7 of this maximum spot size. Electrostatic tubes are in general near the maximum.

As a result of these factors, electrostatic tubes are invariably used for deflection-modulated indicators, but, except in cases of extreme weight limitation, magnetic tubes are used for most intensity-modulated displays.

The size of the tube selected depends upon the particular application. The relative *resolution* is fairly independent of the size, so that from this standpoint alone there is little value in increasing the screen diameter

beyond the point where the resolving power of the eye ceases to be the limiting factor (5- to 7-in. magnetic tubes, 3- to 5-in. electrostatic tubes). On the other hand, if geometrical measurements or estimates are to be made, or if plotting is to be done, it is desirable to have as much *dispersion* as possible; for such use the larger sizes are definitely preferable. This requirement must be balanced against the available space, particularly in airborne equipment.

For many purposes even the largest available tubes (12-in. diameter) are still too small, and it is desirable to provide an enlarged presentation. Since at most scanning rates phosphorescent screens do not provide enough light for projection, it is necessary to devise special methods. Of those used to date, one involves the projection of a photograph of the PPI or other scope, which is developed and projected in a few seconds by special techniques; the other involves opaque projection from the skiatron screen (see Sec. 7-3).

COORDINATION WITH THE SCANNER

Under various circumstances several different sorts of electrical information must be delivered from the scanner to the indicator. (Some of these have already been illustrated in Sec. 12-1.)

1. Signals capable of controlling an electromechanical repeater. The repeater can be used to position a cathode-ray-tube coil or to provide a dummy scanner shaft to which the final data transmitters are attached.
2. Slowly varying voltage proportional to the scan angle, to be used in the cartesian display of angle as in type B and type C displays. The voltages can be obtained directly from a potentiometer of proper characteristics or from the envelope of a carrier which has been properly modulated. In many cases where the angle displayed is small, $\sin \theta$ can be substituted for θ .
3. Signals produced by passing a range sweep voltage (or current) through a resolver to produce an electronic PPI or an RHI.
4. Slowly varying voltages proportional respectively to the sine and cosine of the scanner angle, which are used to control sawtooth generators in such a way that they produce sweep components equivalent to (3). This is spoken of as "pre-time-base" resolution. The voltages themselves can be transmitted by a potentiometer, or a modulated carrier can be used as in (2). The approximations $\sin \theta \approx \theta$ and $\cos \theta \approx 1$ are often used.

13-4. Angle-data Transmitters.—The devices that provide the scanner data are known as "(angle) data transmitters." They in-

clude potentiometers, variable transformers, variable condensers, and generators.

Potentiometers.—Potentiometers are principally used to provide voltages whose only frequency components are those resulting from the scanning. Because even the best potentiometers have a certain amount of brush "jitter" which can be removed only by filtering, they are not very satisfactory for resolving range sweeps. A wide variety of linear potentiometers differing in accuracy, size, ruggedness, and so on have been developed specifically for radar use (see Vol. 17).¹ For many purposes they provide the simplest method of data transmission and often they are more effective than any other device. A difficulty arises, however, in those cases where it is necessary to shift the sector under view. This can be done within limits by adding a fixed voltage to the circuit, but over any extended angle it is necessary to introduce the shift mechanically by using a differential gear or by rotating the body of the potentiometer. This requires either that the potentiometer be near the operator or that a remote mechanical control be provided. Since the second method is costly, some other data-transmission system is usually chosen in preference.

Several varieties of potentiometer have been made with sine or with sine and cosine characteristics, for use as resolvers (Vol. 17). These have been designed with great care and are fairly good at low turning rates and low signal frequencies. As data transmitters they are occasionally used for purposes of pre-time-base resolution on slowly scanning systems, or as a basis of information for computers.

Variable Transformers (Resolvers, Synchros, etc.).—Figure 13-6a illustrates the principle of certain variable transformers called "resolvers," "synchros," "selsyns," "autosyns," and so on, which are widely used as position-data transmitters. An iron-cored coil (rotor) of special shape is mounted on a freely turning axis inside a slotted-iron framework much like a motor stator. Two or three stator coils are symmetrically wound into the slots in such a way that the coupling of each with the rotor is proportional to the sine of the rotor angle measured with respect to a position of zero coupling. The device thus fulfills the requirements of a resolver for a-c signals. An important aspect of this process is that the polarity of the output signal at a given phase of the input signal reverses as the synchro passes through a null position for that particular secondary winding.

If the stators of an excited resolver, called the "transmitter," are loaded with the stators of a second, called the "repeater" or the "receiver," (Fig. 13-6b) the latter will experience currents producing a changing

¹ Many of these potentiometers are useful as control elements for purposes other than scanner-data transmission.

magnetic field in the same direction as the magnetic axis of the transmitter rotor coil. There will be induced in the receiver rotor a voltage proportional to $V \sin (\theta_1 - \theta_2)$ where θ_1 and θ_2 are the respective orientations of the two rotors. If the receiver is provided with a second rotor coil orthogonal to the first, it will experience a voltage proportional to $V \cos (\theta_1 - \theta_2)$. Thus the two devices together "resolve" the original voltage in terms of the difference of the two angles. A resolver with a multiple-phase rotor is known as a "differential" synchro or resolver. It is used in reducing data from relative to true bearing (one synchro

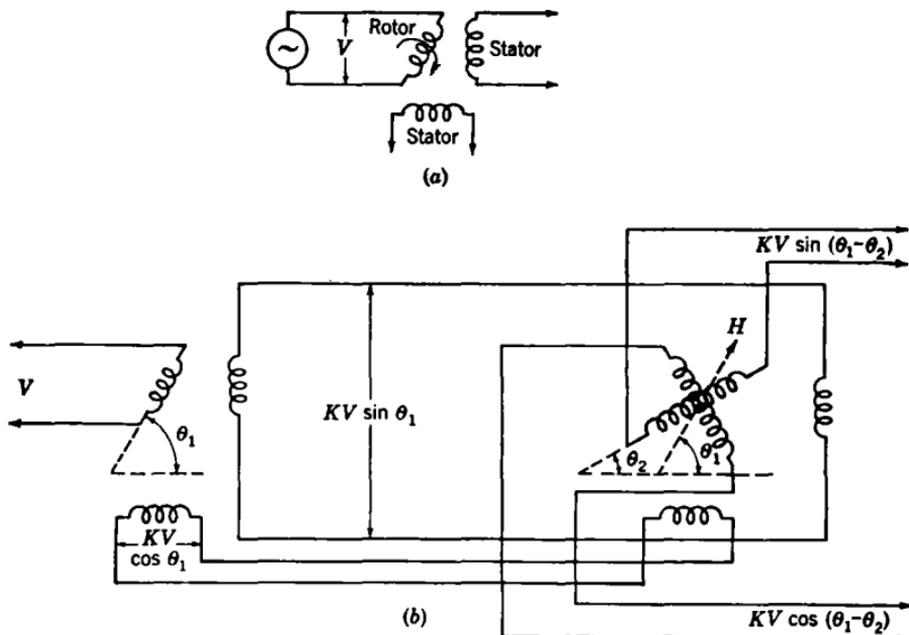


FIG. 13-6.—Two-phase synchros. (a) Two-phase synchro; (b) differential synchro resolver.

being driven by a compass), and in enabling an operator to shift the scanner reference angle for purposes of sector selection.

The 2-phase resolver has been used in nearly all practical applications of the resolved-sweep technique to "electronic" PPI's and RHI's (Fig. 12-3) and in many applications where pre-time-base resolution is applied to these displays.

For many purposes a 3-phase rather than a 2-phase¹ device is used. If equal loads are applied across the three phases the sum of currents in the three stators is zero. Thus, if this restriction is imposed, three brushes and three wires can be eliminated by making a *Y* connection. If a synchro with a 2-phase rotor is used as a receiver, its induced voltages

¹ In this sense the word "phase" refers to rotation of the rotor rather than to the internal phase of any signal being transmitted.

will be proportional to the sine and to the cosine of the difference of the two angles, as in the 2-phase case. If a 3-phase rotor is used, the machine becomes a "differential," which effectively subtracts the two angles in developing its output voltages.

Three-phase synchros are widely used in mechanical repeaters. As resolvers where the voltage itself is to be used, they have the disadvantage that they do not deliver rectangular "components." If only one component is needed, however, as in a type B or type C display (using the approximation $\sin \theta \approx \theta$) they are often used because of their somewhat higher accuracy and their greater availability.

Variable Condensers.—Certain types of rapid scanners (see Secs. 9-14 to 9-16) operate in such a way that, as a driving shaft rotates continuously, the antenna beam scans linearly across a given sector and then "snaps" back and repeats itself. Rotating-coil-indicator techniques cannot be used in such applications; hence the display must be synthesized electrically. The form of signal modulation provided by a synchro is suitable only for those sector widths for which the discontinuity in the scan can be achieved by switching between stator leads (that is, 90° , 180° , etc. for a 2-phase, 120° for a 3-phase synchro). Potentiometers cannot be used for rapid scans, since brush "chatter" becomes excessive and the life is exceedingly short. To fill this need, special variable condensers have been developed. An example of such a condenser and one type of circuit used is shown in Fig. 13-7. The condenser

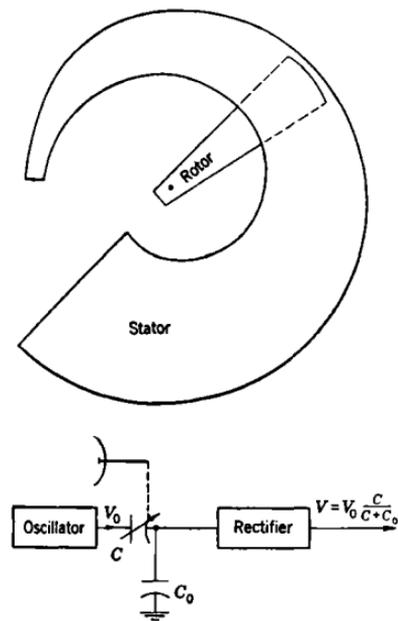


FIG. 13-7.—Special variable condenser.

plates are so shaped that the capacity increases as the rotor turns clockwise, producing the desired potential variation (usually linear with angle) across C_0 . Recycling takes place when the rotor passes through the sector in which there is no stator. The shaft is geared in such a way that the condenser rotates 360° during one scanning cycle. Since C is never zero, the output voltage has a constant term which must be removed in some way.

In addition to specially shaped condensers such as that of Fig. 13-7, more conventional ones are sometimes used with rapid scanners, in order

to permit the use of higher carrier frequencies than can readily be passed through a synchro.

Generators.—In some instances of conical or spiral scanning, part of the scanner data is provided by means of a d-c excited 2-phase generator geared directly to the scanner axis. In the case of conical scanning, the output voltages of this generator are direct measures of the sine and cosine of the phase of the scan. In the spiral scan, the voltages are modulated in terms of the nod angle, either by varying the generator field current by a potentiometer on the nod shaft, or by passing each sinusoidal voltage through such a potentiometer after generation. The signals thus produced can be used directly in the synthesis of type B and type C displays.

13-5. Electromechanical Repeaters.—Two types of electromechanical devices are used to repeat the motion of a rotating shaft at a remote point.

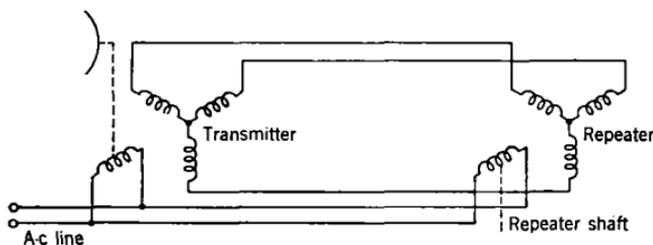


FIG. 13-8.—Synchro-driven repeater.

The Synchro-driven Repeater.—If the stators and rotors are attached respectively in parallel and the rotors are connected to an a-c power source (Fig. 13-8), the two rotors will tend to align themselves in the same direction. Any departure from this condition will result in circulating currents in the stators which will cause a motor action tending to produce alignment. If one rotor is driven, the other will follow, with only enough lag to furnish the necessary power. If the second rotor has little or no mechanical load, this lag will be very small at rotational speeds up to a few revolutions per second; if the load is appreciable, the lag may amount to a few degrees. In order to reduce the effect of lag, the system is often "geared up." For example, the transmitter synchro can be attached to a shaft rotating 10 times as fast as the scanner, and the receiving synchro can then drive a rotating PPI deflection coil through a 10-to-1 gear reduction. Under these conditions, the lag error can be kept to a small fraction of a degree. There is, however, a 10-fold uncertainty in the position of the deflection coil in our example, since any of 10 antenna positions looks the same to the receiver. The necessity for phasing the system manually each time it is turned on can be avoided

either by a system of cams and microswitches which disables the repeater when it is in any but the correct 36° sector, or in more elaborate ways.

Standard-design low-impedance synchros are used. Transmitter and repeater are identical except for the inclusion of a mechanical damper on the shaft of the repeater. The maximum speed of rotation of the synchros is about 400 rpm, restricting a 10-speed system to 40 rpm.

The principal disadvantage of using synchros to provide torque is that power must be transmitted by the primary synchro. This limits the number of repeaters which can be used. Further, in systems with more than one repeater, an error in any tends to throw the system off balance and to affect the accuracy of other repeaters.

The Servomechanism.—Scanner position can be repeated at a distance with an angular error as small as desired by means of a servomechanism.

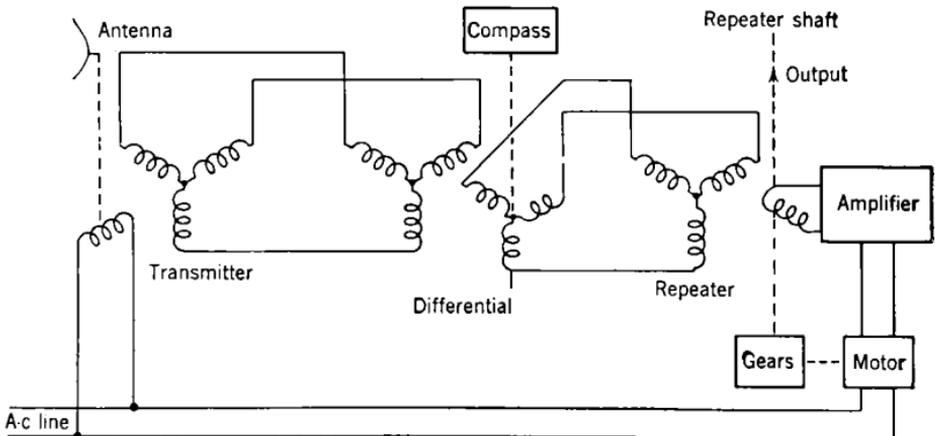


Fig. 13-9.—Servo system including compass-driven differential.

This term refers to a device that is arranged to reduce to zero an "error signal" present when misalignment between transmitter and repeater exists. Synchros can be used to generate such an error signal. The stators of the synchros are connected in parallel as before, with, perhaps, a differential synchro inserted between them (Fig. 13-9) to allow compass or other corrections. Only the transmitter rotor is excited from the line. If the two rotors are oriented at right angles to one another, no voltage is induced on the receiver rotor. Any departure from this orientation results in an error signal whose internal phase is opposite for errors of opposite sign. This error signal is amplified by vacuum-tube circuits and used to drive the motor that turns the load. Proper phasing is determined by reference to the line voltage, usually by using the latter to provide field current for the motor.

As in the double-synchro direct-drive system, "gearing up" is often used. Here, however, the proper rotational phase is usually selected by

a second servomechanism acting at "single speed." Some sort of switch, such as a gas-filled tube, serves to transfer control to the single-speed unit when the error becomes large enough to confuse the higher-speed servo.

For further information on the extensive subject of electromechanical repeaters, the reader is referred to Vol. 21 of this series.

BASIC ELECTRICAL CIRCUITS

The technique of producing radar displays involves the use of many unfamiliar or entirely new types of vacuum-tube circuits. A great many of these use a tube as a nonlinear element in such devices as electronic switches, pulse formers, generators of rectangular and other waveforms, etc. The following sections will describe briefly some of these techniques in order to provide a basis for understanding the methods of display production.

13-6. Amplifiers.—No attempt will be made here to describe the standard forms of amplifier met with in ordinary vacuum-tube circuit

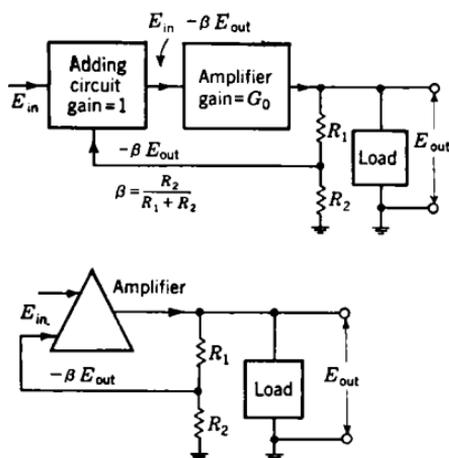


FIG. 13-10.—Negative feedback principle.

the amplifier is then given by $E_{in} - \beta E_{out}$. The output signal can be expressed as

$$E_{out} = (E_{in} - \beta E_{out})G_0,$$

where G_0 is the gain of the amplifier proper, in the absence of feedback. Solving for E_{out} ,

$$E_{out} = \frac{G_0 E_{in}}{\beta G_0 + 1};$$

¹ For a complete discussion of amplifiers see Vol. 18 of this series.

practice, but a number of special forms important in indicator design will be touched upon.¹ The important case of video amplifiers has already been met in Sec. 12-6.

Negative Feedback.—In a large number of instances negative feedback is necessary to insure linearity and adequate frequency response. The principles involved are illustrated in the upper diagram of Fig. 13-10. By any of several methods a fraction β of the output signal is subtracted from the input signal ahead of the amplifier. The effective signal to the amplifier is then given by $E_{in} - \beta E_{out}$. The output signal can be

expressed as

$$E_{out} = (E_{in} - \beta E_{out})G_0,$$

where G_0 is the gain of the amplifier proper, in the absence of feedback. Solving for E_{out} ,

$$E_{out} = \frac{G_0 E_{in}}{\beta G_0 + 1};$$

¹ For a complete discussion of amplifiers see Vol. 18 of this series.

whence the gain of the entire device is

$$G = \frac{E_{out}}{E_{in}} = \frac{G_0}{\beta G_0 + 1} = \frac{1}{\beta + \frac{1}{G_0}} \quad (1)$$

If β is large compared to $1/G_0$, variations in G_0 will have little effect on G . Care must be taken, in using large values of G_0 and β , that phase shifts in some of the frequency components do not result in positive feedback and hence distortion or even oscillation.

The more usual method of representing the circuit is shown in the lower diagram of Fig. 13-10, where the adding and amplifying are indicated together.

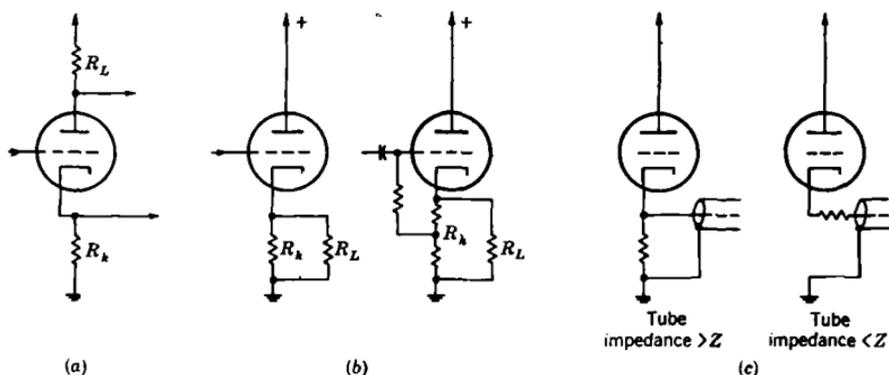


FIG. 13-11.—Cathode feedback. (a) General case; (b) cathode followers; (c) cable-matching.

Several types of negative-feedback amplifiers are used in indicator circuits.

Cathode Feedback.—A simple variety of negative-feedback circuit uses an unbypassed cathode resistor to provide the feedback voltage (Fig. 13-11a). The load may be in the plate, it may be the cathode resistor, or it may be divided between plate and cathode circuits to provide two output points.

For a plate load, the feedback voltage is actually proportional to the current through the load¹ rather than to the voltage across it; this is called "current feedback." If the load current is of interest (as in driving a deflection coil), this is just what is wanted. If, on the other hand, the voltage across the plate load is of interest, a cathode resistor gives the proper type of feedback only if the load is also purely resistive. If

¹ If a pentode or beam-power tube is used, the screen current passes through the cathode resistor but not through the load. In very exacting cases some account must be taken of this fact.

the plate load is reactive, the cathode impedance should be also, with the same sort of frequency response. For pure resistances the feedback ratio is $R_k/(R_L)$ and G_0 is $R_L/(R_k + R_L)$ multiplied by the gain that would obtain if both resistances were in the plate circuit.

In the cathode followers of Fig. 13-11b, the cathode circuit serves as the load and the plate is connected directly to B+, resulting in a negative-feedback amplifier of gain less than 1, characterized by great linearity, excellent frequency response, high input impedance, and low output impedance.

Since the entire output voltage appears across the cathode, the feedback ratio β is unity. Therefore G is always less than 1 [Eq. (1)], approaching, for large values of the cathode resistor, the value $\mu/(\mu + 1)$, where μ is the voltage amplification constant of the tube. Since $1/G_0$ is nearly always considerably less than β , the amplifier has a faithful response.

The internal impedance of the circuit driving the total load is given by $\frac{1}{g_m + 1/R_p}$, where g_m is the mutual conductance and R_p the plate resistance of the tube. This is a little less than $1/g_m$, and is therefore in the range from one to a few hundred ohms.

The input impedance of the tube itself (assuming no grid current) results from the sum of the capacities between the grid and the fixed elements (plate, screen, etc.), plus the grid-cathode capacity divided by $1 - G$, where G is the gain. This impedance is in parallel with the grid resistor and the wiring capacity. The effective impedance of the grid resistor connection can be increased by the method shown on the right in Fig. 13-11b, since the drop in the grid resistor is thereby reduced.

Among the applications of the cathode follower are the following:

1. As a low-impedance source to supply the considerable power required to drive such a load as a transformer or a deflection coil.
2. As a low-impedance source which will provide fairly uniform frequency response even when considerable capacity is associated with the load. The time constant and therefore the high-frequency response will be determined by the capacity in parallel with the internal impedance of the tube and resistance R_k . Since the tube impedance is small, good response can be obtained at frequencies of a few megacycles per second, even with fairly large load capacities. However, the internal impedance of the tube at any instant depends upon the actual tube constants at that instant. A steep negative wavefront may cause the grid to enter into a region of low g_m or to cut off, if the current flowing is too small to discharge the condenser with sufficient rapidity. The current can be increased

by using a smaller value of R_k , which sacrifices gain, or by using a positive grid bias. The latter is usually to be preferred.

3. To match the impedance of a cable. In this use the cable sees R_k in parallel with the internal impedance of the circuit proper. Correct methods of matching are shown in Fig. 13-11c.

Another special case of Fig. 13-11a is one in which the plate and cathode resistors are made equal, thus providing equal signals of both polarities from a unipolar input signal. Since the feedback ratio is unity for each output signal, this "split load" or "phase-splitting" amplifier has many of the desirable properties of the cathode follower, but the single-sided gain is always less and the internal impedance is greater.

Amplifiers for Deflection-coil Currents.—Special problems are involved in producing rapidly changing deflection currents such as those involved in range sweeps, since the voltage across the coil may become very large. For a linearly increasing current (as in a range sweep) this voltage is given by

$$L \frac{di}{dt} + Ri = L \frac{di}{dt} + Rat,$$

where a is a constant. Since di/dt is also a constant the waveform consists of a step plus a linear increase. The sweep is usually produced by an *increasing* current since, except in special cases, this results in the minimum average current. As a result, the coil voltage drop during the sweep results in a decrease in the plate potential of the driver tube. The drop across the coil reaches its greatest value at the end of the fastest sweep, and the power-supply potential must be designed to accommodate this case. Such a sweep need not dissipate much power since the average current is low, but unfortunately the same power supply is used on high-duty-ratio sweeps where the average current is high. The total power dissipation due to the sweep circuit can be minimized by using a large number of turns on the deflecting coil and a correspondingly high supply voltage, since this reduces the losses in the driver tube and in the power supply itself. Usually, however, other circuits derive their power from the same supply, and a high voltage results in unnecessary dissipation in them. In consequence, the system is often designed in such a way that the plate of the driver tube operates on the margin of insufficient voltage on fast sweeps. Even if this were not the case, the amplifier would tend to discriminate against the higher frequencies because of the higher impedance offered by the coils to these frequencies.

As a result of these factors good range sweeps can be obtained only with the use of greater negative feedback than would be necessary merely to correct for nonlinearities in the tube in a normal application.

Special amplifiers satisfying these requirements will be described later in connection with specific applications. Many of these employ multi-stage feedback.

13-7. The Generation of Rectangular Waveforms.—A rectangular waveform is one whose outstanding characteristics are steep, alternately positive and negative wavefronts with a space between (for example, Fig. 13-12). The term will include waveforms whose “tops” and “bottoms” are not completely flat.

In some applications only the timing of one or both wavefronts matters, the particular shape of the whole wave not being of interest;

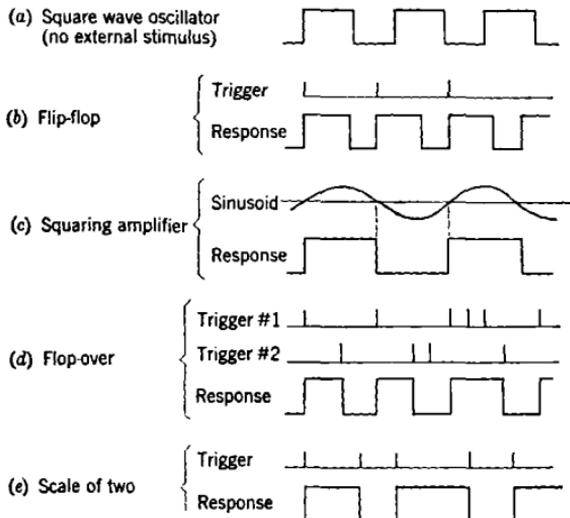


FIG. 13-12.—Response of various waveform generators.

in others the shape is also important. Among the latter are such applications as the intensifying of cathode-ray tubes and the operating of various electronic switches. Rectangular waves can be used for timing purposes—for example, to initiate the generation of sharp pulses or otherwise to initiate some particular event at the instant of occurrence of one of the wavefronts.

The circuits that produce such waveforms can be divided into three categories:

1. Those in which the waveform is produced without external stimulus—that is, free-running oscillators (Fig. 13-13).
2. Those in which an externally induced departure from a stable state with production of a wavefront is later followed by a spontaneous return which produces the wavefront of opposite polarity (Fig. 13-14). Such a device is called a “single stroke” generator or

“flip-flop.” It has wide application both in cases where the waveform is directly used and in timing or “delay” circuits.

3. Those in which both wavefronts are externally induced. In this class are the “squaring amplifier,” which produces square waves from sinusoids or other waveforms, and various triggered devices. The latter can be divided into two classes: (a) the “flopover” or “lockover,” in which an external signal induces a change from one stable state to another, and a reverse signal or one from a second source reverses the operation (Fig. 13-15a); and (b) “scale-of-two” circuits, in which successive triggers from a single source induce alternate transitions between two stable states (Fig. 13-15b). The name “scale-of-two” arises from the fact that if sharp pulses are derived from those wavefronts of one polarity their number will be half that of the original trigger pulses.

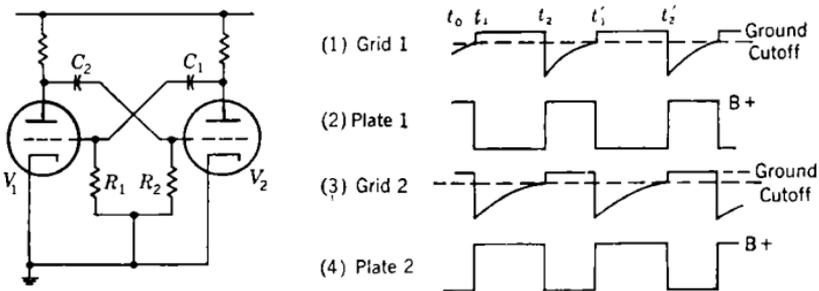


FIG. 13-13.—Eccles-Jordan multivibrator.

Aside from the squaring amplifier, which will not be discussed here, there are three principal forms of rectangular-wave generators.

The Eccles-Jordan Circuit.—Figures 13-13 to 13-15 illustrate a group of two-tube circuits in which the sharp transitions are produced by positive feedback from each plate to the alternate grid. All are based upon a circuit of Eccles and Jordan.¹

Figure 13-13 illustrates the free-running multivibrator, which is a form of relaxation oscillator. In order to understand its action, assume an initial condition (for example, at t_0 on the waveform diagram) with the grid of V_1 beyond cutoff and that of V_2 at cathode potential. Then V_2 is temporarily quiescent with its plate at a low value; the plate of V_1 is at $B+$, and grid g_1 is rising exponentially toward bias potential as C_1 is discharged through R_1 . At time t_1 , g_1 reaches the cutoff point and V_1 starts to amplify. The amplified signal is passed to V_2 where it is further amplified and fed back to V_1 . This regenerative action quickly lifts g_1 to the grid-current point and drives g_2 far past cutoff so that the original condition is reversed. The plate of V_1 is down and that of V_2 is at $B+$.

¹ W. H. Eccles and F. W. Jordan, *Radio Rev.*, 1, 143 (1919).

Grid g_2 now rises, until at t_2 it reaches cutoff and regeneration occurs in the reverse direction, and so on. Thus two square waves of opposite phase are available from the two plates. For given tube types and supply potentials, the time interval $t_1 - t_2$ is primarily determined by R_2C_2 , and the time interval $t_2 - t_1'$ by R_1C_1 ; these furnish convenient constants by means of which control of the time intervals can be exerted.¹ The multivibrator can be used as a completely free oscillator, or it can be synchronized by supplying it with a sine wave or a pulse train at a higher frequency than that of the natural oscillation.

Figure 13-14 shows a single-stroke or "flip-flop" multivibrator. The resistors and the bias voltage of the grid of V_2 are so chosen that when V_1 is fully conducting, g_2 comes to equilibrium beyond cutoff; but when V_1 is cut off, g_2 is pulled hard against the cathode. If left to its

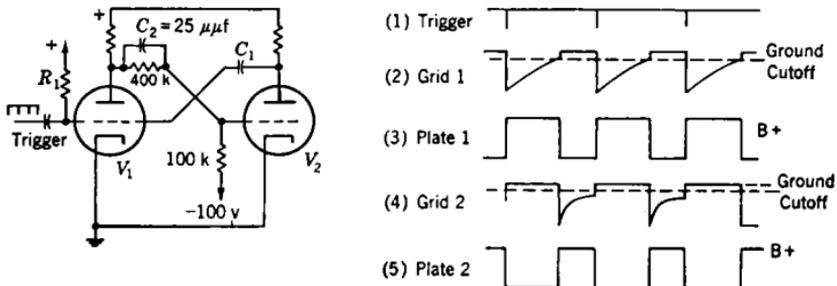


FIG. 13-14.—Eccles-Jordan flip-flop.

own devices, the circuit will come to equilibrium with V_1 on and V_2 off. When an impulse of the correct polarity is applied to g_1 or to g_2 , the regenerative action cuts off V_1 and turns V_2 full on. Grid g_1 then rises until the cutoff point is reached, when the regeneration reverses and the cycle is completed with the original condition restored. The length of the square pulse during the flip-flop action can be easily controlled by varying R_1C_1 or the potential applied to R_1 . This voltage is made positive to assure that g_1 will be rising sharply when the cutoff point is reached; this helps maintain a constant square-wave duration. Typical circuit constants are indicated in the grid circuit of V_2 . The condenser C_2 serves only to quicken the regenerative action and is not essential except for the highest speed operation.

The "flopover" and the "scale-of-two" (Fig. 13-15) are direct-coupled in both directions in such a way as to have two stable states, and will remain in either until disturbed in the proper manner. In the flopover,

¹ For simplicity it can be assumed, as is practically always the case, that the supply voltage is much higher than the cutoff voltage and that the plate resistors are much smaller than the grid resistors but large enough to absorb most of the potential drop when in series with a saturated tube.

triggering signals from different sources, applied to the opposite sides, result in opposite actions. Regardless of the order in which pulses are received from the two triggering sources, the flopover will respond alternately to triggers from the two.

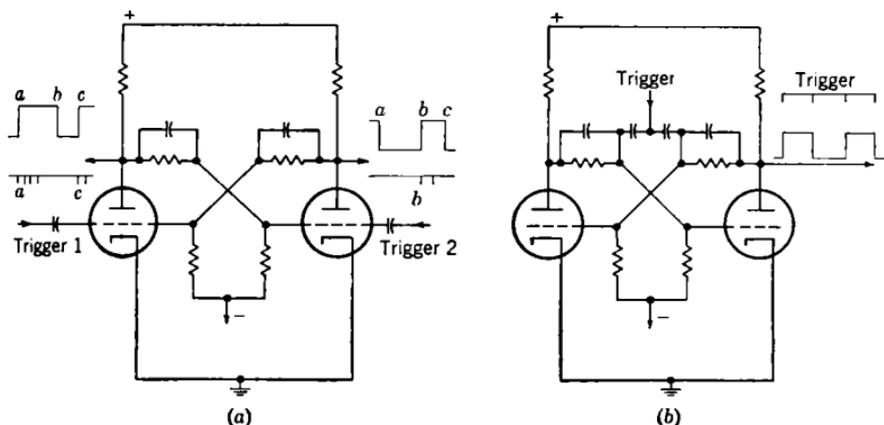


Fig. 13-15.—Double external triggering of Eccles-Jordan circuits. (a) Flopover; (b) scale-of-two.

In the scale-of-two, pulses from a single source are applied to both sides; whenever a pulse arrives the circuit changes from one stable state to the other. Thus two pulses are required to complete a full cycle. The output waveforms can be used directly for switching purposes, or they can serve as a means of generating pulses of half the frequency of the original ones.

In all the above circuits it is feasible to use self-biasing, but care must be taken that the bias voltage is independent of duty ratio. This can be assured by using a common self-bias resistor for both tubes, provided their loads are the same.

Cathode-coupled Multivibrator.—Figure 13-16 illustrates a different type of single-stroke multivibrator which is considerably better than the Eccles-Jordan variety for accurate timing purposes. The two cathodes

are coupled through a common resistor, and plate-grid coupling is used in one direction only. In the normal state, V_2 is conducting by virtue of the positive grid supply voltage. The resulting current produces sufficient potential drop across R_k to cut off V_1 . If, now, a positive trigger pulse of sufficient voltage is applied to g_1 (or a negative pulse to g_2) the amplified pulse is applied to V_2 through C , and V_2 starts to cut off.

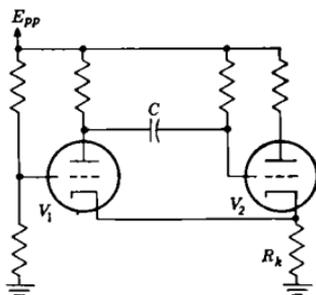


FIG. 13-16.—Cathode-coupled flip-flop circuit.

The drop in its current further reduces the bias on V_1 , and there is a violent regeneration which ends with V_1 full on and the grid of V_2 far past cutoff. The grid of V_2 then experiences an exponential recovery toward $B+$, and when the cutoff point is reached regeneration occurs in the opposite direction and restores the initial condition. The circuit has the marked advantage that the time for the flip-flop action is remarkably linear with the bias voltage of g_1 , so that by means of an accurate potentiometer controlling this voltage a linear timing circuit of rather low precision can be made. The circuit can be somewhat improved over that shown in Fig. 13-16 by the addition of a biased diode in the grid circuit of V_2 to determine the precise limit of the negative excursion of that grid.

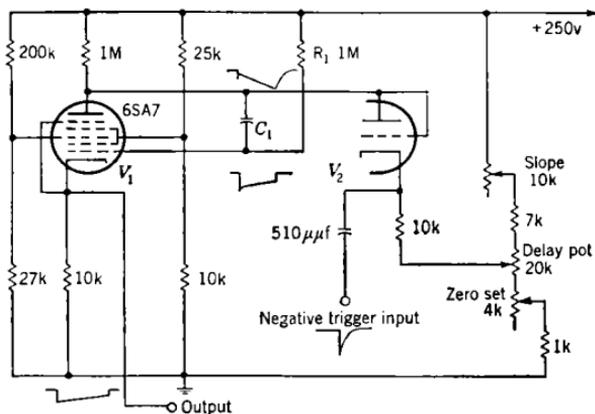


FIG. 13-17.—Phantastron delay circuit.

The cathode-coupled multivibrator can also be used in the form of a freely running oscillator, a flopover, or a scale-of-two, but the Eccles-Jordan form is most common for such purposes.

The Phantastron.—The phantastron of Fig. 13-17 is a flip-flop of a quite different type which serves as a timing circuit maintaining its calibration to about one per cent. In the normal condition, V_1 is quiescent, with the cathode sufficiently positive to cut off the second control grid so that all of the current goes to the screen. The plate potential is determined by the setting of the delay potentiometer. If a sufficiently strong negative trigger is supplied through V_2 to the first grid, the fall in the cathode potential turns on the second control grid. The establishment of plate current further reduces the potential of the first control grid so that the second grid is turned full on. The plate can now fall farther only as the potential of the first grid rises, by discharge of C_1 through R_1 . The condenser C_1 acts as a feedback condenser, assuring linearity of the tube response so that the plate falls linearly with the time. When the plate reaches the potential of the second grid it no

longer falls, the feedback action ceases, and the first grid rapidly increases in potential, pulling the cathode positive. The second grid is thus cut off and the original condition restored. Since the plate falls at a very linear rate to a fixed destination, the delay time is linear with the potential from which it started its downward journey. The critical parts of the circuit are R_1C_1 , the delay potentiometer, and the various supply voltages. A pentode with sharp suppressor cutoff, such as the miniature 6AS6, can be substituted for the more complicated multiple-grid tube with equally good results. The phantastron can also be used as a free-running relaxation oscillator.

13-8. The Generation of Sharp Pulses.—Sharp pulses needed for triggers, range markers, and other indicator uses can be generated in a number of ways. Figure 13-18 indicates three methods by which a step

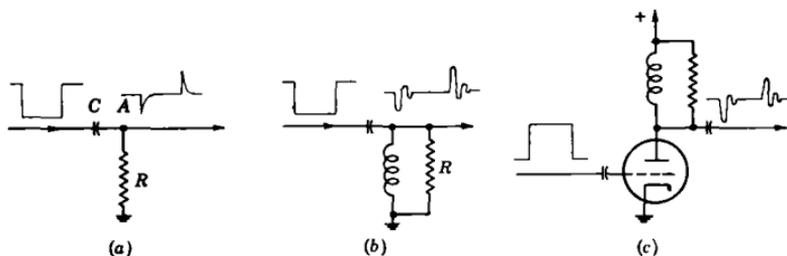


FIG. 13-18.—Simple pulse-generating circuits.

wavefront can be used for the purpose. In Fig. 13-18a the advent of the wavefront abruptly changes the potential of point A. Immediately thereafter C starts to discharge through R, and pulses of the shape shown result. This arrangement is usually spoken of as a “differentiating circuit.” The steepness of the front edge of the pulse is largely determined by that of the wavefront. In order to obtain a steep rear edge, RC should be small, but a point is reached for a given wavefront at which decreasing this product reduces the amplitude. The impedance of the driving source should be small compared to R.

In Fig. 13-18b, a steep wavefront sets up a shocked oscillation which is quickly damped out by R so that essentially only one pulse is produced. As in the case of the RC differentiator, this circuit requires a steep wave from a low-impedance source.

If the wavefront is from a high-impedance source, or if it has insufficient steepness or amplitude, the circuit of Fig. 13-18c is useful, particularly for positive pulses. Turning the tube on or off gives rise to a damped oscillation as in Fig. 13-18b.

A diode can be used in any of the above circuits to remove pulses of an unwanted polarity.

The device can be used as a "single stroke" pulse generator if the grid is biased beyond cutoff. The circuit is then quiescent until an external signal renders the grid conducting. The regenerative action takes place as before, except that after it is over the grid returns to a point beyond cutoff where it remains until a new signal is received. The circuit is often used in this form as a means of generating sharp pulses from poorer ones, from steep wavefronts (which are differentiated to form the triggering impulse), from sine waves, and so on.

13-9. Electronic Switches.—An electronic switch is a device that can change the parameters of the circuit in which it occurs. The principal functions performed by such switches are to control the absolute potential level of a point in the circuit, and to control the transmission or generation of signals by switching them on and off or by choosing between different signals. Activation of the switch can be controlled either by the signals themselves (just as a rectifier or an overload relay is controlled) or by a stimulus independent of the signals. In its action the switch can be either a series element analogous to a valve that opens and closes to block or transmit the signals, or a parallel element that, when closed, holds a point in the circuit in a quiescent state regardless of the presence of signals.

Clamps.—The name "clamp" is applied to a wide variety of electronic switches which, when closed, hold or clamp two circuit points together more or less rigidly. They differ from ordinary mechanical switches in that their impedance is often appreciable and frequently variable, and in that many clamps conduct current in only one direction. In fact, it is often the latter characteristic that enables the device to function as a switch.

The simple diode is a clamp that is closed when the anode attempts to be positive with respect to the cathode, but open when the reverse is true. The opening and closing may result from the waveform applied to one of the electrodes, as in a rectifier or a detector, or one of the electrodes may be changed in potential from time to time by a separate impulse. The use of diodes for switching purposes has been rendered more attractive by the advent of the germanium-crystal type of rectifier, which admirably replaces the vacuum-tube diode in many applications, particularly those involving 50 volts or less of back emf.

The D-c Restorer.—Figure 13-20 illustrates a common method of fixing the absolute potential taken by a point in the circuit during extreme excursions of the signal in one direction. In the absence of the diode, the average potential of point *A* must be V_0 , since *R* furnishes the only d-c connection. When the diode is placed in the circuit, it prevents *A* from swinging more negative than V_0 , so that V_0 may be thought of as a base with respect to which the entire waveform is positive. During each

positive excursion of A the resistance extracts a small charge, which is replaced by the diode on the next negative excursion.¹ The clamping can be done at the positive extreme by reversing the diode. If fidelity of response is of importance, the RC product must be sufficiently large to give proper low-frequency response.

The d-c restorer has many applications, only a few of which can be listed here. It is frequently used to insure that a grid to which unipolar

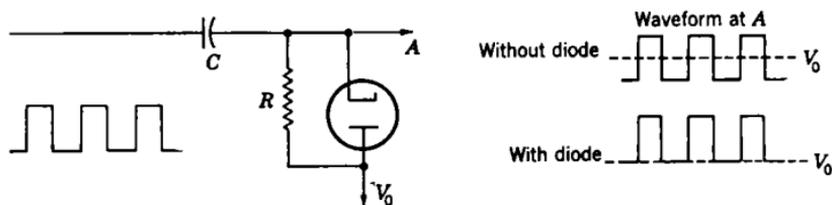


FIG. 13-20.—D-c restorer.

video signals are applied is operating in the correct voltage range. This is particularly important in the intensity modulation of a cathode-ray tube where the tube must be just cut off for zero signal amplitude. The d-c restorer can be used to fix the starting point of a sweep by clamping the deflecting plate of an electrostatic tube, or the grid of a direct-coupled sweep amplifier, between sweeps. In the above examples, the clamping fixes the point from which signals are measured. In others, the clamping fixes the level to which an electrode is switched "on" by a square wave, as for example when intensifying a cathode-ray tube.

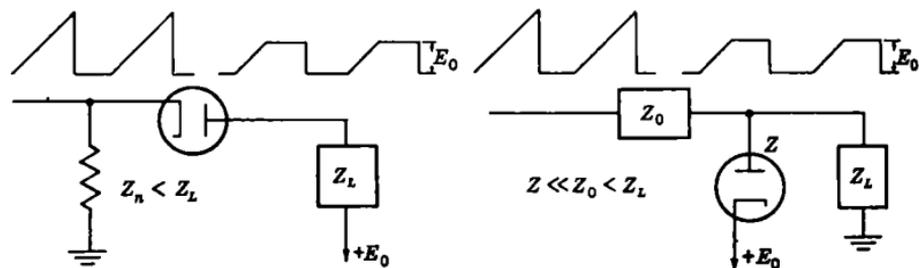


FIG. 13-21.—Diode limiters.

Biased Diodes.—Figure 13-21 illustrates two self-evident methods of limiting the excursion of a signal by clipping off the top at a level determined by a d-c potential difference applied to the circuit. The circuit of Fig. 13-22, on the other hand, transmits only that portion of the signal in which the voltage exceeds the applied potential, so that in effect the bottom is clipped. This arrangement is of considerable importance in time-measuring applications. If a very linear sawtooth is applied to the

¹ Sometimes the resistor is made slightly more negative than V_0 in order to hold A tightly against the diode during the clamping period.

input of the circuit, the time elapsing before an output signal appears is proportional to the bias applied to the diode, thus providing an accurate and easily controllable time delay (see Sec. 13-12).

Switched Clamps.—In a great many applications the clamp must be opened and closed over particular time intervals which may or may not be directly related to the signals. According to whether the clamp can conduct in one or both directions when closed it is classified as a one-way (“single-sided”) or a two-way (“double-sided”) clamp.

In the diagrams, E_0 is the potential of the clamping point, X is the point being clamped, A is the maximum signal amplitude, and T is the time interval over which the clamp is to be opened. If E_0 is not ground potential, the voltage

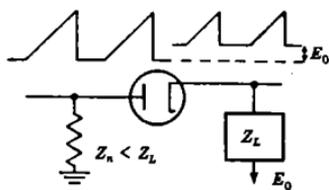


FIG. 13-22.—Bottom clipper (biased diode “pick-off” circuit).

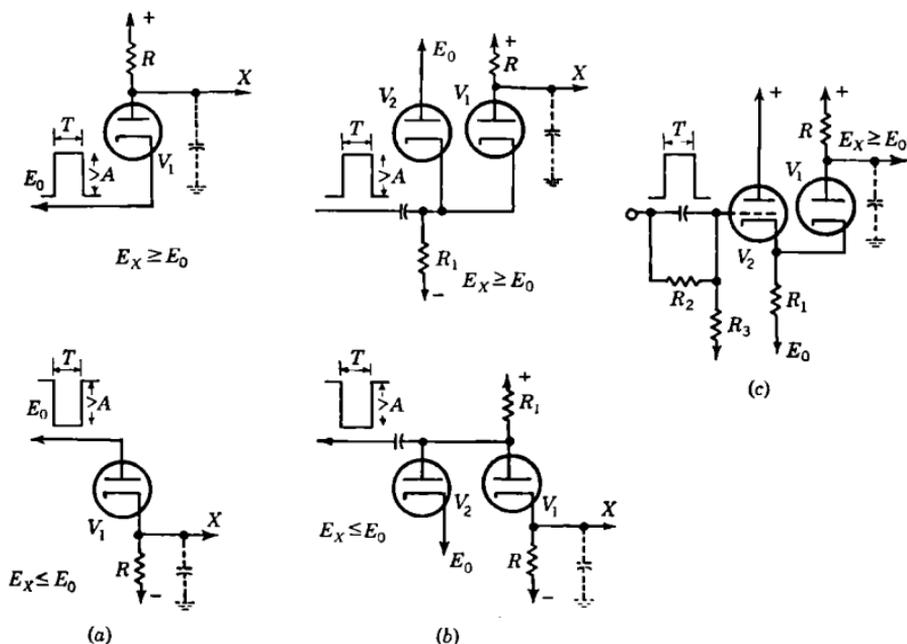


FIG. 13-23.—One-way switched-diode clamps. (a) Switched single diodes; (b) double-diode clamps; (c) diode switched by cathode follower.

X is point being switched.
 A is maximum amplitude of signal at X .
 E_0 is clamping potential.

T is length of open interval.
 R must be very high.
 Plus and minus signs are with respect to E_0 .

source supplying the clamping point must be “stiff” enough to furnish the currents drawn without appreciable potential change.

Among the simple switched clamps are the diodes of Fig. 13-23a, the

two diagrams indicating clamps of opposite polarities. They will, of course, close only if the potential of point X is in the specified direction from E_0 , and even when closed they will offer no impediment to potential changes in the opposite direction. The resistance R is included to counteract any slight leakage charges.

The requirements for the source of the switching waveform are somewhat rigid. It must provide a satisfactory value of E_0 and a square wave of sufficient amplitude. Furthermore, since its impedance adds to that of the clamp, this impedance must be sufficiently low to withstand any "pulling" by the circuit attached to X , and must be capable of discharging sufficiently rapidly any capacity associated with X .

Frequently a satisfactory square-wave source already exists in the equipment. If not, further complexities must be added to the clamp. Figure 13-23*b* shows a modification that can be used when the source has a

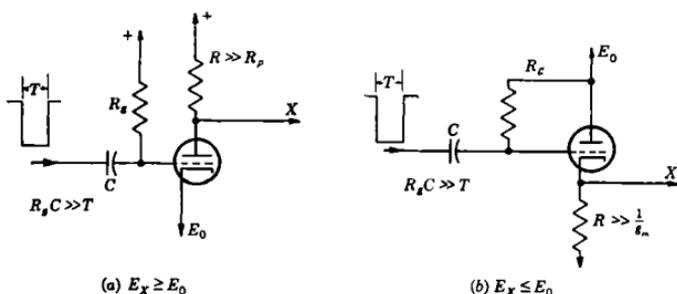


FIG. 13-24.—One-way triode clamps. (a) Positive clamp; (b) negative clamp.

sufficiently low impedance but an improper d-c level. The tube V_1 serves the same function as before and V_2 acts as a d-c restorer. The presence of the resistor R_1 increases the requirement for a low-impedance source.

If the waveform source has high impedance, it is necessary to introduce a cathode follower as shown in Fig. 13-23*c*. The clamping takes place through the diode and R_1 ; thus R_1 must not be too large. This in turn requires considerable current through V_2 during the clamping interval. The proper d-c level for the switching signal can usually be chosen by a proper choice of R_2 , R_3 , and the bias potential. If not, it may be necessary to use condenser coupling and a d-c restorer on the cathode-follower grid. The corresponding negative clamp is not used, since the circuit of Fig. 13-24*b* accomplishes the same results more simply.

Figure 13-24 illustrates one-way clamps using a single triode.¹ In both cases the clamping is done through the tube impedance as viewed from X . In the positive clamp this impedance is unfortunately rather large so that this clamp is not very "tight." The negative one, on the

¹ The well-known Rossi coincidence circuit is an example of a triode clamp.

other hand, is satisfactory in this respect. Care must be taken in the latter that "droop" in the switching-off pulse does not allow the grid to go too far positive when the clamp is closed, for this can produce overshoot through grid-cathode diode action. Square-wave droop can also cause overshoot in the positive clamp if the clamping is not tight.

The relative merits of the various single-sided clamps depend upon the uses to which they are to be put, the type of switching source available, and so on. If a proper driving source already exists, the simple diode (Fig. 13-23a) is cheapest and best. If not, the triode of Fig. 13-24b best combines simplicity and good characteristics among the negative

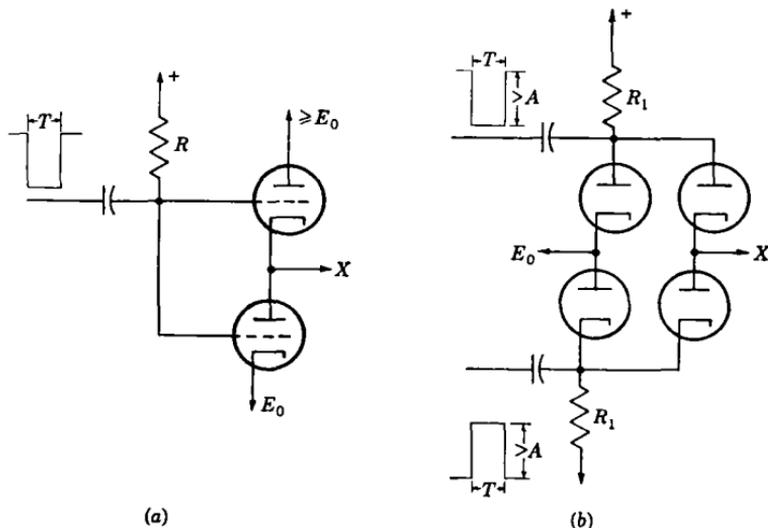


FIG. 13-25.—Two-way clamps. (a) Two-way double-triode clamp; (b) two-way four-diode clamp.

clamps, and Fig. 13-23c is the best and Fig. 13-24a is the simplest of the positive clamps. If the load impedance is high, the latter may be quite satisfactory.

Two-way clamps are usually made up of two one-way units of opposite polarities, as indicated in Fig. 13-25a and b. The requirements and shortcomings of the corresponding one-way clamps apply equally here. The "phase splitter" of Fig. 13-12 provides an excellent driving circuit for Fig. 13-25b if it is necessary to provide a special one. If a proper driving source is used, the four-diode clamp can be much "tighter" than the two triodes in the positive direction. Furthermore, the capacity coupling through the two triodes adds, whereas in the diode circuit the effects of the opposite waves tend to cancel. On the other hand, the double-triode circuit is much the cheaper, particularly if a special driving circuit must be provided for the diodes.

In certain applications where the clamp is used as a series element, it is essential that the current drawn from the reference point be as small as possible. Even such small currents as those drawn by the grid circuits of Fig. 13-25a may be objectionable. In the circuit of Fig. 13-26, the transformer-coupled grid circuits form closed loops which draw no current whatever from the reference point.

This circuit is widely used as a demodulator in cases where the phase of the carrier with respect to the reference signal is of importance. The waveform to be demodulated is applied to the input terminal of the circuit and the transformer is excited by an unmodulated wave which is exactly in or exactly out of phase with the carrier. Because of grid current, the two tubes will bias themselves to such a point that they are turned on only at the peak of the keying waveform; hence the device is a peak detector. The polarity of the rectified signal at the output terminal will depend upon which of the two phase relationships exists. This property is useful in such applications as the demodulation of a carrier wave modulated by a synchro. As pointed out in Sec. 13-4, the output signal at a given phase of the input signal has opposite polarity on either side of a null position of the synchro rotor; in other words, the carrier undergoes a 180° phase change as the synchro passes through a null. Thus, when the circuit of Fig. 13-26 is used as a demodulator, the output voltage is positive or negative depending upon whether the synchro rotor is oriented positively or negatively with respect to the null position.

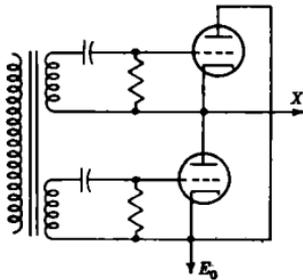


FIG. 13-26.—Two-way double-triode clamp with transformer-coupled switching.

The switching signal always “leaks” through the clamp tube to some extent by capacity coupling, particularly at the “off-point.” For this reason the switching signal should be no steeper or larger than necessary. Because of this, pentodes are sometimes used in such circuits as those of Fig. 13-24a and 13-25a. This increases the impedance of the positive clamp, but is only necessary when the load impedance is also high.

Gated Amplifiers.—An important form of electronic switch of the nonclamping variety is a “gated” amplifier tube, in which the switching signal is applied to an electrode in such a way that the electron current to the plate is interrupted and no signal can pass. Figure 13-27 illustrates various methods which are more or less self-explanatory.¹ All of these are widely used as pulse coincidence circuits² or to select the pulse or

¹ It is possible to combine two or more of these methods so that three or more electrodes have signals.

² In the case of pulse coincidence there is, of course, no distinction as to which electrode is being switched.

pulses occurring during a given time interval. The triode can be used only if both signals are limited, since otherwise it could be turned on by

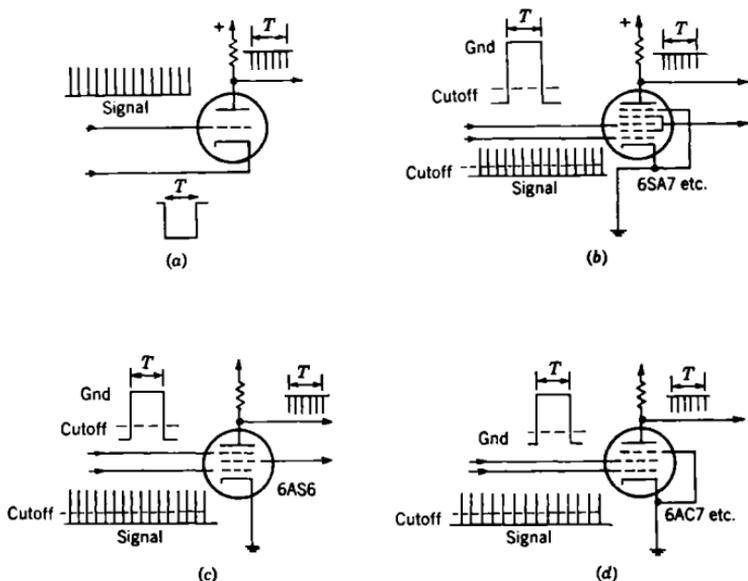


FIG. 13-27.—Gated amplifiers. (a) Switched triode; (b) multicontrol tube switched on second control grid; (c) suppressor switching; (d) screen switching.

one alone. The circuits of diagrams (b) and (c) of Fig. 13-27 require the least power input. The 6AS6, which has a sharp suppressor cutoff, is the most satisfactory tube unless considerable load current is needed. Since none of the tubes that can deliver large currents has either a second control electrode or a sharp suppressor cutoff, it is customary to use screen gating when the current requirements are high. In the case of video signals, the tube must have a very sharp cutoff on the signal grid; even then it is usually necessary to choose between an extreme bias, which sacrifices gain at low signal amplitudes, and a lesser one, which allows a "plateau" to be transmitted. The use of two tubes in the way described in the next paragraph is more satisfactory.

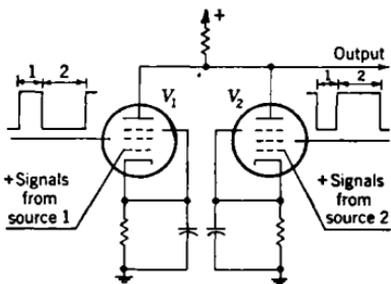


FIG. 13-28.—Switching of alternate signals.

Two such switches can be used in parallel to alternate signals from two sources, as shown in Fig. 13-28. In this case the simultaneous but opposite switching of the two tubes eliminates the plateau if the two cir-

cuits are properly matched, and the signal grids can therefore be in full operation at all times. Since it is essential that the proper d-c levels be maintained when the signal duty interval is large, d-c restorers must follow any a-c coupling. This circuit is often used as a single-channel switch, the signals being applied to one tube, and the other serving merely to cancel the plateau. The circuit may also be adapted to suppressor grid switching.

13-10. Sawtooth Generators.—The sawtooth waveforms used as a basis for range sweeps and in certain varieties of precision timing circuits are practically always generated by a wave-forming network that is switched on and off by some sort of clamp.

The basic action of all such devices can be illustrated by the simple circuit of Fig. 13-29. When S , a single clamp of proper polarity, is closed, an equilibrium state is reached in which X differs from ground potential

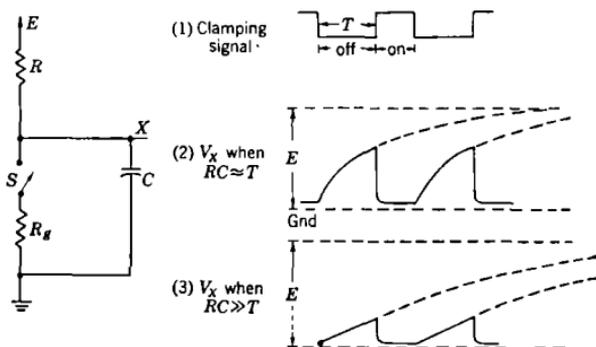


FIG. 13-29.—Basic sawtooth generator.

only by the small drop across R_g . If S is suddenly opened, the current is shifted from S to C , and the latter starts to charge exponentially at a rate determined by R , C , and E . The rising wave of the sawtooth is an exponential of time constant RC , as indicated in Waveform 2 of Fig. 13-29. If the sawtooth has an amplitude small compared to E —that is, if $RC \gg T$ —the rise is approximately linear (Waveform 3). Such sawtooth generators are widely used in cases where no great precision is required.

When various sweep speeds are to be used, it is necessary to change the rate of rise of the sawtooth. This can be done in discrete fashion by switching the condenser, the resistance, or, within limits, the charging voltage. Continuous variation is most satisfactorily accomplished by using a rheostat for R (unless the control is remote) or by varying E through a potentiometer.

Figure 13-30 illustrates both positive and negative sawtooth generators using single-triode clamps. If waves of both polarities are desired, the two circuits of this figure can be combined, or a single tube can be

used to perform both functions in a somewhat less satisfactory manner. More often, however, a single wave is generated and the second obtained from it by a phase-inverting amplifier.

More sophisticated sawtooth generators are necessary when precision is required. All of these endeavor to provide great linearity in one of the

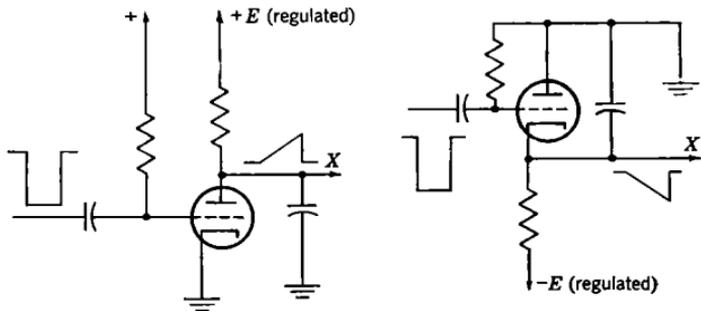


FIG. 13-30.—Sawtooth generators using triode clamps.

following ways: (1) by substituting a “constant current” device for R ; (2) by adding the condenser voltage to E , thereby keeping the current through R constant; (3) by including a high-gain negative-feedback amplifier, so that only an extremely small voltage need be developed across the condenser.

1. A pentode can be used as a constant-current device. This method has been little used in radar, however, partly because of general complications and partly because when a pentode is used in a positive sawtooth generator it is necessary to provide a floating screen-voltage supply.

2. Figure 13-31 illustrates the use of a cathode follower to keep the voltage drop across R nearly constant. That part above the switch indicates three alternatives. Neglecting the a-c drop across C_1 , the positive feedback raises point Y by an amount $E_x G$, where G is the gain of the cathode follower and E_x the potential of the point X . Thus the drop across R is given by

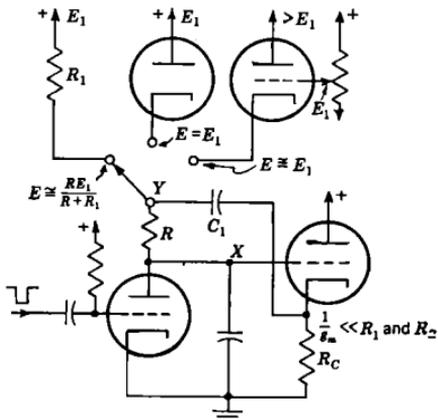


FIG. 13-31.—Bootstrap sawtooth generator. Switch indicates alternate methods.

$$E - E_x + E_x G = E - E_x(1 - G)$$

as compared to $E - E_x$ were the feedback not present; the effect of the amplifier is the same as that of using a supply voltage $E/(1 - G)$. For a triode, $1/(1 - G)$ may be as much as 20, and for a pentode as much as 200.

Use of the resistor R_1 is simplest, but it has important disadvantages. The charge leaking off C_1 through R and R_1 must be returned through them during recycling, and since the time constant of the circuit must be large compared to T this requires appreciable time. If the duty ratio is appreciable, a d-c shift results which depends on that ratio. This circuit can, therefore, be used only when the

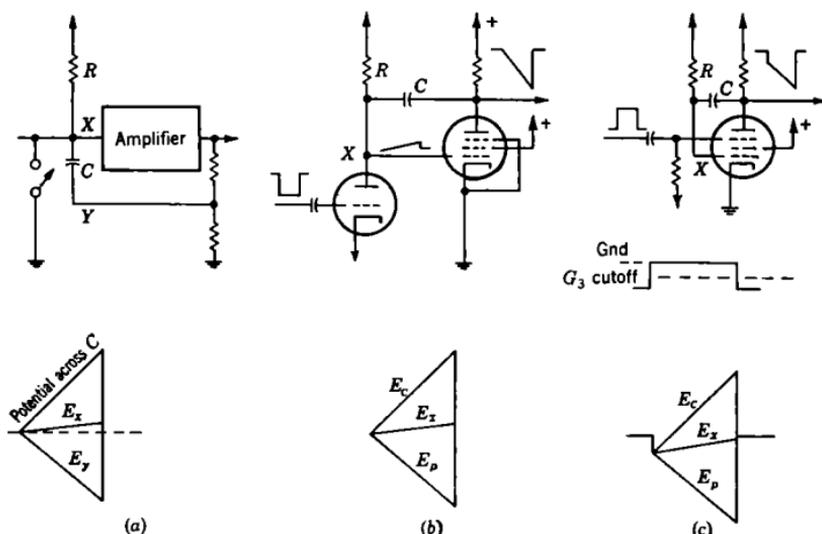


FIG. 13-32.—Sawtooth generators incorporating negative feedback.

ratio is very small or very constant. Furthermore, the presence of R_1 reduces the feedback gain. The diode avoids these difficulties, since it cuts off during the duty interval but closes to give a low impedance during recycling. The triode cuts off almost as quickly as the diode and provides a convenient method of providing voltage control since it draws no current from the variable potential source.

When extreme precision is required, the defects due to the charge drawn from C_1 and the less-than-unity gain of the cathode follower may be important. These errors can be completely eliminated by use of a more complicated amplifier of gain greater than 1, the excess gain compensating for the drop across C_1 . An alternative is to add to the sawtooth a compensating voltage which, should be approximately proportional to the integral of the sawtooth.

This method is used in the precision delay circuit of Fig. 13-36, described in the next section.

3. The most common method of using linear amplifier gain to keep the necessary signal voltage small is illustrated in the diagrams of Fig. 13-32. The amplified signal is applied to the "bottom" of the condenser, so that the potential of X changes only by the difference between the charge across the condenser and that fed back. Thus the change in potential across R is kept very small and the charging current is very constant. The most commonly used of such methods is the Miller "run-down" circuit of Fig. 13-32*b* and *c*. The two circuits shown differ only in the methods of switching. In both cases the amplifier consists of a single tube on whose grid the original signal is formed. The entire plate swing is applied to the condenser, so that the feedback ratio is 1. In the case of suppressor switching, all of the current goes to the screen between sweeps. The control grid is held against the cathode by virtue of its positive bias. When the suppressor is switched on, most of the current transfers to the plate, which experiences a negative surge that is passed on to the control grid. The sawtooth generation then begins as indicated in the waveform diagrams. The step at the beginning of the sawtooth is in some cases detrimental and in others useful.

INDICES

It is always desirable and usually mandatory to provide some form of quantitative indices or markers for the radar indicator. These may consist simply of a gridwork of lines or, when high accuracy is required, of one or more movable indices. In addition, it is often desirable to superpose some form of map or chart on a radar display, in order to provide accurate correlation with fixed echoes for navigational purposes or to show at a glance the geographical position of a ship or aircraft target. Markers can be provided either by placing a surface containing the marks as nearly as possible in optical superposition with the display, or by modulating the electron beam in such a way that the marks appear as part of the display itself.

Indices or charts ruled on a transparency over the tube face are the simplest of all to provide, but their use results in errors due to display inaccuracies, to parallax, and to faulty interpolation. Furthermore, if the origin of the display is to be moved, it is necessary to provide a corresponding motion of the reference system, which is usually cumbersome, or if only a few positions are involved, to furnish multiple sets of marks in such a way that no confusion results. The methods of optical superposition described in Sec. 7-3 largely eliminate parallax and are

helpful in the case of multiple indices, but their principal usefulness is in connection with plotting and the use of charts.

“Electronic” marks eliminate or reduce many of these difficulties. Parallax errors are entirely eliminated. Since the indices are generated by precision methods that are independent of the formation of the display geometry, they automatically fall in their proper place on the display regardless of where it may be centered or how much it may be deformed or distorted. In general, electronic methods of producing interpolation indices are far less cumbersome than mechanical ones. However, in the case of a slow scan the intermittent nature of the display makes it difficult to set an index on the echo, unless rather complicated switching methods are introduced to provide the indices at more frequent intervals.

Electronic indices are invariably used for range determinations, since they can be readily provided with a high degree of precision and since the radar data are inherently capable of providing great accuracy in range. In cases where the fundamental data are not so precise, the choice depends upon circumstances. Electronic marks are always used if the display is to be continuously movable in position or in size, but in many of the cases where only a few discrete changes are involved, external markers are sufficiently simpler to warrant their use.

The following two sections will describe methods of providing directly viewed electronic indices; auxiliary optical aids have already been described in Chap. 7.

13-11. Angle Indices.—Because of their simplicity, fixed angle indices ruled on a transparency are widely used in spite of the inaccuracies described above. The inherent data are usually far less accurate than those of range. Furthermore, in many applications, such as that of homing, the angle of interest changes slowly if at all and there is time for repeated observations. Parallax errors are made as small as possible by placing the scale very close to the tube face, and in many cases by ruling on both sides of a rather thick transparency in order to provide a line of sight.

Movable angle indices of the same type are little used except in the important case of the centered PPI, where the motion is one of simple rotation. The index or “cursor” may consist of a thin metal strip viewed edgewise or of a transparency with a ruled line or a thin slit milled through it. It is supported by a ring bearing larger than the tube face and is usually turned by means of a hand crank to which a data transmitter may be attached if desired. Readings are made from a circular scale at the edge of the tube.

Fixed Electronic Angle Indices.—An electronic angle index can be provided by brightening the cathode-ray tube for a few sweeps so that a

bright narrow line is produced at a given azimuth. The electrical impulse necessary to do this can be provided in any of several ways.

1. The simplest method involves the use of mechanical contactors on the scanner or on any mechanical repeater of the scanner motion. The contact is usually made by a microswitch operated by a cam on a rotating shaft. Since it is impossible to open and close the switch in a short enough angular interval, unless the scan is very slow, some device for producing a short pulse must be introduced. This can consist of a flip-flop triggered by the microswitch although, for medium or slow scans, it is possible to make use of the transit time of a mechanical relay.¹ In order to produce a set of markers, an equal number of cams on a one-speed shaft or

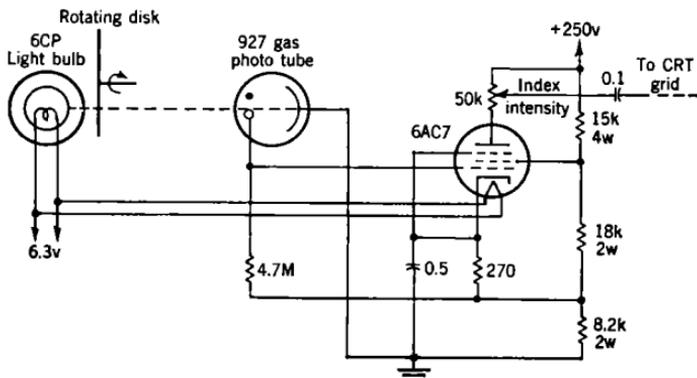


FIG. 13-33.—Photocell index circuit.

fewer cams on a higher-speed shaft can be used. The latter method is preferable because of the cleaner action,

2. Figure 13-33 illustrates a method for generating angle marks by using a photocell mounted behind a slotted rotating disk. Since the slots can be made extremely narrow and no inertia is involved, this method produces extremely clean-cut markers. The disk can turn at the scanner speed and have as many slots as there are markers, or it can be geared up and have fewer slots.
3. A method of using a carrier modulated by a synchro is illustrated in Fig. 13-34. (The figure actually illustrates a movable index. In the case of fixed indices, the signals can be taken from the first synchro and the second one omitted.) The modulation introduced by the rotating synchro is sinusoidal with scanner rotation; if the latter is uniform the modulation envelope obtained from the detector has this waveform. The sharp cusp occurring at the nulls

¹ See, for example, Fig. 4-2 of Vol. 22.

is the desired signal. Since there are two nulls per revolution, the synchro must be geared up by a ratio equal to half the number of markers. In the figure the null is used as a switching signal for an oscillator of frequency approximately one megacycle per second. The train of pulses is "rectified" by the cathode-ray tube, so that

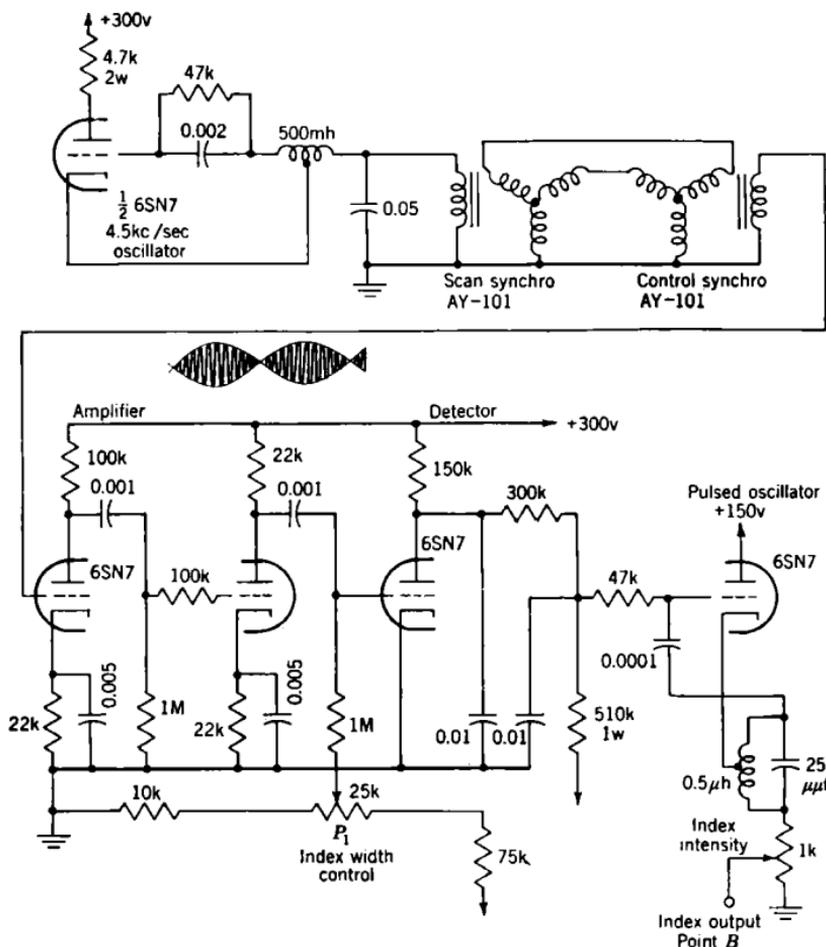


FIG. 13-34.—Synchro null indices.

the marker appears as a row of dots which merge except on fairly fast sweeps. An oscillation of this sort is preferable to a simple rectangular pulse if the signals are to be passed through mixing or other circuits whose low-frequency response is not as good as demanded by the very long (as much as $\frac{1}{10}$ sec) marker pulse.

Movable Electronic Angle Indices.—The methods of producing fixed electronic angle indices can be extended to provide movable ones. This can most easily be done, with the synchro method of Fig. 13-34, by using

the second synchro to introduce the differential setting. The two nulls will result in two signal markers 180° apart. If this is undesirable one of them can be removed by a method involving the use of a multiphase rotor in the second synchro. A second modulation envelope, 90° out of phase with the first, is obtained and used to blank one of the markers (see Fig. 4-5, Vol. 22).

The production of a movable index by the contactor or photocell method requires that the differential setting be made mechanically. This method is seldom used unless the indicator contains a mechanical repetition of the scanner motion.

Movable electronic indices appear on the indicator only when the cathode-ray beam is on a particular part of the tube face. To set an index on an echo, it is essential to be able to observe the results of moving the index. If the scan is rapid, the operator has frequent chances to do so, but if it is slow he cannot correct an error for several seconds. This is not too serious in the case of the movable range marker, because it appears at all azimuths and can be "steered" in the proper direction as it approaches the persistent echo remaining from the previous scan.

An angle index, on the other hand, if formed by the methods of the preceding paragraph will "flash up" only as the scanner passes by. The operator has no chance to make a correction before the next scan, and even then the correction must be an educated guess. It is very difficult to make accurate settings within the space of a few scans, especially if there is relative motion between radar set and target.¹

It is, therefore, essential on slow scans to use some other method. For all displays except rotating-coil PPI's, the so-called "substitution" method illustrated in Fig. 13-35 can be used. From time to time control of the display is switched from the regular angle data transmitter to one that can be set by the operator. During these intervals the direction taken by the range sweep is determined by the setting of the "substitution" transmitter and is thus under the control of the operator. The switching may be made to occur automatically on a certain fraction of the range sweep cycles equally distributed throughout the scan (which requires rapid electronic switching), or it may occur continuously over an

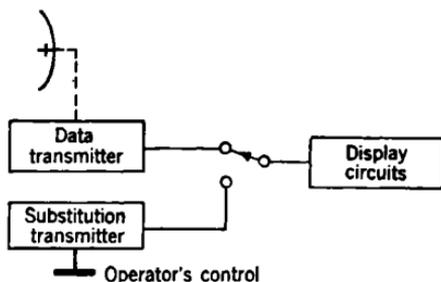


FIG. 13-35.—Substitution method for movable angle index.

¹ The problem is greatly simplified if aided tracking is used—that is, if some sort of computer moves the index in accordance with the relative motion, using knowledge acquired by making the best settings possible in the past.

appreciable sector of the scan remote from the one of interest, in which case the switching can be slower and in some cases mechanical. In the first case the index appears continuously, and in the second it appears during an appreciable fraction of the scan. In either event, the operator is able to make this setting while the index is present. The method is not applicable to a rotating-coil system because of its inertia. Fortunately, however, this type of display is itself exceedingly accurate in angle, so that the mechanical cursor can be used on the centered PPI display with little error except that due to parallax.

13-12. Range and Height Indices; Synchronization.—Since the providing of electronic range indices is frequently intimately associated with synchronization, these topics can best be discussed together. The discussion will be simplified if movable indices are described first.

Methods of Obtaining Movable Markers.—In general, four methods are used for obtaining a continuously movable marker. In increasing order of the precision that can be obtained they are: (1) the cathode-coupled multivibrator; (2) the phantastron; (3) a timing circuit based on a linear sawtooth; (4) the phase-shifting of a precision sinusoid.

The use of the multivibrator or the phantastron for time-delay purposes has been discussed in the sections describing those devices and no further description need be given.

The use of a sawtooth voltage wave for timing depends upon the fact that the time taken for such a wave to reach a given voltage is proportional to the voltage chosen. Figure 13-36 illustrates a circuit by means of which this principle can be very precisely applied. Tubes V_1 , V_2 , V_3 , and V_4 constitute a precision sawtooth generator. The drop across the condenser C and the lack of unity gain in the cathode follower (Sec. 13-10) are compensated by the network composed of C_2 and R_2 which integrates the sawtooth appearing on the cathode of V_3 and thus provides across C_2 a correction proportional to t^2 . The sawtooth waveform is applied to the plate of diode V_5 , whose cathode has a positive bias of an amount determined by the setting of the delay potentiometer. Because of this bias, no signal passes through the diode until the sawtooth has reached a definite amplitude, determined by the bias value. When the critical amplitude is reached (at time t_1 on the waveform diagram), the remainder of the sawtooth appears on the grid of V_6 . This partial sawtooth is amplified by V_6 , differentiated in the plate-loading transformer, further amplified in V_7 , and ultimately used to trigger the single-stroke blocking oscillator circuit of V_8 , which produces the delayed pulse. The slope control determines the range scale and the zero-set resistors balance out the combined effects of the starting time of the switching square wave, the starting voltage of the sawtooth, and the conduction point of the diode. The critical circuit elements in addition to those in

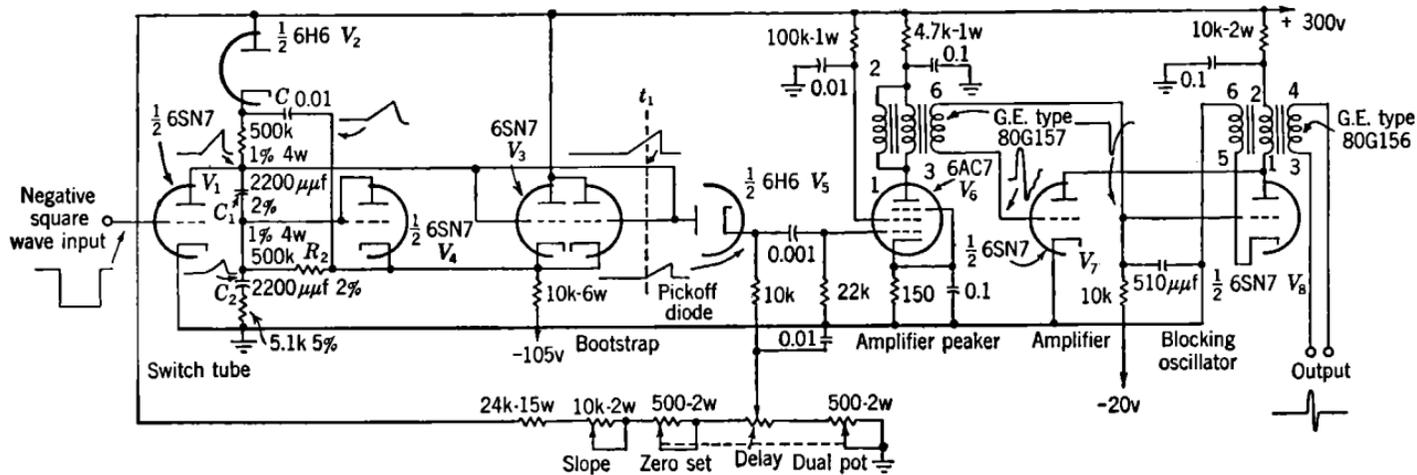


FIG. 13-36.—Sawtooth delay circuit.

the sawtooth generator are the delay potentiometer and its associated resistors and the diode V_5 . The circuit is described in detail in Chaps. 5 and 6 of Vol. 20 and Chap. 13 of Vol. 19.

The most accurate timing device at present available is a precision oscillator. This device can be used for the generation of a movable marker only by some method involving a continuous shifting of its phase. Extremely precise indices can be generated in this way, although the error of most phase-shifting elements is large (0.3 per cent). Space permits no description of this method but the reader is referred to Chap. 4 of Vol. 20 of this series.

Movable Height Marker.—The indices used for determining height on the RHI or the type E display are usually engraved mechanical ones. Often, however, a movable electronic index is provided. Neglecting the earth's curvature, the time t at which an index of constant height h should appear on an individual sweep is defined by the equation

$$h = R \sin \phi = \frac{c}{2} t \sin \phi \quad \text{or} \quad t = \frac{2h}{c \sin \phi},$$

where R is slant range, ϕ is elevation angle, and c the velocity of light. The timing can most easily be done by means of a linear-sawtooth delay circuit. A voltage proportional to $\sin \phi$ is used as the supply potential for the sawtooth generator. Thus, for a given diode bias, the slope of the sawtooth is proportional to $\sin \phi$ and the time delay is proportional to $1/\sin \phi$. The height represented can then be linearly controlled by varying the diode bias. In general, the inherent radar data are not accurate enough to justify circuits as precise as that of Fig. 13-36. At large ranges a correction must be made for the earth's curvature. This can be done with sufficient accuracy by adding a t^2 term to the sawtooth voltage, such a term being obtainable by integrating the sawtooth itself as in Fig. 13-36.

Discrete Timing Markers.—Discrete indices are invariably produced by deriving sharp pulses from a sinusoid that is properly phased with respect to the modulator pulse. The exact methods depend upon circumstances and in particular upon whether or not the marker circuit provides the synchronization. Figure 13-37 gives a typical example in which it does so; Fig. 13-38, one in which it does not.

The oscillator of Fig. 13-37 can be of any type satisfying the particular precision requirements. The original sinusoid is amplified and clipped to produce a symmetrical square wave of the same frequency. The "negative-going" edge of this square wave is then used to trigger a single-stroke blocking oscillator. The frequency of the sinusoidal oscillator, and therefore of the pulses, is made equal to that of the most closely

spaced markers desired.¹ Lower-frequency markers are provided by scaling circuits. A final scaling circuit provides pulses of proper frequency for triggering the modulator, the sweeps, etc. In the figure it is assumed that the desired frequency values are successively integral multiples of the lowest frequency. If this is not the case, multiple counting channels are necessary; for simplicity an integral relationship is usually chosen.

When synchronization is external, the oscillator must be shock-excited. Figure 13-38 illustrates an *LC*-oscillator widely used in medium-precision applications. When V_1 is conducting, the circuit is quiescent and energy is stored in the inductance by virtue of the cathode current.

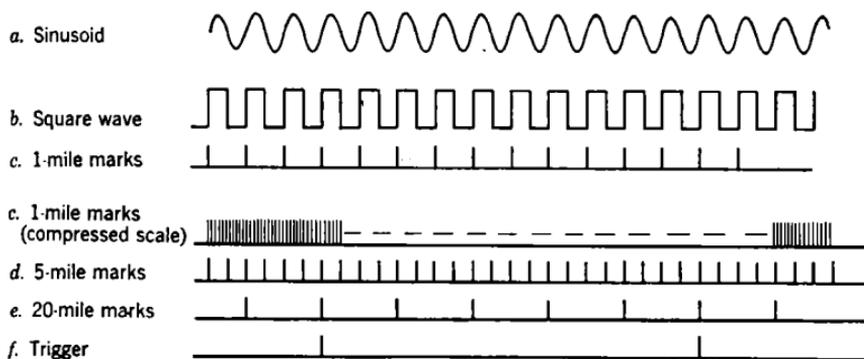
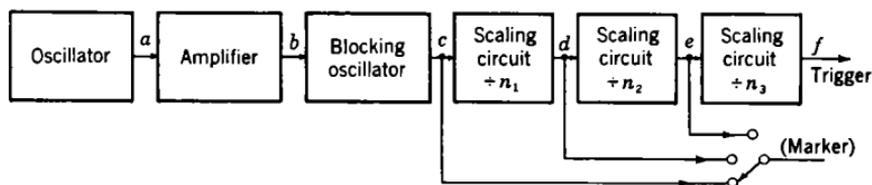


FIG. 13-37.—Derivation of pulses from c-w oscillator.

Interruption of this current by the square wave shock-excites the oscillating circuit. The resulting voltage wave has an initial phase of zero and an approximate amplitude of $I\sqrt{2L/C}$. The constants of the Hartley circuit of V_2 are so chosen that this initial amplitude is just preserved. In the present example the amplification necessary to produce steep wave fronts is provided by the regenerative amplifier formed by V_3, V_4 . Each time that V_4 is turned on, the surge of current through its plate circuit² triggers the single-stroke blocking oscillator formed by V_5 , giving rise to the desired pulse. It should be noted that a slight delay is inherent in the production of the first pulse, since a short but finite time is required to turn on V_4 . Later pulses are all displaced by

¹ For example, in the case of 1-mile markers this frequency is 93.12 kc/sec.

² The current necessary to charge C_3 increases the positiveness of this action.

the same amount. Since no pulse is involved in the reverse regeneration, the exact phase at which it appears is immaterial. The precision of the circuit is largely determined by L and C , which should be temperature-compensated and well shielded. Since R_1 has some effect on frequency, it should be of a reasonably stable type. Frequency adjustment is best made by slug-tuning the coil L . Once aligned, a well-designed unit should maintain its calibration to better than 1 per cent.

Many variations of this circuit, mostly simplifications, have been used. For example, if only a few markers are required on each pulse cycle, V_2 can be omitted, since the high- Q coils in the oscillating circuit

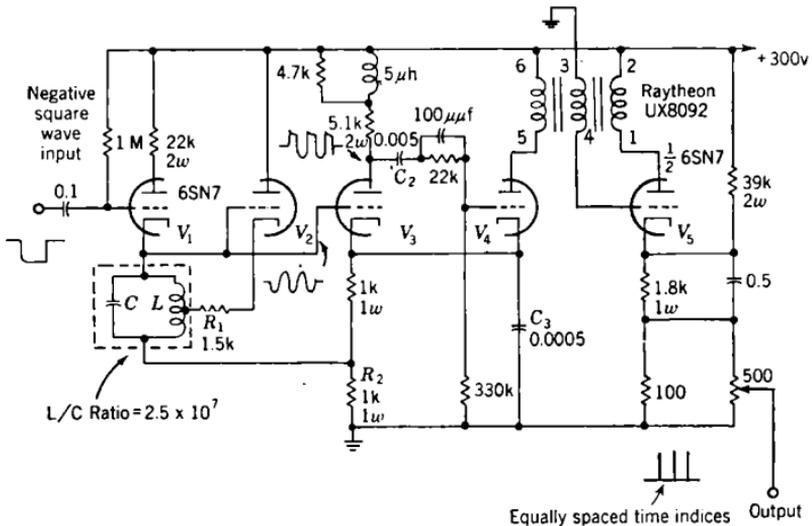


FIG. 13-38.—Switched generator for multiple time indices.

will ring with sufficient amplitude for several cycles. In many cases a simpler pulse generator can be used.

When two or more marker frequencies are needed alternatively, the resonant circuit is switched. Scaling-down is rarely used, since more vacuum tubes would be involved and since it would be necessary to recycle the counting circuits on each pulse cycle.

Delayed Sweeps.—Any of the timing circuits just discussed can be used to provide the trigger for a delayed, expanded sweep, and in fact this function can be combined with that of accurate range determination. The most usual method using continuous delays is illustrated in the block diagram of Fig. 13-39. The precision variable-delay circuit directly provides the trigger for the delayed sweep. An index marker is provided at the center of the sweep by a second fixed-delay circuit, usually a flip-flop, whose action is initiated by the sweep trigger. The timing arrangements are obvious from the waveform diagram. The total range

to the center of the sweep corresponds to the sum of the two delays, but the fixed delay can be included in the calibration.

Range determinations made in this way require that the display pattern be moved as coincidence between the echo and the marker is established. If this is undesirable, the marker-pulse delay can also be made variable and the two added, either by the operator or by an arrangement for mechanically summing the settings of the two control knobs. This method is more convenient than separating the delay and the measuring circuits, since it ensures that the marker will always appear on the sweep regardless of the setting of the two knobs.

A discrete set of pulses is frequently used to provide a stepwise sweep delay. The usual method is illustrated in Fig. 13-40. The principal

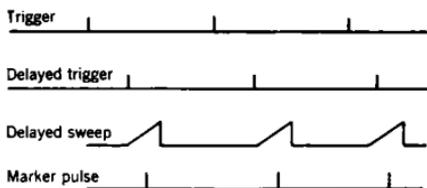
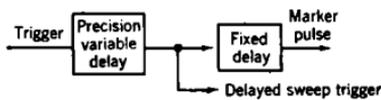


FIG. 13-39.—Synchronization of delayed sweep with precision range index.

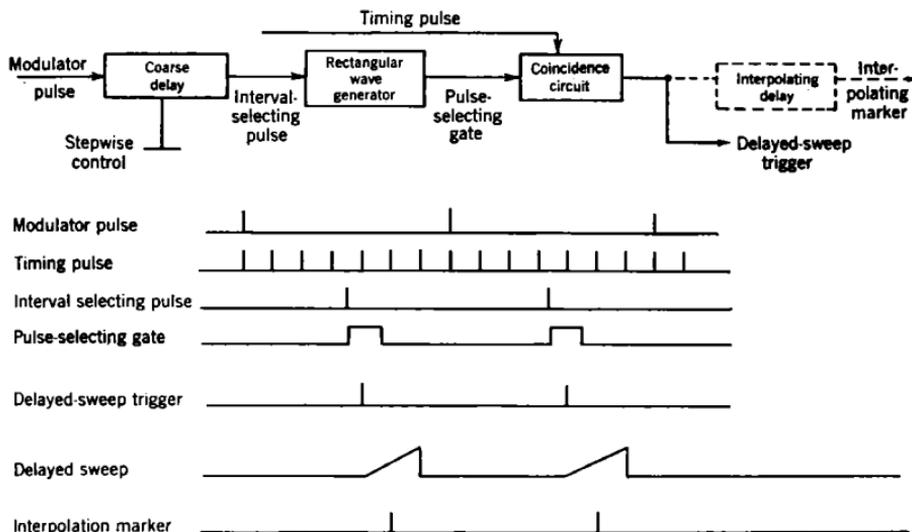


FIG. 13-40.—Discrete sweep delay.

problem is that of selecting the proper pulse. This is done by coincidence between the pulse and a gate initiated by a delay circuit sufficiently accurate to ensure that the gate brackets the proper pulse (see waveform diagrams). The length of the delayed sweep is usually made twice that of the delay steps, in order to ensure sufficient overlap. If desired, the delayed trigger may also be used to initiate an interpolating range circuit

as illustrated by the dotted part of Fig. 13-40. It is quite feasible to provide discrete indices on the expanded sweep by providing pulses of sufficiently high frequency to be useful. These may, for example, arise from an earlier stage of the chain of Fig. 13-37 than that which provides the delaying trigger pulses. Such multiple-scale systems are described in Chap. 6 of Vol. 20 of the series.

DISPLAY SYNTHESIS

Displays are synthesized by combining the components and techniques described in the preceding sections of this chapter. Although space will permit only brief descriptions of some of the more important and characteristic of the methods used, variations to fit particular circumstances and extensions to other applications will be apparent in many cases.

13-13. The Design of A-scopes.—This section will describe two methods for synthesizing type A displays, one of extreme simplicity intended only for test and monitoring purposes, and the other of a more elegant form.

A Simple A-scope.—In Fig. 13-41 simplicity of design has been carried to the extreme in producing an indicator intended only for testing or monitoring applications in which brightness and sweep linearity are not of great importance. The operation of the circuits is as follows. Tube V_{1a} and the first three elements of V_2 form a flip-flop providing a square wave, A , for intensifying the cathode-ray tube and at the same time cutting off the current from the plate of V_2 , on which a positive sawtooth is formed by the usual condenser charging circuit. One of the horizontal deflecting plates of the CRT is driven directly by this sawtooth; the other, by a negative sawtooth provided by the inverting amplifier V_{1b} , which is provided with cathode feedback. Since a rather large sawtooth (≈ 100 volts) is generated without exponential correction, the sweep slows down slightly toward the end.

As is shown, two sweep speeds are provided, one of 300 μsec using C_2 as the integrating condenser, and a second of approximately 5 μsec using the stray capacity at point B . More values can of course be provided by adding more switch points and condensers. In the interest of simplicity the flip-flop is not switched, being left on for the duration of the slower sweep. When the faster sweep is used, the beam is swept entirely off the tube face and remains off until termination of the flip-flop.

The video amplifier illustrated has a bandwidth of about 1 Mc/sec, the exact value depending upon the stray capacity introduced by the connecting circuits.

The power supply should include a conventional rectifier and LC -filter to provide the 300 volts direct current and an RC -filtered negative supply

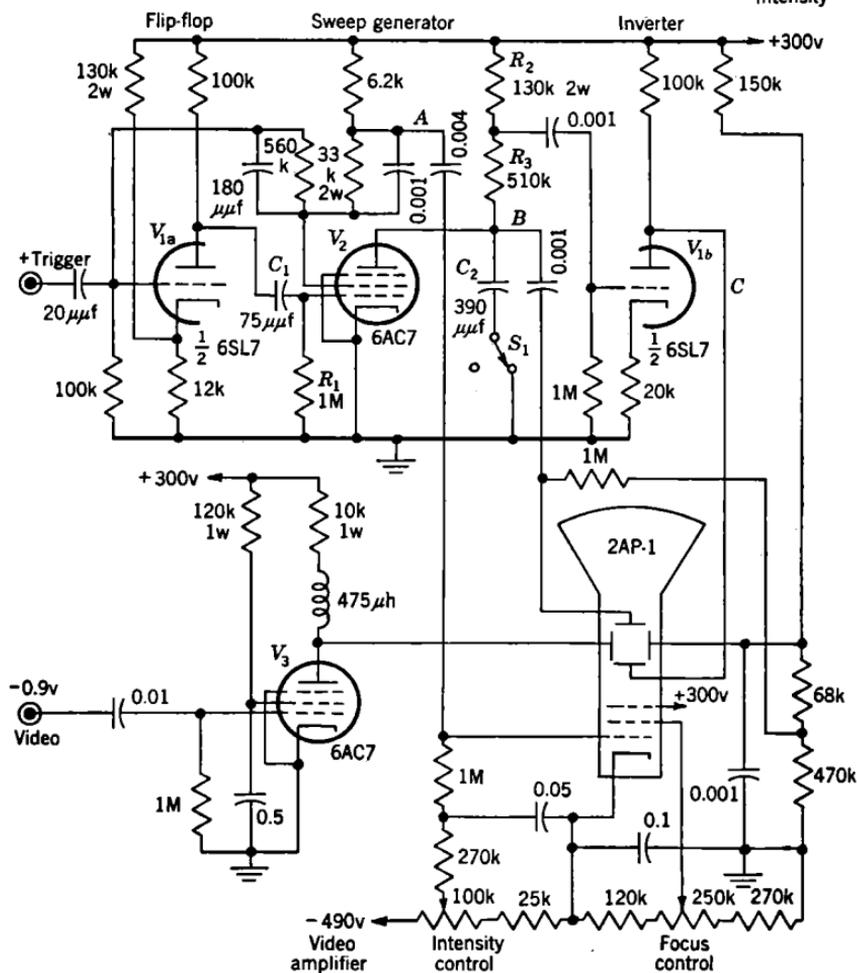
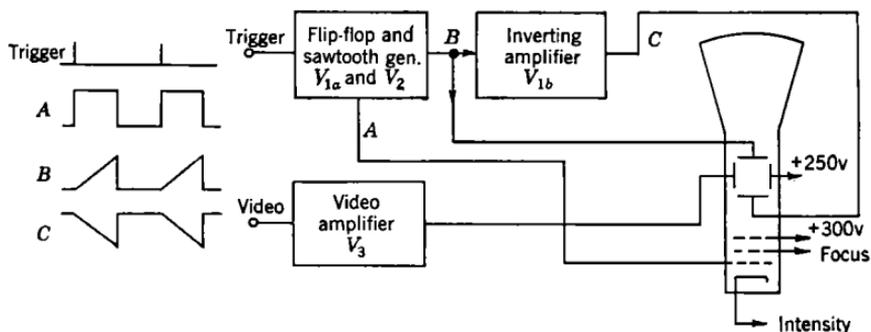


FIG. 13-41.—Simple A-scope.

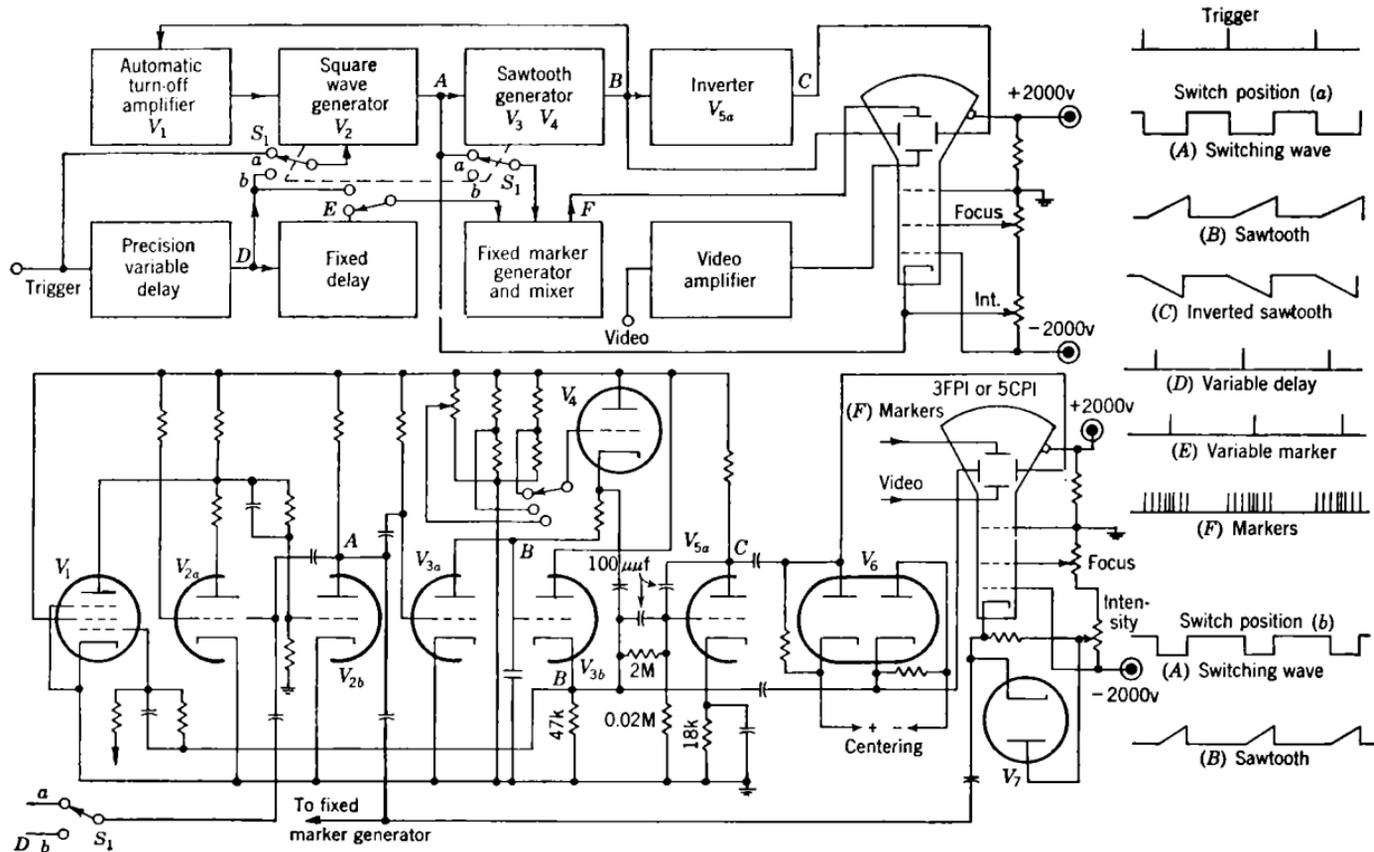


FIG. 13-42.—General-purpose A-scope.

from the same transformer. Since the load on the latter is extremely small, the voltage will correspond to nearly the peak amplitude of the transformer output wave.

The circuit can readily be adapted to larger cathode-ray tubes. For the 3BP1 (3 in.) it is only necessary to alter the sawtooth-generator charging resistor to compensate for a slightly different deflection sensitivity. A much more intense display can be provided by using the split-accelerator type 3JP1. In this case approximately -800 volts should be applied to the cathode and 1000 volts to the third anode. The second grid and the focus electrode are kept at suitable intermediate voltages. Auxiliary facilities, such as range-measuring circuits or sweep delays, can be added, but when these are desired it is usually preferable to use a more elegant basic indicator.

A General-purpose A-scope.—The block diagram of Fig. 13-42 illustrates an A-scope having provisions for both delayed and undelayed sweeps, fixed range markers, and an interpolating range marker. The cathode-ray tube used is of the post-deflection-acceleration type using up to 4000 volts over-all with reasonable deflection sensitivity (Sec. 13.1).

A precision variable delay and a fixed delay are combined to furnish a delayed trigger and a precision range marker as described in connection with Fig. 13-39. The marker can be used to determine the desired delay in advance by setting it at the point of interest. Since the minimum range of the marker is increased by the fixed delay, a switch (S_2) is provided for substituting the earlier pulse when an undelayed sweep is being used at short ranges. The dial calibration must then be changed accordingly, the simplest method being the use of a second scale.

The sweep circuits, which comprise the top row of the block diagram, are shown schematically below. On being triggered, the square-wave generator V_2 brightens the cathode-ray tube, switches the sawtooth generator, and, when the sweep is not delayed, also switches the fixed-marker generator. The sawtooth generator V_3 , V_4 (see Sec. 13-10) furnishes a very linear positive sawtooth which is of sufficient amplitude to drive one of the horizontal deflecting plates of the cathode-ray tube. The inverting amplifier V_{5a} , which drives the opposite CRT plate, is made very linear by plate-to-grid feedback. The equal condensers in the grid circuit ensure that the gain from point B to point C is unity, to keep the mean potential of the deflecting plates constant (Sec. 13-1). The d-c restorers V_6 hold the sweep origin fixed regardless of duty ratio.

Two fixed and one variable sweep speeds are shown. The square-wave generator is automatically turned off at the proper time by the action of V_1 . The positive sawtooth is delivered to the grid of V_1 , which is biased beyond cutoff by such an amount that it starts to amplify just as the sawtooth reaches the desired peak amplitude. The signal on the

plate of V_1 triggers off the flopover. This method is extremely useful when a variable sweep is involved, since it avoids the necessity of "tracking" two potentiometers.

13-14. B-scope Design.—The essentials for providing a type B display have been shown in Fig. 12·1 and described in Sec. 12·2. Either an electrostatic or a magnetic cathode-ray tube can be used, the former being much cheaper but far less satisfactory. Except for the final amplifiers the circuit requirements in a given situation are similar, no matter which tube type is used.

An electrostatic B-scope can be built around the range sweep circuits of the A-scope of Fig. 13·42,¹ the only changes necessary being to move the video and marker signals to the intensity-modulating elements of the CRT and to provide the azimuth sweeps. A potentiometer is almost invariably used as the azimuth data transmitter for an electrostatic tube. It may drive the deflecting plates directly, in which case it must have two brushes delivering peak potentials of from one to two hundred volts. Alternatively, a push-pull direct-coupled amplifier is inserted. If the potentiometer has only one brush, phase inversion is accomplished by common cathode-coupling like that used in amplifier V_4 , V_5 of Fig. 13·43b.

Magnetic B-scopes.—Magnetic B-scopes of varying degrees of complexity have been used. If the center of the display is to remain fixed in angle, a potentiometer is usually employed and the only essential departure from A-scope design is in the amplifiers. In many applications, however, more flexibility is desired, requiring more complicated methods. Figure 13·43 illustrates an example in which both the center of the sector and the scale factor in each direction can be chosen at will.

In each diagram the top row comprises the azimuth sweep circuits. An audio oscillator V_1 excites the first of a pair of differentially connected synchros by using its rotor in an oscillating circuit. The excited synchro is rotated in synchronism with the scanner, the second being manually oriented to select the desired sector (Sec. 13·4). As many as five or six B-scopes with independent sector selection can be operated in this way from a single oscillator and antenna synchro. The modulated signal from the second synchro is passed through an impedance-changing cathode follower V_2 , demodulated by a phase-sensitive rectifier V_3 (Sec. 13·9), then smoothed by an RC -filter. The azimuth deflecting circuit is completed by the push-pull amplifier V_4 , V_5 , in which the latter tube receives its exciting signal by virtue of the common cathode and screen resistors.

If continuous rotation is used, rather than sector scanning, the sweep

¹ The circuit of Fig. 13·41 is not suited for a cathode-ray tube using accelerating potentials as high as those required for a B-scope with a persistent screen.

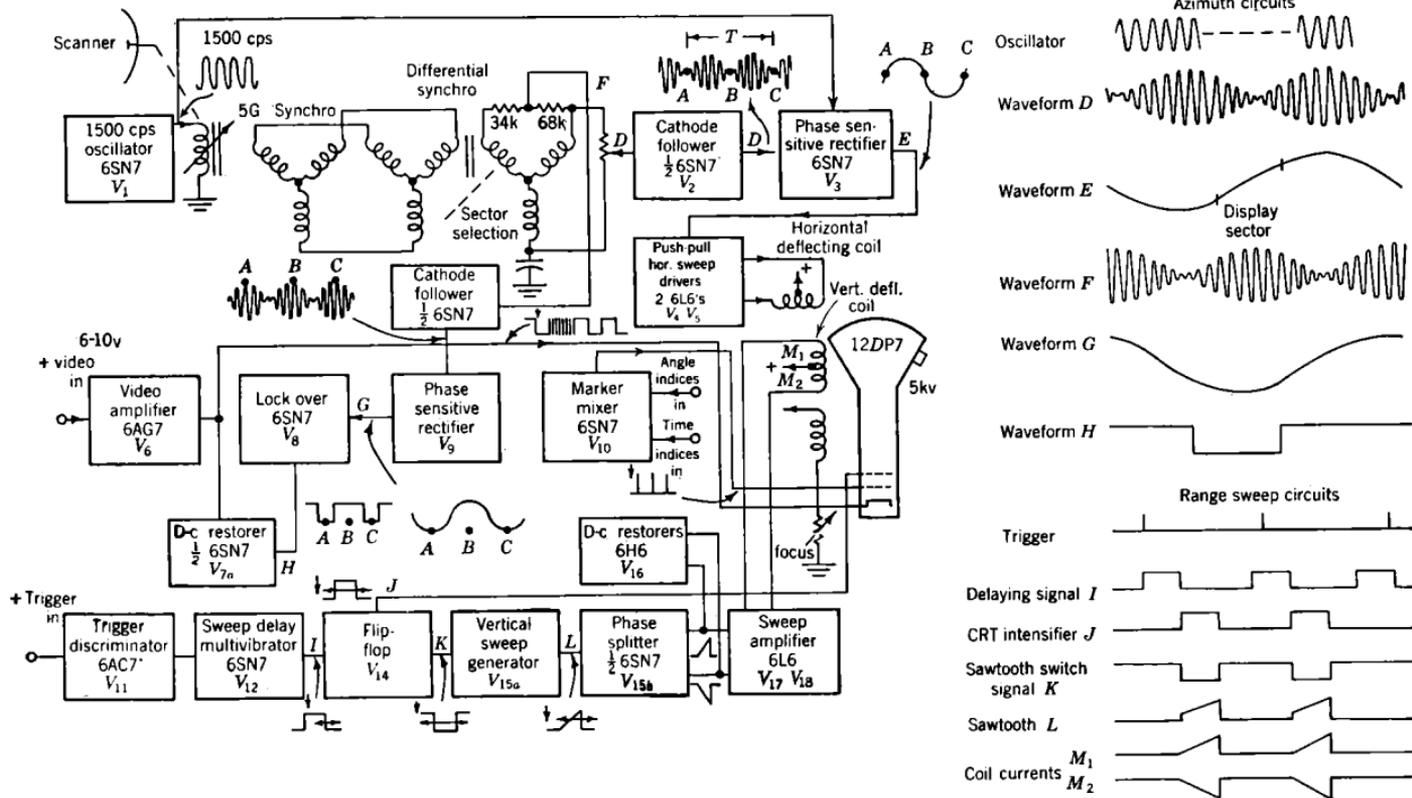


FIG. 13-43a.—Magnetic B-scope, schematic diagram. Small vertical arrows on waveforms at left indicate the timing of the transmitter trigger.

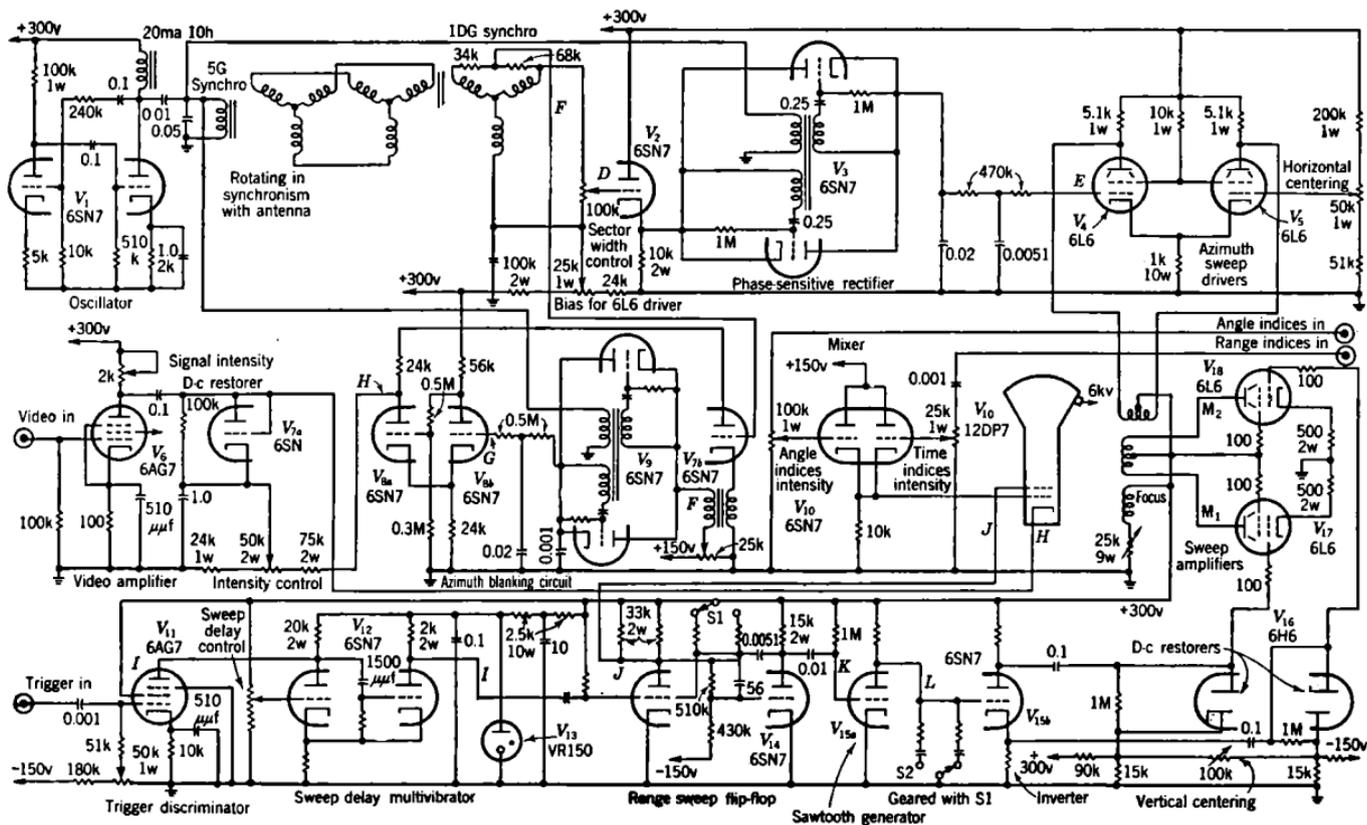


FIG. 13-43b.—Magnetic B-scope, circuit details.

will pass back across the tube when the scanner is pointing 180° away from the sector being viewed; it is thus necessary to blank the tube during that interval. This is accomplished by the circuit of V_{7a} , V_{7b} , V_8 , and V_9 . By means of a resistance network across two phases of the differential synchro a voltage¹ is obtained which has a modulation phase 90° removed from that of the azimuthal sweep voltage. After rectification by V_9 , this signal controls the action of the flopover V_8 . The latter triggers in opposite directions at two points which are symmetrical in scan angle with respect to the center of the sector displayed. The angular interval between these points is adjustable by means of the bias applied through the secondary of the isolating transformer ahead of the rectifier. The plate of V_{8a} controls the potential of the CRT cathode through d-c restorer V_{7a} and its parallel resistor in such a way that the cathode-ray tube is turned on only when V_{8a} is conducting.

The range sweep and associated delay circuits are shown in the bottom row. The simple delay multivibrator is not intended for use in range determinations, but can be replaced by some such arrangement as that of Fig. 13-42 if accurate range measurements are to be made.

Since only discrete sweep lengths are involved, the range-sweep flip-flop is switched along with the sawtooth generator in order to eliminate the vacuum tube otherwise required for an automatic "turn-off" circuit. Since the amplitude of the sawtooth is small compared to the supply voltage, it is sufficiently linear without special precautions. If continuous sweep-length control is desired, the circuits of Fig. 13-42 should be substituted, the feedback voltages for the turn-off circuit and for the sawtooth generator being taken from the cathode of V_{15b} .

The sawtooth passes through the "phase-splitting" amplifier V_{15b} , which provides signals of both polarities for the push-pull sweep amplifiers. The cathode feedback of the final amplifiers is sufficient to give reasonably linear current amplification, since the push-pull action tends to compensate for tube nonlinearities. The high-frequency response is not good enough, however, to provide a linear displacement at the beginning of very fast sweeps. For this reason a "step" is introduced at the beginning of the sawtooth by placing a small resistor in series with the charging condenser of the sawtooth generator. The sudden transfer of current from the switching tube V_{15a} to the condenser circuit results in the abrupt appearance of a voltage drop across the resistor, thus providing the step.

The above circuits can be used to produce a "micro-B" display by the expedient of gearing together the sweep-delay and the sector-width controls, and providing the proper normalization. A given normalization will be correct for only one sweep speed.

¹ Except for this requirement, this synchro could have a single-phase rotor.

In applications involving extremely rapid sector scans (several scans per second), synchros cannot be used and a variable condenser must be substituted for the scanner synchro. The oscillator frequency must then be much higher. No problem of sector selection is involved, since the rapid scanning covers only a narrow sector, which is completely displayed at all times.

A typical narrow-band video amplifier and a marker signal mixer are included for completeness. Incoming video signals of 8 to 10 volts are amplified by V_6 and applied to the cathode of the CRT. The tube V_{7a} provides d-c restoration to assure the proper operating level. Electronic range and angle indices from sources outside of the diagrams can be mixed into the double cathode follower V_9 and applied to the CRT grid. It is assumed that both range and angle indices have positive polarity and that the potential of the angle-index lead is zero in the absence of a signal.

If the center of the sector is fixed in azimuth, a great simplification can be brought about by driving amplifier V_4 , V_5 directly from a potentiometer which replaces the remainder of the top row in the diagrams. The azimuth blanking circuit (V_{7b} , V_8 , V_9) can also be replaced by a simple cam-and-switch arrangement. If the scanner executes only a sector scan, even this is unnecessary.

13-15. Plan-position Indicator.—As has been indicated in earlier sections (e.g., Sec. 12-2), PPI displays can be produced by any of several methods, which, in general, may be classified under three principal headings.

1. The rotating-coil method, in which the range-sweep current is passed through a coil which is rotated about the tube neck in synchronism with the antenna motion. If "off-centering" is desired, it is provided by means of a fixed coil outside of the moving one.
2. The "resolved time base" method, in which a sawtooth waveform is passed through a 2-phase resolver,¹ practically always a synchro, and the resulting sweep "components" are utilized to energize orthogonal deflecting coils or plates.
3. The method of "pre-time-base resolution" in which slowly varying sine and cosine voltages obtained from a d-c excited sinusoidal potentiometer or by rectification of signals from an a-c resolver are used to generate the necessary sawtooth. This method is used principally in cases where the scan is too rapid for the rotating-coil method and the antenna is so far away that the transmission of the

¹ In some cases a 3-phase synchro is used, and the resulting signals are passed through a resistor network which reduces the three phases to two.

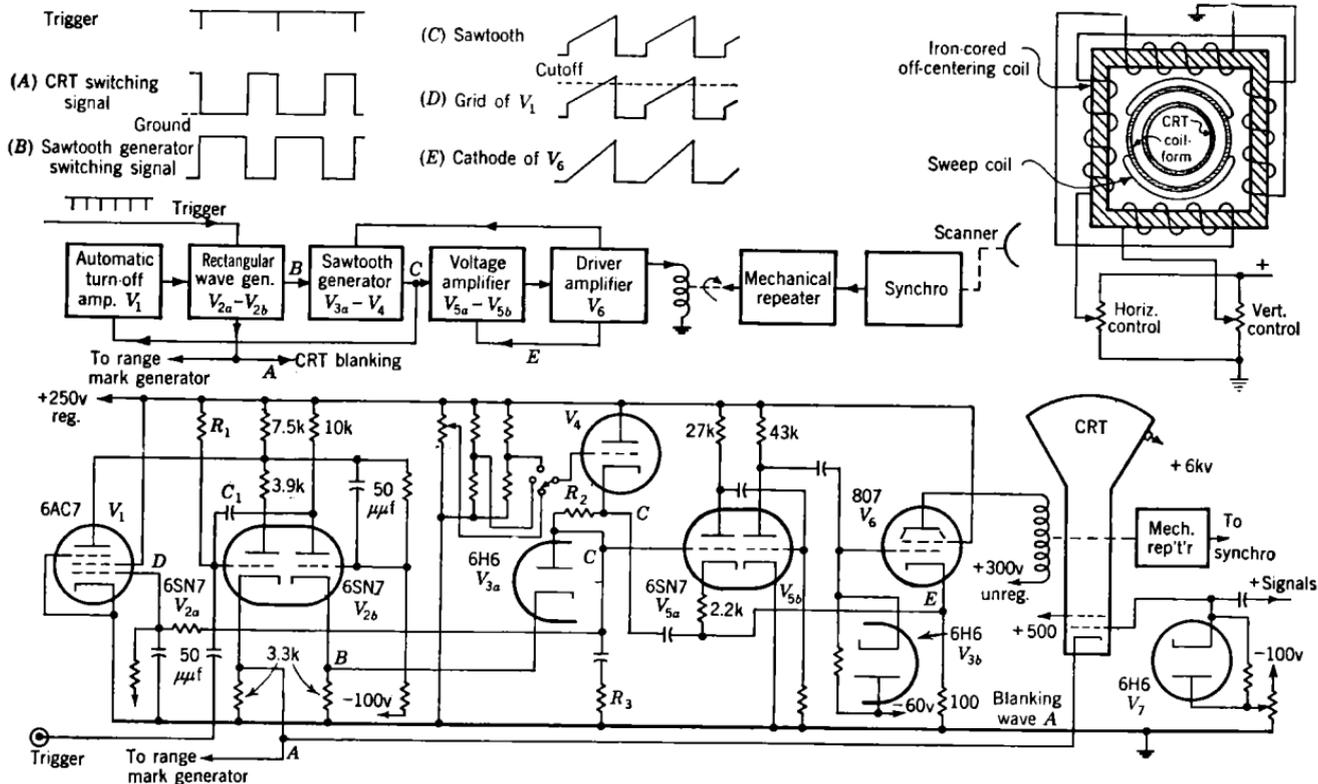


FIG. 13-44.—Rotating-coil PPI.

slowly varying voltages is appreciably easier than would be the case with the sawtooth waveforms involved in the resolved-time-base method.

Methods 2 and 3 can be applied either to magnetic or to electrostatic cathode-ray tubes, but, because of their inferiority for intensity-modulated displays, the latter tubes are seldom used.

The Rotating-coil Method.—Figure 13-44 illustrates the circuits used to produce the sweeps of a rotating-coil PPI, omitting such accessory parts as range-marker circuits, etc. The tubes V_1 , V_2 , V_{3a} , and V_4 constitute a square-wave and a sawtooth generator similar to those of Fig. 13-42 except that a diode clamp (V_{3a}) is illustrated for variety in Fig. 13-44. Tubes V_{5a} , V_{5b} , and V_6 constitute a feedback amplifier whose linear output current drives the deflection coil. The drop across the cathode resistor of V_6 , which is due almost entirely to the deflecting current, is fed back to the cathode of V_{5a} . Because of the high gain in the feedback loop, this voltage faithfully reproduces the form of the original sawtooth, thus assuring the desired linear increase in the current. Since any current through V_6 between sweeps would result in a circular rather than a point origin for the sweep, this tube is biased beyond cutoff. Since there is no feedback so long as V_6 is beyond cutoff, its grid is elevated very rapidly at the beginning of the sawtooth and there is thus no appreciable delay before the current starts.

The deflecting coil is rotated by a mechanical repeater such as those described in Sec. 13-5. The diagram at the upper right illustrates one method of connection for an "off-centering" coil.

This type of PPI can be made to have the greatest dependable accuracy attainable with any type of intensity-modulated display. Mechanical repeater systems can, within certain limits of rotational speed, be made to have accuracies of a small fraction of a degree, 0.1° being a common error limit. Since the range-sweep feedback amplifier is extremely faithful at little cost, the burden of accuracy in range falls on the sawtooth generator, for which precision parts of high constancy can easily be used. The principal disadvantages of the rotating-coil method are: (1) it is limited to continuous scanning rates of 30 to 60 rpm and to sector scanning of comparable angular velocity; (2) off-centering is extremely expensive in power, and cannot be used for displacements of more than two or three tube radii because of excessive distortion and defocusing; (3) for cases demanding minimum weight, even at the expense of accuracy, the rotating-coil PPI is less suitable than some other types.

13-16. The "Resolved Time Base" Method of PPI Synthesis.—A PPI can also be derived by passing a sawtooth waveform through a 2-phase resolver and using the resulting sine-and-cosine-modulated signals, with

or without amplification, to drive an orthogonal deflecting system (Fig. 13-45). When amplifiers are used, or if an electrostatic tube without amplifiers is involved, the resolver must deliver sawtooth voltages into a high-impedance load; when a magnetic tube is used without amplifiers, sawtooth currents must be delivered to an inductive load.

In principle the resolver may be a synchro, a potentiometer, or a condenser; a synchro is far the most satisfactory and, in spite of certain shortcomings, it is universally used.

As in all cases of transmission of a nonsymmetrical signal through an a-c coupling, the d-c component of the signal is lost. This presents one

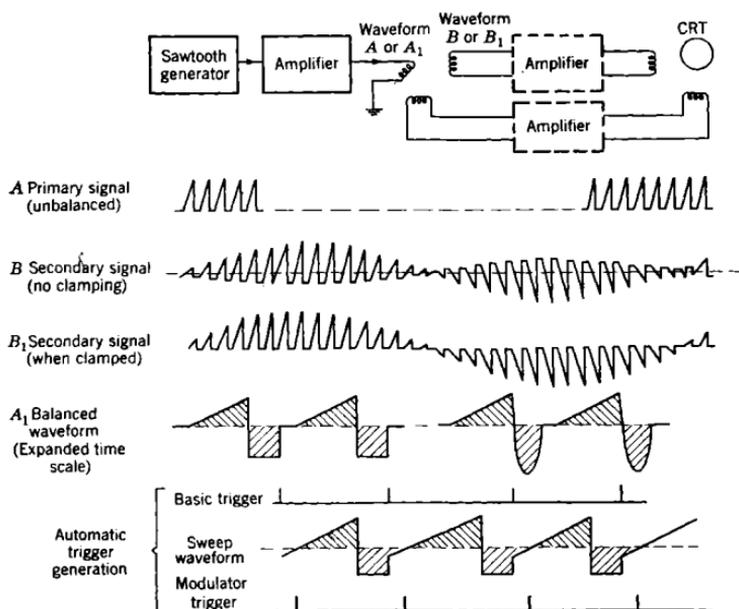


FIG. 13-45.— Resolved-sweep PPI methods.

of the most serious difficulties involved in this technique. With a passive and linear load, the long-time average of both the current through the secondary of the synchro and the voltage across it must be zero (Waveform B, Fig. 13-45). Thus the absolute current or voltage at the beginning of the modulated sawtooth will not be constant, but will change with the modulation amplitude and will reverse sign when the sawtooth does. As a result, if no corrective measures were taken, each individual sweep on the display tube would be displaced in the negative direction by a fixed amount and the locus of the sweep origins would be a "negative" circle. This difficulty has been overcome by two general methods, each of which has variations most suitable to particular circumstances.

The first method involves the use of clamps or switches in the secondary circuit to force the potential or the current to zero in the interval between sweeps. So long as the sweep component has a given polarity, this can be very simply accomplished in the potential case by a d-c restorer (Sec. 13-9), and in the current case by a selenium rectifier in series with the secondary circuit. In cases where the scanning is confined to 90° or less, these variations of the first method are the most satisfactory of all. If the scan covers more than 90° , one of the diodes must be switched in polarity from time to time—for example, by a cam arrangement on the scanner. If the scan covers more than 180° both of them must be so switched. Although it is difficult to switch quickly and smoothly enough to avoid aesthetically displeasing display irregularities, this method is quite satisfactory and is probably the simplest for the case of resolved *currents*, providing the scanning rate is not too high. A better and only slightly more expensive method of *voltage* restoration involves the use of a double clamp (Fig. 13-25) which connects the secondary to a point of proper potential between sweeps but releases it when the sweep begins. This is the most widely used method in the voltage case when polarity reversal is involved. No satisfactory series switches exist for analogous use with resolved currents.

In the second method of correcting the sweep-origin difficulty, the positive and negative waveforms are “balanced” about the sweep origin either by introducing a precisely controlled negative waveform in the primary circuit between sweeps, or by arranging to trigger the transmitter automatically at the precise time when the secondary current or voltage passes through zero. Both variations involve a good deal of precision and are more costly than the switched clamps mentioned in the last paragraph as usable when voltage is involved. However, the present methods have been widely used in connection with resolved currents, since they require no switching elements in the low-impedance secondary circuit.

Applications of these various methods will be described in connection with the following specific examples.

Magnetic PPI Using Amplifiers.—Figure 13-46 shows the secondary circuit of a magnetic PPI of the resolved-time-base type using the method of clamping between sweeps. It is supposed that the synchro is preceded by the sweep generator and amplifier of Fig. 13-44, except that the circuit of V_6 is changed as indicated so that the feedback voltage is that across the synchro itself.

The clamps on each of the amplifier grids of Fig. 13-46 are of the double-triode variety illustrated in Fig. 13-25a. They are switched by the same flip-flop (rectangular-wave generator) that controls the sawtooth generator, and are thus closed during the entire interval between

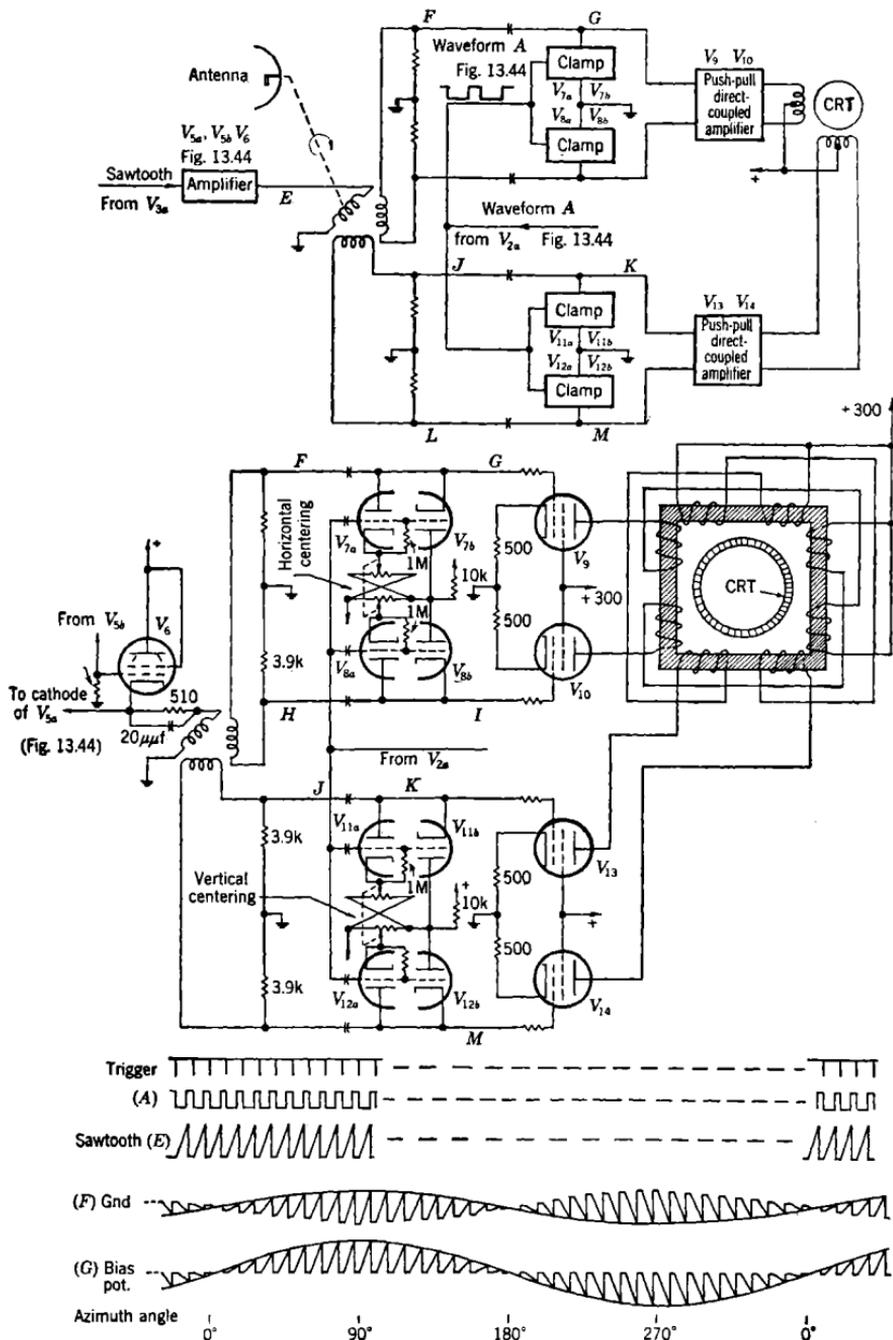


FIG. 13-46.—Resolved-sweep PPI using amplifiers.

sweeps. The switched clamps shown can be replaced by simple d-c restorers if the scan covers less than 90° . The final sweep amplifiers are of the cathode-feedback type used in the B-scope of Fig. 13-43.

The centering of the PPI can be adjusted by setting the potentiometers controlling the clamping points. If extreme off-centering moves the amplifier grids out of their normal operating range, a delay occurs at the beginning of the sweep which gives the effect of a virtual center at the desired point. By using an extremely high-voltage sawtooth, off-centering of as much as six to eight tube radii has been provided. For such extreme amounts of off-centering, it is best to insert cathode followers between the potentiometers and the clamps.

The dependable accuracy obtainable with the resolved-time-base method is not so great as can be obtained with a rotating-coil PPI, although with care it can be made nearly as good except on very fast sweeps. However, the present method has much less severe limitations with respect to scanning speed, and provides a much greater degree of usable off-centering than does the rotating-coil scheme.

Its weight is somewhat less and its power requirements somewhat greater than those typical of the rotating-coil variety without off-centering. When off-centering must be provided, the present method is far lighter and less costly in power.

The technique just described can be applied to electrostatic tubes by substituting voltage amplifiers in the final driving stages, or in some cases by eliminating amplifiers entirely and connecting the clamped points directly to one member of each pair of deflecting plates. Unless weight and power consumption are extremely critical, the saving effected does not justify the sacrifice in quality entailed.

13-17. Resolved-current PPI.—When the deflection coils are driven directly by the synchro, the amplifiers of Fig. 13-46 are eliminated, with the result that the indicator is considerably simplified and errors due to imperfections in the amplifier response are avoided. These advantages are accompanied by certain restrictions which will appear below.

In order to reduce the effects of distributed capacity in the synchro and the deflecting coils, and of the shunt capacity of the cables, low-inductance windings of few turns are used throughout. The synchro rotor is matched to the driving tube by a stepdown transformer. Since the secondary circuit is almost purely inductive (ideally it would be precisely so), its current waveform will resemble that in the primary of the transformer, and a current amplifier similar to that driving the deflection coil of Fig. 13-44 is appropriate. In many cases imperfections in the transformer are overcome by taking the feedback from its secondary circuit.

The low deflection sensitivity of the sweep coils renders it impractical

to use them for off-centering; therefore, when this is desired, separate coils are provided on the same form. Since they are inductively coupled to the sweep coils, the number of turns in the off-centering coils must also be limited. Off-centering is therefore expensive in current, and its extreme amount is usually limited to not more than two tube-radii.

Since vacuum tubes are not appropriate as switches in such low-impedance circuits, it is not possible to fix the sweep origin by methods analogous to that of Fig. 13-46. It has been necessary to use either the method of a series dry-disk rectifier switched (if necessary) by the scanner, or to apply some method of waveform balance. Applications of these methods appear in the following descriptions.

The Use of Series Rectifiers.—Figure 13-47 illustrates those parts that must be added to the sweep generator and amplifier of Fig. 13-44 to produce a complete PPI involving the use of disk rectifiers. By provid-

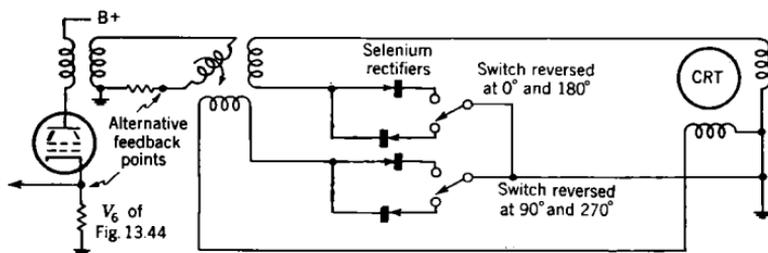


FIG. 13-47.—Resolved-current PPI using switched rectifiers.

ing two rectifiers for each secondary as indicated, only one point need be switched. Optimum smoothness at the instant of switching is accomplished if contact is made through the second rectifier before the first one is opened, so that the circuit is never broken.

This method has the virtue of extreme simplicity and in many circumstances it gives excellent performance. However, it has two principal limitations in addition to the restriction on off-centering mentioned above: the mechanical switching limits the scanning rates at which it can be used, and some distortion is introduced on "long" sweeps (more than 50 miles) by the low but appreciable resistance of the rectifiers. If this resistance were constant, the only effect would be a droop caused by inadequate low-frequency response. This droop could easily be overcome by shaping the primary waveform, but unfortunately the rectifier resistance is nonlinear, being much larger at very low than at intermediate or high currents. As a result, the effect of the rectifier resistance changes with scan angle. "Circles" of equal range are slightly "squared," the "corners" coming at the 90° positions.¹

¹ This effect should not be confused with the "square circles" found on certain early PPI's of the type shown in Fig. 13-46. These errors, which resulted from

Automatic Triggering of the Transmitter.—Waveform balance by triggering the transmitter as the secondary current reaches zero is illustrated in Fig. 13-48. Since the voltages in the secondary circuits are low, the signal from the cathode resistor in the sweep amplifier is used to provide the required information. This signal is passed through a network whose voltage response is equivalent to the current response of the deflection-coil circuit, and from there to the first stage of a high-gain amplifier. The grid of the first amplifier tube is so biased that in the absence of signal it

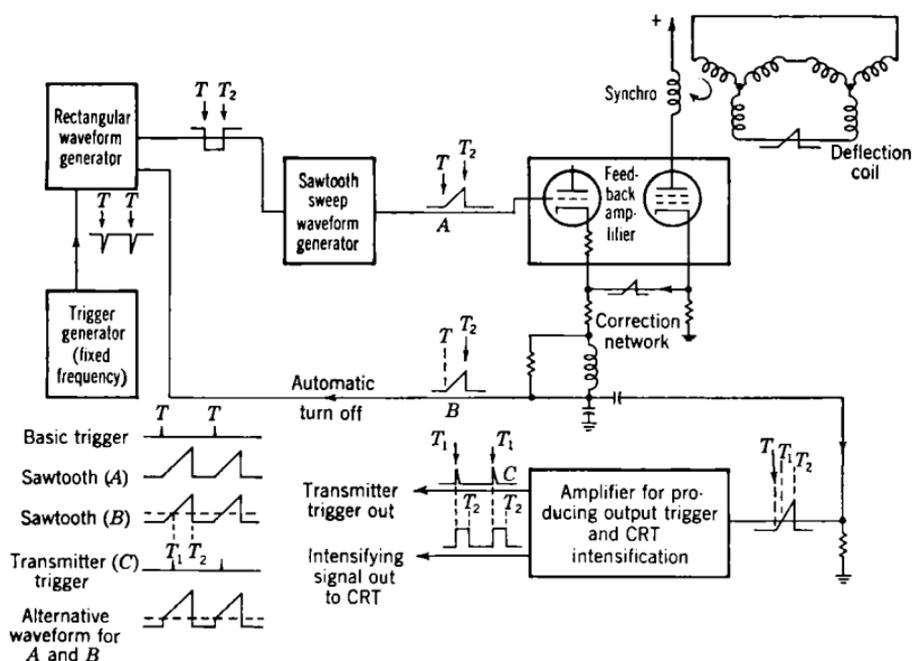


FIG. 13-48.—Resolved-current PPI using automatic transmitter triggering.

is exactly at cutoff. Thus when a sweep is being applied this grid will pass upward through its cutoff point at exactly the same time that the deflecting current is passing through zero. The remainder of the sawtooth is therefore amplified and, by suitable overdriving and clipping, can be turned into a rectangular wave. The transmitter trigger is then derived from the front edge of this rectangular wave, and the wave itself serves to intensify the cathode-ray tube.

If a simple sawtooth is used, the maximum useful sweep duration has an upper limit of half the pulse period; somewhat less can be realized in

inadequate *high-frequency* response in the synchro and amplifiers, and from cross-coupling in the former, were troublesome only on fast sweeps. Furthermore the "corners" were at the 45° rather than the 90° positions.

practice because of the finite time for sweep "recovery." By using a trapezoidal waveform, in order that the time spent at maximum "negative" current become appreciable, sweep duration can be increased to about 60 per cent in practice. This is easily accomplished through the medium of a resistor in series with the integrating condenser of the sawtooth generator, as is explained in Sec. 13-14.

This form of resolved-current PPI has been successfully used in the lightweight airborne equipment (AN/APS-10) described in Secs. 15-12 to 15-14. As is shown in Fig. 13-48, the synchro is driven directly by the amplifier without a transformer. The synchro secondary and the deflection coil (which is wound in a standard synchro stator form) have 3-phase rather than 2-phase windings. The equipment is light in weight, comparatively simple and economical. Its principal drawbacks are the limited duty ratio, the precision required in the triggering circuit, and the limited off-centering attainable. No off-centering is used with the 3-phase system of Fig. 13-48; when it is desired, 2-phase windings are more suitable.

The Use of Special Waveforms.

The method that will now be described is similar to that just treated in that positive and negative signals are "balanced." It differs, however, in that the waveform includes a quiescent period, immediately preceding each sweep, during which the waveform has its average value, so that there is no necessity for precise control of the trigger (Fig. 13-45).

The required waveform in the primary is derived by the introduction, immediately following the sweep proper, of a signal of opposite polarity whose time integral is exactly equal to that of the sweep. Both sinusoidal and rectangular waves have been used, but only the former will be discussed here.

A waveform of the proper type can be generated by the Miller-type circuit (Sec. 13-10) shown in Fig. 13-49. When the grid of V_1 is released by opening the switch, it quickly rises to cutoff and a negative sawtooth begins on the plate. The form of this wave is practically independent of the load impedance, which is in this case made up of the inductance L , V_2 , and R_2 in series, and the various capacities. The a-c component of current in the coil is at any instant given by

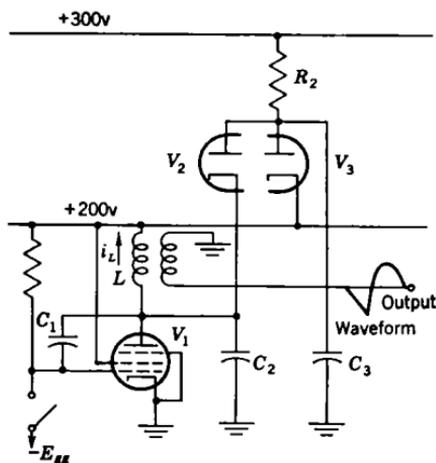


FIG. 13-49.—Balanced-waveform generator using an inductance.

$$I = \frac{1}{L} \int E dt,$$

and is therefore proportional to the time integral of E . When E reaches the desired maximum value, the switch is closed and V_1 is disconnected. The sudden termination of its current results in an oscillation of L and its associated capacity. The plate end of the inductance swings positive, disconnecting V_2 as the potential rises past the 200-volt level. On the down-swing, V_2 becomes conducting again and clamps the inductance to the 200-volt point, resulting in the waveform shown. The amplitude of the oscillation is dependent upon the energy stored in the inductance at the end of the sawtooth, which in turn is determined by the current and hence by the "area" under the sawtooth. It can be rigorously proved by analysis that this amplitude must be such that the average voltage across the inductance (neglecting its resistance) is zero during the waveform. We shall be content with the general statement that this must be so since, in the absence of resistance, the average potential across an inductance during a completely closed cycle must be zero regardless of waveform. Therefore in this case the initial, the average, and the final values of the waveform are identical and it is "balanced" as desired.

A complete circuit using this method is illustrated in Fig. 13-50. The Miller circuit is switched by the double-diode clamp V_3, V_4 , which in turn is controlled by the cathode-coupled flopover V_1, V_2 . The latter is automatically terminated at a fixed sawtooth amplitude by the signal passed through the biased diode V_8 . The Miller circuit is like that described in connection with Fig. 13-49, except that V_7 (V_3 of Fig. 13-49) is biased somewhat below 200 volts, thereby introducing a small correction to take care of inadequacies in the response of the deflection system. The potentiometer P_1 allows a wide variation in the sweep speed used, and P_3 adjusts the pattern to proper size.

The amplifier uses negative feedback of a signal taken across a resistor in the transformer secondary, thereby correcting for deficiencies in transformer response. In contrast to previous cases, the final stage of the amplifier is driven negative during the actual sweep; thus the induced voltage on the plate is positive and the tube does not "bog down." The high potential needed for quick recovery when the current is rising again is provided by the joint action of L_1 and C_1 . During the sawtooth a large positive voltage is induced in L_1 , and C_1 is charged accordingly. When the current derivative reverses at the end of the sawtooth, the voltage is temporarily maintained at a high level by C_1 , and, in spite of the induced voltage across the transformer, the plate of the tube remains at a workable voltage level during the rapid back-swing. The choice of polarities is such that the interval in which C_1 aids the plate supply is short.

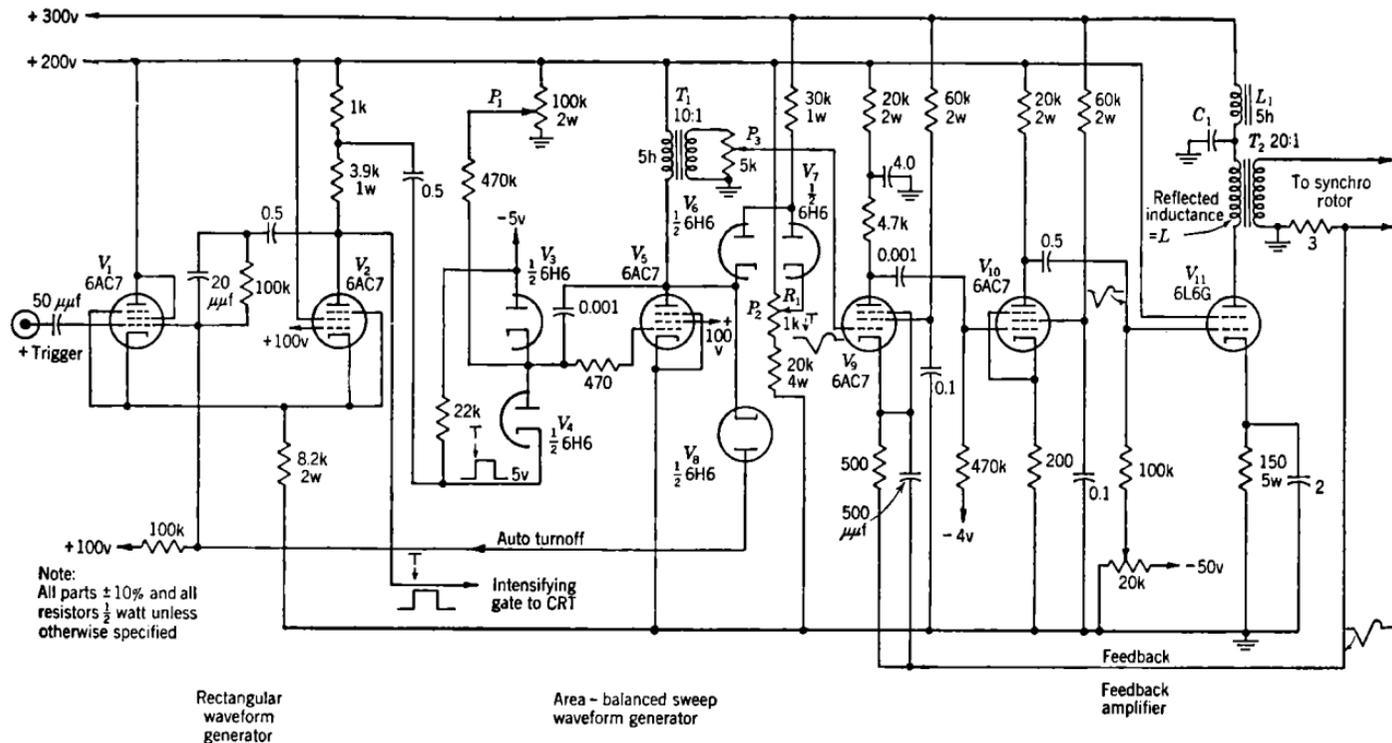


FIG. 13-50.—Balanced-waveform PPI.

The performance of this type of PPI is comparable to that in which the transmitter trigger is generated automatically. Technically, it involves fewer parts, its method of area balancing has somewhat greater reliability, and it has the advantage of being usable with self-synchronous modulators. In common with all balanced-waveform methods, its duty ratio is limited to some 50 to 60 per cent unless excessively large peak currents are used in the balancing waveforms.

13-18. The Method of Pre-time-base Resolution.—In certain cases it is desirable to transmit the scanner data in terms of slowly varying sine and cosine voltages derived from a sine-cosine potentiometer, or by the modulated-carrier technique used in the B-scope of Sec. 13-14. These are used as potential sources for sawtooth generators, as indicated in Fig. 13-51, where the switches represent double-sided clamps that are

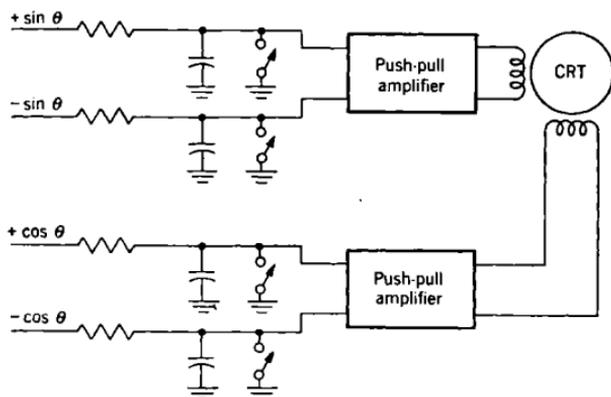


FIG. 13-51.—Pre-time-base resolution method of PPI synthesis.

opened during the sweep interval. Thus the signals at the input to the amplifiers are sine-and-cosine-modulated sawteeth of precisely the same nature as those obtained by passing a sawtooth through a synchro. The amplifiers can, therefore, be identical with those of Fig. 13-46. If it is too inconvenient to provide both polarities, one can be omitted in each circuit and a phase-inverting amplifier inserted between the sawtooth generator and the final amplifiers. In this case, however, it will be necessary either to use d-c coupling (which is difficult) or to provide clamps on the grids of the final amplifiers.

This method can be used at very high scanning speeds, since it is possible to use a variable condenser for carrier modulation. It has advantages over the resolved-time-base methods when very fast sweeps are used, since there is no starting delay such as that introduced in passing sweeps through a synchro. It can also be advantageously used when the scanner is so remote as to make the transmission of sweep signals by cable difficult, and when the scan is at the same time too rapid to permit

the use of mechanical repeaters. An extreme case is that of radar relay (see Chap. 17). Many of these advantages make the technique particularly appropriate for use on range-height indicators.

On the other hand, under average conditions, pre-time-base resolution is considerably less satisfactory than the methods previously described. The sine potentiometer, though simple, is prone to troublesome irregularities which cannot be completely filtered, except on extremely slow scans, without causing phase lags in the basic scanner data. The use of a carrier and of demodulators introduces considerable complexity compared with methods described earlier. The necessity of maintaining the sawtooth generators in proper normalization in order to avoid distortion is somewhat of a burden compared to other cases where only one such generator is involved and where, furthermore, any lack of constancy affects only the size and not the fidelity of the display.

13-19. The Range-height Indicator.—Any off-centered PPI can be used as a range-height indicator by substituting elevation angle for azimuth, providing the indicator can follow the scanner involved. In most cases, it is desirable to present height on a scale that is considerably expanded in comparison to that used for the display of ground range. In consequence the most adaptable PPI techniques are those in which each cartesian coordinate has its own amplifier which permits the desired height expansion to be provided simply by increasing the gain of the proper channel. The methods of Figs. 13-46 and 13-51 can both be used in this way.

Simplification can usually be effected. The horizontal sweep always occurs on the tube face in one direction from the origin. In many cases, this is also true of the vertical component. (Sometimes negative heights are involved when the radar is on an elevated site.) Furthermore, it is often possible to use one or both of the approximations $\sin \theta \approx \theta$ and $\cos \theta \approx 1$. In any case in which a sweep component is unipolar, the clamps of Fig. 13-46 can be reduced to unswitched d-c restorers, and the switched clamps of Fig. 13-51 can become single-sided, though still being switched. The approximation $\cos \theta \approx 1$ allows the use of unmodulated sweeps in the horizontal direction; this is of no particular virtue if the resolved-sweep technique is used, but provides a considerable saving in connection with the use of a modulated carrier. The approximation $\sin \theta \approx \theta$ allows the use of linear data transmitters, such as linear potentiometers and condensers.

Since the applications of Fig. 13-46 and Fig. 13-51 to use with an RHI are fairly obvious, no attempt will be made to discuss them specifically. Instead, a brief description will be given of an RHI for use with rapid scans including the "sawtooth" type developed by certain electrical scanners (Secs. 9-14, 9-15, and 9-16). It is assumed that $\sin \theta \approx \theta$ and

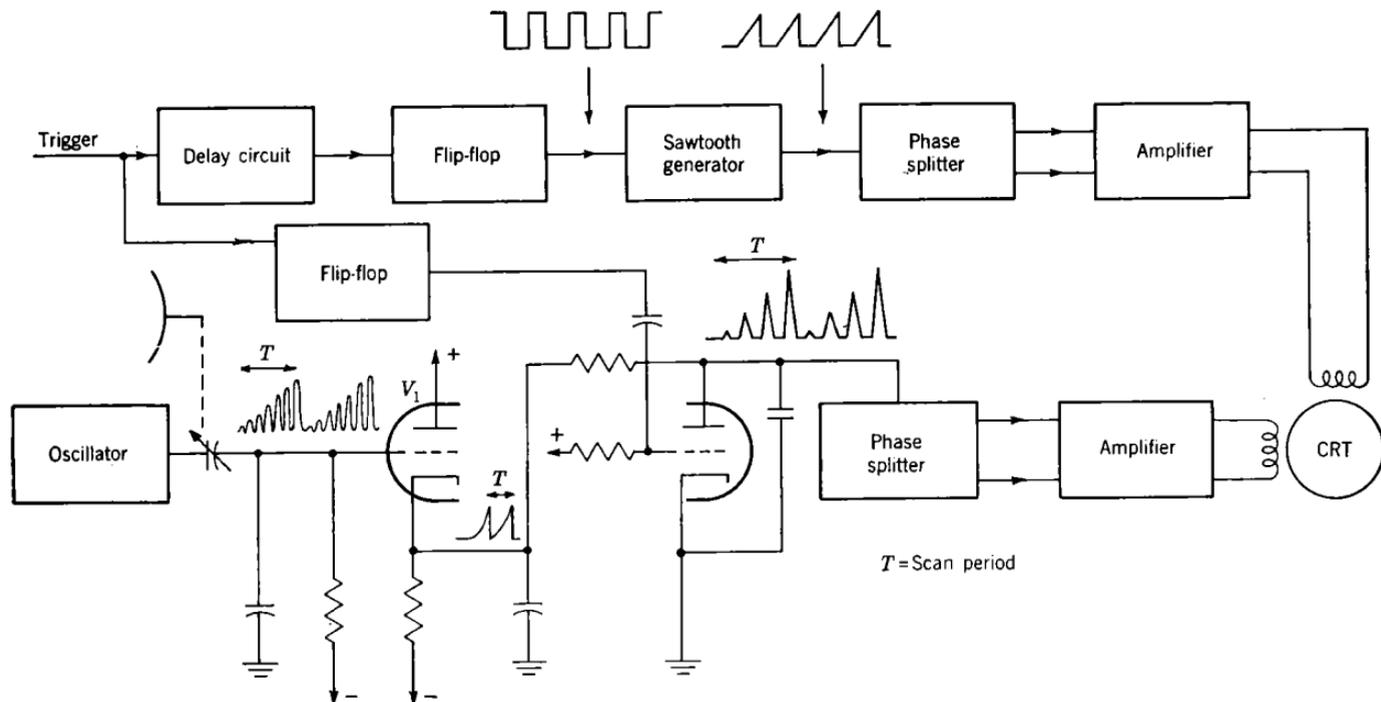


FIG. 13-52.—Range-height indicator for rapid-scan system.

$\cos \theta \approx 1$ are sufficiently good approximations. In Fig. 13-52, information for the vertical sweep is derived by modulating a high-frequency oscillation by a variable condenser, using the voltage-dividing circuit of Fig. 13-7. This is peak-detected by the cathode follower V_1 , whose cathode circuit has a time constant that is long compared with the period of oscillation. The cathode and grid of V_1 are so biased that the former is at ground potential in the absence of signal. The demodulated voltage controls a sawtooth generator of the ordinary type, thus providing sawteeth modulated in proportion to the elevation angle. These are passed through a phase splitter and a push-pull amplifier to the vertical deflecting coils. Since the sweep is always in the same

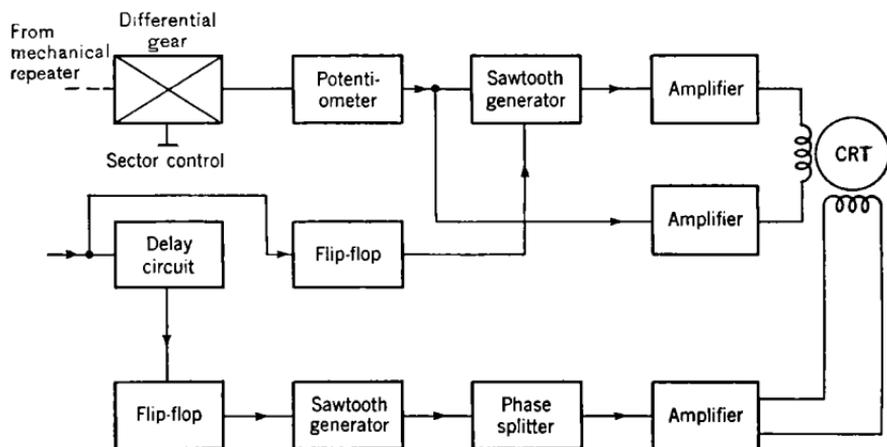


FIG. 13-53.—V-beam height indicator.

direction, d-c restorers provide adequate clamping on the output stages. The horizontal sweep, of constant amplitude, is generated in the ordinary way and amplified as was the vertical one. Off-centering is provided by proper biasing of the driver tubes. A delay can be used in the horizontal circuit to allow a distant range interval to be examined in detail. (If this is not a requirement, the same flip-flop can be used for both sawtooth generators.) A controllable height marker can be derived, if desired, by applying the modulated sawtooth to a biased-diode delay circuit, as explained in Sec. 13-12.

V-beam Height Indicator.—In the *V-beam* height indicator (Fig. 13-53), a linear potentiometer turned by a mechanical repeater provides the azimuth data, a differential gear allowing control of the sector displayed. The vertical deflection is obtained by adding, in the driving amplifiers, a slowly varying voltage proportional to the azimuth angle measured from the edge of the sector chosen and a range sweep modulated

by the same azimuth voltage. The horizontal sweep is unmodulated but arranged to permit delay in discrete steps. Height is read from an engraved scale as explained in Sec. 6·12. If several scales are to be used in connection with delayed sweeps, the reflection method of Fig. 7·4 is used.

SIGNAL DISCRIMINATION, RESOLUTION AND CONTRAST

The factors that influence the discernibility of a simple radar echo signal against thermal noise have been discussed in Chap. 2 and in the early sections of Chap. 12. We have seen that the characteristics of the receiving equipment which influence this matter most profoundly are the noise figure and bandwidth of the receiver, the integrating properties of the cathode-ray tube screen, and the type of display. To a lesser extent the cathode-ray-tube spot size and the scale-factor of the display also play a role. Section 12·8 points out methods of increasing the signal discernibility in the presence of various forms of radiant interference by taking advantage of differences between the desired and the undesired signals, and shows that to some extent these same techniques aid in distinguishing a point target from more diffuse ones. Chapter 16 describes an extremely elegant method by means of which a given target can be readily distinguished from its surroundings, providing it is in motion with respect to those surroundings and that they are at rest or have a uniform mass motion.

The following paragraphs will discuss another aspect of the discrimination problem—that of the presentation of a multiple or complex picture when it is not desired to suppress any of the information but rather to present it in as detailed a manner as possible.

13-20. Resolution and Contrast.—In order to distinguish between objects or among different parts of the same object it is necessary that they be resolved. The characteristics of the receiving equipment which influence resolution are the bandwidth of the signal channel, which affects range resolution, and the spot size and scale factor of the display, which also affect range resolution and to a lesser extent angular resolution as well. Since factors concerned with the display are least under control, and since their limitations often fix the limit of *useful* bandwidth, they will be discussed first.

The spot size of a cathode-ray tube depends upon the tube type, the performance of the individual tube, the voltage employed, the design of focus and deflecting coils, and the intensity level at which the tube is operated. In the series of magnetic tubes used for radar, spot size is roughly proportional to the tube radius and is usually such that between 150 and 200 spots can be resolved along the display radius.

In this discussion, the number of resolvable spots in a radius will

be taken as 180. On a range sweep of length R nautical miles, the number of radar pulse lengths resolvable in principle is $12 \cdot 2R/\tau$, where τ is the pulse length in microseconds. Thus, on a centered PPI, the fundamental pulse-length resolution and the spot-size resolution are equal when $12 \cdot 2R/\tau = 180$, or $R \approx 15$ when $\tau = 1$. Accordingly, on a set with a 1- μ sec pulse, a 15-mile centered PPI will have the full inherent range resolution; a 100-mile PPI will reduce it approximately sevenfold, and so on. This reduction in resolution can be overcome by the use of expanded displays, but often this measure restricts the field of view.

Fortunately for the indicator designer, however, this is not the entire story. For one thing the operator cannot, except on the fastest sweeps, realize the fundamental range resolution without optical aids even if the display were to make it available; indeed, on a 5-in. tube the spot-size resolution for a 100-mile sweep is nearly as good as can be comfortably used, although on a 12-in. tube it is three or four times worse. Second, range resolution alone is of limited usefulness if it is accompanied by poor angular resolution; here the fundamental limitation is usually not the indicator but the azimuthal beamwidth. The angular width of the CRT spot on any PPI is given in radians by $R/180r$, where R is the radius of the tube and r is the distance from the range origin to the spot in question. This width will be equal to the beamwidth θ , measured in degrees, when $57R/180r = \theta$ or $R/r = 3\theta$. Thus, even for a 1° beam, the display resolution exceeds the fundamental resolution for all points farther from the origin than one-third of a tube radius. The cathode-ray tube is usually not the limiting factor in over-all *angular* resolution, although in certain cases of accurate range measurements or of observations on groups of aircraft or ships at long range, better *range* resolution on long-range displays would be useful. Frequently, however, the need for high *dispersion* requires expanded displays quite apart from question of resolution.

The above discussion shows that in general the cathode-ray tube imposes a severe restriction in range resolution on any but the fastest sweeps. It is therefore apparent that unless fast sweeps are to be used there is no point in going to extreme bandwidths in the receiver. In most cases the receiver bandwidth need not be so great as to impair seriously the signal-to-noise discrimination. It should be borne in mind, however, that with too small a bandwidth, a signal far in excess of the limit level has a much greater duration after limiting than has a weak signal.

Contrast.—The contrast of the display depends upon the characteristics of the screen and upon the way in which it is excited. Unfortunately present tubes of the persistent type have serious shortcomings in contrast. The screen material has a natural color much like that of the

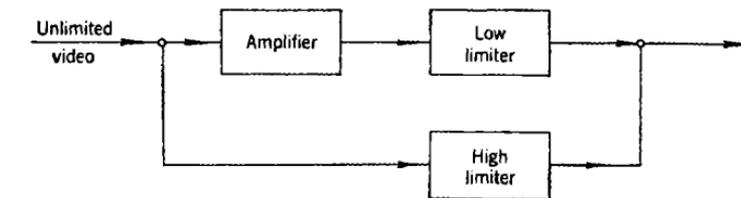
phosphorescent light, so that contrast is reduced by reflected light. The necessary thickness of the screen tends to blur the edges of the signals and thus reduce contrast gradient, as do also limitations in the sharpness of focus of the electron beam. Many screens also have a certain amount of graininess which contributes to the reduction of contrast gradient.

Contrast between signals of different intensities is especially important in overland flying, where it is essential to be able to distinguish land from water and at the same time to have good contrast between weak echoes from ordinary terrain and the much stronger ones from built-up areas. Unfortunately the range of echo intensities involved is so great that, if the receiver has gain high enough to render the terrain visible, the signals from built-up areas must be drastically limited to avoid "blooming" of the cathode-ray tube, and detail is destroyed. If the gain is low enough to preserve detail, the land background is not visible.

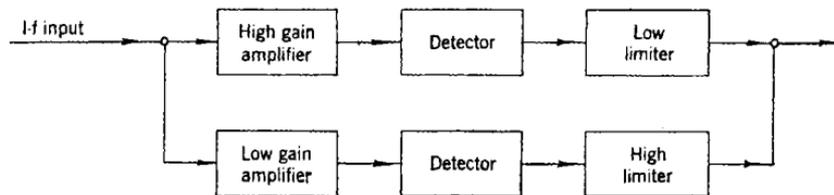
This difficulty is made worse by the nonlinear response of the cathode-ray tube. As was previously mentioned, the beam current in a magnetic tube is proportional to the cube of the grid-voltage swing measured from cutoff, which is the operating point in the absence of signal. This means that a signal of, for example, half the voltage of a limited signal will, in principle, give only one-eighth as much light intensity. In actual fact the difference is not so great as this, because of the tendency of the screen to saturate on strong signals. This nonlinearity reduces considerably the difference in voltage between those signals that are just strong enough to produce a visible result and the level at which strong signals must be limited. It is possible to introduce a nonlinear element in the video amplifier to compensate for the tube characteristic, and this has been done with some success. No really satisfactory circuit of this sort has been devised, no doubt because the need for it was fully appreciated only recently. Even with a linear CRT grid characteristic, the problem of narrow dynamic range would be a serious one. General methods of attacking it will now be described.

13-21. Special Receiving Techniques for Air-to-land Observation.—

Some improvement in the ability to distinguish land from water and at the same time to see detail in built-up areas can be effected by using the anticlutter techniques described in Sec. 12-8. Fast time constants, instantaneous automatic gain control, and other such circuits tend to suppress saturation on long blocks of signals and to make full use of changes in the signal intensity. Unfortunately the use of such measures results in an unnatural appearance of the display which makes interpretation exceedingly difficult. Shorelines, for example, stand out very strongly, as do changes in the terrain. Similarly a solidly built-up area appears chiefly in outline, with other strong signals at points of changing signal intensity. Consequently, although they are extremely



(a) Split video channel



(b) Split i-f channel

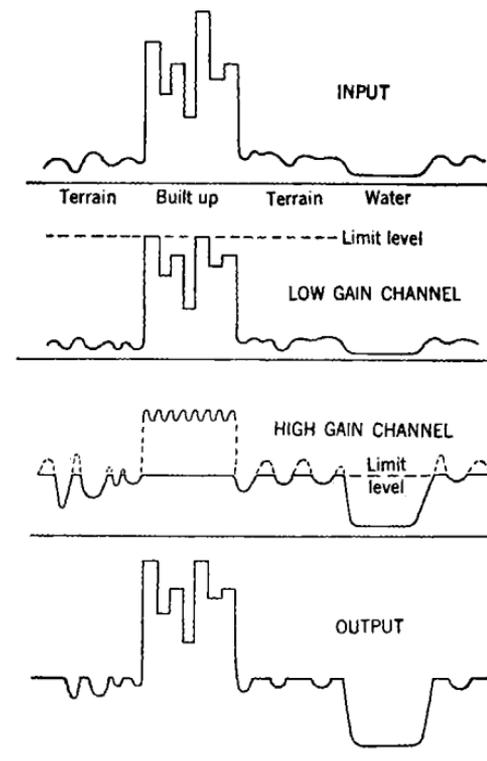


FIG. 13-54.—Three-tone PPI.

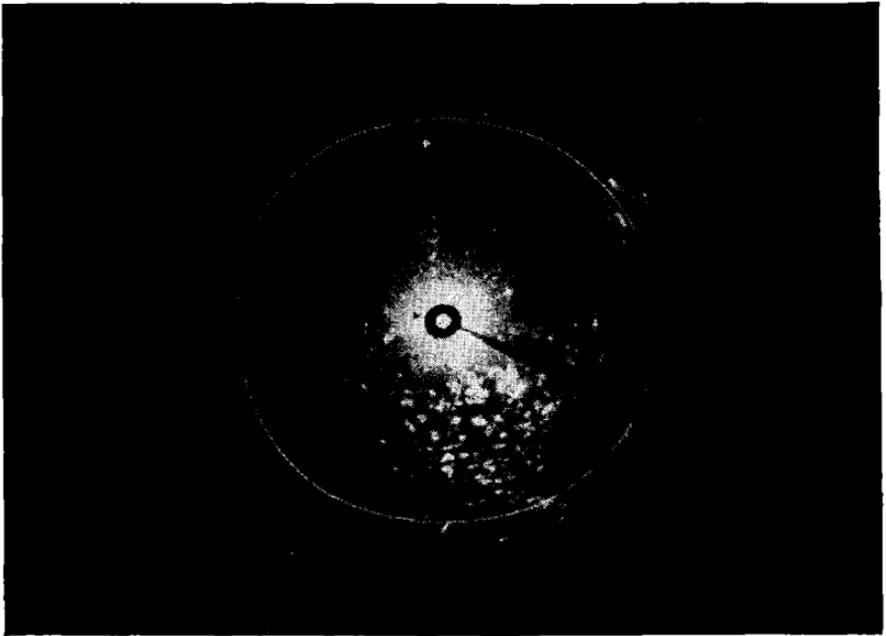
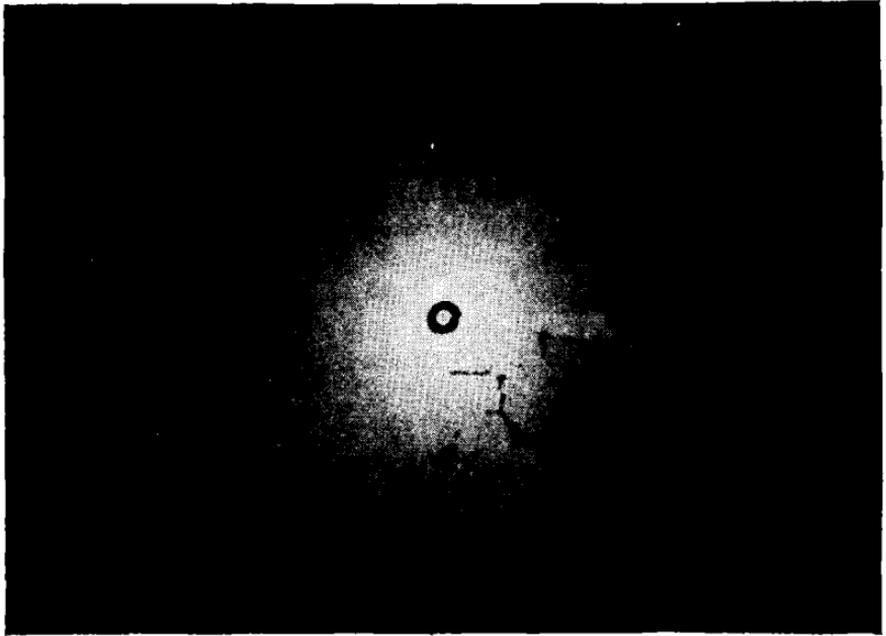


FIG. 13-55.—Effect of three-tone PPI. (a) AN/APS-15 scope photo of Boston without three-tone; (b) AN/APS-15 scope photo of same area with three-tone.

valuable for certain purposes, anticlutter measures leave much to be desired in the situation under discussion.

The "Three-tone" Method.—A simple but extremely effective method of retaining detail on high-intensity echoes while still providing land-to-water contrast is illustrated in Fig. 13-54. At some point in the receiver, ahead of any saturation or limiting, the signals are split into two channels, in one of which they are amplified much more than in the other. If the signals are not already in video form, each channel includes a detector. In any case, each contains a limiter, the one in the low-gain channel being set at approximately twice the level of the other. The two sets of signals are mixed beyond the limiters and then passed through a standard video amplifier. As the waveform diagrams show, the result is to enhance the strength of the land echoes without destroying the detail

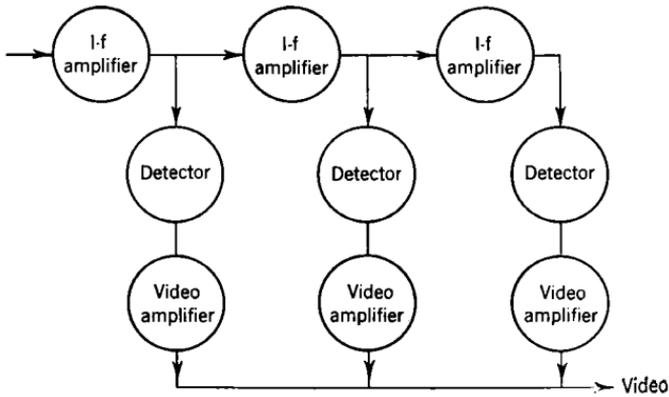


FIG. 13-56.—Logarithmic receiver.

in those from the built-up areas. A "before and after" picture is shown in Figure 13-55.

Many variations of this technique are possible. One, which has been applied to existing sets in which an additional channel is difficult to install, consists of switching the gain up and the limit level down on alternate pulse cycles by means of a scale-of-two multivibrator, the results being added on the cathode-ray-tube screen. The results are almost indistinguishable from those obtained by the use of two channels.

The Logarithmic Receiver.—A more involved method of increasing the dynamic range of the system is by use of the so-called "logarithmic" receiver, whose gain characteristics are such that the response varies approximately as the logarithm of the input signal. One method of attaining such a response, which is actually an extension of the three-tone method, is illustrated in Fig. 13-56. At low signal levels the principal contribution is from the last stage, where the over-all gain is highest.

As the signal level increases, the last stage saturates and can contribute no more. By this time, however, the contribution from the preceding stage has become appreciable and the output continues to increase, but at a lower rate than that occurring when the gain of the last stage was included. Eventually the second stage saturates, and the gain drops to that from the next preceding stage, and so on. Thus the total signal range is compressed, but there is no complete saturation even on very strong signals. Such receivers have not had very extensive tests with conventional airborne radar, but the rather meager results have shown promise. One variant that has not been tried, to the author's knowledge, is to use different saturation (or limit) levels in the various channels, as in the three-tone method, making this a "multi-tone" device.

CHAPTER 14

PRIME POWER SUPPLIES FOR RADAR

BY M. M. HUBBARD AND P. C. JACOBS

Radar equipment requires for its operation high-voltage (1000- to 10,000-volt) direct current, medium voltage (200- to 600-volt) direct current, and power for vacuum-tube heaters which usually can be either alternating or direct current. Normally, the desired d-c voltages are obtained from an a-c supply by the use of transformers and rectifiers, either of the thermionic or of the dry-disk type. Power supply design varies with the application, and differs for aircraft installations, ground system installations, and ship installations.

AIRCRAFT SYSTEMS

Radar power in aircraft should, if possible, be obtained from directly-generated alternating current. Where this is not practicable, means exist to convert low-voltage direct current from the aircraft electrical system to the desired type of supply voltage.

TABLE 14-1.—RECOMMENDED POWER SOURCES

Radar load, watts	Recommended radar power source	Alternate choice
Less than 150.....	Vibrator	Dynamotor
150-250.....	400 cps 1 ϕ motor-alternator	Dynamotor*
250-750.....	Direct engine-driven 1 ϕ alternator 400 cps or higher	400 cps 1 ϕ motor-alternator
750-10,000.....	Direct engine-driven 3 ϕ alternator 400 cps or higher	For ranges up to 2300 watts, 400 cps 1 ϕ motor-alternator

Aircraft electrical systems today are normally 12-volt or 24-volt direct current. A few aircraft have 110-volt d-c supply. Ratings are nominal; a "24-volt" aircraft system usually supplies 27.5 ± 2.0 volts.

14-1. Choice of Frequency.—One of the most controversial points in early radar development was the selection of the a-c frequency to be used as primary radar power. A transformer correctly designed for the frequency f will behave tolerably at frequencies up to $3f$, but will burn

up at a frequency of $f/2$. The general equation for induced voltage in a transformer

$$E = k \cdot n \cdot \frac{d\phi}{dt} \quad (1)$$

gives, for a fixed core area,

$$E = k \cdot n \cdot A \frac{dB}{dt}, \quad (2)$$

where k is a constant, n the number of turns in the winding considered, ϕ the flux linking these turns, A the area of the core, and B the magnetic induction in the core. For sinusoidal voltages of frequency f ,

$$E_{\max} = k \cdot n \cdot f \cdot A \cdot B_{\max}. \quad (3)$$

Thus if f is halved, B_{\max} must double for a fixed input voltage, all other terms being constant in any given case. Since losses in the iron are a function of B_{\max} , doubling B_{\max} produces an immense increase in magnetizing current and in core loss. The core area and thus the weight of the transformer as a whole can be decreased as f increases because of the corresponding reduction in B_{\max} . By going from 60 cps to 400 cps, savings in transformer weight of 50 per cent are attainable. Core area cannot be much further reduced by increasing f from 400 cps to 800 cps, as iron losses per pound for constant B_{\max} increase more than linearly with frequency. Weight savings of 10 per cent at most are obtainable by this change in frequency.

Performance of rotating equipment must be considered, as well as that of transformers. For stable performance and minimum weight of alternators and a-c motors, 400 cps is preferable to 800 cps. (See Sec. 14-2 on wave shape.) The majority of existing airborne radar sets are designed for 400-cps minimum supply frequency, and it appears that this will continue to be the nominal standard. Variable-frequency alternators are also used; they commonly give a 2 to 1 ratio between minimum and maximum frequencies, depending on the speed of the aircraft engine. Ranges often encountered are 400 cps to 800 cps, 800 cps to 1600 cps, and 1200 cps to 2400 cps. Table 14-2 shows the power frequencies used in Allied military aircraft during the past war.

TABLE 14-2.—STANDARD POWER FREQUENCIES

User	Frequency, cps
USAAF.....	400 fixed
USAAF.....	400 to 800 variable
USN.....	800 fixed
USN.....	800 to 1600 variable
USN.....	400 to 800 (3 ϕ) variable
RAF.....	{ 1200 to 2400 variable 1300 to 2600 variable

14-2. Wave Shape.—Since all electric power used for radar is converted to direct current or is used for heating, there is no inherent requirement for sinusoidal wave shape.¹ However, an alternator furnishing voltage of poor waveform may show very marked changes in waveform when its load is changed. Since a stable d-c voltage is the final requirement in most radar power supplies, it is necessary to maintain a fixed relationship between the maximum voltage and the rms voltage, whose ratio is called the “crest” or “amplitude factor.” The output d-c voltage f or rectifier circuits employing condenser input will be a function of the crest voltage. For rectifiers employing choke input, the ratio of rms voltage to average voltage is important; this ratio is called the “form factor.” These two ratios should be as constant as possible under all load conditions. This condition is more easily met by the use of alternators with low subtransient reactance. Practically, it is preferable to employ conventional salient-pole rotating-field synchronous alternators designed for 400 cps, rather than high-impedance inductor alternators operating at higher frequencies. Although the reactance of an inductor alternator can be neutralized by a series capacitance (as is done in the Model 800-1-C Bendix 800-cps motor-alternator) such balance is completely effective only at one value of a load exhibiting fixed power factor. Changes in load cause changes in wave shape. Thus, even if the voltage regulator (Sec. 14-5) behaves perfectly and maintains constant E_{av} or E_{rms} , the output d-c voltage applied to the radar may vary excessively. (See Fig. 14-1.)

Furthermore, individual machines of the inductor-alternator type show wide variation in characteristics. Crest factors measured for different machines of a given type have shown variations in the range from 1.15 to 1.75. When power is to be derived from an alternator of this type, transformers for radar rectifiers should be provided with primary taps, and filament supplies should be obtained from separate transformers.

14-3. Direct-driven Alternators.—Since all power is ultimately derived from the rotation of the aircraft engine, fewer devices that can cause trouble are required if mechanical rotation is used directly to drive an alternator. A direct-driven alternator will normally furnish alternating current whose frequency varies with engine speed. This is usually satisfactory for radar operation. In some complex components, such as equation-solving circuits, and for synchro applications, fixed frequency is necessary. In these cases it is often desirable, particularly where the a-c loads are heavy, to obtain that part of the load which can be allowed to vary in frequency from an engine alternator, using a small inverter set to obtain the fixed-frequency power needed.

¹One exception to this statement should be mentioned. In an a-c resonance charge modulator, the generator must produce a good sine wave.

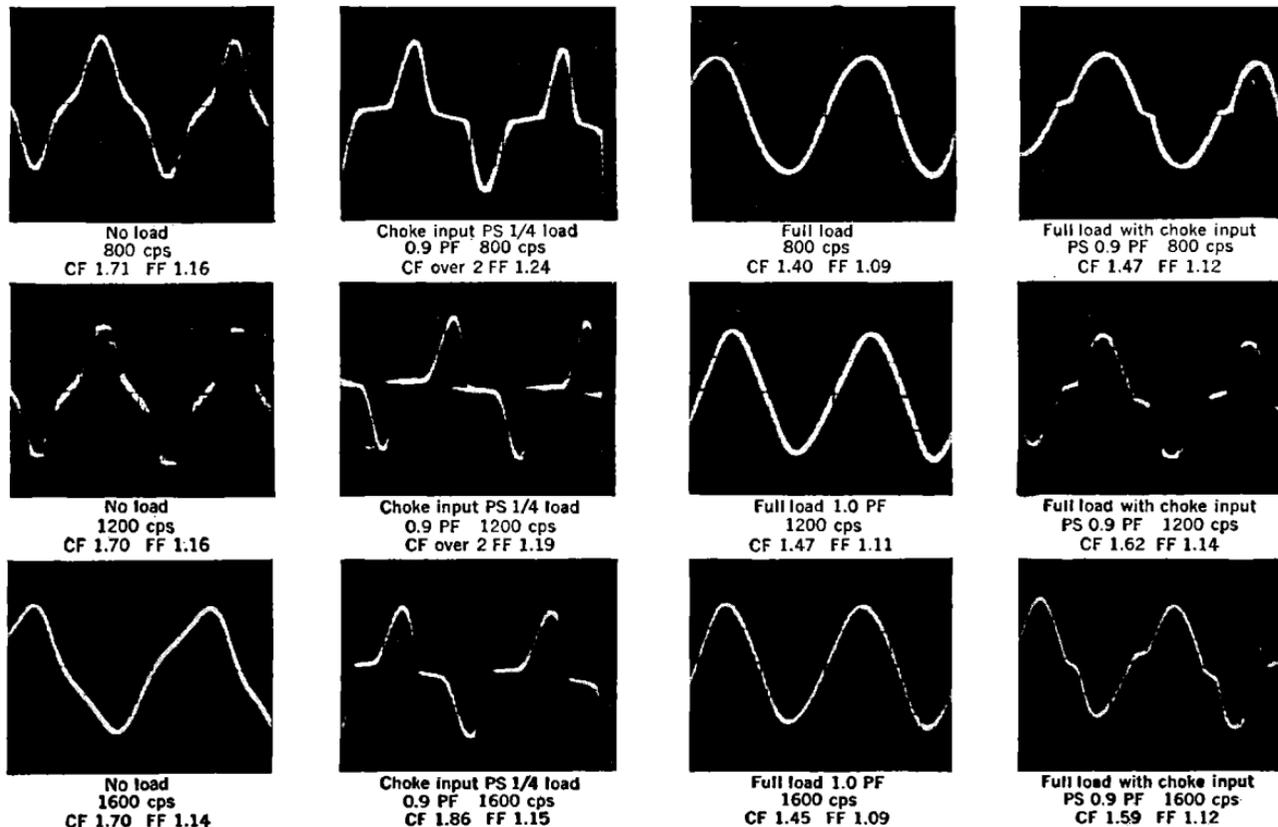


Fig. 14-1.—Voltage wave shape of Eclipse NEA-6 generator. CF—crest factor; FF—form factor.

For loads approaching 1 kw, particularly in the case of systems with heavy current demands, weight savings can be appreciable if 3-phase generation is employed. No saving in transformer weight is usually so

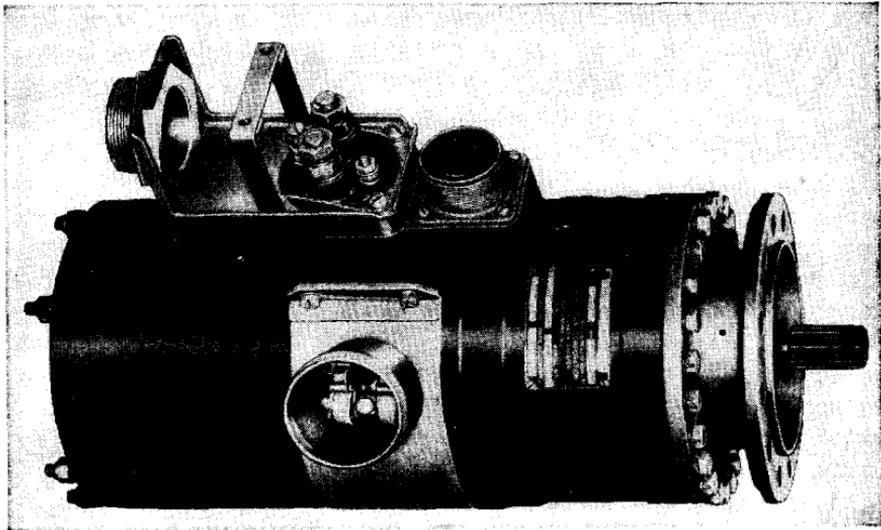


FIG. 14-2.—Eclipse NEA-7 dual-voltage generator.

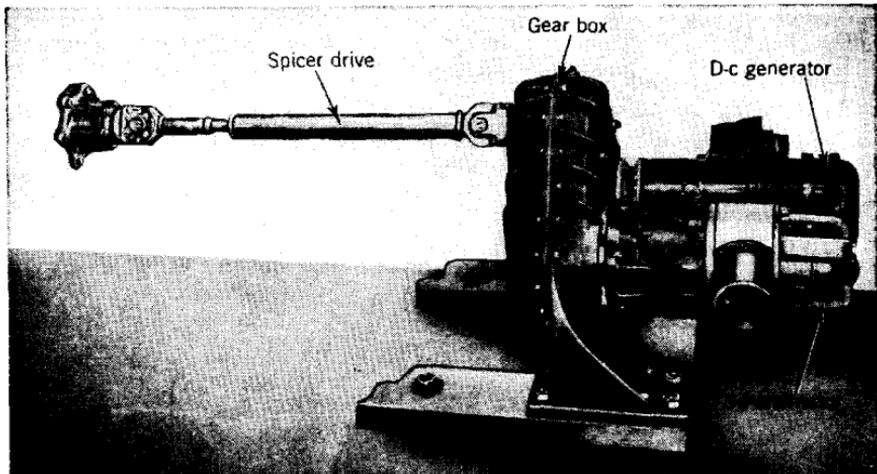


FIG. 14-3.—General Electric dual-output gear box and generators for 3-phase a-c radar power and aircraft d-c supply.

obtained, but the alternator is lighter or more reliable, and filters for the higher ripple frequency are appreciably smaller and lighter. In one specific case, a 3-phase power supply furnishing 3-kw direct current showed a saving of almost 70 lb over the equivalent single-phase unit.

In many cases, a limit on the number of engine mounting pads will force the use of a dual-purpose generator such as the Bendix Aviation Corporation NEA-7, which delivers 125 amp of 27.5-volt direct current and also, independently controlled in voltage, 2500 va of single-phase, 115-volt alternating current at 800 to 1600 cps for shaft speeds of 4000 to 8000 rpm. (See Fig. 14-2.) A typical example of a large 3-phase direct-drive alternator is the General Electric Company 10-kva 3-phase 208/120-volt 400- to 800-cycle machine shown in Fig. 14-3. This photograph shows the installation as designed for use in TBM-3 aircraft, using an auxiliary gear box to obtain dual output from the TBM-3's single engine pad. This expedient is usually much more successful than attempting to convert (with low efficiency) direct current to large "bites" of alternating current. When pressed for output mounting-pad space it is well to consider such "dual outlet" gear boxes. Often such a box can be employed with older engines (2000- to 4000-rpm jack-shaft speed) to utilize high-output 4000- to 8000-rpm alternators by incorporating a step-up ratio in the box. Such a gear box was designed to permit use of the same General Electric Company generator with the low-speed jack shafts of the B-17-G aircraft engines.

Large aircraft of the future will probably have electrical systems furnishing 3-phase 400-cps fixed-frequency power at 208/120 volts (120-volt phase to neutral, 208-volt phase to phase). This will give 200/115 bus voltages. Preliminary tests indicate that such systems, especially those having several large generators in parallel, will be stable, reliable, and directly usable for radar.¹

Summary of Alternatives to Direct-driven Alternators.—If direct-driven alternators cannot be employed in any given case, or if the power requirements are so small as to render a special generator impractical, there are five alternatives:

1. Motor-alternator sets. D-c motor input, 400- or 800-cps output.
2. Inverters. Same input and output as (1) but with a single magnetic circuit—a true inverter.
3. Dynamotors. Low-voltage d-c input, high-voltage (or voltages) d-c output.

¹ H. E. Keneipp and C. G. Veinott, "A 40-kva, 400-cycle Aircraft Alternator," *AIEE Transactions*, **63**, 816-820 (November 1944).

L. G. Levoy, Jr., "Parallel Operation of Main Engine-driven 400-cycle Aircraft Generator," *AIEE Transactions*, **64**, 811-816 (December 1945).

W. K. Boice and L. G. Levoy, Jr., "Aircraft Alternator Drives," *AIEE Transactions*, **64**, 534-540 (July 1945).

M. M. Hubbard, "Investigation of 3-phase 208/120-volt 400-cycle Aircraft Alternators," RL Group Report 56-061545.

4. Mechanical vibrators. Low-voltage d-c input, output direct current or 120-cps alternating current or any combination.
5. Electronic inverters. Low-voltage d-c input, output as desired.

The options (2) and (5) can be discarded at once. Electronic inverters, although theoretically the most desirable, are practical only for power levels of a few watts because of low efficiency arising from tube plate-drop. Inverters are usually undesirable because they cannot be regulated. This leaves motor-alternators, dynamotors, and vibrators. Practical limitations on commutator diameter, voltage per bar, and operation at altitude restrict the use of dynamotors. For example, it is never wise to attempt to use dynamotors with outputs above 1200 volts. Motor-alternators with voltage regulators are the most reliable conversion means.

14-4. Motor-alternator Sets.—Table 14-3 gives common sizes of motor-alternator sets now in use.

TABLE 14-3.—COMMON SIZES OF MOTOR-ALTERNATOR SETS

Capacity, va	Frequency of output, cps	Phases	Weight, lb	va/lb
100	400	3
250	400	1 or 3	13	19.2
500	400	1	23	21.8
750	400	1 or 3	38	19.8
840* (1000)	800	1	32†	26.2 (31.5)
1500	400	1	55	27.2
2500	400	1	75	33.4

* 840 is nameplate rating. For operation below 10,000 ft with blast cooling, 1000 va are obtainable.

† Includes weight of external r-f filter and 12- μ f series condenser. See Figs. 14-8, 14-16, 14-17, 14-21 for pictures of some commonly used 400-cycle motor-alternators.

For heavy-current duty, particularly at high altitude, brush wear becomes a problem. Care should be taken to see that all rotating devices using brushes requiring operation at altitudes above 20,000 ft have suitable "high-altitude" brushes.¹

Starting currents are often a serious problem in the case of motor-alternator sets rated at 750 va or more since they may be 5 to 10 times

¹ L. M. Robertson, "Effect of Altitude on Electric Apparatus," *Electrical Engineering* (June, 1945).

Paul Lebenbaum, Jr., "Altitude Rating of Electric Apparatus," *Transactions AIEE*, **63**, 955-960 (December, 1944).

D. Ramandanoff and S. W. Glass, "High-altitude Brush Problem," *Transactions AIEE*, **63**, 825-829 (November, 1944).

Howard M. Eelsey, "Treatment of High-altitude Brushes," *Transactions AIEE* (August, 1945).

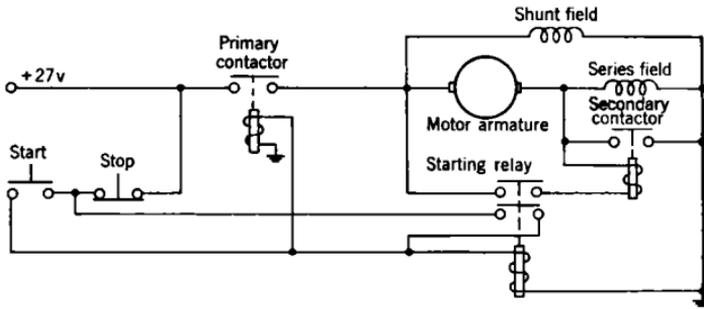


FIG. 14-4.—Motor circuit of 1500-va and 2500-va inverters showing starting relay and contactors.

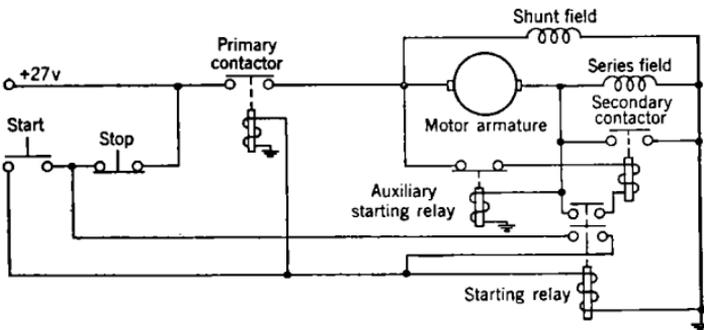


FIG. 14-5.—Addition of auxiliary starting relay to circuit of Fig. 14-4.

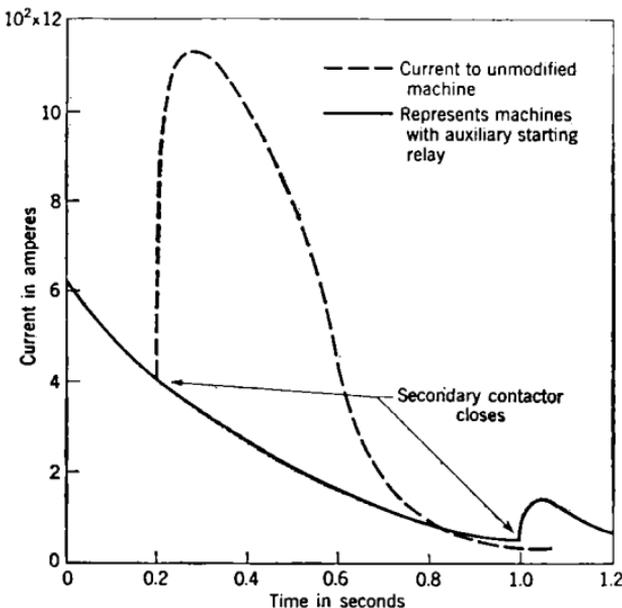


FIG. 14-6.—Starting current of 1500-va motor-alternator, Signal Corps Type PE-218-D. No load on alternator.

the normal full-load current. Such high starting currents cause: (1) excessive brush wear, (2) voltage dip on aircraft electrical system, (3) excessive load on contactor points unless an oversize contactor is used. Extremely high currents may cause laminated brushes to explode. Voltage dips on the electrical system can cause the release of vital relays or contactors. To reduce the initial current surge a series resistor (often an auxiliary series motor field) is added. When the motor is up to speed, this resistor is shunted out by a contactor (see Fig. 14-4.) This reduces the initial starting current, but the current again rises sharply if the secondary contactor closes too soon. Closing of the secondary contactor can be further delayed by an auxiliary starting relay connected as shown in Fig. 14-5. The curve of Fig. 14-6 indicates the reduction in starting current effected by the use of the auxiliary starting relay.

14-5. Voltage Regulators.—Since constant output voltage must be maintained despite changes in load, shaft speed, and d-c input voltage, some form of voltage-regulating device is used with an alternator. For radar applications, an electronic voltage regulator is the most satisfactory type. Although it weighs more than a simple mechanical regulator, an electronic regulator should be employed on systems delivering 750 va or more.

Electronic Voltage Regulators.—An electronic voltage regulator consists of a voltage-sensitive element, an amplifier, and an output stage that supplies d-c excitation for the generator or alternator field.

Three types of voltage-sensitive elements have been used in experimental regulators. One consisted of a VR-tube bridge excited from a transformer and rectifier connected to the output of the alternator to be controlled. A suitable filter could be added to the rectifier so that regulation was performed with respect to the peak value of the output waveform. Similarly, other types of filters could be used to regulate with respect to the average value of output voltage, or to some chosen value between peak and average. The filter introduces a time constant that, in some cases, is too long to achieve the desired rate of response.

The second voltage-sensitive element, developed by Bell Telephone Laboratories, was a bridge network made of thermistors and excited from alternating current. This gives an a-c output error voltage that can be readily amplified, but has the disadvantage of a relatively long time constant. The thermistor bridge, though rather difficult to compensate for wide variations in temperature, regulates to the rms value of the output wave, which is a considerable advantage for some applications.

The third form of voltage-sensitive element was a tungsten-filament diode operated in the region of saturated emission. The filament was heated by the alternator output through a transformer, and the anode

excited from the d-c power supply. The plate current varies in proportion to nearly the fourth power of the rms value of the alternator output voltage. This provides a very sensitive signal voltage with a very short time constant if the filament is of small diameter. Such a diode must be ruggedly constructed in order to maintain the mechanical spacing of the tube elements, and thus the tube characteristics, even under severe mechanical vibration.

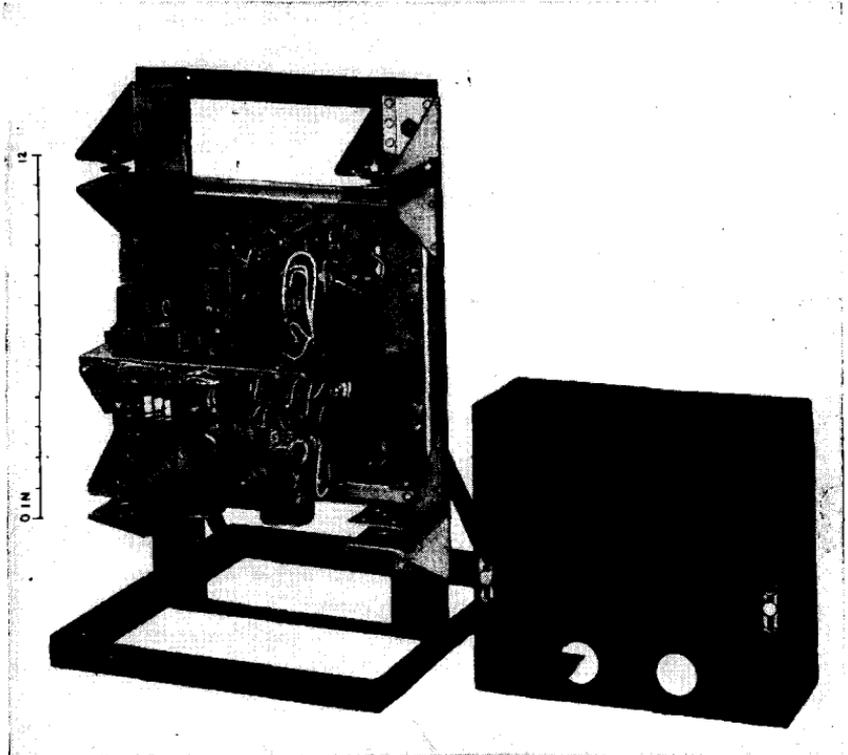


FIG. 14-7.—KS-15055 electronic voltage regulator designed by Bell Telephone Laboratories.

A conventional amplifier is used which must have sufficient gain to provide the required over-all sensitivity. Stable operation of the combination of voltage-sensitive element, amplifier, output stage, and alternator demands that the usual conditions for stability of a servo system be fulfilled.

Direct current excitation for the alternator field is provided by a controlled rectifier fed from the alternator output. The controlled rectifier may be a combination of transformer and rectifier controlled by a saturable reactor, or it may have grid-controlled thyatrons for rectification.

Bell Telephone Laboratories developed a regulated exciter, type KS-15055, which weighed about 12 lb complete and was meant for use with the PE-218 inverter. (See Fig. 14-7.) It regulated the 115-volt output to about ± 0.5 volt rms, but had a rather slow response and only fair temperature compensation. The temperature compensation was improved in a later model.

General Electric Company developed a similar regulator, type 3GVA10BY1, for use with PE-218-D inverters; it used a saturated-diode voltage-sensitive element. (See Fig. 14-8.) This regulated the 115-volt output to about ± 0.1 volt rms and had good response and temperature characteristics, but the diode was sensitive to vibration.

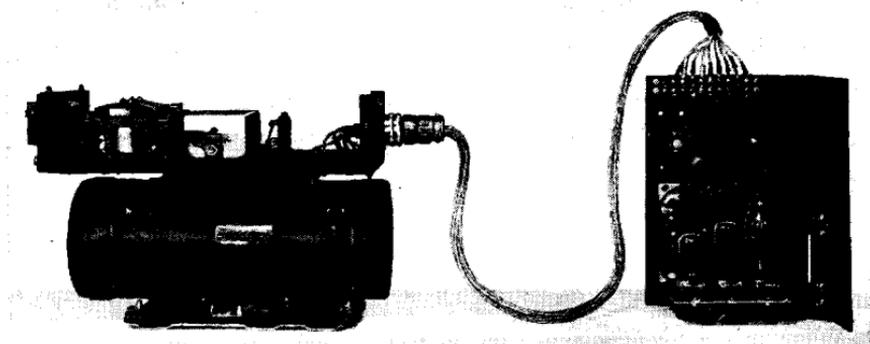


FIG. 14-8.—General Electric 1500-va aircraft inverter type PE-218-D, 115 volts, 400 cps, and electronic voltage regulator.

In the latter part of 1944, more interest was shown in the development of electronic regulators because of a trend toward the use of engine-driven alternators that were too large for control by carbon-pile regulators unless separate exciter generators were used. The ATSC Equipment Division at Wright Field sponsored the development of two regulators to control and excite engine-driven alternators rated 8 kva, 1-phase, 400 to 800 cps, 115 volts a-c. One developed by Bell Laboratories weighed 33 lb; its operation was reported to be very satisfactory.

In 1945 Radiation Laboratory undertook the development of regulators using a saturated-diode voltage-sensitive element and grid-controlled thyratrons in the output. The intention was to perfect a more or less universal regulator which could be applied to different alternators by changing the thyatron output. The development had progressed to the preliminary test model stage for two types; one type for single-phase inverters which had been applied to the PU-7 2.5-kva machine, and one type for 3-phase engine-driven alternators of 6.5- to 12-kva rating. The results were very promising, and it appears quite

certain that satisfactory electronic regulators can be developed for stationary ground generating equipment as well as for airborne equipment.

Mechanical Voltage Regulators.—All mechanical voltage regulators are prone to improper operation in the presence of vibration since they are sensitive devices that balance a spring against a magnetic force. A mechanical voltage regulator consists of two principal elements: (1) a variable resistance in series with the generator field, and (2) a control device, operating from the generator output voltage, which alters the

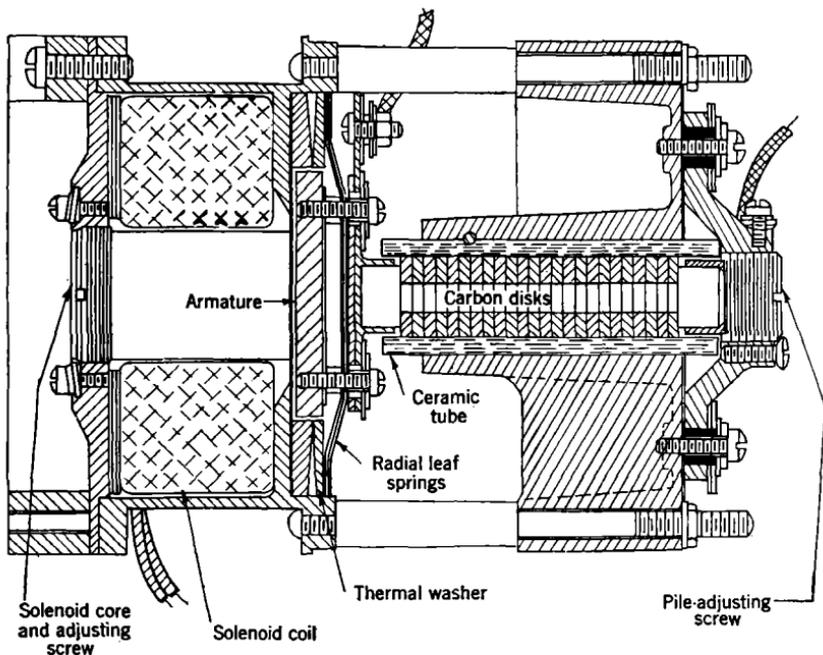


FIG. 14-9.—Cross section of carbon-pile voltage regulator.

variable resistance to maintain constant output voltage. The two types of regulators commonly used in aircraft are the carbon-pile and the finger types.

Carbon-pile Regulators.—The carbon-pile regulator (Fig. 14-9) was first developed in England, where it is known as the Newton regulator after the inventor. The variable resistance element consists of a stack of carbon disks or annular rings placed in a ceramic or tempered-glass cylinder. The cylinder is mounted in a metal housing which serves to dissipate the heat from the pile and gives structural support. One end of the pile rests against a button held in place by radial leaf springs;

the other end is retained by a screw usually referred to as the pile screw or pile-adjusting screw.

The control element of the regulator is a solenoid coil and armature. The current through the solenoid coil is proportional to the voltage to be regulated, and the pull of its core on the armature is proportional to the air gap between the two, and to the coil current. The air gap is adjustable by movement of the core; this constitutes one of the adjustments to be made on the regulator.

With the core in fixed position and the pile-adjusting screw tight, the output voltage of the generator will be high. As the pile screw is loosened the voltage will drop, reach a minimum point, rise, and then drop again as shown in Fig. 14-10. The generator will regulate properly

to the left of the hump or to the right of the dip. Between these two points, that is, on the downward slope of the curve, regulation will be unstable. The slope of the curve to the left of the peak is less than that to the right of the valley, giving better regulation. However, in this region the carbon pile is under less pressure, and is thus more suscepti-

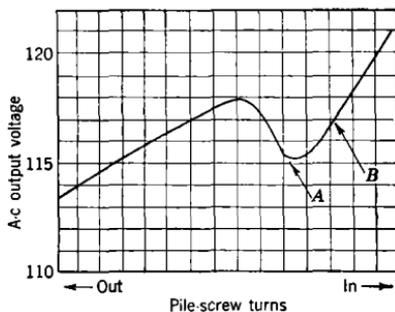


FIG. 14-10.—Carbon-pile regulator adjustment. Point *A*—bottom of dip—115 volts. Point *B*—optimum operating point—117 volts.

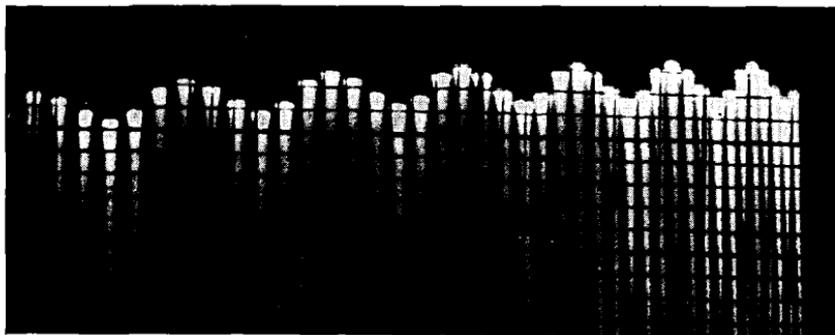


FIG. 14-11.—Output waveform showing amplitude modulation.

ble to mechanical vibration and shock. Most recent practice, therefore, is to adjust the regulator to operate on the right-hand side of the dip.

When carbon-pile regulators were first used on inverters they were mounted directly on the rotating machine. This gave rise to serious troubles because of the susceptibility of the regulator to vibration. Vibration of the carbon pile causes amplitude modulation or "jitter" of

the a-c voltage, which in a radar set shows up as spoking and blurring of the indicator scope (see Fig. 14-11). The frequency of the modulation is normally 30 or 40 cps. The maximum "jitter" which can be tolerated is approximately 1 to $1\frac{1}{2}$ volts.

The regulator should be shock-mounted as shown in Fig. 14-12. This procedure, together with adjustment of the regulator on the tight side of the dip, helps not only in clearing up the troubles mentioned above,

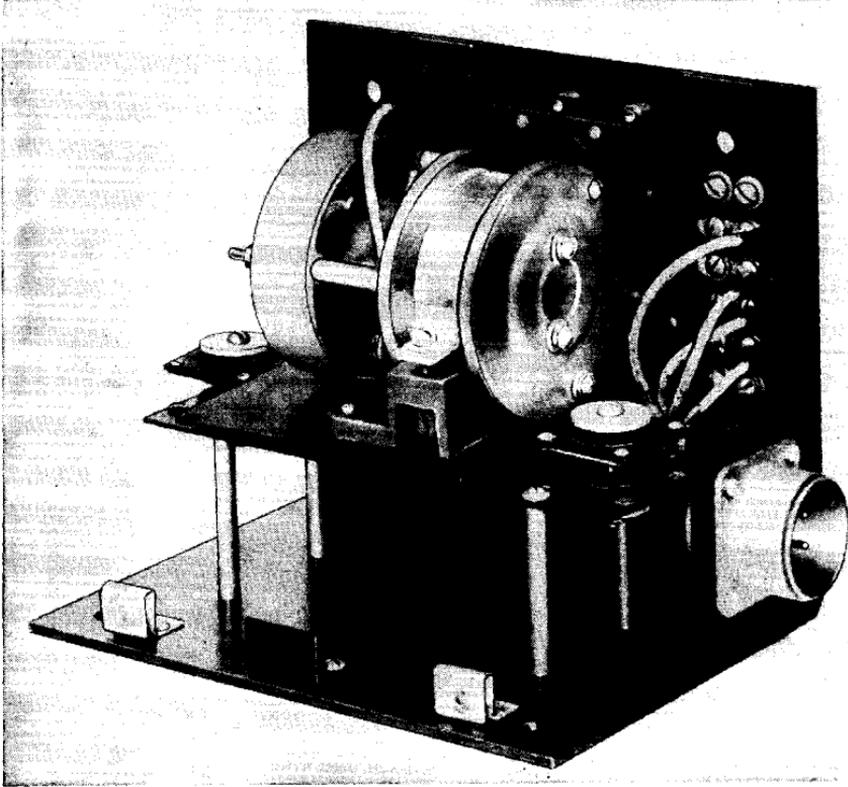


Fig. 14-12.—Typical vibration mount for voltage regulator.

but also in increasing the interval between regulator adjustments, and in reducing the wear of the carbon disks.

The carbon-pile regulator is affected by moisture. Moisture in the carbon pile materially reduces the resistance, and frequently when a motor-alternator is started after having stood for some time in an atmosphere saturated with water vapor the output voltage is high—approximately 135 to 145 volts. After the machine has run for an hour or longer, the moisture is usually driven off and the voltage returns to normal. During this drying-out period the carbon disks are often

burned because, as the film of moisture is gradually reduced, it may break down at one particular point on the face of the disk and the high current-density at that point may be sufficient to burn the disk. In this event, the regulator pile must be removed and the damaged disk replaced.

One fundamental defect of the carbon-pile regulator is the lack of any means for insuring uniform voltage distribution across the stack. Some carbon junctions may be tightly mated, with low voltage drop; others may be loose, with high voltage drop. If the drop per junction exceeds 1 to 2 volts, sparking may occur which will ruin the pile.

Under changing load conditions, the regulator acts to increase or reduce the field current; as in any action of this kind, hunting may result. In order to reduce hunting, a stabilizing transformer (Fig. 14-13) is now used with most of the larger alternators. Such a transformer provides a correction depending on the rate of change of excitation.

The resistance of the carbon pile ranges from 2 to 60 ohms.¹ Its electrical rating is based on the rate of allowable heat dissipation, and is given in watts. The unit often used on aircraft inverters is rated at 35 watts. Smaller inverters, notably the Eclipse 100- and 250-va units, use a smaller regulator rated at 20 watts, while 2500-va inverters use a 75-watt regulator, of the same physical size as the 35-watt unit, but with fins to increase the rate of heat dissipation. For the purpose of using a smaller regulator than would otherwise be possible, the carbon pile can be shunted with a fixed resistor. Such an arrangement increases the range of resistance over which the carbon pile must operate, and care must be taken that the maximum stable range of the pile is not exceeded.²

The coil of the carbon-pile regulator usually used in aircraft inverters has a resistance of 185 ohms and a d-c rating of 115 to 125 ma, which gives a voltage drop of 21 to 23 volts. As the voltage to be regulated is 115 volts alternating current, rectification and voltage-dropping are necessary. Rectification is provided by a selenium dry-disk rectifier, in series with a "globar" resistor which provides some temperature

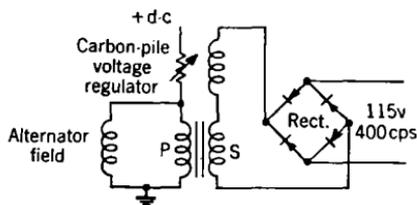


FIG. 14-13.—Schematic diagram of stabilizing transformer in voltage regulator circuit of 400-cps aircraft inverter.

¹ Although the 35-watt carbon-pile regulator can be made to operate over a range of resistance from 2 to 60 ohms, it has been found desirable to use only the range from 2 to 30 ohms since the high-resistance end of the range has a tendency to be unstable.

² W. G. Neild, "Carbon-pile Regulators for Aircraft," *AIEE Transactions*, 63, 839-842 (November, 1944).

compensation. The voltage is reduced either by a small auto-transformer or by a dropping resistor.¹ In any case, a variable resistor is used, either on the a-c or d-c side of the rectifier, to provide a voltage adjustment whose range is approximately 10 volts.

Figure 14-14 shows a carbon-pile regulator for use with a 10-kva, 3-phase, 208-volt engine-driven generator; Fig. 14-15 is a schematic diagram of the circuit. Note the two potential transformers, connected in open delta, which energize the regulating coil, so that the regulated

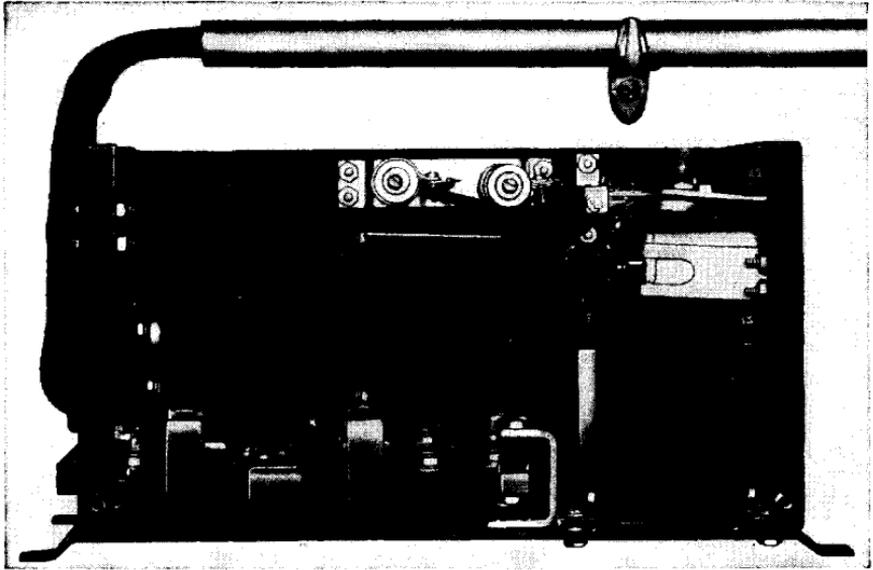


FIG. 14-14.—General Electric Company voltage regulator for 10-kva alternator.

voltage is an average of the three line-to-line voltages. Note also the antihunt circuit and coils on the solenoid.

It is probable that further development of the carbon-pile regulator can greatly extend the usefulness of the device. At the close of World War II, Leland Electric Company, of Dayton, Ohio, had developed experimental models of improved carbon-pile regulators that appeared to have a greatly extended operating range of resistance and improved resistance to humidity as compared to earlier service models.

Finger-type Regulators.—Finger-type voltage regulators are similar in operation to the carbon-pile type, except that the resistance is varied in fixed steps rather than being continuously variable between minimum and maximum resistance. Mechanical construction varies widely with

¹ A third method, employed on the 1500-v-a series PE-218 inverters as well as a number of others, is the use of a low-voltage tap on the armature winding.

different manufacturers. In general, a voltage-sensitive element, consisting of a solenoid coil and a movable armature, is connected to a multiple contactor. Movement of the armature from the closed to the open position progressively short-circuits steps of resistance connected to the contacts. It is necessary to choose the resistance steps with the field-excitation characteristic of the alternator in mind, so that each progressive step of resistance produces an approximately equal change of output voltage. The incremental change in output voltage produced by one step of resistance determines the limit of sensitivity of the regulator. An air dashpot is usually provided to damp any oscillation or hunting of the armature.

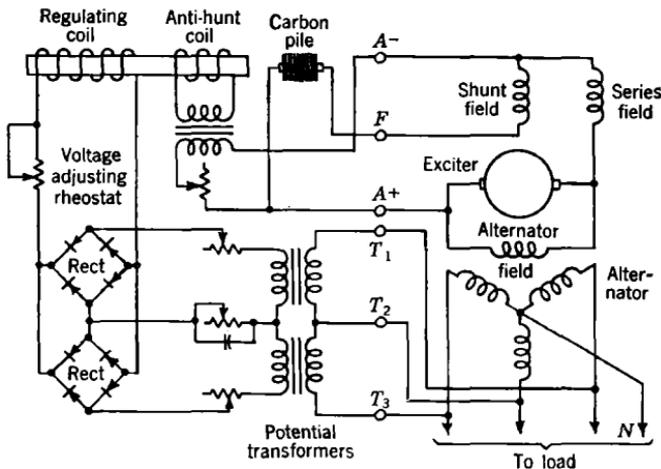


FIG. 14-15.—Schematic diagram of a General Electric voltage regulator Type GEA-2-B2 and 3-phase 230-volt alternator with exciter.

In field service, the usefulness of the finger-type regulators was limited by the tendency of the contacts to stick and burn, and by corrosion of the contacts due to moisture and salt air. In addition, the device is very difficult to repair in the field. Consequently, finger-type regulators have been used on radar systems only rarely.

14-6. Speed Regulators.—The a-c frequency generated by motor-alternator sets varies over a wide range with changes in input voltage, load, altitude, and temperature. At high altitude, low temperatures and air densities are encountered. The low temperature tends to lower the speed by reducing the resistance of the motor field circuit, whereas low air density allows the speed to increase because of lower windage losses. The following test data indicate the conditions for maximum and minimum frequency of the 1500-v_a PE-218-C and PE-218-D inverters.

Maximum frequency at 28.5 volts, $\frac{1}{4}$ load, 35,000 ft., -15°C ,

Leland PE-218-C 485 cps

GE PE-218-D 458 cps

Minimum frequency at 28.5 volts, full load, sea level, -35°C ,

Leland PE-218-C 401 cps

GE PE-218-D 392 cps

The variation in frequency caused by change in input voltage and in temperature is indicated in the following table.

TABLE 14-4.—CHANGES IN FREQUENCY OF INVERTERS WITH TEMPERATURE AND VOLTAGE

Input voltage	Temperature, $^{\circ}\text{C}$	Frequency, cps	
		Leland PE-218-G	GE Co. PE-218-D
26	-45	373	379
	+25	396	396
	+55	406	398
29	-45	411	404
	+25	443	423
	+55	449	428
Over-all regulation		Leland PE-218-C	GE Co. PE-218-D
In cps		76	49
In per cent		17.0	11.5

All data were taken at sea level and with full rated load on inverter. The GE PE-218-D unit has a compensating resistor in series with the field which accounts for its slightly better regulation. It also has an adjustable resistor which allows the frequency to be varied over a range of approximately 40 cps. These inverters are shown in Figs. 14-16 and 14-17.

In most radar applications, such a range in frequency is not objectionable. Certain indicator and computing circuits, however, require much closer regulation of the order of $\pm 2\frac{1}{2}$ per cent. Good speed regulation also considerably reduces the problem of voltage regulation. In accordance with the need for better speed characteristics of inverters, the latest Army-Navy specifications (AN-I-10) stipulate the following frequency limits: For all voltages between 26 and 29 volts and at any load from $\frac{1}{8}$ to full, the allowed frequency limits for a temperature range of -10°C

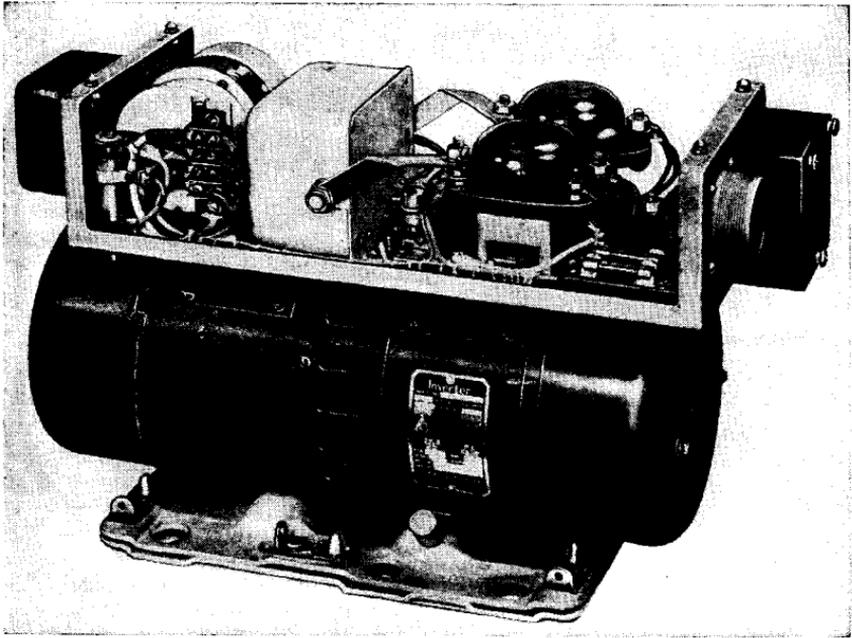


FIG. 14-16.—Leland Electric Company aircraft motor-alternator with regulator.

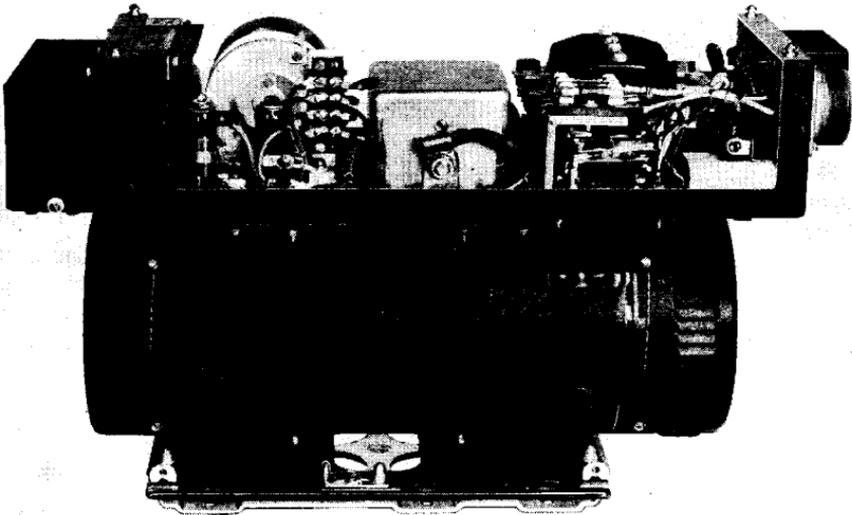


FIG. 14-17.—General Electric Company aircraft motor-alternator with regulator.

to +10°C shall be 390 to 410 cps; over the same range of voltage and load, the frequency limits for a temperature range of -55°C to +70°C shall be 380 to 420 cps.

Table 14-5 gives the frequency range of inverters now in general use, and in the column headed "Remarks" the motor type or method of speed control is indicated. All data refer to a temperature of 25°C and sea-level pressure.

TABLE 14-5.—SPEED CONTROL OF CERTAIN MOTOR-ALTERNATOR SETS

Manufacturer	Rating va phases	Style	Frequency, cps		Remarks
			26 volts full load	29 volts ½ load	
Eclipse.....	100 3φ	12123-1-A	410	410	Lee Regulator
Eclipse.....	250 3φ	12121-1-A	400	410	Lee Regulator
Leland.....	500 1φ	10596	395	475	Compound-wound
Wincharger.....	750 1φ	PU-16	390	403	Carbon-pile regu- lator
Holtzer-Cabot.....	1000 3φ	MG-153	340	395	Resonant
Leland.....	1500 1φ	PE-218-C	395	450	Shunt-wound
GE.....	1500 1φ	PE-218-D	395	440	Shunt-wound
Airways.....	2500 1φ	PU-7	410	430	Regulating field
Eclipse*.....	840 1φ	800-1-C	730	900	Compound-wound

* Nominal rated frequency 800 cps.

Lee Regulator.—The Lee speed regulator is a centrifugal device which alternately opens and closes a short circuit across a fixed resistor in series with the motor field. The short-circuiting contacts are on the centrifugal device and can be set to open and close at any predetermined speeds within 5 per cent of each other. Speed regulation is therefore good; it can be maintained at approximately 5 per cent. This type of regulator creates a good deal of r-f noise that is difficult to suppress, and gives pronounced modulation of the output voltage.

Voltage Regulator Control.—Speed regulation can also be provided by means of a regulating field controlled by the generator voltage regulator. This method is used in the 2500-v Type PU-7 inverter designed by Wincharger and built by Wincharger and Airways. It is shown schematically in Fig 14-18.

As the input voltage or inverter load changes, the resistance of the carbon pile in the voltage regulator varies to maintain constant output voltage. This also serves to vary the current through the motor regulator field in the proper sense to maintain constant speed. The motor is designed to obtain proper division of flux between the main field and the regulating field. The resistances of the regulating field and of the

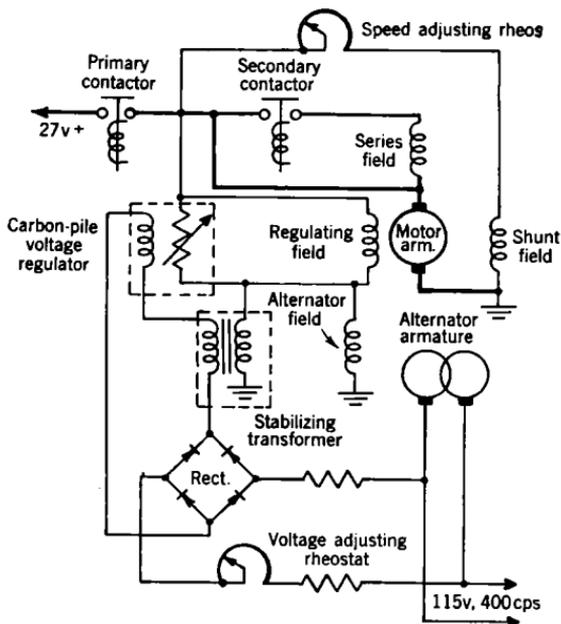


FIG. 14-18.—Schematic wiring diagram of PU-7 aircraft motor-alternator, 2500 va, 115 volts, 400 cps.

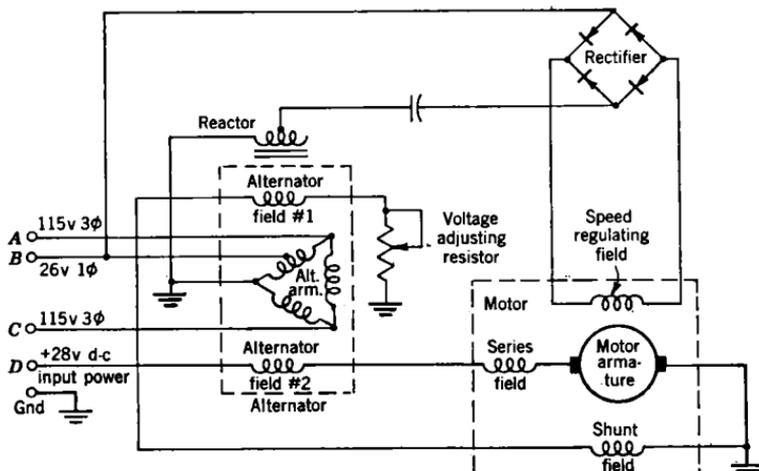


FIG. 14-19.—Schematic circuit of Holtzer-Cabot MG-153 aircraft motor-alternator. Output ratings: 750 va, 115 volts, 400 cps, 3-phase; 250 va, 26 volts, 400 cps, 1-phase. Ground terminal common to all circuits.

alternator field must be kept in proper relation to the normal range of resistance of the carbon pile in order to obtain good speed regulation.

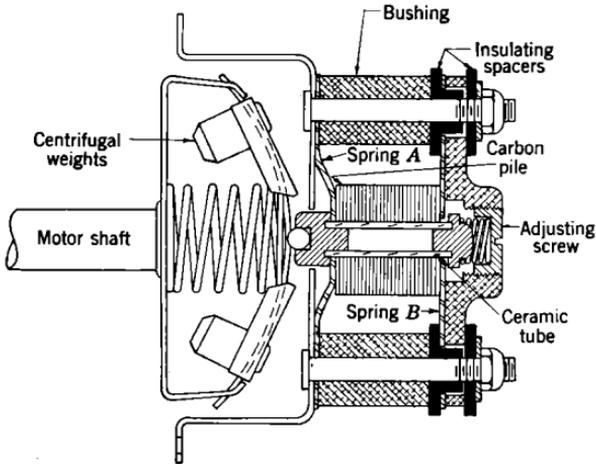


FIG. 14-20.—Holtzer-Cabot Electric Company carbon-pile speed governor.

If the regulator is removed from the circuit, the alternator will still generate a voltage because the field current will flow through the motor-regulating field to the generator field. In the case of inverters not

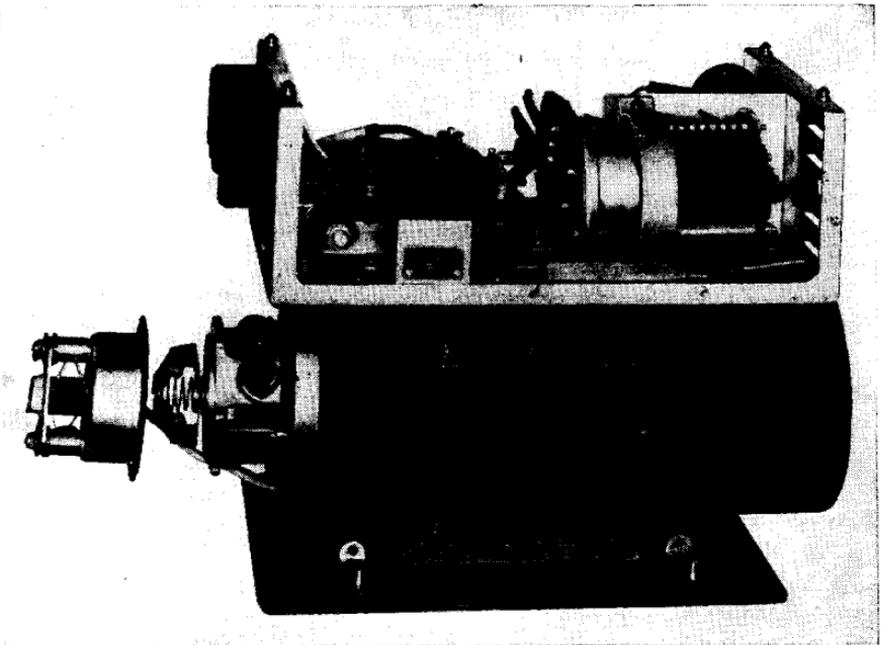
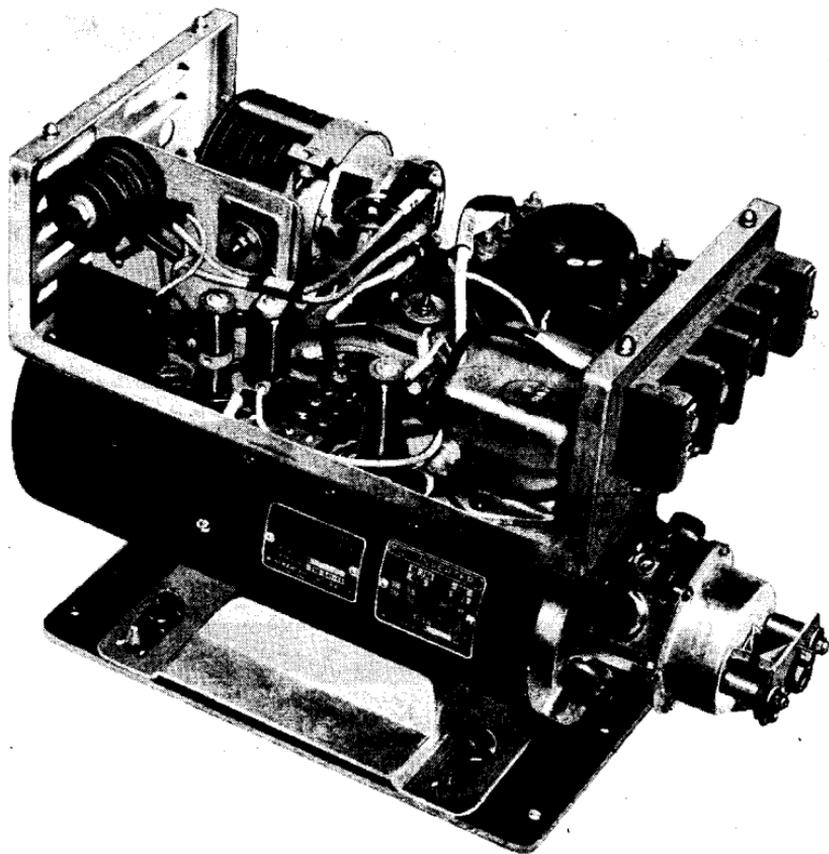


FIG. 14-21.—Wincharger Corporation aircraft

having the regulating field, removal of the regulator from the circuit kills the alternator voltage.

In order to get good results by this method, the motor field must operate below saturation; this adds weight to the magnetic circuit.

Tuned Circuit Control.—An interesting method of speed control has been used by the Holtzer-Cabot Electric Company in their MG-153 inverter, Fig. 14-19. The reactor and condenser connected in series with the rectifier which energizes the regulating field are designed to resonate at a frequency well above the rated frequency of the alternator (in this case, 400 cps). The circuit therefore operates on the low-frequency side of the resonance curve; any change in speed will alter the reactance of the circuit and thus vary the field current in the correct direction to provide speed regulation. Should the motor attain a speed



motor-alternator with Holtzer-Cabot speed control.

that would generate a frequency above resonance, the machine speed would continue to increase, being limited only by friction and windage losses.

Centrifugal Control.—The most recent method of speed regulation to be used on aircraft inverters is a carbon-pile governor as used by the Holtzer-Cabot Electric Company. It is shown in Fig. 14-20 and also in Fig. 14-21 mounted on a 750-v-a PU-16 inverter. The electrical connections are shown in the schematic diagram Fig. 14-22. A centrifugal device on the end of the inverter shaft acts against the carbon pile in such a manner that increased speed will reduce the pressure on the carbon pile and reduced speed will increase the pressure. As shown in Fig. 14-20, when the machine is at rest the carbon pile is held in compres-

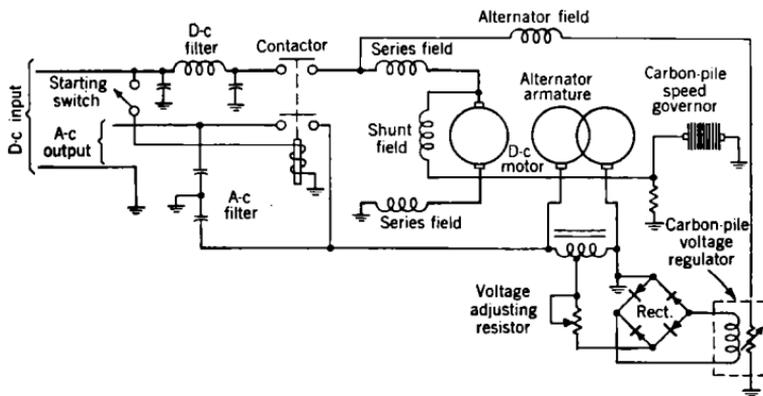


FIG. 14-22.—Schematic diagram of Wincharger PU-16 aircraft inverter, 750 va, 115 volts, 400 cps.

sion by the large coil spring in the rotating element. The resistance is then least, giving maximum field for good starting characteristics. As the machine speed increases, the centrifugal weights compress the spring and reduce the pressure on the pile, increasing its resistance. Should the machine exceed its rated speed, the carbon pile will be compressed by the spring under the adjusting screw. At rated speed, the carbon pile is balanced between these two springs and a slight change in speed will cause a relatively large change in resistance. Springs *A* and *B* position the tube holding the carbon pile.¹

Electronic Speed Controls.—The Eclipse-Pioneer Division of Bendix Aviation has recently developed an electronic speed control for use on motor alternator sets of 1500-v-a capacity and larger. This regulator is designed to hold speed within the limits of Army-Navy Specification AN-I-10.

¹C. T. Button, "A Carbon-pile Speed Governor," *AIEE Transactions*, **65**, 4 (January 1946).

14-7. Dynamotors.—To supply smaller radar sets with moderate power at not over 1000 or 1200 volts, a dynamotor is often more economical in weight and efficiency than an inverter. A dynamotor has conventional d-c shunt- and series-field windings on the stator and one or more armature windings on the rotor. The rotor winding is excited through brushes and a commutator from the input voltage.

A single rotor winding can be tapped to give the required secondary voltage. Alternatively, separate rotor windings can be used, with turns-ratios to the primary winding chosen to provide the correct output voltages. The required d-c voltages are brought out through commutators from the secondary windings. Dynamotors have been built

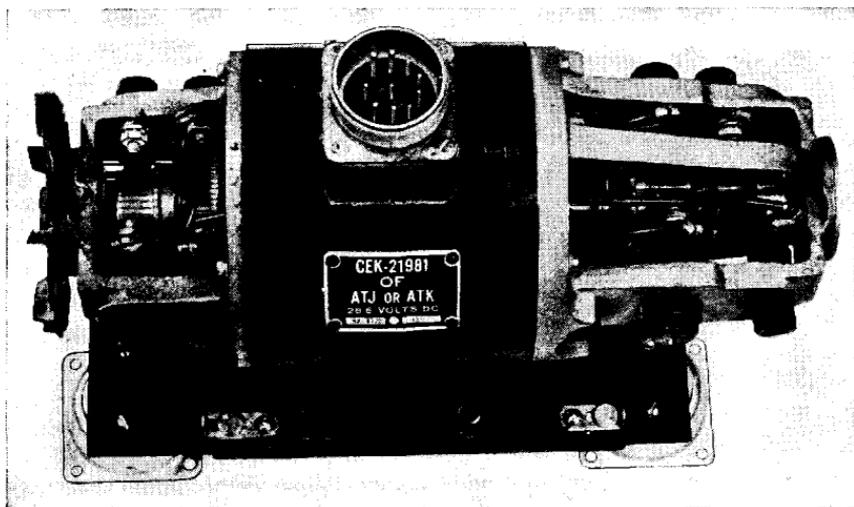


FIG. 14-23.—Eicor Inc. aircraft dynamotor.

to give voltages as high as 2000 volts, but such machines meet severe difficulties with commutation and insulation. Dual- and triple-output dynamotors are very common, and several machines having four outputs (Fig. 14-23) have been built, but it is impractical to go beyond this because of the large number of windings and commutators involved.

Dynamotors for military aircraft are built for an input voltage of 27.5 volts. For automobiles, commercial aircraft, and miscellaneous use, dynamotors are built for input voltages of 6, 12, 32, and 110 volts. Output ratings vary from 14 or less to over 1200 volts, and from 30 ma to 5 amp, with any combination of two, three, or four at a time. Weights vary from 5 lb, for a total output of 20 watts, to 26 lb, for an output of 400 watts. Efficiencies range from 40 per cent to 60 per cent.

The dynamotor suffers from poor regulation. Furthermore, variations in input voltage may be as great as 10 per cent. The dynamotor

can be designed for an input of 19 or 20 volts if it is to be used with an actual line voltage of 27.5, and a variable control resistor put in series with the line. This resistor can be a carbon-pile voltage regulator, with the carbon pile in series with the line and the operating coil across the line

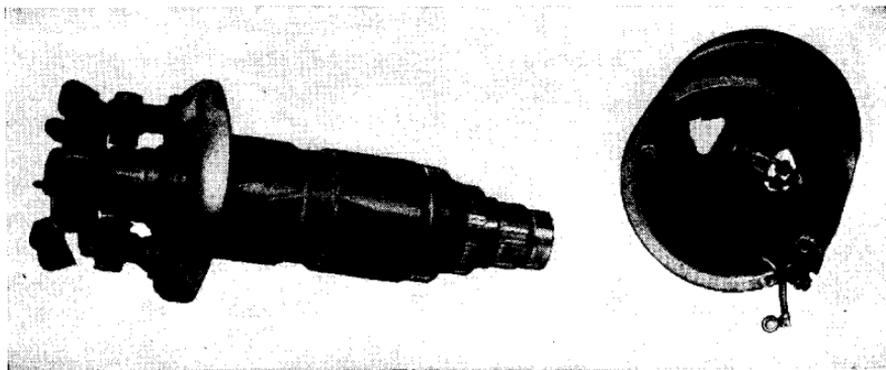


FIG. 14-24.—Bendix dynamotor.

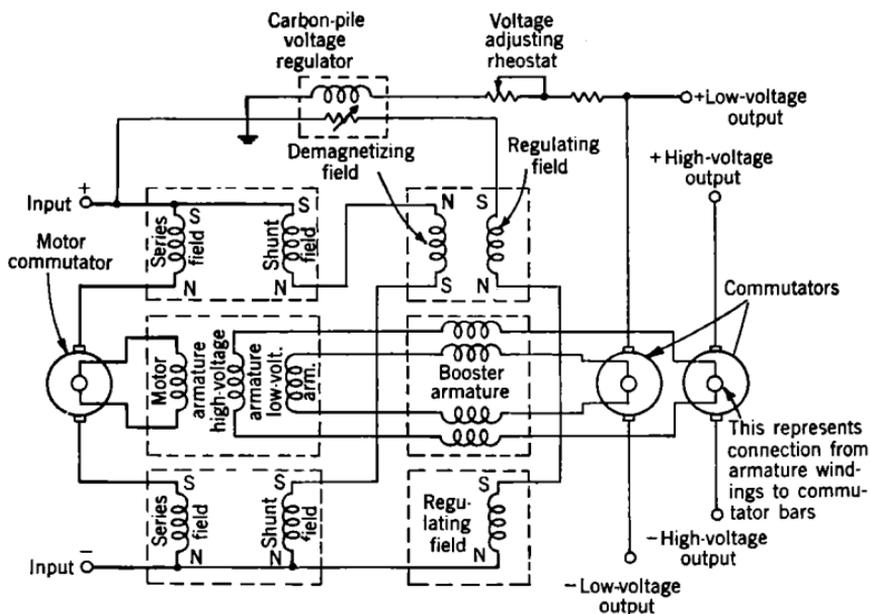


FIG. 14-25.—Schematic diagram of dual-output dynamotor with booster windings and regulating field.

just ahead of the dynamotor. For an input from 25 to 29 volts at constant load, the regulation of the output voltage can be held within a range of 5 to 10 per cent. Regulation becomes poorer with changing load.

Booster Armature Voltage Regulation.—A recent development to improve dynamotor regulation is the construction in the dynamotor

itself of a booster armature winding and regulating field. A three-voltage dynamotor is shown in the photograph of Fig. 14-24 and the circuit of a dual-voltage machine is shown schematically in Fig. 14-25. In any multi-output dynamotor only one voltage can be regulated, but, except in the event of very unevenly loaded circuits, the unregulated voltages will closely follow the regulated voltage. Regulation can be handled in several ways: (1) all output circuits have windings on the booster armature although only one of them is controlled (Fig. 14-25); (2) only the voltage to be regulated has a booster winding; and (3) the voltage to be regulated is generated entirely in the windings of the booster armature. The third method is practical only for low voltages, because of the large number of turns and relatively high field current required to generate a high voltage.¹

Over ranges of input voltage from 25 to 29 volts, from no load to full load on the dynamotor, the output-voltage change will be approximately 3 per cent. The booster armature and regulating field increase the size and weight of the dynamotor, but the regulator itself can be considerably smaller than a series regulator because it carries only the field current rather than the line current.

14-8. Vibrator Power Supplies.—If relatively small amounts of power are required, vibrator power supplies are useful. Conventional vibrators, such as those used commercially in automobile radios, are capable of supplying 50 to 60 watts of square-wave alternating current to a transformer. The transformer output may be rectified and filtered to furnish direct current of appropriate voltage.

Vibrator power supplies can, unlike dynamotors, supply a number of different d-c output voltages without disproportionate increase in the weight of the power supply. For example, it is possible to supply 2500 volts direct current at 5 ma, 250 volts direct current at 100 ma, and — 150 volts direct current at 5 ma, at approximately the same total weight as that required for conventional power supplies designed to run from 115-volt 60-cps a-c lines. The vibrator power supply just described weighs from 6 to 8 lb, depending on how it is packaged and on the degree of filtering required.

The output voltage of a vibrator power supply can be regulated by use of the series-tube type of electronic regulator, which is apt to be wasteful of power. Further, output power is consumed by the rectifier-tube cathodes. Where a number of d-c output voltages are required,

¹ In any case, it is seldom practical to regulate the highest-voltage output since current-limiting resistors are required in the regulator coil circuit, and at high voltages the loss in them is excessive. It is also less important to regulate the higher voltages, for this can readily be done by regulating circuits in the radar set itself.

rectifier cathodes may use up a large part of the 50 to 60 watts of available power.

Two or more vibrator power supplies can be designed for operation in series or in parallel on the d-c output side, so that larger currents or voltages can be obtained from small vibrators. Large vibrators, capable of supplying 200 watts or more from 12- to 14-volt or 24- to 28-volt d-c supplies, are available, but it is usually more convenient to use conventional inverters with transformer-rectifier power supplies for outputs greater than 250 watts.

Vibrator power supplies usually give efficiencies of 50 to 70 per cent, based on d-c output vs. d-c input.

Precautions must usually be taken to reduce radio interference from vibrator power supplies. Small r-f chokes in the power input leads, mica r-f bypass condensers on input and output, and careful shielding and grounding usually provide adequate radio interference suppression for aircraft use. Where extremely low noise level is required it may be necessary to shield the vibrator, transformer, and each d-c ripple filter separately, and use r-f chokes in each transformer connection. However, it is usually less troublesome to suppress radio noise in a vibrator supply than in a small dynamotor.

Vibrators can be obtained in hermetically sealed containers, making it possible to design power supplies for operation under extreme conditions of altitude and humidity.

Production-type small vibrators are available with reed frequencies from about 100 to 115 cps. These vibrators have a useful life of at least 500 hours when used with well-designed power supplies.

Vibrators have recently been developed with reed frequencies of about 180 cps. Smaller transformers can be used at this higher frequency, so that an over-all weight reduction of 20 to 25 per cent may be achieved. The life of such vibrators may be somewhat less than that of low-frequency vibrators. A more complete treatment of vibrator power supplies is given in Vol. 17, Chap. 12 of the series.

14-9. Summary of Recommendations for Aircraft Radar Power.

1. Consider the radar-radio power-supply problem of a specific aircraft as a whole. What dynamotor can be eliminated? What inverters can be combined? Is all equipment uniformly designed to operate from one standard a-c source? Often enough weight can be saved to justify such refinements as an electronic voltage regulator.
2. Use a direct-driven generator whenever possible, either at constant frequency (if airplane is so equipped) or at variable frequency. If

portions of the installation require fixed frequency, consider an auxiliary d-c to a-c motor-alternator.

3. Use electronic voltage regulation for engine generators or motor alternators.
4. Consider vibrators for small loads.

GROUND AND SHIPBOARD SYSTEMS

Power for ground radar systems should be obtained from commercial electric service whenever possible, with certain precautions outlined below. The general problem of power supply for ground radar falls into four main groupings:

1. Fixed locations with commercial electric service available.
2. Large systems (with requirements for highly dependable service under all conditions) at fixed locations without commercial electric service, or for mobile use.
3. Small systems for use in fixed locations without commercial power, or for mobile use.
4. Small ultraportable mobile sets.

In selecting a specific engine-generator set, or any specific alternator, serious consideration must be given to the specifications of the set (the alternator in particular) if satisfactory performance and characteristics are to be obtained. One good tabulation of standard Army power supplies is Signal Corps Manual TM 11-223. However, most of the sets listed therein were not originally designed with radar in mind. Accordingly, their electrical performance may not meet radar needs without modification. A fuller discussion of the factors involved in choosing alternators and engine-generator sets is given in Vol. 17, Chap. 12, Radiation Laboratory Series.

14-10. Fixed Locations.—Commercial power supply for radar installations must be provided with three major considerations in mind: (1) reliability required, (2) voltage stability of supply, (3) interference from common loads.

Reliability.—If extremely reliable service must be provided, commercial service from a first-grade public utility is most satisfactory, provided that it is backed up by emergency feeders duplicating normal service (preferably by an alternate route) to a network station of the public-utility system. An emergency engine-alternator set (which may be made automatic-starting if deemed necessary) should also be provided. This arrangement is preferable to constant reliance on local generation. A regular maintenance schedule must be followed, with the emergency engine operated for a minimum of one hour per week at no load and a

minimum of two hours per month at full load. For such emergency service as that outlined, gasoline-powered prime movers are reliable and satisfactory.

Voltage Stability of Supply.—If the service is obtained from a source unstable in voltage, local voltage stabilization must be employed. If the voltage changes take place slowly, without transients, they can be eliminated by means of induction regulators, with automatic or manual control; tapped autotransformers, manual or servo-driven; or (for small loads) variacs. For example, consider the use of commercial power by a remote beacon station served by a line shared with other power users. As night falls, increasing load on the line lowers the voltage at the beacon. Servo-driven variacs have proved useful in such a situation.

Interference from Common Loads.—If the loads sharing power service with the radar impose high intermittent demands, and especially if the total capacity of the system is low with respect to these loads, serious transient interference may be encountered. An isolating motor-generator is the only satisfactory solution for such difficulties. If synchronous frequency is not essential, induction-motor drive is preferable to the greater complexity of a synchronous motor. It is good practice, where possible, to make input and output services identical, and to provide means for bypassing the isolating set in case of trouble. That is, if the main service is 220 volts, 3-phase, 60 cycles, the motor-generator set should be 220-volt 3-phase input (60 cycles) to 220-volt 3-phase output (58 cycles).

14-11. Large Systems Where No Commercial Power Is Available.—

Where 24-hour operation of a prime mover is required, it is necessary to provide duplicate power generation equipment, regular and emergency. Such equipment should be interchangeable, and the two sets should be used alternately to ensure reliable service. Where extreme reliability is desired, three engines should be furnished; one can then be disassembled for maintenance without risking complete system breakdown if the service engine fails. Generators of a size at least twice the estimated load should always be provided, to permit future additions of auxiliary devices not initially specified. Diesel engines are definitely more reliable for continuous operation than gasoline engines. Diesel fuel and diesel exhaust fumes are less hazardous.

Electric power for loads larger than 7.5 kva should be 208/120-volt 3-phase 60-cycle alternating current. It is desirable to use 3-phase 60-cycle because electric motors of this rating are compact and simple to operate. Weight savings in rectifier-filter supplies can usually be achieved with 3-phase power, if this is important.

For large mobile systems it is well to investigate availability of standard U.S. Army power-supply units. Unfortunately, the voltage

regulators in such systems are almost uniformly unsatisfactory for radar equipment since they have a tendency to produce modulation in the amplitude of the a-c voltage. It is often necessary to resort to hand rheostat control of the exciter field or the alternator field. Radar loads are usually constant or nearly so, and hand control requires only occasional adjustment after the equipment has come to normal operating temperature. The engines are usually provided with flyball-type speed governors which are satisfactory when well maintained, and keep the output frequency within a range of about 58 to 62 cps.

14-12. Smaller Mobile Units.—If high reliability is required, the same considerations must be observed in this case as in the case of large systems, notwithstanding the weight penalty. For example, a navigation aid required to operate 24 hours per day under all conditions must use conservative, long-lived prime movers and be provided with standby equipment, though it may consume only 750 watts.

Much design effort has been expended on packaging units of about 500-lb weight and 5-kw output in a readily portable form. There has also been much development of two-cycle gasoline engines for power units up to about 2-kw output. It is likely that two-cycle engines coupled to 400-cps generators will be available in the future in many ratings; these units are much lighter and more compact than four-cycle engines with 60-cps generators of equivalent rating.

It is true that two-cycle engines have some disadvantages. They are often more difficult to start than the small four-cycle engines, and, because of their tendency to foul exhaust ports and spark plugs, they are likely to require more maintenance. However, as mobile units, they have two advantages over and above their lightness of weight and their compactness. They are easier to repair in the field and require less replacement of parts than the four-cycle engines, which are subject to valve-fouling and burning.

14-13. Ultraportable Units.—Where extremely light weight is required this can be obtained only by sacrificing reliability or by shortening the required period of operation.

Ultraportable ground radar systems usually need very little power. Where only a few hours of operation are required, the systems can be powered by small storage batteries which operate vibrator power supplies or dynamotors.

Where longer operation is required, very small two-cycle engines connected to permanent-magnet-field 400-cps generators have been developed by the Jacobsen Manufacturing Company, Racine, Wis., and the Judson Manufacturing Company, Philadelphia, Pa.

Jacobsen produces a unit rated 125 watts 400 cps weighing about 13 lb. A single-cylinder, air-cooled, two-cycle engine is direct-connected

to a permanent-magnet generator. The speed is controlled by an air-vane governor.

Judson produces several small models, one of which is rated 150 watts 400 cps and weighs about 15 lb. The general design is similar to that of the Jacobsen machine, except that the engine is belt-connected to the generator.

There is no way of regulating the output voltage of permanent-magnet generators, but capacity in series with the output will hold the output voltage constant to about ± 5 per cent despite changing load. The voltage can be altered to compensate for temperature changes by manual adjustment of the speed governor on the engine.

Work has been initiated on the development of small prime movers other than conventional gasoline engines. Small steam engines that were capable of about 300 mechanical watts output were built by Radiation Laboratory, but no satisfactory boiler had been developed at the close of activity. Small steam engines and gas turbines are attractive because of their ease of control and flexibility. Further development would probably produce some worth-while small prime movers.

14-14. Ship Radar Systems.—The problem of providing power for large shipborne radar systems is principally that of stability. Most modern warships have 440-volt, 3-phase, 60-cps power supply which is subject to transient fluctuations of 10 to 20 per cent in voltage and perhaps 5 per cent in frequency. Transients are caused by the sudden application of large loads such as gun turret drives and airplane elevators.

Good results have been obtained by powering the radar system from a 60-cycle motor-generator set. An induction motor is used to drive the alternator; when not heavily loaded, the motor shows a relatively small change in speed for a change in input voltage. Motor speed tends to follow a change in frequency, but the mechanical inertia of the set is usually adequate to hold up the speed during the usual short-period frequency transients. The induction motor rating should be about 50 per cent larger than that required to drive the alternator. It is desirable to have the alternator output the same as the ship's power so that the radar system can be operated directly from the ship's supply in case of breakdown or maintenance shutdown of the motor-alternator set.

The motor may be direct-coupled or belt-connected to the alternator; usually direct coupling is preferred because of its simplicity. The output frequency of a direct-connected alternator will be less than 60 cps by the slip of the induction motor. This results in an output frequency of 57 to 58 cps, which is normally satisfactory for the operation of transformers or other 60-cycle equipment.

There is no really satisfactory means now available for regulating the output voltage of the alternator. Most of the available voltage regu-

lators have a tendency to be unstable, exhibiting rapid hunting of about 20 to 40 cps. This may not be troublesome on lighting or motor loads, but it is one of the major causes of spoking of PPI displays and instability of other forms of radar presentation.

One method of using regulators has been followed with some degree of success. That is to provide a rheostat in series with the regulating resistance of the regulator in the field circuit. Assuming that 115-volt output is desired, the regulator can be adjusted to regulate at 118 to 120 volts and enough resistance cut in, by means of the rheostat, so that the voltage is 115 volts at full radar load with everything at normal operating temperature. Under normal operating conditions, then, the regulator is all out and not functioning. If the voltage should rise, as during warmup period or when the radar is on stand-by, the regulator will cut in when the voltage comes up to 118 to 120 volts and prevent further rise in voltage. If there is a load greater than normal causing the output voltage to drop, the manual rheostat will have to be adjusted. However, with the motor-generator set acting as an electromechanical flywheel to take up transient fluctuations in the ship's power supply, operation is reasonably steady and satisfactory.

CHAPTER 15

EXAMPLES OF RADAR SYSTEM DESIGN

BY R. G. HERB AND R. L. SINSHEIMER

15.1. Introduction.—The development of a new radar system may be called for because of recognition of a new application for which radar has not been used but for which its successful use appears possible. In other cases, a development may be justified for an application where radar is in successful use but where improved results could be expected by the use of more modern equipment.

Early ideas on a new system and initial proposals may come from a development laboratory or from the potential users of the equipment. Before design characteristics can be successfully crystallized, regardless of the origin of a proposal, there must be an extensive interplay of ideas and of information between specialists in the following categories: (1) application specialists representing the potential using organization, who must contribute necessary information on desirable performance characteristics, limitations on size, weight, and power consumption, and limitations on the number and skill of operators and maintenance men; (2) component specialists, who must contribute information on the status of development and on the limitations of the many component parts that make up a complete system; (3) system specialists with experience in the design and operation of radar systems.

Normally a man experienced in system design is best able to coordinate the over-all development project. He must have a supporting group of experienced systems men to assist on the design problem and to carry out operational test work that may be desirable, and he must have the continued advice and support of specialists of the first two categories listed above.

Versatility.—Initial planning for a radar system may be made with one specific application in mind. As plans progress, possibilities of related applications are generally recognized and consideration is given to including facilities that improve the system in versatility. The value of versatility has been amply demonstrated, not only for meeting related applications but for meeting changing conditions frequently encountered in the central application.

Provision for versatility, unless ingeniously made, may cost heavily in increased complexity, size, or weight; further, the utility of the set for

the principal application may suffer. How far to go with such provisions and how best to include them are problems that frequently outweigh in importance any other design problem. No rules can be given for guidance along these lines, but the degree of success to be expected will depend markedly on the extent to which development specialists are familiar with problems of operation in the using organization and on the degree to which representatives of the using organizations have become familiar with the technical problems, the possibilities, and the limitations of radar.

Requirements for Component Development.—With sufficient care in design and a sufficiently conservative policy, a radar set might be designed making use of only those components on which development work has been completed. Ordinarily the improvement to be gained in system performance or in general utility by improvement in certain components is sufficient to warrant component development work. The extent to which such requirements can safely be included must be determined from consideration of the time schedule to be met and from a careful, realistic appraisal of the development time needed. Sound system-design work, together with performance tests, can serve in a valuable way to direct component-development work along the most advantageous lines.

Detailed Problems in Design.—General rules for guidance on the many detailed problems encountered in the development of a radar system would be extremely difficult to formulate because of the multiplicity of applications with radically differing requirements. Information accumulated through experience on a large number of systems can be presented most easily by outlining the development of successful systems from the early-idea stage on through to their operational use.

Two typical system developments are described in this chapter: (1) a high-performance ground-based set for air surveillance and control; and (2) the AN/APS-10, a lightweight airborne set for air navigation. Considerations leading to the selection of these systems are as follows:

1. Substantial development and design effort went into both systems and important advances resulted.
2. Both systems are relatively simple, and design problems encountered in their development are more basic in nature and more generally applicable than in the case of a more elaborate system, where usually a large proportion of the system is very specialized in nature.
3. These two systems represent opposite extremes in requirements. Limitations on factors such as size, weight, and power requirements for the ground-based set were largely subordinated to the requirement for the best possible radar performance. In the case of the airborne set, severe limits were imposed on weight, size, and similar

factors, and the ingenuity of the development and design men was taxed to meet the minimum acceptable radar performance within these limits.

A detailed treatment of further systems beyond these two would be of questionable value, since changing requirements may soon make obsolete all but the more general, basic developments of the past. Radically new requirements will be met only by radically new developments, and to be firmly bound by past experience would stifle progress. Ingenious technical men with good facilities, free in their thinking, working in close touch with using organizations, fully informed on operational problems and planning—these are the important requirements for future advances.

15-2. The Need for System Testing.¹—One general remark concerning system design is sufficiently important to be made before taking up the design examples. No radar system is complete until adequate provisions for performance testing and maintenance have been made and appropriate maintenance procedures prescribed.

The most important function of test equipment in radar maintenance is to permit quantitative measurement of those properties of the system which affect its range performance. Experience has shown that it is essential to determine by reliable quantitative methods how well a radar is operating. Under most circumstances, the important units of test equipment should be built into the radar in a way that permits their most convenient use. If limitations of weight and bulk do not permit the use of built-in test equipment, suitable test points should be provided to permit convenient and adequate tests to be performed frequently on the important parts of the system.

Radar Performance Figure.—The ratio of P , the pulse power of the radar transmitter, to S_{\min} , the power of the minimum detectable signal, is a measure of the radar performance. This is the fundamental quantity studied in tests. The radar equation given as Eq. (2-4b),

$$R_{\max} = \sqrt[4]{\frac{P\sigma A^2 f^2}{S_{\min} 4\pi\lambda^2}}, \quad (1)$$

shows that all factors except P and S_{\min} are either invariable or beyond our control once the radar has been designed.

A quantity closely related to S_{\min} is usually measured rather than S_{\min} itself. This may be the power of the weakest test signal that can be detected, or of a test signal that produces some other easily reproducible effect. Because of the effects of presentation time and pulse width

¹ By R. D. O'Neal and J. M. Wolf.

(Secs. 2-10 and 2-11), such test signals are generally not equal to S_{\min} but differ from it by a constant ratio. The ratio between pulse power and the power of such a test signal, P/S_T , expressed in decibels, is called the "radar performance figure" and is a suitable measure of the ability of the system to see radar targets.

Figure 15-1 shows the range performance on various types of radar targets as a function of the radar performance figure. As explained in Chap. 2, radar range performance does not always follow the fourth-power law of Eq. (1), but the performance figure remains of vital importance in determining what fraction of the maximum radar range can be realized by a given system against a given type of target, regardless of the existing propagation conditions.

The Inadequacy of Guessing Performance.—It has often been wrongly assumed that over-all radar performance can be adequately judged without using test equipment by means of one of the following "rule-of-thumb" criteria: (1) the general appearance of the picture seen on the radar indicator, (2) the maximum range at which target signals can be seen, or (3) the signal strength above noise of the echo from a "standard" target.

Both (1) and (2) are strongly affected by changes in propagation conditions¹ as well as by interference effects resulting from composite targets,² tidal changes,³ etc. Although the effect of anomalous propagation can be practically eliminated by choosing a near-by weak target signal as a standard, such a favorable choice of "standard target" does not necessarily eliminate interference effects. Targets that do not show such effects in some degree are exceedingly rare.

These rule-of-thumb criteria are inadequate for the consistent judgment of radar performance, as has been shown by surveys of radar performance made on radar equipment in the field during the war. Unfortunately during most of the war adequate test equipment was commonly lacking; when it was present, it was often in charge of inadequately trained maintenance personnel who did not appreciate the need for using it regularly. Figure 15-2 shows the combined results of a

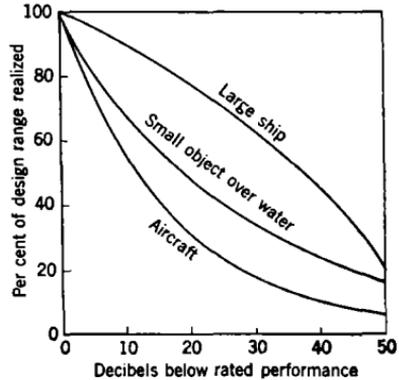


FIG. 15-1.—Relation between radar performance deficit and available radar range for various types of target.

¹ See Sec. 2-14.

² See Secs. 3-8 and 3-9.

³ Sec. 2-12.

number of such surveys of ground, ship, and airborne 10-cm and 3-cm radars during the years 1943 and 1944. The radar performance figure of the average set tested was 14 or 15 db below the rated value for the radar. Since a performance-figure deficit of 12 db results in cutting the maximum range of the set on aircraft by a factor of 2 (Fig. 15-1), poor radar performance was responsible for a loss of more than half the tactical usefulness of the systems tested. Figure 15-3 shows the findings of a more recent (July 1945) survey covering 10-cm ship radars; it indicates that there had been no change for the better.

Such serious deficiencies in actual field radar performance emphasize the fact that the use of test equipment to measure performance and to trace down the causes of impaired performance had not been incorporated into routine maintenance practice for most radar sets at the end of the

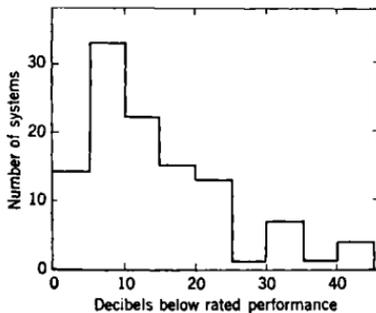


FIG. 15-2.—Radar performance surveys, 1943-1944.

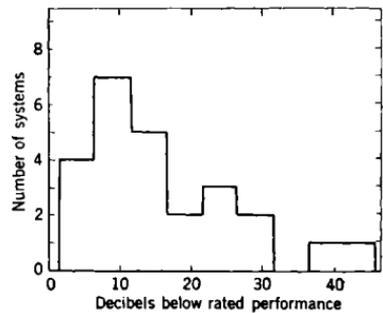


FIG. 15-3.—Radar performance survey, July 1945.

war. Such important losses in performance are glaringly inconsistent with the great effort expended in radar design to attain high performance. In the design of new equipment, it is essential to make careful provision for adequate testing and trouble-shooting; in the use of existing equipment it is equally important to be sure that routine maintenance procedure is sufficient to keep the radar within a few decibels of rated performance. The design and use of test equipment is treated in detail in Vol. 23 of this series.

DESIGN OF A HIGH-PERFORMANCE RADAR FOR AIR SURVEILLANCE AND CONTROL

15-3. Initial Planning and Objectives. *Formulation of Requirements.*

At a relatively early stage in the development of microwave techniques consideration was given to the possibility of applying them to the problem of long-range air surveillance. Greatly improved azimuth resolution, minimization of ground "clutter," and improved low-angle coverage were the principal advantages that microwaves appeared to offer. The best

output pulse powers were then low, receiver sensitivities were poor, and system problems were not well understood. Coverage equal to or better than that provided by the longer-wave equipment then in use or under development appeared to be attainable only by means of extremely high antenna gain.

As these ideas developed, microwave techniques were rapidly advancing. Requirements for antenna gain and size were reduced to moderate values, and finally computations supported by measurement showed that a microwave system could satisfactorily meet the coverage problem. Work was then initiated leading to the production of a ground radar, which proved to be one of the most successful sets ever developed. Its design characteristics were so well chosen that, even with the best techniques available at the end of the war, it could not easily be surpassed for air surveillance and control.

As in several other system developments involving great advances over previous equipment, requirements on this set were not firmly specified during the developmental stage. In many respects this freedom from rigid requirements was advantageous. General objectives were clear, the design men were well situated to draw freely for advice and help on component specialists and on men familiar with operational needs, and they were largely free to use their own ingenuity and judgment to meet the many problems that were presented.

General Objectives.—The principal objectives can be stated simply as follows: Coverage was considered to be important both as to maximum detection range and as to the altitude region included, but resolution in azimuth and in range were thought to be of comparable importance. Where height information was desired, auxiliary height finders were to be used; thus the radar beam could be fixed in elevation and swept continuously in azimuth by rotation of the antenna.

Estimates of detailed requirements were difficult to make in regard to characteristics such as (1) accuracy of range and bearing values as determined by methods of data presentation, (2) interval between "looks" at the target as determined by the revolutions per minute of the antenna, and (3) traffic-handling capacity as determined by the number and type of indicators provided. The designers, however, adopted as a general policy the improvement or extension beyond current practice of equipment characteristics that determined the accuracy and speed of data presentation and use. Care was also taken in the design to permit still greater extension of these facilities by simple changes or additions.

Beam Shape in Elevation.—A large air-surveillance set must be located on the ground or be elevated at most by a low tower to avoid interference by local obstruction. The radar beam in its 360° sweep may look over land, over water, or over both land and water. Figure 15-4

shows the general form of beam shape that gives the most desirable type of coverage under these conditions. Its lower edge is determined by the optical horizon, with corrections for atmospheric refraction and interference due to surface reflection. At a distance d in land miles from the radar set, the horizon height h in feet is given by $h = \frac{1}{2}d^2$ for normal atmospheric refraction (neglecting the interference effect, which is small for microwaves; see Sec. 2-13). Reduction of the blind region below this lower edge by locating the set on high terrain may be advantageous where the coverage is largely over water, but for overland surveillance this recourse is severely limited because of permanent echoes, which can mask out all return in regions where the radar beam is intercepted by land surfaces.

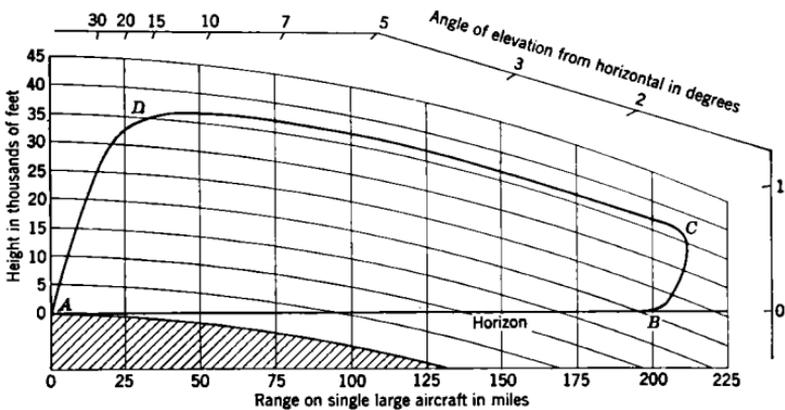


FIG. 15-4.—Desirable coverage for long-range air-surveillance radar.

In setting a desirable limit for the upper contour for detection of a given type of aircraft it was necessary to balance the value of high coverage against coverage in range, since adding to one subtracts from the other unless the designer chooses to increase the complexity of the equipment. For operations during World War II, provision for detection of a four-engine plane at all altitudes up to 35,000 ft was desirable, and this value was taken to determine the upper section CD of Fig. 15-4. Increasing the elevation angle of section AD of the contour does not add greatly to the radar coverage requirements, since the region is small; but practical difficulties of antenna design are encountered. Since close-in regions have been in the past largely masked out by permanent echoes, an angle of 30° was considered acceptable. With the upper and lower limits of the contour determined, the problem was now to push the maximum range out to the best possible value.

A detection range of 180 miles for a four-engine aircraft appeared to be the best that could be expected, and the contour thus determined is

not unduly difficult to fit by beam-shaping. Even greater coverage both in range and height was known to be desirable, but it appeared to be attainable only by excessive antenna size or by an excessive multiplicity of separate radar beams, each with its own transmitter and receiver system.

A microwave set with the coverage shown in Fig. 15-4 presented a very difficult problem. Because of uncertainty in measurements and in the outcome of component development, there was no firm assurance that it could be met. The success of the undertaking, however, was not dependent on meeting this coverage fully, since maximum range and altitude values within 20 per cent of the values of Fig. 15-4 would have represented a major advance over any previous equipment. Failure to detect one heavy plane above 30,000 ft would have limited the utility of the equipment for special problems, but, because of the practice of group flying, the normal problem was much less severe.

Azimuth Beamwidth.—The usefulness of the radar set for surveillance was expected to improve in traffic-handling capacity, in ability to resolve closely spaced aircraft, in accuracy of data, and in reduction of ground clutter, as the azimuth beamwidth was decreased. The desirable lower limit was expected to be set only from considerations of a practical upper limit to antenna size and requirements of mechanical accuracy. Results obtained fully confirmed these conclusions for beamwidths down to 1° at half power, but in a later section of this chapter we shall see that qualifications may be introduced if future development is extended to beams much narrower.

Miscellaneous Requirements.—Although limitations on size, complexity, weight, and power consumption could be subordinated to the requirement for excellence in radar performance, the designers were fully aware of the importance of practical limitations on these characteristics. It was also well understood that without great care and effort to achieve the utmost in simplicity, dependability, convenient facilities for test, and ease of maintenance, the equipment might fail to give satisfactory field service.

In computations leading to the choice of system parameters the designers did not have the advantage of much information that is now available, yet in practically every case their decisions were excellent. The discussion of design now given will make use of up-to-date information not available to the designers. This procedure is helpful for clarity of presentation and will make the analysis more generally applicable.

15-4. The Range Equation.—There will now be undertaken a detailed discussion of considerations leading to the choice of values for the following system parameters: (1) pulse length, (2) pulse recurrence frequency, (3) azimuth scanning rate, (4) azimuth beamwidth, (5) beam shape in

elevation angle, and (6) wavelength. We shall make use of the equation that gives the maximum detection range of a system as a function of the system parameters.

The range equation (see Sec. 2-4) can be written as

$$R_{\max} = \left(\frac{P\sigma A^2 f^2}{4\pi S_{\min} \lambda^2} \right)^{1/4}, \quad (1)$$

where

P = pulse power

σ = radar cross section of target

A = area of antenna aperture

f is a dimensionless constant dependent on the efficiency of the antenna aperture. It is equal to about 0.6 for a simple paraboloidal reflector, and its value is approximately the same for the parabolic cylinder used for the lower beam of the set described here. If the reflector is distorted to spread the beam, as was done for the upper beam of the set under consideration, A and f cannot be used in the normal way.

S_{\min} = minimum detectable signal power. It varies directly with the bandwidth B of the receiver, assuming that B is properly adjusted for the pulse length, and, to a good approximation for the problem under consideration, it varies inversely with the square root of the number of pulses per scan on the target.

Acceptable values for σ and S_{\min} are now reasonably well established, and, since other quantities determining R_{\max} can be readily measured, Eq. (1) can be used with some confidence for an absolute determination of range. The safer practice, however, is to compute coverage relative to the range performance on similar targets of a system with a similar scan, which has been carefully checked as to component performance. This procedure will be followed, and Eq. (1) will be used only for the relationships it establishes among the system parameters.

15-5. Choice of Pulse Length.—The pulse length is convenient for first consideration since its relations with the other parameters are relatively simple. It enters the range equation through its effect on the value of S_{\min} ; this quantity normally varies inversely with pulse length since the receiver bandwidth, if chosen for optimum performance, is an inverse function of pulse length.

Before proceeding further with the analysis, a discussion of other quantities that influence the value of S_{\min} will be useful. To a good approximation for the problem under consideration, the value of S_{\min} varies inversely with the square root of the number of pulses per scan on the target.

$$\text{Since } N_{\infty} = \frac{\Theta_{\nu_r}}{\omega},$$

where

N_{sc} = the number of pulses per scan on the target,

ν_r = pulse recurrence frequency,

Θ = azimuth beamwidth in degrees between half-power points, and

ω = scan rate in degrees per second,

we obtain the relation

$$S_{min} \propto \sqrt{\frac{\omega}{\Theta \nu_r}}. \quad (2)$$

From the preceding section we have

$$S_{min} \propto \frac{1}{\tau}, \quad (3)$$

where τ is the pulse length.

From Eqs. (1), (2), and (3) we obtain

$$R_{max} \propto P^{1/4} \tau^{1/4} \nu_r^{1/4},$$

holding other parameters constant, or

$$R_{max}^4 \propto P \tau \nu_r^{1/2}. \quad (4)$$

The average power output is given by $P \tau \nu_r$.

It is desirable to drive the transmitter of a high-performance ground-based system at the highest pulse power and the highest average power that it will safely withstand. The value of R_{max} can then be increased by an increase in pulse length and a corresponding reduction in the pulse recurrence rate, leaving average power and pulse power fixed at their highest safe values.

Limitations on pulse length are imposed by the following considerations:

1. The minimum range difference at which two targets can be resolved varies directly with pulse length. With a 1- μ sec pulse, aircraft inseparable in azimuth can be resolved if they differ in range by more than 164 yd plus a small distance which depends on the characteristics of the receiver and the indicator. With a 5- μ sec pulse, this limit for resolution becomes 820 yd plus the same additive constant.
2. The intensity of cloud return relative to signal return from planes varies directly with pulse length (Sec. 3-10).
3. The coverage area obscured by ground clutter will increase as pulse length is increased.
4. Magnetron behavior has been found to become more critical, and requirements on the voltage pulse from the modulator to become more exacting, with increasing pulse length.

5. Effects of pulse interference are more difficult to eliminate if the receiver bandwidth is narrow, as required for best sensitivity with a long pulse.
6. If a reduction in the pulse recurrence frequency is required to increase pulse length, the azimuth accuracy may be adversely affected and means for the selective detection of moving targets may be less effective (Chap. 16).

Consideration of all these points leads to the conclusion that a pulse length greater than about $2 \mu\text{sec}$ is to be avoided.

Upper limits imposed on pulse power and on the pulse recurrence frequency have been such that a decrease in pulse length below $1 \mu\text{sec}$ would require a decrease in average power below the maximum safe value, resulting in a decrease of R_{max} . Further, as we shall see, a $1\text{-}\mu\text{sec}$ pulse yields a discrimination in range which is, for most of the area covered by the radar, roughly equal to the azimuth discrimination afforded by the beamwidth chosen. This is desirable from the standpoint of the indicator.

A pulse length of $1 \mu\text{sec}$ was finally chosen for the following reasons:

1. Average power output was pushed to the safe upper limit for the transmitter with a value of 750 kw for pulse power, a pulse recurrence frequency of 400, and a $1\text{-}\mu\text{sec}$ pulse length.
2. Reduction of the pulse recurrence frequency to 200 with a corresponding increase of pulse length to $2 \mu\text{sec}$ would have given an appreciable increase in R_{max} if this change led to no reduction in pulse power or in average power. Since the magnetron then available performed better both in pulse power and in average power at $1 \mu\text{sec}$ than at $2 \mu\text{sec}$, the realization of any improvement in R_{max} would have been doubtful.
3. The many advantages of a short pulse were thought to outweigh any small increase in coverage which a longer pulse appeared to offer.

15-6. Pulse Recurrence Frequency.—The requirement of an adequate time interval for presentation of data usually serves to set the upper limit to the pulse recurrence frequency. Sufficient indicator sweep-time must be provided to present signals without overlapping out to approximately the maximum range of detection. To the longest indicator sweep-time desired must be added sufficient time for recovery of the indicator sweep circuits, in order to arrive at the minimum time interval between pulses. Additional allowance must be made for any departure from regularity in modulator firing.

Presentation of a 200-mile range was desirable, giving a time interval of $2140 \mu\text{sec}$ for the sweep. An additional $360 \mu\text{sec}$ was added for indica-

tor recovery and to allow for irregular firing of the rotary spark-gap modulator, giving an average time interval of 2500 μ sec between pulses, or a ν_r of 400. Thus, the value of ν_r was set close to the permissible maximum for a radar set with a 200-mile range.

For some applications, presentation of greater range would have been helpful. A pulse recurrence frequency of 350 or 300, permitting range presentation to 250 or 300 miles, would have reduced R_{\max} very little, but the characteristics of the modulator chosen restricted ν_r to 400. A value of ν_r of 350 or 300 would be desirable from the standpoint of indicator range presentation for a future radar system of this general type, but two other factors favoring higher ν_r must be considered:

1. The limiting accuracy of azimuth-angle determinations is approximately half of the angular separation between pulses. A scan rate of 36°/sec with a ν_r of 400 gives an angular separation between pulses of 0.09°. The contribution of this effect to azimuth error would be about 0.045°, which is negligible for normal surveillance and control. In operations requiring very precise control of aircraft this error would be appreciable, and if it were doubled by decreasing the ν_r to 200 the resulting error might be a handicap.
2. Cancellation of echoes from stationary objects by MTI means is less effectively achieved as the number of pulses per scan on the target (N_{sc}) decreases (Chap. 16). If the scan rate is 36°/sec, with a beamwidth at half power of 1.0° and a ν_r of 400, N_{sc} will have a value of 11, which is about the lower limit for satisfactory performance of MTI.

15-7. Azimuth Scan Rate.—Although the relation

$$S_{\min} \sim \sqrt{\text{azimuth scanning rate}}$$

was not yet established at the time of the development here described, it was roughly understood; and it put a high premium on slow scanning rates. Mechanical problems in connection with the antenna mount are greatly simplified by a slow scan rate.

Strongly opposing these considerations is the operational need for more complete and up-to-the-second information, in order to follow continuously the movement of high-speed aircraft and to control their movements intelligently. Control problems vary in difficulty from that of a well-formed group of aircraft moving over a simple course to that presented by the requirement for accurate following or control of numerous aircraft moving over complex independent courses.

A plane flying 300 mph moves $\frac{5}{8}$ mile in 10 sec, or $2\frac{1}{2}$ miles in 30 sec. Data at 30-sec intervals, as given by a scan rate of 2 rpm, was considered

adequate for a simple tracking or control problem, but complex problems clearly called for higher scan rates even at the expense of coverage.

Foreseeing a wide range of applications for the new radar equipment, its designers provided an antenna drive by means of which the scanning rate could be adjusted over the range from 1 to 6 rpm. For one wartime application that involved only simple tracking and control problems but put a high premium on range performance, a scan rate of 2 rpm was used exclusively. Other sets, which were used for more complex problems, were commonly run at 4 rpm. Operation at 6 rpm was very limited because of mechanical difficulty.

There is now good evidence that a scan rate of 6 rpm is desirable for complex tracking and control, and even higher rates should be advantageous. Mechanical difficulties at high scan rates and requirements for high azimuth accuracy will probably conspire to set an upper limit on future scanning rates. In fact rates above 6 rpm may be entirely ruled out by these limitations.

15-8. Choice of Beam Shape. *Azimuth Beamwidth.*—The merits of a narrow beam are so widely understood that they will not be discussed in detail here. All microwave development is recognition of the importance of angular resolution. As wavelength is decreased, beamwidths obtainable with antennas of practical size are decreased. With the development of a microwave ground radar, the beamwidth for a long-range air surveillance set was pushed down by an order of magnitude from values previously used for this application, and results were spectacularly successful. The set described here had an antenna aperture 25 ft wide and operated in the 10.7-cm wavelength region, giving a beamwidth of approximately 1.0° .

It will be of interest to see how far development toward increasingly sharper beams might usefully be carried, and to establish as closely as possible the optimum beamwidth for a long-range air-surveillance set. It can readily be shown that a lower limit, and therefore an optimum value, exists, entirely apart from questions of mechanical difficulty.

A scanning rate of $36^\circ/\text{sec}$ gives an angular motion of 0.07° between transmission of a pulse and return of the echo from a target at 200-mile range. If the beamwidth were much less than 0.07° , the antenna gain for reception of the signal would be greatly below normal. For a beamwidth of about 0.15° or 0.2° this effect becomes negligible, and the limit it sets on beamwidth is probably not of practical importance.

A similar and more important fundamental limitation on beamwidth is determined by the requirement that at least one pulse hit the target per scan, with the beam axis off in angle from the target by no more than a small fraction of the beamwidth. At a scanning rate of $36^\circ/\text{sec}$ and a ν_r of 400, the angular separation between pulses is 0.09° , and for

reasonable overlapping of pulses a beamwidth of 0.2° at half power might be considered adequate.

In order to show clearly the effect of beamwidth on R_{\max} , it is necessary to reconsider the range equation [see Eq. (1)],

$$R_{\max} = \left(\frac{P\sigma A^2 f^2}{4\pi S_{\min}\lambda^2} \right)^{1/4}.$$

If we let $A = H \times W$, where

H = height of antenna aperture,

W = width of antenna aperture,

we obtain

$$R_{\max} = \left(\frac{P\sigma f^2 W^2 H^2}{4\pi S_{\min}\lambda^2} \right)^{1/4}. \quad (5)$$

With coverage requirements in height specified, the beamwidth in elevation angle is fixed; therefore, H/λ is a constant. We then have

$$R_{\max} = K_1 \left(\frac{W^2}{S_{\min}} \right)^{1/4}, \quad (6)$$

holding P and σ constant, where

$$K_1 = \left(\frac{P\sigma f^2 H^2}{4\pi\lambda^2} \right)^{1/4},$$

which is now a constant.

From Eq. (2) we see that S_{\min} is inversely proportional to the square root of beamwidth Θ . All other factors that enter into the value of S_{\min} can now be considered as constants and we can therefore write

$$R_{\max} = K_2 W^{3/2} \Theta^{1/2}, \quad (7)$$

where K_2 is another constant. The beamwidth in radians is λ/W to a good approximation; substituting this expression for Θ we have

$$R_{\max} = K_2 W^{3/2} \lambda^{1/2}. \quad (8)$$

From Eq. (8) we see that if the beamwidth is sharpened by decreasing the wavelength and holding the antenna width constant, R_{\max} decreases slowly. The variation of R_{\max} with antenna width is comparatively rapid. If the beamwidth is sharpened to the point where the number of pulses per scan on the target (N_{sc}) is smaller than about 10, Eq. (8) no longer holds. Then R_{\max} becomes a more rapid function of λ and a slower function of W . For values of N_{sc} less than about 5, the variation of R_{\max} with W and λ is given approximately by

$$R_{\max} = K_2 W^{3/4} \lambda^{3/4}. \quad (9)$$

Further decrease of beamwidth below 1° should be advantageous operationally, although improvements due to increased resolution may not be as marked in this region as they were in the region above 1° . If wavelength is decreased below the 10-cm region to sharpen the beam, coverage requirements will become more difficult to meet, even with the assumption that component performance is independent of wavelength. Because of coverage demands and for other reasons to be discussed later, a decrease of wavelength below about 8 cm to sharpen the beamwidth does not appear desirable.

An increase of antenna width beyond 25 ft to sharpen the beam would have the advantage of improving coverage, but the following limitations would be encountered:

1. If the set is to be mobile or easily transportable, an antenna of greater width is cumbersome. With proper construction, an antenna perhaps 35 ft wide might be satisfactory.
2. The requirement (for good gain per unit area) that the reflector shape shall not depart from the shape desired by more than about 0.1λ becomes increasingly difficult to meet as antenna width is increased beyond 25 ft. No difficulty was experienced in maintaining this tolerance with a 25-ft reflector, but the tolerance problem was found to be severe for an experimental 50-ft reflector constructed for a particular application.

These experiences would indicate that an antenna width up to about 35 ft might be desirable. If a wavelength in the region from 8.0 cm to 8.5 cm were used with such an antenna, the beamwidth should be about 0.55° , which appears to be about the practical minimum.

Beam Shape in Elevation.—A beam or a multiple system of beams properly shaped was required to fit as closely as possible the contour of Fig. 15-4. A beam of 3° width in elevation angle with a maximum range of 200 miles fits the desired contour well at long range. For best fit its axis must be elevated about 1.0° above the horizontal. Computations showed that, with antenna width chosen at a practical maximum and with all components pushed to the limit in performance, the coverage as shown in Fig. 15-5 for this 3° beam (Curve A) was the best that could be obtained with a single system. An antenna aperture 8 ft high is required.

Close-in coverage could only be obtained by use of a second system which included transmitter, receiver, and antenna, with antenna width and all components the same as for the first system. Under these conditions an antenna height of 5 ft, with the lower edge of the reflector distorted from its parabolic shape over a small region, gives the coverage shown by Curve B, Fig. 15-5. Improved coverage could have been

provided by addition of a third system, but the additional complexity was thought to be operationally not worth its cost.

Reflection of radiation from the earth's surface greatly influences the radar beam characteristics in elevation angle, and such effects must be given careful consideration. Over water, the reflection coefficient is high for all wavelengths of interest for air surveillance, and for very small angles of incidence it can be assumed to be 100 per cent. Over land, the reflection coefficient varies markedly with detailed surface shape and with the composition of the terrain, and in general increases with increasing wavelength. Adequate data are not available on the reflection coefficient, but at wavelengths of 10 cm or less it is assumed to be normally negligible, and at a few meters it is normally substantial.

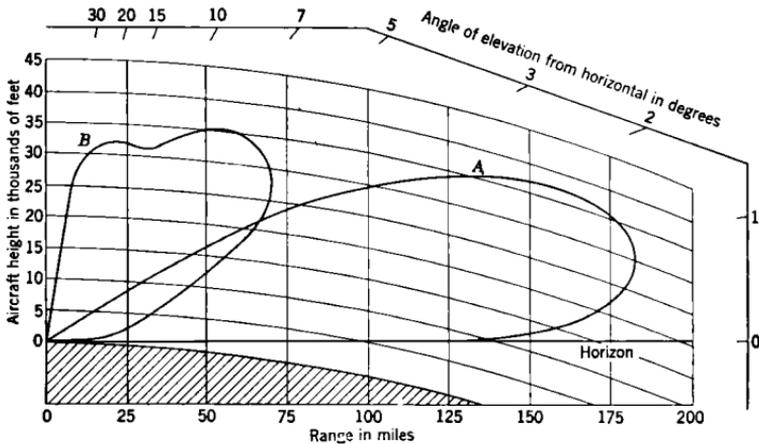


Fig. 15-5.—Calibrated vertical-coverage diagram of actual radar using 7-in. B-scope with antenna rotating at 2 rpm. Effective antenna height is 100 ft. Target is a single B-24.

The detailed effects of surface reflection which influence coverage are very complex. Some of the more important effects and conclusions are summarized below (see also Sec. 2:12).

1. If surface reflection is high, the lowest possible angle of coverage is raised from the horizon to an angle $\alpha = \lambda/2h$, where α is in radians and h is the height of the antenna above the surface.
2. If surface reflection is low, radiation striking the surface is lost. It is clearly desirable to minimize this loss by beam-shaping.
3. Over an angular region in elevation where reflected intensity equals direct intensity, R_{\max} at the center of an interference lobe is doubled, but midway between two lobes $R_{\max} = 0$.
4. At long wavelengths (one meter or greater), the use of reflected radiation reduces requirements on antenna size, but the nulls

between lobes and especially the null region below the lowest lobe are sufficiently broad to handicap detection and tracking greatly. Shaping of a radar beam to reduce to a satisfactorily low value the percentage of energy striking the surface becomes impractical at wavelengths about 25 cm.

5. At short wavelengths (10 cm or below), the angular separation between lobes is small. Loss of tracking due to nulls will normally occur only over short track distances and will not be as serious as it is for long wavelengths. For a system that works at least partially over land it is advisable, however, to minimize the radiation striking the surface by beam-shaping, in order to reduce the waste in energy. Antenna height required for satisfactory beam-shaping is not excessive.

In the region of 25 cm, assessment of the influence of these effects on the general performance of the equipment becomes difficult, but for a high-performance set of general application the above considerations indicate that the best coverage obtainable as a function of elevation angle continues to improve with decreasing wavelength down to wavelengths even below 10 cm.

15-9. Choice of Wavelength.—In the discussion of other parameters, the consideration of wavelength was necessarily included; the many factors influencing choice of wavelength will now be summarized.

Component development has been concentrated on particular wavelength bands. Development of a new wavelength region is so costly in technical effort and time that the wavelength chosen for a new system must ordinarily lie in a region already exploited. Of the wavelength regions in use, development in one may be further advanced than in another, or inherent capabilities in one may be better than in another. These considerations will influence the choice of wavelength.

Excluding the state of component development, the most important factors that influence choice of wavelength are summarized below.

1. Beam characteristics in elevation angle. From previous discussions it was concluded that the best obtainable beam characteristics in elevation angle improve with decreasing wavelength. At wavelengths greater than about 25 cm, beam-shaping to eliminate the undesirable effects of radiation striking the earth's surface is no longer feasible.
2. Coverage. If characteristics other than antenna height are held constant, then maximum detection range decreases slowly as wavelength decreases, but the antenna height required for the maintenance of a given elevation coverage varies directly with wavelength.

3. **Beamwidth.** Regardless of wavelength choice, the maximum antenna width for a transportable air-surveillance set appears to be about 35 ft. The minimum width, as determined by coverage requirements, is probably about 25 ft. A beamwidth of 1° or less is desirable if MTI requirements are ignored, and the most desirable wavelength would then appear to be in the 10-cm region or even shorter. Provision for MTI requires that the azimuth beamwidth be 1° or more, and 10 cm then becomes about the minimum wavelength.
4. **Atmospheric effects.** Atmospheric absorption and variations in refraction can be neglected over the wavelength region under consideration, but the scattering of radiation from water droplets giving rise to echoes from storm centers is an effect of considerable importance (Sec. 3-10). The scattering cross section of the droplets responsible for the signal return varies as $1/\lambda^4$. Return from a large storm center also varies linearly with pulse length and with beamwidth, but these parameters are restricted to relatively narrow regions and the variation of cross section with wavelength is therefore the controlling relation. Dependable operational data are rather limited but the following results appear to be established: (a) At a wavelength of 3 cm, cloud return in some regions frequently obscures a large fraction of the coverage area and seriously restricts the use of the equipment; (b) reports of this difficulty at 10 cm are less frequent, and information furnished on storm centers is valuable for direction of planes around dangerous regions; (c) as wavelength is increased, less of the coverage area is obscured by cloud return; but storm areas that are dangerous to planes may fail to show on the radar at the longer wavelengths.

The wavelength region between 8 cm and 25 cm appears to be best on the basis of all these considerations. Though an optimum wavelength cannot safely be specified, we can make three general remarks concerning the best wavelength for a long-range ground-based system for air surveillance and control:

1. The choice of the 10-cm region for the set described here was excellent, although a wavelength as short as 8 cm might have been more desirable.
2. If MTI is to play an important role in future system performance, and if improvement in MTI performance with increasing beamwidth is sufficiently great, longer wavelengths may be desirable. The best wavelength may then become as great as 15 cm or even 20 cm.

3. The ultimate capabilities of radar components should not differ greatly throughout the wavelength region from 8 cm to 25 cm. Somewhat higher transmitter power will be available at the longer wavelengths. At the time of the development described, performance of existing components was better in the 10-cm region than at any other wavelength suitable for this application.

15-10. Components Design.—The general system parameters chosen as a result of considerations outlined in the last few sections are as follows:

Wavelength.....	10.15 to 11.10 cm
Pulse length.....	1 μ sec
PRF.....	400 pps
Azimuth scan rate.....	1 to 6 rpm
Antenna aperture, lower beam.....	8 by 25 ft
Antenna aperture, upper beam.....	5 by 25 ft
Beamwidth.....	1°

It now remains to describe the components of the radar set.

The Antenna.—The antenna represented a departure from the usual paraboloid of revolution which had been used for earlier microwave antennas. It is described and illustrated in Sec. 9-12.

In the final design of the antenna mount, the reflectors for the upper and lower beams are mounted back-to-back, so that their axes differ by 180°. This complicates somewhat the mechanism for transmission of azimuth-angle data from the antenna mount to the indicators. When a given indicator is switched from upper-beam data to lower-beam data, the mechanism controlling angular orientation of the sweep must be suddenly shifted through a large angle. Operationally, on the other hand, this arrangement was a considerable convenience. A given aircraft could frequently be seen in either beam, and by switching beams every half-revolution of the antenna the operator could double the rate at which he received plots.

More recent advances in antenna development have introduced antenna types, other than the parabolic cylinder fed by a linear array, which give the beam characteristics required for long-range surveillance and control. The method that now appears most promising uses a section of a paraboloid of revolution for the reflector. The periphery of this surface is cut to give an aperture several times greater in width than in height, and the reflector is fed by several horns or dipoles arranged along a vertical line passing through the focus.

The latter antenna system¹ has the following advantages over the linear array:

¹ Similar antennas are described in Chap. 14, Vol. 12 of this series.

1. Azimuth angle of beams does not shift with a change in wavelength.
2. Multiple beams from separate transmitter-receiver systems can be accurately set to the same azimuth angle. This property is very desirable if the system is to be equipped for work with radar beacons.
3. Open mesh or a grid structure can be used for the reflector surface to reduce wind resistance. The linear-array system requires a solid surface for the reflector, since the comparatively high gain of the linear-array feed will give an objectionable back lobe through mesh or gridwork.

The Modulator.—Because of its simplicity and power-handling capacity, an a-c resonance-charging line-type modulator using a rotary

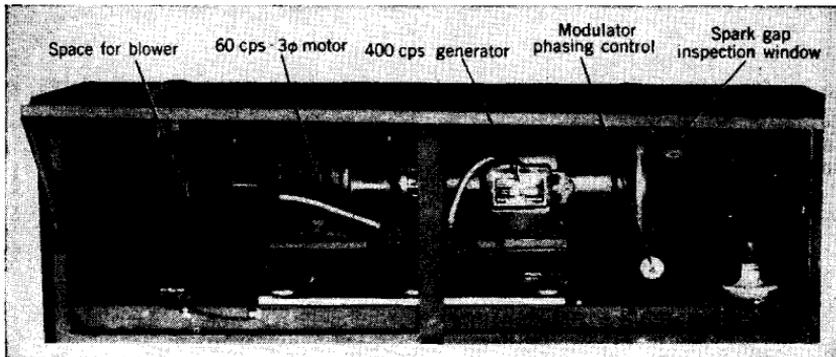


FIG. 15-6.—Modulator of high-power ground radar.

gap as a switching device was chosen. The rotary gap is mounted directly on the shaft of a 400-cps alternator, which excites the network and is properly phased by mechanical adjustment (Fig. 15-6). This modulator choice was important in fixing the value of the PRF, but, as already remarked, the desirable value lies in the neighborhood of 400 pps in any event.

Separate transmitters are used on the upper-beam and the lower-beam systems; both are driven, through pulse transformers, from a single modulator. The modulator is rated at 3000 kw.

R-f Components.—The high pulse powers used in this system demand the use of waveguide for the r-f system. Each magnetron operates in a field of 2800 gauss at 28 kv pulse voltage, drawing a current of about 30 amp and putting 750 kw of r-f power into the line at about 65 per cent efficiency.

Crystal protection at this high power level is difficult, and a scheme of "prepulsing" the TR tube is used. Part of the 8-kv modulator pulse

delivered to the pulse transformer input is capacitively coupled to the keep-alive connection of the TR tube. The delay in the pulse transformer and magnetron is such that this pulse arrives at the TR tube slightly before the r-f pulse arrives. This ensures that the TR tube will fire, improving the crystal protection.

The complete duplexer assembly, which includes the magnetron, its coaxial-output-to-waveguide adapter, the TR switch, the signal mixer, and the AFC mixer, as well as a slotted section for SWR measurements, is shown in Fig. 15-7.

The Receiver.—The receiver is conventional, with an i-f bandwidth of 1.8 Mc/sec and an over-all gain of slightly more than 120 db. Instan-

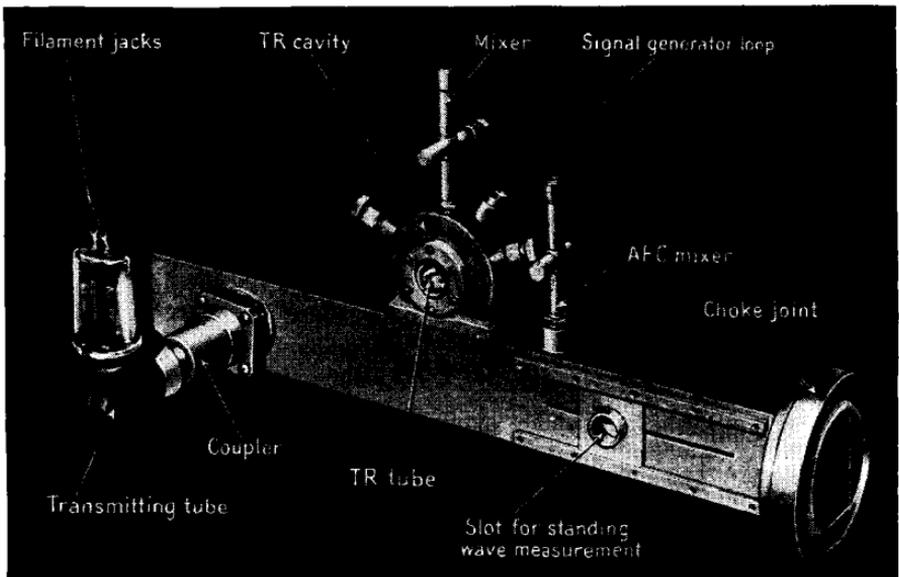


FIG. 15-7.—Duplexing assembly. Two such units are included, one for upper-beam system and one for lower-beam system.

taneous AGC circuits are provided and can be switched in or out at the will of the operator. The AFC operates from a separate mixer and amplifier and is of the hunt-lock type described in Sec. 12.7. Separate receivers are used on the upper-beam and the lower-beam systems.

Indicator Equipment.—The set described here was based on the idea that full flexibility is required in the indicator complement. Both B-scopes and PPI displays can be used, in numbers and with geographical coverage determined by the radar location, the density of targets, and the mission performed by the set.

All data voltages and power voltages for the indicators are supplied from a central point, the so-called "power console," which houses the antenna rotation controls, the transmitter switches, a servo-driven

azimuth gear train that follows the rotation of the antenna mount, and the information generator. The last-named unit contains the range-mark generator (which produces signals 10 statute miles apart, with every fifth one of greater magnitude than the others), a calibrating unit for the range-mark generator, and a unit that generates from an azimuth synchro signal the sawtooth voltage required to provide the azimuth sweep for the B-scopes used with the set.

Azimuth marks 10° apart, with every third mark more intense, are generated in the power console by means of a photoelectric device like that shown in Fig. 13-34.

The B-scopes used are of the magnetic type described in Sec. 13-14 and illustrated in Fig. 13-43. PPI displays are of the rotating-coil type, and are preferably arranged to permit off-centering. The rotating sweep coil is driven by a size-5 synchro energized by a size-6 synchro connected to the azimuth gear train in the power console. Two size-6 synchros are driven by the azimuth gear train; they are set up 180° apart in angle. One provides the drive for PPI displays on the upper antenna beam, and the other for PPI displays on the lower antenna beam.

15-11. Modifications and Additions.—The performance of the radar as finally built was excellent. Some of the PPI photographs in this book, such as Figs. 6-5 and 17-21, show the long range, good definition, and relative freedom from ground clutter which characterize the set.

By the time the set was built and ready for use, the need for it in its originally planned role as a long-range air-warning set had nearly disappeared. The usefulness of such a set in controlling offensive air operations was just beginning to be realized (Sec. 7-6). Accordingly the radar, which had originally been intended for air warning from a fixed site, was provided with facilities, sketched in Sec. 7-6, which made it suitable as a mobile radar control center for use in offensive air warfare.

Apart from the addition of a greater number of indicators than would have been needed for simple air warning, and the provision of off-center PPI's for the use of controllers, the modifications made in the set itself to fit it for its control function were not extensive or important. Another change in the set, also dictated by operational experience, did result in altering it substantially. This will now be described.

Modification for Beacon Use.—The usefulness of ground radar for aircraft control can be greatly extended by modifications that permit it to work with airborne beacons (Chap. 8) mounted in the aircraft to be controlled. Three principal advantages are offered by such a step:

1. Range extension. The useful range against single small aircraft of the set described here is about 90 miles; this can be increased by the use of beacons to the limit of the radar horizon (about 200 miles for an aircraft at 25,000 ft).

2. Positive identification. Aircraft control is frequently complicated by the large number of aircraft operating at the same time within range of a single radar set. Departure of aircraft from planned flight schedules makes reliance on movements information unsatisfactory for identification. A beacon can afford positive identification of the aircraft carrying it.
3. Freedom from ground clutter. By having the beacon reply at a frequency different from that used by the radar for interrogation, the beacon replies can be displayed without any confusing radar echoes.

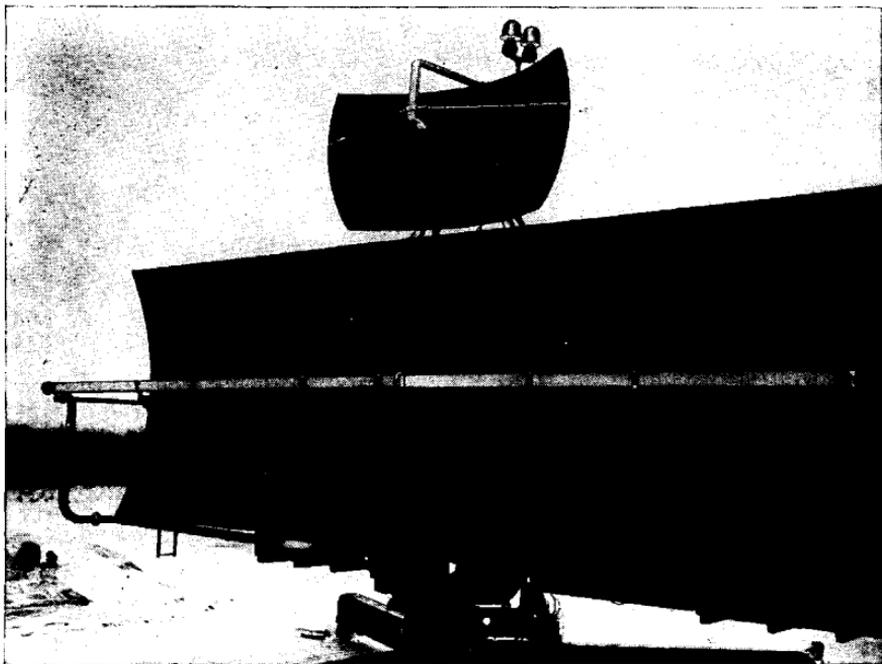


FIG. 15-8.—Beacon receiving antenna mounted above lower-beam radar antenna.

Airborne beacons have been provided, almost without exception, with vertically polarized antennas. They will thus work with the lower beam of the set just described, but not with the upper beam since the latter is horizontally polarized. The beacon is arranged to be triggered by the pulses from the radar, but it replies on a different frequency, to receive which an entirely separate beacon receiving antenna and receiver system are provided. The beacon antenna, mounted on top of the lower-beam radar antenna, is shown in Fig. 15-8. It is a horn-fed paraboloid cut to 8 ft by 4 ft. Despite the greater beamwidth of the beacon antenna than that of the radar antenna, the beacon signals are almost as narrow in

azimuth as radar signals, since the beacon is triggered only over the narrow beam of the radar antenna used for transmitting the beacon challenge.

The beacon receiver has a separate local oscillator and an i-f bandwidth of 10 Mc/sec, to allow for differences in frequency among beacon transmitters in the various aircraft. It is provided with sensitivity-time control (Sec. 12-8). The video output signals of the beacon receiver go to a video mixer with controls so arranged that the indicator will display either radar signals alone, radar and beacon signals together, or beacon signals alone.

DESIGN OF A LIGHTWEIGHT AIRBORNE RADAR FOR NAVIGATION

The next example of radar system design to be described could scarcely be more different. In the design of the ground radar set described, everything was subordinated to attaining the best possible performance; in the AN/APS-10, the set now to be described, performance was important, as always, but it had to be attained within a variety of strict limitations. These limitations dealt with total weight, size, and power consumption, and with the required simplicity of operation, maintenance, and repair. With its communications and height-finding facilities, the total weight of an installation of the ground radar comes to 66 tons; the AN/APS-10 is made up of a few simple units whose total weight is scarcely 120 pounds.

15-12. Design Objectives and Limitations.—By the year 1943, it was clear that airborne radar could offer an extremely important air-navigational facility. Long-wave airborne radar then in use for sea search (Sec. 6-13) could be used for navigation in the vicinity of coastlines, but interpretation of its type L display required long training even under the best circumstances, and was impossible over land, where the multiplicity of echo signals was hopelessly confusing.

Microwave airborne radar with PPI display was just coming into large-scale use, and it was clear that the picture of the ground afforded by such equipment would be a useful navigational aid over any sort of terrain except the open sea. Further, microwave beacons could provide fixed landmarks visible and identifiable at long range. However, the microwave airborne radar sets then in existence had been designed for some specific wartime operational requirement, such as sea search, aircraft interception, or blind bombing. In consequence, their very considerable weight and bulk (arising from the youth of the microwave radar art) were not regarded as serious drawbacks. The radar was necessary in any case to enable the aircraft carrying it to perform its mission, and the navigational use of the radar was only incidental to its main purpose.

Nevertheless, the clarity and convenience of the map-like presentation

of microwave radar suggested that if such facilities could be incorporated in aircraft whose primary mission had little to do with radar—transport aircraft, for example—they could serve an important navigational purpose. Since radar in this application is a very great convenience rather than an imperative operational necessity, it must make the minimum demands on the aircraft in terms of weight, drag, power consumption, and attention required from air and ground crew.

Existing microwave radar scarcely met these requirements in 1943. For example, the AN/APQ-13, a bombing radar whose manufacture was just beginning at the end of that year, has an installed weight of 620 lb without its special bombing facilities, and requires 100 amp of 28-volt aircraft power. It has 24 controls to be adjusted, omitting those concerned with its bombing function. Forty-two cables are required to interconnect its 19 separate component boxes. It is clearly uneconomic to install such an equipment as one of several navigational aids in an aircraft whose primary mission is not high-altitude bombing.

Thus, the design of a new set for this navigational use was indicated; and it appeared that by careful attention to every detail such a set might be made simple enough, light enough, and convenient enough to come into widespread use. The equipment that resulted is the AN/APS-10.

Performance Requirements.—As always, the range performance of the radar was the principal initial requirement. The following ranges (in nautical miles) on various targets were taken as being comfortably over the minimum necessary for navigation even of a high-speed airplane.

Ground painting.....	25 to 30
Cities.....	30 to 50
Storm clouds.....	10 to 40
Mountains.....	25 to 30
5000-ton ship.....	25
Ground beacons.....	150 (if not limited by horizon)
Portable ground beacons (low power) ..	115 (if not limited by horizon)

Since rearward vision is usually as important in radar navigation as forward vision, it was concluded that the set should have a full 360° azimuth scan. In order that the appearance and location of nearby targets should not change too much between scans in an airplane traveling 4 to 6 miles/min, only a few seconds could be allowed per scanner rotation. On the other hand, a reasonable number of pulses per beamwidth had to be allowed.

Requirements on the antenna pattern were relatively simple. The beamwidth in azimuth had to be as narrow as possible, to afford good resolution of targets on the ground. The elevation pattern was required to give the most uniform ground coverage possible throughout the normal flying altitudes: from 1000 to 10,000 ft.

Limitations Imposed by Aircraft Installation.—The performance requirements just mentioned can be simply stated, and it is easy to check whether the completed radar meets them successfully. Neither of these remarks applies to the extremely important set of design limitations that arise from the fact that the radar is to be used in aircraft. The more important of these limitations are mentioned below.

The outstanding requirement is for low total weight. Every part of the design is influenced by the necessity for making the final weight of the radar as small as possible.

Low power consumption is important. Many aircraft types have relatively low electrical-generating capacity, and it is most desirable to be able to install a radar meant primarily as a navigational aid without the necessity of revising the electrical installation of the aircraft. The set must be able to operate either from a fixed-frequency 400-cps motor-alternator set, or from a variable-frequency engine-driven alternator, which may supply power at a frequency as high as 2400 cps.

Aerodynamic drag produced by the antenna housing must be as low as possible. This requirement puts a great premium on as small an antenna as possible. Taken in conjunction with the requirement for a narrow beam, this means that the wavelength on which the radar operates should be as short as is practicable.

Ease of installation in all types of aircraft is important. This suggests that the radar be built in the form of several small components instead of one or two large units, since small units afford greater flexibility in installation. Maintenance is also simplified by dividing the equipment into units that can be replaced for checking and removed for repair. The necessity for connecting such units with cables adds considerably to the total weight of the set, however, and the physical layout must be carefully planned to reduce the number of cables and to keep their lengths as small as possible.

Equipment mounted in aircraft is subject to extreme variations of pressure and temperature. External air pressure can vary from that at sea level to the pressure of less than one-quarter atmosphere found at 30,000 ft. At the reduced pressures of high altitude, clearances required to prevent arc-over from high-voltage points are likely to become excessive. Further, the low temperatures of high altitude are likely to cause condensation, in the r-f line, of the water vapor that is present in warm sea-level air. The simplest solution for these difficulties is to seal the transmitter and the r-f line pressure-tight so that they can be maintained at sea-level pressure at any altitude.

Such pressurization of the modulator and the r-f system increases the problems of cooling. To meet operating conditions on the ground in the tropics, the radar must operate indefinitely at an ambient tem-

perature of 50°C, and must withstand an hour's operation at 71°C. At the same time, because of the low temperatures at high altitude, the set must work satisfactorily at -55°C.

Severe vibration and shock are encountered in aircraft, and some of the weight of an airborne radar must be spent on adequate shock-isolating mountings.

Not only must the rectifier-filter power supplies for the radar circuits operate at any frequency from 400 to 2400 cps; they must also tolerate poor voltage regulation and poor waveform in the primary power supply.

The presence in the same aircraft of high-power radio transmitters, sensitive radio receivers, and a high-voltage ignition system makes it necessary to shield and filter the radar set adequately to prevent it from interfering with other equipment, and to protect it from interference arising in other equipment.

Limitations Imposed by Personnel.—A radar for navigational use in all sorts of aircraft must be manufactured in great quantity with ease and economy, must be so simple to use and understand that only a short period of training is necessary to operate it adequately, and must be so easy to maintain that it can be kept in satisfactory condition with little maintenance effort. Whereas, in principle, the operators of such a specialized radar as a blind-bombing device could be highly trained (though, in point of fact, few were, in the war), the operators of this navigational set were not primarily concerned with learning to operate it properly. Its own simplicity had to be great enough so that its operation was no more difficult to master than that, for example, of the radio compass.

15-13. General Design of the AN/APS-10.—In the framework of these performance requirements and design limitations, the design of the AN/APS-10 was attempted. The most important decision concerned the wavelength to be used. At the outset, experiments were made with a simple PPI radar based on the 10-cm lighthouse-tube transmitter-receiver unit used in the AN/APG-15 (Sec. 6-14), using an antenna of 20-in. diameter. It was found that the range performance was marginal. Worse, the resolution of the set was so low that excessive skill in interpretation was needed to navigate with the set in the absence of well-defined geographical features. Since the larger antenna required to improve both the range and the resolution of this set could not be housed in aircraft without creating excessive drag, it seemed clear that a shorter wavelength would have to be used for general navigation.

The other bands at which components had been developed were around 3.2 cm and 1.25 cm. The 1-cm art was in a very early state, and equipment was still cumbersome and of poor performance. For this reason, 3.2 cm was chosen. This was a fortunate choice; for

although it was not then known, the absorption of 1.25-cm radiation in water vapor is so strong as to remove almost entirely the usefulness of this wavelength for navigational purposes.

Next, a decision had to be taken regarding the pulse power. High pulse power is very costly in the design of radar for aircraft, both in terms of the primary power requirement and in terms of total weight.¹ A comparison with the performance of the AN/APQ-13 served as a guide in determining the pulse power of the AN/APS-10. The AN/APQ-13, using about 50-kw pulse power and a high-altitude cosecant-squared antenna of gain 950, showed a range of about 45 miles for general ground painting and 50 to 100 miles on cities. Its receiver had a sensitivity of about 5×10^{-13} watts for a signal equal to noise.

The 45-mile ground-painting range of the AN/APQ-13, as compared with the 25 miles required of the AN/APS-10, allowed a 10-db reduction of transmitter power for the latter, if the antenna gain and the receiver sensitivity were kept the same. However, it was felt that a 30-in. antenna was too large to be tolerated for the AN/APS-10. Experiments showed that an 18-in. cosecant-squared antenna designed for an altitude of 7000 ft would be useful in the usual altitude range from 1000 to 10,000 ft without unduly overilluminating short-range targets at low altitude or causing too serious a hole in the center of the pattern at altitudes higher than 7000 ft. This antenna proved to have a gain of 700; this meant that the reduction of antenna gain would require about 3 db of the calculated 10-db leeway between the AN/APQ-13 and the AN/APS-10.

The remaining 7 db could be and were used to permit a reduction in pulse power. The lightweight, low-voltage 2J42 magnetron (see Fig. 10-4b) was used in the AN/APS-10, giving an output pulse power of 10 kw. The lower voltage and lower power required by this tube enabled the design of an extremely compact and simple pulser, shown in Fig. 10-47.

Since the whole of the 10-db differential between the two sets had been used up in the smaller antenna and the lower pulse power of the AN/APS-10, it was necessary to make sure that the receiver sensitivity of the latter set did not fall below that of the AN/APQ-13. Further, special pains had to be taken to provide good range performance against microwave beacons. In the AN/APQ-13, the TR-switch tuning was not changed from search to beacon operation, and a narrow-band ATR was used. The loss of beacon signal in the TR and ATR alone could be as much as 15 or 20 db. In addition, no automatic frequency control of the beacon local oscillator was provided. Better facilities for working with beacons were desired in the AN/APS-10.

¹ W. L. Myers, "Weight Analysis of Airborne Radar Sets," RL Report No. 450, Jan. 1, 1945.

15-14. Detailed Design of the AN/APS-10.—With these major design decisions taken, the detailed components design of the AN/APS-10 remained. No attempt will be made to discuss the many interacting decisions that determined the design of the components of the set;

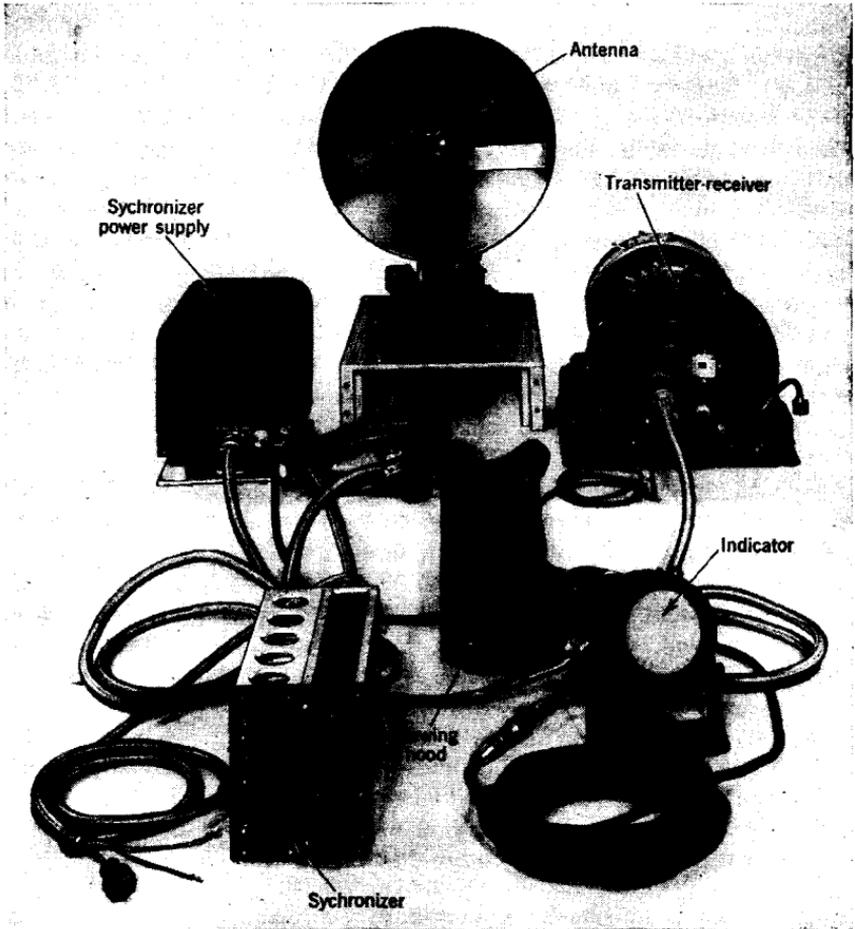


FIG. 15-9.—Units of AN/APS-10.

only the final result will be presented, and an attempt made to rationalize it in terms of the considerations of Sec. 15-12.

A photograph of the components of the set is shown as Fig. 15-9. Figure 15-10 is a schematic diagram of the various units and their major subassemblies. The set consists of five major units: a transmitter-receiver, a synchronizer, a synchronizer power supply, a scanner, and an indicator. There are two minor components, a trim-control box and a

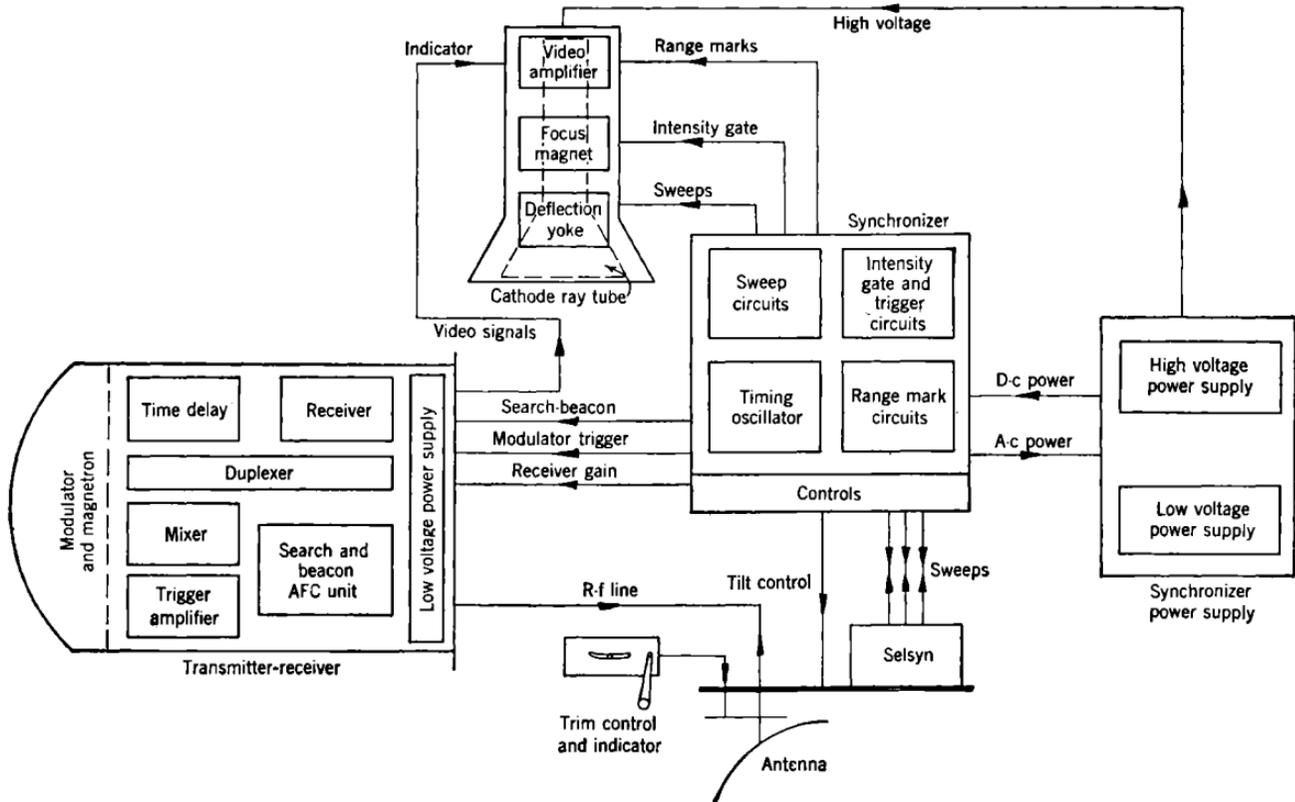


FIG. 15-10.—Block diagram of AN/APS-10 system.

TABLE 15-1.—UNITS OF THE AN/APS-10

Unit	Weight	Weight of mounting	A-c power, watts	D-c power, watts	Maximum over-all dimensions of units, in.			Pressurization
					H	W	D	
Transmitter-receiver.....	46 lb	4 lb	185	15-20	12 $\frac{5}{8}$	12 $\frac{5}{8}$	20 $\frac{9}{16}$	Yes
Synchronizer.....	13 lb 1 oz	13 oz	...	3'	9 $\frac{5}{8}$	6 $\frac{5}{8}$	14 $\frac{3}{8}$	No
Synchronizer power supply.....	19 lb 7 oz	1 lb 4 oz	150	11 $\frac{1}{2}$	9	12	No
Scanner*.....	20 lb	20-30	28	18	18	Feed and r-f line pressurized
Indicator and visor.....	8 lb 5 oz	5 oz	Included with sync.	6 $\frac{3}{4}$	8	13 $\frac{1}{2}$	No
Pressure pump.....	6 lb
Trim control indicator and flexible shaft.....	2 lb 8 oz	3 $\frac{5}{8}$	5 $\frac{1}{8}$	4 $\frac{1}{2}$
	2 lb 6 oz

* These figures apply to the standard AN/APS-10 scanner. A special lightweight (13 lb) scanner, suitable for most applications, was also developed for AN/APS-10. (See Sec. 9-12 and Fig. 9-16.)

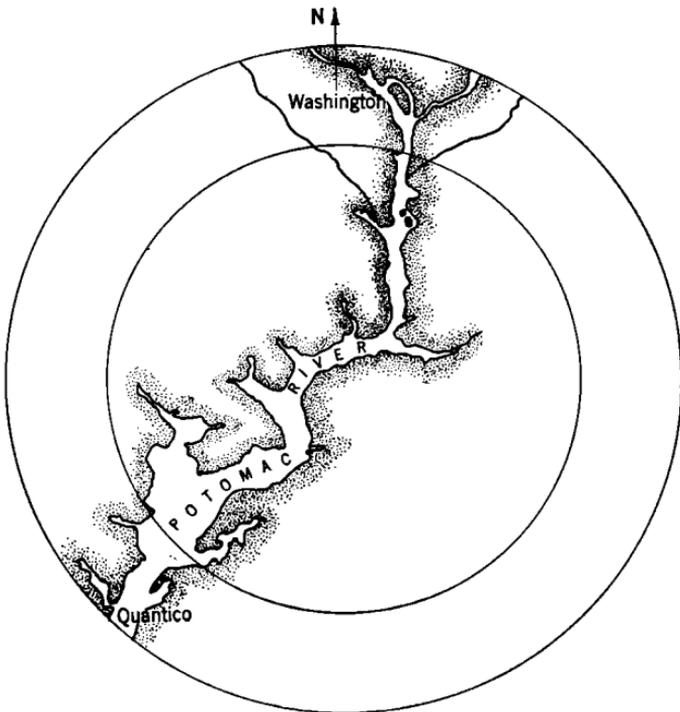
pressure pump. The weight, power consumption, and size of the units are summarized in Table 15-1. Table 15-2 gives more detail on the design of the various units.

The transmitter-receiver unit contains the magnetron, the modulator and its power supply, the duplexer, the TR switch, the mixer, and the local oscillators, as well as a motor-driven time-delay switch for modulator turn-on, and a trigger amplifier for the modulator. The receiver strip,

TABLE 15-2.—DETAILED SPECIFICATIONS OF THE AN/APS-10

Total weight of units and mountings.....	124* lb less cables
Primary power required	
Alternating current.....	115 v. 400-2400 cps, 340 watts
Direct current.....	27.5 v, 80 watts
Over-all system performance	
Transmitter frequency.....	9375 \pm 55 Mc/sec
Pulse power.....	10 kw
Receiver sensitivity, search.....	131 db below 1 watt
Receiver sensitivity, beacon.....	125 db below 1 watt
Pass band of untuned r-f components.....	9280 \pm 185 Mc/sec
Modulator	
Type.....	Line-type with hydrogen thyratron
PRF, search.....	810 pps
PRF, beacon.....	405 pps
Pulse length, search.....	0.8 μ sec
Pulse length, beacon.....	2.2 μ sec
Receiver	
Intermediate frequency.....	30 Mc/sec
Bandwidth of i-f amplifier.....	5.5 \pm 1.0 Mc/sec
Indicator	
Type (Sec. 13-17).....	5-in. resolved-current PPI
Focus.....	Permanent magnet
Cathode-ray tube.....	5FP14 (or 5FP7)
Video bandwidth.....	3 Mc/sec
Antenna	
Reflector.....	18-in. paraboloid; $f = 5.67$ in; 7000-ft cosecant-squared pattern
Feed.....	Two-dipole type
Gain.....	700
Polarization.....	Horizontal
Azimuth beamwidth.....	5°
Scanner	
Scan rate.....	30 rpm
Tilt angle.....	0° to -18°
Trim angle.....	6½° to -10½°
Number of tubes	
Transmitter-receiver.....	31 tubes and 3 crystals
Synchronizer.....	25
Indicator.....	4
Rectifier power supply.....	6

* With standard scanner. Lightweight scanner reduces total weight to 117 lb.



(b)

FIG. 15-11.—AN/APS-10 low-altitude PPI. Range mark at 10 miles; altitude 1600 ft.; Washington, D. C. at 0°, 8 to 13 miles. (a) Photograph of scope. (b) Sketch from map.

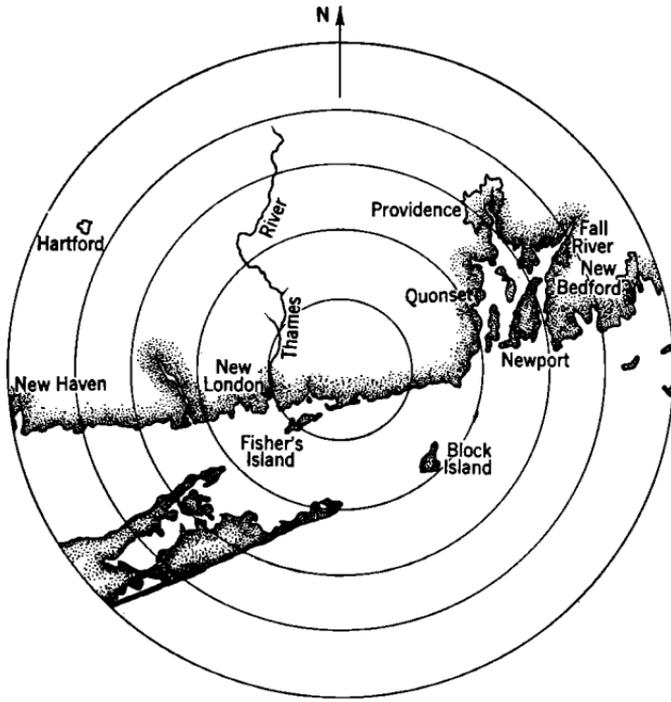
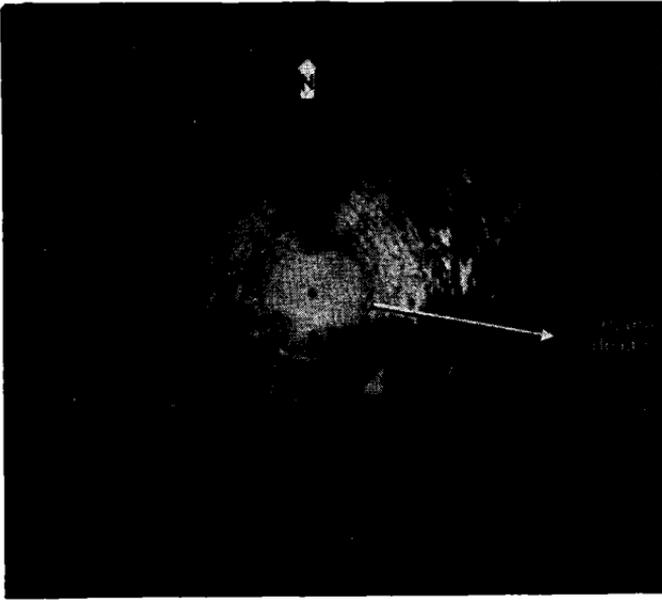


FIG. 15-12.—AN/APS-10 medium-altitude PPI. Special 30-in. antenna used. Range marks 10 miles apart; altitude 6000 ft.; Narragansett Bay at 350° , 20 to 30 miles; Long Island at 160° , 20 to 40 miles. (a) Photograph of scope. (b) Sketch from map.

(containing the i-f amplifier, the detector, and a video output stage), the automatic frequency control unit, and the low-voltage power supply for these circuits, are also mounted in this unit. The detailed design of this receiver is treated in Sec. 12-10.

The synchronizer unit includes the primary timing circuits, and the sweep, range-mark, and intensifier circuits. On its front panel are the operating controls for the system. This unit serves as a cable junction box, and its use in this capacity has enabled the set to be designed with a total of only seven cables. Since the circuits are designed to allow for cable capacitance and for voltage drops, any of these cables can be as long as 25 ft. Separation of the synchronizer power-supply unit from the synchronizer itself permits the latter to be small enough to be conveniently used as the main control box. It can be mounted with the control panel either vertical or horizontal, as each installation may demand.

The indicator unit contains a video amplifier, in addition to the cathode-ray tube and its deflection yoke and focus magnet.

The synchronizer power supply provides the necessary voltages for the indicator and the synchronizer.

Performance of the AN/APS-10.—The completed set shows performance meeting the initial requirements. Scope-map comparisons that display the usual overland performance of the set at low and medium altitude are shown in Figs. 15-11 and 15-12. The blanked-out sectors abeam are due to shadows cast by the engine nacelles of the C-47 aircraft in which the set was installed. Figure 15-12 was taken while a special experimental antenna of 30-in. horizontal aperture was in use; thus the azimuth resolution and range performance are better than those attained with the standard 18-in. antenna.

The performance of the set with 3-cm beacons is especially good because careful attention was paid to the radar-beacon problem in design. The major changes involved in switching from search to beacon operation are:

1. Change of pulse length from 0.8 to 2.2 μ sec.
2. Change of PRF from 810 to 405 pps.
3. Change of TR-switch resonant frequency from the magnetron frequency to the beacon frequency.
4. Change from search to beacon local oscillator.
5. Change from search to beacon AFC. (Sec. 12-7.)
6. Stretching of beacon reply pulses in the video amplifier.

A broadband ATR tube is used primarily to pass beacon signals with minimum loss, and a wider i-f band than necessary for optimal signal-to-

noise ratio is employed to insure reception of signals even from beacons that have drifted off frequency.

The long (2.2- μ sec) pulse necessary for beacon triggering required a reduction in PRF in order to keep the duty ratio of the modulator nearly that used for search. The PRF of 810 used in search was set by the requirements of the indicator circuit; it is halved for beacon operation.

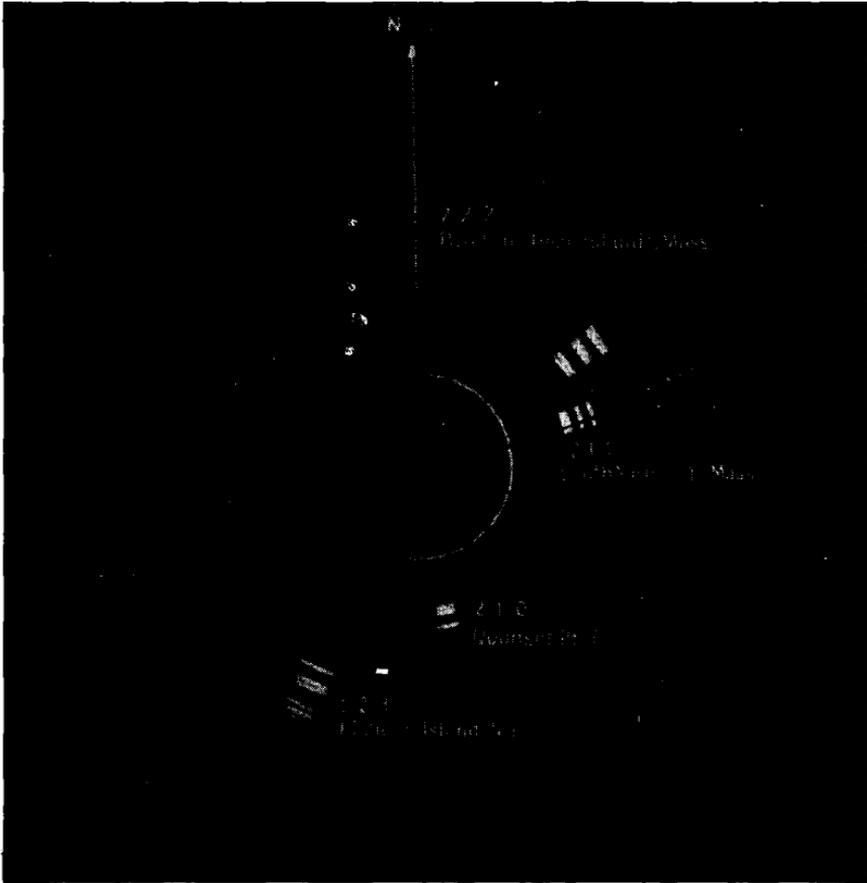


FIG. 15-13.—Beacon replies on AN/APS-10 scope. Range marks 20 miles apart.

At a repetition rate of 405 pps with a 5° beam, five to six effective pulses are sent to a beacon per scan, since only half the pulses incident on the beacon are effective because of its switched local oscillator (Vol. 3, Sec. 8-13). Figure 15-13 shows the responses from four beacons appearing at once on the indicator of an AN/APS-10 radar.

Suitability for Aircraft Installation.—The low power consumption and low weight of the AN/APS-10 are shown in Tables 15-1 and 15-2. The

net weight of the components is about 120 lb. To this must be added 12 lb for cables and brackets, 28 lb for an inverter, and 25 lb for a radome, so that the installed weight of the set is about 185 lb, well under a third that of the AN/APQ-13. Snap fasteners and spring-loaded bands are used as retaining elements to facilitate the removal and replacement of units.

To achieve low weight and simplicity, many attractive design features had to be sacrificed, including antenna stabilization, sector scan, ground-range sweeps, and long pulses for maximum range performance. The possibility of adding to the AN/APS-10 special units to perform special functions has, however, been retained. A trigger pulse and video signals are available for any attachment, as are provisions for an additional azimuth-angle take-off and means for externally reversing the azimuth scan motor.

Adequate filtering and shielding are provided to deal with any ordinary problems of mutual interference between the radar and other electronic equipment installed in the same aircraft.

The transmitter-receiver is protected against changes in external pressure by means of a rubber-gasket seal and pressure-tight cable connectors. The r-f line is also sealed and is connected to a pressure pump with a dryer on its intake.

Internal and external fans are provided for the unit in which there is the greatest heat dissipation—the transmitter-receiver. The internal rise above ambient temperature in this unit is 35°C.

All units except the scanner are shock-mounted. The r-f line between the shock-mounted transmitter-receiver and the rigidly attached scanner has a pressure-tight flexible section. Where possible, as in the case of the transmitter-receiver, center-of-gravity shock mounts are used.

To minimize the effect of variations in supply voltage and waveform, full-wave rectifier power supplies with choke-input filters are used throughout the set. Electronic regulation of voltage has been used where necessary to maintain precise voltage or to remove the low-frequency ripple found in the output voltage of many aircraft alternators.

Operational Suitability.—In contrast to the 24 controls of the AN/APQ-13, the AN/APS-10 has only 10. Four must be used often, four infrequently, and the remaining two are primarily for the convenience of the operator. The four commonly used controls are:

1. A range-selector switch that enables the operator to choose either a continuously variable, 4- to 25-mile sweep or a 50-mile sweep for search operation, or a 0- to 90-mile or 70- to 160-mile sweep primarily intended for beacon operation. Two-mile range marks are provided on sweeps shorter than 14 miles, 10-mile marks on

sweeps between 15 and 50 miles, and 20-mile marks on the beacon sweeps.

2. A receiver gain control.
3. A tilt control for the antenna. This enables adjustment of the depression angle of the radar beam to the optimal value for any altitude. A tilt meter calibrated in altitude is provided; it serves as an approximate guide to the correct setting.
4. A search-beacon switch.

The controls less often used are:

1. The OFF-ON switch.
2. The focus control of the indicator.
3. The brilliance control of the indicator.
4. The trim control. This enables the operator to adjust the scanner mechanism in pitch so as to keep the axis of the scanner vertical despite changing attitude in flight due, for example, to consumption of fuel. A fore-and-aft bubble level and a hand crank connected by flexible shaft to the scanner are provided.

The two "convenience" controls are the adjustments for range-mark intensity and for dial-light brilliance.

The operation of the AN/APS-10 is so simple that an hour's flight experience is sufficient to qualify a navigator to adjust and use the set.

For ease in maintenance, all units of the set have been made independently replaceable, with no need of adjustment for individual units. Even within the major units of the set, subassemblies have been designed for replacement in the event of failure.

External test points have been provided on the transmitter-receiver and the synchronizer power supply to aid in identifying defective units in cases of failure. On the transmitter-receiver these include a test trigger from the modulator, an extra video channel, and a lead from a directional coupler which enables r-f checking. On the synchronizer, power supply pin jacks are provided to permit measurement of power-supply voltages. Other test points are incorporated on many of the subchassis of the transmitter-receiver unit. These external test points permit a routine procedure of inspection and preventive maintenance to detect incipient failure and insure peak performance.

The consensus of its users is that the AN/APS-10 fits its requirements well. Further improvement in lightness, convenience, and modesty of power demand will follow on general advance in the art, and in particular on detailed attention to component design. To mention one example, the substitution of subminiature tubes for larger standard types will reduce bulk, weight, and power consumption.

CHAPTER 16

MOVING-TARGET INDICATION

BY A. G. EMSLIE AND R. A. McCONNELL

INTRODUCTION

16.1. The Role of Moving-target Indication.—The object of moving-target indication (MTI) is to present the signals received by a pulse radar set in such a way that moving targets show up while stationary objects give no response. The most advanced method of doing this allows the moving targets to be presented on a PPI. Figure 16·1 shows two PPI photographs, one with MTI in operation and one without, taken on a ground radar set at Bedford, Mass., using a wavelength of 10.7 cm and a PRF of 300. The removal of the ground clutter is seen to be complete. The photographs in Fig. 16·2 were taken at Boston, Mass., where, because of screening by a ring of low hills, clutter does not extend appreciably beyond 10 miles. In these photographs, taken on a set having a wavelength of 10.7 cm and a PRF of 1650, the effectiveness of MTI in reducing storm echoes is shown.

The problem of MTI is somewhat more difficult and the results less satisfactory when the radar set is carried on a moving ship or airplane because the clutter to be eliminated is itself moving relative to the radar set and the clear-cut distinction between moving and stationary targets is lost. In spite of this, it is possible to arrange MTI so that vehicles moving on the ground can be seen from an airplane. In the case of a shipborne set it is possible to compensate for the ship's own motion and therefore to see other ships and aircraft in the presence of sea-clutter and storm echoes.

16.2. Basic Principles of MTI.—Two fundamental ideas are involved in the solution of the MTI problem: first, a method of reception that responds differently to fixed and to moving targets; second, an arrangement that takes advantage of this difference by selecting only the moving targets.

The method of reception always uses the doppler effect in one form or another. The simplest arrangement is that shown in Fig. 16·3. Power from the transmitter is mixed with the c-w echo from the target. After detection, the beats between the two frequencies f and f' can be

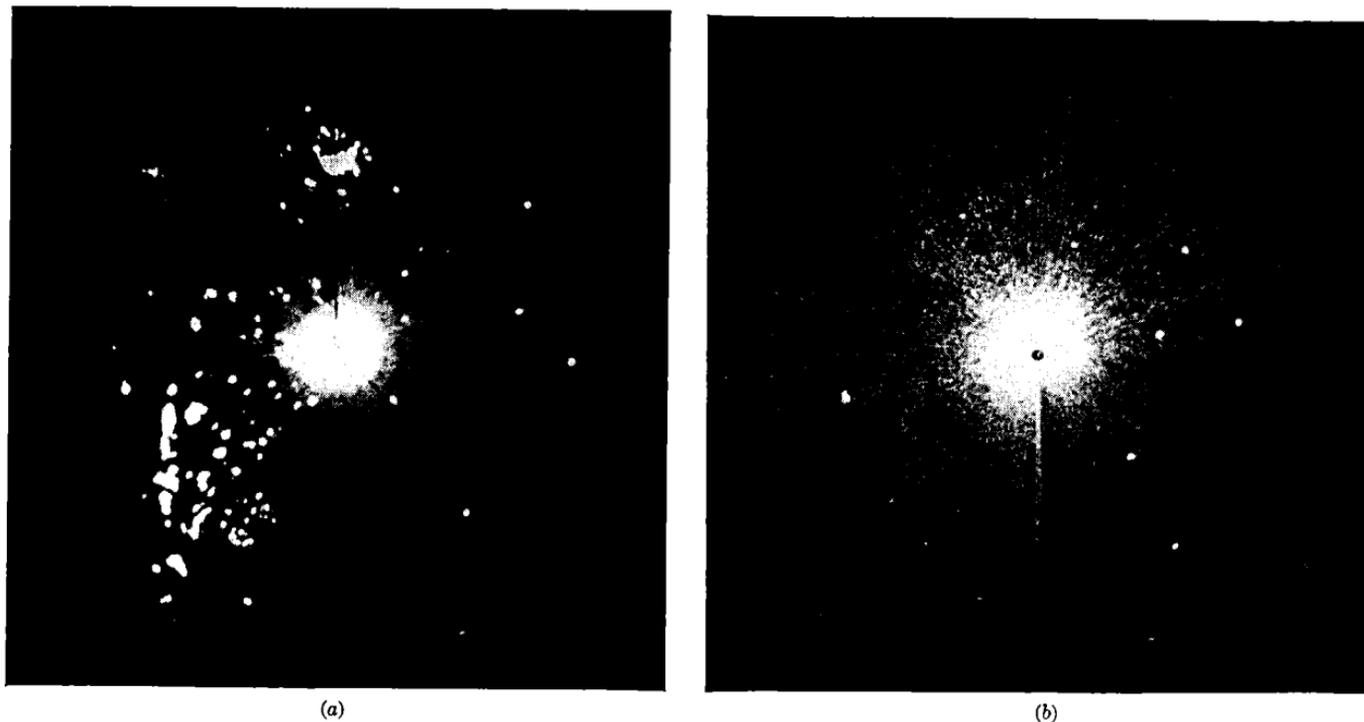


FIG. 16-1.—Comparison of normal and MTI PPI presentations: (a) normal scope, 35-mile sweep; (b) scope with MTI, 35-mile sweep.

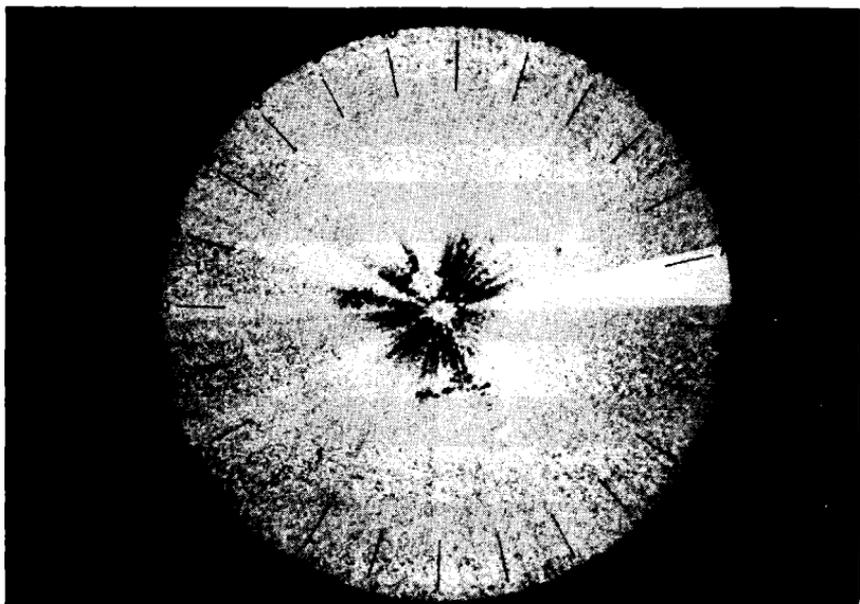


FIG. 16-2.—Successive 30-mile normal and MTI PPI photographs. Note storm echoes at 10° at 25 miles, 60° at 15 miles, and 150° at 30 miles. Seven aircraft are visible on the MTI PPI.

heard in the phones. The echo frequency is given by the well-known doppler formula

$$f' = \frac{c + v}{c - v} f.$$

The beat frequency is then

$$f_d = f' - f = \frac{2v}{c - v} f.$$

Since the target velocity v is very small compared with the velocity of light c , this can be written as

$$f_d = \frac{2v}{c} f = \frac{2v}{\lambda}.$$

For v in miles per hour and λ in cm this becomes

$$f_d = \frac{89v}{\lambda}.$$

See, for example, Fig. 5-12.

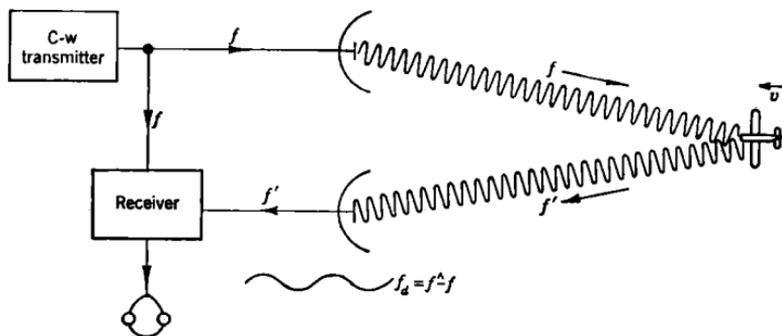


FIG. 16-3.—Doppler effect with c-w system.

The arrangement of Fig. 16-3 gives no range information. This can be remedied by chopping up the outgoing train of waves as in Fig. 16-4. The beats in the telephones now consist of a succession of pulses whose envelope has the doppler frequency f_d . Note that the doppler effect can be viewed as causing a phase shift of the echo from pulse to pulse. It is easy to calculate this phase change and to show that it is equivalent to the frequency shift. The distance traveled by the target between pulses is vT , where T is the repetition period. Hence each pulse travels a distance $2vT$ less than the preceding pulse. This is $2vT/\lambda$ wavelengths, so that the phase change is $2\pi \cdot 2vT/\lambda$ between each pulse and the next. The beat frequency is then $2v/\lambda$, as before.

To obtain range information with the scheme of Fig. 16-4, the phones can be replaced by an A-scope synchronized with the modulator. The appearance of the A-scope is shown in Fig. 16-5, where one moving target

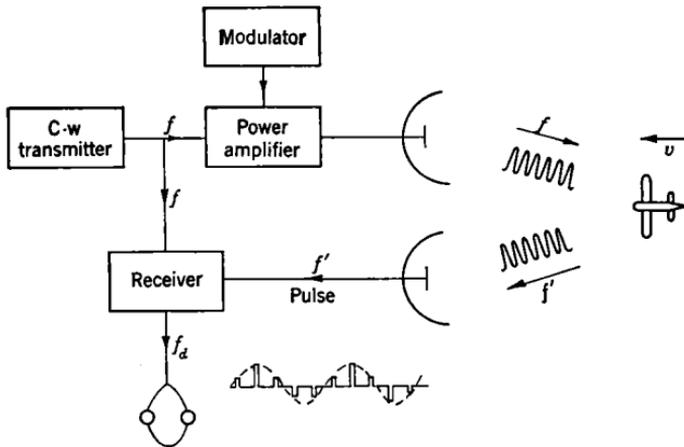


FIG. 16-4.—Doppler effect with pulsed system suitable for low frequencies.

is seen among several stationary ones. The butterfly-like appearance of the moving target is the result of the variation in pulse amplitude as shown in Fig. 16-4. Stationary targets exhibit constant phase from pulse to pulse, and therefore a steady amplitude.

The arrangement of Fig. 16-4 can be used at frequencies of a few hundred megacycles per second but is not practical at microwave frequencies in the absence of suitable power amplifiers. However, the



FIG. 16-5.—Doppler beats on an A-scope.

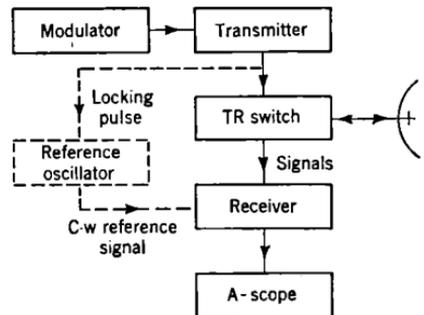


FIG. 16-6.—Doppler system for microwaves. Dotted lines show what must be added to ordinary radar set.

same effect¹ can be obtained by the circuit of Fig. 16-6, which shows the basic diagram of a microwave pulsed radar, with the addition of a reference oscillator. This reference oscillator provides a c-w signal with which to beat the incoming echoes. Since the transmitter starts with random phase from pulse to pulse, it is necessary to match the phase of the reference oscillator to that of the transmitter at each transmitted pulse.

¹ "Second time around" echoes—signals received from the second preceding pulse—will not be canceled by the arrangement of Fig. 16-6.

This can be done by allowing a sufficient amount of power from the transmitter to enter the resonant cavity of the oscillator, which is then forced into step with the transmitter. This process is called "locking" the phase of the oscillator, or making it "coherent." The appearance of the A-scope will be the same as that shown in Fig. 16-5.

Two methods of presenting the doppler information have been mentioned: the aural method (Fig. 16-3), and the A-scope method (Fig. 16-6). These indicators are useful only with a stationary or very slowly rotating antenna. At normal rates of scanning, it is necessary to display signals on a PPI in such a way that only moving targets appear on the scope. This can be done by the method of pulse-to-pulse cancellation illustrated in Fig. 16-7, where the first four traces represent successive sweeps on the A-scope of Fig. 16-5. The scheme is to delay the signals of Sweep 1 for a whole repetition period and then subtract them from the signals of Sweep 2. In the same way Sweep 2 is delayed and subtracted from Sweep 3, and so on. The results are shown in the last three traces

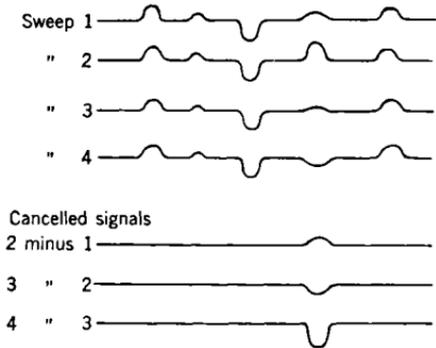


FIG. 16-7.—Pulse-to-pulse cancellation. The first four traces represent successive sweeps of A-scope. The last three traces show the canceled signals.

of Fig. 16-7. This process can be carried out continuously by the arrangement of Fig. 16-8, in which the signals are split into two channels, one of which contains a supersonic delay line, and then brought together again for cancellation. An important practical point in the use of this scheme is the degree to which the output signal of the delay line simulates the input signal, for the fidelity of reproduction of signals by the delay line influences the maximum clutter cancellation possible.

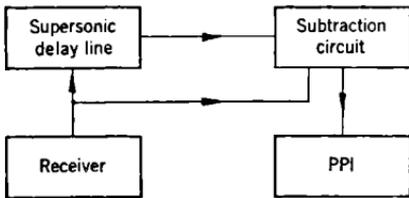


FIG. 16-8.—Cancellation of signals of constant pulse-to-pulse phase.

It is also possible to delay the signals by means of a "storage tube," which works in a manner related to that of the Iconoscope used in television. The supersonic delay line was used as a delay device in the MTI systems that have had the most thorough testing; its use is therefore assumed in what follows. Certain advantages attend the use of a storage tube, notably the ability to apply MTI to a system whose

repetition rate has "time jitter" (see Vol. 19, Chap. 25 of the Radiation Laboratory Series).

16-3. A Practical MTI System.—Figure 16-9 is a block diagram of a practical MTI system. It can be regarded as a refinement of the arrangement of Fig. 16-6; instead of beating the signals with an r-f reference signal, the same thing is now done at intermediate frequency. This is accomplished by applying the superheterodyne principle both to the locking pulse and to the signals. The same stable local oscillator provides a signal to two mixers, one of which reduces the frequency of the locking pulse, the other the frequency of the signals, to 30 Mc/sec. The i-f locking pulse from the first mixer is applied to an i-f oscillator which is thereby rendered coherent in phase with the locking pulse.

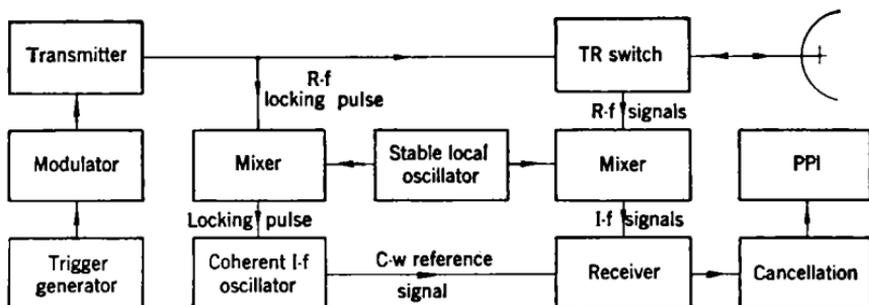


FIG. 16-9.—Practical MTI system.

The circuits of Figs. 16-6 and 16-9 are entirely equivalent as far as the output of the receiver is concerned. The phase of the i-f echo signal from a stationary target depends on the starting phase of the transmitter, the starting phase of the local oscillator, and the range (which determines the number of cycles executed by the local oscillator while the transmitted pulse travels to the target and back). The i-f reference signal provided by the coherent oscillator has a phase that depends on the starting phases of the transmitter and of the local oscillator, and on the range (which determines the number of cycles executed by the coherent oscillator during the echo-time). The starting phases of the transmitter and the local oscillator cancel out when the i-f echo signals and the i-f reference signal beat against each other, so that the phase of the receiver output depends only on the number of cycles executed by the local oscillator and by the coherent oscillator. Both of these oscillators are made to be stable; consequently the beat signal from a stationary target has a steady amplitude from pulse to pulse. When the target is moving, its range will change from pulse to pulse and a fluctuating output signal results from the corresponding change in the phases traversed by the oscillators during the echo-time.

A simple laboratory type of delay line is shown in Fig. 16-10. It consists of a steel tube filled with mercury, with X-cut quartz crystals cemented to its ends by means of lacquer. An end cell filled with mercury is attached to the outside of each crystal by the same method. When an alternating voltage is applied across one of the crystals (that is, between the mercury in the end cell and that in the tube), the crystal undergoes periodic changes in thickness due to the piezoelectric effect. The vibrations of the crystal are communicated to the mercury as a supersonic wave which travels down the tube at a speed of 4700 ft/sec, corresponding to a delay of $17.6 \mu\text{sec}/\text{in.}$ On arriving at the other end of the line the wave causes the receiving crystal to vibrate, and this vibration produces an alternating voltage between the faces of the crystal.

The mercury end cells, in addition to acting as electrodes, serve also to prevent reflection of the supersonic wave by the crystals. When the crystal is driven at its resonant frequency, it acts as a

half-wave acoustical transformer; thus the delay line is properly terminated at each end by the end cells. The wave entering an end cell is broken up and absorbed by the skew back of the cell.

A free quartz crystal has a very high Q and therefore such a narrow bandwidth that it cannot transmit microsecond pulses. In the delay line, the loading of the crystal by the mercury causes the crystal to be almost critically damped—that is, to have a Q near unity. Thus the bandwidth of the crystal, when it is cut, for example, for 15 Mc/sec, is many megacycles per second. The frequency response curve has its maximum at the resonant frequency of the crystal and falls to zero at zero frequency. Thus it is not possible to transmit video pulses directly through the delay line without distortion. Instead, the pulses are used to modulate a carrier whose frequency is that for which the crystal is resonant.

Figure 16-11 shows, in more detail than Fig. 16-8, the arrangement of the cancellation circuits. The video signals from the MTI receiver amplitude-modulate a 15-Mc/sec oscillator and amplifier of conventional television design. Its output signals go to the input crystal of the 15-Mc/sec delay line, and also to the “undelayed” channel of a two-channel amplifier. The “delayed” channel of the amplifier receives the output signal from the delay line and amplifies it. The two channels of

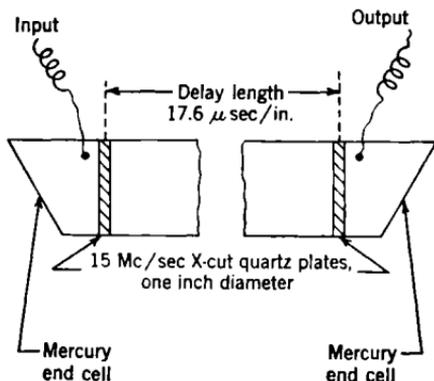


FIG. 16-10.—Laboratory-type supersonic delay line.

the amplifier have separate diode detectors arranged to give opposite signal polarities, so that the delayed and the undelayed signal can be adjusted in amplitude to cancel each other when added. Since the video signal from a moving target is bipolar (cf. Fig. 16-7), a rectifier is included in the circuit ahead of the PPI.

In order to get good cancellation, the signals in the two channels of the amplifier must match very closely in time. This means that the

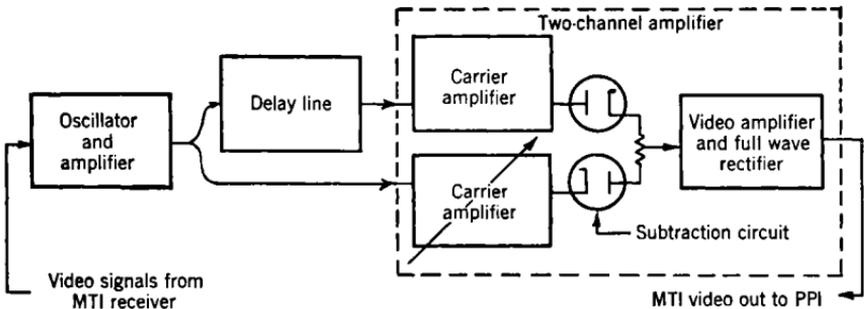


FIG. 16-11.—Block diagram of delay-line circuits.

pulse-repetition period has to match the delay time of the supersonic line with great accuracy. The velocity of sound in mercury varies with the temperature by about one part in 3000 per degree centigrade, so that temperature variations will cause the delay time to drift. This can be compensated for by providing a trigger generator whose PRF is altered to take account of changes in the delay time. A simple way of doing this is to use another supersonic delay line as the timing element in the trigger

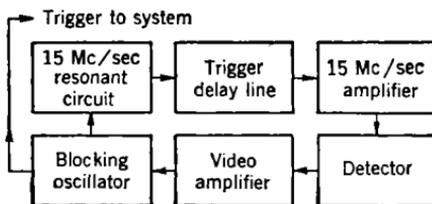


FIG. 16-12.—Trigger generator.

generator, as shown in Fig. 16-12. The circuit works as follows. When the blocking oscillator fires, it delivers a trigger to the system and also shock-excites a 15-Mc/sec resonant circuit. The oscillations so set up are passed through the delay line, and thereafter amplified and detected. The delayed signal is used as a trigger; it is applied to the blocking oscillator, which then fires and starts a new cycle. By making one of the delay lines variable in length (which requires a different design from that of Fig. 16-10) the delays in the trigger and signal circuits can be matched

initially. The two lines are mounted close together in a thermally insulated box so that the temperature changes equally for both.

16-4. Alternative Methods for Obtaining Coherence.—Figures 16-6 and 16-9 show two different ways of producing coherence between echo signals and a reference signal: in the first method, the oscillator whose phase is locked runs at radio frequency and the signals are added at radio frequency; in the second, the locking and adding are both done at the intermediate frequency.

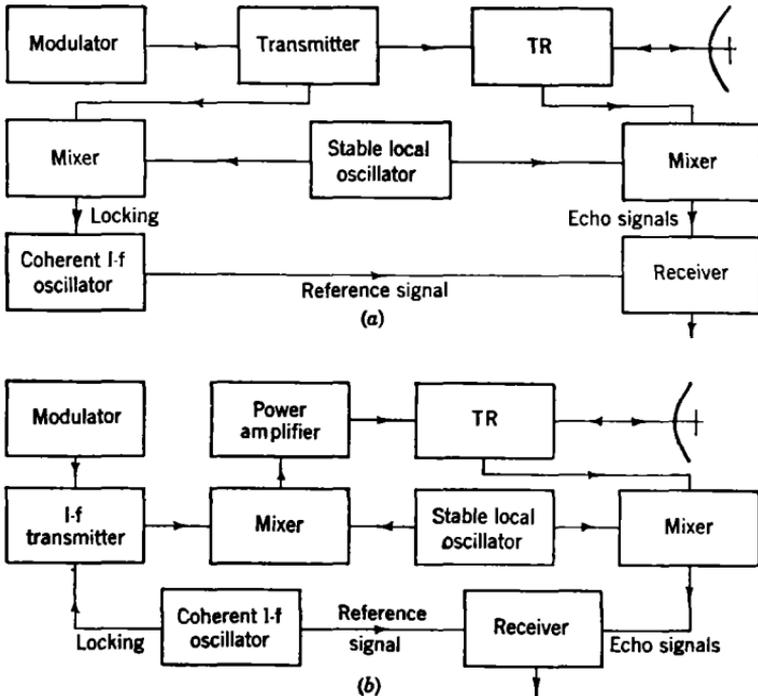


FIG. 16-13.—I-f locking, i-f addition. (a) I-f locking by transmitter, i-f addition. (b) I-f locking by oscillator, i-f addition.

It is also possible to lock at radio frequency and add at intermediate frequency. Further, the oscillator can be made to lock the phase of the transmitter instead of vice versa. The various arrangements can therefore be classified according to—

1. Whether the transmitter locks the oscillator or vice versa.
2. Whether the locking takes place at radio frequency or intermediate frequency.
3. Whether the signals are added at radio frequency or intermediate frequency.

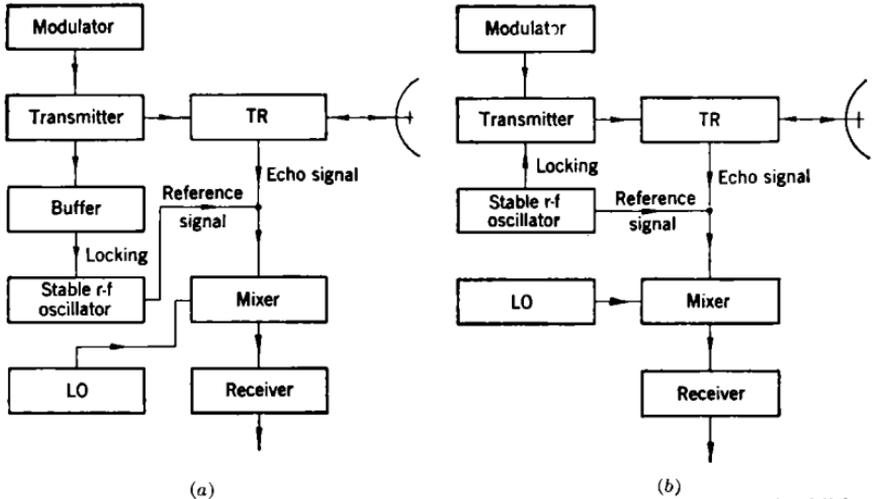


FIG. 16-14.—R-f locking, r-f addition. (a) R-f locking by transmitter, r-f addition. (b) R-f locking by oscillator, r-f addition.

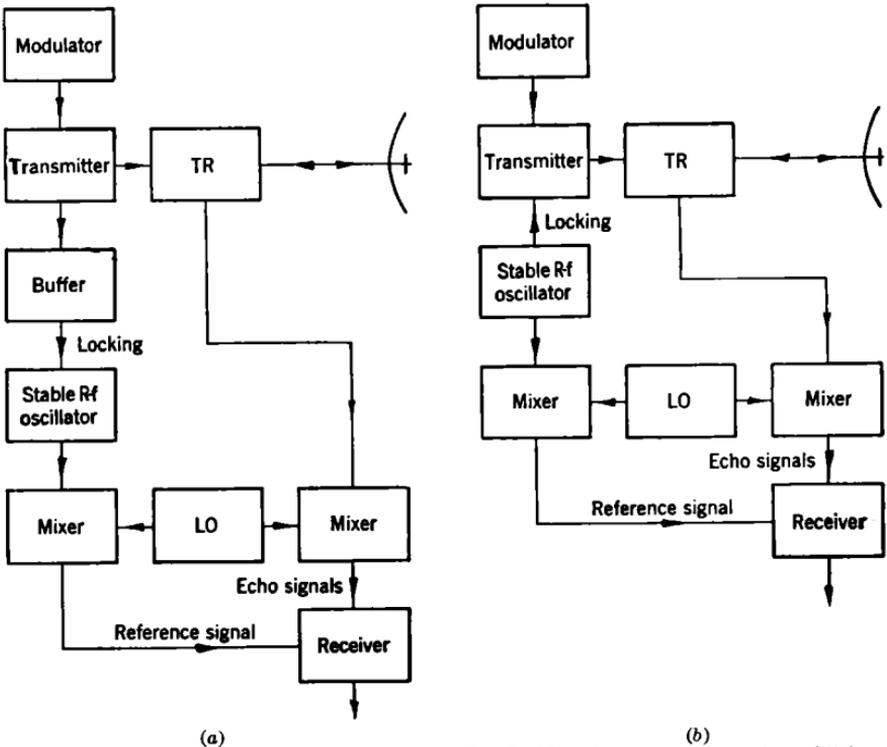
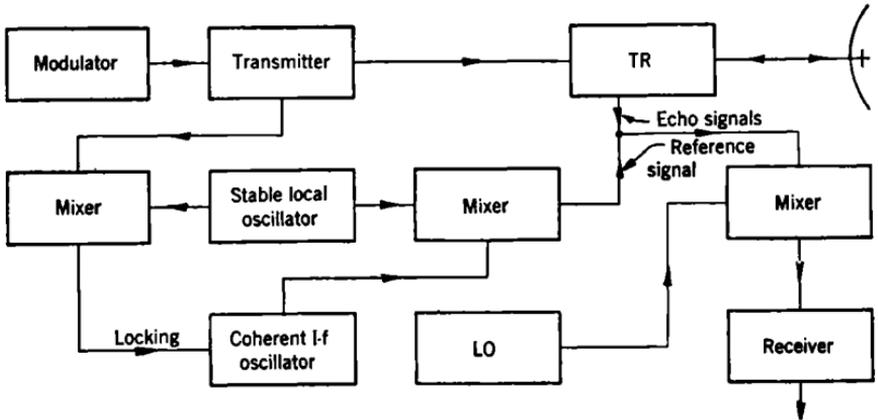


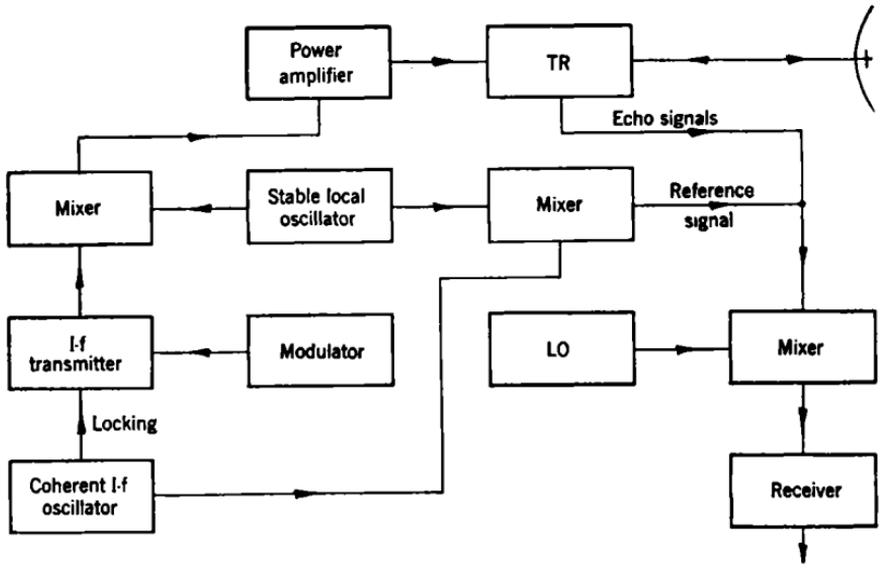
FIG. 16-15.—R-f locking, i-f addition. (a) R-f locking by transmitter, i-f addition. (b) R-f locking by oscillator, i-f addition.

Taking all possible combinations, we get 2^3 different types of circuit. These are shown in Figs. 16·13 to 16·16.

The circuit shown in Fig. 16·13a has advantages over the other schemes. In the first place, the signals are added at intermediate



(a)



(b)

FIG. 16-16.—I-f locking, r-f addition. (a) I-f locking by transmitter, r-f addition. (b) I-f locking by oscillator, r-f addition.

frequency, which allows amplification of the signals before addition. Second, i-f locking is superior to r-f locking. The latter is hard to do in the case of the transmitter, because of the large amount of power required;

it is also difficult if the stable r-f oscillator is to be locked, because of the high Q required to make the oscillator stable. Locking a coherent i-f oscillator is easy since the Q can be smaller by a factor of 100 for the same allowed rate of frequency change. The scheme of Fig. 16-13*b*, which uses i-f locking and i-f adding, is inconvenient because it requires a power amplifier.

PERFORMANCE CRITERIA AND CHOICE OF SYSTEM CONSTANTS

The primary objective of moving-target indication is, of course, to attain a high degree of discrimination in favor of the echoes from moving targets compared to those from fixed ones. In order that this be accomplished it is necessary that echoes, or clutter, from fixed targets be eliminated or greatly reduced and that those from moving targets be retained with optimum sensitivity compared to the residual clutter and the inherent noise of the receiver.

Since the recognition of moving targets is based upon changes in the returning echo from one pulse to the next, any changes present in the signals from the fixed echoes will interfere with their elimination. Such undesirable variations are of two sorts: (1) those inherently present in the returning echoes, principally due to internal motions of the targets themselves or to the effects of scanning; (2) spurious variations due to instabilities or other shortcomings of the radar equipment. The methods used and the constants chosen for the set should be selected in such a way as to minimize the effects of the inherent fluctuations. The elimination of spurious variations is largely a matter of careful engineering design.

Another factor to be considered in the choice of system parameters is that since different radial velocities produce different changes in relative phase from pulse to pulse, certain velocity intervals are much more effective than others in producing large uncanceled signals. Indeed, a phase change of one cycle—that is, a radial motion of one-half wavelength between pulses—is equivalent to no change at all. Thus there are certain “blind” velocity intervals in which the mass motion of the target does not lead to a discernible indication. Since the velocities about which these intervals are centered are determined by the wavelength and the pulse repetition rate of the set, it is possible, within limits, to arrange that they fall as little as possible in the range of velocities to be expected for the targets of interest.

The various factors dealing with clutter elimination and target visibility on a fixed system will be discussed in detail from both the theoretical and the practical standpoint in the next few sections. The additional factors present when the system is moving will be introduced later.

16-5. Stability Requirements. *Frequency of the Beating Oscillators.*—The amplitude of the video signal from a stationary target can be written

as

$$y = y_0 \cos \phi, \quad (1)$$

where y_0 is the i-f amplitude and ϕ is the phase difference between the i-f echo signal and the reference signal. Now ϕ depends, as we have already seen, only on the number of cycles executed by the stable local oscillator and the coherent oscillator during the echo-time t_1 . Thus

$$\phi = (\omega_l \pm \omega_c)t_1,$$

where ω_l and ω_c are the angular frequencies of the stable local oscillator and the coherent oscillator, respectively. The positive sign is to be taken when the local oscillator is tuned below the magnetron frequency, the negative sign when it is above. If either oscillator varies in frequency between pulses by an amount $\Delta\omega$, the phase change produced is then

$$\Delta\phi = t_1 \Delta\omega.$$

If the frequency drifts at a uniform rate, then

$$\Delta\omega = T \frac{d\omega}{dt},$$

where T is the repetition period. Thus

$$\Delta\phi = Tt_1 \frac{d\omega}{dt}. \quad (2)$$

Now, from Eq. (1), we obtain, for the fractional change in beat amplitude,

$$\frac{\Delta y}{y_0} = -\sin \phi \Delta\phi,$$

which has a maximum value equal to $\Delta\phi$. For a high-performance MTI system a value of 6 per cent for $\Delta y/y_0$ could be tolerated since this represents the maximum residue of clutter amplitude left after cancellation. The average residue will then be considerably less than 6 per cent. The corresponding value of $\Delta\phi$ is 0.06 radian or $\frac{1}{100}$ cycle. On substituting this value in Eq. (2) we get for the allowable rate of drift of frequency

$$\frac{df}{dt} = \frac{1}{100Tt_1}. \quad (3)$$

For example, this has the value 20 kc/sec² for $T = 1000 \mu\text{sec}$ (PRF of 1000) and $t_1 = 500 \mu\text{sec}$ (target at range of 50 miles). It will be seen later that such a low rate of drift cannot be obtained in a local oscillator without some special means of stabilization. In the case of the coherent oscillator, on the other hand, the figure is readily attainable with a well-constructed oscillator circuit.

So far we have assumed that the coherent oscillator is exactly in tune with the i-f echo signals. If this is not so, the video signals are no longer square pulses but contain a number of cycles of the beat frequency, as shown in Fig. 16-17. In the case of a moving target the beat cycles shift horizontally from pulse to pulse and give a filled-in appearance on the A-scope. Figure 16-17*b* really represents coherent i-f rather than coherent video signals, with the coherent oscillator acting as a second local oscillator. Thus a loss in signal-to-noise ratio occurs when the signals are passed through the video amplifier, whose bandpass characteristic is wrong for i-f signals. It is interesting to note that a gain in signal-to-noise ratio can be obtained theoretically by detuning the coherent

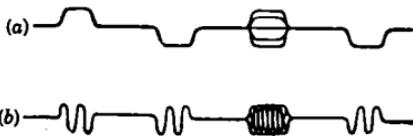


FIG. 16-17.—Effect of detuning the coherent oscillator. (a) Oscillator in tune with i-f signals. (b) Oscillator detuned.

oscillator sufficiently and replacing the video amplifier by an i-f amplifier having a suitable pass band.¹ However, since this scheme requires much greater precision in the cancellation circuits, it is probably better to retain the video amplifier and make sure that the coherent oscillator is always well tuned to the i-f signals. No serious loss in signal-to-noise will occur if there is less than one-quarter of a beat cycle in each pulse. This means, for example, that the detuning should be less than $\frac{1}{4}$ Mc/sec in the case of a 1- μ sec pulse.

Magnetron Frequency.—The frequency stability required of the magnetron can now be considered. The effect of a change Δf_m is to alter the number of beat cycles within the “video” pulse by the amount $\tau \Delta f_m$, where τ is the pulse length. For reasons similar to those used in connection with the beating oscillators, this quantity should not exceed $\frac{1}{100\tau}$ cycle from pulse to pulse. Hence we get for the allowed rate of drift, in the same way as before,

$$\frac{df_m}{dt} = \frac{1}{100T\tau}. \quad (4)$$

For $T = 1000 \mu\text{sec}$ and $\tau = 1 \mu\text{sec}$, the rate is 10 Mc/sec², which is 500 times the permissible rate for the beating oscillators. We conclude that, if it is found necessary to have AFC on a system with MTI, it should be applied to the magnetron rather than the local oscillator.

Variation in magnetron frequency from pulse to pulse causes another kind of effect in the case of extended ground clutter. Consider two clutter signals that just overlap in range. If the magnetron frequency changes by an amount Δf_m , there will be a relative phase shift between the two echoes which amounts to $\tau \Delta f_m$ cycles. This phase shift will cause the

¹ A. G. Emslie, “MTI Using Coherent IF,” RL Group Report No. 104—8/22/45.

echoes to beat in amplitude in the region of overlap, even when no coherent reference signal is present. Since the expression for the phase shift is identical with the previous expression, Eq. (4) also applies here.

Extra cycles will also appear in the video pulses if the magnetron frequency varies during the transmitted pulse. Such variation occurs chiefly at the beginning and end of the pulse. If Δf_1 is the maximum departure in frequency from the value in the central part of the pulse, the number of beat cycles is certainly less than $\tau \Delta f_1$. With the same criterion as before we therefore should have

$$\tau \Delta f_1 < \frac{1}{4}. \quad (5)$$

For a 1- μ sec pulse this gives $\Delta f_1 < \frac{1}{4}$ Mc/sec. However, since most of the effect is concentrated near the ends of the pulse, a much larger variation is probably permissible. Since the frequency pattern within the pulse does not change from pulse to pulse, the quality of the cancellation is not affected. All that is involved is a slight loss in signal-to-noise ratio for moving targets.

Cancellation Equipment.—Consider next the stability requirements on the cancellation equipment. The detectability of a signal on the PPI can be roughly measured by the area under the voltage pulse. Thus if the trigger and signal delay lines differ in delay time by an amount $\Delta\tau_1$, there will be two uncanceled spikes of width $\Delta\tau_1$ for each echo. In order to cancel, for example, to 4 per cent, we must therefore make $\Delta\tau_1$ less than 2 per cent of the pulse length. The delay lines should be capable of matching each other within this tolerance for at least an hour at a time. Likewise the amplitudes of the signals in the delayed and undelayed channels should match to 4 per cent for an hour at a time.

Modulator.—Pulse-to-pulse variation in the repetition rate can produce the same effect as unmatched delay lines. Thus the modulator should fire relative to the trigger with a variation of not more than 2 per cent of the pulse length, or $\frac{1}{250} \mu$ sec in the case of a 1- μ sec pulse.

Finally, we have to consider the effect of variation in pulse length. If this should change by an amount $\Delta\tau$ from one pulse to the next, there will be an uncanceled spike of width $\Delta\tau$. Thus we should not allow $\Delta\tau/\tau$ to exceed 4 per cent.

Variation in pulse length can also cause an indirect effect on the phase of the coherent oscillator when the latter is not in tune with the i-f echo signals. Let us assume, for example, that the locking pulse causes the coherent oscillator to execute forced vibrations at the intermediate frequency for the duration of the pulse, after which it reverts to its natural frequency. Then the variation in phase of the coherent oscillator is given by $\Delta\tau \cdot \Delta f$ (in cycles), where Δf is the amount by which the coherent

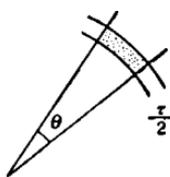
oscillator is off-tune. Taking $\Delta\tau = 0.04 \mu\text{sec}$ (for $\tau = 1 \mu\text{sec}$) and allowing a phase change of $\frac{1}{100}$ cycle, we get $\Delta f = \frac{1}{4} \text{Mc/sec}$.

Summary of Requirements.—Table 16.1 summarizes the stability requirements. It should be mentioned that these figures represent almost ideal conditions, in which the clutter to be canceled shows no fluctuations of its own. If for any reason (such as rapid scanning rate) the clutter should fluctuate by 10 or 15 per cent from pulse to pulse, there would be no sense in canceling to 3 or 4 per cent. In such a case, some of the requirements in the table could be relaxed considerably.

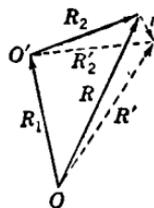
TABLE 16-1.—STABILITY REQUIREMENTS FOR HIGH-PERFORMANCE SYSTEM WITH PRF OF 1000 FPS, PULSE LENGTH $1 \mu\text{SEC}$, GROUND CLUTTER OUT TO 50 MILES

Component	Quantity	Maximum allowable value
Stable local oscillator.....	Frequency drift	20 kc/sec ²
Coherent oscillator.....	Frequency drift	20 kc/sec ²
Magnetron.....	Frequency drift	10 Mc/sec ²
Magnetron.....	Frequency change within pulse	$\frac{1}{4}$ Mc/sec ² (and probably much more)
Modulator and magnetron..	Pulse length variation from pulse to pulse	4 %
Coherent oscillator.....	Detuning from intermediate frequency	$\frac{1}{4}$ Mc/sec
Modulator.....	Stability relative to trigger	$\frac{1}{80} \mu\text{sec}$
Delay lines.....	Relative drift in delay time	$\frac{1}{80} \mu\text{sec/hr}$

16.6. Internal Clutter Fluctuations.—In this section we shall consider echo fluctuations due to internal motions of the clutter—for example, the



(a)



(b)

FIG. 16-18.—Composite nature of ground clutter. (a) Illuminated area of clutter. (b) Contributions to signal: fixed clutter = R_1 ; moving clutter = R_2 ; resultant signal = R ; resultant for next pulse = R' ; variation between pulses = r .

motions of trees in the case of ground clutter. The ground-clutter pattern may include strong echoes from single targets, especially in regions where there are large structures having simple geometrical shapes. Echoes from water towers and building faces may easily equal in intensity the composite echo obtained from a mountainside. However, echoes from structures are generally found to be fairly steady since the targets do not sway much in the wind

and do not present a serious cancellation problem when the antenna is stationary.

Most ground clutter is composite in the sense that the echo amplitude at a given instant is the vector sum of many small echoes from the individual targets scattered over a land area determined by the beamwidth

and half the pulse length, as shown in Fig. 16-18a. The individual targets may consist of rocks, tree trunks, tree branches, leaves, etc. Some of these, such as the rocks, are fixed. Others, such as the branches and leaves, move in the wind. In Fig. 16-18b, the vector R_1 represents the contribution from the fixed targets and R_2 that from the moving targets, and R is the resultant signal. At the next pulse, R_2 changes to R'_2 , and the resultant to R' . The small vector r is the pulse-to-pulse change.

Individual moving targets in the illuminated area are so numerous that even the largest is small in size compared with their sum. Also the phases of their echo signals are completely independent. Under these two conditions, the vector R_2 has a probability distribution like that of statistical noise—that is, the end of the vector has a Gaussian distribution about the point O' . The pulse-to-pulse change r likewise has a Gaussian distribution.

The final result connecting r and R_2 is the following:

$$(r)_{\text{rms}} = \frac{k}{\lambda f_r} (R_2)_{\text{rms}}, \quad (6)$$

where f_r is the PRF, λ the wavelength, and k a factor that depends on the wind speed and type of terrain. Measurements of the ratio of $(r)_{\text{rms}}$ to $(R)_{\text{rms}}$ have been made by H. Goldstein.¹ The following table gives typical values extrapolated from the experimental figures which were obtained using a repetition rate of 333/sec. The voltage ratio of fixed to variable clutter is also given. In the case of sea and storm echoes, there is no fixed component.

TABLE 16-2.—TYPICAL MEASURED VALUES OF CLUTTER FLUCTUATION

Kind of clutter	λ , cm	Voltage ratio of rms fixed component to rms variable com- ponent	Total echo/rms fluctuation, db		
			PRF = 500	1000	2000
Thunderstorm.....	9.2	0	5	11	17
Sea echo.....	9.2	0	13	19	25
Sea echo.....	3.2	0	8	14	20
Wooded terrain (wind 45 mph).....	3.2	0	14	20	26
Wooded terrain (wind 25 mph).....	3.2	$\ll 1$.	22	28	34
Wooded terrain (wind 25 mph).....	9.2	0.9	34	39	46
Wooded terrain (wind 10 mph).....	9.2	3.7	51	54	57

¹ Secs. 6-13 to 6-21 of Vol. 13 of this series.

Several remarks are in order regarding the generality of the above experimental values. Both the amplitude and the velocity of sea clutter can be expected to vary widely with weather. Storm echoes are variable in character, depending perhaps upon their internal turbulence. A reduction of storm echoes by 10 db or more has been frequently observed on a 10-cm MTI system at 2000 pps. On some occasions, however, the reduction of echoes from certain areas of the storm has been negligible despite a tangential wind velocity.

The wooded terrain is that of New England in summer. Limited evidence suggests that there is little change with the seasons. However, some questions of importance remain with regard to ground clutter. On the PPI of a 10-cm MTI system operating near Boston at a PRF of 300, many thousands of uncanceled targets have been observed under conditions that make an explanation difficult. These targets, which might possibly be birds, have been noticed to increase in number just after sunset.

When extrapolating to other wavelengths, it should be borne in mind that the wavelength dependence of Eq. (6) cannot be expected to hold rigorously. Classes of moving reflectors that are negligible at 10 cm may become of importance at 3 cm.

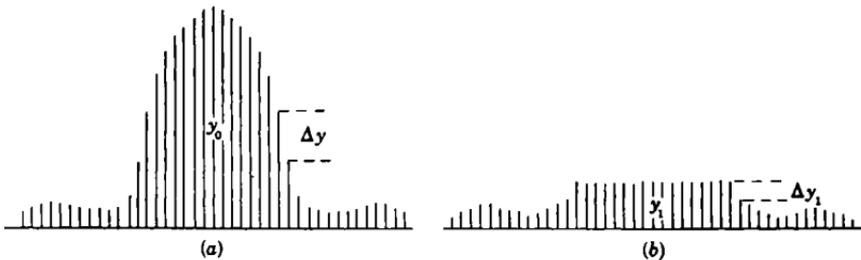


FIG. 16-19.—Scanning fluctuations of the echo from an isolated target. (a) Linear receiver. (b) Limiting receiver.

16-7. Fluctuations Due to Scanning.—When the radar beam sweeps over a simple isolated target, the received echoes vary from pulse to pulse as shown in Fig. 16-19a. The envelope of the echoes is the antenna pattern (voltage two ways), and the maximum pulse-to-pulse variation Δy occurs at the point where the antenna pattern has its maximum slope. Thus Δy can be readily calculated if a Gaussian error curve is used as an approximation to the actual antenna curve. The result can be written in the form

$$\frac{\Delta y}{y_0} = \frac{1.43}{n}, \quad (7)$$

where y_0 is the maximum received voltage and n is the number of pulses transmitted while the antenna rotates through an angle equal to the

beamwidth (as measured between half-power points, one way). For example, in the case of a 1° beam rotating at 6 rpm with a PRF of 1000, n is 28 and Δy is therefore 5 per cent of y_0 . This represents the residue that would be left after cancellation if a linear receiver were used.

In the next section it will be shown that a nonlinear receiver is necessary in order to remove all the clutter from the PPI. Figure 16-19b shows the output of a simple limiting receiver; the maximum pulse-to-pulse variation Δy_1 is less than before and occurs farther out on the antenna pattern. The fluctuation is now given by the expression

$$\frac{\Delta y_1}{y_1} = \frac{k}{n}, \quad (8)$$

where k is no longer a constant but depends on the ratio of maximum input signal y_0 to the limit level y_1 , as shown in Table 16-3.

TABLE 16-3

$\frac{y_0}{y_1}$	k
10	5.1
100	7.2
1000	8.7

It will be noticed that k does not change much as the input signal strength is varied over a wide range. This makes it easy to obtain removal of clutter of varying size.

Now let us consider the fluctuations in the case of extended ground clutter. The echo from such clutter, as we have already seen in Sec. 16-6, consists of the vector sum of the contributions of a large number of scattering elements. As the beam sweeps over the ground, new elements are illuminated and the old ones pass out of the beam. Thus the signal fluctuates in both amplitude and phase, as shown in Fig. 16-20, where R is the signal voltage at a given instant and R' the value after one repetition period. The change from pulse to pulse is represented by the vector r . If it is assumed that the scattering elements are equal in size and randomly distributed, it is not hard¹ to calculate the ratio of the rms values of R and r . The result is

$$\frac{r_0}{R_0} = \frac{\Theta}{n} \left\{ \frac{\int [f'(\theta)]^2 d\theta}{\int [f(\theta)]^2 d\theta} \right\}^{1/2}, \quad (9)$$

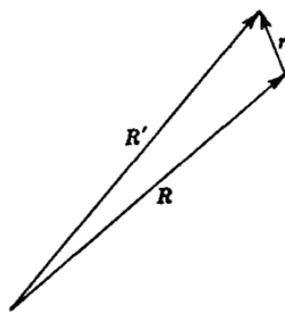


FIG. 16-20.—Scanning fluctuation of the echo from an extended target.

¹ A. G. Emslie, "Moving Target Indication on MEW," RL Report No. 1080, Feb. 19, 1946.

where r_0 and R_0 are the rms values of r and R , θ is the beamwidth, n is the number of pulses per beamwidth, and $f(\theta)$ is the antenna pattern (voltage two ways). For the case of a Gaussian antenna pattern, we obtain the equation

$$\frac{r_0}{R_0} = \frac{1.66}{n}. \quad (10)$$

It is convenient to analyze the fluctuation into an amplitude part and a phase part. To do this we can resolve the vector r into two components, one in the direction of R , the other at right angles to R . The rms value of each of these components is $r_0/\sqrt{2}$. Thus the rms fluctuation in amplitude is $r_0/\sqrt{2}$ and the rms fluctuation in phase is approximately $r_0/(R_0\sqrt{2})$. We can therefore write

$$\begin{aligned} \text{Rms pulse-to-pulse amplitude fluctuation} &= \frac{1.66}{n\sqrt{2}} R_0 \\ \text{Rms pulse-to-pulse phase fluctuation} &= \frac{1.66}{n\sqrt{2}} \text{ radians.} \end{aligned} \quad (11)$$

These equations apply to the case of uniformly distributed scattering elements. It might appear that a sudden discontinuity in the density of scattering elements would cause a larger fluctuation. To see that this is not so we can consider the integral $\{\int [f'(\theta)]^2 d\theta\}^{1/2}$, which is proportional to the rms pulse-to-pulse fluctuation r_0 . In the case of a uniform distribution of scattering elements, the limits of integration are from $-\infty$ to $+\infty$. When there is a discontinuity at an angle θ_1 from the center of the beam, the limits are θ_1 to ∞ , and the integral will be less than before since the integrand is everywhere positive.

16-8. Receiver Characteristics.—Ideally, an MTI system would remove all clutter signals from the scope, maintaining the greatest possible sensitivity for moving targets both in the clear and in the clutter. Practically, this is not always possible, because of clutter fluctuations, the finite perfection of cancellation, etc. Under these circumstances, it appears desirable to adjust the residual clutter, after cancellation, so that it resembles receiver noise in amplitude and texture when it is presented on the scope.

Clutter fluctuations already have an amplitude distribution like that of noise, whether they are due to scanning or to the action of the wind. Thus the only problem left is to reduce the rms fluctuation to the same value everywhere. There are two ways of doing this, the first depending on the nature of the amplitude fluctuations of the clutter, the second on the nature of the phase fluctuations.

As we have seen, the rms fluctuation in amplitude is proportional to the rms clutter amplitude. Thus if x is the size of the clutter at the

input of the receiver, the total increment in amplitude can be written as

$$\Delta x = kx + N,$$

where kx is the amplitude component of the clutter fluctuation and N is the component of the i-f noise vector in the direction of x . The output fluctuation Δy is given by

$$\Delta y = \frac{dy}{dx} \Delta x,$$

and we require it to be independent of the clutter amplitude x . Thus we have

$$\frac{dy}{dx} = \frac{K}{kx + N},$$

where K is the (constant) value of Δy . Integrating this equation with the condition that $y = 0$ when $x = 0$, we get the following expression for the characteristic of the i-f amplifier:

$$y = \frac{K}{k} \ln \left(1 + \frac{kx}{N} \right). \quad (12)$$

For values of x giving a fluctuation kx considerably less than noise, the characteristic is linear. For large values of x it is logarithmic. A receiver with this characteristic is called a lin-log receiver.

The output of the lin-log i-f amplifier has to be added to the reference signal and then detected. Since the clutter varies in phase as well as amplitude, its phase variations are converted by the detector, when the reference signal is present, into amplitude variations, which unfortunately are not independent of the size of the clutter. It can be shown, however, that no serious harm results if the reference signal level is kept somewhat smaller than the maximum amplitude of the linear portion of the receiver characteristic.

The second method of obtaining uniform output fluctuations depends on the fact that the rms phase variation of the clutter is independent of the size of the clutter. Thus all that is required is a limiting i-f amplifier to remove the amplitude variations, followed by a phase-sensitive detector to convert the phase variations into uniform amplitude variations. To obtain the maximum sensitivity for targets moving in the clutter it is necessary that the phase-sensitive detector should have a linear characteristic—that is, the output amplitude variation should be proportional to the input phase variation. This can be achieved by using a balanced detector for mixing the reference signal with the output of the limiting i-f amplifier.

The relative advantages of the two types of receiver will now be considered. If the lin-log receiver is used, it is possible to dispense with

the oscillator and amplifier shown in Fig. 16-11 and to send the i-f signal, together with the reference signal, directly through the delay line. The amplifier then works at the intermediate frequency used in the receiver. These changes make the circuit simpler and avoid the problems of carrier modulation and demodulation. However, the loss of freedom in the choice of frequency for the delay line and the comparison amplifier is a disadvantage. Also, the method is not suitable for "back-of-dish" radar sets, since the i-f signals cannot conveniently be brought out through slip rings. This method of using a lin-log receiver should be used only when simplicity and compactness are of prime importance. In general, such a receiver should be used in conjunction with an oscillator and amplifier (Fig. 16-11) in the usual way.

An important advantage of the limiting receiver is that the output video signals have a range of amplitude extending from noise only up to the limit level. This small dynamic range makes it easy to design the cancellation circuits. For the lin-log receiver, the dynamic range is 6 to 12 db greater under conditions of equal performance.

For general MTI use, the limiting type of receiver is to be preferred. However, there is a type of MTI system in which moving targets are detected in the clutter by the fluctuations which they produce in the clutter amplitude. The lin-log receiver must be used in such a system, since the limiting receiver cannot detect amplitude changes.

Some idea of the magnitudes involved in the receiver problem can be obtained from Table 16-4. The fluctuations due to the wind were

TABLE 16-4.—FLUCTUATIONS DUE TO WIND AND SCANNING

λ , cm	PRF	Beam- width, degrees	Scanning rate, rpm	Pulses per beam- width	Wind effect $\frac{\text{rms fluctuation}}{\text{rms total echo}}$, db	Scanning effect $\frac{\text{rms fluctuation}}{\text{rms total echo}}$, db
9.2	500	1	4	21	-34	-22
9.2	2000	3	4	250	-46	-44

obtained from Table 16-2 for wooded terrain and a wind velocity of 25 mph. The scanning effect was derived from the formula for extended clutter in Sec. 16-7. It will be noticed that the scanning effect predominates in the case of the low PRF and narrow beam, whereas the two effects are about equal for the high PRF and wide beam. The fluctuations in the first case are roughly 10 per cent, considerably greater than the fluctuations due to instability of the components. On the other hand, the fluctuations in the second case are only about 1 per cent and are therefore negligible compared with system instability.

Consider how a limiting receiver should be adjusted in each of these cases. Since only the phase part of the fluctuation is selected, the fluctuation of the video output is $1/\sqrt{2}$ or 3 db less than the total input fluctuation. Thus the figure -22 db becomes -25 db at the output, which means that the rms fluctuation in the first case is about $\frac{1}{18}$ of the limit level. Therefore, in order to obtain a PPI with uniform background, the gain control of the receiver, ahead of the limiting stages, should be adjusted so that rms noise is also $\frac{1}{18}$ of limit level. "Peak" noise is then about $\frac{1}{6}$ of limit level. This adjustment of noise relative to limit level is made while viewing the output of the limiting receiver on an A-scope. When this has been done, the video gain control is adjusted so that the noise and canceled residue just show on the PPI. The MTI photograph in Fig. 16-1 shows everything in correct adjustment so that the residue of the clutter blends with the noise. The corresponding photograph in Fig. 16-2 shows a case of incorrect adjustment: receiver noise is too large compared with the limit level. The result is that the residue does not show on the scope and black "holes" appear where the clutter was. Under these conditions there is an excessive loss of sensitivity for moving targets in the clutter.

In the second case (high PRF and broad beam) the predominating fluctuations will be due to system instability and may be of the order of 5 per cent (peak) if all the stability requirements are met. Thus peak receiver noise should in this case be set at 5 per cent of limit level.

16-9. Target Visibility.—There are two problems to be discussed in this section. The first has to do with the visibility of moving targets when they are clear of the clutter; it includes consideration of undesirable targets, like clouds, as well as desirable targets such as aircraft. The second problem is concerned with the visibility of moving targets that occur at the same range and azimuth as ground clutter.

Targets in the Clear.—We have seen in Sec. 16-2 that the video pulses from the MTI receiver, in the case of a moving target, are amplitude-modulated at the doppler frequency f_d . Thus the pulse amplitude at a given instant t is given by

$$y_1 = y_0 \cos 2\pi f_d t.$$

For the next pulse,

$$y_2 = y_0 \cos 2\pi f_d (t + T).$$

The canceled video pulse is therefore

$$y = y_1 - y_2 = 2y_0 \sin(\pi f_d T) \sin 2\pi f_d \left(t + \frac{T}{2} \right). \quad (13)$$

Thus the canceled pulses also exhibit the doppler modulation frequency and have an amplitude

$$y'_0 = 2y_0 \sin(\pi f_d T). \quad (14)$$

This expression is zero when f_d is a multiple of the PRF and, since $f_d = 2v/\lambda$, the first "blind" speed is given by

$$f_d T = \frac{2v_1 T}{\lambda} = 1,$$

or

$$v_1 = \frac{\lambda}{2T} = \frac{\lambda f_r}{2}, \quad (15)$$

where f_r is the PRF. In other words, the target appears to be at rest when it travels a distance $\lambda/2$ (or a multiple of $\lambda/2$) between pulses. Numerically,

$$\text{First blind speed} = \lambda f_r / 89 \quad (\lambda \text{ in cm, speed in mph}) \quad (16)$$

Going back to Eq. (13) we see that the average canceled signal, after rectification, is given by

$$|\bar{y}| = \frac{4}{\pi} y_0 |\sin(\pi f_d T)|. \quad (17)$$

We have to compare this amplitude with that of noise after cancellation. The noise amplitude in the delayed channel is completely independent of that in the undelayed channel. We must therefore add the noise powers in the two channels, obtaining an increase of $\sqrt{2}$ in noise amplitude after cancellation. Hence the change in signal-to-noise ratio caused by the addition of MTI is represented by the factor

$$\frac{4}{\pi \sqrt{2}} |\sin(\pi f_d T)|. \quad (18)$$

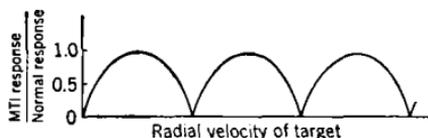


FIG. 16-21.—Response curve for target in the clear. The first blind speed has the value $\lambda f_r / 89$ (λ in cm, speed in mph).

The numerical factor is about 0.9. Thus MTI causes scarcely any loss for a target moving at the optimum speed but does cause a loss at other speeds. Figure 16-21 shows the voltage response for MTI relative to the response for a normal radar set.

It will be observed that the response at small speeds is proportional to the speed v . It can be written as follows:

$$\frac{\text{Response}}{\text{Max. response}} \approx \pi f_d T = \pi \frac{v}{v_1}, \quad (20)$$

where v_1 is the first blind speed. As an example, consider a set with $\lambda = 9.2$ cm and $f_r = 2000$ pps. Then from Eq. (16) we have $v_1 = 206$ mph. For a storm cloud traveling radially at 30 mph, the ratio in Eq. (20) becomes $1/2.2$ or -7 db, whereas Table 16-2 shows that the fluctuation due to internal motion of the cloud is only -17 db. From the MTI point of

view translational motion is more serious than the internal motion of the cloud. It should be noticed that the two effects change together as the PRF is varied.

A question of some interest is the following: If a target of given size can have any radial speed with equal probability, how often will it be detected? Let us suppose that the target, when moving at one of the optimum speeds in Fig. 16.21, gives a signal N times as large as the minimum detectable signal. Then the relative response at other speeds is given by

$$N \sin (\pi f_a T).$$

The target will be seen as long as this quantity is greater than unity; the probability of this being the case is

$$\frac{\frac{\pi}{2} - \sin^{-1} \frac{1}{N}}{\frac{\pi}{2}} \tag{21}$$

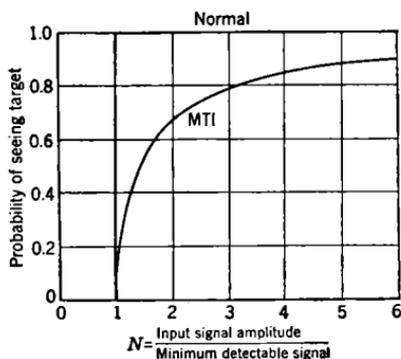


FIG. 16-22.—Probability of seeing a moving target.

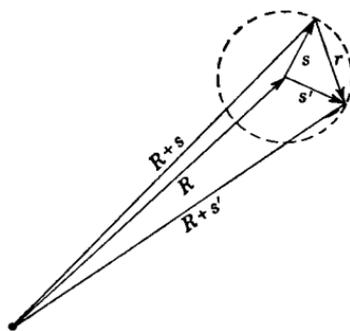


FIG. 16-23.—Moving target in clutter: clutter = R ; target = s ; target at next pulse = s' ; variation between pulses = r .

A plot of this expression is shown in Fig. 16-22. The probability is 0.5 when $N = \sqrt{2}$; consequently in this sense we can say that MTI causes an average loss of $\sqrt{2}$ in voltage or 3 db.¹

Targets in the Clutter.—We have seen in Secs. 16-6 and 16-7 that the fluctuations of extended clutter, whether due to the wind or to scanning, have noiselike character. This “clutter-noise” is what appears on the MTI scope at places normally occupied by the clutter, and it is adjusted to be equal in brightness to ordinary noise. A moving target in the

¹ The above discussion neglects the fact that the phase-sensitive detector reduces the noise voltage by $\sqrt{2}$. Thus Eq. (18) should be multiplied by $\sqrt{2}$, as should the ordinate of Fig. 16.21. The abscissa of the MTI curve of Fig. 16.22 should be divided by $\sqrt{2}$. Use of MTI may thus entail no average loss.

clutter has to compete with the clutter-noise in much the same way as a target in the clear competes with ordinary receiver noise. Figure 16-23 is a vector diagram for a moving target in the clutter. The pulse-to-pulse variation r in the resultant vector has both amplitude and phase components just like the clutter fluctuations. The magnitude of r can easily be obtained from the fact that s rotates with the doppler frequency f_d . Thus the angle between s and s' is $2\pi f_d T$, and accordingly

$$r = 2s \sin(\pi f_d T). \quad (22)$$

This equation, which is the same as Eq. (14), shows that blind speeds also exist for targets in the clutter. If we take the arithmetical average of Eq. (22) for all speeds, we get

$$\bar{r} = \frac{4s}{\pi}. \quad (23a)$$

If, on the other hand, we take the median with respect to velocity, which is the average that we took in the case of targets in the clear, we get

$$\bar{r} = \sqrt{2}s. \quad (23b)$$

Since the two values of \bar{r} differ by only 10 per cent or about 1 db, it does not matter which we take. The value given by Eq. (23a) is also the arithmetical average for the case of a noncoherent signal from a signal generator. Measurements can therefore be made without appreciable error by using the signal from a signal generator instead of that due to an actual target.

A moving-target or signal-generator pip can be expected to be visible on the PPI if \bar{r} is at least three times the rms clutter fluctuation r_0 . Using Eq. (23a), this condition can therefore be written as

$$s = \frac{3\pi}{4} r_0 = 2.4r_0. \quad (24)$$

That is, the signal amplitude should be 8 db larger than the rms clutter fluctuation.

It is convenient to express the visibility condition in terms of the rms clutter amplitude R_0 instead of the rms fluctuation. In the case of scanning effects this leads, with the help of Eq. (10), to the following simple relationship:

$$\frac{R_0}{s} = \frac{n}{4} \quad (\text{scanning}), \quad (25)$$

where n is the number of pulses per beamwidth. In the case of wind effects, we have, from Eq. (6),

$$\frac{R_0}{s} \propto \lambda f_r \quad (\text{wind}). \quad (26)$$

As is pointed out at the end of Sec. 16-6, the dependence on λ is only approximate.

The ratio R_0/s , which can be called the "subclutter visibility," gives a measure of the performance of an MTI system. When scanning and wind effects are present together, along with the effect of system instability, we have to go back to Eq. (24) and write it as

$$s = 2.4(r_{\text{scanning}}^2 + r_{\text{wind}}^2 + r_{\text{system}}^2)^{1/2}, \quad (27)$$

since the three fluctuations are independent of each other. Now each fluctuation is proportional to the size of the clutter. Thus we have

$$\frac{R_0}{s} = \frac{1}{\left(\frac{1}{K_1^2 n^2} + \frac{1}{K_2^2 \lambda^2 f_r^2} + \frac{1}{K_3^2} \right)^{1/2}}, \quad (28)$$

where the terms in the denominator refer respectively to the effects of scanning, wind, and system instability. The quantities K_1 and K_2 are constants, but K_3 may depend somewhat on λ and f_r . For example, it is probably harder to stabilize the local oscillator at 3 cm than at 10 cm. Again, Eq. (3) shows that the repetition rate enters into the stability condition. However, we shall assume, for simplicity, that the *over-all* system instability is independent of λ and f_r .

16-10. Choice of System Constants.—The following are the objectives in the design of a ground radar set with MTI:

1. Maximum sensitivity to moving targets in the clutter.
2. Maximum sensitivity to moving targets in the clear.
3. Maximum cancellation of undesirable moving objects, such as clouds.

Some of these objectives are in direct conflict with the normal aims of a radar set. For example, a broad beam and slow rate of scan will reduce scanning fluctuations and therefore help the first objective, but will decrease resolution and hamper the rate of flow of information. Because of this, MTI should be integrated into the radar set from the start and not added as an afterthought.

Subclutter Visibility.—From Eq. (28) we see that the subclutter visibility can be increased by increasing n and λf_r , but that nothing much is gained by going greatly beyond the point where the scanning and wind fluctuations are equal to those due to system instability. A well-designed system might have an rms instability of 2 per cent, in which case, from Eq. (24), the greatest possible subclutter visibility is 26 db. There is then no need to make the individual subclutter visibilities for scanning and wind greater than 30 db. For the case of scanning, that means, from Eq. (25), that n does not have to be larger than 120 pulses per beamwidth. Again, from Eq. (24), the rms wind fluctuation should be less than 22 db. This happens to be the value for wooded terrain when $\lambda = 3.2$ cm and

$f_r = 500$ and the wind velocity is 25 mph, as is shown in Table 16-2. Thus, λf_r should be about 1600 for this particular wind speed. In general, it is easy to obtain the required figure for λf_r , but not for n , without sacrifice elsewhere. Scanning fluctuations are the principal difficulty in MTI design.

The quantity n is proportional to beamwidth and PRF, and inversely proportional to scanning rate. It has the value 120 for a PRF of 720 with a 1° beam rotating at 1 rpm. Although this set of constants satisfies the condition above, the scanning rate is too slow for many purposes. To attain greater rate without losing subclutter visibility, either the beamwidth or PRF must be raised; the alternative is to take a loss in subclutter

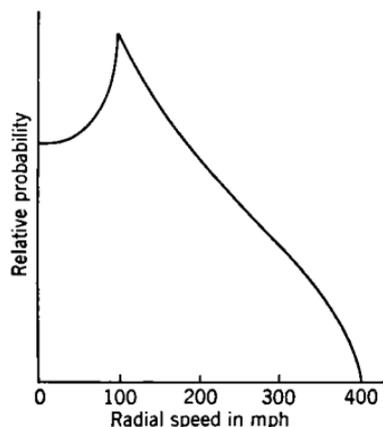


FIG. 16-24.—Distribution of radial speeds for random aircraft, assuming all ground speeds equally likely in range 100 to 400 mph and all directions equally likely.

visibility, which may be the best solution in some cases. For example, if the targets are all larger than the clutter, a subclutter visibility of zero db is sufficient. In that case the scanning rate can be raised to 30 rpm. It is, therefore, of considerable importance to use all possible methods of decreasing the size of the clutter relative to the moving targets, such as shaping the beam in elevation, using short pulses, and choosing the site carefully.

Blind Speeds.—It has already been seen that MTI causes an “average” loss of about 3 db when all radial speeds are equally probable.¹ This is true no matter where the blind speeds happen to fall. We are interested, however, in

targets such as airplanes and clouds which have non-uniform distributions of radial speeds, and would like to know the best values of the blind speeds in these cases. Figure 16-24 shows the distribution of radial velocities for airplanes whose ground speed can have any value between 100 and 400 mph with equal probability and whose direction of flight is random. It is hard to choose the best blind speeds from this graph. There will probably be an average loss¹ of about 3 db, as in the case of a uniform distribution, whether there are many blind speeds in the interval 0 to 400 mph or whether the first optimum speed comes at, for example, 400 mph. The choice of blind speeds is of no particular consequence, *per se*; but there may be special cases, such as airport traffic control, where the radial speed can be regulated. Blind speeds can then be chosen accordingly.

¹See footnote on p. 651.

To get the greatest cancellation of an undesirable moving object, the first blind speed, v_1 , should be made as large as possible. For example, if v_1 is 30 times the radial speed of the object the response will be 10 per cent of the maximum value, by Eq. (20). In the case of clouds moving at 30 mph, that gives us $v_1 = 900$ mph, which is a suitable value as far as Fig. 16-24 is concerned. Since $v_1 = \lambda f_r / 89$, it can be seen that such a high value of v_1 is obtainable, when a normal PRF is used, only by going to a wavelength of the order of 100 cm. For microwave frequencies, therefore, one can scarcely reduce storm echoes by more than about 10 db.

In the detection of ships, whose velocities are comparable with those of clouds, it might be thought that the ships would be canceled along with the clouds. This is not so because ships, when within the horizon, give much larger signals than clouds, and still show up strongly after a 10-db cancellation.

MOVING-TARGET INDICATION ON A MOVING SYSTEM

16-11. Compensation for Velocity of System.—Figure 16-25 shows what has to be done to compensate for the velocity of the station when a system is carried on a ship or airplane. The phase-shift unit shown in *a* changes the phase of the reference signal at the same rate as that at which the phase of fixed echoes is being changed by the motion of the station. The effect is to give the station a virtual velocity which cancels the actual velocity.

Figure 16-25*b* is a block diagram of the phase-shift unit. An oscillator supplies a signal at the doppler frequency which is mixed with the reference signal. The upper sideband is then selected by a crystal filter. The addition of the doppler frequency to the reference signal is equivalent to shifting the phase of the coherent oscillator at the rate of f_d cps. It is to be noted that this is not the same as merely tuning the coherent oscillator to the frequency $f_c + f_d$, since the phase shifter is applied *after* the oscillator is locked. Full details on the phase-shift unit can be found in a report by V. A. Olson.¹

The doppler frequency depends on the radial component of velocity and is therefore proportional to the cosine of the azimuth angle. It is necessary to vary the frequency of the doppler oscillator automatically as the antenna scans.

In an airborne set, there is a further complication due to the fact that the radial velocity depends on the depression angle, which means that the doppler frequency of the ground clutter varies with the range. In consequence, the phase-shift unit cannot be used for clutter at large

¹ V. A. Olson, "A Moving Coho Conversion Unit," RL Report No. 975, Apr. 3, 1946.

depression angles; some other method must be used, such as the noncoherent method now to be described.

16-12. The Noncoherent Method.—We have already seen in Fig. 16-19 that a target moving in the clutter produces a resultant echo that varies in both amplitude and phase. To detect the amplitude variations, all that is required is a receiver that does not limit—that is, one with a linear-logarithmic response. Since no coherent reference oscillator is used, this is called the noncoherent method of detecting moving targets. The ground clutter itself acts as the reference signal.

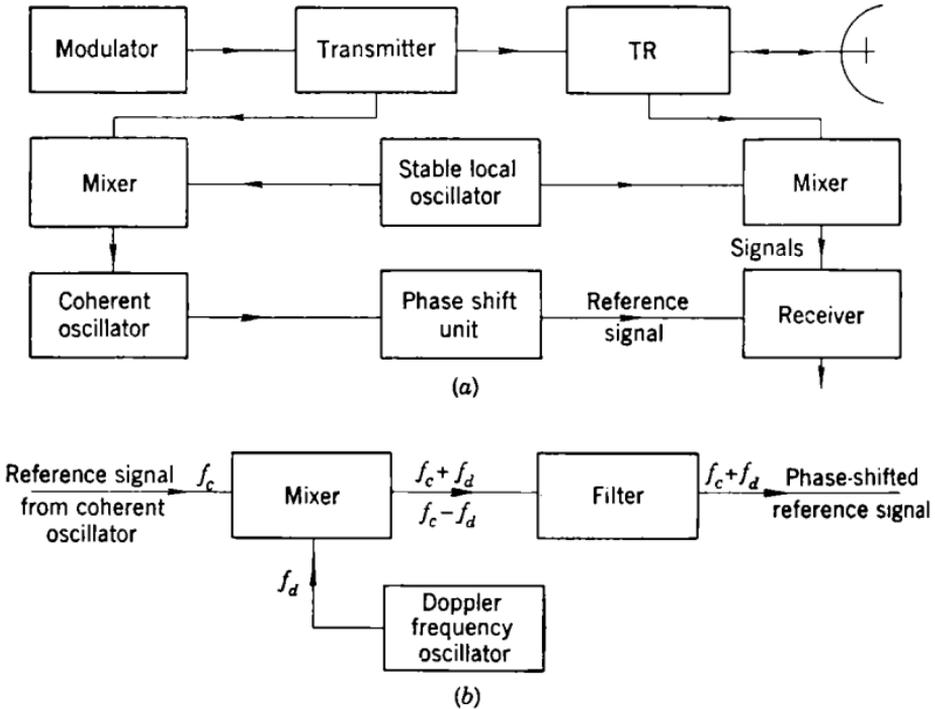


FIG. 16-25.—MTI on a moving system. (a) Block diagram of system showing location of phase-shift unit. (b) Simplified block diagram of phase-shift unit.

In the noncoherent method the local oscillator does not have to be stabilized since phase changes are not to be detected. However, the transmitter must have the same stability as in the coherent method because the overlapping ground clutter beats with itself if the transmitter frequency varies from pulse to pulse. As shown in Table 16-1 the stability requirement is not at all severe.

Another great advantage of the noncoherent method is that it works even when the radar set is carried on a moving station. This is because the set is sensitive only to amplitude fluctuations, whereas the motion of the station causes mainly phase changes in the received echoes. An amplitude effect, due to the finite beamwidth, will be discussed later.

The simplest way of employing the noncoherent method is to use an A-scope presentation and watch for echoes that are fuzzy on top. This type of echo indicates the presence of a moving target in the clutter at the range shown on the scope. The scheme is useful only when very slow scanning can be tolerated. An alternative indication can be obtained by gating off all but a short portion of the A-sweep and putting the output through a pair of headphones. A moving target within the "gate" then produces a musical note in the phones.

A more useful arrangement is to put the video output through the usual cancellation circuits and display the moving targets on a PPI. With such equipment in an airplane, roads will show up on the scope as intersecting lines of dots due to the moving vehicles.

The drawback to the noncoherent method is that a moving target can be detected only when there is ground clutter at the same range and azimuth as the target. Thus although the method works well out to the range where the clutter begins to get patchy, beyond this range moving targets become lost in the clear places. A plan for overcoming this difficulty is to have noncoherent operation for short ranges and coherent operation for long ranges. This can be accomplished by gating off the coherent reference oscillations for the required number of miles at the start of each sweep. Since the station is moving, it is necessary, during the coherent part of the sweep, to compensate for this motion. The method of doing this has been described in the preceding section.

16-13. Beating Due to Finite Pulse Packet.—In addition to the fluctuations due to scanning and wind, there is another kind of fluctuation when the system is moving. This arises from the fact that the clutter elements illuminated by the beam at a given instant do not, because of the spread of radial velocities within the finite illuminated area, all have the same doppler frequency. Signals from clutter will therefore beat with each other at frequencies up to $f_1 = 2 \Delta v / \lambda$, where Δv is the spread in radial velocity.

Let v be the velocity of the system and θ the azimuth angle. Then the radial velocity is $v \cos \theta$, and thus we have

$$\Delta v = -v\Theta \sin \theta \quad \text{for } \theta \gg \Theta,$$

where Θ is the beamwidth. If the beating is represented by a factor $\cos(2\pi f_1 t)$, then the fractional change in amplitude from pulse to pulse is of the order $2\pi f_1 / f_r$, which can be written as $4\pi \Delta v / (\lambda f_r)$ or $2\pi \Delta v / v_1$, where v_1 is the first blind speed as defined in Sec. 16-9. For example, suppose

$$\begin{aligned} v &= 300 \text{ mph}, & v_1 &= 800 \text{ mph}, \\ \theta &= 2^\circ = \frac{1}{30} \text{ radian}, & \theta &= 90^\circ. \end{aligned}$$

Then $\Delta v = 10$ mph and the fluctuation is 8 per cent.

When $\theta = 0$ the above equation for Δv no longer holds and we have instead

$$\Delta v = \frac{1}{8}v\theta^2 \quad \text{for } \theta = 0.$$

For $v = 300$ mph and $\theta_0 = \frac{1}{30}$ radian, as before, we get $\Delta v = \frac{1}{24}$ mph, showing that the fluctuations are negligible when the antenna is pointing straight ahead.

When the system is carried in an airplane there is an effect due to the depression angle as well as the azimuth effect just discussed. Since the beam is usually broad in a vertical plane, the range of depression angles within the illuminated area of clutter is limited by the pulse width. Under these conditions the spread in radial velocity is given by

$$\Delta v = \frac{\tau v}{2h} \frac{\sin^3 \phi \cos \theta}{\cos \phi} \quad \text{for } \phi \ll 90^\circ,$$

where τ is the pulse length (in feet), h the altitude, ϕ the depression angle, and θ the azimuth. For $\tau = 1000$ ft (1 μ sec), $h = 10,000$ ft, $\phi = 45^\circ$, $\theta = 0$, $v = 300$ mph, we get $\Delta v = 8$ mph, which is about the same as the above value for the azimuth effect.

For very large depression angles, the formula becomes

$$\Delta v = v \sqrt{\frac{\tau}{h}}, \quad \phi = 90^\circ$$

For $v = 300$ mph, $\tau = 1000$ ft, $h = 10,000$ ft, we have $\Delta v = 100$ mph, illustrating that very large fluctuations are present in the clutter immediately below the airplane.

COMPONENT DESIGN

The MTI systems designed at the Radiation Laboratory have been at 10 and 3 cm. For that reason, the discussion of typical components given here will be in microwave terms, despite the fact that MTI methods are useful with radar systems operating at longer wavelength. The components that will be described have been chosen to illustrate recent developments or the breadth of possibilities, or to supplement the detailed treatment of the components elsewhere in the Radiation Laboratory Series.

16-14. The Transmitter and Its Modulator.—The transmitter stability requirements, specified in Sec. 16-5, can be met by magnetrons of existing design. The 10-cm magnetrons that have been successfully tried for MTI are the type 2J32, the type 4J47, and the type 706CY. At 3 cm, the type 2J42 has been used. Most experience available up to 1946 has been with the 2J32. With the 2J32 and a hard-tube modulator, frequency modulation within a pulse never reaches an objectionable level

when amplitude modulation is avoided. Information on this point is not available for the other types.

Magnetron frequency pulling with antenna scanning can be objectionable with any type of tube. It may be caused either by imperfect rotary joints or by obstacles in front of the antenna. If the pulling is slow compared with 10 Mc/sec² (see Table 16-1), ground clutter will remain canceled at all azimuths. However, when the absolute detuning becomes excessive, moving targets will be lost owing to their high video spectrum, as is explained in Sec. 16-5. This type of detuning is particularly insidious.

Objectionable modulation of the transmitted pulse by the use of 60-cps power on the heater has been observed with the type 4J47. This is probably caused by the magnetic field of the heater current. No such difficulty has been found with the 2J32, owing apparently to the design and construction of the heater. Three methods of avoiding this modulation are available. At high PRF's the heater current can be turned off after starting the system, since the cathode temperature will be maintained by bombardment. Alternatively the PRF can be made an exact submultiple of the power-line frequency. Since neither of these methods can be generally applied, it will sometimes be necessary to operate the magnetron heater on direct current. Heater hum modulation should eventually be made a matter of tube-manufacturing specification.

The use of a supersonic delay line necessitates a constancy of PRF to $\frac{1}{30}$ of a pulse width. This is the most severe new requirement placed by MTI upon the modulator. It has been successfully met by hard-tube and by hydrogen-thyratron designs. In their current state, triggered-gap modulators were not quite adequate for MTI systems in 1946. Since little effort had been put upon the jitter problem, it seemed reasonable to expect that triggered gaps can be considerably improved. The rotary gap can be used for MTI only if a storage tube serves as delay element.

16-15. The Stable Local Oscillator.—Two types of 10-cm local oscillator tubes have been investigated: the type 417 and type 2K28 reflex klystrons. The 417 has a built-in cavity and is therefore more microphonic than the 2K28, which has a demountable cavity. However, the external cavity gives the 2K28 a lower Q , which causes larger noise fluctuations and greater sensitivity to external 60-cps magnetic fields. Both tubes show frequency modulation when the heaters are run on a-c power.

The 2K28 was chosen for incorporation into the local oscillator unit shown in Fig. 16-26. The outer metal case of this component is lined with sound-proofing material. The parts within are mounted on a metal base-plate which is in turn supported by four circular shear rubber shock mounts of conventional design, not visible in the photograph.

The oscillator tube is mounted, with its axis vertical, in a blackened mu-metal shield visible to the left. A single pickup loop is connected to two lines via a decoupling fitting. A tunable echo box (Johnson Service Company Model TS-270/UP) is coupled to the oscillator by a rotatable

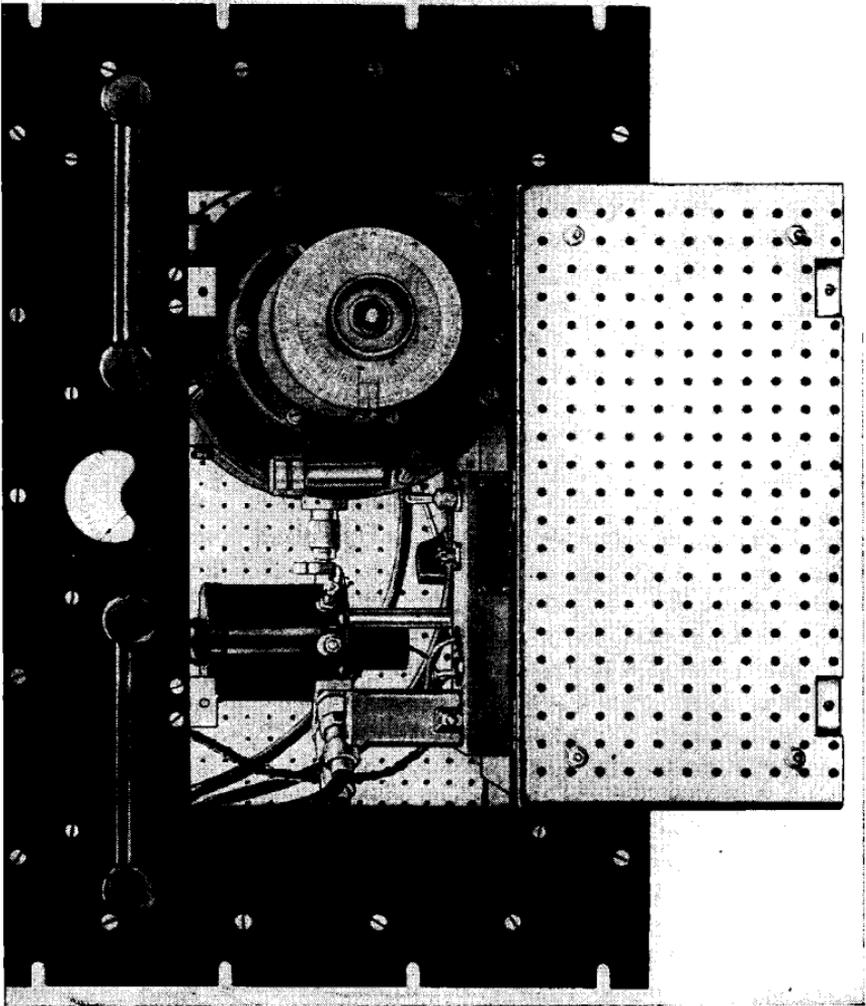


FIG. 16-26.—A 10-cm stable local oscillator, using a 2K28 reflex klystron, stabilized by an echo box.

loop, the lever arm of which is visible close to the oscillator. An echo-box crystal-current meter is of value in making tuning adjustments.

The remote power supply uses conventional electronic regulation. A simple *RC*-filter at the oscillator isolates the reflector electrode. Heater power is furnished by a dry-disk rectifier and high-current filter.

The optimum value of echo-box coupling varies for different fre-

quencies and different oscillators. If this coupling is too loose, frequency instability will result. If the coupling is too tight, tuning will be difficult and the power output from the local oscillator will be reduced. It must be possible to obtain two to three times the mixer current that will actually be used. When restarting a correctly tuned oscillator, it may be necessary to vary the reflector voltage slightly up and down to cause the echo box to lock in.

A microwave discriminator circuit developed by R. V. Pound¹ gives promise of overcoming many of the difficulties of the simple echo-box stabilizing arrangement. Measurements of local-oscillator stability at 3 cm, as reported below, show that this system of stabilization might be adapted to MTI.

TABLE 16-5.—TYPICAL LOCAL-OSCILLATOR STABILITIES

Description of oscillators compared. (cf. Sec. 16-21)	Frequency average of $\frac{df}{dt}$, (Kc/sec ²)	Frequency band to half-power, cps
Two identically shock-mounted type 417 reflex klystrons (10 cm) without stabilizing cavities	79	1300
	61	250
	50	140
Two identical type 2K28 reflex klystrons (10 cm), without stabilizing cavities, in a quiet room	168	1300
	100	250
Same tubes as above, one stabilized	137	1300
	68	250
Same tubes, both stabilized	1.4	250
	1.4	140
Two other stabilized tubes of same type in a quiet room (more typical value)	3	250
Two such oscillators; 60 cps a-c heater power used for one	36	250
Two such oscillators operating in an SCR-584 truck	10	
Two such oscillators operating in the radio compartment of a B-17 aircraft after somewhat improved shock mounting	70	
Two type 2K25 reflex klystrons (3 cm) in the laboratory with echo-box frequency control	20	

¹ Described in detail in Chap. 2 of Vol. 11 of the series.

Equation (3), giving the allowable rate of frequency drift in the local oscillator, was based on the assumption that the frequency was changing at a constant rate. Most frequency deviations are not of this simple sort. They are more likely to be periodic or random, with a superimposed uniform drift. The rate of change of frequency can be expressed by a Fourier spectrum and the total phase error determined by superimposing the contributions resulting from each of the Fourier components. These components are not all equally harmful to MTI performance. A frequency modulation exactly synchronized with the PRF or any multiple thereof would cause no MTI phase error whatsoever. The phase of ground clutter at a particular range would be shifted by a given amount from its value in the absence of the frequency modulation, but the shift would be constant for successive transmitted pulses.

It has been shown¹ that the rate of change of oscillator frequency averaged over all modulation frequencies can be used in place of the simple df/dt of Eq. (3) as a figure of merit for local-oscillator stability. Table 16-5 shows some typical figures on local-oscillator stability.

16-16. The Coherent Oscillator.—Locking at radio frequency has been accomplished at 10 cm, but its feasibility for field service is yet to be investigated. The relative ease and dependability of i-f locking have resulted in its use for all coherent microwave systems of the past. This section will discuss oscillators that receive an i-f locking pulse and deliver an i-f reference signal according to the arrangement of Fig. 16-9.

The Locking Pulse.—If the phase-locking is to be constant from one pulse to the next with a precision of 1° , the carrier-frequency packet that does the locking must be unusually free of random components. These may arise from several sources, of which the most likely is the transmit-receive switch. The gaseous arc discharge in the TR tube generates noise-like oscillations. These may reach the coherent oscillator by passing from the signal mixer, via the local oscillator, to the locking-pulse mixer. The available remedies are the use of loose mixer coupling, lossy local-oscillator cable, and separate pickup loops in the local-oscillator cavity.

Another cause of random components in the locking pulse is excessive r-f pulse amplitude at the locking mixer. If the r-f pulse exceeds the level of the local oscillator, self-rectification will take place, yielding a video pulse that will have noncoherent components at the intermediate frequency. Even though an effort has been made to reduce the r-f pulse amplitude, this trouble may persist in a subtle form. If the attenuator preceding the locking mixer is of the simple "waveguide-beyond-cutoff" variety, it may allow harmonic frequencies, generated by the transmitter,

¹ S. Roberts, "A Method of Rating the Stability of Oscillators for MTI," RL Report No. 819, Oct. 16, 1945.

to pass to the mixer with so little loss that their amplitude is comparable to or greater than that of the fundamental frequency.

A further source of locking trouble is statistical noise. If a simple cable-coupling circuit is used in the absence of a locking-pulse preamplifier, care must be taken to maintain adequate pulse amplitude.

Circuit Design.—Any conventional type of freely running oscillator can be phase-locked by injecting into its tuned circuit a sufficiently large carrier pulse. There is a certain degree of incompatibility between the requirements that the oscillator be extremely stable and yet precisely phasable. A simple method of reconciling these requirements is to stop the oscillator completely before each radar transmission so that the locking pulse is also a starting pulse. By this process, all previous phase information is destroyed and the locking pulse is required to overwhelm merely the statistical noise fluctuations present in the oscillator tuned circuit. The stopping of the coherent oscillator can be accomplished by applying a bias "gate" perhaps 100 μ sec before each transmitted pulse. This gate is released just before, during, or just after the locking pulse, while the resonant circuit of the oscillator is still ringing. The amplifying circuits that precede the oscillator can be made reasonably narrow since pulse shape is unimportant. These circuits must be broad enough to allow for local-oscillator detuning and to maintain a large locking-pulse-to-noise ratio. It is advantageous to use a bias slightly beyond cutoff in a late amplifying stage in order to suppress any spurious low-level oscillations which, by their persistence, might introduce an appreciable phase error either before or after locking. The actual injection of the locking pulse into the oscillator might be accomplished in any of several ways, including the use of an extra oscillator control grid. The most flexible method and the one most widely used is the connection of the plate of a pentode injection-tube across all, or part of, the LC-circuit of the oscillator.

The oscillator itself must be designed with considerable care. The resonant circuit must have a high Q . Input and output loading must be held to a minimum. Heater hum modulation must be minimized by keeping the d-c and r-f impedance between cathode and ground as small as possible. The oscillator tube must be of a nonmicrophonic type. The circuit as a whole must be rigidly constructed and shock-mounted. Power-supply ripple must not exceed a few millivolts. When these precautions are observed, free-running stabilities of around 1 kc/sec² can be obtained at 30 Mc/sec. A typical oscillator with its preceding and following stages is shown in Fig. 16-27.

Phase-shift Unit.—Figure 16-28 shows, in more detail than the simplified diagram of Fig. 16-25*b*, the arrangement of the phase-shift unit required for MTI on a moving system. For station motion up to 400 mph,

the doppler frequency at 10 cm will vary from plus to minus 3500 cps. The way in which this frequency is added to the reference signal is shown in the diagram. The doppler oscillator receives a shaft rotation and

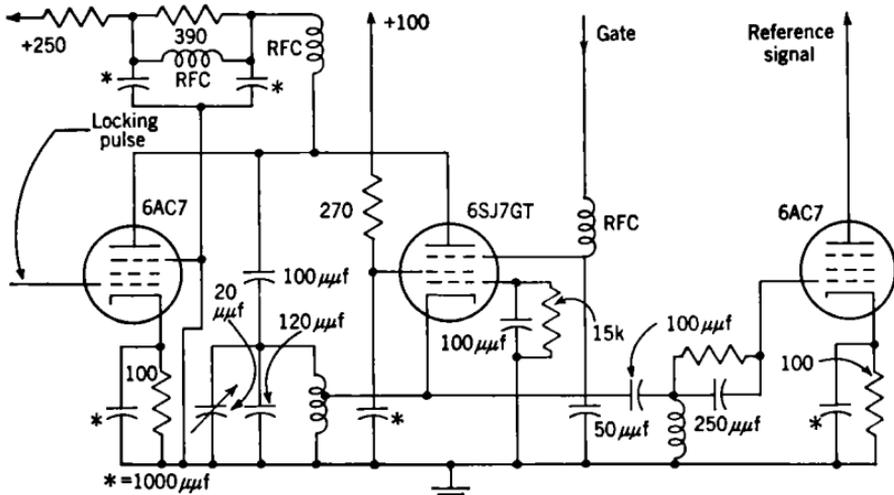


FIG. 16-27.—A typical 30-Mc/sec coherent oscillator circuit.

delivers a linearly related frequency with an absolute accuracy of 2 cps. This oscillator is similar to the General Radio Type 617-*c* Interpolation Oscillator. The single sideband rejection requirements for the carrier and unwanted sideband are very severe. The quartz filter was designed

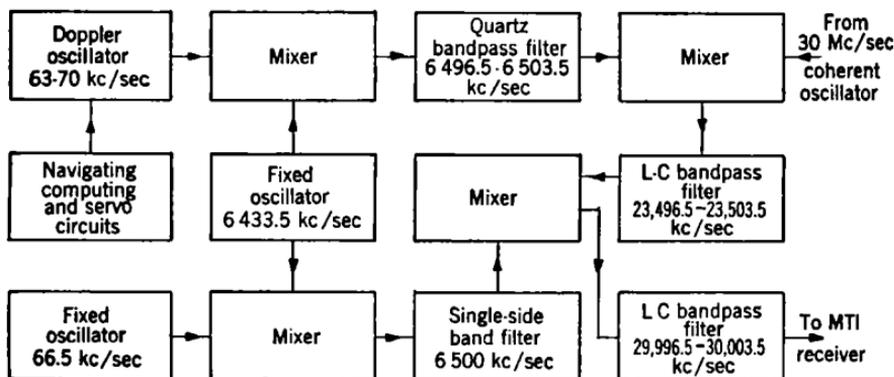


FIG. 16-28.—Block diagram of a phase-shift unit.

by Bell Telephone Laboratories. Pending delivery of the latter, the circuit was tested by means of LC-filters with one additional mixing step, and found to cancel target motion satisfactorily. The absolute accuracy of the circuit from the shaft of the doppler oscillator is about

$\frac{1}{2}$ mph at 10 cm. Full circuit information can be found in a report by V. A. Olson.¹

16-17. The Receiver.—Two types of receiver characteristic have been discussed in Sec. 16-8. The lin-log type is necessary in a noncoherent system, and the limiting type is best for coherent systems.

The Lin-log Characteristic.—A linear-logarithmic characteristic is linear for small signals and limits in a specified gradual fashion for large signals. A method for achieving this response is shown in Fig. 16-29. Output signals from several successive i-f stages are combined. At low signal levels, the last stage delivers a linear signal in the normal fashion, the output from the preceding stages being negligible. As the signal level increases, the output signal of this last stage reaches a saturation level

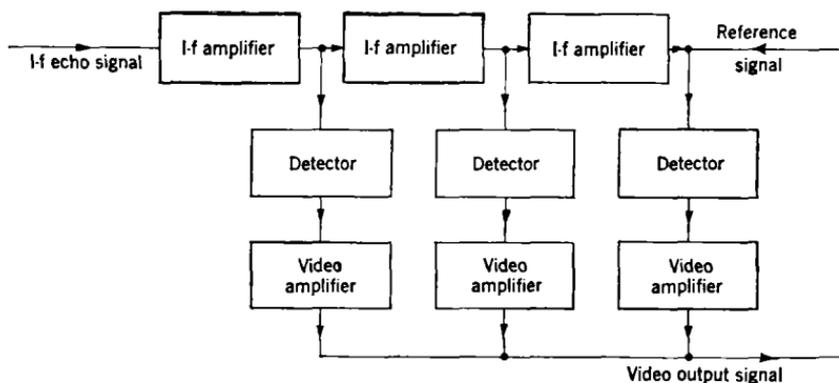


FIG. 16-29.—Circuit to give lin-log response.

above which it cannot rise regardless of further input-signal increases. At this point the output signal from the preceding stage has become appreciable; it continues to increase with increasing input signal. Now, however, the increase in output-signal level is at a slower rate because the gain of the last stage is not available. Eventually the second stage overloads and the third from the last becomes effective. The amplitude response curve of such a circuit can be adjusted to be accurately linear-logarithmic.²

The Limiting Receiver.—The block diagram of a typical limiting receiver is shown in Fig. 16-30. The arrangement is one that might be appropriate to a system employing a 1- μ sec pulse width and an intermediate frequency of 30 Mc/sec. The requirement that the amplitude response show no inversion above the limit level is one not ordinarily met

¹ V. A. Olson, "A Moving Coho Conversion Unit," RL Report No. 975, Apr. 3, 1946.

² The design problems involved in this and other MTI receiver circuits are discussed in Vol. 23, Chap. 22.

by radar receivers. It will be noticed that several extra stages are necessary to obtain this type of limiting. The coherent oscillator circuit has already been described. The function of the "coherent oscillator locking test" channel will be discussed in Sec. 16-21.

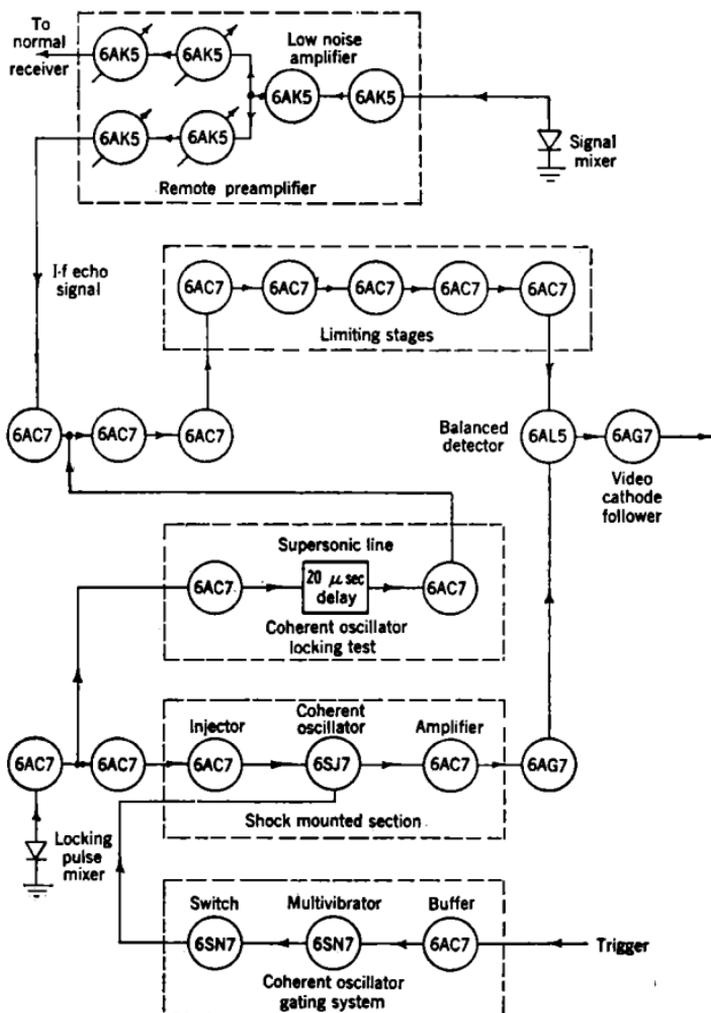


FIG. 16-30.—Block diagram of a typical MTI receiver: i-f limiting; balanced detector; intermediate frequency 30 Mc/sec; 1-μsec pulses.

Figure 16-31 shows the circuit of the balanced detector, which converts phase changes at the input into amplitude changes at the output with a fair degree of linearity.¹

¹ The theory of the circuit is given in Vol. 23, Chap. 22.

16-18. The Supersonic Delay Line.—The principal elements of a delay line are the transmission medium and the electromechanical transducers. Because of the bandwidth required, MTI delay lines have been

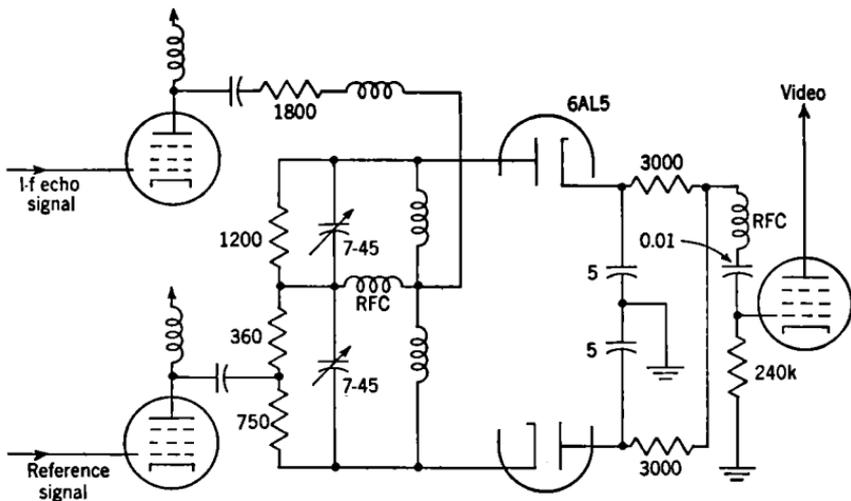


FIG. 16-31.—Circuit of a balanced detector.

operated with amplitude-modulated carrier frequencies in the region of 15 Mc/sec. The transducers consist of piezoelectric quartz, X-cut and optically polished. The two transmission media so far used are mercury

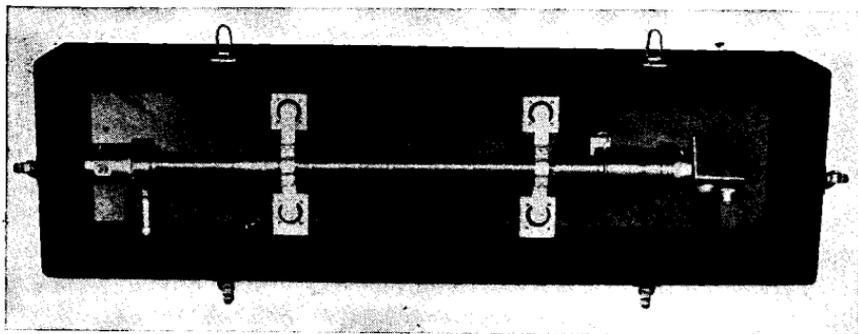


FIG. 16-32.—A supersonic mercury delay line of variable length.

and fused quartz. The principal factors encountered in line design have been treated in detail elsewhere¹ and will be briefly reviewed here.

Liquid Lines.—The simplest delay line is a straight tube with parallel transducers at the end. A laboratory design of mercury line has been shown in Fig. 16-10. A model for field use is shown in Fig. 16-32. The

¹ Chap 6, Vol. 17; Chap 24, Vol. 19, Chaps. 13, 14, Vol. 20.

latter line is variable in length over a small range to compensate for temperature changes or to allow the use of MTI on more than one beam of a synchronized multibeam system. The screw thread and sliding alignment surfaces are external to a stainless-steel bellows whose length is three-quarters the length of the delay line. The mean delay is 586 μ sec.

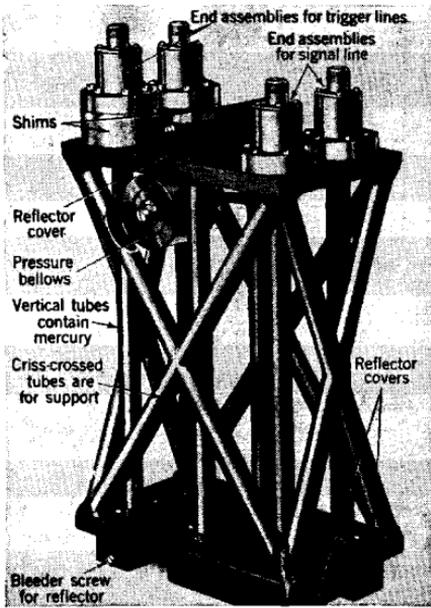


Fig. 16-33.—A folded mercury delay line.

the nearest tube, is reflected twice, passes up the tube next left, etc., until after six reflections it enters the cell at the extreme right. The trigger line is half as long, giving rise to a double-frequency trigger which is divided electrically.

This type of line is relatively easy to make. The two end surfaces of the main assembly are ground flat and parallel after welding. The tolerance in parallelism is eased for the critical signal line by the auto-collimating effect of the three reflector blocks. Freedom from warping is insured by the use of hot-rolled steel and by annealing. The corner-reflector blocks are precision-ground pieces, but of a shape familiar to a tool-making shop. These reflectors bolt without gaskets to the main end plates beneath the visible housings. Thus, the only gaskets that might affect alignment are those beneath the end cells. Air is removed from the line; thermal expansion is accommodated by a bellows within the cylindrical housing on the side of the line. A positive pressure is maintained so that any leakage is outward. The line normally operates with end cells downward so that small remaining air bubbles will collect above the corner

The line, with expansion chamber and electrical end housings, is shown shock-mounted in a thermally insulated box. This line was designed for use at 30 Mc/sec. Quartz-to-quartz parallelism is maintained to 0.03°

A straight tube having the length necessary for low-PRF systems would be inconvenient for field use. Several methods of folding delay lines have been successfully used, one of which is shown in Fig. 16-33. In addition to the intersecting truss members, there can be seen six vertical tubes. Four of these constitute the signal delay line; the other two, the trigger-generating line. Beginning at the end cell on the nearest top corner, the signal passes down

reflectors out of the path of the beam. The tubing has an inside diameter of $\frac{3}{8}$ in. The total signal delay is 1000 μ sec. The carrier frequency is 15 Mc/sec, and the pulse length is $\frac{1}{2}$ μ sec.

Another delay-line arrangement that is feasible but has not seen field service is the mercury tank. A quartz plate oscillating in a large volume of mercury will generate a "free-space" beam whose width can be calculated from diffraction theory. If such a beam is incident upon a wall of the mercury container, it can be specularly reflected without appreciable loss, providing the angle of incidence is sufficiently large. Laboratory delay lines have been constructed consisting of a mercury-filled tray within which a supersonic beam is repeatedly reflected from the side walls in any of a variety of geometrical patterns until the beam reaches the receiving quartz. Designs that have been proposed in the past have appeared inferior to the folded line as regards weight, size, or ease of construction, and have consequently not been pursued to the engineering stage.

The Fused Quartz Line.—Solid media offer difficulties not encountered in liquids. The uncontrolled transfer of energy between compressional and transverse modes will give rise to pulse distortion and multiple echoes. Such difficulties may occur at the boundaries of the material or at points of inhomogeneity. Since sound velocity in solids is greater than that in mercury by a factor of about 4, the beam spread is greater and the delay path longer. However, a 420- μ sec line giving acceptable performance has been made of an annealed fused quartz block 6 by 17 by 18 cm. This line operated at 15 Mc/sec with a pulse length of less than 1 μ sec. An internal reflection pattern was employed which resembled the pattern of the liquid delay tank, but use was made of the specific elastic properties of fused quartz to achieve a controlled transfer of energy between compressional and transverse modes. The result was that the energy in the transverse mode could be reduced to zero after each group of three reflections. Although much engineering remains to be done, it is possible that delay times can be extended to 1000 or 2000 μ sec by the use of larger blocks.

End Cells.—The various ways of mounting the quartz crystals may be classified according to whether the backing behind the crystal is intended to absorb or to reflect. Spurious echoes may reach the receiving crystal in either of two principal ways. If the vibratory signal *enters* the backing material it may emerge later, unless care is taken to absorb or disperse the energy. On the other hand, if the energy is predominantly reflected at the surface of both end cells by the use of a mismatching material in contact with the piezoelectric quartz, the signals may return to the receiving end a second time after reflection, first from the receiving and then from the transmitting quartz.

The end cell shown in Fig. 16-10 is highly efficient as an acoustic absorber but is useless outside of the laboratory because the unsupported quartz plate can be cracked by standing the line on end. Several methods of supporting the crystal from the rear while retaining the mercury as a backing medium have been tried. A more promising approach to the design of an absorbing end cell is the use of lead soldered to the quartz. The lead provides a good acoustic match to mercury, while strongly attenuating the supersonic energy that it receives.

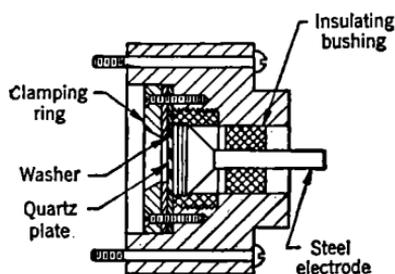


FIG. 16-34.—A reflecting end cell for a mercury line.

The use of absorbing end cells is preferable with mercury lines of delay less than about 500 μ sec. With longer lines, the use of reflecting end cells may be better for two reasons. In the first place, attenuation always takes place within the mercury. Thus, by the choice of carrier frequency and tube diameter, it is possible to reduce the amplitude of unwanted echoes to less than 1 per cent of the desired signals. Second, the use of absorbing end cells costs 12 db (6 db per end). The construction of a reflecting cell is shown in Fig. 16-34. The quartz is supported directly by the steel electrode. The occluded air film provides the necessary acoustic mismatch.

Design Constants for Mercury Lines.—The delay time is presumed to be given. From it, the length of mercury column can be calculated according to the formula

$$D = L(17.42 + 0.0052T), \quad (18)$$

with an estimated probable error of 0.06 per cent between 10°C and 40°C, where D is the delay in microseconds, L the length in inches, T the centigrade temperature.

The two quantities to be chosen, along with the line configuration and the end-cell type, are the carrier frequency and tube diameter. These enter into the design mainly in connection with attenuation, bandwidth, and demodulation. The over-all delay-line attenuation can be divided into two parts: that which occurs in the medium itself, and that which has to do with the efficiency of the quartz crystals in converting electrical energy into acoustic energy.

Both carrier frequency and tube diameter affect the attenuation that occurs within the delay medium. The free-space frictional attenuation in liquids is proportional to the square of the frequency. In mercury this attenuation is about 0.11 db/in. at 10 Mc/sec. The attenuation due to wall effects is less clearly understood but is believed to vary inversely

as the diameter of the tube and directly as the square root of frequency. An experimental value is 0.1 db/in. at 10 Mc/sec with a $\frac{1}{4}$ -in. tube.

The efficiency of the crystals can be stated in terms of the transfer impedance of the delay line, as follows:

$$\frac{V_2}{I_1} = \frac{2R_1R_2}{R_k},$$

where I_1 is the current driving the transmitting crystal, V_2 is the voltage developed across the receiving crystal, R_1 is the load resistance shunting the transmitting crystal and R_2 that for the receiving crystal, R_k is the electrical equivalent of the characteristic impedance of the line. The equation applies to the case of reflecting end cells. Now R_1 and R_2 depend on the bandwidth that is required, since the bandwidth at each crystal is inversely proportional to the product of the loading resistance and the total capacity, including the electrostatic capacity of the crystal and the stray capacity of cable connections. This assumes that single-tuned coupling circuits are used.

The characteristic impedance R_k is inversely proportional to the active area of the crystal and to the square of the frequency. At 10 Mc/sec the value is 22,000 ohms for an active area of 1 in.² With this area the total capacity at each crystal is about 50 $\mu\mu\text{f}$, and the bandwidth is then 3.8 Mc/sec per end if $R_1 = R_2 = 300$ ohms. These values give a transfer impedance of 8.6 ohms. It will be noticed that no account is taken of the inherent bandwidth of the crystals. This is because the bandwidth for both crystals in series is more than $\frac{8}{10}$ of the carrier frequency.

It can be shown that, for a given bandwidth, the transfer impedance is a maximum when the crystal diameter is chosen so that the active crystal capacity equals the total stray capacity.

The carrier frequency for mercury lines may range from 5 to 30 Mc/sec, with the most usual values in the range from 10 to 20 Mc/sec. Choice of frequency depends upon the delay-line attenuation, and on the fact that the number of carrier cycles per pulse must be sufficient to allow accurate demodulation.

It was mentioned earlier that the quartz-mercury combination has adequate inherent bandwidth. In a reflecting end cell the response is actually flat over a frequency range equal to 40 per cent of the carrier frequency. This allows a simple cure for a practical difficulty that is frequently encountered with this type of cell. Because the reflecting end cell is loaded on one side only, a slight smudge or spot of scum will unload a portion of the crystal area without removing the electrical excitation. When this happens, the pass band exhibits sharp spikes at the resonant frequency of the quartz. These seriously impair the quality of cancella-

tion. This difficulty can be avoided by operating at a carrier frequency that is far off crystal resonance but still on the flat part of the loaded characteristic.

16-19. Delay-line Signal Circuits.—The delay line is a low-impedance device with a large attenuation; it is mechanically awkward to mount the line physically close to vacuum-tube circuits. These facts combine to make the delay process electrically expensive. The subject of delay-line signal circuits has been treated at length in another book.¹ It will be reviewed here.

Line-driving Circuits.—The requirement that the signal output shall be well above the statistical noise level generally means that considerable power must be furnished to the transmitting crystal.

If an i-f signal from the receiver is to drive the line, amplification is required. Otherwise, a carrier frequency must be modulated with a video signal. This problem is similar to that of the television transmitter. Only a reasonable degree of linearity is required; the response slope should not vary by more than perhaps 15 per cent. The real difficulty arises from the proximity of the carrier and video frequencies and from the requirement that the delayed and undelayed video signals must cancel to 1 per cent in amplitude.

In simple modulators, modulation components are removed from the modulated carrier channel by frequency discrimination. If the modulation spectrum extends to more than half the carrier frequency, spurious transients in the carrier pass band may arise in two ways. The modulation spectrum may have components that lie in the carrier-channel pass band, or modulation components may add to carrier side-band components to give components lying within the carrier-channel pass band. Further, spurious components may arise from frequency modulation of the carrier oscillator by the video signal. Since the carrier phase is random from pulse to pulse, the envelope of the summation of all such components within the carrier pass band will vary from pulse to pulse so that cancellation will be imperfect. These difficulties can be avoided by the balanced modulation of a carrier amplifier driven by an isolated oscillator. A common procedure is to use parallel-grid modulation of a push-pull amplifier stage. Alternatively, the problem might be solved by phase-locking the carrier oscillator to the transmitted pulse trigger, thereby insuring a constant relative phase between each echo and the carrier. Oscillator isolation would still be necessary, since otherwise a strong moving target might destroy the phase coherence.

The Two-channel Amplifier.—The signal level at the input terminals of the two-channel amplifier must be great enough so that the statistical noise generated therein is considerably smaller than the statistical noise

¹ Vol. 20.

originating in the radar receiver. The high cost of delay-line input power makes necessary the careful conservation of output signal. Maximum signal-to-noise ratio is obtained with a short lead from the delay line to the first amplifier grid. If the amplifier cannot be located within a few inches of the delay line, a preamplifier is necessary. Low-noise input circuits can profitably be used, but critically adjusted circuits should be avoided, if possible, because of the difficulty of maintaining balance between the delayed and undelayed channels.

The bandwidth of the amplifier must be sufficient to avoid serious impairment of system resolution and the loss of system signal-to-noise.¹ In addition, the bandwidth of the delayed and undelayed channels must be sufficient to make the interchannel balance noncritical. Bandpass unbalance will have relatively little influence upon pulse shape if the bandwidth of each split channel is perhaps twice that of the over-all system.

Even though the bandwidth of these channels is large, it may still be necessary to make an approximate compensation for two types of unbalance. The smaller gain in the undelayed channel would ordinarily result in a greater bandwidth. This channel should therefore be narrowed to match the other by the addition of capacity to ground in the low-level stages. The second type of unbalance arises from the frequency dependence of attenuation in the delay-line medium. The square-law factor generally predominates. Compensation can be obtained by inserting in the delayed channel an LC-circuit tuned to a frequency much higher than the carrier frequency.

Amplitude cancellation to 1 per cent implies an extraordinary degree of linearity in the two channels. The carrier level at the canceling diodes must be well above the region of square-law diode response. The final i-f stages must be conservatively operated. A first-order correction for residual nonlinearity will result if the last two i-f stages of the two channels are identical in every respect.

Gain adjustments should be made in low-level stages to avoid the introduction of nonlinearity. A wider range of adjustment will be required in the delayed channel because of possible variations in delay-line attenuation. A first-order balance of transconductance changes due to heater voltage fluctuations will result if the same number of amplifying tubes is used in each channel. Pentodes with high mutual conductance frequently exhibit small sudden changes in gain as a result of either heating or vibration. The gains of the individual channels can be stabilized through feedback of bias derived by detection of the carrier level. A

¹ In coherent MTI, the signal-to-noise ratio depends upon the over-all system bandwidth and not in any special sense upon the bandwidth up to the detector.

more elegant approach is to measure the cancellation residual of a sample pulse and to correct the relative channel gain accordingly.

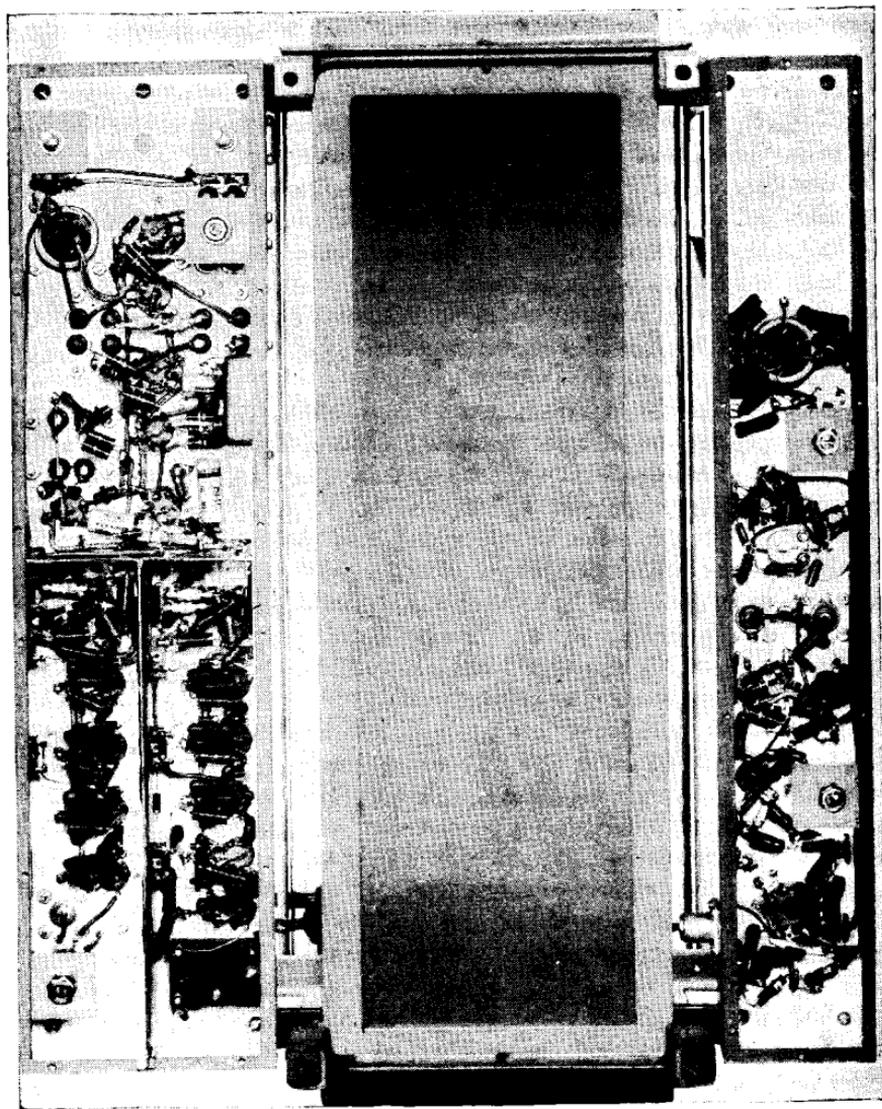


FIG. 16-35.—Signal circuits and case for the line of Fig. 16-33; two-channel amplifier on the wide chassis at left, video-modulated oscillator and amplifier on the narrow chassis at right.

After detection and cancellation of the echoes in the delayed and undelayed channels, the residual video signal must be amplified, rectified, and limited for application to an intensity-modulated indicator such as

a PPI. The most difficult of these operations is amplification. If statistical noise and weak moving-target signals are to be fully visible, they must reach the final video rectifier at a level sufficient to cause linear operation of that detector. Since the noise level at the cancellation detector may be as much as 40 db beneath the signal peak, a gain of several hundred may be required between cancellation and rectification. This video amplification is difficult because of the bipolarity and wide dynamic range of the signals. Unless special precautions are taken, strong signals from moving targets may draw grid current and bias the amplifier to the point where weaker moving-target signals are lost. The bipolar nature of the signals prevents the application of the clamping technique, which is useful in avoiding a similar trouble with unipolar video amplifiers. Blind spots and overshoots can be adequately minimized by careful choice of time constants, grid bias, and plate load resistors, although several additional amplifier tubes may be needed as a result.¹

Full-wave rectification before application to a unipolar device such as a PPI is roughly equivalent to doubling the PRF_v and is correspondingly to be recommended. Either crystal or diode rectifiers may be used. Limiting can take place both before and after rectification.

Figure 16-35 shows the two-channel amplifier on the left, with the delay-line box in the middle, and the video-modulated oscillator and amplifier for driving the delay line on the right.

16-20. Delay-line Trigger Circuits.—For proper cancellation the PRF of the radar transmitter must match the supersonic delay to $\frac{1}{16}$ of a pulse length. Free-running oscillators for generating the trigger have been built to fire the transmitter at a suitably constant rate. Because of the large thermal coefficient of delay in mercury, normal fluctuations in ambient temperature necessitate manual readjustment of such oscillators at intervals ranging from 10 min to 1 hr. Since this amount of attention cannot be tolerated in most applications, several methods have been devised for automatically maintaining time synchronism between the PRF and the supersonic delay.

Time synchronism is maintained by generating trigger pulses at a PRF determined either by the signal delay line or by an auxiliary delay line. The methods for doing this can be divided into two classes: regenerative and degenerative. The regenerative method has already been described briefly in Sec. 16-3.

The Degenerative Trigger Circuit.—A highly stable oscillator generates the transmitter trigger, and a correction is applied to the oscillator if it fails to match the delay line. This correction is obtained as follows. Several microseconds before the transmitted pulse, the echo input to the

¹ Nonblocking amplifier design is discussed in Secs. 5-8 and 10-4 of Vol. 18.

delay-line driving circuit is desensitized by a trigger from the stable oscillator. This trigger does three other things: it generates a sample video pulse which travels down the line; it operates a coincidence circuit which examines the time of arrival from the comparison amplifier of the preceding sample pulse; and finally after suitable delay it fires the transmitter and resensitizes the echo circuits. The coincidence circuit supplies a variable d-c bias to the oscillator which keeps the frequency of that oscillator in step with the supersonic delay time as measured by the sample pulses.

This arrangement has the disadvantage that it takes a number of tubes and, like all multistage feedback circuits, is difficult to analyze in case of trouble. It has the advantage that, with small additional elaborations, the sample pulse residue after cancellation can be used to control the gain of one channel of the comparison amplifier. This type of AGC corrects for changes in both the amplifier and the delay line.

Regenerative Trigger Circuits.—Three distinct methods have been proposed for trigger regeneration. They differ in the way in which the fraction of a microsecond inevitably lost in the trigger amplifier is made up in the delayed signal channel. The compensating delay must be inserted after the point at which the trigger pulse has left that channel but before cancellation.

The first method is to add the required delay either at carrier or video frequency. Unfortunately, no electrical delay lines are available which are capable of reproducing a microsecond pulse with less than 1 per cent distortion.

The second method involves the use of an auxiliary delay-line receiving crystal slightly closer to the transmitting crystal than is the regular receiving crystal. The auxiliary crystal receives the trigger pulse. The final time adjustment can be accomplished in the trigger circuit either mechanically by an actual shift of the crystal or electrically by the use of a short variable delay. The disadvantages of this method lie in the excessive attenuation in the trigger channel, the increased difficulty of delay-line construction, and the loss of design flexibility due to the necessity of mixing the trigger and the video signals at the transmitting end of the delay line while preventing their interaction at the receiving end.

The third method of trigger regeneration, already described in Sec. 16-3, is the use of an extra delay line in close association with the signal delay line as illustrated in Fig. 16-33. This line is thermally lagged by the box of Fig. 16-35. The method allows freedom in the choice of line constants and of electrical coupling, with some resultant circuit simplification as compared with the three-crystal delay line. However, the amount of thermal correction is not as great as in the three-crystal line. The

extra half-length line generates triggers at double the PRF. These are divided by a counting circuit before delivery to the radar transmitter.

16-21. Special Test Equipment. Operating Tests.—The minimum additional equipment required to check the components peculiar to an MTI system can be incorporated in a single A-scope chassis. This chassis should include video amplifiers, expanded and delayed sweeps, and a vacuum-tube voltmeter, together with switches and permanent connections to other parts of the system. In addition, the receiver chassis should include the locking-test channel shown in Fig. 16-30, and the two-channel amplifier may contain a delay-line attenuator.

By means of a rectifying crystal permanently connected to the r-f transmission line, together with an expanded and suitably phased sweep, the transmitted pulse envelope can be inspected for hum modulation and mode jumping. This test can be used whenever the more general coherent oscillator locking test shows trouble.

The locking-pulse mixer current should be monitored. This can be done by switching the meter normally used for the signal-mixer current.

As can be seen from Fig. 16-30, the locking-test channel delays a sample of the locking pulse by means of a short auxiliary supersonic delay line, and then mixes this sample with the received signal. If reflecting end cells are used, multiple reflections within the delay line will give rise to a number of equally spaced locking pulse echoes. These echoes will beat with the reference signal. The presence of cycles within the echoes, when viewed on an expanded A-scope, will show when the local oscillator is out of tune. Since the coherent oscillator is unlikely to drift, the tuning of the local oscillator to give maximum receiver response needs to be checked only rarely.

The principal function of the locking-test channel, as its name implies, is to reveal any unsteadiness in the original locking pulse or any failure of the coherent oscillator to lock properly.

An attenuator and appropriate switches can be built into the input circuits of the two-channel amplifier to permit measurement of delay-line attenuation.

Checking and adjustment of cancellation can be done while the radar is operating by mixing a delayed video pulse with the signals before cancellation. The delay can be chosen to bring the test pulse to a range position near the outer edge of the PPI, and adjustments can then be made for the best cancellation of the video pulse.

Because of the several gain controls in series in the receiving train, a d-c vacuum-tube voltmeter is necessary to check the carrier level at the cancellation detectors.

Testing MTI Oscillator Stability.—It has been shown in Sec. 16-4 that free-running frequency stability of a high order is required of the local

oscillator. A similar type of stability is required of the coherent oscillator. The apparatus designed by S. Roberts¹ for measuring oscillator stability has two functions to perform. The first is the quantitative determination of the rate of change of frequency. Because a precise calculation of phase error requires the assumption that the modulating frequencies are small, the second function is to check the accuracy of this assumption. This is done by changing the time constant of the measuring circuit.

A block diagram of the apparatus used for testing local-oscillator stability is given as Fig. 16-36. The difference frequency of 70 kc/sec is obtained by mixing the power from two oscillators in a crystal rectifier. One of these oscillators can be a reference standard against which the

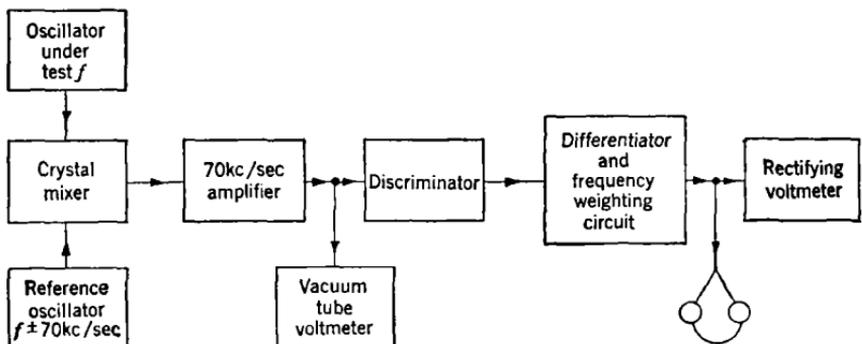


FIG. 16-36.—Block diagram of oscillator stability tester.

other is compared. The output signal from the mixer is amplified and applied to a frequency discriminator circuit. The voltage applied to the discriminator is monitored by means of a vacuum-tube voltmeter and is always adjusted to the same value. The voltage output of the discriminator is linearly related to the difference frequency. This discriminator output voltage is applied to a differentiating circuit consisting of a resistor and condenser. The average voltage across the resistor, proportional to the average rate of change of the difference frequency, is measured by a voltmeter. Headphones assist in the identification of the frequency of any modulation that may exist.

To measure the stability of a coherent oscillator, it is allowed to run freely. For the crystal mixer of Fig. 16-36 is substituted a 6SA7 mixer, whose third grid is connected to the output signal of the coherent oscillator. Signals from a quartz-controlled standard oscillator are applied to the first grid of the mixer. The difference frequency between the coherent oscillator and the standard is chosen as 70 kc/sec.

¹ *Loc. cit.*

Subclutter Visibility Measurement.—A block diagram of a subclutter visibility meter suitable for measuring the internal performance of most MTI systems is shown in Fig. 16-37. An i-f pulse originating either in the locking-pulse mixer or in a separate generator is modulated, delayed and injected into the receiver channel. This same pulse locks the coherent oscillator. The delayed pulse is modulated in phase and in

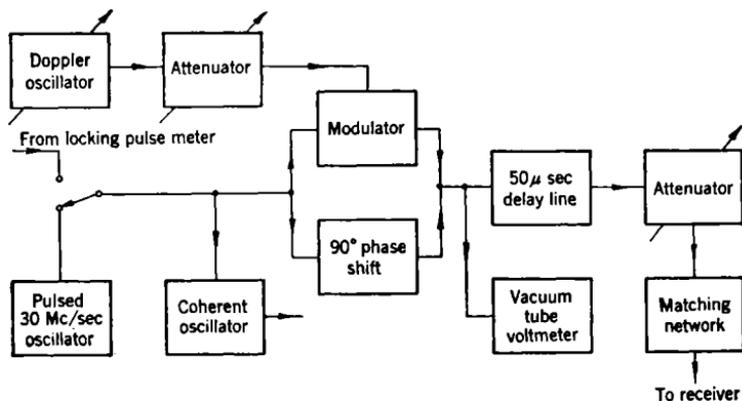


FIG. 16-37.—Subclutter visibility meter.

amplitude by a controlled amount at a controlled doppler frequency, so that it imitates a moving target in clutter. The minimum percentage modulation for threshold signal visibility is a measure of subclutter visibility. Although the simple modulation scheme shown does not provide an exact duplication of a moving-target echo, it is believed to be equivalent for test purposes. The meter does not measure the loss of subclutter visibility due to fluctuations in the clutter produced by scanning or the wind. It measures only the quantity K_3 in Eq. (28).

CHAPTER 17

RADAR RELAY

BY L. J. HAWORTH AND G. F. TAPE

INTRODUCTION

From the standpoint of the effectiveness with which a radar collects information, the location of its antenna is of supreme importance. In the use of the information much depends upon the location of the indicators. "Radar relay" is a means for separating these two components so that each can occupy the most favorable site or so that indicators can be operated at several places simultaneously. As the name implies, the radar data are transmitted from the source at which they are collected to some remote point by means of a radiation link.

17-1. The Uses of Radar Relay.—Control of aircraft in either military or civilian applications requires the review and filtering of a mass of information gathered from many sources, of which radar is only one. Control centers are, therefore, located at sites chosen for their operational convenience, whereas radar locations are chosen mainly for terrain and coverage reasons. Sometimes it is desirable to collect the data at great distances from one or more fixed land stations. Advantages are also gained by obtaining the data at an airborne site with its extended horizon, but displaying and using the data on the ground or on a ship. On the other hand, occasions arise in which an aircraft can usefully employ information collected from another site. In any of these cases, the possibility of multiple dissemination of the data to many points offers attractive possibilities.

Prior to the advent of radar relay such transmission was done by voice or not at all. Since the average operator can pass on only about five data per minute this method is far too slow for any rapidly changing complex situation, in addition to being rather inaccurate. Any really sophisticated use of radar data at remote points therefore demands the use of some sort of relay technique.

Two general types of situation arise: (1) those in which the data are transmitted between *fixed points* on land, in which case it is possible to use fixed, narrow-beam antennas at both ends of the relay link (or, if the information is to be broadcast from a single antenna, to use directive antennas at least on the various receivers); and (2) those in which one or both sites are *moving*, in which case the antennas must be either omni-

directional or controlled in direction. The first situation is by far the simpler from the technical standpoint since large antenna gains can be used in a very simple manner.

17-2. The Elements of Radar Relay.—One obvious method of relaying radar information is to televise an indicator screen at the radar site and to transmit the information by existing television means. This system leads to a loss both in signal-to-noise ratio and in resolution because persistent displays do not televise well. Furthermore, any single televised display would have to be a PPI presenting the maximum radar

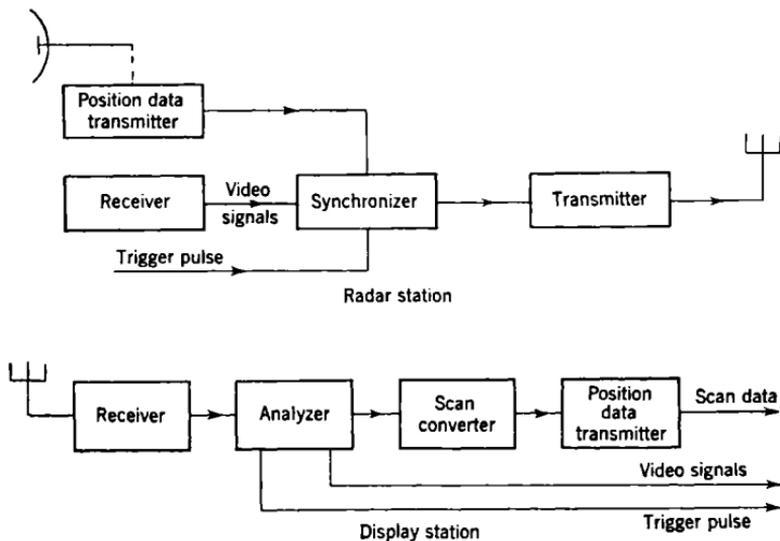


FIG. 17-1.—Elements of a simple radar relay system. In some cases electrical data from the analyzer can be used directly to generate the indicator sweeps.

range, with the result that much inherent resolution is immediately lost even though expanded displays might be used at the receiving end. The first, but not the second, of these difficulties might be overcome by storing the information on an Orthicon or other storage device rather than on a cathode-ray tube.

A far superior method is to transmit the original radar data in such a way that any desired displays can be produced at the receiving location in exactly the same way as can be done at the radar itself. To do this, it is necessary to provide at the receiving station: radar video signals, the modulator trigger (or the pulse itself), and a mechanical motion (or its electrical equivalent) that faithfully reproduces the motion of the scanner.

The elements of a system for transmitting this information are indicated in Fig. 17-1. The radar data are delivered from the set to a "synchronizer" which arranges them in proper form to modulate the

transmitter. At the receiving station, the receiver amplifies and demodulates the incoming signals and delivers the results to an "analyzer." The latter performs the necessary sorting into video signals, trigger pulse, and scanner data. The video and trigger are delivered immediately to the indicator system. The scanner data must usually be modified in form before being passed on either to the indicators for direct use in electrical display synthesis, or to the "scan converter." The scan converter uses these data to construct a duplicate of the scanner motion that can be used to drive a position-data transmitter associated with the indicators.

Since the requirements of the actual transmission and reception are very similar to those of television, slightly modified television transmitters and receivers can be used. Together with the antennas, etc., they will be referred to as the "radio-frequency" (r-f) equipment. The remaining components will be called the "terminal" equipment.

Except, perhaps, when microwave frequencies are used in the radio link, the ultimate limit of sensitivity is usually set by the degree of outside interference rather than by the inherent signal-to-noise ratio of the receiver. Many factors must be considered in trying to minimize the effects of this interference.

1. The strength of the desired signals at the receiver input terminals should be made as high as feasible compared with that of the interference. This is chiefly a matter of making proper choice of frequency, transmitter power, and antenna characteristics.
2. The data signals should, as far as possible, be made unlike the expected interference in signal characteristics, and every advantage should be taken of these differences in the receiving equipment.
3. In certain cases a favorable signal-to-interference ratio can be enhanced by techniques such as the use of wide deviation ratios with frequency modulation.

METHODS OF SCANNER-DATA TRANSMISSION

In even the simplest situation, the relay link must transmit the radar video signals, the trigger pulse, signals descriptive of the scanning, and sometimes range and angle markers.¹ In more complex cases some or all of these items may be duplicated, and additional data such as beacon or Identification of Friend or Foe, IFF, signals may be involved. One of the major problems of radar relay is to find economical methods of carrying all this information at one time. In the next few sections it will be assumed that a single transmitter is to relay one set of data of each

¹ Range markers need be transmitted only if some error is unavoidably present in the timing of the modulator trigger pulse. Angle markers, on the other hand, furnish a convenient check on the accuracy with which the scanner motion is followed.

variety. Complications introduced by multiple sets will be described later.

17-3. General Methods of Scanner-data Transmission.—Simultaneous transmission of video signals and range and angle markers is relatively simple. Marker signals need only be mixed with the video signals; no separation is performed at the receiver. Trigger pulses can also be mixed with the video since there is no conflict in time, but some method of separating them at the receiving station must be provided. If the pulses are transmitted at a higher power level than the video signals, the difference in voltage amplitude can be used as a criterion. This high-level transmission of pulses is usually done with amplitude-modulated r-f equipment, since the brevity of the pulses allows high pulse powers, with attendant signal-to-interference gain, to be obtained cheaply (Secs. 17-11, 17-12), but it is not feasible with frequency-modulated equipment. The alternative is to separate the signals in time, the video signals being excluded from the transmitter during an interval prior to and including the pulse. The trigger channel at the receiving station is blocked by an electronic switch except when the pulse is expected. Means for accomplishing both of these separation methods will be described later.

Relaying the scanner data is much more complex. It is not feasible to transmit the numerical value of an angle by a proportional amplitude or frequency modulation of an r-f carrier. It is possible to devise methods whereby certain functions (e.g., the sine and cosine) of the scan angle can be used to modulate two or more subcarriers¹ directly by means of slowly varying voltages; this has been done in the laboratory. It has been found far more effective to convey the information through the medium of periodic signals whose frequencies or repetition rates are several times those involved in the scanning and whose characteristics are in some way descriptive of it. Since scanning frequencies lie in the interval from zero to a few cycles per second, the scanner data signals usually have periodicities in the audio range.

The signals mentioned above can take either of two forms.

1. Continuous a-f signals can be used, the information being carried in terms of the amplitude, the frequency, or the phase with respect to a fixed reference. Up to the present time this technique has been very little used.
2. Pulse-timing techniques can be applied. Data can be transferred in terms of the frequency of a single train of pulses, the degree of

¹ A subcarrier is a sine wave, usually in the audio- or video-frequency range, which, after it has been modulated by the information-bearing signals, is used to modulate an r-f carrier.

staggering of two trains, or the lapse of time between a "basic" pulse which is one of a train and a second pulse occurring a controlled time later in the same cycle.

Omitting for the moment all questions of amplitude versus frequency modulation, and all problems of external interference, the choice of a data-transmission system involves three intricately related questions: (1) how to avoid interference with the video and trigger signals and mutual interference among the various data signals; (2) whether to use c-w or pulse methods; (3) which geometrical quantities among those descriptive of the scanner motion can best be chosen for transmission.

Since the video signals contain frequencies from nearly zero up to a few megacycles per second it is not feasible to separate scanner data signals and radar echo signals on a basis of their frequency components. This leaves the two alternatives of time-sharing within the radar pulse cycle, or the use of one or more subcarriers.

In the time-sharing method the scanner data are sent during the "rest" interval at the end of the radar cycle when the indicators are idle and the video signals are not useful. The interval can be occupied by as many signals as are necessary to carry the information. Pulse-timing techniques are usually used with time sharing since they are somewhat simpler and probably more accurate than those involving interrupted continuous waves. It is usually possible to allow some of the pulses to take part in the transmission of more than one piece of data. Often the modulator pulse itself is used as part of this timing system.

When the subcarrier method is adopted, it is customary to use continuous waves rather than pulses, partly for reasons of simplicity and partly because of smaller bandwidth requirements. When c-w signals are used, the various components of the scanner data can be distinguished from one another by sending each on a separate subcarrier, or by using a different audio frequency for each signal and transmitting them together on a single subcarrier.

The geometrical quantities transmitted can be chosen in various ways, depending on the application. The most important are:

1. Changes in the orientation of the scanner can be transmitted by means of a wave train in which one cycle represents an advance of the scanner through a given small increment of angle. If the scanner velocity is constant or changes only slowly, the wave train can be a sinusoid, or convertible into a sinusoid; hence a synchronous motor may be used in the scan converter. Some method of adjusting the initial phase of the converter relative to the scanner and of recognizing alignment must be provided. This general

method is satisfactory only for transmitting continuous rotation at a fairly constant rate.

2. The angle itself can be transmitted in terms of the relative phases of two sets of periodic signals which are usually either sinusoids or pulse trains (Sec. 17-5). Since this is a single-valued method, no zero-phase adjustment need be made. It can therefore be used for sector scanning, or for interrupted scanning. The data can be used to position a scan converter by means of a phase-sensitive mechanism, or they can be used directly to provide electrical information for such an indicator as a B-scope.
3. The values of the sine and the cosine of the scan angle can be transmitted in any of several ways (Sec. 17-7). Transmission of sine and cosine is also a single-valued method. The data can be used to control a scan converter or directly in electronic PPI's or related indicators.

17-4. Methods of Combating Interference.—Aside from providing the best possible signal-to-interference intensity ratio at the receiver input, the principal method of minimizing the effects of interference is to take advantage of differences between the desired and the undesired signals. The differences, which should deliberately be made large, can be exploited both by using them as a basis for excluding the undesired signals from the operating device and by making that device as insensitive as possible to interfering signals that are not excluded. Unwanted signals can be rejected by frequency discrimination; the bulk of the interference is excluded in this way. However, since it is always necessary to have a finite bandwidth to admit the necessary information, some interference is likely to get through to the analyzer.

In the case of the scanning data, it is often possible to protect against transient interference (or absence of signals) by exploiting electrical or mechanical inertia. Care must be taken that the inertia does not appreciably inhibit satisfactory response to scanning accelerations, a serious restriction if sector scanning is involved.

A second, and more promising, method of approach is to take advantage of approximate knowledge of what the real signals should do by excluding completely all information that does not closely agree with expectation, just as tuned circuits or filters reject signals outside their pass band. The knowledge on which to base the selection may be available a priori, or it may depend partly on a memory of what has happened in the immediate past.

As an example of such methods consider an information-bearing pulse. Very similar pulses with the same frequency components as the signals are likely to be present in the interference and of course cannot be

excluded by the receiver. If, however, these interfering pulses are not too numerous and strong, it is possible to provide almost complete protection against them by using a "coded" signal to represent the pulse, or by excluding all pulses that do not come within the narrow time interval which follows the last useful pulse by the known repetition period, or by doing both.

The most usual type of coding consists of a group of pulses succeeding each other by precise, unequal time intervals. The responsive circuit at the receiver is arranged to recognize only a group with precisely these spacings. A similar combination is extremely unlikely to occur in the interference.

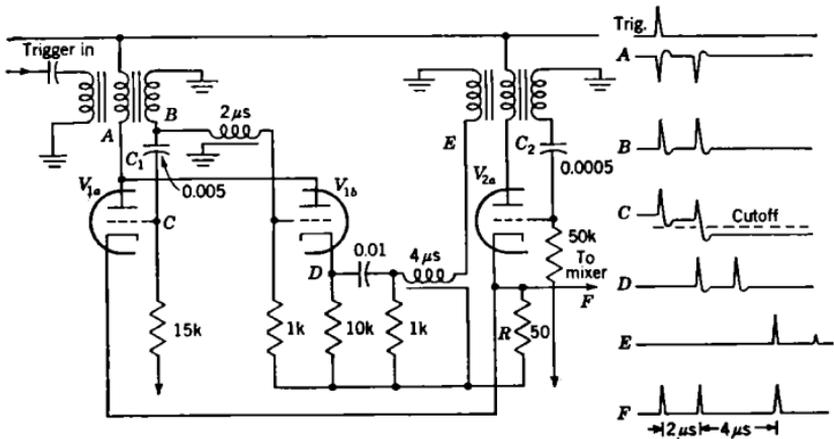


FIG. 17-2.—Triple-pulse coder.

A common type of coding and a simple method of producing it are illustrated in Fig. 17-2. When the blocking oscillator V_{1a} is triggered by an incoming signal, it produces a pulse of less than $1\text{-}\mu\text{sec}$ duration. The sharp positive pulse at B is passed through the $2\text{-}\mu\text{sec}$ delay line to the grid of V_{1b} . The negative pulse on the plate of V_{1b} passes back to V_{1a} and retriggers it to produce a second pulse delayed by $2\ \mu\text{sec}$ (Waveforms A and B). The cumulative effect of the two actions charges C_1 sufficiently to cut off V_{1a} for a time so that further regeneration does not take place. The first pulse is driven down the $4\text{-}\mu\text{sec}$ delay line by V_{1b} and triggers V_{2a} after a total delay of $6\ \mu\text{sec}$. The firing of V_{2a} charges C_2 sufficiently to cut off the tube so that the second pulse, which arrives $2\ \mu\text{sec}$ later, cannot trigger it. Thus three pulses occurring at 0, 2, and $6\ \mu\text{sec}$ appear across resistance R and pass to the mixer to be combined with the video and other signals.

Other combinations of time delay can, of course, be used. The individual delays should not only be unequal, but within reason each should be great enough to prevent radar or other interference pulses from

bracketing pulses. They are not, however, usually made greater than a few microseconds because of the bulkiness of longer delay lines.

A method of decoding the three pulses is shown in the upper part of Fig. 17-3. Signals from the receiver are differentiated by the grid circuit of amplifier V_{1a} so that blocks of signals are not passed. If the incoming pulses are large enough, the tube can be biased past cutoff in order to exclude signals at lower levels (as, for example, when amplitude is to be used as a basis of discrimination between pulses and video signals). Weak interference can also be excluded in this way. V_2 acts as a limiting

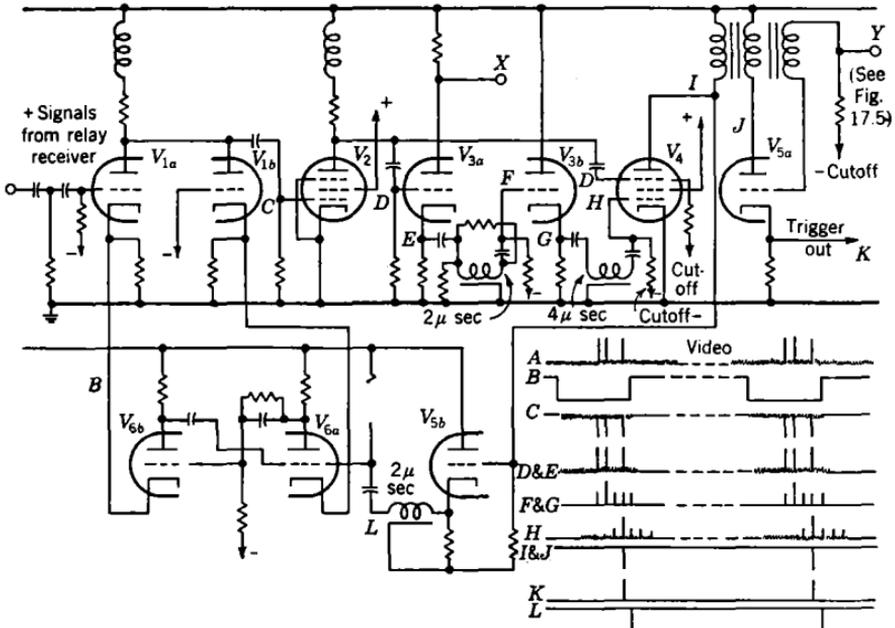


FIG. 17-3.—Interference blander and triple-pulse decoder.

amplifier. After passing through the cathode follower V_{3a} , each pulse arrives twice at the grid of V_{3b} , once with no delay and once with a 2- μ sec delay. Six pulses will therefore arrive at V_{3b} , and two of these, the first delayed pulse and the second undelayed pulse, will be in coincidence, as the waveform diagram shows. Thus five pulses, the second of which has double amplitude, will arrive at the grid of V_4 after a further delay of 4 μ sec. V_4 is so biased that only the large pulse lifts the grid past cutoff. Since the total delay of this pulse is 6 μ sec, its arrival will coincide with that of the third undelayed pulse on the suppressor of V_4 . At this time, and this time only, the plate of V_4 receives a signal and fires V_{5a} . This event cannot occur if any of the pulses is missing; therefore interference can produce a result only when three spurious pulses occur with approximately the correct spacing. Differentiation at the input circuit prevents

long signals from straddling two or more pulse intervals and thus providing false coincidences. It cannot, of course, prevent such blocks of signals from saturating the receiver and excluding the desired pulses.

A brief analysis of the effectiveness of the coding is worth while. Periodic pulses cannot give a response from V_4 unless their frequency is greater than the reciprocal of the delay times in the coding. Random pulses or noise do, however, have a finite probability of initiating a response. To calculate this probability suppose that there are n random pulses per second and let τ be their length. Then the fractional time during which the grid of V_{3b} is receiving signals through the delay line is $n\tau$. The probability that a signal from a given undelayed pulse will overlap that from the delayed pulses is therefore $2n\tau$, the factor 2 entering on the assumption that any overlap at all will produce a result. Thus the number of reinforced signals reaching the grid of V_4 each second is $2n^2\tau$. Since the probability that one of these will coincide with a pulse on the suppressor grid is $n\tau$,¹ the total number of pulses triggering V_4 each second is $2n^3\tau^2$. This is to be compared with the number n of interfering signals originally present, the ratio of improvement being

$$n/2n^3\tau^2 = 1/2n^2\tau^2.$$

If, for example, $\tau = 1 \mu\text{sec}$ and $n = 10,000$, this ratio is 5000. The ratio improves for smaller n and vice versa.

If the signal being relayed is the firing time of the modulator, the firing should, if possible, coincide with the transmission of the third pulse. If the modulator can be triggered with sufficient accuracy, the output connection indicated in Fig. 17-2 can be used for this purpose. If, on the other hand, the modulator is self-synchronous, its pulse must be used to trigger the coder. The synchronizing pulse at the receiving station, which must await the arrival of the third pulse, will be $6 \mu\text{sec}$ late. If this error cannot be tolerated, the video signals must be passed through a $6\text{-}\mu\text{sec}$ delay line before they reach the indicators.

Figure 17-3 also illustrates one method for excluding all signals except during an interval surrounding the expected time of arrival of the desired signal. The coincidence pulse from the plate of V_4 is passed, after buffering in V_{5b} , through a $2\text{-}\mu\text{sec}$ delay line, and is used to trigger the flip-flop circuit formed by V_{6a} and V_{6b} (Sec. 13-7), in which circuit the latter tube is normally off. After the flip-flop has been triggered, the current through V_{6b} lifts the cathode of V_{1a} , cutting this tube off by an amount greater than the signal level. The flip-flop is timed to reverse shortly before the next desired signal is expected. V_{1b} prevents an appreciable pedestal from appearing on the common plate circuit since it is turned on when V_{1a} is turned off and vice versa. If the signals

¹ The factor 2 does not enter again.

are large enough to allow V_{1a} to rest below cutoff even when on, V_{1b} is unnecessary. Alternative types of video switches are described in Sec. 13.9. The sensitive time must be long enough to allow for all uncertainties in the periodicity of the incoming signal and for changes in the flip-flop circuit timing.

This method can also be used to distinguish between different signals which have been transmitted on a time-sharing basis: for example, to separate pulses from video signals. If the coincidence circuit is to be used to decode more than one set of pulses, the flip-flop can be triggered by the last set or by a pulse from a delay circuit which spans the signal interval. More often, a train of switches is necessary to separate the various pulses from one another after decoding so that no switch is necessary at the input circuit. Somewhat more elegant methods which allow narrower open intervals can be used in certain special cases in which the opening of the switch can be controlled from a sequence of events within the cycle. Various arrangements will appear in later sections in connection with specific methods of data transmission.

17.5. The Method of Incremental Angle.—This section and the next three will describe various specific methods of relaying the scanner information, the basis of classification being the geometrical quantities used.

As stated in Sec. 17.2, the method of angular increments usually involves, as a final stage, a synchronous motor driven by a sinusoid. Three methods of relaying the necessary data have been used.

In the first method, the sinusoid itself may be transmitted directly on a subcarrier. The extra expenditure of power involved is usually not justifiable for this rather inflexible method, especially since the pulse methods are extremely simple.

As a second method, the modulator pulses themselves can be used to represent the increments of scanner angle¹ if the scanner motion can be made sufficiently constant to control the modulator triggering satisfactorily. Such a system is represented in Fig. 17.4.² Some form of signal generator—usually electromagnetic—geared to the scanner produces periodic signals of frequency proportional to the scanner velocity and suitable for the pulse recurrence frequency. These signals control a blocking oscillator or some other device to produce sharp pulses which are then coded for transmission over the relay link. The third pulse of the code (from point E of Fig. 17.2) is used for the modulator trigger.

Some method must be provided for separating pulses from video

¹ This is, of course, an example of time sharing in which two of the signals coincide.

² In all the diagrams of this chapter the individual blocks are functional and are intended to include proper input and output circuits including amplifiers, cathode followers, blocking oscillators, etc.

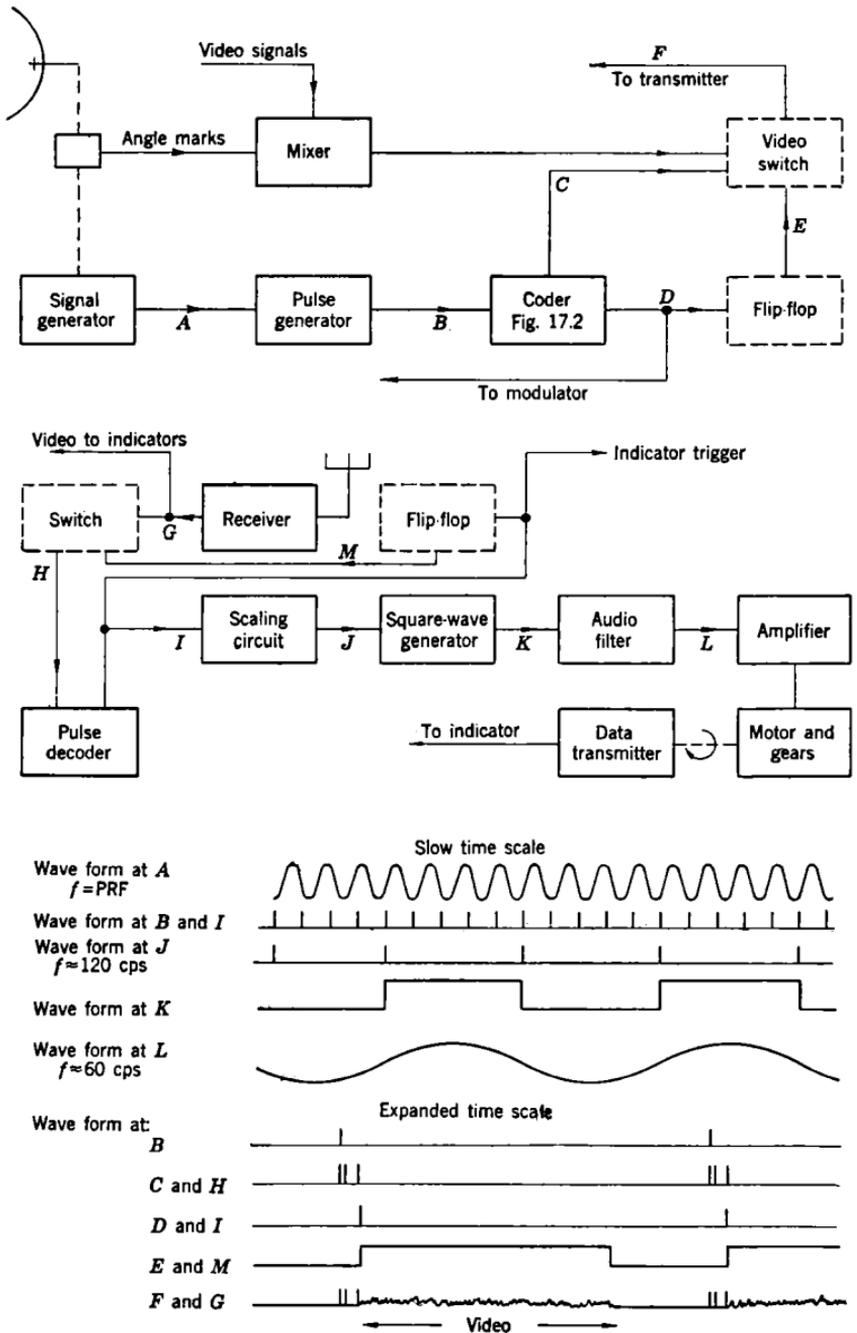


FIG. 17-4.—Simple incremental-angle synchronization.

signals at the receiving station since otherwise video signals might occasionally produce a spurious trigger. If the transmitter provides higher power in the pulses than in the video signals, amplitude selection can be used, and the coded pulses can simply be passed to the transmitter. Otherwise, the coded pulses must be distinguished in time, and the extra equipment indicated by the dotted boxes must be used. The video switch (Sec. 13-9) is arranged to pass video signals only when the flip-flop (Sec. 13-7) is on, and pulses only when it is off. The flip-flop is fired by the modulator trigger, so that video signals are passed until shortly before the next cycle, at which time the flip-flop opens the switch for pulses until the modulator pulse occurs again.

At the receiving station, signals from the receiver pass to a switch which is open when a pulse code is expected. Following this switch is a decoder similar to that of Fig. 17-3. The resulting single pulses serve to trigger the indicators, to control the signal-selecting switch, and to provide the rotation. The switching action is similar to that of V_{1a} , V_{1b} , V_{6a} , and V_{6b} of Fig. 17-3, the length of the flip-flop being slightly less than that of the flip-flop at the transmitter so that video signals are always excluded from the decoder.

To produce the mechanical rotation, the decoded pulses first actuate a scaling circuit, such as that of Fig. 13-20, which reduces their frequency to twice that appropriate to a synchronous motor. The resulting pulses, by triggering a scale-of-two circuit similar to that of Fig. 13-16*b*, produce a symmetrical square wave. A broadband a-f filter removes the higher harmonics, leaving a sine wave at the fundamental frequency which, after amplification, powers a synchronous motor that drives a data transmitter. Proper initial phasing of the data-transmitter shaft can be made by methods analogous to those of display-sector selection (Chap. 13); alternatively, a controlled phase shifter can be inserted between the a-f filter and the amplifier. A convenient index for use in this alignment can be provided by transmitting one or more angle markers along with the video signals, as illustrated in Fig. 17-4.

This method is restricted in its use. The requirement of a continuous scan at a nearly constant speed mentioned in Sec. 17-3 is made even more rigorous by the synchronization with the modulator, which must usually operate at a definite PRF. Any requirement for variation in the scanning rate or the PRF introduces serious complexities because of the fixed relation between these two quantities. Another drawback is that rephasing must be done whenever the signals are interrupted or seriously interfered with.

More flexibility is provided by a third means for relaying incremental angle data. If the scanner and the modulator cannot be synchronized but the scanning is nevertheless reasonably uniform, the periodic signal can be

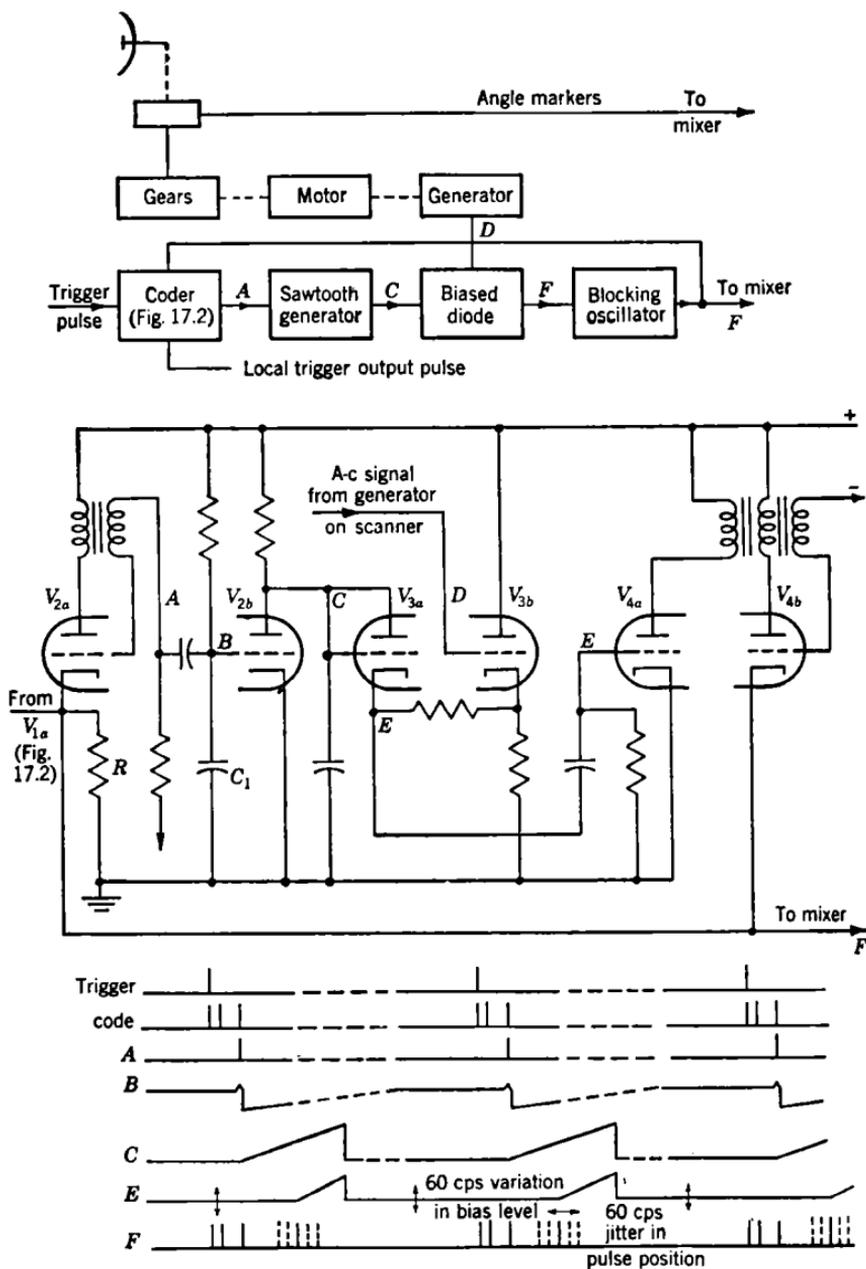


Fig. 17-5a.—Jittered-pulse system, transmitting station.

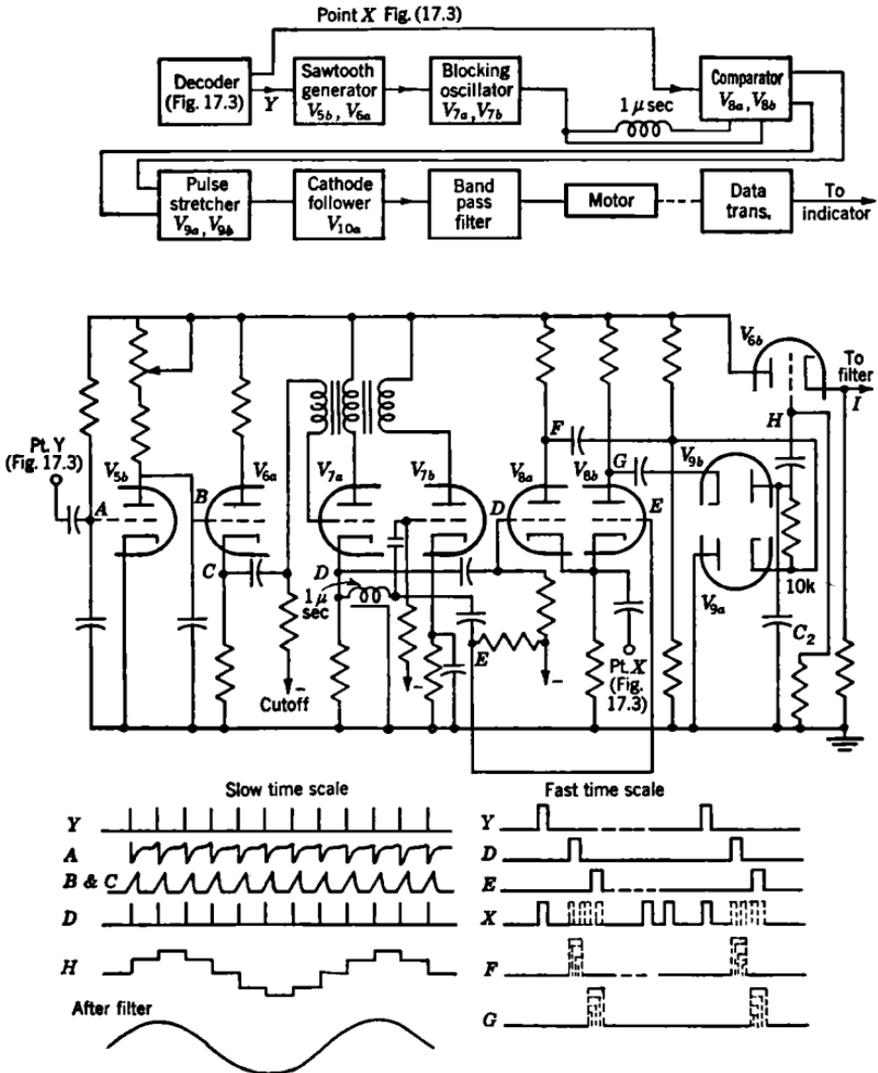


Fig. 17.5b.—Jittered-pulse system, receiving station.

expressed by a sinusoidal variation in the relative timing of a pair of pulses appearing in each pulse cycle. A method for accomplishing this is illustrated in Fig. 17.5. The basic pulse is initiated by an external trigger and coded by the circuit of Fig. 17.2. The third pulse, originating in V_{2a} of that figure (repeated in Fig. 17.5a), initiates the action of a time-delay circuit consisting of a sawtooth generator V_{2b} , a biased diode V_{3a} , an amplifier V_{4a} , and a blocking oscillator V_{4b} . The need for a square-wave

generator to switch V_{2b} is avoided by the action of the positive pulse from V_{2a} . During the pulse the grid current of V_{2b} charges C_1 sufficiently to block off the tube long enough to allow generation of the desired sawtooth. The a-c voltage from a generator on the scanner acts through the cathode follower V_{3b} to vary the bias of the diode sinusoidally and therefore vary the delay by steps which approximate a sinusoidal variation of about one microsecond amplitude. The mean value is set at 5 to 10 μ sec. The delayed pulse is mixed with the coding for the basic pulse on the common cathode resistor R and sent to the master mixer. It is not necessary to code the azimuth pulse since its function is merely to transmit the sinusoidal frequency; any interference strong enough to mask this pulse would render the whole system useless in any case.

At the receiving station the basic pulse is decoded by the coincidence circuit of Fig. 17-3.¹ The resulting pulse, taken from point Y of that diagram, blocks the grid of V_{5b} and initiates the action of the delay circuit comprising the sawtooth generator V_{5b} , the cathode follower V_{6a} , and the blocking oscillator V_{7a} . The tube V_{7a} is delayed in firing by its excessive bias. When V_{7a} fires, its cathode drives the 1- μ sec delay line, and the delayed pulse returns through V_{7b} to turn off the blocking oscillator. Thus successive 1- μ sec pulses, adjacent in time, appear at the grids of V_{8b} and V_{8a} respectively. Their boundary time is made coincident with the mean time of arrival of the azimuth pulse by adjustment of the slope of the sawtooth in the delay circuit. The amplified pulse train from point X (plate of V_{3a}) of Fig. 17-3 is applied to the cathodes of V_{8a} and V_{8b} . Because of the bias the tubes can conduct only during the coincidence of the azimuth pulse with the pulses on the grids. Condenser C_2 is charged negatively by the signal from V_{8b} acting through diode V_{9b} , and positively by the signal from V_{8a} which forces negative charge to ground through V_{9a} . If the coincidence time is equal on V_{8a} and V_{8b} , the net result is zero. However, as the azimuth pulse moves back and forth in time with respect to the switching pulses, the coincidence times vary in an out-of-phase manner so that the potential of C_2 has a stepwise variation with a sinusoidal envelope (Waveform I). This alternating signal is passed through cathode follower V_{6b} and a bandpass filter in order to produce the desired sinusoid for driving the synchronous motor.

In order to minimize the number of transmitted pulses, the basic pulse can be related to the modulator pulse in any of several ways. If the first few microseconds of the radar cycle are not too important and the modulator is not self-synchronous, its trigger can be the third pulse of the coded group. No range error will result, but the azimuth pulse will appear on the indicators at the receiving station at a range corresponding

¹ The pulses can be separated from the video signals by amplitude selection, or by time separation as in Fig. 17-4.

to its delay time. If the modulator is self-synchronous, its firing time must coincide with the first coded pulse. A range error in the indicators results unless a compensating delay line is used in the video channel at one station or the other.

If the appearance of the azimuth pulse at short range on the indicators is not tolerable, a triggered modulator can be fired after the azimuth pulse, the delay from the third coded pulse being precisely fixed. At the receiver a similar delay initiated by the decoded basic pulse can be inserted ahead of the indicators. Correct adjustment of this delay can easily be made by observing the transmitter pulse appearing with the video signals.

No analogous method of azimuth-pulse removal exists for a self-synchronous modulator. It is usually satisfactory to remove the pulse by anticoincidence methods, even though a "hole" is left in the video signal train. Otherwise, it will be necessary to introduce between the modulator pulse and the basic pulse a delay which is longer than the useful video-signal interval and to transmit both pulses.¹

17-6. The Phase-shift Method.—If a sine wave is passed through a linear full-wave phase-shifting device connected directly to the scanner, the resulting phase shift is numerically equal to the scan angle measured from the position of zero phase shift. By transmitting the phase-shifted wave together with a reference signal of commensurate frequency and fixed phase, it is possible to use the relative phases as data from which to reconstruct the scanner angle.

C-w Methods.—A method of accomplishing this by use of c-w signals is illustrated in Fig. 17-6. Separate subcarriers f_1 and f_2 are respectively modulated by the phase-shifted sinusoid and by the reference signal, which is the same sinusoid without phase shift. At the receiving end the video signals and each of the subcarriers are separated by appropriate filters and the phase-shifted and the reference sinusoids are obtained by suitable demodulators. The reference wave is then passed through a phase shifter similar to that on the scanner. This phase shifter is driven by a servomechanism whose error signal is the output of a circuit which compares the phase of the shifted wave and that of the reference wave. A data transmitter geared to the phase shifter provides proper information to the indicators.

The above method is very expensive because of the two subcarriers. Fig. 17-7 illustrates an alternative c-w method which avoids this difficulty. Here the phase-shifted signal is a harmonic of the reference signal and is derived from it by a frequency multiplier. Thus the two signals can be transmitted without confusion on the same subcarrier. At the receiving

¹ It is not satisfactory to use such a long delay between the basic pulse and the azimuth pulse since small percentages of "jitter" or drift become too important.

station the two are separated by appropriate a-f filters, and the reference signal is passed through a frequency multiplier. The resulting sinusoid is then passed through a phase shifter controlled by a phase-sensitive servomechanism as before. Although this method is simple, it places

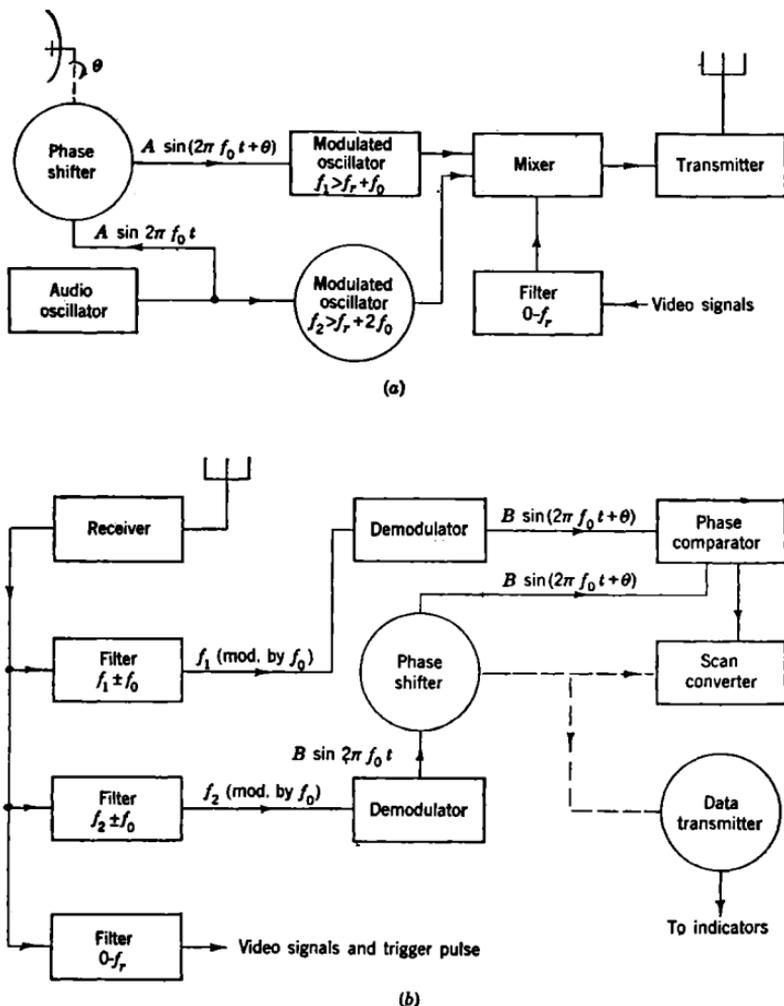


FIG. 17-6.—Phase-shift data transmission using two c-w subcarriers.

severe requirements on the audio filters. Because of their harmonic relationship, the sinusoids must be kept extremely pure, and furthermore any relative phase shifts introduced by the filters (or other circuit components) must be extremely constant with time. This is very difficult to assure in a-f filters if they are subjected to changing temperatures.

In the above systems any one of the data-bearing signals can be conveniently used for automatic gain control purposes.

Pulse Methods.—The pulse methods analogous to those just discussed would consist of the use of two continuous trains of pulses derived from sinusoids of equal or commensurate frequencies, one being shifted in phase with respect to the other. This would require the use of at least one subcarrier with wide sidebands, and therefore would be expensive in the

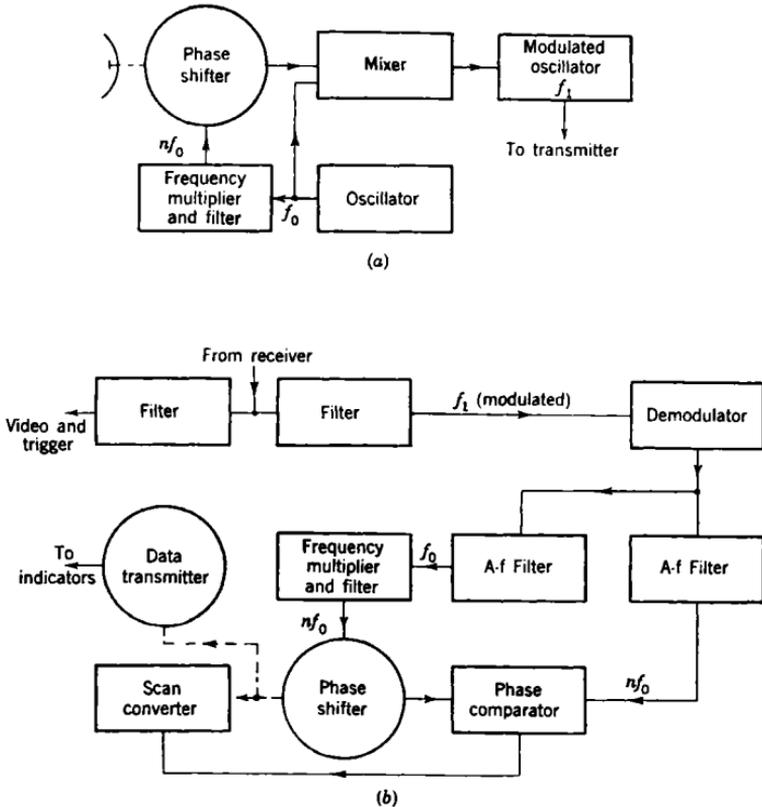
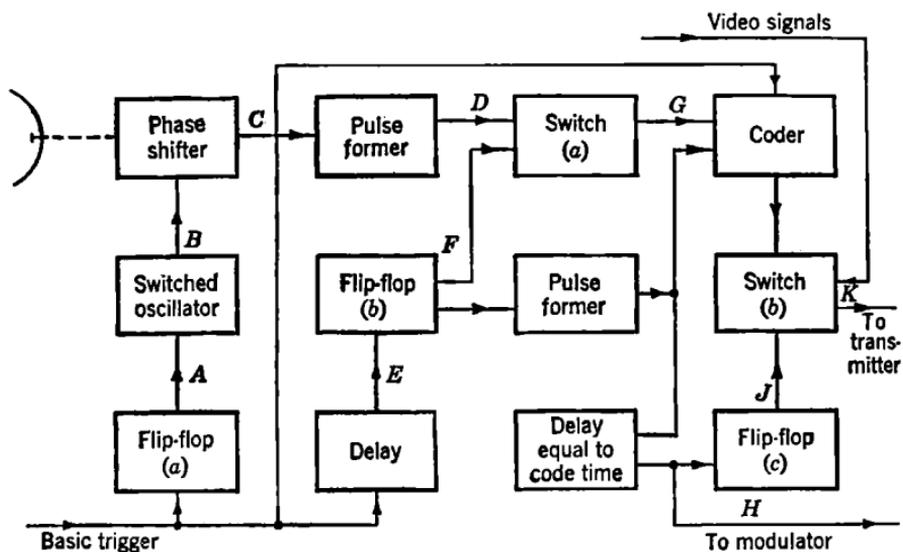


FIG. 17-7.—Phase-shift data transmission using one c-w subcarrier.

transmitter and receiver design. An alternative would be to transmit the pulses on the video carrier during those intervals when the echo signals are not useful. Since the difficulty remains that the two sets of pulses must “ride through” each other, some method would have to be found for distinguishing the two trains. This difficulty can be avoided by transmitting one master pulse and one train of “phase-shifted” pulses on each radar cycle and, at the receiving station, constructing the reference train by shock-exciting an oscillator with the master pulse.



Waveforms

Basic trigger

A

B Interrupted sinusoid

C Phase-shifted sinusoid

D Phase-shifted pulses

E Delayed trigger

F Switching voltage

G Selected pulse train

H Modulator trigger

I Coded pulses

Basic trigger

(contracted time scale)

J Video switch waveform

K Output signals

Pulses Video signal

Fig. 17-8a ---Phase-shifted pulses transmitting station.

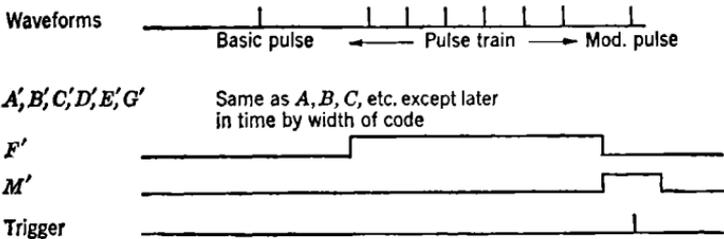
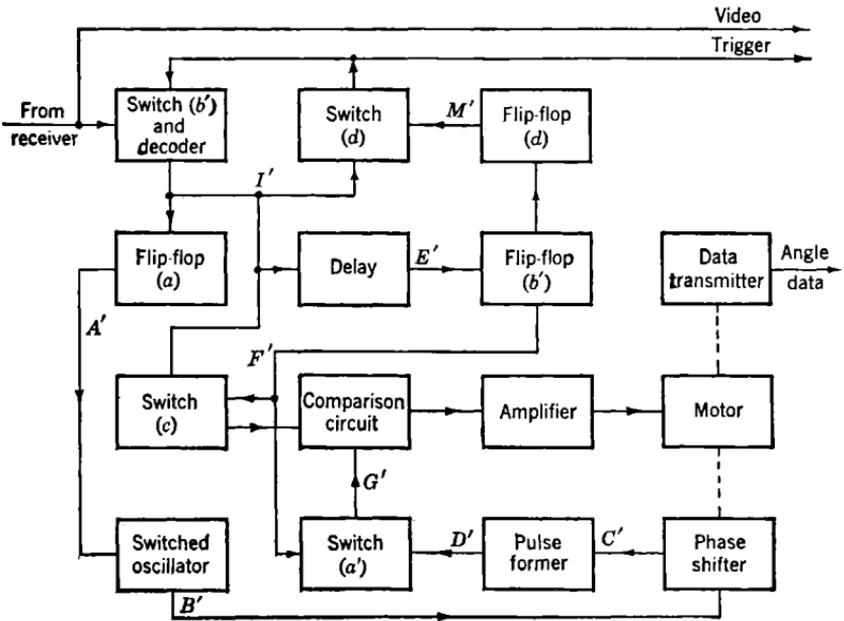


FIG. 17-8b.—Phase-shifted pulses, receiving station.

The method is illustrated in Fig. 17-8. At the transmitter the master pulse triggers a square-wave generator (flip-flop *a*) which switches a shock-excited oscillator (Fig. 13-41). The oscillation *B* is passed through a phase shifter geared to the scanner, and the resulting wave *C* produces a pulse train *D*.

In order to avoid early transients which might lead to confusion with the master pulse (especially after coding) the first two or three pulses in the train are excluded by switch *a*. This switch is controlled by a square wave from flip-flop *b*; it is arranged to turn on late enough to exclude the required number of pulses and to remain on until after the end of the pulse

train. The remainder of the circuit is concerned with coding and mixing the various signals and providing synchronization with the modulator (which has been assumed to be triggerable). The basic pulse and the pulse train are passed through the coder, as is a pulse formed at the end of flip-flop b . This last pulse, delayed by a time equal to that of the coding, serves as the modulator trigger. In order to separate the pulse signals from the video signals at the receiver, if amplitude selection cannot be used, the two are passed alternately through video switch b , which is controlled by flip-flop c . In its normal position, the latter causes the switch to pass signals from the coders. When the flip-flop is triggered by the pulse to the modulator, the switch is reversed and video signals are passed until the flip-flop spontaneously returns to its initial condition, shortly before the next basic pulse, and opens the channel to pulses again. If the transmitter is such that the pulses can be transmitted at several times the level of the video signals, this switching need not be done since amplitude discrimination can then be used at the receiver.

At the receiving station, the various signals must be separated and the data abstracted and put into usable form. The signals from the receiver are received by a switching and decoding circuit similar to that of Fig. 17-3 except for the source of the input signal to V_{5b} , whose function will be explained later. For the moment assume that the switch is open. When the basic pulse is decoded it starts a chain of events through flip-flop a , the switched oscillator, the phase shifter, and the pulse former on the one hand; and another chain through the delay circuit and flip-flop b' on the other. Both culminate in switch a' and produce at its output terminal G' a train of pulses like that at point G at the transmitter. Meanwhile, the transmitted pulse train is decoded and passed through switch c , whose purpose is to exclude the basic pulse and the modulator pulse. The two pulse trains are brought together in a "comparison" circuit which produces a polarized error voltage if they do not coincide. (If the two pulse trains have slightly different frequencies, because of slight differences in the oscillators, the error voltage will refer to their "centers of gravity.") The amplified error voltage controls a motor which turns the phase shifter in such a way that the error voltage is kept very small,¹ and the phase shifter and the data transmitter rotate in synchronism with the radar scanner. The modulator pulse is selected by switch d which is activated by flip-flop d . The latter is triggered at the moment of recovery of flip-flop b' and endures until after the arrival time of the modulator pulse. In addition to its function at the indicators this pulse also triggers the flip-flop (V_{6a} , and V_{6b} , Fig. 17-2) controlling the switch (V_{1a} and V_{1b} , Fig. 17-2) ahead of the decoder. When this switch is

¹ See Vol. 20 of this series for details of the comparison circuit and motor drive.

closed all signals (video and interference) are excluded until the flip-flop returns to its stable condition shortly before the next basic pulse.

17-7. General Methods of Relaying Sine and Cosine.—A complete description of an angle can be given by expressing its sine and its cosine in the same units. Probably the greatest variety of actual and proposed systems for data transmission have involved the relaying of these quantities in one way or another.

If sine and cosine of the scan angle are transmitted, these data can be used for PPI synthesis either with or without mechanical motion. Since

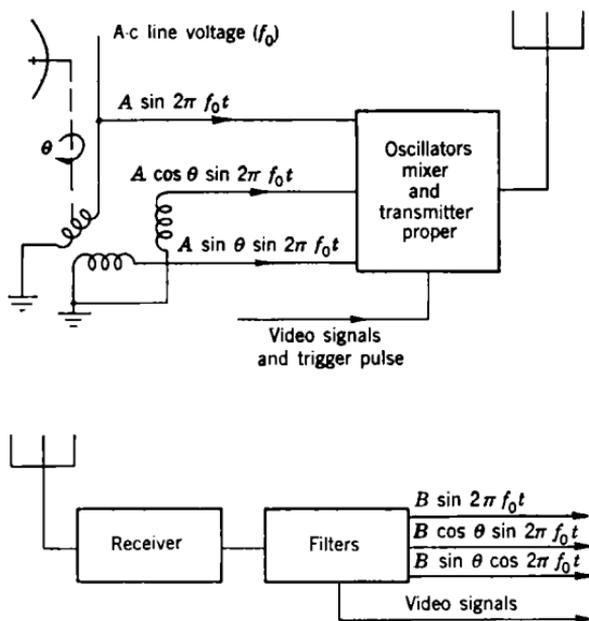


FIG. 17-9.—Hypothetical method of transmitting synchro data by c-w.

the method of synthesis has an important bearing on the choice of a relay method, a brief discussion of the use of the data will be given. Two general methods are possible: mechanical duplication of the scanner motion at the receiving station (as in the previous cases), or use of the sine and cosine voltages to produce the two necessary PPI range-sweep "components."

Derivation of mechanical motion from sine and cosine information requires the use of a servomechanism. Practically speaking, it is necessary, in order to reproduce the mechanical motion, to provide a-c voltages proportional to sine and cosine, as well as a voltage of constant amplitude for reference purposes. All voltages must have the same frequency and phase. From the terminal-equipment standpoint these voltages could

be obtained most simply by exciting a synchro on the radar scanner and relaying each of the resulting voltage "components" together with a constant reference voltage (Fig. 17-9). This method, which would require the use of multiple subcarriers, and would thus involve excessive bandwidth, complex multiple modulation and filtering, or both, has never been used. Two component voltages and one reference voltage can be provided at different a-c frequencies, but this arrangement is not suitable for use with a servomechanism. It can, however, be used to produce slowly varying sine and cosine voltages. These can then be used to modulate alternating current of proper frequency to provide the required voltage components for a synchro-controlled servomechanism. The method

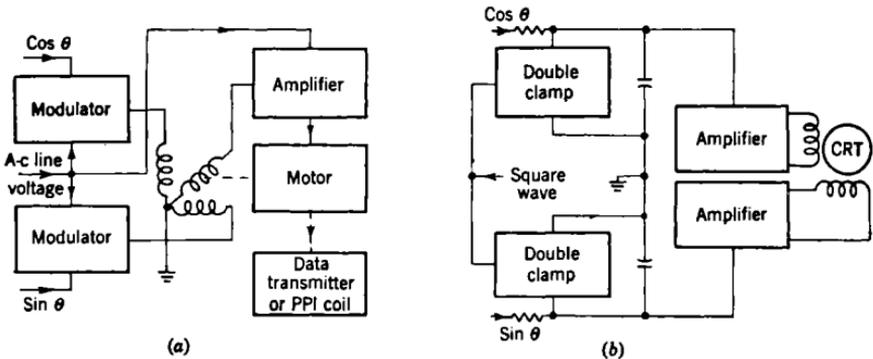


FIG. 17-10.—Use of slowly varying sine and cosine voltages: (a) production of mechanical motion; (b) fixed-coil PPI synthesis.

is illustrated in Fig. 17-10a. It has been successfully used in connection with various types of data transmission resulting in slowly varying sine and cosine voltages.

In contrast to the reproduction of mechanical motion, the fixed-coil PPI can best use slowly varying polarized direct current.¹ Since alternating current can easily be rectified to give such voltages, any of the methods is applicable to this case. As shown in Fig. 17-10b, the sweep components are produced by sawtooth generators using the signal voltages as their source potentials so that the "sawteeth" are modulated in the proper manner as the scanner rotates (Sec. 13-16). Through the provision of low-impedance or multiple-output circuits for the sine and cosine voltages, as many indicators as desired can be used. The sweep length for each can be chosen independently of those of the others.

¹ It is possible, of course, to relay the sweep components themselves. However, a separate subcarrier would be required for each. Furthermore, each "sawtooth" must endure for the time of the longest desired sweep, and fast sweeps must be achieved by amplification of a short segment. Errors would be greatly magnified in such a process.

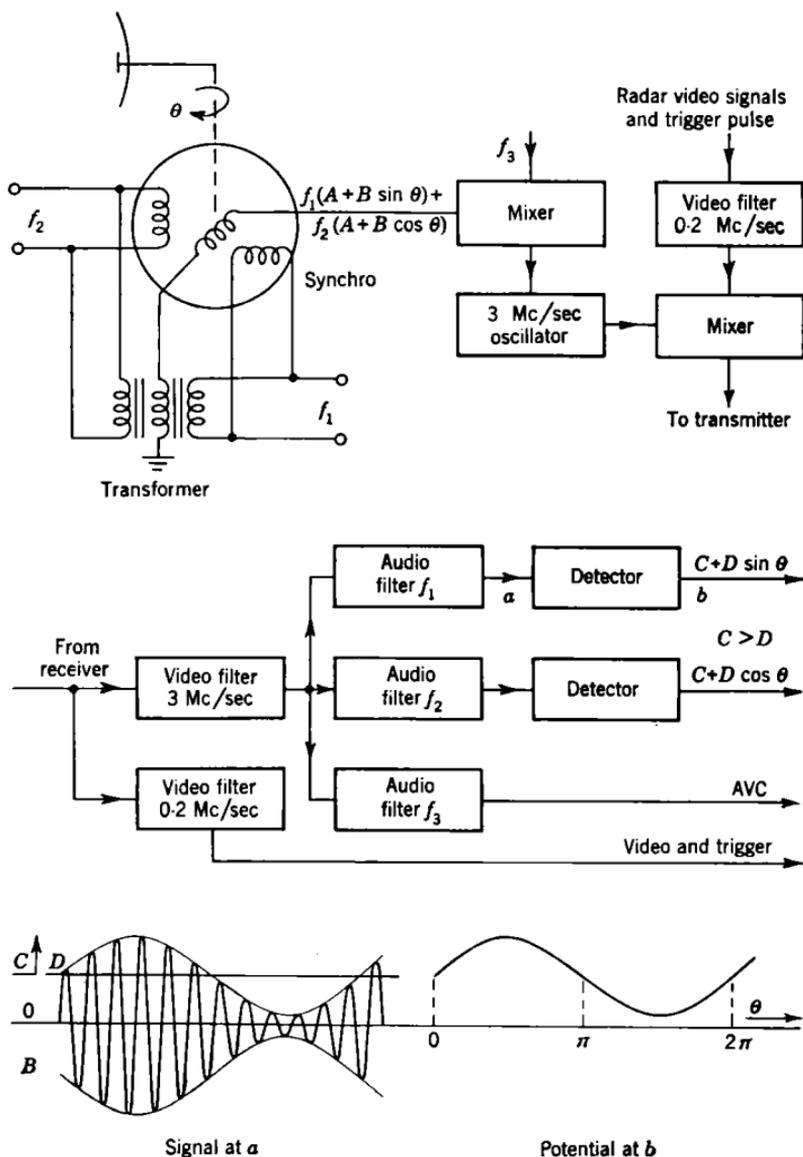


FIG. 17-11.—Method of relaying sine and cosine by c-w on one r-f carrier.

At least two methods of supplying slowly varying sine and cosine voltages have been successfully applied. One is a c-w method; the other involves pulse timing.

A C-w Method.—In order to allow transmission of all the scanner data on one r-f subcarrier, each item of information can use a different audio frequency, the subsequent sorting being done by a-f filters. Figure 17-11

illustrates such a method. The stators of the scanner synchro are excited by equal-amplitude signals of frequencies f_1 and f_2 , as are the two equal primaries of a transformer whose secondary is in series with the synchro rotor. The voltage across the combination is then

$$f_1(A + B \sin \theta) + f_2(A + B \cos \theta),$$

where f_1 and f_2 represent the audio-frequency input voltages and A and B are constants. Since A is greater than B , the amplitude terms are always positive. This voltage is mixed with a constant-amplitude a-f signal of frequency f_3 (which is used for AGC at the receiver), and the combination is used to modulate a subcarrier of somewhat higher frequency than the maximum required for video signals. The modulated carrier, the video signals, and the trigger pulse are then mixed and passed on to the transmitter proper. At the receiving station, the various component signals are separated by appropriate filters, as indicated in Fig. 17-11, and the sine and cosine signals passed through detectors. The resulting voltages contain a d-c component of magnitude C (with respect to the detector bias point) and a varying component of amplitude D , the sum being always positive. The d-c component can be prevented from affecting the final device by relating its bias level properly to that of the detector. The scheme of adding a constant a-f component to the signal is intended to preserve the phase sense of the components varying sinusoidally with the scanner rotation; in this case it is considerably simpler than transmitting a reference signal to be used in keying a phase-sensitive rectifier.

Pulse-timing Methods.—The simplest method of using pulse-timing techniques to relay sine and cosine information can be understood by referring to the timing diagram of Fig. 17-12. A basic pulse occurs once each radar cycle. The delay of a second pulse is varied with respect to it in accordance with the expression $A + B \sin \theta$, where A must be greater than B in order that the delay shall never become negative. A third pulse is delayed with respect to the "sine" pulse by an amount proportional to $A + B \cos \theta$. In the interpretation of the data it is only necessary to provide for each function a circuit that will develop a voltage proportional to the time lapse between the members of a particular pulse pair.

The essentials of such an equipment are illustrated in Fig. 17-12. Slowly varying sine and cosine potentials are furnished by a data transmitter (which can be either a d-c excited sine-cosine potentiometer or a two-phase synchro whose output signals are rectified in a phase-sensitive manner). Each of these potentials controls the operation of a linear delay circuit. The sine delay circuit is triggered by the basic pulse, and the cosine circuit is triggered by the sine pulse produced by the sine delay circuit. Each gives a finite delay A when the controlling voltage is zero

The resulting pulses are mixed with the modulator pulse and coded before being combined with the video signals preparatory to transmission.

At the receiving station, the pulses are decoded and sorted. Each of the pulse pairs (basic-plus-sine and sine-plus-cosine) is fed to an automatic range-tracking circuit which develops the required voltage as indicated in Fig. 17-12. The d-c components can be removed in the way explained in connection with Fig. 17-11.

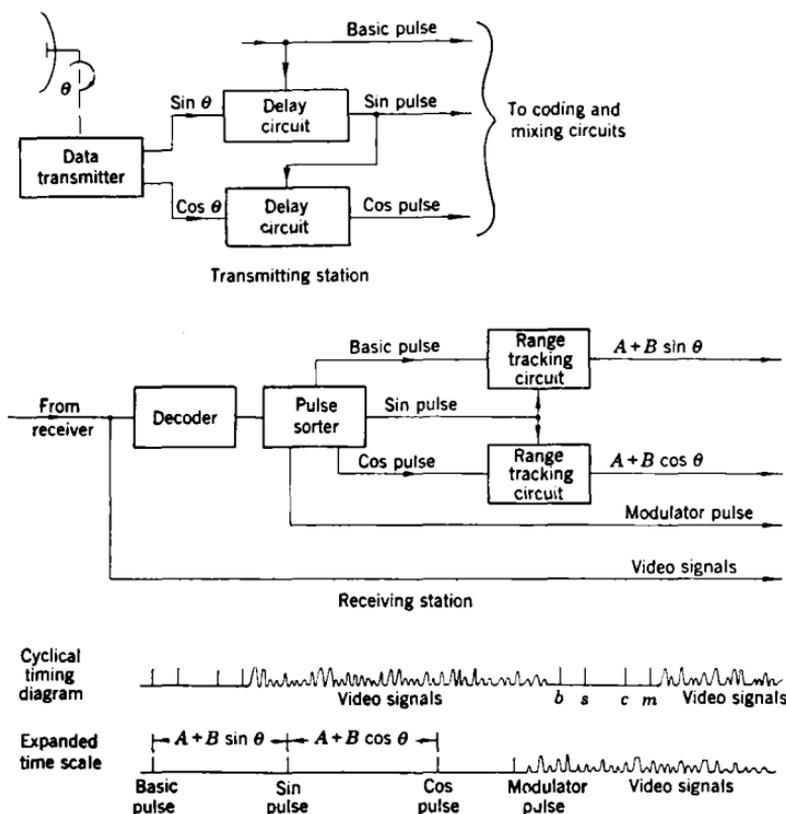


FIG. 17-12.—Essentials for relaying sine and cosine by pulse-timing methods.

Since this general type of synchronization equipment has seen far more service than any other to date, it will be worth while to consider it in more detail. The following section gives an extended treatment of one particular example.

17-8. Pulse Method for Relaying Sine and Cosine.—Electronic means for determining the interval between pairs of pulses will be considered first. In Fig. 17-13 it is assumed that the basic pulse and the sine pulse have been decoded and separated so that each is available on a

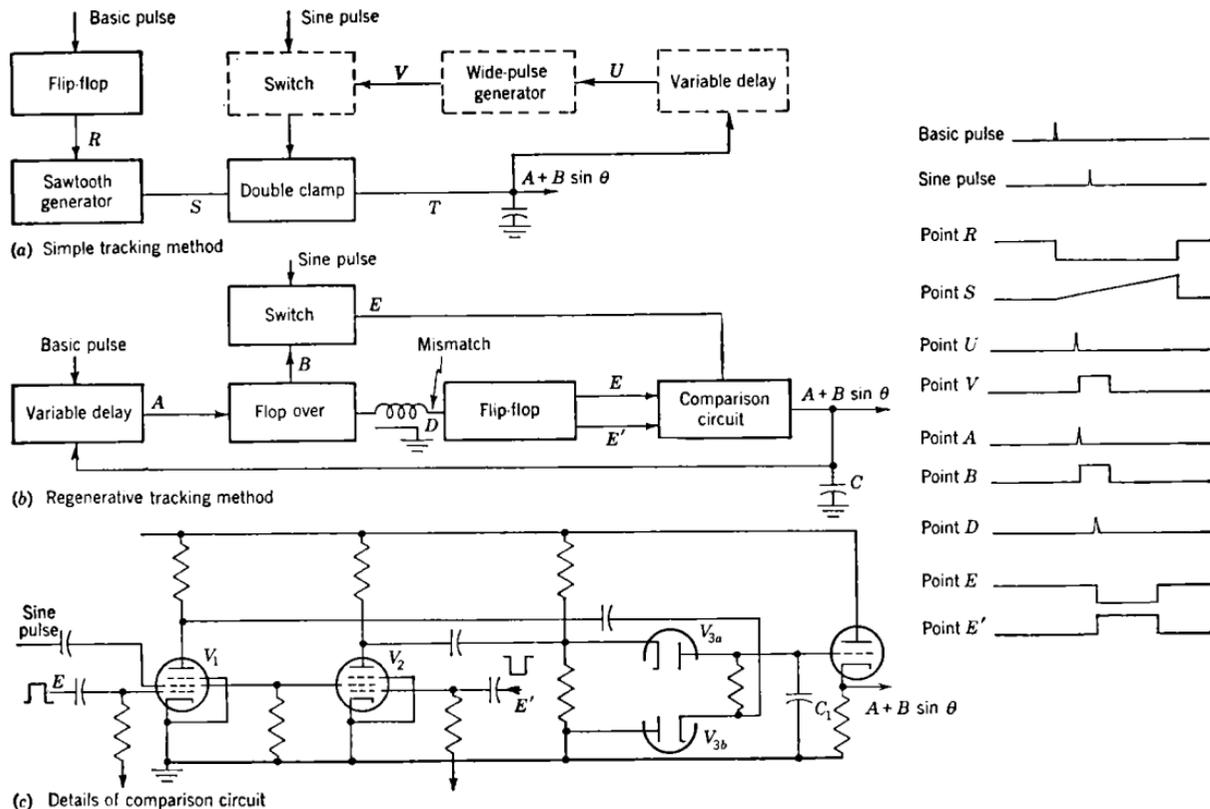


FIG. 17-13.—Tracking methods: (a) simple tracking methods; (b) regenerative tracking method; (c) details of comparison circuit.

separate channel. There must now be developed a d-c voltage whose value is at every instant proportional to the delay of the sine pulse with respect to the basic pulse. Two of the many possible methods of accomplishing this are illustrated. The components of Fig. 17-13a drawn in solid lines illustrate a very simple method. The basic pulse triggers a flip-flop having a lifetime slightly greater than the maximum delay of the sine pulse; the resulting square wave is used to switch a sawtooth generator of very low output impedance. Thus, within the life of the sawtooth, the instantaneous voltage at S is proportional to time elapsed since the occurrence of the basic pulse. When the sine pulse occurs, it momentarily closes the "double clamp" (similar to Fig. 13-26), thus connecting point T tightly to point S so that the condenser is charged to the instantaneous potential of S . The leakage path from T to ground is made to have a high resistance so that the potential at T remains essentially constant until the new cycle, at which time it will take a new value corresponding to the new value of the time delay of the sine pulse. Thus the potential at T will go through the same variations as the time delay and will in the present case have the form $A + B \sin \theta$ as desired. This very simple method is satisfactory provided little or no interference is encountered. An interfering pulse will, however, cause T to take a potential corresponding to its time of appearance. This effect can be reduced by filtering so that no single pulse can cause any very great change, but this filtering may cause troublesome phase lags in the desired output. A better method of protection is to employ the output voltage to control the opening of a switch (dotted circuits in Fig. 17-13a), through which the sine pulse must pass, in such a way that a pulse can be admitted only during a very narrow time interval including the time when the true pulse is expected. (Since this device as described will work only when the tracking is already nearly right, some means must be provided for holding the switch open until correct conditions are once established.) With the addition of this protective switch, the method of Fig. 17-13a is satisfactory. The addition makes it, however, nearly as complex as more elegant methods of regenerative tracking. Regenerative tracking, although more complex, provides certain very definite advantages in compensating for errors occurring later in the circuit.

The elements of a regenerative tracking circuit are shown in Fig. 17-13b. The basic pulse triggers a delay circuit which is controlled by the final output voltage in such a way that, assuming this voltage is right, the delay circuit produces a pulse shortly before the arrival of the sine pulse. The delay-circuit pulse triggers a flopover (Sec. 13-7) which in turn sends a pulse down an improperly terminated delay line. The reflected pulse, of opposite polarity, turns off the flopover. The resulting square

wave from the flopover is used to open the sine pulse switch for 25 μ sec or so at the expected time of arrival of the sine pulse.¹

The delayed pulse reaches the far end of the delay line exactly at the middle of the switching wave. The remainder of the circuit is designed to force this "comparison" pulse into time coincidence with the sine pulse by properly controlling the delay through adjustment of the output voltage. Various methods can be used to accomplish this adjustment. The operation of the method illustrated is described below. The comparison pulse triggers a flip-flop. The square waves from the two plates, one positive and one negative, are fed to the control grids of a pair of pentodes (Fig. 17-13c), turning one on and the other off. The screen grids, normally at cathode potential, receive the positive sine pulse. If the pulse arrives when the control grid of a given tube is on, that tube will produce a negative pulse on its plate. The signals are integrated in opposite polarity on condenser C_1 by means of the diodes. A negative signal from V_2 drives a negative charge through V_{3a} to C_1 . A signal from V_1 drives a negative charge to ground through V_{3b} . The potential on C_1 is the output signal. It is returned (perhaps after amplification) to the comparison-pulse delay circuit which it controls. Thus, if the switching occurs before the arrival of the sine pulse, V_1 conducts, C_1 becomes more positive, and the delay is increased. If the switching is too late, V_2 conducts and the delay is decreased. The process continues on successive cycles until the delay is such that the sine pulse "straddles" the instant of switching. The delay circuit will then follow the variations in the delay of the sine pulse so that the output voltage varies in the desired manner.

Figure 17-14 illustrates a system in which, although it is assumed that the modulator can be triggered, means are provided to distinguish between the cosine pulse and the modulator pulse for other reasons.

The delay circuits in the synchronizer may be any of the varieties described in Secs. 13-7 and 13-12, the most precise and trouble-free results being obtained by the use of circuits similar to but not so refined as the circuit shown in Fig. 13-36. If this type is used, the sine or cosine voltage is used to bias the cathode of the diode, and the bias of the anode is adjusted to give the desired delay A at zero scanner angle. The value of B is determined jointly by the slope of the sawtooth and the amplitude of the sine and cosine voltages from the scanner.

The basic pulse, sine pulse, and cosine pulse are passed to the usual three-pulse coder along with the modulator trigger signal, which is derived from a fixed-delay circuit triggered by the cosine pulse. In order that the trigger signal shall be properly timed when decoded at the receiver, the

¹ Since this interval is too narrow to allow rapid "locking in" initially or after tracking has been lost, the circuits are arranged in a manner not shown so that in the absence of pulses the grid remains on.

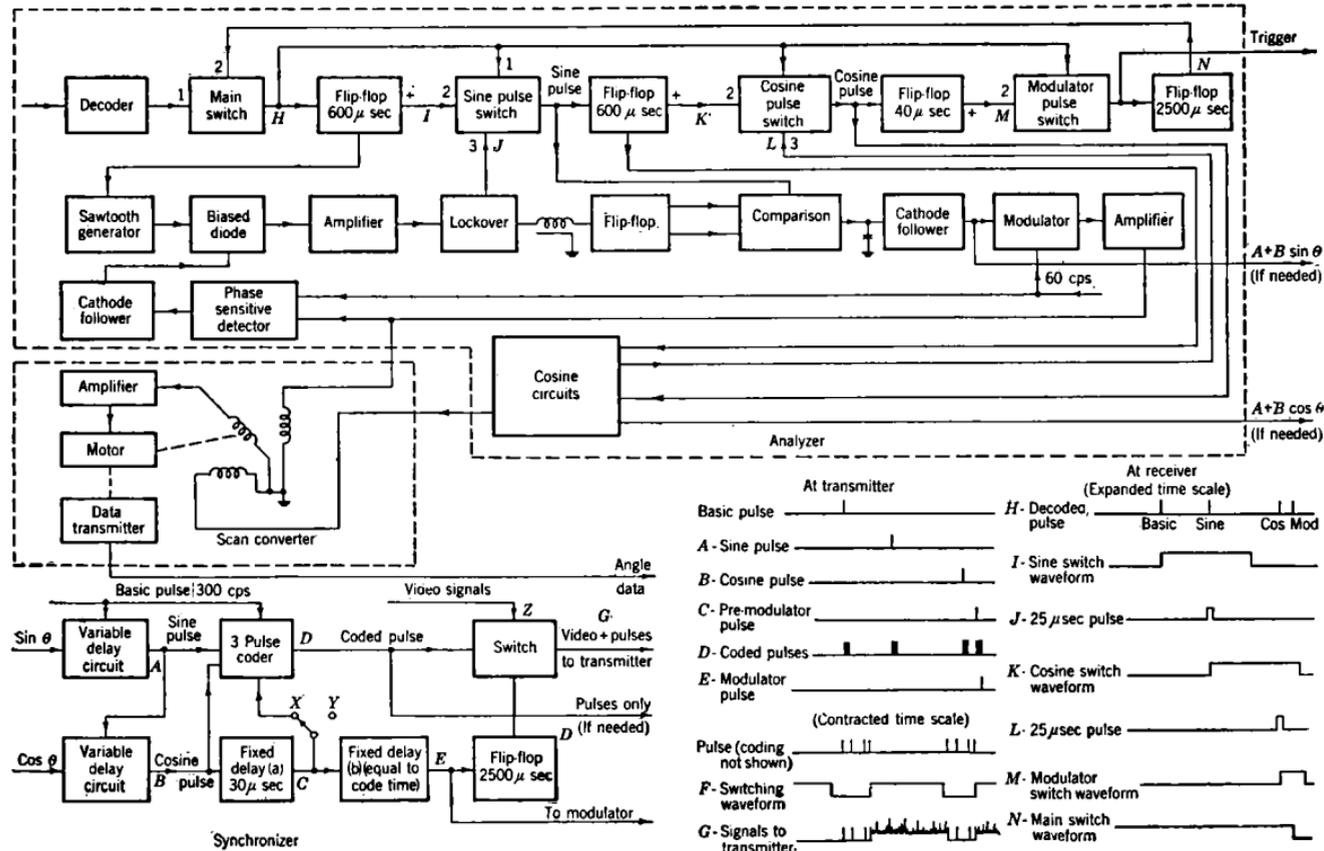


FIG. 17-14.—Specific method for relaying sine and cosine by pulses.

actual modulator trigger is delayed from the signal to the coder by a time equal to that occupied by the code. By closing the relay onto point *Y* and inserting a proper electronic switch between *X* and *Y*, the modulator trigger signal can be used as a switching signal when cyclical time sharing with beacon or other signals is desired. If no such need is involved, the cosine pulse can serve also as the modulator trigger signal so that delay circuit *a* can be omitted and points *B* and *C* made the same.

The coded pulses and the video signals are made to share the radar cycle by a switch operated by a flip-flop keyed by the modulator trigger, as described in previous cases.

At the receiving station, the pulses must first be separated from the video signals and from one another. The "sequencing" circuits which accomplish the latter are shown on the top row of the diagram. After being decoded, the signals pass to several switches of the multiple-grid variety shown in Fig. 13-27. The numbers alongside each indicate to which grid the various signals are applied.¹ The main switch is opened shortly before a basic pulse is expected, in a manner that will be described presently. On passing through this switch, the pulse triggers a flip-flop which remains on for an interval slightly greater than the largest delay time of the sine pulse. This elevates the No. 2 grid of the sine-pulse switch so that the sine pulse can pass through provided the No. 3 grid is on, as explained in connection with Fig. 17-13. Once the system is "locked in," the narrower switch pulse on the latter grid is the deciding factor in rejecting unwanted signals; the longer interval is useful during the locking-in process when the shorter one cannot be used. The sine pulse, in addition to its tracking role, triggers a flip-flop controlling the cosine-pulse switch, and the cosine pulse similarly triggers a flip-flop controlling the modulator pulse switch. The modulator pulse is passed on to the indicators and, in addition, triggers a flip-flop which closes the switch ahead of the decoder while video signals are being received, and opens it shortly before the next basic pulse.

A few words should be said about the process of "locking in" initially or after tracking has been lost through failure of signals or severe interference. Since the main switch is controlled by a flip-flop, it will sooner or later be open. If no pulse occurs during the open interval, this switch will close, but it will be opened again by the next pulse. Sooner or later it will receive a pulse when open. If this is the sine pulse, the sequence is established. If it is not, the cosine-pulse switch will pass no signal, and the process must start over. In practice, the right combination is promptly found, and the various pulses then pass through the sequencing circuits in the proper manner.

¹ The pulses may be separated from the video signals by amplitude selection, or by time separation as in Fig. 17-4.

The tracking circuits are similar to the circuit shown in Fig. 17-13c except for changes made to provide for a scan converter and for the fact that the switching voltages for the sawtooth delay circuits are provided by the same flip-flops that control the sine- and cosine-pulse switches. A-c signals for the scan converter are obtained by modulating the 60-cps line voltage with the sine and cosine voltages, as in Fig. 17-10a. The regenerative tracking assists greatly in overcoming errors in the modulators by rectifying the modulated alternating current with a phase-sensitive detector and using this result to control the tracking delay circuit. Thus the amplitude of the *modulated* alternating current (which is the quantity of primary interest) is forced to vary in the same manner as the delay in the signal pulse, regardless of errors in the comparison circuits, the modulator, any amplifiers, etc.

Equipment essentially like that just described has had a great deal of use and in its final form has proved satisfactory. An entire system using this method is described in Sec. 17-16.

17-9. Comparison of Synchronization Methods.—Not all the synchronization methods described in the previous sections have been used on actual radar relay systems although nearly all of them have been set up and tested in the laboratory. Because of the inadequacy of testing in many of the cases, evaluations are difficult, but the following general statements can be made.

Pulse vs. C-w Methods.—Pulse methods have been much more highly developed than those using c-w transmission, partly because of the pulse-circuit experience available at the Radiation Laboratory, and partly because the r-f equipment available during most of the work was of the amplitude-modulation rather than the frequency-modulation type. However, much can be said in favor of the pulse methods from a fundamental standpoint, regardless of the type of carrier modulation.

1. Since one pulse, the trigger, is involved in any method, adequate protection against pulse interference must be provided in any case.
2. When time sharing is used within the radar cycle, the pulse methods require less complex r-f equipment than do the c-w methods with their requirement for at least one subcarrier. Were time sharing with the video signals abandoned and the pulses sent on a subcarrier, the additional complications would be compensated by considerable simplification and increased effectiveness of the terminal equipment. The synchronization signals would automatically be separated from the video signals. The use of 100 per cent of the time and the freedom from restrictions imposed by the radar PRF would lessen the sensitivity to interference and permit much more flexibility than is available with the time-sharing technique.

Such a method would be extremely advantageous if multiple sets of angular data were to be relayed.

3. Pulse methods are inherently more accurate than c-w methods since they involve either "counting" or the measurement of time—probably the most accurate types of physical measurements that can be made. Furthermore, no dependence whatever is placed on the linearity of r-f modulators, amplifiers, detectors, etc., in contrast to the c-w methods where all such devices must be extremely linear and free from distortion.

On the other hand, c-w methods involve somewhat simpler terminal equipment than do the pulse methods. They are also much less susceptible to pulse interference, which is the type most likely to be met with at most of the radio frequencies involved.

Comparison of Specific Methods.—The specific synchronization methods that have been most thoroughly tested are—

1. The conveying of sine and cosine by pulses.
2. The conveying of angular increments by the modulator pulses.
3. The conveying of angular increments by the sinusoidally varied pulse.
4. The method of phase-shifted pulses.
5. The conveying of sine and cosine by a-f signals of different frequencies.

Certain qualitative comparisons among these are possible.

The method of relaying angular increments by modulator-pulse timing is by far the simplest if it can be used, which is only when the scanner rotates very uniformly and can be exactly synchronized with the modulator. Unfortunately, the tests of this method were made in connection with an airborne radar in which the scanning rate varied considerably; hence appreciable inertia could not be used. Furthermore, multiple-pulse coding had not been adopted at that time. As a result, interference led to somewhat erratic results. Since no opportunity has arisen for testing with a proper radar and with coded pulses, it is difficult to make an accurate assessment, but the comparative success under unfavorable conditions indicates that when properly applied the method can be satisfactory.

If the scanning is fairly uniform but cannot be synchronized with the modulator, the method of the sinusoidally varied pulse (Sec. 17-5) gives very satisfactory results with a minimum of complexity as proved by a reasonable amount of testing (Sec. 17-15). This method is somewhat less susceptible to interference effects than Method 2 since more data are sent per radar cycle and since an extra pulse does not cause an irreversible effect.

The method of relaying sine and cosine data by pulse-timing tech-

niques has given satisfactory results even under very severe interference conditions. It can, of course, follow sector, as well as continuous, scanning. Its principal drawback is its relative complexity, but much of the complication in Fig. 17-14 arises from extreme precautions against interference and from the necessity of operating a servomechanism. Were fixed-coil PPI's used, or were it possible to operate a servomechanism satisfactorily from d-c signals, the equipment would be considerably simplified. A far simpler model operating fixed-coil PPI's and designed for less severe interference conditions has seen considerable successful service.

The phase-shifted pulse method is quite comparable to the sine-cosine method in effectiveness; its method of converting to mechanical motion is somewhat simpler, and it has considerably fewer adjustments. Since more data are transmitted per radar cycle, greater protection against interference is probably afforded. Although this method is of more recent origin than the sine-cosine method, has not been so highly developed, and has not had such thorough tests, it appears to be satisfactory.

The sine-cosine method using multiple a-f signals is in principle quite simple and should prove satisfactory. Although it has been successfully operated in an actual system (not at the Radiation Laboratory), the author is not aware of any extensive tests in the presence of interference. This method requires more complicated and wider-band transmitting and receiving equipment than do the pulsed methods which involve time sharing.

THE RADIO-FREQUENCY EQUIPMENT

In selecting the radio-frequency equipment, the requirements to be met and the operational conditions must be carefully considered. Among the important factors are the station locations (land-, water-, or air-based), the maximum range required, the types of interference to be met, and the nature of the data to be transmitted. These factors affect many of the variables of the design, such as the types of antennas chosen, the r-f power necessary, the carrier frequency most desirable, and the mechanical construction of the equipment.

17-10. Antennas, Frequencies, and the Radiation Path.—The antenna gain should be as high as practicable at both stations. For a given transmitted power, the signal discernibility is always directly proportional to the gain of the transmitting antenna. At the receiving station the signal-to-internal-noise ratio is also proportional to the antenna gain, and considerable interference reduction is accomplished by the directionality that accompanies high gain. Such antennas therefore permit the use of much lower power to achieve a given result than would be required with omnidirectional antennas.

Highly directional antennas can always be used for transmission

between fixed ground stations since they can be permanently oriented in the correct direction. Very high, even microwave, frequencies are desirable for such applications since directional antennas can then be small, rugged, and easily constructed and installed. At least one microwave equipment has been successfully used in experimental tests (Secs. 17-14 and 17-15).

When one or both stations are moving, on the other hand, it is necessary either to use sufficiently wide antenna patterns to provide coverage in all the necessary directions, or else to provide for automatic pointing of the antenna, with the consequent added complexity and weight. (If several receiving stations are involved the transmitting antenna must cover them all.) As a result, lower frequencies (100 to 900 Mc/sec) have predominated in such applications since more power is available and higher antenna gains can be obtained with a given pattern. In cases involving transmission from aircraft, certain interference effects described below are much less troublesome at low frequencies than at very high frequencies.

In systems tested at the Radiation Laboratory for the transmission of data from an aircraft to the ground, elementary antenna arrays have been used. One system, operating at 300 Mc/sec, had a vertically polarized dipole mounted on the tail section of the aircraft and two vertically stacked dipoles at the receiving station. The vertical gain of this latter antenna improved somewhat the ratio of signal to inherent and to local noise, but did not reduce interference appreciably since little interference comes from high angles. In a test of a 100-Mc/sec relay equipment, the transmitting antenna was a quarter-wave vertical radiator mounted on the skin of an aircraft and a corresponding vertical quarter-wave receiving antenna was used at the ground station.

When a link was established at 800 to 900 Mc/sec, it was found very desirable to increase the gain of the two antennas in order to extend the usable range of the equipment. This was accomplished by using stacked dipoles at both the transmitting and receiving stations. These arrays had a uniform horizontal pattern and some gain in the vertical plane.

Associated or near-by equipments often constitute a serious source of interference; consequently both antenna and power-line filters are desirable at the receiving station. Frequently these must be designed to reject a particular frequency being radiated by a near-by antenna. Conversely, it is often necessary to filter the r-f output of the relay transmitter in order to minimize the radiation of harmonics that interfere with near-by receivers. The specific filters required for a given installation must be designed to meet the operational requirements of a specific system involving a given complement of radar, communication, and navigational equipment. No general rules can be given.

Diffraction Phenomena.—Whenever one of the relay stations is in an aircraft, especially if the transmission path is over water, interference

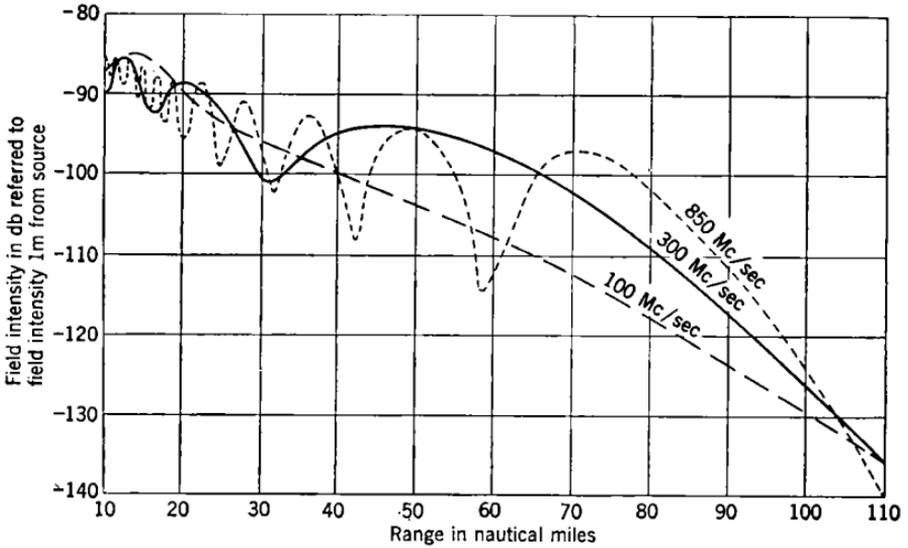


FIG. 17-15.—Field intensity vs. range for propagation of vertically polarized signals over sea water at various frequencies. Transmitter antenna height, 5000 ft; receiver antenna height, 75 ft.

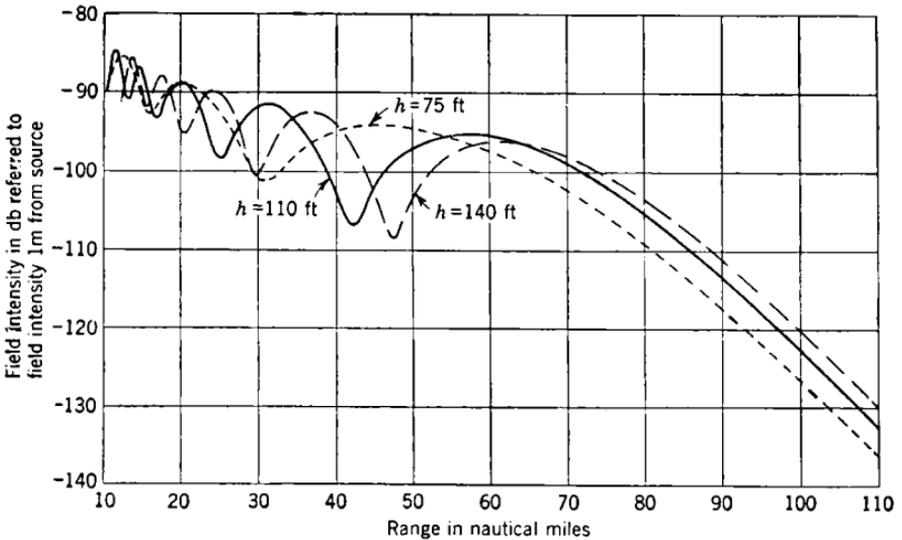


FIG. 17-16.—Field intensity vs. range for propagation of vertically polarized signals over sea water at various receiving antenna heights. Transmitter antenna height, 5000 ft; frequency, 300 Mc/sec.

between direct and reflected rays can occur. This subject is treated in detail in Vol. 13 of this series. Figure 17-15 shows curves of

signal strength as a function of range for frequencies of 100, 300, and 850 Mc/sec, displaying the effect of frequency on the number and the magnitude of the fluctuations in signal strength that are due to interference. Figure 17-16 shows curves of signal strength as a function of range for receiver antenna heights of 75, 110, and 140 ft above the surface of the sea. These curves indicate the manner in which the effects of signal cancellation can be reduced by using diversity reception with antennas at different heights. A simple diversity system might consist of two antennas respectively 75 and 140 ft above the sea, with arrangements for switching the receiver to the antenna providing the greater signal strength at any given moment. In effect, this reduces the depths of the cancellation minima in either of the antennas considered separately.

Future Trends in Frequency.—Frequencies most generally used for radar relay have been in the region from one to a few hundred megacycles per second, partly for reasons of achievable power, higher gains of non-directional antennas, and so on. Future trends appear to lead toward higher frequencies, up to and including microwaves.

Comparatively speaking, the use of such frequencies is characterized by the possibility of using simple, highly directional antennas, by low gain in omnidirectional antennas, by reduction in man-made static (except for pulse interference from radars), by a great increase in the overwater diffraction effect, by a large number of available channels, and by the relatively low power now available for c-w operation.

The powers available for continuous operation with equipment developed by the end of the war are approximately as follows:

100 Mc/sec	
(FM).....	100 watts average
(AM).....	{ 100 watts average
	{ 250 watts peak video level
	{ 500 watts peak in pulses
300 Mc/sec	
(FM).....	30 watts average
(AM).....	{ 30 watts average
	{ 80 watts peak video level
	{ 175 watts peak in pulses
1000 Mc/sec	
(FM).....	25 watts average
(AM).....	{ 25 watts average
	{ 40 watts peak video level
	{ 80 watts peak in pulses
3000 Mc/sec (FM).....	10 watts average
10,000 Mc/sec (FM).....	0.1 watts average

Any of these is sufficient for use in applications where both antennas can be directive, and there is little question that the very highest fre-

quencies will be used for such purposes in the future. Until recently it was not economical to use a-f modulation of microwaves because of the extreme oscillator-stability requirements necessary to avoid excessive bandwidth. However, a highly ingenious application of a microwave discriminator in combination with a feedback amplifier controlling the oscillator frequency¹ appears to overcome this difficulty and to permit the use of c-w synchronization methods.

For applications requiring omnidirectional antennas, the ranges so far obtained at frequencies above 300 Mc/sec have been rather limited, especially in situations involving diffraction effects. As greater power becomes available at the higher frequencies, they will undoubtedly find more and more applications even where it is not possible to use highly directed beams. The diffraction difficulty can largely be overcome by the careful use of diversity antennas.

17-11. General Transmitter and Receiver Considerations.—Although the specific characteristics desirable in the transmitter and receiver depend upon the particular application, certain general statements can be made.

The transmitter should provide sufficient power to ensure clear signals, free from noise and interference, at the maximum required range. In common with all components, it must have sufficient bandwidth to accommodate the band of frequencies present. The receiver should have a satisfactory noise figure, a proper bandwidth, and in many cases must provide special means of distinguishing between desired and undesired signals by methods analogous to the antijamming techniques described in Sec. 12-8. Automatic gain control is necessary to prevent strong signals from overloading the receiver, and to insure that signals are applied to the decoder at the correct level.

Bandwidth varies with the particular characteristics of the radar set and the type of synchronization used. In general, the video sections will have a bandwidth between 1 and 3 Mc/sec, with corresponding r-f and i-f bandwidths from 2 to 6 Mc/sec, when normal search radar systems are used. Since the relay link is only one section of the over-all channel, the bandwidths of its components must be somewhat greater than would be necessary if it alone were involved.

An important decision is the choice between amplitude and frequency modulation. The relative advantages and disadvantages of these two methods are somewhat different for pulsed and for c-w signals, and depend also upon the type of interference expected. The principal advantage of frequency modulation is this: if the carrier power is appreciably greater than that of an interfering signal, the latter tends to be

¹ R. V. Pound, "An Electronic Frequency Stabilization System for CW Microwave Oscillators," RL Report No. 815, Oct. 1, 1945; *Rev. Sci. Inst.* **17**, 490 (1946).

suppressed. Frequency modulation is thus helpful in cases where it is desirable to remove the last traces of low-level interference. Such interference reduction is effective only if a large deviation ratio is used, that is, if the ratio of half the maximum carrier-frequency excursion to the maximum modulation frequency is large. (A deviation ratio of 4 is considered excellent.) This requirement increases the necessary bandwidth of the r-f parts of the transmitter, and of the r-f and i-f stages of the receiver. This increase presents additional transmitter circuit problems, and reduces the gain in the amplifier stages. In the receiver, the greater bandwidth admits more interference in addition to complicating the receiver design.

The most important aspects of the relative virtues of the two types of modulation arise, however, in connection with the consideration of particular types of signals and interference. For example, if pulse synchronization signals are received at a level appreciably above that of interference, amplitude-selection methods can entirely exclude the interference, regardless of the type of modulation. This gives the amplitude-modulation method a definite advantage because the low duty ratio of the synchronization pulses makes it possible to transmit them at peak powers several times higher than the permissible average¹ and thus assists these signals to override the interference. A fairly high ratio of peak to average power can also be maintained for video signals since, except in extreme cases, echoes are received for only a small fraction of the time. In frequency modulation, on the other hand, the carrier operates at a constant power level.

No such advantages exist for amplitude modulation in connection with c-w synchronization methods. Low-level, more or less continuous interference can be very disturbing, and the natural suppressing effect of the limiter and discriminator in an f-m receiver can be of very great advantage. Furthermore, the use of subcarriers and carriers of higher order can be most readily accomplished in a system which is frequency-modulated throughout.

The above discussion, admittedly rather hypothetical in the absence of extended comparative tests, might be summarized as follows:

1. Amplitude modulation methods seem preferable for the relaying of synchronization pulses (of which there is always at least one), the advantage increasing with the strength of the interference.
2. There is probably little to choose between AM and FM with respect to the video signals. If the interference is severe, the higher peak powers and narrower bandwidths possible with ampli-

¹ This is usually accomplished by combining grid modulation by both the video and the pulse signals with plate modulation by pulses only.

tude modulation make it preferable; if, on the other hand, the interference is at low level it can be more completely suppressed by frequency-modulation methods.

3. Frequency-modulation methods are definitely preferable for c-w synchronization signals.

The following sections give brief descriptions of some actual equipments.

17-12. A 300-Mc/sec Amplitude-modulated Equipment.—Largely because of its availability, the type of r-f equipment most used at Radiation Laboratory consists of a modification of an amplitude-modulated television transmitter-receiver combination operating in the 300-Mc/sec region (specifically on any of 10 channels between 254 and 372 Mc/sec). The transmitter provides 90 watts of peak video signal power, and 250 watts of pulse power.

A block diagram of the transmitter is shown in Fig. 17-17a. Provision is made for grid modulation by all of the signals and for additional plate modulation by synchronization pulses. Negative video and synchronization signals are amplified by a three-stage broadband video amplifier. The second stage has a gain control in the cathode to compensate for variations in input-signal amplitude. The amplified signals drive the grids of the 8025 r-f power amplifiers through a cathode follower. Since the bias for the cathode follower must remain constant for all duty ratios, a d-c restorer is used between its grid and the -105 -volt supply. The synchronization pulses are amplified by a 4-stage video amplifier operating into a pulse transformer connected to the cathode of a diode through which the r-f amplifiers draw their power. In the absence of pulses, the diode is conducting and the plates of the r-f amplifiers are connected to the high-voltage supply (800 volts). The arrival of a pulse disconnects the diode and raises the plate potential of the amplifier by several hundred volts, resulting in a very high instantaneous power.

The master oscillator consists of a pair of 8025's in push-pull, the plate and grid circuits being tuned by transmission-line elements of variable length. The amplitude of oscillation is controlled by the capacitive reactance of the filament line, and can also be varied by changing the length of the filament line. Change of channel necessitates retuning of the r-f power amplifier by adjustment of a short-circuiting bar on the parallel line which constitutes the plate load. The electrical length of the coupling loop is also varied with the plate tuning. The monitor diode rectifies a small portion of the output signal, which is displayed on a scope for monitoring purposes.

A reflectometer, or bidirectional coupler, is coupled into the r-f line at all times. This gives a continuous indication of the power output and

an insertion loss of less than 2 db from 290 to 320 Mc/sec, and a loss of more than 40 db above and below this band. The design is conventional: three T -sections are matched to the line with a π -section at each end. Inductors are used as series elements, and combinations of lines as the shunt elements.

Figure 17-17*b* shows a block diagram of the receiver. The r-f amplifier is a miniature triode (6J4) connected as a grounded-grid amplifier. A dual triode (6J6) is used in a push-pull oscillator circuit tuned 30 Mc/sec below the carrier frequency. The 30-Mc/sec i-f signal from the converter is amplified by six stages, gain control being applied to the first three. The i-f bandwidth is about 3 Mc/sec between half-power points. The output of the detector, a 6AC7 connected as a diode, is applied to the first video stage. A choice of two time constants is available in the grid circuit of this stage, 0.47 sec and 2.4 μ sec. The longer one is normally used; it gives good response to very low video frequencies. The short time constant, when used, serves the same function against extended or c-w interference as similar circuits do in a radar receiver (Sec. 12-8).

The automatic gain control is actuated by the synchronization pulses. A small signal is taken from the plate of the final cathode follower, amplified, and passed to another cathode follower. Because of inverse feedback, the output signal of this cathode follower is a sharp spike, rather than a flat-topped 2- μ sec pulse. If the synchronization pulses are coded, they pass through a delay line to a coincidence tube, the combination acting as a decoder. The coincidence tube is so biased that only pulses will actuate it, the video signals being biased out at this point by the video-level control. The output signal of the coincidence tube is applied to the cathode of a diode, whose plate potential is set by the AGC level control. Thus, if the signals from the coincidence tube are sufficiently negative to cause the diode to conduct, the grid of the cathode follower which is also connected to the plate of the diode will change potential and thereby change the grid potential on the first three i-f stages. A long time constant in the cathode-follower grid circuit holds the grid potential essentially constant between pulses. There is thus a loop in which strong pulses produce a more negative potential on the i-f grids to reduce the receiver gain, and vice versa.

It may happen that operation of a radar in the vicinity will overload the receiver during transmission. Such interference can be overcome by introducing a portion of the interfering radar trigger at the interference-suppression terminals shown. This reduces the receiver gain at the instant of radar transmission, with the loss of only one or two microseconds of video-signal reception.

17-13. A 100-Mc/sec Frequency-modulated Equipment.—A second type of equipment which has been used for air-to-ground or air-to-ship

relay links is a 100-Mc/sec frequency-modulated system. The equipment operates at four frequencies between 78 and 116 Mc/sec, channel-changing being accomplished in the receiver by a coil-switching mechanism and in the transmitter by plug-in coils. A block diagram is given in Fig. 17-18. The input signals are amplified and delivered push-pull to the deviator by a phase splitter. The deviator acts as a reactance tube, modulating an oscillator which operates at one-eighth the desired carrier frequency. Three stages of frequency-doubling and power amplification

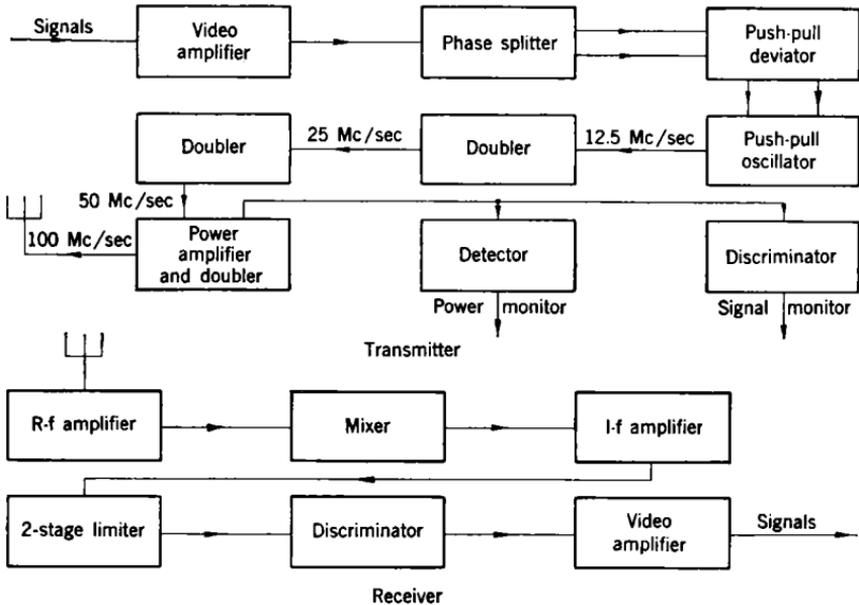


FIG. 17-18.—100-Mc/sec frequency-modulated equipment.

follow. The final stage consists of a pair of 4E27 tubes operated as a Class *C* amplifier with an average r-f output power of about 100 watts.

The equipment is designed to accommodate a maximum video bandwidth of 2 Mc/sec. The oscillator gives a maximum linear frequency deviation of 0.5 Mc/sec. Therefore, in this and the first doubler stage the bandwidth need be only the 2 Mc/sec determined by the video frequencies involved. After a second doubling, the bandwidth is made 4 Mc/sec in order to support two sidebands on each side of the carrier. After the final doubling, a bandwidth of 6 Mc/sec would be required to support all the sidebands above 5 per cent, but it was found experimentally that distortion was not serious if the bandwidth were reduced to 4 Mc/sec. This simplified the amplifier design and resulted in higher r-f power than would otherwise have been available.

Two miniature double diodes are included in the transmitter for monitoring purposes. One of these tubes is connected as a discriminator to provide a video-output test point at the antenna-line connector. During all testing, this video signal gives an accurate over-all check of the r-f and video sections of the transmitter. One half of the remaining double diode is used as an r-f detector to provide a relative power indication on a panel meter which is also used, by means of a rotary selector switch, to measure the grid currents of the several doublers. This diode is also connected to a test point to allow a scope to be used when over-all alignment and bandwidth measurements are being made. Because of the stability of this transmitter, it was found unnecessary to use a bidirectional coupler with it.

The receiver has one tuned r-f stage, a mixer, a local oscillator, seven i-f stages, two limiters, a discriminator, a video amplifier, and three tubes connected in parallel as a cathode-follower output stage. All the tubes in this section of the receiver are 6AK5 miniature pentodes, except the one used in the discriminator which is a 6AL5 miniature double diode.

The r-f, i-f, and video sections, as well as the discriminator, are standard in design, an 8-Mc/sec bandwidth in the i-f amplifier being obtained by double staggering of alternate stages. A two-stage limiter is used to insure constant input-signal voltage at the discriminator. Both limiter tubes operate at reduced plate and screen voltages to reduce the grid swing necessary to cause plate limiting. A fast time constant in the grid circuit of the first limiter was chosen to discriminate against impulse noise by producing the limiting bias quickly. Longer time constants are possible in the second limiter since the signal variations at this tube are neither large nor of short duration.

This equipment has been given extensive airborne tests in conjunction with two different types of synchronization equipment, and has also been operated between fixed ground stations.

17-14. Microwave System for Point-to-point Service.—The two equipments described in Secs. 17-12 and 17-13 were developed primarily for air-to-surface work involving the use of omnidirectional antennas; an upper limit was therefore set to the possible radio frequency. The present section will describe equipment designed for use between fixed ground stations which permit the use of directional antennas.

Safe margins of power are easily attainable in such applications since the maximum range is usually sharply limited by the horizon or by roughness of the terrain. The received signal was specified to be 40 db above thermal noise at maximum range in order to provide a safe operating margin. In the application of this criterion to the selection of frequency and antenna sizes, the following table is illuminating. Paraboloid antennas are assumed at both stations.

TABLE 17-1.—MAXIMUM RANGE OF RELAY SYSTEMS

Frequency, Mc/sec	Power, watts	Beamwidth, degrees		Free-space range, miles	
		9-ft paraboloid	3-ft paraboloid	9-ft paraboloid	3-ft paraboloid
300	50	≈ 27	≈ 80	196	22
1,000	25	8	24	460	51
3,000	0.5	2.7	8	195	22
10,000	0.1	0.8	2.4	292	32.5

When one considers the power involved, the decrease of man-made interference with increased frequency, and the privacy and protection from interference provided by narrow beams, microwave frequencies appear to be the most desirable. The decision between 3000 and 10,000 Mc/sec was based largely upon the fact that in the latter case waveguide of a convenient size could be used, an extremely desirable design feature. Since a 32-mile range is adequate for the uses intended, and since too great sharpness of beam might lead to alignment difficulties, 3-ft paraboloids were used at both stations.

Frequency modulation was chosen, partly because of its advantages when signal-to-noise and signal-to-interference ratios are high, but mostly because it simplified the oscillator design. Large deviations of the oscillator are easily accomplished, and there are no problems of r-f bandwidth. Since the required video bandwidth was about 1.5 Mc/sec, a total deviation of 6 Mc/sec was chosen. The total frequency spectrum involved is then a little more than 9 Mc/sec. In order to minimize the required i-f bandwidth in the receiver, the discriminator was set on one side of the pass band of the receiver. A value of 11 Mc/sec was then chosen for the i-f bandwidth to provide a margin to take care of improper tuning.

The equipment is shown schematically in Fig. 17-19. The oscillator, a 2K39 reflex klystron, is stabilized against a cavity by means of a microwave discriminator.¹ The output of this device is a d-c signal whose voltage is proportional to the deviation of the oscillator frequency from the frequency for which a resonant cavity, used as comparison standard, is set. This error signal is amplified by a push-pull d-c amplifier and used to control the reflector voltage of the klystron in such a way that its frequency is forced into agreement with the resonant frequency of the cavity. Video signals and pulses are applied directly to the reflector to produce the desired frequency modulation. Rapid response is purposely avoided

¹ R. V. Pound, A Microwave Frequency Discriminator, RL Report No. 662, Aug. 4, 1945; *Rev. Sci. Inst.* **17**, 490 (1946).

in the frequency-control circuits so that they will not stabilize against the signal frequencies.

Not shown in the transmitter diagram is a monitor, consisting of a crystal mixer and a video amplifier, which draws power from the main waveguide. In combination with a synchroscope, this provides a very effective means of checking and aligning the transmitter.

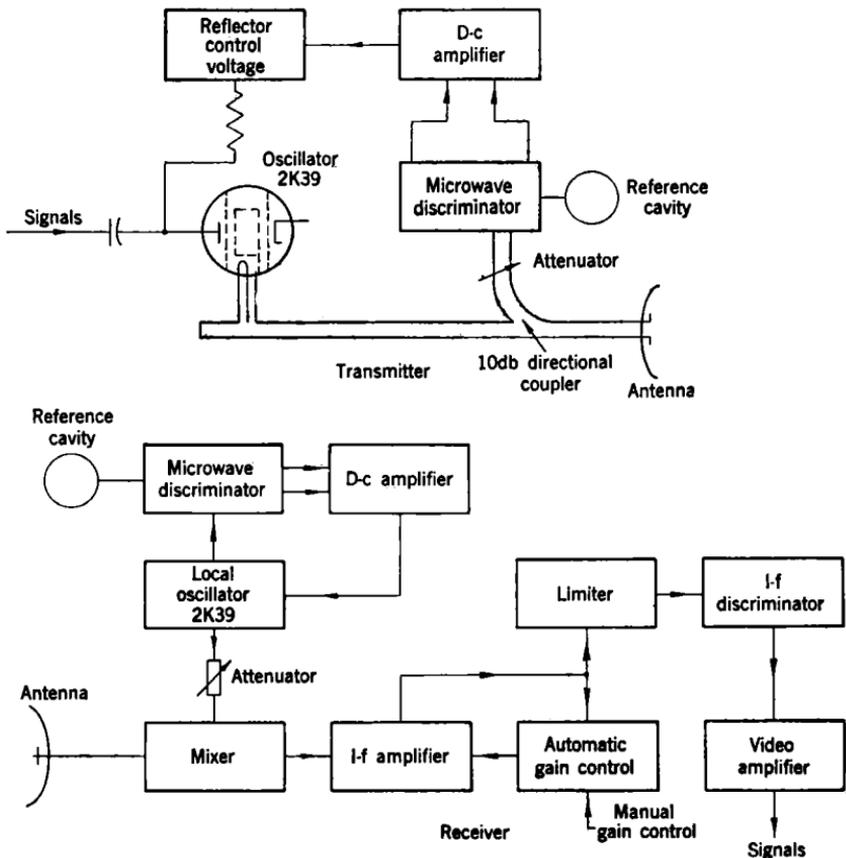


FIG. 17-19.—Microwave equipment.

The local oscillator of the receiver is also frequency-stabilized against a cavity. The circuits and layout of the i-f amplifier and the automatic gain control are similar to those of the receiver described in Sec. 12-11, the bandwidth being 11 Mc/sec. A two-stage limiter is used ahead of the i-f discriminator, care being exercised to provide sufficiently rapid limiting action to reduce impulse noise. The discriminator is similar to that used in radar AFC circuits and has a bandwidth of about 17 Mc/sec between peaks.

RADAR RELAY SYSTEMS

Two complete radar relay systems will now be described; their general features are sufficiently applicable to future requirements to make this worth while, even though better systems could now be designed. The requirements for the two systems are quite different. The first involves the transmission between fixed ground stations of three sets of radar echo signals and one set of beacon signals, all resulting from antennas mounted on a single continuously rotating scanner. Although the number of sets of video signals is large, the use of directional antennas eases the r-f problem, and the continuous scan renders scanner synchronization relatively simple.

In the second example, one set of radar signals and one set of beacon signals are relayed to ground from a long-range airborne radar set arranged to permit sector scanning. In contrast to the former case, the "picture" data are relatively simple, but the requirement of large angular coverage forces the use of low-gain antennas and puts a severe requirement on the r-f system. Sector scan and the turning of the aircraft greatly complicate the scanner-synchronization problem.

17-15. A Ground-to-ground Relay System.—The radar set originating the data in this example is a ground-based microwave set (see Chap. 15) in which the scanner, rotating at either 2 or 4 rpm, carries two radar antennas and one beacon antenna. One of the radar antennas provides long-range low-angle coverage; the other provides coverage at high angles. Both regular video signals and MTI (Chap. 16) video signals are derived from the upper-beam signals; this beam is chosen for MTI because it is the one predominantly used at the shorter ranges where the clutter is worse.

Thus four sets of video signals must be transmitted: lower-beam radar echoes, upper-beam radar echoes, upper-beam MTI video signals, and beacon responses. Time sharing is used to put two sets of video signals on each of two carriers. One channel is shared between the MTI video signals and the lower-beam video signals, MTI video being transmitted for a time interval corresponding to the first 30 to 50 miles of range from the radar, and the lower-beam video for the remainder of the radar cycle. A second channel is shared between the upper-beam radar echoes and the beacon signals; since the data from these two are simultaneous, switching must be done on a whole-cycle, rather than on a fractional-cycle, basis. Two pulse cycles are allotted to radar, then one to beacon, and so on, the unequal division being used because signal sensitivity is more critical in radar than in beacon operation. The resultant loss in sensitivity is 0.5 db for radar signals and 1 db for beacon returns.

Two transmitters, feeding a common antenna through a duplexer, are

used to provide the two channels, since, at the time of design, equipment accommodating subcarriers was not available and weight and power were not crucial items. The remaining data are combined on these same two channels as indicated in Fig. 17-20. The mixing and switching of the signals for the first transmitter is fairly simple. Range markers and the proper set of angle markers are mixed with each set of video signals, and the two sets are fed to a video switch, which passes the MTI signals for the first 50 miles or so and the lower-beam signals thereafter. The switch is like the circuit of tubes V_{1a} and V_{1b} of Fig. 17-3, but has signals on both grids.

The synchronizer is similar in function to that of Fig. 17-5 (including Fig. 17-2),¹ with the addition of a scale-of-three circuit to produce the beacon switch pulse on every third cycle. The third pulse of the modulator code is counted down for this purpose and passed to the signal switching unit. The switching signal to be relayed is delayed 16 μsec (by reflection in an 8- μsec delay line) in order that it be clear of the azimuth pulse at the receiving station.² The three-pulse code, the azimuth pulse, and the beacon pulse are "mixed" by using them all to trigger a blocking oscillator.

The video signals to the second transmitter are switched twice. The upper-beam video and the beacon video are switched cyclically as described above. The switch is controlled by a flopover circuit (Fig. 13-16) which remains in the stable state that causes video switch a to pass the radar signals until a beacon switch-pulse occurs; it then passes to a second stable state that causes the video switch to pass beacon signals. At the next modulator pulse the original condition is restored.

A video switch provides for time sharing between the video signals and the synchronizing pulses in order that the former shall not interfere with the latter. In its normal position, the flip-flop holds switch b in the state that passes pulses. Firing of the flip-flop by a pulse delayed by 30 μsec from the modulator pulse reverses the video switch allowing echo signals to pass. The flip-flop returns spontaneously to its original state shortly before the start of the next radar cycle. Signals from switch b are combined with range and angle marks in a video mixer, from which they pass to transmitter No. 2.

At the receiving station, the signals pass through a duplexer to two receivers. The first delivers the time-shared MTI and lower-beam video signals, together with markers, directly to the indicators. Signals from

¹ The circuit details of the equipment actually tested differ considerably from those of Figs. 17-2 and 17-5.

² In order that pulses passing through the three-pulse coincidence circuit shall not cause false coincidences, they should follow each other by at least the sum of the code length and the pulse length—in this case a total of 8 μsec .

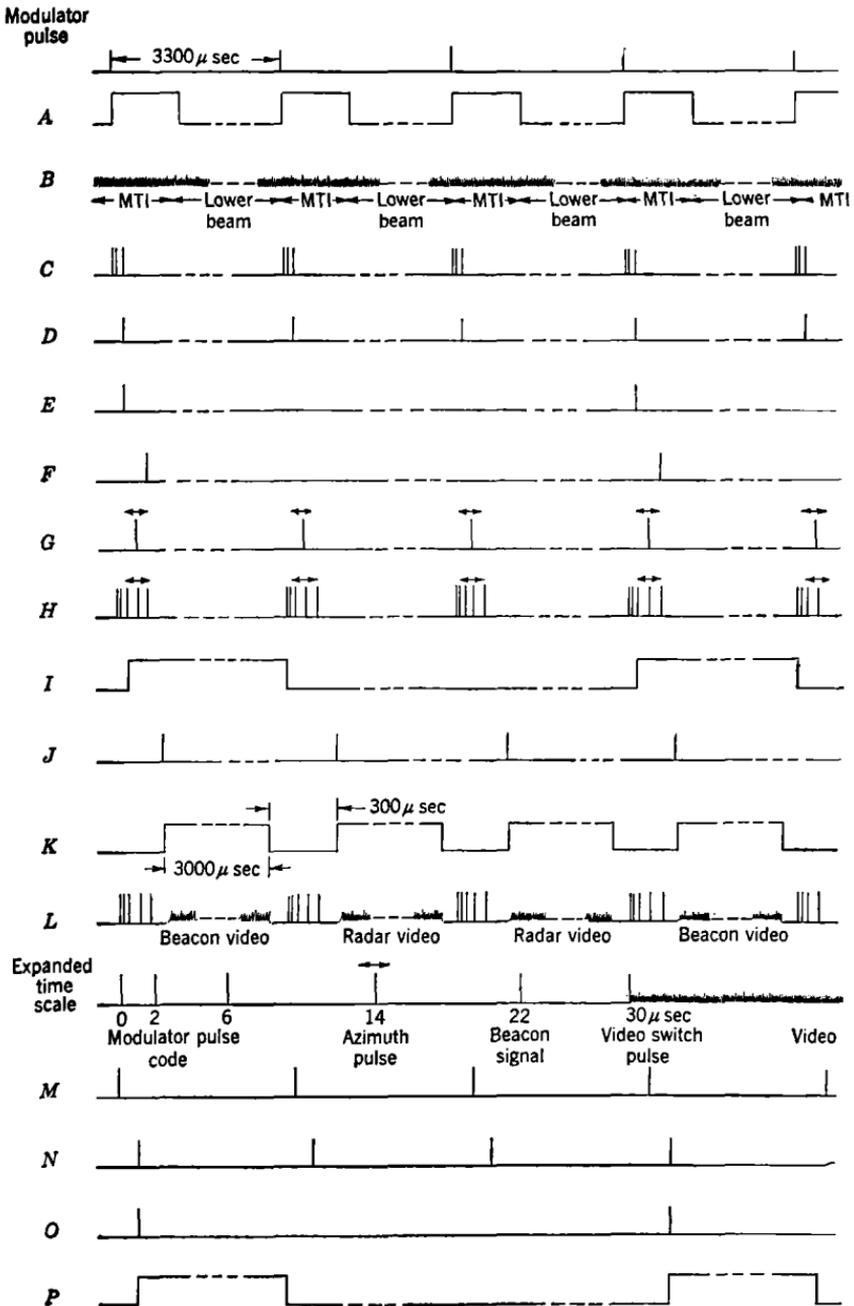
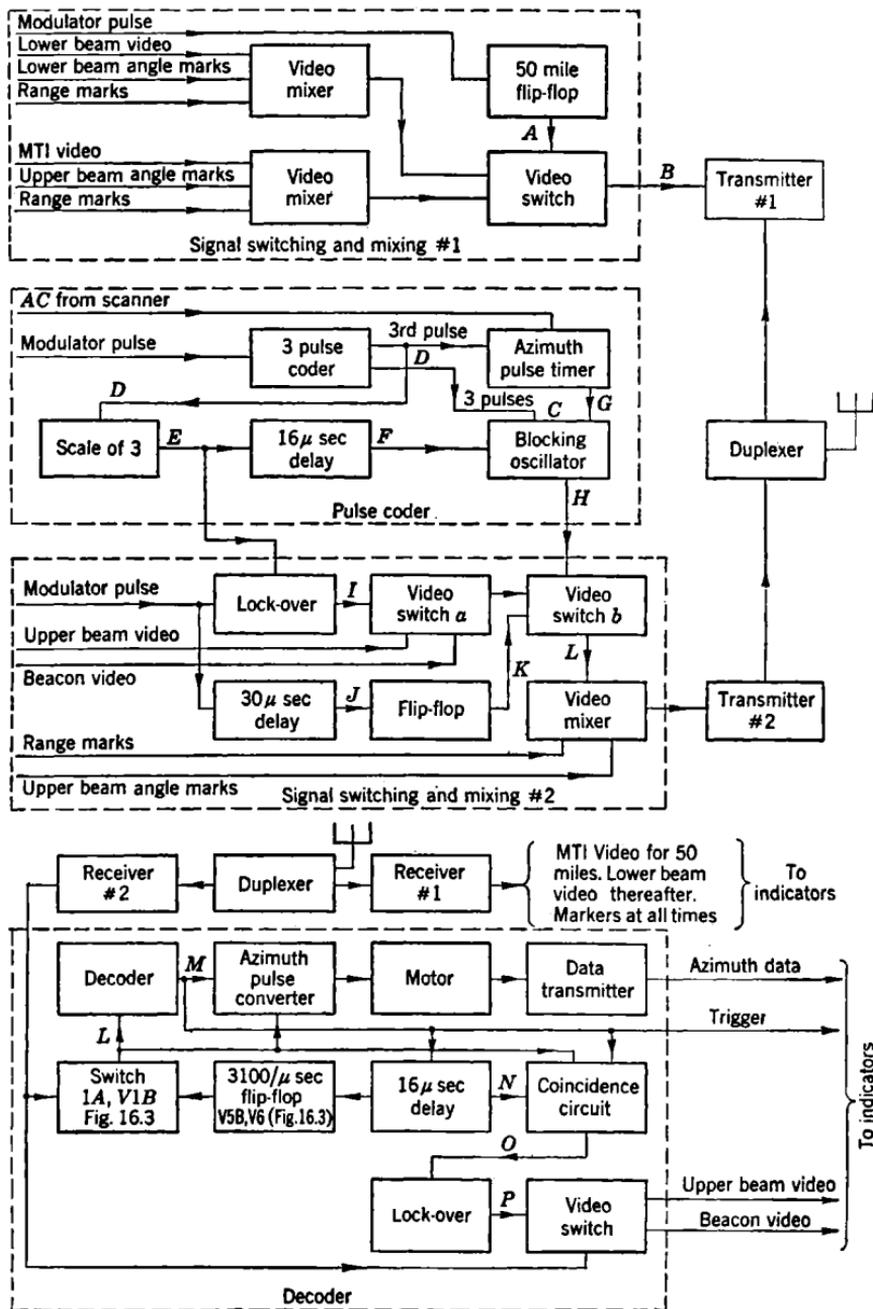


FIG. 17-20.—Ground-to-



ground radar relay.

the second receiver contain the coded modulator pulse, the azimuth pulse, and the beacon switch pulse. They pass through a switch which excludes all signals while video is being transmitted but opens shortly before the arrival of a trigger pulse. This switching, the decoding of the modulator pulse, and the derivation of the scanner information from the azimuth pulse, are done by methods similar to those of Figs. 17-2 and 17-5. The upper-beam video signals and the beacon video signals are separated by a video switch which consists essentially of a pair of out-of-phase switches similar to that made up by tubes V_{1a} and V_{1b} in Fig. 17-3. The video switch is controlled by a flopover which is triggered to the beacon position by the beacon switch pulse and to the radar position by the next modulator pulse. The switch pulse is singled out by coincidence with a pulse derived locally (N) at the proper time by delaying the decoded modulator pulse.

The time occupied by the pulse code results in the triggering of the indicators 8 μ sec too late. This produces a slight distortion in the displays, but there is no error in range measurements since accurate range markers are transmitted with the video signals. The display distortion is unimportant because short-range displays are not used. Similarly, the fact that the indicators must be blanked out for the first 30 μ sec is of no importance since targets at such close range are practically never of interest.

Several r-f equipments, including the three described in earlier sections, were tried experimentally in this application. All operated with reasonable satisfaction, maximum range being limited in every case only by the line of sight. On the whole, the microwave equipment is considerably superior to the others because of its compactness, the small power involved, the narrowness of the beam, and the greater freedom from interference. However, the fact that the use of frequency modulation did not permit the pulses to be transmitted at higher level than the video signals was a definite handicap for reasons described below.

The equipment as a whole operated about as anticipated. Comparative PPI photographs taken simultaneously at the two stations are shown in Fig. 17-21. The only difficulties of consequence involved occasional loss of synchronization, usually because of pulse interference picked up on the radio link or on the radar set. The direct results of spurious triggers on either the angle data or the sweep triggering were not appreciable, but loss of the trigger occasionally upset the sequencing with unfortunate results. Once the proper chain is broken, it can be spuriously started by interference or by video signals and remain in error for several cycles. Both the indicator sweeps and the azimuth data are then in error, sometimes by as much as 5° or 10° . The resulting angular error persists until it is manually removed. When amplitude selection of the



FIG. 17-21.—Comparative PPI photographs taken simultaneously at the transmitting and at the receiving stations.

pulses is possible (with the amplitude-modulated equipment), these effects are greatly reduced. Amplitude selection also provides suppression of weak interference picked up on the link itself. With this protection interruptions are extremely brief and cause little difficulty.

17-16. Relay System for Airborne Radar.—In the system just described, the scanner synchronization was rendered fairly simple by the continuous scanning, and the problem of obtaining adequate signal strength was simplified by the use of directional antennas. The principal complications were those involved in the simultaneous transmission of several sets of video data. The present section will describe briefly the arrangements used to solve a far more difficult problem, in which the data originate from a long-range airborne set equipped for sector scanning. The scanner synchronization, difficult in any case, is rendered far more so by the fact that the omnidirectional antennas required give so little gain that the interference problem is severe. Every possible device must be used to provide a maximum of power from the transmitter, to reject interference in the receiver, and to protect the synchronization pulses by coding, by switching, and so on.

The video data involved are simple, consisting merely of radar signals and of signals from a separate beacon receiver. In order that the two sets of video signals may be accommodated, cyclical time sharing is used during the intervals of beacon use, the modulator trigger serving as the signal that radar is being transmitted on a given cycle.

The design was built around the sine-cosine synchronization method of Sec. 17-9 and the 300-Mc/sec amplitude-modulated r-f equipment of Sec. 17-12. Much experimentation was done, however, with the phase-shifted pulse method of synchronization (Sec. 17-6), and with the 100-Mc/sec frequency-modulated r-f equipment of Sec. 17-13. The former combination is outlined in Fig. 17-22, in which some parts peculiar to this system and not heretofore described are shown.

It is necessary to provide the azimuth data in terms of compass directions rather than aircraft heading. To accomplish this, an a-f wave is passed through a two-phase synchro on the scanner and a two-phase differential synchro controlled by a compass so that the two resulting signals have amplitudes proportional to $\sin \theta$ and $\cos \theta$ respectively, where θ is the scanner orientation with respect to true north (Sec. 13-4). Each of these signals is passed through a phase-sensitive rectifier keyed by the audio oscillation in order to develop the slowly varying voltages necessary to control the sine and cosine delay circuits (Sec. 17-8).

The remainder of the synchronizer is like that shown in Fig. 17-14 except for the provision for radar-beacon switching on alternate cycles. During periods of beacon use, the relay of Fig. 17-14 is to the right, diverting the modulator pulse from the coder to the scale-of-two multi-

vibrator of Fig. 17-22. The square waves from the latter control a switch that alternates between radar and beacon video signals, if the latter are to be relayed. A pulse formed on those cycles in which radar is

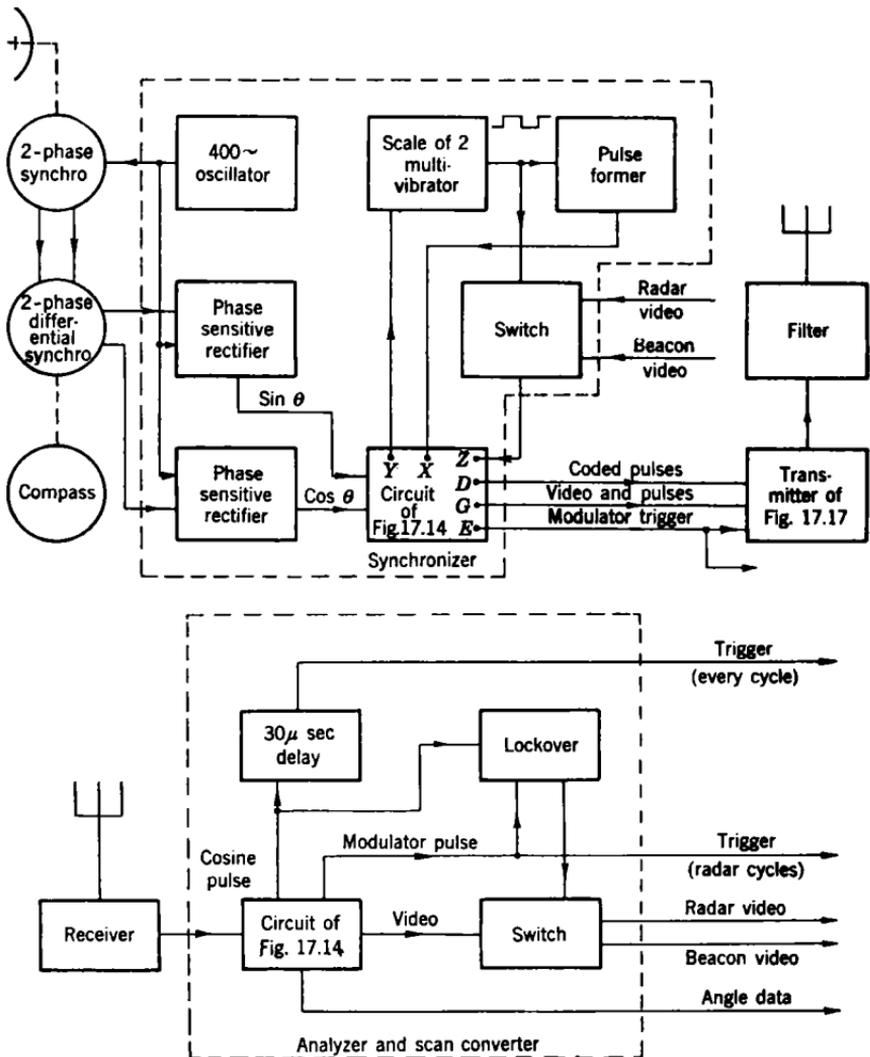


FIG. 17-22.—Air-to-surface relay system.

transmitted is passed to the coder to serve as an identifying pulse in the relay channel. When the beacon signals are not desired, the relay of Fig. 17-14 is reversed and the original pulses go to the coder on every cycle.

The video switch of Fig. 17-14 alternates the video signals (radar or radar alternated with beacon) with the coded pulses, and passes the results to the transmitter (Fig. 17-22), where they ultimately modulate the grid of the r-f power amplifier. The pulses are additionally used to modulate the plates of the same tubes; this arrangement provides a much higher power in the pulses than would otherwise be available. The transmitter power is approximately 50 watts average, 80 watts peak on video signals, and 175 watts peak on pulses.

An r-f filter is provided in the antenna lead. The antenna consists of a vertically polarized dipole mounted on the tail section of the aircraft.

At the receiving station, the energy is received by an antenna consisting of two vertically stacked dipoles and passes through a resonant cavity of loaded Q equal to 100 on its way to the receiver described in Sec. 17-12.

The analyzer and scan converter operate as shown in Fig. 17-14, separating the various signals and providing a simulated scanner motion. The separation of video signals from pulses is aided by an amplitude selector which takes advantage of the higher power in the pulses. This is helpful, especially during the locking-in period. The additional parts necessary to separate the radar and beacon data are shown in Fig. 17-22. The cosine pulse is delayed 30 μ sec to form a trigger available on every cycle, regardless of whether radar or beacon video signals are being transmitted. The modulator pulse indicates those cycles on which radar video is transmitted. A trigger occurring only on the beacon cycles can be formed by an anticoincidence circuit operated by these two pulses. The video switch separating the two types of video is controlled by a flopover which is thrown to the radar position whenever the modulator pulse occurs, and back to the beacon position by the next cosine pulse. The switch then passes beacon signals on every cycle in which the modulator pulse does not occur.

This equipment gave reasonably satisfactory results in its final form. In spite of all the precautions taken, however, the considerable interference from radar and communications equipment, together with the presence of diffraction minima, limited the reliable operating range to about 30 miles when the interference was severe, and to 50 miles or so under reasonably favorable conditions. These figures would be somewhat improved by the use of diversity antennas, but in the absence of tests no figures can be given.

Extensive tests of the 100-Mc/sec frequency-modulated equipment have been made under less severe interference conditions than those faced in tests of the 300-Mc/sec equipment. At 100 Mc/sec, diffraction minima occur only at short ranges where the signal strength is high, and they are much less pronounced than those at higher frequencies, as

indicated by Fig. 17-15. This relative freedom from diffraction effects, together with the higher antenna gains, gave the 100-Mc/sec equipment better performance in the absence of severe interference than that of the 300-Mc/sec system. Under the test conditions, satisfactory results were achieved at ranges up to 100 miles with the airplane flying at 10,000 ft. However, a great deal of interference exists in this frequency band in busy locations. The tests did not give adequate opportunity to observe the effect of this interference in reducing the maximum range since the sites used were relatively isolated.

The phase-shifted pulse method of synchronization (Sec. 17-6) gave results comparable to those of the sine-cosine method with either type of r-f equipment under reasonably interference-free conditions. Since it is slightly simpler and has fewer adjustments, it was therefore somewhat superior under the test conditions. No data are available on its relative performance in the presence of severe interference.

Index

A

- Absorbent materials, 69
- Absorber, of second kind, 69
- Absorbing material, bandwidth of, 71
- Absorption, atmospheric, effect of, 112
- of microwaves, in oxygen, 59
 - in water vapor, 59
- AFC, 453-457
- beacon, 456
 - radar, 454-456
- AI, 200
- AI Mark IV, 201
- AIA, 203
- Air control, high-performance radar for, 592-611
- Air surveillance, high-performance radar for, 592-611
- Aircraft electrical systems, 555
- Aircraft interception (*see* AI)
- Aircraft Radio Laboratory, 15
- Airways, 574
- Alternator, airborne, direct-driven, 557-561
- wave shape of, 557
 - aircraft, alternatives to direct-driven, 560
 - inductor, performance of, 557
 - voltage regulators for (*see* Voltage regulators)
- (*See also* Motor-alternator)
- Altimeter, radio, 143-147
- Altitude signal, 88
- Alvarez, L. W., 291
- A.M.E.S., Type 7, 186
- Type 13, 190
- Amplified back-bias, 459
- Amplifier, gated, 508-510
- i-f, 8
 - design of, 442-449
 - double-tuned, 446
 - single-tuned, 445
- Amplifier, negative feedback, 492
- phase-splitting, 495
 - squaring, 497
 - two-channel, for MTI, 672-675
 - video, 8, 450-453
 - cutoff frequency of, 451
 - rise time of, 450
- Amplitude factor, 557
- Amplitude modulation (*see* Modulation, amplitude)
- AN/APA-15, 306
- AN/APG-1, 202
- AN/APG-2, 202
- AN/APG-15, 206
- AN/APQ-7 scanner, 291-295
- AN/APS-3, 199
- AN/APS-4, 199
- AN/APS-6, 203
- spiral scanner, 290
- AN/APS-10, detailed design of, 616-625
- detailed specifications of, 619
 - general design of, 614
 - performance of, 622
 - scanner, 288
 - units of, 618
- AN/APS-15, 199
- Angle indices, 514-518
- fixed electronic, 514
 - movable electronic, 516
- AN/MPN-1 (*see* GCA)
- Antenna, beamwidth of (*see* Beamwidth)
- coscant-squared, 23-27
 - end-fire, 277
 - gain of (*see* Gain, of antenna)
 - mechanical construction of, 279-280
 - nonscanning, 277-279
 - paraboloid, 272-274
 - polyrod, 278
 - for radar relay, 713-719
 - receiving cross section of (*see* Cross section, effective receiving)

- Antenna, rolled parallel-plate scanning, 304
 Schwartzschild, 295-298
 of SCI height finder, 298-302
 stabilization of (*see* Stabilization; Stabilizer)
 Yagi, 277
- Antenna equation, 271-272
- Antenna feed, 272-274
- Antenna filter for radar relay, 720
- Antenna mount, 271
- Antenna pattern for airborne ground-mapping, 23
- Antenna temperature, 32
- Anticlutler circuits, 460
- Antijamming, 457-460
- AN/TPG-1, 210, 211
- AN/TPS-10, 191
 antenna mount of, 286
- Array, corrugated coaxial line, 303
 of polyrod radiators, 303
- A-and-R-scope, 166
- ASB, 197
- A-scope, 164
 design of, 524-528
- ASD, 199
- ASD-1, 199
- ASE, 196
- ASG, 199
- ASH, 199
- Aspect function, 88
- ASV Mark II, 196
- ATR switch, 7, 407-411
- ATR tube, 407
- Attenuation, by fog or cloud, 60, 62
 of microwaves in atmosphere, 58
 by rain, 61
 by water droplets, 60
 in waveguide (*see* Waveguide, attenuation in)
- Aural detection (*see* Detection, aural)
- Automatic frequency control, 453-457
 (*See also* AFC)
- Autosyns, 487
- Azimuth-pulse removal, 695
- B
- Back-bias, 459
- Back bombardment of magnetron cathode (*see* Magnetron cathode, back bombardment of)
- Back-of-the-dish system (*see* System, back-of-the-dish)
- Baltzer, O. J., 80
- Bandwidth, i-f amplifier, 444
 over-all, of cascaded double-tuned circuits, 448
 of cascaded single-tuned stages, 446
 of receiver (*see* Receiver bandwidth)
- Bartelink, E. H. B., 449
- Beacon, airborne, 246
 azimuth width of reply of, 256
 choice of frequency for, 260
 fixed ground, 248
 with ground radar, 609
 interrogation of, 252-260
 interrogation coding of, 263-264
 frequency, 263
 pulse length, 263
 multiple pulses, 263
 two-frequency interrogation, 264
 overinterrogation of, 265
 portable, 249
 radar, 27, 243-270
 general description of, 243-246
 use for communication, 244, 264
 range performance of, 254
 reply coding of, 264
 gap coding, 264
 range coding, 264
 width coding, 264
 side-lobe interrogation of, 257
 swept-frequency, 262
 traffic capacity of, 265-268
 unsynchronized replies of, 268
- Beacon system, radar, 246-254
- Beam shape, choice of, 600-604
- Beamwidth of antenna, 20, 271
- Bell Telephone Laboratories, 291, 565, 664
- Bendix Aviation, 578, 580
- Blackmer, L. L., 221
- Blind speeds, choice of, 654
 in MTL, 650
- Blocking oscillator, 502
- Boice, W. K., 560
- Breit, G., 13
- British Technical Mission, 15
- Brush, high-altitude, for rotating machines, 561
- B-scope, 171
 design of, 528-532

- B-scope, electrostatic, 528
magnetic, 528
Button, C. T., 578
- C
- Cable, coaxial, 397
Carlson, J. F., 65
Cathode follower, 494
Cathode-ray tube, 475-486
deflection coil of, 477
display projection of, 219
electrostatic deflection of, 476
electrostatic focusing of, 476
magnetic deflection of, 477
magnetic focusing of, 476
types of, 483
Cathode-ray tube screens, 479-483
cascade, 480
dark trace, 483
long-persistence, 480
supernormal buildup in, 482
Cathode-ray tube sweeps, delayed, 522
Cavities, resonant, 405-407
CH (*see* Home Chain)
Chaff, 82
Chain, Home (*see* Home Chain, British)
Channel, signal, 434
Chart projector, 215
Chokes, 397
Chu, L. J., 64
Circuit efficiency, of magnetron, 345
Cities, radar signals from, 101
Clamps, 503-508
switched, 505-508
Close control, 232-240
Clutter, 124-126
rain, 81
Clutter fluctuation, internal, 642-644
measured values of, 643
when radar is moving, 657
Clutter-noise, 651
CMH, 188
Coaxial-type mixer, 417
Coder, triple-pulse, 686
Coding (*see* Beacon, interrogation coding of)
effectiveness of, 688
of pulses, 686
Coherence, ways of producing, between
echo signals and reference signal,
635-638
Coherent oscillator, 632
effect of detuning, 640
for MTI, 662-665
circuit design of, 663
Combined plan and height systems, 192
Complex targets, 73, 75-81
Component, r-f (*see* R-f components)
Compound targets, 73, 81
Computers, dead-reckoning, 215
Conical scan (*see* Scan, conical)
Connectors coaxial-line, 396
type N, 397
Contrast of PPI display (*see* PPI displays, contrast of)
Control (*see* Speed; Voltage; etc.)
Controllers, 235
Corner reflector, 67
Cosecant-squared antenna (*see* Antenna, cosecant-squared)
Counter, V. A., 80
Coupling, for coaxial line (*see* Line, coaxial, coupling for)
waveguide (*see* Waveguide, choke coupling for)
Coverage, high, 50
low, 50
Coverage diagram, 54
Crest factor, 557
Cross section, of aircraft, 76
experimental, 78
corner reflector, 67
cylinder, 66
effective receiving, 20
flat sheet, 65
propeller modulation of, 76
radar, 21, 63
scattering, 21
from sphere, 64
segment of sphere, 66
of ships, 80
CRT (*see* Cathode-ray tube)
Crystal, converter-type, specifications of,
414
for mixer, 412-414
noise temperature of, 413
C-scope, 173
Cut paraboloid, 272
C-w radar systems
(*See also* Radar system, c-w)
CXAM, 180
Cylindrical reflector, 276

D

Dark-trace screens, 483
 Dark-trace tube, 220
 Data stabilization, 311-312
 Data transmission, 283
 potentiometers for, 487
 variable condensers for, 489
 variable transformers for, 487
 autosyns, 487
 resolvers, 487
 selsyns, 487
 synchros, 487
 Data transmitter, angle, 486-492
 D-c restorer, 503
 Deck-tilt error (*see* Error, deck-tilt)
 Decoder, triple-pulse, 687
 Delay line, characteristic impedance of, 671
 folded mercury, 668
 fused quartz, 669
 laboratory type of, 633
 liquid, 667-669
 mercury, design constants for, 670
 supersonic, 667-672
 Delay-line attenuation, 670
 Delay-line circuits, 634
 Delay-line driving circuits, 672
 Delay-line end cells, 669
 Delay-line signal circuits, 672-675
 Delay-line trigger circuits, 675-677
 degenerative, 675
 regenerative, 676
 Delay tank, liquid, 669
 Detection, aural, 134
 Detector, balanced, for MTI, 666
 second, 449
 Dicke, R. H., 32
 Diffraction cross section, 69
 Diffraction phenomena at medium wave-length, 715
 Diode, biased, 504
 charging, 383
 Diode limiters, 504
 Display, double-dot, 174
 one-dimensional, 164-167
 three-dimensional, 174-175
 two-dimensional, 167-174
 pip-matching, 167
 sector, 168
 (*See also* Indicator)

Doppler effect, 125, 629
 Doppler frequency, 128
 Doppler system, bandwidth of, 135
 pulsed, 630
 pulse-modulated, 150-157
 range-measuring, 139-143
 simple, 132-139
 Double-dot display (*see* Display, double-dot)
 Double-tuned circuit, 446
 DuBridge, L. A., 16
 Duct, 56-58
 Dueppel, 82
 Duplexing, 407-411
 Dynamotors, 579-581
 booster armature voltage regulation of, 580
 dual-output, 579
 triple-output, 579

E

Eagle (*see* AN/APQ-7 scanner)
 Eccles, W. H., 497
 Eccles-Jordan circuit, 497
 Echo, from rain, 81
 fluctuations of, 83
 reduction of, 84
 second time around, 117
 from storm, 81
 Eclipse, 559, 574
 Effective height, of ship target, 80
 Eicor Inc., 579
 Eighth-power region, 51
 Electromagnetic energy storage, 356
 Electronic efficiency, of magnetron, 345
 Electronic switches, 503-510
 Electrostatic deflection of beam of CRT
 (*see* Cathode-ray tube, electrostatic deflection of)
 Electrostatic energy storage, 356
 Electrostatic focusing of CRT (*see* Cathode-ray tube, electrostatic focusing of)
 Emsley, Howard M., 561
 Emslie, A. G., 640, 645
 Error, deck-tilt, 309
 in aircraft, 311
 E-scope, 173
 Evans Signal Laboratory, 17

F

- Fairbank, W. M., 80
 Fan beam, 22, 274-277
 Fast time constant circuit, 458
 Feedback amplifiers (*see* Amplifiers, negative feedback)
 Feedback chains, 449
 Feedback pairs, 449
 Ferguson, A. J., 449
 Fighter Control Center, 229-231
 Fittings, type-N (*see* Connectors, type N)
 555th Signal Air Warning Battalion, 229
 Fletcher, R., 353
 Flip-flop, 497
 Flopover, 497, 498-499
 Form factor, 557
 Forward Director Posts, 229-240
 Frequency modulation (*see* Modulation, frequency)
 F-m range measurement, multiple target, 147-149
 F-m range-measuring system, 143-147
 FTC (*see* Fast time constant circuit)

G

- Gain, of antenna, 19
 storage, 44
 Gain-bandwidth product, single-tuned stage, 446
 double-tuned stage, 447
 Gain control, instantaneous automatic, 459
 GCA, 211
 Gear boxes, aircraft, dual-outlet, 560
 General Electric Company, 559, 565, 570, 571, 573, 574
 General Radio Company, 664
 Generator, aircraft, dual-purpose, 560
 400-cps, permanent-magnet-field, 585
 single stroke, 496
 Glass, S. W., 561
 Gordy, W. O., 80
 Goudsmit, S. A., 65
 GPI, 216
 Gray, Marion C., 65
 Ground control of approach (*see* GCA)
 Ground-position indicator (*see* GPI)
 Ground return, 92-96, 154
 Guerlac, H. E., 13

H

- Half-wave line (*see* Line, half-wave)
 Hard-tube pulser (*see* Pulser)
 Heat removal from r-f head (*see* R-f head, heat removal from)
 Height, effective, 52
 Height-finding, 184-196
 by elevation scanning, 189
 by null readings, 184
 by searchlighting, 187
 by signal comparison, 185
 Height-finding system, combined with plan system, 192
 Height indicator, V-beam, 547
 Height indices, 518-524
 Height markers, movable, 520
 Hertz, Heinrich, 13
 Holtzer-Cabot Electric Company, 574, 575, 576, 577, 578
 Home Chain, British (CH), 175-180
 organization of, 226-228
 Homing, 196-203
 on aircraft, 200
 on surface target, 196
 Hubbard, M. M., 560
 Hudspeth, E. L., 80
 Hulsmeyer, 13

I

- IAGC (*see* Gain control, instantaneous automatic)
 Identification systems, (*see* IFF)
 I-f amplifier (*see* Amplifier, i-f)
 IFF, 251
 Impedance, characteristic, 391
 of delay lines, 671
 internal, of pulser, 366
 normalized, of waveguide, 401
 Indication, true-bearing, 311
 Indicator, classification of, 164
 ground position (*see* GPI)
 plan-position (*see* Plan-position indicator & PPI)
 radar, 161-175
 radial time base, 174
 range-height (*see* Range-height indicator)
 spot error, 175
 (*See also* Display)
 Indices, 163

Indices, of range and angle, 513-524
 Inductance, nonlinear, as switch, 381
 Inductance charging, 382
 Integration, of radar information, 38-47
 Interference absorber, 69
 Interference lobes, 50
 Intermediate frequency, choice of, 444
 Intermediate-frequency amplifier (*see*
 Amplifier, i-f)
 Interrogator-responder, 253

J

Jacobsen Manufacturing Company, 585
 Jordan, F. W., 497
 Josephson, V., 77
 J-scope, 166
 Judson Manufacturing Company, 585

K

Katzin, M., 80
 Keep-alive electrode, 410
 Keneipp, H. E., 560
 Klystron, reflex, 414
 Kock, W. E., 85
 K-scope, 167

L

Lawson, J. L., 33
 Lebenbaum, Paul, Jr., 561
 Lee regulator, 574
 Leland Electric Company, 570, 573, 574
 Levoy, L. G., Jr., 560
 LHTR, 208
 Lighthouse tube, 207
 Limiting of video output signal level, 452
 Line, coaxial, 393-398
 coupling for, 396
 rotary joint for, 396
 to waveguide, transition between,
 403
 half-wave, 392
 long, effect of, 393
 matched, 393
 quarter-wave, 392
 stub-supported, 395
 Line-type pulser, (*see* Pulser)
 Linford, L. B., 77
 Loaded Q, 406

Lobe-switching, 203
 Lobes, side, 272
 Local oscillator, 8, 414-416
 MTI, design of, 659-662
 stability of, 638-640
 Local-oscillator stabilities, typical, 661
 Lockover, 497
 L-scope, 167

M

Magnetic focusing of CRT (*see* Cathode-
 ray tube, magnetic focusing of)
 Magnetron, 7, 320-355
 construction of, 321-325
 electron orbits in, 330
 frequency pulling of, 349
 frequency stabilization of, 351
 inductive tuning of, 347
 input impedance of, 346
 modes of operation for, 328
 for MTI, stability of, 640
 output coupling of, 329
 performance chart for, 336
 pulling figure of, 349
 pulse-length limitations on, 346
 pulse power of, 341
 resonant system of, 325
 rising sun, 330
 space-charge distribution in, 335
 strapped, 330
 tunable, C-ring, 347
 tuning of, 347
 wavelength scaling of, 341
 Magnetron cathode, back bombardment
 of, 344
 Magnetron efficiency, 345
 Magnetron instabilities, 353
 sparking, 353
 mode-changing, 353
 Marconi, G., 13
 Markers (*see* Range, Timing, Angle,
 Height markers)
 Massachusetts Institute of Technology,
 16
 Mattress antennas, 274
 Micro-B, 171
 Micro-B display, 531
 Microwave Committee, 15
 Microwave propagation (*see* Propaga-
 tion of microwaves)

- Microwave, reason for use of, 10
attenuation of (*see* Attenuation of microwaves)
- Mixer, 8
coaxial-type, 417
crystal for (*see* Crystal, for mixer)
microwave, 416-418
- Modulation, amplitude, 129, 139
frequency, 130
pulse, 130
from scanning, 136
- Modulator, MTI requirements on, 641
(*See also* Pulser)
- Morse, P. M., 65
- Motor-alternator sets, aircraft, 561-563
speed regulators for, 571-578
starting current of, 561-563
- Motor-alternator sets, speed control of, 574
- Mount, antenna (*see* Antenna mount)
three-axis, 309
- Mountain relief, 96-99
- Moving-target indication (*see* MTI)
- MTI, 626-679
basic principles of, 626-632
blind speeds in, 650
magnetron for, stability of, 640
on moving system, 655-658
noncoherent method of, 656
phase-shift unit for, 655
design of, 663
special test equipment for, 677-679
target visibility in, 649-653
trigger generator for, 634
- MTI cancellation equipment, stability requirements on, 641
- MTI component design, 658-679
- MTI component requirements, summary of, 642
- MTI components (*see* particular component)
- MTI local oscillator, stability of, 638-640
- MTI locking pulse, requirements on, 662
- MTI operating tests, 677
- MTI oscillator stability, testing, 677
- MTI receiver characteristics (*see* Receiver, MTI, characteristics of)
- MTI requirements on modulator, 641
- MTI system, practical, 632-635
- MTI system constants, choice of, 638-655
- MTI targets in clutter, visibility of, 651-653
- MTI transmitter, 658
- Multivibrator, 497-500
cathode-coupled, 499
flip-flop, 498
single-stroke, 498
- Myers, W. L., 283
- N
- National Defense Research Committee (*see* NDRC)
- Naval Research Laboratory, 14
- Navigation, lightweight airborne radar system for, 611-625
radar (*see* Radar navigation)
- NDRC, 15
Division 17, 216
- Neild, W. G., 569
- Network, pulse-forming, 375
- Newton regulator, 566-570
- IX TAC, 229
- Noise, 28-47
- Noise figure, over-all receiver, 32, 441
- Noise fluctuations, 35-40
- Noise power, available, 30
- Noise temperature, of crystal (*see* Crystal, noise temperature of)
- Nonscanning antennas, 277-279
- Nyquist, H., 30
- O
- Oboe system, 246-247
- Odographs, 216
- Olson, V. A., 655, 665
- Optical superposition, of indicator scales, 218
- Oscillator, local (*see* Local oscillator)
phase locking of, 632
- P
- Paraboloid antennas, 272-274
- Peaking, series, 452
shunt, 451
- Pedestal, 216
- Performance chart, for magnetron (*see* Magnetron, performance chart for)
- Phantastron, 500
- Phase locking of oscillators, 632
- π -mode, 326

- Pillbox, 276
 Pip-matching, 203
 Plan-position indicator, 6, 167
 (See also PPI)
 Plotting board, 180°, 238
 vertical, 235
 X-Y, 240
 Polyrod, 278
 Polyrod radiators, array of, 303
 Pound, R. V., 717, 724
 Power, for mobile radar, 585
 prime, supplies for radar, 555-587
 for radar, in aircraft, 555-583
 frequency of, 555
 recommendations for, 582
 at fixed locations, 583
 for large systems, locally generated,
 584
 for shipborne systems, 586
 Power frequencies, in aircraft, standard,
 556
 Power supply, 3-phase a-c radar, 559
 for ultraportable equipment, 585
 vibrator, 581
 PPI, 167
 delayed, 169
 design of, 532-545
 off-center, 168
 open-center, 169
 pre-time-base resolution for, 544
 resolved-current, 538-545
 resolved time base, 534-538
 rotating-coil, 534
 stretched, 170
 three-tone, 553
 using automatic transmitter triggering,
 540
 PPI displays, contrast of, 548-554
 resolution of, 548-554
 Pre-plumbing, 408
 Pressurization, of r-f lines, 283, 420
 PRF, choice of, 598
 Prime movers, small, 586
 Prime power supplies for radar (see
 Power, prime, supplies for radar)
 Projector, chart, 215
 Propagation, free-space, 18
 of microwaves, over reflecting surface,
 47-53
 (see also Microwave propagation)
 PSWR (see Standing-wave ratio, power)
- Pulse, phase-shifted, 697
 sharp, generation of, 501-503
 Pulse cable, 386
 Pulse length, choice of, 596-598
 Pulse modulation (see Modulation, pulse)
 Pulse packet, 122
 Pulse power (see Magnetron, pulse power
 of)
 Pulse-forming network (see Network,
 pulse-forming)
 Pulse-modulated doppler system, 150-157
 Pulse-to-pulse cancellation, 631
 Pulse recurrence frequency (see PRF)
 Pulse transformers, 384-386
 Pulser, 353-390
 basic circuit, 356-360
 driver circuit, 371
 energy sources of, 387
 hard-tube, 367-373
 hard-tube and line-type, comparison
 of, 360-363
 line-type, 358, 374-383
 a-c charging of, 383
 recharging circuit of, 382
 switches for, 377-381
 overload protection of, 364
 Pulser switch, 357
 for line-type pulsers, 377-381
 nonlinear inductance as, 381
 Pulser switch tubes, high-vacuum, 368
- Q
- Q, of cavity, 406
 Quarter-wave line (see Line, quarter-
 wave)
 Quarter-wave plate, 84
- R
- Racons, 246
 Radar, 419
 comparison of, with eye, 1
 c-w, 127-159
 comparison with pulse radar, 123
 ground, use of beacons with, 609
 history of, 13
 limitations of, 116-126
 prime power supplies for (see Power,
 prime, supplies for radar)
 principle of, 3-6

- Radar, pulse, 3
 range performance of, 8-10
 word, 3
- Radar beacon (*see* Beacon, radar)
- Radar components, 6-8
- Radar coverage, effects of surface reflection on, 603
- Radar cross section, 21, 63
- Radar equation, 21-27
- Radar horizon, 53
- Radar indicator, 161-175
 (*See also* Indicator)
- Radar landmarks, 113
- Radar navigation, 108-115
- Radar performance figure, 590
- Radar performance surveys, 592
- Radar relay, 225, 680-735
 AM vs. FM for, 717-719
 antenna for, 713-719
 antenna filter for, 720
 frequencies used for, 716
 100 Mc/sec, equipment, 721-723
 300 Mc/sec, equipment, 719-721
 uses of, 680
 (*See also* Relaying)
- Radar relay analyzer, 682
- Radar relay interference, combating, 685
- Radar relay receiver design, 717
- Radar relay r-f equipment, 682, 713-719
- Radar relay scan converter, 682
- Radar relay synchronizer, 681
- Radar relay system, for airborne radar, 732-735
 ground-to-ground, 726-732
 maximum range of, 724
 microwave, 723-726
 simple, 681
- Radar relay terminal equipment, 682
- Radar relay transmitter design, 717
- Radar scanning patterns (*see* Scanning patterns, radar)
- Radar system, 12
 airborne, lightweight, for navigation, 611-625
 c-w, 127-159
 comparison with pulse radar, 123
 summary of, 157-159
 high-performance, for air surveillance and control, 592-611
 testing, 590
 versatility of, 588
- Radial time base indicator (*see* Indicator, radial time base)
- Radiation Laboratory, 16
- Radio Corporation of America, 240
- Radomes, 314
 airborne, drag of, 315
 electrical transmission of, 316
 examples of, 317-319
 sandwich construction of, 316
 structural design of, 316
- RAF Fighter Command, 226
- Rain echo (*see* Echo, from rain, 81)
- Ramandanoff, D., 561
- Range equation, 595
- Range-height indicator (RHI), 172, 545-547
- Range indices, 518-524
- Range markers, movable, 518
- Range scopes, 165-167
- Range sweep, 162
- Rapid photographic projection, 221
- Ratio, standing-wave (*see* Standing-wave ratio)
- Rayleigh scattering, 63
- Receiver, definition of, 435
 general-purpose, 462-464
 input circuits of, 442
 lightweight airborne, 464-470
 logarithmic, 553
 MTI, adjustment of limit level in, 649
 characteristics of, 646-649
 comparison of types of, 648
 limiting, 647, 665
 lin-log, 647, 665
 over-all noise figure of, 32, 441
 radar, 433-474
 typical, 460-474
 wideband, 470-474
- Receiver bandwidth, 33
- Receiver design, MTI, 665
- Receiving system, typical, 435-441
- Rectangular waveforms, generation of, 496-501
- Reflection, diffuse, 89-92
 specular, 89-92
- Reflection coefficient, 49
- Refraction, 53
- Regulator (*see* Voltage, Speed, etc.)
- Relaxation oscillator, free-running, 496
- Relay radar (*see* Radar relay)
- Relaying, of scanner data, 683

- Relaying, of sine and cosine, general methods of, 701-711
- Repeater, electromechanical, 490-492
synchro-driven, 490
- Resistance, back, of crystal, 413
spreading, of crystal, 413
- Resolution of PPI display (*see* PPI displays, resolution of)
- Resolvers, 487
- Resonance charging, 382
- Resonant cavities (*see* Cavities, resonant)
- Responder beacons, 246
- R-f components, 391-432
- R-f head, 419-432
examples of, 425-432
heat removal from, 421
test points in, 424
- R-f switch, 295
- R-f unit, 420
- Rieke, F. F., 353
- Rieke diagram, 339
- Rising sun magnetron (*see* Magnetron, rising sun)
- Roberts, S., 662, 678
- Robertson, L. M., 561
- Robertson, R. M., 291, 295
- Robinson, C. V., 298
- Rotary joint, for coaxial line (*see* Line, coaxial, rotary joint for)
for waveguide (*see* Waveguide rotary joint)
- Rotary spark gap, 381
- R-scope, 165
- Rubenstein, P. J., 64
- Ryde, D., 62
- Ryde, J. W., 62
- S**
- Sawtooth generators, 510-514
- Scale factor, 162
- Scale markers, 163
- Scale-of-two, 497, 499
- Scan, complex, 281
circular, 281
conical, 188, 205, 281
helical, 281
horizon, 281
Palmer, 282
sector, 199, 281
simple, 280
- Scan, spiral, 281
- Scanner, 271
airborne, installation of, 312
electrical, 291-304
of long-range ground radar, 287
mechanical, 282-291
stabilization of (*see* Stabilization; Stabilizer)
surface-based, installation of, 313
weight of, 283
- Scanning, conical (*see* Conical scan)
signal fluctuations due to, 644-646
- Scanning loss, 43
- Scanning modulation (*see* Modulation, from scanning)
- Scanning patterns, radar, 280-282
- Scanning rate, 116-121
azimuth, choice of, 599
- Schwarzschild antenna, 295-298
- SCI height finder, 298-302
- SCR-268, 203
- SCR-270, 181
- SCR-271, 181
- SCR-521, 196
- SCR-527, 186
- SCR-540, 201
- SCR-584, 207-210
close control with, 238
scanner, 284-286
- SCR-588, 186
- SCR-615, 188
- SCR-627, 186
- SCR-702, 202
- SCR-717, 199
- SCR-720, 201
- Screens, cathode-ray tube (*see* Cathode-ray tube screens)
- Sea return, 92-96
- Second-time-around echo, 117
- Sector control, 230
- Sector display (*see* Display, sector)
- Sector scan (*see* Scan, sector)
- Self-synchronous system, 438
- Selsyns, 487
- Sensitivity-time-control circuit, 460
- Sequencing circuits, 710
- Series gap switch, 377
- Series peaking (*see* Peaking, series)
- Servomechanisms, 491
- Shunt peaking (*see* Peaking, shunt)
- Side lobes, 272

- Signal, interference of, 73
 minimum detectable, 28-47
 minimum discernible, 34
- Signal Corps Laboratories, 14
- Signal fluctuations due to scanning, 644-646
- Signal loss caused by MTI, 651
- Single-tuned circuit, 445
- Skiatron, 220, 483
- Skin depth, 406
- Slug-tuned coil, 442
- SM, 188
- Spark gap, rotary, 381
- Specular cross section, 69
- Speed control, of motor-alternator sets, 574
 motor-, electronic, 578
 of motor, by voltage generator, 574
 tuned-circuit, for motors, 577
- Speed governor, carbon-pile, 578
- Speed regulator, for aircraft motor-alternators, 571-578
 motor-, Lee, 574
- Spencer, R. C., 68
- Stabilization, 304-312
 airborne antenna, 305-308
 frequency, of various components (*see* various components)
 shipborne antenna, 308-312
- Stabilizer, airborne, errors in, 308
 line-of-sight, 306-308
 roll, 306
 shipborne, accuracy of, 311
 stable-base, 306
- Stagger damping, 449
- Stagger-tuned circuits, 448
- Standing waves, 391
- Standing-wave ratio, power, 392
 voltage, 392
- Storage gain (*see* Gain, storage)
- Storage tube, as delay device, 631
- Storm echo (*see* Echo, from storm)
- Straight line charging, 383
- Strapped magnetron (*see* Magnetron, strapped)
- Stratton, J. A., 64
- Straus, H. A., 295
- Structures, radar signals from, 99
- Stub support, broadband, 395
- Subcarrier, phase-shifted, 695
- Subcarrier method, 684
- Superposition, optical, of indicator scales, 218
- Superrefraction, 55
- Switch (*see* ATR, R-f, TR, etc.)
 pulser (*see* Pulser switch)
- Switch tubes, high-vacuum, for pulsers, 368
- Synchronization, 518-524
 incremental angle, method of, 689-695
 jittered-pulse system of, 693
 phase-shift method of, 695-701
- Synchronization methods for radar relay, comparison of, 711-713
- Synchronizer, radar relay, 681
- Synchros, 487
- Synchroscope, 164
- System, back-of-the-dish, 419
 radar (*see* Radar system)
- T
- Targets, complex, 73, 75-81
 compound, 73, 81
 extended surface, 85
- Teleran, 240
- Temperature, noise, of crystal (*see* Crystal, noise temperature of)
- Test equipment for MTI, special, 677-679
- Thyratron, hydrogen, 379
- Time base, 162
- Time-sharing method, 684
- Timing markers, discrete, 520-522
- TR switch, 7, 407-411, 420
- TR tube, 407
 prepulsing of, 607
- Tracking, automatic angle, 207-210
 regenerative, of pulses, 707
- Tracking circuit, locking in of, 710
- Transducers, electromechanical, 667
- Transformer, jog, 295
 (*see also* type of transformer)
- Transmission, c-w, of sine and cosine, 703-711
 data (*see* Data transmission)
 pulse, of sine and cosine, 704
- Transmit-receive switch (*see* TR switch)
- Transmit-receive unit, 420
- Transmitter, data (*see* Data transmitter)
- Transponders, 246
- Trapping, 56-58

- Trigatron, 379
 Trigger generator for MTI, 634
 Tube (*see* Cathode-ray tubes; Klystrons; etc.)
 Tuve, M. A., 13
- U
- Unloaded Q , 406
 U.S. Tactical Air Commands, 229-240
- V
- Van Vleck, J. H., 59, 60
 Variable Elevation Beam (VEB), 189
 VEB (*see* Variable Elevation Beam)
 V-beam (*see* Height indicator, V-beam)
 V-beam radar, 193-196
 Veinott, C. G., 560
 Vibrator power supplies (*see* Power supply)
 Video, output limiting of (*see* Limiting of video output signal level)
 Video amplifier (*see* Amplifier, video)
 Video mapping, 223
 Video mixing, 45
 Visibility, of MTI targets, in clutter, 651-653
 subclutter, 653
 measurement of, 679
 of target, in clear, 649-651
 in MTI, 649-653
 Voltage regulation, booster armature, of dynamotors, 580
 Voltage regulator, for aircraft alternators, 563-566
 carbon-pile, 566-570
 Voltage regulator, carbon-pile, adjustment of, 567
 shock-mounting of, 568
 control of motor speed by, 574
 finger-type, 570
 mechanical, 566-571
 VSWR (*see* Standing-wave ratio, voltage)
- W
- Wallman, H., 448, 449
 Watson-Watt, Sir Robert, 14, 176
 Wave shape of airborne alternators, 557
 Waveform (*see* type of)
 Waveguide, 398-405
 attenuation in, 405
 bends in, 402
 choke coupling for, 401
 to coaxial line, transition between, 403
 cutoff frequency in, 400
 modes in, 400
 power-handling ability of, 404
 of variable width, 291
 wavelength in, 400
 Waveguide rotary joint, 403
 Wavelength, choice of, 604
 in waveguide, 400
 Waves, standing (*see* Standing waves)
 Weinstock, Robert, 64
 Williams, D., 77
 Wincharger Corporation, 574, 576
 Window, 81
 Woodcock, W., 77
- Y
- Yagi antenna (*see* Antenna, Yagi)
 Yaw stabilization, 311