PROPAGATION OF SHORT RADIO WAVES

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—*Marcuvitz*
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*

PROPAGATION OF SHORT RADIO WAVES

Edited by
DONALD E. KERR
ASSISTANT PROFESSOR, DEPARTMENT OF PHYSICS
JOHNS HOPKINS UNIVERSITY

OFFICE OF SCIENTIFIC RESEARCH AND DEVELOPMENT
NATIONAL DEFENSE RESEARCH COMMITTEE

FIRST EDITION

NEW YORK · TORONTO · LONDON
McGRAW-HILL BOOK COMPANY, INC.
1951

TK 6573
.M41
v.13

PROPAGATION OF SHORT RADIO WAVES

COPYRIGHT, 1951, 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.

SCIENCE LIBRARY

MASS. INST. TECH.
APR 17 1951

PROPAGATION OF SHORT RADIO WAVES

EDITORIAL STAFF

DONALD E. KERR LEON B. LINFORD

S. A. GOUDSMIT ALBERT M. STONE

CONTRIBUTING AUTHORS

ARTHUR E. BENT

RICHARD A. CRAIG

WILLIAM T. FISHBACK

JOHN E. FREEHAFER

WENDELL H. FURRY

HERBERT GOLDSTEIN

ISADORE KATZ

DONALD E. KERR

R. B. MONTGOMERY

EDWARD M. PURCELL

PEARL J. RUBENSTEIN

A. J. F. SIEGERT

J. H. VAN VLECK

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 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

Most of the volumes of the Radiation Laboratory Series are devoted to specific radar subjects such as components, systems and their applications, or measurement techniques. This volume, however, treats the phenomena associated with the propagation of short radio waves between terminal points, whether they be the radar antenna serving a dual purpose or the antennas of a communications system. The intention is to present a summary of the state of knowledge in the microwave-propagation field at the close of the war. There has been no attempt to produce either a handbook or textbook, but only an interim report on a rapidly changing subject. An attempt has been made to survey all relevant information that was available, from whatever source, and to summarize as much of it as was feasible.

The preparation of the book was undertaken primarily by the Propagation Group (Group 42), and all of its thirty-odd members contributed either directly or indirectly to the material given here. In addition, substantial contributions have been made by authors who were not members of this group but who worked closely with the group during the war. The division of authorship was to a certain extent arbitrary. The principal criterion was, of course, familiarity with the subject matter, and where possible the people who had made original contributions were favored. There were limiting factors, however, such as the degree of availability of possible authors and the fact that it was impractical to have a large number of writers. Unfortunately, it is impossible to give adequate recognition to all those who have contributed directly or indirectly or even to represent the correct proportion of the contributions of those whose names appear here.

A vast amount of material was available for consideration—much more than could have been presented in one volume. Consequently, some topics have been omitted completely, as, for example, diffraction by trees, hills, and obstacles other than the earth or objects used as radar targets. In this case, as in some others, no significant original work on the subject was done at the Radiation Laboratory, and reviewing work done entirely by others did not appear desirable. Other subjects that have been omitted are the numerous attempts at application of radiometeorology to forecasting of radio and radar propagation performance and the climatological studies needed to make such knowledge useful on

a world-wide scale. In this case, authors were not available to undertake the work. In choosing the meteorological material that was to be presented, it was decided that in the limited time available it was feasible to present only the material considered to have the soundest fundamental background and to eliminate material that involved an appreciable amount of speculation or that would require reworking or further research to put it into the desired form. In general, throughout the book when similar decisions were necessary, they were nearly always made in favor of an exposition of selected material rather than a sketchy, uncritical report of a large amount. We are aware that despite our attempts to include data from many sources our own work tends to predominate; knowing it most thoroughly, we have treated it in greatest detail.

Much of the wartime work was necessarily done in haste without adequate preliminary planning, care in execution, or sufficient analysis of results. If we appear to be overly critical or pedantic here, the reader is asked to understand that this arises, at least in part, from the reaction of the authors to the nature of much of the source material from which the following chapters are formed. We have not hesitated to point out the need for critical examination of the data reviewed here, for such an examination must certainly be one of the first steps in further research in the field. We have also made numerous suggestions for future investigations.

The methods employed in recent propagation research are, we believe, rather important, and we have described them in some detail when it appeared that the description would aid others in future plans. Apparatus details involving radio-frequency techniques are omitted, as most of them are covered in other volumes of this series, but methods of planning experiments and of analyzing results are emphasized. The meteorological instrumentation and new measurement techniques are also emphasized, as they are of utmost importance in investigations of the effects of atmospheric refraction on microwave transmission.

Nomenclature and symbols were matters about which positive decisions were necessary if the book was to be readable. The present choice is the result of considerable deliberation and compromise among several well-established but highly conflicting systems. It embodies as much as possible of the best or of the most firmly established features of each system. A serious attempt has been made to avoid undue overlapping use of symbols but at the same time to adhere to uniform usage throughout the book; some inconsistencies appear inevitable, however.

We have attempted to acknowledge the sources of all our information, even though, unfortunately, these sources are frequently in the form of reports that possibly will never be generally available. Some of the reports cited here are beginning to appear in the literature as this material goes to press, however, and the appropriate footnote references have been inserted wherever possible. When the source of experimental material is not specifically stated, it may be assumed to be the Radiation Laboratory,

but because of the high mobility of ideas, it is not always possible to be certain of their origin. Except for the measurements on oxygen and water-vapor absorption, ship and aircraft cross sections, and a few miscellaneous items, almost all of the Radiation Laboratory material is the work of the Propagation Group or of its close associates.

The information summarized here represents a large investment of effort by many persons and agencies, and it is impossible to acknowledge fully our indebtedness to all of them. Our principal indebtedness is to the remainder of the Propagation Group, whose work contributed so much to this volume. Second, we must acknowledge particular indebtedness to the several authors who at considerable inconvenience to themselves contributed their services long after the termination of the activities of the Radiation Laboratory Office of Publications.

We should like to acknowledge specifically the very great assistance rendered by the several branches of the armed services, who contributed generously in both man power and in equipment such as boats, aircraft, housing facilities, and the many other items necessary to carry on field operations on a large scale. We should like to thank the members of the U.S. Weather Bureau and its several branch offices, whose personnel not only contributed information but in some cases participated in our research program. We are also greatly indebted to Dr. Charles Brooks of the Blue Hill Observatory of Harvard University for his meteorological advice. Most of the aircraft soundings in Chapter 3 were obtained by Robert H. Burgoyne and Earl G. Boardman, who contributed his aircraft and his services as skillful pilot. This work deserves special mention because of its hazardous and highly exacting nature.

In an attempt to ensure accuracy in reporting the work of other groups, we have submitted portions of the manuscript for review to several individuals and organizations. Particular thanks are due to the following people: Sir Edward Appleton, Dr. R. L. Smith-Rose, and the other members of the Tropospheric Wave Propagation Committee in England; Dr. John B. Smyth of the U.S. Navy Electronics Laboratory; A. B. Crawford of the Bell Telephone Laboratories; Professor Paul A. Anderson of Washington State College; Dr. H. H. Beverage of RCA Laboratories; K. A. Norton and Dr. T. J. Carroll of the Central Radio Propagation Laboratory, Bureau of Standards; and Professor C. R. Burrows of Cornell University. The corrections and suggestions offered by these men have been of great value in integrating the descriptions of the work with which they are most familiar. Thanks are also due Norma W. Donelan for her aid in final preparation of the manuscript.

CAMBRIDGE, MASS.
July, 1947

DONALD E. KERR

Contents

FOREWORD vii
PREFACE ix

CHAP. 1. ELEMENTS OF THE PROBLEM 1
(JOHN E. FREEHAFER AND DONALD E. KERR)

EVOLUTION OF THE PRESENT PROBLEMS 1
1·1. The Ionosphere and the Transmission of Long Waves 1
1·2. Optical Properties of Short Waves 3

TROPOSPHERIC REFRACTION 9
1·3. The Effects of Variable Gradients of Refractive Index 9
1·4. The Meteorological Elements and the Modified Index 12
1·5. The Modified Index and Field-strength Distribution 15

ATMOSPHERIC SCATTERING AND ATTENUATION 22
1·6. Radar Echoes from Precipitation 22
1·7. Scattering and Absorption by Particles 23
1·8. Absorption by Gases 25

CHAP. 2. THEORY OF PROPAGATION IN A HORIZONTALLY STRATIFIED ATMOSPHERE 27
(JOHN E. FREEHAFER, WILLIAM T. FISHBACK, WENDELL H. FURRY, AND DONALD E. KERR)

FUNDAMENTAL CONCEPTS 27
2·1. Transmission in Free Space 27
2·2. The Transmission Medium and the Pattern-propagation Factor . . . 34

GEOMETRICAL OPTICS 41
2·3. Ray-tracing Formulas 41
2·4. The Modified Index 50
2·5. Limitations of Ray Methods 53

PHYSICAL OPTICS 58
2·6. The Field from a Dipole in a Stratified Atmosphere near the Earth . . 58
2·7. The Fundamental Theorem 65
2·8. Phase-integral Methods 70

THE LINEAR MODIFIED-INDEX PROFILE 87
2·9. The Properties of Solutions of $d^2y/d\zeta^2 + \zeta y = 0$ 87
2·10. The Field Integral 95
2·11. The Interference Region 98
2·12. The Diffraction Region 109

METHODS FOR CALCULATING FIELD STRENGTH WITH STANDARD REFRACTION . 112
2·13. The Interference Region 113
2·14. The Diffraction Region 122

2·15. The Intermediate Region . . . 125
2·16. Contours of Constant Field Strength . . . 130

THE BILINEAR MODIFIED-INDEX PROFILE . . . 140
2·17. Definition of the Problem and Preliminary Formulation . . . 140
2·18. Methods for Calculating Characteristic Values . . . 146
2·19. Behavior of Characteristic Values and Characteristic Functions for the First Mode . . . 161
2·20. The Problem of Calculating Field Strength for the Bilinear Profile . . . 168

NONLINEAR MODIFIED-INDEX PROFILES . . . 174
2·21. The Linear-exponential and Power-law Profiles . . . 174

CHAP. 3. METEOROLOGY OF THE REFRACTION PROBLEM . . . 181
(RICHARD A. CRAIG, ISADORE KATZ, R. B. MONTGOMERY, AND PEARL J. RUBENSTEIN)

HUMIDITY AND REFRACTIVE INDEX . . . 181
3·1. Vapor Pressure and Saturated Vapor . . . 182
3·2. Water-vapor Concentration . . . 184
3·3. Saturation Temperatures on Isobaric Cooling . . . 186
3·4. Refractive Index of Air at Radio Frequencies . . . 189

VERTICALLY HOMOGENEOUS AIR AND ADIABATIC CHANGES . . . 193
3·5. Adiabetic Temperature Lapse Rate and Potential Temperature . . . 194
3·6. Humidity Lapse in Homogeneous Air . . . 196
3·7. Gradient of Refractive Modulus in Homogeneous Air, Potential Modulus . . . 198
3·8. Characteristic Curves and Mixing . . . 200

REPRESENTATION AND DESCRIPTION OF SOUNDINGS . . . 202
3·9. Approximate Formula for Refractive Modulus . . . 203
3·10. Representation of Soundings . . . 206

EDDY DIFFUSION . . . 208
3·11. Eddy Viscosity and Eddy Diffusivity . . . 208
3·12. Layer of Frictional Influence in Neutral Equilibrium . . . 213
3·13. Logarithmic Distributions in the Turbulent Boundary Layer . . . 215

VERTICAL DISTRIBUTIONS IN NEUTRAL AND UNSTABLE EQUILIBRIUM . . . 219
3·14. Heating from Below . . . 220
3·15. Application of Logarithmic Distribution . . . 223
3·16. Rate of Modification of Unstable Air . . . 226

VERTICAL DISTRIBUTIONS IN STABLE EQUILIBRIUM . . . 228
3·17. Cooling from Below . . . 228
3·18. Shear in Stable Equilibrium . . . 234
3·19. Initially Homogeneous Warm Air over Cold Water . . . 237
3·20. Complex Over-water Modifications . . . 250
3·21. Nocturnal Cooling and Diurnal Cycles . . . 253

OTHER ATMOSPHERIC PROCESSES AND THEIR EFFECT ON M-PROFILES . . . 260
3·22. Subsidence and Subsidence Inversions . . . 260
3·23. Fronts and Frontal Inversion . . . 263
3·24. Sea-breeze Circulations . . . 264

3·25. Horizontal Gradients . . . 267
3·26. Local Variations with Time . . . 268

INSTRUMENTS TO MEASURE TEMPERATURE AND HUMIDITY IN THE LOWER ATMOSPHERE . . . 272
3·27. Psychrograph . . . 272
3·28. Wired Sonde . . . 283
3·29. Aircraft Psychrometers . . . 287
3·30. Resistance Thermometer and Humidiometer . . . 289
3·31. Thermocouples . . . 290
3·32. General Problems Associated with Low-level Soundings . . . 291

METEOROLOGICAL CONSTANTS . . . 292
3·33. Useful Meteorological Constants . . . 292

CHAP. 4. EXPERIMENTAL STUDIES OF REFRACTION . . . 294
(PEARL J. RUBENSTEIN, DONALD E. KERR, AND WILLIAM T. FISHBACK)

ONE-WAY TRANSMISSION OVER WATER . . . 294
Transmission over Massachusetts Bay
4·1. Radio Measurements Program . . . 296
4·2. Meteorological Measurements and Analysis . . . 297
4·3. General Characteristics of Transmission . . . 301
4·4. Comparison with Theory . . . 307
4·5. Transmission under Complex Conditions . . . 315
4·6. Some Statistical Results . . . 319
Transmission Experiments in the British Isles
4·7. The Irish Sea Experiment . . . 322
4·8. South Wales to Mt. Snowdon . . . 328
Transmission along the California Coast
4·9. San Diego to San Pedro . . . 328
Transmission over an Inland Lake
4·10. Flathead Lake . . . 335

ONE-WAY TRANSMISSION OVER LAND . . . 336
4·11. Early Experiments . . . 336
4·12. Summary of General Characteristics . . . 340
4·13. Additional Observations . . . 343
4·14. Discussion . . . 350

RADAR TRANSMISSION . . . 353
4·15. New England Coast . . . 354
4·16. California Coast . . . 361
4·17. Welsh Coast . . . 363
4·18. The English Channel Region . . . 367
4·19. Other Regions . . . 369

SPACE VARIATIONS IN FIELD STRENGTH . . . 373
4·20. Shallow Surface *M*-inversions . . . 374
4·21. Deep Surface *M*-inversions . . . 378
4·22. Elevated *M*-inversions . . . 382

ANGLE MEASUREMENTS ON SHORT OPTICAL PATHS . . . 385
4·23. Measurements of Angle of Arrival . . . 386
4·24. Theoretical Discussion . . . 391

CHAP. 5. REFLECTIONS FROM THE EARTH'S SURFACE . . . 396
(DONALD E. KERR, WILLIAM T. FISHBACK, AND HERBERT GOLDSTEIN)

THEORY OF SPECULAR REFLECTION . . . 396
5·1. Fresnel's Equations for a Smooth Plane Surface . . . 396
5·2. Geometrical Interpretation of the Divergence Factor . . . 404
5·3. Effects of Reflections on Field Strength . . . 406
5·4. Surface Roughness . . . 411

REFLECTION COEFFICIENT OF THE OCEAN . . . 418
5·5. Measurements of Short-time Variations . . . 419
5·6. Interference Measurements over Long Ranges . . . 421
5·7. Interference Measurements at Short Ranges . . . 427
5·8. Interpretation of Measurements . . . 429

REFLECTION COEFFICIENT OF LAND . . . 430
5·9. Measurements over Long Ranges . . . 430
5·10. Measurements at Short Ranges . . . 433
5·11. Measurements of Time Variations . . . 434
5·12. Interpretation of Measurements . . . 435

ERRORS IN RADAR HEIGHT MEASUREMENTS . . . 436
5·13. Qualitative Discussion . . . 437
5·14. Illustrative Examples . . . 441

CHAP. 6. RADAR TARGETS AND ECHOES . . . 445
(DONALD E. KERR AND HERBERT GOLDSTEIN)

THE RADAR CROSS SECTION OF ISOLATED TARGETS . . . 445
6·1. Scattering from a Sphere . . . 445
6·2. Vector Form of Huygens' Principle . . . 454
6·3. Scattering from Planes and Curved Surfaces . . . 456

COMPLEX TARGETS . . . 469
6·4. Radar Cross Section of Aircraft . . . 470
6·5. Radar Cross Section of Ships . . . 472

SEA ECHO . . . 481
6·6. Nature of the Problem . . . 481
6·7. Nature of the Sea Surface . . . 486
6·8. Validity of the Fundamental Assumptions . . . 490
6·9. Frequency Dependence of Sea Echo . . . 494
6·10. Measurements of the Properties of Sea-echo Cross Section . . . 499
6·11. The Fluctuation of Sea Echo . . . 514
6·12. Theories of Sea Echo . . . 518

THE ORIGINS OF ECHO FLUCTUATIONS . . . 527
6·13. The Limitations of System Stability . . . 527
6·14. Atmospheric Variations . . . 531
6·15. Fluctuations in the Space Interference Pattern . . . 535
6·16. Isolated Moving Targets . . . 539
6·17. Interference Phenomena in Complex Targets . . . 547

THE FLUCTUATIONS OF CLUTTER ECHOES . . . 550
6·18. The Nature of Clutter Echoes . . . 550
6·19. The Theory of Clutter Fluctuations . . . 553

6·20. Experimental Techniques in the Study of Clutter Fluctuations 562
6·21. Experimental Results 571

CHAP. 7. METEOROLOGICAL ECHOES 588
(HERBERT GOLDSTEIN, DONALD E. KERR, AND ARTHUR E. BENT)

ORIGIN OF THE ECHO 588
7·1. The Echo from Incoherent Scatterers Distributed in Volume 589
7·2. Evidence of Direct Correlation between Meteorological Echoes and Precipitation 591
7·3. The Approximate Magnitude of Rain Echoes on the Drop Theory . . . 596
7·4. Possible Alternative Theories to Scattering by Drops 598
7·5. Modifications of the Drop Theory 604

THE INTENSITY OF METEOROLOGICAL ECHOES 607
7·6. The Radar Cross Section of Single Drops 608
7·7. Drop-size Distribution 615
7·8. Echoes from Solid Precipitation 618

GENERAL PROPERTIES OF PRECIPITATION ECHOES 621
7·9. Identifying Characteristics 621
7·10. Confusion and Masking of Other Echoes 625

PRECIPITATION ECHO PROPERTIES AND METEOROLOGICAL STRUCTURE 626
7·11. Classification of Echo Types 626
7·12. Thunderstorms 627
7·13. Other Forms of Localized Precipitation 632
7·14. Widespread Precipitation 633
7·15. Cyclonic Storms of Tropical Origin 636

CHAP. 8. ATMOSPHERIC ATTENUATION 641
(J. H. VAN VLECK, E. M. PURCELL, AND HERBERT GOLDSTEIN)

8·1. Properties of the Complex Dielectric Constant 641

THEORY OF ABSORPTION BY UNCONDENSED GASES 646
8·2. Oxygen 648
8·3. Uncondensed Water Vapor 656

MEASUREMENT OF ATMOSPHERIC ABSORPTION 664
8·4. Direct Measurement of Absorption by Oxygen 665
8·5. Measurements of Water-vapor Absorption 666

ATTENUATION BY CONDENSED WATER 671
8·6. Phenomenology of Attenuation by Precipitation 671
8·7. Calculation of Attenuation by Water Drops 674
8·8. Calculation of Attenuation by Precipitation in Solid Form 685
8·9. Measurements of Attenuation by Rain 688

APPENDIX 693
(DONALD E. KERR, A. J. F. SIEGERT, AND HERBERT GOLDSTEIN)

Application of the Lorentz Reciprocity Theorem to Scattering 693
Coherent and Incoherent Scattering from Assemblies of Scatterers 699

NAME INDEX 707
SUBJECT INDEX 713

CHAPTER I

ELEMENTS OF THE PROBLEM

By John E. Freehafer and Donald E. Kerr

This first chapter is intended to serve two purposes: (1) It introduces a number of definitions and concepts that will be useful to orient the reader before he embarks upon the detail of the following chapters. (2) It presents a thumbnail sketch of the type of information that is to follow, but free from the specialized terminology and methods that are frequently necessary in later chapters. We attempt only to indicate a few of the high lights rather than to give a complete outline, as the latter course would lead to undue repetition.

EVOLUTION OF THE PRESENT PROBLEMS

It is interesting to observe that the workers in the present microwave field have returned to the wavelength region in which major features of Maxwell's electromagnetic theory received confirmation through the researches of Hertz. The details of the trend from short to long waves and back again are of no interest here, but some features of the trend are sufficiently relevant to be considered briefly. In addition, the complete identity of many of the problems considered here with well-known problems in optics is emphasized, as experience has shown that familiarity with concepts and methods of optics is of great utility in the microwave field.

1·1. The Ionosphere and the Transmission of Long Waves.[1]—Interest in the propagation of electromagnetic waves over the surface of the earth first became active when Marconi demonstrated in December 1901 that signals could be transmitted across the Atlantic Ocean. During the next 18 years the most eminent mathematicians and physicists of the time contributed to a lively discussion of the quantitative aspects of the four mechanisms that had been proposed to account for the reception of signals beyond the horizon. These were diffraction, reflection from an elevated layer of ionized gases, atmospheric refraction, and transmission by means of surface waves that follow the interface between media of different dielectric properties.

Formulation of the diffraction theory by considering the field due to an oscillating dipole near an imperfectly conducting sphere readily yields an infinite series that unfortunately converges at a hopelessly slow rate

[1] By John E. Freehafer.

when the radius of the sphere measured in wavelengths is large. The attempts of various investigators to approximate this series led to widely divergent results, none of which agreed with the empirically determined Austin-Cohen formula,[1] which seemed adequate to summarize the experimental data available at the time.

Meanwhile other possibilities were being examined. The presence of ionized gases at great heights in the atmosphere was inferred from observations of the spectrum of light from the night sky. These showed the continuous presence of lines associated with the aurora. Furthermore the existence of a conducting layer aloft had been postulated to explain the observed variations of terrestrial magnetism. In 1902 both Kennelly and Heaviside[2] published articles suggesting that long-range radio transmission might be accounted for by reflecting from an elevated conducting surface. Ten years later Eccles[3] contributed an investigation of the mechanism of ionization by solar radiation and presented the fundamental theory of ionic refraction.

The effect of the decrease of refractive index of the air because of the decrease in its density with height was also examined by several investigators.[4] On the assumption that the temperature is constant, it was shown that the radius of curvature of the rays would always be many times the radius of the earth. It was therefore concluded that bending by the atmosphere, although acting in the right direction, is unimportant.

Finally the idea that radio waves could be guided along the surface of the earth much as electromagnetic energy can be conducted along wires was examined theoretically and experimentally. It was observed that high antennas were not necessary to receive signals and that ranges were usually longer over water than over land. Zenneck[5] was able to explain some of the observations by applying to the radio problem the results of his demonstration that Maxwell's equations admit a solution that represents a surface wave guided along the interface between two different

[1] According to the Austin-Cohen formula

$$I \propto \frac{z_1 z_2}{\lambda r} \exp(-\alpha r \lambda^{-\frac{1}{2}})$$

where I is the current induced in a receiving antenna of height z_2 per unit current in the transmitting antenna of height z_1. The range and wavelength are denoted by r and λ. The parameter $\alpha = 1.5 \times 10^{-3}$ (kilometers)$^{-\frac{1}{2}}$ for transmission over sea water. For the data from which this formula was deduced see L. W. Austin, *Nat. Bur. Standards Bull.* 7, 315 (1911).

[2] A. E. Kennelly, *Elec. World*, **39,** 473 (1902); O. Heaviside, *Encyclopaedia Britannica*, 10th ed., Vol. 33.

[3] W. H. Eccles, *Proc. Roy. Soc. A*, **87,** 79 (1912); *Electrician*, **79,** 1015 (1912).

[4] J. A. Fleming, *Principles of Electric Wave Telegraphy and Telephony*, 1919, p. 660.

[5] J. Zenneck, *Ann. Physik*, **23,** 846 (1907); *Lehrbuch der drahtlosen Telegraphie*, 2te. Aufl., Stuttgart, 1913.

media. Further work by Sommerfeld[1] seemed to show that a surface wave of the type discussed by Zenneck would actually be excited by an oscillating dipole at the boundary between a dielectric and imperfectly conducting plane. Sommerfeld's results were later questioned by several investigators,[2] and there is still disagreement concerning the proper nomenclature for and interpretation of the surface wave.

The initial exploratory period was brought to a close with the appearance of two important papers by Watson in 1919. In the first of these[3] he exposed the errors that were responsible for the confusion in the diffraction theory and gave the correct solution. His results showed that diffraction failed completely to account for the long ranges observed. In the second paper[4] he derived the Austin-Cohen formula on the assumption that transmission takes place in the space between two concentric conducting spheres. This was a clear-cut triumph for the reflecting layer hypothesis and directed future effort toward the fruitful field of ionospheric research.

During the next 14 years[5] most of the known phenomena concerned with the transmission of radiation at frequencies up to 30 Mc were observed and satisfactorily explained in terms of measured properties of the ionosphere. Except under conditions of unusually dense ionization, ionic refraction cannot account for the reflection of energy at wavelength much below 9 m. Thus when new developments in experimental technique extended the usable spectrum from 10 down to 1 m and below, interest in the effects of diffraction and atmospheric refraction was revived.

1·2. Optical Properties of Short Waves.[6]—The wavelength range of interest in this book is from about 3 m down to about 1 mm, but no attempt will be made to fix definite limits. The material presented here applies particularly to the region in which the wavelength is small enough that "surface waves" and reflections from the ionosphere are absent but where at the same time it is long enough that large numbers of atomic or molecular resonances in the gaseous components of the atmosphere do not occur in a small wavelength interval. The limitations imposed in this way are flexible and depend, for instance, upon whether the transmission path under consideration is long and near the earth's surface or is short and high in the atmosphere. Much of the material presented here is valid far outside this loosely defined wavelength range, but outside this range it must be applied with a full knowledge of other effects that may mask those under present discussion.

[1] A. Sommerfeld, *Jahrb. drahtlosen Telegraphie*, **4**, 157 (1911); *Ann. Physik*, **81**, 1135 (1926).

[2] See, for instance, F. Noether, *E.N.T.*, **10**, 160 (1933).

[3] G. N. Watson, *Proc. Roy. Soc.*, **95**, 83 (1919).

[4] *Ibid.*, p. 546.

[5] For a review of the activity in this field see H. R. Mimno, *Rev. Mod. Phys.*, **9**, 1 (1937).

[6] By Donald E. Kerr.

Within this wavelength region[1] the relation of the wavelength to the size of material objects at the transmission terminals and along or near the transmission path is of primary importance to the properties of the transmission. In the microwave region most of these objects are comparable in size to the wavelength, and many are much larger than the wavelength. As a consequence it is desirable to discard some of the concepts employed in radio engineering at longer wavelengths and to substitute some of the language and techniques of optics, which have been devised to handle precisely this situation in a different part of the electromagnetic spectrum.

It is convenient to recall that for purposes of classification, optics is divided into the two broad fields of *geometrical optics* and *physical optics*. Geometrical optics is, in general, the simpler of the two, as it predicts propagation of waves along *rays* according to simple geometrical laws without regard to wavelength or phases. Physical optics, on the other hand, results from a solution of the wave equation and thus automatically introduces wavelength, phases, and penetration of waves into shadow regions in which geometrical optics gives no information. As Baker and Copson point out, "There is a general theorem due to Kirchhoff which states that geometrical optics is a limiting form of physical optics. More precisely, the diffuse boundary of the shadow in diffraction phenomena becomes the sharp shadow of geometrical optics as the wavelength of the light tends to zero."[2] We shall find that both physical and geometrical optics are necessary in the study of microwave propagation and that important cases arise in connection with certain types of refraction in which the transition from one to the other may become poorly defined. In general, geometrical optics is used wherever possible because of its comparative simplicity; unfortunately, however, it may easily yield highly erroneous or meaningless results where it is needed most, and in this same region physical optics is likely to be so difficult as to be useless. In such cases one resorts to numerous *ad hoc* artifices and special devices and often bold interpolation in an attempt to fill the gaps left by straightforward theory.

There are many classical experiments of optics that have useful counterparts in the microwave field. One of the best known of these experiments is that of Lloyd's mirror.[3] It is ordinarily performed with a smooth glass plate or a mirror illuminated by a point or line source of light. A screen or photographic plate is arranged to indicate the intensity of the light in the region illuminated both by the source and by reflection at nearly graz-

[1] The range of wavelengths below about 1 m will also be referred to frequently as the *microwave* region.

[2] *The Mathematical Theory of Huygens' Principle*, Oxford, New York, 1939, pp. 79*ff*.; also J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 343. For a very complete discussion see M. Born, *Optik*, Springer, Berlin, 1933, Chaps. 2, 3, and 4.

[3] Jenkins and White, *Fundamentals of Physical Optics*, McGraw-Hill, New York, 1937, pp. 66 and 407.

ing angles from the mirror. A sketch of the experiment is shown in Fig. 1·1. In the shaded region the light intensity results from a superposition of the light from the source and that reflected from the mirror. The light reflected from the mirror appears to come from an image of the source indicated by the broken lines, and the image appears to be out of phase with the source by the amount of the phase shift upon reflection. For grazing incidence this phase shift is just π, as shown by the fact that the intensity is always a minimum (zero for most practical purposes) at the surface of the mirror. It will be shown later that if absorption losses in the mirror are neglected the intensity at a height h on the screen above the surface is given by

$$I = 4I_0 \sin^2 ah,$$

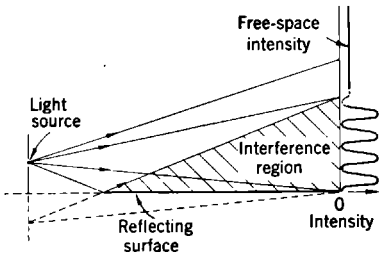


FIG. 1·1.—Interference of light waves as illustrated by the Lloyd's mirror experiment. The vertical scale on the right has been expanded for the sake of clarity.

where I_0 is the intensity from the source in the absence of the mirror (denoted hereafter as the free-space intensity) and a is a constant. This variation is indicated by the curve at the right of the figure.

If the surface of the earth is sufficiently smooth, it produces essentially the effect shown above upon microwave radiation at distances for which the earth's curvature is unimportant. If the surface is so rough that the reflection is diffuse rather than specular, however, the interference pattern disappears and the intensity is essentially I_0, with additional minor irregular variations from point to point in space (for discussion of the details see Chap. 5).

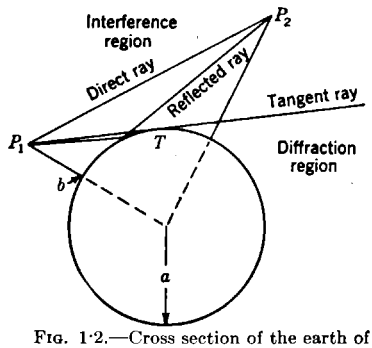


FIG. 1·2.—Cross section of the earth of radius a. A source of radiation is at P_1, and energy reaches P_2 along the two indicated ray paths when P_2 is in the interference region. When P_2 is in the diffraction region, energy reaches it by diffraction around the bulge of the earth.

As the earth is spherical rather than plane, the Lloyd's mirror phenomenon is modified considerably at large distances. Figure 1·2 shows a cross section of the earth, above which a source of radiation at P_1 sends energy along the direct and reflected-ray paths to the point P_2. The *horizon* for the point P_1 is at T, and the *tangent ray* is the ray through P_1 and T, extended indefinitely.[1] The region above bT and above the tan-

[1] Earlier terminology designated P_1T as the *line of sight*. This is not an appropriate term when both terminals are elevated. When *line of sight* is used in this volume, it applies to the direct-ray path P_1P_2.

gent ray beyond T is called the *interference region*, as the intensity can be described in terms of waves following the two paths and adding vectorially to produce a resultant intensity having periodic fluctuations roughly similar to those above the Lloyd's mirror.[1] The region beyond T and below the tangent ray is called the *diffraction region*, as energy penetrates this region by diffraction, which is the process of principal interest in physical optics. The ease with which energy penetrates into this region depends upon numerous factors to be discussed in Chap. 2, but it is of interest here to observe that as the wavelength decreases, the "shadow" cast along the tangent ray by the bulge of the earth becomes more sharply defined. This is another way of saying that the rate of attenuation of intensity with distance or height increases with decreasing wavelength.

As most transmission paths require propagation along a line very close to the tangent ray, where the diffraction shadow begins to be very pronounced, the exact position of the transmission terminals with respect to the tangent ray is important and is usually stated in describing a path. The unobstructed path P_1P_2 in Fig. 1·2 is called an "optical" path; whereas if P_2 is on the tangent ray to the right of T, the common term is a "grazing" path. If P_2 is in the diffraction region, one refers to an "extra-optical" path. An additional correction is almost always made in which an *effective* value of earth's radius of $\frac{4}{3}a$ is used in calculating ray trajectories to allow for refraction effects to be described later. In this case a grazing path may also be referred to as "radio-optical." Both terms will be employed in this book, with an effective earth radius of $\frac{4}{3}$ the true value always being implied.

The effects of the index of refraction of the atmosphere n will receive considerable attention in this book. It will be found that not the absolute value of n (which is roughly 1.0003) but rather its vertical gradient is very important in determining the intensity in the vicinity of the tangent ray and in the diffraction region. In the first steps of analysis of the effects of refractive index gradients it is convenient to resort to geometrical optics, invoking Fermat's principle and Snell's law of refraction derived from it, in order to trace the rays describing the wave paths in the atmosphere. In the interest of precise statement of the problem, *we define the rays as the normals to the surfaces of constant phase of the wavefronts.*[2] When the transmission medium is homogeneous, the rays also give the direction of propagation of the energy of the waves. In the vicinity of sharp corners or in regions in which ray patterns exhibit certain peculiarities as a result of refraction, the energy no longer follows the rays and geometrical optics fails to give meaningful results. Physical optics is then required to continue the analysis. The reader should bear in mind that although

[1] The true situation is considerably more complicated than suggested above. The details will be discussed fully in Chaps. 2 and 5.

[2] See Sec. 2·3 for derivation of the equation for the rays.

raytracing procedures are very useful in the region in which they are valid, they are easily misused and numerical results based on them should be used with caution. This point will recur frequently in later sections.[1]

In order to include the specific problems of radar in a study of the propagation process it is necessary to investigate the scattering of microwave radiation by objects ranging from raindrops to battleships. More specifically, the process of interest is the diffraction of plane waves by these objects. Here again the procedures of physical optics supply the results in the cases in which an exact solution is possible at all, but most practical cases are so complicated that analytical methods usually consist of a series of desperate artifices that lean heavily on geometrical optics where possible. We shall consider here only two simple cases to illustrate the fundamental principles, leaving the details to Chap. 6.

The simplest radar target (and the only one for which the scattering has been calculated with complete rigor) is a sphere. The radius of the sphere is denoted by a, and the wavelength of the incident radiation by λ. The ratio a/λ and the dielectric constant and conductivity of the sphere are sufficient to define the scattering problem. (The effect of the latter two quantities will not be considered here.) We begin with a very small sphere, that is, one for which $a/\lambda \ll 1$. The wave incident upon the sphere excites currents in it which in turn act as the source of a new wave—the scattered wave that we are seeking. These currents radiate a wave that is identical with the field from a classical Hertzian dipole with a suitably chosen electric dipole moment; thus the scattered wave may be computed in a simple manner from well-known formulas. The ratio of the apparent scattering cross section[2] of the sphere to its geometrical cross section (as noted by an observer at the radar measuring the back-scattered wave) is proportional to $(a/\lambda)^4$. This fourth-power dependence is an example of the well-known Rayleigh scattering law, used by Lord Rayleigh to explain the blue color of the sky and in fact applying to any scattering object with dimensions sufficiently small in terms of wavelength.

As a/λ approaches unity, the situation becomes far more complex. The scattered wave no longer behaves like the radiation from a simple dipole, but rather from a group of electric and magnetic dipoles, quadrupoles, and more complicated charge and current distributions within the sphere. The ratio of apparent cross section to geometric cross section increases at a rate less than $(a/\lambda)^4$, finally reaches a maximum, and oscillates with a substantially constant period and with slowly decreasing amplitude of oscillation about the value unity, which it approaches in the limit as a/λ approaches infinity (see Fig. 6·1). This limiting case, for which the back-scattering cross section is precisely equal to πa^2, is the

[1] In particular, see Secs. 1·5, 2·5, and 4·24.

[2] A precise definition of scattering cross section is deferred to Sec. 2·2.

value predicted by geometrical optics, which can give only results independent of wavelength. Because of its simplicity and its resemblance to sections of practical radar targets the sphere has received a great deal of attention.

As an example offering a sharp contrast to the sphere, we consider a flat plate, for convenience assumed to be made of metal and having a fairly simple shape. Whereas the scattering properties of the sphere are independent of orientation of the sphere because of its perfect symmetry, the cross section of the plate depends upon its orientation. If the plate is very small in terms of wavelength, the cross section again follows the $(a/\lambda)^4$ Rayleigh law, where a is now some suitably determined average dimension of the plate. As the plate becomes large, the scattering cross section continues to be a function of wavelength (instead of becoming independent of wavelength as does the sphere), and it also becomes critically dependent upon orientation. For a wave normally incident upon the plate the ratio of back-scattering cross section to geometrical area A is just $4\pi A/\lambda^2$. This expression will be recognized as the gain of an antenna of area A having a uniform field strength over its mouth; thus a large plate viewed at normal incidence behaves like an antenna excited by the currents across its mouth that are actually excited on the plate by the incident wave.[1] As the orientation is changed, the cross section fluctuates rapidly as a result of the multiple lobes of the diffraction pattern of the plate, as shown in Fig. 1·3. The width of the main lobe is roughly λ/a radians. Most of the energy of the incident wave is reflected specularly in the main lobe in a direction such that the angles of incidence and reflection are equal, and the remaining small amount of energy is distributed in the side lobes throughout the remaining solid angle. Similar diffraction phenomena are inherent in every optical instrument and must be considered in detail in design of the instrument.

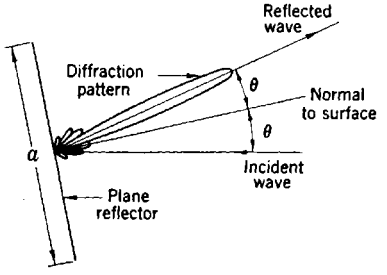


FIG. 1·3.—Scattering from a flat metal plate for which $a/\lambda \gg 1$ (not to scale).

The preceding discussion sketches qualitatively some of the broad features of the behavior of microwaves in relation to their environment, with particular emphasis on the usefulness of concepts and techniques of optics. From here on we consider details of individual sections of the propagation field, borrowing these optical methods freely when it appears advantageous to do so.

[1] This statement is intended to convey a qualitative idea, not to state an exact equivalence (see Sec. 6·2).

TROPOSPHERIC REFRACTION

BY JOHN E. FREEHAFER

1·3. The Effects of Variable Gradients of Refractive Index.—Although waves shorter than about 6 m are seldom and those in the centimeter range never reflected by the ionosphere, there were many observations prior to the opening of hostilities in 1939 to show that the horizon does not always limit the range of ultrahigh-frequency radio facilities. For instance, in 1932 Trevor and Carter reported reception of 69-cm signals over a path from New York to Rocky Point, Long Island, the path being 1.2 times radio-optical. In 1934 Hershberger reported transmission of 75-cm waves over a range of 87 miles, which was five times the sum of the horizon distances from the antennas. During the war, many striking observations of "anomalous" ranges were made. A radar beacon on a frequency of approximately 200 Mc/sec at Bathurst, Gambia, West Africa, was seen on several occasions by aircraft flying below 6000 ft at ranges exceeding 500 miles. Echoes from Nova Scotia were seen at Provincetown, Mass., by a 10-cm radar at a height of 50 ft above sea level. The range in this case was limited to 280 miles by the length of the sweep of the indicator. Perhaps the longest terrestrial ranges observed were obtained on a 200-Mc/sec radar at Bombay, India, which received echoes from points in Arabia 1700 miles away. On the other hand propagation conditions can give rise to strictly limited ranges. For instance, there were occasions when for periods of several hours, centimeter radars in good operating condition on Fisher's Island, New York, were unable to see Block Island 22 miles away although it was optically visible.

As the frequencies involved rule out the ionosphere, attention is directed to refraction by the atmosphere, and in fact, because of the limited vertical extent of the antenna patterns involved in many of the observations, to the lower portion of the atmosphere, the troposphere. It is readily shown, as we shall see in Sec. 1·5, that when meteorological conditions are such that a layer of atmosphere exists in which the index of refraction decreases rapidly with height, radiation of sufficiently short wavelength may be trapped in the layer and guided around the curved surface of the earth by an action analogous to that of a waveguide. Conditions are especially favorable to the formation of such trapping layers or *ducts* in the first few hundred feet of air above the surface of the sea when the temperature of the air is greater than that of the water or aloft at the boundary between air masses of contrasting temperatures.

Experimental evidence of the effect of a surface duct on radio transmission is shown in Fig. 1·4. (The experiment is described in detail in Sec. 4·20.) A 3-watt 10-cm c-w transmitter was located at a height of 25 ft above the surface of the sea. The receiving antenna, a paraboloid with a beamwidth at 10 cm of about 25°, was carried in the nose of an

airplane. At a range of approximately 60 miles the vertical section of field strength shown in Fig. 1·4 was made by descending from several thousand feet to within 20 ft of the surface as rapidly as was consistent with keeping the axis of the paraboloid within a few degrees of horizontal. The total variation in range over the time the data shown in the figure were taken did not exceed 5 miles. The section is noteworthy because of the strong field at an altitude of about 125 ft. Instead of a height-gain effect there is actually a height-loss effect above 125 ft, although over the whole range shown in the figure the field is well above what one would

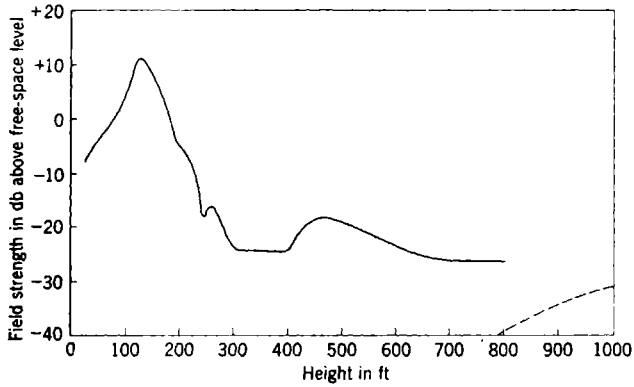


FIG. 1·4.—Vertical field-strength sounding 60 miles from a 10-cm transmitter at a height of 25 ft, June 9, 1943, off the New England coast. The dashed line indicates the expected field strength for standard refraction.

expect under normal refraction conditions. At 100 ft the measured field is 186 db above the field calculated using an effective radius of the earth $\frac{4}{3}$ times the actual radius, and at 800 ft it is still 14 db above. At 60 miles, the height of the tangent ray from the transmitter for a "$\frac{4}{3}$" earth is 1780 ft.

Simultaneously with the radio measurements, low-level meteorological soundings were made by airplane showing that the index of refraction did, indeed, fall off rapidly enough in the surface layer to produce a duct some 200 ft deep (Fig. 4·49).

To support the one-way c-w transmission evidence, a 40-kw 10-cm radar sited near the c-w transmitter was obtaining echoes out to the end of its 110-mile sweep.

Strong refraction effects are often not a transient condition responsible only for sporadic transmission anomalies, but at certain seasons of the year and in certain localities they are present most of the time. This assertion is supported by the following statistics for a 9-cm circuit between Provincetown and Gloucester, Mass. (see Secs. 4·1 to 4·6). The path was 41 miles long with terminals 100 and 136 ft high. It was thus 1.36 radio-optical. For the period covering the months from July through October

1944, the signal averaged over periods of an hour was greater than the free-space level about 15 per cent of the time and greater than that calculated for a "$\frac{4}{3}$" earth 80 per cent of the time. During a particularly favorable week the signal was above the free-space level 25 per cent of the time and above that for a "$\frac{4}{3}$" earth 95 per cent of the time. It must be pointed out that the 4 months from July through October represent the portion of the year when the meteorological conditions over Massachusetts Bay are exceptionally favorable to the formation of surface ducts.

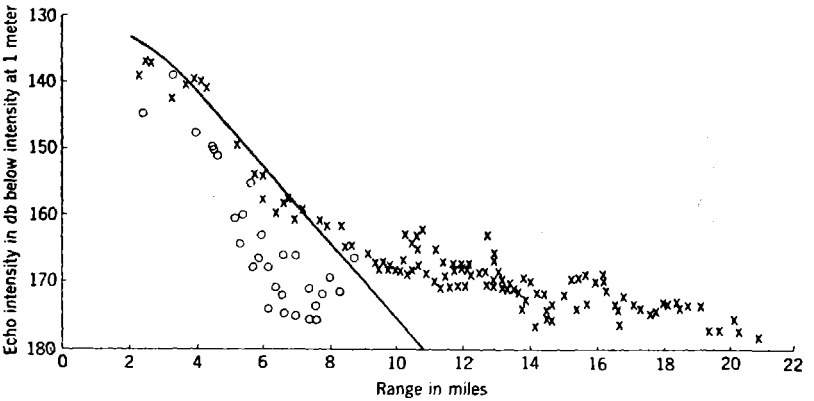


FIG. 1·5.—The effect of refraction over the ocean surface on the relation between echo intensity and range. The crosses and circles represent data taken on two successive days; one day was characterized by a strong surface duct, and the other was not. The wavelength was 10 cm, and the radar and target were at heights of 120 and 6 ft above the ocean surface, respectively. The solid line represents the echo to be expected in the absence of ducts (see text).

The effects of surface ducts are noticed most often as an extension of operating range of both one-way and radar systems operating near the surface. An example is shown in Fig. 1·5, which compares the results of "tracking" a reflector mounted on a boat on two different days. On one day the effects of refraction were negligible, whereas on the second day a strong duct resulted in a large extension of the useful detection range.[1]

The solid line indicates the echo strength to be expected in the absence of ducts and with the reflector perfectly aligned in the vertical plane. The reflector was 10 wavelengths high, however, resulting in a very narrow beam in the vertical plane. As it was impossible to preserve accurate vertical alignment, the discrepancy between absolute values indicated by

[1] The target was made of two square plane reflectors 1 m on a side, mounted at right angles along a vertical center line in such a way that four 90° corners were formed. This reflector (called a *diplane*) was mounted on a small boat with almost no superstructure and was rotated at a constant rate about the common vertical axis of the two planes. The resultant echo consisted of a characteristic sequence of "flashes" that were easily identified because of their timing and relative intensities.

the solid lines and the circles is not surprising. The significant feature of the data is the marked change in the dependence of echo intensity upon range brought about by a duct. This effect will be encountered many times later in our study of the effects of refraction.

1·4. The Meteorological Elements and the Modified Index.—In a uniform atmosphere the rays are straight lines and the trace of the earth on a meridian plane through the transmitter is a circle. The situation is not changed essentially and the analysis is simplified if the earth is regarded as flat and the rays as curved, provided that the difference between the curvatures of the earth and the rays is the same in both cases. To account for the upward bending of the rays in the case of the flat earth, it is necessary to assume a fictitious distribution of refractive index in which the index increases with height. Because, as we shall see in Sec. 2·3, the curvature of the rays is equal to the derivative of the index, it is necessary, in order that the relative curvature of the rays with respect to the earth be the same under flat-earth as under curved-earth conditions, that the gradient of the fictitious index distribution equal the reciprocal of the earth's radius. Thus it seems reasonable that the problem of propagation over a curved earth with uniform atmosphere can be reduced to that of propagation over a flat earth with an atmosphere whose index increases with height. This is in fact true, subject to certain limitations[1] on the height interval in which the field is to be discussed, and the same artifice may be employed to arrive at a flat-earth problem even when the actual index is not constant but varies with height.

The foregoing considerations lead to the concept of the modified index N by which the actual index n is replaced when the earth is treated as flat. The modified index N is related to n by

$$N = n\left(1 + \frac{z}{a}\right) \approx n + \frac{z}{a}, \tag{1}$$

where z is height above the surface and a is the radius of the earth. As N is a number that differs from unity by about 3 parts in 10,000, it is convenient for many purposes to introduce the refractive modulus M, which is related to N by

$$M \equiv (N - 1) \times 10^6. \tag{2}$$

The quantity M has values in the neighborhood of 300. Throughout the following discussion it will be found that N is used in arguments based on electromagnetic theory whereas M is used whenever meteorology and measurements are concerned.

[1] See C. L. Pekeris, *Phys. Rev.*, **70**, 518 (1946). Pekeris concludes that the earth-flattening approximation is valid to within 2 per cent up to ranges of about half the radius of the earth but for wavelengths in the centimeter region the approximation breaks down at altitudes of several thousand feet, the fractional error being proportional to the frequency.

In order to apply the results of meteorological investigations to the study of propagation effects it is necessary to relate the distribution of index of refraction to the distribution of temperature and moisture. For practical purposes at frequencies below 10^4 Mc/sec the index of refraction of a parcel of air at absolute temperature T and pressure p is given by

$$(n - 1) \times 10^6 = \frac{A}{T}\left(p + \frac{Be}{T}\right), \tag{3}$$

where e is the partial pressure of the water vapor present, $A = 79°\text{K/mb}$, and $B = 4800°\text{K}$. The form of this expression is based on theory,[1] the first term on the right being due to the induced dipole moments of all the component molecules and the second being due to the permanent dipole moment of the water vapor molecules. The values of the constants are averages based on published experimental data.[2] That there is no dispersion is confirmed experimentally by the lack of observable amounts of absorption in this region of the spectrum.[3]

From Eqs. (1), (2), and (3), we obtain

$$M = \frac{A}{T}\left(p + \frac{Be}{T}\right) + \frac{z}{a}\,10^6. \tag{4}$$

As the gradients of the meteorological elements have much larger vertical than horizontal components, some justification exists for the usual assumption required by analytic procedure that the refractive modulus and modified index are functions only of height.

By differentiating Eq. (4) we find that

$$\frac{dM}{dz} = \frac{A}{T}\frac{dp}{dz} - \frac{A}{T^2}\left(p + \frac{2Be}{T}\right)\frac{dT}{dz} + \frac{AB}{T^2}\frac{de}{dz} + \frac{10^6}{a}. \tag{5}$$

All of the quantities A, B, T, p, e, and a in Eq. (5) are, of course, positive; hence, the first term on the right is always negative because pressure always decreases with height and the last term is always positive. The two center terms may be either positive or negative depending upon conditions. Thus if there is a temperature inversion, that is, an increase in temperature with height, dT/dz is positive and the term involving dT/dz is negative. Likewise if the water-vapor content decreases with height, the term involving de/dz is negative. Frequently the negative terms contribute more to the sum than the single positive term, and dM/dz becomes negative. For example, when warm dry air passes out over a cool ocean surface, the temperature of the layer of air in contact with the water

[1] Debye, *Polar Molecules*, Chemical Catalog Co., New York, 1929. The formula is discussed in Sec. 3·4.

[2] Sänger, Steiger, and Gächter, *Helvetica Phys. Acta*, **5**, 200 (1932); Stranathan, *Phys. Rev.*, **48**, 538 (1935). See also A. W. Friend, *Jour. Aeronautical Sci.*, **7**, 352 (1940)

[3] This point is discussed in detail in Sec. 8·1.

increases rapidly with height. The resulting stability prevents mixing, and hence the water-vapor content decreases rapidly. Under these conditions dM/dz may be negative over a height interval of several hundred feet above the surface. This, as we shall see in the next section, constitutes a surface duct.

A curve obtained by plotting M as abscissa and z as ordinate is called an *M-profile*. An *N-profile* is obtained in the same way by using N instead of M and differs from the M-profile only in scale. Measurements have shown that on the average M increases with height at the rate of about

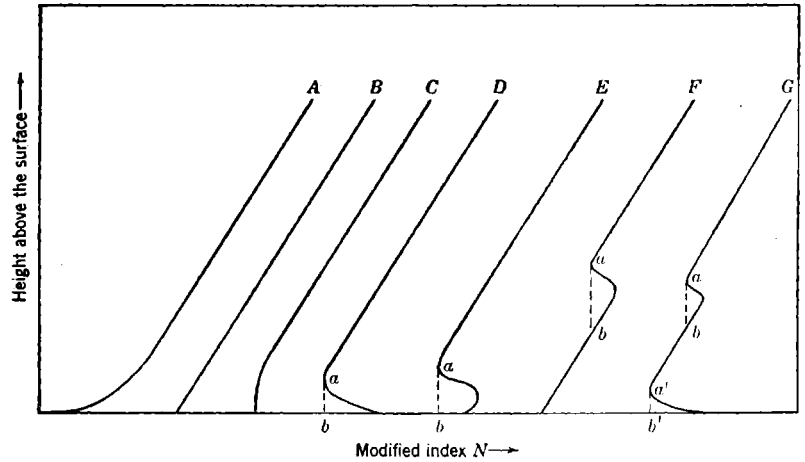


FIG. 1·6.—Idealized modified index profiles: (*A*) Substandard surface layer; (*B*) profile for standard refraction; (*C*) superstandard surface layer; (*D*) superstandard surface layer with surface duct; (*E*) elevated superstandard layer with surface duct; (*F*) elevated superstandard layer with elevated duct; (*G*) surface and elevated superstandard layers with both surface and elevated ducts. In all cases the duct extends from a to b and from a' to b'.

3.6 per 100 ft, and this value has been adopted as the standard M gradient. Likewise a linear M-profile having standard slope is called a standard M-profile, and the field strength calculated on the assumption that the M-profile is standard is defined as *standard* field strength. The latter agrees with values calculated by using an earth's radius equal to $\frac{4}{3}$ its geometrical radius.

Because an M-profile seldom has the standard slope at all heights, it is convenient to divide the atmosphere at a given time into layers such that within each layer the gradient of M is substantially standard or entirely greater than or entirely less than the standard value. Layers in which the gradient is greater than standard are called *substandard*, as in the presence of such layers of sufficient depth the performance of radar and communication facilities is generally poorer than under standard conditions. Similarly, layers for which dM/dz is algebraically less than standard

are called *superstandard* because in their presence performance is generally enhanced. It should be noticed that layers for which $0 < dM/dz <$ (3.6 per 100 ft) as well as those for which dM/dz is negative are all classified as superstandard. The latter, however, are referred to as M or N *inversions* when it is desired to distinguish them from the former. Atmospheric ducts are associated with M inversions. Finally, layers in which the gradient is essentially standard are called standard layers.

Figure 1·6 illustrates the definitions of the three classes of layers and shows in idealized form the types of M-profiles encountered in nature. Curve B represents a standard M-profile. Note that all profiles assume a standard gradient at sufficient height. Profile A shows a substandard surface layer, whereas profile C shows a superstandard surface layer that is not an M inversion. Curves D through G exhibit various kinds of M inversions, resulting, as we shall see in the next section, in both surface and elevated ducts.

1·5. The Modified Index and Field-strength Distribution.—When the N-profile, position of the transmitter, and the electrical characteristics of the earth are prescribed, the problem of determining the distribution of electromagnetic field is completely defined. It is with this problem of interpreting the effect of a given N-profile on the field distribution that we shall be concerned in the present section.

From Snell's law, one can readily show that for rays having inclinations not exceeding a few degrees,

$$\beta = \pm \sqrt{2} \sqrt{N(z) - N(z_1) + \frac{\alpha^2}{2}}, \tag{6}$$

where β is the inclination angle at height z, α is the inclination angle of the ray at the transmitter, and z_1 is the height of the transmitter.[1] For simplicity let us consider first the case of pure diffraction for which $n = 1$. Then Eq. (1) becomes

$$N = \left(1 + \frac{z}{a}\right). \tag{7}$$

Equation (6) suggests a convenient way to visualize the ray family. Refer to Fig. 1·7*a*, which represents the N-profile given by Eq. (7). Through the point on the N-profile at transmitter height z_1 draw a vertical reference line AB. To examine the behavior of a ray leaving the transmitter at angle α draw a line parallel to and at a distance equal to $\alpha^2/2$ to the left of AB. This line will be called the *characteristic* of the ray. In Fig. 1·7*a* lines 1, 2, and 3 are the characteristics of various rays. Then Eq. (6) states that the inclination of the ray at any height is proportional to the square root of the separation between the characteristic and the N-profile

[1] For a derivation of Eq. (6) refer to Sec. 2·3. Equation (6) is Eq. (2·83) with n, the actual index, replaced by N, the modified index, to allow for the curvature of the earth.

at that height. Thus, because the separation between the N-profile and characteristic 1 increases with height, the inclination of the associated ray labeled 1 in Fig. 1·7*b* becomes increasingly steep as the ray arises. On the other hand, the ray that starts downward with the same numerical value of the inclination angle as ray 1 becomes less steeply inclined, until at the height at which the characteristic intersects the N-profile the ray is horizontal and has a turning point. This ray is labeled 1′ in Fig. 1·7*b*.

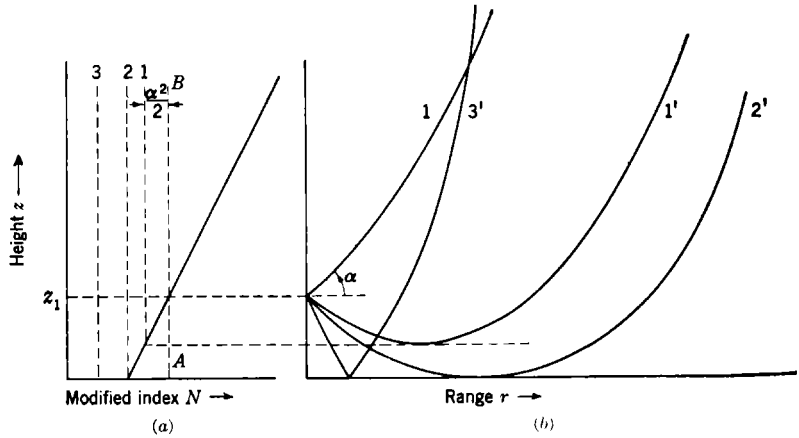


FIG. 1·7.—Ray family associated with a linear N-profile.

Note that a single characteristic corresponds to two rays having slopes at the transmitter that are numerically equal but opposite in sign. In Fig. 1·7*b* the rays that start downward are denoted by primes. Line 2 is the characteristic of the tangent or horizon ray 2′. Line 3 is the characteristic of a ray 3′ that is reflected from the earth's surface. It should be noted that although we have assumed that the index n is constant and equal to unity, the rays are curved. This is as it should be, as the earth has been assumed to be flat. The region to the right of ray 2′ is not reached by the ray family and is therefore the shadow or diffraction zone.

Now let us consider the case shown in Fig. 1·8*a* in which the N-profile has a surface N inversion and z_1, the height of the transmitter, is less than d, the height at which N assumes its minimum value. Let characteristic 2 be drawn tangent to the N-profile at its minimum, and let the distance between characteristic 2 and the vertical reference line AB drawn through the point on the N-profile at the transmitter height be denoted by $\alpha_p^2/2$. Then any ray whose initial angle of inclination is less than α_p has a characteristic that lies between characteristic 2 and line AB. Characteristic 1 is a case in point. As the distance between characteristic 1 and the N-profile decreases with increasing height, the inclination of the associated ray 1 of Fig. 1·8*b* decreases with height. At the height at which the

characteristic intersects the N-profile the ray has zero inclination and hence has a turning point. As the separation between the characteristic and N-profile increases as the height decreases below that of the turning point, the ray becomes more and more steeply inclined as it turns toward the surface. At the surface the ray turns once more and begins another "hop." Thus the ray proceeds by a series of hops and is confined to a layer near the surface. It is easy to see that this same behavior characterizes all rays for which $\alpha < \alpha_p$ and that energy is trapped in the layer of air below height d. It is convenient to refer to such a layer as a *duct*.

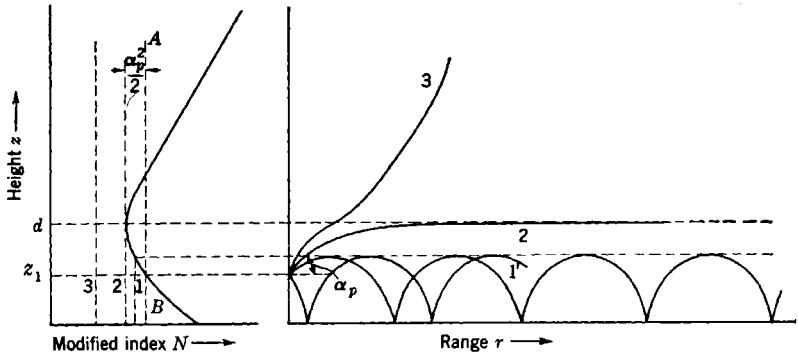


FIG. 1·8.—Modified index profile and ray pattern for a simple surface duct.

Rays whose initial inclinations are greater than α_p have characteristics that lie to the left of line 2, as, for example, characteristic 3. As characteristic 3 does not intersect the N-profile, the associated ray 3 has no turning point. As the ray rises from the transmitter its inclination at first decreases and reaches a minimum but not zero value at height d. From there on the inclination increases monotonically. The ray 3 therefore penetrates the duct. Thus we see that ray 2 whose initial inclination angle is α_p, the so-called *angle of penetration*, divides the family of rays into two groups: those which are trapped and those which penetrate the N-inversion.

It should be noted that in the case of the refraction represented by Fig. 1·8 there is theoretically no region that is not reached by the ray family. There is therefore no shadow region such as is found in Fig. 1·7 and no horizon ray. Instead a region of relatively high energy density near the surface penetrates what is normally the diffraction region and gives rise to abnormally high signal strengths. This is the qualitative explanation of the action of surface ducts.

The two examples just discussed serve to illustrate the general principles of qualitative ray tracing. These same principles may be used to appraise the refraction produced by other types of N-profiles. Thus, for

example, it is readily seen that an "overhanging" elevated inversion such as is shown in profile E of Fig. 1·6 gives rise to a surface duct. On the other hand the elevated inversion of profile F of the same figure produces an elevated duct whose geometrical boundaries extend from the top of the inversion to the height at which the tangent to the minimum intersects the N-profile. Finally it is easy to see that profile G gives rise to both a surface and an elevated duct.

Although ray tracing yields a convenient qualitative picture of the trapping phenomenon, it leads to error in two respects: (1) It gives no hint of the way in which frequency enters into the effect, and (2) it suggests that appreciable effects can be obtained only when the receiver and transmitter are in the duct.

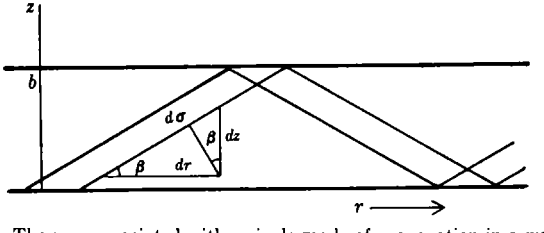


FIG. 1·9.—The rays associated with a single mode of propagation in a waveguide.

Let us consider the frequency dependency first. The physical situation in the presence of a duct is in many ways analogous to that in a waveguide. In both cases the field may be resolved into a sum of elementary waves or modes; but whereas in the case of a guide all of the modes are completely confined to the guide, in the case of a duct all of the modes to some extent penetrate the top of the duct. Those modes for which this penetration or leakage is small may be represented by rays that proceed by a series of hops just as in the case of the guide. To pursue this analogy further let us consider a waveguide filled with material of constant index N. Figure 1·9 shows the mode whose rays are inclined at angle β to the sides of the guide. Let r be the coordinate measured along the axis of the guide and z be the coordinate perpendicular to the axis. Then the phase change due to a displacement $d\sigma$ along a ray inclined at angle β is

$$\frac{2\pi}{\lambda} N\, d\sigma = \frac{2\pi}{\lambda} N\, dr \cos\beta + \frac{2\pi}{\lambda} N\, dz \sin\beta, \tag{8}$$

where λ is the wavelength in vacuo. Because the phase is the sum of a part that depends only on r and a part that depends only on z, it is possible to separate the mode into waves traveling in the r and z directions and to consider the two parts separately. The waves that travel in the z direction bounce back and forth between the surfaces of the guide and interfere with each other. In order that this interference be not chaotic and the

reflections mutually destructive, it is evident that a wave starting upward at a certain height must upon returning to that height after reflection from the top and bottom surfaces have the same phase as it had initially. In this way the interference effects are repeated periodically in each cycle of reflections. Put briefly, this means that the part of the phase depending only on z must be a periodic function of z, and this leads us to write

$$\frac{4\pi}{\lambda}\int_0^b N \sin\beta\, dz - \pi - \pi = 2(m-1), \qquad m = 1, 2, \cdots, \tag{9}$$

where b is the width of the guide. The π's are inserted to represent the phase changes at the reflection points. Thus

$$\int_0^b N \sin\beta\, dz = \frac{m\lambda}{2}; \tag{10}$$

or as N and β are constant,

$$Nb \sin\beta = \frac{m\lambda}{2}, \qquad m = 1, 2, \cdots. \tag{11}$$

This is the familiar relationship among the width of the guide, the angle of inclination of the allowed modes, and the wavelength.

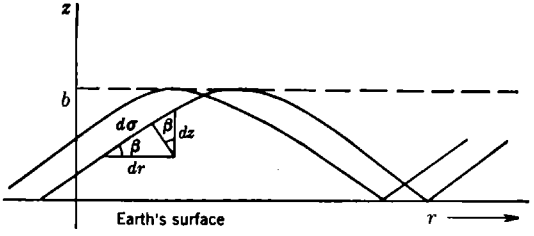


FIG. 1·10.—The rays associated with a single mode of propagation in an atmospheric duct.

In the case of propagation of a single mode in an atmospheric duct as represented by Fig. 1·10, one can use the same arguments with slight modifications. (1) N and β are no longer constants but are functions of z. (2) There is no metal surface producing the reflection at the upper turning point. In fact, at this height the family of rays forms a caustic and there is a change of phase of 90°. This phase change is characteristic of points at which neighboring rays of a family intersect and is well known in optics in connection with rays that pass through a focus.

With these changes we obtain

$$\int_0^b N \sin\beta\, dz = \left(m - \frac{1}{4}\right)\frac{\lambda}{2}, \qquad m = 1, 2, \cdots, \tag{12}$$

where b is the height of the turning point. Now Snell's law states that the product of the index and the cosine of the angle of inclination is con-

stant along a ray.[1] Thus

$$N(z) \cos \beta = N(b). \tag{13}$$

Hence

$$N(z) \sin \beta = \sqrt{N^2(z) - N^2(b)}, \tag{14}$$

and Eq. (12) becomes

$$\int_0^b \sqrt{N^2(z) - N^2(b)}\, dz = \left(m - \frac{1}{4}\right)\frac{\lambda}{2}, \qquad m = 1, 2, \cdots. \tag{15}$$

Equation (15) is a condition on b and shows that there is a discrete set of allowed values of b. This set may be denoted by b_1, b_2, etc., corresponding to $m = 1, 2, \cdots$ and indicated on the N-profile as shown in Fig. 1·11.

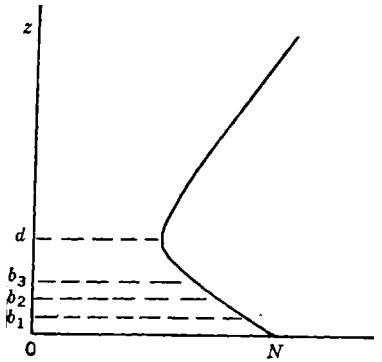


FIG. 1·11.—The discrete set of turning points for the modes of propagation in an atmospheric duct.

As $N(z)$ and $N(b)$ are both nearly equal to unity, Eq. (15) can be simplified to

$$\sqrt{2}\int_0^{b_m} \sqrt{N(z) - N(b_m)}\, dz = \left(m - \frac{1}{4}\right)\frac{\lambda}{2}, \qquad m = 1, 2, \cdots. \tag{16}$$

But $N(z) - N(b_m)$ is just the spacing between the N-profile and the characteristic through the point on the profile for which $z = b_m$; hence, we see that b_m increases both with m and also with λ. Because b_m increases with λ, an upper limit is set to the wavelength of trapped radiation by the fact that the modified index decreases over only a finite height interval d, and b_1 cannot exceed d. Using d for b_1 in Eq. (16) yields the following condition for determining λ_{max}, the longest wavelength trapped:

$$\lambda_{max} = \tfrac{8}{3}\sqrt{2}\int_0^d \sqrt{N(z) - N(d)}\, dz. \tag{17}$$

[1] For a demonstration see Sec. 2·3.

If it is assumed that within the duct $N(z)$ is a linear function decreasing at the rate of 4×10^{-8} per foot, an assumption that experience shows to be reasonable for low ducts, Eq. (17) leads to

$$\lambda_{max} = 0.014 d^{3/2}, \tag{18}$$

where λ_{max} is in *centimeters* and d is in *feet*. Thus we arrive at the following suggestive values.

TABLE 1·1.—MAXIMUM TRAPPED WAVELENGTH AS A FUNCTION OF DUCT HEIGHT

λ_{max}, cm	d, ft
1	20
10	80
100	400
1000	2000

Because atmospheric ducts are generally of the order of several tens of feet in depth and are perhaps never more than 500 or 600 ft, the duct phenomenon is important chiefly for microwave frequencies. As will be pointed out later, *the values of* λ_{max} *given above do not represent cutoff conditions. Although radiation at these or shorter wavelengths is strongly guided, radiation at several times these wavelengths may also be affected strongly by the duct.*

The picture just presented is, of course, oversimplified. When the problem is examined in detail,[1] it is found that the field E at a range r and height z from a dipole at height z_1 in the presence of the earth for which the free-space field is E_0 is given by

$$\left|\frac{E}{E_0}\right| = \sqrt{\lambda r}\left|\sum_{m=1}^{\infty} e^{-i\kappa_m r}\, u_m^{(1)}(z_1)\, u_m^{(1)}(z)\right|, \tag{19}$$

where $u_m^{(1)}$ is a solution of

$$\frac{d^2 u_m^{(1)}}{dz^2} + \left(\frac{4\pi^2}{\lambda^2} N^2 - \kappa_m^2\right) u_m^{(1)} = 0, \tag{20}$$

representing upward traveling waves as $z \to \infty$, and the κ_m's are a set of characteristic values chosen so that $u_m^{(1)}(0) = 0$. The u_m's are normalized so that

$$\int_0^{\infty} [u_m^{(1)}]^2\, dz = 1.$$

Equation (19) is subject to certain simplifying assumptions that need not, however, concern us here.

[1] See Chap. 2.

The expression for $|E/E_0|$ is in the usual form for expressing the diffraction field; and when N^2 is taken as a linear function of z with positive gradient it yields values in agreement with those of the classical treatment of Watson, van der Pol, and Bremmer. When N^2 is not linear, the solution of Eq. (20) is difficult and requires extraordinary techniques involving in general a great deal of arithmetical labor.

The important point here is that the κ's are in general complex with negative imaginary parts. The imaginary part represents attenuation with range and for standard refraction is comparable in size to the real part. The effect of a duct is to reduce the size of the imaginary part of κ and thus to decrease the exponential attenuation with range. If the wavelength is sufficiently short for the depth of the duct, the κ's for the first few modes may be essentially real. In this case it can be shown that[1]

$$\mathrm{Re}\,(\kappa_m) = \frac{2\pi}{\lambda} N(b_m), \tag{21}$$

where the b_m's are calculated by Eqs. (15) or (16). Modes with small attenuation are said to be *trapped*, but the expression is meaningless unless the word "small" is defined. If we assume that an exponential attenuation of less than 5 db per 100 miles is "small," we arrive at essentially the same relationship between the maximum wavelength trapped and the duct height given in Table 1.1.

As the degree of trapping increases, the character of the height-gain function $u_m^{(1)}$ varies as well as the attenuation. For standard refraction the height-gain function is complex with a magnitude that increases exponentially with height. As the trapping increases, the height-gain function of the mode assumes a constant phase throughout the duct and appreciable amplitude within the duct, but essentially zero amplitude without. Thus a strongly trapped mode cannot be excited by a transmitter far from the duct. There are, of course, all sorts of intermediate conditions, and frequently a mode exists that is much less attenuated than standard but nevertheless "leaks" sufficiently to be excited by transmitters above the duct. Transmission may then be considerably improved even though the terminals are not within the duct.

ATMOSPHERIC SCATTERING AND ATTENUATION

BY DONALD E. KERR

1·6. Radar Echoes from Precipitation.—It was pointed out in Sec. 1·2 that echoes from spherical particles much smaller than the wavelength scatter radio waves according to a simple law, but numerical magnitudes were not given. It will be shown in Chap. 7 that both theory and experi-

[1] Equation (21) is readily obtained by comparing Eq. (15) with Eq. (240) of Chap. 2.

ment supply ample evidence that raindrops and other forms of precipitation very often are large enough and exist in sufficient concentration to produce strong radar echoes. Because of the Rayleigh law mentioned in Sec. 1.2 the echoing power of a given rainstorm increases rapidly as the wavelength is decreased. This trend does not continue indefinitely with decreasing wavelength, however.

The mere fact that radar detection of precipitation is possible is of immediate practical value, for it permits advance warning of arrival of storms, often from distances exceeding 100 miles, and is an excellent means for guiding both aircraft and surface vessels around dangerous precipitation areas. The practical experience obtained during the war has built up a large fund of rule-of-thumb information concerning the relation of the characteristics of the precipitation echoes to their effects upon various contemplated activities in the vicinity of the storm.

Of possibly greater long-range importance is the fact that radar storm detection offers the meteorologist a new tool for exploring such poorly understood matters as details of storm structure, turbulence, and the general physical processes connected with formation and growth of precipitation. The subject is extremely complex and thus far has not progressed far beyond the stage of correlation of various characteristic types of meteorological echo with the general types of meteorological structure associated with the precipitation causing the echo. Even these preliminary correlations permit recognition of isolated thunderstorms, cold fronts accompanied by showers or thunderstorms, hurricanes and typhoons, and certain other characteristic meteorological phenomena.

Although radar storm detection is at present only a qualitative research tool, it is reasonable to expect that this is not a permanent condition. A large amount of theoretical work has been done to calculate echoes from precipitation, and the point has been reached where much purely meteorological information on drop sizes, concentrations, and related quantities is needed before great progress is likely. The few wartime attempts at a quantitative experimental approach are discussed in Chaps. 6 and 7, and later work is described in a series of reports issued by the Department of Meteorology of Massachusetts Institute of Technology.

1·7. Scattering and Absorption by Particles.—Precipitation scatters radiation in all directions from a passing wave; and if the particle size and concentration are sufficiently large, this scattering results in an appreciable rate of attenuation of the primary wave. In addition, as the precipitation particles comprise a lossy dielectric, they absorb energy from the wave and convert it into heat. Both phenomena are entirely negligible at wavelengths greater than about 10 cm; but as the wavelength decreases, the scattering and absorption become important, until at wavelengths around 1 cm they place a limitation on transmission over appreciable distances through rain.

When water forms into droplets, as in fog or clouds, each drop acts as a small "blob" of absorbing and scattering material; if the drop is sufficiently small, the dielectric absorption is independent of drop size and is proportional to the total water content per unit volume. Thus, if M is the concentration of liquid water in grams per cubic meter, the absorption coefficient is $0.28\ M$ db/km for a wavelength of 1.25 cm and a temperature of 18°C. As M seldom reaches 1 g/m^3 in fog or clouds, except possibly in very heavy sea fogs, attenuation by fog and clouds is generally negligible.

As the drop size increases, the dielectric absorption also increases, but it now depends upon drop size in a complicated way, increasing at a fairly rapid rate. Also the scattering begins to become important, until its contribution to the total attenuation is comparable to that from absorption when the drops are large and the wavelength is in the neighborhood of 1 cm.

As very little is known of drop sizes and concentrations high above the earth, it is necessary to work with measurable meteorological quantities that are related to the desired quantities in a known manner. One generally uses the rate of precipitation of rain (or snow, etc.) reaching the ground. The theory developed thus far is reduced to a form in which the attenuation is expressed in terms of wavelength and rate of precipitation. This procedure rests upon a number of questionable assumptions and empirical relations (meteorological, not electromagnetic, in nature), but they are probably less important in most cases than the natural inhomogeneity of a rainstorm. Tests have shown that even for path lengths as short as 2 or 3 miles the rainfall is usually so nonuniform in space and time that there is no possibility of obtaining a single value of rate of precipitation which will characterize the rain for our present purpose. Because of this fundamental limitation imposed by the inhomogeneity of the rain no satisfactory method is available for calculating attenuation by precipitation over distances of several miles.[1]

Although attempts at direct calculation of attenuation for long paths is not a fruitful procedure, considerable practical radar experience supports the earlier statement that attenuation at wavelengths greater than about 10 cm is entirely negligible. In the vicinity of 3 cm it begins to be noticeable but is still of little importance, except possibly in the heaviest cloudbursts or tropical rains. At snorter wavelengths it becomes rapidly more important, however, until at wavelengths of 1 cm or less a heavy rainstorm verges on being opaque. For instance, for 1 cm, theory predicts an attenuation in a heavy rain of about 3 db/km and in a cloudburst about 18 db/km. The attenuation increases monotonically with decreasing wavelength and probably reaches a maximum value at about 0.3 cm in

[1] As satisfactory agreement with theory has been obtained over path lengths of a few hundred meters in heavy rain, there is little reason to question the correctness of the existing theory. The difficulty seems to lie in lack of adequate meteorological information. See also Sec. 8·9.

very heavy rain, but the details at this and shorter wavelengths have not yet been explored because of the laborious calculations required.

For frozen precipitation the dielectric loss is considerably smaller than for liquid water. If the particles, such as hailstones, are sufficiently large, the scattering loss can be appreciable at the shorter wavelengths, however. As these two types of loss are interconnected in a fairly complicated manner, we defer further discussion to Chap. 8, where attenuation from all causes is treated in detail.

1·8. Absorption by Gases.—Not only particles but uncondensed gases in the atmosphere are capable of absorbing energy from a passing wave. The mechanism is different from that of the scattering and dielectric absorption described above, however. The reaction occurs in this case because the molecular structure of some gases is such that the individual molecules behave like dipoles; that is, they possess permanent dipole moments, which may be of either electric or magnetic type. These dipole moments furnish the mechanism by which the electric or magnetic field of a passing wave reacts with the molecules, causing them to rotate end over end or to oscillate in many other possible ways. According to quantum theory, to each of the possible modes of vibration there corresponds a possible *energy level* of the molecule.[1] The important features of these energy levels are that in general they are numerous and are separated by very small energy differences. Let us assume that the energy difference between two of these levels is ΔE ergs and that the incident wave has precisely the frequency ν and wavelength λ corresponding to

$$\Delta E = h\nu = \frac{hc}{\lambda},$$

where h is Planck's constant. The result is that the wave delivers to the gas molecules discrete units of energy ΔE that are used to keep the molecules continually in a state of transition from one energy level to the other. They absorb energy from the wave each time the transition is made from the lower to the higher level and reradiate it during the reverse transition. The downward transition is a random process however and does not reinforce the exciting wave, hence the total effect of the continued absorption to reduce the amplitude of the wave. In addition to the pure resonance effect described, damping occurs, and it controls the width and height of the resonance curve just as in familiar mechanical or electrical oscillators.[2]

[1] We ignore degeneracies and other quantum-mechanical details.

[2] All of the effects described here have been investigated in the field of conventional spectroscopy, which because of the enormously higher frequencies involves correspondingly larger energy level spacings. The new field of *microwave* spectroscopy promises to be invaluable in exploring the closely spaced levels and in determining quantities such as the widths of the lines to a much higher degree of accuracy than has been obtainable previously.

Investigation of the principal atmospheric gases reveals that only oxygen and water vapor have both permanent dipole moments and energy level spacings of the appropriate value to be of interest at wavelengths above 1 mm. Oxygen has a magnetic moment, as it is paramagnetic, and water vapor has an electric moment. Together they cause gaseous absorption to become noticeable at a wavelength as small as 1.5 cm and very pronounced below 1 mm.

The spacing of the oxygen energy levels results in resonance peaks at wavelengths of 0.50 and 0.25 cm. The attenuation by atmospheric oxygen at sea level and 20°C is given by theory and experiment as about 14 db/km at the 0.50-cm wavelength and is predicted theoretically to be about 2 db/km at 0.25 cm. There are other resonance peaks at wavelengths less than 0.25 cm, but we shall not consider them here. On the long-wavelength side of the 0.5-cm peak the attenuation drops to about 0.01 db/km at 1.25 cm and 0.0066 db/km at 10 cm.

In the case of water vapor there is a resonance peak at 1.33 cm,[1] as determined by J. M. B. Kellogg and his collaborators at the Columbia Radiation Laboratory (see Sec. 8·5). It was necessary to explore this region in detail, as ordinary spectroscopic data were not capable of predicting the position of the resonance wavelength accurately. The attenuation at the resonance peak was found to be 0.03 db/km per g per m^3 of water vapor. In temperate climates the vapor density on an average day is likely to be around 10 g/m^3, resulting in an attenuation of 0.3 db/km. In the tropics it can be two or three times this value. This is sufficient to produce a noticeable effect on transmission over long distances, particularly in radar applications.

Theory predicts other water-vapor and oxygen absorption peaks at shorter wavelengths with considerably higher values of attenuation. The total effect is to render the atmosphere nearly opaque for wavelengths between 1 mm and the visible region. There are, of course, "windows" in the infrared region, but on the whole this wavelength range is not likely to be useful for transmission over appreciable distances.

[1] More recent measurements place this peak at 1.3481 cm. See C. H. Townes and F. R. Merritt, *Phys. Rev.*, **70**, 558 (1946).

CHAPTER 2

THEORY OF PROPAGATION IN A HORIZONTALLY STRATIFIED ATMOSPHERE

By John E. Freehafer, William T. Fishback, Wendell H. Furry, and Donald E. Kerr

This chapter presents a development of the theory of propagation of very short electromagnetic waves in free space, over a plane earth, and over a spherical earth, including the effects of the electromagnetic properties of the earth and refraction by the atmosphere. In a sense it is two chapters, for it presents detailed specific results and methods for computation of field strength in the absence of refraction or with a linear modified index profile, as well as theory and a few numerical examples for certain types of nonlinear profiles. In order to prevent duplication and to preserve generality as much as possible, however, the presentation has been designed to adhere closely to fundamental electromagnetic theory, from which specific results are developed as special cases. Because of the mathematical complexity of even the simplest cases of nonstandard refraction of practical interest the numerical results for these cases are few.

FUNDAMENTAL CONCEPTS

By Donald E. Kerr

2·1. Transmission in Free Space.—In order to introduce fundamental concepts of propagation in a simple manner and to obtain formulas that will be required throughout the remainder of this volume we begin by considering the properties of transmission in "free space," that is, transmission in vacua between terminals so remote from all other objects that these objects exert no measurable influence on the transmission.

We begin by considering an isolated antenna radiating into free space. In general an antenna radiates more strongly in one direction than in any other; this direction of maximum transmission will be taken as the polar axis of a spherical coordinate system with its origin at the antenna. Any direction from the antenna may be specified by the angles θ and ϕ, where θ is measured as a zenith angle from the polar axis and ϕ is measured about the axis in a plane perpendicular to it. The radiating properties of the antenna are specified in terms of the *antenna pattern function* $f(\theta,\phi)$, *which is defined as the ratio of electric (or magnetic) field strength in the direction* (θ,ϕ) *to the field in the direction of maximum radiation, all fields being deter-*

mined at a fixed distance much larger than the wavelength and the largest dimension of the antenna. In symbols,

$$f(\theta,\phi) \equiv \frac{E(\theta,\phi)}{E_0} \leqslant 1, \tag{1}$$

where E_0 is the field in the direction of maximum transmission, for which $\theta = 0$. In general $f(\theta,\phi)$ is a complex function of the angles, although for most purposes in this volume only real forms for f will be employed. The generalization of the present material for complex f is obvious.

The Poynting vector of the radiated wave $\mathbf{S}(\theta,\phi)$ is directed radially outward from the antenna, and the time average of its magnitude $\overline{S}(\theta,\phi)$ gives the rate of energy flow per unit area in the direction of transmission. It is well known that[1]

$$\overline{S} = \frac{1}{2}\,\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*) = \frac{E^2}{2\eta_0}, \tag{2}$$

where η_0 is 120π ohms, the intrinsic impedance of free space and all field strengths are given in peak values.[2] Combining Eqs. (1) and (2), we have

$$\overline{S}(\theta,\phi) = |f(\theta,\phi)|^2 \overline{S}_0, \tag{3}$$

where S_0 is the value of S in the direction of maximum transmission, corresponding to E_0.

Equation (3) may be used as the basis for a definition of antenna gain. We shall select as our reference antenna with respect to which all gains are defined as an *isotropic* antenna,[3] a hypothetical antenna that radiates equally in all directions and for which consequently $f(\theta,\phi) = 1$. For such an antenna radiating an *average* power P_1 watts,

$$\overline{S}_{\mathrm{iso}} = \frac{P_1}{4\pi R^2}, \tag{4}$$

where R is the distance from the antenna.[4] Let us now find the power P_2 that must be radiated by the directional antenna in order that it may produce the same Poynting vector in its direction of *maximum* transmis-

[1] The distance from the antenna is always large enough that any limited section of the wavefront may be considered plane.

[2] In all developments of electromagnetic theory in this volume the rationalized mks-coulomb system of units will be employed. The permeability and permittivity of free space are $\mu_0 = 4\pi \times 10^{-7}$ henry/m and $\epsilon_0 \approx 10^{-9}/36\pi$ farad/m, respectively, and $\eta_0 \equiv \sqrt{\mu_0/\epsilon_0} \approx 120\pi$ ohms.

[3] See also S. A. Schelkunoff, *Electromagnetic Waves*, Van Nostrand, New York, 1943, Sec. 9·4.

[4] Except where specifically stated otherwise (principally in Chap. 6) all expressions for power denoted by P with appropriate subscripts and all Poynting vectors denoted by $\overline{S}$ will refer to time averages of power (over the period of a single pulse for pulse transmission), whereas all field strengths will be in peak values.

sion that the isotropic antenna produces in *all* directions at the same distance. This power is obtained by integrating Eq. (3) over the surface of a large sphere of radius R, for which the element of area is $R^2 d\Omega$, where Ω is the solid angle viewed from the antenna.

$$P_2 = \int_{(4\pi)} \bar{S}(\theta,\phi)\, R^2 d\Omega = 4\pi R^2 \bar{S}_0 \int_{(4\pi)} |f(\theta,\phi)|^2\, d\Omega. \tag{5}$$

We now *define the absolute antenna gain for transmission* G_t *as* the ratio of the power radiated by an isotropic antenna necessary to produce a given field strength (or $\bar{S}$) at a given distance to the power radiated by the directional antenna producing the same field strength at the same distance in the direction of maximum transmission. In terms of the quantities given above,

$$G_t = \frac{P_1}{P_2} = \frac{S_{\text{iso}} \int_{(4\pi)} d\Omega}{S_0 \int_{(4\pi)} |f(\theta,\phi)|^2\, d\Omega} = \frac{4\pi}{\int_{(4\pi)} |f(\theta,\phi)|^2\, d\Omega}. \tag{6}$$

Thus the antenna gain is uniquely specified by the absolute value of its pattern function.[1]

By previous definition, if $\bar{S}_{\max}$ in Eq. (3) is equal to $\bar{S}_{\text{iso}}$ of Eq. (4), then $P_1 = G_t P_2$, and

$$\bar{S}_{\max} = \frac{P_1}{4\pi R^2} = \frac{P_2 G_t}{4\pi R^2}.$$

Employing this result in Eq. (3) yields the Poynting vector in the direction (θ,ϕ) from the antenna in terms of the radiated power P_2 and the pattern function. Henceforth in this volume P_2 will be called P_t (transmitted power); then

$$\bar{S}(\theta,\phi) = \frac{P_t G_t}{4\pi R^2} |f(\theta,\phi)|^2. \tag{7}$$

Alternatively, the field strength in the direction (θ,ϕ) is, with the aid of Eq. (2),

$$E(\theta,\phi) = \frac{\sqrt{60 P_t G_t}}{R} f(\theta,\phi) = E_0 f(\theta,\phi) \qquad \text{volts/m.} \tag{8}$$

Earlier practice in radio engineering refers antenna gains either to a Hertzian dipole or to a half-wave doublet, for which gains relative to an

[1] The delicate question of the polarization of the radiated wave has been ignored above. There is no difficulty if we assume that all fields are linearly polarized at large distances from the antenna, and we shall tacitly make this assumption until reaching the end of Sec. 6·3; a further discussion is given in Appendix A in connection with scattering by radar targets.

isotropic radiator are $\frac{3}{2}$ and 1.64, respectively. If these antennas are to be used as gain standards, the gain G in all formulas in this volume should be divided by $\frac{3}{2}$ or 1.64. In passing, we note that the field strength in the equatorial plane of a half-wave doublet, obtained from Eq. (8), is

$$E_{\text{rms}} \approx \frac{7\sqrt{P_t}}{R} \qquad \text{volts/m}, \tag{9}$$

where we have shifted temporarily to rms values to obtain the formula commonly used in radio engineering.

Equations (7) and (8) are important relations that will be amplified considerably in later discussion of the coverage problem. It is frequently desired to obtain the locus of all points in space at which E (or $\overline{S}$) has a specified magnitude, this magnitude being selected from requirements to be considered later. The locus, given by Eq (7) or (8), will in general be a three-dimensional surface having a shape determined by the antenna pattern and enclosing a volume determined by the specified value of E and the antenna pattern. This three-dimensional surface is called a *free-space contour diagram*, although for most purposes we shall consider only a cross section in a plane passing through the polar axis, and the contour diagram becomes a polar plot in the coordinates R,θ.

It is usually most convenient for the present purpose to rewrite Eq. (8) in dimensionless form. If we denote by E_s and $\overline{S}_s$ the electric field and the attendant magnitude of the Poynting vector respectively at a reference or standard distance R_s in the direction of maximum transmission, Eq. (8) becomes

$$\left|\frac{E(\theta,\phi)}{E_s}\right| = \sqrt{\frac{\overline{S}}{\overline{S}_s}} = \frac{R_s}{R}\left|f(\theta,\phi)\right|. \tag{10}$$

This equation can be rewritten to give the locus corresponding to a specified *attenuation* of the field:

$$R = R_f\,|f(\theta,\phi)|, \tag{11}$$

where R_f is defined by

$$R_f = R_s\left|\frac{E_s}{E_0}\right| = R_s\sqrt{\frac{\overline{S}_s}{\overline{S}_0}}. \tag{12}$$

The quantity R_f is the numerical scale factor converting the antenna pattern into a free-space contour diagram. It is the greatest distance reached by a contour of the attenuation given by Eq. (12). For instance, if the contour corresponding to a field strength 100 db below the field at a distance 1 m from the transmitter is desired, $R_s = 1$ m, $E_s/E_0 = 10^5$, $R_f = 10^5$ m, and the contour detail is obtained from Eq. (11). Thus Eqs. (11) and (12) afford a means of describing the important features of

the propagation without reference to any property of the transmitter except the antenna pattern, which is of direct interest.[1]

Thus far the only consideration has been the distribution of the radiated energy in free space. In order to be useful for communications purposes part of this energy must be absorbed by a receiving antenna and delivered to the receiver. If the receiver is matched properly to the antenna and is oriented for maximum reception, the received power P_r is given by

$$P_r = \bar{S}A_e, \tag{13}$$

where A_e is the *effective cross section* of the antenna, usually from 50 to 85 per cent of the geometrical cross section. The gain of the receiving antenna G_r is related to A_e by

$$G_r = \frac{4\pi A_e}{\lambda^2}, \tag{14}$$

where λ is the wavelength.[2] If the receiving antenna is oriented in such a way that its direction of maximum reception makes angles of θ',ϕ' with the direction of the incoming plane wave and it is polarized in the same direction as the incoming wave, the expression for received power in Eq. (13) must be multiplied by the square of the antenna pattern function, just as in the case of the transmitting antenna. By combining the receiving antenna pattern with Eqs. (7), (13), and (14) we obtain the fundamental equation for transmission between two antennas in free space:

$$\frac{P_r}{P_t} = \frac{G_t G_r \lambda^2}{(4\pi R)^2} |f_t(\theta,\phi) f_r(\theta',\phi')|^2, \tag{15}$$

where the subscripts t and r designate quantities connected with the transmitter and receiver, respectively. This notation will be used throughout this volume. In particular, if both antennas are aligned for maximum transmission along the line joining them, Eq. (15) becomes

$$\frac{P_r}{P_t} = \frac{G_t G_r \lambda^2}{(4\pi R)^2}. \tag{16}$$

[1] The present method of approach differs from most earlier radio engineering practice, in which contours of constant field strength (in microvolts per meter or a similar unit) are plotted for a specified radiated antenna power. From the point of view of propagation the question of transmitter power and absolute field strength is often irrelevant and at times confusing. In this volume field attenuation will be emphasized; and when apparatus details are introduced, they will appear through the free-space range, to be defined presently. For the microwave region the present method is generally preferable to the older methods. Although the differences are formally little more than a matter of notation, there are numerous practical and conceptual advantages to the methods given here. The force of these remarks is best appreciated after reading Sec. 2·16.

[2] See S. A. Schelkunoff, *op. cit.*, Chap. 9; a paper by H. T. Friis and W. D. Lewis, "Radar Antennas," *Bell System Tech. Jour.*, **26**, 220 (1947), or Vol. 12 of this series.

For Hertzian dipoles Eq. (16) assumes the familiar form $P_r/P_t = (3\lambda/8\pi R)^2$ used commonly in radio engineering practice.

It is highly desirable (and sometimes a vital necessity) in propagation investigations to preserve as clearly as possible a distinction between measured quantities that are properties of propagation alone and those which are properties of the associated equipment. To this end we introduce the *free-space range R_0 for one-way transmission, defined as the largest distance at which a useful signal is received.* Henceforth the minimum useful received power will be designated as $P_{\min}$. From Eq. (16), we obtain R_0.

$$R_0 = \frac{\sqrt{G_t G_r}\,\lambda}{4\pi}\sqrt{\frac{P_t}{P_{\min}}}. \tag{17}$$

The free-space range is the most convenient standard of performance (it represents the *best possible* performance), and it will be used freely throughout the remainder of this book.[1]

If R_0 is now introduced into Eq. (15), we obtain an expression for the *coverage diagram* in terms of R_0 and the patterns of the two antennas:

$$R = R_0\,|f_t(\theta,\phi)f_r(\theta',\phi')|. \tag{18}$$

This equation may be recognized as an extension of Eq. (11), in which the receiving antenna pattern has been added and R_f has the special value of R_0 connecting the system properties with the corresponding value of $|E_s/E_0|$. Usually the receiving antenna is aligned for maximum reception, and the coverage diagram is simply

$$R = R_0\,|f(\theta,\phi)|. \tag{18a}$$

The distinction between a *contour* diagram and a *coverage* diagram is now clear. A *contour* diagram is a three-dimensional surface giving the locus of all points corresponding to an attenuation given by Eq. (12) and enclosing a volume inside which the field strength is always greater than that on the surface. A *coverage* diagram is a special case of a contour diagram, for which $R_f = R_0$, thus introducing the properties of a specific set of apparatus; the coverage diagram encloses a volume inside which the field is always greater than the minimum useful value. As mentioned earlier, in practice these diagrams are almost always used in two rather than three dimensions (in the vertical plane).

[1] Two reminders are appropriate here. (1) $P_{\min}$ and consequently R_0 are in general dependent upon the way in which the apparatus is used, and a single number for a given equipment may not be obtainable. (2) All values of transmitted and received power refer to power actually radiated and to power absorbed by a lossless receiving antenna; hence all the transmission equations given here must be corrected for losses from ohmic resistance or impedance mismatch. Such matters are omitted, as they are of no interest here.

The concepts developed above are easily applied to radar transmission after introducing a parameter describing the properties of the radar target. This parameter is called the *radar cross section*, or alternatively the *back-scattering cross section*, to be denoted by σ, and having the dimensions of area. For reasons that are explained in detail in Chap. 6, σ is in reality only a convenient mathematical fiction, as for most practical targets it cannot be assigned a unique numerical value. Until reaching Chap. 6 we shall ignore difficulties of this kind, however, and we shall formulate the transmission equations on the assumption that somehow a meaningful value of σ has been obtained. The definition of σ is *the area intercepting that amount of power which, when scattered isotropically, produces an echo equal to that observed from the target.*[1]

From this definition, we see that the *back-scattered* Poynting vector $\overline{S}^s$ at a distance from the target R is related to σ and the *incident* Poynting vector $\overline{S}^i$ by

$$\overline{S}^s = \frac{\overline{S}^i \sigma}{4\pi R^2}, \tag{19}$$

or in a form that is a useful analytic definition,

$$\sigma = 4\pi R^2 \frac{\overline{S}^s}{\overline{S}^i}. \tag{20}$$

The power received by the radar antenna is, from the previous discussion,

$$P_r = \frac{A_e \overline{S}^i \sigma}{4\pi R^2} \left| f_r(\theta', \phi') \right|^2 = \frac{\overline{S}^i G_r \lambda^2}{(4\pi R)^2} \left| f_r(\theta', \phi') \right|^2 \sigma. \tag{21}$$

The incident Poynting vector $\overline{S}^i$ can now be obtained from Eq. (7); and when it is employed in Eq. (21) we have the *free-space radar transmission equation*,

$$\frac{P_r}{P_t} = \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \left| f(\theta, \phi) \right|^4. \tag{22}$$

(For convenience we shall always assume identical transmitting and receiving antennas for radar systems.)

Henceforth in this volume in any discussion involving σ, unless there are specific statements to the contrary, we intend it to be the quantity σ appearing in the radar transmission equation in the form of Eq. (22).

In a manner similar to that used for one-way transmission, *we define the free-space radar range R_0 as the maximum value of range that reduces P_r*

[1] It is tacitly assumed in most applications that the wave incident on the target is linearly polarized and only that component of the back-scattered wave with similar polarization is considered in determining σ. More general situations involve complications that will not be considered here. For further discussion see the end of Sec. 6·3 and Appendix A.

to its smallest useful value. From Eq. (22) we obtain

$$R_0 = \sqrt{\frac{G\lambda}{4\pi}} \sqrt[4]{\frac{P_t}{P_{\min}} \frac{\sigma}{4\pi}}. \tag{23}$$

This equation should be compared with its analogue for one-way transmission [Eq. (17)]. If we assume for the sake of simplicity that for a given system the ratio $P_t/P_{\min}$ is comparable for both radar and one-way operation (not an unreasonable assumption for some purposes), Eqs. (17) and (23) show that

$$R_0(\text{radar}) = \sqrt[4]{\frac{\sigma}{4\pi}} \sqrt{R_0(\text{one-way})}. \tag{24}$$

It is clear that increasing radar range by increasing system performance is considerably more expensive for radar than for one-way transmission.

The *free-space coverage diagram for radar* may be employed by combining Eqs. (23) and (22):

$$R = R_0 |f(\theta,\phi)|, \tag{25}$$

which is identical with Eq. (18*a*) after allowance is made for the difference in definition of R_0 for the two cases. The significance of the coverage diagram for radar is identical with that discussed for one-way transmission if proper interpretation is made in terms of the properties of σ (see Chap. 6).

The results of this section apply only to free-space conditions, which in some applications may be approximated very satisfactorily, whereas in others the role of the surroundings may predominate in determining the properties of the transmission. We proceed now to this second case.

2·2. The Transmission Medium and the Pattern-propagation Factor.—Insertion of reasonable numerical values into the equations of the preceding section show that the free-space ranges predicted by the formulas given there are often several orders of magnitude greater than the values obtained in practice. The reason for this discrepancy is, of course, that operation of the equipment in the atmosphere and near the surface of the earth often does not correspond even remotely to operation in free space. We shall now discover that instead of discarding the concepts of the preceding section, however, we shall continue to apply them with appropriate modifications; in fact, it is in obtaining these modifications that we are led naturally into the problems of the several branches of the propagation field.

The presence of the earth complicates the situation in three ways: (1) The earth's surface reflects or scatters radiation and in certain regions produces an interference pattern; (2) the earth casts a shadow and gives rise to diffraction phenomena; and (3) the earth's atmosphere is inhomogeneous and can produce violent refraction effects. For purposes of formulation it is convenient to group all of these factors into a single quantity called the *pattern-propagation* factor.[1]

[1] D. E. Kerr and P. J. Rubenstein, "Introduction to Microwave Propagation," RL Report No. 406, Sept. 16, 1943.

The pattern propagation factor F is defined as the ratio of the amplitude of the electric field at a given point under specified conditions to the amplitude of the electric field under free-space conditions with the beam of the transmitter directed toward the point in question. In symbols,

$$F \equiv \left|\frac{E}{E_0}\right|, \tag{26}$$

where as before E_0 is the magnitude of the free-space field at a given point with the transmitting antenna oriented directly toward the point and E is the field to be investigated at the same point. It will be seen that F involves both the effects of the external features of the transmission path and the pattern function $f(\theta,\phi)$ of the transmitting antenna but is independent of other properties of the source. In accordance with earlier remarks, it is desirable to obtain as complete separation as possible of the source and the transmission medium, and experience has shown this method to be most convenient. Only the component of E having the same polarization as E_0 is generally useful in computing F.

For a given orientation of the transmitting antenna, F is a scalar function of position. Furthermore, F^2 is the ratio of the magnitude of the Poynting vector at the field point under the specified conditions to the magnitude of the Poynting vector at the same distance along the direction of maximum transmission under free-space conditions. In the special case of free-space transmission, $F = |f(\theta,\phi)|$. In the presence of the earth, F involves interference, diffraction, and refraction effects and is not a function of angle alone (although cases will be shown at a later point in which for some purposes F is a function only of angle of elevation, to a good degree of approximation).

It follows directly from the definition of F that for *one-way transmission the generalized transmission equation is*

$$\frac{P_r}{P_t} = \frac{G_t G_r \lambda^2}{(4\pi R)^2} F^2 \left|f_r(\theta',\phi')\right|^2. \tag{27}$$

In the radar case, the power of the isotropic source equivalent to the target is $(P_t G_t \sigma/4\pi R^2)\, F^2$. Because of the reciprocity relationship it turns out that F^2 is the factor to be applied to convert the free-space energy flow into the energy flow that, if it were incident on the antenna in the direction of the maximum of the antenna pattern, would produce the same signal as the target.[1] Thus *the generalized form of the radar transmission equation is*

$$\frac{P_r}{P_t} = \frac{G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} F^4 \tag{28}$$

[1] Consider a transmission path whose terminals consist of an arbitrary antenna having a gain G and an isotropic antenna whose gain is unity and whose cross section

Equations (27) and (28) serve to formulate the problem of field-strength calculation in the presence of the earth and the atmosphere. They are subject to the restriction that the waves incident upon either the antennas or the targets are essentially plane in spite of complications from the presence of the earth or atmosphere. Furthermore, Eqs. (27) and (28) follow from Eqs. (15) and (22) as a result of the definition of F, in which all of the difficulties of propagation theory are hidden.

If the definitions of R_0 given by Eqs. (17) and (23) are employed in Eqs. (27) and (28), respectively, the *fundamental equation for the coverage diagram including the effects of earth and atmosphere* is obtained:

$$R = R_0 F. \tag{29}$$

This, however, is an implicit relation, as F is in general a pronounced function of R, and it may appear that no simplification has been obtained by employing the form of Eq. (29), but it will be seen in later sections that Eq. (29) and its companion equation

$$|E| = |E_0|\, F \tag{30}$$

are, in fact, extremely useful relations.

The remainder of this chapter will, for the most part, be devoted to a discussion of the techniques for obtaining F.

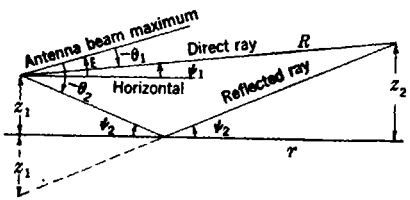


FIG. 2·1.—Geometry of reflection from a plane earth. The convention for signs of the angles is explained in the text.

Simple Properties of the Pattern-propagation Factor.—In order to display the qualitative properties of the pattern-propagation factor in the interference region we shall assume that the earth is a smooth, plane, reflecting surface without an atmosphere. Let a transmitting antenna having a pattern $f(\theta)$ in the vertical plane be located at height z_1 above the earth as indicated in Fig. 2·1.

---

by Eq. (14) is $\lambda^2/4\pi$. Let power P_t be radiated by the arbitrary antenna. Then the power P_{ri} received by the isotropic antenna is

$$\frac{P_{ri}}{P_t} = \frac{G}{4\pi R^2} F^2 \frac{\lambda^2}{4\pi}.$$

If on the other hand the same amount of power is radiated by the isotropic radiator, the power P_r received by the arbitrary antenna can be written

$$\frac{P_r}{P_t} = \frac{1}{4\pi R^2} \frac{G\lambda^2}{4\pi} (F')^2,$$

where F' takes into account the presence of the earth and the fact that the antenna is not necessarily pointing at the isotropic source. From the reciprocity relationship (Appendix A) which holds even in the presence of the earth and attending complications, $P_{ri} = P_r$, and we see that $F^2 = (F')^2$.

The angles θ_1 and θ_2 are measured between the direction of maximum transmission of the antenna pattern (which makes an angle ξ with the horizontal) and the direct and reflected rays, respectively. At a distance R, large compared with the dimensions of the antenna, the field in the direction of maximum transmission under free-space conditions is denoted by E_0.

At a field point at distance R in the direction inclined at angle ψ_1 with the horizontal, the field will be the sum of E_d, the field due to a direct wave, and E_r, that from a wave reflected from the earth and appearing to come from the image of the source at a height $-z_1$. From the definition of f,

$$E_d = E_0 f(\theta_1). \tag{31}$$

From elementary geometry the length of the direct-ray path is

$$\sqrt{r^2 + (z_2 - z_1)^2},$$

and the length of the reflected-ray path is

$$\sqrt{r^2 + (z_2 + z_1)^2}.$$

The *path-length difference* between the two paths ΔR is easily obtained by expanding each path length by the binomial theorem and subtracting; thus,

$$\Delta R = \frac{2z_1z_2}{r}\left(1 - \frac{z_1^2 + z_2^2}{2r^2} + \frac{3z_1^4 + 10z_1^2z_2^2 + 3z_2^4}{8r^4} - \cdots\right). \tag{32}$$

This expression is cumbersome, but fortunately in almost all cases of practical interest ψ_1 and ψ_2 are very small. The customary approximation is to retain only the approximate value

$$\Delta R \approx \frac{2z_1z_2}{r}, \tag{33}$$

which will be used throughout the remainder of this volume for all plane-earth cases (with $r = R$) and in a modified form for the spherical earth.[1] Because of ΔR the reflected ray is retarded in phase behind the direct ray by the angle $k\,\Delta R = 4\pi z_1z_2/\lambda R$.

For the present we assume that the presence of the plane earth can be represented by a quantity called the *plane-wave reflection coefficient*, Γ, which in general is a complex quantity represented in this volume by

$$\Gamma \equiv \rho e^{-i\phi}, \qquad 0 \leqslant \rho \leqslant 1. \tag{34}$$

The quantity ρ is the magnitude of the reflection coefficient, the fraction

[1] An equally common form for ΔR is obtained by considering the direct and reflected rays to be parallel ($\psi_1 = \psi_2$) and by dropping a perpendicular from the antenna position to the dotted line. For this construction $\Delta R \approx 2z_1 \sin \psi$.

of the incident field that is reflected, and ϕ is the angle by which the reflected field is retarded in phase upon reflection.[1]

The total excess phase retardation of the reflected wave is the sum of $k\,\Delta R$ and ϕ, and the amplitude of the reflected wave is $E_0\rho f(\theta_2)$. Thus the total field strength at the second terminal is

$$E = E_r + E_d = E_0\,[f(\theta_1) + \rho f(\theta_2)e^{-i(\phi + k\,\Delta R)}]. \tag{35}$$

An interference pattern is formed in space, the shape of which is given by the quantity in brackets. Referring to Eq. (26), however, we observe that the magnitude of the quantity in brackets is just F for the plane earth.[2]

$$F = |f(\theta_1) + \rho f(\theta_2)e^{-i\alpha}| = \sqrt{f^2(\theta_1) + \rho^2 f^2(\theta_2) + 2\rho f(\theta_1)f(\theta_2)\cos\alpha}, \tag{36}$$

where for convenience we introduce

$$\alpha = \phi + k\,\Delta R \approx \phi + \frac{4\pi z_1 z_2}{\lambda R}. \tag{37}$$

To calculate F one must know the behavior of ρ and ϕ (which are discussed in Chap. 5) and θ_1 and θ_2 must be determined. It is necessary to establish a convention for signs for the angles of Fig. 2·1. The angles ψ_1 and ξ are measured positive upward from the horizontal, whereas θ_1 and θ_2 are measured positive upward from the antenna beam maximum; ψ_2 is always considered positive.[3] (In the example of Fig. 2·1, θ_1 and θ_2 are numerically negative, and ψ_1 and ξ are positive.) The defining equations are

$$\left.\begin{aligned} \theta_1 &= \psi_1 - \xi, \\ \theta_2 &= -(\psi_2 + \xi), \\ \tan\psi_1 &= \frac{z_2 - z_1}{r}, \\ \tan\psi_2 &= \frac{z_2 + z_1}{r}. \end{aligned}\right\} \tag{38}$$

Returning to Eq. (36), it is instructive to investigate the range of variation of F for a plane earth. For the moment we shall assume ρ, ϕ,

[1] See Chap. 5 for a detailed discussion of Γ.

[2] No allowance has been made for the fact that the wavefronts at the second terminal are traveling in slightly different directions. This is of importance only when the wave is polarized in the plane of incidence ("vertical" polarization) and ψ_1 differs appreciably from ψ_2, or when it is desired to know the field in a specified direction (the vertical component, for instance). Also the slight difference in spatial $(1/R)$ attenuation of the direct and reflected waves has been neglected. Such considerations are of negligible importance in most practical cases.

[3] The *signs* of θ_1 and θ_2 are immaterial for a symmetrical antenna pattern, but for an unsymmetrical pattern, such as that of a cosecant-squared antenna, they may be important. The right-hand side of Eq. (36) has been written for a real antenna-pattern function f.

θ_1, and θ_2 to be constant[1] and observe that as $k\,\Delta R$ changes by π, F varies from a maximum to a minimum through an average value:

$$\left.\begin{aligned} F_{\max} &= f(\theta_1) + \rho f(\theta_2), \\ F_{\min} &= f(\theta_1) - \rho f(\theta_2), \\ F_{\text{ave}} &= f(\theta_1). \end{aligned}\right\} \tag{39}$$

Thus the field strength as a function of position in space varies (over sufficiently small intervals) between symmetrical limits above and below its arithmetic mean, which is the fraction of maximum free-space field radiated in the direction under investigation. Several special cases are of interest.

As ρ approaches zero, the interference pattern disappears and free-space transmission conditions are reached. This is the case for microwave transmission over very rough land at large values of ψ_2.

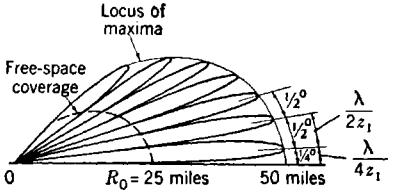


FIG. 2·2.—Coverage diagram for a free-space range of 25 miles over a plane earth with reflection coefficient of -1 and beam-width of 6° with the antenna pointed horizontally. The angular scale has been exaggerated for the sake of clarity. The lobe spacing of 0.5° corresponds to $z_1/\lambda = 57.3$. As drawn, the lobes are slightly too thick.

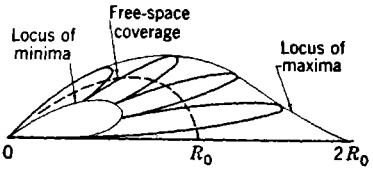


FIG. 2·3. Qualitative appearance of coverage diagram over a plane earth for which ρ is unity for zero grazing angle but decreases as grazing angle increases. The angular scale has been exaggerated for the sake of clarity and is not directly comparable to that of Fig. 2·2.

The second and more interesting case occurs when $\rho = 1$ and $\phi = \pi$, a condition well approximated for horizontally polarized waves over smooth sea water. For this case F is determined solely by the antenna pattern, varying between the limits $f(\theta_1) \pm f(\theta_2)$. Let us assume further that the antenna is pointed horizontally ($\xi = 0$) and that the pattern is symmetrical, that is, $f(\theta) = f(-\theta)$. Equation (36) then becomes

$$F = \left| f(\psi_1)\left(1 - e^{-ik\,\Delta R}\right)\right| = 2\left| f(\psi_1) \sin\left(\frac{2\pi z_1 z_2}{\lambda R}\right)\right|. \tag{40}$$

Usually ψ_1 is sufficiently small that except for very sharp beams $f(\psi_1)$ is essentially unity, and we have the well-known plane-earth formula

$$F = 2\left|\sin\left(\frac{2\pi z_1 z_2}{\lambda R}\right)\right|. \tag{41}$$

The expression for F is converted into a coverage diagram by R_0, as shown in Fig. 2·2. The "solid" coverage in free space, shown by the

[1] If $z_1/\lambda \gg 1$ and ψ_2 is small, this condition may be approximately satisfied over several neighboring lobes of the interference pattern.

dotted line, is now broken up into a series of *lobes* given by Eq. (40). As $0 \leqslant F \leqslant 2\ |f(\psi_1)|$, the detection range in some directions is reduced to zero, whereas in others it is twice the range in the same direction under free-space conditions. The lobes in the figure are drawn somewhat too thick; but even after allowance is made for this fact the gaps in coverage inside the range R_0 are fairly small. In terms of total volume no coverage has been lost, for it can be shown that the area inside the lobes is equal to the area inside the total free-space antenna pattern, but it has been redistributed in space. This redistribution of field strength may be an advantage or not, depending upon the problem at hand.

An important feature of Eq. (40) is the fact that for small angles the first maximum of the interference pattern occurs at an elevation angle viewed from the base of the transmitter by

$$\gamma = \frac{z_2}{R} = \frac{\lambda}{4z_1}, \tag{42}$$

which predicts that for fixed transmitter height z_1 the low-angle coverage improves as the wavelength λ is decreased. On the other hand, because the spacing between consecutive maxima (or minima) is given by $\lambda/2z_1$, the number of lobes in the interference pattern increases.

The preceding ideal case is modified considerably if ρ and ϕ are not constant but vary with the grazing angle ψ_2. For vertically polarized waves this dependence is pronounced (even for a smooth surface), and the effect is indicated qualitatively in Fig. 2·3. Instead of varying between zero and 2, F now varies as shown by Eq. (39) except that in this case ρ, θ_1, and θ_2 all vary between the maxima and minima. The result is that the coverage fluctuates less than in the preceding example but the extreme extension of range is not obtained. As the direct wave is not so strong as the reflected wave, interference is not complete in the minima nor is the field strength doubled in the maxima.

The curvature of the earth introduces complications that are discussed in detail in Secs. 2·13 to 2·16, 5·2, and 5·3. It is sufficient at this point to mention that because the reflection occurs at a spherical rather than at a plane surface, the wave diverges at an increased rate after reflection. This is taken into account by multiplying the reflection coefficient for a plane surface ρ by a divergence factor D which is unity at high angles and decreases toward zero as the angle decreases. This has the double effect of diminishing the range at low angles and of filling in the minima.

Another matter that has not yet been mentioned is that in addition to the direct and reflected waves there is also a *surface wave*, which at short distances and low angles can become important.[1] At frequencies above

[1] By *surface wave* we imply, not the well-known surface wave of Sommerfeld, but rather the wave complex given this name by K. A. Norton in *Proc. IRE*, **25**, 1192 (1937). See also C. R. Burrows, *Proc. IRE*, **25**, 219 (1937).

about 100 Mc/sec, however, the surface wave is so rapidly attenuated with range that for most applications it may be neglected at ranges of more than a very few miles. For an excellent discussion of transmission over a plane earth that gives much more detail than we have room for here see C. R. Burrows, "Radio Propagation over Plane Earth—Field Strength Curves," *Bell System Tech. Jour.*, **16.** 45 (1937).

Other forms for F will be derived in later sections in connection with specific problems (see particularly Secs. 2·13 through 2·16).

GEOMETRICAL OPTICS

BY JOHN E. FREEHAFER

2·3. Ray-tracing Formulas.—It happens many times that a solution of Maxwell's equations can be made to depend upon a suitable solution of the scalar wave equation. This is, in fact, the case when dealing with propagation through the atmosphere, as we shall see in Sec. 2·6. Accordingly we shall need to discuss means for dealing with the scalar wave equation

$$\nabla^2\psi + k^2n^2\psi = 0. \tag{43}$$

A time-honored way to find approximate solutions of Eq. (43) is to use the methods of geometrical optics. These methods yield in many cases useful graphical results which, however, may lead readily to false conclusions unless they are carefully interpreted. For this reason we shall review the method at some length.

Let us seek solutions of Eq. (43) in the form

$$\psi = Qe^{-ikS}, \tag{44}$$

where we assume that Q and S are real functions of position. Substituting Eq. (44) into Eq. (43) and equating the real and imaginary parts of the resulting equation to zero yield

$$(\nabla S)^2 - \frac{\nabla^2 Q}{Qk^2} - n^2 = 0, \tag{45}$$

$$\nabla^2 S + \frac{2(\nabla S \cdot \nabla Q)}{Q} = 0. \tag{46}$$

Equations (45) and (46) are exact. To obtain a tractable simplification, we assume that $\nabla^2 Q/Qk^2 \ll n^2$ and write Eq. (45) as

$$(\nabla S)^2 = n^2. \tag{47}$$

Provided that $\nabla^2 Q/Q$ remains finite, this becomes a better and better approximation as k and therefore the frequency increase. We shall return to this point later. Equation (47) is frequently referred to as the equation of the eikonal.

Let us now consider how to construct solutions of Eq. (47). The problem is to find a function the magnitude but not the direction of whose gradient is prescribed. Accordingly we may choose a surface Σ at each point of which we arbitrarily assign a constant value S_0 to S. Just how this surface is chosen depends on the boundary conditions and the kind of wave that we are trying to represent. This choice need not concern us at present. The gradient of S is then perpendicular to Σ at each point; and hence through Eq. (47), ∇S is known completely. To find the surface along which $S = S_0 + \Delta S$, we erect at each point of Σ perpendiculars whose lengths are equal to $\Delta S/n$, where n is the value of the index at the foot of each perpendicular. The tips of these perpendiculars form the desired surface. This process may, of course, be repeated to produce the whole family of surfaces along which S is constant. In the limit as $\Delta S \to 0$, the scalar field obtained in this way satisfies Eq. (47).

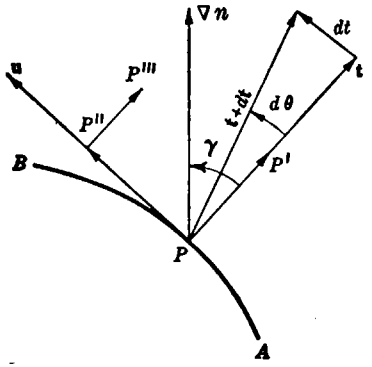


FIG. 2·4.—Process for continuing a ray. This figure is drawn in the plane of $\mathbf{t}$ and ∇n.

Instead of constructing the level surfaces of the field S, we may adopt the more usual alternate procedure, namely, to start at each point on the initial surface and draw a family of lines to which the surfaces of constant S are perpendicular. These lines are the rays, and the surfaces of constant S are the wavefronts of the solution of the wave equation with which we are dealing.

The differential equation of the rays may be obtained in the following way. In Fig. 2·4, P is a point on the wavefront AB on which S has a constant value S.

Here $\mathbf{t}$ is a unit vector normal to AB at P and is in the direction of the ray through P. With P as origin, the position vector of a point P' on the ray and also on the wavefront along which S equals $S + dS$ is $\mathbf{t}\, d\sigma$ where $d\sigma = dS/n$. This follows from the fact that n is the magnitude of ∇S and $d\sigma$ is measured in the direction of ∇S. The problem is to find $\mathbf{t} + d\mathbf{t}$, the direction of the ray at P'. To do this, it is necessary to construct a segment of the wavefront at P' and to draw $\mathbf{t} + d\mathbf{t}$ as a unit vector perpendicular to this segment. Let $\mathbf{u}$ and $\mathbf{v}$ be unit vectors perpendicular to $\mathbf{t}$, $\mathbf{u}$ lying in the plane of $\mathbf{t}$ and ∇n and making an acute angle with ∇n. By definition, $\mathbf{v} = \mathbf{t} \times \mathbf{u}$. If now we proceed for a short distance $d\rho$ in the direction $\mathbf{u}$ to the point P'' and then a distance equal to $dS/(n + \nabla n \cdot \mathbf{u} d\rho)$ in the direction of $\mathbf{t}$ we arrive at the point P''' which lies on the segment of the wavefront through P'. As the index does not change in the direction of $\mathbf{v}$, which is perpendicular to ∇n, all points near P' on a line through

P' in the $\mathbf{v}$ direction also lie on this segment. Therefore $\mathbf{t} + d\mathbf{t}$ is perpendicular to the plane determined by $\overrightarrow{P'P'''}$ and $\mathbf{v}$. Thus

$$\overrightarrow{PP'''} = d\rho\,\mathbf{u} + \frac{dS}{n}\left(\frac{1}{1+\dfrac{\nabla n\cdot\mathbf{u}\,d\rho}{n}}\right)\mathbf{t} = d\rho\,\mathbf{u} + d\sigma\left(1 - \frac{\nabla n\cdot\mathbf{u}}{n}\,d\rho\right)\mathbf{t} \quad (48)$$

and

$$\overrightarrow{P'P'''} = \overrightarrow{PP'''} - \overrightarrow{PP'} = d\rho\left(\mathbf{u} - \frac{\nabla n\cdot\mathbf{u}}{n}\,d\sigma\,\mathbf{t}\right). \quad (49)$$

Making use of the fact that $\mathbf{u}\times\mathbf{v} = \mathbf{t}$ and $\mathbf{t}\times\mathbf{v} = -\mathbf{u}$, we obtain to the first order in $d\sigma$

$$\mathbf{t} + d\mathbf{t} = \frac{\overrightarrow{P'P'''}\times\mathbf{v}}{\left|\overrightarrow{P'P'''}\right|} = \mathbf{t} + \frac{\nabla\mathbf{n}\cdot\mathbf{u}}{n}\,d\sigma\,\mathbf{u}, \quad (50)$$

whence

$$\frac{d\mathbf{t}}{d\sigma} = \frac{\nabla n\cdot\mathbf{u}}{n}\,\mathbf{u}. \quad (51)$$

Equation (51) is one form of the differential equation of the ray. In principle a ray can be drawn piece by piece in the following way. At point P, n and ∇n are given. If, in addition, the direction of the ray through P is specified, $\mathbf{t}$ and therefore $\mathbf{u}$ are both known. If we choose an element of length $d\sigma$, the position of a new point P' on the ray is given relative to P by

$$\overrightarrow{PP'} = \mathbf{t}\,d\sigma \quad (52)$$

and the direction of the ray at P' is given by

$$\mathbf{t}(P') = \mathbf{t}(P) + \frac{d\mathbf{t}}{d\sigma}(P)d\sigma, \quad (53)$$

where $d\mathbf{t}/d\sigma$ is calculated by Eq. (51). The process may now be repeated, and the ray continued. In the limit as $d\sigma \to 0$, the chain of line segments so obtained approaches the ray. This argument shows that given the scalar field n, a ray is determined when one point on it and its direction at that point are specified.

When the rays have been determined, S is obtained by the relationship

$$S(\rho) = \int_{\rho_1}^{\rho} |\nabla S|\,d\sigma + S_0 = \int_{\rho_1}^{\rho} n\,d\sigma + S_0, \quad (54)$$

where the integral is to be evaluated along the ray from the point with position vector ρ_1 on Σ to the point with position vector ρ at which it is desired to evaluate S.

A useful relationship can be obtained by equating the magnitudes of both members of Eq. (51). Thus

$$\left|\frac{d\mathbf{t}}{d\sigma}\right| = \frac{|\nabla n| \sin \gamma}{n}, \tag{55}$$

where γ is the angle between $\mathbf{t}$ and ∇n. As $\mathbf{t}$ is a unit tangent and σ is the arc length measured along the ray, $|d\mathbf{t}/d\sigma|$ is the curvature of the ray. Thus if R denotes the radius of curvature,

$$\frac{1}{R} = \frac{|\nabla n| \sin \gamma}{n}. \tag{56}$$

In radio propagation problems ∇n is nearly vertical and the rays are nearly horizontal. Hence γ and $\sin \gamma$ are approximately 90° and unity respectively. As n is also nearly unity, a convenient approximation to Eq. (56) is

$$\frac{1}{R} = \left|\nabla n\right|. \tag{57}$$

If, therefore, the gradient is constant, the rays are approximately arcs of circles. In applying Eq. (55) or (56) it must be borne in mind that the position vector of the center of curvature relative to a point on the ray makes an acute angle with ∇n. Thus, for instance, if in the atmosphere n decreases with height, ∇n points downward and the ray is concave downward.

Equation (51) can be written in a different form which exhibits certain important geometrical properties of the rays. As $\mathbf{t}$ is a unit vector in the direction of the ray, we have

$$\frac{dn}{d\sigma} = \nabla n \cdot \mathbf{t}. \tag{58}$$

Multiplying Eq. (58) by $\mathbf{t}$ and Eq. (51) by n and adding, we obtain

$$n\frac{d\mathbf{t}}{d\sigma} + \mathbf{t}\frac{dn}{d\sigma} = (\nabla n \cdot \mathbf{u})\mathbf{u} + (\nabla n \cdot \mathbf{t})\mathbf{t} \tag{59}$$

or

$$\frac{d}{d\sigma}(n\mathbf{t}) = \nabla n. \tag{60}$$

Let us consider now the integral $\int_{\rho_1}^{\rho_2} n \, d\sigma$ evaluated along a curve C joining the two points whose position vectors are ρ_1 and ρ_2. A fundamental problem of the calculus of variations is to determine the necessary and sufficient conditions on C in order that the integral have a stationary value. It turns out that these conditions may be expressed in terms of

Euler's equation which is just Eq. (60). Thus if the line integral of the index of refraction be evaluated along the ray joining two points, the integral has a stationary (usually minimum) value. Conversely, if a curve C be found connecting two points such that the line integral of the index of refraction evaluated along it has a stationary value, then C is a ray. This statement of the stationary properties of the rays is *Fermat's principle* and is frequently used as a starting point in developing the theory of geometric optics.

There are two special cases in which Eq. (60) is readily interpreted. Suppose first that n is such that ∇n is constant in direction. Let $\mathbf{j}$ be a unit vector in the direction of ∇n. The vector product of $\mathbf{j}$ with both members of Eq. (60) yields

$$\frac{d}{d\sigma}(n\mathbf{j} \times \mathbf{t}) = 0. \tag{61}$$

From this it follows that $n\mathbf{j} \times \mathbf{t}$ is a constant along the ray. Equation (51) shows, however, that $d\mathbf{t}$ always lies in the plane of $\mathbf{t}$ and ∇n. Thus if ∇n has a constant direction, the rays are plane curves. Hence $\mathbf{j} \times \mathbf{t}$ is perpendicular to the plane of the ray and is therefore constant in direction. Equation (61) then reduces to

$$\frac{d}{d\sigma}(n \sin \gamma) = 0, \tag{62}$$

where as before γ is the angle between the ray and ∇n. Equation (62) means that if ∇n has a fixed direction, the product of the index by the sine of the angle between the ray and the gradient of the index is constant along the ray. This is *Snell's law.*

Suppose next that ∇n is always in a radial direction from a fixed center. An example of this is the case of the atmosphere in which ∇n is ideally always vertical. Let $\mathbf{p}$ be the position vector of a point on the ray referred to the center as origin. As $\mathbf{p}$ and ∇n are parallel, the vector product of Eq. (60) by $\mathbf{p}$ yields

$$\mathbf{p} \times \frac{\mathrm{d}}{d\sigma}(n\mathbf{t}) = 0. \tag{63}$$

Now

$$\frac{d}{d\sigma}(\mathbf{p} \times n\mathbf{t}) = \frac{d\mathbf{p}}{d\sigma} \times n\mathbf{t} + \mathbf{p} \times \frac{d}{d\sigma}(n\mathbf{t}). \tag{64}$$

But $d\mathbf{p}/d\sigma$ is tangent to the ray and hence parallel to $n\mathbf{t}$. Therefore

$$\mathbf{p} \times \frac{d}{d\sigma}(n\mathbf{t}) = \frac{d}{d\sigma}(\mathbf{p} \times n\mathbf{t}), \tag{65}$$

and Eq. (63) becomes

$$\frac{d}{d\sigma}(\mathbf{p} \times n\mathbf{t}) = 0. \tag{66}$$

As in the previous case, the rays are plane curves and indeed lie in planes containing the center. Thus for any ray $\mathbf{p} \times \mathbf{t}$ is constant in direction and Eq. (66) can be written

$$\frac{d}{d\sigma}(|\mathbf{p}|\, n \sin \gamma) = 0, \tag{67}$$

where γ is the angle between $\mathbf{p}$ (and therefore ∇n) and the direction of the ray. Thus $|\mathbf{p}|n \sin \gamma$ is a constant along any ray. This is the generalization of Snell's law to the case of a radially directed index gradient.

The Stratified Atmosphere.—In treating atmospheric refraction, it is convenient and indeed almost necessary to assume that the properties of the atmosphere vary only with height. Such an idealized atmosphere is said to be stratified, and it is with stratified atmospheres that we shall be concerned in the remainder of this chapter.

To simplify the problem further and to derive equations that can later be generalized, we shall assume at first that the surface of the earth can be treated as a plane and its curvature neglected. Let us introduce a cylindrical coordinate system (r, ϕ, z) with the z-axis in the vertical direction and the plane $z = 0$ coincident with the surface of the earth. We shall treat the source of radiation as a point on the z-axis at height z_1 above the earth and assume cylindrical symmetry so that all quantities are independent of ϕ. Because ∇n is everywhere in the z-direction, it follows from Eq. (51) that the rays are curves lying in planes passing through the z-axis. It is therefore necessary to examine only one of these planes.

Consider now a ray that as shown in Fig. 2·5 is inclined at angle α to the r-axis at the transmitter height. By Snell's law

$$n(z) \cos \beta = C, \tag{68}$$

where β is the angle of inclination of the ray at any height z and C is a constant characteristic of the ray (C can be evaluated whenever the angle of inclination is known at a given height).
Thus

$$C = n(z_1) \cos \alpha; \tag{69}$$

or because the angle of inclination at the turning point P_0 (r_0, z_0) is zero,

$$C = n(z_0). \tag{70}$$

From Eq. (68)

$$n(z) \sin \beta = \pm \sqrt{n^2(z) - C^2}. \tag{71}$$

Dividing Eq. (71) by Eq. (68) yields

$$\tan \beta = \frac{\pm \sqrt{n^2(z) - C^2}}{C} = \frac{dz}{dr}, \tag{72}$$

where z and r are the coordinates of a point on the ray. As dr is always positive, the radical must be assigned the same sign as dz.

Equation (72) is a convenient form of the differential equation of the ray. It may be separated, and the solution reduced to a quadrature. Thus

$$r = C \int_{z_1}^{z} \left| \frac{dz}{\sqrt{n^2(z) - C^2}} \right|. \tag{73}$$

The absolute value signs are included to emphasize that dz and the radical

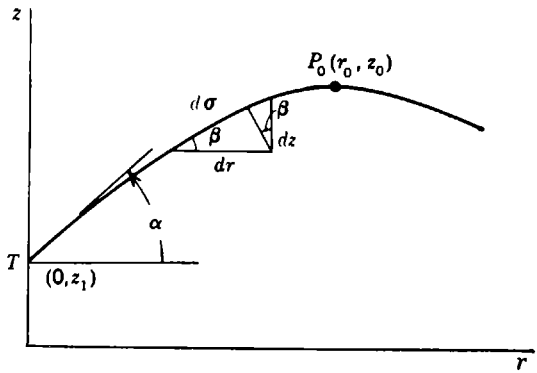


FIG. 2·5.—Geometry of a ray for the case of a stratified atmosphere. T is the transmitter, and P_0 is the turning point.

have like signs. Thus in case the point (r,z) lies to the right of P_0 of Fig. 2·5, Eq. (73) becomes

$$r = C\left(\int_{z_1}^{z_0} \frac{dz}{\sqrt{n^2(z) - C^2}} + \int_{z}^{z_0} \frac{dz}{\sqrt{n^2(z) - C^2}} \right). \tag{74}$$

When the ray family has been determined from Eq. (72), S, the optical length from the source to any point (r,z), is obtained by using Eq. (54). From Fig. 2·5, we see that

$$d\sigma = \cos \beta \, dr + \sin \beta \, dz. \tag{75}$$

Hence, from Eqs. (68) and (71),

$$dS = n \, d\sigma = C \, dr + |\sqrt{n^2 - C^2} \, dz|. \tag{76}$$

Therefore

$$S = \int n \, d\sigma = Cr + \int_{z_1}^{z} |\sqrt{n^2 - C^2} \, dz|. \tag{77}$$

For a qualitative discussion of the ray family such as is given in Chap. 1, it is convenient to introduce a further simplification in Eq. (72). In many applications n and C are nearly unity, and then we can write

$$n^2 - C^2 = (n - C)(n + C) \approx 2(n - C), \tag{78}$$

and Eq. (72) becomes

$$\frac{dz}{dr} \approx \pm \sqrt{2}\sqrt{n - C}. \tag{79}$$

This shows qualitatively how the inclination of the ray depends upon the difference between n and C. By making use of Eq. (69), Eq. (79) can be transformed into

$$\frac{dz}{dr} = \pm \sqrt{2}\sqrt{n(z) - n(z_1)\cos\alpha}. \tag{80}$$

Now for rays inclined at small angles,

$$\frac{dz}{dr} \approx \beta \tag{81}$$

and

$$n(z_1)\cos\alpha \approx n(z_1)\left(1 - \frac{\alpha^2}{2}\right) \approx n(z_1) - \frac{\alpha^2}{2}; \tag{82}$$

consequently Eq. (80) may be written

$$\beta \approx \pm \sqrt{2}\sqrt{n(z) - n(z_1) + \frac{\alpha^2}{2}}. \tag{83}$$

Squaring Eq. (83) yields the symmetrical relationship

$$\beta^2 - \alpha^2 = 2[n(z) - n(z_1)]. \tag{84}$$

Up to this point, the surface of the earth has been treated as plane. Equations will now be developed that take the curvature into account. Let us set up a spherical coordinate system (p,ϕ,θ) with center at the center of the earth. A point source is located on the line for which $\theta = 0$ at a height z_1 above the surface of the earth, the radius of which is denoted by a. For a stratified atmosphere, ∇n is always parallel to the direction in which p varies, and the rays are therefore curves lying in the planes containing the source and the center of the earth. Figure 2·6 shows a typical plane of this kind; and since as before we shall assume cylindrical symmetry about the vertical line through the transmitter, all functions are independent of ϕ and we can limit our attention to a single meridian plane.

Under these conditions Eq. (67) applies, and we can write

$$pn \cos \beta = \text{const. along a ray.} \tag{85}$$

By dividing both sides of Eq. (85) by a, we obtain

$$\frac{pn}{a} \cos \beta = C, \tag{86}$$

where C is a constant characteristic of the ray.
Equation (86) can be made to resemble Eq. (68) by introducing the abbreviation $N = pn/a$; whence

$$N \cos \beta = C. \tag{87}$$

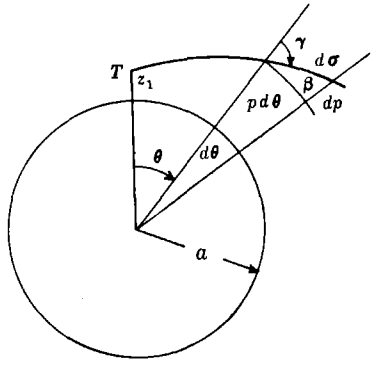


FIG. 2·6.—Path of a ray in a vertical plane through the source. T represents the transmitter.

Just as Eq. (68) leads to Eq. (72), Eq. (87) leads to

$$\tan \beta = \frac{\pm \sqrt{N^2 - C^2}}{C} = \frac{dp}{p\,d\theta}. \tag{88}$$

Now let us change variables from p and θ to z and r where

$$p = a + z, \tag{89}$$

$$\theta = \frac{r}{a}. \tag{90}$$

The coordinate r represents distance measured along the earth's surface. Eq. (88) then becomes

$$\frac{\pm \sqrt{N^2 - C^2}}{C} = \frac{dz}{\left(1 + \frac{z}{a}\right) dr}. \tag{91}$$

Equation (91) is exact; but because in problems concerning atmospheric

refraction z/a is small compared with unity, we arrive at the approximation

$$\frac{dz}{dr} = \frac{\pm\sqrt{N^2 - C^2}}{C}. \tag{92}$$

This equation is just Eq. (72) with n replaced by N. Thus, subject to the provision that $z/a \ll 1$, the effect of the curvature of the earth can be taken into account in computing ray paths by replacing the index of refraction by the quantity $N = pn/a$ and using the same differential equation that applies when the earth is treated as flat.

Next let us consider the expression for the optical length. From Fig. 2·6

$$d\sigma = p\, d\theta \cos\beta + dp \sin\beta, \tag{93}$$

and from Eqs. (89) and (90)

$$n\, d\sigma = \frac{np}{a}\, dr \cos\beta + n\, dz \sin\beta. \tag{94}$$

Application of Eq. (86) yields

$$n\, d\sigma = C\, dr + \left|\sqrt{\left(\frac{np}{a}\right)^2 - C^2}\, \frac{dz}{1 + \frac{z}{a}}\right|. \tag{95}$$

If as before we neglect z/a in comparison with unity, we obtain

$$dS = C\, dr + |\sqrt{N^2 - C^2}\, dz|. \tag{96}$$

By comparing Eq. (96) with Eq. (76) it is evident that subject to the condition $z/a \ll 1$ the optical length can be calculated from the plane-earth formula if n is replaced by N.

Equations (92) and (96) suggest on the basis of geometrical optics that the discussion of propagation near the surface of the spherical earth may be reduced to a discussion of propagation over a plane earth by the artifice of replacing the index n by the modified index

$$N = \frac{pn}{a} = n\left(1 + \frac{z}{a}\right). \tag{97}$$

This matter will be examined from a different point of view in the next section.

2·4. The Modified Index.—If the earth had no atmosphere the rays would be straight lines which, with increasing distance from the source, would become separated farther and farther from the earth's surface. The rays would be straight, and the earth curved. Actually what is of importance is not the separate curvatures of the ray family and the earth but the relative curvature of the rays with respect to the earth. It turns

out that it is sometimes convenient to regard the earth as flat and the rays as having sufficient curvature to maintain the same relative curvature as before. To account for the curvature of the rays we must assume that the refractive index instead of being constant increases with height, and indeed at such a rate that the derivative of the index with respect to height equals the reciprocal of the earth's radius. Thus the earth may be treated as flat if the index is taken equal to $1 + (z/a)$. In this way diffraction is treated as a problem in refraction.

The general idea of regarding the earth as flat and modifying the index distribution appropriately is a useful one, and we shall now examine the procedure more closely. The method is due to M. H. L. Pryce in an unpublished manuscript. Let r, θ, and ϕ be the spherical coordinates of a point, the origin of the system being at the center of the earth, and let us introduce new coordinates x_1, x_2, and z, defined by

$$x_1 = a \sin \theta \cos \phi, \tag{98a}$$

$$x_2 = a \sin \theta \sin \phi, \tag{98b}$$

$$z = r - a, \tag{98c}$$

where θ is to be measured from the line connecting the transmitter to the center of the earth.

The element of arc $d\sigma$ is given by

$$d\sigma^2 = \left(1 + \frac{z}{a}\right)^2 \left[dx_1^2 + dx_2^2 + \frac{(x_1\, dx_1 + x_2\, dx_2)^2}{a^2 - x_1^2 - x_2^2} \right] + dz^2. \tag{99}$$

In general, we are interested only in the region for which x_1/a and x_2/a are sufficiently small to justify neglecting their squares and higher powers, and we can therefore write

$$d\sigma^2 = \eta^2(dx_1^2 + dx_2^2) + dz^2, \tag{100}$$

where

$$\eta \equiv 1 + \frac{z}{a}. \tag{101}$$

Maxwell's equations for free space become, in the new coordinate system,

$$\frac{\partial}{\partial x_2} E_3 - \frac{\partial}{\partial z} (\eta E_2) + \eta\mu \frac{\partial H_1}{\partial t} = 0, \tag{102a}$$

$$\frac{\partial}{\partial z} (\eta E_1) - \frac{\partial}{\partial x_1} E_3 + \eta\mu \frac{\partial H_2}{\partial t} = 0, \tag{102b}$$

$$\frac{\partial E_2}{\partial x_1} - \frac{\partial E_1}{\partial x_2} + \eta\mu \frac{\partial H_3}{\partial t} = 0, \tag{102c}$$

$$\frac{\partial}{\partial x_2} H_3 - \frac{\partial}{\partial z} (\eta H_2) - \eta\epsilon \frac{\partial E_1}{\partial t} = 0, \tag{102d}$$

E. G. & G. LIBRARY
LAS VEGAS BRANCH

$$\frac{\partial}{\partial z}(\eta H_1) - \frac{\partial}{\partial x_1} H_3 - \eta\epsilon \frac{\partial E_2}{\partial t} = 0, \tag{102e}$$

$$\frac{\partial}{\partial x_1} H_2 - \frac{\partial}{\partial x_2} H_1 - \eta\epsilon \frac{\partial E_3}{\partial t} = 0, \tag{102f}$$

$$\frac{\partial}{\partial x_1}(\eta\mu H_1) + \frac{\partial}{\partial x_2}(\eta\mu H_2) + \frac{\partial}{\partial z}(\eta^2\mu H_3) = 0, \tag{102g}$$

$$\frac{\partial}{\partial x_1}(\eta\epsilon E_1) + \frac{\partial}{\partial x_2}(\eta\epsilon E_2) + \frac{\partial}{\partial z}(\eta^2\epsilon E_3) = 0. \tag{102h}$$

The terms $(\partial/\partial z)(\eta E_2)$, $(\partial/\partial z)(\eta E_1)$, $(\partial/\partial z)(\eta\eta\epsilon E_3)$ and similar terms involving H complicate the demonstration. They are all of the form $(\partial/\partial z)(\eta F)$. Now

$$\frac{\partial}{\partial z}(\eta F) = \frac{\partial F}{\partial z} + \frac{\partial}{\partial z}\left(\frac{zF}{a}\right). \tag{103}$$

The field points in which we are interested are all at low angles as viewed from the transmitter, and in this region the field components vary much more rapidly in the horizontal than in the vertical direction. For this reason we are justified in neglecting $(\partial/\partial z)(zF/a)$ in comparison with derivatives with respect to x_1 and x_2. With this simplification and the introduction of the notation $\mu' = \eta\mu$ and $\epsilon' = \eta\epsilon$, we obtain

$$\frac{\partial E_3}{\partial x_2} - \frac{\partial}{\partial z} E_2 + \mu' \frac{\partial H_1}{\partial t} = 0, \tag{104a}$$

$$\frac{\partial}{\partial z} E_1 - \frac{\partial E_3}{\partial x_1} + \mu' \frac{\partial H_2}{\partial t} = 0, \tag{104b}$$

$$\frac{\partial E_2}{\partial x_1} - \frac{\partial E_1}{\partial x_2} + \mu' \frac{\partial H_3}{\partial t} = 0, \tag{104c}$$

$$\frac{\partial H_3}{\partial x_2} - \frac{\partial}{\partial z} H_2 - \epsilon' \frac{\partial E_1}{\partial t} = 0, \tag{104d}$$

$$\frac{\partial H_1}{\partial z} - \frac{\partial}{\partial x_1} H_3 - \epsilon' \frac{\partial E_2}{\partial t} = 0, \tag{104e}$$

$$\frac{\partial H_2}{\partial x_1} - \frac{\partial}{\partial x_2} H_1 - \epsilon' \frac{\partial E_3}{\partial t} = 0, \tag{104f}$$

$$\frac{\partial}{\partial x_1}(\mu' H_1) + \frac{\partial}{\partial x_2}(\mu' H_2) + \frac{\partial}{\partial z}(\mu' H_3) = 0, \tag{104g}$$

$$\frac{\partial}{\partial x_1}(\epsilon' E_1) + \frac{\partial}{\partial x_2}(\epsilon' E_2) + \frac{\partial}{\partial z}(\epsilon' E_3) = 0. \tag{104h}$$

These equations, however, are just Maxwell's equations in rectangular coordinates in which the surface of the earth is given by $z = 0$ and the permeability and permittivity of the medium are μ' and ϵ'. We have

thus reduced the problem of propagation around a spherical earth to propagation over a flat earth with an atmosphere whose electromagnetic properties are represented by μ' and ϵ'.

The index of refraction of the atmosphere modified to account for the curvature of the earth is called the *modified index of refraction* and is denoted by N. From the preceding discussion, it follows that the relationship between N and the true index n is given by

$$N = \eta n = \left(1 + \frac{z}{a}\right) n \approx n + \frac{z}{a}. \tag{105}$$

It should be noted that whereas n usually decreases with height, the modified index, except for limited regions, increases with height. In particular when n decreases linearly with height, that is

$$n = n_0 + \left(\frac{dn}{dz}\right) z, \tag{106}$$

where n_0 and dn/dz are constants and $(dn/dz)z$ is small compared with n_0, we can write

$$N = n_0 \left(1 + \frac{z}{a_e}\right), \tag{107}$$

in which

$$\frac{1}{a_e} = \frac{1}{n_0} \frac{dn}{dz} + \frac{1}{a}. \tag{108}$$

When a_e is a positive constant, it may be considered as an effective radius[1] of the earth. Measurements have shown that the average gradient of n over height intervals not including layers within several hundred feet of the surface of the earth is usually such that $a_e = \frac{4}{3}a$. (This point will be considered in detail in the next chapter.) It is customary, therefore, to refer to the condition in which $a_e = \frac{4}{3}a$ as *standard refraction*. For standard refraction

$$\frac{dN}{dz} = 3.6 \times 10^{-8}\ \text{ft}^{-1},$$

$$\frac{dn}{dz} = -\ 1.2 \times 10^{-8}\ \text{ft}^{-1}.$$

It is sometimes convenient in numerical work to take $a_e = 5280$ miles.

2·5. Limitations of Ray Methods.—The ray-tracing formulas of Sec. 2·3 may be integrated analytically in simple cases and numerically or mechanically in more complicated cases so that it is possible to obtain the ray patterns for a large variety of index distributions. It is necessary, how-

[1] The concept of an equivalent or effective earth's radius was introduced in a somewhat different way by Schelleng, Burrows, and Ferrell, *Proc. IRE*, **21**, 427 (1933).

ever, to examine each case to determine if the conditions are fulfilled upon which the validity of the technique depends. These conditions are to be examined in this section.

The first condition to be satisfied is that the index must not change appreciably in a distance equal to a wavelength of the radiation. This requirement is expressed by the inequality

$$\frac{1}{kn}\frac{|\nabla n|}{n} \ll 1. \tag{109}$$

Insofar as inequality (109) is concerned, a given ray pattern becomes a better and better approximation as the frequency and therefore k is increased. On the other hand, inequality (109) certainly does not represent a sufficient condition, for even in a homogeneous medium for which $\nabla n = 0$, ray methods fail to account for diffraction phenomena. As we shall see, the additional requirement is that the fractional change in the spacing between neighboring rays must be small in a wavelength. A ray pattern therefore leads to questionable results in regions where rays either diverge or converge rapidly and especially where neighboring rays cross to form foci or caustics.

It is to be noted that whereas the first condition depends alone upon the frequency of the radiation and the structure of the medium, the second condition depends upon what sort of wave the ray pattern is to represent. Whether or not the first condition is satisfied can be determined in advance once the structure of the medium and the frequency of the radiation are specified, but whether or not the second condition is satisfied can be determined only after the ray pattern has actually been traced. In this connection, it is well to recall that *a single ray has no significance by itself but has meaning only when related to the other members of the family to which it belongs.*

To develop the argument that supports the preceding assertions, it is first necessary to integrate Eq. (46):

$$\nabla^2 S + \frac{2(\nabla S \cdot \nabla Q)}{Q} = 0. \tag{110}$$

Now

$$\nabla \cdot (Q^2 \nabla S) = 2Q \nabla Q \cdot \nabla S + Q^2 \nabla^2 S \tag{111}$$

so that Eq. 110 can be written

$$\frac{1}{Q^2} \nabla \cdot (Q^2 \nabla S) = 0 \tag{112}$$

or

$$\nabla \cdot (Q^2 \nabla S) = 0. \tag{113}$$

Equation (113) states that the vector $Q^2\,\nabla S$ is solenoidal. Therefore a solution of Eq. (113) is obtained by writing

$$Q^2\,\nabla S = \nabla \times \mathbf{U}, \tag{114}$$

where $\mathbf{U}$ is any vector field whose curl is parallel to ∇S.

Now let us suppose that Eq. (47) has been integrated by constructing the family of rays normal to the initial surface Σ at each point of which S has the value S_0. The family of rays may then be given in parametric form

$$\left.\begin{aligned} x &= F_1(\xi,\eta,S),\\ y &= F_2(\xi,\eta,S),\\ z &= F_3(\xi,\eta,S),\end{aligned}\right\} \tag{115}$$

where ξ and η are the parameters that determine the point on Σ through which the ray passes. Thus if S is put equal to S_0, Eqs. (115) are the parametric equations of the surface Σ and x, y, and z are the coordinates of the intersection of the ray with the surface. Let the result of solving Eqs. (115) for ξ, η, and S be represented by

$$\xi = f_1(x,y,z), \tag{116a}$$

$$\eta = f_2(x,y,z), \tag{116b}$$

$$S = f_3(x,y,z). \tag{116c}$$

Equations (116a) and (116b) are two families of surfaces whose intersections are the rays. It should be noted that these families are not uniquely determined by the ray pattern, as it is possible to set up an infinite variety of families having the same intersections.

As the rays are perpendicular to the wavefronts, ∇f_1 and ∇f_2 are perpendicular to ∇S and $\nabla f_1 \times \nabla f_2$ is parallel to ∇S. Hence the identity

$$\nabla \times (f_1\,\nabla f_2) = \nabla f_1 \times \nabla f_2 + f_1\,\nabla \times \nabla f_2 \tag{117}$$

shows that $\nabla \times (f_1\,\nabla f_2)$ is parallel to ∇S. It follows therefore that a suitable choice for $\mathbf{U}$ in Eq. (6) is

$$\mathbf{U} = f_1\,\nabla f_2. \tag{118}$$

Accordingly Eq. (114) becomes

$$Q^2\,\nabla S = \nabla \times f_1\,\nabla f_2 = \nabla f_1 \times \nabla f_2. \tag{119}$$

Taking the scalar product of each member of Eq. (119) with ∇S and making use of Eq. (47), we obtain

$$Q^2 = \frac{\nabla f_1 \times \nabla f_2 \cdot \nabla S}{n^2}. \tag{120}$$

Even on the assumption that S is known, Eq. (120) does not determine Q uniquely, for as was pointed out previously, f_1 and f_2 are not determined uniquely by the ray family. The significance of f_1 and f_2 may be established in the following way. We note first that

$$\nabla f_1 \times \nabla f_2 \cdot \nabla S = \begin{vmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{vmatrix} = \frac{\partial(f_1,f_2,f_3)}{\partial(x,y,z)} = \frac{1}{J}, \tag{121}$$

where J is the Jacobian of x, y, and z with respect to f_1, f_2, and f_3. If $d\tau$ denotes the volume element bounded by the surfaces $f_1 = \xi$, $f_1 = \xi + d\xi$, $f_2 = \eta$, $f_2 = \eta + d\eta$, $f_3 = S$, $f_3 = S + dS$, then it is a property of the Jacobian that

$$d\tau = J(f_1,f_2,f_3)\, d\xi\, d\eta\, dS. \tag{122}$$

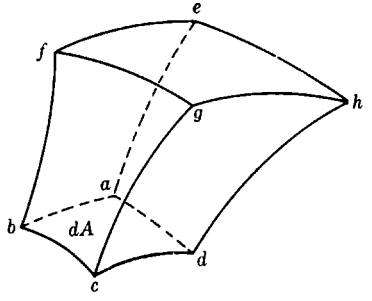


FIG. 2·7.—The volume element $d\tau$ bounded by rays and surfaces of constant phase. The face *abcd* lies in the surface $f_3 = S$ and the face *efgh* lies in the surface $f_3 = S_1 + dS$. Similarly *abfe* and *dcgh* are portions of the surfaces $f_2 = \eta$ and $f_2 = \eta + d\eta$, respectively. The remaining two faces *adhe* and *bcgf* lie in the surfaces $f_1 = \xi$ and $f_1 = \xi + d\xi$, respectively.

Let dA be the element of area bounded by the traces of the surfaces f_1 = constant and f_2 = constant on the wavefront $f_3 = S$ as shown in Fig. 2·7.

As the rays are perpendicular to the wavefronts, it follows that

$$d\tau = dA\, d\sigma, \tag{123}$$

where, as before, σ is the arc length measured along the ray. Equations (122) and (123) yield

$$J = \frac{dA\, d\sigma}{d\xi\, d\eta\, dS}. \tag{124}$$

But

$$\frac{dS}{d\sigma} = |\nabla S| = n. \tag{125}$$

Hence

$$J = \frac{dA}{n\, d\xi\, d\eta}. \tag{126}$$

From Eq. (120) and (121)

$$Q^2 = \frac{1}{Jn^2}; \tag{127}$$

Eq. (126) and (127) yield

$$d\xi\, d\eta = Q^2\, n\, dA = Q^2\, \nabla S \cdot \frac{\nabla S}{n}\, dA. \tag{128}$$

As $\nabla S/n$ is a unit vector normal to dA, we see that $d\xi\, d\eta = df_1\, df_2$ is the element of flux through a wavefront of the vector field $Q^2\, \nabla S$. In many applications of the wave equation $Q^2\, \nabla S$ is either proportional to the power density or approximately so. Hence Eq. (128) shows that the functions f_1 and f_2 are determined by the radiation pattern of the source.

As pointed out in Sec. 2·3, the development of ray-tracing formulas depends upon the assumption that

$$\frac{\nabla^2 Q}{k^2 n^2 Q} \ll 1. \tag{129}$$

It is now possible to examine this inequality more carefully. From the identity

$$\nabla \cdot \left(\frac{\nabla Q}{nQ}\right) = \frac{\nabla^2 Q}{nQ} + \left(\nabla \frac{1}{nQ}\right) \cdot \nabla Q, \tag{130}$$

it follows that

$$\frac{\nabla^2 Q}{k^2 n^2 Q} = \frac{1}{kn}\, \nabla \cdot \left(\frac{\nabla Q}{knQ}\right) + \frac{1}{kn}\, \frac{\nabla n}{n} \cdot \frac{\nabla Q}{knQ} + \frac{\nabla Q}{knQ} \cdot \frac{\nabla Q}{knQ}. \tag{131}$$

From Eq. (127)

$$\frac{\nabla Q}{knQ} = -\left(\frac{\nabla J^{1/2}}{knJ^{1/2}} + \frac{1}{kn}\, \frac{\nabla n}{n}\right). \tag{132}$$

Equations (131) and (132) yield

$$\frac{\nabla^2 Q}{k^2 n^2 Q} = \frac{\nabla J^{1/2}}{knJ^{1/2}} \cdot \frac{1}{kn}\, \frac{\nabla n}{n} + \left(\frac{\nabla J^{1/2}}{knJ^{1/2}}\right)^2 - \frac{1}{kn}\, \nabla \cdot \left(\frac{\nabla J^{1/2}}{knJ^{1/2}}\right) - \frac{1}{kn}\, \nabla \cdot \left(\frac{\nabla n}{knn}\right). \tag{133}$$

From Eq. (133) it is readily seen that inequality (129) is satisfied provided that the vector fields $(1/kn)(\nabla n/n)$, and $\nabla J^{1/2}/knJ^{1/2}$ have magnitudes small compared with unity and are smooth so that the divergence terms are likewise small. Thus we arrive at the two ray-tracing conditions

$$\begin{aligned} \frac{1}{kn}\, \frac{|\nabla n|}{n} &\ll 1, \\ \frac{1}{kn}\, \frac{|\nabla J^{1/2}|}{J^{1/2}} &\ll 1. \end{aligned} \tag{134}$$

The meaning of the first inequality is clear and has been discussed previously. To interpret the second condition we consider the pencil of rays bounded by the surfaces $f_1 = \xi$, $f_1 = \xi + d\xi$, $f_2 = \eta$, and $f_2 = n + d\eta$. Equation (126) shows that J is proportional to dA, the cross section of

this pencil, and hence $J^{1/2}$ is essentially the spacing between neighboring rays. Then condition (134) states that the fractional change in the spacing between neighboring rays in a wavelength must be small compared with unity. Obviously condition (134) is violated at those points where the cross section of the pencil vanishes as it does when the pencil comes to a focus or passes through a caustic.

PHYSICAL OPTICS

BY JOHN E. FREEHAFER

2·6. The Field from a Dipole in a Stratified Atmosphere Near the Earth.—In this section we shall show that the calculation of the field from either a vertical electric or magnetic dipole near the earth can be reduced to finding a suitable solution of the scalar wave equation

$$\nabla^2\psi + k^2N^2\psi = 0. \tag{135}$$

Here ψ is the z-component of a Hertz vector $\mathbf{\Pi}$, whose x_1- and x_2-components are zero.

It will also be shown that a solution of Eq. (135) which satisfies the appropriate boundary conditions is given by

$$\psi = \int_0^\infty \kappa J_0(\kappa r)v(z,\kappa)\, d\kappa, \tag{136}$$

where r and z are the cylindrical coordinates of the field point and

$$v = \frac{2}{W}\left[\frac{u_1'(0) - pu_1(0)}{u_2'(0) - pu_2(0)} - \frac{u_1(z_1)}{u_2(z_1)}\right] u_2(z_1)u_2(z), \qquad z \geq z_1, \tag{137a}$$

$$v = \frac{2}{W}\left[\frac{u_1'(0) - pu_1(0)}{u_2'(0) - pu_2(0)} - \frac{u_1(z)}{u_2(z)}\right] u_2(z_1)u_2(z), \qquad z \leq z_1. \tag{137b}$$

The symbols used in the expressions for v have the following meanings:

1. z_1 is the height of the dipole. (The dipole is located at $r = 0$, $z = z_1$.)
2. v, u_1, and u_2 are all solutions of the equation

$$\frac{d^2v}{dz^2} + (k^2N^2 - \kappa^2)v = 0; \tag{138}$$

u_1 and u_2, however, are those solutions of Eq. (138) which for large z represent waves traveling in the direction of decreasing and increasing z respectively; W is the Wronskian of u_1 and u_2 and is independent of z.

3. p is a function of κ that depends upon the polarization. For a vertical electric dipole

$$p(\kappa^2) = i\frac{k_0^2}{k_1^2}\sqrt{k_1^2 - \kappa^2}, \tag{139}$$

where k_0 is the propagation factor of the atmosphere at the surface and k_1 is the propagation factor of the earth. For a vertical magnetic dipole

$$p(\kappa^2) = i\sqrt{k_1^2 - \kappa^2}. \tag{140}$$

The radicals are to indicate the root with the *negative* imaginary part.

Although the results stated above apply strictly to dipole sources, they may be used to discuss more elaborate sources. It is frequently a good approximation to assume that the ratio of the field strength to the free-space field strength at a point is the same for an actual transmitter as it is for the dipole. When this is not permitted, we can in principle approximate the source by a suitable array of dipoles. The field is then the sum of the fields from the dipole elements.

From this point on we shall suppress the time factor $e^{i\omega t}$. Let us consider the Hertz vector $\mathbf{\Pi} = (0,0,\psi)$ which satisfies the equation

$$\nabla \times \nabla \times \mathbf{\Pi} - \nabla\nabla \cdot \mathbf{\Pi} - \mu'\epsilon'\omega^2 \mathbf{\Pi} = 0 \tag{141}$$

and derive a field of the electric type in the usual way:

$$\mathbf{B} = i\mu'\epsilon'\omega\nabla \times \mathbf{\Pi}; \qquad \mathbf{E} = \nabla \times \nabla \times \mathbf{\Pi}. \tag{142}$$

Because $\mathbf{\Pi}$ has only a z-component and μ' and ϵ' vary only in the z-direction, it follows that

$$\nabla(\mu'\epsilon')\cdot\nabla \times \mathbf{\Pi} = 0, \qquad \nabla(\mu'\epsilon') \times \mathbf{\Pi} = 0. \tag{143}$$

From Eqs. (141), (142), and (143), we find that

$$\nabla \times \mathbf{E} = -i\omega\mu'\mathbf{H}, \tag{144a}$$

$$\nabla \times \mathbf{H} = i\omega\epsilon'\mathbf{E} + \frac{\nabla\epsilon'}{\epsilon'} \times \mathbf{H}, \tag{144b}$$

$$\nabla\cdot\mathbf{D} = \nabla\epsilon'\cdot\mathbf{E}, \tag{144c}$$

$$\nabla\cdot\mathbf{B} = 0. \tag{144d}$$

Equations (144b) and (144c) show that the fields obtained from Eq. (142) do not satisfy Maxwell's equations exactly. For the atmosphere, however, the gradient of ϵ' is so small that the terms involving it can be neglected. Thus the ratio of the magnitude of the second to the first term on the right-hand side of Eq. (144b) can be written

$$\frac{\lambda}{2\pi}\frac{|\nabla\epsilon'|}{\epsilon'}\frac{\sqrt{\mu_0\epsilon_0}}{\epsilon'}\frac{|\mathbf{H}|}{|\mathbf{E}|}.$$

Except quite close to the source, the field will be almost a plane wave for which $|\mathbf{H}|/|\mathbf{E}|$ is approximately $\sqrt{\epsilon'/\mu'}$. Furthermore, $|\nabla\epsilon'|/\epsilon'$ is of the order of $1/a$, and therefore the ratio of the second to the first term is of

the order of λ/a, which is always very small compared with unity. From Eqs. (144*a*) and (144*c*) it follows that

$$\frac{\nabla\cdot\mathbf{E}}{|\nabla\times\mathbf{E}|}\approx\frac{\lambda}{2\pi}\frac{|\nabla\epsilon'|}{\epsilon'}\frac{\sqrt{\mu_0\epsilon_0}}{\mu'}\frac{|\mathbf{E}|}{|\mathbf{H}|},$$

and this again is of the order of λ/a. It follows that Eq. (144) differs from Maxwell's equations by a negligible amount, and we may regard Eq. (142) as a valid way to obtain the electromagnetic field.

From a solution of Eq. (141), a field of the magnetic type can be constructed by using the equations

$$\mathbf{D}=-i\mu'\epsilon'\omega\nabla\times\mathbf{\Pi},\qquad\mathbf{H}=\nabla\times\nabla\times\mathbf{\Pi}. \tag{145}$$

The **D** and **H** obtained in this way satisfy the equations

$$\nabla\times\mathbf{E}=-i\omega\mu'\mathbf{H}+\frac{\nabla\mu'}{\mu'}\times\mathbf{E}, \tag{146a}$$

$$\nabla\times\mathbf{H}=i\omega\epsilon'\mathbf{E}, \tag{146b}$$

$$\nabla\cdot\mathbf{D}=0, \tag{146c}$$

$$\nabla\cdot\mathbf{B}=\nabla\mu'\cdot\mathbf{H}. \tag{146d}$$

It follows from the same arguments applied to Eq. (144) that the field given by Eq. (145) is an acceptable approximate solution of Maxwell's equations.

Because ψ is the only component of $\mathbf{\Pi}$ that differs from zero, and because moreover ψ is a rectangular component of $\mathbf{\Pi}$, Eq. (141) reduces to Eq. (135). We have thus shown that the solution of our problem depends upon finding a suitable solution of the scalar wave equation.

We now seek a solution of Eq. (135) that satisfies the proper boundary conditions at the source, at the surface of the earth, and at great heights.

If we introduce the variable $r=\sqrt{x_1^2+x_2^2}$, and make use of the cylindrical symmetry of the field, we can write Eq. (135) as

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right)+\frac{\partial^2\psi}{\partial z^2}+k^2N^2\psi=0. \tag{147}$$

Integration by the usual separation procedure leads to elementary solutions of Eq. (147) given by $Z_0(\kappa r)\ v(z,\kappa)$ where Z_0 is any integral of the Bessel equation of zero order, κ is the separation constant, and v is a solution of Eq. (138). From the elementary waves we can construct a solution independent of κ by integration over κ. Thus

$$\psi=\int D(\kappa)\ Z_0(\kappa r)\ v(z,\kappa)\ d\kappa, \tag{148}$$

where $D(\kappa)$ is to be regarded as a function of κ that is at present undetermined.

Our procedure is now guided by the fact that near the source, ψ must behave like e^{-ikR}/R. The expansion of a spherical wave in terms of cylindrical waves is well known and is given by[1]

$$\frac{1}{R}e^{-ikR} = \int_0^\infty \kappa J_0(\kappa r)\frac{\exp(\sqrt{\kappa^2-k^2}\,z)}{\sqrt{\kappa^2-k^2}}\,d\kappa, \qquad z \leqslant 0, \tag{149a}$$

$$\frac{1}{R}e^{-ikR} = \int_0^\infty \kappa J_0(\kappa r)\frac{\exp(-\sqrt{\kappa^2-k^2}\,z)}{\sqrt{\kappa^2-k^2}}\,d\kappa, \qquad z \geqslant 0, \tag{149b}$$

where $R = \sqrt{r^2+z^2}$. The source is at $z = 0 = r$. When $\kappa > k$, the radical stands for the positive root and when $\kappa < k$, $\sqrt{\kappa^2-k^2}$ is $i|\sqrt{k^2-\kappa^2}|$.

Comparison of Eq. (148) with Eq. (149) suggests integration from zero to infinity and choosing

$$D(\kappa) = \kappa,$$

$$Z_0(\kappa r) = J_0(\kappa r).$$

Finally we note that

$$\frac{d}{dz}\left[\frac{\exp(-\sqrt{\kappa^2-k^2}\,z)}{\sqrt{\kappa^2-k^2}}\right]_{z=0} - \frac{d}{dz}\left[\frac{\exp(\sqrt{\kappa^2-k^2}\,z)}{\sqrt{\kappa^2-k^2}}\right]_{z=0} = -2.$$

Thus the part of the integral that depends upon z is continuous at the height of the source, but its derivative suffers a discontinuity equal to -2 at that height. It is important, as we shall see, that the discontinuity in derivative is independent of κ. Accordingly we shall require that v satisfy the conditions

$$\left.\begin{aligned} v\Big|_{z_1+0} &= v\Big|_{z_1-0}, \\ \frac{dv}{dz}\Big|_{z_1+0} &= \frac{dv}{dz}\Big|_{z_1-0} - 2. \end{aligned}\right\} \tag{150}$$

It has now been established tentatively that Eq. (148) takes on the form of Eq. (136). To verify that ψ of Eq. (136) does indeed have the proper behavior, we need only to evaluate the volume integral $\int (\nabla^2 + k^2N^2)\psi\, d\tau$ over a small region containing the source. This integral is essentially the moment of the dipole; and for a unit source such as is represented by e^{-ikR}/R, it should be equal to -4π. The verification follows:

$$\int (\nabla^2 + k^2N^2)\psi\, d\tau = \int\int_0^\infty \kappa(\nabla^2 + k^2N^2)J_0(\kappa r)v\, d\kappa\, d\tau$$

$$= \int\int_0^\infty \kappa J_0(\kappa r)\left[\frac{d^2v}{dz^2} + (k^2N^2 - \kappa^2)v\right]d\kappa\, d\tau.$$

[1] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 575.

Now $d\tau = 2\pi r\, dr\, dz$, and from Eqs. (138) and (150) it follows that $(d^2v/dz^2) + (k^2N^2 - \kappa^2)v$ is independent of κ. Therefore

$$\int (\nabla^2 + k^2N^2)\psi\, d\tau = \lim_{\alpha\to 0}\int_{z_1-\alpha}^{z_1+\alpha}\left[\frac{d^2v}{dz^2} + (k^2N^2 - \kappa^2)v\right]dz \cdot \lim_{r\to 0}\int_0^r 2\pi r \int_0^\infty \kappa J_0(\kappa r)\, d\kappa\, dr.$$

But

$$\int_{z_1-\alpha}^{z_1+\alpha}\left[\frac{d^2v}{dz^2} + (k^2N^2 - \kappa^2)v\right]dz = \frac{dv}{dz}\bigg|_{z_1-\alpha}^{z_1+\alpha} + \int_{z_1-\alpha}^{z_1+\alpha}(k^2N^2 - \kappa^2)v\, dz = -2.$$

$$\lim_{r\to 0}\int_0^r 2\pi r \int_0^\infty \kappa J_0(\kappa r)\, d\kappa\, dr = 2\pi \lim_{r\to 0}\int_0^\infty \frac{d\kappa}{\kappa}\int_0^r \kappa r J_0(\kappa r)\, d(\kappa r)$$

$$= 2\pi \lim_{r\to 0}\int_0^\infty J_1(\kappa r)\, d(\kappa r) = 2\pi\, J_0(\kappa r)\Big|_\infty^0 = 2\pi.$$

Hence

$$\int (\nabla^2 + k^2N^2)\psi\, d\tau = -4\pi,$$

and this confirms that ψ given by Eq. (136) and subject to conditions (150) has the proper behavior in the neighborhood of the source.

We shall next examine the boundary conditions at $z = 0$. Consider first the case of the electric-type field given by Eq. (142). Using the cylindrical coordinates (r,ϕ,z) and making use of the cylindrical symmetry of the field, we obtain

$$\left.\begin{aligned}\mathbf{B} &= i\mu'\epsilon'\omega\left(0, -\frac{\partial\psi}{\partial r}, 0\right),\\ \mathbf{E} &= \left(\frac{\partial^2\psi}{\partial z\,\partial r}, 0, -\frac{\partial^2\psi}{\partial r^2}\right).\end{aligned}\right\} \tag{151}$$

Let ψ be used to denote the z-component of the Hertz vector within the earth, then the boundary conditions for $z = 0$ are

$$\frac{\partial}{\partial r}\frac{\partial\psi}{\partial z} = \frac{\partial}{\partial r}\frac{\partial\psi_1}{\partial z},$$

$$k_0^2\frac{\partial\psi}{\partial r} = k_1^2\frac{\partial\psi_1}{\partial r},$$

where k_1 is the complex propagation factor of the earth. Integration with respect to r yields

$$\frac{\partial\psi}{\partial z} = \frac{\partial\psi_1}{\partial z};\quad k_0^2\psi = k_1^2\psi_1. \tag{152}$$

Because of the conductivity of the earth, the field within the earth is rapidly attenuated with depth. For this reason it is not necessary to take into account the curvature, and we shall assume that the properties of the

earth are independent of the coordinates. If now ψ_1 is expressed as an integral

$$\psi_1 = \int_0^\infty \kappa J_0(\kappa r) v_1(z, \kappa)\, d\kappa, \tag{153}$$

the boundary condition is satisfied if

$$\left(\frac{\frac{dv}{dz}}{v}\right)_{z=0} = \frac{k_0^2}{k_1^2}\left(\frac{\frac{dv_1}{dz}}{v_1}\right)_{z=0}, \tag{154a}$$

$$v_1(0, \kappa) = \frac{k_0^2}{k_1^2} v(0, \kappa). \tag{154b}$$

Equation (154*b*) is a condition on v_1 and is of interest only if the field within the earth is to be calculated; v_1 must be a solution of the equation

$$\frac{d^2 v_1}{dz^2} + (k_1^2 - \kappa^2) v_1 = 0. \tag{155}$$

Therefore

$$v_1 \propto \exp\left(i\sqrt{k_1^2 - \kappa^2}\, z\right), \tag{156}$$

where, in order to ensure attenuation as z becomes larger negatively, the branch of the square root having a negative imaginary part must be taken. Then Eq. (154*a*) becomes

$$\left(\frac{\frac{dv}{dz}}{v}\right)_{z=0} = p(\kappa^2), \tag{157}$$

where $p(\kappa^2)$ is given by Eq. (139). In a similar way one can show that the boundary condition for the case of the magnetic-type field represented by Eq. (145) leads to Eq. (157) where $p(\kappa^2)$ is given by Eq. (140).

Now κ/k is essentially the cosine of the angle of inclination of the normal to the wavefront of the elementary wave characterized by that value of κ. For the field points in which we are interested, the values of κ near k will contribute most strongly to the integral of Eq. (136), and therefore for purposes of estimating the order of magnitude of p we may regard $\kappa \approx k$. Thus for horizontal polarization

$$|p| \approx k\left|\sqrt{\frac{k_1^2}{k^2} - 1}\right|.$$

As is shown in Sec. 5·1, $|k_1/k|^2$ is much greater than unity except in the case of very dry ground; for example, at 3000 Mc/sec, $(k_1/k)^2 = 70 - i40$ for sea water. Hence $|p|$ is of the order of ten times k, and $|1/p|$ amounts to about 0.02λ The same argument holds also for vertical

polarization. Equation (157), together with Fig. 2·8, tells us that $|1/p|$ is essentially the distance by which a node fails to occur at the surface of the earth. Again for those values of κ which contribute most to the integral of Eq. (136), λ is small compared with the vertical distance in which v varies significantly. Thus, for practical purposes, it is a convenient approximation to use as the boundary condition for both polarizations that v have a node at the surface of the earth, that is,

$$v(0,\kappa) = 0. \tag{158}$$

It should be noted that the assumption that the earth is a perfect con-

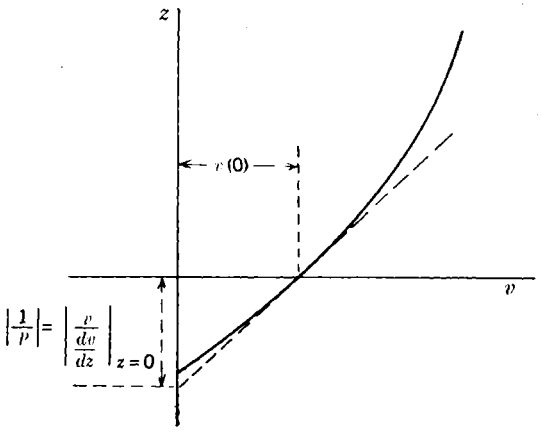


FIG. 2·8.—The geometrical interpretation of $|1/p|$.

ductor leads to Eq. (158) for horizontal polarization. For vertical polarization, however, it leads to

$$\frac{dv}{dz}(0,\kappa) = 0;$$

hence Eq. (158) is not equivalent to assuming infinite conductivity for the earth.

The final step is to examine conditions for large z. Although it is true that at sufficiently great heights the unmodified index of refraction approaches unity and N approaches $1 + z/a$, it is an experimental fact that at the moderate heights to which our approximations limit us anyway, the unmodified index n has a small essentially constant gradient. Accordingly we shall assume that for large z, N approaches $1 + z/a_e$ where a_e is a constant that may be regarded as the effective radius of the earth.

For any real value of κ, the expression $k^2N^2 - \kappa^2$ eventually becomes and remains positive. Let z_0 denote the height at which the last zero of $k^2N^2 - \kappa^2$ occurs. Then Eq. (138) has two independent solutions that can

be represented asymptotically for $z \gg z_0$ by

$$u_1 = (k^2N^2 - \kappa^2)^{-\frac{1}{4}} \exp\left[i \int_{z_0}^{z} (k^2N^2 - \kappa^2)^{\frac{1}{2}}\, dz \right], \tag{159}$$

$$u_2 = (k^2N^2 - \kappa^2)^{-\frac{1}{4}} \exp\left[- i \int_{z_0}^{z} (k^2N^2 - \kappa^2)^{\frac{1}{2}}\, dz \right]. \tag{160}$$

The fractional powers are to mean the positive roots. The functions u_1 and u_2 represent waves traveling in the direction of decreasing and increasing z, respectively. Accordingly we shall require that for large z, v shall behave like u_2. As the transmitter height is the only height at which it is possible to change from one linear combination of solutions to another, we write

$$v = Au_2, \qquad z \geqslant z_1, \tag{161}$$

$$v = Bu_1 + Cu_2, \qquad z \leqslant z_1. \tag{162}$$

Equations (150) and (157) permit the determination of A, B, and C, and we obtain Eqs. (137a) and (137b).

When Eq. (158) is substituted for Eq. (157), the expressions for v given in Eqs. (137a) and (137b) are replaced by

$$v = \frac{2}{W}\left[\frac{u_1(0)}{u_2(0)} - \frac{u_1(z_1)}{u_2(z_1)} \right] u_2(z_1)u_2(z) \qquad z \geq z_1, \tag{163a}$$

$$v = \frac{2}{W}\left[\frac{u_1(0)}{u_2(0)} - \frac{u_1(z)}{u_2(z)} \right] u_2(z_1)u_2(z) \qquad z \leq z_1. \tag{163b}$$

2·7. The Fundamental Theorem.—In Sec. 2·6 it was shown that the field from a dipole near the earth can be expressed in the form of a generalized Sommerfeld integral. The next step is to devise means of evaluating this integral. It turns out that there are two ways of doing this, one way being appropriate to the diffraction region and the other to the interference region. In this section we shall discuss the diffraction region and show that subject to certain restrictions the potential ψ may be expressed in the form

$$\psi = -\pi i \sum_{m=1}^{\infty} H_0^{(2)}(\kappa_m r) u_m^{(1)}(z) u_m^{(1)}(z_1) + \frac{1}{2} \int_B \kappa H_0^{(2)}(\kappa r) v(z,\kappa)\, d\kappa, \tag{164}$$

where (1) κ_m stands for the roots of

$$u_2'(0,\kappa) - p(\kappa^2) u_2(0,\kappa) = 0, \tag{165}$$

(2) the "normalized" height-gain functions $u_m^1(z)$ are defined by

$$u_m^{(1)}(z) = \frac{u_2(z, \kappa_m)}{\sqrt{\displaystyle\int_0^{\infty} u_2^2(z,\kappa_m)\, dz + \frac{u_2^2(0,\kappa_m)}{2\kappa_m} \left(\frac{dp}{d\kappa} \right)_{\kappa^2}}}, \tag{166}$$

the integral being along a ray in the fourth quadrant of the z plane,

(3) $\int_B$ denotes integration around the branch cuts in the κ-plane of the function $v(z,\kappa)$.

The expansion represented by Eq. (164) is referred to as the fundamental theorem. This theorem has been established only subject to certain assumptions that must be justified along lines to be explained below for each special form of N that it is desired to discuss. When the simplification given in Eq. (158) is used, v has no branch points and the integral is missing. In the general case, however, the presence of the double-valued function p in the expression for v produces branch points at $\kappa = \pm k_1$. The integral part of Eq. (164) is then present and represents a "surface wave," which at high frequencies is rapidly attenuated both with range and height. From a practical standpoint, however, in the region for which z/R and z_1/R are very small compared with unity, ψ is given by a few terms of the series of Eq. (164).

To begin the demonstration we shall first transform the integral expression for ψ into one that involves $H_0^{(2)}(\kappa r)$ instead of $J_0(\kappa r)$.

Because

$$J_0(\kappa r) = \frac{1}{2}\left[H_0^{(1)}(\kappa r) + H_0^{(2)}(\kappa r)\right],$$

Eq. (136) can be written

$$\psi = \frac{1}{2}\left[\int_0^\infty \kappa H_0^{(1)}(\kappa r)v(z,\kappa)\,d\kappa + \int_0^\infty \kappa H_0^{(2)}(\kappa r)v(z,\kappa)\,d\kappa\right]. \qquad (167)$$

It is now convenient to introduce an assumption that is difficult to establish in complete generality but can be verified for all cases of practical interest. This assumption is that $v(z,\kappa)$ has no singularities in the first quadrant of the κ-plane.

That v has poles in the first quadrant is ruled out because it may be shown that the presence of such poles would lead to terms in the final expression for ψ that represent waves converging toward the source. This is physically impossible.

The question of branch points is more difficult to decide. When $N(z)$ is an integral single-valued function of z, it may be shown that $u_1(z,\kappa)$ and $u_2(z,\kappa)$ are integral single-valued functions of κ for all z. Likewise, when the real positive axis of z can be divided into a finite number of segments in each of which $N(z)$ is free of singularities and is so defined that a fundamental set of solutions of Eq. (138) exists, then for all real positive values of z, u_1 and u_2 are integral single-valued functions of κ. In this case the solution having the proper behavior in the outer segment is continued inward by joining a linear combination of solutions in one segment to a linear combination in the next and requiring that the function and its first

derivative be continuous at the joints. The fundamental solutions and their coefficients are integral functions of κ and also, therefore, are u_1 and u_2. Accordingly, we shall assume that for the values of z with which we are concerned, u_1 and u_2 have no singularities in the finite portion of the κ-plane. Subject to this assumption, the branch points of p are the only branch points of v. These occur at $\kappa = \pm k_1$. But k_1^2 lies in the fourth quadrant, and therefore the branch points lie in the second and fourth quadrants.

Because of this assumption concerning the singularities, we can deform the path of integration and write

$$\int_0^{\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = \int_0^{i\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa + \int_{\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa, \quad (168)$$

where $\int_{\infty}$ means integration over the quarter circle of infinite radius lying in the first quadrant. The path of integration of the first integral on the right is the positive axis of imaginaries. Set $\kappa = iu$; then

$$\int_0^{i\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = -\int_0^{\infty} u H_0^{(1)}(iur) v(z,iu)\, du.$$

But

$$H_0^{(1)}(iur) = -H_0^{(2)}(-iur);$$

and as the equations defining v contain κ^2 only,

$$v(z,iu) = v(z,-iu).$$

Therefore

$$\int_0^{i\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = \int_0^{\infty} u H_0^{(2)}(-iur) v(z,-iu)\, du.$$

Now if we put $\kappa = -iu$, we obtain

$$\int_0^{i\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = \int_{-i\infty}^{0} \kappa H_0^{(2)}(\kappa r) v(z,\kappa)\, d\kappa. \quad (169)$$

As $H_0^{(1)}(\kappa r) \to 0$ as $\kappa \to \infty$ along any ray in the first quadrant, we introduce another assumption that must be verified for each specific case, namely,

$$\int_{\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = 0. \quad (170)$$

Using Eqs. (169) and (170) in Eq. (168), we obtain

$$\int_0^{\infty} \kappa H_0^{(1)}(\kappa r) v(z,\kappa)\, d\kappa = \int_{-i\infty}^{0} \kappa H_0^{(2)}(\kappa r) v(z,\kappa)\, d\kappa. \quad (171)$$

Finally, Eqs. (167) and (171) lead to

$$\psi = \frac{1}{2}\int_L \kappa H_0^{(2)}(\kappa r)v(z,\kappa)\,d\kappa, \tag{172}$$

where the path of integration L extends from $-i\infty$ to ∞ along the negative imaginary axis and the positive real axis as shown in Fig. 2·9.

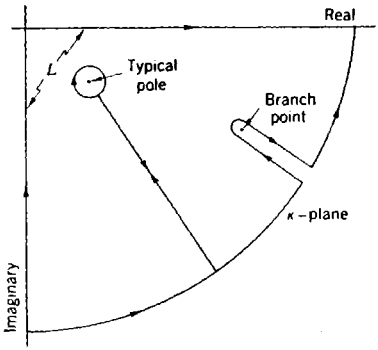


FIG. 2·9.—Equivalent paths of integration in the κ-plane.

The next step is to reduce the integral to a sum by deforming L into a quarter circle of infinite radius in the fourth quadrant plus loops around the poles and branch points as indicated in Fig. 2·9. Now because of the same assumption involved in Eq. (170), the integral over the quarter circle is zero, and Eq. (172) becomes

$$\psi = -\pi i \sum R_m + \frac{1}{2}\int_B \kappa H_0^{(2)}(\kappa r)v(z,\kappa)\,d\kappa,$$

where ΣR_m means the sum of the residues at all the poles of v lying in the fourth quadrant. The integral is along the branch cut.

Reference to Eqs. (137*a*) and (137*b*) shows that the poles of v are the roots of Eq. (165). The values of κ that satisfy Eq. (165) are the characteristic values of the problem and will be designated by κ_m. They are the values of κ for which the solution of Eq. (138) representing an outgoing wave itself satisfies Eq. (157). On the assumption that the poles of v are simple poles, we obtain

$$\psi = -\pi i \sum_{m=1}^{\infty} \kappa_m H_0^{(2)}(\kappa_m r) \frac{2}{W(\kappa_m)} \frac{u_1'(0,\kappa_m) - p(\kappa_m^2)u_1(0,\kappa_m)}{\left\{\frac{\partial}{\partial\kappa}\left[u_2'(0,\kappa) - pu_2(0,\kappa)\right]\right\}_{\kappa=\kappa_m}} \times u_2(z,\kappa_m)u_2(z_1,\kappa_m) + \frac{1}{2}\int_B \kappa H_0^{(2)}(\kappa r)v\,d\kappa. \tag{173}$$

We shall order the κ_m's so that $\mathrm{Im}(\kappa_{m-1}) > \mathrm{Im}(\kappa_m) > \mathrm{Im}(\kappa_{m+1})$. Now the

Wronskian is given by

$$W(\kappa_m) = u_1(0,\kappa_m)u_2'(0,\kappa_m) - u_2(0,\kappa_m)u_1'(0,\kappa_m). \tag{174}$$

Because for $\kappa = \kappa_m$, Eq. (165) is satisfied

$$W(\kappa_m) = -u_2(0,\kappa_m)[u_1'(0,\kappa_m) - p(\kappa_m^2)u_1(0,\kappa_m)]. \tag{175}$$

The use of Eq. (175) reduces Eq. (173) to

$$\psi = 2\pi i \sum_{m=1}^{\infty} \kappa_m H_0^{(2)}(\kappa_m r) \frac{u_2(z,\kappa_m)u_2(z_1,\kappa_m)}{u_2(0,\kappa_m)\left\{\frac{\partial}{\partial\kappa}\left[u_2'(0,\kappa) - pu_2(0,\kappa)\right]\right\}_{\kappa=\kappa_m}} + \frac{1}{2}\int_B \kappa H_0^{(2)}(\kappa r)v\, d\kappa. \tag{176}$$

In the discrete sum only u_2 appears; accordingly we shall adopt the shorter notation

$$u_2(z,\kappa_m) \equiv u_m(z), \tag{177}$$

$$u_2(z,\kappa) \equiv u(z). \tag{178}$$

We shall next obtain an alternate form for the expression within the braces of Eq. (176). The procedure is the usual one for proving the orthogonality of solutions of the Sturm-type equation. Thus by multiplying the differential equation for u_m by u and the equation for u by u_m, we obtain

$$u\frac{d^2u_m}{dz^2} + (k^2N^2 - \kappa_m^2)u_m u = 0, \tag{179}$$

$$u_m\frac{d^2u}{dz^2} + (k^2N^2 - \kappa^2)uu_m = 0. \tag{180}$$

Subtracting Eq. (180) from Eq. (179) and integrating with respect to z from 0 to ∞ along a ray in the z-plane, we have

$$\left(u\frac{du_m}{dz} - u_m\frac{du}{dz}\right)_0^{\infty} - (\kappa_m^2 - \kappa^2)\int_0^{\infty} uu_m\, dz = 0. \tag{181}$$

It turns out that a ray can be chosen (one in the fourth quadrant is satisfactory) such that $u(\infty)$ and $u_m(\infty)$ both vanish. Accordingly

$$(\kappa^2 - \kappa_m^2)\int_0^{\infty} uu_m\, dz = u(0)\left(\frac{du_m}{dz}\right)_0 - u_m(0)\left(\frac{du}{dz}\right)_0. \tag{182}$$

But

$$\left(\frac{du_m}{dz}\right)_0 = p(\kappa_m^2)u_m(0);$$

hence Eq. (182) can be written

$$(\kappa + \kappa_m)\int_0^\infty uu_m\,dz = \frac{u_m(0)\left[p(\kappa_m^2)u(0) - \left(\frac{du}{dz}\right)_0\right]}{(\kappa - \kappa_m)}$$

$$= u_m(0)\left\{\frac{p(\kappa^2)u(0) - \left(\frac{du}{dz}\right)_0 + u(0)\left[p(\kappa_m^2) - p(\kappa^2)\right]}{\kappa - \kappa_m}\right\}. \quad (183)$$

Now if we let κ approach κ_m, we obtain

$$2\kappa_m\int_0^\infty u_m^2\,dz = u_m(0)\left[\left\{\frac{\partial}{\partial\kappa}\left[p(\kappa^2)u(0) - u'(0)\right]\right\}_{\kappa=\kappa_m} - u_m(0)\left(\frac{\partial p}{\partial\kappa}\right)_{\kappa_m^2}\right] \quad (184)$$

and

$$u_m(0)\left\{\frac{\partial}{\partial\kappa}\left[u'(0) - p(\kappa^2)u(0)\right]\right\}_{\kappa=\kappa_m} = -2\kappa_m\int_0^\infty u_m^2\,dz - u_m^2(0)\left(\frac{dp}{d\kappa}\right)_{\kappa_m^2}. \quad (185)$$

Substitution of Eq. (185) in Eq. (176) yields

$$\psi = -\pi i\sum_{m=0}^{\infty} H_0^{(2)}(\kappa_m r)\frac{u_2(z,\kappa_m)u_2(z_1,\kappa_m)}{\int_0^\infty u_m^2\,dz + \frac{u_m^2(0)}{2\kappa_m}\left(\frac{dp}{d\kappa}\right)_{\kappa_m^2}} + \frac{1}{2}\int_B \kappa H_0^{(2)}(\kappa r)v\,d\kappa. \quad (186)$$

Equation (186) is equivalent to Eqs. (164) and (166).

2·8. Phase-integral Methods.—In order to apply the fundamental theorem to a given case of atmospheric refraction, it is necessary to find the normalized characteristic functions $u_m^{(1)}$ and the characteristic values κ_m. This is in general a difficult task, as the detailed behavior of solutions of Eq. (138) is known for only a very limited number of index distributions. Fortunately means exist for constructing quite general approximate solutions of this differential equation, because the same equation was encountered early in the development of quantum mechanics. It is from this field that the techniques discussed in this section are borrowed.

The first step in obtaining approximate expressions for the characteristic functions is to find a pair of asymptotic representations of solutions of the differential equation. Next it is necessary to examine how the same exact solution is approximated by different linear combinations of the asymptotic expressions in different parts of the complex plane. This leads

to a discussion of the Stokes phenomenon and the derivation of the connection formulas of quantum mechanics. By means of the connection formulas it is possible to show how the function that satisfies the boundary condition at great heights can be represented at the surface. The boundary condition at the surface then determines the characteristic values. The final step is to normalize the functions so that the various terms or modes in the expansion given by the fundamental theorem have the correct relative magnitudes. Two separate formulations result depending upon whether the characteristic value is nearly real and the mode is strongly "trapped" or the imaginary part of the characteristic value is comparable to the real part and the mode is "leaky." Trapped and leaky modes are sometimes referred to as Gamow and Eckersley modes respectively.[1]

It is convenient for our present purpose to write Eq. (138) as

$$\frac{d^2u}{dz^2} + k^2su = 0, \tag{187}$$

where

$$s = N^2 - \frac{\kappa^2}{k^2} \tag{188}$$

and u, the dependent variable, is an abbreviation for $u_2(z,\kappa)$. We shall retain the notation adopted in Sec. 2·7 of attaching the subscript m to any function of κ in which κ is replaced by κ_m. Thus s_m means $N^2 - \kappa^2{}_m/k^2$.

Let us seek solutions of Eq. (187) in the form

$$u = e^{ikP} \tag{189}$$

and assume that P can be expanded in a series of descending powers of k. Thus

$$P = kw(z) + w_0(z) + k^{-1}w_1(z) + \cdots, \tag{190}$$

where the w's are to be determined. Substitution of Eq. (189) into Eq. (187) yields

$$i\frac{d^2P}{dz^2} - \left(\frac{dP}{dz}\right)^2 + k^2s = 0. \tag{191}$$

When Eq. (190 is substituted into Eq. (191) and the coefficients of k^2 and k equated to zero, we obtain

$$s - \left(\frac{dw}{dz}\right)^2 = 0, \tag{192}$$

and

$$i\frac{d^2w}{dz^2} - 2\frac{dw}{d}\frac{dw_0}{dz} = 0. \tag{193}$$

[1] This treatment of phase-integral methods follows closely that given by W. H. Furry in RL Report No. 680.

Integration of Eq. (192) and (193) leads to

$$w = \int s^{1/2}\, dz, \tag{194}$$

$$w_0 = \frac{i}{2} \ln w' = \frac{i}{2} \ln s^{1/2}. \tag{195}$$

Constants of integration are unimportant, as additive constants in the expression for P are equivalent to constant factors multiplying u. Thus neglecting all but the first two terms of Eq. (190), we obtain from Eqs. (189), (190), (194), and (195)

$$u \approx s^{-1/4} \exp\left(ik \int_{z_0}^{z} s^{1/2}\, dz\right). \tag{196}$$

To be definite we have chosen z_0, a zero of s, as the lower limit of the integral. Obviously, a similar expression with the sign of the exponential changed is also an approximate solution of the differential equation so that the general solution can be written

$$u = Af_1 + Bf_2\,, \tag{197}$$

where for the sake of brevity we have put

$$f_1 = s^{-1/4}\, e^{ikw}, \tag{198}$$

$$f_2 = s^{-1/4}\, e^{-ikw}, \tag{199}$$

$$w = \int_{z_0}^{z} s^{1/2}\, dz. \tag{200}$$

Note that the zeros of s are singularities of f_1 and f_2 but not of u; hence Eq. (197) fails in the vicinity of zeros of s.

Equation (197) is the so-called Wentzel-Kramers-Brillouin, or WKB, approximation to the solution of Eq. (187). If functions f_1 and f_2 were exact solutions of the differential equation, A and B would be constants, independent of the value of z. Because of the asymptotic nature of the approximation, however, A and B take on different values in different regions of the complex plane and change abruptly at the boundaries between these regions. This is known as the *Stokes phenomenon*, and to this matter we must next give careful attention.

In the following discussion it is assumed that s has a simple zero z_0 which is sufficiently isolated that there exists a region around it free of other zeros in which the functions f_1 and f_2 are good approximations to solutions of Eq. (187). In the neighborhood of z_0

$$s \approx (z - z_0)\left(\frac{ds}{dz}\right)_{z_0}. \tag{201}$$

By Eq. (200)

$$w \approx \frac{2}{3}\left[\left(\frac{ds}{dz}\right)_{z_0}\right]^{\frac{1}{2}}(z - z_0)^{\frac{3}{2}}. \tag{202}$$

Equation (202) shows that there are three directions at z_0 separated by 120° along which w is real. Starting in any one of these directions we can construct a line along which w is real by choosing increments dz such that $s^{\frac{1}{2}}dz$ is real. There are three lines of this nature which will be labeled arbitrarily I, II, and III as shown in Fig. 2·10. Similarly there are three lines, making angles of 120° with each other and 60° with the first set of lines at the point z_0, along which w is pure imaginary. These are labeled S_1, S_2, and S_3 in Fig. 2·10 and are referred to as the Stokes lines.

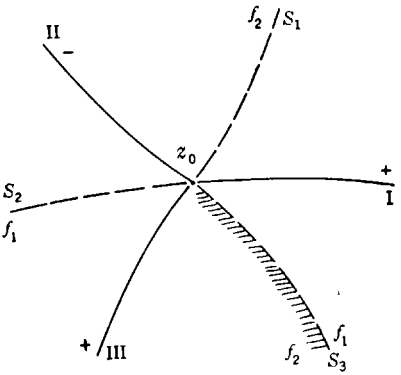


FIG. 2·10.—S_1, S_2, and S_3 are the Stokes lines associated with the point z_0. I, II, and III are the anti-Stokes lines. Algebraic signs refer to w. Dominant terms are indicated on each Stokes line.

It is now essential that the branches of w be carefully defined. We shall establish a cut along the line S_3 and require that w be positive on Line I. Then w will be positive imaginary on S_1, negative on II, negative imaginary on S_2, positive on III, and positive imaginary on shaded side of S_3.

Along the anti-Stokes lines I, II, and III, the functions f_1 and f_2 are oscillatory and of equal absolute values. On the Stokes lines, however, the functions are real exponentials; and as w increases with increasing distance from z_0, the positive exponential is very much larger than the negative. Thus f_2 is dominant on S_1; f_1 is dominant on S_2; and f_2 is again dominant on the shaded side of S_3.

Now although z_0 is a branch point of the functions f_1 and f_2, it is not a singular point of the differential equation. Exact solutions of the differential equation are, in fact, single valued. Therefore, if an exact solution u is to be represented by a linear combination of the multiple-valued functions f_1 and f_2 as in Eq. (197), it is necessary that the coefficients A and B change values as the independent variable z makes a circuit about z_0. This takes place in the following way. In the neighborhood of the Stokes line the dominant term is so much larger than the subdominant that the coefficient of the latter may change without having any appreciable effect on the value of the linear combination. The coefficient of the dominant term may not change, however, nor may the coefficient of the subdominant term change if the dominant term is missing, because in either case the

value of the linear combination would be obviously affected. Thus the Stokes lines divide the complex plane into three regions in each of which the coefficients A and B are constant except near the boundaries, where the coefficient of the subdominant term changes rapidly. These regions will be designated I, II, or III according to whether they contain the lines I, II, or III. The values of the coefficients A and B in Region I will be designated A_{I} and B_{I}, with similar designations in the other regions.

As the differential equation is linear and any solution can be represented as a linear combination of two independent solutions, it follows that the values of A and B in one region are linear combinations of their values in any other region. Thus, for instance, using a's to represent undetermined coefficients we can write

$$A_{\mathrm{II}} = a_{11} A_{\mathrm{I}} + a_{12} B_{\mathrm{I}}, \tag{203}$$

$$B_{\mathrm{II}} = a_{21} A_{\mathrm{I}} + a_{22} B_{\mathrm{I}}. \tag{204}$$

Now A is the coefficient of f_1, which is subdominant on S_1. Accordingly A changes when z crosses the Stokes line from Region I into Region II and, unless B_{I}, the coefficient of the dominant term, is zero, A_{II} differs from A_{I}. Furthermore, under no circumstances may the coefficient of the dominant term change. B_{II} must equal B_{I}. Thus $a_{11} = a_{22} = 1$ and $a_{21} = 0$. This leads to

$$A_{\mathrm{II}} = A_{\mathrm{I}} + \alpha B_{\mathrm{I}}, \tag{205}$$

$$B_{\mathrm{II}} = B_{\mathrm{I}}, \tag{206}$$

where α has been written as an abbreviation for the undetermined coefficient a_{12}.

By similar reasoning, we may justify writing the equations

$$A_{\mathrm{III}} = A_{\mathrm{II}}, \tag{207}$$

$$B_{\mathrm{III}} = \beta A_{\mathrm{II}} + B_{\mathrm{II}}, \tag{208}$$

in which β is to be determined. Substitution of Eqs. (205 and (206) into Eqs. (207) and (208) yields

$$A_{\mathrm{III}} = A_{\mathrm{I}} + \alpha B_{\mathrm{I}}, \tag{209}$$

$$B_{\mathrm{III}} = \beta A_{\mathrm{I}} + (1 + \alpha\beta) B_{\mathrm{I}}. \tag{210}$$

Now w has been given a single-valued definition by cutting the plane along the line S_3. This same cut may serve to make $s^{-1/4}$ single-valued. Consider the expression for u on the unshaded edge of the cut:

$$u = s^{-1/4} (A_{\mathrm{I}} e^{ikw} + B_{\mathrm{II}} e^{-ikw}). \tag{211}$$

On the shaded edge

$$u = s^{-1/4} e^{-i\pi/2} (A_{\mathrm{III}} e^{-ikw} + B_{\mathrm{III}} e^{ikw}). \tag{212}$$

In both Eqs. (211) and (212), s and w stand for the values of the functions on the unshaded edge of the cut where w is negative imaginary. Hence e^{ikw} is the dominant term. As u is single-valued, the dominant terms in Eqs. (211) and (212) must be equal. Therefore

$$A_{\mathrm{I}} = -iB_{\mathrm{III}}. \tag{213}$$

From Eqs. (210) and (213),

$$A_{\mathrm{I}}(1 + i\beta) + i(1 + \alpha\beta)B_{\mathrm{I}} = 0. \tag{214}$$

Equation (214) must hold for all values of A_{I} and B_{I}, and this requires the coefficients to vanish. Inspection shows that the coefficients all vanish if $\alpha = \beta = i$. Thus Eqs. (205), (206), (207), and (208) become

$$A_{\mathrm{II}} = A_{\mathrm{I}} + iB_{\mathrm{I}}, \tag{215}$$

$$B_{\mathrm{II}} = B_{\mathrm{I}}, \tag{216}$$

$$A_{\mathrm{III}} = A_{\mathrm{II}}, \tag{217}$$

$$B_{\mathrm{III}} = iA_{\mathrm{II}} + B_{\mathrm{II}}. \tag{218}$$

These are the fundamental connection formulas which show how an asymptotic representation in one Stokes region must be transformed to represent the same solution of Eq. (187) in another Stokes region.

We are now ready to consider the case of a surface duct for which the N-profile is as shown in Fig. 2·11. By referring to the fundamental theorem and recalling that the asymptotic representation of $H_0^{(2)}(\kappa_m r)$ contains a factor $e^{-i\kappa_m r}$, we see that $\mathrm{Im}(\kappa_m)$ determines the rate at which the mode is attenuated with distance. Thus, strongly "trapped" modes for which the attenuation is small have characteristic values that are almost entirely real, and it is with these modes that we shall be concerned for the present.

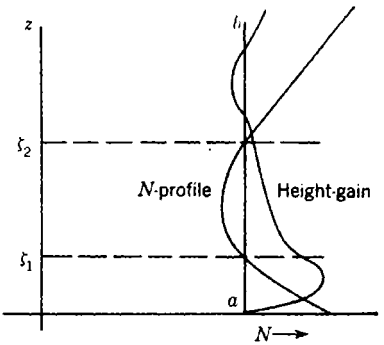


FIG. 2·11.—The height-gain function of a trapped mode. The region for which $\zeta_1 < z < \zeta_2$ serves as a barrier to confine the mode to the duct.

Because N is very nearly unity, Eq. (188) can be approximated by

$$s = 2\,(N - \sigma - i\tau), \tag{219}$$

where

$$\sigma + i\tau = \frac{k^2 + \kappa^2}{2k^2}. \tag{220}$$

If, in accordance with our assumption that the mode is strongly trapped, we neglect τ in comparison with α as a first approximation, s is real and has real zeros ζ_1 and ζ_2, which are the heights at which N equals σ. Then ζ_1 and ζ_2 can be obtained graphically as shown in Fig. 2·11 by finding the heights at which the line ab drawn at a distance σ from the z-axis intersects the N-profile. Under these conditions u is also real, and it follows from Eq. (187) that u is oscillatory or exponential according to whether s is positive or negative. This is shown in Fig. 2·11 where the line ab is used as the axis for which $u = 0$. Thus when $0 < z < \zeta_1$, $s > 0$ and u is oscillatory, the case illustrated being the first mode, which has only one loop. For $\zeta_1 < z < \zeta_2$, s is negative and u behaves like an exponential. Finally, for $\zeta_2 < z$, s is positive and u is again oscillatory.

Although the differential equation permits of solutions that both increase and decrease exponentially with height in the region for which $\zeta_1 < z < \zeta_2$, it turns out that the solution which satisfies the boundary conditions always decreases. For this reason the region $\zeta_1 < z < \zeta_2$ is referred to as a *barrier*, a term borrowed from quantum mechanics. This barrier may be thought of as a lid that tends to confine the mode to the duct.

Now τ is never strictly zero, nor is u strictly real. It is better therefore to regard Fig. 2·11 as representing the real part of u rather than u itself. As we shall see, the imaginary part of u is essentially zero until z becomes greater than ζ_2, after which it gradually increases in oscillatory fashion. Well above the barrier u spirals about the origin of the complex plane in a clockwise direction with increasing magnitude as z increases toward infinity.

The qualitative description of the behavior of u may now be made quantitative by the application of the asymptotic approximations. Consider first the behavior of the asymptotic solutions in the neighborhood of ζ_2. As the characteristic value is essentially real, s is real on the real axis and Line I becomes that portion of the real axis for which $\zeta_2 < z$. Line S_2 is that portion of the real axis lying to the left of ζ_2. Because the boundary condition at infinity requires a wave traveling upward, we must choose f_2 as the asymptotic representation of u on Line I. Accordingly we have

$$A_{\mathrm{I}} = 0, \tag{221}$$

$$B_{\mathrm{I}} = c, \tag{222}$$

where c is a factor independent of z to which a convenient value is to be assigned later. Hence from Eqs. (215) and (216)

$$A_{\mathrm{II}} = ic, \tag{223}$$

$$B_{\mathrm{II}} = c. \tag{224}$$

Along S_2, f_1 is dominant, and except in the immediate neighborhood of ζ_2 the subdominant term can be neglected. Thus we are led to

$$u = c\bar{s}^{-\frac{1}{4}} \exp\left(-ik \int_{\zeta_2}^{z} s^{\frac{1}{2}}\, dz\right), \zeta_2 < z, \tag{225}$$

$$u = ice^{-i(\pi/4)}|\bar{s}|^{-\frac{1}{4}} \exp\left(k \int_{z}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\, dz\right), \zeta_1 < z < \zeta_2. \tag{226}$$

In Eqs. (225) and (226) $\bar{s}$ is written to represent s when the imaginary part of κ has been neglected. Thus

$$\bar{s} = N^2 - \frac{\mathrm{Re}(\kappa^2)}{k^2}.$$

The reason that s and not $\bar{s}$ appears in the exponential of Eq. (225) is that for large values of z the imaginary part of the product of k times the integral, though small compared with the real part, is nevertheless large compared with unity and hence affects the magnitude of u appreciably.

Since

$$\int_{z}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\, dz = -\int_{\zeta_1}^{z} |\bar{s}|^{\frac{1}{2}}\, dz + \int_{\zeta_1}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\, dz, \tag{227}$$

Eq. (226) can be written

$$u = ce^{i(\pi/4)}|\bar{s}|^{-\frac{1}{4}} \left(\exp k \int_{\zeta_1}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\, dz\right) \exp\left(-k \int_{\zeta_1}^{z} |\bar{s}|^{\frac{1}{2}}\, dz\right), \ \zeta_1 < z < \zeta_2. \tag{228}$$

Next it is necessary to examine the behavior of the asymptotic approximations in the neighborhood of ζ_1. In this case Line I extends along the real axis to the left of ζ_1 and the Stokes line S_2 extends to the right. If f_1 and f_2 refer to the asymptotic expressions associated with the point ζ_1, then along S_2

$$f_1 = |\bar{s}|^{-\frac{1}{4}} e^{-i(\pi/4)} \exp\left(k \int_{\zeta_1}^{z} |\bar{s}|^{\frac{1}{2}}\, dz\right), \tag{229}$$

$$f_2 = |\bar{s}|^{-\frac{1}{4}} e^{-i(\pi/4)} \exp\left(-k \int_{\zeta_1}^{z} |\bar{s}|^{\frac{1}{2}}\, dz\right). \tag{230}$$

Comparison of Eqs. (229) and (230) with Eq. (228) shows that

$$A_{II} = 0, \tag{231}$$

$$B_{II} = ce^{i(\pi/2)} \exp k \int_{\zeta_1}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\, dz. \tag{232}$$

From Eqs. (231), (215), and (216)

$$A_I = -iB_{II}, \tag{233}$$

$$B_I = B_{II}; \tag{234}$$

and hence, along Line I

$$u = B_{II}\,\bar{s}^{-\frac{1}{4}}\,(-ie^{ikw} + e^{-ikw}) \tag{235}$$

$$= e^{-i(\pi/4)}\,B_{II}\,\bar{s}^{-\frac{1}{4}}\,[e^{ikw-i(\pi/4)} + e^{-ikw+i(\pi/4)}] \tag{236}$$

$$= 2e^{-i(\pi/4)}\,B_{II}\,\bar{s}^{-\frac{1}{4}}\cos\left(kw - \frac{\pi}{4}\right). \tag{237}$$

Account is taken of the fact that w is positive along Line I by writing

$$w = \int_z^{\zeta_1} \bar{s}^{\frac{1}{2}}\,dz. \tag{238}$$

Equations (238), (237), and (232) yield

$$u = 2ce^{i(\pi/4)}\left(\exp k\int_{\zeta_1}^{\zeta_2} |\bar{s}|^{\frac{1}{2}}\,dz\right)\bar{s}^{-\frac{1}{4}}\cos\left(k\int_z^{\zeta_1}\bar{s}^{\frac{1}{2}}\,dz - \frac{\pi}{4}\right),$$
$$0 \leq z \leq \zeta_1. \tag{239}$$

For trapped modes it is proper to use the simplified boundary condition, which requires merely that the height-gain function vanish at the surface. Thus by equating the argument of the cosine to an odd multiple of $\pi/2$ we are led to the following relationship for determining the characteristic values κ_m:

$$\int_0^{\zeta_1^{(m)}} \sqrt{k^2N^2 - \mathrm{Re}\,(\kappa_m^2)}\,dz = (m - \tfrac{1}{4})\pi, \qquad m = 1,2,\cdots, \tag{240}$$

where $\zeta_1^{(m)}$ is the smaller root of the equation

$$k^2N^2 - \mathrm{Re}(\kappa_m^2) = 0. \tag{241}$$

In terms of the simplification represented by Eqs. (219) and (220), the phase-integral condition given by Eq. (240) becomes

$$\sqrt{2}k\int_0^{\zeta_1^{(m)}} \sqrt{N - \sigma_m}\,dz = (m - \tfrac{1}{4})\pi, \qquad m = 1,2,\cdots, \tag{242}$$

where $\zeta_1^{(m)}$ is the smaller root of the equation

$$N - \sigma_m = 0. \tag{243}$$

Analytic application of the phase-integral method is limited to those index distributions for which the integrations indicated in Eq. (240) or (242) can be carried out. It is to be noted that the conditions just derived determine only the real parts of the characteristic values. Means for determining the imaginary parts will be discussed later.

The unnormalized height-gain functions and the real part of the characteristic values have been obtained. Next, we shall normalize the height-gain functions. The first step is to obtain the fundamental theorem for

the simplified boundary condition by letting p become infinite in Eq. (173). Thus

$$\psi = -\pi i \sum_{m=1}^{\infty} H_0^{(2)}(\kappa_m r') \frac{2}{u_2'(0,\kappa_m)} \frac{\kappa_m}{\left[\frac{\partial}{\partial\kappa} u_2(0,\kappa)\right]_{\kappa=\kappa_m}} u_2(z,\kappa_m) u_2(z_1,\kappa_m) + \tfrac{1}{2} \int_B \kappa H_0^{(2)}(\kappa r) v \, d\kappa. \tag{244}$$

To obtain Eq. (244) use is made of the fact that for the simplified boundary condition

$$\text{Wronskian } W = u_1(0,\kappa_m) u_2'(0,\kappa_m). \tag{245}$$

By comparing Eqs. (244) and (164), we see that the normalized height-gain functions are given by

$$u_m^{(1)}(z) = u_2(z,\kappa_m) \sqrt{\frac{2\kappa_m}{u_2'(0,\kappa_m)\left[\frac{\partial}{\partial\kappa} u_2(0,\kappa)\right]_{\kappa=\kappa_m}}}. \tag{246}$$

Examination of Eq. (246) shows that the right-hand side is unchanged when u_2 is multiplied by an arbitrary function of κ. Thus if $f(\kappa)$ denotes an arbitrary function of κ and use is made of the fact that $u_2(0, \kappa_m) = 0$, we obtain

$$\left[\frac{\partial}{\partial\kappa} f(\kappa) u_2(0,\kappa)\right]_{\kappa=\kappa_m} = f(\kappa_m) \left[\frac{\partial}{\partial\kappa} u_2(0,\kappa)\right]_{\kappa=\kappa_m}. \tag{247}$$

From this it is obvious that if u_2 is replaced by $f(\kappa)\, u_2$ in Eq. (246), the equation is unchanged. It is therefore proper and convenient to choose c so that

$$2ce^{i(\pi/4)} \exp\left(k \int_{\zeta_1}^{\zeta_2} |\bar{s}|^{1/2}\, dz\right) = 1. \tag{248}$$

Then Eq. (239) becomes

$$u = \bar{s}^{-1/4} \cos\left(k \int_z^{\zeta_1} \bar{s}^{1/2}\, dz - \frac{\pi}{4}\right), \qquad 0 \le z < \zeta_1. \tag{249}$$

Now u given by Eq. (249) may be used for u_2 in Eq. (246). Then neglecting the imaginary part of κ completely, we obtain

$$\frac{\partial}{\partial\kappa} u(0) = \bar{s}^{1/4}(0) u(0) \frac{\partial}{\partial\kappa} \bar{s}^{-1/4}(0) - k\bar{s}^{-1/4}(0) \sin\left(k \int_0^{\zeta_1} \bar{s}^{1/2}\, dz - \frac{\pi}{4}\right) \frac{d}{d\kappa} \int_0^{\zeta_1} \bar{s}^{1/2}\, dz. \tag{250}$$

But

$$\frac{d}{d\kappa} \int_0^{\zeta_1} \bar{s}^{1/2}\, dz = \frac{1}{2} \int_0^{\zeta_1} \bar{s}^{-1/2} \frac{\partial \bar{s}}{\partial\kappa}\, dz + \frac{d\zeta_1}{d\kappa} \bar{s}^{1/2}(\zeta_1). \tag{251}$$

Because ζ_1 is a zero of $\bar{s}$, the second term on the right vanishes. From Eq. (188)

$$\frac{\partial s}{\partial \kappa} = \frac{-2\kappa}{k^2}, \tag{252}$$

and Eq. (251) simplifies to

$$\frac{d}{d\kappa}\int_0^{\zeta_1} \bar{s}^{1/2}\, dz = -\frac{\kappa}{k^2}\int_0^{\zeta_1} \bar{s}^{-1/2}\, dz. \tag{253}$$

When $\kappa = \kappa_m$, $u(0) = 0$ and because of Eq. (240), the argument of the sine becomes an odd multiple of $\pi/2$. Thus from Eqs. (250) and (253)

$$\left[\frac{\partial}{\partial \kappa} u(0)\right]_{\kappa=\kappa_m} = \frac{\kappa_m}{k}\bar{s}_m^{-1/4}(0)(-1)^{m-1}\int_0^{\zeta_1^{(m)}} \bar{s}_m^{-1/2}\, dz, \tag{254}$$

where $\bar{s}_m$ is used to denote the value of $\bar{s}$ when $\kappa = \kappa_m$. Similarly from Eq. (249)

$$\frac{\partial u}{\partial z}(0,\kappa_m) = k\bar{s}_m^{1/4}(0)(-1)^{m-1}. \tag{255}$$

Hence from Eqs. (254) and (255),

$$\frac{2\kappa}{u_2'(0,\kappa_m)\left\{\dfrac{\partial}{\partial \kappa} u_2(0,\kappa)\right\}_{\kappa=\kappa_m}} = \frac{2}{\displaystyle\int_0^{\zeta_1^{(m)}} \bar{s}_m^{-1/2}\, dz}. \tag{256}$$

Equation (256) may be used in Eq. (250) to obtain the normalized height-gain functions.

The results so far obtained may be summarized as follows

$$u_m^{(1)}(z) = u_m(z)\sqrt{\frac{2}{\displaystyle\int_0^{\zeta_1^{(m)}} \bar{s}_m^{-1/2}\, dz}}, \tag{257}$$

where

$$u_m(z) = \bar{s}_m^{-1/4}\cos\left(k\int_z^{\zeta_1^{(m)}} \bar{s}_m^{1/2}\, dz - \frac{\pi}{4}\right), \qquad 0 \le z < \zeta_1^{(m)}, \tag{258}$$

$$u_m(z) = \frac{1}{2}|\bar{s}_m|^{-1/4}\exp\left(-k\int_{\zeta_1^{(m)}}^{z} |\bar{s}_m|^{1/2}\, dz\right), \qquad \zeta_1^{(m)} < z < \zeta_2^{(m)}, \tag{259}$$

$$u_m(z) = \frac{e^{-i(\pi/4)}}{2}\bar{s}_m^{-1/4}\exp\left(-k\int_{\zeta_1^{(m)}}^{\zeta_2^{(m)}} |\bar{s}_m|^{1/2}\, dz\right) \times \exp\left(-ik\int_{\zeta_2^{(m)}}^{z} s_m^{1/2}\, dz\right), \qquad \zeta_2^{(m)} < z \tag{260}$$

$$\bar{s}_m = N^2 - \frac{\mathrm{Re}(\kappa_m^2)}{k^2}, \tag{261}$$

$$s_m = N^2 - \frac{\kappa_m^2}{k^2}. \tag{262}$$

Here $\zeta_1^{(m)}$ and $\zeta_2^{(m)}$ are zeros of $\bar{s}_m$ and $\zeta_1^{(m)} < \zeta_2^{(m)}$; and $\mathrm{Re}(\kappa_m^2)$ is determined by Eq. (240).

Equations (258) to (260) are good approximations to u_m except in the neighborhood of $\zeta_1^{(m)}$ and $\zeta_2^{(m)}$, where the expressions on the right have singularities. Although it is possible to find other analytic expressions that are good approximations in the neighborhood of $\zeta_1^{(m)}$ and $\zeta_2^{(m)}$, it is usually sufficient to bridge the gaps by graphical interpolation.

In order to complete the phase-integral formulation for the strongly trapped modes, it is necessary that an expression be obtained for the imaginary parts of the characteristic values. This may be done in the following way.

As mentioned above, u is essentially real for real z up to values somewhat greater than ζ_2, after which it acquires an imaginary part and spirals about the origin with increasing magnitude. Let H be a value of z greater than ζ_2 for which u is still essentially real. From Eq. (187),

$$\frac{d^2u^*}{dz^2} + k^2 s^* u^* = 0, \tag{263}$$

where the asterisk is used to denote the complex conjugate of any quantity to which it is affixed. Subtracting the result of multiplying Eq. (263) by u from the result of multiplying Eq. (187) by u^* yields

$$u^* \frac{d^2u}{dz^2} - u \frac{d^2u^*}{dz^2} + k^2(s - s^*)u^* u = 0. \tag{264}$$

By integrating from 0 to H by parts, we obtain

$$\left(u^* \frac{du}{dz} - u \frac{du^*}{dz}\right)_0^H + k^2 (s - s^*) \int_0^H u^* u \, dz = 0. \tag{265}$$

But from Eq. (188),

$$k^2(s - s^*) = -2i \, \mathrm{Im} \, (\kappa^2). \tag{266}$$

If the normalized characteristic function $u_m^{(1)}$ be substituted for u in Eq. (265), the result is

$$\mathrm{Im} \, (\kappa_m^2) = -\frac{i}{2} \frac{\left[u_m^{(1)*}(H) \frac{du_m^{(1)}}{dz}(H) - u_m^{(1)}(H) \frac{du_m^{(1)*}}{dz}(H) \right]}{\int_0^H |u_m^{(1)}|^2 \, dz}. \tag{267}$$

Now H is in the region where Eq. (260) is valid even with s_m replaced by $\bar{s}_m$ in the exponential. Using Eq. (260) with this modification, we can

transform Eq. (267) into

$$\text{Im}\,(\kappa_m^2) = \frac{\dfrac{-k}{2}\exp\left(-2k\int_{\zeta_1^{(m)}}^{\zeta_2^{(m)}}|\bar{s}_m|^{1/2}\,dz\right)}{\displaystyle\int_0^{\zeta_1^{(m)}}\bar{s}_m^{-1/2}\,dz\int_0^H|u_m^{(1)}|^2\,dz}. \tag{268}$$

For the simplified boundary condition, Eq. (166) becomes

$$u_m^{(1)}(z) = \frac{u_2(z,\kappa_m)}{\sqrt{\displaystyle\int_0^\infty u_2^2(z,\kappa_m)\,dz}}. \tag{269}$$

From this it follows that

$$\int_0^\infty [u_m^{(1)}(z)]^2\,dz = 1, \tag{270}$$

where the path of integration goes to infinity along a ray in the fourth quadrant. Now the path may extend along the real axis from 0 to H and then go to infinity in the fourth quadrant. Thus

$$\int_0^H [u_m^{(1)}(z)]^2\,dz + \int_H^\infty [u_m^{(1)}(z)]^2\,dz = 1. \tag{271}$$

By careful examination of Eq. (260) it can be shown that $u_m^{(1)}$ decreases rapidly as z departs from the real axis and moves into the fourth quadrant. It is therefore a good approximation to neglect the second integral in Eq. (271) and write

$$\int_0^H [u_m^{(1)}(z)]^2\,dz = 1. \tag{272}$$

Because $u_m^{(1)}$ is essentially real for $0 < z < \text{H}$, we have

$$\int_0^H |u_m^{(1)}|^2\,dz = \int_0^H [u_m^{(1)}(z)]^2\,dz = 1. \tag{273}$$

Then Eq. (268) becomes

$$\text{Im}\,(\kappa_m^2) = \frac{-k\exp\left(-2k\displaystyle\int_{\zeta_1^{(m)}}^{\zeta_2^{(m)}}|\bar{s}_m|^{1/2}\,dz\right)}{2\displaystyle\int_0^{\zeta_1^{(m)}}\bar{s}_m^{-1/2}\,dz}. \tag{274}$$

This is the phase-integral expression for the imaginary part of the characteristic value and completes the set of formulas for the trapped modes.

There is a close relationship between phase-integral and ray-tracing methods. To illustrate this point we shall show how a physical meaning can be attached to the denominator of the right-hand member of Eq. (274).

Let us define α_m by

$$N^2(0) \cos^2 \alpha_m = \frac{\mathrm{Re}(\kappa_m^2)}{k^2}. \tag{275}$$

From Fig. 2·11 it is evident that for a Gamow mode, $\cos \alpha_m$ is less than unity and α_m is, in fact, a real angle. We may therefore draw a ray using α_m as the initial angle of inclination at the earth's surface. The differential equation for this ray is by Eq. (92)

$$\frac{dz}{dr} = \frac{\sqrt{N^2 - \dfrac{\mathrm{Re}(\kappa_m^2)}{k^2}}}{\dfrac{\mathrm{Re}(\kappa_m)}{k}}. \tag{276}$$

In writing Eq. (276) κ_m^2 is assumed to be so nearly real that $\mathrm{Re}(\kappa_m)$ is a good approximation to $\sqrt{\mathrm{Re}(\kappa_m^2)}$. From Eq. (276) it follows that $\zeta_1^{(m)}$ is the height of the turning point of the ray. Accordingly, if r_m be used to denote the distance from the starting point of the ray at the surface to the point at which the ray returns to the surface again, we can write

$$r_m = \frac{2 \,\mathrm{Re}(\kappa_m)}{k} \int_0^{\zeta_1} \overline{s}_m^{\frac{1}{2}} \, dz. \tag{277a}$$

Now $\mathrm{Re}(\kappa_m)/k$ is very nearly unity so that to a good approximation

$$r_m = 2 \int_0^{\zeta_1} \overline{s}_w^{\frac{1}{2}} \, dz. \tag{277b}$$

Thus the denominator of the right-hand member of Eq. (274) is just the length of the "first hop" of the ray that has been associated with the mode. This is the physical meaning of the integral.

Making use of Eq. (277*b*) and the fact that $|\kappa_m| \approx k$, we find that for a Gamow mode

$$\left| \frac{\mathrm{Im}\,(\kappa_m^2)}{\kappa_m^2} \right| \approx \frac{\exp\left(-2k \displaystyle\int_{\zeta_1^{(m)}}^{\zeta_2^{(m)}} |\overline{s}_m|^{\frac{1}{2}} \, dz \right)}{kr_m}. \tag{278}$$

The numerator of the right-hand member of Eq. (278) is never greater than unity, and the denominator, which is essentially the length of the "first hop" in wavelengths, is always much greater. In this connection it should be pointed out that the length of a first hop is often measured in miles. Accordingly Eq. (278) confirms that for a Gamow mode, the imaginary part is a very small fraction of the total characteristic value.

It should be noticed that for a mode to be treated as a Gamow mode it is necessary that there be a barrier. This means that the coefficient of the term involving u in Eq. (187) must be essentially real and negative

over an interval of values of z. For an N-profile such as is shown in Fig. 2·11, this condition may be stated in terms of the inequality

$$N_{\min}^2 < \frac{\mathrm{Re}(\kappa_m^2)}{k^2}, \tag{279}$$

where $N_{\min}$ is the minimum value of N. The inequality (279) is frequently used as a criterion for determining if a given mode is to be considered trapped, a trapped mode being one that satisfies the inequality. It should be emphasized that this is an arbitrary criterion and that physically there is no sharp demarcation between trapped and untrapped modes.

The phase-integral method, which is now to be examined for dealing with strongly leaking or "Eckersley modes," is most readily introduced by discussing a simple example. Let us consider the case of a linear N-profile for which

$$N = n_0(1 + \tfrac{1}{2} qz) \tag{280}$$

or

$$N^2 = n_0^2(1 + qz), \tag{281}$$

where n_0 is the value of the index at the surface and q is twice the reciprocal of the effective radius of the earth. Then

$$s = n_0^2(1 + qz) - \frac{\kappa^2}{k^2} = n_0^2 q(z - z_0), \tag{282}$$

where z_0, the zero of s, is related to κ by

$$z_0 = \frac{\kappa^2 - k_0^2}{k_0^2 q}. \tag{283}$$

In Eq. (283), k_0 is an abbreviation for kn_0. Then

$$\int_{z_0}^{z} s^{1/2}\, dz = \frac{2}{3} n_0 q^{1/2}(z - z_0)^{3/2}. \tag{284}$$

From Eq. (284) it is obvious that the anti-Stokes lines make angles of 0, $2\pi/3$, and $4\pi/3$ with the positive axis of reals. We shall choose the line parallel to the axis of reals as Line I of the Stokes pattern. The solution representing a wave traveling outward for large real values of z must be given on the real axis by

$$u = f_2 = [n_0^2 q(z - z_0)]^{-1/4} \exp\left[\frac{-2i}{3} kn_0 q^{1/2}(z - z_0)^{3/2}\right]. \tag{285}$$

Thus $A_{\mathrm{I}} = 0$, $B_{\mathrm{I}} = 1$ and from Eqs. (215), (216), (217), and (218)

$$A_{\mathrm{II}} = i,$$
$$B_{\mathrm{II}} = 1,$$
$$A_{\mathrm{III}} = i,$$
$$B_{\mathrm{III}} = 0.$$

Consider now the simplified boundary condition which requires u to vanish when $z = 0$. This can occur only if the asymptotic representation of u contains both f_1 and f_2, as f_1 and f_2 cannot vanish individually. Thus the origin of the z-plane must fall in Stokes region II because only in that region are both A and B different from zero. Making use of the values of A_{II} and B_{II} as given before we find that in region II

$$u = 2\,[n_0^2 q(z - z_0)]^{-\frac{1}{4}} e^{i(\pi/4)} \cos\left[\frac{2}{3} k n_0 q^{\frac{1}{2}} (z - z_0)^{\frac{3}{2}} + \frac{\pi}{4}\right]. \tag{286}$$

The simplified boundary condition is satisfied if the argument of the cosine for $z = 0$ is an odd multiple of $\pi/2$. Thus

$$\frac{2}{3} k n_0 q^{\frac{1}{2}} [-z_0^{(m)}]^{\frac{3}{2}} + \frac{\pi}{4} = \left(m - \frac{1}{2}\right)\pi, \tag{287}$$

whence

$$z_0^{(m)} = \left[\frac{3}{2} \frac{(m - \frac{3}{4})\pi}{k n_0 q^{\frac{1}{2}}}\right]^{\frac{2}{3}} e^{-\pi i/3}. \tag{288}$$

From the fact that the origin must fall near Line II, it is obvious that of the possible values $\pi/3$, π, and $-\pi/3$ for arg $z_0^{(m)}$, $-\pi/3$ is the proper one to choose.

The simple case just treated points the way to a generalized discussion of Eckersley modes. The procedure is to assume that for a characteristic value κ_m^2, a zero $z_0^{(m)}$ of s_m exists for which the Stokes pattern is similar to that encountered in the example above; namely, Line I extends to the right roughly parallel to the real axis, and the origin of the z-plane falls near Line II. Then because u_m must represent a wave traveling outward for large z,

$$u_m(z) = \tfrac{1}{2} e^{-i(\pi/4)} s_m^{-\frac{1}{4}} \exp\left(-ik \int_{z_0^{(m)}}^{z} s_m^{\frac{1}{2}}\, dz\right), \qquad z \gg z_m', \tag{289}$$

where z_m' is the point at which the real axis crosses the Stokes line S_1. The constant factor is simply for convenience. Making use of the connection formulas we can establish that in Region II

$$u_m(z) = s_m^{-\frac{1}{4}} \cos\left(k \int_{z}^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4}\right), \qquad z \ll z_m'. \tag{290}$$

As before, it is only in Region II that u_m can possibly vanish, and hence $z = 0$ must lie in Region II. For the simplified boundary condition, it is necessary that the argument of the cosine be an odd multiple of $\pi/2$ when $z = 0$, and hence

$$k \int_{0}^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz = (m - \tfrac{1}{4})\pi. \tag{291}$$

Equation (291) determines $z_0^{(m)}$ from which κ can be obtained from the relationship

$$\kappa_m^2 = k^2 N^2(z_0^{(m)}). \tag{292}$$

It is possible to modify Eq. (291) to include the general boundary condition. From Eq. (290)

$$u_m'(z) = s_m^{-\frac{1}{4}} \sin\left(k \int_z^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4}\right) k s_m^{\frac{1}{2}}. \tag{293}$$

In obtaining Eq. (293), $s_m^{-\frac{1}{4}}$ is treated as a constant according to the rule for differentiating asymptotic expressions.

From Eqs. (290) and (293), it follows that

$$\begin{aligned} &p u_m(0) - u_m'(0) \\ &\quad = s_m^{-\frac{1}{4}}(0) \sqrt{p^2(\kappa_m^2) + k^2 s_m(0)} \cos\left(k \int_0^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4} + \delta_m\right), \end{aligned} \tag{294}$$

where

$$\cos \delta_m = \frac{p(\kappa_m^2)}{\sqrt{p^2(\kappa_m^2) + k^2 s_m(0)}}; \qquad \sin \delta_m = \frac{k s_m^{\frac{1}{2}}(0)}{\sqrt{p^2(\kappa_m^2) + k^2 s_m(0)}}. \tag{295}$$

The boundary condition is now

$$k \int_0^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz = (m - \tfrac{1}{4})\pi - \delta_m. \tag{296}$$

Equations (292) and (296) are a pair of simultaneous equations for the determination of κ_m^2 and $z_0^{(m)}$. It is evident that solving these equations in any given case is far from a trivial problem.

The final step is to determine the normalized height-gain functions $u_m^{(1)}$. By comparing Eqs. (164) and (176), it is evident that

$$u_m^{(1)} = u_2(z,\kappa_m) \sqrt{\frac{2\kappa_m}{u_2(0,\kappa_m)\left\{\frac{\partial}{\partial \kappa}\left[p u_2(0,\kappa) - u_2'(0,\kappa)\right]\right\}_{\kappa = \kappa_m}}}. \tag{297}$$

Let us use $u_m(z)$ as given in Eq. (290) for $u_2(z,\kappa_m)$ in Eq. (297). Making use of Eq. (294) and dropping the subscript m because κ is not required to be a characteristic value, we obtain

$$\begin{aligned} &\frac{d}{d\kappa}\left[p u(0) - u'(0)\right] \\ &\quad = \frac{d}{d\kappa}\left[s^{-\frac{1}{4}}(0) \sqrt{p^2 + k^2 s(0)}\right] \cos\left(k \int_0^{z_0} s^{\frac{1}{2}}\, dz - \frac{\pi}{4} + \delta\right) \\ &+ \left[s^{-\frac{1}{4}}(0) \sqrt{p^2 + k^2 s(0)}\right] \sin\left(k \int_0^{z_0} s^{\frac{1}{2}}\, dz - \frac{\pi}{4} + \delta\right) \times \left[\frac{\kappa}{k} \int_0^{z_0} s^{-\frac{1}{2}}\, dz - \frac{d\delta}{d\kappa}\right]. \end{aligned} \tag{298}$$

Because of Eq. (296) when κ is put equal to κ_m in Eq. (298), the term involving the cosine vanishes and the sine becomes $(-1)^{m+1}$. Thus

$$\left\{\frac{d}{d\kappa}\left[pu(0) - u'(0)\right]\right\}_{\kappa=\kappa_m} \tag{299}$$
$$= (-1)^{m+1} s_m^{-\frac{1}{4}}(0) \sqrt{p^2(\kappa_m^2) + k^2 s_m(0)}\, \frac{\kappa_m}{k}\left[\int_0^{z_0^{(m)}} s_m^{-\frac{1}{2}}\, dz - \frac{k}{\kappa_m}\left(\frac{d\delta}{d\kappa}\right)_{(\kappa_m)}\right].$$

Now from Eq. (290)

$$u_m(0) = s_m^{-\frac{1}{4}}(0) \cos\left(k\int_0^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4}\right). \tag{300}$$

But from Eq. (296),

$$k\int_0^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4} = \left(m - \frac{1}{2}\right)\pi - \delta_m, \tag{301}$$

and hence,

$$\cos\left(k\int_0^{z_0^{(m)}} s_m^{\frac{1}{2}}\, dz - \frac{\pi}{4}\right) = (-1)^{m+1} \sin\delta_m. \tag{302}$$

Combining Eqs. (295), (297), (299), (300), and (302), we obtain finally

$$u_m^{(1)} = u_m(z) \sqrt{\frac{2}{\displaystyle\int_0^{z_0^{(m)}} s_m^{-\frac{1}{2}}\, dz - \frac{k}{\kappa_m}\left(\frac{d\delta}{d\kappa}\right)_{\kappa_m}}}. \tag{303}$$

To summarize, Eq. (303) gives the normalized height-gain functions for an Eckersley mode where u_m is given in turn by Eqs. (289) and (290). The characteristic values are determined by Eqs. (292) and (296).

The validity of the phase-integral method for the Eckersley modes has not been established in general. Each application must stand on its own merits. In particular, the method is based on the tacit assumption that $z_0^{(m)}$ is an isolated zero. It is likely in the case of modes bordering on the Gamow type that the formulas developed for Eckersley modes fail because of another zero approaching too close to $z_0^{(m)}$.

THE LINEAR MODIFIED-INDEX PROFILE

BY JOHN E. FREEHAFER

2·9. The Properties of Solutions of $d^2y/d\zeta^2 + \zeta y = 0$.—In the special cases that follow, use must be made of the solutions of the differential equation

$$\frac{d^2y}{d\zeta^2} + \zeta y = 0. \tag{304}$$

The nature of the integrals of this equation must therefore be discussed briefly.

The only singular point of Eq. (304) is an irregular singular point at infinity. Appropriate solutions will therefore be sought in the form of a Laplace integral

$$y = \int_{\alpha}^{\beta} e^{\zeta t} v(t)\, dt. \tag{305}$$

To determine $v(t)$, Eq. (305) is substituted into Eq. (304), leading to

$$\frac{d^2y}{d\zeta^2} + \zeta y = \int_{\alpha}^{\beta} [t^2 e^{\zeta t} v(t) + \zeta e^{\zeta t} v(t)]\, dt.$$

But

$$\zeta e^{\zeta t} v = \frac{d}{dt}(v e^{\zeta t}) - e^{\zeta t}\frac{dv}{dt};$$

hence,

$$\frac{d^2y}{d\zeta^2} + \zeta y = \int_{\alpha}^{\beta} \left(t^2 v - \frac{dv}{dt}\right) e^{\zeta t}\, dt + \left(v e^{\zeta t}\right)_{\alpha}^{\beta}$$

Thus, Eq. (305) is a solution of Eq. (304) provided that v is any solution of

$$t^2 v - \frac{dv}{dt} = 0, \tag{306}$$

and provided that the limits of integration are so chosen that

$$v e^{\zeta t}\Big|_{\alpha}^{\beta} = 0. \tag{307}$$

A solution of Eq. (306) is

$$v = e^{t^3/3}, \tag{308}$$

whence Eq. (307) becomes

$$e^{\zeta t + t^3/3}\Big|_{\alpha}^{\beta} = 0. \tag{309}$$

Equation (309) is evidently satisfied if β and α are infinite in magnitude having arguments that are the arguments of the cube roots of -1. Figure 2·12 shows suitable contours in the t-plane along which the integral in Eq. (305) can be evaluated to obtain a solution of Eq. (304).

Three functions are now defined which are solutions of Eq. (304),

$$y_j = \frac{1}{\sqrt{\pi}} \int_{j} e^{\zeta t + (t^3/3)}\, dt. \tag{310}$$

The subscript j takes on the values 1, 2, and 3, and the corresponding j adjacent to the integral sign indicates which path of integration in Fig. 2·12 is to be used.

A simple additive property of the solutions follows immediately:

$$y_1 + y_2 + y_3 = 0.$$

The following rotation properties may also be demonstrated readily by replacing ζ in Eq. (310) by $\zeta e^{2\pi i/3}$ and changing the variable of integration to $te^{2\pi i/3}$:

$$\left.\begin{aligned} y_1(\zeta e^{2\pi i/3}) &= e^{-2\pi i/3} y_2(\zeta), \\ y_2(\zeta e^{2\pi i/3}) &= e^{-2\pi i/3} y_3(\zeta), \\ y_3(\zeta e^{2\pi i/3}) &= e^{-2\pi i/3} y_1(\zeta). \end{aligned}\right\} \tag{311}$$

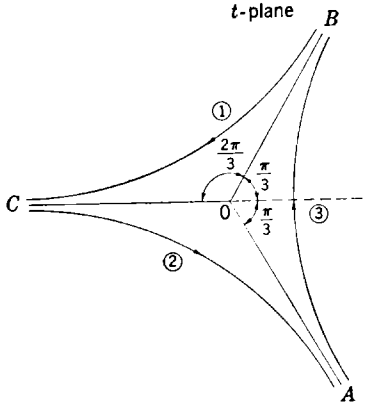


FIG. 2·12.—Paths of integration in the t-plane.

In addition, the symmetry of Fig. 2·12 gives rise to the reflection property

$$y_1(\zeta^*) = -y_2^*(\zeta). \tag{312}$$

If in Eq. (310) e^{t^3} is expanded in a power series in ζt and the integration is carried out term by term, one obtains

$$\left.\begin{aligned} y_1 &= e^{-5\pi i/6} Y_1(\zeta) + e^{-\pi i/6} Y_2(\zeta), \\ y_2 &= e^{-\pi i/6} Y_1(\zeta) + e^{-5\pi i/6} Y_2(\zeta), \\ y_3 &= i Y_1(\zeta) + i\, Y_2(\zeta), \end{aligned}\right\} \tag{313}$$

where

$$\left.\begin{aligned} \frac{\sqrt{\pi}(3)^{1/6}}{\Gamma(\frac{1}{3})} Y_1 &= 1 + \sum_{m=1}^{\infty} (-1)^m (3m-2)(3m-5)\cdots(1)\, \frac{\zeta^{3m}}{(3m)!}, \\ \frac{\sqrt{\pi}}{\Gamma(\frac{2}{3})(3)^{1/6}} Y_2 &= \zeta + \sum_{m=1}^{\infty} (-1)^m (3m-1)(3m-4)\cdots(2)\, \frac{\zeta^{3m+1}}{(3m+1)!}. \end{aligned}\right\} \tag{314}$$

The expressions given in Eq. (314) are the basic fundamental set of solutions at the origin, and Eq. (313) enables us to evaluate the functions for small values of $|\zeta|$. To obtain expressions that are convenient when $|\zeta|$ is large, the integrals will be evaluated by the method of steepest descent. This procedure will be presented in some detail because it contributes to a thorough understanding of the important properties of the functions.

Let

$$w = \xi + i\eta = \zeta t + \frac{t^3}{3},$$

and

$$t = \sigma + i\tau.$$

Imagine a surface constructed by plotting ξ against (σ,τ), the coordinate ξ being perpendicular to the t-plane. In the directions A, B, and C of Fig. 2·12, there are deep valleys in this surface that become deeper as $|t|$ increases. The structure for large $|t|$ is independent of ζ, but the detailed structure for ζt comparable to $t^3/3$ depends in an important way on ζ. In particular at points where $dw/dt = 0$, it is also true that $\partial\xi/\partial\sigma = \partial\xi/\partial\tau = 0$; and because ξ is a solution of Laplace's equation and can have neither maxima nor minima, the surface has saddle points. These saddle points may be thought of as passes that connect the three main valleys. If the paths of integration in Eq. (310) proceed from valley to valley along the lines of steepest descent through the passes, the region that contributes appreciably to the integral can be confined to an interval near the saddle points. This is the essence of the method of steepest descent.

In the following discussion fractional powers of ζ will always denote the principal roots of ζ; that is, $-\pi < \arg\zeta \leqq \pi$. Further restrictions will be imposed from time to time and will be specifically stated. As the roots of $dw/dt = 0$ are $t = \pm i\zeta^{1/2}$, there are, for each value of ζ, two saddle points symmetrically placed with respect to the origin. There is always, therefore, one pair of valleys not directly connected by a pass. Thus, for instance, when ζ is real, the saddle points are on the imaginary axis. It is then possible to pass directly from Valley C to Valley A and obtain y_2 by evaluating the integral in the neighborhood of $-i\zeta^{1/2}$. There is, however, no pass that directly connects Valley A to Valley B, and the path of integration for determining y_3 must therefore first go through the pass from A to C and then through the pass from C to B. Hence the integral y_3 receives contributions from the neighborhood of two saddle points. If, on the other hand,

$$\frac{\pi}{3} < \arg\zeta < \pi,$$

there is no direct pass between Valleys C and A, but y_2 may be evaluated by passing from C to B and then to A. It is therefore given by the sum of two contributions near two saddle points. Thus, depending upon the argument of ζ, there are different approximations to y_2 or to any of the other solutions of the differential equation. This interchange of the connections between valleys and the attendant representation of the same exact solution by different approximate solutions for different ranges of the argument of the independent variable is known as *Stokes' phenomenon.*[1]

Let $f_1(\zeta)$ and $f_2(\zeta)$ represent the approximations obtained by evaluating the integrals in the neighborhood of $i\zeta^{1/2}$ and $-i\zeta^{1/2}$ respectively, and let us consider for the present only f_1. Expanding w in a power series about $i\zeta^{1/2}$ yields

$$w = \tfrac{2}{3}\, i\zeta^{3/2} + i\zeta^{1/2}(t - i\zeta^{1/2})^2 + \tfrac{1}{3}(t - i\zeta^{1/2})^3, \tag{315}$$

and changing variables from t to

$$u = e^{-\pi i/4}\, \zeta^{1/4}(t - i\zeta^{1/2}) \tag{316}$$

leads to

$$w = \tfrac{2}{3}\, i\zeta^{3/2} - u^2 + \tfrac{1}{3}\, \frac{e^{3\pi i/4}}{\zeta^{3/4}}\, u^3. \tag{317}$$

Then

$$f_1(\zeta) = \frac{e^{i\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)}}{\sqrt{\pi}\, \zeta^{1/4}} \int_{-\infty}^{\infty} \exp\left(-u^2 + \frac{1}{3}\, \frac{e^{3\pi i/4}}{\zeta^{3/4}}\, u^3\right) du. \tag{318}$$

Integration along the path of steepest descent is assured by requiring that u be real. By expanding $\exp\left[\frac{1}{3}(e^{3\pi i/4}/\zeta^{3/4})u^3\right]$ in a series of powers of the exponent, integrating the individual terms by parts, and making use of the fact that

$$\int_{-\infty}^{\infty} e^{-u^2}\, du = \sqrt{\pi}. \tag{319}$$

we obtain

$$f_1(\zeta) = \frac{e^{i\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)}}{\zeta^{1/4}} \left[1 + \frac{5\cdot 3}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) + \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2i\zeta^{3/2}}\right)^2 + \cdots\right]. \tag{320}$$

Equation (316) leads to

$$dt = \frac{e^{i\pi/4}}{\zeta^{1/4}}\, du; \tag{321}$$

[1] *Trans. Cambridge Phil. Soc.*, **10**, 106 (1864).

and as arg $(du) = 0$, it follows that

$$\arg(dt) = \frac{\pi - \arg(\zeta)}{4}. \tag{322}$$

But arg (dt) gives the direction and sense of the path of integration in the neighborhood of the saddle point. The portion of Fig. 2·13 lying in the first and second quadrants indicates by means of arrows the direction, in accordance with Eq. (321), of the path of integration at the saddle point whose position is marked by the initial point of the arrow. Thus if the saddle point lies in the sector between the positive real axis and the line OB, the sense of integration is the same as the sense of Path 3 of Fig. 2·12; if it lies in the sector between OB and OC, the sense of integration is opposite that of Path 1. If now $-\pi/3 \leqq \arg \zeta \leqq \pi$, a saddle point lies in the sector BOC; therefore, with reference to the sense indicated in Fig. 2·13, we have

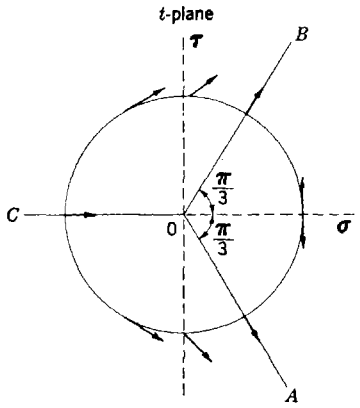


FIG. 2·13.—Directions at the saddle points of the path of steepest descent.

$$y_1 \approx -f_1, \qquad -\frac{\pi}{3} \leqq \arg \zeta \leqq \pi. \tag{323}$$

The symbol $\approx$ is used in this section to mean that the function on the left is represented asymptotically by the one on the right. The nature of this representation becomes clear when we consider how the expression in Eq. (320) was derived. For a given value of ζ there is a definite limit to the number of terms of the series that it is profitable to carry. For as u increases and requires the use of more terms, the path of integration deviates more and more from the straight line in the direction of steepest descent at the saddle point. No matter how many terms are carried, this deviation is not corrected, and the situation is improved only when $|\zeta|$ becomes larger and confines the portion of the path that contributes to the integral to a smaller interval about the saddle point.

By expanding w about the point $-i\zeta^{1/2}$ and using

$$u = e^{i\pi/4}\zeta^{1/4}(t + i\zeta^{1/2}),$$

we obtain the expansion

$$f_2(\zeta) = \frac{e^{-i\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)}}{\zeta^{1/4}}\left[1 - \frac{5\cdot 3}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) + \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2i\zeta^{3/2}}\right)^2 - \cdots\right]. \tag{324}$$

The direction of the path of steepest descent at the saddle points is given by the portions of Fig. 2·13 that lie in the third and fourth quadrants.

When $-\pi < \arg \zeta \leqq -\pi/3$, a saddle point lies in the sector between the positive real axis and the direction OB and another one in the sector AOC. To reach Valley C from Valley B, we must first go through the pass from B to A, which yields a contribution $-f_1$; and then from A to C through a pass, which adds a further contribution $-f_2$. Thus

$$y_1 \approx -(f_1 + f_2), \qquad -\pi < \arg \zeta \leqq -\frac{\pi}{3}. \tag{325}$$

Equations (323) and (325) give the complete asymptotic representations of y_1. When $\arg(\zeta) = -\pi/3$, there are two different expressions. This difficulty is not serious, however; for when $\arg(\zeta) = -\pi/3$ and $|\zeta|$ is large, f_1 behaves like an exponential with large positive exponent and f_2 like an exponential with large negative exponent. The function f_2, therefore, contributes only a small amount to the sum in Eq. (325) and in fact contributes less than the uncertainty already inherent in the asymptotic representation.

By the methods already described one can readily obtain asymptotic representations of $y_2(\zeta)$ and of $dy_1/d\zeta$ and $dy_2/d\zeta$. A summary of the results follows.[1]

$$-\pi < \arg \zeta < -\frac{\pi}{3},$$

$$y_1 \approx -\frac{2}{\zeta^{1/4}}\left\{\cos\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)\left[1 - \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2\zeta^{3/2}}\right)^2 + \cdots\right] + \sin\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)\left[\frac{5\cdot 3}{216}\left(\frac{3}{2\zeta^{3/2}}\right) - \cdots\right]\right\}, \tag{326}$$

$$y_1' \approx 2\zeta^{1/4}\left\{\sin\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)(1 + \cdots) + \cos\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)\left[\frac{21}{216}\left(\frac{3}{2\zeta^{3/2}}\right) + \cdots\right]\right\}. \tag{327}$$

$$-\frac{\pi}{3} \leqq \arg \zeta \leqq \pi,$$

$$y_1 \approx \frac{e^{i\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)}}{\zeta^{1/4}}\left[1 + \frac{5\cdot 3}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) + \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2i\zeta^{3/2}}\right)^2 + \cdots\right], \tag{328}$$

$$y_1' \approx -\zeta^{1/4} i e^{i\left(\frac{2}{3}\zeta^{3/2} + \frac{\pi}{4}\right)}\left[1 - \frac{21}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) + \cdots\right]. \tag{329}$$

$$-\pi \leqq \arg \zeta \leqq \frac{\pi}{3},$$

[1] In the statement of the limits on $\arg(\zeta)$, the equality signs have been inserted in such a way that when two expressions are equally valid, as pointed out in the discussion following Eq. (325), the simpler one is indicated. Thus when ζ is real and negative, it turns out that y_2 is given either by Eq. (330) with $\arg(\zeta) = -\pi$ or by Eq. (332) with

$$y_2 \approx \frac{e^{-i\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)}}{\zeta^{1/4}}\left[1 - \frac{5\cdot 3}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) + \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2i\zeta^{3/2}}\right)^2 - \cdots\right], \quad (330)$$

$$y_2' \approx -i\zeta^{1/4} e^{-i\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)}\left[1 + \frac{21}{216}\left(\frac{3}{2i\zeta^{3/2}}\right) - \cdots\right]. \quad (331)$$

$$\frac{\pi}{3} < \arg\zeta < \pi,$$

$$y_2 \approx \frac{2}{\zeta^{1/4}}\left\{\cos\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)\left[1 - \frac{11\cdot 9\cdot 7\cdot 5}{(216)^2 2!}\left(\frac{3}{2\zeta^{3/2}}\right)^2 + \cdots\right]\right.$$
$$\left. + \sin\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)\left[\frac{5\cdot 3}{216}\left(\frac{3}{2\zeta^{3/2}}\right) - \cdots\right]\right\}, \quad (332)$$

$$y_2' \approx -2\zeta^{1/4}\left\{\sin\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)\left[1 + \cdots\right]\right.$$
$$\left. + \cos\left(\frac{2}{3}\zeta^{3/2}+\frac{\pi}{4}\right)\left[\frac{21}{216}\left(\frac{3}{2\zeta^{3/2}}\right) + \cdots\right]\right\}. \quad (333)$$

In passing it should be noted that

$$y_1 y_2' - y_2 y_1' = \text{Wronskian of } y_1 \text{ and } y_2 = 2i. \quad (334)$$

Two functions h_1 and h_2 are now introduced, which are related to y_1 and y_2 by

$$h_1(\zeta) = \frac{(12)^{1/6}}{\sqrt{\pi}} e^{\pi i/3} y_1(\zeta), \quad (335)$$

$$h_2(\zeta) = \frac{(12)^{1/6}}{\sqrt{\pi}} e^{2\pi i/3} y_2(\zeta). \quad (336)$$

The numerical factors are chosen so that

$$h_j(\zeta) = \left(\frac{2}{3}\zeta^{3/2}\right)^{1/3} H_{1/3}^{(j)}\left(\frac{2}{3}\zeta^{3/2}\right). \quad (337)$$

The h's, unlike the y's, have been tabulated[1] for complex values of the argument and are used instead of the y's when numerical values are needed.

The zeros of h_2 are the zeros of y_2. Let them be denoted by ζ_n. Then it can be shown that

$$\zeta_n = |\zeta_n| e^{2\pi i/3},$$

---

arg $(\zeta) = \pi$. Equation (330) is the simpler expression, for it does not contain the subdominant function; hence the limits are stated so that Eq. (330) and not Eq. (332) is used to obtain y_2, although this is a slight departure from the convention concerning principal roots used in this chapter.

[1] "Tables of the Modified Hankel Functions of Order One-third and of Their Derivatives," *Ann. Harvard Univ. Computation Laboratory*, **2**, 1945. The functions $h_1(\zeta)$, $h_2(\zeta)$, and their derivatives are tabulated to eight decimal places at the points of a square lattice of spacing 0.1. The lattice covers roughly a semicircle of radius 6 in the upper half plane with center at the origin.

where the first three values of $|\zeta_n|$ are given to five figures by

$$|\zeta_1| = 2.3381, \qquad |\zeta_2| = 4.0879, \qquad |\zeta_3| = 5.5206.$$

By requiring the dominant term in Eq. (332) to vanish we are led to the approximate formula

$$|\zeta_n| \approx [\tfrac{3}{2}(n - \tfrac{1}{4})\pi]^{2/3}, \qquad n = 1,2, \cdots. \tag{338}$$

This expression is good to about 1 per cent when $n = 1$, and its accuracy improves with increasing n.

In addition to $h_1(\zeta)$ and $h_2(\zeta)$, the Airy functions $\mathrm{Ai}(\zeta)$ and $\mathrm{Bi}(\zeta)$ have been tabulated.[1]
These are related to the y's by

$$\mathrm{Ai}(-\zeta) = \frac{i}{2\sqrt{\pi}}[y_1(\zeta) + y_2(\zeta)], \tag{339}$$

$$\mathrm{Bi}(-\zeta) = -\frac{1}{2\sqrt{\pi}}[y_1(\zeta) - y_2(\zeta)]. \tag{340}$$

For reference the following additional relationships among the various solutions of Eq. (1) are listed:

$$h_1(\zeta) = -i\,2^{1/3}3^{1/6}\,\mathrm{Ai}(\zeta e^{-\pi i/3}) = -h_2(\zeta e^{-2\pi i/3}), \tag{341}$$

$$h_2(\zeta) = i\,2^{1/3}3^{1/6}\,\mathrm{Ai}(\zeta e^{\pi i/3}) = -h_1(\zeta e^{2\pi i/3}), \tag{342}$$

$$\mathrm{Ai}(-\zeta) = \frac{1}{2^{1/3}3^{1/6}}[e^{\pi i/6}h_1(\zeta) + e^{-\pi i/6}h_2(\zeta)], \tag{343}$$

$$\mathrm{Bi}(-\zeta) = \frac{i}{2^{1/3}3^{1/6}}[e^{\pi i/6}h_1(\zeta) - e^{-\pi i/6}h_2(\zeta)], \tag{344}$$

$$h_2(\zeta^*) = h_1^*(\zeta). \tag{345}$$

2·10. The Field Integral.—Although the formulation of the theory of propagation in a stratified atmosphere can be outlined with considerable generality, the degree to which the details can be treated and numerical results obtained depends upon the specific index distribution assumed. The case that can be carried through most completely with regard to both mathematical detail and numerical results is the important one where the square of the modified index is a linear function of height. This corresponds physically to the situation where the unmodified index is either constant or falls off linearly with height; it therefore includes the special case of standard refraction.

[1] J. C. P. Miller, "The Airy Integral," *British Association Mathematical Tables*, Cambridge, London. These tables give values of $\mathrm{Ai}(\zeta)$ and $\mathrm{Ai}'(\zeta)$ to eight decimal places for $-20 \leqq \zeta \leqq 2$ at intervals of 0.01 in ζ and values of $\mathrm{Bi}(\zeta)$ and $\mathrm{Bi}'(\zeta)$ to eight decimal places for $-10.0 \leqq \zeta \leqq 2.5$ at intervals of 0.1. See also, P. M. and A. M. W. Woodward, "Four-figure Tables of the Airy Function in the Complex Plane," TRE Report No. T1800, Feb. 14, 1945. $\mathrm{Ai}(\zeta)$, $\mathrm{Bi}(\zeta)$ and their derivatives are given to four figures for $-2.4 \leqq \mathrm{Re}(\zeta) \leqq 2.4$ and $-2.4 \leqq \mathrm{Im}(\zeta) \leqq 0$ at the points of a square lattice of spacing 0.2.

Let us assume that

$$N = n_0(1 + \frac{q}{2}z) \tag{346}$$

and

$$N^2 = n_0^2(1 + qz), \tag{347}$$

where n_0 is the value of the index at the surface and q is twice the reciprocal of the effective radius of the earth.

If Eq. (347) is inserted in Eq. (138) and the independent variable is changed from z to $Z = (k^2 n_0^2 q)^{1/3}z$, we obtain

$$\frac{d^2v}{dZ^2} + (Z + A)v = 0, \tag{348}$$

where

$$A = \frac{k_0^2 - \kappa^2}{(k_0^2 q)^{2/3}} \tag{349}$$

and

$$k_0 = kn_0. \tag{350}$$

The quantity $(k_0^2 q)^{-1/3}$ has the dimensions of a length and occurs with sufficient frequency to warrant introducing the abbreviation

$$H \equiv (k_0^2 q)^{-1/3}. \tag{351}$$

Numerical values of H, the natural unit of height, are given for standard refraction in Fig. 2·17. The dimensionless quantity $Z = z/H$ is the height in natural units, and A is the dimensionless separation constant.

The solutions of Eq. (348) that represent incoming and outgoing waves respectively are

$$u_1(z,\kappa) = y_1(Z + A), \tag{352}$$

$$u_2(z,\kappa) = y_2(Z + A). \tag{353}$$

The y's are the functions defined by Eq. (310) and discussed in detail in Sec. 2·9. The substitution of Eqs. (352) and (353) in Eq. (137*b*) yields

$$v = iH\left[\Gamma\frac{y_1(A)}{y_2(A)} + \frac{y_1(Z + A)}{y_2(Z + A)}\right]y_2(Z_1 + A)y_2(Z + A), \tag{354}$$

where

$$\Gamma = -\frac{\dfrac{y_1'(A)}{y_1(A)} - Hp}{\dfrac{y_2'(A)}{y_2(A)} - Hp} \tag{355}$$

and

$$Z_1 = \frac{z_1}{H}.$$

It will appear later that Γ is the reflection coefficient. For the simplified boundary condition, $p \to \infty$ and $\Gamma = -1$.

Only the case in which Z_1 is greater than Z need be considered, for by the reciprocity theorem it does not matter whether Z_1 refers to the position of the transmitter or to the position of the field point. For the sake of symmetry in the equations that follow, Z and z will be replaced by Z_2 and z_2.

It may be verified that assumptions (1) and (2) of Sec. 2·7 and Eq. (170) are satisfied, and we can start the present argument with Eq. (172),

$$\psi = \frac{1}{2} \int_L \kappa H_0^{(2)}(\kappa r) v(z,\kappa)\, d\kappa.$$

By means of Eq. (349), the variable of integration can be changed from κ to A. Thus making use of Eqs. (351 and (354), we obtain

$$\psi = \frac{i}{4H} \int_{-\infty}^{\infty} H_0^{(2)}(k_0 r \sqrt{1 - qHA}) \left[\Gamma \frac{y_1(A)}{y_2(A)} + \frac{y_1(Z_2 + A)}{y_2(Z_2 + A)} \right] y_2(Z_1 + A) y_2(Z_2 + A)\, dA, \qquad (356)$$

where the integration is entirely along the real axis. The square root in the argument of the Hankel function is a positive real number when $(1 - qHA) > 0$, and is a negative imaginary number when $(1 - qHA) < 0$.

Equation (356) is an exact solution of the wave equation (135) in which N^2 is given by Eq. (347) and the boundary conditions are appropriate to a dipole near a plane surface. The question of convergence will be ignored, and various methods for approximating this integral considered.

A useful transformation of the integral of Eq. (356) is based on the assumption that over the region for which the integral is essentially different from zero, qHA is small compared with unity. Then when r is larger than a few wavelengths, the asymptotic expansion for the Hankel function may be used:

$$H_0^{(2)}(k_0 r \sqrt{1 - qHA}) \approx \sqrt{\frac{2}{\pi k_0 r}}\, e^{-i(k_0 r - \pi/4)} e^{ik_0 qHrA/2}. \qquad (357)$$

Now let us introduce a natural unit of range L which is analogous to H and is defined by

$$L \equiv \frac{2}{k_0 qH} = 2(k_0 q^2)^{-1/3}. \qquad (358)$$

The natural unit of range L is given as a function of λ for standard refraction in Fig. 2·18. The range in natural units will be denoted by X; that is

$$X = \frac{r}{L}. \qquad (359)$$

Equation (356), by virtue of Eqs. (357) to (359), becomes

$$\psi \approx \frac{i}{2L} \frac{e^{-i\left(k_0 r - \frac{\pi}{4}\right)}}{\sqrt{\pi}} \Phi(X, Z_1, Z_2), \tag{360}$$

where

$$\Phi(X, Z_1 Z_2) = \frac{1}{\sqrt{X}} \int_{-\infty}^{\infty} e^{iAX} \left[\Gamma \frac{y_1(A)}{y_2(A)} + \frac{y_1(Z_2 + A)}{y_2(Z_2 + A)} \right] y_2(Z_1 + A) y_2(Z_2 + A)\, dA. \tag{361}$$

It is $|\psi|$ with which we are ultimately concerned. From Eq. (360),

$$|\psi| \approx \frac{1}{2\sqrt{\pi} L} |\Phi(X, Z_1, Z_2)|,$$

or, as $F = r|\psi|$,

$$F \approx \frac{X}{2\sqrt{\pi}} |\Phi(X, Z_1, Z_2)|. \tag{362}$$

Thus the ratio of the field in the presence of the earth to the free-space field can be expressed entirely as a function of the natural coordinates X, Z_1, and Z_2. In the case of the simplified boundary condition, when $\Gamma = -1$, Φ does not depend explicitly upon frequency. Equation (362) then expresses the field without explicit mention of the frequency, and the use of natural coordinates serves to summarize all of the cases for which q, z_1, z_2, r, and k_0 are so related that X, Z_1, and Z_2 have the same values.[1] It is important to note that whereas, at frequencies of 100 Mc/sec and above, $\Gamma = -1$ is a good approximation for both polarizations in the diffraction region, it is not an acceptable approximation for vertical polarization in the interference region.

2·11. The Interference Region.—To obtain the field in the interference region, the integral of Eq. (356) will be evaluated by replacing the functions in the integrand with the appropriate asymptotic representations. The method of steepest descent will then be applied, and it will be found that each value of A for which the phase of the integrand is stationary

[1] For generalized coverage diagrams in the interference region, see B. E. Howard, H. Dodson, and J. Gill, "Field Intensity Contours in Generalized Coordinates," RL Report No. 702, May 2, 1945. The coordinates and parameters used for cases where $h/h_t < 1$ are related to X, Z_1, and Z_2, by

$$\frac{h}{h_t} = \frac{Z_2}{Z_1}, \quad \frac{d}{d_0} = \frac{X}{\sqrt{Z_1}}, \quad J = \frac{Z_1^{3/2}}{2\pi}.$$

When $h/h_t > 1$, subscripts 1 and 2 are interchanged. See also C. Domb and M. H. L. Pryce, "Interim Report on Propagation Within and Beyond the Optical Range," ASE Report No. M448, Sept. 1942, and *Jour. IEE*, **94**, Part III, 325 (1947).

corresponds to a possible ray through the transmitter and field point. The fact that significant contributions to the integral occur only in the neighborhood of stationary values of the phase is just Fermat's principle. The resulting expression for ψ is given essentially by

$$\psi = \frac{1}{r} e^{-i\phi_1}[1 + \Gamma D e^{-i(\phi_2 - \phi_1)}], \tag{363}$$

where ϕ_1 is the phase of the direct wave, ϕ_2 is the phase of the reflected wave and D is the divergence factor that takes into account the fact that a pencil of parallel rays incident upon the convex side of a spherical surface diverges after reflection because of the curvature of the surface (see Sec. 5·2).

Recalling that

$$F = |\psi| r = |1 + \Gamma D e^{-i(\phi_2 - \phi_1)}|,$$

we need concern ourselves, in so far as numerical results are concerned, only with establishing that

$$\phi_2 - \phi_1 = 2k_0 \frac{\left(z_2 - \frac{qr_1^2}{4}\right)\left(z_1 - \frac{qr_2^2}{4}\right)}{r}, \tag{364}$$

$$D = \frac{1}{\sqrt{1 + \frac{2qr_1r_2}{r \sin 2\psi_2}}} \approx \frac{1}{\sqrt{1 + \frac{qr_1r_2}{r\psi_2}}}, \tag{365}$$

$$\Gamma_v = \frac{k_1^2 \sin \psi_2 - k_0 \sqrt{k_1^2 - k_0^2 \cos^2 \psi_2}}{k_1^2 \sin \psi_2 + k_0 \sqrt{k_1^2 - k_0^2 \cos^2 \psi_2}} \quad \text{for vertical polarization.} \tag{366}$$

$$\Gamma_h = \frac{k_0 \sin \psi_2 - \sqrt{k_1^2 - k_0^2 \cos^2 \psi_2}}{k_0 \sin \psi_2 + \sqrt{k_1^2 - k_0^2 \cos^2 \psi_2}} \quad \text{for horizontal polarization.} \tag{367}$$

In these equations r_1 and r_2 are the distances from the terminal points to the reflection point and ψ_2 is the grazing angle of incidence at the reflection point (see Fig. 2·14).

In the development of Eq. (363) from Eq. (356) the ray picture will be used as a guide. First to be obtained, therefore, are the equations of the rays and the expressions for the optical lengths. From Sec. 2·3 it follows that the equation for the rays in terms of r and Z_2 is

$$\frac{r}{\cos \chi} = H\left(\int_{Z'}^{Z_2} \frac{dZ}{\sqrt{qHZ + \sin^2 \chi}} + \int_{Z'}^{Z_1} \frac{dZ}{\sqrt{qHZ + \sin^2 \chi}}\right). \tag{368}$$

The symbols in Eq. (368) have the following meanings:

1. The symbol χ is defined by

$$\cos \chi \equiv 1 + \frac{qHZ_0}{2}, \tag{369}$$

where Z_0 is the height at which the ray has its turning point. If the ray intersects the earth, Z_0 is negative, $\cos \chi < 1$, and χ is a real angle. If the ray approaches the earth but turns around before reaching it, Z_0 is positive, $\cos \chi > 1$, and χ is imaginary. In either case $\cos \chi$ and $\sin^2 \chi$ have perfectly definite meanings, and the trigonometric notation is retained as a matter of convenience.

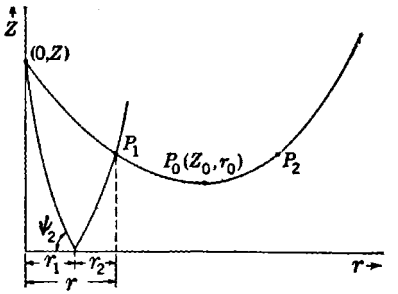


FIG. 2·14.—The geometry of the ray pattern. P_1 and P_2 represent Case *a* and Case *b* geometry respectively. Because of the scale factor involved in Z, ψ_2 is distorted.

2. The coordinate of the lowest point reached by the ray between the terminal points is denoted by Z'. If the ray does not turn between the two terminal points, then $Z' = Z_2$. If the turning point does occur between the terminal points, then, $Z' = Z_0$. In the case of a reflected ray, $Z' = 0$.

The optical length S is given by

$$S = \int N \, d\sigma = n_0 \left[r \cos \chi + H \left(\int_{Z'}^{Z_2} \sqrt{\sin^2 \chi + qHZ} \, dZ + \int_{Z'}^{Z_1} \sqrt{\sin^2 \chi + qHZ} \, dZ \right) \right]. \quad (370)$$

There are obviously always two, and only two, rays between any two terminal points above the horizon. Let us consider first the direct ray. There are two cases, as shown in Fig. 2·14, according to whether the far terminal is to the left of the turning point P_0, as at P_1, or to the right of P_0, as at P_2. They will be designated as Case *a* and Case *b*, respectively.

The relationship among Z_1, Z_2, r, and χ, obtained from Eq. (368) by taking $Z' = Z_2$ for Case *a* and $Z' = Z_0$ for Case *b* is

$$\frac{r}{\cos \chi} = \frac{2}{q} \left[(qHZ_1 + \sin^2 \chi)^{1/2} \mp (qHZ_2 + \sin^2 \chi)^{1/2} \right], \quad (371)$$

and the optical length S_1, given by Eq. (370), is

$$S_1 = n_0 \left\{ r \cos \chi + \frac{2}{3q} \left[(qHZ_1 + \sin^2 \chi)^{3/2} \mp (qHZ + \sin^2 \chi)^{3/2} \right] \right\}. \quad (372)$$

The upper sign refers to Case *a*, and the lower to Case *b*.

To be consistent with the notation commonly accepted, ψ_2 will be used to designate the angle of inclination of the reflected ray at the point of reflection. In contrast to the angle χ for the direct ray, ψ_2 is always real. For the reflected ray, Eqs. (368) and (370) lead to

$$\frac{r}{\cos\psi_2} = \frac{2}{q}\left[(qHZ_1 + \sin^2\psi_2)^{1/2} + (qHZ_2 + \sin^2\psi_2)^{1/2} - 2\sin\psi_2\right], \quad (373)$$

$$S_2 = n_0\left\{r\cos\psi_2 + \frac{2}{3q}\left[(qHZ_1 + \sin^2\psi_2)^{3/2} + (qHZ_2 + \sin^2\psi_2)^{3/2} - 2\sin^3\psi_2\right]\right\}, \quad (374)$$

where S_2 is the optical length along the reflected ray.

Let us now return to a consideration of the field integral. To avoid unimportant complications that hide the essential parts of the discussion, no attempt to present a completely rigorous argument will be made; the purpose is rather to outline the procedure and make it plausible.

To begin with, the Hankel function may be replaced by its asymptotic expression

$$H_0^{(2)}(k_0 r\sqrt{1 - qHA}) \approx \sqrt{\frac{2}{\pi k_0 r}}\,\frac{e^{-i(k_0 r\sqrt{1-qHA} - \pi/4)}}{\sqrt[4]{1-qHA}}. \quad 375$$

In calculating the roots of $(1 - qHA)$ for use in Eq. (375) we must take

$$\arg(1 - qHA) = 0, \qquad qHA < 1 \quad (376)$$

$$\arg(1 - qHA) = -\pi, \qquad qHA > 1. \quad (377)$$

Equation (375) is then usable everywhere except in the neighborhood of $qHA = 1$. If r exceeds a few wavelengths, this region is small and can be dodged by means of a short detour in the complex plane.

The next step is to replace the y's by their asymptotic expressions. In so doing, we must bear in mind that the asymptotic expressions break down when the arguments of the functions become too small. As the arguments are A, $Z_1 + A$, and $Z_2 + A$, the asymptotic expansions fail in the neighborhood of 0, $-Z_1$, and $-Z_2$. These points, together with the point in the vicinity of which the expansion of the Hankel function breaks down, are indicated in Fig. 2·15.

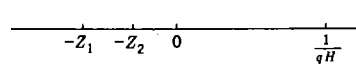


FIG. 2·15.—Points on the real axis of the A-plane in whose neighborhood the asymptotic expansions fail.

Because the functions themselves are well behaved in the critical regions of the asymptotic approximations, these regions will be ignored; they are unimportant when they do not include saddle points, and our simple analysis does not apply when they do.

By using the asymptotic expansions we can now make plausible the assertion that contributions to the integral outside of the region

E. G. & G. LIBRARY
LAS VEGAS BRANCH

$-Z_2 < A < 1/qH$ may usually be neglected. Let us consider first the region $A < -Z_2$. To the accuracy to which the asymptotic expressions are valid

$$y_1(\zeta) = -y_2(\zeta), \text{ when } \zeta \ll 0; \tag{378}$$

hence, for $A \ll 0$ by Eq. (355),

$$\Gamma = -1, \tag{379}$$

and Eq. (354) becomes

$$v = iH\left[y_1(Z_2 + A)y_2(A) - y_2(Z_2 + A)y_1(A)\right]\frac{y_2(Z_1 + A)}{y_2(A)}. \tag{380}$$

Therefore when $A \ll -Z_2$, we conclude from Eqs. (378) and (380) that v is zero. The accuracy of this approximation improves as Z_2 increases. Now let us consider the region $A > 1/qH$.

For $A \gg 0$, Eqs. (328) to (331) and (355) lead to

$$\Gamma \approx \frac{iA^{1/2} - Hp}{iA^{1/2} + Hp}. \tag{381}$$

From Eqs. (139), (140), (349), and (351), it follows that

$$p_v = i\frac{k_0^2}{k_1^2}\sqrt{k_1^2 - k_0^2(1 - qHA)} \qquad \text{for vertical polarization,} \tag{382}$$

$$p_h = i\sqrt{k_1^2 - k_0^2(1 - qHA)} \qquad \text{for horizontal polarization.} \tag{383}$$

The square roots indicated above are to have negative imaginary parts.

From Eqs. (381) to (383), we obtain

$$\left.\begin{aligned} \Gamma_v &\approx \frac{k_1^2 - k_0^2}{k_1^2 + k_0^2} && (384a)\\ \Gamma_h &\approx \frac{H^2(k_0^2 - k_1^2)}{4A} && (384b) \end{aligned}\right\} qHA > \left|\frac{k_1^2 - k_0^2}{k_0^2}\right|.$$

Making use of the fact that $A > 1/qH$ implies that A is large compared with 0, Z_1, and Z_2, we obtain from Eq. (354) and the asymptotic expansions for the y's

$$v = \frac{-iHe^{-iA^{1/2}(Z_1 + Z_2)}[\Gamma + e^{2iZ_2A^{1/2}}]}{A^{1/2}}. \tag{385}$$

When the exponential behavior of the Hankel function is taken into account, it is obvious from Eqs. (384a), (384b), and (385) that the integrand of the field integral approaches zero rapidly when k_0r is large compared with unity. Accordingly contributions to the integral for $A > 1/qH$ will be neglected.

By virtue of the preceding discussion the range of integration will be limited to $-Z_2 < A < 1/qH$. It should be noted that by so doing we are discarding the surface wave, which is unimportant for most applications at frequencies above 100 Mc/sec.

Now let us define three integrals I_1, I_2, and I_3 by

$$I_1 \equiv \int_{-Z_2}^{\frac{1}{qH}} \frac{e^{-ik_0 r\sqrt{1-qHA}}}{\sqrt[4]{1-qHA}}\, y_1(Z_2+A)y_2(Z_1+A)\,dA, \tag{386}$$

$$I_2 \equiv \int_{0}^{\frac{1}{qH}} \frac{e^{-ik_0 r\sqrt{1-qHA}}}{\sqrt[4]{1-qHA}}\, \Gamma\frac{y_1(A)}{y_2(A)}\, y_2(Z_1+A)y_2(Z_2+A)\,dA, \tag{387}$$

$$I_3 \equiv \int_{-Z_2}^{0} \frac{e^{-ik_0 r\sqrt{1-qHA}}}{\sqrt[4]{1-qHA}}\, \Gamma\frac{y_1(A)}{y_2(A)}\, y_2(Z_1+A)y_2(Z_2+A)\,dA. \tag{388}$$

In Eqs. (386) to (388), it is to be understood that integration is not to be carried all the way to the indicated limits but stopped at a sufficient distance from them to permit the use of the asymptotic expansions.

Then Eqs. (356), (375), and (386) to (388) lead to

$$\psi \approx \frac{i}{2^{3/2}H}\frac{e^{i\pi/4}}{\sqrt{\pi k_0 r}}(I_1 + I_2 + I_3). \tag{389}$$

Let us consider I_1 first. If the expressions for y_1 and y_2 as given by the dominant terms of Eqs. (328) and (330) are inserted, then

$$I_1 = -\int_{-Z_2}^{\frac{1}{qH}} \frac{e^{-i\phi_1}\,dA}{[(1-qHA)(Z_1+A)(Z_2+A)]^{1/4}}, \tag{390}$$

where

$$\phi_1 = k_0 r\sqrt{1-qHA} - \tfrac{2}{3}(Z_2+A)^{3/2} + \tfrac{2}{3}(Z_1+A)^{3/2}. \tag{391}$$

To obtain the values of A for which ϕ_1 is stationary we next equate the derivative to zero:

$$\frac{d\phi_1}{dA} = \frac{-qHk_0 r}{2\sqrt{1-qHA}} - (Z_2+A)^{1/2} + (Z_1+A)^{1/2} = 0. \tag{392}$$

By recalling that $1/H^2k_0^2 = Hq$, we can rewrite Eq. (392) as

$$\frac{r}{\sqrt{1-qHA}} = \frac{2}{q}[(qHZ_1+qHA)^{1/2} - (qHZ_2+qHA)^{1/2}]. \tag{393}$$

Comparison of Eqs. (371) and (393) shows that qHA for stationary phase is related to r, Z_1, and Z_2 in the same way that $\sin^2\chi$ is related to them in Case a. Accordingly, when the geometry is that of Case a, A_1, the value

of A for stationary phase [that is, the solution of Eq. (392)], is given by

$$A_1 = \frac{1}{qH} \sin^2 \chi = k_0^2 H^2 \sin^2 \chi. \tag{394}$$

As A_1 is a solution of Eq. (393), differentiation of Eq. (392) yields

$$\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1} = \frac{-k_0 qHr}{4\sqrt{(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)}} \left[1 + \frac{qH\sqrt{(Z_1 + A_1)(Z_2 + A_1)}}{1 - qHA_1}\right]. \tag{395}$$

Thus $(d^2\phi_1/dA^2)_{A_1}$ is negative.

Let ϕ_1 be expanded in a power series about $A = A_1$, and let the terms involving higher powers of $(A - A_1)$ than the second be discarded. Then

$$I_1 \approx -e^{-i\phi_1(A_1)} \int_{-Z_2}^{\frac{1}{qH}} \frac{\exp\left[-i\,\frac{(A - A_1)^2}{2}\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right] dA}{[(1 - qHA)(Z_1 + A)(Z_2 + A)]^{1/4}}. \tag{396}$$

Now the chief contribution to the integral is in the neighborhood of $A = A_1$. Furthermore, except for the exponential, the functions in the integrand vary slowly. Accordingly,

$$I_1 \approx \frac{-e^{-i\phi_1(A_1)}}{[(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)]^{1/4}} \sqrt{\frac{2}{\left|\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right|}} \times e^{+\pi i/4} \int_{-\infty}^{\infty} e^{-s^2}\, ds, \tag{397}$$

where

$$s = e^{-\pi i/4} \frac{A - A_1}{\sqrt{2}} \sqrt{\left|\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right|}.$$

As the integral equals $\sqrt{\pi}$, Eq. (397) becomes

$$I_1 \approx \frac{-\sqrt{2\pi}\, e^{-i[\phi_1(A_1) - \pi/4]}}{\left[(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)\left|\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right|^2\right]^{1/4}}. \tag{398}$$

When A_1 from Eq. (394) is substituted into Eq. (391), we find that $\phi_1(A_1)/k$, the optical length, is given exactly by the right-hand member of Eq. (372) with the upper sign. Thus for the geometry of Case a, I_1 is interpreted as representing the contribution to the field from the direct ray.

Let us next consider I_2. When the asymptotic expansions for the y's as given by Eqs. (328) and (330) are inserted in Eq. (387), we obtain

$$I_2 = -\int_0^{1/qH} \frac{\Gamma e^{-i\phi_2}\, dA}{[(1 - qHA)(Z_1 + A)(Z_2 + A)]^{1/4}}, \tag{399}$$

where

$$\phi_2 = k_0 r \sqrt{1 - qHA} + \tfrac{2}{3}[(Z_1 + A)^{3/2} + (Z_2 + A)^{3/2} - 2A^{3/2}]. \tag{400}$$

Proceeding as before to locate the values of A for which ϕ_2 is stationary by equating $d\phi_2/dA$ to zero, we obtain

$$\frac{r}{\sqrt{1 - qHA}} = \frac{2}{q}[(qHZ_1 + qHA)^{1/2} + (qHZ_2 + qHA)^{1/2} - 2(qHA)^{1/2}]. \tag{401}$$

If A_2 be used to designate the value of A for which ϕ_2 is stationary, a comparison of Eqs. (373) and (401) shows that

$$A_2 = \frac{\sin^2 \psi_2}{qH} = k_0^2 H^2 \sin^2 \psi_2. \tag{402}$$

Expansion of ϕ_2 about A_2 and application of the standard procedure of steepest descent yields

$$I_2 \approx \frac{-\sqrt{2\pi}\, \Gamma(A_2) e^{-i[\phi_2(A_2) - \pi/4]}}{\left[(1 - qHA_2)(Z_1 + A_2)(Z_2 + A_2) \left|\left(\frac{d^2\phi_2}{dA^2}\right)_{A_2}\right|^2\right]^{1/4}}. \tag{403}$$

We note that when Eq. (402) is used in Eqs. (381) to (383), we obtain Eqs. (366) and (367). The expressions on the right of Eqs. (366) and (367) are readily identified as the plane reflection coefficients for two media whose propagation factors are k_1 and k_0. The reflection coefficients are discussed further in Sec. 5·1.

When Eq. (402) is substituted into Eq. (400) we obtain for $\phi_2(A_2)/k$ just the right-hand member of Eq. (374). Thus I_2 is interpreted as the contribution to the field from the reflected wave.

Finally let us consider I_3. Because of Eqs. (378) and (379), Eq. (388) becomes

$$I_3 \approx -i \int_{-Z_2}^{0} \frac{e^{-i\phi_3}\, dA}{[(1 - qHA)(Z_1 + A)(Z_2 + A)]^{1/4}} \tag{404}$$

where

$$\phi_3 = k_0 r \sqrt{1 - qHA} + \tfrac{2}{3}[(Z_1 + A)^{3/2} + (Z_2 + A)^{3/2}]. \tag{405}$$

The equation

$$\frac{d\phi_3}{dA} = 0$$

leads to

$$\frac{r}{\sqrt{1 - qHA}} = \frac{2}{q}\left[(qHZ_1 + qHA)^{1/2} + (qHZ_2 + qHA)^{1/2}\right]. \quad (406)$$

Comparison of Eqs. (406) and (371) shows that if the geometry is such that we are dealing with Case b, ϕ_3 has a stationary value at $A = A_1$, where A_1 is given by Eq. (394). Thus it appears that whereas in Case a, ϕ_1 has a stationary value at $A = A_1$ and ϕ_3 does not, in Case b, ϕ_3 has a stationary value at $A = A_1$ and ϕ_1 does not. Hence, either I_1 or I_3 represents the direct ray, depending upon whether the geometry corresponds to Case a or Case b.

By continuing the method of steepest descent and noting that in this case $d^2\phi_3/dA^2$ is positive, we obtain

$$I_3 \approx \frac{-\sqrt{2\pi}\, e^{-i[\phi_3(A_1) - \pi/4]}}{\left\{(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)\left[\left(\frac{d^2\phi_3}{dA^2}\right)_{A_1}\right]^2\right\}^{1/4}}. \quad (407)$$

The term $\phi_3(A_1)/k$, as expected, is just the right-hand member of Eq. (372) with the positive sign.

In Case a, Eq. (389) together with Eqs. (398) and (403) and the fact that I_3 is negligible leads to

$$\psi = \frac{1}{2H\sqrt{k_0 r}} \frac{e^{-i\phi_1(A_1)} + \Gamma D e^{-i\phi_2(A_2)}}{\left[(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)\left|\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right|^2\right]^{1/4}}. \quad (408)$$

where

$$D = \sqrt[4]{\frac{(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)\left|\left(\frac{d^2\phi_1}{dA^2}\right)_{A_1}\right|^2}{(1 - qHA_2)(Z_1 + A_2)(Z_2 + A_2)\left|\left(\frac{d^2\phi_2}{dA^2}\right)_{A_2}\right|^2}}. \quad (409)$$

By differentiating Eq. (400) twice with respect to A and making use of the fact that $(d\phi_2/dA)_{A_2} = 0$, we find

$$\left(\frac{d^2\phi_2}{dA^2}\right)_{A_2} = \frac{-qHk_0 r}{4\sqrt{(Z_1 + A_2)(Z_2 + A_2)(1 - qHA_2)}} \times \left[1 + \frac{qH(Z_1 + A_2)^{1/2}(Z_2 + A_2)^{1/2}}{1 - qHA_2} + \frac{2Hk_0 r_1 r_2}{A_2^{1/2}(1 - qHA_2)^{1/2} r}\right]. \quad (410)$$

In Eq. (410), r_1 and r_2 are the distances to the reflection point from the terminal points.

$$r_1 = \frac{2(1 - qHA_2)^{1/2}}{qHk_0}\left[(Z_1 + A_2)^{1/2} - A_2^{1/2}\right]; \quad (411)$$

$$r_1 = \frac{2(1 - qHA_2)^{1/2}}{qHk_0}\,[(Z_2 + A_2)^{1/2} - A_2^{1/2}]; \tag{412}$$

$$r = r_1 + r_2 \tag{413}$$

By making use of Eqs. (395) and (410), we can write Eq. (408) as

$$\psi = \frac{1}{r}\,\frac{e^{-i\phi_1(A_1)}\{1 + \Gamma De^{-i[\phi_2(A_2) - \phi_1(A_1)]}\}}{\left(1 + \dfrac{qH\sqrt{(Z_1 + A_1)(Z_2 + A_1)}}{1 - qHA_1}\right)^{1/2}}. \tag{414}$$

Eqs. (395) and (410) reduce Eq. (409) to

$$D = \sqrt{\frac{1 + \dfrac{qH\sqrt{(Z_1 + A_1)(Z_2 + A_1)}}{1 - qHA_1}}{1 + \dfrac{qHk_0r_1r_2}{A_2^{1/2}(1 - qHA_2)^{1/2}r} + \dfrac{qH\sqrt{(Z_1 + A_2)(Z_2 + A_2)}}{1 - qHA_2}}}. \tag{415}$$

In practical cases,

$$\frac{qH\sqrt{(Z_1 + A_i)(Z_2 + A_i)}}{1 - qHA_i} \ll 1, \qquad i = 1, 2;$$

and[1] when these terms are neglected, Eq. (414) becomes Eq. (363) and Eq. (415), by virtue of Eq. (402), becomes Eq. (365).

In Case *b*, we obtain Eqs. (408) and (409) with ϕ_1 replaced everywhere by ϕ_3. Now

$$\left(\frac{d^2\phi_3}{dA^2}\right)_{A_1} = \frac{k_0qHr}{4\sqrt{(1 - qHA_1)(Z_1 + A_1)(Z_2 + A_1)}}\left[1 - \frac{qH\sqrt{(Z_1 + A_1)(Z_2 + A_1)}}{1 - qHA_1}\right]. \tag{416}$$

The difference in the distribution of negative signs in Eqs. (395) and (416) is such as to change the sign of a term that is neglected in the final result. We then arrive as before at Eqs. (363) and (366), where ϕ_1 is now to be interpreted as the phase of the direct ray obtained from the optical length

[1] The expression on the left of this inequality can be written

$$\frac{\sqrt{\left(\dfrac{2}{a_e}z_1 + \psi_3^2\right)\left(\dfrac{2z_2}{a_e} + \psi_2^2\right)}}{1 - \psi_2^2} \qquad \text{for the reflected wave}$$

and

$$\frac{\dfrac{2}{a_e}\sqrt{(z_1 - z_0)(z_2 - z_0)}}{1 + \dfrac{2z_0}{a_e}} \qquad \text{for the direct wave.}$$

In practice ψ_2 usually does not exceed 0.1 and is generally much less. The other terms involved are at most of the order of the ratio of a few miles to the radius of the earth.

given by Eq. (372) with the proper sign to suit the geometry, that is,

$$\phi_1 = kS_1. \tag{417}$$

Now let us obtain an expression for $\phi_2 - \phi_1$. For brevity, only Case a will be considered. If we solve Eq. (371) for $\cos\chi$ and discard the higher-order terms, we obtain

$$\cos^2\chi = 1 + qH\frac{(Z_1 + Z_2)}{2} - \frac{q^2r^2}{16} - \frac{H^2}{r^2}(Z_1 - Z_2)^2. \tag{418}$$

Equations (417), (372), and (418) lead to

$$\frac{\phi_1}{k_0} = r + \frac{H^2}{2r}(Z_1 - Z_2)^2 + \frac{qH(Z_1 + Z_2)\,r}{4} - \frac{q^2r^3}{96}. \tag{419}$$

By applying Eq. (419) separately to the two portions of the reflected ray between the terminals and reflection point, we obtain

$$\frac{\phi_2}{k_0} = r + \frac{H^2}{2}\left(\frac{Z_1^2}{r_1} + \frac{Z_2^2}{r_2}\right) + \frac{qH}{4}(Z_1r_1 + Z_2r_2) - \frac{q^2}{96}(r_1^3 + r_2^3), \tag{420}$$

whence

$$\frac{\phi_2 - \phi_1}{k_0} = \frac{H^2}{2}\left[\frac{Z_1^2}{r_1} + \frac{Z_2^2}{r_2} - \frac{(Z_1 - Z_2)^2}{r}\right] - \frac{qH}{4}(Z_2r_1 + Z_1r_2) + \frac{q^2r_1r_2r}{32}. \tag{421}$$

If we define[1]

$$z_1' = HZ_1 - \frac{qr_1^2}{4} = z_1 - \frac{qr_1^2}{4} = z_1 - \frac{r_1^2}{2a_e}, \tag{422a}$$

$$z_2' = HZ_2 - \frac{qr_2^2}{4} = z_2 - \frac{qr_2^2}{4} = z_2 - \frac{r_2^2}{2a_e}, \tag{422b}$$

Eq. (421) becomes

$$\frac{\phi_2 - \phi_1}{k_0} = \frac{1}{2}\left[\frac{z_1'^2}{r_1} + \frac{z_2'^2}{r_2} - \frac{(z_1 - z_2)^2}{r}\right]. \tag{423}$$

From Eqs. (411) and (412), it can be shown that

$$\tan\psi_2 = \frac{Z_1H - \dfrac{qr_1^2}{4\cos^2\psi_2}}{r_1} \approx \frac{Z_2H - \dfrac{qr_2^2}{4\cos^2\psi_2}}{r_2}. \tag{424}$$

As $\cos\psi_2$ is approximately unity, Eq. (424) yields

$$\frac{z_1'}{r_1} \approx \frac{z_2'}{r_2} \tag{425}$$

[1] The quantities z_1' and z_2' are the terminal heights above a plane tangent at the reflection point to an earth of radius a_e. In the conventional treatments of the interference region problem they are derived by a geometrical argument.

Equations (425) and (413) then reduce Eq. (423) to

$$\phi_2 - \phi_1 = \frac{2z_1'z_2'}{r} k_0. \tag{426}$$

from which Eq. (364) follows.

2·12. The Diffraction Region.—In Sec. 2·11 there was outlined a means of evaluating the integral of Eq. (356) when the terminal points lie above the horizon. In this section it will be shown that when one terminal point is below the tangent ray drawn from the other, ψ is given, for frequencies of 100 Mc/sec or above, by

$$\psi \approx \frac{2\sqrt{\pi}}{L} e^{-i(k_0 r + \pi/4)} \frac{1}{\sqrt{X}} \sum_{m=1}^{\infty} e^{iA_m X} U_m(Z_1) U_m(Z_2) \tag{427}$$

or

$$F \approx 2\sqrt{\pi X} \left| \sum_{m=1}^{\infty} e^{iA_m X} U_m(Z_1) U_m(Z_2) \right|, \tag{427a}$$

where

$$U_m(Z) = i \frac{y_2(Z + A_m)}{y_2'(A_m)} = i \frac{h_2(Z + A_m)}{h_2'(A_m)}. \tag{428}$$

For horizontal polarization the A_m's may be replaced by ζ_m's, the zeros of $y_2(\zeta)$. Numerical values for the first few ζ_m's are listed in Sec. 2·9. The antenna pattern is neglected in Eq. (427*a*).

In Eq. (427) it is tacitly assumed that the transmitter is at such a height that the distance to the horizon is several miles. At this range and at frequencies exceeding 100 Mc/sec, the surface wave may be neglected and Eq. (427) represents the essential part of the field.

To obtain Eq. (427) from Eq. (356), the path of integration is deformed into an infinite semicircle lying in the upper half plane with its center at the origin; ψ is then expressed as the sum of three parts:

1. An integral along the branch cut associated with the branch point at $A = H^2(k_0^2 - k_1^2)$. Physically this represents the surface wave, which for the cases envisaged here can be neglected.
2. An integral over the infinite semicircle. This can be written by reference to Eq. (356) as

$$\frac{i}{4H} \lim_{\eta \to \infty} \int_C H_0^{(2)}(k_0 r \sqrt{1 - qHA}) \left[y_1(Z_2 + A) y_2(Z_1 + A) - \frac{y_1'(A) - Hpy_1(A)}{y_2'(A) - Hpy_2(A)} y_2(Z_1 + A) y_2(Z_2 + A) \right] dA,$$

where C is a semicircle of radius η lying in the upper half plane and centered on the origin. It can be shown by using the asymptotic representations of the functions together with Jordan's lemma that this integral is zero.

3. The sum of the residues at the poles in the upper half plane. Obviously, only the second term of the bracketed expression in the integrand displayed above has poles, and these are the roots of

$$\frac{y_2(A_m)}{y_2'(A_m)} = \frac{1}{Hp}. \tag{429}$$

The order of the A's, all of which are distinct, is to be such that

$$\mathrm{Im}(A_{m-1}) < \mathrm{Im}(A_m) < \mathrm{Im}(A_{m+1}).$$

By the usual procedure for determining the residues at the simple poles of a function, we obtain

$$\psi = (2\pi i)\frac{i}{4H}\sum_{m=1}^{\infty} H_0^{(2)}(k_0 r\sqrt{1 - qHA_m}) \times \frac{y_1'(A_m) - Hp(A_m)y_1(A_m)}{\left\{\frac{d}{dA}[y_2'(A) - Hp(A)y_2(A)]\right\}_{A=A_m}} y_2(Z_1 + A_m)y_2(Z_2 + A_m). \tag{430}$$

The terms of the sum in Eq. (430) can be written in a more convenient form. From Eqs. (334) and (429), it follows that

$$y_1'(A_m) - Hp(A_m) = \frac{-2iHp(A_m)}{y_2'(A_m)}, \tag{431}$$

whence

$$\psi = \frac{-i\pi}{H}\sum_{m=1}^{\infty} \frac{H_0^{(2)}(k_0 r\sqrt{1 - qHA_m})y_2(Z_1 + A_m)y_2(Z_2 + A_m)}{\frac{y_2'(A_m)}{Hp(A_m)}\left\{\frac{d}{dA}[y_2'(A) - Hp(A)y_2(A)]\right\}_{A=A_m}}. \tag{432}$$

As we shall see later, A_m is nearly equal to ζ_m, and hence for all values of m with which we are concerned, qHA_m is small compared with unity. Equation (357) may therefore be used with Eqs. (358) and (432) to obtain Eq. (427), in which, however, $U_m(Z)$ is given by

$$U_m(Z) = \frac{iy_2(Z + A_m)}{y_2'(A_m)\left\{1 + \frac{1}{H^2p^2(A_m)}[A_m + Hp'(A_m)]\right\}^{1/2}}, \tag{433}$$

and not by Eq. (428).

Now consider Eq. (429), from which we see that $A_m = \zeta_m$ when $1/p = 0$. It is reasonable, however, to assume as a first approximation

that A_m is still equal to ζ_m when p is large though finite. Equations (382) and (383) then show that p is nearly independent of A. By neglecting the variation of p with A it is possible to estimate the correction required to take into account the finiteness of p. Expansion of the left-hand member of Eq. (429) about the points ζ_m yields

$$(A_m - \zeta_m) - \frac{y_2''(\zeta_m)(A_m - \zeta_m)^2}{2y_2'(\zeta_m)} + \cdots = \frac{1}{Hp}. \tag{434}$$

But from Eq. (304), $y_2''(\zeta_m) = -y_2(\zeta_m) = 0$. Therefore, as a second approximation,

$$A_m = \zeta_m + \frac{1}{Hp}. \tag{435}$$

By neglecting qHA in comparison with unity in Eqs. (382) and (383), we obtain

$$\frac{1}{Hp_h} = \frac{-i\left(\frac{q}{k_0}\right)^{1/3}}{\sqrt{\left(\frac{k_1}{k_0}\right)^2 - 1}} \tag{436}$$

$$\frac{1}{Hp_v} = \frac{-i\left(\frac{q}{k_0}\right)^{1/3}\left(\frac{k_1}{k_0}\right)^2}{\sqrt{\left(\frac{k_1}{k_0}\right)^2 - 1}} = \left(\frac{k_1}{k_0}\right)^2 \frac{1}{Hp_h}. \tag{437}$$

Representative values are given in Sec. 5·1 for $(k_1/k_0)^2$, which is there called ε_c. Inspection of Table 5·1 shows that $(k_1/k_0)^2 - 1$ is smallest and consequently $1/Hp_h$ takes on its greatest value for dry land. Using $(k_1/k_0)^2 = 2 - i1.62$ the appropriate value for very dry sandy loam at 9 cm we find that $|\sqrt{(k_1/k_0)^2 - 1}| = 1.4$. For dry ground at 1 m, $(k_1/k_0)^2 = 4 - i0.006$, and $|\sqrt{(k_1/k_0)^2 - 1}| = 1.7$. As $(q/k_0)^{1/3}$ varies from 1.5×10^{-3} to 4.7×10^{-3} as λ varies from 10 cm to 3 m, we conclude that $|1/Hp_h|$ is of the order of 10^{-3} in the region below 3 m. From a practical standpoint $1/Hp_h$ is thus a negligible correction to ζ_m.[1]

The largest value of $1/Hp_v$ in the frequency range below 3 m is obtained for the constants appropriate to sea water at 3 m. Using $(k_1/k_0)^2 = 80 - i774$, and $\lambda = 3$ m in Eq. (437), we obtain $1/Hp_v \approx 0.13e^{-3\pi i/4}$; the corresponding value for horizontal polarization is $1/Hp_h \approx 1.8 \times 10^{-4}\, e^{-\pi i/4}$. The departure of A_m from ζ_m is negligible for horizontal polarization but not for vertical. In both cases, the correction term reduces the attenuation with range, the reduction in the case of vertical polarization being very much greater than in the case of hori-

[1] See also C. L. Pekeris, *Jour. Applied Phys.*, **19**, 102 (1948).

zontal. On this basis, the field for vertical polarization tends to be somewhat greater than for horizontal. To be sure that it actually is greater, we must examine the effect of the departure from the ideal boundary conditions on the height-gain functions.

From Eqs. (382) and (383), it follows that

$$p_v' H = -\frac{1}{2}\left(\frac{k_0}{k_1}\right)^4 \frac{1}{Hp_v}, \tag{438}$$

$$p_h' H = -\frac{1}{2}\frac{1}{Hp_h}. \tag{439}$$

Equations (438) and (439), together with the estimates already made of $|1/pH|$, show that the denominator of Eq. (433) is given to within 1 per cent for the first few modes by $y_2'(A_m)$. Thus Eq. (433) becomes Eq. (428). In addition, we have already justified using ζ_m for A_m when the polarization is horizontal.

By replacing A_m by $\zeta_m + (1/Hp)$ in Eq. (428) and expanding about ζ_m it is readily shown that for small Z

$$U_m(Z) \approx Z + \frac{1}{pH}.$$

It follows, therefore, from the examples given that at least near the surface the height-gain function is greater for vertical than for horizontal polarization. As it has already been shown that the attenuation with range acts in the same direction, theory indicates that at wavelengths of 3 m and below the field close to the surface in the diffraction zone is greater for vertical than for horizontal polarization. The difference between the polarizations decreases with increasing frequency.

METHODS FOR CALCULATING FIELD STRENGTH WITH STANDARD REFRACTION

BY WILLIAM T. FISHBACK

Methods of calculating field strengths for an essentially linear modified-index profile have been well established for some time.[1] Unfortunately they are not amenable to easy calculation, and computing field strengths involves considerable labor. Recently simpler methods have been evolved for both the interference and diffraction regions.[2]

[1] C. R. Burrows and M. C. Gray, "The Effect of the Earth's Curvature on Ground-wave Propagation," *Proc. IRE*, **29**, 16 (1941); K. A. Norton, "The Calculation of Ground-wave Field Intensity over a Finitely Conducting Spherical Earth," *Proc. IRE*, **29**, 623 (1941).

[2] W. T. Fishback, "Simplified Methods of Field Intensity Calculations in the Interference Region," RL Report No. 461, Dec. 8, 1943; P. J. Rubenstein and W. T. Fishback,

2·13. The Interference Region.—The determination of F for a point in space over a plane earth can be carried out in a perfectly straightforward manner using Eq. (36). The geometrical quantities required are the path difference ΔR and the grazing angle ψ_2 (to obtain the value of ρ for a given ψ_2); both of these can be calculated directly from the two given heights and the distance between the terminals. When calculations of field strength for a spherical rather than a plane earth are desired, we find that we must introduce a new geometrical quantity, the divergence factor, and that to compute ΔR and ψ_2 we must know the reflection point, or how r is broken up[1] into r_1 and r_2. It is possible to find r_1 and r_2 if r, z_1, and z_2 are given, but to do so we must solve the cubic equation,

$$2r_1^3 - 3rr_1^2 + [r^2 - 2a_e(z_1 + z_2)]r_1 + 2a_e z_1 r = 0.$$

This equation has the formal solution

$$r_1 = \frac{r}{2} + p\cos\left(\frac{\Phi + \pi}{3}\right),$$

where

$$p = \frac{2}{\sqrt{3}}\sqrt{a_e(z_1 + z_2) + \left(\frac{r}{2}\right)^2},$$

$$\Phi = \cos^{-1}\left[\frac{2a_e(z_2 - z_1)r}{p^3}\right],$$

and it is assumed that $z_1 \geqslant z_2$ and thus $r_1 \geqslant r_2$. The labor involved in determining r_1 and r_2 is excessive and can be avoided by using the technique described below.[2] The simplified method presented here consists of the use of a set of graphs with which ΔR, D, and ψ_2 can be determined from r, z_1, and z_2 alone. The process of determining r_1 and r_2 has been completely absorbed in the graphs.

From Eq. (426) we find

$$\Delta R = \frac{2z_1' z_2'}{r}. \tag{440}$$

If we let $\psi_2 \approx z_1'/r_1$, we can write Eq. (365) as

$$D = \left(1 + \frac{2r_1^2 r_2}{a_e r z_1'}\right)^{-1/2}. \tag{441}$$

"Graphs for Computing the Diffraction Field with Standard and Superstandard Refraction," RL Report No. 799, Aug. 13, 1945; C. Domb and M. H. L. Pryce, "Interim Report on Propagation within and beyond the Optical Range," ASE Report No. M448, September 1942; and *Jour. IEE*, **94**, Part III, 325 (1947).

[1] Figure 2·16 should be consulted for definitions of the quantities involved.

[2] A graphical solution of the cubic equation is also given by R. A. Hutner, H. W. Dodson, J. R. Gill, F. D. Parker, and B. E. Howard, "Field Intensity Formulas," RL Report No. 23, Sept. 28, 1943.

Finally, by using Eqs. (424) and (425), we find

$$\tan \psi_2 = \frac{z_1' + z_2'}{r}. \tag{442}$$

These expressions can be calculated more easily if we introduce the dimensionless height and distance parameters

$$S_{1,2} = \frac{r_{1,2}}{\sqrt{2a_e z_{1,2}}} \leqslant 1, \tag{443}$$

$$S = \frac{r}{\sqrt{2a_e z_1} + \sqrt{2a_e z_2}} \leqslant 1, \tag{444}$$

$$T = \sqrt{\frac{z_2}{z_1}} \leqslant 1. \tag{445}$$

The requirement that S_1, S_2, and S be less than unity is fulfilled automatically if we stay within the total horizon range $r_L = \sqrt{2a_e z_1} + \sqrt{2a_e z_2}$. As in Sec. 2·10 advantage is taken of the reciprocity theorem, and z_2 is defined as the smaller of the two terminal heights in the remainder of this section in order to keep T less than unity.

By means of these new parameters Eqs. (440) to (442) can be rewritten in the form

$$\Delta R = \frac{2z_1 z_2}{r}(1 - S_1^2)(1 - S_2^2), \tag{446}$$

$$\tan \psi_2 = \frac{z_1 + z_2}{r}\left[\frac{(1 - S_1^2) + T^2(1 - S_2^2)}{1 + T^2}\right], \tag{447}$$

$$D = \left[1 + \frac{4S_1^2 S_2 T}{S(1 - S_1^2)(1 + T)}\right]^{-\frac{1}{2}}. \tag{448}$$

From Eq. (448) it follows that the divergence factor can be expressed as a function of the family of S and T variables alone. The expressions for the path difference and grazing angle have been reduced to the product of two factors, of which one is just the expression for the desired quantity under flat-earth conditions and the other is a function of the S and T family of variables alone. The part of the expressions for ΔR and $\tan \psi_2$ that is a function only of the S and T variables is thus a correction factor that converts the plane-earth expressions for ΔR and $\tan \psi_2$ into the correct values for a spherical earth with a linear N-profile. For any point in space the plane-earth quantities may be easily calculated and the parameters S and T determined; the quantities S_1 and S_2 cannot be determined directly, but it will be shown that they are functions of S and T alone. We know that

$$\frac{z_1'}{r_1} = \frac{z_2'}{r_2}$$

and

$$r = r_1 + r_2.$$

These expressions lead to the equations

$$S_2(1 - S_1^2) = S_1T(1 - S_2^2),$$

$$S(1 + T) = S_1 + S_2T,$$

which can be solved for S_1 and S_2 in terms of S and T.[1] Thus it has been shown that any quantity which is a function of S_1, S_2, S, and T is actually a function of S and T alone. Let

$$J(S,T) = (1 - S_1^2)(1 - S_2^2)$$

and

$$K(S,T) = \frac{(1 - S_1^2) + T^2(1 - S_2^2)}{1 + T^2},$$

so that Eqs. (446) to (448) become

$$\Delta R = \frac{2z_1z_2}{r} J(S,T), \tag{449}$$

$$\tan \psi_2 = \frac{z_1 + z_2}{r} K(S,T), \tag{450}$$

$$D = D(S,T). \tag{451}$$

Contours of constant J, K, and D are plotted as functions of S and T in Figs. *A* to *D*, included in the envelope in the rear cover of this book.

We can now determine the path difference, grazing angle, and divergence factor for any point in space, given z_1, z_2, r, and a_e. In practice we may use any positive value of a_e, that is, any positive slope for a linear M-profile, but this model is usually of practical importance only for the standard case,[2] where

$$a_e = 8.5 \times 10^6 \text{ m} = 5{,}280 \text{ statute miles}.$$

Having determined ψ_2, we can make use of theoretical curves of ρ and ϕ vs. ψ_2 to determine these quantities. Figures 5·4 to 5·6 are graphs of ρ and ϕ for smooth sea water for several wavelengths. For a given wavelength and isotropic radiator we can now find F from

$$F = |1 + D\rho e^{-i\alpha}| = |1 + D\rho e^{-i(k\,\Delta R + \phi)}|. \tag{452}$$

[1] In the construction of the graphs of J, K, and D, values of S_1 and T were assumed and used to determine values of S_2 and S.

[2] In this case the total horizon range conveniently becomes

$$r_L = \sqrt{2z_1} + \sqrt{2z_2},$$

where r_L is in statute miles and z_1 and z_2 are in feet.

If we know the antenna pattern function $f(\theta)$ and the angle of tilt ξ of the beam maximum, we can find F from the general expression

$$F = |f(\theta_1) + f(\theta_2)D\rho e^{-i(k\,\Delta R+\phi)}|, \tag{453}$$

where θ_1 and θ_2 are the angles between the antenna beam maximum and the direct and reflected rays, respectively, given by

$$\theta_1 = \psi_1 - \xi,$$
$$\theta_2 = -\psi_2 - \eta - \xi.$$

The various angles are indicated in Fig. 2·16. The angles θ_1 and θ_2 are positive if measured upward from the beam maximum and negative if measured downward. Similarly γ, ψ_1, and ξ are positive if measured

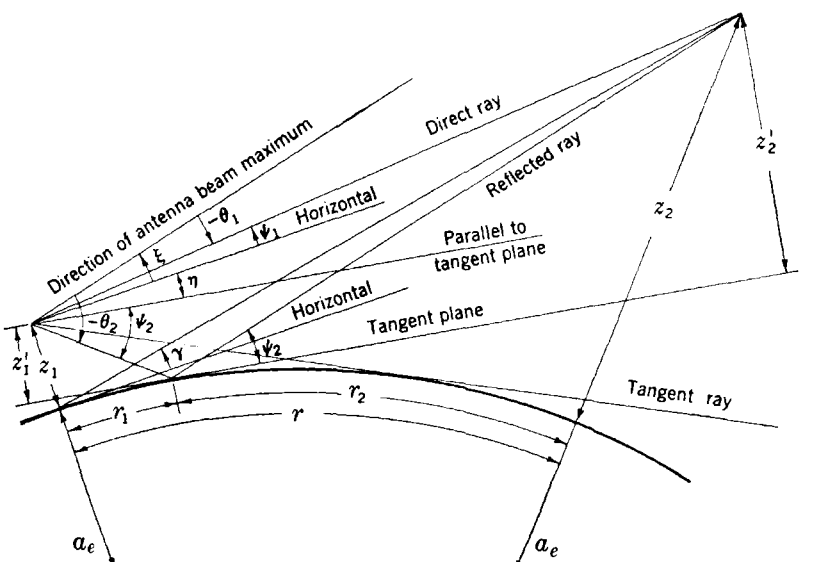


FIG. 2·16.—Geometrical parameters for standard propagation in the interference region. See the text for explanation of the conventions for signs of the angles.

upward from the horizontal at the transmitter and negative if measured downward. As they can never change sign, ψ_2 and η will always be considered positive. In most practical cases,

$$\psi_1 \approx \psi_2,$$

and

$$\eta = \frac{r_1}{a_e} \ll \psi_2,$$

so that we can rewrite Eq. (453) in the simplified form of

$$F = |f(\psi_2 - \xi) + f(-\psi_2 - \xi)D\rho e^{-i(k\,\Delta R+\phi)}|. \tag{454}$$

This expression for the pattern-propagation factor is not valid all the way to the total horizon range for reasons discussed in Sec. 2·11. No exact analytic expression is known for the limit of validity of Eq. (454), but it is believed that it will always be valid for

$$\Delta R \geqslant \frac{\lambda}{4}.$$

For very low antenna heights at microwaves or for moderately low antenna heights at longer wavelengths, ΔR declines to $\lambda/4$ at only a small fraction of the total horizon range.[1] In this case Eq. (454) will remain valid well beyond this point and may on occasion be used until the phase $k\,\Delta R$ is less than 1°. If there is any doubt about the validity of the result obtained using Eq. (454), it is recommended that the result be checked by the methods described below for the intermediate region.

As an example let us calculate F for a 10-cm radar system sited at a height of 100 ft and viewing a target at a range of 80 statute miles and height of 11,000 ft over sea water. The set uses vertical polarization and a paraboloidal antenna with a symmetrical beam 3° in width between half-power points[2] tilted up 0.5°. We have

$$\begin{aligned} r_L &= \sqrt{200} + \sqrt{22{,}000} \\ &= 162 \text{ statute miles,} \end{aligned}$$

so that

$$S = \frac{80}{162} = 0.494$$

and

$$T = \sqrt{\frac{100}{11{,}000}} = 0.095.$$

Now from Figs. A, B, and D we find that

$$\begin{aligned} J &= 0.713, \\ K &= 0.72, \\ D &= 0.990. \end{aligned}$$

These values of J and K can now be substituted along with the given z_1, z_2, and r in Eqs. (449) and (450) to get

$$\Delta R = \frac{2 \times 11{,}000 \times 100}{5280 \times 80} \times 0.713 = 3.72 \text{ ft}$$

[1] In the atmosphere the index of refraction n is slightly different from unity and has different values at different points in space. As a result, wavelength and wave number will vary throughout space and will nowhere be equal to the free-space value. The variation is so slight, however, that it may be ignored in the types of calculations being discussed here. Throughout this section subscripts will be discarded and the free-space wavelength and wave number referred to as λ and k, respectively.

[2] In all cases the total width between the angles at which $f^2 = \frac{1}{2}$ will be meant.

and

$$\tan \psi_2 = \frac{11{,}100}{5280 \times 80} \times 0.72 = 0.0189,$$

or

$$\psi_2 = 1.04°.$$

The values of ρ and ϕ for vertical polarization and $\lambda = 10$ cm are, from Figs. 5·4 and 5·5,

$$\rho = 0.722,$$
$$\phi = 175°.$$

Finally, we find

$$\cos [-(k \Delta R + \phi)] = \cos (-4082° - 175°) = 0.454.$$

We must now consider the antenna-pattern effects. We know that $\xi = 0.5°$ and $\psi_2 = 1°$, and we must determine $f(0.5°)$ and $f(-1.5°)$ for the direct and reflected rays, respectively. For a paraboloidal antenna the form

$$f(\theta) = 1 - 0.293 \left(\frac{2\theta}{\theta_0}\right)^2, \tag{455}$$

where θ_0 is the beamwidth between half-power points, is a good approximation for $2\theta < \theta_0$. From this expression we find that

$$f(0.5) = 0.967,$$
$$f(-1.5) = 0.707.$$

We know all the parameters involved in Eq. (454) and can substitute to get

$$F = 1.28.$$

It is more convenient to express this in the logarithmic form

$$20 \log_{10} F = +2 \text{ db}.$$

This indicates that the field strength incident on the target is 2 db greater than it would be under free-space conditions with the antenna pointing directly at the target.

So far we have been concerned with finding the field strength at an arbitrary point in space. It is often desirable to find the point in space at which a given value of α occurs. We can solve this type of problem by using the S- and T-parameters provided $\phi \approx \pi$. This is always true for horizontal polarization and is a good approximation on vertical polarization for microwaves at small grazing angles. If this restriction is made, it is possible to find the height at which a given maximum or minimum occurs, given the transmitter height and range, or to find the distance at which a given maximum or minimum occurs, given the termi-

nal heights. Because of its great utility in the construction of coverage diagrams, the latter procedure will be outlined here.[1]

If $\phi = \pi$, the phase between the direct and reflected waves is

$$\alpha = \frac{2\pi \Delta R}{\lambda} + \pi.$$

The first maximum will occur when $\alpha = 2\pi$, and the nth maximum will occur when $\alpha = 2n\pi$, so that

$$2n\pi = \frac{2\pi n \Delta R}{\lambda} + \pi,$$

or

$$\Delta R = \frac{(2n - 1)\lambda}{2}.$$

Similarly, the condition for the nth minimum is

$$\Delta R = n\lambda.$$

This technique can be extended for any arbitrary value of the phase difference α. It has been shown that when $\phi = \pi$, the problem of finding a given point in the interference pattern is equivalent to finding the point at which a specific value of ΔR occurs. Because of this equivalence a direct solution of this type of problem is possible, using the S- and T-parameters. Let us replace r by $r_L S$ in Eq. (449) and rewrite the equation as

$$Q(S,T) = \frac{r_L \Delta R}{2z_1 z_2}, \tag{456}$$

where

$$Q(S,T) = \frac{(1 - S_1^2)(1 - S_2^2)}{S}. \tag{457}$$

We know the terminal heights and consequently the total horizon range; we also know the value of ΔR corresponding to the desired value of α; hence, Q can be determined. Knowing Q and T, we can determine S from the relation implicit in Eq. (457). Figures E to G, in the envelope in the rear cover, give contours of constant S as a function of Q and T,[2]

[1] The former problem can be solved in an analogous manner by replacing S by S', where $S' = r/\sqrt{2a_e z_1}$, and writing Eq. (449) in the form

$$\frac{r \Delta R}{2z_1^2} = \frac{J(S',T)}{T^2}.$$

A statement of the technique employed and the graphs necessary for its use may be found in W. T. Fishback, "Simplified Methods of Field Intensity Calculations in the Interference Region," RL Report No. 461, Dec. 8, 1943.

[2] For values of Q larger than those shown on Fig. G, the desired distance may be found with sufficient accuracy by using the flat-earth formula,

$$r = \frac{2z_1 z_2}{\Delta R}.$$

which may be used to find the S corresponding to the given Q and T. The unknown distance can then be determined from the total horizon range and the value of S previously determined.

As an example, let us consider the case of a horizontally polarized 3-cm radar set sited at an elevation of 20 ft and tracking an airplane flying over water at a height of 1500 ft. We want to find the range at which the plane will be in the maximum of the lowest lobe. For a wavelength of 3 cm and horizontal polarization, $\phi = \pi$, and we can use this method. We know that

$$z_2 = 20 \text{ ft},$$
$$z_1 = 1500 \text{ ft},$$

so that

$$r_L = \sqrt{40} + \sqrt{3000} = 61.0 \text{ statute miles},$$

$$T = \sqrt{\frac{20}{1500}} = 0.115.$$

To be at the first maximum, we must have

$$\Delta R = \frac{\lambda}{2} = 0.0492 \text{ ft}.$$

We find

$$Q = \frac{5280 \times 61.0 \times 0.0492}{2 \times 20 \times 1500} = 0.264.$$

From Fig. F we find that for $Q = 0.264$ and $T = 0.115$, we have $S = 0.817$. Hence,

$$r = r_L S = 61.0 \times 0.817 = 49.8 \text{ statute miles},$$

which is the desired range.

Although the examples have been presented using physical units, the methods previously developed are directly applicable to problems with either physical or natural units of height and distance. When only "spot" calculations for the interference region, such as the previous examples, are desired, there is no need to convert to natural units. It is often necessary, however, to make plots of field strength that include both the interference and the diffraction region, particularly in the determination of coverage contours. The calculation of field strengths in the diffraction region requires the use of natural units, and in such cases it is often preferable to do all work, including the plotting of the coverage contours, in natural units. The conversion factors H and L for changing from physical to natural units are plotted in Figs. 2·17 and 2·18 as a function of wavelength. They have been drawn for standard refraction, for which $a_e = 5280$ statute miles.

In terms of natural units, the fundamental relations expressed in Eqs. (444), (445), (449), (450), and (457) can be written as

$$S = \frac{X}{X_L} = \frac{X}{\sqrt{Z_1} + \sqrt{Z_2}} \leqq 1, \tag{458}$$

$$T = \sqrt{\frac{Z_2}{Z_1}} \leqq 1, \tag{459}$$

$$k\,\Delta R = \frac{Z_1 Z_2}{X} J(S,T) \tag{460}$$

$$\tan \psi_2 = \frac{1}{2}\left(\frac{\eta}{k}\right)^{1/3} \frac{Z_1 + Z_2}{X} K(S,T), \tag{461}$$

and

$$Q(S,T) = \frac{X_L k\,\Delta R}{Z_1 Z_2}, \tag{462}$$

where

$$k\,\Delta R = (2n - 1)\pi$$

at the nth maximum and

$$k\,\Delta R = 2n\pi$$

at the nth minimum.

The problem of where it is necessary to use the spherical-earth formulas for interference region calculations and where it is possible to use the flat-earth formulas has not at present been answered precisely. Subject to certain limitations to be discussed presently, a rough criterion can be developed with the aid of Fig. *A* in the back of the book. The most critical difference between the spherical and the flat-earth formulas is probably contained in the expressions for the path difference ΔR, for a change in the value of ΔR will have a profound effect upon the value of F. Figure *A* can be used essentially as a percentage error graph in the value of ΔR obtained from the flat-earth formula. Fortunately the values of the flat-earth correction factor J are only slightly dependent upon T and can be considered a function of S alone to a first approximation; thus the requirement that the value of ΔR calculated from the flat-earth formula differ from the spherical-earth quantity by less than a specified percentage can be fulfilled by requiring that S remain less than some value determined from Fig. *A*. This is equivalent to saying that ΔR can be calculated to a given accuracy provided one stays within a certain fraction of the total horizon range.

This procedure should prove helpful in determining the regions of validity of the flat-earth formula for F, particularly in the microwave region; the resulting limit of validity will probably be conservative. In many cases the flat-earth formulas can be used with impunity over a

much greater range of distances than would appear possible from an inspection of Fig. A. The extent to which this is possible is a function of the terminal heights in natural units. As the values of Z_1 and Z_2 are decreased, the flat-earth formulas will give accurate values of F throughout an increasing proportion of the interference region; indeed, for sufficiently small values of Z the flat-earth formulas will be reliable out to the total horizon range.[1] Calculations made with both Z_1 and Z_2 equal to 0.01 have indicated that the flat-earth formulas gave reliable values of F for a short distance beyond the total horizon range into the diffraction region. Such low values of the terminal heights in natural units will not often be encountered in the microwave region but will often be present for somewhat longer wavelengths. As Γ is often very closely -1 for horizontal polarization on these longer wavelengths, the familiar flat-earth approximations for small values of $4\pi z_1 z_2/\lambda R$ will be valid. This will often not be true, however, on microwaves, where the terminal heights in natural units are greater for a given physical height. In this case it will usually be necessary to use the spherical-earth formulas outlined in this section.

2·14. The Diffraction Region.—At points sufficiently far below the tangent ray the diffraction field can be represented by one mode alone, and the expression for F is given by Eq. (427a). In the case of a standard N-profile, $C_1 = 2.02$, and Eq. (427a) becomes, for a dipole antenna with its direction of maximum transmission in the direction of the horizon,

$$F = 2\sqrt{\pi X}\, e^{-2.02X} |U_1(Z_1) U_1(Z_2)|.$$

For convenience we shall let

$$V_1(X) = 2\sqrt{\pi X}\, e^{-2.02X},$$

so that

$$F = V_1(X) |U_1(Z_1) U_1(Z_2)|. \tag{463}$$

In the regions where Eq. (463) is valid, it is very simple to determine F. Figures 2·17 and 2·18 are plots of H and L as functions of wavelength for standard refraction and may be used to convert the terminal heights and distance into natural units. Once these quantities are known in natural units, Figs. 2·19 and 2·20 can be used to find F or 20 $\log_{10} F$. Figure 2·19 is a plot of the attenuation function, 20 $\log_{10}$ $[V_1(X)]$, as a function of X, the range in natural units, for the first mode with standard refraction. This figure also gives 20 $\log_{10} [V_1(X)/X]$ as a function of X and should be used when it is desired to find 20 $\log_{10} (F/X)$ rather than 20 $\log_{10} F$ (as is usually the case in drawing coverage diagrams). Figure 2·20 is a plot of the corresponding height-gain function 20 $\log_{10} |U_1(Z)|$ as a function of Z. It was computed by means of a series solution and the exact solution in terms of the tabulated values of $h_1(x)$ and $h_2(x)$ for small values of Z

[1] See footnote, Sec. 2·15, p. 125.

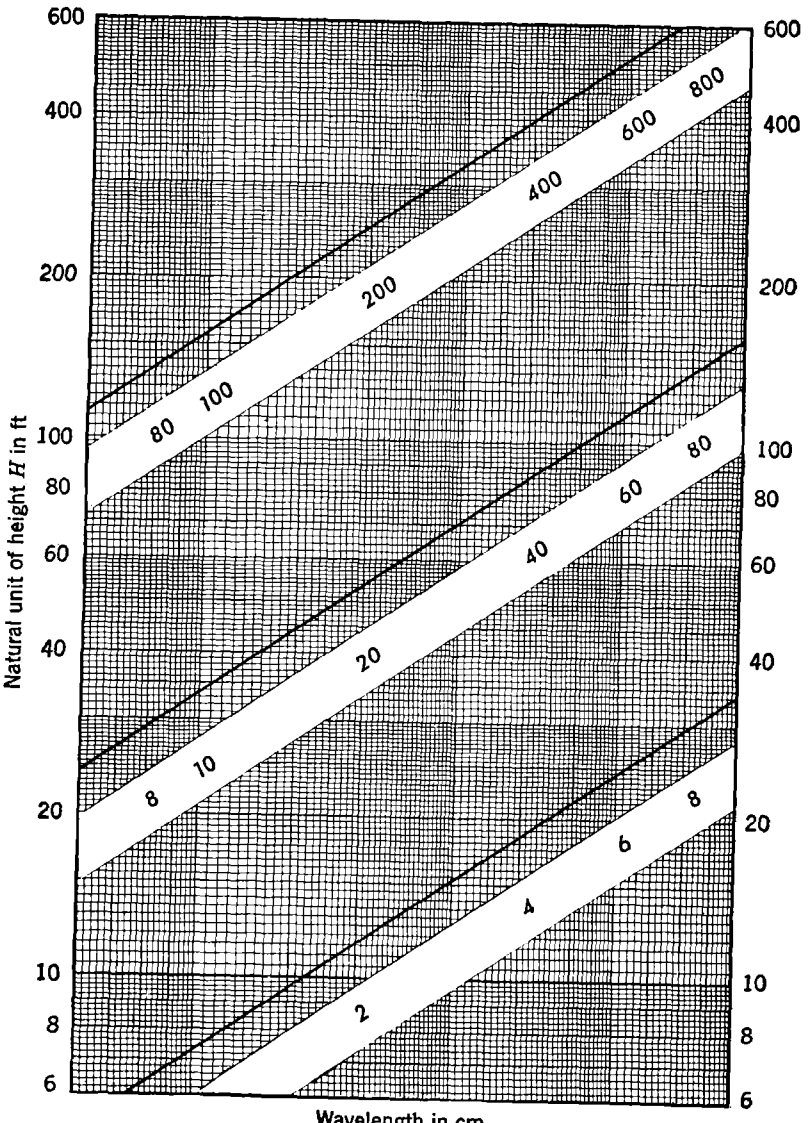


FIG. 2·17.—Natural unit of height H as a function of wavelength for $a_e = 5280$ statute miles.

and by means of the asymptotic expansions for these functions for large values of Z (see Sec. 2·9). These graphs have been computed for the simplified boundary condition $\Gamma = -1$, strictly satisfied only for horizontal polarization and a perfectly conducting earth but generally applicable at wavelengths below 3 m (see Secs. 2·12 and 5·1).

Calculations in the diffraction region with the aid of these graphs may be made very rapidly; for example, suppose we wish to find how far

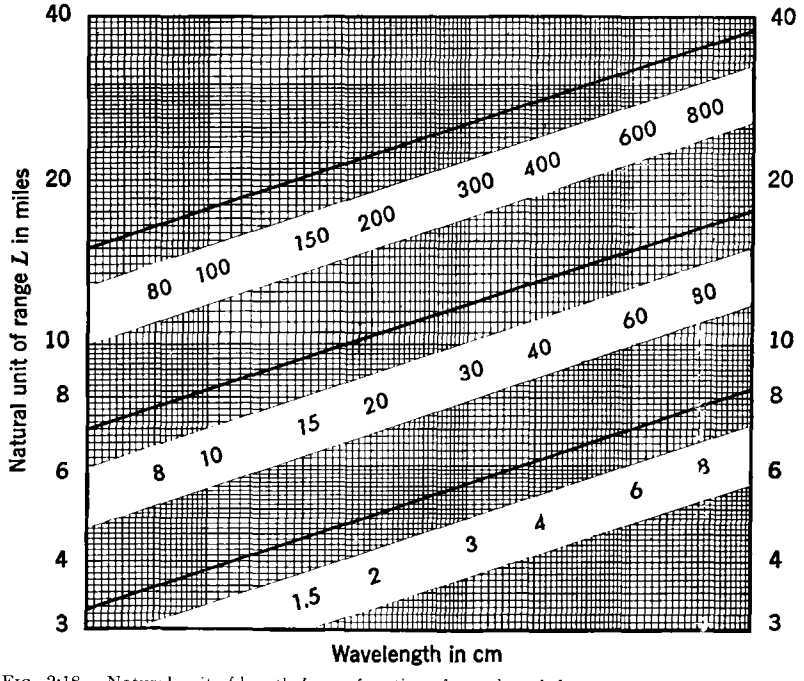


FIG. 2·18.—Natural unit of length L as a function of wavelength for $a_e = 5280$ statute miles.

below the free-space level the signal will be on a 10-cm, 80-statute-mile link between terminals at 200 and 300 ft. From Figs. 2·17 and 2·18, we see that for 10 cm

$$H = 33.3 \text{ ft},$$
$$L = 8.35 \text{ statute miles}.$$

In terms of natural units, the geometry of this path is given by

$$Z_1 = \frac{200}{33.3} = 6.0,$$

$$Z_2 = \frac{300}{33.3} = 9.0,$$

$$X = \frac{80}{8.35} = 9.6.$$

Finally we find from Fig. 2·19

$$20 \log_{10}[V_1(9.6)] = -148$$

and from Fig. 2·20

$$20 \log_{10}|U_1(6.0)| = +27,$$
$$20 \log_{10}|U_1(9.0)| = +37,$$

and thus

$$20 \log_{10} F = -148 + 27 + 37 = -84 \text{ db}.$$

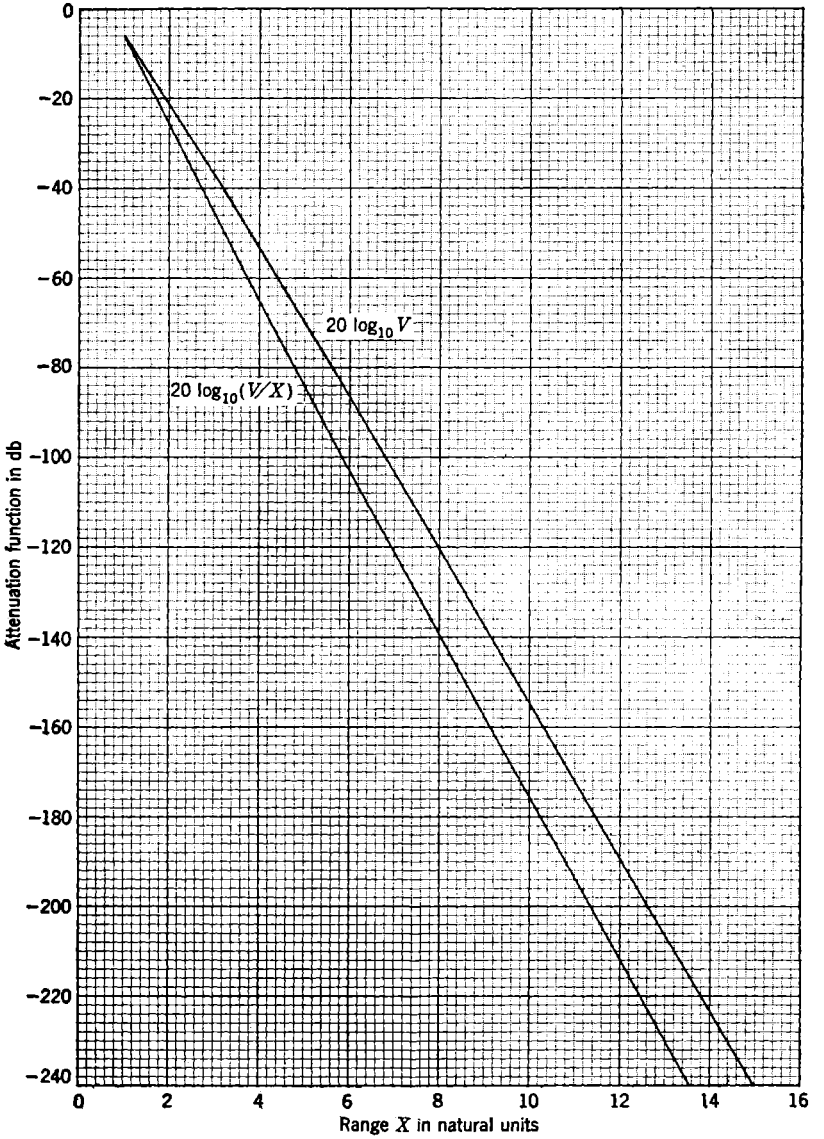
At the total horizon range and for a slight distance beyond it, one mode is not sufficient to obtain a correct value of F. No simple condition is known for finding the distance beyond which the 1-mode representation is valid but for standard conditions the total horizon range needs to be increased by only a small fraction of itself to make the 1-mode representation valid. Care should be taken in determining field strengths near the horizon range, and it is suggested that in such cases the calculations be checked by the method for finding field strengths in the intermediate region, to be discussed in the following section.

2·15. The Intermediate Region.—We have seen that the ray-tracing methods by which the field strength is determined in the interference region break down before the total horizon range is reached. In a like manner, we must start the use of a 1-mode expression for the diffraction field not at but beyond the total horizon range. As a result, there is a region known as the intermediate region near the horizon for which no generally applicable method of determining field strength is yet available,[1]

[1] A method of finding the field at the total horizon range X_L is outlined by C. Domb and M. H. L. Pryce, "Interim Report on Propagation within and beyond the Optical Range," ASE Report No. M448, Sept. 1942, and *Jour. IEE*, **94**, Part III, 325 (1947). For large values of Z_1 and Z_2 it was found that at this particular range the first few terms of the series $\sum_m e^{iA_m X_L} U_m(Z_1) U_m(Z_2)$ were closely in geometric progression and rapidly decreasing. The ratio of the first two terms was taken as the ratio of a geometric series and was used in conjunction with the first term to sum the series. For very small heights ($Z < 0.05$), it was possible to find the field at the total horizon range by using flat-earth ray theory. The field for intermediate values of Z was found by interpolation. The resulting curves have been used frequently at the Radiation Laboratory as an aid to drawing field-strength curves in this region and have proved useful, although the possible error seems to be slightly greater than the 1-db maximum error claimed for the method.

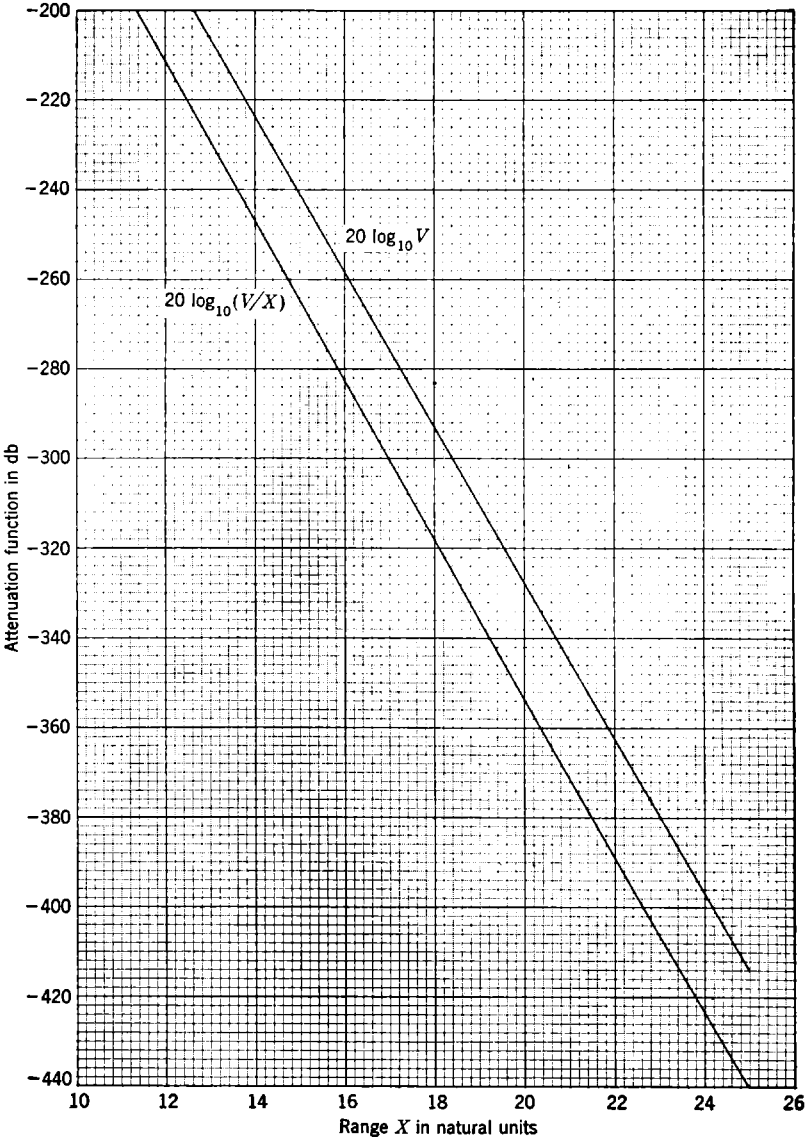
Burrows and Gray [C. R. Burrows and M. C. Gray, "The Effect of the Earth's Curvature on Ground-wave Propagation," *Proc. IRE*, **29**, 16, (1941)] have extended the diffraction formula so that it can be used at somewhat shorter ranges than can the 1-mode solution for the special case where

$$Z_1 > 20,$$
$$Z_2 < 0.5.$$

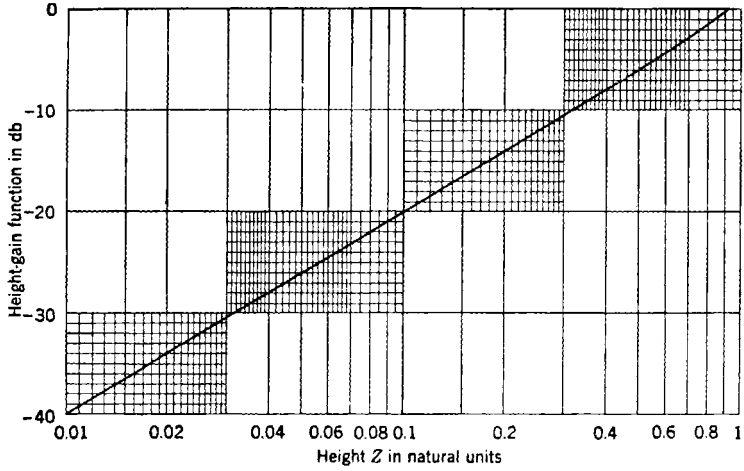


(*a*)

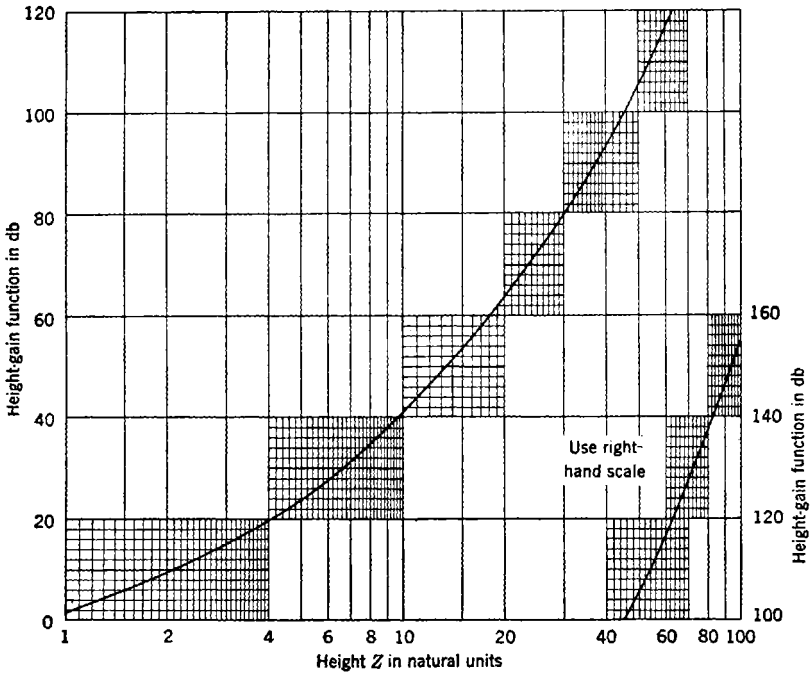
FIG. 2·19.—Attenutation functions for the first mode, $20 \log_{10} V_1$ and $20 \log_{10} (V_1/X)$, as a function of distance in natural units for (*a*) $X \leqslant 15$ and (*b*) $X \geqslant 10$. (Standard refraction.)


−200
−220
−240
−260
−280
−300
−320
−340
−360
−380
−400
−420
−440
Attenuation function in db
20 log10 V
20 log10 (V/X)
10
12
14
16
18
20
22
24
26
Range X in natural units

(*b*)



(*a*)



(*b*)

FIG. 2·20.—Height-gain function for the first mode, $20 \log_{10} |U_1(Z)|$, as a function of height in natural units for (*a*) $Z \leqslant 1$ and (*b*) $Z \geqslant 1$.

aside from summation of the series of Eq. (427). To determine the field strength in this region, it is in general necessary to prepare a plot of field strength as a function of distance for the given terminal heights, extending the plot from the region in which interference region methods are valid through the intermediate region into the region of validity for diffraction methods by bold interpolation. The value of the field strength at the desired distance can then be read on the plot. Because the interference-region points needed to make such a graph will occur at very small grazing angles, it is usually true that $\phi = \pi$, hence Figs. E to G can be used. In most cases, the best procedure is to find the distance at which the first maximum[1] ($\Delta R = \lambda/2$ and $F = 1 + D\rho$) and the first point of quadrature ($\Delta R = \lambda/4$ and $F = \sqrt{1 + D^2\rho^2}$) occur. Once these distances are known, D, ρ, and consequently F can be determined. Next, about three ranges at a sufficient distance in the diffraction region should be chosen, and the value of F for these ranges and the given terminal heights should be determined. The points thus determined for the interference and diffraction regions should be plotted and connected with a smooth curve. The value of F for the desired range in the intermediate region may then be determined from the curve. The procedure is shown

---

In this case an asymptotic approximation to $h_2(Z_1 + \zeta_m)$ can be used in conjunction with the relation

$$\frac{h_2(Z_2 + \zeta_m)}{h_2'(\zeta_m)} \approx Z_2,$$

so that

$$F = 2^{\frac{3}{2}} X^{\frac{1}{2}} Z_1^{-\frac{1}{4}} Z_2 F_L,$$

where

$$F_L = 2^{\frac{3}{2}} 3^{\frac{1}{6}} \sum_m \frac{e^{it\zeta_m (X - Z_1)^{\frac{1}{2}}}}{h_2'(\zeta_m)}.$$

The function F_L is plotted as a function of $L = 2^{\frac{1}{3}}(X - Z_1^{\frac{1}{2}})$ in Fig. 10 of this paper (taking $\delta \ll 1$ for a perfectly conducting earth). The technique described is valid as long as the series for F_L converges rapidly enough so that for all modes used

$$|\zeta_m| \ll Z_1$$

and

$$\arg (Z_1 + \zeta_m) < \frac{\pi}{3}.$$

It will not be valid at the horizon ($X = Z_1^{\frac{1}{2}}$), for $\sum_m [h_2'(\zeta_m)]^{-1}$ converges extremely slowly.

[1] In calculations for the intermediate and diffraction regions, the effective radiation all comes from the same part of the beam, and the expressions given above have neglected the antenna pattern. Thus, when the beam maximum is pointed horizontally, the expression for F is as shown. When the beam is elevated so that $|f(-\xi)| < 1$, the transmission equations (27) and (28) require the additional factors on their right-hand sides of $|f_t(-\xi_t) f_r(-\xi_r)|$ and $|f(-\xi)|^2$, respectively.

schematically in Fig. 2.21. Point A is the first maximum; Point B is the first point of quadrature; and Points C, D, and E were calculated for the diffraction region using one mode. This process is frequently used in determining coverage contours and will be illustrated in detail in the following section.

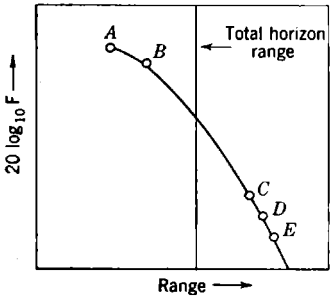


FIG. 2·21.—Example of technique for finding 20 $\log_{10} F$ in the intermediate region.

2·16. Contours of Constant Field Strength.—Heretofore we have been concerned with methods for finding the field strength at any point in space, either above or below the tangent ray. In many cases the desired information is not the field strength at a given point but rather contours of constant field-strength values —in short, a contour or a coverage diagram. In general these loci cannot be determined directly; the methods previously discussed must be employed in what is essentially a graphical solution for the desired contour. This method of determining coverage can be used anywhere in space and will give a complete picture of the detailed lobe structure. As the fine detail is often of no interest, one frequently uses this method to obtain the bottom of the first lobe, whereas above the maximum of the first lobe only the locus of maxima of the lobes is plotted. This is called a *coarse coverage diagram*. Both of these techniques will be discussed in detail.

The fundamental equation upon which coverage calculations are based is (see Secs. 2·1 and 2·2),

$$R = R_f F, \tag{464}$$

where

$$R_f = \frac{G\lambda}{4\pi}\sqrt{\frac{P_t}{P_r}} \tag{465}$$

for one-way transmission, and

$$R_f = \sqrt{\frac{G\lambda}{4\pi}}\sqrt[4]{\frac{P_t}{P_r}\frac{\sigma}{4\pi}} \tag{466}$$

for radar. The quantity R_f is the range in free space at which the received power is P_r with antennas aligned for maximum transmission; if P_r is the minimum detectable power of a system, R_f is just R_0, the free-space detection range. Equation (464) then becomes $R_{\max} = R_0 F$, where $R_{\max}$ is the maximum detection range at the given angle of elevation ψ_1. The quantity R_0 is useful when discussing a complete system, but R_f has greater generality and will be used in discussing field-strength contours (see Secs. 2·1 and 2·2).

Equations (464) to (466) may be related to the field strength in the following way. In free space,

$$\left|\frac{E_0}{E_s}\right| = \left|\frac{R_s}{R}\right|,$$

where E_0 and E_s are field strengths at ranges of R and R_s, respectively, in the direction of maximum transmission and R_s is a reference unit of length. Under actual physical conditions the field strength E is related to E_0 through the defining equation for the pattern-propagation factor $F = |E/E_0|$. Making this substitution in the free-space equation above yields

$$\left|\frac{E}{E_s}\right| = F\frac{R_s}{R}. \tag{467}$$

Combining Eq. (467) with Eq. (464) gives the additional relation

$$R_f = R_s\left|\frac{E_s}{E}\right|, \tag{468}$$

which, in conjunction with Eqs. (465) and (466), gives the connection between transmitted and received power, system parameters, and field strength. A contour of constant field strength, chosen by fixing either $|E/E_s|$ or the quantities in Eq. (466), determines R_f, and the contour is found by solving Eq. (464) or its equivalent [Eq. (467)]. Thus

$$20\log_{10}\left|\frac{E}{E_s}\right| = -20\log_{10}\frac{R_f}{R_s} = 20\log_{10}\frac{FR_s}{R}, \tag{469}$$

from which it can be seen that the field is given directly in decibels relative to the maximum free-space field at reference distance from the source. The term on the far right is the sum of the free-space attenuation[1] $20\log_{10}(R_s/R)$ and the attenuation relative to the free-space field $20\log_{10}F$. This method of expression allows separation of the system parameters from propagation phenomena to the greatest possible extent, the only implicit relation being the role of the antenna pattern function in F.

The preceding sections have treated methods of finding F that permit direct calculation of $20\log_{10}(FR_s/R)$. Points on a contour for a chosen

[1] In the interference region the choice of R_s is largely a matter of preference. For our present purposes, $R_s = 1$ m. In the diffraction region it is most convenient to use $R_s = L$, the natural unit of length. Then $20\log_{10}(FR_s/R)$ gives the field in decibels above the field at 1 m or at a distance L from the antenna. When $R_s = L$, Eq. (469) can be written in the useful form

$$20\log_{10}\left|\frac{E}{E_s}\right| = -20\log_{10}X_f = 20\log_{10}\frac{F}{X}.$$

It is convenient to remember that

$$20\log_{10}(R_{\text{meters}}) = 20\log_{10}(R_{\text{statute miles}}) + 64.13.$$

value of $|E/E_s|$ are obtained by finding values of R that satisfy Eq. (467) or (469). These values of R will be denoted by R_c; the locus of all values of R_c forms the desired contour. If $|E/E_s|$ in Eqs. (467) and (469) is fixed, thus determining the contour to be investigated, the values of R must be found that satisfy these equations. These values, denoted by R_c or X_c, determine the form of the contour.

For high-performance communications systems the desired contour may penetrate well into the diffraction region; hence for such systems the

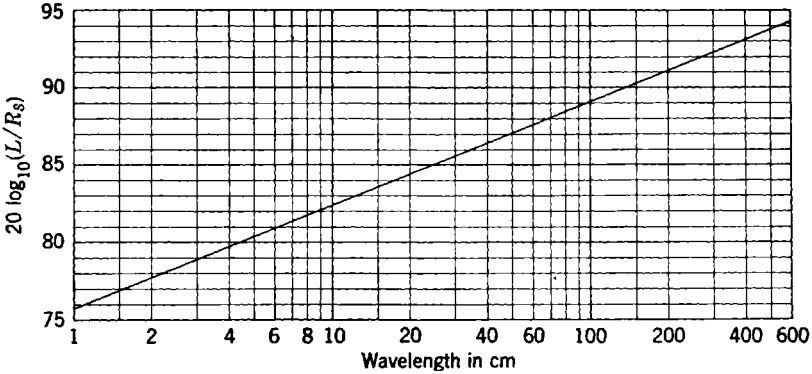


FIG. 2·22.—The function $20 \log_{10}(L/R_s)$ as a function of wavelength, where R_s is 1 m.

difficulties peculiar to the intermediate region are not encountered. When this is true, the desired coverage contour may be obtained by a direct application of the relation

$$20 \log_{10} \frac{F}{X_c} = 20 \log_{10} \frac{V_1}{X_c} + 20 \log_{10} |U_1(Z_1)| + 20 \log_{10} |U_1(Z_2)|. \quad (470)$$

From $X = R/L$ and knowledge of $20 \log_{10} (R_f/R_s)$, $20 \log_{10} (FR_s/R_c)$ is known, and we can find $20 \log_{10} (F/X_c)$ from

$$20 \log_{10} \frac{F}{X_c} = 20 \log_{10} \frac{FR_s}{R_c} + 20 \log_{10} \frac{L}{R_s}. \quad (471)$$

A plot of $20 \log_{10} (L/R_s)$ as a function of wavelength is given in Fig. 2·22. To find X_c one needs only to determine $20 \log_{10} |U_1(Z_1)|$ and $20 \log_{10} |U_1(Z_2)|$ and to use these values along with the determined value of $20 \log_{10} (F/X_c)$ to solve Eq. (470) for $20 \log_{10} (V_1/X_c)$. Figure 2·19 may then be used to find X_c in natural units. By repeating this process for several receiver heights, one obtains sufficient points on the coverage contour to draw that part of the contour which is well within the diffraction region.

Unfortunately, for all radar systems and many one-way circuits the bottom of the lowest lobe lies very near the horizon in the intermediate

region, where in general no direct analytical solution is possible. The procedure to be followed is very similar to that outlined in the previous discussion of the intermediate region. To find X_c for a given transmitter and target or receiver height, a graph of 20 $\log_{10}(F/X)$ must be drawn as a function of distance X through the intermediate region. This can usually be accomplished by finding 20 $\log_{10}(F/X)$ at the first maximum ($\Delta R = \lambda/2$), at the first point of quadrature ($\Delta R = \lambda/4$), and at several points in the diffraction region and drawing a smooth curve through the points so obtained. The range X_c at which Eq. (469) is satisfied is the maximum range for the given terminal heights. By repeating the procedure for various target or receiver heights, enough points can be obtained to plot X_c as a function of target or receiver height.

This technique can best be illustrated by means of an example. Consider a vertically polarized 9-cm radar whose free-space detection range R_0 on a certain type of airplane is 63 statute miles. Then,

$$-20 \log_{10} \frac{R_0}{R_s} = -100 \text{ db}.$$

At what ranges will the plane be detected over smooth sea water for altitudes between 50 and 1000 ft if the radar is at a height of 100 ft? We must plot 20 $\log_{10}(F/X)$ as a function of X for heights between 50 and 1000 ft for a transmitter height of 100 ft. Let us choose heights of 50,

TABLE 2·1.—INTERFERENCE-REGION CALCULATIONS FOR A SAMPLE COVERAGE PROBLEM

(*a*) Determination of 20 $\log_{10}(F/X)$ at the first maximum ($k\,\Delta R = \pi$ and $F = 1 + D\rho$) for $Z_1 = 3.23$.

Z_2	X_L	T	Q [Eq. (456)]	S (Figs. E to G)	X	D (Figs. C and D)	K (Fig. B)	$\psi_2°$ [Eq. (450)]	ρ [Eq. (5·6)]	20 $\log_{10} F$	20 $\log_{10} \frac{F}{X}$
1.62	3.06	0.707	1.84	0.394	1.21	0.867	0.83	0.14	0.96	+5	+ 3
3.23	3.59	1.000	1.08	0.510	1.83	0.767	0.74	0.11	0.97	+5	0
9.69	4.90	0.577	0.492	0.667	3.27	0.650	0.53	0.09	0.97	+4	− 6
16.15	5.80	0.447	0.350	0.724	4.20	0.620	0.43	0.09	0.97	+4	− 8
24.2	6.70	0.365	0.269	0.763	5.11	0.600	0.36	0.08	0.98	+4	−10
32.3	7.48	0.316	0.225	0.789	5.90	0.590	0.31	0.08	0.98	+4	−11

(*b*) Determination of 20 $\log_{10}(F/X)$ at the first point of quadrature ($k\,\Delta R = \pi/2$ and $F = \sqrt{1 + D^2\rho^2}$) for $Z_1 = 3.23$.

1.62	3.06	0.707	0.919	0.548	1.68	0.750	0.69	0.09	0.97	+2	− 2
3.23	3.59	1.000	0.540	0.643	2.31	0.643	0.59	0.07	0.98	+1	− 6
9.69	4.90	0.577	0.246	0.760	3.72	0.540	0.40	0.06	0.98	+1	−10
16.15	5.80	0.447	0.175	0.802	4.65	0.513	0.32	0.06	0.98	+1	−12
24.2	6.70	0.365	0.135	0.833	5.58	0.493	0.26	0.05	0.99	+1	−14
32.3	7.48	0.316	0.113	0.854	6.39	0.480	0.22	0.05	0.99	+1	−15

100, 300, 500, 750, and 1000 ft. In natural units the transmitter height Z_1 is 3.23 and target heights Z_2 are 1.62, 3.23, 9.69, 16.15, 24.2, and 32.3. The angles are low enough that $\phi = \pi$. We can find the distance at which the first maximum and first quadrature points occur for each of these height combinations by determining Q and T and using Figs. E to G to find S and X. With S known we can use Figs. B to D to find D and K (and consequently ψ_2). Finally we find the value of ρ for the

TABLE 2·2.—DIFFRACTION-REGION CALCULATIONS FOR A SAMPLE COVERAGE PROBLEM

Z_1	Z_2	X	$20\log_{10}\lvert U_1(Z_1)\rvert$	$20\log_{10}\lvert U_1(Z_2)\rvert$	$20\log_{10}\left(\frac{V}{X}\right)$	$20\log_{10}\left(\frac{F}{X}\right)$
3.23	1.62	4.0	+16	+7	−65	−42
		5.0			−83	−60
		6.0			−102	−79
3.23	3.23	4.5	+16	+16	−74	−42
		6.0			−102	−70
		7.0			−121	−89
3.23	9.69	6.0	+16	+39	−102	−47
		7.0			−121	−66
		8.0			−139	−84
3.23	16.15	7.0	+16	+56	−121	−49
		8.0			−139	−67
		9.0			−157	−85
3.23	24.2	8.0	+16	+71	−139	−52
		9.0			−157	−70
		10.0			−175	−88
3.23	32.3	8.5	+16	+82	−148	−50
		9.0			−157	−59
		10.0			−175	−77

given ψ_2 from Fig. 5·4, or more accurately from Eq. (5·6). The values of ρ listed in Table 2·1 were computed from Eq. (5·6) with $\varepsilon_c = 69 - i39$, as given in Table 5·1 for $\lambda = 10$ cm. It is then possible to find $20\log_{10}F$ and $20\log_{10}(F/X)$ directly. The numerical work required in this process is given in Table 2·1.

The calculations must now be made for the diffraction region. For each set of heights let us choose three points below the horizon separated by about one X unit and calculate $20\log_{10}(F/X)$ by means of Figs. 2·19 and 2·20. The calculations are shown in Table 2·2.

We have now made all necessary field-strength calculations and can draw curves of $20\log_{10}(F/X)$ as a function of X for each set of heights chosen. The resulting curves are shown in Fig. 2·23. Since we know

that for the given system and target $-20 \log_{10} (R_0/R_s) = -100$ db, we use Eqs. (469) and (471) with $R_f = R_0$, and then Fig. 2·22 to obtain a figure of merit in natural units. We find

$$20 \log_{10} \frac{F}{X_c} = -18 \text{ db.}$$

The line corresponding to $20 \log_{10} (F/X_c) = -18$ db is also drawn on Fig. 2·23. The range at which this line intersects a given field-strength

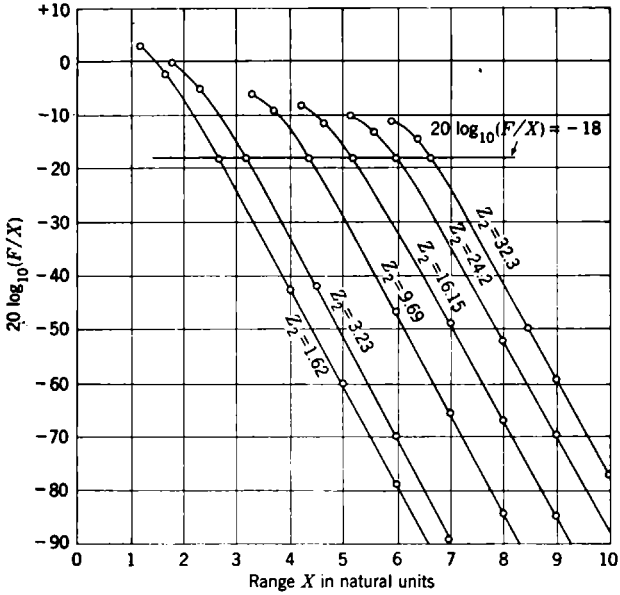


FIG. 2·23.—Field-strength curves employed to determine the coverage contour shown in Fig. 2·24.

curve is the maximum range for the target height for which the curve was drawn. By using all six such points of intersection we may construct the desired coverage diagram shown in Fig. 2·24 with the height and distance scales marked both in natural units and in physical units.

On occasion it will be found that the given line of $20 \log_{10} (F/X_c)$ will lie above the first maximum for the highest or higher heights chosen. This indicates that the lowest lobe will have been closed before this height is reached; for example, if

$$20 \log_{10} \frac{F}{X_c} = -9 \text{ db}$$

for the given system and target, this would mean that the first lobe would have been closed between 500 ft ($Z_2 = 16.15$) and 750 ft ($Z_2 = 24.2$).

Figure 2·23 can actually be used for many more coverage problems than the one worked out here, particularly because it has been drawn in

terms of natural units. In the first place, we may choose as many different figures of merit as we wish for any circuit as long as $Z_1 = 3.23$; the desired contour diagram is obtained by drawing the appropriate 20 $\log_{10}$ (F/X_c) line on the diagram and finding the points of intersection. It is a customary procedure to draw many such contours on a coverage chart such as Fig. 2·23, thereby increasing its applicability. The second great advantage of this type of coverage diagram is the result of using natural units throughout the computations. Although we have worked a coverage problem for $z_1 = 100$ ft and $\lambda = 9$ cm, the resulting diagram shown in Fig. 2·24 is equally applicable to another wavelength if the proper physical unit coordinates are added and the transmitter height is changed consistent with the relation between physical units, wavelength, and natural units. Thus a coverage diagram for $z_1 = 100$ ft and $\lambda = 9$ cm can equally well be a coverage diagram for $z_1 = 48$ ft and $\lambda = 3$ cm, $z_1 = 500$ ft and $\lambda = 1$ m, etc. In view of this fact, it can be seen that a set of coverage charts, each for a different Z_1 and each having many values of 20 $\log_{10}$ (F/X_c) drawn, can be used to determine coverage for circuits with a wide range of system performance, operating wavelength, and transmitter height.[1]

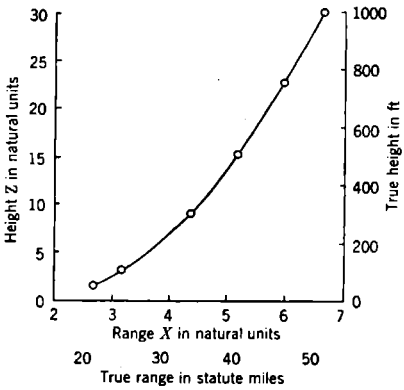


FIG. 2·24.—Coverage contour of a vertically polarized 9-cm radar at a height of 100 ft viewing an elevated target over smooth sea water. The free-space range is 63 statute miles.

The method of determining coverage contours previously discussed can be extended as far into the interference region as desired by extending the graphs of 20 $\log_{10}$ (F/X_c) vs. X back into the interference region. Although we used the method only to locate the bottom of the first lobe, by extending the plots we could have located the top of the first lobe, the bottom of the second lobe, and so on. For microwaves the lobe structure is usually so fine that there is no practical need for determining the detailed lobe structure, and it is sufficient to draw the locus of the lobe maxima. In the case of radar transmission, no moving target will be able to escape detection for a long period within this locus. For fixed point-to-point transmission, diversity reception may often be a sufficient safeguard against being caught in a minimum.[2]

[1] *Cf.* B. E. Howard, H. W. Dodson, and J. R. Gill, "Field Intensity Contours in Generalized Coordinates," RL Report No. 702, May 2, 1945.

[2] Diversity reception is accomplished by combining the output of two or more receivers separated by an appreciable distance.

If only the locus of maxima is desired, we can write

$$R_{\max} = R_f[f(\psi_1 - \xi) + f(-\psi_2 - \eta - \xi)\, D\rho(\psi_2)], \tag{472}$$

where the quantity in brackets will be recognized as the expression for F in a lobe maximum. For transmitter heights of less than 1000 ft a rapid construction of the coarse coverage can be made in terms of γ, the elevation angle measured from the base of the transmitter. For ranges of more than 20 miles, it will almost always be true that $\gamma \approx \psi_1 \approx \psi_2$ and $\eta \ll \gamma$; hence Eq. (472) becomes

$$R_{\max} = R_f[f(\gamma - \xi) + f(-\gamma - \xi)\, D\rho(\gamma)]. \tag{473}$$

The maximum range can thus be found as a function of γ if the divergence

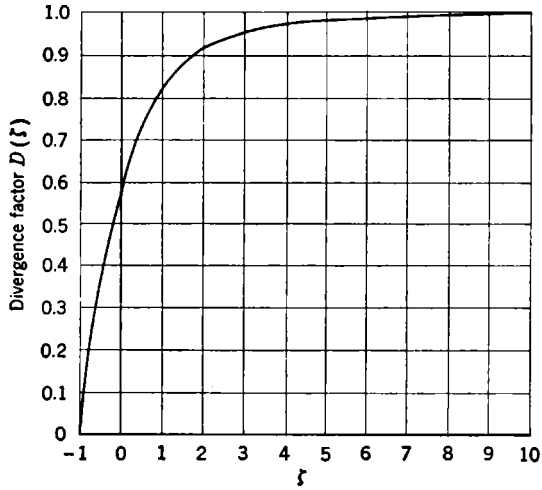


FIG. 2·25.—Divergence factor $D(\zeta)$ as a function of ζ, where $\zeta = \dfrac{a_e \tan \gamma}{\sqrt{2a_e z_1}}$.

factor can be expressed as a function of γ. This can be done[1] with Figs. 2·25 and 2·26.

If we let

$$\zeta = \frac{a_e \tan \gamma}{\sqrt{2a_e z_1}},$$

where z_1 is the transmitter height and $a_e = 5280$ statute miles for standard refraction, then D is given approximately by

$$D(\zeta) = \left[\frac{1}{3}\left(1 + \frac{2\zeta}{\sqrt{\zeta^2 + 3}}\right)\right]^{1/2} \tag{474}$$

[1] R. A. Hutner *et al.*, "Field Intensity Formulas," RL Report No. 23, Sept. 28, 1943, pp. 31–33.

Equation (474) is plotted in Fig. 2·25. A better approximation is given by

$$D(\gamma) = D(\zeta)\left(1 + \frac{\epsilon}{100}\right), \tag{475}$$

where the percentile error factor ϵ can be determined approximately from Fig. 2·26. The contours of ϵ in Fig. 2·26 have been determined in terms of the quantity $\eta = \sqrt{z_1/200}$, where z_1 is in feet. The contours of Fig. 2·26 are accurate only for $z_1 = 200$ ft but are reasonably accurate for

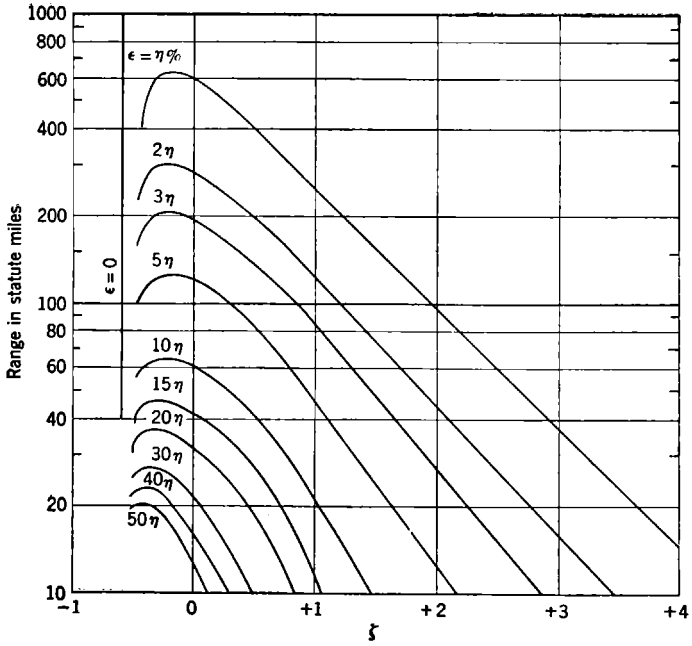


FIG. 2·26.—Contours of per cent error ϵ in $D(\zeta)$ as a function of ζ and approximate range. These contours are correct for $z_1 = 200$ ft and are approximate for other heights.

heights of less than 1000 ft and $\zeta > 0$. Calculations of D for $\zeta < 0$ are apt to be appreciably in error. As the free-space range of the set is known, it will usually be possible to estimate the range well enough to use Fig. 2·26. Although subject to numerous limitations, this method of determining the divergence factor approximately for transmitter heights of less than 1000 ft is in general very satisfactory. Coarse coverage diagrams can be prepared rapidly from Eq. (473) by this method. Such coverage diagrams are usually drawn on specially prepared coordinate paper on which curves of constant γ are straight lines. On this paper, height is measured vertically, whereas distance is measured along a parabola. The parabola representing the earth's surface falls off $r^2/2$ ft at a range of r statute miles. Figure 2·27 is drawn on such a graph.

The reader is advised to be thoroughly familiar with Secs. 5·5 to 5·12 before using Eq. (473). It is pointed out in these sections that for microwaves over land, specular reflection may be very small. If reflections are neglected, Eq. (473) reduces to

$$R_{\max} = R_f f(\gamma - \xi). \tag{476}$$

For horizontally polarized microwave radiation over sea water, measured values of ρ are often considerably less than unity, and values of ρ used in

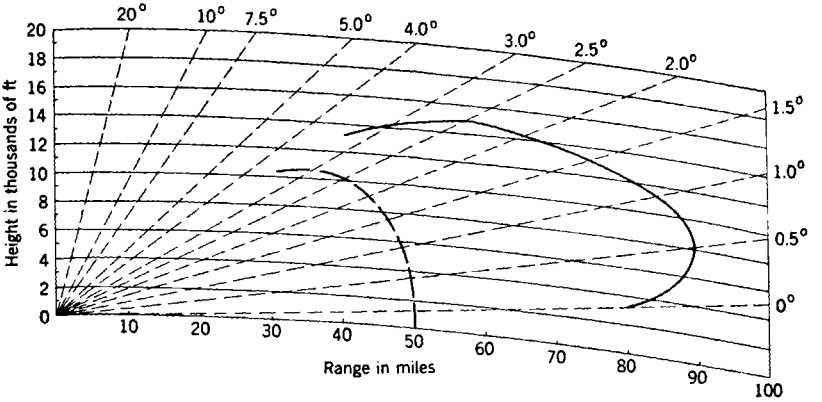


FIG. 2·27.—Coarse coverage diagram for a 10-cm radar employing vertical polarization and sited 200 ft above a smooth sea. The free-space range is 50 miles, and the half-power beamwidth is 6°. The sloping dashed lines indicate elevation angle, and the curved dashed line shows the free-space coverage diagram.

determining the coverage contour should be chosen with the material in Sec. 5·5 to 5·8 in mind.

As an example the coarse coverage diagram will be found for a 10-cm radar system with a horizontally pointed, symmetrical paraboloidal

TABLE 2·3.—CALCULATIONS FOR COARSE COVERAGE DIAGRAM OF VERTICALLY POLARIZED RADAR SYSTEM WITH $\lambda = 10$ CM, $z_1 = 200$ FT, AND $R_0 = 50$ STATUTE MILES

$\gamma°$	ξ	$D(\xi)$ (Fig. 2·25)	ϵ, per cent (Fig. 2·26)	$D(\gamma)$ [Eq. (475)]	$\rho(\gamma)$ (Fig. 5·4)	$f(\gamma)$ [Eq. (455)]	F	$R_{\max}$, statute miles
0	0	0.58	7	0.62	1.00	1.00	1.62	81
0.5	2.3	0.93	1	0.94	0.86	0.99	1.79	90
1.0	4.6	0.98	0	0.98	0.73	0.97	1.66	84
1.5	6.9	0.99	0	0.98	0.62	0.93	1.49	74
2.0	9.2	1.00	0	1.00	0.53	0.87	1.33	66
2.5	.....	.....	.....	1.00	0.45	0.80	1.16	58
3.0	.....	.....	.....	1.00	0.37	0.71	0.97	48
3.5	.....	.....	.....	1.00	0.31	0.60	0.79	40

antenna having a beamwidth of 6°. The system uses vertical polarization and has a free-space range of 50 statute miles on a given type of aircraft target. It is sited at an elevation of 200 ft and looks out over sea water. We can find D from Figs. 2·25 and 2·26 and ρ from Fig. 5·4. For such an antenna Eq. (473) reduces to

$$R_{\max} = R_0 f(\gamma)(1 + D\rho), \tag{477}$$

where $f(\gamma)$ can be found from Eq. (455) by setting $\theta_0 = 6°$. The necessary calculations are shown in Table 2·3 and the coarse coverage diagram may be found in Fig. 2·27.

THE BILINEAR MODIFIED-INDEX PROFILE

By Wendell H. Furry

2·17. Definition of the Problem and Preliminary Formulation.—In the treatment of modified-index profiles that are more complicated than the linear profile the following compensatory simplifications will be used:

1. Only the diffraction region will be considered. No attempt will be made to accomplish a reduction of the field integral for the interference region, as was done in Sec. 2·11 for the linear modified-index profile.
2. The fundamental theorem of Sec. 2·7 will be applied immediately; the distributions to be considered are such that the necessary assumptions are certainly valid.
3. Corresponding to the fact that only the diffraction region is to be considered, the boundary condition at the earth's surface will always be taken in the simplified form given in Eq. (158). The expression for ψ [Eq. (164)] then reduces to the series term only. Also, there is no remaining dependence of the results on polarization.

The formula for ψ will now be written in terms of the natural units defined in Sec. 2·10 and shown in Figs. 2·17 and 2·18. Substitution of

$$r = XL = 2(k_0 q^2)^{-1/3} X, \tag{478}$$

$$\kappa_m = k_0 \sqrt{1 - qHA_m} = k_0 \sqrt{1 - \left(\frac{q}{k_0}\right)^{2/3} A_m} \approx k_0 - \frac{A_m}{L}, \tag{479}$$

in the series of Eq. (164) and the use of the asymptotic expansion of the Hankel function [Eq. (357)] lead to the result

$$\psi = e^{-i(k_0 r + \pi/4)} \sqrt{\frac{\pi}{X}} \left(\frac{q}{k_0}\right)^{1/3} \sum_{\text{all } m} e^{iA_m X} u_m^{(1)}(z_1) u_m^{(1)}(z_2). \tag{480}$$

The only approximations involved in this transformation are the use of the asymptotic expansion for the Hankel function, which is strongly justified, and the approximation indicated in Eq. (479). The latter

approximation depends on the smallness of $(q/k_0)^{2/3}|A_m|$. For the first few terms of the series, the values of $|A_m|$ are not large numbers, and k_0/q is roughly the number of wavelengths in the earth's circumference; thus this approximation is strongly justified for the terms of the series that are of importance.

Equation (480) still contains the heights z_1, z_2 in ordinary units. Because the simplified boundary condition of Eq. (158) is to be used, Eq. (166) can be written

$$\int_0^\infty [u_m^{(1)}(z)]^2\, dz = 1, \tag{481}$$

the integral being along a ray in the fourth quadrant of the z-plane. The definition

$$U_m(Z) \equiv H^{1/2} u_m^{(1)}(z) \tag{482}$$

gives a function of Z normalized; hence

$$\int_0^\infty U_m^2(Z)\, dZ = 1, \tag{483}$$

the integral being along a ray in the fourth quadrant of the Z-plane. Since

$$\left(\frac{q}{k_0}\right)^{1/3} \frac{1}{H} = \frac{2}{L}, \tag{484}$$

Eq. (480) now becomes

$$\psi = \frac{2\sqrt{\pi}}{L} e^{-i(k_0 r + \pi/4)} \frac{1}{\sqrt{X}} \sum_m e^{iA_m X} U_m(Z_1) U_m(Z_2). \tag{485}$$

This is the same as Eq. (427). The pattern-propagation factor F is equal to the magnitude of $r\psi$; accordingly,

$$F = 2\sqrt{\pi X} \left| \sum_m e^{iA_m X} U_m(Z_1) U_m(Z_2) \right|. \tag{486a}$$

If only one mode makes an appreciable contribution,

$$F = 2\sqrt{\pi X}\; e^{-C_{m_1} X} |U_{m_1}(Z_1) U_{m_1}(Z_2)|. \tag{486b}$$

Here the index m_1 refers to the significant mode, and C_{m1} is the attenuation constant, equal to $\mathrm{Im}(A_{m1})$. The method of determining the numbers A_m and the functions $U_m(Z)$ for a given distribution of the modified index has still to be specified.

The numbers A_m and functions $U_m(Z)$ are determined by the solution of a boundary-value problem. The problem in question is obtained by translating into natural units the boundary-value problem that determines κ_m^2 and $u_m^{(1)}(z)$. When Eqs. (479), (137), (482) and the expression

$$z = HZ = (qk_0^2)^{-1/3} Z \tag{487}$$

are substituted in Eq. (138), the resulting differential equation for $U_m(Z)$ is

$$\frac{d^2U_m}{dZ^2} + [Y(Z) + A_m]U_m = 0. \tag{488}$$

Here

$$Y(Z) = \left(\frac{k_0}{q}\right)^{2/3}(N^2 - 1) = \frac{N^2 - 1}{qH} \tag{489}$$

and is a function of Z that expresses the distribution of modified index. Since

$$Y'(Z) = H\frac{d}{dz}Y = \frac{1}{q}\frac{d}{dz}N^2, \tag{490}$$

it is seen that $Y'(Z) = 1$ wherever the curve $N^2(z)$ has the standard slope q.

The boundary conditions on $U_m(Z)$ follow directly from those on $u_2(z,\kappa_m)$:

$$U_m(Z)\, e^{i\omega t} \rightarrow \text{upward-moving wave}, \quad Z \rightarrow \infty; \tag{491}$$

$$U_m(0) = 0. \tag{492}$$

The condition of normalization has already been stated in Eq. (483). An alternative expression of the normalization condition, which is sometimes more convenient, can be obtained from Eqs. (479), (482), (483), (165), and (185) by using stipulation 3, as a result of $p(\kappa^2) \rightarrow \infty$. The condition so expressed is

$$-U_m'(0)\left[\frac{\partial U(0)}{\partial A}\right]_{A = Am} = 1. \tag{493}$$

Here the function $U(Z)$, whose partial derivative is taken, satisfies Eqs. (488) and (491) but does not satisfy Eq. (492) for $A \neq A_m$.

This formulation applies to all cases of modified-index profile, subject to stipulations 1 to 3. The remainder of this section and the three that follow treat the problem of a distribution given by a curve made up of portions of two straight lines; hence

$$\frac{dN^2}{dz} = q = \text{const.}, \qquad z > d, \tag{494}$$

$$\frac{dN^2}{dz} = p = \text{const.}, \qquad z < d, \tag{495}$$

$$\lim_{z \rightarrow d+} N^2 = \lim_{z \rightarrow d-} N^2. \tag{496}$$

This is called the *bilinear distribution*. For $p < 0$, a duct exists, and d is the duct height; in general, d is the depth of the nonstandard layer. By

using Eq. (490) a function $Y(Z)$ can be determined for use in calculations made in natural units:

$$Y(Z) = Z - g, \qquad Z > g; \tag{497}$$

$$Y(Z) = s^3(Z - g), \qquad Z < g. \tag{498}$$

Here

$$g = \frac{d}{H} \tag{499}$$

is the depth of the nonstandard layer in natural units, and

$$s^3 = \frac{p}{q} \tag{500}$$

is the ratio of the slope of the curve of N^2 in the nonstandard layer to the standard slope of the curve above the layer. Equations (497) and (498) contain the convention, adopted for the sake of convenience and definiteness, that $Y = 0$ for $z = g$. It is clear from Eq. (488) that if Y is changed by an additive constant, the quantities A_m are simply changed by the negative of the same constant.

The definition [Eq. (500)] gives meaning only to s^3, but s and s^2 occur often in the calculations. There is no loss in generality and a considerable gain in convenience in stipulating that s shall always be regarded as real.

Figure 2·28*a* shows the curve of Y against Z for a case with a negative value of s; in such cases there is a duct, and g is the duct height in natural units. Figure 2·28*b* shows a case in which s is positive and greater than unity; this gives a substandard layer. The curves shown are of course the same in shape as the M-profiles for such cases, with a particular choice of units and of the axis $Y = 0$.

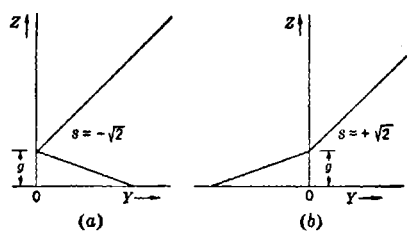


FIG. 2·28.—Examples of bilinear N-profiles, drawn to scales expressed in natural units.

The formal solution of the boundary-value problem for the bilinear profile can be obtained in terms of the functions discussed in Sec. 2·9. In order to have the equations in a form that facilitates numerical calculation of the functions $U_m(Z)$, most of the equations are written in terms of the functions $h_1(\zeta)$ and $h_2(\zeta)$, which have been tabulated for complex arguments.

When Eqs. (497) and (498) are substituted into Eq. (488), it is at once apparent that the solution has the general form

$$U_m(Z) = y_\alpha(Z - g + A_m), \qquad Z \geqq g, \tag{501}$$

$$U_m(Z) = y_\beta(sZ - sg + s^{-2}A_m), \qquad Z < g. \tag{502}$$

Here $y_\alpha(\zeta)$ and $y_\beta(\zeta)$ are solutions of Eq. (304). From Eq. (491) it follows that y_α must be a multiple of h_2:

$$U_m(Z) = K_m h_2(Z - g + A_m), \qquad Z \geqq g. \tag{503}$$

The factor K_m must eventually be determined from Eq. (483) or (493). The function y_β must be a linear combination of h_1 and h_2 and must be so chosen that $U_m(Z)$ and $U_m'(Z)$ are continuous at $Z = g$. Then,

$$U_m(Z) = \frac{K_m}{sW}\begin{vmatrix} h_1(s^{-2}A_m) & h_2(A_m) & h_2(s^{-2}A_m) \\ sh_1'(s^{-2}A_m) & h_2'(A_m) & sh_2'(s^{-2}A_m) \\ h_1(sZ - sg + s^{-2}A_m) & 0 & h_2(sZ - sg + s^{-2}A_m) \end{vmatrix}, \tag{504}$$

Here

$$W = h_1h_2' - h_2h_1' = -1.457495i. \tag{505}$$

The boundary condition [Eq. (492)] must now be applied. This step gives a transcendental equation whose roots determine the characteristic values A_m. For convenience in writing many formulas in the following section one can introduce the symbol

$$w = s^{-2}A_m. \tag{506}$$

From Eqs. (492) and (504) there follows the transcendental equation for w

$$\phi(w) \equiv \begin{vmatrix} h_1(w) & sh_2(s^2w) & h_2(w) \\ h_1'(w) & h_2'(s^2w) & h_2'(w) \\ h_1(w - sg) & 0 & h_2(w - sg) \end{vmatrix} = 0. \tag{507}$$

Methods of solving this equation are discussed in the following section.

The simplest formula for the constant K_m is obtained from Eq. (493). From Eqs. (504) and (507) it is seen that

$$U_m'(0) = -\frac{K_m}{sW}\frac{\partial\phi}{\partial g}. \tag{508}$$

Then from Eqs. (493), (504), (507), and (508), it follows that

$$\left(\frac{K_m}{sW}\right)^2\left(\frac{\partial\phi}{\partial g}\frac{\partial\phi}{\partial A}\right)_{A = A_m} = 1. \tag{509}$$

Since, by the rule for implicit differentiation,

$$\frac{dA_m}{dg} = -\frac{\dfrac{\partial\phi}{\partial g}}{\left(\dfrac{\partial\phi}{\partial A}\right)_{A = A_m}}, \tag{510}$$

Eq. (509) can be written

$$K_m^2 = -(sW)^2\left(\frac{\partial\phi}{\partial g}\right)^{-2}\frac{dA_m}{dg}. \tag{511}$$

Then, by Eq. (508),

$$[U_m'(0)]^2 = -\frac{dA_m}{dg}. \tag{512}$$

In all these equations s is, of course, to be regarded as a given constant.

Whenever values of A_m have been determined for a sequence of values of g with fixed s and dA_m/dg is not too small to be readily evaluated and used, Eq. (511) or (512) provides a convenient rule for normalizing $U_m(Z)$. Another formula is needed, however, for cases in which A_m has been found for only one value of g or in which dA_m/dg is very small.

The required formula can be obtained from Eq. (483). For any function $y(\zeta)$ that satisfies Eq. (304), the formula of indefinite integration

$$\int^{\zeta} [y(\zeta)]^2\, d\zeta = \zeta\,[y(\zeta)]^2 + [y'(\zeta)]^2 + \text{const.}$$

can be verified by differentiation. Then from Eqs. (483), (501), and (502), it follows that

$$(s^{-3} - 1)\{A_m\,[U_m(g)]^2 + [U_m'(g)]^2\} - s^{-3}[U_m'(0)]^2 = 1. \tag{514}$$

The first term can be evaluated from Eq. (503), and the second from Eqs. (504) and (507); the result is

$$K_m^{-2} = (s^{-3} - 1)\{A_m\,[h_2(A_m)]^2 + [h_2'(A_m)]^2\} - s^{-5}W^{-2}\left(\frac{\partial\phi}{\partial g}\right)^2. \tag{515}$$

By comparing Eqs. (512) and (514), it is seen that when the condition

$$\left|\frac{dA_m}{dg}\right| \ll |s^3| \tag{516}$$

is satisfied, the last term of Eq. (515) can be neglected.

Most of the preceding equations cannot be used directly for the case $s = 0$. The simplest way to get the results for this case is to begin again with the differential equation [Eq. (488)]. Equation (503) still holds for $Z > g$, but for $Z < g$ the differential equation is not of the type of Eq. (304); instead it is

$$\frac{d^2U_m}{dZ^2} + A_m U_m = 0, \qquad Z < g, \quad s = 0. \tag{517}$$

The solution that vanishes at $Z = 0$ and makes $U_m(Z)$ continuous at $Z = g$ is

$$U_m(Z) = K_m \frac{h_2(A_m)}{\sin\,(A_m^{1/2}g)} \sin\,(A_m^{1/2}Z), \qquad s = 0. \tag{518}$$

The requirement that $U_m'(Z)$ be continuous at $Z = g$ gives the transcendental equation

$$\phi_0(A_m) \equiv \frac{h_2(A_m)}{h_2'(A_m)} - A_m^{-1/2} \tan\,(A_m^{1/2}g) = 0, \qquad s = 0. \tag{519}$$

The formula that replaces Eq. (511) for $s = 0$ is obtained by substituting Eq. (518) into Eq. (512):

$$K_m^2 = -\frac{\sin^2 (A_m^{\frac{1}{2}} g)}{A_m [h_2(A_m)]^2} \frac{dA_m}{dg}, \qquad s = 0. \tag{520}$$

A formula to replace Eq. (515) can be obtained by direct evaluation of the integral in Eq. (483):

$$K_m^{-2} = -\left[\left(1 - \frac{g}{2A_m}\right) \sec^2 (A_m^{\frac{1}{2}} g) + \frac{1}{2} A_m^{-\frac{3}{2}} \tan (A_m^{\frac{1}{2}} g)\right] [h_2'(A_m)]^2. \tag{521}$$

2·18. Methods for Calculating Characteristic Values.—In this section a number of formulas are obtained for the approximate calculation of characteristic values A_m for various values of s and g. These formulas cover cases in which g is either small or large. When the value of g is neither small nor large, none of these formulas applies, and A_m must be computed by direct numerical solution of the transcendental equation [Eq. (507) or (519)].

The numerical solution of the transcendental equation may be accomplished in either of two ways:

1. The method of mapping. Values of ϕ are computed numerically for several complex values of w, and the results are plotted on the complex ϕ-plane. From the resulting pattern of points the value of w that makes $\phi = 0$ can be estimated approximately.
2. The Newton-Raphson method. Successive estimates $w_0, w_1, \cdots$ of the value of a root are connected by the formula

$$w_n = w_{n-1} - \frac{\phi(w_{n-1})}{\phi'(w_{n-1})}. \tag{522}$$

The details of the application of these two methods need not be discussed here.[1] The attempt to reduce the amount of numerical work required has affected the choice of cases to be treated. Whenever values of any of the functions h_1, h_2, h_1', h_2' have to be obtained for an argument not given directly in the existing tables, a rather laborious process of interpolation is necessary. For this reason the values of s that are chosen are usually such that s^2 is a simple rational number, and values of g are chosen so that $10sg$ is an integer. At least the earlier part of the computation can then be accomplished with values taken directly from the tables.

The first approximate formula that will be discussed is the one for small values of g. For $g = 0$ the bilinear case reduces to the linear case; and

[1] They are discussed in W. H. Furry, "Methods of Calculating Characteristic Values for Bilinear M-curves," RL Report No. 795, Feb. 6, 1946. The appendices to this report contain further details of the derivation of the approximate formulas given in this section.

because of the use of the simplified boundary condition [Eq. (158) or (492)], the characteristic values A_m are just the numbers ζ_m that satisfy

$$h_2(\zeta_m) = 0. \tag{523}$$

This may be seen by using Eqs. (492) and (503) for $g = 0$. The numbers ζ_m all have the same phase,

$$\zeta_m = |\zeta_m| e^{2\pi i/3}. \tag{524}$$

Values of the first few ζ_m and an asymptotic formula for large $|\zeta_m|$ are given in Eq. (338).

For $s = 1$, $g \neq 0$, the distribution is still essentially the linear one; but as shown by Eqs. (497) and (498), Y has been decreased by the amount g; hence A_m must exceed ζ_m by g. For $s \neq 1$, further corrections appear, all proportional to $s^3 - 1$ and to higher powers of g. A_m is thus expressed as a power series in g.

In the derivation of this series, use is made of the Riccati equation corresponding to the linear differential equation [Eq. (304)]. If $y(\zeta)$ is a solution of Eq. (304), then the function

$$r(\zeta) = \frac{y(\zeta)}{y'(\zeta)} \tag{525}$$

is a solution of the Riccati equation

$$\frac{dr}{d\zeta} = 1 + \zeta r^2. \tag{526}$$

The zeros of $r(\zeta)$ are the same as those of $y(\zeta)$, as $y'(\zeta)$ is an integral function. We denote by ζ' the correction to be applied to ζ_m to obtain A_m,

$$s^2 w = A_m = \zeta_m + \zeta'. \tag{527}$$

Then if $r_\alpha(\zeta)$ is a solution of Eq. (526), which also satisfies the requirement

$$r_\alpha(s^2 w - \zeta') = 0, \tag{528}$$

it follows from Eqs. (523), (525), and (501) that

$$r_\alpha(Z - g + A_m) = \frac{y_2(Z - g + A_m)}{y_2'(Z - g + A_m)} = \frac{h_2(Z - g + A_m)}{h_2'(Z - g + A_m)} = \frac{U_m(Z)}{U_m'(Z)}, \qquad Z > g. \tag{529}$$

Similarly, if $r_\beta(\zeta)$ satisfies Eq. (526) and the requirement

$$r_\beta(w - sg) = 0, \tag{530}$$

then, by Eqs. (525), (492), and (502),

$$s^{-1} r_\beta(sZ - sg + w) = \frac{U_m(Z)}{U_m'(Z)}, \qquad Z < g. \tag{531}$$

The requirement that $U_m(Z)$ and $U'_m(Z)$ be continuous at $Z = g$ reduces to the equation

$$r_\alpha(s^2w) = s^{-1}r_\beta(w). \tag{532}$$

Equations (526), (527), (528), (530), and (532) now define the problem of determining ζ' as a function of s, g, and ζ_m. The symmetry of Eqs. (528), (530), and (532) greatly simplifies the calculation. The procedure used is first to find ζ' as a power series in g with coefficients involving s and s^2w and then to use Eq. (527) to obtain from this series a series whose coefficients involve s and ζ_m. The final series could be obtained in a single step but only with much greater effort.

The series,

$$r_\beta(\zeta) = \sum_{n=1}^{\infty} \alpha_n(\zeta)(\zeta - w + sg)^n \tag{533}$$

obviously satisfies Eq. (530). By substituting it in Eq. (526) and equating coefficients of successive powers of $(\zeta - w + sg)$, we can determine the quantities $\alpha_n(\zeta)$:

$$\begin{aligned} \alpha_1(\zeta) &= 1, & \alpha_4(\zeta) &= -\tfrac{1}{12}, \\ \alpha_2(\zeta) &= 0, & \alpha_5(\zeta) &= \tfrac{2}{15}\zeta^2, \\ \alpha_3(\zeta) &= \tfrac{1}{3}\zeta, & & \cdots\cdots\cdots. \end{aligned} \tag{534}$$

Then

$$r_\beta(w) = \sum_{n=1}^{\infty} \alpha_n(w)(sg)^n. \tag{535}$$

Similarly,

$$r_\alpha(s^2w) = \sum_{n=1}^{\infty} \alpha_n(s^2w)\zeta'^n. \tag{536}$$

We can express ζ' formally as a power series in g,

$$\zeta' = \sum_{n=1}^{\infty} \beta_n g^n. \tag{537}$$

The coefficients β_n can be determined by using Eqs. (534) to (537) in Eq. (532) and equating coefficients of successive powers of g:

$$\beta_1 = 1;\ \beta_2 = \beta_3 = 0;\ \beta_4 = \frac{1 - s^3}{12};\ \beta_5 = 0;\ \beta_6 = \frac{s^3 - 1}{90}\, s^2w;\ \cdots. \tag{538}$$

Use of Eqs. (537) and (538) in Eq. (527) gives the result

$$A_m = \zeta_m + g + \frac{1 - s^3}{12} g^4 + \frac{s^3 - 1}{90} A_m g^6 + \cdots. \tag{539}$$

An algebraic process is now required to obtain a series in which the coefficients depend on ζ_m, not on A_m. For the calculation of terms to and

including g^{12} it suffices to replace A_m, where it appears in the coefficients, by the approximate value

$$A_m \approx \zeta_m + g + \frac{1 - s^3}{12} g^4 + \frac{s^3 - 1}{90} \zeta_m g^6. \tag{540}$$

The final result is

$$A_m = \zeta_m + g + \frac{1}{12} S g^4 - \frac{1}{90} S \zeta_m g^6 + \frac{13S^2 - 5S}{1260} g^7 + \frac{1}{1260} S \zeta_m^2 g^8$$
$$+ \frac{7S - 47S^2}{11{,}340} \zeta_m g^9 + \left(- \frac{1}{28{,}350} S \zeta_m^3 + \frac{119S^3 - 190S^2 + 11S}{90{,}720} \right) g^{10}$$
$$+ \frac{17S^2 - S}{23{,}100} \zeta_m^2 g^{11}$$
$$+ \left(\frac{1}{935{,}550} S \zeta_m^4 + \frac{-15{,}709S^3 + 10{,}826S^2 - 265S}{14{,}968{,}800} \zeta_m \right) g^{12} + \cdots, \tag{541}$$

where

$$S = 1 - s^3. \tag{542}$$

This series predicts the initial behavior of A_m as a function of g for given s, but it does not allow the behavior of A_m to be followed very far. The power series in Eqs. (535) and (536) has a fairly small radius of convergence, and the same is presumably true of the series of Eq. (541). The labor of obtaining additional terms of this last series mounts very rapidly.

For large values of g, separate formulas must be used for the three cases $s < 0$, $s = 0$, and $s > 0$. The first case $s < 0$ gives strongly trapped modes when g is large. The values of A_m are very nearly purely real negative numbers. The functions in terms of which Eq. (507) is written cannot in such cases be reliably and unambiguously evaluated by means of the asymptotic expansions [Eqs. (326) to (333)], because the arguments in question fall very near the boundaries of the ranges stated for the use of the expansions. For this reason it is best to rewrite the equation in terms of the Airy integrals, which are defined by Eqs. (339) and (340). The Airy integral $\mathrm{Ai}(-\zeta)$ is proportional to the integral $y_3(\zeta)$, described in Sec. 2·9. From Eqs. (311), (328), and (329) it is seen that for values of ζ near the negative real axis the functions $\mathrm{Ai}(-\zeta)$ and $\mathrm{Ai}'(-\zeta)$ have a uniquely defined asymptotic behavior: It is a strong exponential decrease as $|\zeta|$ increases. When the asymptotic expansions given in Sec. 2·9 are applied to the functions $\mathrm{Bi}(-\zeta)$ and $\mathrm{Bi}'(-\zeta)$ for such values of ζ, ambiguities are found in connection with the exponentially decreasing terms. These terms can, however, be ignored in evaluating $\mathrm{Bi}(-\zeta)$ and $\mathrm{Bi}'(-\zeta)$, for the exponentially increasing terms are enormously larger and the use of $\mathrm{Ai}(-\zeta)$ and $\mathrm{Ai}'(-\zeta)$ as the other basic solutions takes proper account of the possibility of the negative exponential behavior for a solution.[1]

[1] This procedure gives formulas that agree with the asymptotic expansions given in the introduction to the *British Association Tables of the Airy Integrals.*

The first and third columns of the determinant in Eq. (507) can be rewritten by replacing the h_1 and h_2 functions by the Airy integrals, since the function y_β of Eq. (502) could equally well have been regarded as a linear combination of the latter. The middle column must be translated by using Eqs. (343) and (344). The resulting transcendental equation is

$$\begin{vmatrix} \mathrm{Ai}(-w) & s[\mathrm{Bi}(-s^2w) - i\mathrm{Ai}(-s^2w)] & \mathrm{Bi}(-w) \\ -\mathrm{Ai}'(-w) & -[\mathrm{Bi}'(-s^2w) - i\mathrm{Ai}'(-s^2w)] & -\mathrm{Bi}'(-w) \\ \mathrm{Ai}(-w+sg) & 0 & \mathrm{Bi}(-w+sg) \end{vmatrix} = 0. \quad (543)$$

This equation can be rewritten in the form

$$\frac{\begin{vmatrix} s[\mathrm{Bi}(-s^2w) - i\mathrm{Ai}(-s^2w)] & \mathrm{Ai}(-w) \\ -\mathrm{Bi}'(-s^2w) + i\mathrm{Ai}'(-s^2w) & -\mathrm{Ai}'(-w) \end{vmatrix}}{\begin{vmatrix} s[\mathrm{Bi}(-s^2s) - i\mathrm{Ai}(-s^2w)] & \mathrm{Bi}(-w) \\ -\mathrm{Bi}'(-s^2w) + i\mathrm{Ai}'(-s^2w) & -\mathrm{Bi}'(-w) \end{vmatrix}} = \frac{\mathrm{Ai}(-w+sg)}{\mathrm{Bi}(-w+sg)}. \quad (544)$$

For strongly trapped modes, w and s^2w are sizable negative numbers. Then in the left member of Eq. (544) the values of Bi and Bi′ are very large compared with those of Ai and Ai′. Accordingly, the roots of the equation are values of w that make $\mathrm{Ai}(-w + sg)$ take very small values. From Eqs. (342), (523), and (524), it is seen that the zeros of $\mathrm{Ai}(-\zeta)$ are the absolute values of the zeros of $h_2(\zeta)$. Hence a zeroth approximation to the roots of Eq. (544) is

$$w_0 = |\zeta_m| + sg. \quad (545)$$

As s is negative, w_0 is negative for large values of g.

When the value

$$w = w_0 + \Delta w \quad (546)$$

is substituted in the right-hand member of Eq. (544), this member becomes

$$\frac{\mathrm{Ai}(-w + sg)}{\mathrm{Bi}(-w + sg)} = \frac{\mathrm{Ai}(-|\zeta_m| - \Delta w)}{\mathrm{Bi}(-|\zeta_m| - \Delta w)} = -\frac{\mathrm{Ai}'(-|\zeta_m|)}{\mathrm{Bi}(-|\zeta_m|)}\Delta w + O[(\Delta w)^2]. \quad (547)$$

For large g the correction Δw is exponentially small. In the left member of Eq. (544) we can replace w by w_0 and use the asymptotic expansions of the Airy integrals. The asymptotic value of Δw can then be found. The resulting formula is

$$A_{(m)} \approx s^2w_0 + s^2\alpha_m \left\{ -\frac{1}{16} e^{-\frac{4}{3}|w_0|^{3/2}}(1 - s^{-3})|w_0|^{-3/2} \left[1 + \left(-\frac{17}{24} - \frac{3}{4s^3} \right)|w_0|^{-3/2} + \cdots \right] + \frac{i}{4} e^{-\frac{4}{3}(1-s^3)|w_0|^{3/2}} \left[1 + \frac{1}{24}(1 - s^{-3})|w_0|^{-3/2} + \frac{19}{1152}(1 - s^{-3})^2|w_0|^{-3} + \cdots \right] \right\}. \quad (548)$$

Here w_0 is given by Eq. (545), and

$$\alpha_m = -\frac{\mathrm{Bi}(-|\zeta_m|)}{\mathrm{Ai}'(-|\zeta_m|)}. \tag{549}$$

The first few values of α_m are $\alpha_1 = 0.6474$, $\alpha_2 = 0.4935$, $\alpha_3 = 0.4252$. The asymptotic formula

$$\alpha_m \approx [\tfrac{3}{2}(m - \tfrac{1}{4})\pi]^{-\frac{1}{3}} \tag{550}$$

gives the values of further α_m's very well and gives even α_1 with an error of less than 1.5 per cent.

If the first square bracket in Eq. (548) is neglected and the second is replaced by unity, and if the approximate formulas of Eqs. (338) and (550) are used throughout, Eq. (548) becomes just the result given by the Gamow phase-integral method. This is the only case for the bilinear model in which a phase-integral method gives correctly the first term of a valid asymptotic series.[1]

Equations (545), (548), and (549) provide a sequence of solutions for $m = 1, 2, 3, \cdots$; Eq. (548) can be used only when w_0 is negative and not too small numerically; hence there is a limitation on the number of solutions that can be obtained in this way for given values of s and g. Attention must be called to the fact that use of the same value of m in Eqs. (541) and (548) does not necessarily give values of A_m that actually correspond to each other in the sense that one goes over into the other continuously as g increases from small values for which Eq. (541) holds to large values for which Eq. (548) holds. Actually this is not the case in general, but there is good reason to believe that the correspondence does always hold for $m = 1$. The use of the subscript (m) instead of m is a reminder of this lack of general correspondence.

Some of the values A_m given by Eq. (541) for small g and negative s do not go over for large g into any of the values given by Eq. (548). Instead they approach certain constant values, functions of s, as $g \to \infty$. These modes should ordinarily make no important contribution to the field strength, as they are always more strongly attenuated than the first mode of standard propagation, whereas trapped modes have small attenuations. These cases will not be discussed at this point; formulas for these values of A_m in the limit $g \to \infty$ can be stated more conveniently after the case of positive s has been treated.

For $s = 0$, $g \to \infty$, there are two possibilities: Either $A_m \to 0$, or A_m remains finite. The case $A_m \to 0$ is the important one in practice, for it gives modes with small attenuations. The first step in the solution of

[1] The Eckersley type of phase-integral method cannot be used because $Y(Z)$ is not an analytic function. The failure of attempts to use asymptotic expansions to obtain solutions of this sort is discussed in Furry, *op. cit.*, Appendix B.

Eq. (519) for $g \gg 1$, $|A_m| \ll 1$ is to write one of its terms as a Maclaurin's expansion,

$$\frac{h_2(A_m)}{h_2'(A_m)} = a + bA_m + cA_m^2 + \cdots. \tag{551}$$

The coefficients are found to be

$$a = \frac{h_2(0)}{h_2'(0)} = 1.3717\, e^{2\pi i/3} = -0.6859 + 1.1880i, \tag{552}$$

$$b = 1, \qquad c = \tfrac{1}{2}\, a^2, \qquad d = \tfrac{2}{3}\, a, \cdots. \tag{553}$$

The differentiations required in obtaining the coefficients are most readily performed by means of Eqs. (525) and (526). The radius of convergence is equal to the magnitude of the first zero of $h_2'(\zeta)$, or 1.0188.

When Eqs. (551) to (553) are substituted into Eq. (519), the resulting equation can be written

$$\tan\,(A_m^{1/2} g) = aA_m^{1/2} + A_m^{3/2} + \tfrac{1}{2}\, a^2 A_m^{3/2} + \tfrac{2}{3}\, aA_m^{5/2} + \cdots. \tag{554}$$

For $g \gg a$ it is clear that the zeroth approximation to the roots must be

$$A_m^{1/2} \approx \frac{m\pi}{g}. \tag{555}$$

If

$$A_m^{1/2} = \frac{m\pi}{g} + \gamma, \tag{556}$$

then

$$\tan\,(A_m^{1/2} g) = \tan\,(\gamma g) = \gamma g + \tfrac{1}{3}\, \gamma^3 g^3 + \tfrac{2}{15}\, \gamma^5 g^5 + \cdots. \tag{557}$$

The quantity γ can be assumed to have the form,

$$\gamma \approx \frac{\gamma_2}{g^2} + \frac{\gamma_3}{g^3} + \cdots. \tag{558}$$

When Eqs. (556) to (558) are substituted in Eq. (554) and the coefficients of successive powers of g are equated separately, the constants γ_2, γ_3, $\cdots$ can be evaluated. The final formula for A_m, found by squaring Eq. (556), is

$$\begin{aligned} A_{(m)} \approx \frac{(m\pi)^2}{g^2}\Bigg[1 + \frac{2a}{g} + \frac{3a^2}{g^2} + \frac{4a^3 + 2(m\pi)^2(1 - \frac{1}{3}\, a^3)}{g^3} \\ + \frac{5a^4 + 10(m\pi)^2 a(1 - \frac{1}{3}\, a^3)}{g^4} \\ + \frac{6a^5 + 30(m\pi)^2 a^2(1 - \frac{1}{3}\, a^3) - (m\pi)^4 a^2(1 - \frac{2}{5}\, a^3)}{g^5} + \cdots \Bigg]. \end{aligned} \tag{559}$$

Here a is the complex number given in Eq. (552).

Use of the same value of m in Eq. (541) (with $s = 0$) and Eq. (559) does not necessarily give values of A_m that actually correspond to each other; the situation is similar to that of Eqs. (541) and (548) for negative values of s. For $m = 1$, however, numerical calculation of values of A_1 for intermediate values of g shows that it actually does go over into the function $A_{(1)}$ given by Eq. (559).

If A_m remains finite, with positive imaginary part, for $g \to \infty$, then $\tan(A_m^{\frac{1}{2}} g) \to i$. Equation (519) shows that finite limiting values of A_m for $s = 0$, $g \to \infty$ must satisfy the equation

$$h_2'(A_m) + i A_m^{\frac{1}{2}} h_2(A_m) = 0. \tag{560}$$

Roots of this equation will be denoted by $A_{[m]}^{(\infty)}(0)$, to indicate that they are limiting values for $g \to \infty$ and for $s = 0$. The two smallest roots are

$$\begin{aligned} A_{[1]}^{(\infty)}(0) &= -0.446 + 2.962i, \\ A_{[2]}^{(\infty)}(0) &= -1.334 + 4.380i, \end{aligned} \tag{561}$$

and an approximate formula for large m is

$$A_{[m]}^{(\infty)}(0) \approx \left(\frac{3}{2} m\pi\right)^{\frac{2}{3}} e^{2\pi i/3} \left(1 - \frac{i \ln(12m\pi)}{3m\pi}\right). \tag{562}$$

This equation gives the value of $A_{[2]}^{(\infty)}(0)$ correct within 1 per cent.

It would not be difficult to obtain an expression for large finite values of g to correspond to the foregoing results for $g \to \infty$, but this has not been done. The imaginary part of $A_{[1]}^{(\infty)}(0)$ is 1.46 times as large as that of ζ_1, the smallest characteristic value for the standard case; hence all such modes are strongly attenuated and are of little actual importance.

The case of positive values of s and large values of g can be discussed most readily by using the symmetrically defined functions $y_1(\zeta)$, $y_2(\zeta)$, $y_3(\zeta)$, introduced in Sec. 2·9, to express the solutions written in Eqs. (501) and (502). The transcendental equation [Eq. (507)] can then be written

$$\begin{vmatrix} y_1(w) & sy_2(s^2w) & y_3(w) \\ y_1'(w) & y_2'(s^2w) & y_3'(w) \\ y_1(w-sg) & 0 & y_3(w-sg) \end{vmatrix} = 0. \tag{563}$$

The most natural assumption, on general principles, is that $\mathrm{Re}(w) \to +\infty$ for $g \to \infty$; the reasons why this is an attractive hypothesis will appear in the next section, where the consequences of its falsity are studied. On considering the elements of the first two rows in Eq. (563), one sees from Eqs. (311) and (320) to (325) that for $\mathrm{Im}(w) > 0$, $\mathrm{Re}(w) \gg 1$, these elements in the first column become very small whereas these elements in the other two columns are very large. Then the last element of the first column must be very small; and as $g \to \infty$ and $\mathrm{Re}(w) \to \infty$, $w - sg$ must approach a zero of $y_1(\zeta)$. Because these zeros, namely, ζ_m^*, have negative imaginary parts and this would violate the general theorem that $\mathrm{Im}(A_m) > 0$, the hypothesis that $\mathrm{Re}(w) \to \infty$ for $g \to \infty$ leads to a contradiction.

It is established, accordingly, that $\mathrm{Re}(w - sg) \to -\infty$ as $g \to \infty$. Then $y_3(w - sg) \to 0$, and $|y_1(w - sg)| \to \infty$. Thus the characteristic values approach certain finite values, determined by the roots of the equation

$$\begin{vmatrix} sy_2(s^2w) & y_3(w) \\ y_2'(s^2w) & y_3'(w) \end{vmatrix} = 0 \qquad (s > 0,\ g \to \infty). \tag{564}$$

The limiting characteristic values will be denoted by $A_{(m)}^{(\infty)}(s)$ or $A_{[m]}^{(\infty)}(s)$ and are equal to s^2 times the roots of Eq. (564). The use of subscripts (m) and $[m]$ will be adjusted so that the notation is consistent with that already used in Eqs. (559) and (562) for $s = 0$.

The form of Eq. (564) makes it possible to establish a useful general relation between $A^{(\infty)}(s)$ and $A^{(\infty)}(1/s)$. From Eqs. (311) and (312) and the results of differentiating those equations with respect to $\mathrm{Re}(\zeta)$, one readily obtains

$$y_2(\zeta^* e^{\pi i/3}) = [-y_1(\zeta e^{-\pi i/3})]^* = [e^{-\pi i/3} y_3(\zeta e^{\pi i/3})]^*, \tag{565}$$

$$y_2'(\zeta^* e^{\pi i/3}) = [-y_1'(\zeta e^{-\pi i/3})]^* = [e^{\pi i/3} y_3'(\zeta e^{\pi i/3})]^*. \tag{566}$$

When the substitutions

$$w = s^{-2} e^{2\pi i/3} z^*; \quad s = u^{-1} \tag{567}$$

are made in Eq. (564), the two columns are interchanged, the first row is multiplied by $ue^{-\pi i/3}$, and the second row by $e^{\pi i/3}$, the resulting equation is

$$\begin{vmatrix} ue^{-\pi i/3} y_3(u^2 z^* e^{\pi i/3} e^{\pi i/3}) & e^{-\pi i/3} y_2(z^* e^{\pi i/3} e^{\pi i/3}) \\ e^{\pi i/3} y_3'(u^2 z^* e^{\pi i/3} e^{\pi i/3}) & e^{\pi i/3} y_2'(z^* e^{\pi i/3} e^{\pi i/3}) \end{vmatrix} = 0. \tag{568}$$

The complex conjugate of this equation is just

$$\begin{vmatrix} uy_2(u^2 z) & y_3(z) \\ y_2'(u^2 z) & y_3'(z) \end{vmatrix} = 0. \tag{569}$$

Comparison of Eqs. (564) and 569) shows that z is the same function of u that w is of s. Then by Eq. (567) it follows that

$$s^2 w(s) = e^{2\pi i/3} \left[w\left(\frac{1}{s}\right) \right]^* = s^2 e^{2\pi i/3} \left[\frac{1}{s^2} w\left(\frac{1}{s}\right) \right]^* \qquad (g \to \infty,\ s > 0). \tag{570}$$

That is,

$$A^{(\infty)}(s) = s^2 e^{2\pi i/3} \left[A^{(\infty)}\left(\frac{1}{s}\right) \right]^* \qquad (s > 0). \tag{571}$$

This relation holds whether the subscript is (m) or $[m]$. It can be written in the form

$$A^{(\infty)}(s) = s^2 e^{\pi i/3} \left[e^{-\pi i/3} A^{(\infty)}\left(\frac{1}{s}\right) \right]^* \qquad (s > 0). \tag{572}$$

This shows that the complex point $A^{(\infty)}(s)$ can be obtained by reflecting the point $A^{(\infty)}(1/s)$ in the line making the angle 60° with the real axis and multiplying by s^2.

According to Eq. (559), $A_{(m)}^{(\infty)}(s)$ approaches zero for $s \to 0$. As $y_2(0)$ is finite and neither $y_3(0)$ nor $y_2'(0)$ vanishes, it is clear from Eq. (564) that w must not vanish for $g \to \infty$, $s \to 0$, $s^2 w \to 0$. Indeed, w must approach one of the zeros of $y_3(\zeta)$, which are the numbers $|\zeta_m|$. Equation (564) can be written

$$\frac{s y_2(s^2 w)}{y_2'(s^2 w)} = \frac{y_3(w)}{y_3'(w)}. \tag{573}$$

As $y_2(\zeta)$ is proportional to $h_2(\zeta)$, Eqs. (551) to (553) provide an expansion of the left member of Eq. (573);

$$\frac{s y_2(s^2 w)}{y_2'(s^2 w)} = sa + s^3 w + \frac{1}{2} a^2 s^5 w^2 + \frac{2}{3} a s^7 w^3 + \cdots. \tag{574}$$

The right-hand member of Eq. (573) can be expressed by setting

$$w = |\zeta_m| + \delta \tag{575}$$

and making use of Eqs. (525) and (526) to obtain the Taylor's series

$$\frac{y_3(w)}{y_3'(w)} = \delta + \frac{1}{3}|\zeta_m|\delta^3 + \frac{1}{4}\delta^4 + \frac{2}{15}|\zeta_m|^2\delta^5 + \cdots. \tag{576}$$

The quantity δ can be assumed to be a power series

$$\delta = \delta_1 s + \delta_2 s^2 + \delta_3 s^3 + \cdots. \tag{577}$$

By the use of Eqs. (573) to (577), the coefficients $\delta_1, \delta_2, \cdots$ can be determined. The resulting formula for $A_{(m)}^{(\infty)}(s)$, valid for small s, is

$$A_{(m)}^{(\infty)}(s) = |\zeta_m| s^2 + a s^3 + (1 - \tfrac{1}{3} a^3)|\zeta_m| s^5 + a(1 - \tfrac{1}{4} a^3) s^6 + a^2(-\tfrac{1}{2} + \tfrac{1}{5} a^3)|\zeta_m|^2 s^7 + (1 - \tfrac{4}{3} a^3 + \tfrac{7}{18} a^6)|\zeta_m| s^8 + \cdots. \tag{578}$$

The notation used is consistent with that of Eq. (559). The complex quantity a is defined in Eq. (552); a^3 is real. By means of Eq. (571), one can at once obtain from Eq. (578) an expression valid for large s; in doing this we make use of the fact that

$$\arg(a) = \arg(\zeta_m) = \frac{2\pi}{3}. \tag{579}$$

The resulting formula for large s is

$$A_{(m)}^{(\infty)}(s) = \zeta_m + |a| s^{-1} + (1 - \tfrac{1}{3} a^3)\zeta_m s^{-3} + |a|(1 - \tfrac{1}{4} a^3) s^{-4} + |a|^2(-\tfrac{1}{2} + \tfrac{1}{5} a^3)\zeta_m^2 s^{-5} + (1 - \tfrac{4}{3} a^3 + \tfrac{7}{18} a^6)\zeta_m s^{-6} + \cdots. \tag{580}$$

Expressions will now be obtained for $A_{(m)}^{(\infty)}(s)$ for small s and for large s. For $s \to 0$, $A_{(m)}^{(\infty)}(s)$ remains finite, and w must become infinite. Equa-

tions (561) and (562) show that $\arg(w)$ is nearly equal to but somewhat less than $2\pi/3$. From Eqs. (311), (328), and (329) we find that for w becoming extremely large with this phase angle,

$$\frac{y_3(w)}{y_3'(w)} = iw^{-\frac{1}{2}} + O(w^{-2}). \tag{581}$$

When Eq. (581) is used in Eq. (573), Eq. (560) is again obtained as the equation satisfied by $A_{[m]}^{(\infty)}(0)$. A more accurate asymptotic expression than Eq. (581) can be obtained either from the formalism of Sec. 2·9 or more simply from the Riccati equation [Eqs. (525) and (526)]. The result is

$$\frac{y_3(w)}{y_3'(w)} = iw^{-\frac{1}{2}} - \frac{1}{4}w^{-2} - \frac{7}{32}iw^{-\frac{7}{2}} + \cdots. \tag{582}$$

When the quantity ϵ is defined by

$$A_{[m]}^{(\infty)} = s^2 w = A_{[m]}^{(\infty)}(0) + \epsilon, \tag{583}$$

a Taylor's expansion gives

$$\frac{sw^{\frac{1}{2}}y_2(s^2w)}{y_2'(s^2w)} = i + \frac{i}{2A_{[m]}^{(\infty)}(0)}\epsilon + \left\{-\frac{1}{2[A_{[m]}^{(\infty)}(0)]^{\frac{1}{2}}} - \frac{i}{8[A_{[m]}^{(\infty)}(0)]^2}\right\}\epsilon^2 + \cdots. \tag{584}$$

The quantity ϵ can be assumed to have the form

$$\epsilon = \epsilon_1 s^3 + \epsilon_2 s^6 + \cdots. \tag{585}$$

The coefficients $\epsilon_1, \epsilon_2 \cdots$ can be found from Eqs. (573) and (582) to (585). The final result, valid for small s, is

$$A_{[m]}^{(\infty)}(s) = A_{[m]}^{(\infty)}(0) + \tfrac{1}{2}i\,[A_{[m]}^{(\infty)}(0)]^{-\frac{1}{2}}(s^3 + s^6) - \tfrac{1}{8}[A_{[m]}^{(\infty)}(0)]^{-2}s^6 + \cdots. \tag{586}$$

The formula for large s, found by the use of Eq. (571), is

$$A_{[m]}^{(\infty)}(s) = s^2 e^{2\pi i/3}[A_{[m]}^{(\infty)}(0)]^* + \tfrac{1}{2}\,e^{\pi i/6}\{[A_{[m]}^{(\infty)}(0)]^{-\frac{1}{2}}\}^*(s^{-1} + s^{-4}) - \tfrac{1}{8}\,e^{2\pi i/3}[A_{[m]}^{(\infty)}(0)]^{*-2}\,s^{-4} + \cdots. \tag{587}$$

The formulas obtained show that the loci in the complex plane of the numbers $A_{(m)}^{(\infty)}(s)$ start from the origin for $s = 0$ and reach the values ζ_m for $s \to \infty$, whereas the loci of $A_{[m]}^{(\infty)}(s)$ start from the values $A_{[m]}^{(\infty)}(0)$ for $s = 0$ and recede to infinity, in the asymptotic directions given by rays from the origin through the points $e^{\pi i/3}[e^{-\pi i/3}A_{[m]}^{(\infty)}(0)]^*$, for $s \to \infty$.

For values of s very nearly equal to unity, the limit points are given approximately by the roots of the equation

$$\Phi(w) \equiv h_2(w)h_2'(w) + 2w\{[h_2'(w)]^2 + w[h_2(w)]^2\} = \frac{We^{i\pi/3}}{s-1}. \tag{588}$$

This equation is obtained from Eq. (564) by replacing $y_2(s^2w)$ by $h_2(s^2w)$ and $y_3(w)$ by a linear combination of $h_1(w)$ and $h_2(w)$, by means of the formulas of Sec. 2·9, and expanding the functions of s^2w in terms of those of w to first order in $s - 1$. It should be noted that W is the Wronskian defined in Eq. (505). As $h_2(w)$ has no singularities for finite w, it is clear that $s \to 1$ means $A^{(\infty)}(s) \to \infty$. From the formulas of Sec. 2·9, one finds that for large w,

$$\Phi(w) \approx \tfrac{3}{8}\, We^{\pi i/3} w^{-3/2} e^{-(4i/3)w^{3/2}} [1 + O(w^{-3/2})], \tag{589}$$

for $-\pi < \arg(w) < \pi/3$; actually this result is valid beyond these limits, up to the vicinity of the line on which ζ_m, the zeros of $h_2(\zeta)$, are located. From Eq. (588) it is seen that loci of real s, with s very nearly equal to unity, are curves on which $\Phi(w)$ is large and has constant phase. By Eq. (589), these curves are asymptotic to one of the lines $\arg(w) = -\pi$, $\arg(w) = \pi/3$. The requirement that these lines shall be related as shown in Eq. (571) and shall all be in the upper half plane shows that the asymptote is $\arg A^{\infty}(s) = \arg(w) = \pi/3$. From Eqs. (588) and (589) one finds that the branches for $s > 1$ and $s < 1$ that approach the asymptote most rapidly do so approximately symmetrically from the two sides and that for large $|w|$ the perpendicular distance between adjacent branches lying near the asymptote and corresponding to different signs of $s - 1$ is approximately

$$|w| \left| (\arg w)_{s>1} - (\arg w)_{s<1} \right| \approx \frac{\pi}{2|w|^{1/2}}. \tag{590}$$

The approximate formulas that have been obtained for small s, large s, and s nearly equal to unity, together with a number of exact results obtained in connection with numerical calculations of characteristic values, make it possible to draw fairly accurately the loci of $A_{(m)}^{(\infty)}(s)$ and $A_{[m]}^{(\infty)}(s)$ in the complex plane. The result is shown in the solid curves of Fig. 2·29. The labels on the curves are the subscripts (m) and $[m]$. The branches below the line $\arg A = \pi/3$ correspond to $s < 1$ for (m) and $s > 1$ for $[m]$; the opposite is true for the upper branches. A few exactly calculated points are marked on the curves for future reference.

No attempt has been made to map the location of values of s along the curves, apart from labeling end points and asymptotes. Considerable portions of Curve (1) will be presented in more detail in the next section. One important general feature of the distribution of values of s is that the part of the lower branch of any curve that lies fairly within the system of roughly parallel asymptotes to the ray $\arg A = \pi/3$ corresponds to the values of s very near to unity. This is exemplified by the points marked A' and A'' for $s = 1.02$) and B'' (for $s = \sqrt{0.98}$); indeed, these points fall on parts of the curves that are scarcely within the system of asymptotic portions of the curves.

Equation (563) can be written

$$\frac{\begin{vmatrix} y_3(w) & sy_2(s^2w) \\ y_3'(w) & y_2'(s^2w) \end{vmatrix}}{\begin{vmatrix} y_1(w) & sy_2(s^2w) \\ y_1'(w) & y_2'(s^2w) \end{vmatrix}} = \frac{y_3(w - sg)}{y_1(w - sg)}. \tag{591}$$

If g is large but not infinite, the right member is small but not zero and w must depart slightly from its limiting value. This difference is important only for the numerator of the left member, which vanishes when w has its limiting value. The value w_g that w has when g is large but finite is accordingly given by the approximate formula

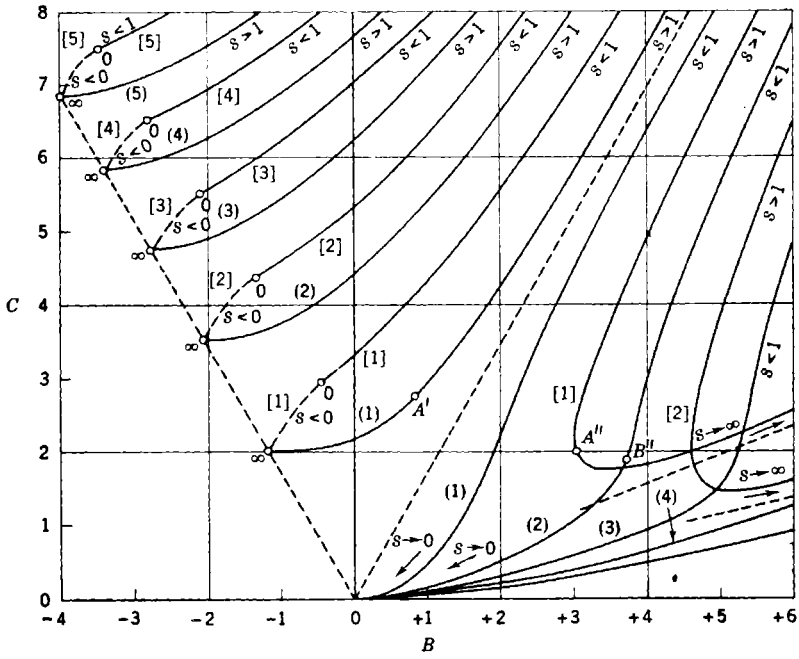


FIG. 2·29.—Loci of $A_{(m)}^{(\infty)}(s)$ and $A_{[m]}^{(\infty)}(s)$ in the complex plane.

$$w_g - w \approx \frac{\begin{vmatrix} y_1(w) & sy_2(s^2w) \\ y_1'(w) & y_2'(s^2w) \end{vmatrix}}{\frac{d}{dw}\begin{vmatrix} y_3(w) & sy_2(s^2w) \\ y_3'(w) & y_2'(s^2w) \end{vmatrix}} \frac{y_3(w - sg)}{y_1(w - sg)}. \tag{592}$$

When the differentiation is carried out and the expression is put into form for numerical computation by replacing the y-functions by their equivalents in terms of the h-functions, the formula obtained for large

but finite values of g is

$$w_g - w \approx \frac{\begin{vmatrix} h_1(w) & sh_2(s^2w) \\ h_1'(w) & h_2'(s^2w) \end{vmatrix}}{(s^3-1)\begin{vmatrix} h_1(w)+e^{-\pi i/3}h_2(w) & h_2'(s^2w) \\ h_1'(w)+e^{-\pi i/3}h_2'(w) & -wsh_2(s^2w) \end{vmatrix}} \left[1 + e^{-\pi i/3}\frac{h_2(w-sg)}{h_1(w-sg)}\right]. \quad (593)$$

Here w means the limiting value for $g \to \infty$; only the last factor depends on g; it is small because $h_2(\zeta)/h_1(\zeta) \cong e^{-2\pi i/3}$ for $-\mathrm{Re}(\zeta) \gg 1$.

The case in which w approaches a constant value for large g and negative s can be discussed by means of the sort of argument that has been given for positive s. For negative s the more important solutions are those for the trapped modes, whose characteristic values, denoted by $A_{(m)}$, are given by Eq. (548). The characteristic values for negative s that approach finite limits for $g \to \infty$ are, accordingly, denoted by $A_{[m]}$, and their limiting values by $A_{[m]}^{(\infty)}(s)$. The equation for the $A_{[m]}^{(\infty)}(s)$ for negative s is obtained by applying to Eq. (507) the sort of argument that, for positive s, was applied to Eq. (563). Because $\mathrm{Im}(w) > 0$, as $g \to \infty$ for fixed w and negative s, $h_1(w - sg)$ becomes small and $h_2(w - sg)$ becomes large, by Eqs. (328) and (330). The equation for the limiting value of w is, accordingly, in the y-notation

$$\begin{vmatrix} y_1(w) & sy_2(s^2w) \\ y_1'(w) & y_2'(s^2w) \end{vmatrix} = 0 \quad (s < 0, g \to \infty). \quad (594)$$

The symmetry condition analogous to Eq. (572) is readily proved; it involves reflection in the real axis and thus does not actually show a valid relationship between acceptable solutions, as Eq. (594) is derived by using the fact that $\mathrm{Im}(w) > 0$. The fact that for negative s the symmetry condition cannot be used to obtain a further solution for $g \to \infty$ from any given solution can be regarded as connected with the fact that one of the sets of characteristic values for negative s, the $A_{(m)}$, does not approach finite limits for large g.

A formal connection between the solutions of Eq. (594) and those of Eq. (564) can be obtained by setting

$$s^2w = A, \qquad s = -ve^{\pi i/3}. \quad (595)$$

Equation (594) then becomes, on multiplying the first row by $e^{2\pi i/3}$ and the first column by $e^{2\pi i/3}$ and using Eq. (311),

$$\begin{vmatrix} y_3(v^{-2}A) & vy_2(A) \\ y_3'(v^{-2}A) & y_2'(A) \end{vmatrix} = 0. \quad (596)$$

On the other hand, Eq. (564) can be written

$$\begin{vmatrix} y_3(s^{-2}A) & sy_2(A) \\ y_3'(s^{-2}A) & y_2'(A) \end{vmatrix} = 0. \quad (597)$$

The solution of Eq. (596) is the same function of $v = -se^{-\pi i/3}$ that the solution of Eq. (597) is of s. When s is replaced by $-se^{-\pi i/3}$ in Eqs. (578) and (587), the resulting quantities have negative imaginary parts and must be rejected. When this replacement is made in Eqs. (580) and (586), the resulting formulas, valid for *negative* s, are

$$A_{[m]}^{(\infty)}(s) = \zeta_m + a^* s^{-1} + (1 - \tfrac{1}{3}a^3)\zeta_m s^{-3} + a^*(1 - \tfrac{1}{4}a^3)s^{-4} + a^{*2}(-\tfrac{1}{2} + \tfrac{1}{5}a^3)\zeta_m^2 s^{-5} + (1 - \tfrac{4}{3}a^3 + \tfrac{7}{18}a^6)\zeta_m s^{-6} + \cdots, \quad (598)$$

$$A_{[m]}^{(\infty)}(s) = A_{[m]}^{(\infty)}(0) + \tfrac{1}{2}i[A_{[m]}^{(\infty)}(0)]^{-\frac{1}{2}}(s^3 + s^6) - \tfrac{1}{8}[A_{[m]}^{(\infty)}(0)]^{-2}s^6 + \cdots. \quad (599)$$

Equation (599) is the same in form as Eq. (586) but applies for $s < 0$ instead of $s > 0$. Thus the loci of $A_{[m]}^{(\infty)}(s)$ are continuous and smooth through $s = 0$.

For $s = -1$, $g \to \infty$ we have from Eq. (594), using the h-notation for purposes of numerical calculation,

$$h_1(A)h_2'(A) + h_1'(A)h_2(A) = 0. \quad (600)$$

A root of this equation, found by Newton's method, is

$$A_{[1]}^{(\infty)}(-1) = -0.591 + 2.830i. \quad (601)$$

By the use of asymptotic formulas, an approximate formula for large m is found to be[1]

$$A_{[m]}^{(\infty)}(-1) \approx e^{2\pi i/3}\left(\frac{3}{2}m\pi\right)^{\frac{2}{3}}\left(1 - \frac{i\ln 6m\pi}{3m\pi}\right). \quad (602)$$

Application of the standard procedure for implicit differentiation to Eq. (594) gives an approximate expression for values of s near -1:

$$A_{[m]}^{(\infty)}(s) = A_{[m]}^{(\infty)}(-1) + (s+1)\left[\frac{h_1'(A)h_2(A)}{2h_1'(A)h_2'(A) - 2Ah_1(A)h_2(A)} - A\right]_{A = A_{[m]}^{(\infty)}(-1)} + \cdots. \quad (603)$$

Equations (598) to (603) make it possible to plot the loci of $A_{[m]}^{(\infty)}(s)$ for negative s with reasonable accuracy. These loci are shown as dashed curves in Fig. 2·29.

An argument like that used in the derivation of Eq. (593) gives a formula for large but finite g and *negative* s:

$$w_g - w \approx \frac{\begin{vmatrix} h_2(w) & sh_2(s^2w) \\ h_2'(w) & h_2'(s^2w) \end{vmatrix}}{(s^3-1)\begin{vmatrix} h_1(w) & h_2'(s^2w) \\ h_1'(w) & -wsh_2(s^2w) \end{vmatrix}} \frac{h_1(w - sg)}{h_2(w - sg)}. \quad (604)$$

As in Eq. (593), w_g means the value for the given finite value of g, and w means the limiting value $s^{-2}A_{[m]}^{(\infty)}(s)$.

[1] The derivations of Eqs. (602) and (562) are given in Furry, *op. cit.*, Appendices E and D.

2·19. Behavior of Characteristic Values and Characteristic Functions for the First Mode.—The characteristic value A_1 has been computed over a range of values of g for each of the following values of s: -3, -2, $-\sqrt{2}$, -1, 0, $\sqrt{\frac{1}{2}}$, $\sqrt{\frac{2}{3}}$, $\sqrt{\frac{4}{5}}$, $\sqrt{\frac{9}{10}}$, $\sqrt{2}$, 2. The quantity A_1 is defined by using the series of Eq. (541) with $m = 1$ for small values of g and continuing the calculation numerically for intermediate values of g until it becomes clear which formula holds for large g. In all of the cases in question it

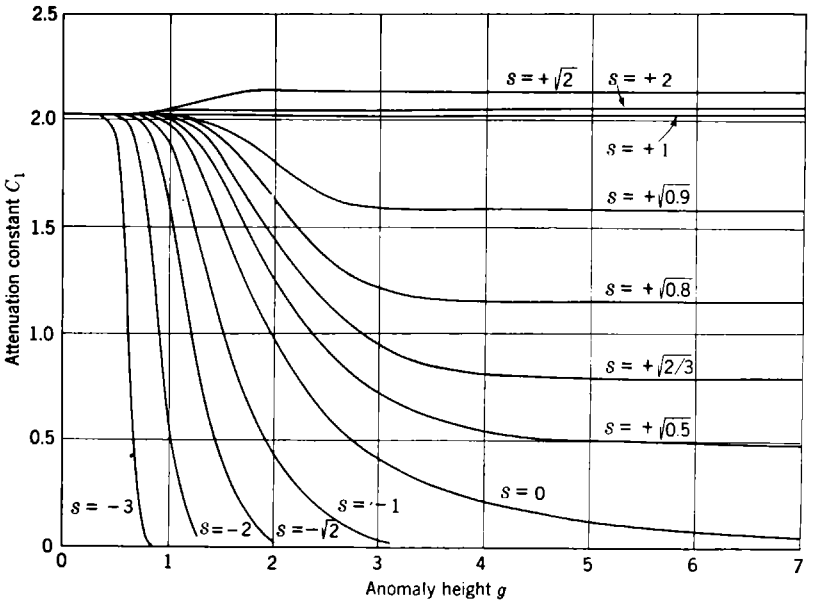


FIG. 2·30.—Attenuation constant of the first mode as a function of anomaly height for a bilinear N-profile.

was found that for large g the values were those given by the proper one of the various formulas for A_1.

The results are given in Fig. 2·30, which shows $C_1 = \mathrm{Im}(A_1)$ as a function of g for the various values of s, and in Fig. 2·31, which shows $B_1 = \mathrm{Re}(A_1)$ in the same way.[1]

In Fig. 2·32 the paths traversed by the characteristic value A_1 are shown in the complex A-plane. All of the curves start from the value ζ_1 for $g = 0$, and at first they practically coincide, because only the first two terms of the series in Eq. (541) are important for small values of g. For negative s the curves have as asymptote the negative real axis, and for positive s they end at finite limiting values. Figure 2·32 also shows a

[1] Graphs of these functions on a larger scale are contained in W. H. Furry, RL Report No. 795. These graphs can be read to an accuracy of about 0.01, which is also about the accuracy of the computations.

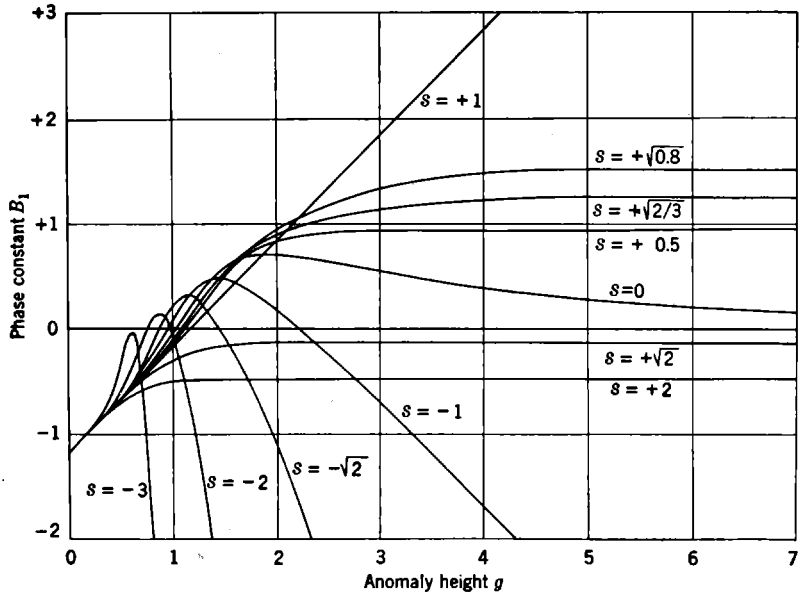


FIG. 2·31.—Phase constant of the first mode as a function of anomaly height for a bilinear N-profile.

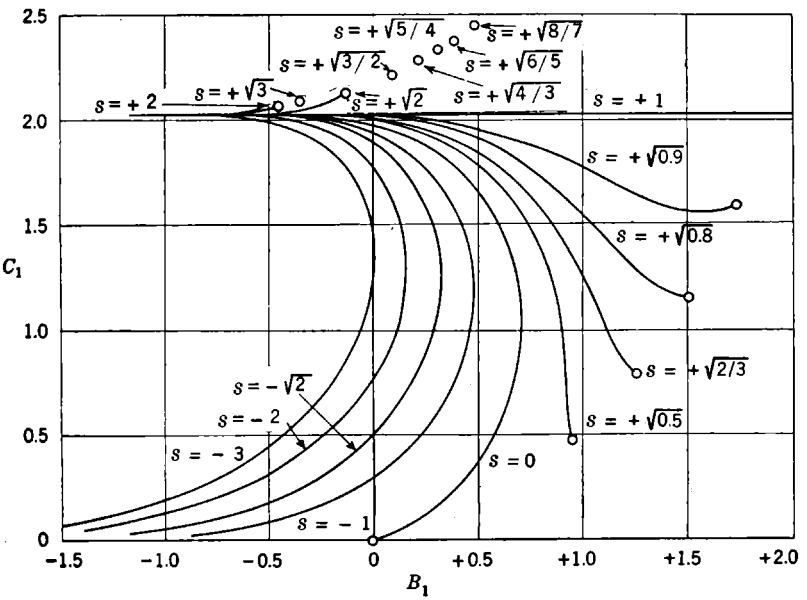


FIG. 2·32.—Characteristic values $A_1 = B_1 + iC_1$ for a bilinear N-profile.

number of other limit points which belong on Curve (1) of Fig. 2·29, labeled with the corresponding values of s.

These three figures also include lines for $s = 1$. For this one value of s, which corresponds to the mere introduction of a new convention into the description of the case of the linear distribution of modified index, the characteristic values A_m are given exactly by the series in Eq. (541), which reduces to its first two terms. It is physically evident that for values of s very near unity the behavior of A_1 should closely resemble that for $s = 1$, at least up to rather large values of g. For this reason computations of A_1 were carried out for $s = 1.02$ and for $s = \sqrt{0.98}$.

The results are shown in Fig. 2·33. For $s = 1.02$ the characteristic value A_1, which starts at ζ_1, for $g = 0$, does not approach the limiting

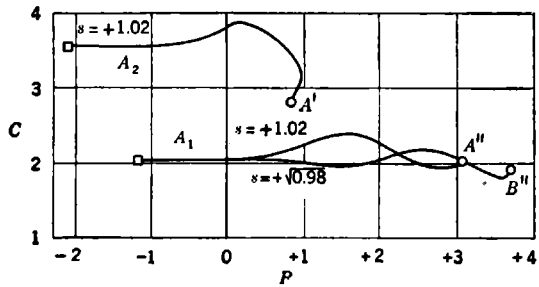


FIG. 2·33.—Characteristic values for s nearly unity. Squares give points for $g = 0$, and circles for $g \to \infty$.

point A', which lies on Curve (1) of Fig. 2·29. Instead, it continues to lie closer to the straight line $A = \zeta_1 + g$, which it would follow for $s = 1$, until it reaches the limiting point A'', which lies on Curve [1] of Fig. 2·29. For $s = \sqrt{0.98}$ the point representing A_1 follows a somewhat longer and straighter path than that for $s = 1.02$ and ends at the limiting point B'', which lies on Curve (2).

In order to learn the actual significance of the limiting point A' for $s = 1.02$, values of A were computed numerically, starting from the value A' for g infinite and continuing down to smaller values of g. The resulting locus is also shown in Fig. 2·33, and leads for $g = 0$ to the *second* standard value ζ_2.

Thus it is evident that for values of s very near unity the limiting values of A_1 do not lie on Curve (1) of Fig. 2·29 as they do for other positive values of s. From Eqs. (588) and (589) it can be seen that for any given numerically very small value of $s - 1$ values of $A^{(\infty)}(s)$ can be found such that their imaginary parts are not very different from $\mathrm{Im}(\zeta_1)$; moreover, the smaller $|s - 1|$ is made the larger the real parts of such values will be. This fits in exactly with the idea of the behavior of A_1 for $|s - 1| \ll 1$ which is suggested by the results for $s = 1.02$ and $s = \sqrt{0.98}$.

For $|s - 1| \ll 1$, A_1 is roughly equal to $\zeta_1 + g$ up to values of g, at which it approaches its limiting value; and as $|s - 1|$ is made smaller, the approximate relation $A_1 \approx \zeta_1 + g$ is more closely correct for given g and also the value of g at which it breaks down becomes larger.

The behavior of the first mode's characteristic value A_1 is, accordingly, acceptably stable against very small departures of the M-profile from the

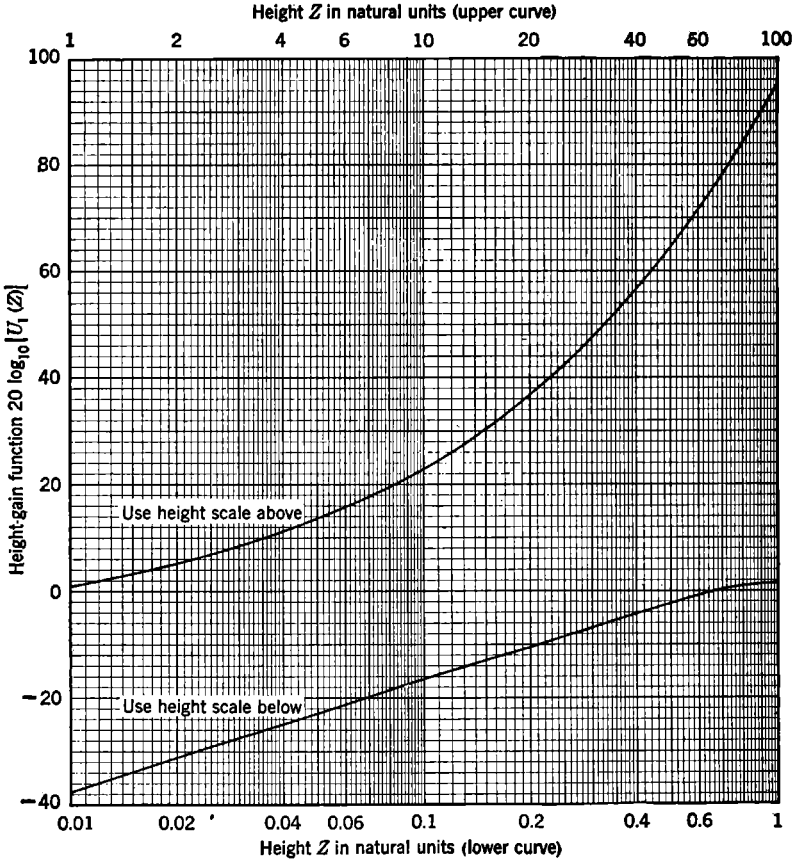


FIG. 2·34.—Height-gain function of the first mode for a bilinear N-profile with $s = -1$, $g = 1.37$, and $C_1 = 1.3$.

standard form. This is not necessarily the case for the other modes, as is illustrated by the curve of A_2 given in Fig. 2·33. The first mode here makes the dominant contribution to the field strength in the diffraction zone. The practical conclusion is that any observed M-profile which corresponds to a value of s nearly equal to unity should simply be treated as a case of standard refraction.

The characteristic value A_1 being known, the height-gain function $U_1(Z)$ can be computed exactly from the tables of $h_1(\zeta)$ and $h_2(\zeta)$ by means of Eqs. (503) and (504). The normalization can in most cases be accomplished by means of Eq. (512). For cases of positive s and fairly large g, however, A_1 will be so close to its limiting value that Eq. (515), with the last term omitted, should be used.

For negative s, both the attenuation constant $C_1 = \mathrm{Im}(A_1)$ and the behavior of $U_1(Z)$ depend essentially on the degree to which the mode is

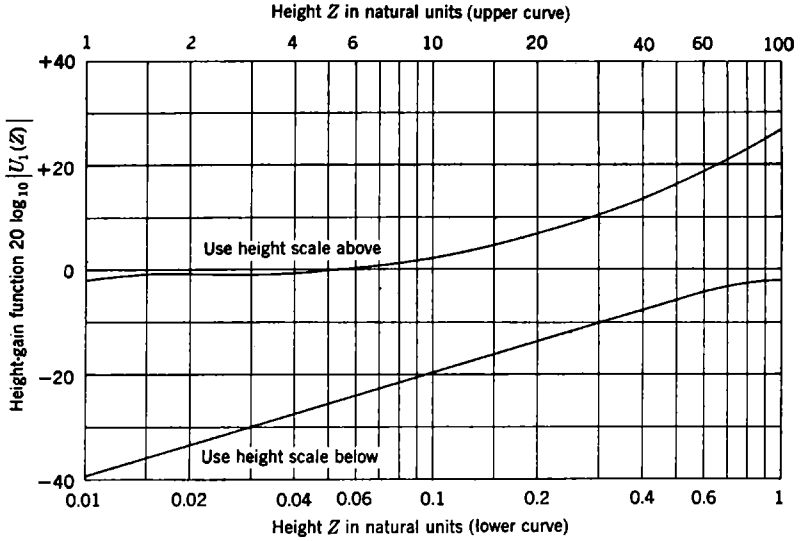


FIG. 2·35.—Height-gain function of the first mode for a bilinear N-profile with $s = -1$, $g = 1.93$, and $C_1 = 0.5$.

trapped. What this degree is can be learned at once from the value of C_1 read from Fig. 2·30. The qualitative effect of trapping on the shape of the height-gain function has been discussed in Sec. 2·8. Figures 2·34 through 2·37 show the actual curves of $20 \log_{10} |U_1(Z)|$ for four values of g with $s = -1$. In two of these cases the mode is untrapped, and in the other two it is trapped, according to the criterion given by the phase-integral method [Eq. (242)] which here becomes

$$\int_0^g Y^{1/2}\, dZ = \tfrac{2}{3}(-sg)^{3/2} > \tfrac{3}{4}\pi = 2.356. \tag{605}$$

The actual values of the integral for the three cases shown are

$$\left.\begin{aligned} &\text{Figure 2·34: } g = 1.37,\ \tfrac{2}{3}(-sg)^{3/2} = 1.07,\\ &\text{Figure 2·35: } g = 1.93,\ \tfrac{2}{3}(-sg)^{3/2} = 1.69,\\ &\text{Figure 2·36: } g = 2.68,\ \tfrac{2}{3}(-sg)^{3/2} = 2.93,\\ &\text{Figure 2·37: } g = 4.34,\ \tfrac{2}{3}(-sg)^{3/2} = 6.03. \end{aligned}\right\} \tag{606}$$

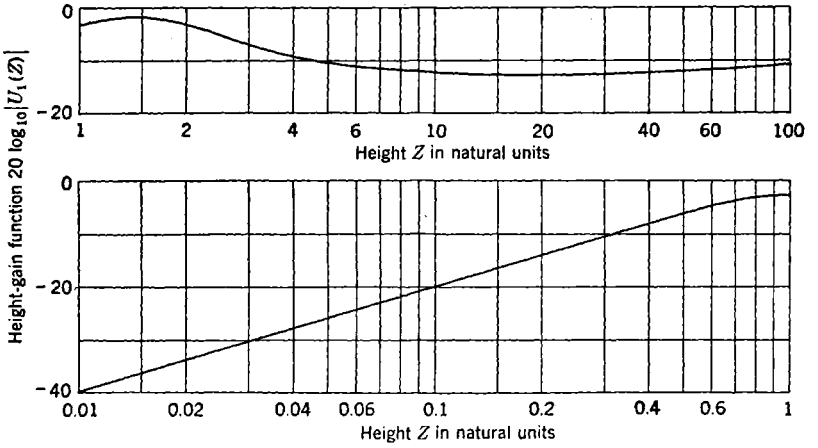


FIG. 2·36.—Height-gain function of the first mode for a bilinear N-profile with $s = -1$, $g = 2.68$, and $C_1 = 0.1$.

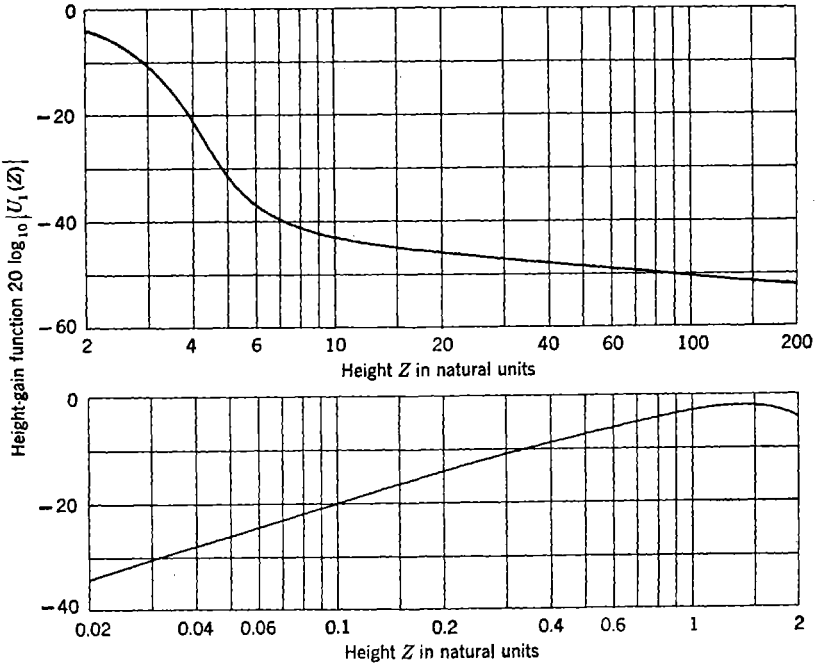


FIG. 2·37.—Height-gain function of the first mode for a bilinear N-profile with $s = -1$, $g = 4.34$, and $C_1 = 0.000089$.

In Fig. 2·35, even though there is no trapping according to Eq. (605), the value of 20 $\log_{10} |U_1|$ has a flat minimum at the top of the duct. In Fig. 2·36, which represents a case of very weak trapping, the minimum has shifted to greater heights and becomes much broader; even at $Z = 100$, $|U_1|$ has not regained the value that it had at the top of the duct. Figure 2·37 shows a case of strong trapping; the decrease in $|U_1|$ above the duct is very pronounced and for practical purposes continues indefinitely, ending actually near $Z = 3 \times 10^7$. For this case another

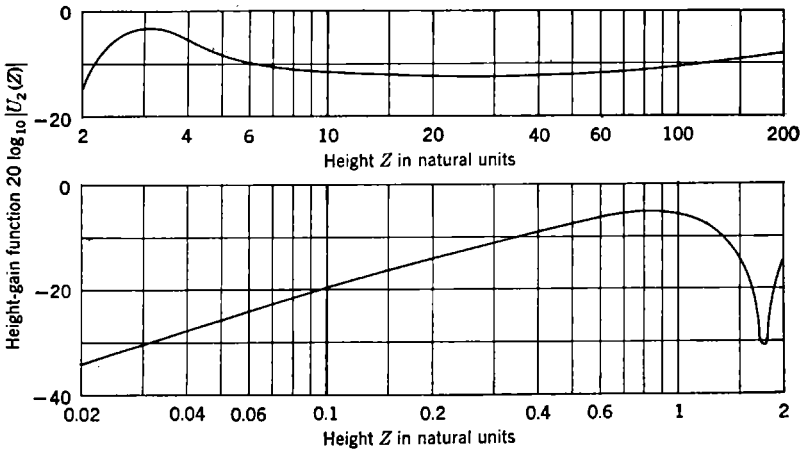


FIG. 2·38.—Height-gain function of the second mode for a bilinear N-profile with $s = -1$, $g = 4.34$, and $C_2 = 0.1$.

mode is trapped according to the phase-integral criterion, which requires $\frac{2}{3}(-sg)^{3/2} > 5.50$ for two modes to be trapped. The characteristic function of this second trapped mode is shown in Fig. 2·38. The sharp minimum at $Z = 1.78$ corresponds to the node of the curve of $\mathrm{Re}[U_{(2)}]$; the small imaginary part, which does not vanish at this point, becomes extremely small when the mode is more strongly trapped.

The general behavior of trapped modes illustrated here is not, of course, a special feature of this model but will appear in trapped modes for other shapes of the M-profile as well.

For positive values of s, qualitative arguments such as those given in the explanation of the Gamow phase-integral method can be used to predict the general behavior of the function U_1. The most important factor affecting this behavior is the fact that $B_1 = \mathrm{Re}(A_1)$ does not increase steadily with g but approaches a limiting value. The qualitative situation in regard to the first modes of two substandard layers of different thicknesses, but the same value of s, is shown in Fig. 2·39. It is supposed that in both cases A_1 is very near its limiting value, $-B_1$ therefore

having about the same value in the two cases. It is seen that the thick nonstandard layer acts as a formidable "barrier" between the upper region and the earth's surface. In both cases the solution in which $|U|$ decreases downward predominates throughout most of the thickness of the barrier; the term that increases as Z decreases becomes appreciable only near the earth's surface and makes possible the fulfillment of the boundary condition $U_1(0) = 0$. The solutions in the two cases differ practically only in the size of this term, and U_1 has almost precisely the same values near and above the top of the barrier in both cases. The contribution to the normalization integral of Eq. (483) from the region near the earth's surface is negligible.

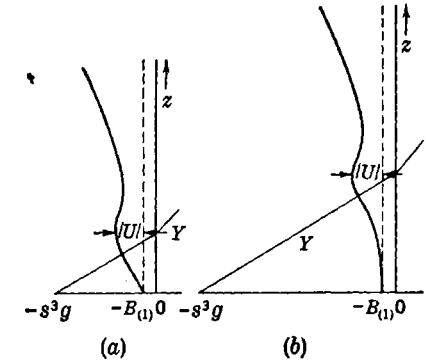


FIG. 2·39.—Substandard layers, (*a*) moderately thick and (*b*) thick with the same values of s. Horizontal scale reduced.

An increase of the thickness of an already fairly deep substandard layer does not change the horizontal attenuation of the mode. It transfers the region in which U has given values of order of magnitude unity or greater to a greater height, and it reduces the values of $|U|$ at small heights by a strong exponential factor. Because for given s, and g larger than some modest value, the height-gain function always has the same value at the top of the anomaly and decreases strongly downward from a point somewhat below this, it seems appropriate to speak of a "depth-loss" effect.

The same phenomenon occurs also, of course, in transitional cases for which $0 < s < 1$. It is seen from Fig. 2·30 or 2·32 that in substandard cases the range attenuation does not rise much above the standard value, but in transitional cases it can fall to much lower values. In both cases the depth-loss effect will often be very pronounced.

2·20. The Problem of Calculating Field Strength for the Bilinear Profile.—The bilinear modified-index profile was chosen for study for two reasons. (1) In many cases the measured profiles are approximated more closely by this model than by any other that could readily be considered, as the gradient dM/dz is frequently roughly constant throughout most of the thickness of the nonstandard layer and changes markedly only in a small region near the top of the layer (except for a very thin layer at the surface). (2) With this model the solution of the boundary-value problem can be formulated explicitly in terms of a few functions having simple properties. This circumstance was regarded as offering the possibility that it might be possible to obtain a fairly complete account of propagation

into the diffraction zone for this model without excessive numerical labor.[1]

In cases for which the possibility of trapping is the important feature, that is, for negative values of s, the bilinear model works reasonably well. It makes possible the fairly rapid determination of characteristic values and functions, both for trapped modes and for cases in which a mode is affected by the presence of the duct but cannot be said to be trapped.

Figures 2·34 to 2·36 illustrate cases in which a single mode makes the dominant contribution to the field strength, and hence Eq. (486*b*) can be used for practically all heights and distances. The effect of the factor $e^{-C_1 x}$ is to lower 20 $\log_{10} F$ by $8.68C_1$ db for each natural unit of range. In these three cases the advantage of the value of this factor for the first mode over its value for any other mode will in general be decisive.

A case in which no single mode is sufficient to determine the whole pattern is illustrated by Figs. 2·37 and 2·38. We may consider, for example, the case of a transmitter at a height of 1 natural unit. For receiver heights less than about 4 natural units and distances larger than 20 or 30 natural units, and also for greater receiver heights and distances greater than perhaps 60 natural units, the first mode suffices because of its much smaller attenuation. At receiver heights greater than 6 natural units, however, 20 $\log_{10} |U_1|$ is less than 20 $\log_{10} |U_{(2)}|$ by 30 or 40 db, and in spite of its larger attenuation the second mode will be dominant for distances less than 20 or 30 natural units. If both transmitter and receiver are at considerable heights, the range for which the second mode is dominant extends to 60 or more natural units. There are, of course, considerable regions in which contributions from both of these modes must be taken into account; for this purpose, one needs to know the phases arg U_1 and arg U_2, which are not shown in the figures but can readily be computed. In particular, when more than one mode is strongly trapped and both transmitter and receiver are at fairly small heights—"within the duct," as it is called—all of the trapped modes must be taken into account.

The calculation of field strengths for general cases in which s is positive is much less simple, and the account given here of these cases of the bilinear model must be regarded as far from complete. The main obstacle that appears consists of the appearance of characteristic values that approach finite limiting values as the thickness of the anomalous layer increases. For deep layers, the phenomenon of depth-loss, associated with the existence of these limits which the characteristic values approach, introduces very serious complications.

In discussing in a qualitative way the effects of depth-loss, one can use the rough rule that the square of the contribution of the mode in question

[1] The linear-exponential profile often represents the measured profiles considerably better, but the numerical calculations are more difficult (see Sec. 2·21).

to the field strength at a height of one natural unit, that is, for $Z = 1$, is affected by a depth-loss factor that is about equal to the distance $|A - A^{(\infty)}|$ measured in the complex A-plane. This follows from Eq. (512) and the fact that the approach to the limiting value $A^{(\infty)}$ is roughly of exponential type.[1] Application of this rule to Figs. 2·30 to 2·32 shows that for the first mode an anomaly of a few natural units in height produces a depth-loss effect of 20 db or more; the height required is especially small for the more pronouncedly substandard cases. It is accordingly clear that cases can easily occur in which the first mode can scarcely make the dominant contribution to the field strength. Some other mode, which may have greater horizontal attenuation but which does not suffer so strongly from depth-loss, will presumably make a larger contribution. This shift of the dominant role from one mode to another clearly complicates greatly the task of obtaining a complete account of the results to be expected with the bilinear model.

How one mode replaces another as the dominant one can be discussed more readily for the transitional case than for the substandard. For values of s appreciably less than unity—let us say $s < \frac{1}{5}$—the important limiting points are the $A_{(m)}^{(\infty)}(s)$, which lie on the lower branches (1), (2), (3) $\cdots$ of Fig. 2·29. It can be seen that these $A_{(m)}^{(\infty)}(s)$ all have roughly equal imaginary parts; for as a^3 is real, the first four terms of the right member of Eq. (578) give an imaginary part that is independent of m. Actually the imaginary part decreases slightly with increasing m. The real parts are fairly evenly spaced, with the intervals gradually decreasing along the sequence of points $A_{(m)}^{(\infty)}(s)$.

For $g = 0$, the propagation is standard and the first mode is the dominant one. For moderately small values of g, the attenuation falls considerably below standard and the contribution of this mode increases. For larger values of g, the depth-loss effect sharply reduces the contribution to the field strength at small or moderate heights. Meanwhile, however, as g increases, another characteristic value is approaching the limiting value $A_{(2)}^{(\infty)}(s)$ and thus is coming to have a rather small attenuation without, as yet, suffering as sharply from depth-loss as the first mode does. It accordingly makes a sizable contribution to the field strength and replaces the first mode as the dominant one. For still larger values of g, this mode is in turn strongly affected by depth-loss, and the mode whose characteristic value has the limit $A_{(3)}^{(\infty)}(s)$ replaces it as the dominant mode, and so on. It is natural to assign the terms second mode, third mode, etc., on the basis of this succession, which goes according to the subscript (m). The question of the correlation of this subscript with the subscript m of the formula for small g [Eq. (541)] is difficult and unimportant.

[1] Further discussion of this rule, which is admittedly very crude, is contained in W. H. Furry, RL Report No. 795, Appendix B.

When g has a value in a critical range such that the contributions of two modes to the field strength are of comparable magnitudes, interference between the two terms in the series of Eq. (486a) must, of course, be taken into account.

Some evidence as to the way in which this replacing of one dominant mode by another takes place is provided by the computations summarized graphically in Fig. 2·40. The two curves are both for $s = \sqrt{\frac{1}{2}}$ and show the approaches of the characteristic values of the first and second

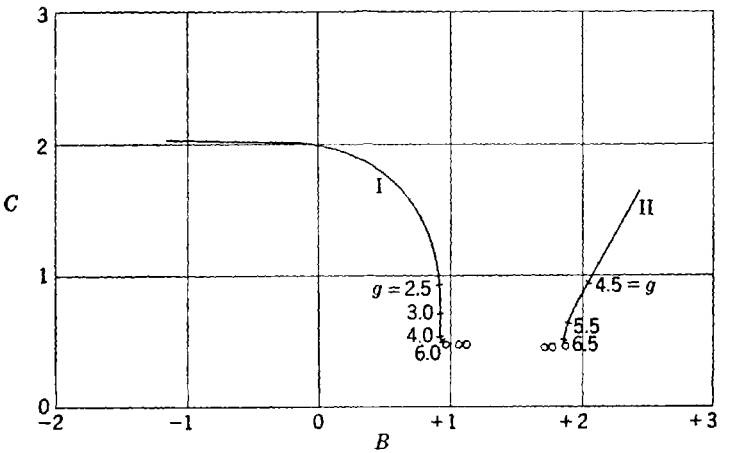


FIG. 2·40.—Behavior of characteristic values for large values of g and $s = +\sqrt{\frac{1}{2}}$. The circles are for $g = \infty$.

modes to their limiting values $A_{(1)}^{(\infty)}(\sqrt{\frac{1}{2}})$ and $A_{(2)}^{(\infty)}(\sqrt{\frac{1}{2}})$. By using $|A - A^{(\infty)}|$ as an estimate of the depth-loss factor, one can discuss the general effects of attenuation and depth-loss for these two modes. From Curve I it is seen that for $g = 2.5$ the horizontal attenuation of the first mode has fallen to less than half the standard value and at $g = 4$ it is about one-fourth of standard. At $g = 4$, however, the depth loss is about 10 db, and it grows stronger fairly rapidly, reaching 20 db at about $g = 6$. One sees also from Curve II that at $g = 4.5$ the second mode has less than half of standard attenuation and that the depth-loss does not amount to 10 db until g is 6.5 or more. Thus as g increases, the second mode replaces the first as the dominant mode but will in its turn, upon a further increase in g, fall a victim to depth-loss and give way to the third mode. The nature of the minimum of field strength between the ranges of values of g in which the two modes dominate depends on the particular values of the heights and the horizontal distance. Certainly in many cases the minimum will be rather marked.

In the case of substandard layers the situation is less clear. For s appreciably greater than unity—say $s \geqslant \frac{5}{4}$—the limiting values $A_{(2)}^{(\infty)}(s)$,

$A_{(3)}^{(\infty)}(s)$, $\cdots$, which lie on the upper branches of Curves (2), (3), $\cdots$ of Fig. 2·29, are probably approached about as rapidly as is the value $A_{(1)}^{(\infty)}(s)$. The modes whose characteristic values approach these limiting points may, accordingly, be supposed to have no great advantage over the first mode in regard to the onset of depth-loss. As they have much larger attenuations, it is reasonable to assume that in general they do not become dominant modes for any values of z. The modes that can become dominant and may be called "second mode," "third mode," and so on, presumably must have characteristic values that approach the points $A_{[1]}^{(\infty)}(s)$, $A_{[2]}^{(\infty)}(s)$, $\cdots$, which lie on the lower branches of Curves [1], [2], $\cdots$ of Fig. 2·29.

The limiting values in question lie on the parts of these branches that extend toward the asymptotes of rather small inclination, for the parts that extend toward the asymptote arg $A = \pi/3$ correspond to very small values of $s - 1$. For practical purposes one is concerned with values of s that are always smaller than 2. From Eqs. (587) and (562) and the appearance of Fig. 2·29, it is evident that the imaginary parts $\mathrm{Im}[A_{[m]}^{(\infty)}(s)]$ do not greatly exceed $\mathrm{Im}(\zeta_1)$ for small $[m]$ and decrease very appreciably with increasing $[m]$.

One source of complications affecting the propagation through thick substandard layers is the fact that the limiting values of the attenuation are not practically the same, as they are in transitional cases, but instead decrease markedly along the series of higher modes after perhaps showing an increase for the second mode as compared with the first. Another complicating feature of the situation is the fact that the real parts of the successive limits $A_{(1)}^{(\infty)}(s)$, $A_{[1]}^{(\infty)}(s)$, $A_{[2]}^{(\infty)}(s)$, $\cdots$ differ by rather large amounts; the difference $\mathrm{Re}[A_{[1]}^{(\infty)}(s)] - \mathrm{Re}[A_{(1)}^{(\infty)}(s)]$ is particularly large. If, as seems rather likely, the limiting points $A_{[m]}^{(\infty)}(s)$ are approached essentially from above, as is apparently the case with the limiting points $A_{(2)}^{(\infty)}(s)$, $A_{(3)}^{(\infty)}(s)$, $\cdots$ considered in discussing deep transitional layers, then these large differences between the real parts probably mean that the minima encountered, for values of g intermediate between those for which one mode is dominant and those for which the contribution of the next mode becomes appreciable, will be extremely pronounced. It may be, however, that the limiting points $A_{[m]}^{(\infty)}(s)$ are approached in the same way as $A_{(1)}^{(\infty)}(s)$, that is, essentially from the left. If this should be so, the transitions from one dominant mode to the next, as g increases, may not be accompanied by pronounced minima.

It is obvious that reliable predictions of the sort of behavior to be expected in transmission through deep substandard layers cannot be made without much more numerical evidence than has now been obtained. For transitional layers the general situation seems to be clearer, but the bulk of the detailed information remains to be filled in.

This rather distressing situation with the bilinear model may be due in part to the fact that for this model it has been possible to study analyt-

ically the whole variety of possible solutions. The main difficulties, however, seem to come from the presence of the finite limits that characteristic values approach as g becomes infinite, and it does not seem likely that this phenomenon will appear for other models. Its occurrence is presumably to be attributed to the effects of waves reflected from the point at which $Y'(Z)$ is discontinuous. The question thus arises as to whether or not the complications that are encountered may originate mainly in the use of an analytic formulation involving this discontinuity.

The simplest model having the principal qualitative characteristic of the bilinear distribution—that the gradient $Y'(Z)$ is constant throughout most of the anomalous layer, the change from nonstandard to standard gradient being localized in a small thickness at the top of the layer—is that in which the change of gradient occurs in an intermediate layer in which Y is a quadratic function of Z. This can be realized by setting

$$Y(Z) = Z - g, \qquad Z > g + \tau, \tag{607}$$

$$Y(Z) = Z - g + \frac{1 - s^3}{4\tau}(Z - g - \tau)^2, \qquad g - \tau < Z < g + \tau, \tag{608}$$

$$Y(Z) = s^3(Z - g), \qquad Z < g - \tau; \tag{609}$$

$Y(Z)$ and $Y'(Z)$ are then everywhere continuous, and the change of gradient takes place in a layer of thickness 2τ. As in Eqs. (501) to (503), the solutions of the differential equation for the regions above and below the transitional layer are

$$U_m(Z) = K_m h_2(Z - g + A_m), \qquad Z > g + \tau, \tag{610}$$

$$U_m(Z) = y_\beta(sZ - sg + s^{-2}A_m), \qquad Z < g - \tau. \tag{611}$$

The solution in the transitional layer $g - \tau < Z < g + \tau$ has been obtained in terms of power series[1] in τ. The requirements that $U_m(Z)$ and $U_m'(Z)$ be everywhere continuous and that $U_m(0) = 0$ suffice to determine the characteristic values A_m.

Because of the use of power series, the computations were feasible only for modest values of the transition thickness 2τ and involved considerable labor. The method of mapping was used exclusively. The results are shown in Table 2·4. They indicate that the results are not sensitive to changes in the precise shape of the joint. The bilinear model may accordingly be regarded as the appropriate one to use in obtaining theoretical information for application to the whole class of actual profiles in which dN/dz is roughly constant throughout almost the whole thickness of the nonstandard layer.

It appears that analytic features of the shape of the modified-index profile do not in themselves have any strong effect on the results, at least

[1] Details in Furry, *op. cit.*, Appendix F.

TABLE 2·4.—EFFECT OF ROUNDING THE JOINT IN THE BILINEAR MODEL BY INSERTING A PARABOLIC SEGMENT

s	g	2τ	A_1
$\sqrt{\frac{1}{2}}$	∞	0	$0.96+0.48i$
$\sqrt{\frac{1}{2}}$	∞	0.283	$0.96+0.48i$
$\sqrt{\frac{1}{2}}$	∞	0.849	$0.96+0.50i$
$\sqrt{\frac{1}{2}}$	2	0	$0.84+1.28i$
$\sqrt{\frac{1}{2}}$	2	0.849	$0.79+1.26i$
2	∞	0	$-0.47+2.06i$
2	∞	0.400	$-0.51+2.09i$

for the dominant mode. Although the analytic structure of the curve is all-important in determining the methods that must be used, the results depend essentially only on the numerical values of N for real z. This statement is confirmed by the evidence obtained by considering a modification of the bilinear model. If it were not true, any theoretical approach to the problem of propagation would be of little actual use, for in practice only approximate numerical values of N can be known.

NONLINEAR MODIFIED-INDEX PROFILES

BY JOHN E. FREEHAFER

2·21. The Linear-exponential and Power-law Profiles.—Although N-profiles composed of one or more linear segments provide attractive idealizations because they lead to characteristic value problems that may be discussed in terms of the tabulated functions h_1 and h_2, it is by no means necessary for analytical reasons to exclude from consideration models not having this property. Two such models have been extensively studied: the linear-exponential and power-law profiles.

In the linear-exponential profile the modified index is given by

$$\frac{N^2}{n_0^2} - 1 = qz + \alpha' e^{-\gamma' z}, \tag{612}$$

where α' and γ' are parameters that determine its detailed form. The second term on the right-hand side of Eq. (612) may be regarded as a perturbation term that represents the departure from a linear profile. If q has the standard value, as will be assumed throughout this section, then the surface layer is superstandard or substandard according to whether α' is positive or negative. As the parameter γ' is always taken as positive, it follows that in either case the perturbation term approaches zero with increasing height and the N-profile approaches the standard profile asymptotically.

The linear-exponential profile was introduced by T. Pearcey[1] in connection with a ray-tracing investigation of surface ducts. The properties of the modes associated with this model of the atmosphere have been studied principally by Hartree,[2] Pearcey, and Pekeris.[3]

The power-law profile has an N-distribution defined by

$$\frac{N^2}{n_0^2} - 1 = qz - \frac{qd}{n}\left(\frac{z}{d}\right)^n. \tag{613}$$

The parameter d is always taken as positive, and the parameter n is limited to values between zero and unity. The N-profile in this case is horizontal at the surface regardless of the values of the parameters and has a surface inversion extending to height d. It therefore always provides a surface duct. Although the ratio of the perturbation term to the linear term in the case of the power law approaches zero as z becomes infinite, the perturbation term itself, unlike that of the linear-exponential profile, becomes infinite.

The power-law profile was proposed by H. G. Booker[4] originally for the special case for which $n = \frac{1}{2}$, because for this form of index distribution it is possible to evaluate analytically the integrals encountered in phase-integral methods. He later proposed the generalized form for reasons based on a special theory of the structure of the lower atmosphere.[5] The power law was the basis of an extensive cooperative program in the United Kingdom during the war involving groups at Telecommunications Research Establishment, Air Defense Research and Development Establishment, the University of Manchester, and Cambridge University.

When either the power-law or linear-exponential index distribution is substituted into the second-order propagation equation and the height in natural units is introduced, the resulting equation can be written:

$$\frac{d^2u}{dZ^2} + [Y(Z) + A]u = 0. \tag{614}$$

Here $Y(Z)$ is a two-parameter function of Z given by

$$Y(Z) = Z + \alpha e^{-i\gamma Z} \tag{615}$$

[1] "The Calculation of Field Strength near the Surface of the Earth under Particular Conditions of Anomalous Propagation," ADRDE Research Report No. 203, October 1943.

[2] D. R. Hartree, P. Nicholson, N. Eyres, J. Howlett, and T. Pearcey, "Evaluation of the Solution of the Wave Equation for a Stratified Medium," (I), ADRDE, M.R. No. 47, May 1944; (II), RRDE Research Report No. 279, March, 1945. See also the article by D. R. Hartree, *et al.*, in *Meteorological Factors in Radio-Wave Propagation*, published by the Physical Society and the Royal Meteorological Society, London (1946).

[3] C. L. Pekeris, *Jour. Applied Phys.*, **17**, 678 (1946); *Proc. IRE*, **35**, 453 (1947).

[4] "The Theory of Anomalous Propagation in the Troposphere and its Relation to Wave-guides and Diffraction," *Meteorological Factors in Radio-Wave Propagation.*

[5] "Elements of Radio Meteorological Forecasting," TRE Report No. T1621, February 1944; "Application of Diffusion Theory to Radio Refraction Caused by Advection," TRE Report No. T1647, April 1944.

for the linear-exponential profile and by

$$Y(Z) = Z - \varepsilon Z^n \tag{616}$$

for the power law. The parameters are α and γ in Eq. (615) and ε and n in Eq. (616).

Because in the case of the power law there is always a surface duct whose thickness is d, it is also possible to reduce the propagation equation to dimensionless form by using d as the unit of height. Thus if $s = z/d$ be introduced as independent variable, the propagation equation becomes

$$\frac{d^2u}{ds^2} + \left[G\left(s - \frac{s^n}{n}\right) + A'\right]u = 0, \tag{617}$$

where G and n are the parameters and A' denotes the dimensionless separation constant. Equation (617) is of the same form as Eq. (614) with

$$Y(Z) = G\left(Z - \frac{Z^n}{n}\right). \tag{618}$$

For some purposes, it is more convenient to discuss the power law by using Eq. (618) rather than Eq. (616).

The problem is threefold: (1) the determination of the characteristic values of A as functions of the two parameters appearing in $Y(Z)$ for which the solution of Eq. (614) which vanishes at $Z = 0$ represents an upward-traveling wave for large Z, (2) the evaluation of the integral necessary for normalizing the characteristic functions associated with the characteristic values, and (3) the calculation of the characteristic functions for real values of Z.

The analytic expression of the condition that a solution represent an upward-traveling wave needs to be carefully examined. Because the characteristic values of leaking modes are complex and enter into the height-gain functions in such a way that the upward-traveling wave dominates the downward-traveling wave as Z becomes large, it is evident that any linear combination of the two waves becomes, at sufficient height, indistinguishable from the pure upward-traveling solution. Careful analysis shows that when $Y(Z)$ is analytic, there exists an arc of the circle at infinity along which the condition $u \to 0$, together with the condition $u(0) = 0$, defines a set of discrete characteristic values of A. It turns out that these are the values sought and that the solution for which $u(Z) \to 0$ as $Z \to \infty$ along the arc mentioned above is in fact the solution that represents the pure upward-traveling wave. For the power law the arc of the circle along which $u(Z) \to 0$ extends from a distance $\mathrm{Im}(A)/G$ below the real axis to a point at about arg $Z = -2\pi/3$. The characteristic values of A are independent of which point on this arc the variable Z approaches.

Similarly, the integral upon which the normalization of the characteristic functions depends can be expressed by $\int_0^{\infty \exp i\theta} u^2 dz$, where the path of integration extends from the origin in such a direction θ that the integral converges. The permissible values of θ are such that $u \to 0$ as $Z \to \infty$ along the arc previously mentioned. The value of the integral is independent of the precise point of the arc that Z approaches.

The characteristic value problem of radio wave propagation, when expressed in terms of the vanishing of the functions at the origin and at infinity, is formally similar to problems encountered in quantum mechanics, and the means used to solve it are similar. The only innovation necessary is to modify suitably the formal procedures in order that they may be applied when both Z and A are complex. The principal methods that have been used are

1. Mechanical integration of the differential equation by means of the Manchester and Cambridge (England) differential analyzers.[1]
2. Numerical integration of the associated Riccati equation.[1]
3. Perturbation calculations starting with the standard N-profile.[2]
4. Phase-integral methods.[3]
5. Variational methods.[4]

The differential analyzer in the hands of Professor Hartree's group has proved to be a powerful tool for the solution of problems of this type. The exact method of using the analyzer depends upon the degree to which the mode leaks. In the case of strongly trapped modes, the analyzer is set up to plot the path of u in the complex plane as Z increases along the real axis, the initial condition being $u(0) = 0$. The value of A is adjusted by trial and error until the path of u in the complex plane becomes circular as Z becomes large. If the path is elliptical, suitable adjustments must be made in both the real and imaginary parts of A and checked by running out another solution on the machine. The effective determination of characteristic values by a trial-and-error method of this sort depends upon devising means for analyzing the behavior of solutions for values of A near the characteristic values in order that the magnitude of the correction to be applied to A may be reliably estimated from the behavior of u. In the case of strongly leaking modes the procedure is varied to permit integration along a path lying in the complex Z-plane.

[1] D. R. Hartree, *et al*, in *Meteorological Factors in Radio-Wave Propagation.*

[2] C. L. Pekeris, *Jour. Applied Phys.*, **17**, 678 (1946).

[3] *Ibid.*, p. 1108.

[4] G. MacFarlane, "Variational Method for Determining Eigenvalues of Wave Equation of Anomalous Propagation," TRE Report No. T1756, November 1944; also *Proc. Cambridge Phil. Soc.*, **43**, 213 (1947); C. L. Pekeris and S. Ament, "Characteristic Values of the First Normal Mode in the Problem of Propagation of Microwaves through an Atmosphere with a Linear-exponential Modified Index of Refraction," *Phil. Mag.*, Ser. 7, **38**, 801 (1947).

Two paths have been used. One path is a ray in the fourth quadrant, and the other is a step-type contour formed by three connected linear segments, the first starting at the origin and lying along the positive real axis, the second lying parallel to the negative axis of imaginaries, and the third lying below and parallel to the positive real axis. In this case the criterion for a characteristic value is that $u(Z) \to 0$ as $Z \to \infty$.

The Riccati equation is the differential equation that expresses the dependency of u'/u on Z. For modes that are not strongly leaking, u'/u

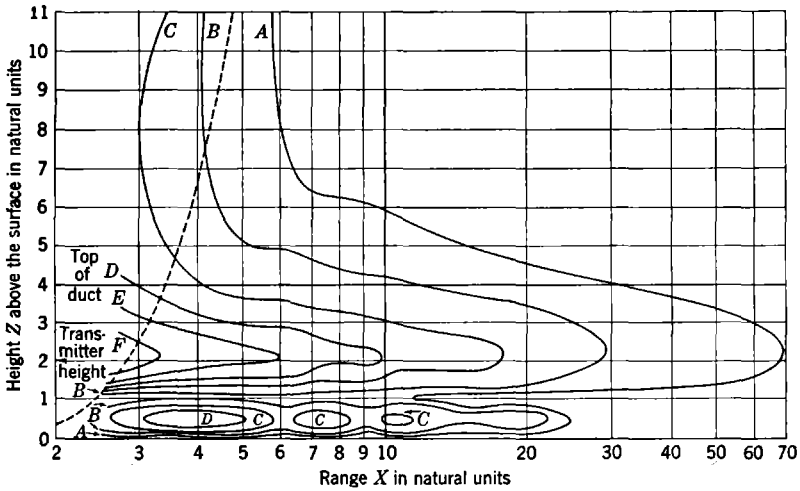
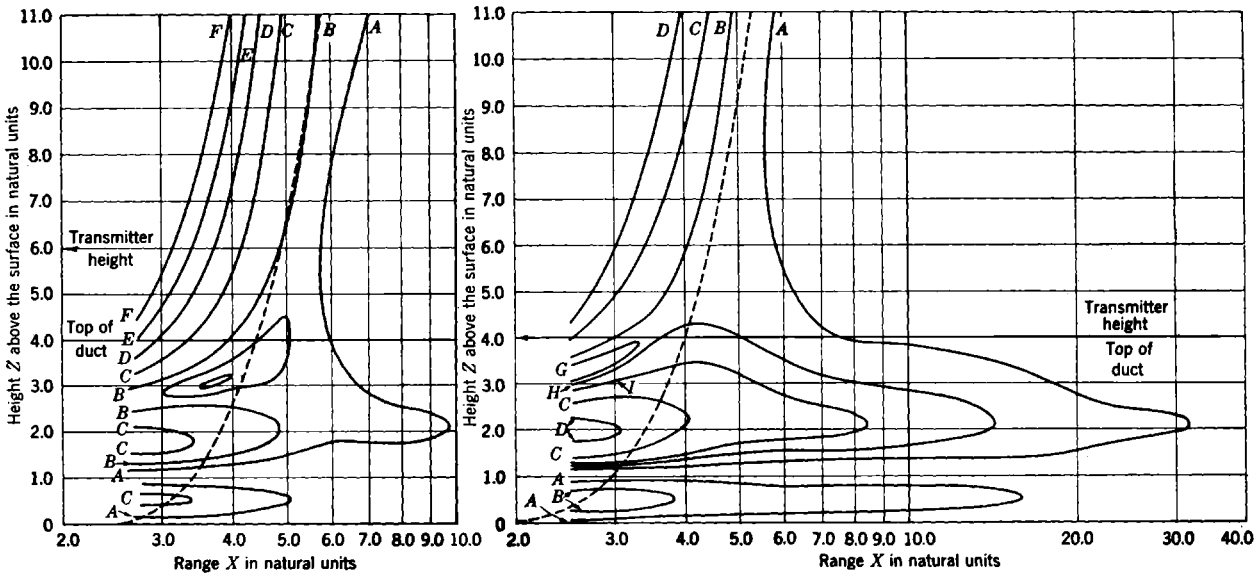


FIG. 2.41.—Contours of constant field strength in the presence of a surface duct with the transmitter within the duct. The modified index is given by $N^2 - n_0^2 = (q/k)_0^{2/3}(Z + 20e^{-0.0356Z})$. The broken line is the tangent ray. (*After Pearcey.*)

is known for large Z for the outward traveling wave. The Riccati equation may then be integrated inward along the axis of reals by numerical means and at an intermediate point compared with the result of integrating Eq. (614) outward by an iteration process. The value of A is adjusted until the values of u'/u at the intermediate point obtained by both outward and inward integration agree. For the case of strongly leaking modes, the integration is carried out along a suitable ray in the complex plane.

It is beyond the scope of this discussion to describe in detail the techniques available for solving the characteristic value problem by mechanical and numerical integration. The two procedures that have been mentioned briefly are intended merely to indicate in a very general way the method of attack that has been used.

Perturbation methods lead to an interesting formulation of the problem. From a practical standpoint, however, their use is limited to strongly leaking modes that differ only slightly from those of the standard at-



(*a*) Transmitter above the duct. (*b*) Transmitter at the top of the duct.

FIG. 2·42.—Contours of constant field strength in the presence of the surface duct shown in Fig. 2·41. (*After Pearcey.*)

mosphere. The application of the variational method is likewise somewhat limited, as the method becomes unwieldy for other than the first mode.

When the characteristic values and normalization factors have been obtained, it is still necessary to calculate the complex values of u for real values of Z unless these values have been obtained as by-products of the procedure used to solve the characteristic value problem. To present the data in the most useful form, it is desirable to draw contours of constant field strength in the vertical plane containing the transmitter. The first step in this process is to combine the various modes in proper phase for a fixed range and variable height. When a number of vertical sections have been obtained in this manner, a contour map may be constructed by interpolation.

Three field patterns obtained by Pearcey for the linear exponential atmosphere defined by $\alpha = 20.0$ and $\gamma = 0.6356$ are reproduced in Figs. 2.41 and 2.42. For this choice of parameters there is a duct extending to a height of 4 natural units. The range is measured in the natural units defined in Sec. 2·10. The contours A, B, C, D, E, and F are for values of F/X proportional respectively to 1, 2, 4, 6, 8, and 10. The set of three figures shows how the pattern varies as the transmitter height is increased. In Fig. 2·41 the transmitter is at a height of 2 natural units and is therefore at the center of the duct. Under these conditions contour A extends to a maximum range of 70 natural units. When the transmitter is placed at the top of the duct, the range of contour A is reduced to about 32 natural units, as shown in Fig. (2·42*b*). Finally, when the transmitter is raised above the duct to a height of 6 natural units, contour A is shrunk to a maximum range of only 10 units, as shown in Fig. (2·42*a*). The closed contours representing peaks in the field-strength pattern of Fig. 2·41 are produced by interference between the first two modes in the region within which they are of comparable magnitude. At ranges beyond about 25 units the first mode dominates and the pattern is free of interference effects.[1]

[1] Further examples of this type of diagram are given by C. L. Pekeris, *Proc. IRE*, **35**, 453 (1947).

CHAPTER 3

METEOROLOGY OF THE REFRACTION PROBLEM

By Richard A. Craig, Isadore Katz, R. B. Montgomery, and Pearl J. Rubenstein

The purpose of this chapter is to describe the types of distribution of refractive index in the atmosphere and to discuss them in terms of the processes producing them and of the circumstances in which they occur. The technical meteorological terms are explained as they are introduced, and the presentation is intended to be understandable to anyone concerned with microwave propagation. Because the subject of this chapter is essentially meteorological, however, the primary intention is to make the present knowledge of the subject available to persons having the meteorological training and experience necessary to apply it and to develop it further.

Microwave propagation has focused attention on meteorological phenomena previously neglected. Although conditions at the surface of the earth have always been of direct concern, and although the wealth of upper-air data of recent years has made common knowledge of the gross vertical structure of the atmosphere to heights of 8 to 10 miles, comparatively little attention has heretofore been devoted to the details of the vertical structure of the surface layer up to about 2000 ft—a layer that is of primary importance in microwave propagation. Knowledge of this layer, although it has increased during the war, remains scant and inadequate, and much promising investigation remains to be undertaken.

This chapter begins with a discussion of certain meteorological variables and processes and their relation to refractive index.[1] A discussion of the vertical variation of refractive index follows. As secondary material the horizontal and time variations of this quantity are also considered. The chapter concludes with a discussion of new observational instruments and techniques.

HUMIDITY AND REFRACTIVE INDEX

The refractive index of air depends on pressure, temperature, and humidity in a manner that will be described presently. First of all, it is

[1] For further background in meteorology a useful book is H. R. Byers, *General Meteorology*, McGraw-Hill, New York, 1944.

suitable to discuss the subject of humidity because there are diverse quantitative representations of it.

3·1. Vapor Pressure and Saturated Vapor.[1]—The subject of humidity may suitably be introduced by mention of certain variables that depend on the amount of water vapor present and that lend themselves to direct measurement[2] by common methods under conditions found in nature.

1. The *water-vapor density* may be determined by measuring the change in weight of a hygroscopic material through which a known volume of air is passed; this method is not adapted to rapid use.
2. The dimensions of some solid substances can be calibrated to indicate *relative humidity*; a strand of human hair is in common use. The hair hygrometer is an extremely simple and popular instrument, but it does not afford quick response or high accuracy. Another method for measuring relative humidity is based on the electrical conductivity of a specially prepared surface; the conductivity usually depends not only on relative humidity but also to some extent on temperature.
3. *Wet-bulb temperature* and air temperature are both indicated directly by a ventilated psychrometer, an instrument consisting of two thermometers, one with its surface wet and the other with its surface dry.
4. *Dew point* is indicated directly by a dew-point hygrometer, a thermometer that can be cooled just to the point where it begins to "sweat."

Other humidity parameters in common meteorological use, although not directly measured, are *vapor pressure* and quantities giving the concentration of water vapor. All these will be defined, and their relationship explained.

Equation of State for Water Vapor.—The relation among pressure, density, and temperature of water vapor in the atmosphere is very nearly independent of the dry-air components present and conforms closely to the equation of state for a perfect gas. If e is the partial pressure of the water vapor or *vapor pressure*, if ρ_v is the *water-vapor density* (or, as it is frequently called, *absolute humidity*), and if T is the absolute temperature, the relation is

$$e = \frac{R}{m_v} \rho_v T. \tag{1}$$

The symbol R is the universal gas constant, and m_v is the molecular weight of water vapor. Their values are listed in Sec. 3·33.

[1] By R. B. Montgomery.

[2] A discussion of methods for measuring humidity is contained in W. E. K. Middleton, *Meteorological Instruments*, 2d ed., University of Toronto Press, 1943, Chap. 4, "The Measurement of Atmospheric Humidity."

Saturated Vapor.—A vapor in equilibrium with a plane surface of the liquid or solid phase of the same substance in pure form is defined as *saturated.* The pressure and density of the vapor depend on the temperature alone, and usually, as is the case in all common problems involving atmospheric water vapor, the influence of the other gases present may be neglected. Thus, for any substance there is a function of temperature alone called *saturation vapor pressure* which is equal to the pressure of saturated vapor at the same temperature.

At temperatures below freezing, saturation with respect to the solid phase rather than the liquid phase is the natural choice. In the atmosphere, however, droplets of supercooled water exist so frequently in clouds that for some meteorological purposes there is frequent use of saturation with respect to liquid water at all temperatures.

There is a useful thermodynamic relation between the variation of saturation vapor pressure e_s with temperature and the latent heat L for the change from the solid or liquid phase to the vapor phase. If α_v is the specific volume (reciprocal of density) of the vapor phase and α_l is the specific volume of the solid or liquid phase, the relation known as *Clapeyron's equation* is

$$\frac{de_s}{dT} = \frac{L}{T(\alpha_v - \alpha_l)}. \tag{2}$$

As the specific volume of atmospheric water vapor is of the order of 10^5 times the specific volume of ice or water, the latter may be neglected in comparison. Substituting from the equation of state [Eq. (1)] and using $\alpha\rho = 1$ give

$$\frac{de_s}{dT} = \frac{m_v}{R}\frac{Le_s}{T^2}, \tag{3}$$

which will be used later.

A convenient tabulation of saturation vapor pressure over ice for temperatures below freezing and over water for temperatures above freezing is given in Tables 78 and 79 of the *Smithsonian Meteorological Tables.*[1] Values with respect to supercooled water are given by Harrison.[2] In both cases the vapor pressure is tabulated for every 0.1°C and is given in millibars, the pressure unit in common meteorological usage;

$$1 \text{ mb} \equiv 10^{-3} \text{ bar} \equiv 10^3 \text{ dyne/cm}^2 \equiv 10^2 \text{ newton/m}^2.$$

One *atmosphere* equals 1013.25 mb.

[1] *Smithsonian Meteorological Tables,* 5th rev. ed., Vol. 86, Smithsonian Miscellaneous Collections, Washington, D.C., 1939.

[2] L. P. Harrison, "Tables (in Millibars) of the 'Pressure of Saturated Aqueous Vapor over Water' at Temperatures from 0° to −50°C," *Monthly Weather Rev.,* **62,** 247–248 (1934).

The relation between saturation vapor pressure and temperature may be shown graphically by means of *saturation curves* drawn on a diagram that has vapor pressure as abscissa and temperature as ordinate. This diagram, which is called a *characteristic diagram* for reasons that will be explained in Sec. 3·8, is illustrated schematically in Fig. 3·1. The saturation curve for water continues without break through the freezing point, below which the water is supercooled. The saturation curve for ice branches from it at a finite angle.

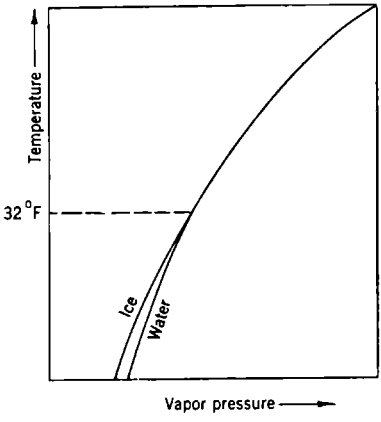


FIG. 3·1.—Schematic representation of characteristic diagram.

From this basic arrangement has been developed the large-scale characteristic diagram, shown by two separate graphs in folded plates in the rear pocket, Figs. *H* and *I*; each graph covers different ranges of the basic coordinates, vapor pressure and temperature. Above 28°F, the freezing point of most sea water, the large-scale diagram shows a vapor-pressure curve for sea water. This curve is constructed for the salinity of 35 per thousand, and is 0.9812 of the saturation vapor pressure over fresh water. Other coordinates have been added, as will be pointed out upon introduction of the parameters that they represent in the following three sections. In Sec. 3·8 will be found a complete descriptive summary of this information, together with a discussion of the considerations that make the characteristic diagram useful to the radio meteorologist.

Relative Humidity.—This quantity is defined as the ratio of vapor density to saturation vapor density. Because the vapor may be regarded as a perfect gas, relative humidity equals the ratio of vapor pressure to saturation vapor pressure. The relative humidity is

$$f \equiv \frac{\rho_v}{\rho_{vs}} = \frac{e}{e_s};$$

it is commonly expressed in per cent. The relative humidity of saturated vapor is 100 per cent, whereas vapor in equilibrium with sea water has a relative humidity of 98 per cent.

3·2. Water-vapor Concentration.[1]—The humidity variables discussed above, namely, vapor density and pressure and relative humidity, depend essentially on the water vapor only. In contrast, the amount of dry air present affects the water-vapor concentration, the dew point, and the

[1] By R. B. Montgomery.

wet-bulb temperature. The concentration is expressed by either of two quantities, specific humidity or mixing ratio.

Specific humidity is the ratio of water-vapor density to total density and is usually designated by q. If ρ and p are the total density and pressure respectively and m is the mean molecular weight of the moist air,

$$q = \frac{\rho_v}{\rho} = \frac{m_v}{m}\frac{e}{p}, \tag{4}$$

which follows from the equations of state for the water vapor and for the moist air. In order to avoid use of a variable mean molecular weight, it is convenient to introduce the mean molecular weight of dry air m_d, the density of the dry components of the air ρ_d, and the partial pressure of the dry air p_d. Hence, because $\rho = \rho_d + \rho_v$ and $p = p_d + e$,

$$q = \frac{\rho_v}{\rho_d + \rho_v} = \frac{m_v}{m_d}\frac{e}{p_d + \left(\frac{m_v}{m_d}\right)e} = \frac{m_v}{m_d}\frac{e}{p - \left(1 - \frac{m_v}{m_d}\right)e}. \tag{5}$$

With this equation the specific humidity can be computed from vapor pressure and total pressure. The value of the constant m_v/m_d is 0.622.

Mixing ratio is the ratio of the water-vapor density to the density of the dry components of the air and is usually designated by w. Hence

$$w = \frac{\rho_v}{\rho_d} = \frac{m_v}{m_d}\frac{e}{p_d} = \frac{m_v}{m_d}\frac{e}{p - e}. \tag{6}$$

A solution of Eqs. (5) and (6) leads to the conversion equations

$$q = \frac{w}{1 + w}, \qquad w = \frac{q}{1 - q}.$$

The results of upper-air soundings taken by radiosonde are customarily reported in terms of mixing ratio rather than in terms of specific humidity.

The water-vapor concentration in the atmosphere is so small, of the order of 10^{-2} and only in exceptional cases exceeding 4×10^{-2}, that specific humidity and mixing ratio are always nearly the same, and for practical purposes they are interchangeable in most cases. It is convenient to express both of these quantities in "parts per thousand" or "per mille"; for example, an observed value might be 12.8 per mille = 12.8×10^{-3}. The expression "grams per kilogram" also is frequently used.

On the characteristic diagram shown in Figs. *H* and *I* in the rear pocket there is a secondary abscissa scale, giving the equivalent mixing ratio for the standard total pressure of 1000 mb. *Saturation specific humidity* and *saturation mixing ratio* are the values for which the vapor would be saturated at the given temperature and total pressure. For 1000 mb the saturation mixing ratio is given by the curve on the characteristic diagram.

Water-vapor concentration is a *conservative* property of air to the extent that it varies only when the composition is changed by evaporation (positive or negative) or diffusion (molecular or eddy). In contrast, the vapor pressure of an air parcel varies also with total pressure, although it is conservative for an isobaric heating without change in composition. Water-vapor density varies with air density. Relative humidity is greatly changed by adiabatic or isobaric changes, except in the special case where the vapor remains saturated. Although they are not directly observed, both mixing ratio and specific humidity are widely used quantities in meteorology because they are conservative.

3·3. Saturation Temperatures on Isobaric Cooling.[1]—If moist air is cooled sufficiently at constant pressure, the water vapor in it becomes saturated. Two particular methods of cooling, both occurring naturally under common conditions, lead to definite temperatures for the occurrence of saturation, namely, dew point and wet-bulb temperature.

Dew point.—This is the temperature at which saturation would occur if the air were cooled without change of pressure or composition. The dew point is realized in nature by a dry surface which is cooled just to the point where moisture begins to condense on it from the air. Dry ground serves this purpose when cooled by nocturnal radiation until dew begins to form. The polished metal surface of the dew-point hygrometer becomes clouded when cooled just below the dew point and clear again when warmed just above the dew point. This affords a simple means of determining dew point directly within narrow limits. Because vapor pressure is constant for an isobaric change with constant composition, dew point is the temperature for which the vapor pressure represents saturation. Hence, the relation between vapor pressure and dew point is identical with that between temperature and saturation vapor pressure, so that conversion between vapor pressure and dew point is given directly by a table of saturation vapor pressure or by the saturation curve on the characteristic diagram.

Wet-bulb Temperature.—This may be defined as the temperature at which saturation would occur if the air were cooled isobarically and adiabatically by means of contact with a water surface. In addition to this decrease in temperature there is an increase in dew point because of evaporation. The process is adiabatic in that the latent heat used in evaporation is specified as being supplied by the cooling of the air. This implies that the water surface itself, in order that the process may be achieved strictly as specified, must be at the wet-bulb temperature; otherwise, the final temperature of the air, after saturation is reached, would not be in equilibrium with the water—a physical impossibility.

If the water temperature is different from the wet-bulb temperature, it is evident that an adiabatic process is impossible and the water body gains or loses heat from the air so that its temperature becomes more

[1] By R. B. Montgomery.

nearly equal to the wet-bulb temperature of the air. This leads to the important result that any body of water in contact with air tends to assume the wet-bulb temperature of the air, provided, of course, that (1) the water body is otherwise insulated and (2) the cooling of the air and its moistening are distributed to an equal degree. These processes occur naturally when rain falls through vertically isothermal air. The temperature of the drops approaches the wet-bulb temperature of the air as they fall, and the air temperature and dew point also approach the wet-bulb temperature. By its definition, wet-bulb temperature is peculiar in that it is conservative during isobaric changes involving evaporation.

Conditions (1) and (2) above are well fulfilled by a wetted thermometer bulb that is sufficiently ventilated. Without sufficient ventilation the heat conducted through the thermometer stem and the radiative heat exchange are not negligible compared with the heat exchange between the wet surface and the air; therefore condition (1) is not fulfilled. The ventilated psychrometer, either a sling psychrometer or one employing mechanical ventilation, is the commonest means of measuring humidity with high accuracy.

Psychrometric Formula.—To obtain a relation between wet-bulb temperature and mixing ratio, consider the isobaric adiabatic transformation of air at pressure p from its actual temperature T to its wet-bulb temperature T_w. For unit mass of dry air with specific heat c_{pd} mixed initially with a mass of water vapor equivalent to the mixing ratio w with specific heat c_{pv}, the heat released by the cooling is $(c_{pd} + wc_{pv})(T - T_w)$. This heat is used to increase the mixing ratio to w_w, which represents saturation at T_w and p, by evaporation at the wet-bulb temperature, for which the latent heat is L_w, so that

$$L_w (w_w - w) = (c_{pd} + wc_{pv})(T - T_w).$$

If the small term containing c_{pv} is neglected,

$$w = w_w - \frac{c_{pd}}{L_w}(T - T_w). \tag{7a}$$

To the same degree of approximation the specific humidity is

$$q = q_w - \frac{c_{pd}}{L_w}(T - T_w) \tag{7b}$$

and the vapor pressure, on substitution from Eq. (5), is

$$e = e_w - c_{pd}\frac{m_d p}{m_v L_w}(T - T_w), \tag{7c}$$

where q_w and e_w represent saturation at the wet-bulb temperature. For wet-bulb temperatures below freezing, because the wet bulb becomes ice-covered, the saturation values refer to ice and L is the latent heat of sublimation. Any form of Eq. (7) is called the *psychrometric formula*, and the quantity $T - T_w$ is called the *wet-bulb depression.*

There are definitions of wet-bulb temperature other than the one given here, but the numerical differences in the values derived from them are never large, and they approach zero as the wet-bulb depression $T - T_w$ approaches zero. A similar statement applies to the approximations made in deriving the psychrometric formula, namely, that the errors introduced are never large and that they approach zero as the wet-bulb depression approaches zero. Because high precision is especially desirable when the vapor is nearly saturated, and because errors of the wet-bulb thermometer are normally greater with a large wet-bulb depression, it is evident that errors in the psychrometric formula are greatest when they are least important.

To find mixing ratio or specific humidity from the psychrometric formula it is necessary to know the saturation mixing ratio or saturation specific humidity. These may be computed from saturation vapor pressure and total pressure. Mixing ratio may be read approximately from any one of several meteorological diagrams (pseudoadiabatic chart, tephigram, emagram, and Rossby diagram). The coefficient of wet-bulb depression depends on the specific heat of dry air, $c_{pd} = 1.004$ joule g^{-1} $°C^{-1}$, and on the latent heat of vaporization of water L, which at the freezing point has the value $L_f = 2500$ joule g^{-1}. For atmospheric temperatures, expressed in degrees centigrade, L may be represented by

$$L = L_f[1 - 0.00094(T - T_f)],$$

where T_f is the freezing point.[1] The coefficient is therefore

$$\frac{c_{pd}}{L_w} = 0.402 \times 10^{-3}[1 + 0.00094(T_w - T_f)]$$

for temperatures expressed in degrees centigrade, and

$$\frac{c_{pd}}{L_w} = 0.223 \times 10^{-3}[1 + 0.00052(T_w - T_f)],$$

for temperatures expressed in degrees Fahrenheit.

In use of the psychrometric formula to find vapor pressure the saturation value depends on wet-bulb temperature alone whereas the coefficient of wet-bulb depression depends on pressure. The coefficient has the value

$$c_{pd}\frac{m_d}{m_v}\frac{p}{L_w} = 0.646 \times 10^{-3}p[1 + 0.00094(T_w - T_f)]$$

for temperatures expressed in degrees centigrade, and

$$c_{pd}\frac{m_d}{m_v}\frac{p}{L_w} = 0.359 \times 10^{-3}p[1 + 0.00052(T_w - T_f)]$$

[1] The constant in this equation is determined from data of Osborne, Stimson, and Ginnings, "Measurements of Heat Capacity and Heat of Vaporization of Water in the Range 0° to 100°C," *Jour. Research Nat. Bur. Standards*, **23**, 197–260 (1939).

for temperatures expressed in degrees Fahrenheit.[1] For a standard pressure of 1000 mb and any common wet-bulb temperature the coefficient is therefore close to 0.36 mb/°F.

This fixed value assumed for the coefficient, it follows from Eq. (7*c*) that for any constant wet-bulb temperature

$$\left(\frac{de}{dT}\right)_{T_w} \approx -0.36 \text{ mb/°F}.$$

Dashed lines with this slope are drawn on the characteristic diagram shown in Figs. *H* and *I* in the rear pocket. These lines represent constant wet-bulb temperature at 1000 mb, the value for each being the same as the temperature at which it intersects the saturation curve. For 1000 mb this family of broken lines together with the two sets of coordinate lines relates the temperature, vapor pressure, and wet-bulb temperature at any point. The mixing ratio is shown by a supplementary horizontal scale; the dew point is the temperature at which the vertical line through the given point intersects the saturation curve.

The deviation of sea-level pressure from 1000 mb does not usually exceed 5 per cent, and the vertical change in pressure from sea level to 1500 ft is also about 5 per cent. To compute mixing ratio or specific humidity from psychrometric observations it is therefore always necessary to include pressure as a third observation in order to avoid errors as large as about 5 per cent. In finding vapor pressure, on the other hand, only the coefficient of wet-bulb depression depends on pressure. For a pressure change of 5 per cent this coefficient is changed 5 per cent, an amount comparable to the observational error in the wet-bulb depression. Unless extreme accuracy is desired, it is permissible to use, as on the characteristic diagram, a fixed psychrometric coefficient of 0.36 mb/°F for measurements made within 1500 ft of sea level. For greater heights than this, the variation with pressure should be considered.

3·4. Refractive Index of Air at Radio Frequencies.[2]—The refractive index n of a gas obeys a relationship of the form[3]

$$n - 1 \propto \rho\left(A + \frac{B}{T}\right),$$

[1] The empirical coefficient given in the Smithsonian Meteorological Tables is

$$0.660 \times 10^{-3}p[1 + 0.00115(T_w - T_f)]$$

for temperatures expressed in degrees centigrade and

$$0.367 \times 10^{-3}p[1 + 0.00064(T_w - T_f)]$$

for temperatures expressed in degrees Fahrenheit. These are probably more nearly correct than the values from the approximate theory, but in most cases the difference is not of practical importance.

[2] By Pearl J. Rubenstein.

[3] P. Debye, *Polar Molecules*, Chemical Catalog Co., New York, 1929; Van Vleck, *Theory of Electric and Magnetic Susceptibilities*, Clarendon Press, Oxford, New York, 1932.

where ρ is the gas density and T the absolute temperature. The term proportional to density alone arises from induced polarization of the gas molecules in an external field; the other term results from the permanent dipole moment of the molecules. The index of refraction of a mixture of gases is generally assumed to obey the additivity rule; that is, the total value of $(n - 1)$ is equal to the sum of the contributions of the individual gases weighted in proportion to their partial pressures.

It is convenient to treat air as a mixture of dry gases and water vapor. Because none of the dry gases of air possesses a permanent dipole moment, only the water vapor contributes a term of the form ρ/T. The numerical value of the constant A thus depends on both the dry gases and the water vapor, whereas B is determined only by the permanent dipole moment of the water molecules.

There is a large amount of experimental information[1] on the dielectric constant[2] and refractive index of dry air taken over a wide wavelength range. The data show that some dispersion exists; the value in the visible region is approximately

$$(n - 1) \times 10^6 = 270 \text{ at } 0°\text{C and pressure of 1 atm.}$$

whereas, at static or radio frequencies the measurements give values from 294 to 296. The lowest of the r-f values, given as an average in the *Smithsonian Physical Tables*, is adopted here. This value determines the constant c to be 79°K/mb in the equation for index of refraction of dry air

$$(n - 1) \times 10^6 = \frac{cp_d}{T}, \tag{8}$$

where p_d is the dry gas pressure.

Experimental data on water vapor are less complete. In this case only values determined at radio frequencies are of interest because of the water absorption bands in the infrared region. Within the radio region dispersion may be neglected for wavelengths greater than about 2 cm.[3]

The determination of both A and B for water vapor requires absolute measurements of the dielectric constant or index of refraction for a series

[1] L. Boltzmann, *Wien. Ber.*, **69,** 795 (1874), and *Pogg. Ann.*, **155,** 403 (1875); K. Tangl, *Ann. Physik*, (4) **23,** 559 (1907), and **26,** 59 (1908); A. R. Jordan, J. W. Broxon, and F. C. Walz, *Phys. Rev.*, **46,** 66 (1934); L. G. Hector and H. L. Schultz, *Physics*, **7,** 133 (1936); Landolt-Börnstein, *Physikalisch-chemische Tabellen, 5te Auflage*, Springer, Berlin, 1923–1936; *Smithsonian Physical Tables*, 8th rev. ed., Washington, D.C., 1933.

[2] If ϵ and ϵ_0 are the permittivities of air and of free space, respectively, and the conductivity of the air is neglected,

$$n = \sqrt{\frac{\epsilon}{\epsilon_0}}.$$

The ratio ϵ/ϵ_0 is commonly called dielectric constant or specific inductive capacity.

[3] See Sec. 8·1.

of temperatures. The value of B, determined from the temperature dependence of the results, is fairly well known. Absolute values are needed if A is to be obtained, and this constant is therefore much less accurately known. From a consideration of all the available measurements[1] and their errors a simplified expression has been adopted for numerical use:[2]

$$(n - 1) \times 10^6 = \frac{c}{T}\left(p_d + e + \frac{be}{T}\right) = \frac{c}{T}\left(p + \frac{be}{T}\right), \tag{9}$$

where n = the index of refraction of moist air,
$c = 79°\text{K/mb}$,
$b = 4800°\text{K}$,
T = temperature in °K,
p = total pressure in mb,
e = the partial pressure of water vapor in mb.

On the characteristic diagram shown in Figs. H and I in the rear pocket, isopleths (contours of constant value) of $(n - 1) \times 10^6$ are given for a total pressure of 1000 mb. They are the curves that slope to the right. It will be noted that for air in which the vapor is saturated with respect to fresh water, salt water, or ice (represented by the saturation curves as described heretofore), the quantity $(n - 1) \times 10^6$ increases with increasing temperature. Despite the fact that the increase in temperature alone, according to Eq. (9), acts to decrease this quantity, the increase in saturation vapor pressure with increasing temperature is sufficient to cause the increase in $(n - 1) \times 10^6$.

In the solution of the wave equation in Chap. 2 a transformation of coordinates has been introduced that effectively flattens the earth. As a result a modified form of the refractive index becomes convenient. In Sec. 2·4 this modified index N was defined to be

$$N \equiv n\left(1 + \frac{z}{a}\right) \approx n + \frac{z}{a}, \tag{2·105}$$

where z is the height above the earth's surface and a is the earth's radius.

For numerical convenience a new parameter M is adopted, called *refractive modulus*, which is defined by the following equation:

$$M \equiv (N - 1) \times 10^6 \approx \left(n - 1 + \frac{z}{a}\right) \times 10^6. \tag{10}$$

[1] C. P. Zahn, *Phys. Rev.*, **27**, 329 (1926); A. C. Tregidga, *ibid.*, **57**, 294 (1940); R. Sänger and O. Steiger, *Helvetica Phys. Acta*, **1**, 369 (1928); R. Sänger, *Physik. Z.*, **31**, 306 (1930); R. Sänger, O. Steiger, and K. Gächter, *Helvetica Phys. Acta*, **5**, 200 (1932); J. D. Stranathan, *Phys. Rev.*, **48**, 538 (1935).

[2] Recent measurements at a frequency of 9340 Mc/s [C. M. Crain, *Phys. Rev.*, **74**, 691 (1948)] give slightly different values for the constants in this equation.

Whereas the modified index N is used in the radio theory of Chap. 2, M is more convenient in other applications. For purposes of qualitative discussion M and N may be used interchangeably, as their profiles have the same shape.

The radio-meteorologist is primarily concerned with the effects of atmospheric variations on the refractive modulus, particularly in the

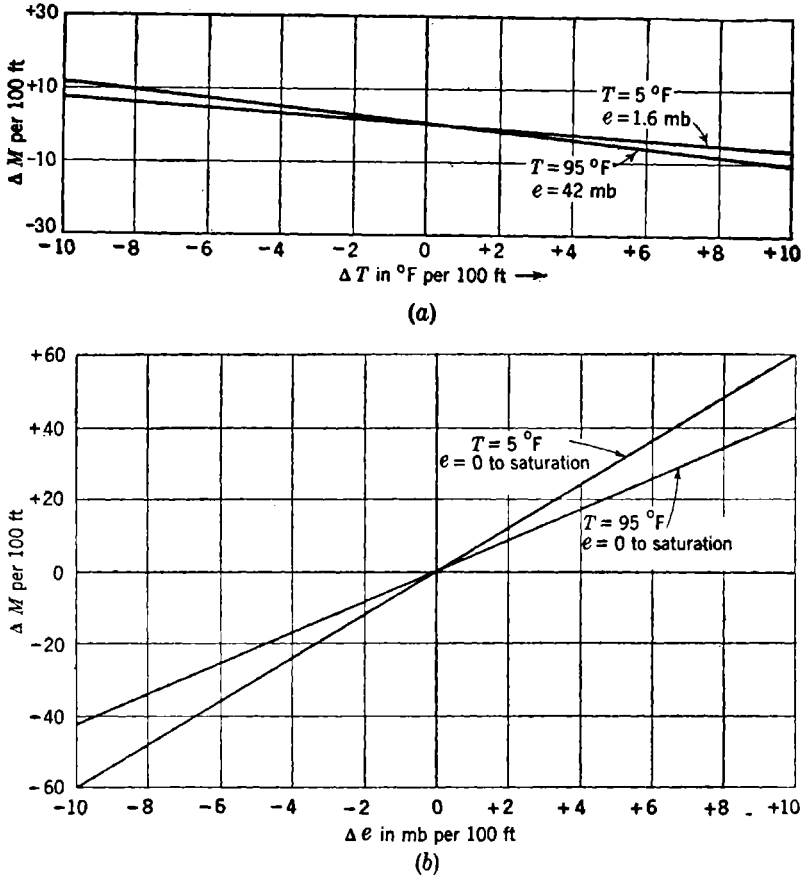


FIG. 3·2.—M-gradients produced by gradients of (*a*) temperature and (*b*) vapor pressure.

case of vertical variations in an air column. It follows from Eqs. (9) and (10) that in the absence of any vertical gradient of vapor pressure, a temperature inversion (temperature increasing with height) tends to cause M to increase less rapidly with height than in the standard case. If the temperature inversion is strong enough, it causes an actual decrease of M with height. In the absence of any vertical temperature gradient, the

vertical M-gradient varies directly with the vertical gradient of vapor pressure. It is desirable to have clearly in mind the amounts contributed by temperature and humidity gradients to the gradient of M. Figure 3·2 shows how much each factor contributes to changes in M. The top curves indicate how much the temperature gradient alone contributes to the M-gradient, and the lower pair of curves does the same for vapor-pressure gradients. These facts form the basis for meteorological refraction studies, and enlargements upon them make up the subject matter of the meteorology of the refraction problem.

VERTICALLY HOMOGENEOUS AIR AND ADIABATIC CHANGES

A vertical column or layer of air that has been completely mixed or stirred will be called *homogeneous*.[1] In such air the vertical variation of temperature is caused solely by changes in pressure; the water-vapor concentration is independent of height, although other humidity parameters vary because of the pressure variation. The lapse rate[2] of temperature in vertically homogeneous air will be discussed in Sec. 3·5, and the lapse rates of humidity variables in Sec. 3·6. In Sec. 3·7 the gradient of refractive modulus in homogeneous air will be considered.

The concept of homogeneity is important in the study of atmospheric refraction, for homogeneous air serves as a convenient reference for actual conditions. Soundings of the atmosphere in this chapter will be presented with reference lines (broken lines) showing conditions that would occur in homogeneous air. The M-gradient in homogeneous air is very nearly the gradient that is accepted as *standard* in its effect on the propagation of radio waves, as was explained in Sec. 1·4.

In addition to its function as a reference, vertical homogeneity is the only simple distribution that exists often in the atmosphere. It is particularly frequent in the lowest part of the atmosphere, the part of primary concern in the refraction problem. It occurs when the air is heated from below and stirred by convection (Sec. 3·14); hence homogeneous air often exists over land in the daytime when the air is warmed because of solar heating of the ground, and it is also found when an air mass moves over a warmer surface, whether ground or water. In these cases homogeneity may exist between some lower point that is about 50 ft above the earth's surface and some higher point that may vary widely

[1] The word "homogeneous" will be used in this chapter only as defined here. It should not be confused with the term "homogeneous atmosphere," which is used by some meteorologists to denote a theoretical atmosphere in which air density is independent of height. Homogeneous atmosphere as defined for this chapter is not homogeneous in the electromagnetic sense, as its electrical properties (refractive index) vary with position.

[2] *Lapse rate* is the decrease per unit increase of height. *Vertical gradient*, as used in this book, is the increase per unit increase of height.

but is usually less than 10,000 ft. Homogeneity is also approximated in some cases when the air near the surface of the earth is mechanically stirred by the wind.

3·5. Adiabatic Temperature Lapse Rate and Potential Temperature.[1]— The lapse rate of temperature that occurs in homogeneous air is called the *adiabatic lapse rate* because it represents the decrease of temperature of an air parcel which rises and cools adiabatically. If the air is unsaturated, the temperature decrease with height is properly called the *dry-adiabatic lapse rate* but in practice is often referred to simply as "adiabatic lapse rate." However, if condensation or evaporation processes occur, latent heat is a factor; the lapse rate under this condition is called the *moist-adiabatic lapse rate* or *saturation-adiabatic lapse rate* and will be discussed in the next section in connection with wet-bulb temperature.

The dry-adiabatic lapse rate follows from the first law of thermodynamics and the hydrostatic equation. The former, for an adiabatic process, can be written as

$$c_p\, dT = \alpha\, dp, \tag{11}$$

where c_p is the specific heat of air at constant pressure, T is the temperature, α the specific volume, and p the pressure.

According to the hydrostatic equation, which assumes no vertical acceleration of the air, the effect of gravity on a unit volume of air is balanced by the difference in pressure above and below the unit volume, or

$$\alpha\, dp = -g\, dz, \tag{12}$$

where g is the acceleration of gravity and z is the vertical coordinate. Strictly speaking, the α in Eq. (11) is somewhat different from the α in Eq. (12). In the former it means the specific volume of the displaced air parcel; in the latter it is the specific volume of the air through which the parcel is moving. The two may be different if the parcel is displaced very much from its original height. In general, however, the variation is small, and they may be considered equal to a good degree of approximation.

Combining Eqs. (11) and (12) gives

$$-\frac{dT}{dz} = \frac{g}{c_p} = 0.98°\text{C per 100 m} = 0.54°\text{F per 100 ft.} \tag{13}$$

A quantity that is often convenient to use is the *potential temperature* θ. It is defined as the temperature that an air parcel under any pressure would assume if its pressure were changed by a dry-adiabatic process to some standard pressure. The potential temperature, in terms of the original

[1] By Richard A. Craig.

temperature T, the original pressure p, and the standard pressure p_0, follows from Eq. (11). This equation becomes, upon substitution for α from the equation of state,

$$\frac{dT}{T} = \frac{R}{mc_p}\frac{dp}{p}, \tag{14}$$

where R is the universal gas constant, and m is the mean molecular weight of air. It follows from integration between the limits p, p_0 and T, θ that

$$\theta = T\left(\frac{p_0}{p}\right)^{\frac{R}{mc_p}} = T\left(\frac{p_0}{p}\right)^{0.286} \tag{15}$$

The standard pressure p_0 is commonly chosen as 1000 mb in meteorology. In meteorology associated with the refraction problem, however, it has proved convenient to take as the standard pressure the surface pressure. In practice, the potential temperature at any height can be found by adding to the temperature the product obtained by multiplying the adiabatic lapse rate by the height. The potential temperature determined in this manner differs from that defined by Eq. (15) by a negligible amount.

The practice of referring potential values to surface pressure rather than to some standard pressure will be followed in this and the next two sections. The primary reason for this practice is that in the refraction problem it is often desirable to compare temperature or humidity at some elevated point with values at the surface; reduction of the former to the surface pressure makes the comparison possible without any adjustment of the surface values. Moreover, the procedure is much simpler because in the refraction problem soundings are commonly plotted with height above the surface as ordinate.

In any one sounding, all values are referred to one pressure, but absolute values of potential temperature (and other potential values) from different soundings are not strictly comparable unless the surface pressures corresponding to the different soundings are identical. This is an important point to keep in mind but it is not a serious consideration at the present stage of refraction work. Such comparisons of absolute value are commonly not made except over small intervals of time or horizontal distance when variations due to variations of surface pressure are negligible. One important exception may occur in the comparison of soundings made at points where the heights above mean sea level differ. In this case the two reference surface pressures may be quite different, and all should be referred to one height for purposes of meteorological comparisons.

Certain considerations with respect to potential temperature may be stated. It is a conservative property for dry-adiabatic changes. In homogeneous air, the potential temperature is independent of height, and the air is said to be in *neutral equilibrium*. If in an air column the potential

temperature increases with height, the air is said to be in *stable equilibrium*; if it decreases with height, the air is said to be in *unstable equilibrium*. The importance of stability and instability will be discussed later in connection with heating from below (Sec. 3·14) and cooling from below (Sec. 3·17).

3·6. Humidity Lapse in Homogeneous Air.[1]—In homogeneous air, the water-vapor concentration is independent of height, but the vapor pressure, dew point, and wet-bulb temperature vary with height because of changes in total pressure of the air. The homogeneous lapse rates that will be derived for the parameters depending on pressure are analogous to the adiabatic lapse rate of temperature; they indicate the changes in the parameters caused by a change in height when no water vapor is gained or lost by the air parcel.

Equation (4) gives the vapor pressure in terms of specific humidity and total pressure. If the specific humidity is constant, differentiation of this equation yields

$$\frac{1}{e}\frac{de}{dz} = \frac{1}{p}\frac{dp}{dz}. \tag{16}$$

Upon substitution from the hydrostatic equation [Eq. (12)] and the equation of state, this becomes

$$\frac{de}{dz} = -\frac{mg}{RT}e. \tag{17}$$

The lapse rate of vapor pressure thus depends inversely on the temperature and directly on the vapor pressure. For a temperature of 0°C,

$$\begin{aligned} -\frac{de}{dz} &= 0.0125e \text{ per } 100 \text{ m} \\ &= 0.0382e \text{ per } 1000 \text{ ft}. \end{aligned} \tag{18}$$

The lapse rate of the dew point in homogeneous air can also be derived. Equation (3) can be rewritten as

$$\frac{1}{e}\frac{de}{dT_s} = \frac{m_v}{R}\frac{L}{T_s^2}, \tag{19}$$

where T_s is the dew point. When de is written in terms of dz according to Eq. (17), this becomes

$$-\frac{dT_s}{dz} = \frac{m}{m_v}\frac{g}{L}\frac{T_s^2}{T} \tag{20}$$

Hence the lapse rate of dew point in homogeneous air varies with the square of the dew point and inversely as the temperature. For the temperature

[1] By Richard A. Craig.

and dew point both equal to 0°C,

$$-\frac{dT_s}{dz} = 0.17°\text{C per 100 m}$$

$$= 0.94°\text{F per 1000 ft.}$$

The lapse rate of wet-bulb temperature in homogeneous air is more complicated. It follows from a differentiation of the psychrometric formula, the form given in Eq. (7*a*) being used here. In these equations, in accordance with the convention introduced in Sec. 3·3, the subscript w refers to conditions at the wet-bulb temperature. If the concentration of water vapor is constant and the coefficient of the wet-bulb depression is assumed to be independent of height,

$$\frac{dw_w}{dz} = \frac{c_{pd}}{L_w}\left(\frac{dT}{dz} - \frac{dT_w}{dz}\right). \tag{21}$$

From Eq. (6), if the small vapor-pressure term in the denominator is neglected,

$$\frac{dw_w}{dz} = \frac{m_v}{m_d}\frac{1}{p}\left(\frac{de_w}{dz} - \frac{1}{p}\frac{dp}{dz}\right).$$

After substitution from Eq. (3) for de_w/dz, from Eq. (12) for dp/dz, from Eq. (6) for e_w (again neglecting the e_w in the denominator), and from the equation of state, the last equation becomes

$$\frac{dw_w}{dz} = \frac{m_v}{R}\frac{L_w}{T_w^2}w_w\frac{dT_w}{dz} + \frac{w_w mg}{RT}.$$

When this is inserted in Eq. (21), together with the adiabatic value for dT/dz from Eq. (13), the wet-bulb lapse rate in homogeneous air is seen to be

$$-\frac{dT_w}{dz} = \frac{g}{c_p}\frac{1 + \dfrac{L_v m}{RT}w_w}{1 + \dfrac{m_v}{T_w^2}\dfrac{L_w^2}{Rc_p}w_w}. \tag{22}$$

The lapse rate depends on the wet-bulb temperature, the temperature, and the mixing ratio at the wet-bulb temperature. It may, of course, be written in several other forms using other parameters such as specific humidity, vapor pressure, and pressure. The form used here has the advantage that it can be compared easily with the dry-adiabatic temperature lapse rate. Assuming a temperature and wet-bulb temperature both equal to 0°C,

$$-\frac{dT_w}{dz} = \frac{g}{c_p}\frac{1 + 32w_w}{1 + 180w_w}.$$

This equation shows that the lapse rate of wet-bulb temperature is *always less* than the dry-adiabatic lapse rate g/c_p. For a mixing ratio of 10 per mille, its value is about one-half the dry-adiabatic value.

The lapse rate of wet-bulb temperature in homogeneous air is to a good degree of approximation the same as the moist-adiabatic lapse rate mentioned in the previous section. Equation (22) holds for the moist-adiabatic lapse rate if the T in the numerator and the T_w in the denominator are identical and all other parameters refer to that temperature.[1]

Quantities called *potential vapor pressure*, *potential dew point*, and *potential wet-bulb temperature* may be defined as the values of the parameters at some standard pressure. In refraction work, as in the case with potential temperature, it is more convenient to refer them always to surface pressure rather than to an invariant standard pressure. The potential values are found in practice by adding to the values at any height the product of the respective lapse rate and the height. Although the lapse rates of all three vary with height, it is sufficiently accurate to assume a mean lapse rate over the height range of 1000 to 2000 ft, the region that is usually considered when potential values are desired in refraction work.

3·7. Gradient of Refractive Modulus in Homogeneous Air, Potential Modulus.[2]—Refractive modulus has been given in terms of pressure, temperature, vapor pressure, and height in Sec. 3·4. Its variation with height can be written in terms of the vertical gradients of potential temperature and specific humidity, these parameters being desirable here because of their constancy in homogeneous air. It follows from Eqs. (9) and (10) that

$$\frac{dM}{dz} = -r_1 \frac{d\theta}{dz} + r_2 \frac{dq}{dz} + r_3. \tag{23}$$

In this equation

$$r_1 = 79 \frac{p}{T^2} \left(\frac{p}{p_0}\right)^{0.286} \left(1 + 15.5 \times 10^3 \frac{q}{T}\right) \quad \text{per degree,}$$

$$r_2 = 61 \times 10^4 \frac{p}{T^2},$$

$$r_3 = 4.78 - 59 \frac{p}{T^2} \left(1 + 4.6 \times 10^3 \frac{q}{T}\right) \quad \text{per 100 ft,}$$

where temperatures are in degrees Kelvin, pressures in millibars, and heights in hundreds of feet; p_0 is the standard pressure, usually the surface pressure, to which the potential temperature is referred.

[1] A graph showing the saturation-adiabatic lapse rate and hence the wet-bulb temperature lapse rate as a function of temperature and pressure may be found in D. Brunt, *Quart. Jour. Roy. Meteorol. Soc.*, **59**, 351 (1933).

[2] By Isadore Katz.

In homogeneous air, the potential temperature and specific humidity are independent of height; therefore the vertical gradient of M is simply equal to r_3. This value increases somewhat with increasing temperature and decreases with increasing pressure and specific humidity. However, within the range of variation of these quantities near the surface of the earth, the vertical gradient of M is affected very little and the value 4.0 per 100 ft is accurate in nearly all cases.

It is often convenient, from the meteorologist's point of view, to deal with a parameter that represents the index of refraction and that is vertically constant in homogeneous air. Such a parameter is the *potential refractive modulus* Φ. It may be referred to simply as *potential modulus*.[1]

The potential refractive modulus is the value of refractive modulus that an air parcel would have at some standard pressure. If the standard pressure chosen is the surface pressure, so that $z = 0$,

$$\Phi \equiv M_{\text{potential}} = \frac{c}{\theta}\left(p_0 + b\frac{e_p}{\theta}\right). \tag{24}$$

The constants c and b are those that entered into Eq. (9), namely, $c = 79°\text{K/mb}$ and $b = 4800°\text{K}$; θ is the potential temperature and e_p is the potential vapor pressure, both referred to surface pressure; and p_0 is the surface pressure. As the gradient of M in homogeneous air is approximately 4.0 per 100 ft, Φ can also be found by subtracting from M at any level the product of 4.0 and the height in hundreds of feet.

In homogeneous air, all the factors on the right in Eq. (24) are independent of height; and accordingly Φ is independent of height. If the temperature lapse rate is less than the dry-adiabatic rate, the potential temperature increases with height and Φ, because of this factor, decreases with height. If the lapse of vapor pressure with height is greater than the homogeneous rate, potential vapor pressure and Φ decrease with height; if it is less than the homogeneous rate, potential vapor pressure and Φ increase with height.

Definitions corresponding to those in Sec. 1·4 for M may be stated in terms of Φ. The vertical gradient of Φ in a standard layer is nearly zero, having the value -0.4 per 100 ft. The difference between a standard and a homogeneous layer is therefore very small. The homogeneous vertical distribution of Φ may, for practical purposes, be regarded as the dividing line between substandard distributions sloping to its right and superstandard distributions sloping to its left. A decrease in Φ of more than 4.0 per 100 ft corresponds to an M-inversion.

[1] A concept quite similar to this was introduced by George D. Lukes of the Camp Evans Signal Laboratory at a conference on wave propagation in Washington, D.C., November 1944.

3·8. Characteristic Curves and Mixing.[1]—It has been seen that a characteristic diagram is basically a graph which has vapor pressure as abscissa and temperature as ordinate, and on which the relationships between temperature and vapor pressure over water and ice are shown by saturation curves. It has been found convenient to add to this simple graph certain other information needed for practical use in connection with the refraction problem. Figures *H* and *I* in the rear pocket represent the characteristic diagram as it may be used in such practical work. In addition to the saturation curves for ice and water, there is a curve for vapor pressure over sea water, as was explained in Sec. 3·1. There is also a supplementary abscissa scale for mixing ratio (Sec. 3·2), a family of broken straight lines showing the approximate wet-bulb temperature at a pressure of 1000 mb (Sec. 3·3), and a family of curves showing $(n - 1) \times 10^6$ for a pressure of 1000 mb (Sec. 3·4).

In studying mixing, the coordinates of the characteristic diagram are better taken to be potential temperature and potential vapor pressure, which are constant during adiabatic ascent or descent of an air parcel not containing saturated vapor. As was mentioned in the previous sections, potential values here are referred to the surface pressure rather than to 1000 mb. Then the saturation curve is correct for the air reduced to the surface, but not for the air at any other height. The scale for mixing ratio becomes approximately correct for any height and exactly correct for any height if the surface pressure is 1000 mb. The lines sloping -0.36 mb/°F become correct for potential wet-bulb temperature referred to the surface (the slope of these lines is strictly correct for certain conditions only, Sec. 3·3). A vertical line through any point intersects the saturation curve at the temperature that is the corresponding potential dew point. The fourth family of lines becomes potential modulus Φ according to the approximate formula in Sec. 3·9.

With these coordinates the diagram becomes very like the Rossby (or equivalent potential temperature) diagram, in common use, on which the coordinates are logarithm of potential temperature and mixing ratio. The characteristic diagram is intended to fill much the same purpose but is drawn on a larger scale and is especially designed for low-level soundings rather than for the common meteorological soundings that extend into the stratosphere.

On either diagram a homogeneous mass of air is represented by a point, because potential temperature and mixing ratio or potential vapor pressure are constant throughout homogeneous air. These quantities are conservative; therefore the point remains fixed during vertical displacements of the air as long as there is no exchange of heat or water by radiation, diffusive processes, or evaporation. If the air is not homogeneous but has continuously varying properties, a sounding through the air

[1] By R. B. Montgomery.

forms a continuous curve on the diagram. This curve remains fixed during changes of the type described. Such a curve is called a *characteristic curve*, and from this the name *characteristic diagram* is derived.

The characteristic diagram has particular advantages in the study of mixing. Consider two samples A and B of homogeneous air. If equal quantities of these are mixed together, the potential temperature of the resulting mixture is the average of the potential temperatures of A and B, and its specific humidity is the average of the specific humidities of A and B. Consequently the potential vapor pressure is very nearly the average of the potential vapor pressures of A and B, so that on the diagram the mixture is represented almost exactly by the midpoint of the straight line joining the points representing A and B, as shown in Fig. 3·3. If A and B are mixed in any other proportion, the mixture is represented by another point on the same straight line. If the two masses of air are separated by a zone of gradual transition, the characteristic curve for the transition zone is the straight line joining the points representing A and B.[1]

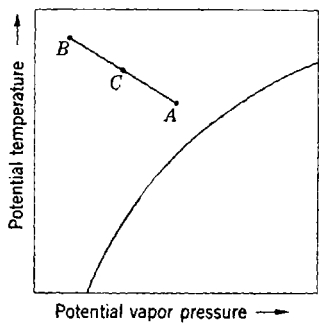


FIG. 3·3.—Characteristic diagram. Point C represents mixture of air masses A and B in equal quantities.

The isopleths of potential refractive modulus on the characteristic diagram are nearly straight, parallel, and equally spaced lines; thus the mixture of A and B in equal quantities has a potential modulus nearly equal to the average of the potential moduli of A and B. A general statement of this approximate result is as follows. In the transition zone resulting from the mixing of two homogeneous masses of air A and B, at any point P the ratio of the change from A to P to the change from A to B is the same for potential temperature, for potential vapor pressure or mixing ratio, and for potential modulus. This property is designated as that of *similarity* in distribution of the quantities mentioned or of others like them.

Mixing tends to shorten the portion of a characteristic curve where the mixing occurs, as shown schematically in Fig. 3·4.

The air in immediate contact with a boundary has the same temperature as the boundary. In addition, if the boundary is pure water, the vapor at the boundary is saturated; if the boundary is ocean water, the

[1] The statements of this paragraph are not strictly correct for the Rossby diagram because the ordinate is the logarithm of potential temperature. Arakawa ["A New Type of Rossby Diagram," *Bull. Am. Meteorol. Soc.*, **21**, 111, (1940)] has modified the Rossby diagram by substituting a linear scale of potential temperature; for this modified version the statements are correct, because, as he says, "The potential temperature of the mixture is equal to the mean potential temperature."

vapor is 98 per cent saturated (Sec. 3·1). Thus the temperature and composition of a water surface fix on the characteristic diagram the point representing the air at the surface. If originally homogeneous air comes into contact with a water surface of constant temperature, and if the transition zone is produced by mixing only, the characteristic curve for the transition zone is the straight line joining the point representing the original air with the point representing the air at the boundary, as shown schematically in Fig. 3·5. It will be seen that this ideal case some-

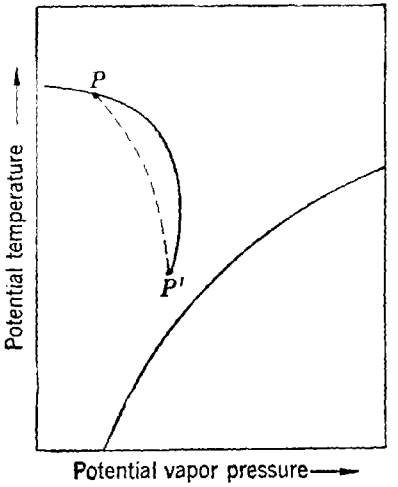


FIG. 3·4.—Characteristic diagram showing the shortening of original characteristic curve caused by mixing within the region P–P'.

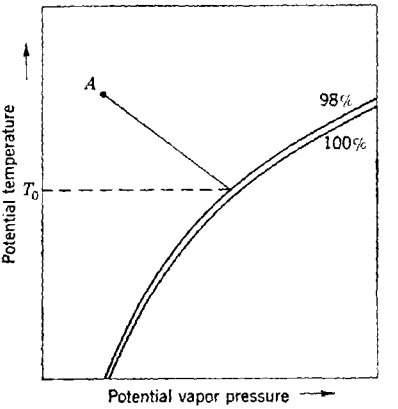


FIG. 3·5.—Characteristic curve produced by mixing of air A in contact with salt water of temperature T_0.

times occurs when air blows from a warm land surface to a cold sea surface. For air over water the characteristic curve in all cases, even when not straight, terminates at the point corresponding to the temperature and composition of the water surface.[1]

An important result is that in many cases the characteristic curve of a measured sounding over water may be extrapolated to intersect the saturation curve, the point of intersection determining the water temperature.

REPRESENTATION AND DESCRIPTION OF SOUNDINGS

BY ISADORE KATZ

The meteorological questions raised in connection with the propagation of short radio waves in the troposphere have necessitated the develop-

[1] The temperature of the water surface, however, is not always the same as the water temperature measured a short distance below the surface; therefore this rule must be applied with caution.

ment of new techniques for measuring the vertical structure of the atmosphere in fine detail. Some of the new instruments and methods for using them will be discussed at the end of this chapter. At this point it is desirable to consider the methods by which these data may be conveniently reduced and presented in a form useful to all concerned.

The common ground between meteorology and the propagation problem is found by use of the refractive modulus M. An approximate formula whereby measurements of temperature and humidity can easily be converted to values of M will be presented in Sec. 3·9. In Sec. 3·10 there will be a discussion of various parameters that may be used in presenting the data graphically.

3·9. Approximate Formula for Refractive Modulus.—In order to facilitate the reduction of raw data, certain simplifications of the formula for M may often be made. This formula, which follows from Eqs. (9) and (10), is

$$M = \frac{c}{T}\left(p + b\frac{e}{T}\right) + \frac{z}{a}10^6. \tag{25a}$$

An approximate expression for the pressure p follows from Eq. (12), integrated under the assumption of constant density. Making use of the equation of state, this becomes

$$p = p_0 - g\rho_0 z = p_0 - g\frac{m}{R}\frac{p_0}{T_0}z,$$

where the subscript 0 refers to values at the surface. On substitution of this expression, the formula for M takes the form

$$M \approx \frac{c}{T}\left(p_0 + b\frac{e}{T}\right) + \left(\frac{10^6}{a} - cg\frac{m}{R}\frac{p_0}{TT_0}\right)z. \tag{25b}$$

If it is assumed that T and T_0 in the final term are equal, the coefficient of z depends only on surface pressure and temperature. A further simplification follows from the assumption that $p_0 = 1000$ mb and $T_0 = 289°$K. Then the approximate formula is

$$M \approx \frac{79}{T}\left(1000 + 4800\frac{e}{T}\right) + 3.8z, \tag{25c}$$

where z *is in hundreds of feet*, T in degrees Kelvin, e in millibars.

This approximation obviates the measurement of pressure and simplifies the computation of M from temperature, vapor pressure, and height. The characteristic diagram, described in Sec. 3·8, contains isopleths of M for zero height according to Eq. (25b) and is easily used for computation purposes.

The assumptions made in obtaining the approximate formula for refractive modulus lead to errors that will now be examined. The

difference between the true value of M and the approximate value given by Eq. (25c) is

$$\delta M = \frac{79}{T}(p - 1000) + z, \tag{26}$$

where the quantities are in the units indicated for Eq. (25c).

The error at the surface ($z = 0$) for $T = 273°\text{K}$ is given by

$$\delta M = \frac{79}{T}(p_0 - 1000) = 0.29(p_0 - 1000).$$

Sea-level pressure may vary between 950 and 1050 mb; thus the absolute error may be as much as ± 14.5 or 4 to 5 per cent of the usual value of M near the surface. Moreover, if the surface is far from sea level, so that the surface pressure varies even more from 1000 mb, correspondingly larger errors result.

Errors slightly larger than those at the surface may occur at a point above the surface. For example, for $z = 1000$ ft, $p = 914$ mb, and $T = 273°\text{K}$ the error would be about -15. These errors are large compared with the error introduced by instrumental inaccuracies, which is about 1 or 2 in M when careful measurements are made.

However, for propagation work the *shape* of the M-profile rather than the absolute values of M is of most importance. Therefore it is desirable to find the error in the gradient of M introduced by using the approximate formula. The gradient is obtained by differentiating the formula for M [Eq. (25a)]:

$$\frac{dM}{dz} = -\frac{c}{T^2}\frac{dT}{dz}\left(p + 2b\frac{e}{T}\right) + \frac{bc}{T^2}\frac{de}{dz} + \frac{10^6}{a} - cg\frac{m}{R}\frac{p}{T^2}.$$

If the gradient is obtained from Eq. (25b),

$$\frac{dM}{dz} \approx -\frac{c}{T^2}\frac{dT}{dz}\left(p_0 + 2b\frac{e}{T}\right) + \frac{bc}{T^2}\frac{de}{dz} + \frac{10^6}{a} - cg\frac{m}{R}\frac{p_0}{T_0^2},$$

the error in M-gradient $\delta(dM/dz)$ is obtained by subtracting this approximate equation from the preceding one:

$$\delta\left(\frac{dM}{dz}\right) = \frac{c}{T^2}\frac{dT}{dz}(p_0 - p) + \frac{cgm}{R}\frac{p_0}{T_0^2}\left(1 - \frac{pT_0^2}{p_0T^2}\right).$$

This equation reduces to a convenient numerical form if $p_0 = 1000$ mb, $T_0 = 289°\text{K}$, and z is measured in hundreds of feet. Then

$$\frac{cgm}{R}\frac{p_0}{T_0^2} \approx 1 \qquad \text{per 100 ft},$$

and

$$\delta\left(\frac{dM}{dz}\right) = \frac{79}{T^2}\frac{dT}{dz}(1000 - p) - \frac{p}{1000}\left(\frac{289}{T}\right)^2 + 1, \qquad \text{per 100 ft.} \qquad (27)$$

The error resulting from the approximation may be divided into two parts:

1. An error depending on p, p_0, T, T_0 given by the last two terms. As an example, for $T = 300°\text{K}$ and $p = 900$ mb, the error amounts to 0.16 per 100 ft. This is about 5 per cent of the standard gradient and for practical purposes is negligible. A negative departure of about the same magnitude would occur in very cold air at 1000 mb.
2. Another error proportional to dT/dz and to $(1000 - p)$. It is greatest in a high temperature inversion. For an estimate of the error under rather extreme conditions of stability assume $T = 273°\text{K}$, $dT/dz = 20°\text{C}$ per 100 ft, $p = 914$ mb. Then, from Eq. (27), the error

$$\delta\left(\frac{dM}{dz}\right) = \frac{79}{T^2}\frac{dT}{dz}(1000 - p) = 1.82 \qquad \text{per 100 ft.}$$

This term is zero in an isothermal atmosphere and under average conditions is only about one-tenth the value computed above. It should be clear that an error of even 2 per 100 ft in the M-gradient is not large under the extreme conditions stated; for under the influence of such stability, the actual M-gradient would be likely to be in the neighborhood of 20 per 100 ft, and the error would be only 10 per cent. Note in this case also that if the surface is not near sea level, the errors may be larger than stated.

The approximate formula for the potential refractive modulus assumes also that the surface pressure is 1000 mb. Thus

$$\Phi \approx \frac{79}{\theta}\left(1000 + 4800\,\frac{e_p}{\theta}\right), \qquad (28)$$

where θ is potential temperature in degrees Kelvin and e_p is potential vapor pressure in millibars.

The absolute difference between the true value of Φ and the value computed from the formula above may be large. It is given by

$$\delta\Phi = \frac{79}{\theta}(p_0 - 1000), \qquad (29)$$

where p_0 is the actual surface pressure, with units as above.

The difference between the surface pressure and 1000 mb may be as much as 5 per cent, and the error in Φ may be as much as 14 (about 4 to 5 per cent) or more if the surface is near sea level.[1] With Φ as with M,

[1] For a station where the surface is not near mean sea level, a pressure near the average surface value should be chosen rather than 1000 mb.

the absolute value is unimportant and the slope of the curve is the important variable. The error in slope is given by

$$\delta\left(\frac{d\Phi}{dz}\right) = -\frac{79}{\theta^2}\frac{d\theta}{dz}(p_0 - 1000) \qquad \text{per 100 ft.} \tag{30}$$

This error is very small in absolute magnitude. It is zero for homogeneous air ($d\theta/dz = 0$) and increases with the stability of the air. For a rather extreme case where the potential temperature increases at the rate of 10°F per 100 ft and the surface pressure differs from 1000 mb by 50 mb, the error in slope is only about 0.2 per 100 ft This may be compared with possible variations in gradient in the neighborhood of 100 per 100 ft because of changing atmospheric conditions.

In the remainder of this chapter M and Φ will be used as defined in Eqs. (25*b*) and (28) respectively.

3·10. Representation of Soundings.—In conventional meteorological work soundings are most frequently plotted on a temperature-pressure chart. Only the temperature curves are drawn; humidity values are merely written in at the proper height. This method of designating a vertical sampling of the atmosphere is totally inadequate for the low-level meteorology of microwave propagation.

In plotting low-level soundings, it is customary to use a linear height scale as ordinate, instead of pressure. It is desirable to show at least three variables, one representing temperature, one representing humidity, and one representing some variant form of the index of refraction.

There are two possible parameters that might be used to represent temperature, namely, potential temperature and temperature itself. The former has the advantage that it is conservative for adiabatic changes. Temperature, however, is in many cases simpler to use, as it is measured directly and is included in all standard weather reports. For this reason it has been chosen for use in the soundings in this chapter. The temperatures are all plotted on a Fahrenheit scale. In many ways the centigrade scale may be superior for scientific work. However, in refraction work it is frequently necessary to use standard weather reports, and because temperatures are customarily reported in degrees Fahrenheit in the United States and in the British Empire, experience has shown that considerable time is saved by conforming to this custom.

Humidity could be represented by any one of several quantities. Of these, relative humidity and wet-bulb temperature are immediately seen to be undesirable because they do not specify the amount of water vapor in the air, except in conjunction with a value of temperature.

Vapor pressure has much to recommend it because it enters simply into the equations for the various forms of refractive index and is directly proportional to the amount of water vapor in the air. Water-vapor concentration (specific humidity or mixing ratio) is proportional to vapor

pressure and, in addition, is conservative for adiabatic changes. Any one of these parameters could well be used under certain conditions.

Dew point, however, has been found to have many advantages that make it a desirable parameter and prompt its use in the soundings in this chapter:

1. It can be plotted on a single scale along with air temperature.
2. Saturated layers are clearly visible from the curves of temperature and dew point.
3. It affords immediate comparison with the temperature of a water surface over which air may be traveling, to indicate the direction of the humidity gradient.
4. All surface weather reports in the United States include dew point. As in the case of temperature, conformity to customary usage saves much labor if standard weather reports are to be used.
5. It is nearly conservative for height changes over an interval of a few thousand feet.

A disadvantage of dew point is the fact that it does not vary linearly with vapor pressure and water-vapor concentration.

Potential modulus Φ and the refractive modulus M are both presented as the product of meteorological soundings in this chapter. Each has its own advantages. Potential modulus Φ is superior to M in that a homogeneous air column is represented by a vertical line, deviations toward substandard or superstandard thus being immediately recognizable. It has the advantage also that its curve is ordinarily very similar to the dew-point curve. For these two reasons, Φ is more readily acceptable to the meteorologist. On the other hand, the numerical convenience of M and the fact that the M-profile is identical in shape with that of the modified-index profile furnish sufficient reason to employ both Φ and M.

One thing more may be said about the sounding presentation. The accuracy far exceeds that of the widely used radiosonde New instruments are in existence that are designed primarily for measuring, in minute detail, the lower 1000 ft of the atmosphere. These will be described in Secs. 3·27 to 3·31. Some of these instruments attain an accuracy of nearly ± 0.2°F, better than the radiosonde by a factor of 5 or more. An expanded temperature scale on the graph is called for in view of the instrumental accuracy.

The soundings in this chapter, then, will uniformly include temperature, dew point, potential modulus, and refractive modulus plotted against height above the earth's surface. The height scale will be in feet; temperature in degrees Fahrenheit. Reference lines (dashed lines) will be given showing distributions in homogeneous air. Arrows will indicate values of temperature, humidity, and refractive modulus at the surface. Wherever practicable, one scale unit will represent 100 ft, 1.8°F, four units of Φ and M. With this scale, the M-profile in homogeneous air slopes

45° to the right of vertical, and the Φ-curve represents an M-inversion if it is inclined more than 45° to the left of vertical.

EDDY DIFFUSION

By R. B. Montgomery

The processes of diffusion in the lower part of the troposphere are intimately associated with boundary friction and wind distribution and fall within the general field of fluid mechanics. There is a large fund of information with bearing on atmospheric processes of diffusion, both from fluid mechanics in general and from meteorological studies in particular, although the latter have not progressed sufficiently to answer most practical questions that arise. The material under this heading consists of some fundamental considerations and a review of the meteorological information available.

There are two books that afford especially extensive additional material. The general one is *Modern Developments in Fluid Dynamics*[1] edited by S. Goldstein, and the one on meteorological aspects is H. Lettau's *Atmosphärische Turbulenz*.[2]

3·11. Eddy Viscosity and Eddy Diffusivity.—Viscosity, conductivity, and diffusivity will precede the eddy counterparts. The simplest sort of fluid friction occurs when the direction of motion is everywhere the same and the speed u depends only on a coordinate z normal to this direction. In many problems the coordinate z is the distance from a solid wall, and in meteorological problems it is the vertical coordinate. Because of friction there is a tangential stress τ on any plane normal to z. The *dynamic viscosity* μ may be defined as the ratio of the stress to the shear, so that

$$\tau = \mu \frac{\partial u}{\partial z}. \tag{31}$$

The positive sign in this equation conforms to the convention that a stress in the direction of positive u is positive if momentum in the direction of positive u is transported in the direction of decreasing z. The viscosity, except for unusual cases, is a physical property of the fluid and is independent of the fluid's motion. A gradient of stress, in the absence of other forces, produces an acceleration

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial z^2}, \tag{32}$$

where the ratio of dynamic viscosity (dimensions $L^{-1}MT^{-1}$) to density, $\nu = \mu/\rho$, is by definition the *kinematic viscosity* (dimensions L^2T^{-1}).

[1] Clarendon Press, Oxford, New York, 1938.

[2] Akademische Verlagsgesellschaft, Leipzig, 1939; reprint, Edwards Brothers, Ann Arbor, 1944.

In the same way, suppose that temperature is constant over every plane normal to z but varies in the z-direction. Because of conduction, heat is transported in the direction of increasing z at the rate per unit area

$$H = -k\frac{\partial T}{\partial z}, \tag{33}$$

where k is the *thermal conductivity*, a quantity dimensionally equivalent to the product of dynamic viscosity and specific heat, $c_p\mu$. If radiation and production of heat are negligible, a nonuniform temperature gradient results in a rate of heating equal to

$$\frac{\partial T}{\partial t} = \nu_c\frac{\partial^2 T}{\partial z^2}, \tag{34}$$

where $\nu_c = k/c_p\rho$ is called *thermometric conductivity* and is dimensionally equivalent to kinematic viscosity.

As the third case, suppose that specific humidity q is constant over any plane normal to z. Because of diffusion, water vapor is transported in the direction of decreasing concentration, the rate per unit area of mass transfer in the direction of increasing z being

$$E = -\rho\nu_d\frac{\partial q}{\partial z}, \tag{35}$$

where ν_d is the *diffusivity* of water vapor in air. The resulting rate of change of specific humidity is

$$\frac{\partial q}{\partial t} = \nu_d\frac{\partial^2 q}{\partial z^2}. \tag{36}$$

Diffusivity also is dimensionally equivalent to kinematic viscosity.

For a gas the magnitudes of the three dimensionally equivalent quantities are of the same order but are usually unequal. At 0°C and 1000 mb the kinematic viscosity and thermometric conductivity of air and the diffusivity of water vapor in air are, respectively,

$$\nu = 0.1346 \text{ cm}^2/\text{sec}, \quad \nu_c = 0.189 \text{ cm}^2/\text{sec}, \quad \nu_d = 0.226 \text{ cm}^2/\text{sec}.$$

The effects of friction, conduction, and diffusion are important where the fluid flow is comparatively steady. This is the case close to any solid in the region called the *laminar sublayer*. Steady flow in the atmosphere occurs in a laminar sublayer of the order of 1 mm in thickness next to the ground or water surface and in temperature inversions that are sufficiently intense in relation to the shear. Although the laminar sublayer is thin, it is by no means negligible, because a large part of the heat gained or lost by the atmosphere must be conducted through the layer, and the water vapor gained by the atmosphere must be diffused through it.

Most of the lower atmosphere, however, is in unsteady flow, and the mixing effect of turbulence accomplishes a vertical transfer of momentum, heat, and water vapor. In comparison with this eddy transfer, as it is called, the effect of true viscosity, conductivity, or diffusivity is of negligible magnitude.

In this unsteady motion let u and w be the velocity components in the directions x (a horizontal coordinate) and z respectively. Let the mean value of u be $\bar{u}$, and let $u = \bar{u} + u'$, and assume that the mean value of w is $\bar{w} = 0$ so that $w = w'$. Because of the component w there is an instantaneous rate of mass transfer per unit area in the direction of decreasing z equal to $-\rho w$ and a corresponding transfer of x-component of momentum

$$-\rho wu = -\rho w'(\bar{u} + u').$$

In the mean the first term on the right vanishes, leaving only $-\rho\overline{w'u'}$. This mean rate of momentum transfer per unit area in the direction of decreasing z is called a component of *eddy stress* on the imaginary surface normal to z. The ratio ρK_v of eddy stress to the appropriate shear of the mean motion is called *dynamic eddy viscosity*, and thus

$$-\rho\overline{w'u'} = \rho K_v \frac{\partial \bar{u}}{\partial z}. \tag{37}$$

This equation is similar in form to Eq. (31). It expresses the eddy stress as the product of the shear of the mean motion and of an eddy viscosity depending in a specific way on the degree of turbulence or mixing. Unlike the true viscosity, eddy viscosity is not a property of the fluid but is a property of the unsteady flow. Eddy viscosity varies widely from case to case and spatially within any one case.

Differentiating the last equation leads to the acceleration of the mean flow if forces other than the eddy stress are neglected:

$$\frac{\partial \bar{u}}{\partial t} = \frac{\partial}{\partial z}\left(K_v \frac{\partial \bar{u}}{\partial z}\right). \tag{38}$$

This equation differs in form from Eq. (32) because of the inherent variability of K_v. The quantity K_v, dimensionally the same as kinematic viscosity, is commonly called simply *eddy viscosity*. It is customary to omit the bars in expressions like $\partial\bar{u}/\partial t$ and $\partial\bar{u}/\partial z$ when the context shows without ambiguity that u represents the mean velocity component.

In the same way, the instantaneous transfer of water vapor in the direction of increasing z is, in mass per unit area and per unit time,

$$\rho wq = \rho w'(\bar{q} + q').$$

The mean rate of transfer is therefore $\rho\overline{w'q'}$. Writing this as

$$\rho\overline{w'q'} = -\rho K_d \frac{\partial \bar{q}}{\partial z} \tag{39}$$

defines the *eddy diffusivity* K_d, also of the same dimensions as kinematic viscosity. If water vapor is transferred by eddy diffusion alone,

$$\frac{\partial \overline{q}}{\partial t} = \frac{\partial}{\partial z}\left(K_d \frac{\partial \overline{q}}{\partial z}\right). \tag{40}$$

Eddy diffusivity is defined above in terms of the particular ratio of eddy diffusion of water vapor to the mean gradient of water vapor. However, it is physically clear that eddy diffusivity is a property of the unsteady flow and is essentially identical for water vapor and any other component of the air such as carbon dioxide or for the eddy transfer of heat. Consequently the eddy heat transfer is given by a relation similar to Eq. (39),

$$c_p \rho \overline{w'T'} = -c_p \rho K_d \frac{\partial \overline{T}}{\partial z}, \tag{41}$$

and the resulting rate of temperature change is

$$\frac{\partial \overline{T}}{\partial t} = \frac{\partial}{\partial z}\left(K_d \frac{\partial \overline{T}}{\partial z}\right). \tag{42}$$

The value of the eddy diffusivity K_d in these two equations is essentially the same as its value in the preceding two equations for the same time and place.

Equations (41) and (42), given in this form for simplicity, are correct when the absolute value of the lapse rate is large compared with the adiabatic lapse rate g/c_p. In general, $\partial \overline{T}/\partial z$ in these equations is to be replaced by $\partial \overline{T}/\partial z + g/c_p$.[1]

At first sight the same identity might be expected to exist between eddy viscosity and eddy diffusivity, but in some cases they differ considerably. A hypothetical example will clarify this statement. Suppose that the turbulence consists of random motion back and forth in the z-direction of air parcels contained in balloons that have no mass, are impermeable to dry air, but are completely permeable to water vapor and are thermally conductive. If the motion back and forth is repeated at a sufficiently rapid rate in relation to the size of the parcels, the eddy transfer of heat and water vapor is inefficient because in each excursion there is time for only partial equalization of temperature and water-vapor concentration by conduction and diffusion within the parcels. No matter how rapid the rate, however, the equalization of momentum can be good because it does not depend on diffusion within the parcels but is effected throughout each parcel by pressure gradients transmitted from the walls of the balloon. For rapidly repeated motion it is evident, therefore, that the ratio of eddy viscosity to eddy diffusivity, K_v/K_d, is greater than

[1] D. Brunt, *Physical and Dynamical Meteorology*, 2d ed., University Press, Cambridge, 1939, p. 225.

unity, with no apparent limit on how large the ratio may be. Also, because the eddy transfer of heat and that of water vapor depend on conduction and diffusion within the parcels, the ratio of eddy diffusivity for water vapor in air to eddy diffusivity for heat can be somewhat greater than unity. The upper limit for this ratio is given by the ratio of diffusivity of water vapor in air to thermometric conductivity, $\nu_d/\nu_c = 1.2$, but highly idealized conditions are required for this upper limit to be reached and little error can be introduced by assuming the ratio always to be unity.

Another hypothetical case is also of special interest. Suppose that the random motion of the parcel in each balloon is discontinuous in such a way that a relatively long period of constant z is followed by a rapid finite change of z, and so on. If these periods of changing z are sufficiently short, no exchange of heat or water vapor can occur during them, but complete equilibrium can be reached during the long periods of rest. The transfer of heat and water vapor accomplished in each change of z approaches complete efficiency because the parcel starts in a condition identical with its surroundings, moves without suffering any change, and delivers its entire excess or deficit of heat and water vapor at the end of its journey. In this process the eddy diffusivity of heat and water vapor are, of course, identical. But during any change of z the parcel experiences a pressure gradient that, at least on the average, reduces the difference in x-component of momentum between the parcel and the air through which it is traveling. The parcel therefore completes its journey after losing some of its excess or lack of momentum, with the result that the eddy transfer of momentum fails of complete efficiency. In this case the ratio of eddy viscosity to eddy diffusivity is less than unity, with no apparent limit on how small the ratio may be.

Although these examples are hypothetical, their less idealized counterparts, which lack the balloons to confine each parcel completely, appear realistic enough, and the examples serve to give some physical understanding of the relation between eddy viscosity and eddy diffusivity.

From observed conditions it appears permissible as a rough first approximation to assume that the ratio of eddy viscosity to eddy diffusivity is unity if the hydrostatic stability is close to zero or is negative. Most of the evidence in support of this assumption is indirect, namely, the fact that use of the assumption in a number of problems has led to answers that are not unreasonable. There is also a little direct evidence[1] that the ratio is unity in certain meteorological situations.

[1] Especially the following: H. U. Sverdrup, "The Eddy Conductivity of the Air over a Smooth Snow Field," *Geofysiske Publikasjoner*, **11**, No. 7 (1936); "On the Influence of Stability and Instability on the Wind Profile and the Eddy Conductivity Near the Ground," *Proc. Fifth Intern. Congr. Applied Mech.*, Wiley, New York, 1938, pp. 369–372.

In case of great stability it is usually recognized that the ratio of eddy viscosity to eddy diffusivity is considerably greater than unity. This is reasonable in view of the first example above. In stable air a parcel displaced vertically from its origin experiences a hydrostatic force toward its origin; hence the average parcel returns quickly after transferring only a small part of its excess or deficit of heat and water vapor, whereas the momentum transfer can be more efficient.

Friction and diffusion are very different at a rough boundary, even though the ratio of eddy viscosity to eddy diffusivity may be unity within the outer part of the turbulent boundary layer. The momentum transfer between the boundary and the fluid is partly by means of a true shearing stress, but it is largely accomplished by normal stresses between the protuberances of the boundary and the eddies in the fluid. Eddy viscosity is therefore effective right up to the boundary, despite the fact that the boundary itself is coated with a laminar sublayer. Heat transfer, however, occurs by true conduction through this layer, and water-vapor transfer occurs by true diffusion. Eddy diffusivity is therefore effective only up to the laminar layer.

In spite of these two major exceptions, namely, stable layers and a region close to any rough boundary, there are important cases where eddy viscosity and eddy diffusivity are of approximately the same magnitude. As rather more basic information on velocity distribution and eddy viscosity is available, they too must be included in a consideration of the distribution of eddy diffusivity in the atmosphere.

3·12. Layer of Frictional Influence in Neutral Equilibrium.—The layer of frictional influence is the part of the atmosphere in which the wind distribution is influenced directly by ground friction, as revealed by a deviation in magnitude and direction from the free-air wind. Close to the ground the wind is usually weaker than the free-air wind and usually has a component of considerable magnitude in the horizontal direction of decreasing pressure. The height at which the deviation from free-air wind ceases is indistinct and varies between a few hundred and a few thousand feet.

The simple theoretical approach in studying this layer is to assume steady flow and a uniform horizontal pressure gradient. Above the layer of frictional influence, geostrophic wind prevails, that is, the wind is normal to the horizontal pressure gradient and its speed is such that, for straight horizontal (geodesic) flow relative to the rotating earth, the necessary horizontal absolute acceleration is produced by the pressure gradient. Close to the ground, friction retards the wind, which consequently blows with a component in the horizontal direction of decreasing pressure; the horizontal pressure gradient is balanced partly by the absolute acceleration of the reduced wind, partly by friction. The hodograph of the complete wind distribution for steady flow, called the *Ekman spiral,* is pictured in

Fig. 3·6. According to this theory, the height at which the wind becomes approximately geostrophic, or the thickness of the layer of frictional influence, is a few feet, in contrast with the observed height of a few hundreds or thousands of feet. Nevertheless, except for the absolute difference in height, the hodograph of the actual mean wind forms a spiral that appears somewhat similar to the Ekman spiral.[1]

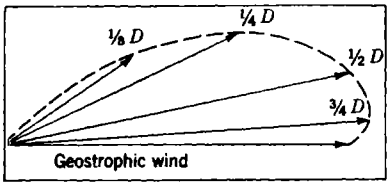


FIG. 3·6.—Ekman spiral in the northern hemisphere. Arrows represent wind at successive levels.

The Ekman spiral follows from the assumption of steady flow. For turbulent flow, if the eddy viscosity is assumed to be independent of height, the acceleration of the mean flow is given by an expression of the same form as that which appears in the case of steady flow [see Eqs. (32) and (38)]. If the kinematic viscosity is replaced by a much larger eddy viscosity, the shape remains the same but the theoretical thickness of the layer of frictional influence can be increased to agree with actuality. The resulting spiral cannot, however, be regarded as a useful approximation to actual conditions, because the eddy viscosity has been assumed constant throughout the layer of frictional influence.

The wind distributions that would accompany certain other mathematically simple distributions of eddy viscosity have been investigated theoretically. The one that best represents natural conditions is Rossby's.[2] For convenience the layer of frictional influence is divided into two parts: a *turbulent boundary layer*, which, according to the theory, constitutes about 10 per cent of the entire layer, and the remaining part, which has been called the *outer layer of frictional influence*. Eddy viscosity is maximum at the interface. Within the turbulent boundary layer it increases linearly with distance from the ground; and within the outer layer, according to Rossby's theory, it is proportional to the square of the distance below the top of the layer. An example is shown in Fig. 3·7.

Within the turbulent boundary layer Rossby assumes the eddy stress to be constant, which is equivalent to assuming the horizontal pressure gradient and the earth's rotation to be negligible. As a consequence, the direction of the mean wind is constant within the layer and the speed, ac-

[1] Two hodographs of average observed conditions may be found in H. Lettau, *Atmosphärische Turbulenz*, Akademische Verlagsgesellschaft, Leipzig, 1939, reprint, Edwards Brothers, Ann Arbor, 1944, p. 106.

[2] C.-G. Rossby, "A Generalization of the Theory of the Mixing Length with Applications to Atmospheric and Oceanic Turbulence," *Meteorol. Papers*, Mass. Inst. of Technol., **1**, No. 4 (1932); C.-G. Rossby and R. B. Montgomery, "The Layer of Frictional Influence in Wind and Ocean Currents," *Papers in Phys. Oceanog. and Meteorol.*, Mass. Inst. of Technol. and Woods Hole Oceanog. Inst., **3**, No 3 (1935).

cording to Eq. (37), is proportional to the logarithm of height. In the outer layer of frictional influence the eddy stress veers and decreases with increasing height. The hodograph of the mean wind consists of a straight line representing the turbulent boundary layer and, without break, a spiral representing the outer layer.

Rossby's distribution of eddy viscosity has the essential feature that it is maximum within the layer of frictional influence and small at its upper and lower boundaries. This agrees with the fact that turbulence is small in the free air and necessarily small close to a solid boundary. One supposition of the theory is that the air is in neutral hydrostatic equilibrium, so that the turbulence is purely mechanical (not affected by heating from below) and is not subject to restraint by stability. Because exact neutral equilibrium rarely occurs, the details of the solution, especially in the higher levels, cannot be expected to be regularly reproduced in nature. As long as the equilibrium is not far from neutral, however, it is reasonable to assume that eddy viscosity and eddy diffusivity are of approximately the same magnitude throughout the layer of frictional influence.

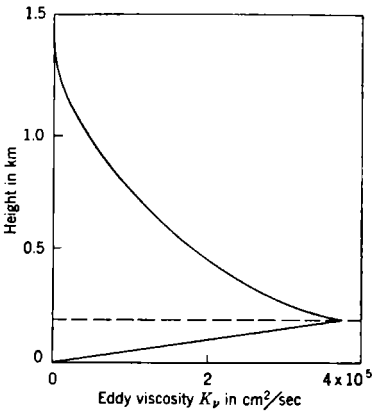


FIG. 3·7.—Theoretical distribution of eddy viscosity in the layer of frictional influence (geostrophic wind 20 m/sec, latitude 43°, roughness length 3.2 cm). (*After Rossby and Montgomery.*)

The turbulent boundary layer, which is separated from the ground or water surface by a very thin laminar sublayer, is the region of important vertical gradients and will be considered in further detail. The outer layer is of less direct concern.

3·13. Logarithmic Distributions in the Turbulent Boundary Layer.—The turbulent boundary layer of the atmosphere is similar to turbulent boundary layers in the ocean, in channels, in pipes, and along flat plates. The basic phenomena are common to all these problems, and the available information stems from them all. The most precise information comes from controlled experiments on the flow of liquids through circular pipes, partly because of the simple boundary conditions. This information forms the foundation for much of what follows.

The three basic assumptions in treating the atmospheric turbulent boundary layer have already been mentioned. One is that the eddy stress is vertically constant, which amounts to neglecting the horizontal pressure gradient within the layer as well as the earth's rotation. A consequence of this assumption is that the mean wind direction is inde-

pendent of height. It results in a definite but small error and in great simplification. The second assumption is that over each horizontal surface the mean flow and eddy stress, as well as such quantities as the mean humidity, are constant. The third is that neutral hydrostatic equilibrium exists. Some effects of departures from neutral equilibrium will, however, be discussed.

It is found quite generally, both in laboratory experiments on the flow of fluids through pipes and along flat plates and in observations of wind, that close to a boundary, with turbulent flow, the mean velocity increases linearly with the logarithm of distance z from the boundary. In other words, the mean velocity is proportional to $\ln(z/z_0)$, where z_0 is a constant length. In order that the relation may hold right up to the boundary, the more general one

$$u \propto \ln \frac{z + z_0}{z_0}$$

is preferable. Because observations agree with the first form, it is seen that $z_0 \ll z$ for values of z within the observational range. It is to be understood here and in what follows that u represents the magnitude of the mean velocity, not an instantaneous value.

Controlled experiments on flow along a rough boundary show that the eddy stress, here written τ, is proportional to density and to the square of the mean velocity at any chosen distance z; thus

$$\tau = \rho\gamma^2 u^2. \tag{43}$$

The nondimensional quantity γ defined by this equation is the *resistance coefficient*;[1] it depends on z and on the character of the boundary but is independent of the speed for turbulent flow past a rough boundary.

The last equation can be written

$$u = \frac{1}{\gamma}\sqrt{\frac{\tau}{\rho}}.$$

The complete expression for mean velocity is therefore

$$u = \frac{1}{\kappa}\sqrt{\frac{\tau}{\rho}} \ln \frac{z + z_0}{z_0}. \tag{44}$$

The nondimensional quantity κ appears to be the same for a wide range of problems and is therefore called the *universal turbulence constant.* A recent analysis[2] of experiments on water flowing through pipes yields the value $\kappa = 0.45$. Solving the last two equations shows how the

[1] This particular definition is peculiar to meteorological usage.

[2] R. B. Montgomery, "Generalization for Cylinders of Prandtl's Linear Assumption for Mixing Length," *Ann. New York Acad. Sci.*, **44**, Art. 1, "Boundary-layer Problems in the Atmosphere and Ocean," pp. 89–104 (1943).

resistance coefficient γ depends on z and z_0:

$$\gamma = \frac{\kappa}{\ln \dfrac{z + z_0}{z_0}}. \tag{45}$$

It follows from the preceding statements that the length z_0, called the *roughness length,* depends on the nature of the boundary only.

By differentiating Eq. (44) and substituting from Eq. (43), one finds that

$$\tau = \rho\kappa\gamma u(z + z_0) \frac{\partial u}{\partial z}.$$

From its definition the eddy viscosity is therefore

$$K_v = \kappa\gamma u(z + z_0), \tag{46}$$

or, according to Eq. (43),

$$K_v = \kappa\sqrt{\frac{\tau}{\rho}}\,(z + z_0).$$

Except very close to the boundary $z_0 \ll z$, and thus the eddy viscosity corresponding to the logarithmic velocity distribution is directly proportional to distance from the boundary. Furthermore it is directly proportional to the mean velocity and to the resistance coefficient. The latter is, in turn, a function of the roughness length and of the distance to the boundary from the point where the velocity is measured.

The preceding discussion of the turbulent boundary layer is applicable when the boundary is hydrodynamically *rough.* The eddy stress is transmitted to such a boundary (as was described in Sec. 3·11) in part directly by normal stresses against the sides of the protuberances. A laminar sublayer exists, and some of the eddy stress is transmitted to the boundary by true shearing stress. However, the size of the protuberances is greater than the thickness of the laminar layer, so that the structure of the boundary is not changed by the laminar layer. In contrast, a hydrodynamically *smooth* boundary is one with protuberances smaller than the thickness of the laminar layer. The protuberances are submerged in the laminar layer and the only tangential stress on the boundary is the one transmitted by true shearing stress. Just as with a rough boundary, the velocity distribution is logarithmic and the eddy viscosity is proportional to distance from the boundary. For a smooth boundary, however, the resistance coefficient is not entirely independent of velocity, the roughness length takes on another significance, and there are other differences. The theoretical treatment depends on separate solutions for the laminar and turbulent layers.[1]

[1] For further details see R. B. Montgomery, "Observations of Vertical Humidity Distribution Above the Ocean Surface and Their Relation to Evaporation," *Papers in Phys. Oceanog. and Meteorol.,* Mass. Inst. of Technol. and Woods Hole Oceanog. Inst., **7,** No. 4 (1940).

According to Rossby's theory, the thickness of the atmospheric turbulent boundary layer is proportional to the wind speed and depends on the character of the surface.[1] As a rough, round number, a thickness of 100 ft is indicated for a wind speed of 10 mph. A good approximation to logarithmic wind distribution and linear eddy-viscosity distribution cannot be expected, however, throughout the entire layer. These distributions and the vertically constant eddy stress are closely maintained in the bottom part of the layer, presumably in the lower half if the hydrostatic equilibrium is neutral. In the upper part, the eddy stress is relatively less constant, so that the deviations from theoretical distributions are more pronounced. With stable equilibrium, these distributions are found in a surface layer having only a small fraction of the thickness for neutral equilibrium. It is, of course, important not to assume these simple distributions to extend higher than is permissible in any case under consideration.

The vertical distribution of water vapor in the turbulent boundary layer is much like that of wind speed. The upward transport of water vapor was written [Eq. (39)] as

$$E = -\rho K_d \frac{\partial q}{\partial z}.$$

In the turbulent boundary layer, if the hydrostatic equilibrium is not far from neutral, the eddy diffusivity and eddy viscosity may be assumed numerically equal; therefore Eq. (46) may be used for eddy diffusivity also. Furthermore, just as eddy stress is assumed constant in the layer, so may the eddy transport of water vapor E be assumed constant and equal to the rate of evaporation from the sea surface. In view of the comparatively small thickness of the turbulent boundary layer compared with the entire frictional layer throughout which water vapor from the surface is distributed, this assumption of constant transport is a good first approximation.

From Eqs. (39) and (46),

$$\frac{\partial q}{\partial \ln \dfrac{z + z_0}{z_0}} = -\frac{E}{\rho\kappa\gamma u}, \tag{47}$$

where γ depends on the level chosen for measuring u [the product γu, furthermore, is independent of z, *cf.* Eq. (43)]. Integration from $z = b$ to $z = z$ gives

$$q = q_b - \frac{E}{\rho\kappa\gamma u} \ln \frac{z + z_0}{z_b + z_0}, \tag{48}$$

showing that specific humidity is a linear function of the logarithm of height within the turbulent boundary layer.

[1] Rossby and Montgomery, *loc. cit.*

Water evaporated from the surface diffuses through the laminar sublayer before reaching the turbulent boundary layer. Equation (47) cannot, therefore, be applied all the way down to the surface.[1] However, the humidity gradient must clearly be proportional to the difference in specific humidity $q_0 - q_b$ between the surface and any fixed level b. The quantity Γ defined by

$$\frac{\partial q}{\partial \ln \dfrac{z + z_0}{z_0}} = -\Gamma(q_0 - q_b) \tag{49}$$

is therefore independent of $q_0 - q_b$ but, of course, depends on b and on the distribution of eddy diffusivity. Because Γ is defined in terms of the vertical humidity distribution, it is easily determined from measurements. From Eqs. (47) and (49)

$$E = \rho\kappa\gamma\Gamma u(q_0 - q_b). \tag{50}$$

Because of its importance in this equation for evaporation, Γ is called the *evaporation coefficient.*

The preceding discussion of water vapor applies equally well to the distribution of temperature and the vertical transport of heat by conduction and eddy diffusion. Since the conductivity and diffusivity are somewhat different in magnitude in the laminar sublayer (see Sec. 3·11), the value of Γ would, strictly, be slightly different for water vapor and for heat; however, to a first approximation, this difference may be neglected.

Observations at sufficiently low levels to study the turbulent boundary layer over the ocean have been confined to various heights on ships, which cause some disturbing effects. In addition to this limitation, individual series of observations are very irregular, so that the logarithmic distribution shows clearly only when the conditions are averaged.

Humidity distributions analyzed by Montgomery[2] yielded values of the evaporation coefficient largely in the range from 0.04 to 0.10. An analysis by Sverdrup[3] of observations made by Black shows values within the same range for wind speeds up to about 15 mph. For greater wind speeds, ranging up to 30 mph, these observations give a value of about 0.15.

VERTICAL DISTRIBUTIONS IN NEUTRAL AND UNSTABLE EQUILIBRIUM

The structure of the air near the ground or water depends on a number of factors, such as wind speed, character of the surface, and past

[1] For a combined solution of the two layers see Montgomery, *loc. cit.*

[2] *Ibid.*

[3] H. U. Sverdrup, "The Humidity Gradient over the Sea Surface." *Jour. Meteorol.*, **3**, 1–8 (1946).

history of the air. A factor of special importance is the temperature difference between the surface and the air a short distance above. Types of vertical structure occurring when the air temperature is less than or the same as that of the surface are discussed in the following three sections. This condition is also called *negative temperature excess*, unstable hydrostatic equilibrium, or heating from below. Types of vertical structure occurring when the air is warmer than the surface are discussed in Secs. 3·17 to 3·21. The latter condition is also referred to as *positive temperature excess*, stable hydrostatic equilibrium, or cooling from below.

3·14. Heating from Below.[1]—To see the typical effect of heating from below, consider what may be expected to happen on a calm, clear morning as a result of the ground's being warmed by solar radiation. At daybreak the ground is cold as the result of nocturnal radiation, and the air temperature increases with height up to an assumed height of 1000 ft. The air is initially in stable equilibrium and resistant to turbulence or mixing. When the ground surface becomes warmer than the air a few feet above, the structure becomes radically different. Unstable equilibrium is produced; air warmed in contact with the ground rises in patches through the heavier air above it, and the heavier air sinks to be warmed in turn. This atmospheric process is called *convection*. The rising parcels tend to be accelerated upward as long as the surrounding air is colder, but they are decelerated as soon as they reach high enough to come into warmer air. Thus there is a definite top to the convection. Within the central part of the unstable layer the thermal turbulence is relatively intense, because there the ascending and descending parcels have their greatest speed. Close to the ground the turbulence is slight, because an ascending parcel can attain little speed in a short distance. Consequently, the central part of the unstable layer is well mixed, and the temperature lapse rate is large at the ground. The average temperature distribution therefore changes in the manner shown schematically in Fig. 3·8.

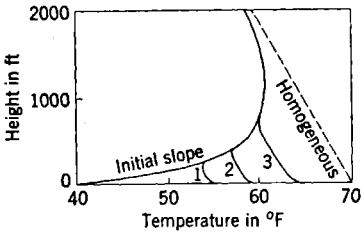


FIG. 3·8.—Successive temperature distributions over land because of heating from below on a calm, clear morning.

Heating from below in calm or light wind thus results in a type of distribution characterized by three layers: (1) a surface layer with a definitely superadiabatic (greater than adiabatic) lapse rate, which increases as the ground is approached; (2) a thick central layer in which the lapse rate is very nearly adiabatic; and (3) an upper layer that is stable.

[1] By Richard A. Craig and R. B. Montgomery.

If the heating from below stops, the superadiabatic surface layer disappears immediately, because its maintenance depends on the heating process. The central layer changes to stable equilibrium but may remain nearly homogeneous for some time before being destroyed by some other process.

If there is wind, mechanical turbulence resulting from boundary friction is present in addition to the thermal turbulence or convection. The mechanical turbulence may predominate to such an extent that the central homogeneous layer does not occur, especially if ground heating is reduced by a cloud cover. A schematic representation of the warming that takes place on a morning when mechanical turbulence predominates over thermal turbulence is given in Fig. 3·9. The initial inversion is, of course, less intense than on a calm, clear morning.

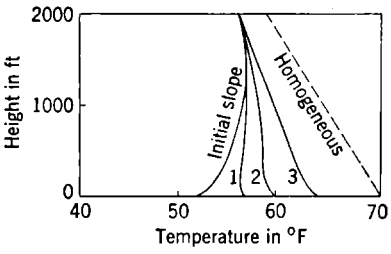


FIG. 3·9.—Successive temperature distributions over land because of heating from below on a windy morning.

With respect to the three layers that are the typical result of heating from below, the central one is nearly homogeneous and is therefore characterized by an approximately standard M-profile. In the stable layer above it there is commonly a lapse of water-vapor concentration. The stable layer is therefore usually superstandard, but occasionally there may be a sufficiently intense inversion of vapor concentration to make the layer substandard. The superadiabatic surface layer is superstandard if the surface is water or wet ground, for the refractive modulus of the air at the surface containing saturated vapor is always greater than the potential refractive modulus of the colder air above, regardless of the humidity of the colder air. If the surface is dry ground, there is no gradient of vapor concentration; therefore the superadiabatic layer is substandard. If the ground is damp, neither dry nor completely wet, the superadiabatic surface layer may be either substandard or superstandard.

A common situation that gives heating from below occurs over the ocean when there is an offshore wind and the surface air on land is colder than the water as the result of nocturnal cooling. The nearly homogeneous layer grows higher as the air moves farther offshore in the same way that it does on land as the morning progresses. An example is shown by the sounding in Fig. 3·10. This sounding has the typical standard homogeneous layer with the superstandard stable layer above it and the superstandard unstable layer below it. The last does not appear in entirety in the sounding, which starts at 20 ft, but is known to exist from the measurement of water temperature. As has already been explained, this is the

common structure resulting from heating from below.[1] However, the stable top layer may be substandard, and the unstable surface layer is substandard over dry ground. Furthermore, the structure may be very different if the wind is sufficiently strong.

Heating from below is very widespread, especially over land in daytime and over tropical oceans. The essentially homogeneous layer is frequently several thousand feet in thickness. At the top of the layer the water vapor is often saturated in patches, giving cumulus or stratocumulus clouds. An example from tropical regions is shown in Fig. 3·11. The water temperature was not measured at the point above which the

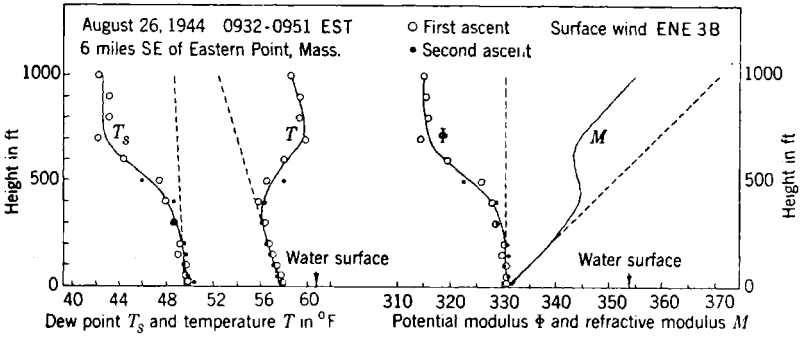


FIG. 3·10.—Heating from below; a sounding in air that has been cooled nocturnally over land and has then passed over warmer water.

sounding was made, but the structure is clearly the result of the air's being warmed from below.

The surface superadiabatic layer is characteristically *shallow*.[2] If it is superstandard, the M-inversion is usually between 5 and 50 ft thick. Thus it may have either no effect or a very large effect on propagation, depending on the thickness of the M-inversion in relation to the wave length and on the antenna height.

Certain terms may be used in describing how the properties of the essentially homogeneous layer differ from those of the surface. The potential temperature of the homogeneous layer less the temperature at the surface is the *temperature excess*, necessarily negative for heating from below. The humidity at the surface less the humidity in the homogeneous layer is the *humidity deficit*, always positive for heating from below over water. The refractive modulus at the surface less the potential refractive modulus in the homogeneous layer is the *modulus deficit*, or M-deficit, always positive for heating from below over water. These

[1] Other similar soundings are presented by R. A. Craig, *Ann. N.Y. Acad. Sci.*, **48**, 783 (1947).

[2] Measurements in the Irish Sea indicate that this layer may be only a few feet thick (see Sec. 4·17).

quantities are especially useful over water, because the temperature of the water surface can often be determined, and it in turn determines the humidity and M at the surface.

The depth of the M-inversion occurring over water depends primarily on the M-deficit, on the temperature excess, and on the wind speed. These relationships will be discussed in the next section.

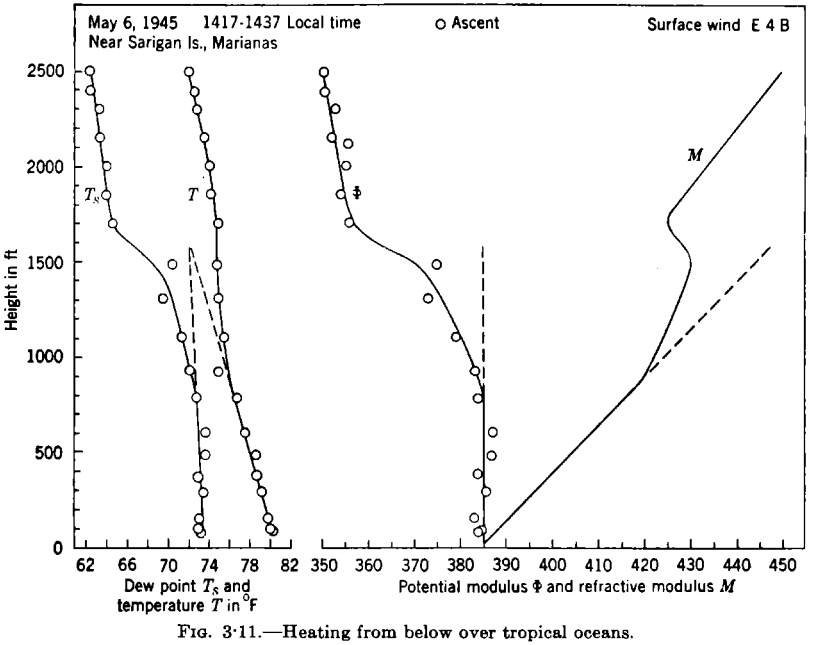


FIG. 3·11.—Heating from below over tropical oceans.

3·15. Application of Logarithmic Distribution.[1]—The logarithmic distribution of temperature and water-vapor concentration near the surface, when the air is in or near neutral equilibrium, can be used to determine theoretically the shapes of M-profiles.[2]

Air in neutral equilibrium relative to the surface, with a temperature excess of exactly zero, rarely occurs in nature. For practical purposes, however, the air may be considered to be in neutral equilibrium when the temperature excess is small (absolute value less than 1°F for the purposes of this chapter). In exact neutral equilibrium the temperature lapse rate is adiabatic from the surface to the top of the layer which is in neutral equilibrium, and the Φ-profile is similar to the humidity distribution.

[1] By Richard A. Craig.

[2] R. B. Montgomery and R. H. Burgoyne, "Modified Index Distribution Close to the Ocean Surface," RL Report No. 651, Feb. 16, 1945.

The case of air over water or saturated ground is the most interesting because of the simple boundary condition; that is, the humidity and M-values at the surface are determined by the temperature of the surface. This condition is also widespread, particularly over the open ocean. Because the vapor cannot be saturated in the lower part of a homogeneous layer, this case is characterized by a lapse of humidity near the surface and a shallow M-inversion.

It has been seen in Sec. 3·13 that under these conditions the specific humidity is distributed logarithmically near the surface and can be expressed by Eq. (49),

$$\frac{\partial q}{\partial \ln \dfrac{z + z_0}{z_0}} = -\Gamma(q_0 - q_b).$$

It has also been shown (see Sec. 3·7) that

$$\frac{dM}{dz} = -r_1 \frac{d\theta}{dz} + r_2 \frac{dq}{dz} + r_3,$$

where r_1, r_2, and r_3 were given in Eq. (23). In this case $d\theta/dz = 0$, so that, if z_0 is neglected in comparison with the values of z that are of interest,

$$\frac{dM}{dz} = -r_2 \frac{\Gamma \Delta q}{z} + r_3. \tag{51}$$

The function $\Delta q = (q_0 - q_b)$ is the humidity deficit if b is sufficiently high to be in the homogeneous layer.

A convenient variable is the *duct height* d (the height of the M-inversion), at which $dM/dz = 0$.

This value of d is given by

$$d = \frac{r_2}{r_3} \Gamma \Delta q. \tag{52}$$

It has been seen that Γ may vary widely (Sec. 3·13). Observational determinations of its value by Montgomery and Sverdrup vary between 0.04 and 0.18. It appears that an average value for neutral equilibrium from these observations is about 0.09. Reducing this to a height of $b = 50$ ft gives $\Gamma = 0.08$. Accordingly,

$$\frac{d}{\Delta q} = 16 \times 10^3 \text{ ft}.$$

If the humidity deficit is expressed in terms of vapor pressure,

$$\frac{d}{\Delta e} = 10 \text{ ft/mb}.$$

It is convenient to express d in terms of the modulus deficit ΔM. For constant potential temperature, integration of Eq. (23) yields

$$\Delta M = r_2 \Delta q.$$

If this is substituted in Eq. (52), it follows that

$$\frac{d}{\Delta M} = \frac{\Gamma}{r_3} = 2 \text{ ft.} \tag{53}$$

It should be borne in mind that this value may conceivably vary between 1 and 4 ft because of the variation of Γ with wind speed and character of the surface. With strong winds (greater than 15 mph) it is apt to be nearer the larger figure, according to Sverdrup. The ratios $d/\Delta q$ and $d/\Delta e$ may vary similarly. It will also be observed that these ratios are independent of the length of time that the air has been over the water.

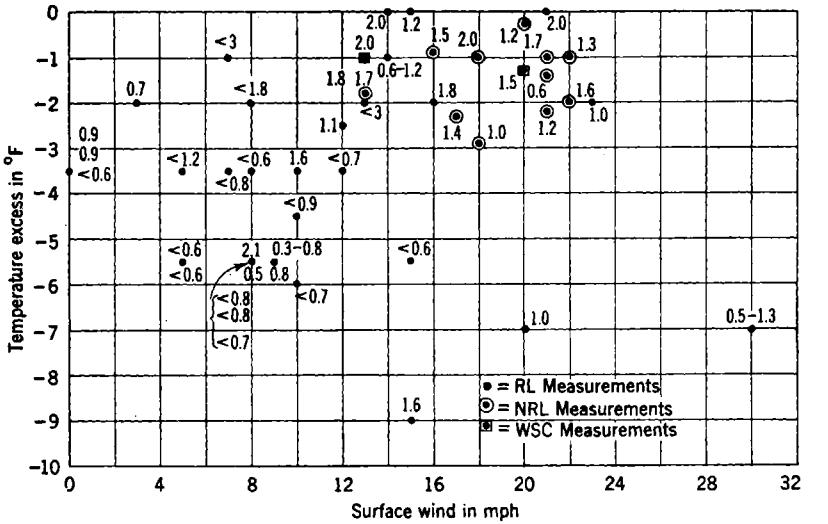


FIG. 3·12.—Some observed values of $d/\Delta M$, in feet, for neutral and unstable equilibrium. The point associated with each number is plotted so that its coordinates indicate the wind speed and temperature excess appropriate to the observation.

When the air is in unstable equilibrium, convective mixing is present in addition to purely mechanical turbulence. Qualitatively this increases the amount of mixing at all levels. However, because the convection is small near the boundary, the mixing is increased relatively more at some distance from the boundary than near it. Hence, large gradients of temperature, humidity, and M tend to be concentrated in an even more shallow layer than is the case in neutral equilibrium. The duct height is thus less in unstable equilibrium than in neutral equilibrium.

Under conditions of unstable equilibrium logarithmic distributions do not occur in a sufficiently thick layer and the ratio $d/\Delta M$ may not be strictly independent of ΔM. The ratio, nevertheless, is a convenient variable in this case also. Some experimental determinations of its value are shown in Fig. 3·12, where $d/\Delta M$ in feet is given by numbers near points that are plotted as a function of temperature excess and

surface wind speed. Circled values were obtained during a measurement program conducted in Antigua,[1] B. W. I. by the Naval Research Laboratory. Points enclosed by squares are derived from measurements made in the Southwest Pacific area by the Washington State College group,[2] and the remainder were obtained in Massachusetts Bay by the Radiation Laboratory (see Sec. 4·2).

These values do not disprove the expectation, based on theoretical considerations, that $d/\Delta M$ should have a value of about 2 ft for neutral equilibrium. They show quite definitely that it is less than 2 ft for unstable equilibrium. The dependence of the ratio on the temperature excess and wind speed is not entirely clear from the data available, but there is a well-marked tendency for the ratio to decrease with increasing instability.

The scatter of the points may be due in part to the fact that the ratio for unstable equilibrium is not strictly independent of ΔM. Moreover, estimation of the duct height is difficult from a measured sounding; and in addition, it may vary rapidly with time because of random motion of the air (Sec. 3·26).

3·16. Rate of Modification of Unstable Air.[3]—A useful estimate can be made of the rate at which the average humidity and refractive modulus change in a case where air in neutral equilibrium moves over the water without condensation within the layer.[4] Because water vapor evaporated from the sea surface is rapidly diffused through the entire homogeneous layer, only the amount of evaporation and the thickness of the layer need be known. The latter can be found from a sounding. The mass of water evaporated from unit surface in unit time for air in neutral equilibrium has been written in Eq. (50) as

$$E = \rho\kappa\gamma\Gamma u(q_0 - q),$$

where q_0 and q are the specific humidities at the water surface and in the air respectively.

The mass of water vapor in any air column of unit area and height z is ρzq. The air travels a distance dx in time dx/u, during which time the mass of water vapor added to the column is $\rho z\, dq$. It follows from Eq. (50) that

$$E\frac{dx}{u} = \rho z\, dq = \rho\kappa\gamma\Gamma\, dx(q_0 - q). \tag{54}$$

[1] M. Katzin, R. W. Bauchman, and W. Binnian, "Three-and nine-centimeter Propagation in Low Ocean Ducts," *Proc.* IRE, **35**, 891 (1947). This report also discusses other empirical information on shallow surface M-inversions.

[2] P. A. Anderson *et al.*, "Results of Low-level Atmospheric Soundings in the Southwest and Central Pacific Oceanic Areas," Washington State College Report No. 9, NDRC Project No. PDRC-647, Contract No. OEMsr-728, Feb. 27, 1945.

[3] By Richard A. Craig.

[4] R. B. Montgomery and R. H. Burgoyne, "Modified Index Distribution Close to the Ocean Surface," RL Report No. 651, Feb. 16, 1945.

Upon integration of Eq. (54), the specific humidity after any trajectory x is seen to be

$$q = q_0 - (q_0 - q_i)e^{-\kappa\gamma\Gamma x/z}, \tag{55a}$$

where q_i is the initial value of the specific humidity. The difference between the specific humidity at the surface and the specific humidity in the homogeneous column decreases exponentially with distance.

After insertion of the values $\kappa = 0.45$, $\gamma = 0.05$, and $\Gamma = 0.08$, Eq. (55a) becomes

$$q = q_0 - (q_0 - q_i)e^{-0.10x/z}, \tag{55b}$$

or

$$\Delta q = \Delta q_i e^{-0.10x/z},$$

where z is measured in units of 100 ft and x in miles.

Because ΔM is directly proportional to Δq in neutral equilibrium,

$$\Delta M = \Delta M_i e^{-0.10x/z}. \tag{56}$$

When air is being heated from below, these equations cannot be expected to serve as more than rough estimates. In the first place, the rate of evaporation is even more doubtful when the air is unstable than when it is in neutral equilibrium. In the second place, the height of the homogeneous column changes with time, as is shown schematically in Fig. 3·13, its upper limit being that height at which the air has initially the same potential temperature as the water. If the initial lapse rate is known, this fact can be considered in the integration, but so many initial distributions of temperature are possible that no general result can be given. In this case, it is best to determine the modification graphically by using the following approximate rule that follows from Eq. (55b): The temperature deficit (or the humidity deficit) is reduced 10 per cent after the air has traveled z miles, where z is the height of the homogeneous layer in hundreds of feet. After each successive application of the rule, a new value of z may be observed on the new temperature and humidity curves. As

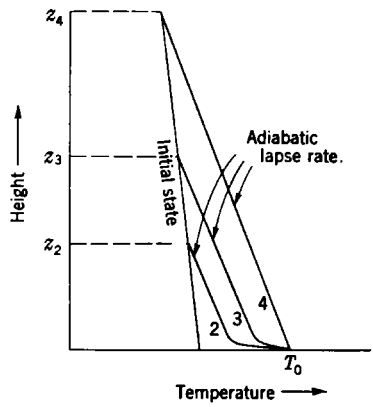


FIG. 3·13.—Heating from below by water at temperature T_0. Curves 2 and 3 represent conditions after the air has been partially heated; Curve 4 represents the final condition.

$$\Delta M = -r_1 \Delta\theta + r_2 \Delta q,$$

the same rule applies to the M-deficit.

It is important to remember that because of the doubt attached to Eq. (50) in unstable conditions, this result should be relied upon for an approximate answer only.

In a somewhat different approach to the problem, Burke[1] has prepared graphs for determining the temperature over the ocean of an air column that is being heated from below. This method has been applied to several actual cases with trajectories of 350 to 1200 miles and temperature deficits of 10° to 35°F. The forecast temperatures were generally within 4°F of the observed temperatures. The "10 per cent rule" for temperature has been applied to several of these cases and appears to yield approximately the same results.

VERTICAL DISTRIBUTIONS IN STABLE EQUILIBRIUM

By Richard A. Craig

The theory for eddy diffusion has not been developed to the point where it can be applied quantitatively when the air is in stable equilibrium. Consequently, the discussion of vertical distributions in stable equilibrium must be essentially qualitative or empirical. Some of the information that makes possible such a discussion is available in standard meteorological texts, but a great part of it is based on low-level soundings of the atmosphere that were made during the war.

3·17. Cooling from Below.—An air column is said to be cooled from below when the temperature of the air a short distance above the underlying land or water surface is greater than that of the surface. Such air is in stable equilibrium because its potential temperature increases with height. Vertical motions are greatly inhibited in air in stable equilibrium, as can be seen from the following simplified argument. Consider an air parcel at any given level in the atmosphere, and let it rise to some higher level. Cooling adiabatically, the parcel retains its original potential temperature and finds itself among other parcels which have higher potential temperatures and are correspondingly lighter. Hence it tends to sink to its original level. Similarly if it descends to some lower level, it finds itself in cooler and heavier air and again tends to return to its original position.

One essential difference between heating from below and cooling from below becomes apparent. In the former case the modification of the air may extend to great heights, limited only by the height at which the air is potentially warmer than the modifying surface. In the case of cooling from below, where convection is by definition impossible, the modification can be effected only by mechanical turbulence and does not extend above

[1] C. J. Burke, "Transformation of Polar Continental Air to Polar Maritime Air," *Jour. Meteorol.*, **2**, 94 (1945).

the layer of frictional influence. Indeed, it is often confined to the lowest few hundreds of feet.

A second consideration with respect to cooling from below is that large temperature, humidity, and M-gradients can extend to relatively high levels. The previous section shows that under conditions of neutral or unstable equilibrium the mixing by eddy processes or convection increases in the lower part of the atmosphere from a value of zero at the surface, and large gradients of the meteorological elements such as superadiabatic lapse rates of temperature are confined to approximately the lowest 50 ft of the atmosphere. However, in stable equilibrium with cooling from below, the mixing is decreased at all levels to values approximating those near the surface, and large gradients comparable to ones existing in unstable equilibrium may exist as high as a few hundred feet above the ground.

A third essential difference between the two cases is in the effect of the time element on the depth of the surface M-inversion over water. It was pointed out in Sec. 3·14 that this depth is primarily determined for heating from below by the wind speed, the temperature excess, and the M-deficit.[1] In the case of cooling from below a fourth variable must be added, namely, the length of time that the air has been subjected to the modification process. A discussion of the effect of this variable will be found in Sec. 3·19.

The M-deficit is particularly important, as it determines whether the modification of the air by the underlying surface will produce a superstandard, standard, or substandard surface layer. Over *dry* ground, it is necessarily positive for cooling from below; thus a superstandard layer is always formed. Over water or *saturated* ground, it may be positive, zero, or negative. These three possibilities are illustrated in Figs. 3·14, 3·15, and 3·16, which show Φ- and M-profiles with superstandard, standard, and substandard[2] surface layers. All three soundings were made in air

[1] The terms temperature excess, humidity deficit, and M-deficit in the case of heating from below were defined in terms of the difference between the surface values and values in the homogeneous layer. They cannot be so rigidly defined for cooling from below if there is no homogeneous layer as a reference point. They are, however, convenient terms to use. The reference point in stable air is usually taken above the layer to which modification has extended, sometimes immediately above and sometimes at some standard level such as 1000 ft. If the air being cooled was initially homogeneous, potential values in this homogeneous layer are often used.

[2] No fog accompanies the substandard layer pictured in Fig. 3·16. It is frequently stated that fog, called advection fog, forms whenever warm, moist air is cooled by contact with the underlying water surface because its temperature falls to its dew point. According to available data this is a misleading statement. This question has been discussed by R. B. Montgomery and G. Emmons, *Jour. Meteorol.*, **4**, 206 (1947).

Examples of warm, moist air are shown in Figs. 3·15 and 3·16. The air close to the water surface is losing not only heat but water vapor also. The water vapor is condensing on the ocean surface, a process exactly opposite to that in Fig. 3·14, where the air is gaining water vapor through evaporation. If advection fog were to be produced by diffusion and mixing in warm, moist air, the examples in Figs. 3·15 and 3·16 would appear to be ideal situations. Actually no fog occurs at these times.

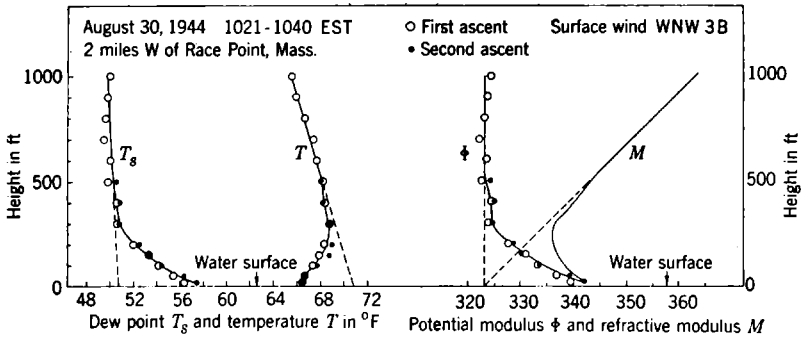


FIG. 3·14.—Superstandard surface layer produced by cooling from below over water with positive M-deficit.

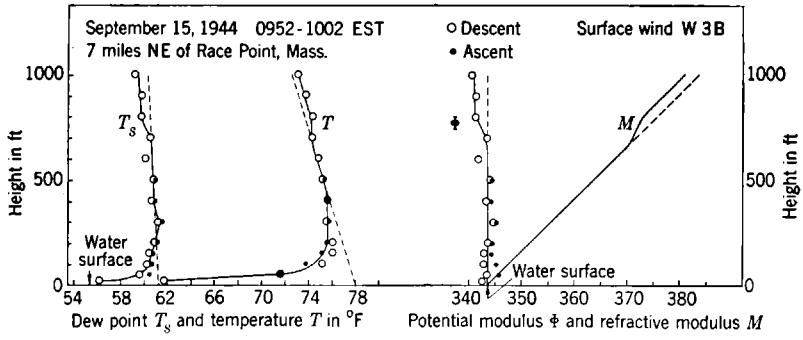


FIG. 3·15.—Standard surface layer produced by cooling from below over water with zero M-deficit.

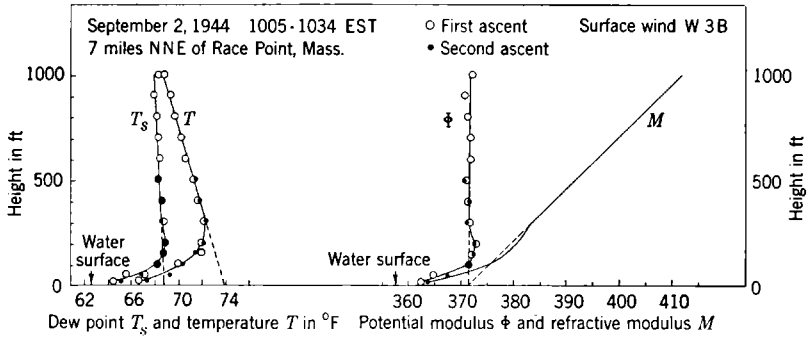


FIG. 3·16.—Substandard surface layer produced by cooling from below over water with negative M-deficit.

In warm air, whether or not the humidity deficit is negative, the stability permits high relative humidity to be produced throughout a layer of appreciable thickness close to the cold water. The explanation of the formation of advection fog appears to

that had been heated over land and had then traveled 25 to 50 miles over the colder waters of Massachusetts Bay.

Both temperature excess and wind speed affect the amount of mechanical mixing that can occur. Light winds and a large temperature excess lead to a small amount of mixing; strong winds and a small temperature excess lead to a relatively large amount of mixing. In the former case, it may be expected that the effect of the modification will be confined to a very low layer of the atmosphere with a temperature inversion occurring

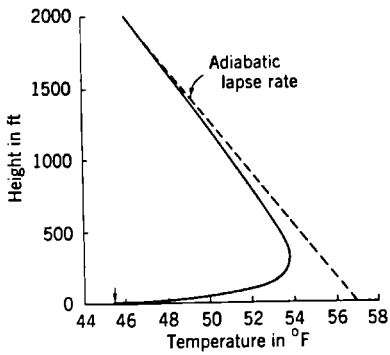


FIG. 3·17.—Cooling from below with light winds and large temperature excess.

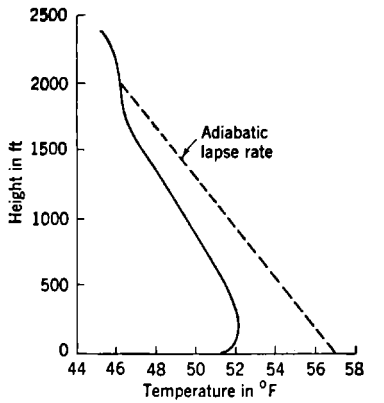


FIG. 3·18.—Cooling from below with strong winds and small temperature excess.

in this layer. In the latter case the mixing is comparatively large and the height to which the air is affected is correspondingly greater, the ultimate limit being the height of the layer of frictional influence. With this large amount of mixing, conditions are more nearly like those which occur during heating from below. There is usually a shallow surface temperature inversion, a thick layer characterized by nearly homogeneous conditions because of the mechanical mixing, and above this another stable layer or temperature inversion. Such an elevated inversion is commonly referred to as a *turbulence inversion*. For moderate amounts of mixing there is ordinarily a temperature inversion at some intermediate height. Significant gradients of humidity and refractive modulus are generally found at the same heights as the large gradients of temperature. Figures 3·17 and 3·18 indicate schematically the extreme cases above. The

be that if a relatively small amount of further cooling by radiation, probably requiring 12 to 36 hr, will bring the air to the dew point, then fog may be produced under favorable conditions.

It follows from the preceding reasoning, verified by many actual measurements, that fog is neither a necessary nor a sufficient condition for the occurrence of a surface substandard layer over water. See R. B. Montgomery, "Modified Index Distribution Close to the Ocean Surface," RL Report 651, Feb. 16, 1945.

relative importance of wind speed and temperature excess is not yet known, but some observations with bearing on the subject will be presented in Sec. 3·19.

There are two important cases in practice that involve cooling from below. The first occurs near the coasts of continents when air is warmed over land by a process of convection as a result of solar heating and is subsequently blown over cooler coastal waters. The second occurs over land at night because of the cooling of the land surface by nocturnal radiation. These may be discussed separately, as there are some considerations that are not common to the two problems.

Cooling from Below over Water.—When a warm air column is blown over cooler water, there are three important processes that always act to change the initial vertical distributions of temperature, humidity, and refractive modulus. The first is radiation from the air column; the second is a shearing effect caused by the change of wind velocity with height; and the third is an eddy diffusive process transferring heat from the air to the cooler water and transferring water vapor in the direction of decreasing humidity.

Radiation acts to decrease the temperature of the air column. The amount of radiation that occurs is variable, depending mainly on the amount of water vapor in the air; however, the effect on the temperature is of the order of magnitude of 0.2°F per hr.[1] The order of magnitude of the changes because of contact with the cooler water is 2°F per hr. It is safe to neglect the effects of radiation when the interval of time involved is reckoned in hours; when the time interval is a matter of days, the effect becomes appreciable.

Shear is a complicating factor that is different from the other two processes in that it involves no loss or gain of heat by the air. It is mentioned here because it has a very real effect on the types of vertical distribution found over water. It will be discussed more fully in the following section.

The third process, that involving eddy diffusion, is the most important. There are two concepts that should be recalled in this connection. (1) The air at the common boundary between the air column and the water must assume the temperature of the water surface and must assume as well the vapor pressure and refractive modulus that are determined by the water temperature (Sec. 3·4). (2) The diffusive processes that are most important in affecting the modification are the same for the vertical transport of both heat and water vapor (Sec. 3·11). This leads to the principle of similarity, which, as was pointed out in Sec. 3·8, tends to produce a straight line when the values within the modified layer are plotted on a characteristic diagram. In particular, if the air was initially homogeneous, the straight line joins the point representing the unmodified

[1] B. Haurwitz, *Dynamic Meteorology*, McGraw-Hill, New York, 1941, pp. 105-107.

air and the point representing the water temperature. It is important to remember that radiation, shearing processes, and changes in the water temperature act to destroy this straight line; hence the rule must be used with care. Figure 3·19 is a sounding made over water in air that had been cooled from below and illustrates the principle of similarity. Figure

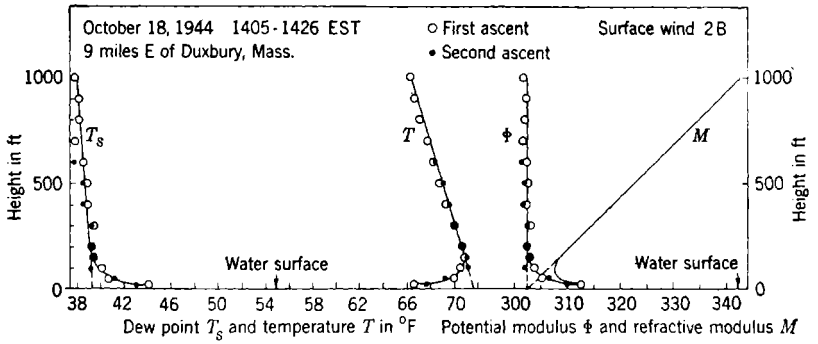


FIG. 3·19.—Similarity in an overwater sounding.

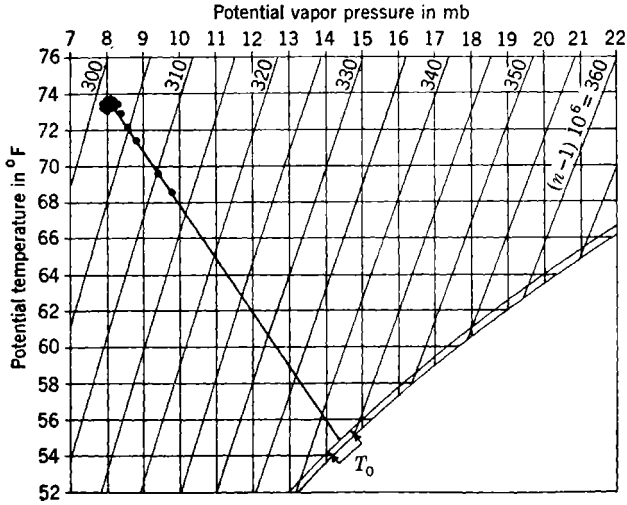


FIG. 3·20.—Similarity in an overwater sounding; a characteristic diagram. The family of curves represents Φ for a surface pressure of 1000 mb.

3·20 is the corresponding characteristic diagram. The indicated water temperatures were measured 3 hr after the sounding was made.

Even if radiation and shearing effects are neglected, the problem of the progressive modification of warm air over cold water is very complicated. It is convenient to consider first the simplest case, where the air is initially homogeneous and where the temperature of the water over

which the air moves remains constant. This problem has been studied rather extensively at the Radiation Laboratory for overwater trajectories of less than 35 miles, and some empirical data will be presented in Sec. 3·19.

Cooling from Below over Land.—The cooling of air over land at night results from the cooling of the earth's surface caused by *nocturnal radiation.* Nocturnal radiation is the difference between the radiation emitted by the earth's surface and the radiation absorbed from the atmosphere. It is present during the day as well as the night, but the name "nocturnal" is given because at night, when solar radiation is absent, it represents the entire radiative transfer at the earth's surface.

The transfer of heat from the air to the cooler surface is largely effected by the same physical process of turbulence as in the overwater case. Radiation and shearing effects are also present. The problem is complicated by the fact that the temperature of the ground does not fix the surface values of humidity and M as does the temperature of a water surface. If the modifying surface is saturated, a superstandard, standard, or substandard layer may result from the modification, as in the overwater case. If the ground is damp, the layer may again be superstandard, substandard, or standard, depending on the M-deficit. If the ground is dry, so that there is no gradient of vapor pressure, a superstandard layer always forms as a result of the temperature inversion.

A further discussion of the problems involved in nocturnal cooling will be found in Sec. 3·21.

3·18. Shear in Stable Equilibrium.—The variation with height of the horizontal component of the wind velocity, which will be referred to here as *shear*, is a process that involves no transfer of heat or water vapor from or to the atmosphere but which may affect the M-profile. It occurs regularly in the atmosphere under all conditions of stability; the present discussion is, however, limited to cases of stable equilibrium because only then can the effects of shear be detected and assessed in the analysis of low-level soundings.

One cause of shear is the variation of horizontal pressure gradient with height. In the layer of frictional influence, however, this factor is usually outweighed by the more regular effect of friction. The effect of friction on the wind in this layer under conditions of neutral equilibrium has been discussed in Sec. 3·12. It may be recalled here that the wind ordinarily veers and increases with height between the surface and the top of the layer of frictional influence. The effect of thermal stability is not accurately known, but qualitatively it appears to increase the two effects described above; it is likely also that shear tends to be concentrated in any temperature inversions that may exist.

Shear is particularly important in the study of the modification of warm air over cold water. If the wind speed and direction change with height, the air at different levels in any air column over the water must

have varying overwater trajectories and land origins. Thus the initial properties of the air may vary with height.

An interesting and regular effect on overwater M-profiles arises from the differences in the time at which the air at various heights leaves the land. Consider a day when the air is being convectively heated over land in the manner described in Sec. 3·14. The potential temperature of the air in the homogeneous layer over land increases steadily throughout the morning, and usually the humidity decreases because of mixing with drier air aloft as the mixed layer grows in height. At any distance offshore the air at progressively higher levels will have left land later because of the increase of wind speed with height and will be correspondingly warmer and drier. This will increase the thermal stability of the air column, a process that may be referred to as *shearing stratification.* This process forms a superstandard refracting layer. Although this is the usual case, it is important to remember that the lower air may on occasion be drier than the air above, a condition that could result in a standard or substandard layer.

If, on the other hand, the air is cooling over land during the late afternoon or evening, so that cooler air overruns warmer air offshore, unstable equilibrium may be set up, resulting in convection and a homogeneous or nearly standard layer. The amount of shear in this case is more difficult to predict because of the various types of cooling that may occur over land.

The variables that govern the amount and type of shearing stratification are complicated. The important ones include

1. Variation of wind velocity with height, which probably varies with distance from the coast because of changes in stability.
2. Orientation of wind directions with respect to the coast line.
3. Variations in initial properties of the air at the various sources because of variations in time or distance.
4. Distance from shore.

Shearing stratification, aside from its effect on the Φ- and M-profiles, greatly increases the difficulties involved in studying the eddy-diffusion processes that are effective in the cooling of warm air over cold water. (1) It tends to increase the stability of the air and to lessen the mixing that may occur. (2) It makes difficult an estimate of the original properties of the modified air and, hence, of the extent of the change that the air has undergone because of modification from below. Although shear is always present to some extent, the analyst needs to estimate its effect in any given case and, for quantitative results, to work with cases where that effect is a minimum.

A method that may be used for detecting shearing stratification, aside from a detailed consideration of the wind structure and variations of temperature and humidity with time and space over land, involves the

use of the characteristic diagram. It was pointed out in Secs. 3·8 and 3·17 that for initially homogeneous air over colder water of uniform temperature, in the absence of radiation and shear, eddy-diffusion processes lead to a characteristic curve that is a straight line. Shear processes, except for rare coincidences, lead to a characteristic curve that is not a straight line. This criterion is helpful in analyzing a sounding made over water when the air is known to have been initially homogeneous and radiation is negligible.

Shear may be expected to be an especially important factor when the temperature excess is large, causing greater stability, or when the trajectory

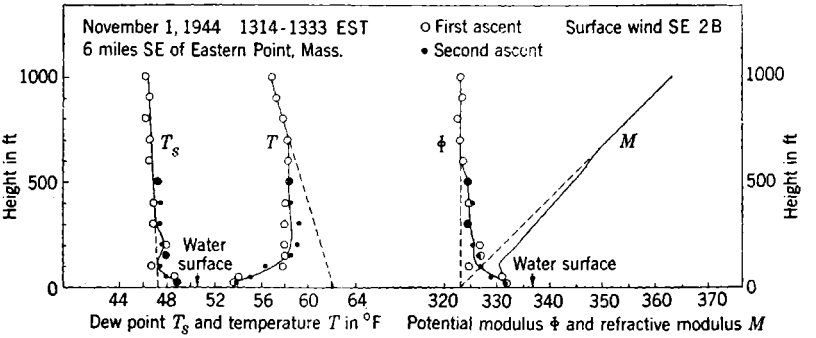


FIG. 3·21.—Shear in stable equilibrium producing a stable distribution of temperature above 100 ft.

is long so that a small variation in wind speed or direction has a marked effect. Radiation Laboratory measurements at some distance from the coastline bear out this expectation.[1] It was found in general that with temperature excesses greater than about 14°F, shear became noticeable for trajectories greater than about 5 miles. With trajectories greater than 50 miles shear was usually found to be important for any positive temperature excess.

Some examples of soundings influenced by shear near coast lines are given in Figs. 3·21 and 3·22. In both cases the air is known to have been homogeneous before leaving land. Figure 3·21 is a case where the effect of shear is observed mainly in the temperature curve, the initial potential dew points of the air at various levels having been nearly identical. The effect of modification by the water extends to about 100 ft; the stability above that level is the result of shearing stratification. In Fig. 3·22 the effect of the water extends to about 150 ft, and the departure from homogeneous distributions above that level is the result of shear. In this case,

[1] Gardner Emmons, "Vertical Distribution of Temperature and Humidity over the Ocean between Nantucket and New Jersey," *Papers in Phys. Oceanog. and Meteorol.*, Mass. Inst. of Technol. and Woods Hole Oceanog. Inst., **10**, No. 3 (1947).

the combined effect on the temperature and dew-point distributions produces a noticeably superstandard layer between 150 and 400 ft. An example of the effect of shear after a considerably longer trajectory will be shown in Fig. 3·39.

There are obviously times when shear is of importance in affecting M-profiles over land. In particular, it can play a considerable role at night when the air is being cooled from below if the land surface has varying characteristics leading to essentially different modifications from point to point. However, these variations seldom approach the distinct change in modifying surface that occurs at a coast line.

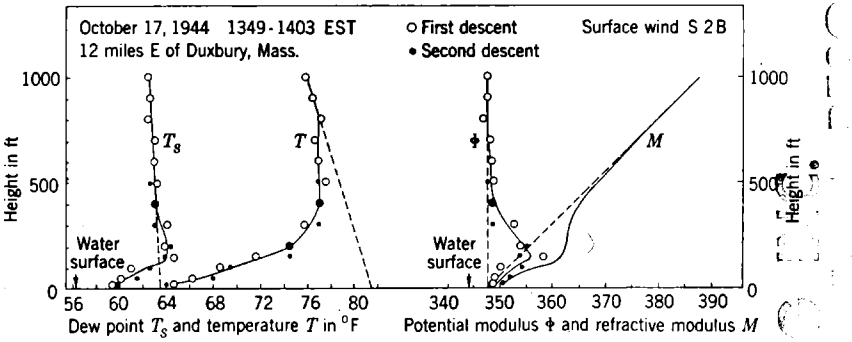


FIG. 3·22.—Shear in stable equilibrium producing a superstandard layer between 150 and 400 ft.

3·19. Initially Homogeneous Warm Air over Cold Water.—In order to study empirically the rate and manner of modification of warm air by colder water as a result of eddy diffusion, it is convenient to consider simple meteorological conditions. The simplest case for this study is the one where initially homogeneous air blows over colder coastal waters of uniform temperature, where the wind remains constant during the modification, and where effects of radiation and shearing processes are small compared with those of diffusion. The resulting changes in temperature, humidity, and M-profiles are then caused solely by diffusive processes. Furthermore, because of the fixed initial distributions, the effect of the water is obvious and comparisons are possible among measurements made on different days. This *ideal case* is often approximated near the coasts of continents with an offshore wind on days when solar heating raises the air temperature above that of the water.

Certain independent variables may be chosen for convenience to describe the meteorological conditions that affect the modification. Two of these are temperature excess and M-deficit, hereafter referred to as ΔT and ΔM (or $\Delta\Phi$). They are defined here in terms of values at the water surface and the potential values in the homogeneous layer that is being modified. A third variable is the wind speed, which influences the

amount of mechanical mixing. It may be measured at a level of 1000 ft, corresponding to the lowest level commonly reported by pilot-balloon stations, and is called W_{1000}. The fourth important variable is the length of trajectory, designated by x. Other independent variables that might be chosen can be expressed essentially in terms of these four. For example, humidity deficit is indicated by ΔT and ΔM, surface wind speed by W_{1000} and ΔT, and time over water by the product of x and W_{1000}.

There are many dependent variables that could be used to describe the modification. In the results that will be presented here, the following have been studied:

1. The ratio of the actual change of parameters at any height to the initial difference between air and water, a function of ΔT, W_{1000}, x, and height above the surface z. According to the principle of similarity, which applies in this case, the *per cent change* is the same for temperature, humidity, and Φ and M.
2. The height of the temperature inversion z_i—a function of ΔT, W_{1000}, and x.
3. The height to which the air is affected by the modification z_m—also a function of ΔT, W_{1000}, and x.
4. The duct height[1] d—a function of ΔT, W_{1000}, x, and ΔM.

In neutral equilibrium d is directly proportional to ΔM (Sec. 3·15), so that the ratio $d/\Delta M$ may be studied. In stable equilibrium the ratio is still convenient and will be used but is not strictly independent of ΔM.

Application of Diffusion Theory.—It has been seen that diffusion theory at the present time is inadequate to handle the problems arising when the air is in stable equilibrium, the case under present discussion. The dependence of the eddy diffusivity on height, wind speed, wind shear, and stability is not known. Initially, when the air is homogeneous, the eddy diffusivity presumably increases approximately linearly with height in the turbulent boundary layer and then decreases with height throughout the rest of the layer of frictional influence. After the air passes over the colder water and stratifies in its lower layers, the turbulence decreases at all levels. This occurs in the stable layer because of the thermal stability and occurs more slowly above it because of the lack of supply of turbulent energy from ground friction. Moreover, as the modification progresses and the thermal structure of the air changes, the eddy diffusivity and wind

[1] For the study of eddy diffusion in the substandard case, there is a height analogous to the duct height, namely, the one at which the M-gradient is twice the value for homogeneous air. Consider two cases of initially homogeneous air, both with the same independent variables except that the two values of ΔM are equal in magnitude but opposite in sign. The eddy diffusivity will be identical in both, with the result that at each level M will deviate from its value in homogeneous air by an amount equal in magnitude but opposite in sign. At the level where the top of the M-inversion occurs in the case of positive M-deficit, the gradient of M will be twice the homogeneous gradient in the case of a negative M-deficit (substandard case).

shear change also. It is apparent that there is a complex relationship among eddy diffusivity, wind shear, and thermal stability and that it is totally inadequate to assume that the eddy diffusivity has the same vertical distribution over land as it has during the entire trajectory over water. It is equally inadequate to assume that the eddy diffusivity is at any time a simple function of height such as a constant or a power law[1] throughout the modified layer.

However, one deduction from eddy-diffusion theory is important and should be recalled: the identity of the eddy diffusivities for the vertical transport of heat and water vapor (discussed in Sec. 3·11), resulting in the principle of similarity in this case.

Observations for Trajectories Less Than 50 *Miles.*—The modification problem for the ideal case could be solved if one could determine the per cent change at any height above the surface and distance from shore for any combination of ΔT and W_{1000}. In principle this could be determined from a large number of soundings made at varying distances from shore for the same values of ΔT and W_{1000}. In practice, however, this is difficult because of the large number of possible combinations of these parameters and the difficulty of obtaining measurements under ideal conditions.

A limited number of such observations are, however, available from the Radiation Laboratory propagation program.[2] From these measurements it has been possible to obtain only a partial and approximate answer to the problem. Figure 3·23, called a *modification cross section,* shows the per cent change for one combination of ΔT and W_{1000}. The ordinate is height above the surface z, and the abscissa is distance from shore x. The plotted figures are the observed per cent change of temperature and vapor pressure, the two changes (and the per cent change of M and Φ) being

[1] An extensive diffusion theory based on such an assumption has been developed at the Telecommunications Research Establishment, Malvern, England, where considerable work has been done attempting to simplify this highly complicated problem to the point where formulas can be used to forecast propagation conditions. See, for instance, "Elements of Radio Meteorological Forecasting," TRE Report No. T1621, Feb. 14, 1944; "Application of Diffusion Theory to Radio Refraction Caused by Advection," TRE Report No. T1647, Apr. 6, 1944; "Outline of Radio Climatology in India and Vicinity," TRE Report No. T1727, Sept. 12, 1944. The electromagnetic phase of this method is discussed in Sec. 2·21.

[2] In the summer of 1944 low-level soundings were made in Massachusetts Bay, generally within 50 miles of the coast (see Sec. 4·2). These soundings were made to a height of 1000 ft by aircraft according to the technique described in Sec. 3·27. See "Measurements of Temperature and Humidity in the Lowest 1000 Feet of the Atmosphere over Massachusetts Bay," by Richard A. Craig, *Papers in Phys. Oceanog. and Meteorol.,* Mass. Inst. of Technol. and Woods Hole Oceanog. Inst., **10,** No. 1 November, 1946. A more detailed analysis of the eddy diffusive effects revealed by these soundings has been published in *Jour. Meteorol.* **6,** 123 (1949), under the title "Vertical Eddy Transfer of Heat and Water Vapor in Stable Air," by R. A. Craig. This analysis, completed too late to be included here, gives additional empirical information about the modification of air under the specified conditions.

theoretically identical according to the principle of similarity and nearly identical in the soundings used here. Each vertical column represents a "sounding" that consists of two measurements of temperature and humidity at each level. Each plotted figure is then the average of four values of per cent change, one each for temperature and humidity according to two different measurements. Isopleths are drawn approximately for every 10 per cent of change.

The soundings used in this cross section were, in general, made on different days. The temperature excesses in the soundings used varied

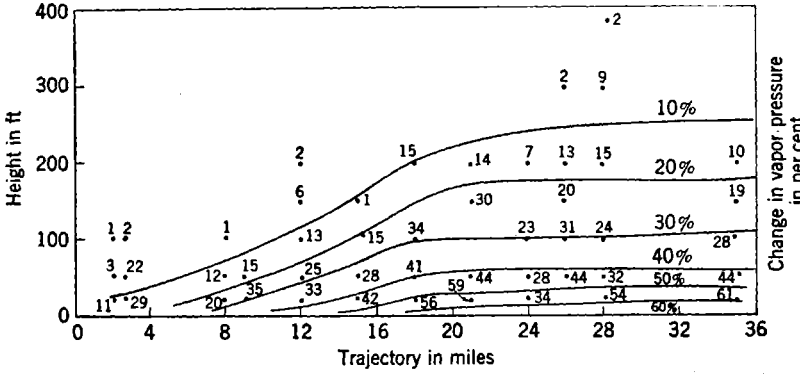


FIG. 3·23.—Modification cross section for ΔT = 6.5 to 11.5°F, W_{1000} = 21 to 25 mph. The numbers near the points are the observed changes in vapor pressure in per cent. The solid curves are the corresponding isopleths.

between 6·5° and 11.5°F, the wind speed at 1000 ft varied between 21 and 25 mph

It is apparent from an inspection of Fig. 3·23 that the isopleths as drawn differ somewhat from the observed values through which they are drawn, the difference seldom exceeding 10 per cent. They are drawn, as far as possible, to indicate a reasonable average distribution of per cent change in accordance with the observations. The discrepancies occur partly because of uncertainties in observation and scarcity of general weather data. It is difficult, for example, to determine without an extremely close network of weather stations, the trajectory, wind, and temperature excess associated with a particular sounding, and some error must be expected. Moreover, the soundings, as has been stated, represent a finite spread in the variables ΔT and W_{1000}, which variations, particularly those in temperature excess, would be expected to make them somewhat less comparable.

However, there is no question but that Fig. 3·23 shows to a good degree of approximation the type of modification that occurs under these wind and temperature conditions and shows it more accurately than any other available meteorological knowledge.

It is desirable to make clear the variation with distance of the other dependent variables found from the same series of soundings. Figures 3·24 to 3·26 show the changes with distance of the height to which modification extends z_m, the height of the temperature inversion z_i, and the ratio $d/\Delta M$. These graphs follow, of course, from Fig. 3·23, which is the best method

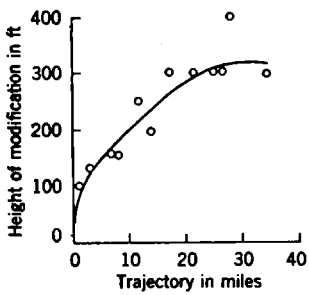


FIG. 3·24.—Variation of height of modification with trajectory for $\Delta T = 6.5°$ to 11.5°F, $W_{1000} = 21$ to 25 mph

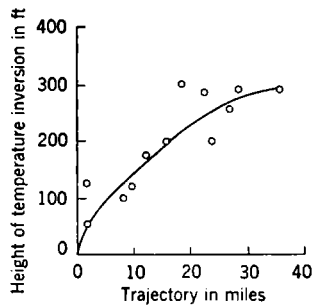


FIG. 3·25.—Variation of height of temperature inversion with trajectory for $\Delta T = 6.5°$ to 11.5°F, $W_{1000} = 21$ to 25 mph.

of describing the modification, and they will prove useful later in this section.

We next consider the effects of variations of wind speed and temperature excess. For an adequate study of these parameters, considerably more information would be required than is now available. A few cross sections similar to those of Fig. 3·23 but for different combinations of ΔT and W_{1000} have been given by Craig (*loc. cit.*). Here we shall use data from the Radiation Laboratory program to discuss the problem from a somewhat different point of view.

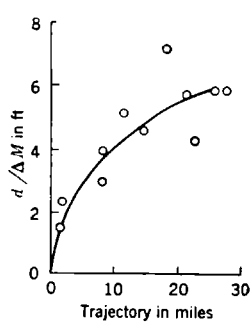


FIG. 3·26.—Variation of $d/\Delta M$ with trajectory for $\Delta T = 6.5°$ to 11.5°F, $W_{1000} = 21$ to 25 mph.

Figures 3·27 to 3·29 show the height of modification, height of temperature inversion, and ratio $d/\Delta M$ plotted on diagrams having ΔT as the abscissa and W_{1000} as the ordinate. In general, of course, these parameters are functions of the length of trajectory also. However, reference to Figs. 3·24 to 3·26 show that at least in the case where those graphs are applicable, z_m, z_i, and $d/\Delta M$ become more or less independent of distance after about 15 miles of rapid increase. There are observational reasons to believe that this is also true for other values of ΔT and W_{1000}, except that the rapid increase may take place in a shorter distance for larger ΔT and in a longer distance for smaller ΔT.

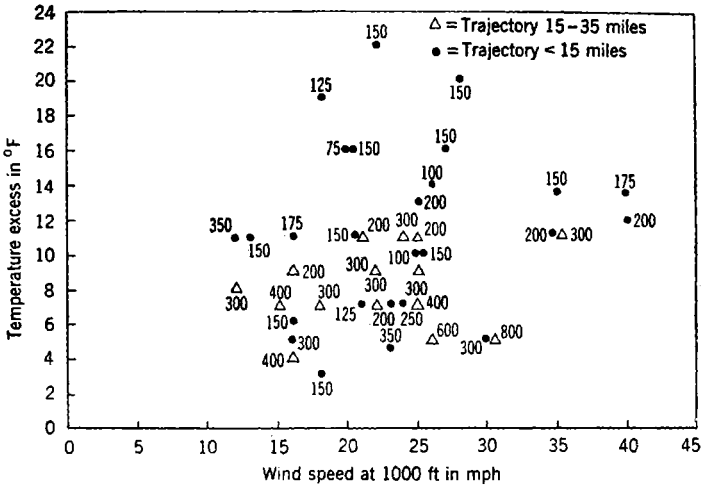


FIG. 3·27.—Some observed values of the height of modification, in feet, for trajectories less than 35 miles.

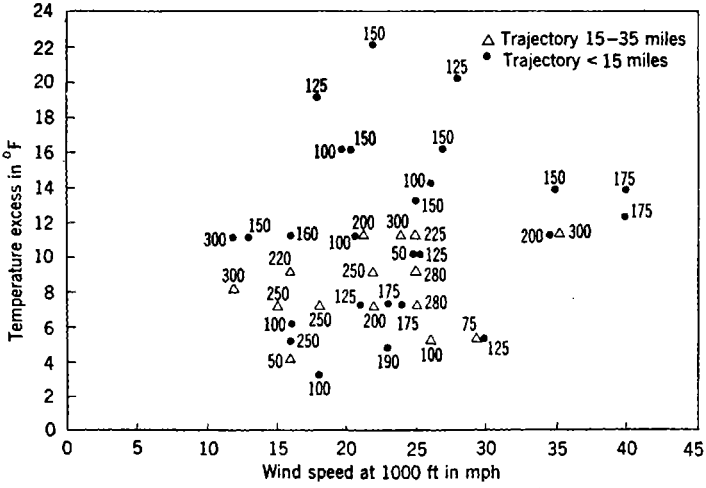


FIG. 3·28.—Some observed values of the height of the temperature inversion, in feet, for trajectories less than 35 miles

Taking 15 miles as an average value, the dots in Figs. 3·27 to 3·29 represent trajectories of less than 15 miles. They should in general be less than the other values, indicated by triangles, which are obtained for trajectories of 15 to 35 miles.

The height of modification is seen to be less than 500 ft except for strong winds and small temperature excess; it is lowest with large ΔT

and with light wind. The wind appears to have a more important effect on z_m when ΔT is small than when it is large.

The temperature inversion with the greatest height is found for some intermediate values of ΔT and W_{1000}. With less mixing the height of modification and hence the temperature inversion are lower. With more mixing, large gradients are confined near the surface, as the eddy diffusivity vanishes at the boundary. In some cases there is also an upper turbulence inversion (see Sec. 3·17 and Fig. 3·18), but the lower one is referred to here.

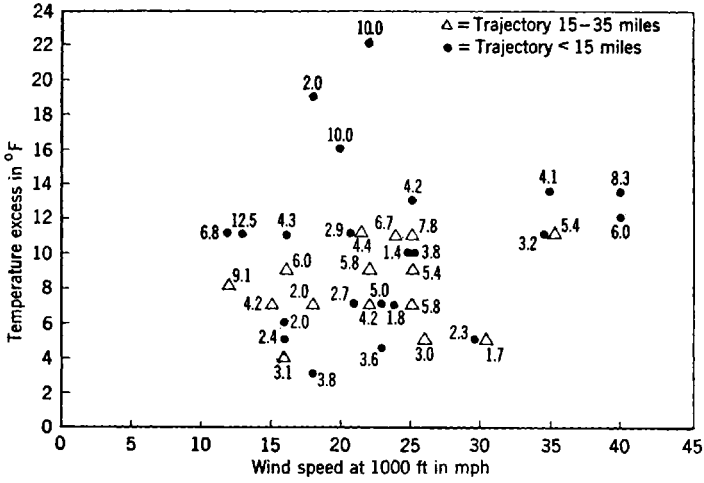


FIG. 3·29.—Some observed values of $d/\Delta M$, in feet, for trajectories less than 35 miles.

The height of the temperature inversion is generally less than 400 ft within 35 miles of the shore.

The graph of $d/\Delta M$ is the most unreliable one. This is probably because of the fact that the ratio is to some extent dependent on ΔM as well as on length of trajectory. It was shown in Sec. 3·15 that the ratio is about 2 ft for neutral equilibrium, and it is undoubtedly larger for a positive temperature excess It seems to increase with increasing ΔT up to about $\Delta T = 12°\text{F}$. Its behavior beyond that point is not clear.

Because duct height is an important variable that the radio-meteorologist is often called upon to predict, further discussion seems necessary. One important rule is that the duct height is usually less than the height of the temperature inversion z_i. This is important, for the latter is easier to predict, because it depends on one less variable than does duct height.

According to the principle of similarity,

$$\frac{M_h - M}{\Delta M} = -\frac{T_h - T}{\Delta T},$$

where M_h and T_h are the unmodified values at a given height and M and T are the modified values at the same height. The values ΔM and ΔT are M-deficit and temperature excess. When this expression is differentiated with respect to height and the substitutions $dM_h/dz = 4.0$ per 100 ft and $dT_h/dz = -0.54$°F per 100 ft, which apply to initially homogeneous air, are made, it becomes

$$\frac{dT}{dz} = \left(4 - \frac{dM}{dz}\right)\frac{\Delta T}{\Delta M} - 0.54 \qquad \text{°F per 100 ft.}$$

At the duct height $z = d$, $dM/dz = 0$. Then

$$\left.\frac{dT}{dz}\right|_{z=d} = 4\,\frac{\Delta T}{\Delta M} - 0.54 \qquad \text{°F per 100 ft.}$$

For the substandard case ($\Delta M < 0$) the same equation holds if d is defined as the height at which $dM/dz = 8.0$ per 100 ft, twice the M-gradient in homogeneous air.

The comparative heights of the temperature inversion and duct follow from this equation. For the case where $\Delta M/\Delta T < 7.4$ per °F, temperature increases with height at $z = d$; hence $z_i > d$. Similarly for $\Delta M/\Delta T = 7.4$ per °F, $z_i = d$; and for $\Delta M/\Delta T > 7.4$ per °F, $z_i < d$.

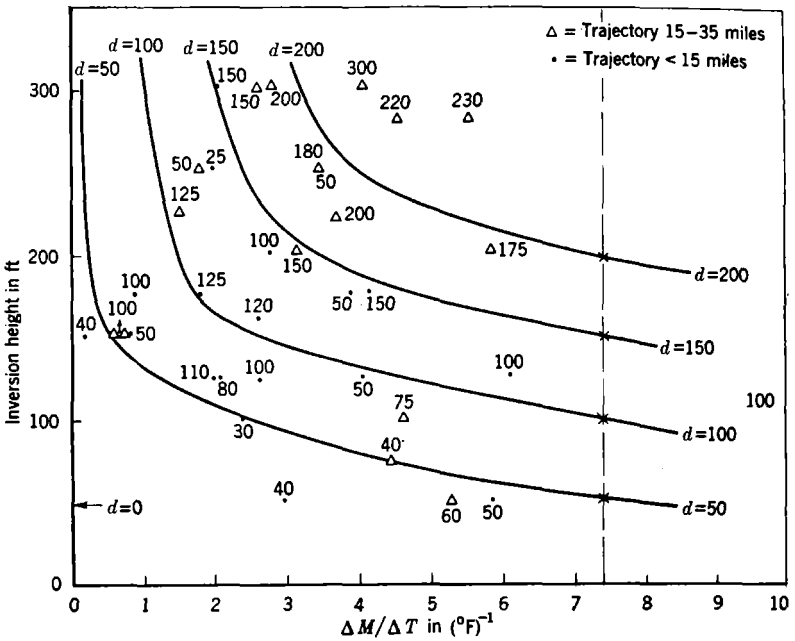


FIG. 3·30.—Variation of duct height with height of temperature inversion and with $\Delta M/\Delta T$. The crosses represent theoretical points (see text).

Consider a diagram with $\Delta M/\Delta T$ as the abscissa and z_i as the ordinate, and let values of d be plotted as points on this chart. The isopleth $d = 0$ is a vertical line through $\Delta M/\Delta T = 0$, for here $\Delta M = 0$. Moreover, any isopleth of d must cross the point where $\Delta M/\Delta T = 7.4$ per °F and z_i has the same value as the isopleth of d. Furthermore, all the isopleths must approach the horizontal line $z_i = 0$ as $\Delta M/\Delta T$ becomes large; for in the

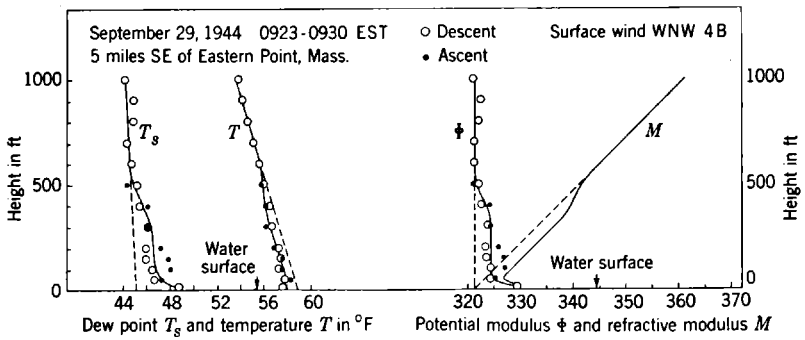


FIG. 3·31.—Ideal case with $\Delta T = 3.4$°F, $\Delta M = 23$, $W_{1000} = 28$ mph, and $x = 8$ miles.

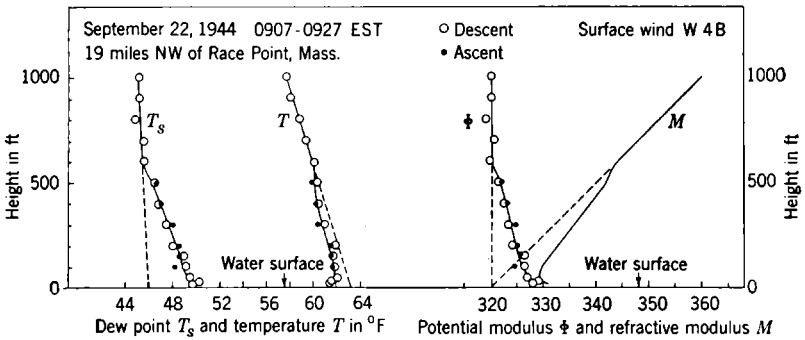


FIG. 3·32.—Ideal case with $\Delta T = 5.6$°F, $\Delta M = 28$, $W_{1000} = 25$ mph, and $x = 25$ miles.

limit when $\Delta T = 0$, $z_i = 0$ and d can have any value depending only on ΔM. Figure 3·30 is such a diagram with values plotted from the same soundings that were used to plot Figs. 3·27 to 3·29.

In Figs. 3·31 to 3·38 are presented a group of selected soundings that correspond closely to the ideal conditions mentioned at the beginning of this section. They are intended to illustrate the types of simple modification that can occur with some combinations of the parameters ΔT, W_{1000}, x, and ΔM. The difference between the distributions in the surface layer and the initial homogeneous distributions (indicated by the broken lines) is presumably due to modification by the eddy-diffusive processes.

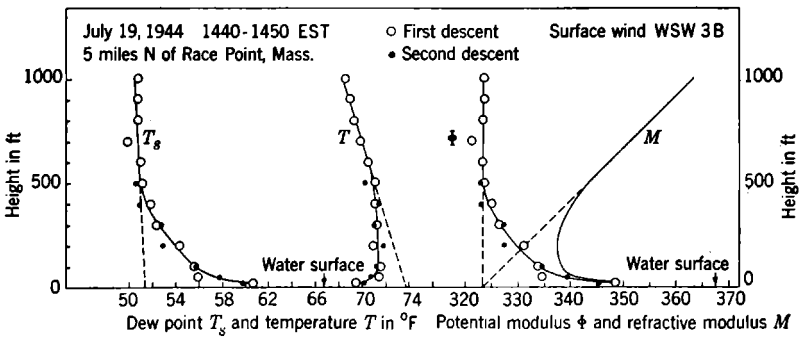


FIG. 3·33.—Ideal case with $\Delta T = 7.0°F$, $\Delta M = 44$, $W_{1000} = 25$ mph, and $x = 28$ miles.

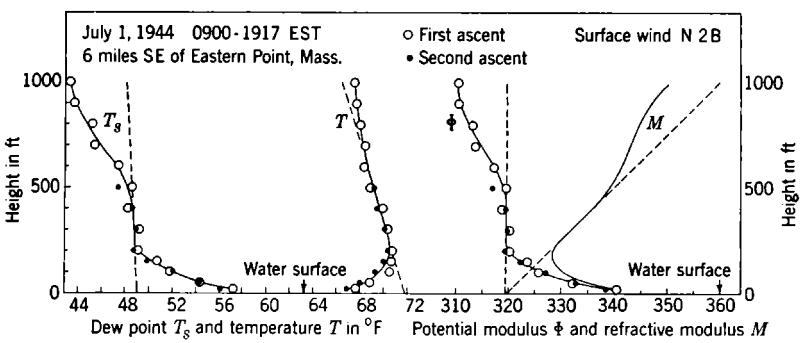


FIG. 3·34.—Ideal case with $\Delta T = 8.5°F$, $\Delta M = 40$, $W_{1000} = 15$ mph, and $x = 30$ miles.

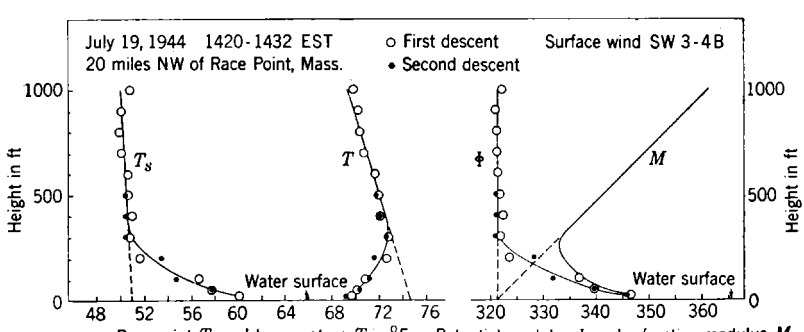


FIG. 3·35.—Ideal case with $\Delta T = 8.8°F$, $\Delta M = 44$, $W_{1000} = 21$ mph, and $x = 21$ miles.

The soundings are arranged in order of increasing temperature excess in order that the general effect of this very important variable may be apparent. However, the effect of other variables should not be overlooked. Figures 3·36 and 3·37 differ essentially only in trajectory. In Fig. 3·34 there is less modification than might be expected from a con-

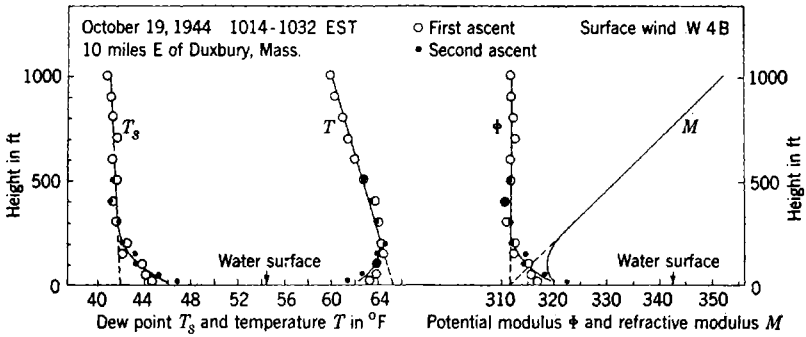


FIG. 3·36.—Ideal case with $\Delta T = 10.8°F$, $\Delta M = 31$, $W_{1000} = 35$ mph, and $x = 9$ miles.

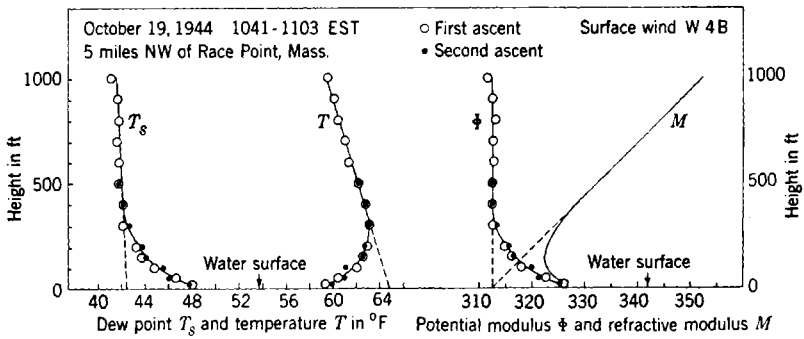


FIG. 3·37.—Ideal case with $\Delta T = 11.0°F$, $\Delta M = 30$, $W_{1000} = 35$ mph, and $x = 23$ miles.

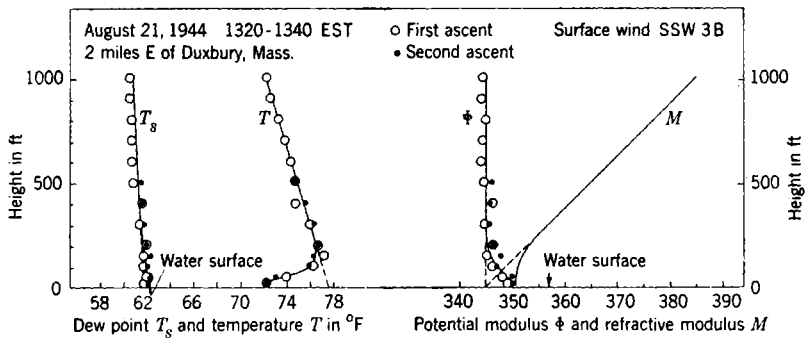


FIG. 3·38.—Ideal case with $\Delta T = 15.3°F$, $\Delta M = 12$, $W_{1000} = 40$ mph, and $x = 8$ miles.

sideration of ΔT alone; this is presumably because of the rather light wind. The effect of ΔM on the Φ- and M-profiles has already been shown in Figs. 3·14 to 3·16. These cases furnish, of course, only a few examples of the many combinations of the various independent variables that can exist.

The modification problem for trajectories longer than 50 miles has been comparatively unstudied to date. Some airplane soundings were made by the Radiation Laboratory south of Long Island in the summer of 1945 in air that had traveled 50 to 300 miles over cooler water. These data have been analyzed by Emmons.[1]

First of all, it appears from these observations that the ideal case that was outlined at the beginning of this section is unlikely for trajectories greater than 100 miles. Even if the air is initially homogeneous, varia-

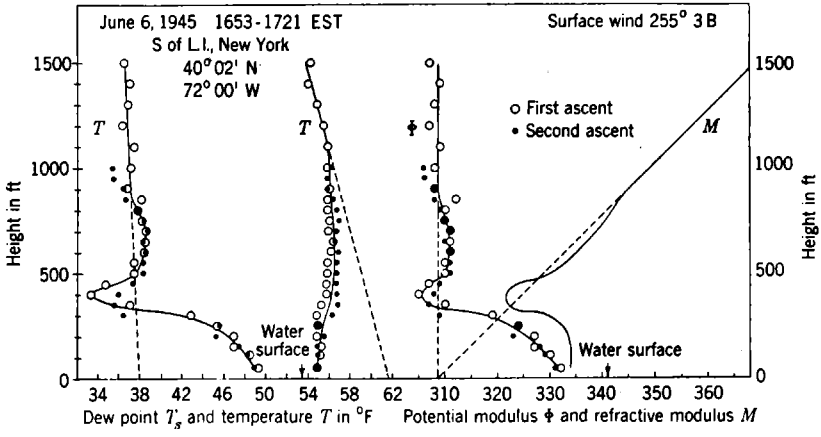


FIG. 3·39.—Overwater sounding showing the effect of shear after a trajectory of about 100 miles.

tions of water temperature and wind speed along the trajectory, radiation, and shear must all be considered in addition to diffusion. Shear is particularly important. Of 29 soundings made in this program, only 3 may have been free of obvious shearing effects in the lowest 1000 ft; these had trajectories of 70 to 90 miles with a small temperature excess and fresh wind. The effect of shear in all cases was to produce shearing stratification and accompanying erratic and unpredictable superstandard or substandard layers above the level of modification.

Figure 3·39 is an example[2] of the effect of shear for an overwater trajectory of about 100 miles (for the air at 1000 ft). The overland distributions were homogeneous during the time when any of the air in this sounding began its travel over the water. The effect of modification by the water probably extends to about 250 ft and certainly no higher than 400 ft. The stable temperature distribution above the level of water modification is the result of shearing stratification as is the substandard layer between 400 and 600 ft and the superstandard layer between 700

[1] Gardner Emmons, "Vertical Distribution of Temperature and Humidity over the Ocean between Nantucket and New Jersey," *Papers in Phys. Oceanog. and Meteorol.*, Mass. Inst. of Tech. and Woods Hole Oceanog. Inst., **10**, No. 3, December 1947.

[2] This sounding is Fig. 14 in Emmons' paper, and is also discussed in the paper.

and 900 ft. If the level of water modification is only 250 ft, which is likely, the structure of the M-inversion between 250 and 400 ft must also be influenced by shear.

It is possible to indicate the order of magnitude of z_m, z_i, and d found in these long-trajectory soundings. The height of modification was between 200 and 500 ft in all the soundings except three, where it was about

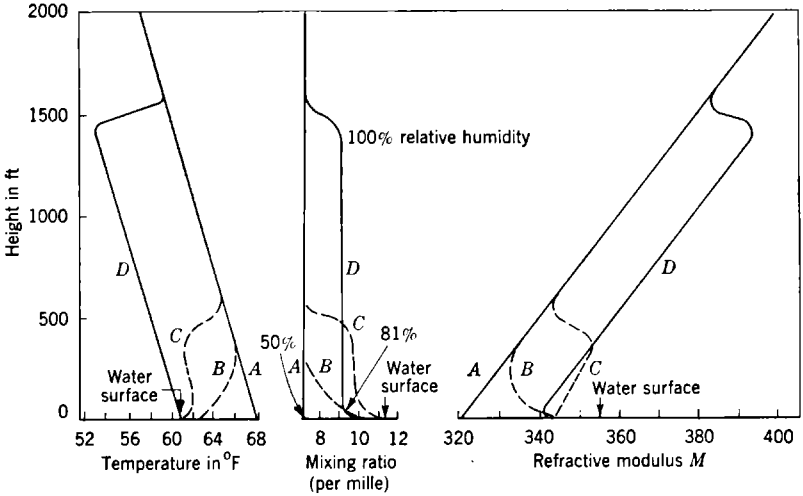


FIG. 3·40.—Schematic representation of probable stages in modification of initially homogeneous warm air over cold water.

1000 ft. The temperature inversions were usually between 200 and 1000 ft; the surface ducts below 500 ft. It must be remembered that because of the complicating factors these values are not strictly comparable to those mentioned before in the study of the ideal case. It is to be expected that in the absence of shear and shearing stratification, z_m would have been somewhat higher and z_i somewhat lower than the values found. Moreover, not all these soundings were made in air that was initially homogeneous at all levels; the surface air in some cases, because of shear, left land before the beginning of convective mixing.

Probable Stages in the Modification of Initially Homogeneous Warm Air over Cold Water.[1]—It is desirable to sum up this section by discussing qualitatively, in the light of theoretical reasoning and these data, the modification history of initially homogeneous warm air over cold water. Consider for the moment the ideal case where shear, radiation, and variations in water temperature and wind may all be neglected. It is believed that during modification this air will go through the stages indicated in Fig. 3·40.

[1] See R. B. Montgomery and R. H. Burgoyne, "Modified Index Close to the Ocean Surface," RL Report No. 651, Feb. 16, 1945.

Initially homogeneous, as indicated by the lines *A*, the air is at first affected so that rather large gradients of temperature and moisture exist in the lowest few hundred feet, as shown by the curves *B*. Additional turbulent mixing leads to a nearly homogeneous layer with large gradients only near the surface and at the top of the layer, as indicated by the curves *C*. This layer continues to grow until it reaches the height of frictional influence as shown by the distributions *D*. In this final state, the potential temperature of the homogeneous layer is equal to the water temperature, whereas the humidity is somewhat less than the value at the water surface, since the humidity value in the layer is limited by the saturation value at the top of the layer.

The empirical information given in this section relates primarily to Stage *B*. This stage is well developed shortly after the air leaves land and persists for some distance which must vary with temperature excess and wind speed. This distance according to available observations is usually greater than 50 miles and often even greater than 300 miles. No well-defined case of air in the state illustrated by *C* has been observed over water. However, such a development has been observed over land (turbulence inversions, Sec. 3·17). Another example in nature is the common occurrence of a surface layer of homogeneous water in the ocean overlying a sharp temperature gradient (thermocline), both produced by wind stirring. Stage *C* is probably not usually well developed until the overwater trajectory is a matter of hundreds of miles, but it must be emphasized that this distance depends on the amount of mechanical mixing that can occur and hence on the temperature excess and wind speed.

Shear complicates this picture by producing shearing stratification when warm air overruns cold air or by producing convective mixing when cold air overruns warm air. The final state *D* is, however, not affected by shear, radiation, or other complicating factors, for at this time all the air in the modified layer, whatever its source, has been mixed by turbulence.

3·20. Complex Overwater Modifications.—Certain factors affecting overwater modification associated with stable equilibrium must be considered in addition to those entering into the ideal case discussed in the previous section. Two of these, shear and radiation, have already been mentioned, and their effects assessed. In this section the additional complications of lack of initial homogeneity, variations in wind speed, and variations in water temperature will be considered.

If the air is not initially homogeneous, two important facts follow. (1) The air is more resistant to turbulent mixing because of the initial stability. This factor tends to keep the level of modification lower than in the case of initially homogeneous air. (2) The distribution over the water depends, above the level of water modification, on the initial distribution. This is, in general, not so simple as in the homogeneous case, where the *M*-profile is a straight line with nearly standard slope. A forecast of the

distribution requires either a sounding in the air before or after leaving land or a complete solution of the overland problem. Figures 3·41 and 3·42 are soundings made in air that was not initially homogeneous. In both cases the air, as determined approximately with the aid of the characteristic diagram, has been modified by the water up to about 200 ft. The distributions above that level depend on initial conditions that have probably been complicated somewhat by shear.

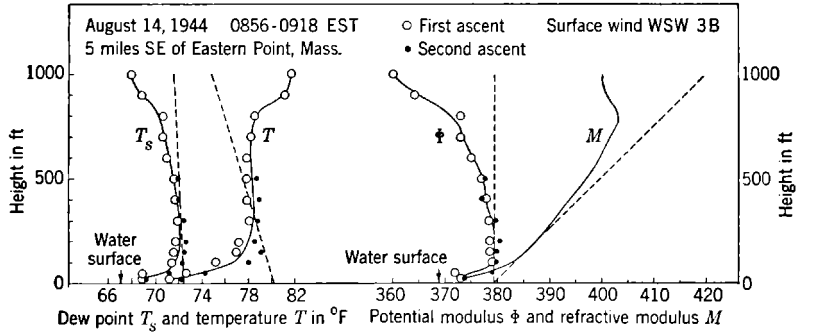


FIG. 3·41.—Cooling of initially stratified air over water, Case 1.

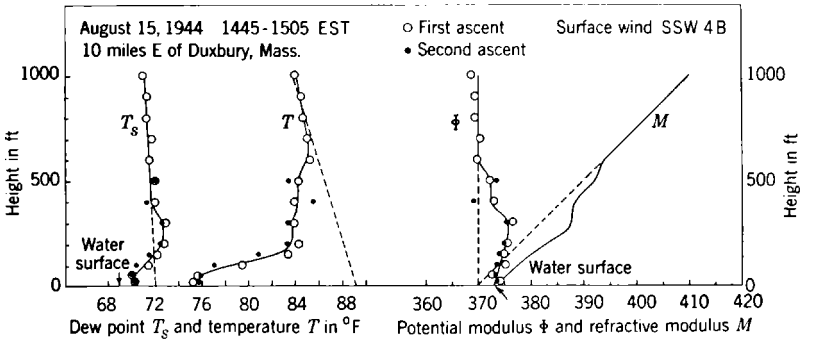


FIG. 3·42.—Cooling of initially stratified air over water, Case 2.

Variation in wind speed along the trajectory is a subject about which little is known. Its primary effect, of course, is to cause a variable amount of mechanical mixing along the trajectory. This effect, superimposed on the several other factors involved in cooling from below, must be small and is very difficult to separate from the others.

Variations in water temperature during the trajectory may be of several types with varying effects. Consider first the case where the air is being heated from below. If the water temperature changes so that the air is still in unstable equilibrium but with a larger or smaller temperature excess, the temperature, humidity, and M-profiles can be expected to

readjust themselves quickly to meet the new conditions. Subsequent soundings will show no obvious effect of the change in water temperature. If, however, the air passes over water that is sufficiently cooler so that the temperature excess becomes positive, cooling from below proceeds in the same manner as it does in the case where initially homogeneous air passes from warm land to cooler water.

Another possibility is that the air is originally over colder water and therefore in stable equilibrium. It has been seen that in this case the temperature and humidity of the air at the boundary are determined by the water temperature and that the characteristic curve in the modified layer is a straight line, or nearly so. If the air then passes over still colder water, the air at the boundary assumes new temperature and humidity values corresponding to the new water temperature. Diffusion processes then act to modify the air above in accordance with these new values and the new temperature excess. The readjustment process is much slower in this case than in the case where the temperature excess is negative and, although it changes in value, still remains negative. Until all the air in the modified layer has been mechanically mixed with air at the new surface values, the distributions of temperature and humidity are not similar. The air in the lower part of the modified layer is adjusted to the new water temperature, but some air above remains as it was when mixed at the former water temperature. Changes in the characteristic curve are shown schematically in Fig. 3·43. Figure 3·44 is the characteristic diagram of a sounding where this effect was actually observed. Starting from shore the trajectory for this sounding passed over water with a temperature of 54°F measured shortly after sunrise the following morning (53° to 54°F measured in the afternoon two days before the sounding). Note that part of the characteristic curve still points toward this temperature. The temperature measured at the same time and place as the sounding is about 52°F, as shown by the arrows in the figure.

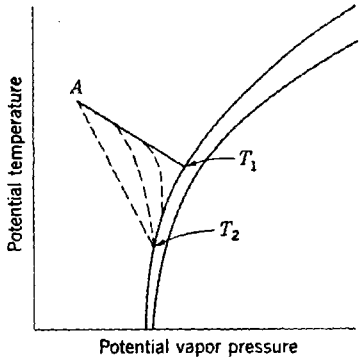


FIG. 3·43.—The characteristic curve is originally $\overline{AT_1}$. When the water temperature changes to T_2, the characteristic curve gradually changes to the straight line $\overline{AT_2}$.

Changes of this type in water temperature may be particularly important when the dew point of the air is close to the water temperature. Figure 3·45 shows a sounding where the air is being cooled and moistened from below. The water temperature then decreases until it is less than the dew point of the air in the lowest 100 ft of the modified air so that the

air proceeds to lose some moisture by condensation on the water. What must have been an M-inversion at 100 to 200 ft changes to the measured distribution when the water temperature changes.

If air in stable equilibrium passes over warmer water, convection immediately sets in. This convection cannot be effective much above the height where the air as originally modified has a potential temperature equal to the new water temperature. If the new water temperature is

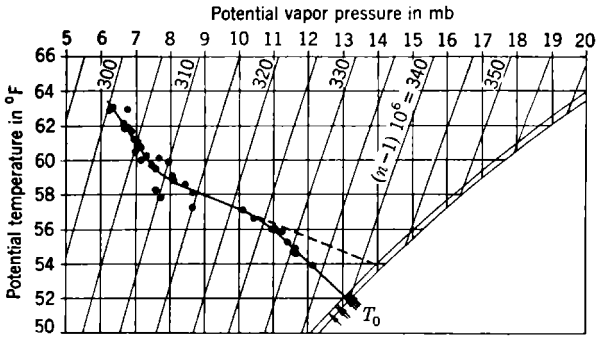


FIG. 3·44.—Characteristic curve of a measured sounding showing the bend caused by changing water temperature. The air has passed over water of 54°F, is now over water of 52°F. The family of curves represents Φ for a surface pressure of 1000 mb.

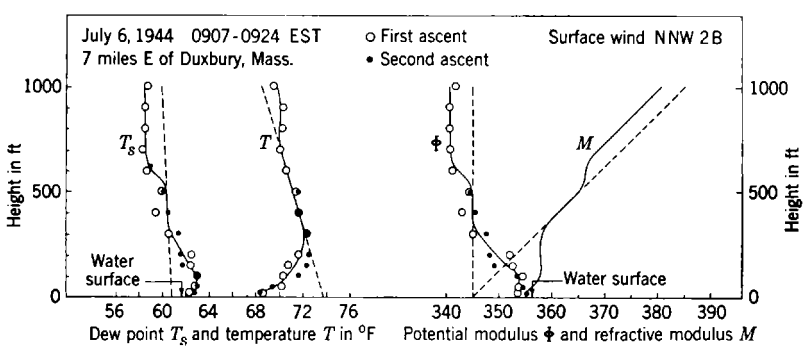


FIG. 3·45.—Measured overwater sounding. The bend in the dew-point and modulus curves at 100 ft is the result of changing water temperature.

not much warmer than the old, this may be only a few feet above the surface, but sometimes the change is sufficiently great that the effect is noted in an airplane sounding at 20 or 50 ft or higher. Figure 3·46 is an example of such a sounding. In the layer affected by the new temperature, conditions are nearly homogeneous because of the convective stirring; the original surface M-inversion has become elevated.

3·21. Nocturnal Cooling and Diurnal Cycles.—At night, in the absence of solar radiation, the temperature of a land surface ordinarily decreases

because of nocturnal radiation. The lower part of the atmosphere is cooled in turn by contact with the cool surface as heat is transferred downward by turbulence. The resulting thermal stratification and the accompanying humidity and M-profiles are of considerable interest.

The problem of forecasting the M-profile under given conditions is twofold. It is necessary to forecast (1) the rate at which the ground surface will cool and (2) the manner in which conditions at the surface will affect the air above. The first question is one that has been studied empirically and theoretically by meteorologists; the second has generally

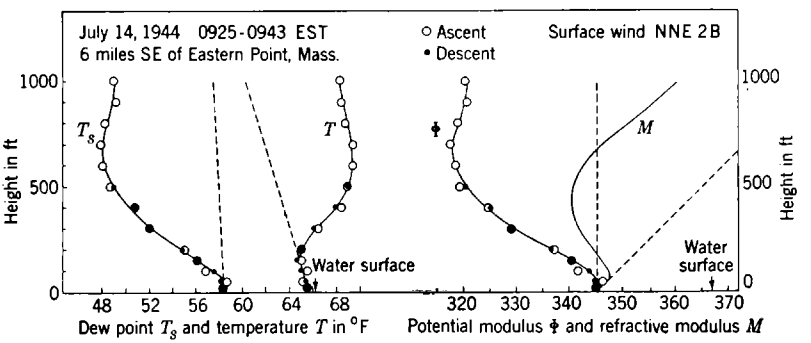


FIG. 3·46.—Overwater sounding in air that was cooled from below and then passed over warmer water.

been discussed only qualitatively and is dependent upon recently developed low-level sounding techniques.

The rate and amount of cooling of the earth's surface at night is dependent upon the value of the nocturnal radiation. This was defined in Sec. 3·17 as the net outgoing long-wave radiation.[1] The earth radiates approximately as a black body, but part of this radiation is absorbed and reradiated by certain atmospheric gases, principally water vapor. Because water vapor absorbs only in certain spectral regions, much of the radiation emitted by the earth escapes through the atmosphere. Cloud surfaces, however, radiate approximately as black bodies; therefore when clouds are present, the nocturnal radiation is reduced considerably.

The temperature of the earth's surface depends not only on the nocturnal radiation but also on the specific conductivity of the surface, which determines the amount of heat that is conducted to the surface from below. Surfaces with a high conductivity do not cool so much (or heat so much during the day) as surfaces with a low conductivity. This factor may

[1] The energy radiated by the earth or the atmosphere is contained mainly between wavelengths of 3 and 100μ, with a maximum at about 10μ; the sun's radiation (after passing through the atmosphere) between 0.3μ and 4μ, with a maximum at 0.475μ. Hence the former is spoken of as "long-wave" (terrestrial) radiation, whereas the latter is often called "short-wave" (solar) radiation.

vary considerably for different types of soil. For dry soil it has a value of 3.3×10^{-4} cal per sec per cm per °C. For dry sand it is about three times as large. Because water at 0°C has a value of 1.2×10^{-3}, moist ground is a better conductor than dry ground.[1]

The dependence of the amount of cooling on factors such as amount of water vapor in the air, amount and height of clouds, and type of surface has been studied extensively. Summaries of results are presented by Haurwitz and Byers.[2]

Even if the temperature of the surface is known as a function of time, there remains the question as to how the air above is affected. This is essentially the same type of modification problem as that of warm air over cold water, but it is in many ways a more difficult one. (1) The temperature of the modifying land surface varies greatly, both with time as the radiative cooling proceeds during the night and with horizontal distance if there is a variation in the type of surface. Accordingly the air is subjected to a modification where the temperature excess varies over wider ranges than are generally found in the overwater problem. (2) The temperature of the land does not uniquely determine the boundary values of humidity and refractive modulus, except in the special case where the ground is wet. If the ground is perfectly dry, however, there is no humidity gradient and the M-profile near the surface varies according to changes in the temperature distribution.

If it were not for these facts essentially the same independent variables could be used to describe the nocturnal cooling of the air over flat terrain as were employed in the overwater problem, namely, temperature excess, M-deficit, wind speed at 1000 ft, and the length of time that cooling has progressed.[3] Because of the inherent variations and uncertainty in the first two, it may be desirable, when careful analysis of overland modification is undertaken, to substitute some other variables. For example, some parameter that takes into consideration cloud cover, amount of water vapor in the air, and specific conductivity of the soil could be used in place of temperature excess inasmuch as the time of cooling is another variable. Special consideration of whether the ground is dry, damp, or wet would be necessary in finding a variable analogous to M-deficit.

Detailed soundings, as well as other types of meteorological data, will be necessary in any satisfactory study of nocturnal cooling. The ordinary meteorological soundings made by radiosonde are inadequate for a careful study because they do not report in fine detail the structure of the at-

[1] These values are from the *Handbook of Chemistry and Physics*, Chemical Rubber Publishing Co., Cleveland, 1933, pp. 1192, 1193.

[2] B. Haurwitz, *Dynamic Meteorology*, McGraw-Hill, New York, 1941, pp. 106-110; H. R. Byers, *General Meteorology*, McGraw-Hill, New York, 1944, pp. 474-477.

[3] Over land the problem is complicated by trees, buildings, marshes, and other irregularities in the surface.

mosphere in as shallow a layer as the lowest 1000 ft. Special low-level soundings, made by instruments described in Secs. 3·27 to 3·31, are sufficiently accurate, and many such soundings have been made during the war, but no thorough analysis of them has been carried out. Despite the lack of analysis, some of these soundings will be presented here as illustrations of certain basic considerations.[1]

Figures 3·47 and 3·48 contain soundings made at Leesburg, Fla. Figure 3·47*a* shows conditions in the lower atmosphere just before sunset

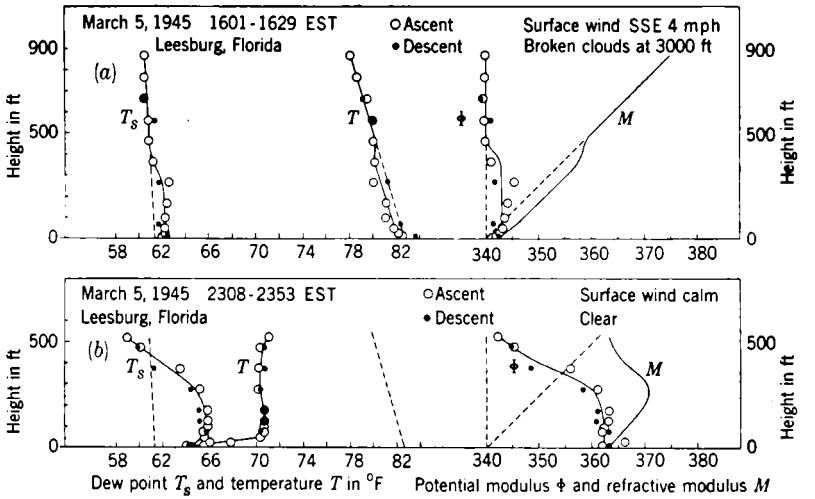


FIG. 3·47.—Soundings indicating effect of nocturnal cooling with clear sky. (*a*) Just before sunset; (*b*) midnight.

on Mar. 5, 1945, and Fig. 3·47*b* shows the effect of nocturnal cooling on the air up to midnight. Figure 3·48*a* and *b* shows similar measurements made on the night of Feb. 19, 1945.

Comparison between these two pairs of soundings is of interest. In both cases the air was nearly homogeneous when the cooling began. In the first case weather conditions at the time of the later sounding included clear skies and calm at the surface, whereas on the night of Feb. 19 the second sounding was made with an overcast at a height of 2000 ft and a surface wind of 7 mph. In the first case a considerably greater amount of cooling occurred than in the other case; presumably this was to a great extent caused by the difference in cloud conditions. Note also that with the smaller amount of cooling and the surface wind of 7 mph in Fig. 3·48*b* the temperature lapse rate was nearly adiabatic, whereas with

[1] The particular soundings shown in this section were made either at Duxbury, Mass., by the Radiation Laboratory (unpublished) or at Leesburg, Fla., by the AAF Tactical Center, Orlando, Fla., under AAF Board Project No. 3767 B000.93.

a larger amount of cooling and no wind there was a sharp inversion with a nearly isothermal layer above.

It is also interesting to compare the humidity and M-profiles of these figures. In Fig. 3·48 no gradient of moisture was in evidence on the soundings, indicating that the ground was dry. This fact, along with the adiabatic lapse rate of temperature, led to a homogeneous, or nearly standard, M-profile. The profiles in Fig. 3·47 are significant because they illustrate a sequence that is common when air is over damp or wet ground

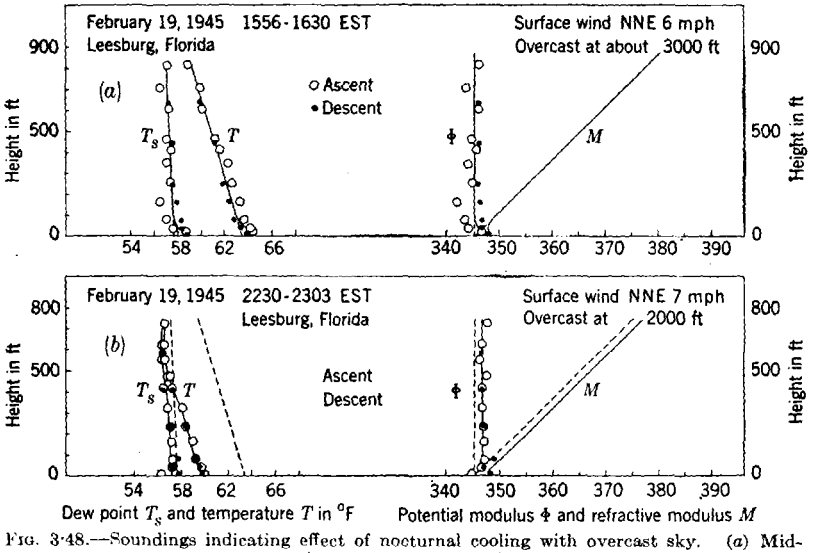


FIG. 3·48.—Soundings indicating effect of nocturnal cooling with overcast sky. (*a*) Mid-afternoon; (*b*) late evening.

such that there is a positive humidity deficit. Before sunset, if the air is heated from below, moisture evaporated from the ground is rapidly carried aloft by convection. However, after nocturnal cooling begins, the air becomes thermally stratified and the moisture evaporated from the ground tends to be confined to a low layer of the atmosphere. This leads to a relatively large humidity gradient near the ground and an accompanying superstandard layer, either based at the surface or, as in this case, elevated. Later during the night the ground may cool below the dew point of the moistened air so that the moisture is condensed on the ground and a decrease of moisture with height occurs in the lowest layer. This was apparently beginning to occur when the sounding in Fig. 3·47*b* was made, according to the observations made at 10 ft. This new humidity distribution tends to cause a substandard layer based at the ground, but on occasion the temperature inversion may be strong enough to keep the layer standard or superstandard.

Figure 3·49*a* is a sounding made at Duxbury, Mass., on the night of Nov. 2, 1944. It is intended to illustrate that appreciable substandard layers may occur along with strong nocturnal temperature inversions when the humidity gradient near the ground is positive. In Fig. 3·49*b*, which shows a sounding made at Duxbury, Mass., on the night of Sept. 25, 1944, there is no gradient of dew point and the temperature inversion has led to a superstandard layer that includes a slight M-inversion. Finally, Fig. 3·49*c* represents a sounding made at Duxbury on Sept. 27, 1944 under

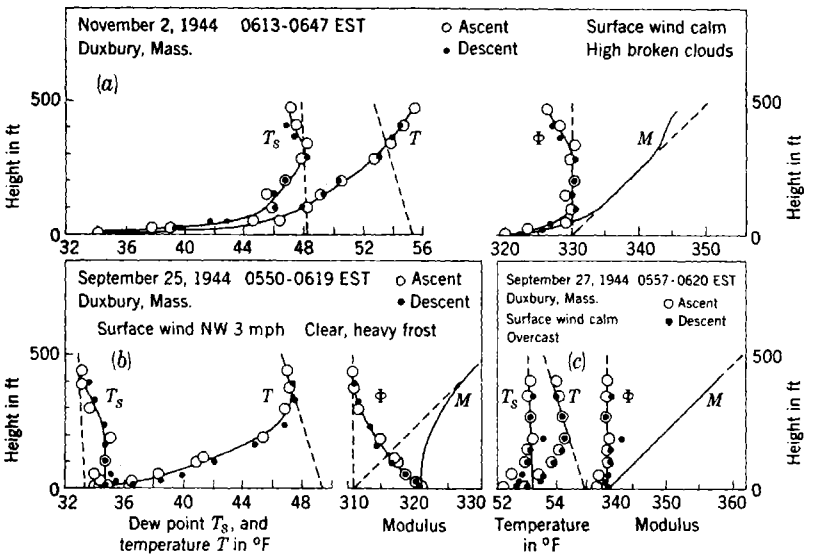


FIG. 3·49.—Examples of substandard, superstandard, and standard surface layers over land at night.

conditions of temperature excess and humidity deficit such that the M-profile was standard in the lowest 400 ft.

These examples have been intended to show the type of data that has been obtained and the kind of information that is now available from such observations. A complete and careful study will have to be made of a large number of similar low-level soundings before the problem will be near a satisfactory solution.

The information contained in this section and in that part of Sec. 3·14 relating to heating from below over land may be used to form a picture of a normal diurnal cycle over land in the M-profile. At midday, when the air is being convectively mixed by contact with the warm surface, the M-profile is homogeneous, or approximately standard. As sunset approaches, the ground must cool to some extent, as the sun's rays become more nearly horizontal, so that the temperature excess becomes positive

and stratification begins at the surface. Whether or not this can produce any important changes in the M-profile before sunset is uncertain, but it is obvious that any such changes must become progressively more pronounced during the night as the radiative cooling of the surface proceeds. It has been seen that a variety of M-profiles may occur during the night, superstandard or substandard anomalies being possible either at the surface or aloft. The important consideration in such a generalized discussion as this is the fact that these effects usually occur at night. In the morning, after sunrise, the ground is warmed by insolation and the air is heated by convective processes, as described in Sec. 3·14. Figure 3·8 shows the mechanism by which anomalies at the surface are carried aloft, usually decreasing in intensity, as this heating proceeds. The M-profile is nearly standard in the convectively mixed layer, the height of which gradually increases during the morning until the cycle is completed.

The sequence discussed here must, in its general features, be a very common one, but it should be remembered that other factors, with no dependence on time of day, may operate to minimize or mask it entirely. Cloudiness and strong winds tend to decrease the amount of change; an extensive snow surface may cause the irregular M-profiles to persist during the day, particularly in high latitudes; and such large-scale meteorological phenomena as frontal passages and wind shifts may alter the picture entirely.

This diurnal cycle over land must be present to some degree at all seasons of the year. The transmission experiments described in Chap. 4 indicate that it is more pronounced in summer than in winter. The reasons for this probably depend upon a great many factors, but a few simple points are worth mentioning. (1) In summer the ground is usually considerably drier than it is in winter. Because moist ground is a much better conductor than dry ground (see page 255), the surface of the ground may be expected to cool considerably more in summer than winter, inasmuch as less heat can be brought from below the surface to replace that lost by radiation. Brunt[1] concludes that this factor causes more intensive nocturnal temperature inversions in summer than in winter, despite the longer nights that occur in winter. (2) The water-vapor content of the air is generally considerably higher in summer than in winter; consequently stronger gradients of humidity are likely to develop at night to contribute to irregularities in the M-profile. Moreover, when the vapor pressure of the air is higher, a smaller temperature gradient is required to produce an M-inversion. This may be illustrated using the graphs in Fig. 3·2; for the extremes shown there, the temperature gradient necessary to produce an M-inversion ($\Delta M = -4.78$ per 100 ft on the graphs) when the gradient of vapor pressure is zero varies between 6.8°F per 100 ft

[1] D. Brunt, *Physical and Dynamical Meteorology*, Cambridge, London, 2d ed., 1939, p. 145.

(winter conditions) and 4.4°F per 100 ft (summer conditions). (3) It can be said that in temperate latitudes, cloudiness and strong winds are more apt to occur in winter than in summer and consequently reduce the over-all effect of the diurnal cycle in winter.

OTHER ATMOSPHERIC PROCESSES AND THEIR EFFECT ON *M*-PROFILES

Thus far, Φ- and M-profiles have been discussed only in relation to interaction between air masses and the earth's surface and attendant phenomena such as shear. Under this heading will be discussed the effects of certain other common meteorological processes.

3·22. Subsidence and Subsidence Inversions.[1]—Thus far, only small-scale vertical motions, such as those connected with eddy diffusion, have been considered, but there are also some large-scale processes of interest. One such type of vertical motion and shear of vertical motion familiar to meteorologists is known as *subsidence*. Subsidence is a sinking or lowering of large masses of air from high to low levels. The causes of its occurrence are essentially dynamic, such as the frictional outflow of air near the surface associated with high-pressure systems, the outflow of air near the centers of pressure rise, and the divergence of air due to change in latitude as an air mass moves from north to south. It is sufficient to say here that it generally is associated with high-pressure systems.

Subsidence has an important effect on the temperature lapse rate in the subsiding air. In a layer of air that subsides without loss of mass, an originally stable lapse rate becomes more stable, a dry-adiabatic lapse rate remains unchanged, and an originally unstable lapse rate becomes more unstable. The first case is by far the most frequent one in the atmosphere. Moreover, if, as is frequently the case, the subsiding layer diverges or loses mass, it becomes even more stable.

There is obviously a lower limit to which the air can descend before it spreads out horizontally. The ultimate limit is the surface of the earth, but for practical purposes the top of the layer of frictional influence usually serves as the base of subsidence. Below this, the air may be mixed by convection or turbulence. The base of the subsidence is usually characterized by a *subsidence inversion*. The original cause of the inversion may be subsidence itself or differential radiation from haze and moisture layers in the atmosphere. However, there are several factors that may intensify it, such as further subsidence, a shear in vertical motion at the subsidence surface such that the downward motion is greater above the inversion, and radiation from haze and smoke that may rise from the earth's surface up to the base of the stable layer.[2]

[1] By Richard A. Craig.

[2] For a more detailed discussion of these factors, the reader is referred to J. Namias, *Subsidence within the Atmosphere*, Harvard Meteorological Studies No. 2, 1934, pp. 10-13.

In a subsidence inversion the humidity frequently decreases rapidly with height. The air above the inversion, having descended from higher levels, is usually very dry. Furthermore, if the surface is moist and the subsidence inversion is low enough, water vapor may be carried up to the base of the inversion by convection or turbulence, intensifying the humidity contrast. Thus, the specific humidity usually decreases with height through the inversion, and the relative humidity decreases even more because of the additional factor of temperature increase in the layer.

Because of the temperature increase, which is always present, and because of the humidity decrease, which is frequently present, there is practically always a superstandard M-gradient in a subsidence inversion. There is often an M-inversion, particularly over the oceans.

It has been mentioned that subsidence is usually associated with high-pressure systems. These include the migratory highs over land, the migratory highs over water, and the subtropical anticyclones over the oceans; they are listed in order or increasing probability that the superstandard layer will contain an M-inversion. The semipermanent high-pressure systems over the oceans are known to exhibit the most marked subsidence in their eastward and equatorward portions.

As a general indication of the frequency, strength, height, and thickness of M-inversions associated with subsidence, Tables 3 1 and 3·2 present statistical data based on aerological ascents.

TABLE 3 1.—RADIOSONDE ASCENTS AT SAN DIEGO AND OAKLAND DURING JULY AND AUGUST 1943

Station	San Diego	Oakland
Number of observations	110	110
Fraction of all 8 P.M. PST observations which showed an M-inversion through a subsidence inversion	55%	31%
Fraction of all 8 A.M. PST observations which showed an M-inversion through a subsidence inversion	53%	44%
Fraction of all observations with an M-inversion through a subsidence inversion	54%	37%
Average elevation in ft of base of all subsidence inversions which showed M-inversions	c. 1800	
Average elevation in ft of top of all inversions which showed M-inversions	c. 2700	
Average decrease of M through the M-inversion	12	

TABLE 3·2.—KITE ASCENTS FROM METEOR EXPEDITION IN NORTH AND SOUTH ATLANTIC OCEAN, 1925 TO 1927

The North Atlantic soundings were made between 6° and 20°N latitude.
The South Atlantic soundings were made between 9° and 30°S latitude.

	North Atlantic	South Atlantic
Number of cases studied	49	46
Fraction of all daytime observations which showed M-inversions	57%	44%
Average elevations in ft of base of M-inversions	3000	4500
Average decrease of M through subsidence inversions	6	6
Average thickness in ft of M-inversions	375	400
Season during which observations were made	Winter and spring	Fall and winter

Table 3 1 contains an analysis of radiosonde data obtained at Oakland and San Diego, Calif.,[1] in July and August 1943.

These stations at this time of year are on the eastern edge of one cell of the Pacific anticyclone Table 3·2 contains a similar analysis of data obtained by kite on the *Meteor* expedition.[2] A word of caution should accompany the presentation of these tables. Because of the lag of the thermometric and humidity elements, particularly the latter, regions of rapid change in the atmosphere are often not accurately measured by sounding equipment in common use. In the case of a subsidence inversion, the thickness of the layer may be overestimated and the changes in temperature and humidity are underestimated. Both factors make the decrease of M through the inversion appear to be less than it actually is. Accordingly, the figures giving percentage of times when an M-inversion is present and average decrease of M through the inversion should be considered as minimum values for the data analyzed.

Because subsidence inversions are more persistent than most meteorological phenomena, the forecasting problem is simplified Particularly over the ocean in the subtropical highs, changes are gradual and daily changes may often be extrapolated for predictions of future developments.

[1] A large amount of observational work has been done in connection with elevated M-inversion at San Diego by the Navy Electronics Laboratory. The program has included meteorological soundings and observations of associated radio phenomena. This work is discussed in Sec. 4 9, which includes a typical sounding made with the wired sonde (Sec. 3 28).

[2] Deutsche Atlantische Expedition *Meteor*, 1925–1927.

3·23. Fronts and Frontal Inversion.[1]—An *air mass* is a large body of air that has approximately uniform temperature and water-vapor concentration horizontally. The boundary between two such air masses is called a *frontal surface*, and the line of intersection between the frontal surface and the earth's surface is called a *front*.[2] The warmer air mass, being lighter, rises above the colder one. The frontal surface is, in general, not parallel to the surface of the earth because of forces other than gravity, especially the deflecting force caused by the earth's rotation.

The slope of the frontal surface is dependent upon the temperature difference between the two air masses and the wind shear across the frontal surface. Although a frontal surface is always nearly horizontal, it tends to be more nearly horizontal for a small wind difference and a large temperature difference. The slope is also directly proportional to the sine of the geographical latitude at which the front is located. Frontal surfaces that occur in the atmosphere have slopes lying between 0 and $\frac{1}{25}$.

When the front moves so that the cold air replaces the warm air over a portion of the earth's surface, it is called a *cold front*. When the warm air moves over regions previously covered by the cold air, the front is said to be a *warm front*. Cold fronts are usually more nearly vertical than warm fronts, the slope of the latter being generally between $\frac{1}{100}$ and $\frac{1}{300}$ and of the former between $\frac{1}{50}$ and $\frac{1}{150}$.

Frontal analysis is an important aspect of weather forecasting. The subject of fronts is dealt with much more fully in most standard meteorological texts than is necessary here.[3]

Fronts are important in a consideration of temperature, humidity, and M-profiles. At the frontal surface, because the overriding air mass is the warmer one, there is always a stable layer and there is often a temperature inversion.[4] The water-vapor concentration in this layer ordinarily, but not always, increases with height because the warmer air mass is more moist. No general statement can be made concerning the shape of the M-profile; an M-inversion is probably rare because of the usual humidity increase with height, but it is a definite possibility, as are standard or substandard layers. This should be contrasted to subsidence inversions, where superstandard layers occur almost without exception. In the case of any particular front it is necessary to obtain accurate meteorological soundings to be sure of the M-profile, although a comparison of surface values of humidity in the two air masses may often give some indication.

[1] By Richard A. Craig.

[2] The entire frontal surface is often referred to as a front.

[3] For example, S. Petterssen, *Weather Analysis and Forecasting*, McGraw-Hill, New York, 1940.

[4] The frontal surface is not a sharp dividing line between the two air masses, but rather a zone of finite width where temperature, humidity, and wind change abruptly. "Abruptly" is used here in the meteorological sense; the layers are many wavelengths thick for microwaves.

The height of the frontal surface above the earth varies widely. It can be found approximately either from a radiosonde observation or from a consideration of the position of the front and the slope of the frontal surface.

A second and indirect consideration with respect to fronts is that weather conditions may change radically at a given place after a frontal passage. The accompanying change of air mass may cause large changes in temperature, humidity, and cloudiness and in the M-profile. As an example, consider a cold-front passage at a point over water. Prior to the frontal passage, the air mass may be warm and moist with a deep substandard layer based at the surface. The new air mass may be cool enough to be heated from below with the attendant shallow duct. With a warm frontal passage the sequence of events would be reversed. These relatively abrupt changes in shape of the M-profile are caused by the change in air mass only, not by the frontal surface itself. Care should be taken to distinguish between these two factors.

Close to fronts, it is known that vertical motions are particularly intense. Little is known about the effect of such vertical motions, but they are mentioned here by way of caution against any assumption of strictly horizontal advection in the vicinity of fronts.

Fronts may vary widely in such characteristics as temperature and humidity contrast between the two air masses, slope, wind discontinuity, and width of the transition layer. Consequently, it is dangerous and misleading to make general statements about the effects of fronts on propagation. In some parts of the world and at some seasons of the year, certain types of front may recur often enough to make general statements appear applicable. However, it is always safer to consider first the effects of fronts on the M-profiles and attempt generalities in terms of the latter.

3·24. Sea-breeze Circulations.[1]—Many radars are sited along coast lines, where the sea breeze is a frequent phenomenon, often causing striking nonstandard propagation effects.

The sea breeze is a complex combination of horizontal and vertical motions caused by distinct differences in temperature between land and water. On clear, warm days the contrast is built up along the coast by unequal heating of land and the adjoining water body. The warm air over land rises, and there is a flow at the surface of denser oceanic air toward land. Although a sea-breeze tendency can exist regardless of the prevailing wind, the most marked effects occur when there is an offshore wind above the surface onshore flow. For this case, the sea breeze is that local circulation cell in which there is a rising mass of air over the land

[1] By Isadore Katz.

that is carried out to sea by the prevailing offshore wind, a subsiding motion at some distance from shore, and a returning cool mass of air that is again heated over land. Often the sea breeze is confined to within 10 miles of the coast line, but cases are reported in which it extends as much as 100 miles from the coast.

It is difficult at present to describe the temperature, humidity, or M-profiles in a general way because of the paucity of soundings in sea-breezes.[1] Most sea-breeze data are restricted to surface observations and a few visual observations of conditions aloft. Those soundings which do exist show considerable variation among themselves. Figure 3·50 is

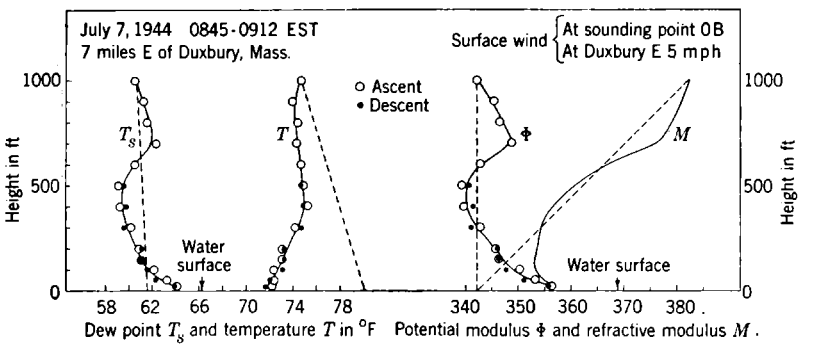


FIG. 3·50.—Overwater sounding in a sea breeze.

a sounding in a sea-breeze in Massachusetts Bay about 8 miles from the coast. Figure 3 51 is a sounding made over land in Kaikoura, New Zealand, in the presence of a sea-breeze.[2] Possibly the difference between these two measurements illustrates the amount of variation in sea-breeze characteristics in various parts of the world.

It is best to caution against assuming a unique type of propagation associated with a sea-breeze. Each case must be treated separately, with consideration being given to the properties of the original air mass and the temperature of the water. It should also be stressed that the sea-breeze circulation is clearly local in character, being confined for the most part to a narrow coastal strip, and should not be confused with winds coming from the sea with long overwater trajectories.

[1] Some observations are presented in the form of vertical cross sections by R. A Craig, *et al.*, "Sea-breeze Cross Sections from Psychrometric Measurements," *Bull. Am. Meteorol. Soc.*, December 1945.

[2] F. E. S. Alexander, "Temperature and Humidity Measurements Made with the Washington State College Wired Sonde Equipment at Kaikoura, New Zealand, between September 23, 1944, and October 19, 1944," Radio Development Laboratory, Wellington, New Zealand.

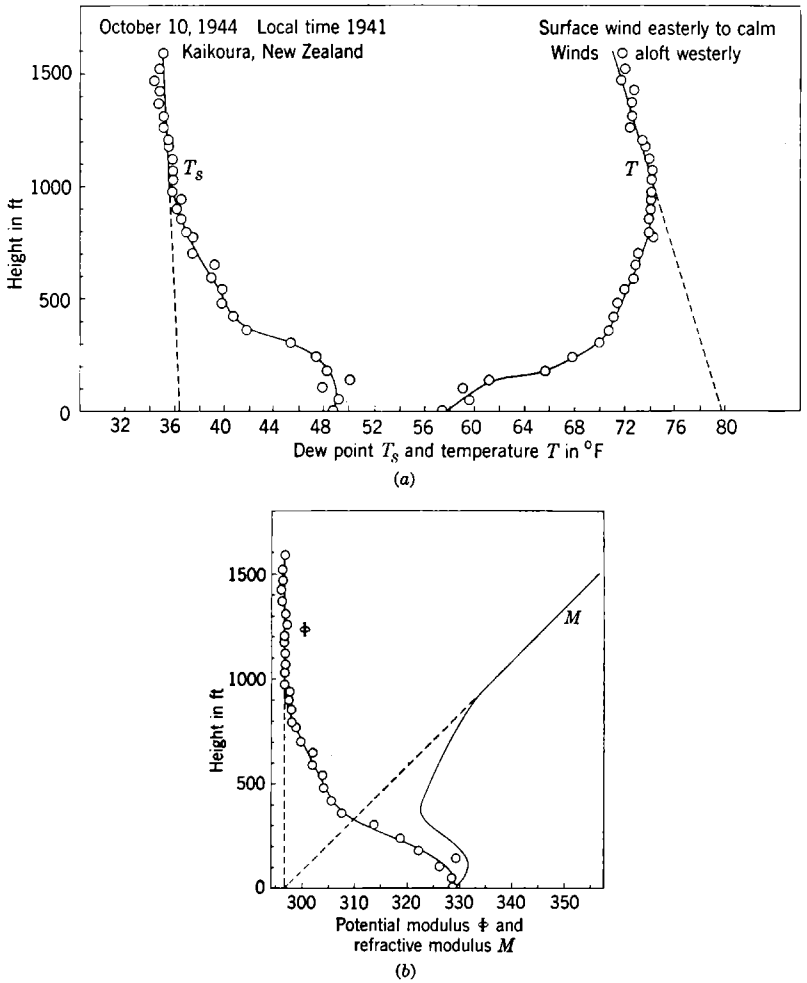


FIG. 3·51.—Overland sounding in a sea breeze.

HORIZONTAL AND LOCAL VARIATIONS

BY RICHARD A. CRAIG

Emphasis thus far has been placed on vertical gradients, as the present theoretical basis for study of the effects of these gradients requires the assumption of purely horizontal stratification. Moreover, variations with time at a given place have been generally neglected, both in the theoretical discussion and in the observational techniques. However, it is of importance to consider possible deviations from these ideal conditions.

3·25. Horizontal Gradients.—It is apparent from the complexity of atmospheric phenomena and the nonuniformity of the earth's surface that horizontal gradients of temperature, humidity, and refractive modulus must occur frequently in the atmosphere. In the large-scale study of meteorology, it is known that these are minimum within well-defined air masses and maximum near frontal surfaces. In radio-meteorology this is correspondingly true for the parts of the atmosphere higher than a few hundreds of feet above the earth's surface. For the surface layers, however, which assume more importance in the specialized study here considered than in ordinary meteorological work, significant gradients may occur also within air masses because of the rapid reaction of these surface layers to changes in the modifying surface. A precise definition of what constitutes a significant variation of M-profile has not been attempted, but it would necessarily involve the wavelength, the geometry of the transmission path, and the heights at which the changes in the M-profile occur. Large changes in the M-profile can occur frequently and fairly rapidly. These occur across frontal surfaces and also in surface layers in the absence of fronts. Changes across frontal surfaces tend to be abrupt, their amount depending on the air-mass contrast. Because a frontal surface has a small but finite slope, the boundary layer between the two air masses may vary in height by as much as 10,000 ft in 100 miles and usually varies at least 1000 ft in 100 miles perpendicular to the front. Moreover, the changes in temperature and humidity across a front are generally accompanied by changes in temperature excess and M-deficit.

The vertical distribution of temperature and humidity in an air column follows to a great extent from the type of surface by which the air has been recently modified. When a large body of air is modified by a nearly uniform surface for a sufficiently long time, an air mass is formed. Of importance here is the fact that in the surface layers variations resulting from changes in the modifying surface may occur in the absence of such prolonged and uniform modification.

In general, such variations occur most frequently at coast lines because of the abrupt change in modifying surface. The horizontal gradients may be of several types. For example, in the ideal case of warm air over cold water (Sec. 3·19) a nearly standard M-profile over land may become substandard or superstandard in the surface layer as soon as it crosses the coast; and furthermore, the shape of the curve may vary considerably during its trajectory. Moreover, because of irregularities in the coast line and wind differences, other changes may occur. This was of importance in the experiment in Massachusetts Bay where significant horizontal gradients were often found (Sec. 4 2).

Horizontal gradients may occur to a lesser extent over continents away from the influence of coastal waters. Land surfaces may vary

considerably in humidity and in heating or cooling properties, and their effect on air columns may produce significant variations (Sec. 3·21).

Probably the open ocean is the region where horizontal gradients are the least important in the absence of fronts. If all the air over a region has traveled hundreds of miles over water, such gradients are less likely to exist. Even here, however, one must be very careful to allow for variations in water temperature, which can produce significant differences.

Obviously, no general rules can be given regarding horizontal gradients. It is important that their existence be borne in mind in choosing any site for experimental purposes and in analyzing any data. In particular, if a measured M-profile is to be correlated with radio transmission phenomena, considerable thought should be given as to whether or not the meteorological data are representative of the entire transmission path.

3·26. Local Variations with Time.[1]—Previous discussions in this chapter relative to values of temperature, moisture concentration, and potential or refractive modulus at a point in the atmosphere have referred to average and not to instantaneous values. Theoretical considerations refer to averages; and furthermore, as will be described in the next section, the observational material was gathered according to a technique that tended to measure average conditions over a time interval of minutes. Nevertheless, certain theoretical and observational material leads to the conclusion that variations about these averages may be considerable, particularly under conditions of instability.

In the case of unstable equilibrium, vertical motions result from convection and mechanical turbulence. An air parcel displaced vertically may travel an appreciable distance before being mixed with its surroundings; and if there are vertical gradients of potential temperature, specific humidity, or potential modulus, such displaced air parcels cause variations in the parameters in their new surroundings. In order to measure the resultant fluctuations in M, one would need simultaneous and rapid measurements of both temperature and humidity. These are difficult to obtain, but it is possible to estimate the fluctuation in M from measurements of temperature alone. Over dry ground, of course, there is no gradient of specific humidity, and the fluctuations in M depend on temperature variations alone. Over wet ground or water, if the temperature excess and M-deficit are known, the estimate can be made on the basis of the principle of similarity, namely, that the ratio of temperature fluctuations to temperature deficit is identical with the ratio of fluctuations in M to M-deficit.

Some indication of the fluctuations encountered in unstable equilibrium may be obtained from the work of Gerdel,[2] who measured temperature

[1] Part of the material in this section is based on an unpublished memorandum by R. B. Montgomery (1945).

[2] R. W. Gerdel, "Temperature-gradient Observations," *Nat. Research Council Am. Geophys. Union Trans.*, Part I, 1943, pp. 182–189.

with thermocouples over land at heights up to 24 ft. He found the standard deviation of temperature in unstable air to have a maximum value of 3°F at 4 ft, decreasing to 0.9°F at 24 ft. Over dry ground, with no gradient of vapor pressure, this would correspond to standard deviations of about 3 and 1 respectively in M. These represent deviations of less than 1 per cent of the usual value of M near the surface of the earth.

Information concerning deviations over water is available from psychrometer measurements made at the masthead of a ship by Woodcock.[1] One series made on Feb. 24, 1945, over the Gulf Stream represents a temperature excess of −13.5°F, a humidity deficit of 17.2 mb, and an

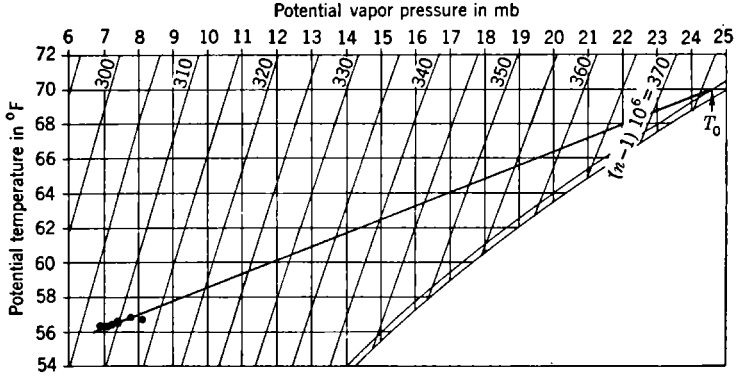


FIG. 3·52.—Fluctuations in unstable air.

M-deficit of 68. The first seven readings of this series, made at ½-min intervals with a wind of 14 knots, are plotted on the characteristic diagram in Fig. 3·52. Note that the points fall reasonably well along a straight line through the observed water temperature of 70°F. The maximum recorded deviation in M was 3, or about 4 to 5 per cent of the M-deficit.

Further information comes from some soundings made in Florida by the Army Air Forces[2] using an aerated wired sonde (Sec. 3·28). Measurements were made over water in the lowest 50 ft of the atmosphere. The instrument was exposed at each level for 2 to 3 min, and the maximum and minimum observed values of humidity were recorded. Because the deviations were small, only average values of temperature were recorded; this was unfortunate. Figure 3·53 is one such sounding made in a wind of about 8 knots. The sea temperature was 75.2°F, and the temperature excess, humidity deficit, and M-deficit were, respectively, about −6°F, 16 mb, and 64. The deviations above 25 ft were 5 to 8 per cent of the deficits, although fluctuations from the average of as much as 17 per cent

[1] Alfred H. Woodcock, Woods Hole Oceanographic Institution, during passage between the West Indies and an eastern United States port in February 1945.

[2] AAF Board Project No. H4298.

of the deficit occurred nearer the surface. Note that with the irregular fluctuations at all levels the shape of the M-profile, including the duct height, must vary considerably in an irregular manner.

It seems reasonable, on the basis of the foregoing information, to estimate the maximum possible deviation in M above 25 ft in unstable air. The Florida sounding indicated a value of about 8 per cent of the M-deficit. Woodcock's series showed a 4 to 5 per cent maximum deviation, a figure that must be somewhat low because of lag in the response of the sling psychrometer, which was exposed at $\frac{1}{2}$-min intervals. If Gerdel's value of 0.9°F for temperature is assumed to be more nearly correct for

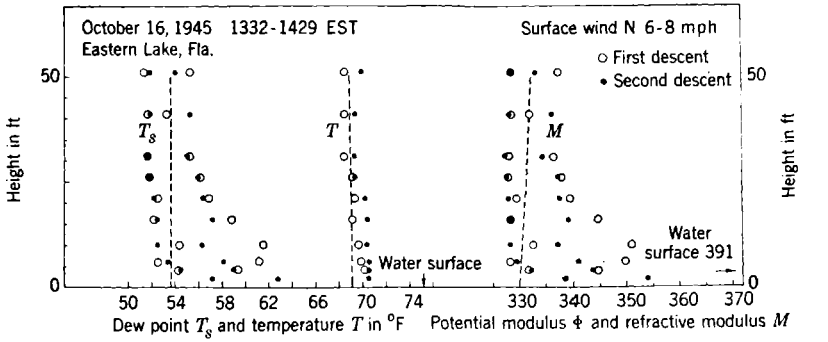


FIG. 3·53.—Fluctuations in dew point and resulting fluctuations in M in unstable air.

this series, the deviation becomes nearly 7 per cent. Eight per cent appears as an acceptable maximum figure. According to these figures deviations of M from the average value above 25 ft because of turbulence and convection could scarcely exceed ± 5.

Another series of Woodcock's measurements serves to throw light on variations in air near neutral equilibrium. Made on Feb. 20, 1945, in the Windward Passage, this series represents a temperature excess of -0.5°F, a humidity deficit of 11.3 mb, and an M-deficit of 51. If there is no vertical gradient of potential temperature, deviations in temperature cannot occur. This is the case for neutral equilibrium except near the top of the mixed layer. However, over water there are always vertical gradients of humidity and M near the surface (Sec. 3·14) so that fluctuations in these quantities can occur. In this case, maximum recorded deviations were about 8 per cent of the deficits. The fluctuations observed by Woodcock in the unstable case were less than those derived from other data. Accordingly, it is reasonable to expect that this figure of 8 per cent for neutral equilibrium may be exceeded at times. Thus it is possible that the variations may sometimes be as much as ± 8 in M.

When the air is in stable equilibrium, vertical displacements can be only temporary (Sec. 3·17). The fluctuations then are associated with

wave motion in the stable fluid rather than with discrete turbulent eddies. Considerable theoretical work has been done by dynamic meteorologists concerning the physical characteristics of such wave motion; but unfortunately, most of it is not directly applicable to the immediate problem. Certain principles may, however, be stated. At temperature inversions in the atmosphere, which often correspond to large vertical M-gradients, wind shear may supply enough energy to set up wave motion in much the same way that waves are formed at the surface of the ocean. Gravitation acts as a stabilizing or restoring effect provided the wavelength of the motion is sufficiently large. The critical wavelength, above which stable wave motion can occur, increases with increasing wind shear and decreasing stability.[1] Observations of billow clouds have indicated that the wavelengths which occur in nature are near these critical values.[2] The vertical velocities associated with such wave motions are negligible at a distance above and below the inversion greater than 37 per cent of the wavelength. This may be assumed to be the same as the critical wavelength.

The inferences that may be drawn from this meteorological knowledge are

1. Fluctuations must be less in stable equilibrium than in neutral or unstable equilibrium, as vertical motions are inhibited.
2. Deviations must decrease with increasing stability and decreasing wind shear in a stable layer, as critical wavelength decreases and the effective range of vertical motion decreases.[3]

Gerdel's measurements included some cases of stable equilibrium and included no temperature deviations greater than 0.5°F over dry land. This would indicate maximum fluctuations in M of about half the value found for unstable equilibrium, about 4 per cent or 2 or 3 per cent in M. This is admittedly a rough figure and must vary somewhat with stability and wind shear.

There is room for much additional research, both theoretical and observational, in the subject of local variations, particularly in stable equilibrium. One important aspect of wave motion, in addition to variations at a point, is the attendant possibility that a duct may vary in height horizontally. It has been suggested that this phenomenon may be of importance in causing the deep fades observed during the guided propagation of microwaves.[4]

[1] A table of critical wavelengths in terms of wind shear and temperature difference is given in B. Haurwitz, *Dynamic Meteorology*, McGraw-Hill, New York, 1941, p. 287. Note that these hold only if the height of the inversion above the ground is greater than 40 per cent of the wavelength.

[2] *Ibid.*, p. 288.

[3] It should not be inferred that most stable radio transmission occurs under stable atmospheric conditions; in Chap. 4 it will be shown that the reverse is usually true.

[4] W. M. Elsasser, "Comments on Duct Transmission and Fading," NDRC Propagation Memorandum, December 1945 (unpublished).

INSTRUMENTS TO MEASURE TEMPERATURE AND HUMIDITY IN THE LOWER ATMOSPHERE

BY ISADORE KATZ

The study of microwave propagation places new emphasis on meteorological instrumentation.[1] For the usual meteorological purposes temperature and humidity are measured with instruments unsuited to the fine detail necessary for low-level soundings. The development of several new instruments and variations of some older types arose out of the needs of this novel sounding problem. In the following sections some of these devices, the sounding methods, and problems arising from their use are discussed.[2]

3·27. Psychrograph.—The psychrograph is a wet- and dry-bulb recording electrical resistance thermometer.[3] It employs two thermistors for measuring wet- and dry-bulb temperature. The resistance of each thermistor is measured with the aid of a Wheatstone bridge and amplifier, the output current of which operates a recorder. The thermistors are mounted in a housing designed for the particular type of service desired. The psychrograph was designed originally for making aircraft soundings and when used in this form is called an *aeropsychrograph*. It has also been modified for use as a ground-based sounding instrument described in the following pages.

Thermistors used as temperature elements for the psychrograph are made of a ceramic material that has been fired at a high temperature, yielding a glasslike, hard semiconductor.[4] They are used in radiosondes for measuring temperature. They are shaped into cylindrical rods of various diameters and lengths. Those used in the psychrograph are 0.06 in. in diameter and 1.9 in. long. Figure 3·54 shows a typical temperature-resistance curve.[5]

[1] For a text on standard instruments, see W. E. K. Middleton, *Meteorological Instruments*, University of Toronto Press, Toronto, Canada, 1941.

[2] Further details on the instruments described here and some others may be found in W. M. Elsasser, "Meteorological Equipment for Short Wave Propagation Studies," Wave Propagation Group Report No. WPG-3 Columbia University, August 1944.

[3] This instrument was developed during the course of propagation research at the Radiation Laboratory and was used almost exclusively to obtain the RL data given in preceding sections. It has also been used by several other groups. See I. Katz, "Instruments and Methods for Measuring Temperature and Humidity in the Lower Atmosphere," RL Report No. 487, Apr. 12, 1944; "An Airplane Psychrograph," *Bull. Am. Meteor. Soc.*, **28**, 363 (1947).

[4] The thermistors described here were obtained from the Washington Institute of Technology, College Park, Md., and from the Julien P. Friez Co., Baltimore, Md. There are several essentially similar types now available commercially. For a general description of thermistors see J. A. Becker, C. B. Green, and G. L. Pearson, *Trans.* AIEE, **65**, 711 (1946); also Vol. 10 of the Radiation Laboratory Series.

[5] Compensating networks to correct the nonlinearity of the temperature-resistance curve are described in P. A. Seay and W. E. Gordon, "Temperature Recording with Thermistors," Electrical Engineering Research Laboratory Report No. 16, June 1, 1948, University of Texas, Austin, Tex.

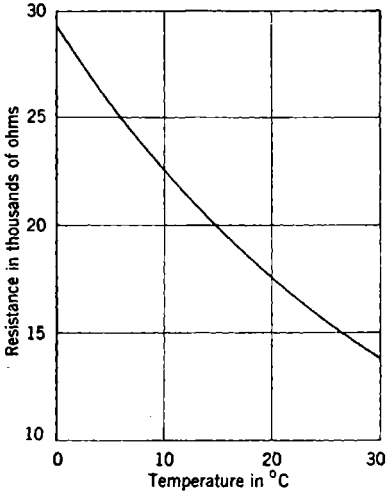


FIG. 3·54.—Temperature-resistance curve for a typical thermistor.

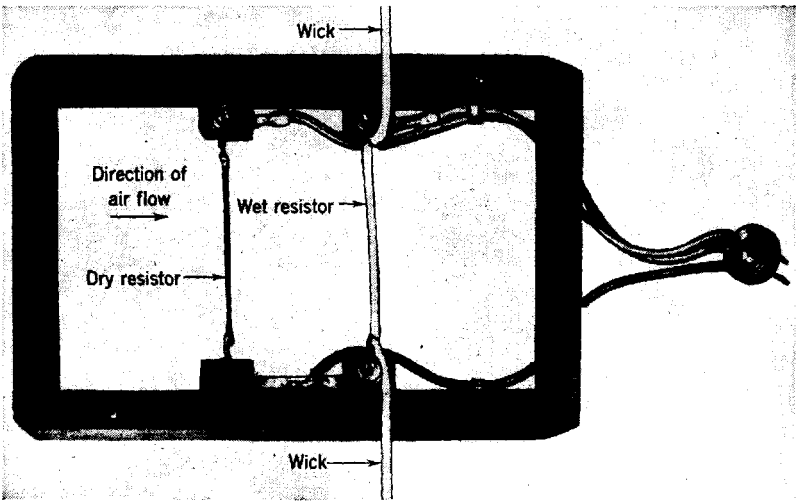


FIG. 3·55.—Aeropsychrograph thermistor rack.

When used in the aeropsychrograph the thermistors are mounted on an open rectangular bakelite rack as shown in Fig. 3·55. This rack provides a rigid support for the thermistors, which are very brittle. The rack, thermistors, and leads are coated several times with polybutylmethacrylate, an insulating plastic that renders them water-resistant and eliminates surface leakage. No uncoated surfaces remain other than those of the connector prongs.

One of the thermistors is tightly covered with a wick and acts as a wet-bulb thermometer. Water is fed from both sides to assure adequate wetting. Large reservoirs, part of the housing to be described in the following paragraphs, provide the water supply for the wick.

The thermometer housing is shown in Fig. 3·56. It consists of an inner aluminum cylinder, an outer radiation shield, and two reservoirs for distilled water. The inner shield has two open conical ends that reduce the airflow through the instrument. This reduction amounts to about 60 per cent in the construction shown. A reduction in ventilation speed

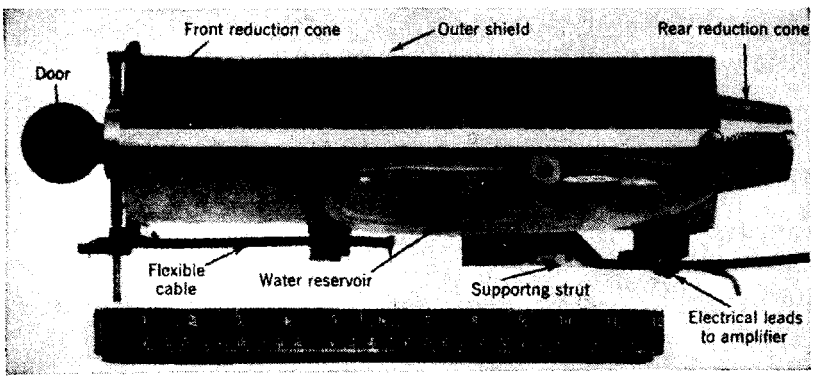


FIG. 3·56.—Thermometer housing of aeropsychrograph.

is desirable in aircraft instruments in order to have the wick remain completely wet at all times. This is not easily accomplished at the usual aircraft speeds.

The outer radiation shield is a hexagonal cylinder made of aluminum. There are attached to the shield the water reservoirs, the rear cone, the connector for the leads to the bridge, a special extension that fits into the supporting strut, and a door with a flexible steel control cable. The two Plexiglas water tanks are fastened on the under side of the outer shield. Each tank has a wick-protecting Plexiglas tube running to within $\frac{1}{2}$ in. of the wet-bulb resistor. These tubes act as shields for the wick to prevent excess evaporation. The wick is long, covering the wet thermistor and reaching into the two water tanks from the ends of the element. This, in effect, provides two wicks to feed the wet bulb, each wick having to wet only half the element.

The door mounted on the front of the housing provides protection for the wick when the instrument is not in use. When the plane is in flight but measurements are not being made, the door can be closed to prevent evaporation of water from the wick. This preserves the water supply and also prevents dust and hygroscopic particles from accumulating on the wick. The door is a thin brass disk that fits over the front

cone when closed and opens to a position parallel to the airflow. It is opened and closed from inside the aircraft by means of the flexible steel cable. A rear view of the housing is shown in Fig. 3·57. The rear cone has been removed to show details of the inside. The thermistors are mounted perpendicular to the air stream, for reasons discussed later in this section.

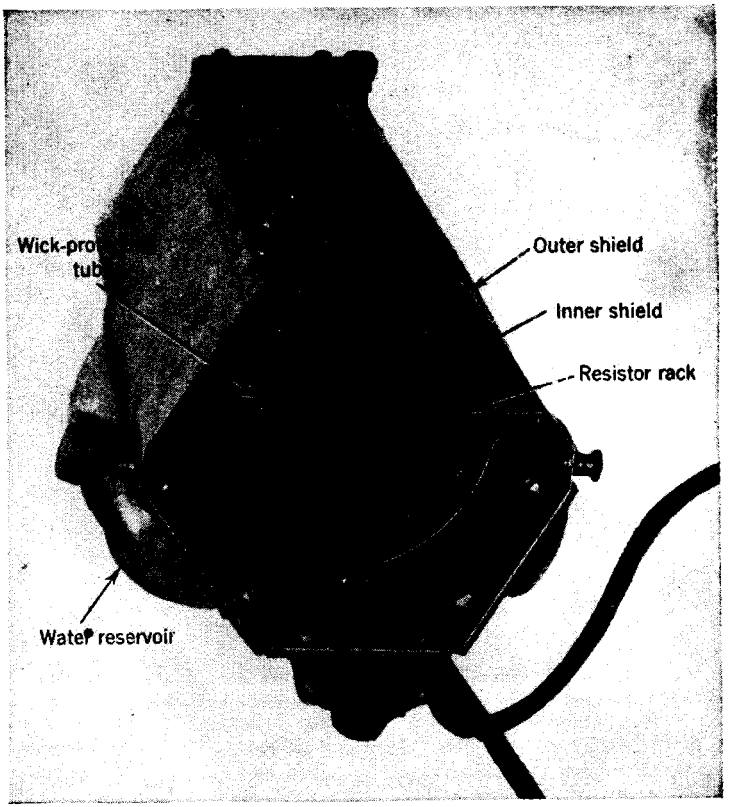


FIG. 3·57.—Aeropsychrograph housing, rear view, with rear cone removed and rack in operating position.

It is desirable to have the housing mounted in the free air stream, well away from any turbulent flow close to the aircraft. Such a position can be found at a point above and forward of the nose. Figure 3·58 is a photograph of a mounting on a PBY aircraft in which a good approach to free-air conditions was obtained.

A highly simplified circuit diagram of the pyschrograph is shown in Fig. 3·59. The wet and dry thermistors are switched alternately (automatically or manually, as desired) into one arm of a Wheatstone bridge.

This bridge operates off balance; the unbalance voltage is applied to the grids of a vacuum-tube voltmeter with high mutual conductance; and the resulting flow of current from one triode plate to the other operates the recording milliammeter M.[1] The switch S_2 selects the temperature range to be covered by the instrument. A range of about 45°F is generally covered by the full scale of the recorder, and sufficient overlapping of the temperature ranges is provided to ensure that both wet- and dry-bulb

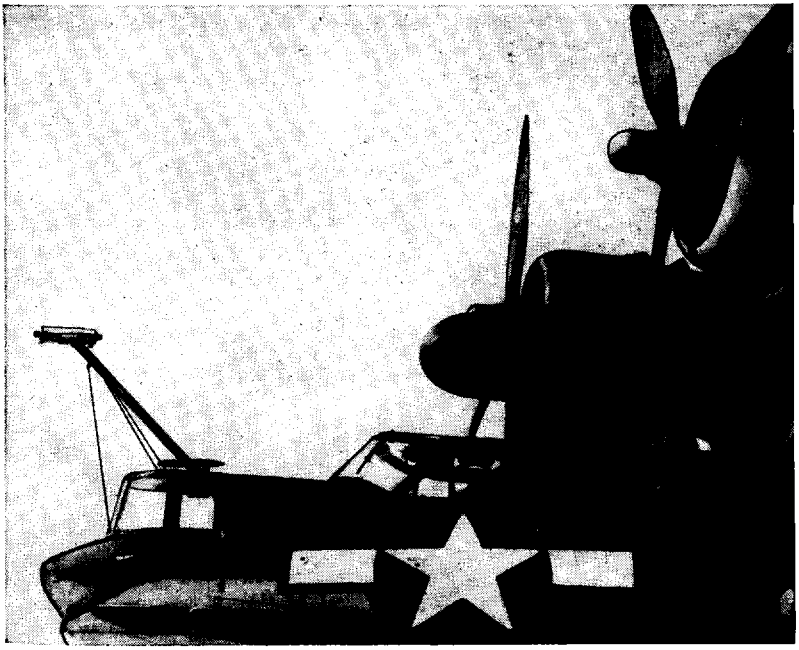


FIG. 3·58.—Aeropsychrograph mounted on a PBY aircraft. (*Courtesy of ASDEVLANT.*)

temperatures can be recorded without resetting S_2, even for large wet-bulb depressions. Because the bridge is operated off balance, the power-supply voltages and tube characteristics must be held to close tolerances. In the present case simple voltage-regulation methods employing gaseous voltage-regulator tubes sufficed to limit the errors to at most one-half of 1 per cent of full-scale deflection for ± 12 per cent variation of line voltage. To provide a quick check on correct amplifier performance the calibrating resistors can be switched into the circuit at will.

Speed Correction.—When an object is placed in a moving air stream, the air at the boundary is brought to rest. As a consequence, a velocity gradient occurs near the surface and there is local heating, resulting in an

[1] The recording milliammeter is the 0-1 ma Model AW Instrument of the Esterline-Angus Co.

increase of the temperature of the boundary above that of the passing air stream. Thus, when the resistor in the aeropsychrograph is exposed to a moving air stream, there is a masking increase in temperature, which varies with the speed of the aircraft. A correction must be applied to the indicated temperature to obtain the true air temperature.

The correction may be estimated by assuming steady flow in an incompressible fluid and no friction. Let u be the velocity along a streamline, s the distance along a streamline, and A the cross section of the tube

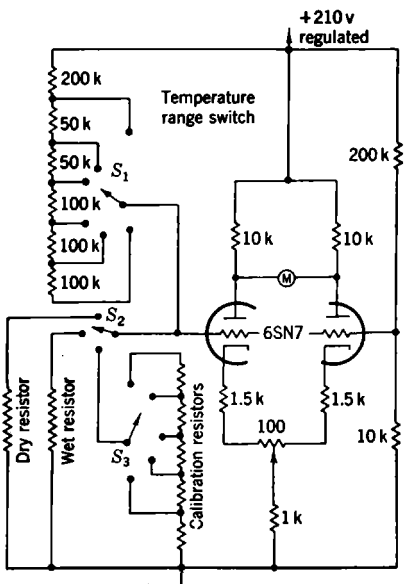


FIG. 3·59.—Simplified circuit diagram of the psychrograph.

considered. Let ρ be the air density and p the pressure. Steady flow implies that u, p, ρ, and A may vary with s but not with time. Because steady flow is assumed,

$$\rho u A = \text{const.}$$

The force from the pressure gradient on an infinitesimal volume $A\ \delta s$ in the direction of flow is $-(dp/ds)A\ \delta s$. The mass of that volume is $\rho A\ \delta s$, and its acceleration is du/dt. From Newton's second law,

$$-\frac{dp}{ds} A\ \delta s = \rho A\ \delta s \frac{du}{dt}.$$

But

$$\frac{du}{dt} = \frac{du}{ds}\frac{ds}{dt} = u\frac{du}{ds};$$

hence

$$\frac{1}{\rho}\frac{dp}{ds} + u\frac{du}{ds} = 0.$$

This can be written as

$$\int \frac{dp}{\rho} + \int u\, du = 0, \tag{57}$$

which is a form of Bernoulli's equation.

From the first law of thermodynamics,

$$dQ = c_p\, dT - \frac{dp}{\rho}, \tag{58}$$

where Q is heat transferred and c_p is the specific heat at constant pressure. For an adiabatic process, $dQ = 0$; hence Eq. (58) becomes

$$c_p\, dT = \frac{dp}{\rho}.$$

If this result is combined with Eq. (57), it follows that

$$\int_T^{T+\Delta T} c_p\, dT = -\int_u^0 u\, du,$$

or

$$\Delta T = \frac{1}{2}\frac{u^2}{c_p}. \tag{59a}$$

The increase in temperature is, then, proportional to the square of the speed. If the speed u is given in miles per hour,

$$\Delta T = 1.0\left(\frac{u}{100}\right)^2 °\text{C} = 1.8\left(\frac{u}{100}\right)^2 °\text{F}. \tag{59b}$$

Equation (57) holds for a full adiabatic compression. Temperature rise may be caused by adiabatic compression or by friction, but the combined effect may be treated as one.[1] The rise in temperature will usually be a fraction of the full adiabatic compression, the exact amount depending on the geometry of housing and thermistors. If this fraction is denoted by ϵ_d,

$$\Delta T = 1.8\,\epsilon_d\left(\frac{u}{100}\right)^2 °\text{F}. \tag{60}$$

Table 3·3 gives values of ΔT in degrees Fahrenheit for various values of u and ϵ_d. When ϵ_d is found for a particular thermometer, the table gives the value to be *subtracted* from indicated temperature to obtain true temperature.

[1] This treatment follows closely that of A. F. Spilhaus, "Atmospheric Pressure, Temperature and Humidity, the Corrections for Aircraft Speed," August, 1943 (unpublished).

TABLE 3·3.—SPEED CORRECTION (ΔT IN °F) FOR VARIOUS VALUES OF ϵ_d AND AIR SPEEDS IN MPH

True air speed, mph, u	ϵ_d									
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
80	0.1	0.2	0.3	0.5	0.6	0.7	0.8	0.9	1.0	1.2
90	0.1	0.3	0.4	0.6	0.7	0.9	1.0	1.2	1.3	1.5
100	0.2	0.4	0.5	0.7	0.9	1.1	1.3	1.4	1.6	1.8
110	0.2	0.4	0.7	0.9	1.1	1.3	1.5	1.7	2.0	2.2
120	0.3	0.5	0.8	1.0	1.3	1.6	1.8	2.1	2.3	2.6
130	0.4	0.6	0.9	1.2	1.5	1.8	2.1	2.4	2.7	3.0
140	0.4	0.7	1.1	1.4	1.8	2.1	2.5	2.8	3.2	3.5
150	0.4	0.8	1.2	1.6	2.0	2.4	2.8	3.2	3.6	4.0
160	0.5	0.9	1.4	1.8	2.3	2.8	3.2	3.7	4.1	4.6
170	0.5	1.0	1.6	2.1	2.6	3.1	3.6	4.2	4.7	5.2
180	0.6	1.2	1.7	2.3	2.9	3.5	4.1	4.7	5.2	5.8
190	0.6	1.3	1.9	2.6	3.2	3.9	4.5	5.2	5.8	6.5
200	0.7	1.4	2.2	2.9	3.6	4.3	5.0	5.8	6.5	7.2
210	0.8	1.6	2.4	3.2	4.0	4.8	5.6	6.3	7.1	7.9
220	0.9	1.7	2.6	3.5	4.4	5.2	6.1	7.0	7.8	8.7
230	1.0	1.9	2.9	3.8	4.8	5.7	6.7	7.6	8.6	9.5
240	1.0	2.1	3.1	4.1	5.2	6.2	7.3	8.3	9.3	10.4
250	1.1	2.2	3.4	4.5	5.6	6.8	7.9	9.0	10.1	11.2
260	1.2	2.4	3.6	4.9	6.1	7.3	8.5	9.7	10.9	12.2
270	1.3	2.6	3.9	5.3	6.6	7.9	9.2	10.5	11.8	13.1
280	1.4	2.8	4.2	5.6	7.1	8.5	9.9	11.3	12.7	14.1
290	1.5	3.0	4.5	6.1	7.6	9.1	10.6	12.1	13.6	15.1
300	1.6	3.2	4.9	6.5	8.1	9.7	11.3	13.0	14.6	16.2

Because the air surrounding the wet bulb is saturated, it is assumed that the wet-bulb speed correction will bear the same ratio to a dry-bulb correction as the moist-adiabatic lapse rate of temperature bears to the dry-adiabatic rate. If a speed correction is obtained experimentally on the wet element with a dry wick, it will yield a value of ϵ_w usually different from ϵ_d. Then

$$\Delta T_w = \gamma \epsilon_w 1.8 \left(\frac{u}{100}\right)^2 \text{°F}, \tag{61}$$

where γ is the ratio of the moist-adiabatic lapse rate to the dry-adiabatic lapse rate. The equation for the moist-adiabatic lapse rate is[1]

$$-\frac{dT}{dz} = \frac{g}{c_p} \frac{1 + \frac{m_v}{m}\frac{e_s}{p}\frac{Lm}{RT}}{1 + \frac{m_v}{m}\frac{e_s}{p}\frac{L^2 m_v}{Rc_p T^2}}. \tag{62}$$

[1] B. Haurwitz, *Dynamic Meteorology*, McGraw-Hill, New York, 1941, p. 55. Equation (62) has been obtained from Haurwitz' equation by use of the Clausius-Clapeyron equation and the ideal gas law. Haurwitz' R is R/m in our notation.

The dry-adiabatic lapse rate, from Eq. (13), is

$$-\frac{dT}{dz} = \frac{g}{c_p};$$

hence the ratio becomes

$$\gamma = \frac{1 + \frac{m_v}{m}\frac{e_s}{p}\frac{Lm}{RT}}{1 + \frac{m_v}{m}\frac{e_s}{p}\frac{L^2 m_v}{Rc_p T^2}}. \tag{63}$$

Table 3 4 gives values of γ for several temperatures at $p = 1000$ mb.

TABLE 3·4.—CORRECTION FACTOR γ AS A FUNCTION OF TEMPERATURE

T, °F	γ
30	0.67
50	0.54
70	0.43
90	0.36

The factors ϵ_d and ϵ_w are determined experimentally for each installation. The wet-bulb resistor is flown with a dry wick. A flight is made in a layer of the atmosphere free from strong convective currents or large temperature inversions, preferably on a day with complete cloud cover (except cumuliform). The pilot should select an easily navigable course for about a 2-min run. He then flies alternately back and forth over the course at several speeds between minimum and maximum aircraft speeds, keeping aspect constant. About six round trips are usually sufficient. The resulting curve of indicated temperature vs. $(u/100)^2$ should be a straight line, the slope of which is a measure of ϵ.

Because air temperature rarely remains constant, it is best to use only two successive temperature-speed determinations for a single value of ϵ; then all six values of ϵ should be averaged. After the correct values of ϵ are obtained for a thermometer, an accurate record of indicated air speed during a sounding must be kept in order to make the proper temperature correction.

Response Time.—Time of response of a thermometer is usually stated in terms of its lag coefficient,[1] λ, which is the time required for that thermometer to indicate $(1 - 1/e)\ \Delta T$, or $0.63\ \Delta T$, when it is subjected to a sudden change in temperature ΔT. Determination of λ for a sounding instrument is useful in determining optimum ascent rates.

One way of determining λ for an aeropsychrograph thermistor in air is to insert the thermistor into an air stream whose speed is known, heat the thermistor by an alternating current (to, let us say, 10° higher than the air stream), switch the thermistor into the circuit while at the same instant

[1] See W. E. K. Middleton, *Meteorological Instruments*, University of Toronto Press, Toronto, Canada, pp. 56–58.

turning off the heating current, and record the temperature decrease as a function of time.

A typical lag coefficient-ventilation curve obtained in this manner is shown in Fig. 3·60. Two curves are presented, the lower one for a thermistor as it arrived from the manufacturer, the upper one for the same thermistor with three applications of polybutylmethacrylate. It can be seen that both curves rise sharply as the speed falls below 10 mph. This indicates that caution is necessary when using unaerated thermometers of this type.

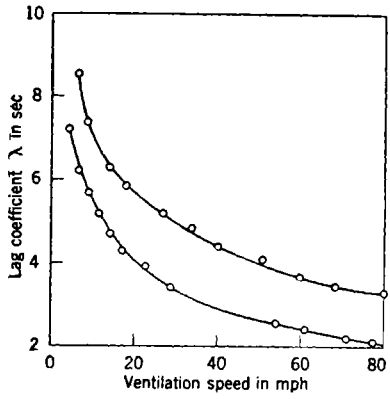


FIG. 3·60.—Lag coefficient as a function of air speed with the thermistor perpendicular to the air stream. The thermistor has a length of 1.9 in. and a diameter of 0.06 in.

The curves were obtained for a thermistor whose long axis was perpendicular to the air stream. The angle between the thermistor axis and the direction of the air stream affects the response time materially; the slowest response occurs when the thermistor is parallel to the air stream. Figure 3·61 shows the relation between lag coefficient and the angle between the axis of the thermistor and the direction of the air stream for an air speed of 11 mph. This relation was obtained for a thermistor with two coats of plastic and is somewhat larger than that used in obtaining the data in Fig. 3·60. The curve clearly illustrates the superiority of perpendicular over parallel mountings.

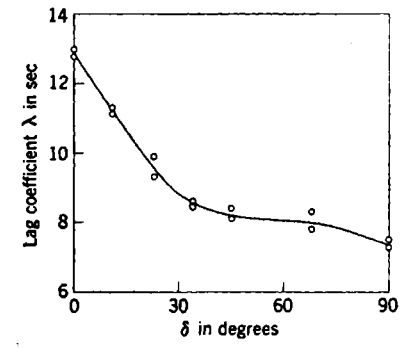


FIG. 3·61.—Lag coefficient as a function of δ, the angle between the axis of the thermistor and the direction of the air stream. Ventilation speed is 11 mph. The thermistor has a length of 2¼ in. and a diameter of 0.065 in. and has two dip-coats of polybutylmethacrylate.

Finding the lag coefficient of a wet-bulb thermometer is difficult. As one factor in the response time of a wet bulb is the rate of evaporation of water from the wick, a rate that varies according to the moisture content of the air, any wet-bulb lag determinations require additional information on speed and humidity. Such determinations should not be made by the simple heating method described above because the entire mass of thermistor, wick, and water cannot be heated equally by supplying a current to the thermistor alone.

No lag coefficients are available for the wet thermistors used in the aeropsychrograph.

Sounding Methods.—The method of making an aircraft sounding usually followed by the Radiation Laboratory consisted of descending to as low an altitude as possible, 20 ft or less if over water, and then ascending in a tight spiral at a rate of approximately 100 ft/min. The instrument automatically switches from dry- to wet-bulb temperature at intervals of about 7 sec, and at selected altitudes the operator records the altimeter indication.

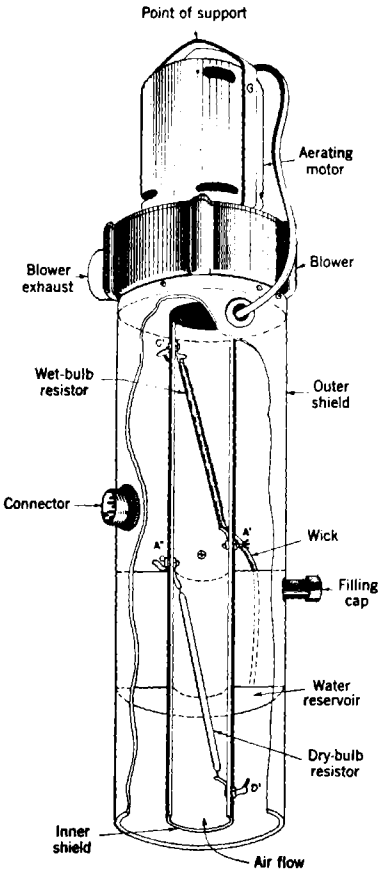


FIG. 3·62.—Modification of the psychrograph housing for ground-base use.

The sounding should be made to a sufficiently high level to ensure measurement of pertinent temperature and humidity distributions. This height is dependent on the geometry and wavelength involved in the particular propagation problem and on the meteorology of the region. A level of 1000 to 1500 ft is often high enough, but, as will be seen in the following chapter, the sounding must sometimes be extended to several thousand feet. Repeat soundings should be made through the layers showing marked deviation from standard conditions. Because wind velocities play such an important role in tracing trajectories and in helping to understand the physical basis of the modification of air, it is usually desirable to estimate surface winds and to measure winds at some higher level, using drift methods or their equivalent.

Other flight techniques have been employed by various groups. Sometimes the aircraft is flown back and forth over a fixed course while the altitude is varied in steps, or the aircraft may fly a straight course, descending at an essentially constant rate. Blimps, which have also been used successfully, require special techniques that depend upon their operating characteristics. In general, the sounding procedure must be worked out in terms of so many different factors, involving the nature of the instruments

employed, characteristics of the aircraft, the prevailing weather situation, and other quantities, that no general rules can be given for the "best" methods of making a sounding.

Figure 3·62 shows a modification of the psychrograph housing for ground-based soundings. The blower provides more than adequate ventilation for the wet-bulb thermistor and assures that the thermistors come to equilibrium quickly with the surrounding air. The unit shown is carried aloft by captive balloons, and specially designed lightweight, moisture-proof cables connect the housing to the remainder of the equipment. In the present instance five wires are required in the cable: two to carry the 60-cps 110-volt power for the blower and three for the electrical connections to the thermistors.

Balloons used for lifting the housing and cable are the 400- or 700-g Darex[1] radiosonde types. In light winds heights of 400 ft or more may be reached with these balloons and the particular unit shown. In winds exceeding 15 mph barrage balloons or kites have been employed. At intermediate wind speeds two 400-g balloons provide satisfactory lift. The principal limitation is, of course, the weight of housing and cable, which must be kept at the lowest possible value.

The psychrograph can be used to measure wet- and dry-bulb temperature to $\pm 0.2°F$ as long as the wet-bulb temperature is above freezing.[2] (It is difficult to use as an ice-bulb thermometer below freezing without elaborate means for proper wetting of the wick.) The lag is sufficiently low to permit satisfactory operation at low rates of climb in aircraft and at fairly low ventilation speeds. The instrument can readily be adapted to either airborne or ground-based sounding techniques. It is capable of high accuracy in the hands of skilled operators, but its complexity and weight render it unsuitable for general field use in the form shown here.

3·28. Wired Sonde.—The wired sonde[3] is a modification of the radiosonde in which the radio transmission circuit is replaced by metallic conductors carrying only direct currents. It employs the temperature and humidity elements of the radiosonde enclosed in a protective radiation

[1] Dewey and Almy Chemical Co., Cambridge, Mass.

[2] The accuracy of measurement (aside from the aerodynamic heating) can be increased by expanding the temperature scales and improving the stability of the associated electronic equipment, as the thermistor characteristics appear to be very stable.

[3] The wired sonde was developed in 1943 at Washington State College by P. A. Anderson, C. L. Barker, K. E. Fitzsimmons, and S. T. Stephenson, under NDRC Project No. PDRC-647, Contract OEMsr-728. It has been widely used by Washington State College and by several Army and Navy groups in various parts of the world. See "The Captive Radiosonde and Wired Sonde Techniques for Detailed Low-level Meteorological Sounding," Washington State College, Report No. 3, Oct. 4, 1943; "Notes on Operational Use of Low-level Meteorological Sounding Equipment," Washington State College Report No. 7, June 15, 1944; "Operating Instructions for the WSC Low-level Atmospheric Sounding Equipment," Washington State College Report No. 8, July 10, 1944.

shield. A commutated direct voltage is applied simultaneously to the temperature and humidity elements, the resistance of which is indicated by the current through a microammeter. This instrument is usually operated as a ground-based sounding tool for which appropriate balloons, cables, and reels are necessary, but it has also been used for making aircraft soundings. Unlike the radiosonde or the aeropsychrograph, the wired sonde is, electrically, extremely simple. The temperature element is the thermistor described in the preceding section. The humidity element, an electrolytic hygrometer,[1] measures relative humidity directly (after a correction for the effect of temperature) rather than wet-bulb temperature. It consists of a plastic strip dipped in a solution of lithium chloride in polyvinyl alcohol. Electrical contact is provided by sheet metal strips that are molded over the edges of the plastic strips before the coating is applied. The conductivity of the lithium chloride coating increases with increasing relative humidity.

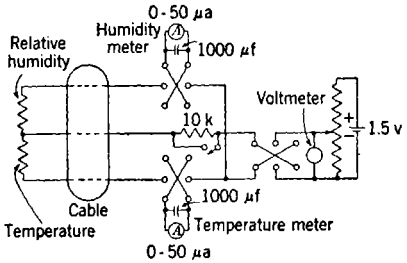


FIG. 3·63.—Circuit diagram of wired sonde.

A circuit diagram of the wired sonde used by Washington State College is shown in Fig. 3·63. A constant voltage is applied across the relative humidity and temperature elements, whose resistances essentially control the flow of current through the microammeters; the meter scales may be calibrated directly in terms of temperature and humidity. Because the electrolytic strip is easily polarized, the direction of the applied voltage is automatically reversed every half second by the motor-driven reversing switches. The 1000-μf condensers shunting the meters provide a sufficiently long time constant to render the meter deflections insensitive to the switching. The 10,000-ohm resistor in series with the elements protects the electrolytic strip and the meters; it is short-circuited during measurement.

A radiation shield used with the wired sonde is shown in Fig. 3·64. It consists of a double-walled aluminum cylinder, suspended vertically, inside of which are mounted the two measuring elements. A cone, mounted several inches above the vertical cylinder, shields the elements from radiation from above. With this arrangement, aeration of the elements must be accomplished by airflow through the radiation shield in a vertical direction. This is adequate down to wind speeds or towing speeds of about 2 mph.[2] For soundings carried out in a dead calm there is some uncertainty as to the degree of equilibrium between the elements and the surrounding air.

[1] F. W. Dunmore, *Nat. Bur. Standards Jour. Res.*, **23**. 701 (1939).

[2] Private communication from Dr. P. A. Anderson, Washington State College.

Rapid ascent or descent from one sounding altitude to the next or jerking the cable is sometimes employed to reduce the uncertainty, although perhaps the most satisfactory solution has been to provide forced ventilation is a manner equivalent to that shown in Fig. 3·62.[1]

The temperature and humidity elements are connected to the meter by a light, waterproof cable, which is taken up on a reel. Development of a satisfactory cable for this instrument (as for all ground-based instruments) has been a major problem too involved to describe here.

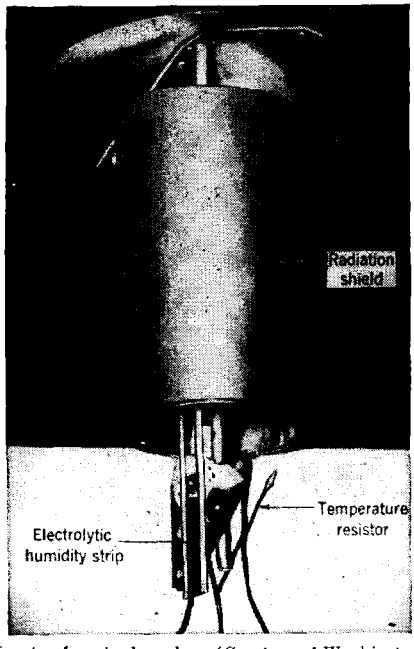


FIG. 3·64.—Housing for wired sonde. (*Courtesy of Washington State College.*)

Figure 3·65 shows the latest version of the wired sonde ground unit, including the reel and meter box.

Balloons to lift the housing and cable are similar to the 350-g radiosonde type but are of neoprene and equipped with a reinforced neck to support the extra tension needed for tethered balloon work. Under ideal conditions altitudes as high as 4000 ft have been reached with the unaerated wired sonde, using three balloons. Each balloon has a free lift of 3 lb when inflated with helium to a diameter of about 5 ft.

[1] L. J. Anderson, S. T. Stephenson, and A. P. D. Stokes, "Improvements in U.S.N.R.S.L. Meteorological Equipment," NRSL Report No. WP-21, San Diego, July 1945, and L. J. Anderson, "Captive-Balloon Equipment for Low-Level Meteorological Soundings," *Bull. Am. Meteor. Soc.*, **28**, 356 (1947).

The wired sonde has the advantage of being very light (the unit shown in Fig. 3·64 weighs 6.5 oz, but provision of forced ventilation increases the weight to several times this value), and it is free from complicated electrical circuits. These features recommend it for field use by personnel with a minimum of training other than meteorological. The principal limitation of the accuracy of the instrument occurs in the electrolytic hygrometer,

FIG. 3·65.—Ground unit of wired sonde showing level-wind winch, cable reel, and shock-mounted meter box. (*Courtesy of Washington State College.*)

which has neither the stability nor the general reliability of the thermistor. Its useful life ranges from 24 hr in the tropics to several weeks in dry, temperate climates.[1] In cases where high accuracy is not necessary it is found that the individual temperature-relative humidity curves of individual elements in a given production lot are enough alike in shape that it is sufficient to adjust an external resistor-compensating network to indicate the correct relative humidity at one point. The technique of manufacture of electrolytic hygrometers is still being modified, and further improvements of the characteristics of the hygrometers for low-level sounding purposes are to be expected.

[1] Private communication from Dr. P. A. Anderson, Washington State College.

3·29. Aircraft Psychrometers.—Another instrument that can be used for low-level soundings is the aircraft wet- and dry-bulb psychrometer, as exemplified by the ML-313.[1] It consists of a pair of matched thermometers, mounted in a streamlined ventilator, for use in measuring wet- and dry-bulb temperatures outside an airplane. Right-angle mercury-in-glass thermometers are mounted on a portable support as shown in Fig. 3·66. The thermometers are made with the 90° bend in order that they may be read from inside the cabin, while the bulbs are inserted into the ventilator,

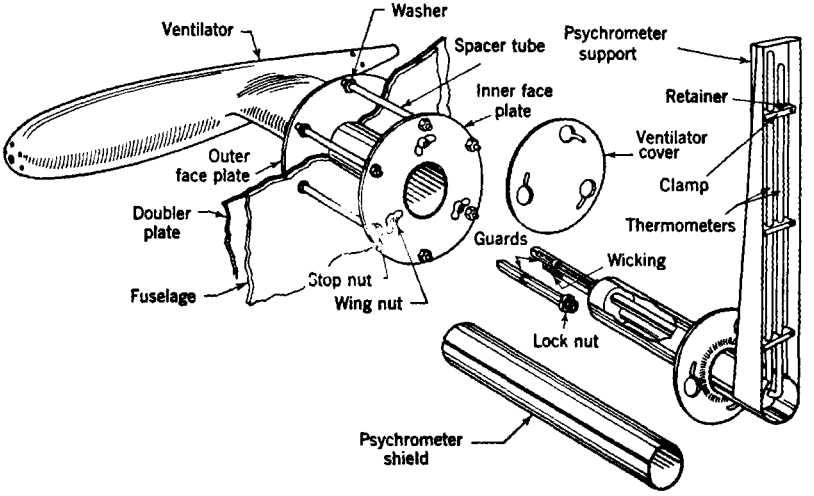


FIG. 3·66.—ML-313 thermometers and housing. (*Courtesy of U.S. Army Signal Corps.*)

which has a current of outside air flowing through it. The cylindrical shield is provided to prevent breaking the thermometers when not in use.

The ventilator is a polished, streamlined housing, which reduces the flow of air and supports the thermometers. It is designed to be clamped to the fuselage, through which a hole has been cut; it should be placed in a position from which observations can easily be made. A place as far forward in the airplane as possible should be selected to avoid heating by the motors or by the disturbed air behind the propellers. Dynamic corrections, which are included in the operating instructions, have been made for the ML-313.

Because of the large lag coefficient of the ML-313, of the order of 45 sec at normal aircraft speeds, it is necessary to fly at a constant level for 2 min to obtain a reading. The ascent or descent therefore must be made in steps. The wick is wetted before each reading; the aircraft flies at constant altitude; and the thermometers are watched for 2 min; if there are

[1] Developed by Army Service Forces, Signal Corps Ground Signal Agency, Eatontown Signal Laboratory, Fort Monmouth, New Jersey.

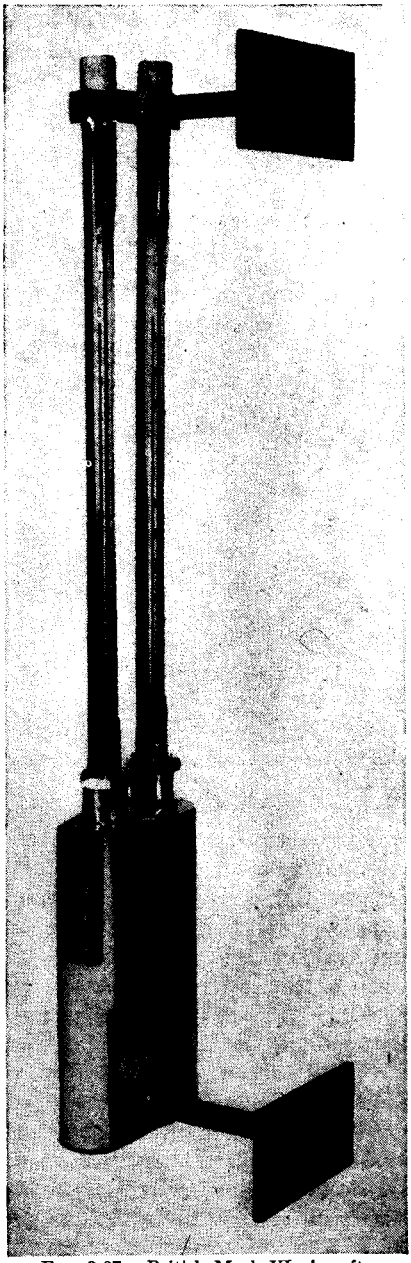


FIG. 3·67.—British Mark VI aircraft psychrometer. (*Courtesy of British Meteorological Office.*)

no horizontal gradients present, the mercury columns will have become nearly stationary in this time. The altimeter and temperatures are then read simultaneously, and the airplane thereafter ascends (or descends) to its next altitude, the same procedure being repeated.

The major advantage of the aircraft psychrometer, as of the wired sonde, is its simplicity. The thermometers are calibrated and the dynamic correction is made prior to their being put into use. There are no electronic circuits; there is no possibility of electrical leakage; and little or no servicing is necessary, except for changing the wick. A second advantage is that observations can be made even though the wet-bulb temperature is below freezing, as the wick is wetted manually between readings. An accuracy of about $\pm 0.2°F$ may be expected with this instrument.

The disadvantages (for this purpose) of the ML-313 are its high lag coefficient and the fact that it is nonrecording. Also, the reduction of data for the wet-bulb readings involves considerable labor.

The aircraft psychrometer has not been used to any great extent as a sounding instrument for low-level work but is included in this discussion because of its potentialities. For the most part, it has been used as a standard instrument for checking other aircraft thermometers.

An instrument similar in principle to the ML-313 is the Mark VI

aircraft thermometer designed by the British Meteorological Office and used in England. The entire unit is mounted outside the observer's window, through which direct readings of dry- and wet-bulb temperatures are made. Figure 3·67 shows the thermometer mounting.

3·30. Resistance Thermometer and Humidiometer.—The Meteorological Office of the British Air Ministry operated for some time an experimental station at Rye, in southwest England, where high towers were available on very flat terrain.[1] Measurements of atmospheric structure as a function of height and time were made by operating recording instruments distributed along the towers. Temperature and relative humidity were recorded continuously at heights of 4, 50, 155, and 350 ft, and other supplementary measurements were made at various times.

Temperature was measured by nickel resistance thermometers, each of which formed one arm of a Wheatstone bridge operating off balance. A recording microammeter indicated the degree of unbalance of the bridge and, consequently, the temperature. Instead of measuring the absolute value of temperature at each height directly, it was found more convenient to measure temperature at one reference height and the difference between temperature at reference height and the other heights. The temperature differences were obtained by operating the thermometers at two different levels as adjacent bridge arms, thus making a differential bridge. Each thermometer element was encased in a projecting cylindrical cover that was housed, along with a humidity instrument, in a conventional meteorological instrument shelter. Because of the high thermal capacity and lack of forced ventilation, the thermometer lag coefficient with low wind speeds was several minutes; consequently, only relatively long-period temperature changes could be measured.

Relative humidity was measured by an instrument known as the *Gregory humidiometer*, which is similar in principle to the electrolytic hygrometer of the wired sonde. Instead of a thin film of hygroscopic material deposited on a strip, the humidiometer employs a piece of cloth impregnated with a hygroscopic salt and stretched tightly around a cylindrical framework of metal rods that are wired alternately in parallel. The electrical resistance measured between the two sets of rods is a direct measure of relative humidity (after application of a temperature correction). The resistance of a sample experimental element at 15°C was 113,500 ohms at 30 per cent relative humidity and 68 ohms at 95 per cent, and the most rapid change of resistance occurred at low humidities. In this installation, an alternating voltage was applied to the element (to avoid polarization effects), and the resulting current was rectified and applied to a recording microammeter similar to those employed for temperature measurement. The lag of the humidiometer is considerably

[1] "Meteorological Investigations at Rye, Part 1: Instrumental Layout for Recording Gradients of Temperature and Relative Humidity," Instruments Branch, Meteorology Office, Air Ministry, May, 1944.

smaller than that of the thermometers at high temperatures, but it varies with both temperature and humidity in a manner not yet ascertained.

A considerable quantity of data has been obtained and analyzed at the Rye station, although it is, at the present time, neither sufficiently complete nor available in sufficient detail to allow discussion of the results.[1]

An advantage of the tower method of measurements is the fact that continuous recording, sometimes unattended, is possible. Also, measurements may be made simultaneously at the various levels. An inherent difficulty in the tower method is the limitation imposed by the height of the tower; frequently changes important to propagation occur above the top of even the highest tower and cannot, therefore, be measured in this way. Also, when data are obtained at fixed levels, there is the possibility that important information between those heights is omitted.

3·31. Thermocouples.—Because of their extensive use in scientific work for accurate temperature measurements, thermocouples would appear to be a logical choice for low-level soundings, but relatively few research groups used them during the war. This is probably attributable in part to the fact that strong emphasis was laid on aircraft soundings, which placed mechanical limitations on the equipment, thus barring such sensitive instruments as the necessary galvanometers. For ground-based soundings, however, this limitation is absent, and it is reasonable to suppose that in the future, thermocouples may be employed to a considerably greater extent than in the early developments reported here.

Two groups in England have reported using thermocouples as wet- and dry-bulb thermometers. One type was designed specifically for use with a tethered balloon, and an early model operated at heights up to about 200 ft,[2] but no report is available on the use of a modified version.[3]

A large amount of meteorological data for propagation purposes has been gathered over the Irish Sea by specially designed thermocouples arranged to measure air temperature and wet-bulb depression.[4] These thermocouples were made of copper and constantan, giving an emf of about 22 μv/°F, and had a very small mass, hence a relatively low lag

[1] During the war, the Meteorological Office issued a number of reports, including tables of values of temperature and humidity for limited periods, and several analyses of particularly interesting situations, such as marked radiation inversions. These reports had limited circulation under numbers issued by the Joint Meteorological Radio Propagation Sub-committee. Examples are JMRP Reports, Nos. 5, 6, 7, 18, 20, 23, and 31.

[2] S. M. Doble, "Balloon Psychrometer for the Measurement of Relative Humidity of the Atmosphere at Various Heights," Imperial Chemical Industries, Ltd., Apr. 1, 1943.

[3] *Addendum* by S. M. Doble and S. Inglefield, Sept. 25, 1943.

[4] These instruments were designed and built by the British Naval Meteorological Service.

coefficient of about 1 min at a ventilation speed of 10 knots (normally provided by motion of the ship). The cold junction was placed in a Dewar flask containing a mixture of melting paraldehyde, the temperature of which is constant at 50°F. Air temperature was indicated by a galvanometer driven by the difference in thermal emf between the cold junction and the junction in free air, and wet-bulb depression was obtained by the difference in emf between the latter junction and a similar one covered by a wick and fed from a small reservoir by capillary action. Two junctions were placed in series to increase sensitivity for both dry- and wet-bulb measurements. The external junctions were enclosed in appropriate radiation shields mounted at several heights on a ship, covering a height range from about 4 to 50 ft.

A sounding was made by reading temperature and wet-bulb depression at one level, then switching to the next level and repeating the procedure. Sea temperature was also measured by another thermocouple which was switched in when desired. Readings up to 200 ft were obtained with a supplementary thermocouple psychrometer raised by a barrage balloon from the stern of the ship.

This method of obtaining low-level soundings is simpler than most other methods. The instrument measures basic properties of the air; untrained observers may be employed; no source of power is required; no complex electronic circuits are used; and the sounding may be made quite rapidly or as slowly as desired.

This psychrometer, however, in the form described is nonrecording, and measurements below wet-bulb freezing temperatures are impossible. The method suffers also from those disadvantages usually associated with making measurements on board ship. It is difficult to find instrument exposures that are not subject to radiation or convection from the ship, and the roll and pitch of the ship make heights above the water uncertain.

3·32. General Problems Associated with Low-level Soundings.—In the preceding five sections there were discussed five of the methods most widely used in attempts to solve the complex problem of obtaining accurate and meaningful soundings in the lower atmosphere. Each instrument and method of use has certain advantages but is also limited in some manner. With an aircraft psychrometer, for instance, soundings can be made quickly over a large area, whereas with balloon-borne equipment soundings are usually limited to a very small area and each sounding requires considerable time to cover an appreciable height range. On the other hand, instruments carried by tethered balloons or towers may be used under conditions of visibility, weather, and terrain in which airplane flights are impossible, and they experience no complications from high-speed airflow over the thermometer. Some instruments are subject to difficulty from induced static charges in an air stream, however. As the science of

low-level soundings stands at present, no one sounding instrument fulfills all the requirements of propagation research.

It is apparent that further research is necessary on meteorological instruments suitable for use in propagation work. On the basis of experience acquired thus far, we list tentatively some of the features that a sounding instrument of the future should possess. It should

1. Have a lag coefficient less than 5 sec.
2. Record all data.
3. Be usable in the field by untrained observers.
4. Have an accuracy in temperature of ± 0.1°F and in humidity of ± 0.1 mb of vapor pressure.
5. Be readily used either as an aircraft or a balloon-borne instrument.
6. Measure a range of temperature from $-60°$ to 120°F and a range of vapor pressures from near 0 to 55 mb.

It is possible that none of the methods described above permits refinement to the point where these specifications can be met, and new methods of approach should be considered. A newly developed dew-point thermometer[1] has recently come into use which may help solve some of the difficulties attached to measuring humidity, especially below freezing.

Some success has been achieved to date with temperature-speed corrections on aircraft using the methods described in Sec. 3 27. This should be regarded more as a first step toward a solution, however. Further thought and experiment are needed on speed corrections before the problem can be considered solved.

Fully as important as the instrument are the methods of using it and of analyzing the data obtained with it in the light of the characteristics of the instrument and of local meteorological conditions and their relation to the synoptic situation. It should be evident from this and the following chapters that any attempts to conduct meaningful meteorological measurements and analysis for propagation purposes without participation of highly skilled meteorologists are indeed pointless.

METEOROLOGICAL CONSTANTS

By R. B. Montgomery

3·33 Useful Meteorological Constants:

Absolute temperature of the ice point: 273.16°K.
1 atmosphere = 1013.25 mb.
Density of dry air at 0°C and 1 atmosphere: 1.293×10^{-3} g cm^{-3}.
Universal gas constant: $R = 8.314 \times 10^{7}$ erg °C^{-1} mole^{-1}.
Mean molecular weight of dry air: $m_d = 28.982$.

[1] D. N. Brissman, *A Method for the Continuous Measurement of Dew Point Temperatures*, Univ. of Chicago, April 1945.

Molecular weight of water: $m_v = 18.016$.

$m_v/m_d = 0.622$.

Isopiestic specific heat of dry air: $c_{pd} = 1.004 \times 10^7$ erg g^{-1} $°C^{-1}$.

Latent heat of vaporization of water: at 0°C, $L_f = 2500 \times 10^7$ erg g^{-1}; at any other temperature T, in °C, $L = L_f[1 - 0.00094\ T]$.

Kinematic viscosity of air at 0°C and 1000 mb: $\nu = 0.1346$ cm^2 sec^{-1}.

Thermometric conductivity of air at 0°C and 1000 mb: $\nu_c = 0.189$ cm^2 sec^{-1}.

Diffusivity of water vapor in air at 0°C and 1000 mb: $\nu_d = 0.226$ cm^2 sec^{-1}

CHAPTER 4

EXPERIMENTAL STUDIES OF REFRACTION

By Pearl J. Rubenstein, Donald E. Kerr, and William T. Fishback

This chapter summarizes the principal experiments that have been performed to investigate the effects of atmospheric refraction on microwave transmission. Many of these experiments (particularly the earlier ones) were highly varied in nature and were not performed or analyzed by methods now in general use. They often revealed transmission properties that now appear to fit into a broad general pattern, although at the time this unity was not apparent. Consequently we do not follow a chronological approach, but instead we have arranged the material in a way that seems most appropriate in the light of present information. The point of view adopted here is that the modified-index distribution is the property of the transmission medium that is of major interest and usefulness.

The marked dependence of the transmission phenomena on meteorological conditions makes it imperative that radio and meteorological observations be made simultaneously and that the weather regime of the experimental location be considered. Because the radio results must be interpreted in terms of meteorological conditions, statistical studies of even large amounts of radio data have little general significance and will be treated only briefly in this chapter. Our main concern will be the qualitative aspects of transmission as influenced by tropospheric refraction.

ONE-WAY TRANSMISSION OVER WATER

Because of the relative simplicity of the surface boundary, the problem of propagation over water has received the major emphasis during the war years. After a period of preliminary trials an experimental procedure was evolved which was then followed in most essentials by all the groups concerned.

As an aid to understanding the plan of operation, let us consider the requirements for an ideal propagation experiment. These may be enumerated as follows:

1. The circuit or path should have simple geometrical properties, so that it may be handled theoretically. In this case, for example, it should be completely over water, with no long stretches of land before the terminals or large islands obstructing the line of sight. In so far as possible, a location should be chosen where the meteoro-

logical conditions are uniform at least a large part of the time; a unique interpretation of the results is otherwise difficult. To facilitate the comparison of experiment and theory, the location should be sufficiently beyond the horizon that under standard conditions the field may be adequately represented by the first mode[1] alone. Calculations would otherwise not be feasible with super-standard conditions.

2. A continuous record of reception is desirable, on which *absolute* field-strength measurements can be made over the entire observed range. A wide range is generally best obtained by means of receivers with logarithmic response.
3. At least three frequencies should be transmitted simultaneously to provide a clear check with theory.
4. Several terminal heights and if possible several path lengths should be used simultaneously to obtain height-gain relationships and range attenuation. Both horizontal and vertical field-strength sections by airborne receivers are highly desirable.
5. Meteorological data must be collected along with the radio data. The type of data required varies with the general climatological features of the region, but detailed soundings over some height interval are generally necessary. Because no way has so far been found of making continuous records of the necessary meteorological quantities, one must be able to interpolate in space and time between scattered measurements. Generally the fulfillment of this requirement necessitates a rather extensive network of meteorological measurements and, especially, expert analysis.

Transmission over Massachusetts Bay

By Pearl J. Rubenstein

A reasonably satisfactory approximation of these ideals was obtained in transmission tests made in Massachusetts Bay at intervals over a 3-year period.[2] The tests were intended mainly as a study of the refraction effects that result when air flowing from a large land mass is modified by a relatively short trajectory over water. The longest transmission path, from Race Point on Cape Cod to Eastern Point in Gloucester, Mass. (see Fig. 4·1), was 41 miles long and nearly parallel to the coast line, about 15 miles from shore.

[1] See Sec. 2·7 "The Fundamental Theorem." The first mode is defined as that with the smallest attenuation.

[2] The detailed descriptions to follow apply to the work done in 1944. For a description of the work of 1942 and 1943 see P. J. Rubenstein, I. Katz, L. J. Neelands, and R. M. Mitchell, "Microwave Transmission over Water and Land under Various Meteorological Conditions," RL Report No. 547, July 13, 1944.

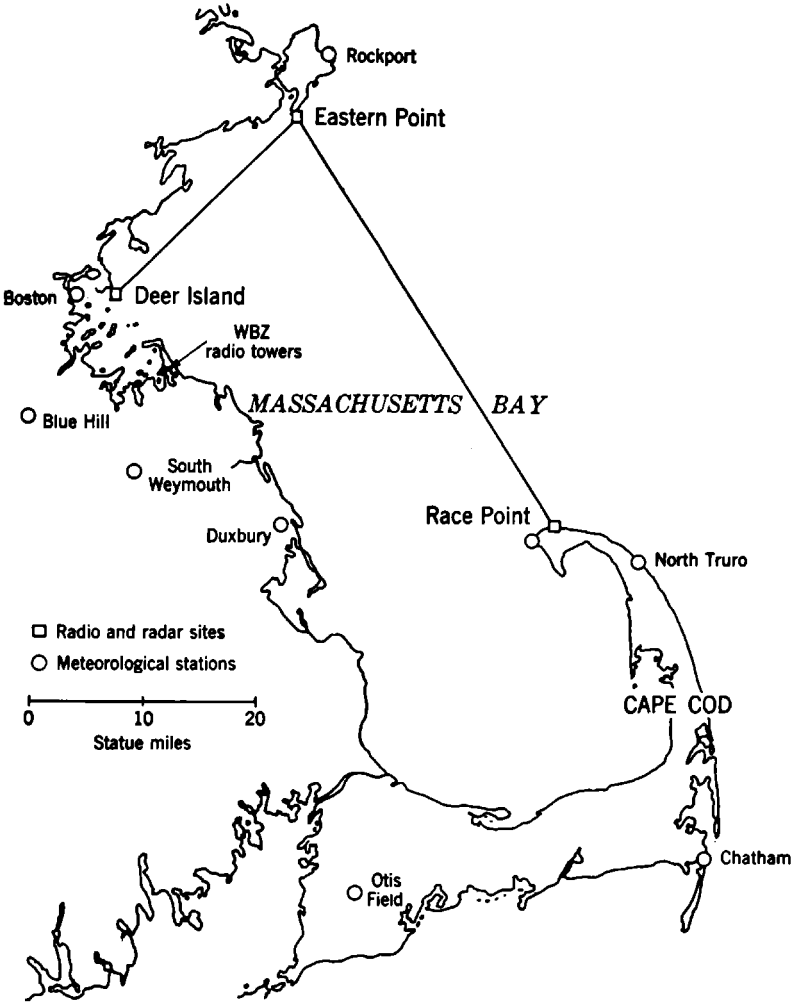


FIG. 4·1.—Massachusetts Bay region. The squares represent radio and radar sites, and the circles represent meteorological stations.

4·1. Radio Measurements Program.—Transmitting antennas were placed 100 ft above mean sea level on a tower close to the water's edge; on a tower at the receiving station, well below the horizon, duplicate installations were made at heights of 28 and 136 ft. Three wavelengths—3, 9, and 256 cm—were used simultaneously. Only the high receiving location was operated at 256 cm. A 1.25-cm circuit was also operated, with a receiver at the 28-ft level only, but for too short a time to provide conclusive results.

At 2.5 m ample signal was obtained with a c-w transmitter and Yagi antenna arrays. At the shorter wavelengths, however, pulsed magnetrons and high-gain paraboloidal antennas were necessary. The powers and antenna gains were so chosen that under standard refractive conditions the received signal would with one exception be at or above the detection level of the receiver on each of the circuits. At 1.25 cm the minimum detectable signal was 20 db above standard level so that only super-standard conditions could be studied.

Carefully monitored radar sets operating at 3 and 9 cm were set up beside the radio transmitting station. Hourly records were kept of signal strengths on selected coastal targets and of the maximum ranges obtained, and PPI photographs were taken frequently. The radar program and its results will be discussed in Sec. 4·18.

An additional one-way circuit at 10 cm was operated from Deer Island in Boston Harbor to Eastern Point. The terminals, 125 and 30 ft above mean sea level, were 22 miles apart, making the path nearly optical. Calculations could therefore be made for this path only under standard conditions; it was nevertheless useful as a check on the longer path and in the study of local meteorological effects.

Horizontal polarizarion was used for the most part, but no difference was observed when a change to vertical polarization was made.

4·2. Meteorological Measurements and Analysis.—It has been indicated previously that meteorological information adequate for microwave propagation research is not available through the normal weather channels. Weather stations are ordinarily spaced too sparsely and lack the equipment needed for detailed soundings of the lower atmosphere. Their data can therefore be used only to supplement special measurements.

For optimum results an especially planned program of accurate low-level soundings is necessary, in addition to closely spaced surface data. As it is impossible to make measurements continuously both in space and in time, a large percentage of time inevitably remains when the vertical structure of the atmosphere in the desired locations is not precisely known. Although much progress has been made toward the understanding of the meteorological processes concerned, the departures of actual conditions from the idealized postulates of the theory prevent the setting up of any fixed methods of interpolation and extrapolation of the data. Only a combination of careful measurements and detailed analysis by competent meteorologists can provide dependable results in any specific case at present.

In the meteorological program carried out in the Massachusetts Bay region along with the radio transmission tests described in Sec. 4·1, low-level soundings measuring temperature and humidity were made in aircraft and with tethered balloons. Two aircraft, a Curtiss-Wright (small single-engine monoplane) and an AT-11 (Army twin-engine trainer), equipped with aeropsychrographs (see Sec. 3·27) sounded in tight spirals usually from a low point of about 20 ft to about 1000 ft above the water

surface. These soundings were made only in the daytime, usually from 0900 to 1100 and 1300 to 1500 hours, Eastern Standard Time.[1] Balloons or kites bearing modified versions of the aeropsychrograph (see Sec. 3·27) were flown at land stations at Duxbury and Race Point and from a 50-ft motor sailer. These soundings were carried to heights of about 300 ft, depending upon winds and visibility, and were usually made at 4-hr intervals at the land stations. The boat was also equipped to make detailed soundings in the lowest 50 ft from a side arm swung from the windward side of the boat; the psychrograph was raised to the top of the mast along the hypotenuse of the triangle formed by the side arm and the mast. Altogether, nearly 1000 soundings were made during the course of the investigation.

Water temperatures were measured from the boat as it plied a course across the Bay. Whenever possible the boat was located so as to give water temperatures near the base of the aircraft soundings.

Hourly surface observations of temperature, humidity, wind speed and direction, and sky conditions were obtained from special stations at Duxbury and Race Point, as well as from the regular weather channels. Most of the stations used are shown on the map of Fig. 4·1. In addition use was made of the observations obtained at Portland, Maine, Bedford and Nantucket, Mass., and Providence, Quonset Point, and Point Judith, R.I. The purpose of the special stations was to sample the air just before it left the land at Duxbury and immediately after it passed over the Bay at Race Point; this sampling was possible only with westerly winds.

Winds aloft were obtained from pilot balloon observations at Bedford, Boston, South Weymouth, Point Judith, Nantucket, and North Truro at 0430, 1030, 1630, and 2230 Eastern Standard Time. For the most part only the winds up to 2000 ft were considered in tracing the air motion; almost invariably the major emphasis was placed on the 1000-ft wind.

Radiosondes released twice daily from Boston and Portland and occasional airplane ascents with aerometeorographs at Quonset provided all the upper-air data available from regular channels. As the radiosonde ascents originated over land, they were used only to determine the structure of the unmodified air.

Synoptic charts for use in the analysis were obtained from the district forecast center at the Weather Bureau Station at the East Boston Airport.

The data were analyzed by meteorologists prior to any correlations with radio data, and the results presented to the radio analysts in the form of "synthetic soundings" for the mid-point of the path at 0430, 1030, 1630, and 2230 Eastern Standard Time, daily These synthetic soundings represented, to the best of the meteorologists' judgment, the vertical distributions of temperature, humidity, and refractive modulus at the

[1] All times in connection with this work will be given in Eastern Standard Time using the 24-hr system.

given time and place. The presence of marked horizontal gradients and the time and nature of any significant change that might have occurred in the interval between synthetic soundings were also noted.

How the meteorological data were coordinated to yield the synthetic distributions will now be outlined briefly. Although the analysis was carried out according to a general uniform pattern, the details of individual cases varied so greatly that no specific rules can be established. Rather, each case must be considered individually to determine the necessary modifications of the general procedure.

Consideration of the Synoptic Situation.—Before the localized weather was studied, the synoptic maps for the day were surveyed to determine the general situation as well as to learn about fronts in the vicinity of the paths, cloud conditions and precipitation, the type of air mass, and significant wind shifts during the day.

Winds.—Winds were next scrutinized in detail to determine the points on land where the air over the circuits had originated and to estimate the amount of mechanical mixing that might be expected. Winds at the surface and at 1000 ft were studied for variations in time and space, and representative 1000-ft winds were estimated for the mid-point of the path at the times of the desired synthetic soundings.

Trajectories.—Trajectories based on the 1000-ft winds were constructed for all soundings, whether measured or synthetic. They were extrapolated from the sounding location back to the land origin, and the distance and time over water estimated as well as the point of land origin. The surface conditions at this point could then be found by interpolation of the hourly weather observations.

Determination of Initial Vertical Distributions and Estimate of Water Temperatures.—Before the modification of the air in its trajectory over water can be investigated, both the initial vertical distributions of temperature and humidity and the water temperature must be known. The initial vertical distributions over land were deduced with the help of the radiosondes at Boston and Portland, the airplane observations at Quonset, and overland soundings at Duxbury. With conditions of surface heating a simple homogeneous distribution exists during the daytime; surface observations then give approximate[1] values of the parameters in the convectively mixed layer. At night, on the other hand, the lack of both adequate measured soundings over land and a thorough understanding of the low-level meteorology of nocturnal radiation prevents any but a very crude estimate of the vertical distribution.

The modification of this initial distribution in its passage over the water depends strongly on the water temperature. Some measurements of water temperature were available from the boat and from the Graves Light and the Boston Lightship in Boston Harbor. Such measurements

[1] Approximate because of possible gradients in the lowest 50 ft.

however, were often unreliable, especially on calm sunny days when the water surface was strongly heated; furthermore the temperatures usually varied with both time and position.

The observations therefore had to be supplemented with values obtained from the characteristic diagram (see Sec. 3·8) by extrapolation of the characteristic curves of the measured soundings to the saturation curve for salt water.

Construction of Synthetic Soundings.—Whenever a measured sounding was found to be representative of conditions near the path mid-point at the required time, it was used as the synthetic sounding for that time. In all other cases the synthetic soundings had to be constructed from the initial vertical distribution over land, the temperature excess, and M-deficit. From these parameters it was immediately clear whether the resulting distribution would be superstandard or substandard near the surface and, in the former case, whether a shallow or deep surface M-inversion would result. The height of the duct or substandard layer could then be determined approximately.

In the case of heating from below (see Sec. 3·14) the amount of heating and moistening that the air would undergo was estimated on the basis of the 10 per cent rule (see Sec. 3·16). The height and strength of the superstandard layer which almost invariably occurred at the top of the convectively stirred layer followed from a knowledge of the initial distribution, after application of the 10 per cent rule.

With cooling from below, the duct height was determined roughly from empirical data.[1] The detailed effects of variable water temperatures, shearing stratification, and initial heterogeneity had to be taken into account on the basis of the analyst's experience in the study of measured soundings.

Qualitative Discussion.—In addition to the synthetic soundings, which described conditions at the path mid-point four times daily, the meteorologists provided discussions of the estimated accuracy of the synthetic soundings, the uniformity of conditions over the circuit at any time, and the time and nature of any significant change that occurred in the interval between synthetic soundings. These qualitative discussions were fully as important in the radio analysis as the synthetic soundings themselves.

Of interest here are some special situations that occurred frequently. With SSW winds, for example, the air near Race Point would have come over the relatively warm water south of Long Island and across Cape Cod Bay. On many occasions this air had picked up so much moisture from the warm water that it produced a substandard layer close to the surface near Race Point. Near Eastern Point, on the other hand, the air would have come directly from the land and might thus give rise to a superstandard surface layer.

[1] A $d/\Delta M$ chart similar to Fig. 3·29, constructed in a preliminary analysis, was used during the main analysis. See R. B. Montgomery and R. H. Burgoyne, "Modified Index Distribution Close to the Ocean Surface," RL Report No. 651, Fig. 8, Feb. 16, 1945.

A second common situation was that of air temperature differing only slightly from the average water temperature. In that case it was possible to have a positive temperature excess over part of the bay and a negative temperature excess over the remainder, with the result that the surface layer might vary from substandard to weak superstandard along the radio circuit.

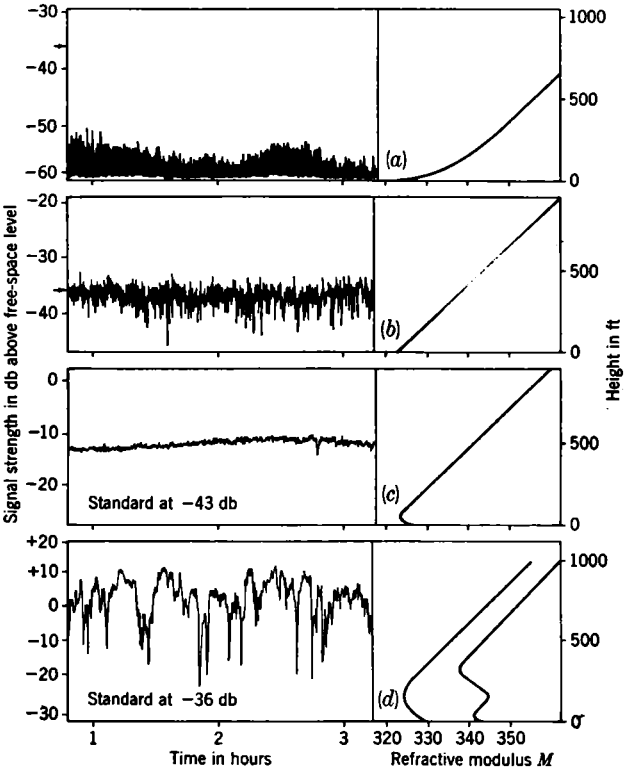


FIG. 4·2.—Signal types and associated M-profiles for microwaves. (*a*) Substandard. (*b*) standard; (*c*) partial trapping; (*d*) strong trapping. The arrows indicate standard signal level.

Finally, with NE winds the air over the entire path might be heated from below after a long trajectory over water; shallow surface ducts might then occur uniformly over the entire region, with only minor variations in duct height and M-deficit caused by changes in water temperatures.

4·3. General Characteristics of Transmission.—The study of the continuous records of radio reception obtained simultaneously at several wavelengths and several terminals can be carried out in many ways. Among these are, at one extreme, the statistical analysis of the entire period as a whole and, at the other, instantaneous comparisons of radio and mete-

orological data. Some of the general characteristics of the transmission found by inspection of the radio records together with the meteorological data and analysis are described here. The results obtained by the other methods will be discussed in later sections.

An interesting feature of the microwave records obtained on both the 41- and 22-mile paths is their natural division into four separate types, as shown in Fig. 4·2. Each of them is associated with one or more distinctive M-profiles, shown in the same figure beside the appropriate radio record sample.

1. When the M-profile is substandard at the surface, the signal is below standard in average level and is characterized by rapid fading known as "scintillation." Occasionally slower variations are superposed on these rapid fades. This is the *substandard* type.
2. With standard or nearly standard[1] M-profiles, the average signal level is standard to within about ± 5 db. Scintillation is almost always present, commonly of amplitude 5 to 10 db but on rare occasions as great as 20 db. This is the *standard* type.
3. With shallow surface M-inversions, or ducts, the signal level is above standard and increases with increasing duct height. It never rises above the free-space value. Only minor short-period variations occur, and the signal is remarkable for its steadiness. This will be called the *partial-trapping* type.
4. With deep surface M-inversions or with overhanging elevated M-inversions the signal is near the free-space value on the average and is unsteady, with "roller" fading of amplitudes as large as 40 db and periods ranging from a few minutes to an hour. The amplitude of the fading increases and its period decreases with increasing intensity of the duct. The fading is characterized by broad flat maxima and sharp deep minima. This will be called the *strong-trapping* type.

These microwave signal types are clear-cut and well separated from each other; transitions between types take place rather rapidly. At 2.5 m, on the other hand, a classification into types can be made only with considerable difficulty, and some question remains as to its validity. A tentative form of these types is shown in Fig. 4·3. Here two different signal types may occur, apparently with a single M-profile. The situation will doubtless be clarified when more and better experimental data are available. In comparing Figs. 4·2 and 4·3, the difference in the ordinate scales should be considered. In many cases what appears to be a highly variable signal record at 2.5 m would seem steady if replotted on the compressed scale of

[1] Within the accuracy of the experiment the standard and homogeneous distributions (see Sec. 3·9) may be used interchangeably. The question of what constitutes an appreciable departure from standard will be discussed in Sec. 4·4. It is a function of wavelength.

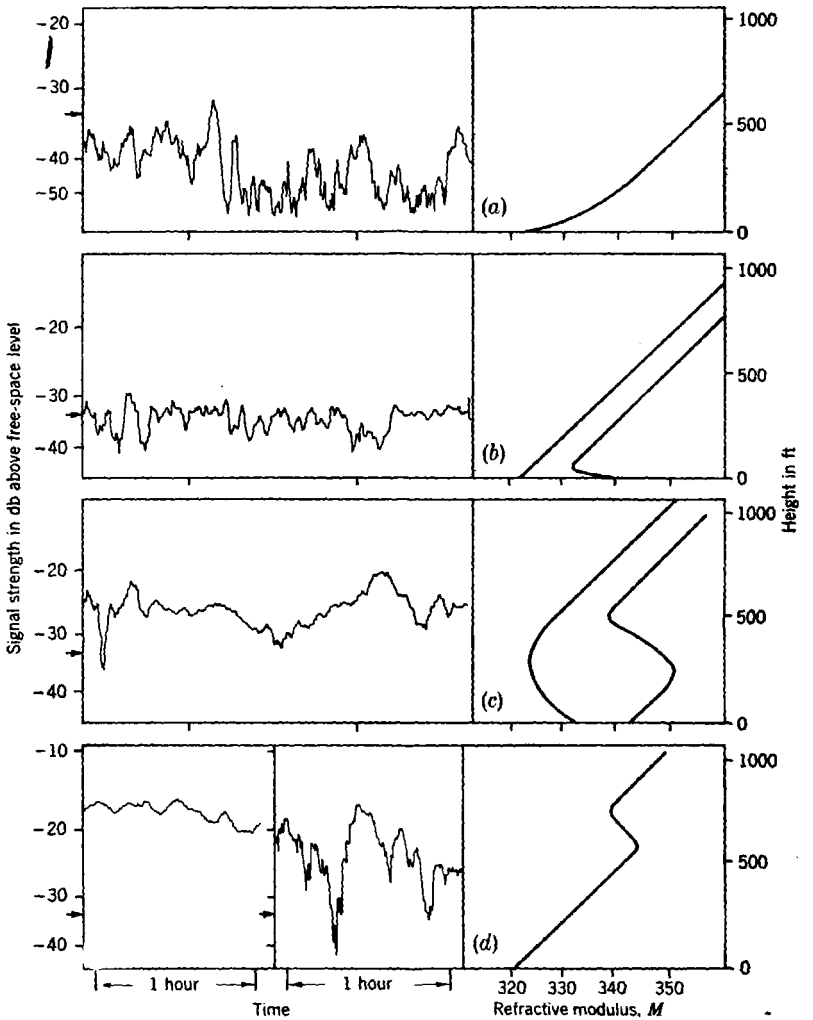


FIG. 4·3.—Signal types and associated M-profiles for $\lambda = 2.5$ m. (*a*) Substandard; (*b*) standard; (*c*) surface ducts; (*d*) elevated ducts. The arrows indicate standard signal level.

the microwave records. This difference in scale serves to emphasize the extreme variability of signal that occurs at microwavelengths.

The following four cases have been considered in some detail:

1. When a substandard layer exists at the surface, the average signal level at 2.5 m is below the standard value; variations in received power occur of about 5 to 10 db with periods of 5 to 15 min. Fairly intense substandard layers are required to affect the transmission.

2. When the M-profile is standard or departs from standard only at the surface and below about 200 ft, the signal is near the standard level; it may be either steady or unsteady, with variations usually 10 db or less.
3. With surface ducts more than 200 ft deep the signal is 5 to 15 db above the standard value and usually fairly steady or with slow roller fading. The same is true with overhanging elevated M-inversions.
4. With elevated superstandard layers 500 to 2000 ft high that are not overhanging, the signal level is above standard, approaching but not reaching the free-space value. (This limitation of the upper value reached is probably caused by the particular meteorological situation.) Steady signal or roller fading occurs about equally often at such times. Fading periods of 10 min to 1 hr have been observed.

Visual inspection of the signal records, besides establishing signal types and their association with distinct modified-index distributions, shows the dependence of transmission results on wavelength and position, the so-called space and frequency diversity. The data obtained in the Massachusetts Bay tests permit comparisons on a given path of wavelengths from 2.5 m to 3 cm, with additional fragmentary information at 1.25 cm. As for space variations, we can study the effect on three circuits at 9 and 10 cm; at 3 cm data are available on two circuits at a fixed range. We can thus obtain some information on height-gain effects at two wavelengths. The comparison of 9- and 10-cm results on the 22- and 41-mile paths does not, however, provide information on range attenuation because of the meteorological differences between the two paths (see Sec. 4·5).

The most striking feature of the transmission is the remarkable similarity, under nearly all conditions, of certain features on all microwave circuits. On most occasions all these circuits show the same signal types. Figure 4 4 illustrates the similarity of reception on various circuits. The average[1] signal level is plotted hourly (solid line) for each of the circuits; the vertical lines indicate the extreme values of the signal during the hour. Not only changes in average level but also those in signal type (indicated by the range of variation) are seen to occur together. The magnitudes of these changes, however, increase with decreasing wavelength and with decreasing terminal height for constant range.

The rather sudden changes from one type to another are practically simultaneous on all paths. In Fig. 4·5 this simultaneity is further illustrated; samples of records obtained at the same time on three different

[1] The averages were obtained by visual estimation from the original recording meter trace; and as the response of the receivers is nearly logarithmic, they are essentially geometrical means of intensity. The visual estimation process was found to give results differing negligibly from those obtained by a more elaborate point-by-point averaging method.

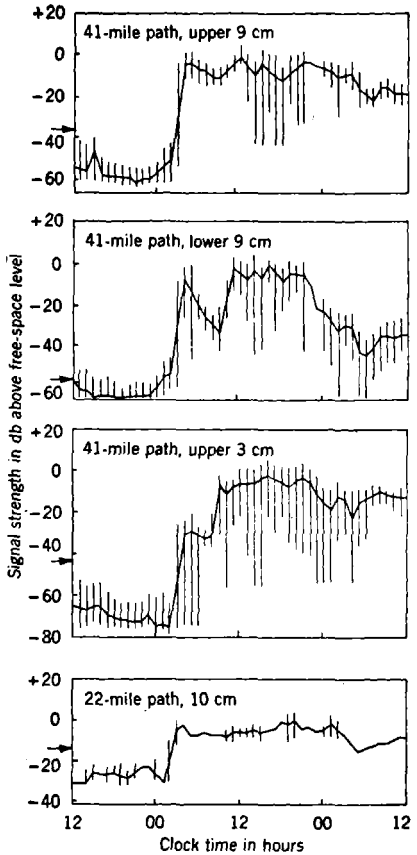


FIG. 4·4.—Average signal strength and extremes, plotted hourly. Arrows indicate standard signal level.

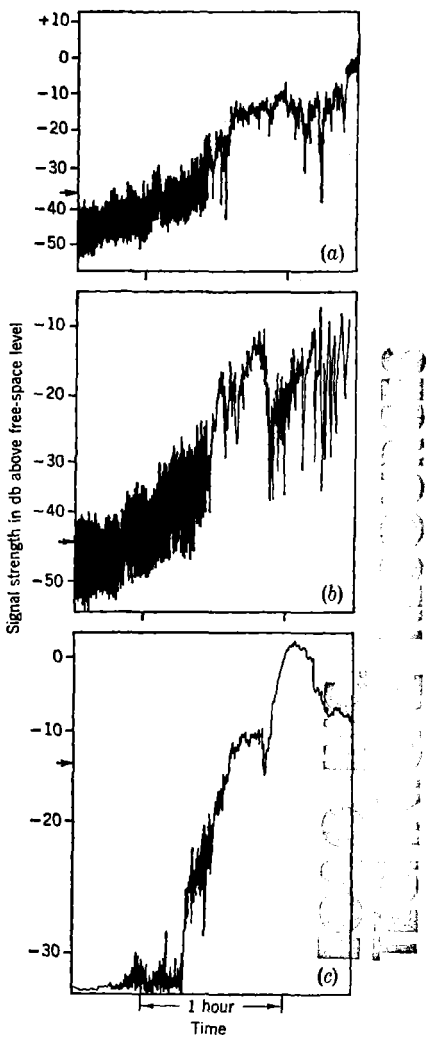


FIG. 4·5.—Comparison of signal records on three microwave circuits during transition from substandard to superstandard conditions. (*a*) Upper 9-cm circuit, 41-mile path; (*b*) upper 3-cm circuit, 41-mile path; (*c*) 10-cm circuit, 22-mile path. The arrows indicate standard signal level.

paths are shown for a period of transition from substandard to superstandard conditions.

In the following discussion it will be convenient to treat separately the different signal types into which the results have been divided, as wavelength and positional dependency vary from one type to another.

With substandard surface layers, scintillating signal occurs on all microwave circuits. The variations are rapid (periods about 30 sec) and irregular, with amplitudes of at least 15 db. The records are too com-

pressed to show details (see, for example, Fig. 4·2*a*), but no apparent relation was noted among the variations on the different circuits. It was not possible to compare fading ranges or periods of scintillation under these conditions.

During periods of substandard surface layers the depressions, from standard, of the average signal levels on the upper 9- and 3-cm circuits are roughly equal. A comparison of the upper and lower 9-cm signals during the same periods shows that the lower the terminal the greater the decrease below standard. The behavior of the lower 3-cm circuit at such times is peculiar (see Sec. 4·5).

Essentially all cases of substandard signal level at 2.5 m correspond to substandard reception on the microwave links, but the reverse is not true. The depth or intensity of the substandard layer required to produce appreciable effects at 2.5 m is clearly greater than that for microwaves, but no quantitative values are as yet available because of the extreme nonuniformity of the meteorological conditions common at such times (see Sec. 4·2). Irregular variability, with periods 5 to 15 min and amplitudes of 5 to 10 db, is characteristic at 2.5 m when the level is substandard. This variability appears to persist at times when the substandard layer is not strong enough to affect the average level appreciably.

As is clear from the meteorological discussion in Chap. 3, a standard M-profile occurs only rarely over water. In transitions from sub- to superstandard conditions, however, there is frequently a short interval of standard conditions. The transition of signal strengths through the standard level is found to occur very nearly simultaneously on all microwave circuits. The 2.5-m signal strengths, on the other hand, generally become standard before the microwaves and usually become superstandard only after a very deep surface duct has built up. Similarly, the signal levels at 9 cm remain standard on some occasions, while the 3-cm levels rise under the influence of very shallow ducts. The infrequency of these occasions reflects the rarity of such shallow ducts in the region.

As the duct height increases from 0 to about 50 ft, the 9-cm levels at first remain standard while the 3-cm levels rise; the scintillation amplitudes decrease. With values of duct height between 20 and 50 ft the signal on both 9- and 3-cm circuits is above the standard level and very steady; the 3-cm level is nearer the free-space value than that at 9 cm. The absolute signal level increases with terminal height, but the lower terminals may experience a greater increase above the standard. A further increase in duct height causes a change in the 3-cm record from the steady partial-trapping type to the roller fading typical of strong trapping of Fig. 4·2*d*, while the average signal level remains near the free-space value. The 9-cm signal continues to be steady and to increase until the duct height reaches 60 or 70 ft.

With deep ducts and the consequent strong trapping, both frequency and space diversity appear to decrease. All microwave circuits have

signals of average level close to the free-space value. At a given wavelength, the period and amplitude of the variations observed at any time on the upper and lower circuits are nearly equal. The fading is far from synchronous, however, as shown in Fig. 4·6. Usually the shorter the wavelength the shorter the fading period; the amplitude of the fading tends to increase with decreasing wavelength. The disappearance of a height-gain effect under these conditions, so clearly shown in the present data, must not be interpreted as a general result; if one terminal were well above the duct, the situation might be completely different.

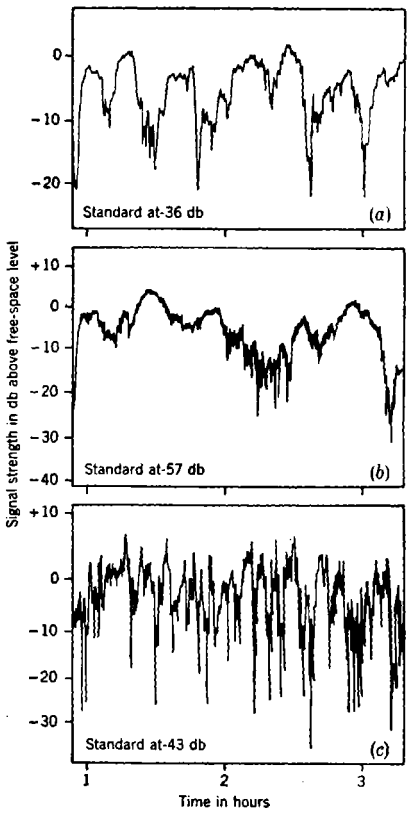


FIG. 4·6.—Comparison of microwave signal fading on three circuits, in the presence of deep ducts. (*a*) 9 cm, upper; (*b*) 9 cm, lower; (*c*) 3 cm, upper.

4·4. Comparison with Theory. —The problem of propagation in an inhomogeneous medium has been discussed from the theoretical standpoint in Chap. 2. For convenience the pertinent results of the solution are summarized here. The discussion is limited to the bilinear-index model.

The bilinear model (see Fig. 2·28) is described by two parameters: d the duct height (g in natural units[1]), and s^3, the ratio of the slopes of the lower and upper segments. The upper segment is assumed to have the standard slope, which in natural units is unity.

Characteristic values of the first mode of the equation $(d^2y/d\zeta^2) + \zeta y = 0$, with appropriate boundary conditions, have been found as a function of g for a series of values of s less than $+1$, that is, for superstandard M-profiles. The measured or synthetic M-profile, superstandard at the surface and standard above, is approximated by a bilinear curve, and field strengths are then calculated. Numerical work is in general feasible only for terminal separations sufficiently great that at most a few terms of the series expansion are required.

[1] See Sec. 2·6 for the definition of natural units.

An arbitrary criterion must be established for fitting observed M-profiles with a bilinear model, for it is not generally possible to draw a bilinear curve that appears to fit the true curve. The criterion adopted in the present work is the equality, for the true and the bilinear profiles, of the phase integrals (see Sec. 2·8) taken from the surface to the minimum of the curve. Thus,

$$\int_0^g \sqrt{[Y(z) - Y(g)]_{\text{true}}}\, dz = \int_0^g \sqrt{s^3 z - g)}\, dz, \tag{1}$$

a relation that serves to determine the value of s.

For the shallow M-inversions that exist with heating from below, the true M-profile is logarithmic (Sec. 3·14). In this case the value $s = -1$ has been found to give an acceptable fit. On the other hand, a visual approximation can be made for the deeper ducts associated with temperature inversions. In the Massachusetts Bay region values of s between -1.0 and -1.5 have usually been found suitable.

Let us suppose that $s = -1$, and let g increase from zero. From Figs. 2·30 to 2·32 we see that two major effects occur: The attenuation of the first mode (imaginary part of the characterictic value) decreases from 2.02 when g is zero to 0.1 when g is 2.7; at the same time the height-gain function decreases in absolute magnitude. These two effects tend to counteract each other, so that at any specified range a value of g may be determined below which no appreciable increases in field strength will occur. For the moderate range (40 miles) of interest in the Massachusetts Bay observations, this value of g is close to 1. Duct heights of less than one natural unit cause negligible effects at such distances. Similarly, for g between 1 and 2 the first mode alone is sufficient to represent the field. Field strength in this interval is a monotonic function of duct height, increasing from the standard to near the free-space value. Such ducts will be termed "shallow." When g is greater than 2, contributions of two or more modes must be summed vectorially, and the ducts are termed "deep."

Although it is strictly true that whether a duct is shallow or deep depends on wavelength, it will be convenient in the following discussion to adopt a terminology appropriate for microwaves (9 and 3 cm). We can, thus, somewhat loosely call surface ducts of height less than 50 ft "shallow" and those more than 50 ft "deep."[1] No ducts were observed that would be termed deep at 2.5 m. On the other hand, at 1.25 cm the value $g = 2$ corresponds to a duct height of only about 15 ft, so that for practical purposes any observed duct will be deep at this wavelength.

[1] This dividing line between shallow and deep ducts was chosen partly as a compromise between 9- and 3-cm values and partly to represent the meteorological conditions of instability and stability (Secs. 3·14 and 3·17) for which a 50-ft dividing line is a rough approximation.

When g is greater than 2, the numerical work increases so rapidly that general calculations become prohibitive. If both terminals are within the duct, it is frequently permissible to neglect all modes that are not strongly "locked," and in that case the calculations can be carried out by the Gamow procedure (see Sec. 2·17). From results of special cases carried through in this fashion some general conclusions may be drawn.

A total of 18 cases was treated in which three or more modes were necessary. The resultant field strengths varied between 14 db above and 7 db below the free-space field. In the following discussion the free-space field will be taken as the reference level. If the parameters g and s are changed by small amounts,[1] the amplitudes of the individual modes change only slightly but the phases vary so greatly as in some cases to produce extreme variations in the resultant.

Because the meteorological parameters are so variable and the adopted values of g and s are inexact as a result of both the meteorological uncertainties and the approximation of the bilinear model, the field-strength calculations for strong trapping may be subject to much greater inaccuracies than those for partial trapping. In place of specific values for each case, therefore, an arbitrary value of 0 db has been adopted as the nominal field strength for all cases of surface M-inversions with g greater than 2.

Theoretical values of field strength on a given transmission circuit may now be determined for any case of surface M-inversions, and the validity of the theory may be tested quantitatively. An examination of the synthetic soundings (see Sec. 4·2) and qualitative discussions for each day of the transmission period permitted the selection of those times when, with surface M-inversions, the theoretical assumption of horizontal stratification was most closely realized. The synthetic M-profiles for these times are then fitted with the appropriate bilinear index models; field strengths are calculated and compared with the observed hourly average values of signal strength.

To facilitate the test of the theory, for each of the 41-mile circuits across Massachusetts Bay graphs were prepared of field strength as a function of duct height for various values of the parameter s. Two such graphs are shown in Fig. 4·7 with $s = -1$. In accordance with the convention described above, for all cases of g greater than 2 the theoretical value was taken as 0 db, although it was understood that variations of at least ± 10 db were to be expected.

Figure 4·8 shows the comparison of theoretical and observed variations of signal strength with duct height for the lower of the two 9-cm circuits. Although the scatter is appreciable, the general theoretical trend is obeyed.

[1] A variation in g alone produces only a rotation of the vector diagram of the sum of modes and thus leaves the resultant unchanged unless there is a change in the number of modes trapped.

An error in the duct height of as little as 10 ft would in many cases suffice to account for the deviations. Not only are the probable errors of the synthetic values large, but the variations in duct height along the path

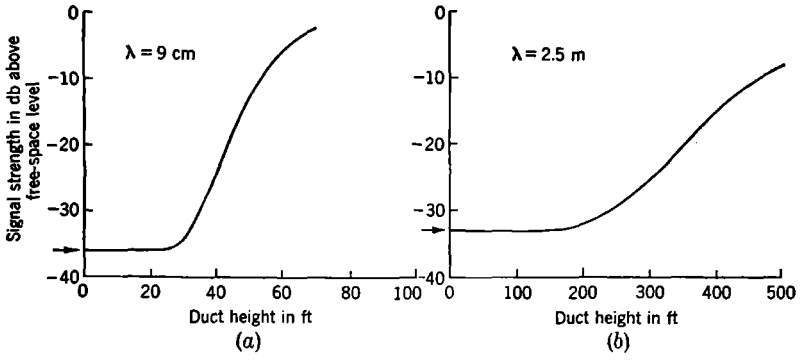


FIG. 4·7.—Calculated variation of received power as a function of duct height, based on a bilinear index model with the slope of the bottom segment the negative of standard. (*a*) λ = 9 cm; (*b*) λ = 2.5 m.

caused by changing water temperatures must frequently exceed 10 ft. Additional scatter should result from the approximation of the true M-profile by a bilinear distribution.

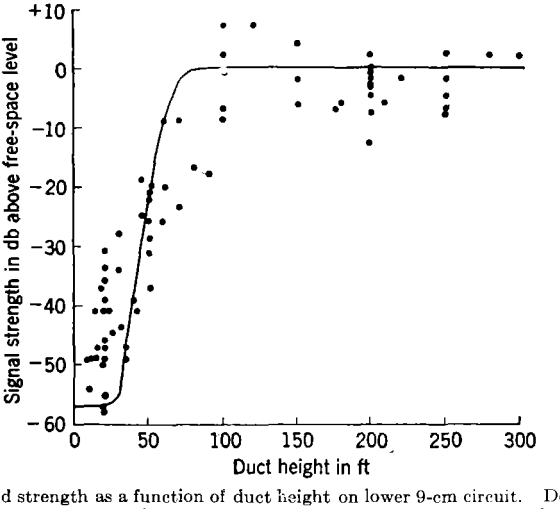


FIG. 4·8.—Field strength as a function of duct height on lower 9-cm circuit. Dots represent measurements, and the solid curve is calculated for a bilinear model with $s = -1$.

Figure 4·9 is a similar graph of theoretical and observed field strengths for the 2.5-m transmission. Again the theoretical trend is followed roughly, though in this case the scatter is large. The lack of field strengths

as high as the free-space value results from the meteorological characteristics of the circuit; surface M-inversions deeper than 300 ft occur only infrequently, and none was measured more than 400 ft deep.

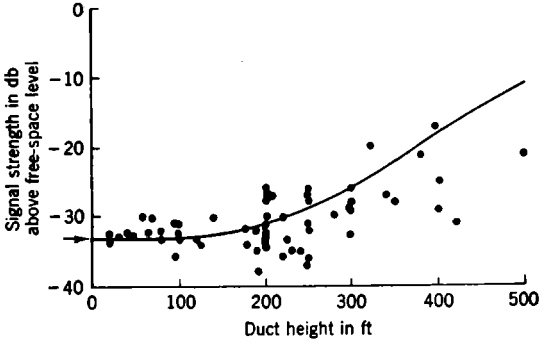


FIG. 4·9.—Field strength as a function of duct height at 2.5 m for surface M-inversions. The dots represent measurements, and the solid curve is calculated for a bilinear model with $s = -1$. The arrow indicates standard signal level.

As an additional test of the theory Figs. 4·10 and 4·11 were prepared. In Fig. 4·10 the observed field strengths on the two 9-cm circuits are compared. The solid line represents the theoretical variation of field strengths

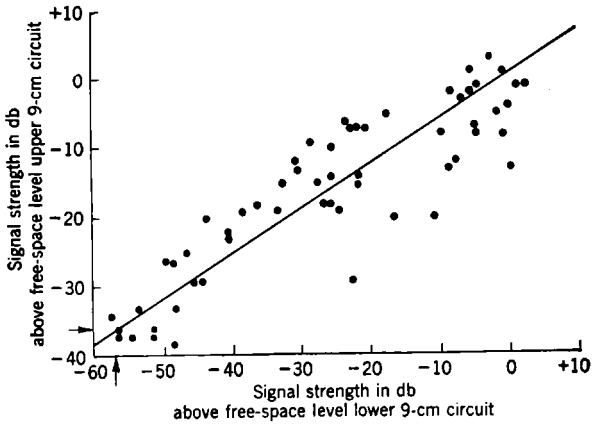


FIG. 4·10.—Observed relation of field strengths on upper and lower 9-cm circuits (dots) compared with the theoretical relation (curve).

on the two circuits calculated for a series of values of g, with $s = -1$. The observed values (dots) fall closely along the line. A similar agreement is obtained at 3 cm.

In Fig. 4·11 the 9- and 3-cm observations on the higher of the two circuits are compared. Again the curve shows the calculated variation, and

the dots represent the observations. Here also, the calculations are based on a bilinear index model with $s = -1$. In this case the observations fall systematically below the theoretical curve. If s is increased algebraically from the value -1, the sharp "corner" on the calculated curve is somewhat rounded but the curve remains a poor approximation to the observations. A rounded curve is obtained also with a power-law model (Sec. 2·21), but here the maximum field strength is found to be 15 to 20 db below the free-space value, in disagreement with observations.[1] Better results

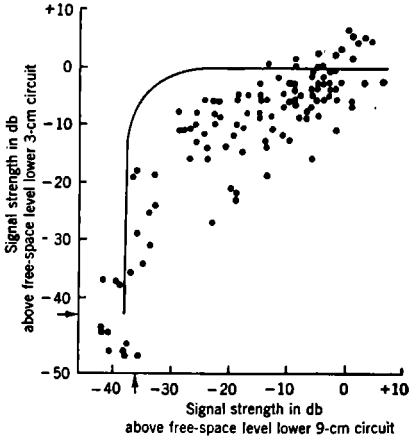


FIG. 4·11—Observed relation of field strengths at 3 and 9 cm (dots) compared with the theoretical relation (curve).

would probably be obtained with the linear-exponential model (Sec. 2·21), which more closely approximates the true M-profile.

Any surface M-inversion more than 15 ft deep is effectively a deep duct for a wavelength of 1.25 cm, so that no shallow-duct effects should be apparent at this wavelength. We should therefore expect to find signal of average level near the free-space value whenever either shallow or deep ducts exist for 3 or 9 cm As only field strengths stronger than 20 db above the standard value could be detected, the 1.25-cm signal should have been detected only with superrefraction. Furthermore, from an extrapolation of the frequency effect observed in fading of the strong-trapping type, we expect to find violent fluctuations with periods of 2 min or less. This picture is modified by the presence of appreciable atmospheric absorption at this wavelength. The water-vapor absorption band with center at about 1.33 cm is strong enough to cause an attenuation in 41 miles of 1.4 db for each gram per cubic meter of water-vapor in the air, or roughly

[1] Recent information from England indicates that this disagreement may not be general. See B. J. Starnecki, *Jour. IEE*, **93**, III *A*, 106 (1946).

1.2db/mb vapor pressure (see Chap. 8). During the period of the transmission tests at 1.25 cm, vapor pressures of 5 to 15 mb were measured in the first 100 ft above the water surface, which would produce total attenuations of 6 to 18 db. Actually the average signal level practically never reached a value stronger than 20 db below the free-space value. Thus, it appears that the water-vapor attenuation is probably not sufficient to explain the entire discrepancy between the observations and the values calculated with surface ducts. This discrepancy is not surprising in view of the difficulty of making accurate absolute measurements on 1.25 cm at the time these experiments were performed.

The 1.25-cm signal was detected on most occasions when the values on the other microwave circuits were superstandard, but the data are not conclusive on this point because operating difficulties at this wavelength prevented continuous information. Except for the observation that strong rapid fluctuations were nearly always present, signal character could not be studied with the equipment available at 1.25 cm.

In the preceding discussion only average values of field strength have been considered. A completely satisfactory theory must, however, account not only for large-scale variations in the field strength and the dependence of these changes on wavelength and circuit geometry but also for the character of the received signal, especially as regards short-period variations, or fading.

Let us consider what signal types are expected in terms of the variations of the modified-index distributions in time and space. Suppose there is a surface duct such that $g = 1.5$ at 9 cm. For such a shallow duct only the first transmission mode is important. Small changes in the value of g cause small variations in the amplitude of this mode and therefore in the resulting field strength. Small changes in the parameter s have a similarly small effect. One may conclude then that relatively steady signal should be obtained as long as the given meteorological situation continues.

For the same meteorological conditions, g for 3 cm would be 3. In this case consideration of the first mode alone is no longer sufficient. Instead, contributions of several modes must be summed vectorially. Small changes in the parameters s and g may now result in large variations of the resultant field, as has already been described. Thus unsteady signal may be expected, with periods that depend on both the nature of the atmospheric variability and the sensitivity of the electric field to such atmospheric variations. There is no reason to expect synchronous fading at two wavelengths or on two paths.

Some insight into the kind of fluctuations in field strength that may result from interference among many modes can be obtained from a statistical analysis based on the theory of random processes (see Sec. 6·19). It is assumed that the condition responsible for the interference effect

(in our cases the existence of a uniform surface duct over the path) is stationary in time.

For a stationary process it can be shown (see Sec. 6·19) that given N vectors of arbitrary amplitudes and random phases, the probability W that the resultant power P is in the range P to $P + dP$ is given by

$$W(P)\,dP = \frac{1}{P_0}\,e^{-P/P_0}\,dP, \tag{2}$$

where P_0 is the mean power. This expression is valid and independent of N when N is greater than about 10. For smaller values of N the proba-

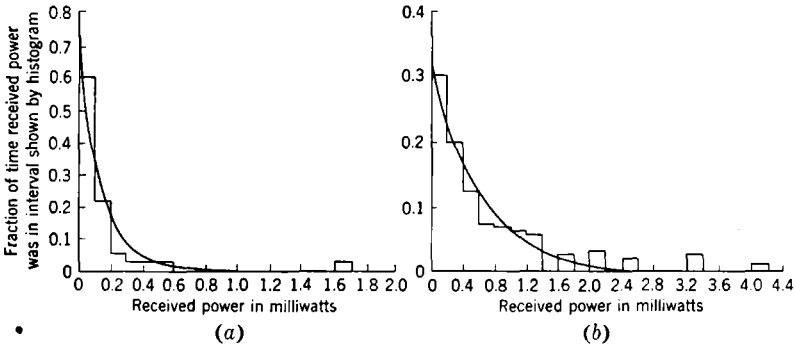


FIG. 4·12.—First probability distribution of power received during periods of strong trapping and histograms of observations. (*a*) $\lambda = 9$ cm; (*b*) $\lambda = 3$ cm.

bility depends on the distribution in time of the amplitudes of the individual terms. For a Gaussian distribution of the amplitudes, however, Eq. (2) holds even when N is 2.

The results of an analysis of fading data obtained at both 3 and 9 cm on the 41-mile path are shown in Fig. 4·12 where the histograms represent the observations and the smooth curves the relation of Eq. (2). The analysis is open to question in two respects. (1) The phases of the various modes are not strictly independent, as each is as function of the wavelength, the modified-index distribution, and the path geometry. For practical purposes, however, the probability distribution of the resultant powers should be closely similar to that for purely random quantities, because a very large change in the phase of any one mode relative to the others results from a very small change in the M-profile. (2) More important, a truly stationary situation probably never exists when air flows out from a large land mass over the sea.

Because the individual amplitudes are assumed to be independent of each other, points must be chosen that are separated by a time interval at least not smaller than the periodicity of the variations observed. The results will be significant only if the number of values is large. Thus the

total period of the study must be long compared with the time interval selected. This imposes a serious limitation on the applicability of the method to our data, because stationary conditions seldom persisted for time intervals sufficiently long compared with the fading periods.

Despite these difficulties, the histograms of Fig. 4·12 are in fair agreement with the theoretical curves.

Thus the hypothesis is favored that fading results from the interference of many modes, whose contributions change rapidly because of minor changes in the modified-index distributions. The observed excess of high values might result either from a non-Gaussian distribution of individual amplitudes, if N is less than about 10, or from departures from stationary conditions.

4·5. Transmission under Complex Conditions.—With a few minor exceptions our discussion to this point has been limited to the consideration of simple and idealized cases of surface M-inversions that are theoretically tractable and for which the theory appears to be qualitatively correct. This section will treat somewhat more fully the observational information pertaining to transmission with elevated M-inversions, substandard layers, nonuniform meteorological conditions, and some further details concerning the effects of surface M-anomalies not covered in previous sections.

No exact solution of the wave equation has yet been obtained for the case of an elevated anomaly of modified refractive index. The attempted approximate solutions expressed in terms of reflection from an elevated layer[1] are not considered satisfactory (see Sec. 4·9). Furthermore, the idealized model of the refractive-index distribution used in these solutions, although it may be useful under certain meteorological conditions, has little resemblance to the distributions that actually occur in the Massachusetts Bay region. Thus, we shall record only empirical observations and leave to the future their integration into a conclusive theory.

The observations of superstandard signal reception at 2.5 m shown in Fig. 4·9, which are limited to cases of surface M-inversions, account for only 40 per cent of the total cases of strong signal at this wavelength. The theoretical curve of Fig. 4·7*b* indicates that with surface M-inversions duct heights 300 ft or greater are necessary to produce appreciable increases above standard in 2.5-m field strength whereas meteorological conditions in Massachusetts Bay are such that surface ducts 100 to 200 ft deep occur commonly, but greater depths are very rare.

A qualitative correlation between strong signal at 2.5 m and the presence of elevated ducts was found that served to account for nearly all remaining hours of superstandard reception at this wavelength. Similar

[1] See J. B. Smyth, "Transmission of Plane Waves through a Single Stratum Separating Two Media," NRSL Reports Nos. WP-9 and 13, Parts I and II, Dec. 22, 1943, and June 23, 1944.

observations made in the vicinity of San Diego[1] showed a definite relationship between the value of the field strength and the height of the base of the inversion. No such height-strength relationship is apparent in the Massachusetts Bay data.

The approximate constancy of thickness and strength of the M-inversions in the San Diego region, which makes possible the description of the M-profiles there by a single parameter, that is, the height of the base of the inversion, results from the special meteorological conditions of that locality and is doubtless essential to the simple relation found experimentally between field strength and the M-profile.

No such simplicity of form of M-profiles is found in Massachusetts Bay. The elevated ducts that occur there originate mainly in the modification of initially inhomogeneous air, a phenomenon occurring most commonly when a nocturnal radiational inversion exists over land. As has been mentioned (see Secs. 3·21 and 4·2), the meteorological analysis of such cases is not yet well understood, and the results are therefore less reliable than in other instances. Thus it was frequently impossible for the meteorologists to specify more than the mere existence of an elevated duct, and it was only rarely that it was possible even to estimate the height of its base, the depth, or the strength because the initial distributions of temperature and moisture were not known.

The failure to secure measurements on elevated ducts is again intimately related to their frequent origin in radiational inversions over land. The elevated ducts usually occurred along the radio paths during the night and early morning and had generally been dissipated before the morning airplane soundings could be made.

Except in the case of elevated layers that are overhanging and thus produce surface ducts, no evidence was found that such layers have an appreciable effect on transmission of the shorter wavelengths considered here (but see Sec. 4·18).

Some additional discussion of transmission through substandard surface layers should also be included here, but solutions for the substandard case are not yet available.[2]

We have seen in the bilinear index model that for surface M-inversions the slope of the inverted portion may be treated as a constant, the only variable being the duct height. For substandard surface layers we may, analogously, neglect variations in slope of the lower segment and plot average field strength on a given circuit as a function of the depth of the

[1] See Sec. 4·9; also, L. J. Anderson, F. P. Dane, J. P. Day, *et al.*, "Atmospheric Refraction—A Preliminary Quantitative Investigation," NRSL Report No. WP-17, Dec. 28, 1944.

[2] Since the above material was written a report by T. Pearcey and M. Tomlin has come to our attention. See "The Effect of a Subrefracting Layer on the Propagation of Radio Waves," RRDE Memorandum No. 83, Feb. 12, 1945.

substandard layer. The results are shown in Fig. 4·13*a* and *b* for the upper 3- and 9-cm circuits respectively. Only cases with approximate meteorological uniformity have been included. The points cannot be said to define the variation of field strength with substandard layer depth, but they suggest a decrease in field with increasing layer depth. Values of d smaller than 100 ft do not occur because shallow substandard layers cannot form; substandard layers exist only under stable meteorological conditions (see Sec. 3·20).

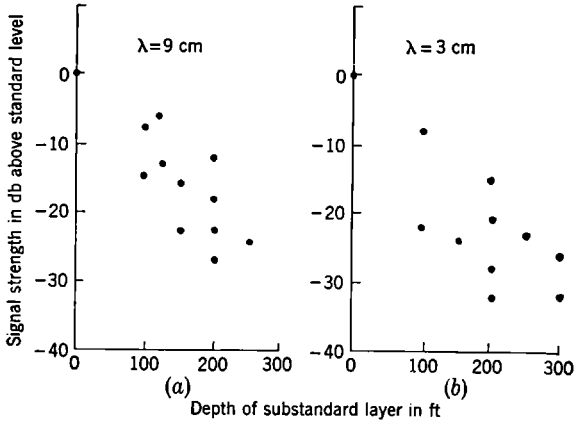


FIG. 4·13 —Signal strength as a function of depth of a surface substandard-layer. (*a*) 9 cm; (*b*) 3 cm.

A peculiar anomaly has been observed in the performance on the lower 3-cm circuit under substandard conditions. The signal level on this circuit *never* dropped below the minimum detectable level, although the standard value was only a few decibels above the detection limit. Whenever the signal on the other microwave links became low and scintillating with peak values near or below standard level, scintillation also began on the lower 3-cm circuit; but even the minimum signal level reached was often not below standard. The average value was usually considerably above standard. On these occasions the absolute power levels on the upper and lower 3-cm circuits are roughly equal. The 9-cm data on the same circuit agree with those on the two upper circuits rather than with these anomalous results.[1]

[1] *Editor's note added in proof.* At the time these results were obtained they caused us considerable concern, for it was clear that other processes were occurring that were not included in the theory then available. On the basis of these results and those obtained by NRL in Antigua (Sec. 4·20), C. L. Pekeris and others suggested that the unexpected signal might arise from scattering by atmospheric irregularities not previously considered. At the date of writing this note (1948) it appears, on the basis of

The 41-mile path across Massachusetts Bay was chosen, at least in part, with a view of obtaining uniformity of meteorological conditions in the region. Such uniformity would exist with westerly winds, provided the air was well mixed upon leaving the shore and was modified only by its passage over the water. Under most circumstances, the departures from this uniformity caused by the variation in water temperatures are not serious.

Fairly uniform conditions were actually found on many occasions, but the deviation of the wind direction from west, the variation of water temperatures, and the heterogeneity of the air before passage over water, especially at night, all contributed to the development of strong nonuniformity. In addition, on a few occasions a stationary front straddled the transmission path.

Despite the complexity of the resulting situations and the consequent impossibility of coping with the problem theoretically, observations under these conditions have disclosed some interesting features that may have general significance. These observations cover combinations of shallow and deep surface ducts along the path and of substandard and superstandard surface layers (usually deep ducts). The surface layer of the M-profile may, for example, be substandard at one end of the circuit and superstandard at the other or substandard at both ends and superstandard near the mid-point.

The outstanding feature is the close similarity in all circumstances of the results on the four 41-mile microwave circuits. The strong wavelength dependence of the attenuation constants of the individual transmission modes and the accompanying sensitivity to change in wavelength of the height-gain functions might lead one to expect that the over-all field distributions in cases of nonuniformity should also show a wavelength effect. No such effect was observed.

With conditions of variable surface duct height along the path, the values of d changing over an appreciable range, the resultant signal generally exhibited the irregular variability characteristic of deep ducts. No general behavior was found in the average signal level at such times, however; the values varied from very high (at or slightly above the free-space value) to -15 or -20 db.

With surface layers varying from substandard to superstandard along the path, results ranged from clearly substandard in both level and character to the strong-trapping type, often with average level somewhat depressed from the free-space value. As has been stated, in each case

---

growing evidence that cannot be reported here, that such scattering may indeed be an important factor in causing transmission beyond the horizon and between low terminal heights. See H. G. Booker and W. E. Gordon, *Proc. I.R.E.*, **38**, 401 (1950); and a series of reports by N. G. Parke entitled "Microwaves in an Irregular Atmosphere" issued by the Air Force Cambridge Research Laboratories, Cambridge, Mass., beginning Sept. 30, 1949.

the balance between the extremes of modified-index distributions apparently affected both 3- and 9-cm circuits similarly. The nonreproducibility of such conditions prevented the determination of the extent to which superstandard and substandard surface layers counteract each other. On several occasions substandard layers over the circuit were displaced through the passage of a cold (dry) front parallel to the transmission path, so that in the course of a few hours (the time of passage of the front over both terminals) the M-profiles at the surface were successively everywhere substandard, partially superstandard and partially substandard, and finally completely superstandard. Results obtained on these occasions indicate that the presence of a surface duct over only a small fraction of the path may suffice to produce transmission indistinguishable from that with uniform ducts. Figure 4·5 illustrates such a situation. Meteorological data show that it took 2 hr for the front to pass from the northern to the southern end of the path. Yet the striking change in transmission shown in the figure occurred within $\frac{1}{2}$ hr of the frontal passage over the northern terminals.

We have seen that the results on the two paths (Deer Island to Eastern Point and Race Point to Eastern Point) were almost always similar. Any exceptions resulted mainly from meteorological nonuniformity. It is obvious, for example, that reception on the two paths might differ radically if there were a superstandard surface layer throughout the northern part of Massachusetts Bay but an intense substandard surface layer over the southern part. Equally striking differences have been explained by the fact that the average water temperature close to the shore is lower than that out in the bay. When the air and water temperatures were nearly equal, a 2° or 3° water-temperature variation with position could cause a change from unstable to stable meteorological conditions (see Secs. 3·14 and 3·17). Similarly, even when surface conditions were substandard throughout the bay, the substandard layer close to the shore was frequently more intense than that farther away. Thus, the occurrence of substandard signal was more pronounced on the 22-mile path than on the longer path. Finally, some cases of dissimilarity between the two paths occurred when with westerly winds the air leaving the coast was not homogeneous. In that case, the original anomaly in the modified-index distribution may in some cases have been only slightly changed in the first few miles of the air passage over the ocean but almost completely altered by the time it reached the 41-mile path.

4·6. Some Statistical Results.—A statistical analysis of the reception has been made for the entire 4-month period (1944) of the transmission test from Race Point to Eastern Point. It must be emphasized that the results of this analysis give merely a convenient condensed description of the conditions on the specific circuit during the specific period. They should not be generalized to another circuit or another time period unless

the meteorological situations of the two are closely similar (and, of course, only if the circuits are similar from a radio point of view).

The test period, covering the months July through October, provided a varied sample of weather conditions. The analysis was carried out by weeks (168 hourly values in each) and finally summed for the total period of more than 2500 operating hours. Figure 4·14 shows the results obtained on the 9-cm upper link for two extreme weeks and also the over-all averages. The ordinate is the fraction of time during which the average level

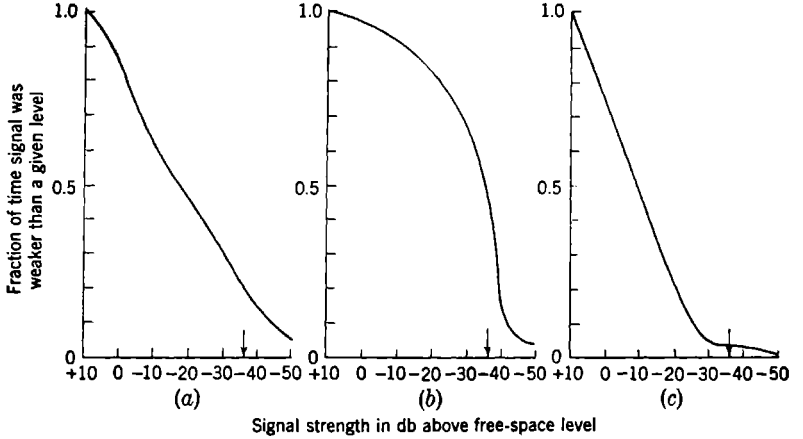


FIG. 4·14.—Statistical analysis of reception on the upper 9-cm circuit. (*a*) Average over entire period; (*b*) week of extremely poor transmission; (*c*) week of extremely good transmission.

was weaker than the given value; the abscissa is the signal level itself, referred to the free-space level. The standard level, −36 db for this circuit, is indicated by an arrow. The graph shows that the field strength on this circuit was above standard more than 80 per cent of the total transmission time. Similar numerical results were obtained on the other microwave circuits, as would be expected in view of their parallel behavior. Even at 2.5 m the level was above standard 75 per cent of the time, whereas at 1.25 cm the signal was detected (and therefore at least 20 db above standard) 79 per cent of the total operating time.[1] The comparison of times above standard implies a weaker wavelength dependence of transmission than actually exists. The difference is demonstrated by the following figures: At 2.5 m, the average level was 10 db or more above the standard 18 per cent of the time; at 9 cm (upper circuit), 65 per cent of the time; at 1.25 cm, at least 79 per cent of the time.

Some seasonal variation was found, although the effect is somewhat masked by the large day-to-day changes typical of New England. For

[1] This is not a good statistical sample, however, because of the irregular periods of the circuit's operation.

example, the incidence of substandard signal tended to drop with the change from summer to fall. Similarly, the incidence at 9 and 3 cm of partial trapping as compared with strong trapping increased as the weather became colder. These tendencies can easily be explained in terms of the average value of the difference between the water and air temperatures at different seasons and the resulting M-deficits. Because the water temperature in Massachusetts Bay drops very slowly at the end of the summer while the average air temperature decreases rapidly, the average value of ΔT changes from positive to negative. With negative ΔT (see Sec. 3·15) the air is heated from below; neither deep surface ducts nor substandard surface layers form, but shallow ducts are almost always present. It is

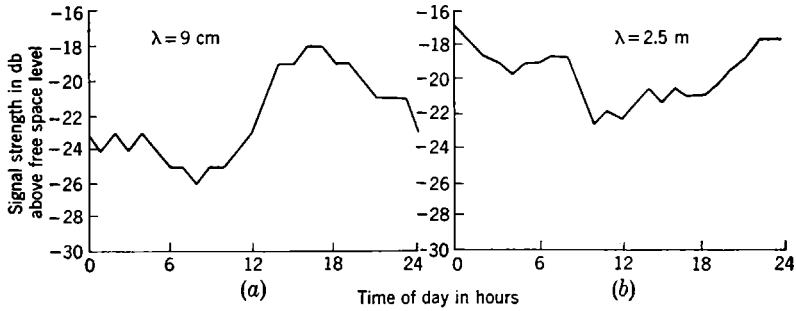


FIG. 4·15.—Average signal level as a function of the time of day, showing diurnal variation. (*a*) 9 cm; (*b*) 2.5 m.

only because of the large deviations from the average conditions in this region that the seasonal effect is not more striking. If the tests had been continued into the winter, we should expect to have found a complete disappearance of substandard reception and only infrequent periods of strong trapping, all of short duration. This expectation was, in fact, borne out in some winter tests on the 22-mile path from Deer Island to Eastern Point (1942-1943), when on only 3 days out of a total of 54 a fairly deep duct could be said to exist. No substandard signal was observed. Indeed, the signal was almost always very steady and not far from standard, the major variation noticed being the change in standard level produced by the tidal variations in terminal height.

A diurnal effect was noted by British observers in the transmission across the Irish Sea (see Sec. 4·7). Our records show no such effect upon visual inspection. A statistical study of the average field recorded at a given time of day in the total period results, however, in the curves shown in Fig. 4·15*a* and *b* for 9 cm and 2.5 m respectively, both of which indicate the existence of a diurnal variation. At 9 cm the spread in the average values is 8 db, the peak occurring near 1700 hours and the minimum near 0800 hours Eastern Standard Time. Because at 9 cm variations in average

signal level of 30 to 60 db within an hour are not uncommon (see Fig. 4·4), it is not surprising that this small diurnal variation is not noticeable. At 2.5 m the total spread of the averages is found to be 6 db, with a peak at 0100 and a minimum at 1100. The curve is less symmetrical than that at 9 cm, and the difference in times of the maxima and minima at the two wavelengths is further evidence that the causes of the strong signal are different. Because at 2.5 m changes in average level of more than 15 db in a short interval are rather rare, the diurnal variation found at this wavelength is relatively more significant than that at 9 cm.

TRANSMISSION EXPERIMENTS IN THE BRITISH ISLES

BY DONALD E. KERR

4·7. The Irish Sea Experiment.—One of the most extensive experimental overwater transmission investigations was carried out in England under the auspices of the Ultra Short Wave Propagation Panel of the Ministry of Supply in cooperation with several other government agencies. The major objectives and the plan of investigation were similar in many

TABLE 4·1.—STATION AND PATH DESIGNATIONS

Station	Location	Height, ft	Path used	
			3 cm	9 cm
A	South Wales	540	*AC*	*AC*
B	South Wales	90	....	*AD*
C	North Wales	825	*AE**	*AE**
D	North Wales	95	*BD*	*BE*
E	Scotland	375	*BF**	*BF**
F	Scotland	95	....	....

* These paths were operated for only part of the test period.

respects to those of the work described in the preceding sections and were part of a broader program of propagation research in England.[1]

This experiment employed one-way transmission over the Irish Sea on several wavelengths, but only the results obtained on 3 and 9 cm have been reported in sufficient detail to discuss here. There were two transmission paths with two transmitting stations in South Wales, two receiving stations in North Wales, and two receiving stations in Scotland. The locations, heights, and letter designations of the stations are given in Table 4·1, and their locations are shown in Fig. 4·16. The transmission tests began in the summer of 1943 on a partial basis and operated nearly

[1] Descriptions of the various aspects of this program are distributed throughout a large number of British reports. A summary of the principal features of the program appears in a paper presented by E. C. S. Megaw at the third conference on propagation,

continuously (with the exceptions noted) from the latter part of 1943 until the war came to an end.

The path length from South Wales (AB) to North Wales (CD) was 57 statute miles, and that to Scotland (EF) was 200 statute miles. The path AC was optical, but all others were nonoptical.

Meteorological soundings were made by ships and aircraft on or near the transmission paths, employing instruments and techniques somewhat different from those used by workers in the United States (see Sec. 3·31). The soundings were either made from the ships by wet and dry thermocouples in radiation shields at fixed heights or carried by a captive balloon. Aircraft soundings were made by wet and dry thermocouples with flight techniques of several kinds. Vertical spirals at fixed locations were sometimes employed, but more often "shuttle" flights at a series of fixed levels and simultaneous horizontal flights of three or four aircraft spaced vertically by several hundred feet (see Sec. 3·32). Each of these techniques may give a different kind of information, depending upon the degree of atmospheric homogeneity, and the data require careful analysis with this fact in mind.

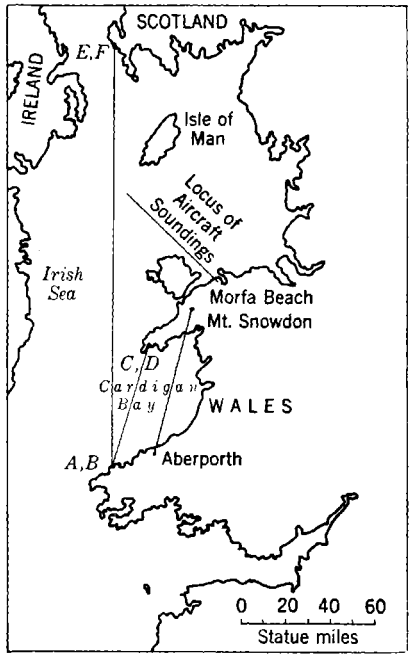


FIG. 4·16.—Map of Irish Sea region showing location of British one-way and radar transmission experiments.

There were several pronounced qualitative features of the transmission over the 57-mile circuit, some of them similar to those of other paths, Massachusetts Bay in particular. These features were

1. The maximum signal level maintained for any appreciable time on the nonoptical paths was of the order of the free-space value, but the maximum for short periods was higher by 10 to 15 db.
2. The range of variation of signal level on the nonoptical paths was very large and increased with excess of path length over horizon

Nov. 16–18, 1944, at Washington, D.C. See *Radio Wave Propagation*, Academic Press, New York, 1949; also *Jour. IEE*, **93**, III *A*, 79 (1946), *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London, 1946, and J. S. McPetrie, B. Starnecki, H. Jarkowski, and L. Sicinski, *Proc. IRE*, **37**, 243 (1949).

distance. It was also greater for 3 cm than for 9 cm by about the difference between free-space and standard signal levels for the two wavelengths. The signal on the optical path AC was much steadier than on the other paths.

3. The major changes in signal level on the two wavelengths and on the three nonoptical circuits AD, BC, and BD usually occurred simultaneously, although the fine details, that is, fading and scintillation, did not.
4. Very high average signal level was accompanied by deep fading; intermediate levels had somewhat less deep fades. There were also periods (12 to 48 hr) when the signal varied steadily and slowly, and fading was not normally experienced.
5. In general, both the rapidity and depth of fading were greater for the shorter wavelength.
6. Pronounced diurnal variations of signal strength occurred on the nonoptical circuits when the signal level was low or moderate, with an increase of average signal level in the late afternoon or early evening and a decrease in the very early morning.[1]
7. Periods of low signal strength were found to coincide, in some cases, with poor visibility from light fog at the terminals, and with passage of fronts across Cardigan Bay.[2] Low signal strength did not ordinarily coincide with reduction of visibility from fine-weather haze, however.
8. In general, periods of high signal strength tended to coincide with anticyclonic periods.
9. A seasonal trend was also found, with high signal level for a greater fraction of the time in summer than in spring or winter.[3]

On the 200-mile circuits the standard signal levels were far below the free-space levels—about 305 db for path AE on 10 cm and 520 db for path BF on 3 cm—but when the signals on any of these circuits were detectable, they tended to rise somewhat abruptly to the vicinity of the free-space level. The general character of the signal behavior was strikingly similar

[1] Similar diurnal variations were observed in Massachusetts Bay (see Fig. 4·15*a*), but in general they were so much outweighed by short-period variations as to require a statistical analysis to establish their presence.

[2] From the results in Massachusetts Bay one would suspect that fog and poor visibility might often be symptoms of meteorological conditions giving substandard refraction, although, as pointed out in Sec. 3·17, the presence of fog is neither a necessary nor sufficient condition for this to occur. Frontal passages over Massachusetts Bay produced no unique signal trend, but in each case the reaction of the water surface upon the new air brought in by the front produced a change in the M-profile sufficient to explain the behavior of the signal.

[3] In Massachusetts Bay, measurements were not made over a sufficient time to establish this point definitely, but from Sec. 4·6 it will be seen that such a seasonal trend is not unlikely.

on the two paths and on the two wavelengths, but at times there were noticeable differences in details. In general, the signals were measurable for a greater fraction of the time on 9 cm at the high site E than on 3 cm at the low site F, and at a given site were stronger at 9 than at 3 cm. As the minimum detectable signal was roughly 30 db below the free-space level for all the receivers, this meant that the signals tended to be somewhat nearer the free-space level on 9 cm at E than on 3 cm at F. During periods of very strong superrefraction, however, this situation was reversed. It was also found that, in general, signals were detected on the 200-mile paths only when the 3-cm signals on the 57-mile BD circuit

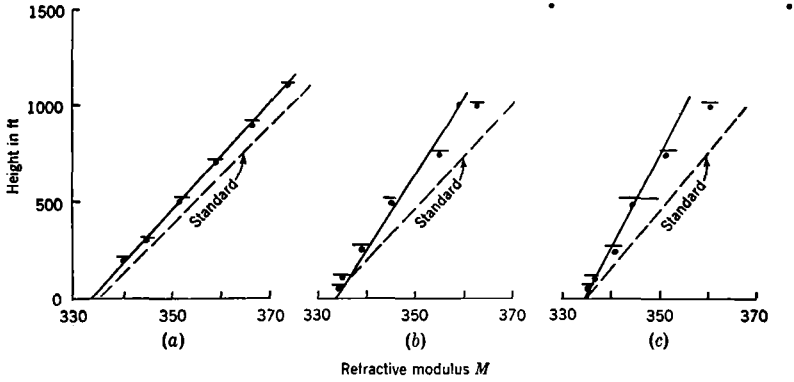


FIG. 4·17.—Mean values and observed variations of M along the BC Irish Sea path. (*a*) at 0600, Mar. 15, 1945; (*b*) at 1200, Mar. 16, 1945; (*c*) at 1800, Mar. 16, 1945.

were stronger than about 12 db below the free-space level. Successful reception was obtained during warm weather, but 9- and 3-cm signals were never detected on the long paths during the winter.

Analysis of the transmission records for the 57-mile path was made at the National Physical Laboratory.[1] Correlations were first sought between the M-profiles in the lowest 50 to 100 ft obtained from the ship soundings and the signal behavior on the transmission path. Such correlations were either lacking or were poorly defined, however, and attention was shifted from low-level soundings to gross phenomena such as fronts

[1] For further details, see R. L. Smith-Rose and A. C. Stickland, "Centimetre Wave Propagation over Sea. A Study of Signal Strength Records taken in Cardigan Bay, Wales, between February and September 1944," RRB/C114, Feb. 28, 1945; "Centimetre Wave Propagation over Sea. Correlation of Radio Field Strength Transmitted Across Cardigan Bay, Wales, with Gradient of Refractive Index Obtained from Aircraft Observations," I, RRB/C121, May 10, 1945, and II, RRB/C127, Oct. 1, 1945; also a paper in *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London, 1946; and a similar paper by E. C. S. Megaw in *Radio-Wave Propagation*, Academic Press, New York 1949; and R. L. Smith-Rose, *Jour. IEE*, **93**, III *A*, 98 (1946).

or fogs, and to aircraft soundings up to heights of several hundred feet. The profile used in each case was obtained by fitting a straight line through the measured points; the resulting slope was then used in calculation of the signal expected at the receivers, which was compared with the observed value.[1] A sample set of linear profiles is shown in Fig. 4·17, where the dots represent the mean values of M obtained in a single horizontal flight along the 57-mile path and the lines represent the range of variation encountered during the flight.

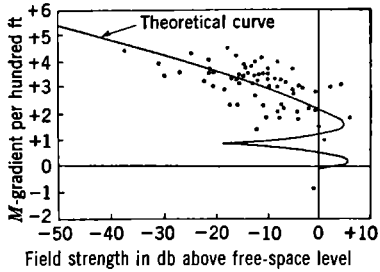


FIG. 4·18.—Correlation of observed field strengths with those predicted using M-gradients of the lowest 500 ft on the 9-cm *BC* Irish Sea path.

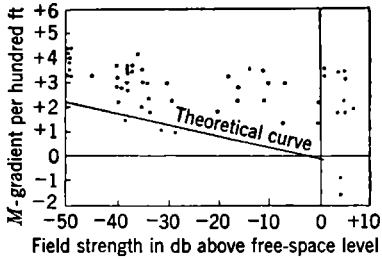


FIG. 4·19.—Correlation of observed field strengths with those predicted using M-gradients of the lowest 500 ft on the 9-cm *BD* Irish Sea path.

A comparison of predicted and observed field strengths for the *BC* circuit on 9 cm is shown in Fig. 4·18, in which the gradient has been chosen as an average for the lowest 500 ft. Figures 4·19 and 4·20 show similar results for the *BD* circuit on 9 and 3 cm.[2] This procedure for calculation of field strength gives results that agree reasonably well with experiment on the *BC* circuit, which has one high terminal, but the disagreement is pronounced for the *BD* circuit, for which both terminals are low. This disagreement is hardly surprising in view of the extreme oversimplication resulting from ignoring the irregularities of the M-profile.[3]

The fact that the signals on the *BD* circuit were so much stronger than calculated, together with the preponderance of evidence from other re-

[1] For details of how the fitting may be done in particular cases, see E. C. S. Megaw and F. L Westwater, "Radio-meteorological Relationships," AC 8336/U.S.W. 141, Ministry of Supply; also E C. S. Megaw, *Jour. IEE*, **93**, III *A*, 79 (1946).

[2] Similar graphs resulting from using the average gradient over the lowest 250 ft gave an appreciably greater scatter of points for both *BC* and *BD* circuits.

[3] It is interesting to note that if measurements at only a few heights had been obtained in the Massachusetts Bay soundings instead of a large number giving considerable detail, it would also have been possible to represent limited sections of these soundings by single straight lines of various slopes. This, however, would have seriously misrepresented the true nature of the profiles, a predominant characteristic of which was that deviations from standard slope usually occurred in well-defined layers, outside of which the slope quickly returned to a value very near standard (see also Sec. 4·9). But

see also the footnote on p. 312

gions, suggests that these high signals must have resulted from M-inversions that were not apparent from the soundings.[1] It is interesting to note that the aircraft soundings of the vertical spiral type made for the KXS trials (Sec. 4·17) in a region not far from that under discussion showed strong surface inversions similar to those shown in Chap. 3. The behavior of the circuits with low terminals is, as would be expected, a more critical test than with high terminals, and on the basis of available evidence it appears that the use of a linear M-profile with variable slope does not meet this test satisfactorily.

It is not clear, at present, whether the apparent differences between the Irish Sea and the Massachusetts Bay data result from greatly different meteorological situations in the two locations or whether the differences in methods of carrying out the experiments and analyzing the resulting data have resulted in unwitting emphasis of some aspects of the problem and neglect of others. It is difficult to see how the issue can be clarified on a sound basis without the equivalent for the Irish Sea data of the detailed analysis described in Secs. 4·2 and 4·4.

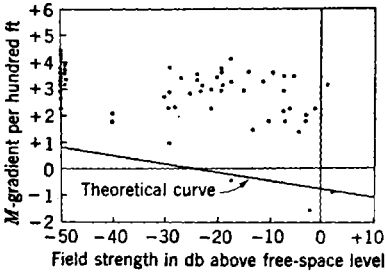


FIG. 4·20.—Correlation of observed field strengths with those predicted using M-gradients of the lowest 500 ft on the 3-cm *BD* Irish Sea path.

It should be pointed out that at the present time there are some differences of interpretation and method of application of the technique of representing the whole M-profile by a single straight line approximating a section of the M-profile having appreciable curvature.[2] This technique is usually easier to apply numerically than the wave-equation methods discussed previously; but in the opinion of the author it tends to obscure the true importance of M-profile curvature and is unsatisfactory as a theoretical basis for interpretation of experimental data with which we are familiar. It is hoped that future investigations will find a way of integrat-

[1] It should be noted that signal strengths up to 6 db above the free-space level can be explained by a linear M-profile, assuming unity reflection coefficient of the ocean surface, but in general the measured M-profiles were not sufficiently steep to predict such high values. Stronger fields could be obtained by using a negative slope, which is equivalent to assuming a concave earth surface, but the physical interpretation of the resulting multiple ray paths and caustics requires investigation (see Sec. 2·5). The equivalent earth radius calculated by these methods is different for each wavelength and set of terminal heights, although it refers to the same atmospheric structure.

[2] In recent private communication Megaw has pointed out that our interpretation of his use of the linear-gradient technique goes somewhat beyond his original intentions. Further discussion of this matter may be found by consulting B. J. Starnecki, *Jour. IEE.* **93,** III *A*, 106 (1946).

ing the two methods of approach into a procedure that retains the desirable features of both.

4·8. South Wales to Mt. Snowdon.—An optical overwater transmission path was operated from June to October 1944, between transmitters 120 ft above sea level in Aberporth, South Wales, and receivers at 3600 ft on Mt. Snowdon, about 70 miles away (see Fig. 4·16). Transmission was carried out on 9 and 10 cm simultaneously.

For about 80 per cent of the time the transmission on both wavelengths was characterized by a regular fading pattern, which, when plotted as a function of time, consisted of a succession of broad, flat maxima and deep, sharp minima, with maximum signal equal to approximately twice the free-space field. This is the kind of pattern which would be expected from interference of two signals nearly equal in amplitude but differing in phase; it was interpreted to be the result of interference between the direct ray and the specularly reflected ray from the surface of the bay.[1] The proposed explanation is that under these conditions the M-profile was essentially linear but varied slightly in slope, changing the path-length difference and consequently the phase difference between direct and reflected rays. For standard refraction this phase difference was about 4.5π on 10 cm and 5π on 9 cm, whereas for "flat-earth" conditions (vertical M-profile) these phase differences were about 12π and 13π respectively. Thus a steepening of the M-profile from the standard slope would tend to cause the two wavelengths to fade antisynchronously, as they were observed to do experimentally.

During the remaining 20 per cent of the time the regular interference pattern was replaced by a different form of fading. The field strength in peaks of the fading cycle was 10 to 12 db above the free-space field; the peaks were sharp and the minima were broad. This is the type of signal to be expected from superposition of a large number of components with random phase relationships, as might be expected on an optical path over which the index distribution varies with time, changing the amplitude and phase relationships between the several important transmission modes.

As no meteorological soundings were made over this path, further investigation into the actual mechanism is not promising.

TRANSMISSION ALONG THE CALIFORNIA COAST

BY DONALD E. KERR

4·9. San Diego to San Pedro.[2]—Another highly interesting experiment is that carried out by the U.S. Navy Radio and Sound Laboratory (now U.S. Navy Electronics Laboratory), San Diego, Calif. It is of particular

[1] F. Hoyle, "Fading in a Line of Sight Experiment in England," in *Radio-Wave Propagation*, Academic Press, New York, 1949.

[2] L. J. Anderson *et al.*, "Atmospheric Refraction—A Preliminary Quantitative Investigation," NRSL Report No. WP-17, December, 1944.

importance because of the striking difference between the meteorological situation (and consequently the M-profile) along this part of the California coast and the conditions in the situations so far described. A strong elevated M-inversion persists over great distances for months at a time, and there seem to be few departures from standard at very low levels except for a low evaporation duct.[1] No attempt will be made here to discuss the meteorological origin of these elevated inversions, for there seems to be some difference of opinion in the matter; the experimental facts concerning the M-profile and radio signal behavior are sufficiently definite to merit discussion.[2]

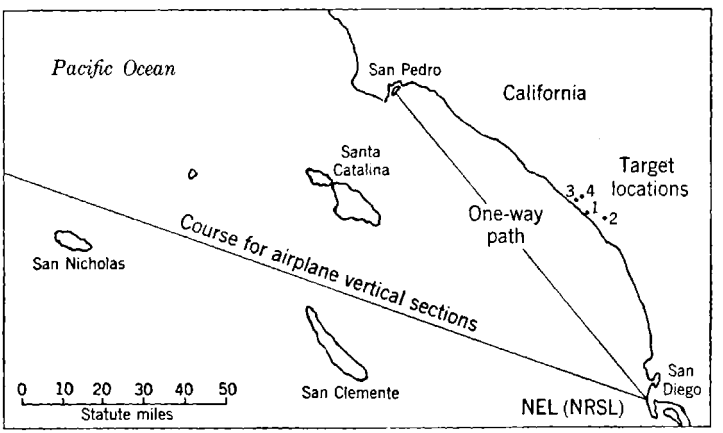


FIG. 4·21.—Southern California coast and location of the NEL (NRSL) propagation experiments.

One-way transmission was carried out from June 1944 to July 1945 between a transmitting terminal 100 ft above sea level at San Pedro and a receiving terminal 100 ft high at Point Loma, San Diego, a distance of 92 statute miles. The path is shown in Fig. 4·21. Wavelengths of 6 m, 3 m, and 55 cm were employed, and the combination of path geometry and equipment was such that standard signal level was well below minimum detectable signal on all three circuits. Meteorological soundings were made by captive balloons and aircraft carrying modified versions of the wired sonde.[3]

[1] S. T. Stephenson, "Some Measurements of Humidity, Temperature and Refractive Index Gradients Immediately Above the Ocean," NRSL Report No. WP-22, September 1945.

[2] See Sec. 4·22 for further data from the same location, obtained with one airborne terminal.

[3] L. J. Anderson, "Wired Sonde Equipment for High-altitude Soundings," NRSL Report No. WP-16, Nov. 17, 1944; also L. J. Anderson, S. T. Stephenson, and A. P. D. Stokes, "Improvements in U.S. NRSL Meteorological Sounding Equipment," NRSL Report No. WP-21, July 3, 1945.

The meteorological situation in this location changes slowly in both time and space in comparison with that in Massachusetts Bay or the Irish Sea; hence relatively fewer soundings and less detailed meteorological analysis are necessary for study of the transmission phenomena. The relevant facts available to date are as follows:

1. For several months (June through October 1944), an elevated temperature inversion persisted. The air below the inversion was

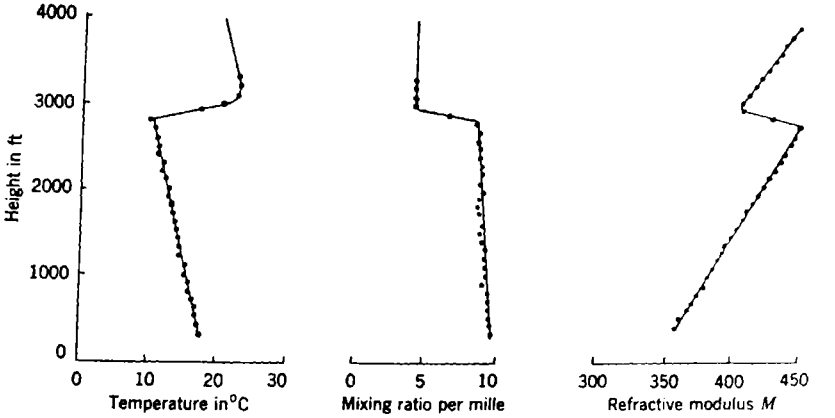


FIG. 4·22.—A sounding 45 miles northwest of San Diego, Calif., on Sept. 29, 1944, showing a typical high M-inversion.

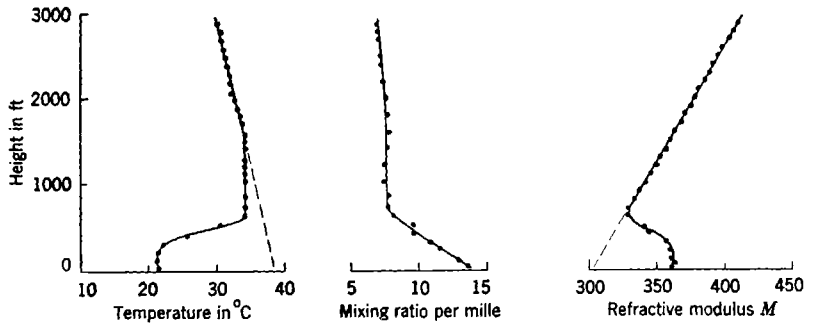


FIG. 4·23.—A sounding 38 miles northwest of San Diego, Calif., on Sept. 7, 1944, showing a typical low M-inversion.

moist and well mixed; that above it was warmer and much drier. The rapid change in index of refraction in a thin stratum between these contrasting types of air caused a marked M-inversion, as illustrated in Figs. 4·22 and 4·23. The change in index n was about 50 to 60 $\times$ 10^{-6} on the average. The principal variation in this structure was that of height of the inversion, which varied from near the surface to around 4000 ft. There was a general tendency

toward a diurnal variation, with a maximum height in the morning at 0800 and a minimum at 1600 hours. Aircraft measurements on a path out to sea showed that surfaces of constant index were sometimes warped in shape, as shown in Fig. 4·24.

2. The thickness of the inversion layer and the strength of the inversion also varied considerably, but the parameter most easily correlated

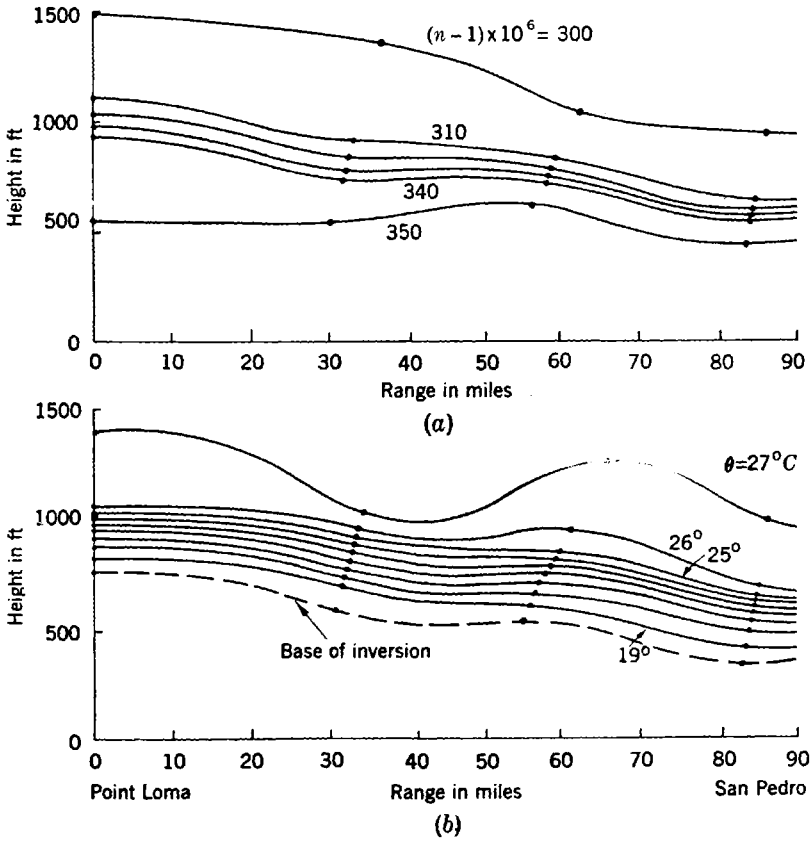


FIG. 4·24.—Atmospheric cross section on the San Diego–San Pedro transmission path showing (*a*) lines of constant $(n-1) \times 10^6$ and (*b*) lines of constant potential temperature.

with the radio transmission was the height of the base of the inversion, as shown in Fig. 4·25. The 55-cm signal fell below the minimum detectable level as the inversion height increased above about 3000 ft, but the 6- and 3-m signals were affected to a much smaller extent. All signals fell below minimum detectable level when frontal action destroyed the inversion, however, indicating that the

layer was responsible for the abnormally strong signal far below the horizon.

3. When the M-inversion was high and did not "overhang," that is, when the minimum value of M in the inversion was greater than that at the transmitter, the average signal level on all these frequencies was above the standard level and the signal was unsteady. Figure 4·26 shows a section of signal record for this case. The signal on 55 cm was barely detectable at such times but was apparently somewhat above standard.

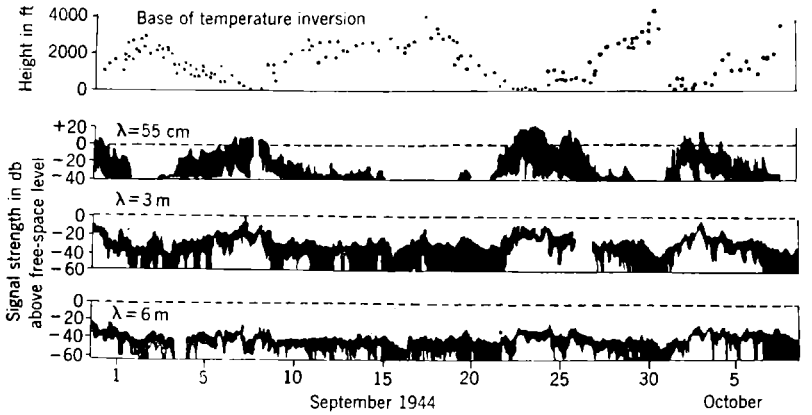


FIG. 4·25.—Correlation of the height of the base of the temperature inversion with signal levels and fading ranges on the San Diego–San Pedro transmission path.

4. When the M-inversion was low, the average signal on all three circuits increased markedly, as shown in Fig. 4·27, the highest frequency rising by the greatest amount; the signal became much steadier on the two lower frequencies.
5. Under all circumstances the signal levels on the two lower frequencies remained below the free-space value, but on the highest frequency it exceeded this value by 10 db or more.

These facts have not yet been explained on a rigorous theoretical basis; presumably such an explanation would require, as a starting point, evaluation of the integral in Eq. (2·172) for M-profiles with elevated irregularities. The ray theory cannot supply information on field strength below the horizon for this case any more than it can for a standard M-profile. The oversimple criterion for existence or nonexistence of "trapping" depending upon whether or not the M-inversion overhangs the transmitter is also completely inadequate. Instead, the concept of "trapping" requires precise definition; this can probably best be done by adoption of a convention regarding the amount of decrease in the attenuation of the transmission modes given by the factor $\exp(iA_m X)$ of Eq. 2·486a. This factor is independent of height and expresses the effect of M-profile and

wavelength, which the ray theory cannot do. Furthermore, the implication is, when the simple concept of trapping based on ray tracing is applied, that the signal strength within the layer does not increase noticeably until rays begin to be turned downward. This is not true, of course, since leaking modes affect the signal strength in or near the nonstandard layer,

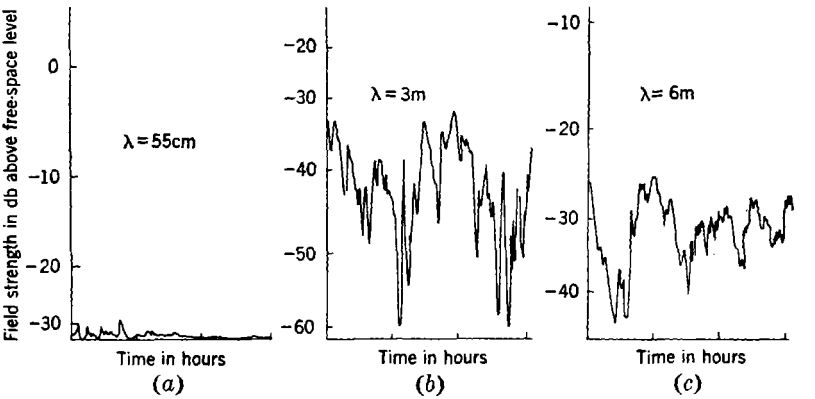


FIG. 4·26.—Typical signal record on the San Diego–San Pedro path when the *M*-inversion was high.

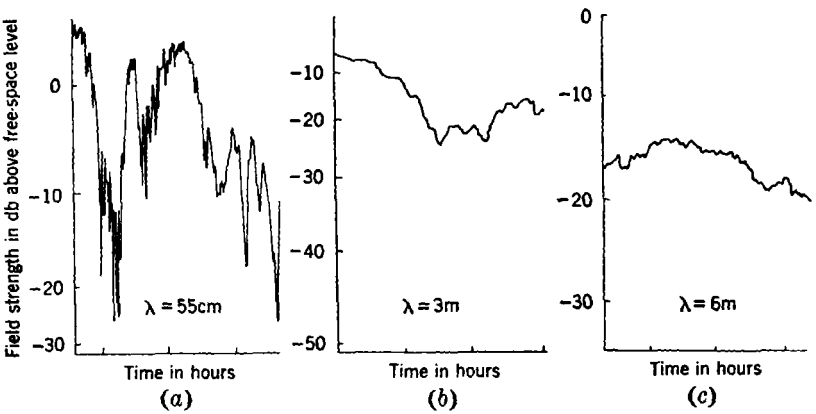


FIG. 4·27.—Typical signal record on the San Diego–San Pedro path when the *M*-inversion was low.

and the increase in signal strength as the layer moves up from the surface is gradual rather than abrupt; the rate of increase is influenced by the magnitude of the irregularity, its height *in natural units*,[1] and heights of transmitter and receiver expressed similarly.

[1] The italics are used to emphasize the wavelength-layer height dependence, which is responsible for the fact that a change in layer height from, say, 100 to 200 ft can produce a very large change in the degree of trapping of wavelengths in the centimeter region but relatively little in the meter region (see Secs. 4·4 and 2·19).

When the elevated layer in this case dropped to near the surface, it was strong enough to trap the 55-cm signal strongly, while the 3-m signal was probably in the region of incipient trapping, and the 6-m signal was affected relatively little.

The wave equation has been applied to obtain exact expressions for the transmission and reflection coefficients of plane layers of finite thickness in which the index of refraction varies with position.[1] The atmosphere in vertical section is divided into three plane strata, the first with constant index of refraction, the second with the index changing continuously in a way permitting analysis with simple functions, and the third with constant index which may differ from that in the first region. The modified-index concept is discarded. A plane wave in the first medium is considered to be incident on the layer at a small grazing angle; and if this angle is sufficiently small, a substantial fraction of the incident energy will be reflected back into the first medium. The amount of reflection depends upon the grazing angle, the ratio of thickness of the transition layer to the wavelength, and the form of variation of index through the layer.

The results of this analysis have been used to explain the characteristics of the signal at the receivers at San Pedro far below the horizon.[2] Qualitative agreement is claimed between the predicted and observed variation of field strength as the height of the layer varies, although there is no check on absolute field strengths.

The application of this solution for the plane layer to the actual physical situation cannot be considered entirely satisfactory, as several factors are ignored and no simple way of evaluating their importance has yet appeared. It is not clear that the assumptions of a plane source wave and plane reflecting layer are reasonably valid for a spherical system. The boundary conditions are ignored at the surface of the earth but are fulfilled at the reflecting layer and at infinity. For a qualitative explanation of why extra energy can be propagated below the horizon this method of approach is useful, but it should not be relied upon to explain the physical facts in detail. The formulation in terms of a series of transmission modes obtained from the fundamental theorem of Eq. (2·172) affords a mathematically difficult but nevertheless more trustworthy procedure.[3] A

[1] J. B. Smyth, "Transmission of Plane Waves through a Single Stratum Separating Two Media," Part I, NRSL Report No. WP-9, Dec. 22, 1943; Part II, NRSL Report No. WP-13, June 23, 1944.

[2] L. J. Anderson *et al.*, "Atmospheric Refraction, a Preliminary Quantitative Investigation," NRSL Report No. WP-17, Dec. 28, 1944; L. G. Trolese, "One-way Transmission Experiments over the Sea between Los Angeles and San Diego," in *Radio Wave Propagation*, Academic Press, New York, 1949.

[3] T. Pearcey has suggested that evaluation of this integral or its equivalent might lead to a form of solution similar to that described above for the layer; but until this is done, there is no direct way of testing the validity of the plane-wave-plane-layer

considerable number of modes may be important, each leaking out of the elevated duct at a rate depending on the characteristics of the inversion and each being excited more or less efficiently, depending upon the position of the transmitter with respect to the height-gain curve of the mode.

TRANSMISSION OVER AN INLAND LAKE

BY DONALD E. KERR

4·10. Flathead Lake.—Some of the first overwater transmission measurements to be correlated in detail with low-level soundings were made from Sept. 9–19, 1943, over Flathead Lake, Montana, by Washington State College.[1] The lake extends in a north–south direction, is 25 miles long and 6 to 14 miles wide. It is located in arid country about 2900 ft above sea level. A radio transmission circuit on 10 cm was set up from a transmitting terminal 15 ft above lake level to receiving terminals approximately 20 miles distant across the length of the lake, at heights of 16, 57, and 180 ft. The 180-ft location was used for 3 days; the 57-ft terminal for the remainder of the period; and the 16-ft terminal for the entire period. The 16- and 180-ft combination of terminal heights gave a just-optical path, while the lower terminals were slightly below the horizon.

Low-level soundings were made up to heights of about 300 ft at intervals of 2 to 4 hr. The prevailing surface winds were southwest, and the soundings were made within 3 miles of the eastern shore, thus giving an overwater trajectory of several miles to the sounding point.

The experiment revealed the following facts:

1. The M-profiles ranged from near standard to strong surface ducts. The near-standard profiles tended to occur between 0200 and 1200 hours when conditions were calm; the ducts occurred in the afternoon and evening not only under calm conditions but with surface wind speeds up to 12 or 15 mph. Considerable lack of uniformity was apparent under nonstandard conditions.
2. When the M-profile was approximately standard, the signal variations at the two receivers at 16 and 57 or 180 ft were in general similar, and the average level was not far from the standard level.
3. When M-inversions occurred, the average signal level at both receivers rose to near the free-space value, reaching a maximum of about 10 db above free space, with the signal changing to the "strong-trapping" type with scintillations of varying amplitude superposed on rollers up to 50 db in amplitude.

method. There is some difference of opinion in this matter, and the statements given here represent the views held by the authors of this volume.

[1] P. A. Anderson, K. E. Fitzsimmons, and S. T. Stephenson, "The Propagation of 10 cm Waves over an Inland Lake. Correlation with Meteorological Soundings," NDRC Report No. 14-212, Washington State College, Nov. 12, 1943.

4. When the average signal was near free-space level, the fades at the two receivers occurred independently, at times fading out of phase. Short, deep fades occurred more frequently at the higher terminal, dropping to 30 db or more below free-space level. The total range of signal variation was 50 to 55 db.

These results appear to be qualitatively in agreement with those obtained on other comparable overwater circuits, although detailed comparisons cannot be made. There is no disagreement with the general conclusions given in Sec. 4·3 concerning the correlation between M-profile near the surface and signal behavior; the soundings support the plausibility of the expectation of that time that the dry, warm air from the land should be modified by cooling (or heating) and evaporation from the lake surface by processes similar to those occurring over the ocean. Presumably there should at times have been an evaporation duct, but the data published thus far do not permit detailed analysis. These measurements were particularly interesting because they aided in the early development of the understanding of overwater transmission. All other experiments at that time had been done over the ocean.[1]

ONE-WAY TRANSMISSION OVER LAND

BY DONALD E. KERR AND PEARL J. RUBENSTEIN

In contrast to the situation in most of the overwater transmission experiments described in the preceding pages, meteorological soundings and related analysis have played a very small part in most of the overland experiments made thus far. Therefore no basis for satisfactory quantitative explanation of the observed phenomena is available; instead, the observed facts must at present be considered largely empirical. Consequently, the overland experiments will not be described individually in detail, but rather the general features that appear to be common to all will be enumerated, and the distinguishing features of each discussed.

4·11. Early Experiments.—Since about 1931 various aspects of ultra-high-frequency propagation over land have been under investigation, the effects of tropospheric refraction being one of the principal features to receive attention. Much of this work has been carried out by the Bell Telephone Laboratories over the gently rolling terrain of New Jersey.[2]

[1] For example, see P. J. Rubenstein, I. Katz, L. J. Neelands, and R. M. Mitchell, "Microwave Transmission over Water and Land under Various Meteorological Conditions," RL Report No. 547, July 13, 1944.

[2] J. C. Schelleng, C. R. Burrows, and E. B. Ferrell, "Ultra-short-wave Propagation," *Proc. IRE*, **21,** 427 (1933); C. R. Burrows, A. Decino, and L. E. Hunt, "Ultra-short-wave Propagation over Land," *ibid.*, **33,** 1507 (1935); and "Stability of Two-meter Waves," *ibid.*, **26,** 516 (1938); C. R. Englund, A. B. Crawford, and W. W. Mumford, "Ultra-short-wave Transmission over a 39-mile Optical Path," *ibid.*, **28,** 360 (1940).

Over a period of 10 years tests were made at wavelengths from about 17 m to 30 cm on circuits varying in length from a few hundred feet to 60 miles. The first tests were designed to give information on the optical reflecting and diffracting properties of the terrain; the observation of fading led to further investigation and to the conclusion that tropospheric refraction is responsible for the fading. A series of measurements made throughout the year 1936 on a wavelength of 2 m over a 37-mile nonoptical path showed the presence of a diurnal effect, with fading and high average signal level at night as contrasted to steady signal near midday. This diurnal effect was more pronounced in summer than in winter.

Attempts were made to correlate transmission characteristics with atmospheric stratification in a later 2-year study in which wavelengths of 2 and 4 m were transmitted simultaneously on a 39-mile optical path. Recordings of temperature, relative humidity, and pressure at one transmission terminal, in combination with upper-air data from U.S. Weather Bureau aircraft soundings, revealed diurnal changes in refractive index at the surface and indicated the possible existence of appreciable variations in index gradient. The airplane meteorographs used for the soundings had too great a lag coefficient and were not sufficiently accurate to permit making soundings of the type now considered necessary.

As the experiments progressed, certain qualitative features of the transmission became noticeable. In most of the later tests these features reappeared with some modifications, usually becoming more pronounced as the wavelength was decreased. These features were

1. The occurrence of a diurnal cycle, usually including two distinct signal types. For several hours near the middle of the day the signal was relatively steady, but at night (on calm, clear nights particularly) it became disturbed, with marked fading; at the same time the average signal level increased.
2. Much more pronounced diurnal effect in summer than in winter. Practically all of the very deep fades occurred during the summer.
3. Almost entire disappearance of fading in rainy, stormy, or windy weather.
4. Marked similarity between major variations in signal on the 2- and 4-m wavelengths, with difference in fine details of fading on the two wavelengths. Deep fades did not occur simultaneously.
5. Little difference between the behavior of vertically and horizontally polarized radiation.
6. Increase of range of variation of signal strength and amount of scintillation or fading as one terminal was decreased in altitude and shifted slightly so as to be screened by a range of hills.
7. Range of variation of field strength from about 13 db above free-space field to about 20 db below for the 2-m signal on the optical path and slightly less for the 4-m signal. The average midday levels

showed a slow seasonal drift of 4 or 5 db about the free-space level, being above in summer and below in fall and winter.[1]

The plot of signal as a function of time in Fig. 4·28 illustrates the diurnal and seasonal effects and the similarity in gross aspects of the signals on the two wavelengths.

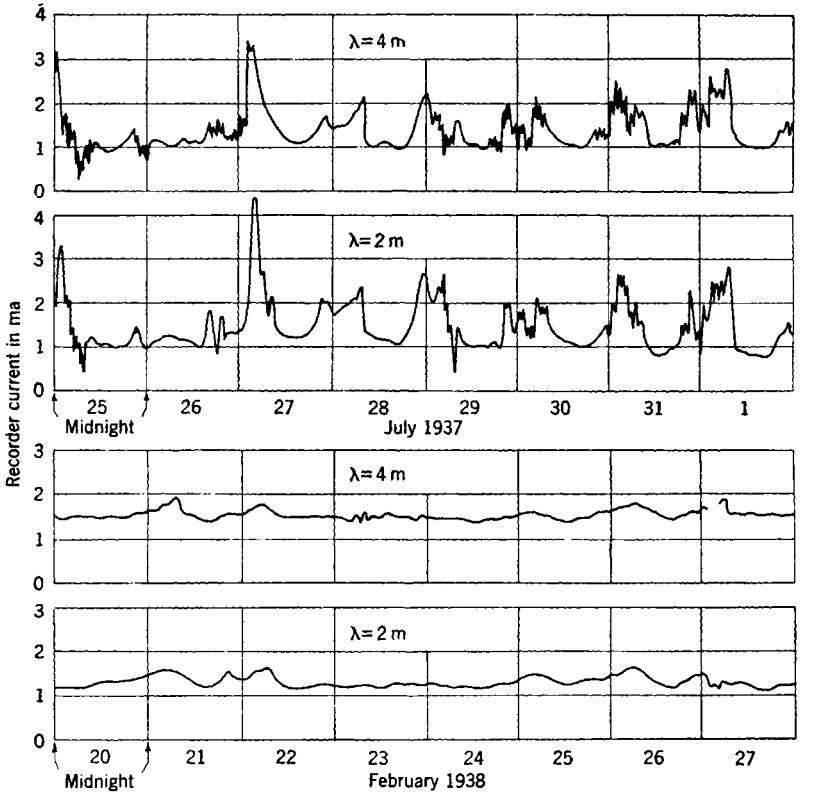


FIG. 4·28.—Diurnal and seasonal effects observed on wavelengths of 2 and 4 m on early BTL transmission path.

Unpublished data for wavelengths of 2 m and 30 cm taken on this same optical path in autumn weather show that the diurnal effect noted in Item 1 occurred again, except that the 2-m average signal occasionally dropped to a low level during nocturnal fading. On the other hand, such a decrease in average signal level was usual rather than exceptional on 30 cm during fading periods. The fading range on 30 cm was appreciably

[1] Seasonal drift of midday levels has not been noticed by those few observers in a position to make such observations, except for RCA Laboratories on the paths described in the next section.

greater than on 2 m, and small scintillations sometimes occurred superposed on the steady midday 30-cm signal.

The observed signal variability was interpreted as being caused by one of the two following mechanisms.[1] (1) Changes in the refractive-index gradient would produce shifts in the interference pattern of the direct and reflected rays. Calculations of this effect were carried out by introducing a variable effective earth radius (see Sec. 2·4). (2) In the presence of an elevated reflecting layer, or "air-mass boundary," a third ray reflected from this boundary might contribute to the interference.

From 1934 to 1937 transmission tests in the wavelength region from 5 to 8 m were carried out by Ross A. Hull,[2] with the aid of other radio amateurs, the U.S. Weather Bureau, and Harvard University. The death of Mr. Hull prevented completion of the work, but the preliminary reports[3] and a recent summary[4] indicate that the tests yielded considerable valuable information of a qualitative nature. Hull appears to have appreciated early in the course of the work the importance of both atmospheric water-vapor content and temperature inversions but lacked the means of expressing their effects quantitatively. He studied particularly the characteristics of transmission between Boston, Mass., and Hartford, Conn., a 100-mile nonoptical path with intervening ridges 1000 ft high. Over this path the 5-m signals showed several of the same characteristics that now appear to be common to overland transmission at the shorter wavelengths. The signal at times approached free-space level; strong diurnal and seasonal variations were apparent, with strong signals prevalent at night and in the summer; a definite relationship was noticed between types of signal fading and the atmospheric structure. It was noticed in particular that strong signal with deep fades tended to occur just before the onset of precipitation resulting from overrunning tropical air. Aircraft soundings revealed gross departures from atmospheric homogeneity, such as pronounced nocturnal surface inversions or elevated inversions associated with overrunning tropical air, and provided the basis for correlations with signal strength variations. These correlations could not be made in detail, but it appears that very large effects were obtained from both types of inversions.

Despite the lack of completeness and the qualitative nature of the results, Hull's work was remarkable, and it is indeed unfortunate that it could not be continued. Some of the questions raised by his results are beyond the scope of most of the other work described in this book and serve to emphasize that a satisfactory over-all understanding of refractive

[1] C. R. Englund, A. B. Crawford, and W. W. Mumford, *Bell System Tech. Jour.*, **18,** 489 (1938) and *Proc. IRE*, **28,** 360 (1940).

[2] Then editor of *QST*, the official publication of the American Radio Relay League

[3] R. A. Hull, *QST*, **19,** 3 (1935), and **21,** 16 (1937).

[4] A. W. Friend, *Proc. IRE*, **33,** 358 (1945).

effects certainly cannot be obtained until the gap has been bridged between wavelengths of several meters and a few centimeters and a satisfactory means obtained for estimating the effects of both surface and elevated M-inversions.

4·12. Summary of General Characteristics.—In this section we shall discuss the characteristics that appear to be common to the overland transmission tests carried out from 1942 to 1945. These tests have included wavelengths from 3 cm to about 7 m and terrain ranging from extremely flat desert to mountains. The principal physical features of the transmission paths are shown in Table 4·2. The terminal heights indicated are above mean sea level instead of above surrounding terrain (with the exceptions noted), for it is usually impossible to determine a meaningful terminal height above rough terrain. Probably the best simple way of describing the transmission path geometry is by means of a profile prepared from topographic maps, but it is not feasible to include the profile of each path here.

Some transmission characteristics of the various circuits appear to recur consistently in all cases given here; others are contradictory or cannot be compared. Space limitations prevent discussion of results for each circuit individually, but the following list presents those characteristics which appear to be broadly representative of the experiments listed in the table. All of these characteristics did not appear on every circuit; usually on any circuit only a fraction of those listed were prevalent (or were reported). The specific disagreements with this list and additional features that cannot be compared directly are discussed in the next section.

The broad features common to all the circuits are the following:

1. On optical and near-optical circuits the signal was steady at a level near the standard level in the early afternoon (see Fig. 4·29*a*). On the fully optical circuits this midday signal level for the microwave circuits was close to the free-space level (an indication of poor specular reflection from rough ground at the shorter wavelengths).
2. Pronounced fading, accompanied by deviation of the average signal level from the daytime value, occurred on *clear, calm nights*. On optical paths the *average* level varied between -12 and $+5$ db,[1] and a variety of fading types was observed (see Fig. 4·29*b*, *c*, and *d*). On nonoptical circuits, on the other hand, the average level invariably increased, sometimes reaching the free-space value and occasionally higher, but the fading range varied greatly from one circuit to another.
3. The time of day at which transitions from steady signal to variable or the reverse occurred was subject to large variations, but the only time when transmission was almost certain to be steady was in the

[1] The reference level is always the free-space level unless stated otherwise.

TABLE 4·2.—OVERLAND TRANSMISSION CIRCUITS

Terminal locations	Path length, miles	λ, cm	Standard value of 20 log F	Heights of terminals, ft†		Nature of terrain	Degree of obstruction	Period of operation	Organization and reference number
Vicinity of Mt. Spokane, Wash.	13.5	10	•	5250	2400	Mountainous	Optical	3 weeks, winter	WSC: 1, 2
	52	10		5250	2630		Optical	8 weeks, winter, spring, summer	
	92	10		5250	2550		One terminal 300 ft below tangent ray	5 days, summer	
	111	10		5250	3450		One terminal 100 ft above tangent ray	10 days, early spring	
	111	10		5250	3650		One terminal 300 ft above tangent ray	9 days, early spring	
	112	10		5250	3950		Obstructed by mountain	8 days, early spring All 1942–1943	
Wembley, Haselmere, England	38	9	~ − 30	725	170	Urban at one end; wooded and hilly at the other. Across Thames Valley.	Obstructed near one terminal	March 1943–July 1946	GEC (England) ASE, NPL: 3, 4
New York City, Hauppauge, N.Y.	42.5	665	− 21	1300	280	Gently rolling farm land and urban territory	Optical	March 1943–March 1944	RCA: 5
		63	− 4		298				
		11	− 2		288				
New York City, Riverhead, N.Y.	70	663	− 35	1300	154	Gently rolling farm land and urban territory	Nonoptical	March 1943–March 1944	RCA: 5
		63	− 32		145				
		11	− 50 to − 60		87				
New York City, Neshanic, N.J.	41	3	0	500	600	Wooded, industrial and urban, partly across and near water	Ray clearance of 50 to 100 ft at hill near midpoint	July 1943–March 1945‡	BTL: 6
		6.5							
		10							
Great Blue Hill, Mt. Wachusett, Mass.	45	10	~ − 2	650	2000	Wooded and urban, hilly	Ray ½° above horizontal; 600 ft clearance	Sept. 29–Oct. 16, 1943	RL: 7
Gila Bend, Datelan, Ariz.	47	9	< − 60	53	35	Flat desert with general rise of 400 ft	Nonoptical	Dec. 13–20, 1944	NEL (NRSL): 8
Blue Ridge Lookout, Mt. Hamilton, Calif.	170	6	•	5860	4300	Across San Joaquin Valley	Optical	Several weeks, summer 1945	OCSigO: 9
Blue Ridge Lookout, Mt. Frazier, Calif.	106	6	•	8000	5800	Mountainous	Optical	2 days, summer 1945	OCSigO: 9

*Information not known or not available. †Terminal heights are above mean sea level except for the Gila Bend—Datelan circuit, for which the heights are above the flat terrain near the respective terminals. ‡Reference 6 covers the period Aug. 1, 1943, to Aug. 1, 1944.

1. P. A. Anderson *et al.*, "The Propagation of 10-cm Waves on a 52-mile Optical Path over Land. The Correlation of Signal Patterns and Radiosonde Data," NDRC Report No. 14-151, Washington State College, June 10, 1943. 2. P. A. Anderson *et al.*, "The Propagation of 10-cm Waves over Land Paths of 14, 42, and 112 Miles," Report No. 14-202, Washington State College, Oct. 26, 1943. 3. A. C. Stickland (NPL) and R. L. Smith-Rose (NPL), "Centimetre Wave Propagation over Land. A Preliminary Study of the Field Strength Records between March and September 1943," DSIR Report No. RRB/S13 Nov. 15, 1943; *Jour. IEE*, **93** III *A*, 98 (1946); *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London, 1946. 4. A. C. Stickland (NPL) and R. W. Hatcher (Meteorol Office), "Centimetre Propagation over Land. A Study of Field Strength Records Obtained during the Year 1943–1944," DSIR Report No. RRB/S18, May 11, 1944; see also Reference 3. 5. G. S. Wickizer and A. M. Braaten, *Proc. IRE*, **35**, 670 (1947). 6. A. L. Durkee, *Proc. IRE*, **36**, 197 (1948). 7. P. J. Rubenstein *et al.*, "Microwave Transmission over Water and Land under Various Meteorological Conditions," RL Report No. 547, July 13, 1944. 8. L. J. Anderson *et al.*, "Atmospheric Refraction under Conditions of a Radiation Inversion," NRSL Report No. WP-19, Apr. 21. 1945; Footnote 1, page 342. 9. T. J. Carroll, "Complementary Diversity Reception of Microwaves," OC Sig O Report No. RPS-4, Jan. 3, 1946; R. Bateman, *Proc. IRE*, **34**, 662 (1946).

early afternoon for a very few hours (except for Item 4). The periods of transition between fading and steady signal were usually relatively short—sometimes 30 min but generally somewhat longer. In the afternoon the transition was likely to occur some hours before sunset, but in the morning it usually occurred long after sunrise.[1]

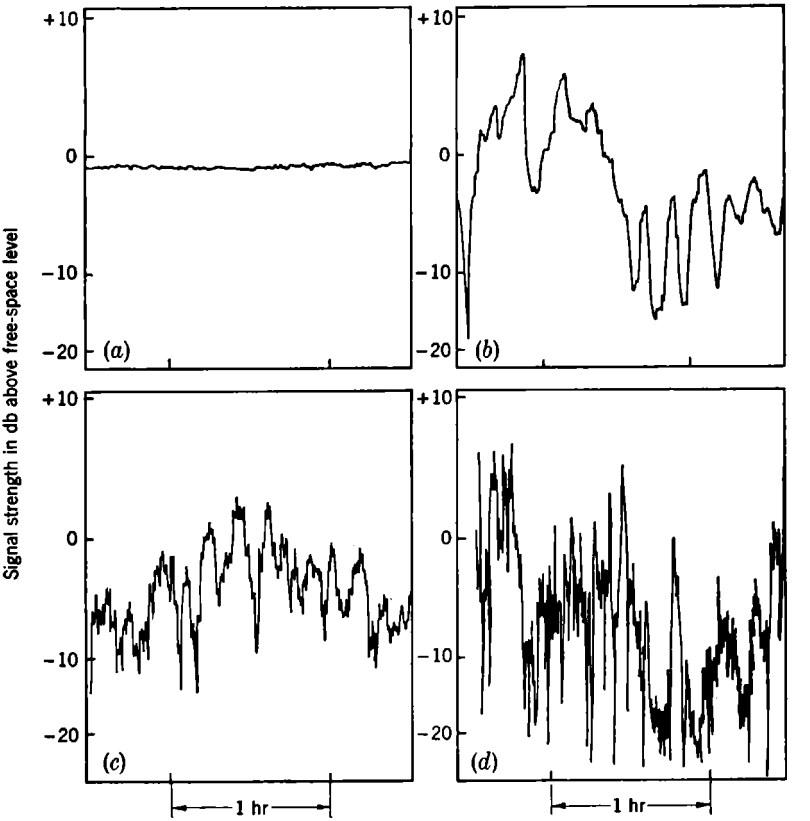


FIG. 4·29.—Typical signal records in overland transmission on an optical or near-optical path. The abscissa is time: (*a*) early afternoon; (*b,c,d*) clear nights. (Blue Hill to Mt. Wachusett circuit.)

4. In cloudy, stormy weather, microwave transmission, both day and night, was standard and steady, as mentioned in Item 1.[2]

[1] The measurements on the Arizona desert showed a considerably greater degree of regularity than any others reported (see Fig. 4·33). On this circuit the fading tended to be fairly slow, as indicated in Fig. 4·29*b*. For a more recent report see J. P. Day and L. G. Trolese, *Proc. IRE*, **38**, 165 (1950).

[2] In the 5- to 8-m band, however, Hull's data show cases of strong signal associated with the onset of precipitation. This phenomenon has also been reported on 9 cm. See E. C. S. Megaw, *Jour. IEE*, **93**, III *A*, 79 (1946).

5. The diurnal change in average signal level and in fading range from day to night was more marked in summer than in winter.
6. The fading range generally increased when one of the terminals was lowered or shifted behind an obstacle.
7. Greatest variability in signal level was observed on the shortest wavelengths.
8. The strongest signal observed was about +15 db, but in most cases it was between 0 and +10 db. The depth of fading varied greatly with the wavelength and the geometry of the circuit but was frequently as large as 40 db.
9. Horizontally and vertically polarized radiation showed similar behavior.

Meteorological soundings were available for only the Mt. Spokane and Arizona circuits.[1] Detailed correlation of the radio transmission with meteorological parameters cannot, therefore, be made.

Two general conclusions can be drawn from these investigations.

1. Steady transmission occurs when the air along the transmission path is well mixed, as on sunny afternoons or in stormy, windy weather.[2] The M-profile is then nearly standard, except that under strong insolation the temperature gradient very near the ground may be superadiabatic, causing a slightly substandard M-gradient. (It should be recalled that homogeneous air, as defined in Chap. 3, with an *adiabatic* lapse rate also gives a slightly substandard M-gradient.)
2. Signal fading occurs in the presence of irregularities in the M-profile. On nonoptical circuits a marked change in average signal level invariably accompanies these irregularities, but for optical circuits such a general rule does not appear to be justified. The irregularities in the M-profile may occur either at the surface (frequently as a result of radiation cooling) or at some altitude that cannot be specified on the basis of existing information. (This point is discussed further in Sec. 4·13. See also Sec. 3·21.) The humidity gradient can, of course, be very important, as well as the temperature gradient. In particular, a positive vapor-pressure gradient through the inversion can cause a substandard M-layer, as shown in Fig. 3·49.

4·13. Additional Observations.—Because of the wide range of physical characteristics and of the periods of operation that mark the circuits described in Table 4·2, intercomparison of the data is often difficult or meaningless. This diversity of conditions, moreover, gives rise to some

[1] See the following section for a description of the soundings.

[2] The accuracy of this statement appears to depend upon the degree of obstruction of the transmission path. It is apparently correct for optical and near-optical paths, but is not necessarily correct for highly nonoptical paths for which the standard diffraction loss is very large. See the following sections for further discussion.

apparently unrelated or even inconsistent observations. These points which appear to be reasonably well defined are discussed below.

A qualitative picture of the range of variation experienced on typical overland circuits can be obtained from Table 4·3, which shows the standard, the maximum, and the minimum field strengths observed on the two RCA circuits over Long Island on three wavelengths. The maximum signal

TABLE 4·3.—SIGNAL LEVELS ON RCA CIRCUITS OVER LONG ISLAND

Wavelength, cm	Receiving terminal location	Field strengths, db above free-space value		
		Standard	Maximum	Minimum
665*	Hauppauge	− 21	− 13	− 29
63	Hauppauge	− 4	+ 10.5	< − 30
11	Hauppauge	− 2	+ 12	< − 25
665*	Riverhead	− 35	> − 18	< − 50
63	Riverhead	− 31 to − 35	+ 10.5	< − 44
11	Riverhead	− 50 to − 60	+ 13	.....

* The 665-cm circuit was in operation only from 1400 to 2200 hours; consequently its value is limited.

level for the shorter wavelengths was 10 to 15 db above the free-space values, but the level on the 6.6-m circuit never exceeded −13 db. The shorter wavelengths had the greater fading range, but the difference in this range between 63 and 11 cm was small. (A similar situation was found on the New York–Neshanic optical path, where the fading ranges on 3, 5.5, and 9 cm were only slightly different.) Table 4·3 also shows the relative conditions on optical and nonoptical paths. (The Hauppauge paths are optical, and the Riverhead paths nonoptical.) The fading range increases as the degree of obstruction of the circuit increases.

On the Long Island circuits the average 11-cm signal level was higher in summer than in winter, both by day and night, as illustrated in Figs. 4·30 and 4·31.

Both the diurnal and seasonal effects are illustrated in Fig. 4·32, which shows the average diurnal variation of the fading range on the 6.5-cm New York–Neshanic circuit in February and August.

On the New York–Neshanic circuit the fading was reported generally to be synchronous at two 6.5-cm terminals spaced 25 ft apart vertically, as well as on 3, 6.5, and 10 cm at a single location. This is in contradiction to the results of 6-cm tests in California, where vertical separation of two receiving antennas operating separate receivers by the calculated maximum-to-minimum spacing in the interference pattern resulted in complementary fading; that is, the signal was strong in one receiver when weak in the other. The difference probably arises from the fact that the reflected wave on the New York–Neshanic circuit was very weak, as it

originated primarily from a narrow ridge screening the region that might otherwise have provided stronger reflection, whereas on the California circuit both direct- and reflected-ray paths were unobstructed and the land reflection was much stronger than usual. The synchronism of the three wavelengths on the BTL circuits is somewhat surprising, as it is

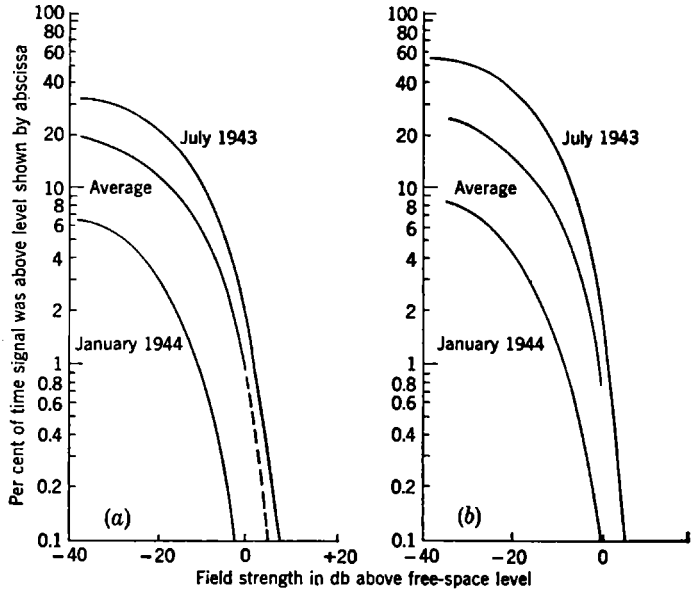


FIG. 4·30.—Seasonal effect on 11-cm New York-Riverhead path, as indicated by percentage of time during which the signal was above the abscissa on the average, during July and during January. (*a*) When all times were considered; (*b*) when only darkness hours were considered.

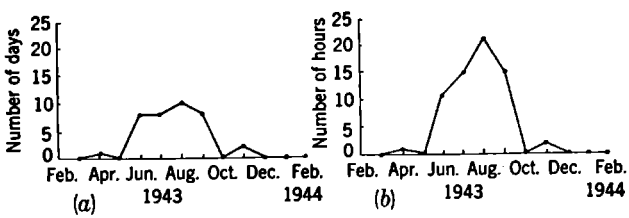


FIG. 4·31.—Seasonal effect on the RCA 11-cm New York-Riverhead path, as indicated by (*a*) number of days and (*b*) number of hours during which the signal was more than 6 db above the free-space level each month.

difficult to visualize a refraction mechanism capable of producing marked fading but insensitive to wavelength.

In further tests of space diversity a comparison was made between the 41-mile New York–Neshanic circuit and the two 20-mile optical links obtained by the use of an additional terminal at the mid-point. The

depth of fading on the long path was greater than that on either of the short ones, and no synchronism was observed.

On several of the circuits, smoke and haze layers were observed to occur simultaneously with fluctuating signal. Such layers generally mark a region of atmospheric stability and, consequently, a tendency toward an M-inversion. The gradual destruction of a nocturnal inversion after sunrise can be followed visually as the bottom of the haze layer moves upward and is eventually dissipated. It has been frequently noticed that the final stages of this process, perhaps the last hour, are marked by extremely large and rapid signal fluctuations, terminating abruptly in the steady midday signal when the layer disappears. These layers also permit visual observation of the well-known "gravitational waves," which were suggested by Helmholtz and have been investigated theoretically at some length.[1] It is frequently suggested that these atmospheric waves may be connected with signal fading,[2] but no experiments have yet been performed to investigate the matter.

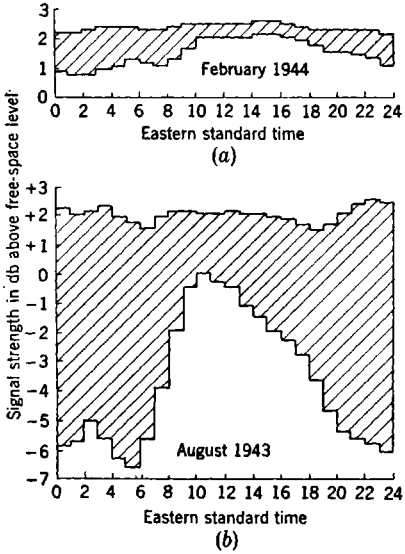


FIG. 4·32.—Maximum and minimum signal levels during 1-hr intervals averaged over the month of (a) February 1944 and (b) August 1943. These figures show the diurnal and seasonal effects observed on the BTL 6.5-cm New York-Neshanic path.

The only overland transmission circuit on which consistent effects of fog have been reported is that in England, between Wembley and Haselmere. In general, the increase of average signal on an initially clear night was arrested by the development of low cloud or fog. In the case of fog formation the signal level then usually dropped below the standard level, strong fading continuing. Although no soundings are available for this circuit, a reasonable qualitative explanation is that radiational cooling

[1] H. Lamb, *Hydrodynamics*, Cambridge, London, 6th ed., 1932, p. 363; B. Haurwitz, *Dynamic Meteorology*, McGraw-Hill, New York, 1941.

[2] For instance, W. M. Elsasser, "Comments on Duct Transmission and Fading," Dec. 5, 1944 (unpublished); J. B. Smyth and L. G. Trolese, "Atmospheric Waves—Fluctuations in High Frequency Radio Waves," NRSL Report No. WP-18, Feb. 1, 1945; R. A. Craig, "Gravitational Waves Associated with Temperature Inversions," in *Radio-Wave Propagation*, Academic Press, New York, 1949.

early in the evening produces a surface temperature inversion with a possible consequent steepening of the base of the M-profile and increase of signal strength. At some stage of this process the surface temperature drops below the dew point, and the resulting positive humidity gradient tends to counteract the effect of the temperature inversion, producing a substandard layer if the humidity gradient is sufficiently steep, with a consequent decrease in average signal level. (See Sec. 3·21 for a more detailed discussion of the possible meteorological processes involved.)

The data from the Wembley-Haselmere circuit have also been examined for the effects of fronts. A statistical summary was made of 40 cases[1] for the period July through September 1944. Twelve warm-type fronts were found to cause a low signal, and one a rise and fall of 9 db in signal level. Of the remainder,

Twenty-seven cold fronts or occlusions gave following results: four a low signal with no rise within five or six hours (but one of which had still warm air at the surface). *Twenty-three gave either a rise in signal during passage or a rise within five or six hours after passage*, including some very marked rises, associated with, in ten cases, a low signal during passage.

From a forecasting point of view during that period the following rules emerge:

1. *Passage of warm fronts* is invariably associated with a low signal, or if the signal is high immediately before, passage will cause it to lower.
2. *Passage of cold fronts* and five or six hours afterwards. It can safely be said that sometime during that period a rise in signal can be expected in about nine cases out of ten, but from purely synoptic considerations it is unsafe to say whether actually during passage over the path or a few hours later.
3. *High wind conditions* (winds of Beaufort F6 or above) normally give a low signal, but there was one exception to this on 24th September, the wind being N.W.F. 8 (fresh gale).
4. *Anticyclonic conditions* give rise to a variable signal with a marked diurnal trend (but there are a few exceptions). Good signal strength is most likely during the afternoon with a gradual fade during the early hours of the morning. The effect would seem to be most marked with light or moderate off-shore winds.

Because the meteorological information obtained by soundings coincident with overland transmission is limited, detailed correlations are impossible. Several investigators attempted to use conventional radiosonde data and aircraft soundings made by aerometeorographs, which, for reasons discussed in Chap. 3, are usually unsatisfactory; nevertheless, the Washington State College group operating near Mt. Spokane found correlations between atmospheric stratification, as revealed by the radiosonde ascents from the Spokane station of the Weather Bureau, and variations in transmission stability. During the later phases of the investigation the radiosonde data were supplemented by low-level sound-

[1] This information was obtained in private communication, through the kindness of R. L. Smith-Rose and others of the Tropospheric Wave Propagation Committee in England.

ings made with early models of the wired sonde. The correlations, although few in number, appeared to indicate consistently that marked fading occurred on all but the 13.5-mile path when the soundings revealed either surface or elevated layers in which the gradient of index of refraction deviated markedly from the standard value.[1] The elevated layers were usually found at heights comparable to the terminal heights and were not noticed to exhibit any diurnal variation, but the surface layers were caused by nocturnal inversions, as discussed in Chap. 3. Despite the very irregular terrain, the effect of the surface inversion appeared to persist over distances of at least 100 miles.

Transmission on the 13.5-mile path was extremely steady, showing only very slow variations of about $\frac{1}{2}$ db. This is not surprising in view of the short path length and the large inclination of the line of sight to the horizontal.[2]

These early experiments by the Washington State College were the first to combine successfully microwave transmission measurements over land with intensive meteorological soundings, which were made first by radiosonde ascents and later by the wired sonde apparatus then under development. Most of the prevailing transmission characteristics described in Sec. 4·12 were found at this time to some degree.

The most detailed program of meteorological measurements was carried out in connection with the transmission tests over the Arizona desert. The meteorological soundings were made with a modified wired sonde (see Sec. 3·28) at both terminals and near the path mid-point at frequent intervals throughout most of each night of the test. They showed that strong temperature and M-inversions formed at the surface at night, increasing in altitude through the night, which produced marked irregularities in the M-profile at heights of several hundred feet.

The radio signal, which was not detected during the daytime on this path, began to rise around sunset. It remained strong throughout the night, reaching a maximum value of about 12 db below the free-space level. The signal again dropped below the detection limit (60 db below free-space level) 2 to 3 hr after sunrise, when the temperature and M-inversions were destroyed by solar heating. This marked diurnal cycle was repeated consistently throughout the course of the tests. Figure 4·33

[1] At the time this work was reported the quantity M was not in use, but the data published in terms of n indicate that strong M-inversions must have occurred in these layers.

[2] Even such short paths are not necessarily free from atmospheric effects, however. Unpublished information from Bell Telephone Laboratories on 10-, 20-, and 30-cm transmission on a 10.5-mile optical path in the summer and fall of 1940 shows that strong fading may occur. In general the fading on 20 cm was slightly more noticeable than on 30 cm, but the fading on 10 cm was very much more noticeable than on 30 cm. The maximum fading range was 20 db on 30 cm and 30 db on 10 cm.

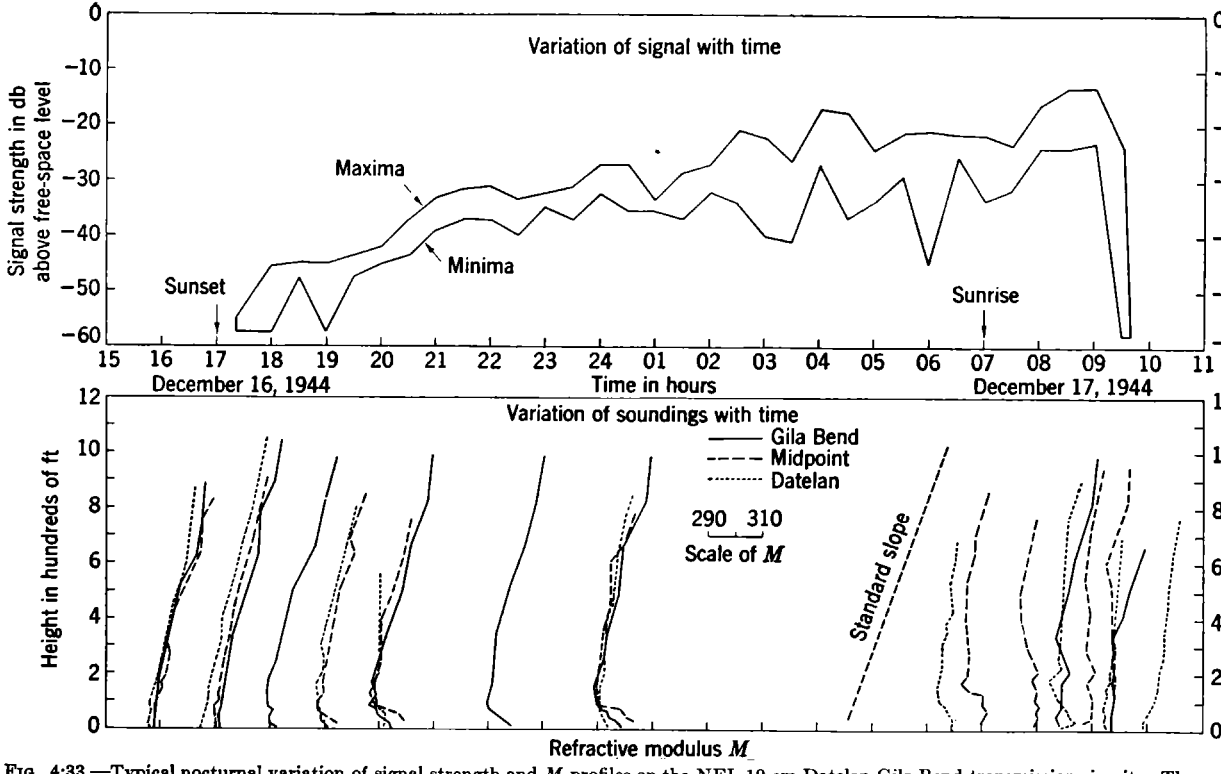


FIG. 4·33.—Typical nocturnal variation of signal strength and M-profiles on the NEL 10-cm Datelan-Gila Bend transmission circuit. The soundings are shown at the approximate times at which they were made.

shows the variation of signal and the accompanying soundings throughout a night with a strong inversion.

Attempts to correlate the measured signal strengths with such parameters as the decrease in M between the transmitter height and top of the duct were unsuccessful.[1] Additional difficulties were possibly introduced by the duration of the soundings (20 to 30 min) as well as departures from the assumed horizontal stratification over the sloping terrain of the circuit.

An interesting and unexplained feature of the desert experiment is that the type of signal fading was noticeably different from that commonly observed on other nonoptical circuits. In the presence of ducts the fading was much slower than that commonly found. It lacked the usual broad maxima and deep, sharp minima, and the fading range was not large.[2]

4·14. Discussion.—Perhaps the most striking similarity among all of the observations for paths that are not too long is the diurnal change in signal behavior, in which the signal changes from a relatively steady type at midday to a type characterized by pronounced fading at night. As shown in Sec. 3·21, the M-profile varies, in general, from approximately

[1] Attempts to correlate signal strengths and such M-profile parameters on the overwater circuits in Massachusetts Bay were also negative. Only when the complete wave solution was used was reasonable agreement obtained between predicted and measured field strengths.

[2] *Editor's note added in proof.* Since the foregoing material was written, further tests have been made by the U.S. Navy Electronics Laboratory at the same location, using a number of wavelengths between 12 m and 1.25 cm. (Personal communication from J. B. Smyth.) The data show a remarkable difference between optical or nearly-optical circuits having small standard diffraction loss and those having large loss. By way of illustration, the first type would be represented by a wavelength of 30 cm and a transmitter height of 58 m, whereas the second would employ a wavelength of 3 cm and a transmitter height of 1 m. Under *midday* conditions, when the measured M-profile was essentially standard, except for a substandard foot because of intense soil heating, the two types of circuits showed a startling difference in performance. The signal on the lower-loss circuit was steady at about the standard level, and variation of the height of the receiving terminal showed a normal height-gain function. For the other type of circuit, however, the signal showed strong scintillation, continually fading rapidly both above and below the average level. Most surprising, the average level was many decibels *above* the level for standard refraction, and the height-gain effect was absent, the signal strength being essentially independent of height. [See J. P. Day and L. G. Trolese, *Proc. IRE*, **38**, 165 (1950).]

This scintillating midday signal, far above standard level, appears to be common on long paths with very low terminal heights. Recordings of signals in the frequency-modulation and somewhat higher frequency bands have now been made for considerable periods over path lengths of over 100 miles by the Federal Communications Commission, the Central Radio Propagation Laboratory of the Bureau of Standards, and other groups. (See K. A. Norton's chapter in *Advances in Electronics*, Vol. 1 Academic Press, New York, 1949.) Presumably this phenomenon is similar to that found over water in Massachusetts Bay (Sec. 4·5) and at Antigua (Sec. 4·20), and is possibly to be explained, in part at least, by scattering from atmospheric irregularities.

standard around midday under the influence of solar heating to either substandard or superstandard near the surface at night as a result of radiation cooling. The irregularities in M-profile undoubtedly vary in height and shape along a transmission path over rough terrain during different parts of the diurnal cycle; hence the observed phenomena represent a combination of "surface trapping" and "reflection from layers" (the quotation marks are used to emphasize that these terms are crudely descriptive in nature rather than significant from a mathematical point of view).

In addition to the fading, the variation of the average signal level upon which the fading is superposed is of some interest. It is usually observed that on nonoptical circuits the average signal level rises when fading occurs,[1] as it does in the overwater case, but on optical circuits it may either rise or fall for considerable periods. At the longer wavelengths the reflection from the ground may be sufficiently strong to introduce a lobe structure that can be distorted by refraction, thus causing some fading, but this explanation is untenable for the shorter wavelengths, where the preference for free-space signal under standard refraction conditions on several wavelengths and direct measurements of reflection coefficient indicate that ground reflection must be essentially absent.

All the experiments that were operated in a way to disclose seasonal variations show essential agreement in the conclusion that the diurnal signal variations are considerably more pronounced in summer than in winter. This is presumably attributable, in part at least, to the high water-vapor content of the air in the summer, which enhances the effectiveness of temperature gradients in producing strong index gradients; at the same time it increases the probability of occurrence of strong humidity gradients, which also affect the index gradients markedly, as pointed out in Sec. 3·21. In addition to the exaggeration of the diurnal fading, an increase in average signal level in the summer was found on the RCA circuits on Long Island, possibly another manifestation of the enhancement of the effects of temperature gradients.

In all overland experiments in which comparisons have been made between vertical and horizontal polarization, no appreciable differences have been observed in transmission stability. This result is not surprising, as those measurements in which reflection from the earth might be expected to play a part (optical circuits on wavelengths of several meters) are usually made at such small grazing angles that the reflection coefficient should not be substantially different for the two polarizations. The theoretical treatment of Chap. 2 predicts no dependence on polarization in the diffraction region, as the second-order terms that were discarded in

[1] This statement is not always true, as may be seen from the transmission data from England, which indicate that substandard layers occur at times, decreasing the average signal level.

formulation of the theory contained the only vehicle for expressing polarization dependence.

The tendency toward decreased stability of transmission as the degree of obstruction of a transmission path is increased is fairly well established. The variation of signal under nonstandard refraction conditions occurs between an upper limit, roughly 10 to 15 db above free-space level, and a lower limit that is often not measurable but is well below the standard signal level.

If the nocturnal cooling first affects the refractive index gradient at the surface (and presumably this would usually be the case on clear nights), surface ducts may form. How their shapes, stability, and rate of growth are affected by hilly slopes and varying types of surface are open questions, as is also the mechanism by which they affect transmission stability.

A general explanation of overland transmission in any but the crudest of theoretical terms is, for several reasons, impossible at present. The boundary conditions for most land surfaces have not been formulated satisfactorily for either the meteorological or the electromagnetic phase of the propagation problem. The usual streams, vegetation, bare earth, buildings, and other irregularities constitute sources (or sinks) of heat and moisture, and calculation of the exchange of these quantities with the atmosphere and the resulting modification of the atmosphere has not been reduced to a practical procedure. As can be seen from the discussion in Chap. 3, even for the vastly simpler case of a smooth sea at constant temperature this problem is amenable only to a crude estimation. The roughness of most land surfaces and lack of knowledge of their conductivity and permittivity preclude application of a reflection coefficient in the customary way (see Sec. 5·1), and no satisfactory procedure has been evolved as a substitute.[1] Finally, the concept of modified index, which has performed valuable service in the analysis of overwater propagation, loses its precise meaning over land. It will be recalled that in the definition of modified index, purely horizontal stratification is assumed; that is, surfaces of constant index of refraction form spherical shells concentric with the earth. This condition is not fulfilled over water as well as one would wish and is even less satisfactorily met over rough land with varying surface properties. The overland problem thus lacks a single unifying parameter comparable to modified index. If the land is very flat and uniform in character, such as that in the Arizona desert, this difficulty may perhaps be somewhat less important.

Further progress in understanding the stability of overland transmission requires emphasis on two main points: (1) an extensive program of detailed

[1] Under trapping conditions the guided waves strike the earth at extremely small angles with the horizontal, with the result that even a rather rough surface may give appreciable specular reflection.

meteorological sounding and analysis under conditions of varying degrees of complexity, designed to obtain basic information on the mechanism of duct formation, and (2) theoretical and experimental study of the electromagnetic features of the problem, including the effects of irregular ground. The simpler corresponding problem for overwater transmission has yielded definitive results only in proportion to the application of such methods.[1]

RADAR TRANSMISSION

It is clear from both theory and experimental data given thus far that nonstandard refraction can produce startling departures from the coverage diagrams discussed in Secs. 2·13 to 2·16. Because radar can explore large volumes in space in a very short time, by comparison with the limited information obtainable with fixed transmission terminals, it might appear at first glance that radar would be the more useful tool in the investigation of refraction. It will become clear from the next five sections and from Chaps. 5 and 6, however, that this is a much too naïve assumption. It is possible to use radar to considerable advantage in studying refraction, but quantitative information comparable to that obtained from one-way transmission is, except under special circumstances, essentially impossible to obtain; on the other hand, some useful information of a qualitative nature can be fairly easily obtained, particularly in conjunction with one-way transmission measurements and a program of meteorological investigation.

To appreciate the reasons for the foregoing statements, let us recall the radar equation [Eq. (2·28)],

$$\frac{P_r}{P_t} = \frac{G^2\lambda^2}{(4\pi)^3}\frac{\sigma F^4}{R^4}.$$

In this equation the target is considered to be concentrated at a *point*, and its radar cross section is given by σ. The term F is also a point function. The value of σ for a real target of finite extent depends upon the distribution of the incident field, that is, upon F, and upon the geometrical features of the target. Near the earth's surface the field incident on an extended target, even for standard refraction, is likely to vary appreciably with position; and if the shape of the target is complicated, it becomes impossible to specify a single value of σ independent of range, radar height, etc. In addition, atmospheric refraction causes the incident field distri-

[1] *Editor's note added in proof.* In view of the new results on the high midday fields on long paths, it appears that a still greater degree of complexity may prevail than has been visualized earlier, and that the M-profile may be only an intermediate step in the study of atmospheric refraction. The practical consequences of the later results may be very important for such applications as television, aircraft, traffic control, and other services in which cochannel interference is a serious problem.

bution to vary with time, resulting in large fluctuations of echo strength. Consequently, radar measurements on random surface targets are by nature rough and require adoption of some kind of arbitrary procedure to make possible comparisons between observations. As it is difficult to estimate "average" signal strengths on most radars with any accuracy, it is usually more useful to measure the peak value of the echo occurring in some chosen time interval, this interval depending upon conditions. In view of the arbitrary method of measurement and the wide variation of refraction effects, then, it is not possible to define a single meaningful value of σ for surface targets such as buildings, cliffs, or ships under conditions of nonstandard refraction.[1]

Calculation of σ for extended targets for nonstandard conditions is in general not feasible, because in addition to the complications introduced by the geometry of the target, the incident field distribution consists of the summation of field from several transmission modes. From Secs. 2·17 to 2·23 it can be seen that this is not a promising prospect.

Lacking the means of describing radar transmission by simple parameters comparable to those of one-way transmission, we are left with only the ratio of P_r/P_t as a phenomenological statement of echo from the target; and as this is devoid of functional relationships, there is no possibility of exploring the detailed nature of the transmission. Even without the detail, there is much to be learned from rough observations, as shown in the following sections.[2]

4·15. New England Coast.[3]—As part of the program of overwater transmission described in Sec. 4·1, experimental 3- and 9-cm radars[4] were set up beside the transmitting terminal tower at Race Point (see Fig. 4·1) with antennas about 50 ft above mean sea level. In addition to random shipping in Massachusetts Bay, coastal targets along the New England coast were available to these radars, the targets being hills, towers, buildings, and a few islands. Local obstructions obscured most near-by land targets, so that except for occasional ships no echoes could be expected between 2 and 20 miles.

The observational program had two parts. A general picture of transmission conditions was obtained from PPI photographs taken hourly; the range and azimuth of the most distant detectable echo were also

[1] The other factors affecting σ in the absence of refraction make determination of σ difficult but somewhat less ambiguous (see Chap. 6).

[2] It is not difficult to visualize an experiment in which the difficulties enumerated above are greatly minimized. If, for instance, measurements were made of echoes from a series of corner reflectors of known cross section, located at various ranges in such a way as to be free from interference from other targets, the results could possibly be interpreted as well as those obtained from one-way transmission measurements.

[3] By Pearl J. Rubenstein.

[4] These special experimental radars are described in Sec. 6·20.

recorded hourly.[1] The echo strengths of four selected overwater targets at ranges of 22, 41, 62, and 73 miles were measured hourly. Only the first of these was visible under standard conditions. These targets were chosen for their isolation and the relative steadiness of their echoes.[2] Absolute measurements of the target echo strengths were made by matching to the echo the pulse on the A-scope from a pulsed signal generator connected to the receiver through a directional coupler. The peak echo value reached in an interval ranging from 30 sec to several minutes was recorded, the time interval being at the observer's discretion.

Because no simple targets were available, standard values of maximum range and of the signal strengths of the selected targets were determined empirically from a preliminary series of measurements.

The patterns of radar echoes observed on the PPI fell into several categories, each associated with a distinctive M-profile; these patterns, shown in Fig. 4·34, are analogous to the signal types observed in one-way transmission.

In Fig. 4·34*a* the PPI is practically devoid of echoes because of the absence of permanent targets between 2 and 20 miles. This is the pattern observed when M-profiles were substandard at the surface and the microwave signal level in one-way transmission was below the standard level. During such periods small ships were occasionally tracked to 10 or 12 miles. Both 3- and 9-cm radars showed symptoms of substandard conditions at the same time.

The standard pattern, established empirically, is shown in Fig. 4·34*b*. A few scattered echoes are seen at ranges near 20 miles. They appear as arcs because of the 6° width of the 9-cm antenna pattern; the corresponding echoes at 3 cm (beamwidth 2°) are essentially dots. The standard patterns at 9 and 3 cm are very similar.

Only a slight improvement over the standard condition is represented in the pattern of Fig. 4·34*c*, which occurred with shallow surface ducts. Maximum ranges on surface targets were extended 10 to 30 per cent; some additional targets appear within the standard maximum range; and the echo signal strengths increase over their standard values. The effects were more pronounced at 3 than at 9 cm.

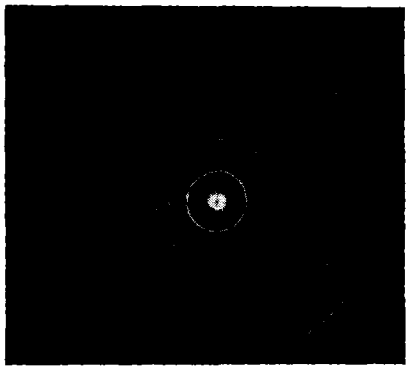
With deep surface ducts the pattern changes radically, as shown in Fig. 4·34*d* (note the change in range scale). Mapping of the coast line becomes fairly complete, and the echo strengths are high. The echoes

[1] Separate recording was necessary because delayed sweeps provided to obtain data at ranges greater than about 140 miles were available only with the A-scopes and not with the PPI's.

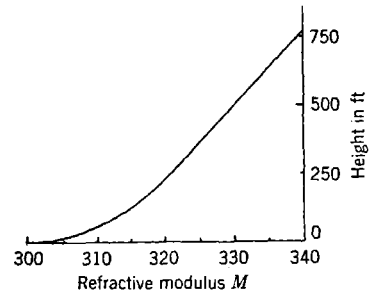
[2] They were in no sense point targets, but the complexity of their echoes was considerably less than for the rest of the coast line, and their physical separation from other prominent targets reduced the identification problem.

are characterized by more severe fluctuations than those normally observed.

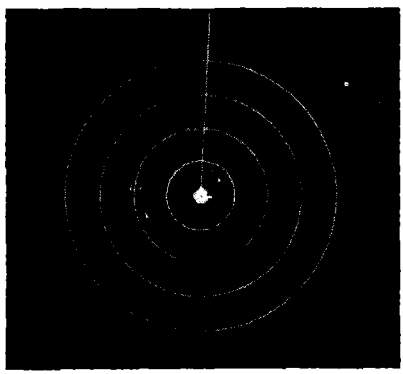
Under these conditions maximum detection ranges at 9 cm were generally greater than those observed simultaneously at 3 cm, as shown in Fig. 4·35. This observation, however, cannot be generally valid in view of the complexity of the transmission. Not only do the transmitted



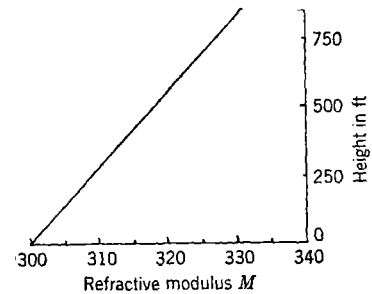
10-mile range marks.



(*a*) Substandard.



10-mile range marks



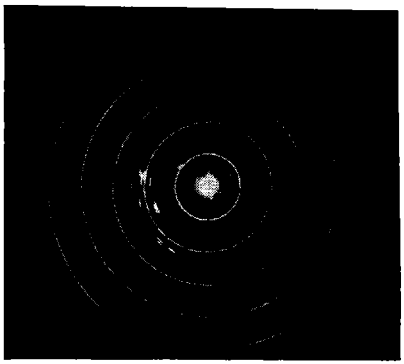
(*b*) Standard.

FIG. 4·34.—PPI patterns and associated M-profiles.

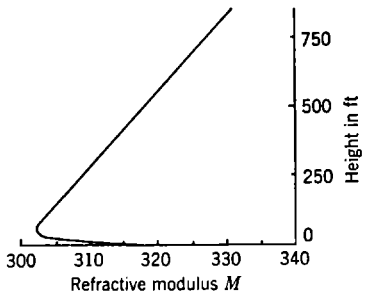
powers and minimum detectable levels of the two radars differ slightly, but the difference in beamwidths is also important in this connection because of the extended nature of the targets. The major factor that cannot be compared in the two cases is, of course, the variation with height and range of the factor F which covers the departures from free-space propagation. At the longest ranges water-vapor and oxygen absorption will, in addition, be noticeable at 3 cm.

An interesting feature of these patterns is the strict limitation of the targets to the coastal zone, although numerous hills and tall buildings a few miles inland form potential targets. This illustrates the delimitation to the water area of these surface ducts formed by cooling of the air from below. The field is apparently attenuated rapidly after it leaves the duct.

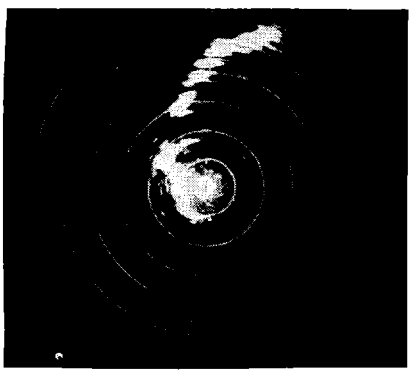
Despite the complexity of the available targets and their nonuniform



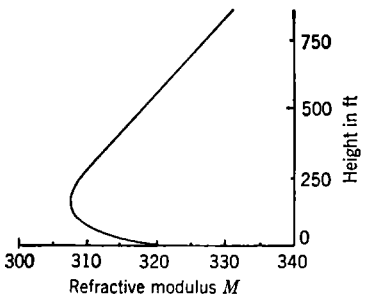
10-mile range marks.



(*c*) Shallow ducts.



20-mile range marks.



(*d*) Deep ducts.

FIG. 4·34.—PPI patterns and associated *M*-profiles.—(*Continued*)

distribution in range, the maximum detection range varies smoothly with increasing duct strength. This is illustrated in Fig. 4·36, in which maximum range is plotted as a function of duct height as obtained from synthetic soundings (see Sec. 4·2.) for cases of surface *M*-inversions. The maximum range observed during the tests was 280 miles, a value determined by the limitations of the available indicator sweeps rather than propagation conditions. The echoes were received from the coasts of

Nova Scotia. The indicated duct heights were determined only for the Massachusetts Bay region, and there is no way of knowing the atmospheric structure at large distances.

Although Fig. 4·36 might be considered a rough means of predicting maximum range on coastal targets if the duct height is known, it should

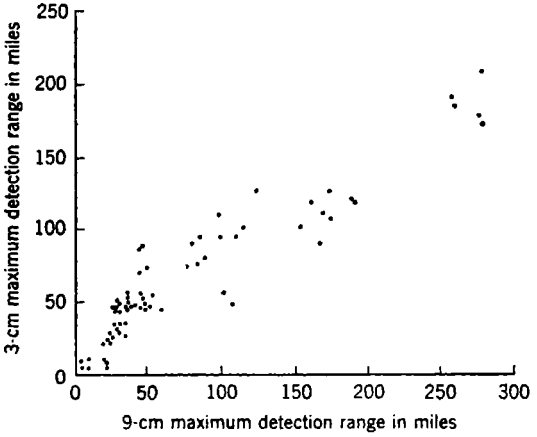


FIG. 4·35.—Comparison of 3- and 9-cm maximum radar ranges observed under conditions of surface ducts.

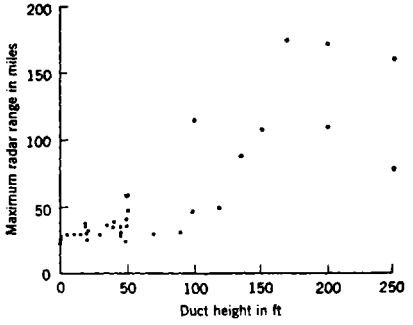


FIG. 4·36.—Maximum observed radar ranges, at 9 cm, as a function of height of surface duct.

be borne in mind that such an empirical calibration is valid only for the specific radar, in its specific site. Under less favorable conditions, such as a less uniform distribution of possible targets, even an extended series of observations may not provide a suitable coverage calibration. (Duct height alone is not, of course, sufficient to specify field distributions, and it is surprising that there is as much regularity as is shown in the figure.)

Two additional patterns of radar echoes were occasionally observed. The first, shown in Fig. 4·37, resembles that obtained with deep ducts

(Fig. 4·34*d*), but the complete mapping of the coast line is interrupted by extensive blank regions. The apparent analogy of this phenomenon of "skip effect" with that produced at longer wavelengths by "reflection" from the ionosphere leads one to seek a possible explanation in terms of strongly refracting elevated layers, probably overhanging.[1] What little meteorological data are available support this hypothesis, but the evidence is far from conclusive. The *M*-profile of Fig. 4·37 is therefore included tentatively. An alternative suggestion is that elevated *M*-inversions having appreciable slope may produce the skip effect.

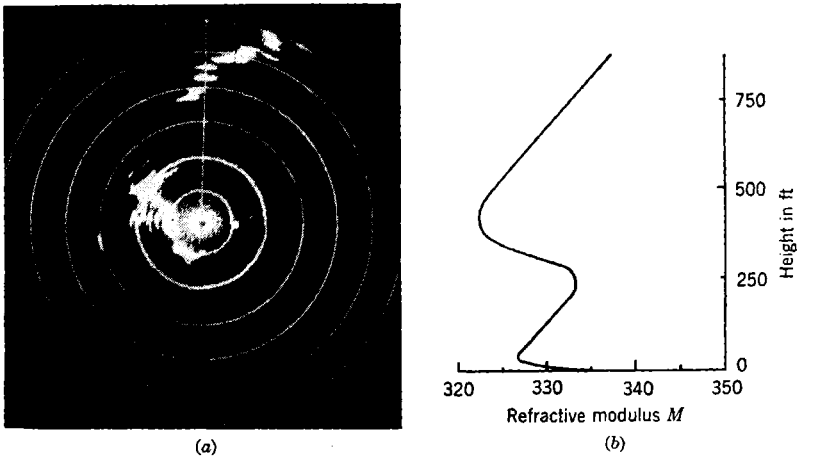
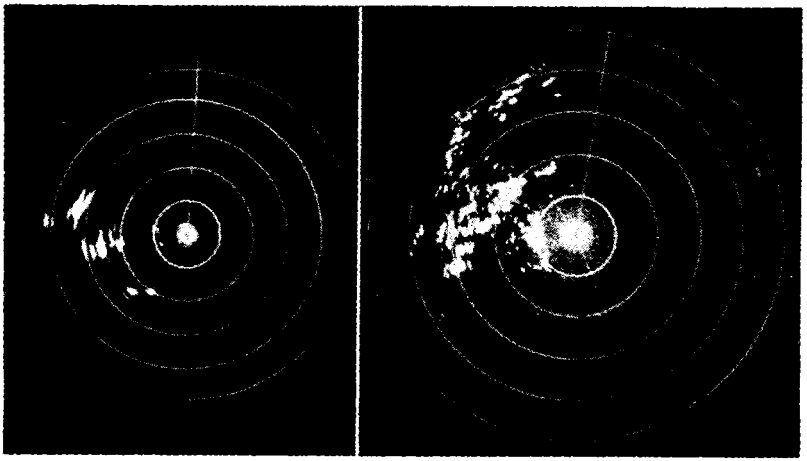


FIG. 4·37.—PPI pattern illustrating (*a*) radar skip effect and (*b*) accompanying *M*-profile. Range marks are 20 miles apart.

The pattern of Fig. 4·38*a* differs from all the others in that echoes are obtained from inland targets rather than those along the coast. An obvious interpretation is that this pattern results when ducts exist over land but not over water. From the meteorological discussion of Chap. 3 we find that ducts over land occur mainly at night, whereas conditions over water are almost always nonstandard. Thus we should expect overland echoes to be obtained mainly during the night and generally in combination with those from coastal targets. These expectations were completely fulfilled. The pattern of Fig. 4·38*a* was rare; in most cases the echoes over land occurred together with those from coastal regions, as shown in Fig. 4·38*b* and frequently were associated with skip effect on coastal targets. Because ducts over the water frequently become complex during the night, with the development of elevated *M*-inversions as the nocturnal radiational inversions travel from land out over the water, these observations form a consistent pattern.

[1] See Sec. 4·9, on the experiments near San Diego.

In Sec. 4·5 the absence of any appreciable effect of elevated ducts on microwave one-way transmission was noted. The remarks of the preceding paragraphs are in apparent contradiction to this point. No convincing explanation of this paradox has been found, but a few suggestions may be made. During periods of strong superrefraction the radar patterns are extremely variable; hence although a skip effect may persist for a period of some hours, the specific region skipped is not con-



(*a*) Echoes from inland targets only. Wavelength 9 cm. (*b*) Echoes from both inland and coastal targets. Wavelength 3 cm.

FIG. 4·38.—PPI patterns showing echoes from land targets; 20-mile range marks.

stant. At a fixed point, then, such a condition might result in a general depression of the average signal level together with irregular variability of signal. Several such periods were observed in one-way microwave transmission, when with strong-trapping type signal the average level was 15 to 20 db below the free-space value for several hours at a time. Only a few of these periods have been accounted for by the presence of known nonuniform meteorological conditions over the path. Existing data do not permit the establishment of a conclusive correlation between these unexplained periods of low-level strong-trapping type signal and radar skip effects, but they do not refute such a correlation.

More quantitative data on radar transmission were furnished by the comparison of echo strengths of radar targets with the one-way field strengths. The target used for this comparison was that at a range of 41 miles, close to the receiving terminal of the one-way circuit, a choice that eliminates any possible ambiguity from variation of refractive conditions in azimuth. If the radar cross section of the target were independent of the incident field distribution, the ratio of target echo intensity

to one-way signal intensity at a constant range should be simply proportional to F^2. The experimental variation is shown by the points in Fig. 4·39; the square-law line is included for comparison. It should be noted that the target did not become visible until the one-way signal strength had risen about 10 db above the standard value, an illustration of the requirement of strong fields for radar.

Measured signal strengths on the remaining selected targets were closely related to those of the 41-mile target chosen for comparison with

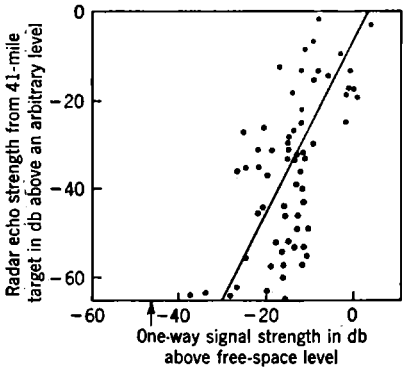


FIG. 4·39.—Comparison of radar and one-way signal strengths on 3 cm. Radar detection level is baseline. Arrow denotes one-way standard signal level on 3 cm.

the one-way results. In so far as could be determined, the variations increased with increasing target range from the radar and with decreasing wavelength. These large echo fluctuations serve to emphasize the fallacy of using echoes from natural targets as a measure of radar set performance.

4·16. California Coast.[1]—Early in 1942, at the United States Navy Radio and Sound Laboratory (now United States Navy Electronics Laboratory) on Point Loma, near San Diego, Calif., observations were begun of the effects on radar operation of the elevated M-inversion described in Sec. 4·9.[2]

The initial investigations were made with radar operating on 43 cm sited 425 ft above sea level in a location permitting an unobstructed view along the California coast line and out to sea. Figure 4·21 shows the location of the radar and the surrounding features of principal interest.

The islands and prominent features of the coast line provided natural targets convenient for qualitative radar investigations which first demon-

[1] By Donald E. Kerr.

[2] L. J. Anderson, J. B. Smyth, F. R. Abbott, and R. Revelle, "Radar Wave Propagation," NRSL Report No. WP-2, Nov. 30, 1942; L. J. Anderson and J. B. Smyth, "Atmospheric Refraction—A Qualitative Investigation," NRSL Report No WP-5, May 7, 1943; L. J. Anderson and L. G. Trolese, "Measurement of Radar Wave Refraction and Associated Meteorological Conditions," NRSL Report No. WP-7, Dec. 10, 1943.

strated the remarkable effects of the elevated M-inversions in this region in causing radar echoes from below the horizon. A detailed study of echoes from the islands and coastal regions indicated on the map made over several months during the winter of 1942 showed variations in echo strengths at times greater than 20 db in a few minutes, but neither amplitude nor rate of fading was comparable to those ordinarily encountered on shorter wavelengths under nonstandard refraction conditions. Echoes from adjacent localities were observed to fluctuate independently over short time intervals.

Natural targets, consisting of cliffs, buildings, and the like, have the disadvantage that echoes from them cannot be calculated, and they are often inconveniently distributed in space. To obviate this difficulty a series of targets was set up at the locations indicated in Fig. 4·21 by the numbers 1, 2, 3, and 4, at heights of 85, 550, 1300, and 1700 ft, respectively, above sea level.[1] For a period of a month in the fall of 1943 the strength of the echoes from these targets was observed on a radar operating on 1.5 m at the 430-ft Point Loma site. For standard refraction target No. 4 was at the maximum of the lowest lobe and No. 1 was just below the horizon; targets 2, 3, and 4 produced an echo well above noise. The predominant M-profile was characterized by a strong elevated irregularity of variable characteristics. The sounding techniques then available were not capable of revealing the detail shown in Sec. 4·9, and it was possible to conclude only that elevated superstandard layers occurred frequently and produced the greatest effect on the radar echoes when below 2000 ft. The details of the layer found later are described in Sec. 4·9.

The relative echo amplitudes varied from about 30 db above noise to values below the minimum detectable level. For about one-third of the time target No. 1 produced a strong echo, the other echoes being usually, but not always, detectable. This occurred when the soundings indicated strong M-gradients in the layer. There were a few occasions when the layer was very weak or was above 2000 ft and targets 2, 3, and 4 were visible, but during most of the time the combination of "visibilities" was highly erratic; not infrequently only targets 2 and 3; 1, 2, and 4; or 1, 2, and 3; and so on, were visible. On some occasions all four targets were invisible for several hours at a time.[2] Early interpretation of these effects was that radar "blind zones" from these layers could be serious; but in the light of later findings on the lack of homogeneity of the index distri-

[1] These targets were made of tightly strung wires, spaced so closely as to give essentially a plane reflector, and were oriented accurately to face the radar. The geometrical area of each was 1600 ft². In siting the targets an effort was made to choose locations that would permit the echo from the target to overpower that from the surrounding terrain and at the same time be free of local diffraction effects. No check is available on the degree to which these undesirable effects were avoided.

[2] The performance of the radar set was reported to be under careful check during the course of the tests.

bution in space and other observations on diversity reception and on the rapid and independent fading of echoes from adjacent targets, it seems that positive statements about blind zones are somewhat difficult to make.

4·17. Welsh Coast.[1]—During the summer of 1944 a series of Interservice radar trials was carried out on the north coast of Wales. This section concerns those results which bear on atmospheric refraction.[2]

Experimental radars on wavelengths of 1.25, 3.2, and 10 cm were located on Morfa Beach at heights above sea level ranging between 20 and 30 ft (see Fig. 4·16). Measurements were made of echo strength at times of nonstandard propagation on ships and on such random fixed targets as the coast of Ireland, the Isle of Man, buoys, and lighthouses. Soundings were made simultaneously by aircraft and ships along the line indicated in Fig. 4·16, employing instruments similar to those described in Secs. 3·29 and 3·31. Considerably more detail is reported in these aircraft soundings than is given in those for the measurements in the Irish Sea (see Sec. 4·7); these soundings are strikingly similar to those illustrated in Chap. 3, as contrasted to those reported for the Irish Sea transmission path and described in Sec. 4·7.

The effects of nonstandard M-profiles on 3- and 10-cm transmission were to extend markedly the detection ranges on ships and to cause echoes to appear from the Isle of Man and Irish coast, far below the horizon.[3] A prominent feature of these echoes was that they consisted of a large group of signals, each usually fading independently of the others at a fairly rapid rate; hence no single echo could be used alone as a measure of the intensity of superrefraction. In the present instance it was necessary to deal with a "composite" signal strength, taking into account both the number and strength of echoes received on bearings a few degrees apart. Occasionally the echoes were steady enough to make direct comparisons, however, and at these times the echo strengths from the Isle of

[1] By Donald E. Kerr.

[2] J. R. Atkinson, "General Summary Covering the Work of the Interservice Trials, Llandudno, 1944," TRE Report No. T1770, May 1945; W. A. Hayes, J. Ingham, and F. L. Westwater, "KXS Llandudno Interservice Trials Summer 1944. Correlation of Radar and Meteorological Conditions during Periods of Very Weak Superrefraction," TRE Report No. T1865, Aug. 27, 1945; W. A. Hayes, J. Ingham, and R. W. James, "KXS Interservice Trials Summer 1944. Investigation of Anomalous Propagation (Correlation of Meteorological Conditions and Radar Signals)," TRE Report No. T1803, Mar. 29, 1945.

[3] The 1.25-cm radar was rarely able to detect long-range surface echoes, even when they came from only slightly beyond horizon range, and at no time did it detect the Isle of Man. This fact was attributed to absorption by water vapor, which for a 1.25-cm echo from the Isle of Man would be somewhere between 20 and 30 db (see Chap. 8). This, in addition to the fact that the over-all set performances on $\lambda = 1.25$ cm was about 35 db lower than that on 10 cm, represents a serious handicap, even for very strong superrefraction.

Man on the two wavelengths differed by about the difference in over-all set performance.

The meteorological soundings revealed three principal deviations from standard M-profiles: elevated ducts 200 or 300 ft thick with bases at roughly 500 ft, surface ducts 100 to 200 ft thick, and a semipermanent duct in the lowest 50 ft. The thick surface ducts were ascribed to modification by advection (see Sec. 3·17). Figure 4·40 shows a sounding illustrating this situation. The low evaporation duct results from a strong humidity

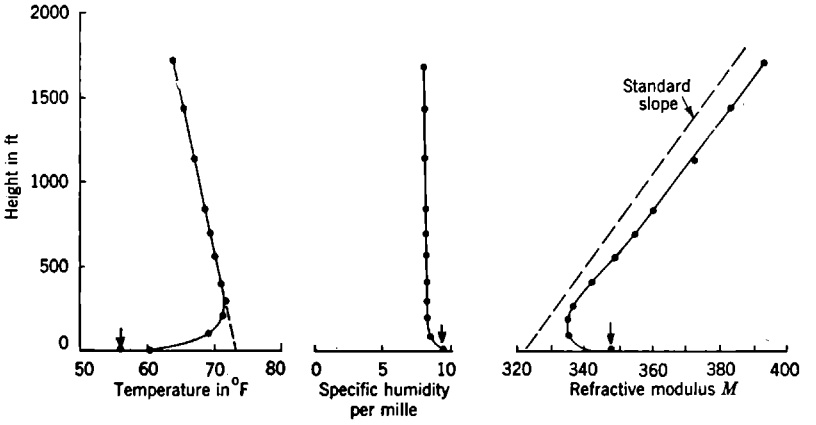


FIG. 4·40.—Sounding in the Irish Sea showing a deep surface duct ascribed to modification by advection. The arrows indicate values at the water surface.

deficit that outweighs a superadiabetic temperature gradient resulting from heating from below, as discussed in Secs. 3·14 to 3·16. Figure 4·41 shows a combination of an evaporation duct and an elevated duct. The meteorological origins of the elevated ducts were not determined in individual cases, but the data available suggest that they were caused in a number of different ways. Occasionally they occurred at very low altitudes, overhanging the surface.

Aircraft soundings made at a number of different locations indicated that in a homogeneous air mass a single sounding was broadly representative of conditions over a wide area, although details of the profiles varied with location. Very near the surface, however, ship soundings showed such marked fluctuations in temperature and humidity that considerable difficulty was experienced in the interpretation of the soundings, particularly when compared with simultaneous aircraft soundings at the same location; the principal emphasis was placed on the latter in correlation with the radar data. Despite the fluctuations shown by the ship soundings, one striking feature appeared. Under conditions of heating from below, practically the entire temperature, humidity, and M-deficits occurred

between the sea surface and a height of about 5 ft.[1] Between 5 and 50 ft the temperature and humidity profiles varied relatively little with height and the top of the duct was poorly defined.

A study was made of 50 days' records of radar performance, variations of signal level on the one-way transmission circuit described in Sec. 4·7, and measured M-profiles. The degree of superrefraction, as judged by strength of echoes from fixed targets and by maximum ranges on ships, could be correlated qualitatively with the irregularities of the M-profiles,

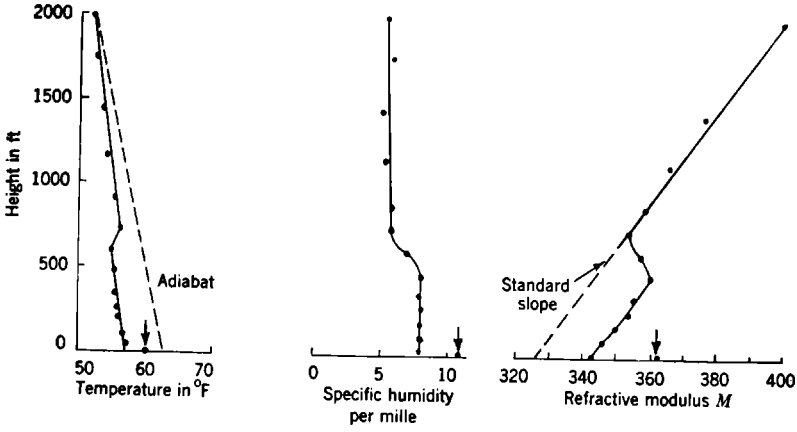


FIG. 4·41.—Sounding in the Irish Sea showing a combination of an evaporation duct and an elevated duct. The arrows indicate values at the water surface.

but the variability of the latter very near the surface and the frequent occurrence of combinations of surface and elevated ducts precluded quantitative conclusions.

The strongest superrefraction occurred with combinations of surface and elevated ducts, often of rather irregular form. Figure 4·42 shows a sounding[2] made at a time of very strong superrefraction. At about the time this sounding was made, echoes were obtained from Scotland, Ireland, and the Isle of Man on the 3- and 10-cm sets on Morfa Beach, and a near-by 1.5-m radar received echoes from Ireland—a rare phenomenon. Considerable superrefraction was also present in situations illustrated by Fig. 4·40. There were periods of apparent inconsistency, when the temperature excess was positive and the humidity deficit was zero or slightly positive, apparently accompanied by appreciable local fluctuations. At

[1] The process of eddy diffusion cannot operate efficiently very near the surface, and large gradients are necessary to accomplish the transfer of heat and moisture required by the temperature and humidity deficits (see Sec. 3·11 to 3·13 and 3·26).

[2] This sounding was made in Cardigan Bay, but others made northwest of Morfa Beach showed roughly similar shapes.

these times the M-deficit tended to change sign, resulting in a surface layer whose slope fluctuated between superstandard and substandard.

Periods of noticeably weaker superrefraction occurred when the elevated ducts were weak or fairly high, as shown in Fig. 4·43, and least effect was observed when soundings and analysis of the synoptic situation indicated only a shallow evaporation duct at the surface and standard slope of the M-profile above. These conditions of weak superrefraction were particularly interesting because of the marked dependence of performance on

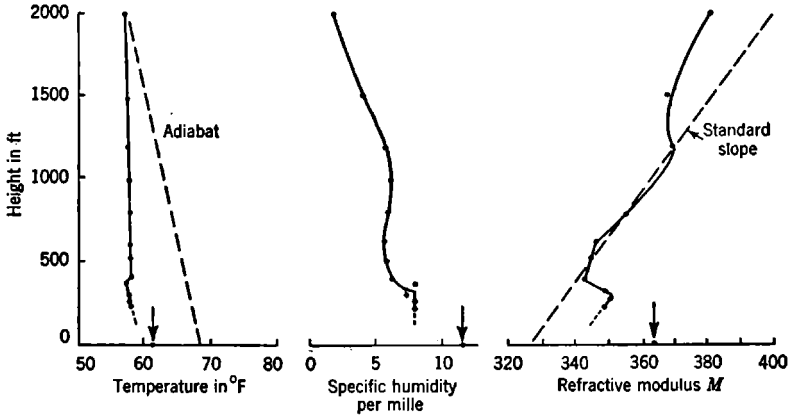


FIG. 4·42.—Sounding in the Irish Sea showing an M-profile associated with the strongest superrefraction. The arrows indicate values at the water surface.

wavelength; the 3-cm performance on shipping and the nearer natural targets was noticeably better than on 10 cm (cf. Sec. 4·15), and decreasing the set height from about 30 to 20 ft further increased this advantage for 3 cm.

In addition to these explorations of surface coverage variations, coverage up to a height of about 300 ft was investigated at short ranges (15,000 to 25,000 yd) by measurement of echoes from balloon-borne corner reflectors and from an autogyro. Well-defined lobes resulting from reflection from the water were found even on 1.25 cm. At times of weakest ducts the positions and shapes of the lobes near the maxima on 10 cm agreed with simple calculations from plane-earth theory, but the finer lobe structure on the shorter wavelengths made a close check impossible. At times of strong superrefraction, however, noticeable changes in lobe shape and position were evident. The lowest 10-cm lobe showed a tendency to split in two, with a somewhat weakened top section, a slight gap in the middle, and a low "finger" very near the surface. The changes in lobe structure were not so clearly defined on 3 and 1.25 cm, but there appeared to be a general lowering of the lowest two or three

lobes and irregular distortion of their shapes, rendering them individually unidentifiable.[1]

The radar behavior and measured M-profiles were also compared with transmission data from the Irish Sea experiment described in Sec. 4 7. Good general agreement was found between periods of high signal strength on the *BD* circuit, long radar ranges, and M-profiles with strong surface ducts alone or combined with elevated ducts. Detailed comparisons have not been made, but there is no evidence in the data thus far published

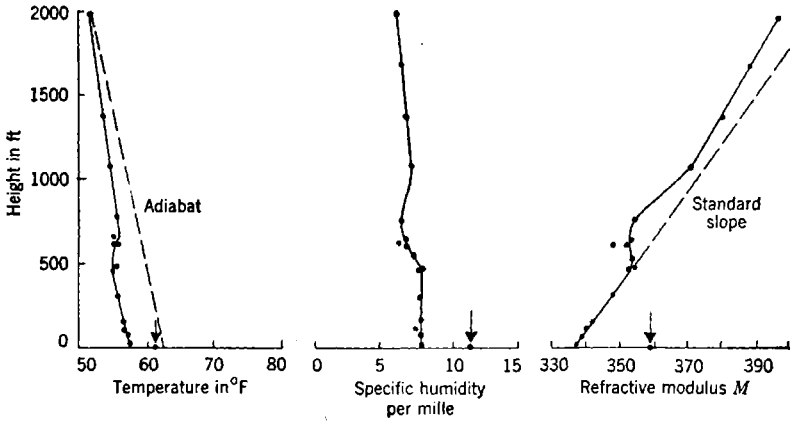


FIG. 4·43.—Sounding in the Irish Sea showing a weak elevated duct. The arrows indicate values at the water surface.

to suggest that there should be essential disagreements between these results and those described in Secs. 4·1 to 4·6 and 4·15 for Massachusetts Bay.

4·18. The English Channel Region.[2]—Because of the early use of radar along the south coast of England, propagation vagaries in and near the region of the English Channel were probably the first to be explored to an appreciable extent by radar means. Some of the first observations were made in 1941 with an early 10-cm radar on the south coast of England near the Isle of Wight.[3] Echoes from the coast and from inland targets in France were observed at ranges up to 170 miles, and study of short-period variations revealed the now familiar independent fading of echoes from targets fairly close together.

Following installation of a microwave radar chain along the south and southeast coasts of England, procedures were initiated for systematic

[1] See also the article by J. A. Ramsay in *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London (1946).

[2] By Donald E. Kerr.

[3] A. E. Kempton, "Anomalous Echoes Observed with 10 cm C.D. Set," ADRDE Report No. 119, Oct. 8, 1941. (Probably the first such observations were made at Christchurch in 1940 using a wavelength of 50 cm.)

study of the effects of refraction by observation of echoes from near-by objects such as buoys, from surface vessels, and from landscape targets across the Channel in France, Holland, and Belgium. During the early experiments[1] made by a high-power set (300 kw, 10-ft paraboloid) at a height of 780 ft it became clear that temperature inversions at low altitudes were important in producing superrefraction, and attempts were made to forecast periods of nonstandard propagation, using the occurrence of temperature inversions as a basis for the predictions. Although this procedure ignored the difference in dew point of the air and the water temperature, it proved to have some value in tactical applications of coast-watching radar.

A statistical study was made of the echoes obtained on a number of 10-cm and 1.5-m radar sets along the south and southeast English coasts from landscape targets across the Channel.[2] Over very short time intervals individual installations were likely to experience different refraction effects in a manner indicating a considerable degree of meteorological nonuniformity. Over periods of hours, however, the general features were usually rather similar. The echoes were divided into rough classifications according to "normal" (presumably occurring with standard refraction) or "anomalous."[3]

An average of the performances during periods showing nonstandard conditions on four of the 10-cm and on the four 1.5-m stations for the period May to August 1944 revealed the diurnal variation[4] shown in Fig. 4·44. The tendency toward greater superrefraction on the shorter wavelength is evident, suggesting predominantly surface-type ducts as the cause. In addition to the diurnal variation, there was a strong seasonal variation. Superrefraction was frequent and intense in warm weather but very much reduced during the winter months, as would be expected in view of the variations in air- and water-temperature contrasts. (One interesting exception occurred when strong superrefraction was observed for several hours at the climax of a winter gale.) Substandard performance was also observed over a wide area on a number of occasions, at

[1] "Anomalous Propagation of 10 cm RDF Waves over the Sea," AORG Report No. 87, Feb., 1942; Supplement, July 26, 1943.

[2] D. Lack, "Anomalous Propagation at English Coastal Radar Stations, Oct. 1944 to Mar. 1945," AORG Report No. 286, Oct. 1, 1945, and AORG Report No. 258, May 30, 1945, of the same title.

[3] As radar heights from 170 to 780 ft and a wide range of target heights and horizon distances were involved, such a division necessarily involved a certain amount of arbitrary choice, but changes brought about by superrefraction were usually sufficiently large to make the procedure not too difficult.

[4] The 10-cm set at Ventnor at a height of 780 ft persistently experienced an opposite diurnal variation, with a peak between 0400 and 0800 hours. No explanation has been offered. Also there were occasions, usually in the early morning, when the 1.5-m set showed superrefraction but the 10-cm set did not.

times when it would be expected from the temperature and humidity contrasts between air and water.

One of the tasks of microwave radar in the Channel was to estimate the size of a surface vessel from the strength of the echo and the way in which it varied with range. This required for each radar height, set performance, and ship type preparation of theoretical curves of echo strength vs. range (see Sec. 6·5) and calibration procedures to ensure the required degree of accuracy. The principal effect of superrefraction was to distort the signal strength-vs.-range curve, beginning as the target was very near the horizon, extending the bottom of the curve out to several times the normal maximum detection range. (This behavior is illustrated for a point target in Fig. 1·5.) A series of tests was made with a 3-cm set at a height of 50 ft and a 3-cm set at 780 ft, in which the theoretical curves were investigated in detail.[1] Under standard conditions the high set tracked an "average" ship to 80,000 yd and the low set to 35,000 yd, but strong superrefraction increased both of these figures considerably; moreover, the low set produced three times as many plots beyond 80,000 yd as did the high set. During periods of weaker superrefraction (in September, whereas the more striking effects occurred in August) the low set often plotted ships at 40,000 to 59,000 yd, whereas the high set experienced no extension of ranges. It was also noticed that under synoptic conditions conducive to strong overwater temperature inversions, strong superrefraction was accompanied by deep fading of the echoes, although with conditions conducive to atmospheric instability over the water the weaker superrefraction produced noticeably steadier echoes. Unfortunately no soundings are available, but these results strongly suggest that the advantage of the low set over the high set is explained by the proximity of the low set to surface ducts, which became weaker and took on the characteristics of the "evaporation duct" type with the onset of fall weather (see Sec. 4·6).

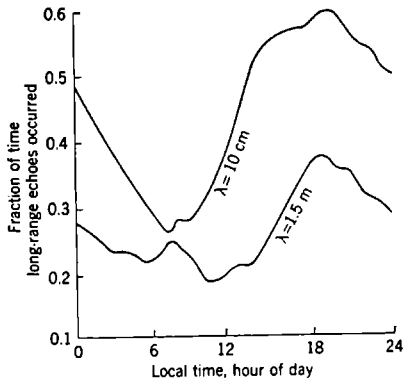


FIG. 4·44.—Fraction of time long-range echoes occurred on four 10-cm and four 1.5-m radar stations on the Channel coast.

4·19. Other Regions.[2]—A list of the many occasions when nonstandard refraction is reported to have produced a marked effect on radar per-

[1] G. C. Varley, "Anomalous Propagation with High and Low Sited 3-cm Ship-watching Radar Sets," AORG Report No. 250, Mar. 20, 1945.

[2] By Donald E. Kerr.

formance would be very long indeed and would reveal that such effects have occurred at some time in substantially every region where radar has been used. The examples given in the previous sections were chosen because they were based for the most part on large numbers of observations made according to reasonably consistent procedures and in regions where the meteorological situation had been subjected to some degree of scrutiny. Unfortunately these conditions were not fulfilled in most of the remaining large number of cases, but some of them are interesting, although practically no quantitative data are available that would permit an analysis like that of the foregoing sections.

Detection ranges of several hundred miles have been obtained in the Mediterranean Sea on wavelengths from 1.5 m to 10 cm by shipborne and coast-watching radar at low heights. Tracking surface vessels out to 200 miles is not uncommon; and landscape targets near the coast have been detected at ranges up to 500 miles. Comparable ranges have also been reported in the Atlantic Ocean off the northwest coast of Africa.

The western coast of Australia and the vicinity of New Zealand are the source of many reports that superrefraction on both meter and centimeter wavelengths causes large extensions of surface-detection ranges and serious gaps in the coverage diagram, with consequent loss of reliability in aircraft detection. The available information is so incomplete that no assessment of the real magnitude of the difficulty is possible, however. The limited meteorological information available suggests that this region should furnish widely varied types of nonstandard propagation, for large-scale meteorological processes involving strong contrasts between continental air and the ocean are present, and intense local circulation giving the effect of a sea breeze is a predominant feature along the coasts.[1] The greatest effects from superrefraction have been found in the summer, and the least in late fall and early winter.[2]

In the Strait of Hormuz, between the Gulf of Oman and the Persian Gulf, a 10-cm radar sited at a height of about 170 ft has observed interesting seasonal variations in superrefraction[3] (see Fig. 4·45). From May to October 1943, strong superrefraction was encountered almost continuously on looking up into the Persian Gulf, and shipping was customarily tracked far beyond the horizon. Looking into the Gulf of Oman, however, a seasonal trend was observed. During May and June the performance was essentially similar to that in the Persian Gulf, but during July "skip"

[1] See Fig. 3·51 for an example of a sounding under sea-breeze conditions. See also, H. G. Booker, "A Qualitative Outline of the Radio Climatology of Australia," TRE Report No. T1820, Apr. 19, 1945. See also the article by F. E. S. Alexander in *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London (1946).

[2] For instance, see "Abnormal Radar Propagation in the South Pacific," RN AZAF Report No. 119, May 4, 1944; F. J. Kerr, *Austral. Jour. Sci. Res.*, A, **1**, 443 (1948).

[3] Naval Officer in Charge, "Anomalous Propagation in the Persian Gulf," Hormuz, A. C. 5975 USW.

effects became marked; inside the horizon coverage was undisturbed, but beyond the horizon was a "skip zone" in which shipping could not be detected, and then beyond the skip zone it was again possible to track shipping to fairly large ranges. The inner boundary of the skip zone moved out so far in August that detection was essentially limited to horizon range. In September the reverse trend occurred, and by the end of October conditions had returned to essentially the same as those observed in May and were again comparable to those in the Persian Gulf.

An explanation has been offered for these phenomena, based on the general climatology of the region.[1] During the spring, summer, and fall months the wind across the Persian Gulf comes predominantly from the northwest and is dry and warmer than the water; consequently ducts of some magnitude are formed. Presumably this condition also prevails over the Gulf of Oman in the spring and fall[2] but is quite different in the summer when the southwest monsoon from the Arabian sea thrusts a surface layer of cool moist air into the Gulf of Oman up to the vicinity of Hormuz and gradually forces the surface duct aloft, producing the skip effect and approach toward standard refraction in August as the duct moves aloft or is destroyed. The monsoon then recedes, and by October[3] the warm dry air from the northwest again predominates.

Perhaps the most phenomenal ranges yet reported are those from 1.5-m radar sited 255 ft above the sea at Bombay, India.[4] Figure 4·45 shows the coastal regions from which echoes were received under conditions of very strong superrefraction. (Somewhat similar performance was noted on a radar set at Karachi, although the data are not so complete.) At any one time echoes were usually obtained from only a limited region of the Arabian coast, but at one time or another echoes were obtained from most of the Arabian and Persian coasts bordering on the Arabian Sea and Gulf of Oman, at distances up to about 1700 miles. These echoes sometimes persisted for hours and at the height of the hot season even for days, aside from the usual fading. Ships were not infrequently plotted at ranges from 100 to 300 miles and occasionally up to 700 miles. During the period for which detailed records are available (February to May 1944), superrefraction increased very rapidly from February to April, when the peak appears to have been reached. Installations on the east coast of India also observed echoes from the coast of Burma and the

[1] H. G. Booker, "Radio Climatology of the Persian Gulf and Gulf of Oman with Radar Confirmation," TRE Report No. T1642, Mar. 15, 1944.

[2] In the winter months the air in both regions is somewhat cooler than the water, and presumably weak superrefraction results.

[3] In the light of the discussion in Chap. 3 it seems unlikely that the mechanism in either gulf is as simple as this explanation suggests, but more detailed explanations do not appear possible without soundings.

[4] H. G. Booker, "Outline of Radio Climatology in India and Vicinity," TRE Report No. T1727, Sept. 12, 1944.

Andaman Islands, up to 700 miles across the Bay of Bengal, but incomplete records preclude comparisons. Superrefraction of this order farther south apparently did not occur, since no such effects were observed by 1.5-m radars at Cochin and Ceylon. The small amount of 10-cm radar information available indicates no extreme ranges of this nature, although a 10-cm radar near the 1.5-m radar at Bombay often plotted ships at 100 to 200 miles.[1]

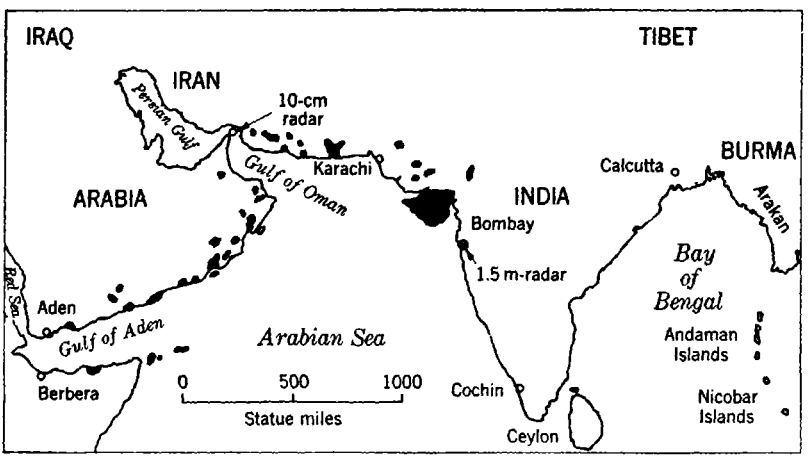


FIG. 4·45.—India and vicinity, showing long-range fixed echoes observed by 1.5-m radar at Bombay during February and March 1944.

The mechanisms responsible for the phenomenal 1.5-m superrefraction can hardly even be surmised, in view of the lack of sufficient knowledge of the atmospheric structure in the vicinity of India.[2] The very long ranges obtained on the long wavelength suggest elevated M-inversions acting as efficiently reflecting layers, but this alone does not furnish a very useful clue, as several meteorological processes may cause such inversions. One meteorological effect does seem clear, however: the southwest monsoon causes cool, damp air to invade the region in the summer, when extremely long ranges are not observed, presumably because the monsoon destroys whatever stratification is present.

Strong superrefraction also occurs over land in India and has been noticed particularly by 1.5-m radars near Calcutta, on the alluvial plain

[1] Over these very great ranges the 10-cm radar suffers an additional handicap from atmospheric absorption, which, for radar ranges of about 1000 miles, might be as much as 40 db (see Chap. 8).

[2] The explanation suggested in Booker, *op. cit.*, postulates very thick ducts, whose shapes are governed by a power law of height. It is based on a theory of diffusion shown in Sec. 3·19 to be inapplicable to a large height range.

at the mouth of the Ganges River,[1] where radar ground clutter from ranges up to 200 miles has been observed at night. Some effect was also noted on 7-m radar installations.

An interesting example of the effects of an elevated duct was observed on a 10-cm early-warning radar located at a height of 1540 ft on Saipan Island in the Marianas group in the western Pacific.[2] A series of soundings, a representative sample of which was shown in Fig. 3·11, revealed the presence of a weak elevated duct that persisted over distances of many miles. The characteristics of the duct varied with local weather conditions; in particular, the height of the duct coincided very closely with the base of the prevailing cumulus clouds. It was observed that when this duct occurred at or very near the height of the radar, echoes from islands within the horizon[3] increased greatly in strength and strong echoes appeared from below the horizon at ranges up to 250 miles (the greatest range permitted by the pulse recurrence frequency and indicator system), but usually only from those islands with peaks about 1500 ft or more high. This suggests the existence of a fairly well-defined elevated region of high field strength, resulting from some degree of trapping in the duct, causing strong illumination of island peaks.

From the examples given so far in this chapter it can be seen that nonstandard propagation can produce remarkable and, at times, perhaps useful effects in both radar and point-to-point applications of very short waves. The question of possible harmful effects arising from sufficient distortion of the coverage diagram to reduce field strength excessively in important regions is still unanswered. It is discussed further in the following section.[4]

SPACE VARIATIONS IN FIELD STRENGTH

BY WILLIAM T. FISHBACK

The transmission measurements previously discussed have been concerned primarily with time variations, whereas space variations were confined to measurements at different fixed heights or ranges. It has been recognized for some time that measurements of space variations over a wide range of heights and distances are desirable for a thorough check on theory; indeed, the Bell Telephone Laboratories were studying

[1] Booker, *op. cit.*

[2] D. E. Kerr and J. P. Nash, "Investigation of Performance of AN/CPS-1A (MEW) on Saipan," Operational Research Section, G-3, Headquarters U.S. Army Forces, Middle Pacific, APO 958, July 15, 1945.

[3] "Horizon" is the total horizon range resulting from the height of the radar and the height of the highest point on the island.

[4] See also E. C. S. Megaw, *Jour. IEE*, **93**, III A, 79 (1946).

space variations with airborne receivers as early as 1932.[1] Unfortunately, experimental difficulties are extreme in this type of measurement, and few data have been obtained. We shall be concerned here with observations made in the diffraction region under nonstandard refraction conditions. Measurements in the interference region with standard refraction are discussed in Secs. 5·5 and 5·9.

4·20. Shallow Surface *M*-inversions.—*Antigua, B.W.I.*—The most comprehensive measurements of space variations of field strength were made by the Naval Research Laboratory at Antigua, British West Indies, in the spring of 1945.[2] The path chosen was over water in a northeasterly direction from the island. Measurements of field strength as a function of distance were made at 3 and 9 cm, with transmitters at 16 and 46 ft and with receivers at various heights in the lowest hundred ft. The receivers were placed on a tower at the water's edge, and the transmitters installed in a small U.S. Navy patrol vessel which traveled on a radial course either toward or away from the shore installation. Meteorological soundings using the wired sonde (Sec. 3·28) were made aboard ship and at the water's edge.

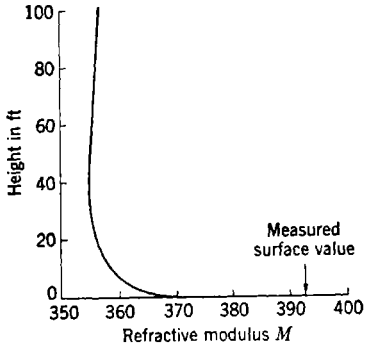


FIG. 4 46.—*M*-profile measured off Antigua, B.W.I., on Mar. 20, 1945, showing a typical evaporation duct.

Antigua is in the region of the northeast trade winds and has no land northeast of it for many miles. The air on the transmission path had therefore had a long overwater trajectory, unmodified by any land mass. All soundings revealed the shallow evaporation duct expected under these conditions (see Secs. 3·14 to 3·16). On the average, duct heights were about 40 ft but tended to be lower with low winds and higher with high winds. As the *M*-deficit is a function of the air-mass properties and water temperature, which varied only slightly, it shows no dependence on wind speed. Figure 4·46 shows a typical *M*-profile measured at the edge of the water. Measured soundings showed no striking nonuniformity along the path, nor did they vary much from hour to hour; consequently, in spite of the fact that a boat run took as long as 30 hr, time variations should not be an important factor in the data obtained. The very rapid

[1] J. C. Schelleng, C. R. Burrows, and E. B. Ferrell, "Ultra-short-wave Propagation," *Proc. IRE*, **21**, 427 (1933); C. R. Englund, A. B. Crawford, and W. W. Mumford, "Further Results of a Study of Ultra-short-wave Transmission Phenomena," *Bell System Tech. Jour*, **14**, 369 (1935).

[2] M. Katzin, R. W. Bauchman, and W. Binnian, *Proc. IRE*, **35**, 891 (1947); C. L. Pekeris, *Proc. IRE*, **35**, 453 (1947).

fluctuations typical of unstable air were pronounced, however (see Sec. 3·26). The trade-wind inversion existed at Antigua throughout the period of observations but occurred at heights between 5000 and 10,000 ft and presumably did not affect the measurements.

Considerable data were obtained on the signal levels associated with shallow ducts, but it was not possible to study signal types during this experiment. The roll and pitch of the ship was sufficient to preclude perfect alignment of the antennas, with the result that considerable non-atmospheric variation was obtained in the record. The envelope of maximum field strength recorded was used in the analysis. Figure 4·47 shows the results obtained for 9 cm on Mar. 2 and 3, 1945, with the 16- and 46-ft transmitters. (The ordinate $20 \log_{10} (FR_s/R)$ is the field strength in decibels above that at 1 m from the transmitting antenna.) Although the attenuation rates and height-gain effects varied noticeably from run to run, the data shown are typical.

In general, for 9 cm it was found that the field was always above the standard level, that the highest terminal heights of those available gave the strongest signal, and that lowering a terminal gave a progressively weaker signal. This finding suggests that only one leaky mode was responsible for the observed field. This height-gain effect was present only out to ranges of 65 to 90 miles. At greater ranges the field was essentially independent of height and the observed rate of attenuation was very small. The attenuation rate at ranges of less than 65 to 90 miles was on the average about 1 db/mile, although it varied slightly with the terminal combination used.

Figure 4·48 shows a similar set of data for 3 cm taken from Feb. 27 to Mar. 1, 1945. At this shorter wavelength all height combinations gave superstandard field strengths; furthermore, the highest signal strength was obtained with the *lowest possible* combination of terminal heights, which for some runs was as low as 6 and 8 ft. The height-gain effect in this case was different from that obtained on 9 cm. It showed maxima at about 6 and 60 ft and a minimum at about 30 ft. The interpretation of the minimum in the height-gain curve is not entirely clear, as the measurements were made at a range so near the horizon range for the high terminals, but it is likely that several modes were present and happened to combine in this particular way. The difference between maxima and the minimum was 8 db or less. There was no division of the diffraction region into two sections with different attenuation rates as on 9 cm. The observed attenuations were about 0.4 db/mile for the lower height combinations and from 0.4 to 0.8 db/mile for the higher height combinations.

In general the observations seem compatible with current theoretical predictions of the effects of shallow ducts. A correlation worth noting is that obtained with wind speeds. As indicated earlier, somewhat higher ducts were formed with higher than average wind speeds. The radio

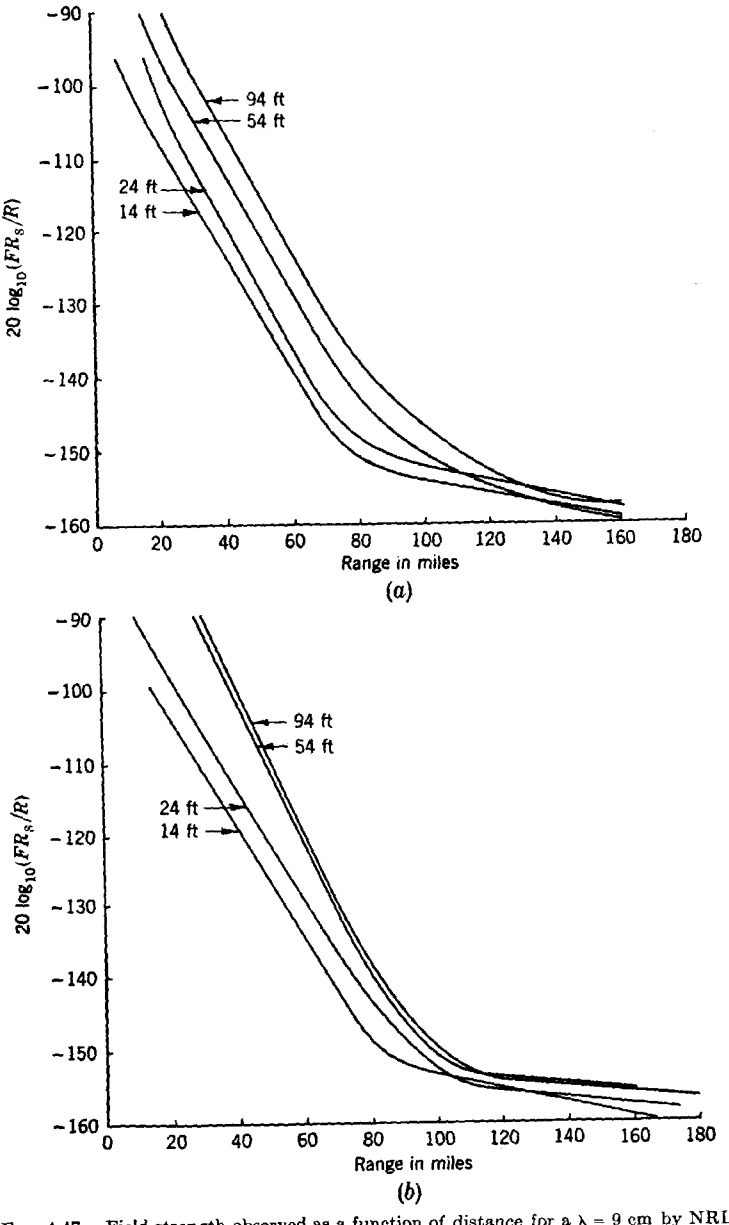


FIG. 4·47.—Field strength observed as a function of distance for a $\lambda = 9$ cm by NRL off Antigua, B.W.I., for indicated receiver heights. (*a*) With transmitter at 16 ft; (*b*) with transmitter at 46 ft.

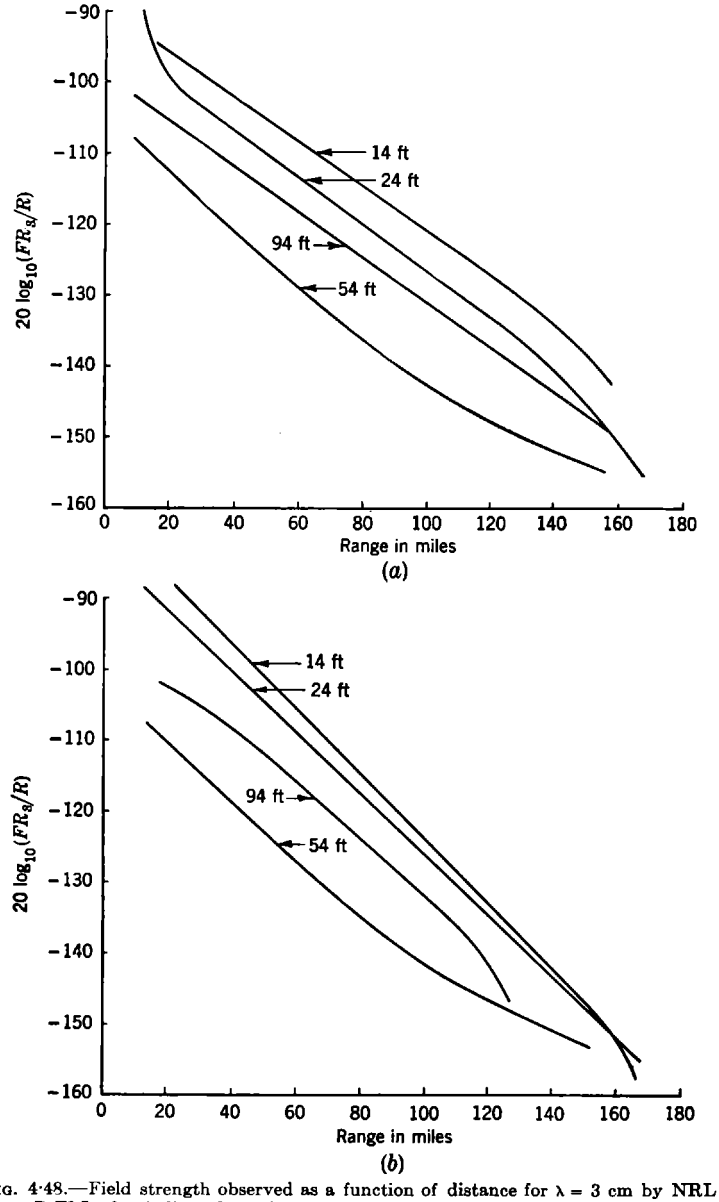


FIG. 4·48.—Field strength observed as a function of distance for $\lambda = 3$ cm by NRL off Antigua, B.W.I., for indicated receiver heights. (*a*) With transmitter at 16 ft; (*b*) with transmitter at 46 ft.

data confirm this observation in that lower attenuation rates prevail when the wind speeds and consequently the ducts are higher. Propagation on 9 cm seems to be accomplished by means of a single leaky mode, in view of the height-gain observed effects. The height-gain effect on 3 cm indicates that the first mode is either completely or almost completely trapped and that other modes may be of some importance.

A reconciliation of the attenuation rates with existing theory is not so simple, however. The range attenuation at great ranges for 9 cm is lower than would be expected on the basis of the bilinear M-profile, and enough information is not at present available for the linear-exponential profile to make a satisfactory analysis.[1] Finally, the observed attenuation on 3 cm might be expected to be somewhat smaller, in view of the indication of complete or nearly complete trapping given by the height-gain effects.

Measurements off the northwest Gulf Coast of Florida by the U.S. Army Signal Corps also show results in essential agreement with the observations made at Antigua.[2] Transmission on 6.6 cm was carried out over ranges of 19, 32, and 50 miles at heights in the lowest 50 ft. With a shallow duct 30 to 35 ft deep the observed attenuation was about 0.67 db/mile (standard attenuation would be 2.5 db/mile, a value between the 9- and 3-cm attenuation rates observed at Antigua). The 50-ft terminals gave the highest signal, but there was little change until the terminals were placed lower than 20 ft, and in this case the signal level decreased. One might expect from the Antigua results that such an intermediate wavelength would be on the verge of complete trapping and would have an essentially constant height-gain curve, with range attenuation appreciably less than standard. The Florida observations bear out this prediction.

4·21. Deep Surface M-inversions. *New England Coast.*—On June 8 to 10, 1943, the Radiation Laboratory made a series of measurements on a path along the New England coast between Rockport, Mass., and Portland, Maine. A 10-cm transmitter using vertical polarization was placed 25 ft above sea level at Rockport, and a receiver carried in an airplane was used to explore the field distribution.[3] During this time strong M-inversions were found offshore with duct heights of about 200 ft on the transmission path. Figure 4·49 shows a sample M-profile illustrating the type of profiles occurring during this period.

Figure 4·50 shows a vertical section of the field-strength pattern made on June 9, 60 miles from the transmitter. To obtain this section, the

[1] For later analysis, see C. L. Pekeris, *loc. cit.*

[2] J. A. Engstrom, J. R. Gerhardt, and G. D. Ramspeck, "Comparative Tests of Radio Relay Equipment, Phase II, Low-level Overwater Propagation, Northwestern Florida, August–October, 1945," SCEL Development Detachment, Ft. Monmouth, N.J., December 1945.

[3] See Sec. 5·5 for a more complete description of the experimental procedure followed in these tests.

plane descended rapidly from 800 to 25 ft and then climbed back to 800 ft; the horizontal travel during the descent and ascent was 7 miles. The measured vertical sounding agrees roughly with field-strength calculations based on a bilinear M-profile and with observations on the Race Point-to-Eastern Point path. Within the duct the field strength was within 12 db of the free-space value.[1] Several modes were completely trapped and were responsible for the high field within the duct. The field above the duct was superstandard and was presumably caused by the first partially trapped mode, which had less range attenuation than that of the first standard mode.

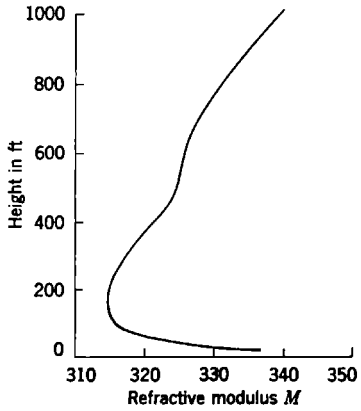


FIG. 4·49.—Typical M-profile on the Rockport-Portland path, June 9, 1943. This M-profile illustrates the type occurring during this period.

Figures 4·51 and 4·52 show horizontal sections made at heights of 20 and 500 ft, respectively, on the same day over the same path. Although the measurements show a very striking height advantage in favor of 20 ft, the flight at 500 ft does not in general indicate substandard fields at this height. On this occasion a detectable signal was found out to a range of at least 60 miles; in the

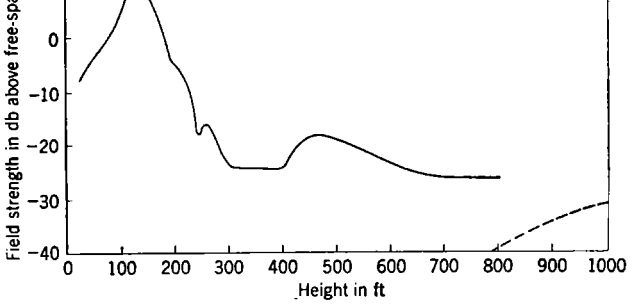


FIG. 4·50.—Vertical field-strength sounding 60 miles from a 10-cm transmitter at a height of 25 ft, June 9, 1943, off the New England coast. The dashed line indicates the expected field strength for standard refraction.

[1] The free-space level for these observations cannot be determined accurately from the equipment characteristics. The value chosen is in agreement with the measured value in the interference region and is 6 db lower than that calculated for the probable values of the equipment characteristics.

standard case signals should not have been detected at ranges greater than 42 miles. The field strength just inside the total horizon range was profoundly affected in the 500-ft flight. The first maximum was observed at 7 miles instead of at 19 miles for standard refractive conditions. Two lobes were lost or displaced in the region between 7 and 19 miles. Although a moderate field strength was observed in this region,

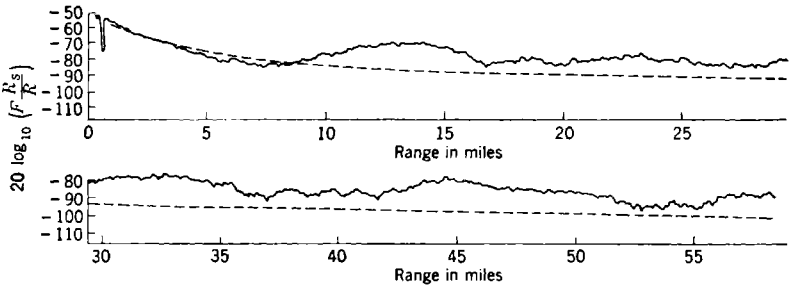


FIG. 4·51.—Field strength at a height of 20 ft over the ocean from a 10-cm transmitter at a height of 25 ft on the Rockport-Portland path, June 9, 1943. The dashed line indicates the free-space field.

it would have been higher, on the average, under standard conditions. Within 7 miles the field strength at the maxima obeyed the inverse-first-power law very well.

The observed field strengths on the 20-ft flight were superstandard at all ranges. Although the field measured on the 500-ft flight beyond the

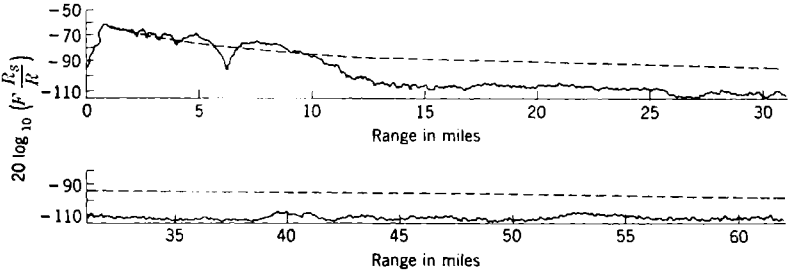


FIG. 4·52.—Field strength at a height of 500 ft over the ocean from a 10-cm transmitter at a height of 25 ft on the Rockport-Portland path, June 9, 1943. The dashed line indicates the free-space field.

horizon was caused by leaking high-order modes, the 20-ft flight shows high field strengths caused by strongly trapped, strongly excited low-order modes. For this combination of terminal heights and standard refraction the first maximum should occur at 1.1 miles.

It is difficult to say just how successful this experiment was in separating time and space variations. Experience on the Race Point-to-Eastern

Point path indicates that there is considerable short-time variation present under this type of meteorological condition, and there is no doubt that considerable time variation was also present in these measurements. This should be especially true for the observations within the duct. It would probably have been impossible to reproduce exactly the lowest 300 ft of the vertical sounding on a repeat flight, although high field strengths would have been observed. The effects of short-time variations, as well as those of nonuniformity, were also rather pronounced in the 20-ft flight. The resulting erratic variations in the field strength below the horizon make it impossible to determine a rate of attenuation. In spite of their limitations, the data further confirm the prediction that with a deep surface duct ranges within the duct are greatly extended and that even above the duct extended ranges and superstandard fields prevail.

Vertical soundings at short ranges were made at times when a deep surface duct was prevalent off the Welsh coast (see Sec. 4·17) and near Sydney, Australia.[1] The former observations indicated that in the 1- to 10-cm wavelength range, the presence of a duct tended to shift the bottom lobes toward the surface and to break up the lowest lobe into two parts, the lower of which remained very close to the surface and presumably represented trapping by the duct. The Australian observations consisted of vertical field-strength sections up to 10,000 ft at a range of 25 miles using a wavelength of 1.5 m. When a deep surface duct was presumably present,[2] the lowest two lobes appeared to be shifted downward in height from 10 to 20 per cent. Higher lobes were unaffected.

One of the most important practical questions arising from the existence of nonstandard refraction is whether or not it can produce serious gaps in the standard coverage diagram. When this question was first considered, it was treated by ray-tracing methods before the limitations of ray tracing described in Sec. 2·5 were generally known and before the wave solution was available. The existence of discrete values of critical angle of penetration of a superrefracting layer indicated that serious loss of coverage might occur at sufficiently low elevation angles. In the light of later developments it is clear that this simple ray concept is totally inadequate and that the field should be considered as the superposition of several transmission modes.[3] Unfortunately the numerical labor involved in

[1] "Investigation of Propagation Characteristics of A. W. Stations," ORG (Australia) Report No. 17, Mar. 9, 1943.

[2] Temperature measurements alone were available to correlate with the radio data. The situation referred to corresponds to observations of strong temperature inversions. The experiment was carried out prior to the development of satisfactory sounding equipment.

[3] It should be remembered, however, that the region in which refraction can produce noticeable effects occurs at very small elevation angles; consequently only those parts of the coverage diagram in the bottom of the interference region and in the intermediate and diffraction regions are of interest.

applying the wave solution in this region is so great that it has not been explored. As the theory has not yet been applied numerically, the only information available consists of the experiments described in the preceding sections, which must certainly be considered fragmentary evidence.

Both the British and Australian observations were made at very short ranges, and it is impossible to extrapolate the results to greater ranges with assurance. Although the measurements show displacement of the lowest lobes and the building up of an extra lobe near the surface, they do not reveal any conspicuous gaps in the interference pattern. The vertical section made off New England (Fig. 4·50) is at the other extreme from these short-range sections, having been made at so great a range that the entire section is below the horizon. Certainly leaky modes are important in this case; for strongly superstandard fields prevail to at least four or five times the duct height, and no "blind zone" is evident. On the other hand the 500-ft horizontal section (Fig. 4·52) shows lower than standard fields through the first two lobes. Extensive measurements remain to be made for gaining adequate knowledge of field-strength distributions in the presence of a deep duct, but in view of this (small) amount of information it does not seem probable that serious gaps occur in this situation. Theoretical coverage diagrams in the diffraction region for certain deep duct conditions are shown in Sec. 2·21.

4·22. Elevated *M*-inversions. *California Coast.*—In conjunction with their investigation of propagation affected by an elevated *M*-inversion, the Navy Electronics Laboratory at San Diego made vertical soundings of field strength along the path indicated in Fig. 4·2.[1] The soundings were made over water at heights up to 5000 ft and at ranges up to 143 miles. The land terminals were 100 ft above the surface, and wavelengths of 4.8 and 1.8 m, and 57 and 9 cm were used. Meteorological soundings, made immediately before or after each radio sounding, showed the elevated *M*-inversion characteristic of this region, but in every case the *M*-profile was decidedly nonuniform along the path.

When the subsidence inversion was low enough to cause a surface duct, the performance on the various wavelengths was similar to that predicted from a bilinear index model. The vertical sections are illustrated in Fig. 4·53, which shows the field strength observed on the four wavelengths at 37, 69, 110, and 143 miles; the measured *M*-profiles are also indicated at the appropriate range along the path. The 4.8-m radiation was partially trapped; the signal level was independent of height above about 500 ft and was slightly below the free-space value. The signal level decreased below 500 ft. The high field strength above 500 ft persisted well below the horizon, out to at least 143 miles, indicating that

[1] L. J. Anderson *et al.*, "Atmospheric Refraction, A Preliminary Quantitative Investigation," NRSL Report No. WP-17, Dec. 28, 1944.

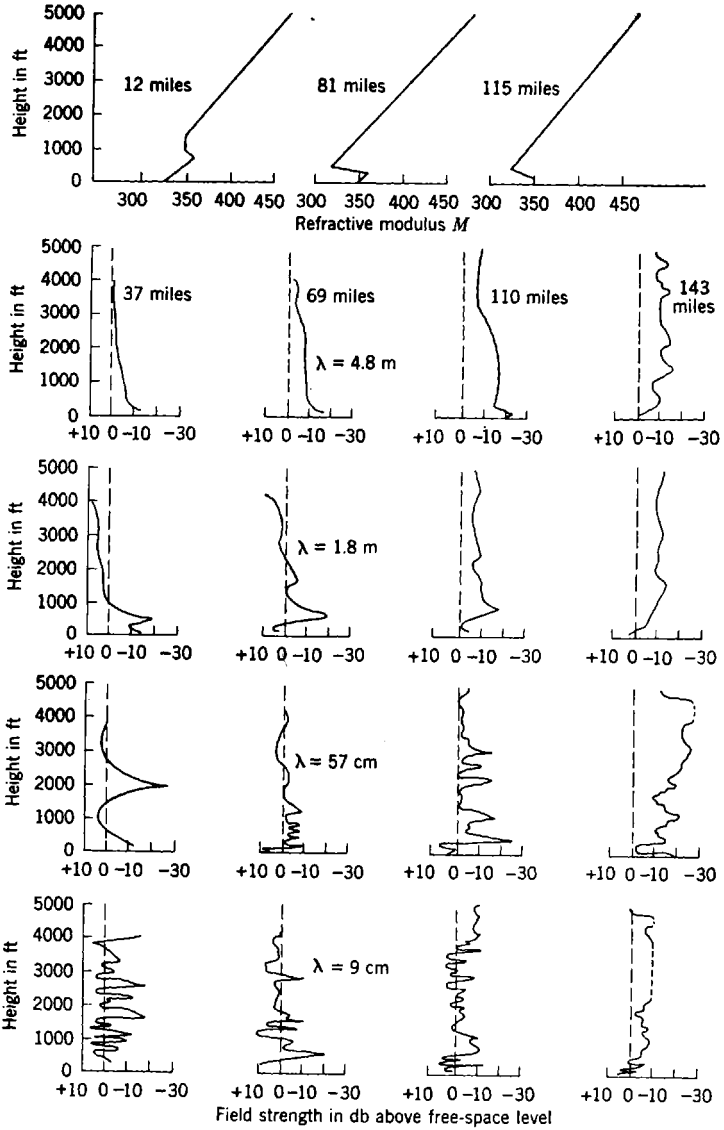


FIG. 4·53.—Vertical field-strength soundings and associated M-profiles measured off San Diego, Calif., by NEL on Oct. 2, 1944. The ranges at which the soundings were made are indicated on the M-profiles and on the top field-strength sections.

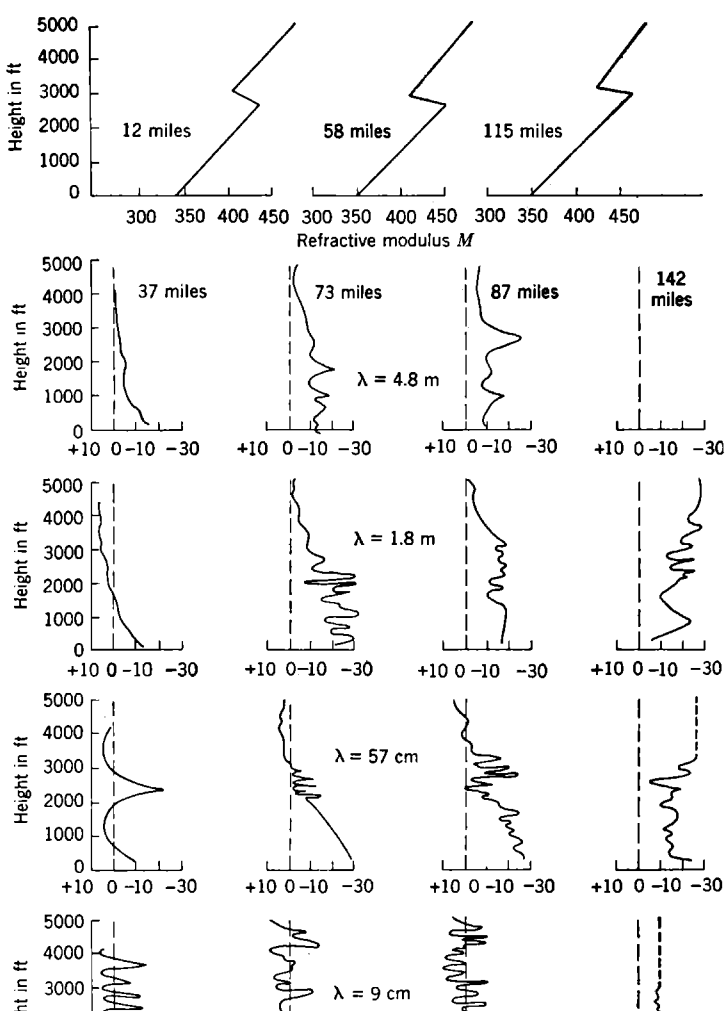


FIG. 4·54.—Vertical field-strength soundings and associated M-profiles measured off San Diego, Calif., by NEL on Sept. 29, 1944. The ranges at which the soundings were made are indicated on the M-profiles and on the field-strength sections.

the attenuation factor was small. It is uncertain whether or not the 1.8-m radiation was completely trapped. On the average the signal level was near the free-space level, but there was a tendency toward fading and variable signal below the horizon, particularly in the first 1000 ft. Complete trapping apparently occurred on 57 and 9 cm. The signal was in general near the free-space level on the average but was characterized by fading that could not be reproduced on repeat soundings. The highest signal levels were found close to the surface at great distances.

Figure 4·54 shows the field strengths observed when the duct was aloft, between 2000 and 3000 ft. In this case the lowest frequency in general gave the greatest field strength below the horizon. The 4.8-m signal level remained 10 to 20 db below free-space level on the average, even well below the horizon, and was characterized by fading. The other circuits showed the signal level decreasing with decreasing height and often falling below minimum detectable level well above the surface at sufficient distance beyond the horizon. The signal was superstandard on the shorter wavelengths, however. On 57 cm, the signal was still detectable at 142 miles below 3000 ft, and the lowest lobe of the interference pattern was depressed below the horizon at 9 cm.

The data presented in Figs. 4·53 and 4·54 are consistent with the observations on the San Pedro–San Diego one-way path (see Sec. 4·9). In the case of a duct aloft, the lowest frequencies were affected most, and unsteady signal was present on all frequencies below the horizon. Similarly, the one-way data corroborate the observation of trapping on wavelengths up to about 3 m and steady signal on longer wavelengths when a surface duct is present.[1]

ANGLE MEASUREMENTS ON SHORT OPTICAL PATHS

By Pearl J. Rubenstein and Donald E. Kerr

The experiments described thus far have dealt with the field strength at a specified position or series of positions and have considered only the amplitude and phase of the elementary waves that combine to give the total field. Additional information on the effects of refraction may be obtained by study of the apparent direction of arrival of these waves at a fixed receiving point. The angle of arrival is of interest because it may aid in interpretation of refractive effects in the region where wave theory is hopelessly complicated and ray theory is inadequate. It also has a bearing on possible refraction errors in radar height measurements.[2]

[1] For an analysis of some later experiments see C. L. Pekeris, *J. Applied Phys.*, **18**, 838 (1947).

[2] The work described here deals with angle of arrival in the vertical plane; for studies of horizontal angle of arrival see A. W. Straiton and J. R. Gerhardt, *Proc. IRE*, **36**, 916 (1948).

4·23. Measurements of Angle of Arrival.—Perhaps the most serious technical problem involved in measuring the very small angular differences involved in this problem is that of building an antenna with the necessary resolving power. The first successful solution was obtained in 1944 by Bell Telephone Laboratories.[1] A series of measurements was made on wavelengths of 3.25 and 1.25 cm, with receiving antennas giving fan beams that were very narrow in the plane in which the angle of arrival was measured. The 3.25-cm antennas employed modified "pillbox" parabolic reflectors and had half-power beamwidths of 0.36° and 15° in the planes perpendicular to and parallel to the large and small dimensions, respectively. For the 3.25-cm measurements two of these antennas were mounted at right angles to each other, one being used to measure the elevation angle of arrival and the other the azimuthal angle. The antenna for 1.25 cm was of the "metal-lens" type,[2] having principal beamwidths of 0.12° and 1.2°. The transmitted wave on both wavelengths was polarized at an angle of 45° and was radiated from antennas having beamwidths of 4° and 2.7° on 3.25 and 1.25 cm, respectively.

The angle of arrival of incoming waves was measured by rocking each antenna through an arc of $\pm 0.75°$ and recording the signal output of an associated receiver as the antenna beam scanned the incoming wave pattern. The angle of arrival was determined from the position of the resulting signal peak on the record. The length of a complete scanning cycle was 20 sec. Differences in angle of arrival of about 0.02° could be measured on 3.25 cm, but absolute accuracy in angular position was limited to about 0.05°. On 1.25 cm the corresponding accuracies were slightly better.

The 3-cm antennas were mounted on a rotating platform 350 ft above sea level at Beer's Hill, N.J., and could be directed toward transmitters in New York City, 24 miles distant, and in Deal, N.J., 12.6 miles distant. The New York circuit was principally over water, with the point of reflection in Raritan Bay. For standard refraction the reflected ray cleared the top of Staten Island by less than the radius of the first Fresnel zone, a fact that was probably at least partly responsible for the small amplitude of the reflected wave. The direct ray from transmitter to receiver was well away from obstructions, however. The Deal circuit was free from obstructions, over gently rolling land having an effective reflection coefficient on 3 cm of about 0.18 (see Sec. 5·9).

Signal was also received on 3.25 cm by a paraboloidal reference antenna with a beamwidth of 4°, allowing a comparison of the combined

[1] A detailed description of the equipment and the two series of measurements summarized here may be found in *Proc. IRE*, **34** (1946): W. M. Sharpless, "Measurement of the Angle of Arrival of Microwaves," p. 837, and A. B. Crawford and W. M. Sharpless, "Further Observations of the Angle of Arrival of Microwaves," p. 845.

[2] W. E. Kock, "Metal-lens Antennas," *Proc. IRE*, **34**, 828 (1946).

waves received over a large angular range with those received by the narrow-beam antennas in a much smaller range.

The 1.25 cm equipment was operated only over the 12.6-mile path from Beer's Hill to Deal. The receiving antenna was used to scan in either the vertical or horizontal plane, as it could be rotated through an angle of 90°.

The first series of measurements (Sharpless, *loc. cit.*) was made with the 3.25-cm equipment over both the 24- and 12.6-mile paths, primarily in the daytime in the summer and fall of 1944. The second series, in the summer and fall of 1945, was made with both the 3.25- and the 1.25-cm equipment but over only the 12.6-mile path. The periods of observation in the second case ranged from $\frac{1}{2}$ to 8 hr in length on most of the days and on 20 clear, calm nights when it appeared that nonstandard propagation was likely to occur.

No meteorological soundings were available in 1944, but during the 1945 measurement period 220 soundings were made at Holmdel, N.J., near Beer's Hill. In addition, other meteorological data were obtained on a 400-ft tower at Oakhurst, N.J., near Holmdel.[1]

The principal results were as follows:

1. Under conditions of standard refraction (indicated by steadiness of signal from the reference antenna), the measured azimuth and elevation angles were steady and coincided with their predicted values.[2] On both the New York and Deal circuits the ground-reflected waves were weak in comparison with the direct waves, and their elevation angles were so near those of the direct waves that they could be only partially resolved.
2. When measurable fluctuations in elevation angle of arrival occurred, they were usually much more persistent and much larger than those in azimuth. No correlation was found between deviations in the horizontal and vertical planes.
3. Deviations from true azimuth occurred much less frequently than deviations from true elevation, and the magnitude of the deviations was much smaller. The greatest azimuth deviation observed was 0.1° in the 1944 tests. In the 1945 tests on the short path the largest deviation observed was 0.03°.
4. Under conditions of nonstandard refraction the direct and reflected waves appeared to be affected differently, but there was no apparent correlation between the observed variations in their angles of

[1] The meteorological results and their relation to the measured angles of arrival are given in detail in W. E. Gordon and A. T. Waterman, Jr., "Angle-of-arrival Aspects of Radio-meteorology," Electrical Engineering Laboratory Report No. 1, University of Texas, Austin, Tex., Mar. 23, 1946.

[2] On the New York circuit the difference between the true elevation angle (for no refraction) and that for standard refraction was 0.04°, which was less than the probable accuracy of measurement for 3.25 cm, and it is therefore neglected.

arrival. On the New York circuit the reflected wave at times *increased* in strength to a value approximating that of the direct wave, and its angle of arrival *decreased* below its standard value. On one occasion this decrease was as much as 0.17°. At such times the phase between direct and reflected waves appeared to vary slowly from in phase to out of phase; on one such occasion the signal from the broad-beam reference antenna varied by as much as 20 db, but the output of the sharp-beam antenna varied less than 4 db when that antenna was pointed directly toward the incoming direct wave.

5. In a second type of nonstandard behavior, observed on both New York and Deal circuits, the angular elevation of the direct wave usually increased markedly above its standard value but sometimes decreased to a value slightly below standard. On two nights in 1945 the signal on the broad-beam antennas and the apparent angle of arrival given by the sharp antennas fluctuated violently, and the scanning record showed that the received signal resulted from several components arriving simultaneously at different elevation angles. The largest upward deviation, observed on the second of these nights, was 0.75° on 1.25 cm, and on several other occasions in 1944 deviations of half as much were observed on the New York path. On the first of the unusual nights there appeared to be two components of incoming signal; one had the amplitude and direction of the normal daytime signal, but the other fluctuated in strength by about 20 db, and its angle of arrival varied from 0.21° to 0.46° above the normal component. Soundings at Holmdel revealed an elevated M-inversion near the line of sight. On the second night, "two, three, and, at times, four separate transmission paths were observed. All of these varied in angle of arrival and all had large, rapid variations in signal level; the normal or daytime path sometimes was missing." Soundings showed evidence of two elevated inversions near the line of sight.
6. Except for the two nights of multipath transmission, the variation of vertical angle of arrival in 1945 was from $-0.04°$ to $+0.11°$, occurring mostly during the nighttime hours. A rough correlation was found on both wavelengths between angle of arrival and slope of (presumably relevant) sections of the M-profiles occurring at or about the time of the observations.[1]
7. There was good agreement between angles of arrival measured simultaneously on 1.25 and 3.25 cm, within the range $-0.04°$ to $+0.12°$.

[1] See Crawford and Sharpless, *op. cit.*, Fig. 2. The theoretical basis for this procedure is considered in the next section.

8. "Fading ranges observed with the broad-beam antenna were usually less than 6 db, except on the nights of multiple-path transmission. Scintillation fading (rapid fluctuations of from $\frac{1}{4}$ to $1\frac{1}{2}$ db about a steady average signal level) was usually present during the daytime and on windy nights, and was more severe at 1.25 cm than at 3.25 cm. Observations made at 1.25 cm showed that the scintillation fading was generally less on the large narrow-beam scanning antenna (held in a fixed position for this test) than on the small broad-beam antenna; this observation is analogous, perhaps, to the optical one in which star scintillation is less when viewed through a large telescope than when seen by the unaided eye."
9. On the New York path, for which there was an appreciable reflected wave, when the reflected wave was strongly deviated but the direct wave was not affected, the very narrow antenna pattern resulted in a material improvement in stability of the received signal, as the antenna could be fixed in position to receive the relatively steady direct wave and discriminate against the fluctuating reflected wave. On the other hand, when the direct wave experienced upward angular deviations, these deviations were so large that no single antenna position would allow reception free from marked signal fading. This behavior suggests that there is an upper limit to antenna directivity for one-way systems employing fixed antennas and also that a slight upward tilt of a sharp-beam antenna may be desirable to improve signal stability.
10. Changing the polarization of the transmitted wave failed to reveal a measurable dependence of the fading on polarization, and rotating the reference receiving antenna polarization until it was at right angles to the transmitted polarization resulted in a decrease in signal of more than 20 db even during periods of marked refraction, indicating that the wave suffers negligible depolarization in the refraction process.

These experiments suggest a new and powerful approach to the study of atmospheric refraction, and it is to be hoped that they will eventually yield information that will aid in filling the large gaps in existing knowledge of refraction effects.

A different method of attacking the angle-of-arrival prob.em is now on trial by the University of Texas.[1] The site for their experiments was the

[1] E. W. Hamlin, W. E. Gordon, and A. H. LaGrone, "X-band Phase Front Measurements in Arizona during April 1946," Electrical Engineering Research Laboratory Report No. 6, University of Texas, Austin, Tex., Feb. 1, 1947. This report was received just before the foregoing material was sent to press, and only a brief summary of the work is possible here. See also E. W. Hamlin and W. E. Gordon, Proc. I.R.E., **36,** 1218 (1948), and E. W. Hamlin and A. H. LaGrone, *Jour. Applied Phys.*, **21,** 188 (1950).

Gila Bend terminal of the NEL transmission experiment listed in Table 4·2 of Sec. 4·12. This site offered numerous advantages, such as good physical facilities for making the desired measurements, meteorological conditions favorable to formation of strong surface-temperature gradients, and sufficiently flat terrain that ground reflection could be studied without undue effort.

Instead of using a single antenna with a very narrow beam, this second method employs two small 3-cm antennas spaced many wavelengths (10 ft) apart on a nearly vertical axis; the signals from the two antennas are combined in appropriate phasing networks and receiving and indicating systems. The antennas are oriented in the same direction, perpendicular to this axis, which can then be tilted in a plane perpendicular to the earth to place the antenna mouths in an equiphase plane. If the incident wave is strictly a plane wave, in this position a line perpendicular to the axis and in the direction of orientation of the antennas yields the direction of the normal to the incident wavefront and hence tilt of the wavefront or the angle of arrival. In practice this axis remains fixed in position, and variations in phase difference between the two antennas is interpreted in terms of variations of the angle of arrival. The two antennas are moved together in the vertical direction on a tower.

This method assumes that the incident wavefront is indeed plane; if it is not, the angle of arrival of the wave loses its meaning. If the phase front is more complicated, as, for instance, in the case of a direct and an earth-reflected wave, this two-antenna method measures the phase difference between two points separated by the antenna spacing in the total wavefront resulting from superposition of the two waves; this phase difference may be very different from that of either the direct or reflected wave alone.[1] Certain auxiliary measurements are necessary to yield the angle of arrival of the direct (or other) rays. If, moreover, refraction (or other processes) causes multiple-path transmission to occur, the two-antenna method may yield results that are very difficult to interpret. On the other hand, with appropriate elaboration the method may serve as the foundation for the investigation of both amplitude and phase over an extended section of a complicated wavefront, from which the response of any desired antenna or antenna system may be predicted, in principle at least. The experiments described in the report cited were undertaken primarily to investigate the rapidity of variation of the phase front. They have not yet progressed to the point of obtaining direct information on angle of arrival, except by interim methods mentioned later.

The results available at present, although preliminary, are interesting and suggestive. Briefly, they are as follows:

[1] A derivation of the field for this case (including phase) is given in Sec. 6·5 in connection with scattering from a vertical plane sheet. See also E. W. Hamlin and A. W. Straiton, *Proc. IRE*, **36**, 1538 (1948).

1. In the interference region at points not too close to the interference minima, the phase difference between the two antennas 10 ft apart vertically varied less than 15° on fair days with a period of several minutes. On fair nights a rapid oscillation of about 5° was superposed on a drift of as much as 5° per min. During rain and near sunrise large and rapid variations occurred, and the data were often not reproducible because of the rapid fluctuations.
2. When the M-profile was essentially linear, the measured phase difference in the interference region agreed satisfactorily with that computed by ray theory.
3. When ducts occurred with depths of 100 ft or less and with M-gradients of -3 to -5 per hundred ft, fairly good agreement was obtained between the measured angle of arrival and that computed by tracing rays into the diffraction region. Estimations of the angle of arrival of the direct ray were made for four nights on which weak surface ducts occurred, using an approximate scheme for absolute phase calibration. The angles of arrival calculated in this way were between $-0.08°$ to $+0.08°$ relative to the direct ray for standard refraction.[1]
4. The apparent reflection coefficient of the earth showed a marked diurnal variation between limits of 0.85 in the afternoon and 0.2 in the early morning. This variation showed a definite correlation with measured low-altitude M-gradients. It was attributed to a combination of the effects of extreme refraction (both super and subrefraction) in a very thin layer of air at the surface and a Fresnel pattern of reflection peculiar to the path profile.

4·24. Theoretical Discussion.—As existing wave theory for the field in the interference region requires summation of an inordinately large number of terms, the alternative is to employ ray theory, with appropriate regard for the limitations discussed in Sec 2·5. Unfortunately these limitations are encountered in all but the simplest phases of the present problem because of the occurrence of foci and caustics, but it is nevertheless instructive to examine the ray-tracing equations.

In Sec. 2·3, Snell's law expressed in spherical coordinates was shown to be

$$(a + z)n \cos \beta = \text{const.}, \tag{3a}$$

where n is the index of refraction at the height z, and a, z, and β are as shown in Fig. 4·55. If the expression above is divided by a, one obtains

$$N \cos \beta = N_1 \cos \beta_1, \tag{3b}$$

where N is modified index and the subscript denotes values at a fixed

[1] For details the original report should be consulted. This procedure probably violates some of the restrictions on ray tracing, but the matter has not been investigated fully

reference point, which for convenience may be one transmission terminal. This expression can be rewritten as

$$\tan \beta = \frac{\sqrt{\left(\frac{N}{N_1}\right)^2 - \cos^2 \beta_1}}{\cos \beta_1} = \frac{dz}{dr}, \tag{3c}$$

one possible form for calculation of ray paths through an atmosphere in which N is an arbitrary function of z.

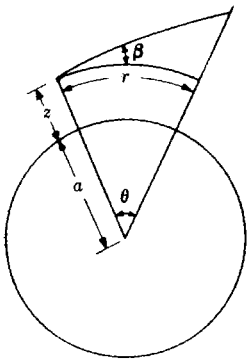


FIG. 4·55.—Path of a ray at an angle β with the horizontal.

In Sec. 2·4 it was shown that a positive value of effective earth's radius a_e can be used if N^2 varies linearly with z; specifically, in terms of the present problem, if

$$\left(\frac{N}{N_1}\right)^2 = 1 + \frac{2(z - z_1)}{a_e},$$

Eq. (3) becomes

$$\tan \beta = \frac{dz}{dr} = \frac{\sqrt{\sin^2 \beta_1 + \frac{2(z - z_1)}{a_e}}}{\cos \beta_1}, \tag{4}$$

which can be integrated directly to give

$$z - z_1 = \frac{r^2}{2a_e \cos^2 \beta_1} + r \tan \beta_1. \tag{5}$$

As the angles of interest here are very small, the usual small-angle approximations give, for Eq. (5),

$$z - z_1 = \frac{r^2}{2a_e} + r \beta_1, \tag{6a}$$

or

$$\beta_1 = \frac{z - z_1}{r} - \frac{r}{2a_e}. \tag{6b}$$

This value of β_1 can be inserted into Eq. (4), which becomes

$$\beta = \frac{z - z_1}{r} + \frac{r}{2a_e} = \beta_1 + \frac{r}{a_e}. \tag{7}$$

If z_1 is the height of one terminal, β_1 is the angle of arrival (or departure) of a ray from the other terminal at a height z and a distance r away, where β was the angle of departure (or arrival).

Equations (4) through (7) constitute no more than a restatement of the ray-tracing equations for the assumed variation of N, which is equivalent to a linear N- or M-profile. The relation between the equivalent

earth radius a_e and the slope of the M-profile is

$$\frac{1}{a_e} = \frac{1}{a} + \frac{dn}{dz} = \frac{dN}{dz} = \frac{dM}{dz}\, 10^{-6}, \tag{8}$$

where dN/dz is independent of z (and of r, as horizontal stratification is assumed in any discussion involving N). These restrictions on dN/dz must be fulfilled if rays traced by Eqs. (4) through (7) are to give a complete description of the field in the interference region. If they are not fulfilled, rays may still be traced, but the resulting distortion of wavefronts and lack of strict coincidence of ray paths with the direction of energy flow pose theoretical problems in interpretation of the ray patterns that are as yet unsolved. Physical intuition suggests that if the M-profile is essentially linear over a layer roughly coinciding with the transmission terminals but with a slope differing from that outside the layer, Eq. (6) should give an estimate of apparent angle of arrival of rays traveling in the layer. In this calculation a_e should be obtained from Eq. (8), using the M-gradient in the layer. This argument is undoubtedly correct in principle if the wavelength is sufficiently short that geometrical optics is valid, but no general quantitative tests are available to determine its accuracy in specific cases. (This point has also been discussed in Sec. 4·7.)

It has been emphasized that after rays pass through caustics, cusps, or foci, the ray paths and direction of energy flow in general do not coincide. In regions traversed only by rays that have not yet encountered these singularities in the ray pattern, however, the pattern is a reliable measure of field distribution, even with negative slope of the M-profile, if reflections from the earth's surface are negligible. In this case the angle of arrival may be calculated from Eq. (6). If the value of a_e obtained from Eq. (8) is constant throughout the height range covered by the ray, Eq. (6) may be used directly; otherwise step-by-step integration of Eq. (3) is necessary.

The angular deviations to be expected may be illustrated by considering a bilinear M-profile in which the slope of the bottom section varies from superstandard to substandard and the upper section has a positive slope. The critical height at which the slope changes will be z_c, and the values of effective earth's radius [from Eq. (8)] for the lower and upper sections will be a_1 and a_2, respectively. Assuming that the receiver at which angle of arrival is to be determined is at a height z_1, less than z_c, rays may be traced from the receiver into the refracting region.

In the region below z_c, the path of a ray departing at an angle β_1 is, from Eq. (6),

$$z = \frac{r^2}{2a_1} + r\,\beta_1 + z_1. \tag{9}$$

When it reaches the height z_c, it has traveled a horizontal distance r_1, which can be obtained from Eq. (9),

$$r_1 = -a_1\left(\beta_1 - \sqrt{\beta_1^2 + \frac{2(z_c - z_1)}{a_1}}\right). \tag{10}$$

Upon reaching z_c the slope of the ray is, from Eq. (6),

$$\beta = \beta_1 + \frac{r_1}{a_1}. \tag{11}$$

This value of β serves as the angle of departure β_1 for the ray as it progresses above z_c, where the path of the ray is given by

$$z - z_c = \frac{(r - r_1)^2}{2a_2} + (r - r_1)\left(\beta_1 + \frac{r_1}{a_1}\right). \tag{12}$$

If r_1 from Eq. (10) is inserted into Eq. (12), one finds, after some algebraic manipulation, that

$$z = \frac{a_1^2}{2a_2}\left[\frac{r}{a_1} + \beta_1 + \left(\frac{a_2}{a_1} - 1\right)\sqrt{\beta_1^2 + \frac{2(z_c - z_1)}{a_1}}\right]^2 - \frac{a_2}{2}\beta_1^2 - \left(\frac{a_2}{a_1} - 1\right)z_c + \frac{a_2}{a_1}z_1. \tag{13}$$

An alternative form is useful for the case in which the lower segment of the M-profile has negative slope and a_1 is negative. In this case trapping may occur at angles below a critical angle β_c given by

$$\beta_c^2 = -\frac{2(z_c - z_1)}{a_1},$$

and Eq. (13) becomes

$$z = \frac{a_1^2}{2a_2}\left[\frac{r}{a_1} + \beta_1 + \left(\frac{a_2}{a_1} - 1\right)\sqrt{\beta_1^2 - \beta_c^2}\right]^2 - \frac{a_2}{2}(\beta_1^2 - \beta_c^2) + z_c. \tag{14}$$

Equations (13) and (14) are written in terms of the actual angle of arrival (or departure) β_1, which differs from β_1', the angle of arrival of a ray traveling between the same terminals with *no* refraction. The path of the latter ray is given by

$$z = \frac{r^2}{2a} + r\beta_1'. \tag{15}$$

The refraction error $\delta = \beta_1 - \beta_1'$ is most easily obtained by equating Eqs. (13) and (15) and solving for β_1', which is then expressed in terms of the actual angle of arrival, terminal height, and properties of the M-profile.

Figure 4·56 shows the angular error δ caused by several types of M-profiles.

Two special cases are of interest. If $a_2 = a_1$, it is easily shown that

$$\delta = \frac{r}{2a_2}\left(\frac{a_2}{a} - 1\right). \tag{16}$$

For standard refraction this gives a value for δ of 1.81×10^{-3} degrees per mile. The second special case is that for which

$$\beta_1^2 \gg \left|\frac{2(z_c - z_1)}{a_1}\right|,$$

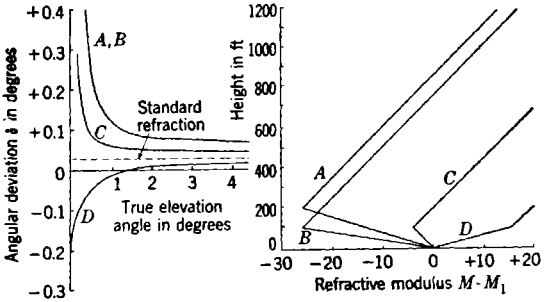


FIG. 4·56.—Angular error as a function of true elevation angle (for no refraction) for various M-profiles. The range is 15 miles, and M_1 is value of M at the receiver, which is at zero height.

corresponding to the nearly horizontal sections of the curves of Fig. 4 56. Then δ becomes

$$\delta = \frac{r}{2a_2}\left\{\frac{a_2}{a} - \left[1 + \left(\frac{z_c - z_1}{r\beta_1}\right)\left(\frac{a_2}{a_1} - 1\right)\right]^2\right\}. \tag{17}$$

The relations developed above apply only to untrapped rays penetrating regions above the critical heights of the simplified M-profiles of Fig. 4·56. It is probable that in practical situations the refraction conditions are more complicated than those pictured above and that observed increases in angle of arrival may arise from either untrapped rays or strongly trapped rays from surface or elevated layers. Trapped rays, with their attendant complex ray patterns, are a possible cause of the multiple-ray paths that have been observed occasionally. Further analysis of the angle of arrival problem without considerable extension of existing theoretical methods and without more detailed meteorological information does not appear hopeful, however.

CHAPTER 5

REFLECTIONS FROM THE EARTH'S SURFACE

By Donald E. Kerr, William T. Fishback, and Herbert Goldstein

The effects of reflection of waves from the earth's surface were discussed qualitatively in Sec. 2·2. In Sec. 2·11 formulas were given for the reflection coefficient and divergence factor in terms of the earth's electromagnetic properties and of the transmission-path geometry. In this chapter these quantities and their effect on the coverage diagram will be discussed further. Surface roughness will be shown to cause marked departures from the results predicted theoretically for a smooth surface; the extent of this departure is related to the size of the surface-roughness elements relative to wavelength and to the grazing angle. Reflections from the surface will be shown to limit accuracy of radar height measurements at angles of elevation so small that the antenna illuminates the surface strongly and specular reflection from the surface is large.

THEORY OF SPECULAR REFLECTION

By Donald E. Kerr

5·1. Fresnel's Equations for a Smooth Plane Surface.—Equations (2·366) and (2·367) for the reflection coefficient Γ can be rewritten as

$$\Gamma_v = \rho_v e^{-i\phi_v} = \frac{\left(\frac{k_1}{k_0}\right)^2 \sin\psi_2 - \sqrt{\left(\frac{k_1}{k_0}\right)^2 - \cos^2\psi_2}}{\left(\frac{k_1}{k_0}\right)^2 \sin\psi_2 + \sqrt{\left(\frac{k_1}{k_0}\right)^2 - \cos^2\psi_2}} \tag{1}$$

and

$$\Gamma_h = \rho_h e^{-i\phi_h} = \frac{\sin\psi_2 - \sqrt{\left(\frac{k_1}{k_0}\right)^2 - \cos^2\psi_2}}{\sin\psi_2 + \sqrt{\left(\frac{k_1}{k_0}\right)^2 - \cos^2\psi_2}}. \tag{2}$$

These expressions are well enough known not to require derivation here; they may be found in any standard work on electromagnetic theory.[1]

[1] For instance, see J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, Secs. 9·4 and 9·9. Stratton's ρ is equivalent to the ρ used here, but he defines reflection coefficient as ρ^2.

This section will consider the dependence of Γ upon the electromagnetic properties of the earth, grazing angle ψ_2, and polarization.[1]

For horizontal polarization Γ_h is the ratio of reflected to incident electric field or the ratio of the vertical components of the magnetic field, but the ratio of reflected to incident horizontal components of the magnetic field is $-\Gamma_h$. Similarly, for vertical polarization Γ_v is the ratio of reflected to incident magnetic field, or the ratio of the vertical components of the electric field; but the ratio of the horizontal components of electric field is $-\Gamma_v$. As long as ψ_2 is small, the field is practically vertical and Γ_v may be considered the ratio of electric fields, but for large ψ_2 the horizontal and vertical components of field must be considered separately in relation to Γ_v.

The way in which the electromagnetic properties of the earth affect Γ are revealed through the propagation factor $k_1^2 = \omega^2\mu\epsilon - i\omega\mu\sigma$. For the present purpose k_0^2 for the air at the surface differs negligibly from the value for free space, $\omega^2\mu_0\epsilon_0$, and μ for both the air and the earth is essentially μ_0. Then

$$\left(\frac{k_1}{k_0}\right)^2 = \frac{\epsilon}{\epsilon_0} - i\frac{\sigma}{\omega\epsilon_0} = \varepsilon_1 - i\varepsilon_2 = \varepsilon_c. \tag{3}$$

This quantity is ε_c called the *complex dielectric constant,* and ϵ/ϵ_0 is the usual dielectric constant of the earth, denoted by ε_1. In mks units $\varepsilon_2 = \sigma/\omega\varepsilon_0$ becomes $60\lambda\sigma$ when λ is expressed in meters and σ in mhos per meter. Then

$$\varepsilon_c = \varepsilon_1 - i60\lambda\sigma = \varepsilon_1\left(1 - i\frac{1}{Q}\right), \tag{4}$$

where $Q = \varepsilon_1/\varepsilon_2$, the ratio of displacement current to conduction current. (The factor Q is analogous to the "Q" of a parallel combination of a lossless capacitor and a loading resistor.) It is a useful index of the degree to which a transmission medium tends to behave as a high-quality dielectric: The higher Q the lower the loss of energy by heating. The reciprocal of Q is also the tangent of the argument of Eq. (4) and is frequently used in describing the properties of transmission media.

For wavelengths greater than roughly 20 cm the properties of sea water are essentially independent of wavelength. Both ε_1 and σ vary noticeably with temperature, ε_1 decreasing and σ increasing as temperature increases. The value of ε_1 for 20°C is commonly taken to be 80, but σ is subject to considerable additional variation because of local variations in salinity. The value 4.3 mhos/m will be used here. Below about 20 cm, ε_1 decreases with decreasing wavelength, and σ increases roughly as shown in Table 5·1.

[1] The terminology of polarization is not precise but is adhered to because of usage in radio engineering. "Horizontal" and "vertical" polarization as used here mean that the electric vector is perpendicular to or lies in the plane of incidence.

For land surfaces both ε_1 and σ are in general much lower than for water. For dry soil ε_1 may fall as low as 2 but is more likely to lie in the range from 3 or 4 to 20. The conductivity of land may vary from roughly 10^{-5} to 10^{-2} mhos/m, the smaller values tending to occur simultaneously with low values of ε_1 for dry, rocky, or sandy soil, whereas the higher values of both σ and ε_1 occur with rich soils. This wide range of conductivity and dielectric constant results in a considerable variation in ε_c, which in turn produces a marked effect on the reflection coefficient, particularly for vertical polarization.

The values shown in Table 5·1 are intended to serve only as a guide rather than as an accurate set of earth parameters. They represent a composite of values obtained from widely varying sources and reflect both the variability to be found in nature and the incomplete state of knowledge in the short-wavelength region.

TABLE 5·1.—APPROXIMATE ELECTROMAGNETIC PROPERTIES OF SOIL AND WATER*

Medium	λ	σ mho/m	ε_1	ε_2	Q
Sea water	3 m–20 cm	4.3	80	774	0.10
				52	1.5
20°–25°C	10 cm	6.5	69	39	1.8
28°C	3.2 cm	16	65	30.7	2.1
Distilled water, 23°C	3.2 cm	12	67	23	2.9
Fresh-water lakes	1 m	10^{-3}–10^{-2}	80	0.06	1330
				0.60	133
Very dry sandy loam	9 cm	0.03	2	1.62	1.23
Very wet sandy loam	9 cm	0.6	24	32.4	0.74
Very dry ground	1 m	10^{-4}	4	0.006	670
Moist ground	1 m	10^{-2}	30	0.6	50
Arizona soil	3.2 cm	0.10	3.2	0.19	17
Austin, Tex., soil, very dry	3.2 cm	0.0074	2.8	0.014	200

* L. H. Ford and R. Oliver, "An Experimental Investigation of the Reflection and Absorption of Radiation of 9-cm Wavelength," DSIR, RRB/C107, Oct. 27, 1944, and *Proc. Phys. Soc.*, **58**, 265 (1946); C. R. Burrows, "Radio Propagation over Plane Earth-Field Strength Curves," *Bell System Tech. Jour.*, **16**, January 1937; F. E. Terman, *Radio Engineering Handbook*, McGraw-Hill, New York, 1943; *Tables of Dielectric Materials*, Laboratory for Insulation Research, Massachusetts Institute of Technology, Vol. 1, February 1944; Vol. 2, June 1945; A. W. Straiton and C. W. Tolbert, "Measurement of Dielectric Properties of Soils and Water at 3.2-cm Wavelength," Electrical Engineering Research Laboratory, Report No. 4, University of Texas, July 10, 1947. The last report was received too late to quote in detail. The 3.2-cm values in the table are taken from this report.

It can be seen that except for very dry ground,

$$|\varepsilon_c| \gg 1. \tag{5}$$

Introducing this approximation into Eqs. (1) and (2) yields

$$\Gamma_v = \rho_v e^{-i\phi_v} \approx \frac{\sqrt{\varepsilon_c}\sin\psi_2 - 1}{\sqrt{\varepsilon_c}\sin\psi_2 + 1} \tag{6}$$

and

$$\Gamma_h = \rho_h e^{-i\phi_h} \approx \frac{\sin \psi_2 - \sqrt{\varepsilon_c}}{\sin \psi_2 + \sqrt{\varepsilon_c}} \tag{7}$$

As pointed out in Sec. 2·6, the radicals in Eqs. (1) and (2) must have negative imaginary parts; that is, the principal root must be taken in calculating $\sqrt{\varepsilon_c}$ from Eq. (4).

The behavior of the reflection coefficient near grazing incidence can be ascertained by inspection of Eqs. (1) and (2) or (6) and (7). At grazing incidence, $\psi_2 = 0$ and

$$\Gamma_v = \Gamma_h = -1,$$

independent of the properties of the surface, that is, of ε_c, as long as ε_c remains finite. It is instructive, however, to investigate the behavior of Γ when ε_c is allowed to approach infinity. It is clear from Eq. (7) that as ε_c increases without limit, Γ_h approaches -1 for all values of ψ_2. For vertical polarization the situation is not so simple, however. If we assume that ε_c is real (the earth is a perfect dielectric, thus ε_2 is zero), it will be seen from Eq. (6) that $\rho_v = 0$ when $\sin \psi_2 = 1/\sqrt{\varepsilon_1}$. (This is just the complement of Brewster's polarizing angle.) When ψ_2 is less than this critical value, $\phi_v = \pi$, but at the critical angle it decreases abruptly to zero, where it remains for all larger values of ψ_2. When ε_c is complex, however, ρ_v can decrease only to a *minimum* value greater than zero, reaching the minimum at an angle smaller than $\sin^{-1}(1/\sqrt{\varepsilon_1})$. At the same time ϕ_v is no longer discontinuous at this angle, but it decreases rapidly from somewhat less than π to a small value and decreases slowly thereafter. As ε_c becomes very large, the critical angle approaches zero, the minimum value of ρ_v approaches unity, and $\Gamma_v \to +1$ for all values of ψ_2. This is the basis of the familiar statement that for a perfect conductor $\Gamma_v = +1$. In the presence of finite conductivity this condition can be approached only at grazing angles well above the critical angle, and it will be seen that in the wavelength region considered here it cannot be considered as a practical possibility. Usually only small values of ψ_2 are of interest for very short waves, for in general beamwidths are so narrow that the surface is not strongly illuminated at large angles of depression; but more important, the effects of surface roughness, discussed in later sections, are frequently severe enough to reduce markedly the effectiveness of the specular reflection at high angles discussed so far.

The reflection coefficient at normal incidence can be obtained directly from either Eqs. (1) and (2) or (6) and (7). It is

$$\Gamma_v = -\Gamma_h = \frac{\sqrt{\varepsilon_c} - 1}{\sqrt{\varepsilon_c} + 1}$$

It should be noted that at normal incidence a "vertically polarized" field becomes horizontally polarized; and, because the ratio of the horizontally

polarized electric field components is $-\Gamma_v$, the difference between the two polarizations disappears, as it obviously must, and the ratio of reflected to incident field for both cases is

$$\Gamma = \frac{1 - \sqrt{\varepsilon_c}}{1 + \sqrt{\varepsilon_c}}.$$

The variation of Γ with ψ_2 can be visualized easily with the aid of Figs. 5·1 and 5·2, which portray the numerator and denominator of

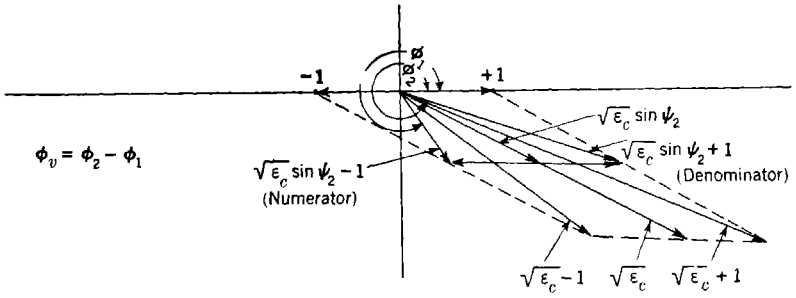


FIG. 5·1.—Behavior of the components of Γ_v in the complex plane.

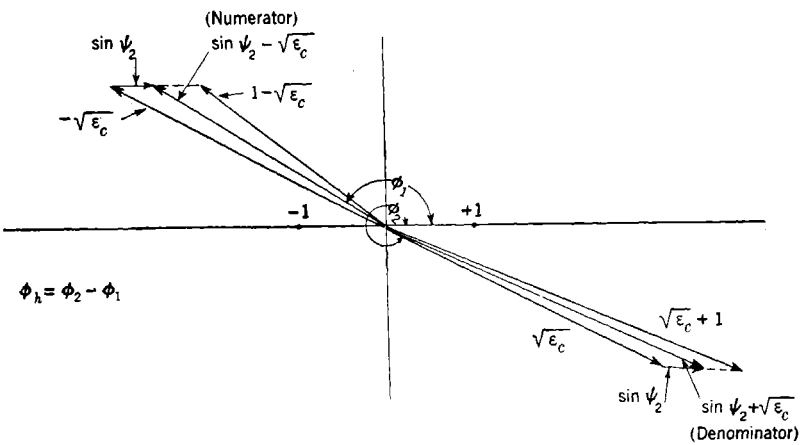


FIG. 5·2.—Behavior of the components of Γ_h in the complex plane.

Eqs. (6) and (7) in the complex plane.[1] These quantities, constructed from the appropriate combinations of ε_c and $\sin\psi_2$, vary with ψ_2 in such a way that the tips of their vectors follow the dotted lines. It is obvious that ρ_v is unity

[1] The value of ε_c used here was chosen to make the diagrams readable rather than to conform to the approximations leading to Eqs. (6) and (7). This does not limit the usefulness of the information obtainable from the figures.

for $\psi_2 = 0$ and that it passes through a minimum and rises to the value given by Eq. (8) at normal incidence. The phase lag ϕ_v, because of the sign in the definition in Eq. (1), is given by the argument of the denominator minus that of the numerator. It can be seen to start at $+\pi$ for $\psi_2 = 0$, to vary most rapidly near the minimum of ρ_v, and to fall to the value given by Eq. (8) for normal incidence. The simpler diagram of Fig. 5·2 shows the corresponding information for horizontal polarization. The qualitative trend of these quantities is shown in Fig. 5·3.

The details of the variation of ρ and ϕ for ψ_2 up to 10° and for wavelength from 3 m to 3 cm are shown in Figs. 5·4 to 5·6. The values of ε_1 and σ used for these figures were taken from Table 5·1, except for 3 cm, for which the information in the table was not available. (The phase for horizontal polarization is not given, as its total increase over 180° is less than 4°.) In these curves the predominant features are the lack of sensitivity of Γ_h to the grazing angle in contrast to the marked sensitivity of Γ_v.

In Fig. 5·4 the arrows indicate the critical values of grazing angle $\sin^{-1}(1/\sqrt{\varepsilon_1})$ to be expected in the absence of conductivity. In Fig. 5·5 the dotted lines indicate the behavior of ϕ_v under the same assumptions. The contrast between the dotted and solid curves and the displacement upward and to the left of the minimum in ρ_v show the trend in the behavior of the water, which is predominantly a dielectric at 3 cm and an indifferent conductor at 3 m.

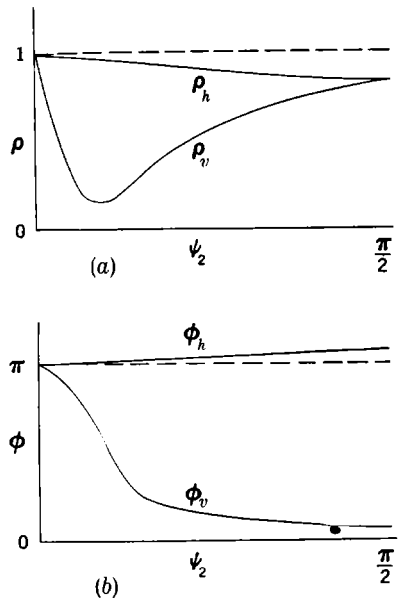


FIG. 5·3.—Variation of ρ and ϕ with grazing angle ψ_2.

The discussion thus far has been confined to the properties of Γ as defined by Eqs. (1) and (2), which apply to a plane wave reflected from a plane surface, whereas the physical situation to which the equations are to be applied involves a spherical wave reflected from a spherical surface. It turns out that the spherical shape of the reflecting surface can be taken into account by the divergence factor, given by Eq (2·365) and discussed further in the next section. But the spherical nature of the wavefront begins to dominate the reflection phenomenon as the point of reflection approaches the horizon, that is, as the grazing angle ψ_2 approaches zero.

This matter has been discussed by van der Pol and Bremmer,[1] who derive a value for the true spherical reflection coefficient for the limiting case of grazing incidence.[2] It will be recalled from Sec. 2·11, however, that the

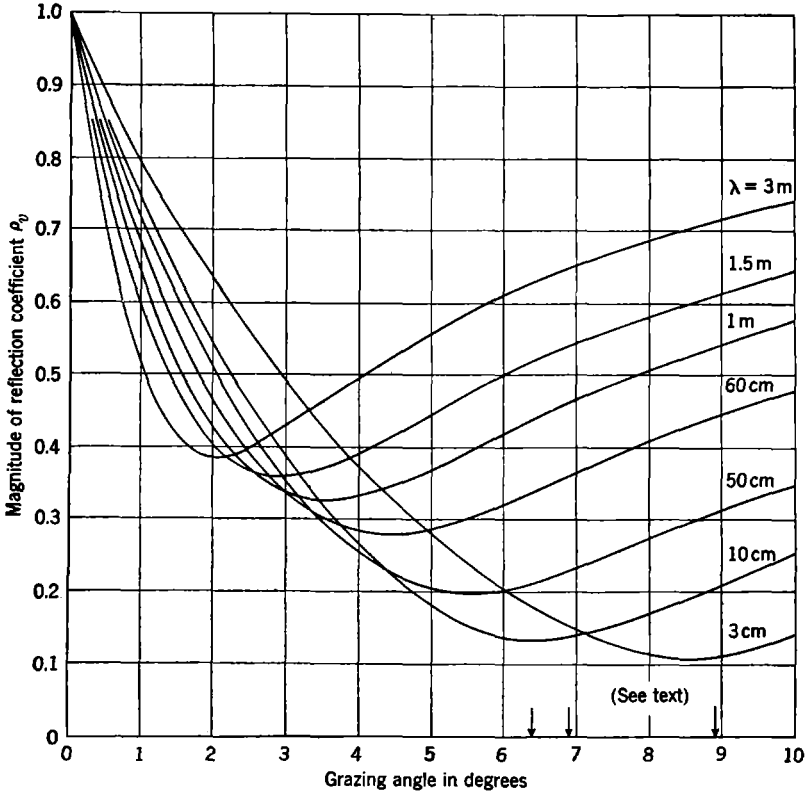


FIG. 5·4.—Magnitude of the reflection coefficient ρ_v as a function of grazing angle for a smooth sea and vertical polarization. Parameters on curve refer to wavelength. See text for meaning of arrows.

field integral of Eq. (2·346) has been evaluated by an asymptotic expansion that excludes values of ψ_2 near zero, and the simple formula

$$F = |1 + \rho De^{-i\alpha}|$$

is no longer valid very near the tangent ray. Without a complete exploration of the intermediate region there is no apparent justification for attempting refinement of the expression for Γ alone.

[1] *Phil. Mag.*, **27**, No. 182, 261, (1939).

[2] Their formula may also be derived from Eq. (2·355) by setting $A = 0$ and using the properties of the y functions given in Eqs. (2·313 and 2·314).

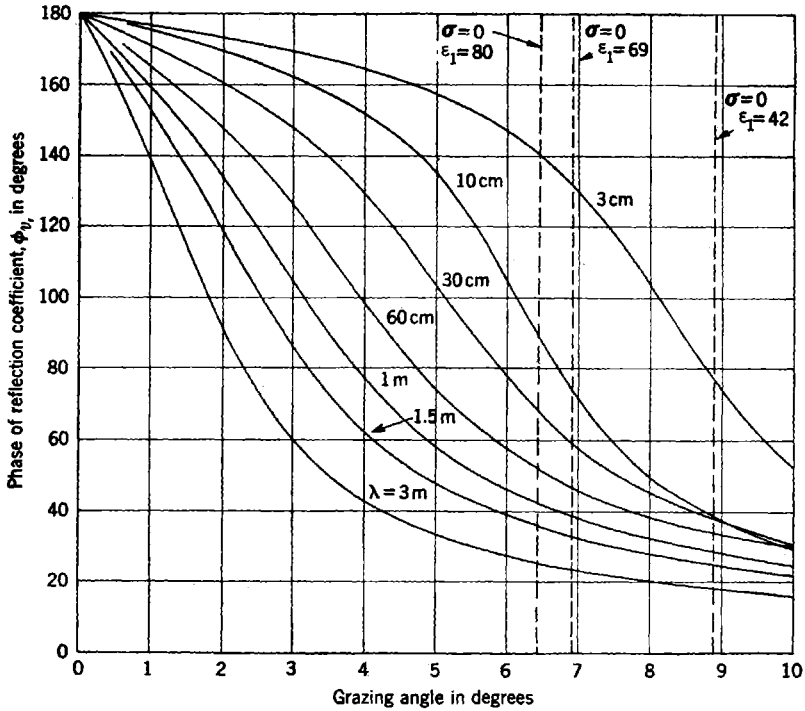


FIG. 5·5.—Phase of the reflection coefficient ϕ_v as a function of grazing angle for a smooth sea and vertical polarization.

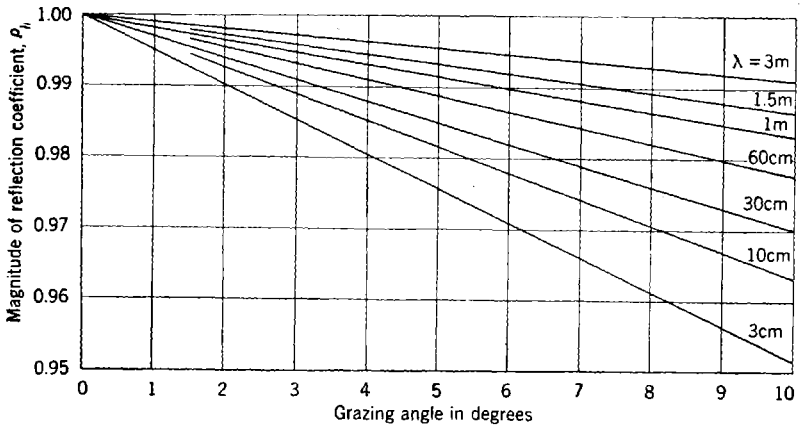


FIG. 5·6.—Magnitude of the reflection coefficient ρ_h as a function of grazing angle for a smooth sea and horizontal polarization.

5·2. Geometrical Interpretation of the Divergence Factor.—The divergence factor D has been mentioned earlier as the quantity that in the interference region represents the spherical shape of the earth in weakening the energy reflected from the surface. This meaning of D is not evident from the derivation of the field-strength formulas in Chap. 2, however. It is profitable to derive the formula for D from purely geometrical considerations and to show what degree of approximation is required to reduce it to the commonly used form given in Chap. 2. The demonstration given here was suggested by that of van der Pol and Bremmer, who derived the same result as part of an elaborate analysis of diffraction of waves by a conducting sphere.[1]

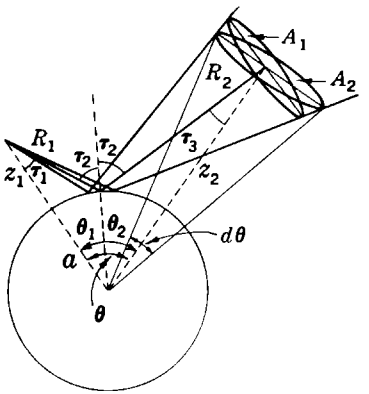


FIG. 5·7.—Geometry for the divergence factor D. (*After van der Pol.*)

Figure 5·7 shows the cross section of the earth, with an isotropic source at a height z_1. It is desired to compare the density of the rays in a small cone reflected from the sphere near the principal point of reflection with the density that the rays of the same cone would have if they were reflected from a plane reflector at the same point. The field strength is proportional to the square root of the ray density. More specifically, we shall compare the cross section of the cone after reflection from a sphere with that after reflection by a plane. D is the square root of the ratio of the latter to the former cross section.

The cross section of the bundle of rays leaving the source is $R^2 \sin \tau_1 \, d\tau_1 \, d\phi$, where ϕ is measured perpendicular to the plane of the paper and R is slant range. After reflection from a plane its cross section would be $(R_1 + R_2)^2 \sin \tau_1 \, d\tau_1 \, d\phi$, as the rays appear to have traveled a distance $R_1 + R_2$ from the image of the source below the plane tangent at the point of reflection. After reflection from the sphere the cross section is A_1 of Fig. 5·7. It is also equal to $A_2 \cos \tau_3$, where $A_2 = (a + z_2)^2 \sin \theta \, d\theta \, d\phi$. The divergence factor is then given by

$$D = \sqrt{\frac{(R_1 + R_2)^2 \sin \tau_1 \, d\tau_1 \, d\phi}{(a + z_2)^2 \sin \theta \cos \tau_3 \, d\theta \, d\phi}} = \frac{R_1 + R_2}{a + z_2} \sqrt{\frac{\sin \tau_1 \, d\tau_1}{\sin \theta \cos \tau_3 \, d\theta}} \tag{8}$$

In order to obtain $d\theta$, we require several relations that can be obtained from inspection of Fig. 5·7:

[1] Van der Pol and Bremmer, "Further Note on the Propagation of Radio Waves over a Finitely Conducting Spherical Earth," *Phil. Mag.*, **27**, Ser. 7, No. 182, March 1939.

$$a \sin \tau_2 = (a + z_1) \sin \tau_1 = (a + z_2) \sin \tau_3; \tag{9}$$

$$R_1 \sin \tau_1 = a \sin \theta_1, \qquad R_2 \sin \tau_3 = a \sin \theta_2; \tag{10}$$

$$R_1 = (a + z_1) \cos \tau_1 - a \cos \tau_2; \tag{11}$$

$$R_2 = (a + z_2) \cos \tau_3 - a \cos \tau_2; \tag{12}$$

$$R_1^2 = a^2 + (a + z_1)^2 - 2a(a + z_1) \cos \theta_1; \tag{13}$$

$$R_2^2 = a^2 + (a + z_2)^2 - 2a(a + z_2) \cos \theta_2. \tag{14}$$

By manipulation of Eqs. (9) to (12) it is easy to show that θ can be expressed in terms of a, z_1, z_2, and τ_1. As τ_1 is the only variable quantity, we can then write

$$d\theta = \frac{\partial \theta}{\partial \tau_1} d\tau_1.$$

Then

$$D = \frac{R_1 + R_2}{a + z_2} \sqrt{\frac{\sin \tau_1}{\sin \theta \cos \tau_3 \dfrac{\partial \theta}{\partial \tau_1}}}, \tag{15}$$

where $\partial\theta/\partial\tau_1$ must now be evaluated. Differentiation of Eqs. (13) and (14) with respect to τ_1 and use of Eq. (10) yields

$$\frac{\partial (R_1)^2}{\partial \tau_1} = 2R_1 \frac{\partial R_1}{\partial \tau_1} = 2a(a + z_1) \sin \theta_1 \frac{\partial \theta_1}{\partial \tau_1} = 2R_1 (a + z_1) \sin \tau_1 \frac{\partial \theta_1}{\partial \tau_1},$$

$$\frac{\partial (R_2)^2}{\partial \tau_1} = 2R_2 \frac{\partial R_2}{\partial \tau_1} = 2a(a + z_2) \sin \theta_2 \frac{\partial \theta_2}{\partial \tau_1} = 2R_2 (a + z_2) \sin \tau_3 \frac{\partial \theta_2}{\partial \tau_1},$$

from which

$$\frac{\partial \theta_1}{\partial \tau_1} + \frac{\partial \theta_2}{\partial \tau_1} = \frac{\partial \theta}{\partial \tau_1} = \frac{1}{(a + z_1) \sin \tau_1} \frac{\partial R_1}{\partial \tau_1} + \frac{1}{(a + z_2) \sin \tau_3} \frac{\partial R_2}{\partial \tau_1}.$$

From Eqs. (9) to (12) one can obtain

$$\frac{\partial R_1}{\partial \tau_1} = R_1 \tan \tau_2,$$

$$\frac{\partial R_2}{\partial \tau_1} = R_2 \cot \tau_1 \tan \tau_2 \tan \tau_3.$$

Combination of the last three equations yields

$$\frac{\partial \theta}{\partial \tau_1} = \frac{(a + z_2) R_1 \cos \tau_3 + (a + z_1) R_2 \cos \tau_1}{a(a + z_2) \cos \tau_2 \cos \tau_3},$$

which with the aid of Eqs. (15) and (9) gives van der Pol and Bremmer's expression for D:

$$D = \frac{a(R_1 + R_2)\sqrt{\sin \tau_2 \cos \tau_2}}{\sqrt{[(a + z_2)R_1 \cos \tau_3 + (a + z_1)R_2 \cos \tau_1]\,(a + z_1)(a + z_2) \sin \theta}}. \quad (16)$$

Equation (16) is unwieldy and unnecessarily refined for most practical computation purposes. It may be simplified to have the form of Eq. (2·365), which is commonly used. The angle τ_2 is the complement of the grazing angle ψ_2, which is always very small in the region in which D has an appreciable effect. Hence we write $\sin \tau_2 = \cos \psi_2 = 1$ and $\cos \tau_2 = \sin \psi_2$. If we also substitute $(a + z_1) \cos \tau_1$ and $(a + z_2) \cos \tau_3$ from Eqs. (11) and (12) and rearrange terms, Eq. (16) can be written

$$\frac{1}{D^2} \approx \frac{\left(1 + \frac{z_1}{a}\right)\left(1 + \frac{z_2}{a}\right) a \sin \theta}{R_1 + R_2}\left[1 + \frac{2R_1R_2}{a(R_1 + R_2) \sin \psi_2}\right]. \quad (17)$$

We restrict terminal heights to values very small by comparison with the earth's radius, thus permitting z_1/a and z_2/a to be discarded. Also ranges are so small that $a \sin \theta \approx a\theta \approx R_1 + R_2$. The coefficient of the quantity in brackets is then approximately unity. To a very satisfactory approximation R_1 and R_2, the slant ranges from the terminals to the point of reflection, may be replaced by the corresponding ground ranges r_1 and r_2. The final expression for D then becomes

$$D \approx \frac{1}{\sqrt{1 + \frac{2r_1r_2}{a(r_1 + r_2) \sin \psi_2}}}. \quad (18)$$

The geometrical interpretation of D given previously is thus justified, and the degree of approximation involved in the final expression is made clear without regard to the earth-flattening procedure of Chap. 2. The approximations of small angles, low terminal heights, and short ranges (in terms of earth radius) are consistent with those already imposed by the simplified expressions for path-length difference ΔR given in Chap. 2.[1]

The value of earth's radius given here is, of course, meant to be the effective earth's radius a_e, obtained as described in Chap. 2.

Further insight into the physical importance of D can be obtained from the examples of the next section.

5·3. Effects of Reflections on Field Strength.—In Secs. 2·13 and 2·14 methods were given to permit calculation of field strength in the interference and diffraction regions. The purpose of this section is to present a few further details of the way in which the earth's surface affects the cov-

[1] Since the foregoing material was written a much more elaborate and complete divergence formula for surfaces of arbitrary (smooth) shapes has appeared. See H. J. Riblet and C. B. Barker, *J. Applied Phys.*, **19**, 63 (1948).

erage diagram, on the assumption that the surface is smooth and homogeneous.

As mentioned earlier, there are three common methods of presenting field-strength information. These are (1) plots of signal strength vs. range for fixed terminal heights; (2) plots of signal strength vs. height of one terminal, for fixed range and height of second terminal, (3) a map of

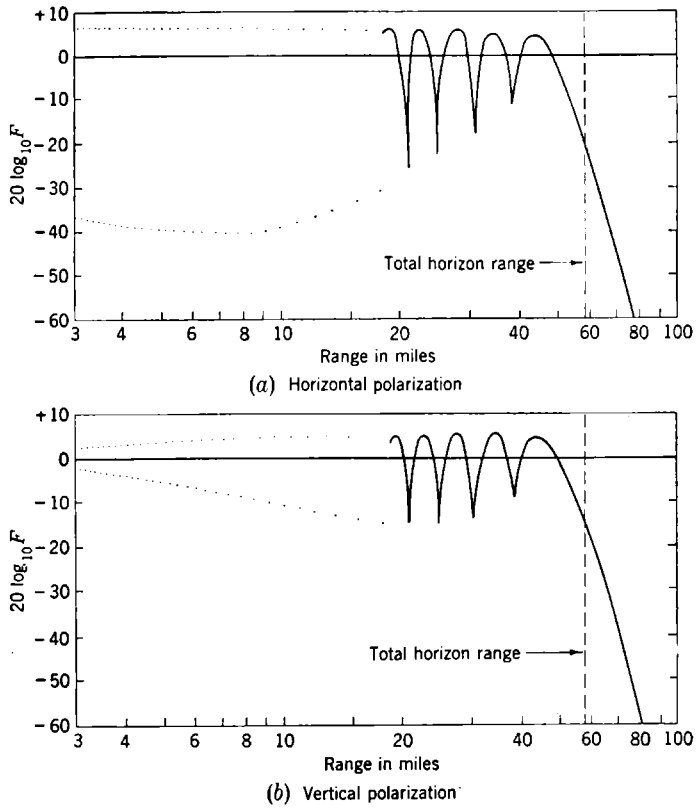


FIG. 5·8.—Variation of field strength relative to free-space field (20 $\log_{10}F$) from an isotropic antenna for a wavelength of 10 cm and terminal heights of 90 and 1000 ft over a smooth sea.

the field, consisting of contours of constant field strength plotted on a graph whose coordinates are range and height (a contour diagram). Each of these methods has advantages in presenting information in certain regions, and examples of each will be given.

Examples of the first type of diagram are shown in Fig. 5·8, which shows the field strength in decibels relative to the free-space field (20 $\log_{10}F$) as a function of range for 10-cm transmission over sea water, with terminal

heights of 90 and 1000 ft and an isotropic radiation pattern. These curves reveal an interesting property of the product $D\rho$. The quantity D is unity high in the interference region and falls to zero rather abruptly along the tangent ray, whereas ρ is unity at grazing incidence and decreases with increasing grazing angle, as shown in Figs. 5·4 and 5·6. As a consequence

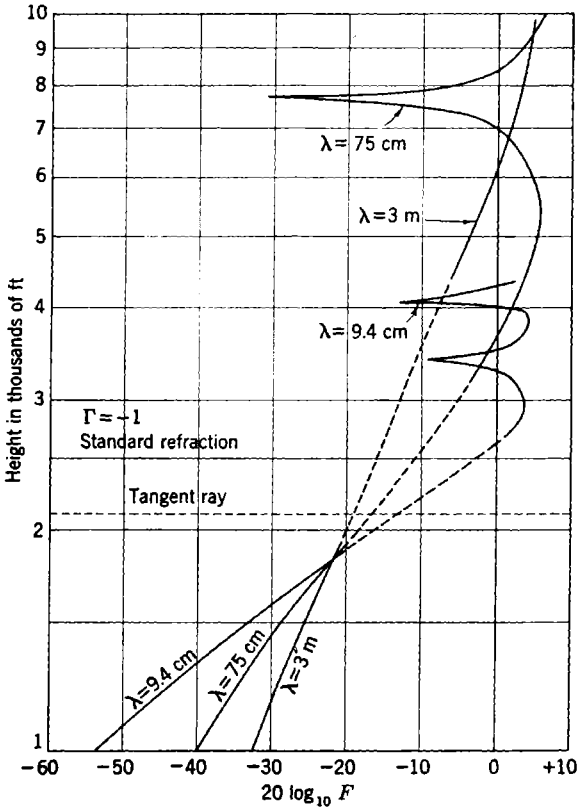


FIG. 5·9.—Field strength relative to the free-space field (20 $\log_{10}F$) as a function of height at a range of 80 miles from a transmitter at a height of 110 ft. The broken curve indicates the intermediate region (Sec. 2·15).

$D\rho$ reaches a maximum at a small grazing angle; for horizontal polarization the maximum value of $D\rho$ is larger and occurs at a larger angle than for vertical polarization. This fact is responsible for the deeper minima and stronger maxima in the interference pattern for horizontal polarization (revealed by the shape of the envelope of maxima and minima) and for the fact that the minimum of the envelope occurs at a shorter range. This phenomenon will be noticed only for large terminal heights, because for

small heights the maximum in $D\rho$ occurs at ranges outside the first few minima.[1]

The depths of the minima and heights of the maxima give an indication of the magnitude of $D\rho$ and, as will be shown in Sec. 5·6, can be used in measurements of ρ.

It is worth mentioning again that with standard refraction the field strength along the tangent ray (at the total horizon range) is *always* considerably less than the free-space field by an amount that depends on wavelength, heights, and range; moreover, the rate of decrease of field strength with range at and beyond the horizon is large.[2] Thus, the horizon constitutes an upper range limit for radar detection under standard refractive conditions, although point-to-point *communication* beyond the horizon is far less difficult because the transmission loss does not occur twice as it does in radar.

The behavior of the field in the intermediate region and near the lowest lobe of the interference pattern portrays the superiority of very short waves in concentrating a large field near the surface, that is, of providing good low coverage. Such information can be presented on plots of either type 1 or 2 described above. Fig. 5·9 shows an example[3] of type 2 for wavelengths of 3 m and 75 and 9·4 cm. This trend for the interference region is obvious from the discussion of F given in Sec. 2·2 for a plane earth, but it is illustrated in more detail here. This and the previous figure point clearly to the desirability of using short wavelengths in applications requiring detection as near the surface as possible.

The continuation of the trend of wavelength dependence into the interference region is shown in Fig. 5·10, which shows only the two lowest lobes of the interference pattern for the same wavelengths and transmitter height used in the preceding figure. The marked increase in the fine structure of the lobes results from the increase in effective transmitter height, measured by $z_1/\lambda^{2/3}$, as the wavelength is decreased.[4] If, on the other hand, this ratio had been held constant by changing the transmitter height as the

[1] Surface roughness, discussed in later sections, is an even more important factor in destroying the shape of the bottom of the envelope in Fig. 5·8.

[2] The presence of surface ducts reduces this trend, as shown in Chaps. 1, 2, and 4.

[3] As these curves are plotted for $\Gamma = -1$, they represent the behavior of a horizontally polarized field. If the earth constants are taken into account, the field in the diffraction region is slightly greater for vertical than for horizontal polarization on all wavelengths and is slightly less in the interference region at the maxima of the interference lobes. The difference in the diffraction region decreases rapidly with decreasing wavelength because of the decrease in ϵ_c (see Sec. 2·11).

[4] It will be recalled from Chap. 2 that once a choice is made of the contour to be plotted (R_f or $|E/E_s|$) and the height of the transmitter in natural units is specified, the coverage diagram is uniquely determined (for $\Gamma = -1$ and standard refraction). The transmitter height z_1 is related to height in natural units by

$$Z_1 = \frac{z_1}{H} = 2z_1 \left(\frac{\pi^2}{a_e \lambda_0^2}\right)^{1/3}.$$

wavelength was changed, the lobes for all three cases would have coincided. [An increase in the height of the 3-m transmitter to $110(300/9.4)^{\frac{2}{3}} = 1100$ ft would be necessary to place its lobes in the position of the 9.4-cm lobes from the transmitter at 110 ft.] It is interesting to note that whereas the lobe positions are governed by the ratio $z_1/\lambda^{\frac{2}{3}}$ for the spherical earth, they

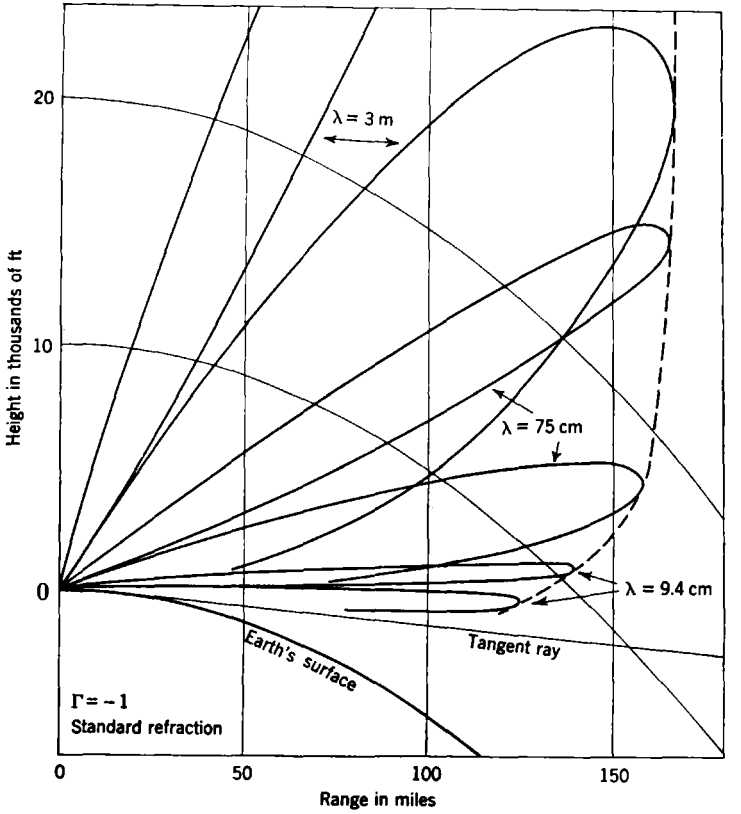


FIG. 5·10.—Contours of field strength 103 db below that at 1 m from an isotropic antenna 110 ft above the surface (R_f = 88 miles). Only the two lowest lobes are shown for each wavelength.

are governed by z_1/λ for the plane earth. As shown in Sec. 2·2, for a plane earth the angular elevations of the successive lobe maxima are proportional to λ/z_1, but no such simple relation exists for the spherical earth.

The dotted line in Fig. 5·10 is the envelope of maxima, which in the present case is just a contour of $1 + D$, as we have assumed $\Gamma = -1$. At high angles, $D \approx 1$ and $F_{max} \approx 2$, as can be seen from the doubling of R_f at high angles.

The assumption that $\Gamma = -1$ is, of course, an oversimplification. For wavelengths of the order of 1 m over the ocean it represents a good approxi-

mation at small grazing angles even for vertical polarization; but as shown in the following sections, it may become very poor for ordinary terrain and wavelengths of a few centimeters. The idealized cases are worthy of detailed study, however, because of the insight that they give into the relations between the various parameters and because they are approached reasonably well in many practical cases.

5·4. Surface Roughness.—The discussion thus far has assumed that the earth's surface is smooth. We now consider qualitatively the way in which surface roughness modifies the results for a smooth surface. Perhaps the simplest experiment that one can perform to obtain an idea of the effects of roughness is to observe the light reflected from a sheet of paper having a rough surface. At large grazing angles the surface appears dull, but at sufficiently small angles it appears shiny and acts sufficiently like a mirror to produce a crude image of the light source.

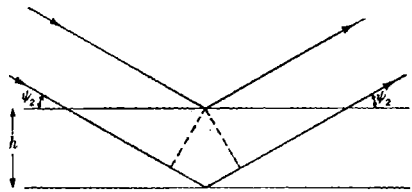


FIG. 5·11.—Phase difference between rays reflected from two levels.

It seems reasonable from intuition that the efficiency of specular reflection from a rough surface should involve the ratio of the dimensions of the surface roughness elements to the wavelength as well as the grazing angle. Lord Rayleigh suggested a way of formulating the relation between these quantities. In Fig. 5·11 let h represent the height of an irregularity of the surface and ψ_2 the grazing angle. The path difference between the two rays shown is just $\Delta R = 2h \sin \psi_2$, and the phase difference is

$$k\,\Delta R = \frac{4\pi h}{\lambda} \sin \psi_2.$$

If this phase difference is very small, the effect of the roughness is small and the surface is effectively smooth. Increasing h/λ or ψ_2 increases the path difference and weakens the resultant reflected field; when $k\Delta R = \pi$, the direct and reflected rays are in phase opposition. Between these two extremes it is desirable to establish a value of phase difference to divide "rough" from "smooth" surfaces. We shall *arbitrarily* choose $\pi/2$ as the limiting value,[1] which gives as a criterion

$$h \sin \psi_2 < \frac{\lambda}{8} \tag{19}$$

for the surface to be considered smooth.

[1] There appears to be no well-established practice in this matter, although Schuster and Nicholson (*Theory of Optics*, 1928, p. 166) make the definite suggestion given above. In view of the crudity of this method of approach there seems to be little point in being more precise.

It should be clear from the nature of this argument that it cannot be considered a quantitative statement of the effects of roughness. The use of the simple ray picture automatically excludes from consideration the local diffraction effects that occur when the roughness elements are comparable in size to the wavelength. If this simple concept is applied to the ocean surface, additional questions arise because of the irregular distribution of size and shape of the waves, because of foam and bubbles, etc. A rigorous solution of the problem would require techniques as yet undeveloped for satisfying electromagnetic boundary conditions over a surface with a statistical distribution of irregularities. Calculations have been made of effective reflection coefficients using simple models of the sea surface, such as sinusoidal[1] or wedge-shaped[2] waves. Because the surface in reality cannot be represented well by any simple geometrical configuration, it seems preferable to limit the present discussion to consideration of the effects of roughness in terms of simple Fresnel diffraction theory.[3]

In the derivation in Chap. 2 of the expressions for field strength in the interference region, it was shown that the field could be described in terms of interference between two waves following the paths of a "direct" and "reflected" ray. The location of the point of reflection was determined from simple image theory.[4] The ray reflected from a single point is in reality a means of expressing the fact that the energy source illuminates a large region of the surface, inducing currents that radiate in all directions (Huygens' principle); at any point in space the "reflected" field is the sum of the radiation from these induced currents.

As a first step in analyzing the reflected field from this point of view, we shall locate the areas on the earth's surface from which the secondary radiation has a specified phase relation to radiation proceeding directly from the primary source to the field point.

Figure 5·12 shows the source and the field point over a reflecting plane surface.

The dotted line indicates the path of the reflected ray and the point of reflection. The ground range between source and field point is r, and the slant range is R_1. Radiation traveling between the two points by any other path than R_1 will travel a greater distance, which will be denoted by R_2. The difference between any other arbitrary path length and R_1 will be denoted by δ,

$$R_2 - R_1 = \delta.$$

[1] F. Hoyle and M. H. L. Pryce, "The Limiting Ranges of R.D.F. Sets over the Sea," ASE Report No. M395; G. G. MacFarlane, "Sea Returns and the Detection of Schnorkel," TRE Report No. T1787, Feb. 13, 1945.

[2] S. O. Rice, BTL Memorandum No. MM43-210-6, Oct. 13, 1943.

[3] Section 6·7 represents a discussion of the nature of the sea surface.

[4] This discussion will be confined to a plane earth, as all the essential facts can be obtained without recourse to the much more difficult theory for the spherical earth.

The locus of all points in space corresponding to a specified value of δ will be an ellipsoid of revolution with the two terminals $(0,0,z_1)$, $(r,0,z_2)$ as foci and R_1 as the axis of revolution. If now δ is allowed to increase in steps, of, let us say, $\lambda/2$, a family of ellipsoids results. The intersections of these ellipsoids with the xy-plane is a family of ellipses, each ellipse corresponding to a given value of δ. The average phase of the radiation from each zone bounded by adjacent ellipses differs from that of the neighboring zone by just π; that is, the elementary waves from successive zones are in phase

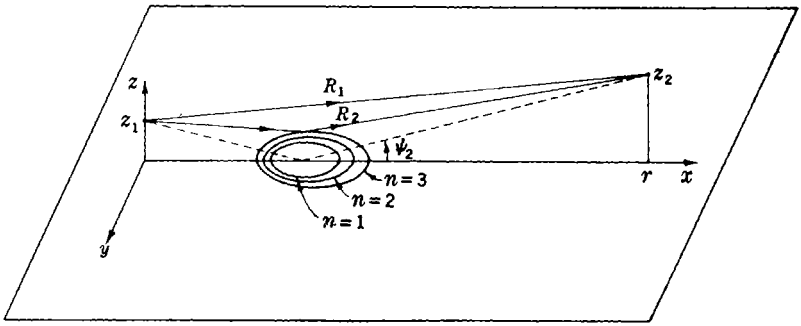


FIG. 5·12.—Half-period zones in the xy-plane.

opposition. [They do not quite cancel completely, as the amplitudes change very slowly from zone to zone, but the first zone ($n = 1$ in the figure) is the principal contributor.] It is of interest to calculate the dimensions of the zones in order to discover how large an area must be sensibly smooth if the smooth-earth phase formulas are to be applicable. The algebraic details of the following development are tedious and uninteresting; consequently the steps will only be outlined briefly.

The equation for the ellipsoids is most conveniently written relative to a new set of axes that have the new x'-axis tilted up at an angle θ with the x-axis and falling along R_1. The y'-axis is parallel to the y-axis, but the origin of the new axes is shifted to the mid-point of R_1. From analytic geometry one finds that

$$x' = \left(x - \frac{r}{2}\right)\cos\theta + \left(z - \frac{z_1 + z_2}{2}\right)\sin\theta,$$

$$y' = y,$$

$$z' = -\left(x - \frac{r}{2}\right)\sin\theta + \left(z - \frac{z_1 + z_2}{2}\right)\cos\theta,$$

and the equation of the ellipsoid corresponding to a given value of δ is

$$b^2x'^2 + a^2(y'^2 + z'^2) = a^2b^2, \tag{20}$$

where

$$a = \frac{r \sec \theta + \delta}{2}, \tag{21}$$

$$b = \frac{1}{2}\sqrt{\delta^2 + 2r\delta \sec \theta}, \tag{22}$$

$$\tan \theta = \frac{z_2 - z_1}{r}. \tag{23}$$

If the expressions for the primed coordinates are substituted into Eq. (20), the resulting equation gives the family of ellipsoids in terms of x, y, z. As only the ellipses in the xy-plane are of interest here, we set $z = 0$ in the transformed equation and obtain, after some manipulation,

$$(x - \tfrac{1}{2}r - p)^2 \ (b^2 \cos^2\theta + a^2 \sin^2\theta) + a^2y^2 = \frac{a^2b^2(b^2 \cos^2\theta + a^2 \sin^2\theta - c^2)}{b^2 \cos^2\theta + a^2 \sin^2\theta}, \tag{24}$$

where

$$p = -\frac{c(a^2 - b^2) \sin\theta \cos\theta}{b^2 \cos^2\theta + a^2 \sin^2\theta},$$

$$c = \frac{z_1 + z_2}{2}.$$

Equation (24) represents an ellipse in the xy-plane, with its center on the x-axis at the position x_0, given by

$$x_0 = \frac{r}{2} + p = \frac{r}{2}\left\{1 - \frac{\left(\frac{z_2 - z_1}{r}\right)^2}{\left[\frac{\delta}{r} + \sqrt{1 + \left(\frac{z_2 - z_1}{r}\right)^2}\right]^2 - 1}\right\}. \tag{25}$$

The x- and y-intercepts of the ellipse may be found from Eq. (24). The y-intercepts (and length of semiminor axis) are found by setting $x = x_0$, yielding

$$y_1 = \pm\, b\sqrt{1 - \frac{c^2}{b^2 \cos^2\theta + a^2 \sin^2\theta}} \tag{26a}$$

or

$$y_1 = \pm\, \frac{r}{2}\sqrt{\left(\frac{\delta}{r}\right)^2 + \frac{2\delta}{r}\sqrt{1 + \left(\frac{z_1 - z_2}{r}\right)^2}} \times \sqrt{1 - \frac{\left(\frac{z_1 + z_2}{r}\right)^2}{\left[\frac{\delta}{r} + \sqrt{1 + \left(\frac{z_2 - z_1}{r}\right)^2}\right]^2 - 1}}. \tag{26b}$$

Similarly, the x-intercepts are found by setting $y = 0$ and then using Eq. (26).

$$x_1 = x_0 \pm y_1 \frac{a}{\sqrt{b^2 \cos^2 \theta + a^2 \sin^2 \theta}}, \tag{27a}$$

or

$$x_1 = x_0 \pm y_1 \sqrt{1 + \frac{1}{\left[\frac{\delta}{r} + \sqrt{1 + \left(\frac{z_2 - z_1}{r}\right)^2}\right]^2 - 1}}. \tag{27b}$$

The length of the semimajor axis is the second term of Eq. (27). It will be denoted by ξ_n. Then $x_1 = x_0 + \xi_1$.

These equations are exact but are so complicated that they do not lend themselves to simple analysis. They can be greatly simplified by introducing the approximations used in most practical cases, namely, that z_1, z_2, and δ are all much smaller than r. This assumption leads to

$$\left[\frac{\delta}{r} + \sqrt{1 + \left(\frac{z_2 \pm z_1}{r}\right)^2}\right]^2 - 1 \approx \frac{2\delta}{r} + \left(\frac{z_2 \pm z_1}{r}\right)^2 \ll 1. \tag{28}$$

We also make an additional substitution

$$\delta = \delta_0 + \frac{n\lambda}{2}, \tag{29}$$

where

$$\delta_0 = \frac{2z_1 z_2}{r} \tag{30}$$

will be recognized as the approximate form for the path difference ΔR between R_1 and the length of the ray path via the point of reflection. This is done because it is the excess of path difference over that through the point of reflection which is of immediate interest rather than the total path length. Incorporating these approximations and the transformation of δ, we have finally

$$x_{0n} \approx \frac{r}{2} \frac{1 + \frac{2z_1(z_1 + z_2)}{n\lambda r}}{1 + \frac{(z_1 + z_2)^2}{n\lambda r}}, \tag{31}$$

$$x_{1n} \approx x_{0n} \pm \frac{r}{2} \frac{\sqrt{1 + \frac{2\delta_0}{n\lambda}}}{1 + \frac{(z_1 + z_2)^2}{n\lambda r}}, \tag{32}$$

$$y_{1n} \approx \pm \frac{\sqrt{n\lambda r}}{2} \sqrt{\frac{1 + \frac{2\delta_0}{n\lambda}}{1 + \frac{(z_1 + z_2)^2}{n\lambda r}}}. \tag{33}$$

The subscript n refers to the particular ellipse being considered. The area of the ellipse for a given n is $A_n = \pi \xi_n y_n$.

$$A_n \approx \frac{\pi r \sqrt{n\lambda r}}{4} \frac{1 + \dfrac{2\delta_0}{n\lambda}}{\left[1 + \dfrac{(z_1 + z_2)^2}{n\lambda r}\right]^{3/2}} \tag{34}$$

Further simplification is possible for a special case of some practical importance. If we impose the additional restriction

$$\frac{4z_1}{z_2} \ll 1, \tag{35}$$

and if we consider only the first ellipse ($n = 1$) at the reflection point for the bottom lobe of the interference pattern ($\delta_0 = \lambda/2$), we have

$$x_{01} \approx \frac{12 z_1^2}{\lambda}, \tag{36}$$

$$x_{11} \approx x_{01} \pm \frac{8\sqrt{2} z_1^2}{\lambda} \approx (1 \pm 0.943) x_{01}, \tag{37}$$

$$y_{11} \approx 2\sqrt{2} z_1 \approx \frac{x_0}{\dfrac{3\sqrt{2} z_1}{\lambda}}. \tag{38}$$

These formulas are useful in cases of transmission between a very low terminal and a high terminal, such as an aircraft, and are useful for radar siting calculations.[1]

A numerical example is helpful to illustrate the sizes of the zones involved. We shall compare two wavelengths in two types of problems. In the first case the path geometry (z_1, z_2, and r) will be fixed, and it is required to determine the position and dimensions of the first zone. In the second case z_1 and r will be fixed, and it will be required to find the position and dimensions of the first zone producing radiation reflected to form the lowest lobe of the interference pattern, for which $\delta_0 = \lambda/2$. The calculations are summarized in the following table. The grazing angle ψ_2 at the

[1] Since the foregoing material was written, a paper entitled "Maximum Range of a Radar Set," by K. A. Norton and A. C. Omberg, has appeared in *Proc. IRE*, **35**, 4 (1947). Formulas corresponding to Eqs. (25) to (27) are derived, but the results are written in such a way that direct comparison is difficult. These formulas reduce to Eqs. (36) to (38) under the assumption stated. This paper gives some discussion of the siting problem. In discussing the Rayleigh criterion for roughness the value of $k \Delta R$ suggested to mark the limit between rough and smooth surfaces is $\pi/4$, thus reducing the allowable height of surface roughness elements to half the value given in Eq. (19). In a recent communication to the author, Norton states that he now believes the use of $\pi/8$ provides a more realistic criterion for deciding when the smooth-earth theory may be used with confidence.

point of reflection and the maximum wave height from Eq. (19) are also given.

TABLE 5·2.—ANALYSIS OF REFLECTION ZONES
(Distances in meters)

λ	Case 1		Case 2	
	0.1	1.0	0.1	1.0
z_1	100	100	100	100
r	10,000	10,000	10,000	10,000
z_2	1000	1000	2.5	25
δ_0	20	20	$0.05 = \frac{\lambda}{2}$	$0.5 = \frac{\lambda}{2}$
x_{01}	910	943	9,350	6,830
y_{11}	9.1	29	6.6	44
ξ_1	83	262	615	2,760
A_1	2,370	23,900	128,000	382,000
ψ_2	6.3°	6.3°	0.59°	0.72°
h_{max}	0.11	1.1	1.22	1.25

The highly elongated shape of the ellipse, particularly for the lowest lobe, is very striking. [It is most exaggerated in cases covered by Eqs. (36) to (38), which apply to Case 2 for $\lambda = 0.1$ m.] The large area of the first zone, particularly for the longer wavelength, illustrates the need for a large, smooth, elongated reflecting area if a well-developed lobe pattern is to be formed. For the first lobe considered in Case 2 a crest-to-trough wave height of about 1 m gives a surface that is effectively smooth for either wavelength, according to Eq. (19), but in Case 1 the same surface is far from smooth for the shorter wavelength.

Thus far only the phase of the reflected radiation has been considered. We have derived formulas for the zones from which the average phase varies by units of π from one zone to the next. In this respect these zones are the Fresnel zones familiar in physical optics.[1] But there is one essential difference. The usual Fresnel zones are constructed in such a way that they are concentric and that their amplitudes of excitation (whether through surface currents or fields in an advancing wave front) are symmetrical about the direction of propagation. The amplitude of excitation varies slowly from one zone to the next, with the result the contributions of successive zones almost completely cancel, leaving the radiation from approximately half the area of the first zone as the end product. In the present instance, however, the zones are inclined at a large angle to the direction of propagation, and their eccentricity is so large, for small grazing angles,

[1] J. C. Slater and N. H. Frank, *Introduction to Theoretical Physics*, McGraw-Hill, 1933; F. A. Jenkins and H. E. White, *Fundamentals of Physical Optics*, McGraw-Hill, 1937.

that the amplitude of excitation varies markedly over the zone. The total contribution from each zone is given by integrating the excitation over the zone. It is not evident, a priori, that the contribution of the first zone alone predominates in the present case. For example, it does not seem entirely reasonable in the numerical examples given above that there is not considerable contribution from the higher order zones from a direction along the minor axes of the ellipses near the reflection point; presumably at the same time the contributions from the ends of the ellipses would decrease in importance. If the antenna radiation pattern is sufficiently narrow, it can also enter into the problem. It would still appear that the principal part of the reflected wave arises from an area near the point of reflection, but whether or not the relative importance of the series of half-period zones illustrated above is the same as for the usual Fresnel zones is not clear. To the author's knowledge this problem has not been carried beyond the qualitative arguments presented here.

The problem of time variations in the effective reflection coefficient of a rough surface such as the ocean is particularly interesting. Explanation in terms of fluctuation of average altitude of the half-period zones, with resulting variable "zone-plate" effects, does not appear promising because of the large size and elongated shape of the zones. It would appear more likely that the fluctuations arise from random superposition of reflections from areas smaller than these zones and depending for their effects upon details of shape and spacing of the roughness elements as well as upon their average height. Intuitively one suspects that there should be a close connection between the reflection coefficient of the ocean and sea echo. No simultaneous measurements of both quantities are available, but the sea-echo measurements described in the next chapter indicate that sea echo tends to be caused (for 10 and 3 cm at least) by very small elements, such as droplets. A parallel conclusion regarding reflection coefficient would not be reasonable, however.

Although the effects of surface roughness have not been formulated in a quantitative manner, this qualitative approach probably indicates the trend of variation of some of the most important factors. The few experiments thus far made that are sufficiently clear-cut to give relevant information are discussed in following sections. The general problem of the effects of surface roughness merits serious study from both theoretical and experimental sides. Lacking much better information than is now available, prediction of microwave coverage rests on insecure ground.

REFLECTION COEFFICIENT OF THE OCEAN

In the preceding sections the theory of specular reflection and its effect upon field strength has been discussed, on the assumption that the ground constants of a perfectly smooth spherical earth are known, and an indication has been given of possible effects of surface roughness. In the labora-

tory the conductivity and dielectric constant for any type of surface at any given wavelength can be measured with some accuracy but for lack of adequate theory we must resort to experiment to investigate how surface roughness modifies the Fresnel laws. The reflection coefficient for both ocean and land have been measured by several different techniques, and about as many different sets of values as measurement techniques have been obtained. When the circumstances under which the measurements were made are properly weighed, the variation in values of ρ can be reconciled and some understanding obtained of the reflection phenomenon from a rough surface. The observations made to date have been limited in scope, but they are of considerable value in estimating field strengths and coverage patterns.

5·5. Measurements of Short-time Variations.[1]—There are often large and rapid fluctuations in the energy reflected from the earth's surface.

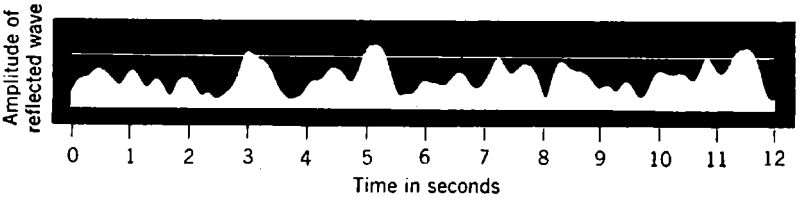


FIG. 5·13.—Variation with time of 10-cm horizontally polarized radiation reflected from a rough sea at a grazing angle of 11°. The white line indicates the amplitude of the direct wave.

This is particularly true of a rough sea, where the continually changing position and size of the surface irregularities cause variations of the reflected field. Two experiments have been performed that demonstrate this phenomenon admirably. In the first,[2] one-way transmission measurements were made on a wavelength of 9 cm between two cliffs separated by an inlet easily invaded by waves from the open sea. The path length was 231 yd. and the grazing angle was about 11°. The antenna patterns were sufficiently narrow to allow the transmitting and receiving antennas to be directed to measure either the direct or reflected waves separately. The transmitter was pulsed at a recurrence frequency of 1000 cps with a pulse length of 500 μsec. The receiver output, roughly proportional to voltage, was displayed on an oscillograph without a sweep and recorded on a continuously moving film. Except for slight line ripple the direct wave was very steady. Figure 5·13 shows the reflected wave on horizontal polarization. During the interval shown the amplitude of the reflected wave varied by more than a factor of 3. It should be repeated that there is no question

[1] By Donald E. Kerr and Herbert Goldstein.

[2] H. Archer-Thomson, N. Brooke, T. Gold, and F. Hoyle, "Preliminary Report on the Reflection of 9-cm Radiation at the Surface of the Sea," ASE Report No. M542, September 1943.

here of motion of an interference pattern; the reflected wave was examined separately. It is significant that for periods of about 0.5 sec occasionally the reflected wave was stronger than the direct wave.

The average values of ρ over periods of several seconds were approximately 0.5 for horizontal and 0.15 for vertical polarization. The estimated height of the water ripples was 20 cm, and the waves were traveling nearly at right angles to the transmission path.

In the second experiment, also in England,[1] airplane targets were simultaneously viewed by a radar set having two antennas and receiving channels; one antenna had a narrow beam in the vertical plane and was used both to transmit and receive in the conventional fashion; the other antenna had a broad vertical beam and was used for receiving only. The sharp beam was sufficiently narrow that the sea surface was not illuminated appreciably at angles greater than 2° from the antenna axis. The wide beam received both direct and reflected waves. The output voltages of both receivers were photographed simultaneously, and the gains were so adjusted that respective deflections corresponded to the same r-f input intensity. The ratio of the two output voltages then eliminated effects due to the variation in target aspect. A complete report of the results has not yet been made, but with grazing angles less than 5° very deep fluctuations of the reflected wave were observed with periods in the neighborhood of 0.02 sec. Above 5° the reflected wave was described as "sporadic in character." The experiment was performed off the coast of Scotland at a season when the sea was rough and hence probably represents fairly extreme conditions.

If the surface is rough, in the sense of the Rayleigh criterion, one would expect large space-diversity effects. (Such effects have in fact been observed for points separated by only a few feet.) Hence with any sort of extended target the amplitude and phase of the incident field on the target should be a rapidly varying function of time and position. It is natural that under such conditions the radar echo would show violent fluctuations. Even when the sea was not quite so rough as required by the Rayleigh criterion, a 10-cm radar echo from a lighthouse at a range of 8000 yd in Boston Harbor has shown time fluctuations of 5 to 10 db.

The effect of the varying reflection from the sea surface is most pronounced when a point target is in the minimum of an interference pattern that shifts in position with the tidal cycle, for then a relatively small change in the reflected field will cause a large percentile change in the total signal. The scatter of points at the minimum in the plot shown in Fig. 6·24 for the reflection from a corner reflector can be ascribed to this cause. It has been reported[2] that at the minimum of the tidal cycle, the echo from a corner

[1] "Reviews of Progress of U.S.W. Propagation," Ultra Short Wave Propagation Panel, No. 11, A.C. 7027, Oct. 14, 1944.

[2] Interservice KXS Radar Trials, "Overwater Performance against Surface Targets," ASE Report No. M688, February 1945.

reflector observed on 1.25 cm wavelength at a grazing angle of 0.3° suffered violent "beating" of 20 db even for a moderate swell. Figure 5·14 presents a series of photographs of an A-scope showing the echo from a lighthouse in Boston Harbor at the "tidal" minimum. The signal on the left is an artificial echo from a signal generator, and some outlying rocks show as echoes immediately surrounding the lighthouse signal. The photographs are separated in time by 0.25 sec. During a period of less than 2 sec the echo strength changed by more than 10 db.

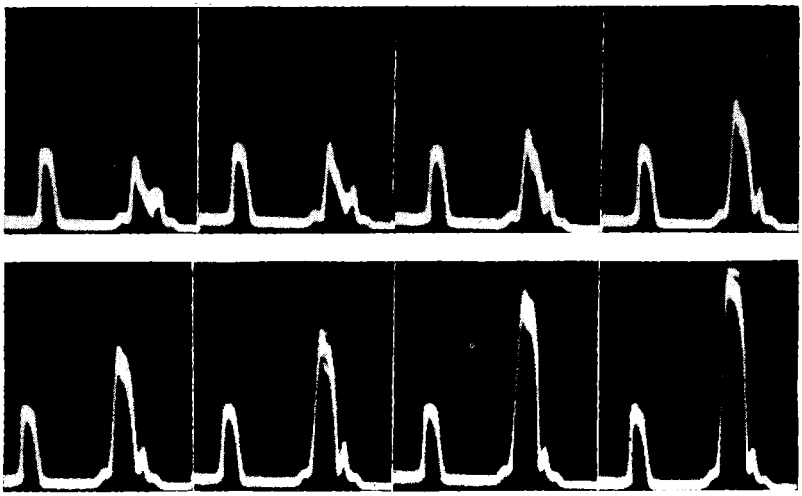


FIG. 5·14.—Photographs of A-scope, showing the echo from a lighthouse at a tidal minimum.

5·6. Interference Measurements over Long Ranges.[1]—Measurements of the reflection coefficient of sea water at wavelengths of 3 and 10 cm were made by the Radiation Laboratory on a one-way transmission path in Massachusetts Bay.[2] One terminal of the path was fixed on land; the other was carried in an airplane that flew toward the fixed terminal with as nearly constant bearing, altitude, and ground speed as possible. In this manner a horizontal cross section of the interference pattern was obtained, from which ρ was determined. Such a record is shown in Fig. 5·15. This type of measurement technique was employed to simulate the practical situations in which the data are to be applied; that is, to conditions over the ocean at considerable heights and distances. The heights and distances at which the observations were made are much greater than those frequently employed in reflection-coefficient measurements and are

[1] By William T. Fishback and Donald E. Kerr.

[2] W. T. Fishback and P. J. Rubenstein, "Further Measurements of 3- and 10-cm Reflection Coefficients of Sea Water at Small Grazing Angles," RL Report No. 568, May 17, 1944.

the major distinguishing characteristics of the experiment. Flights were made at heights between 500 and 5000 ft, and useful data were obtained between 5 and 60 statute miles.

In the interpretation of a record of the type shown in Fig. 5·15 a number of considerations are involved. The formula for the ratio of actual to free-space fields valid for isotropic radiation patterns and a smooth spherical earth is $F = |E/E_o| = |1 + D\rho e^{-i\alpha}|$ [Eq. (2·452)], where α is the phase difference between direct and reflected rays, D is the divergence factor, and

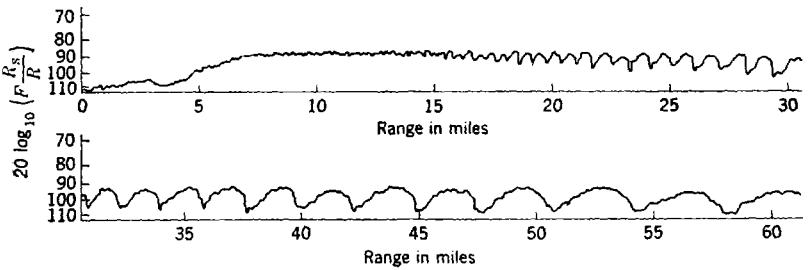


FIG. 5·15.—Reflection-coefficient field-strength flight of Apr. 7, 1943, using vertically polarized 10-cm radiation with terminal heights of 122 and 5000 ft over sea water.

E and E_o are the actual field strength and free-space field strength at a point in space. In a region of fine lobe structure α changes rapidly with position; in particular, for the cases under consideration, the fractional change in range and the change in grazing angle between a minimum and an adjacent maximum of the field pattern in Fig. 5·15 are so small that $D\rho$ and E_o may be considered constant for this section of the pattern. If E_{max} and E_{min} are the field strengths at a maximum and an adjacent minimum, at which points α is $2n\pi$ and $(2n + 1)\pi$, respectively, the preceding equation yields

$$D\rho = \frac{|E_{max}| - |E_{min}|}{|E_{max}| + |E_{min}|} = \frac{F_{max} - F_{min}}{F_{max} + F_{min}}, \tag{39}$$

from which ρ can be determined by using the observed values of E_{max}, E_{min}, and D calculated from Eq. (2·441). This assumes D to have its theoretical value for a smooth surface and requires all departures from theory to be represented by the behavior of ρ. This is necessary to compare results for a wide range of transmission-path geometry, as D and ρ depend in such different ways on the grazing angle ψ_2.

This technique is not suitable for measurement of the phase shift on reflection, since extremely accurate determination of the positions of maxima would be required and in addition the inherent fluctuations of the signal would probably make such determination impossible.

Figure 5·16 shows the values of ρ, measured by the technique outlined above, at 10 cm for horizontal and vertical polarizations; Fig. 5·17 shows the same results for 3 cm. Each dot represents a value of ρ determined by a measurement of field strengths at adjacent maxima and minima. At both wavelengths the agreement with theory for a smooth sea is good for vertical polarization, but the experimental values for horizontal polariza-

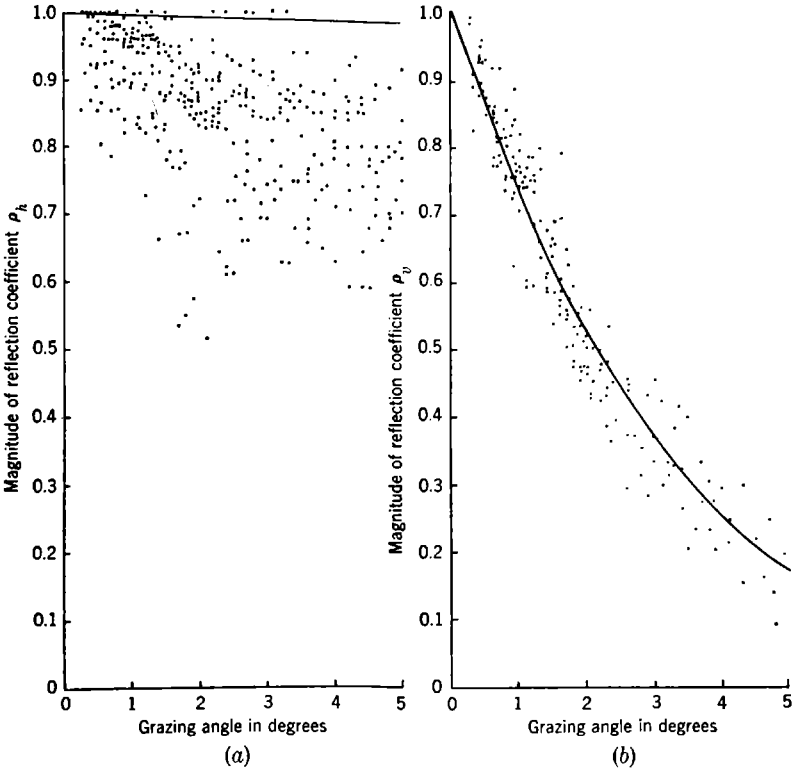


FIG. 5·16.—Values of the magnitude of the reflection coefficient observed by the Radiation Laboratory at a wavelength of 10 cm over sea water. (*a*) On horizontal polarization; (*b*) on vertical polarization.

tion are appreciably less than the theoretical values of almost unity and show a marked scatter.

The departure of the results from theory is attributed to surface roughness, but a satisfactory quantitative explanation of the effects of roughness has not yet been found. Let us consider in a qualitative way the possible effects of roughness, however, particularly as it affects Eq. (39), which served as a means of translating the interference pattern measurements of

Fig. 5·15 to the indicated experimental values of ρ in Figs. 5·16 and 5·17. It is desirable to seek a reason for the large scatter and the deviation from theoretical values of ρ for horizontal polarization, by contrast to the apparent better agreement with theory for vertical polarization.

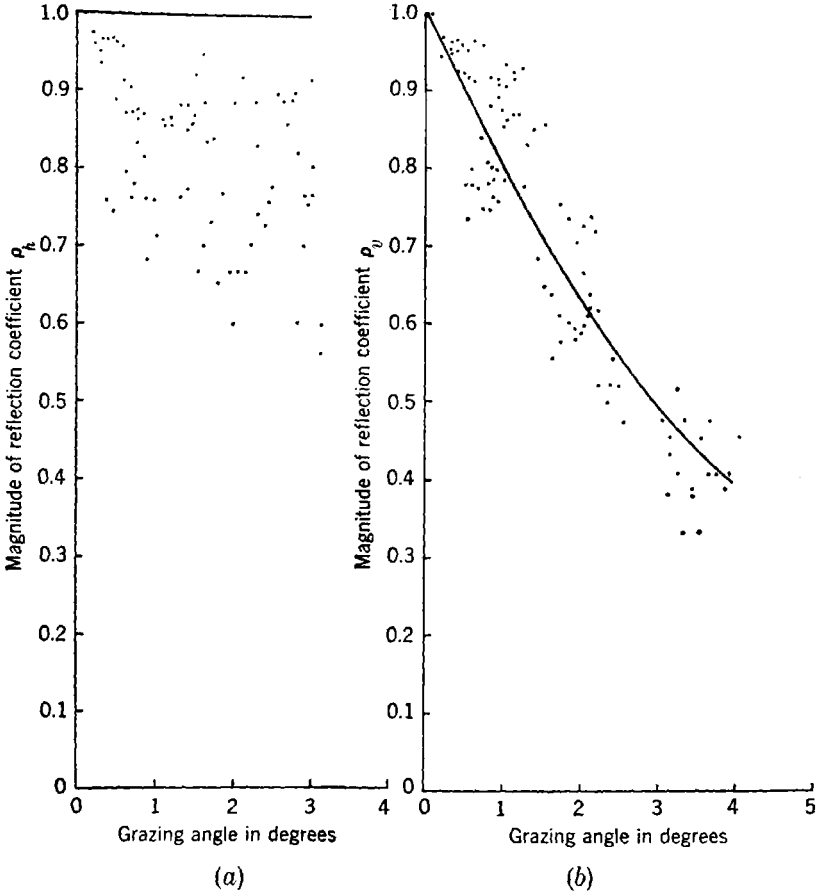


FIG. 5·17.—Values of the magnitude of the reflection coefficient observed by the Radiation Laboratory at a wavelength of 3 cm over sea water. (*a*) On horizontal polarization; (*b*) on vertical polarization.

From Eq. (39) it can be seen that when $D\rho$ is very nearly unity, $|E_{\min}| \ll |E_{\max}|$. Accurate measurement of $D\rho$ in this region requires very accurate determination of $E_{\min}$. If the field is fluctuating at the time of measurement, $E_{\min}$ in the vicinity of the interference minimum is subject to a relatively large random variation. Let us attempt to represent this effect crudely in the following way. Assume that $D\rho$ has

its theoretical value for a smooth surface at the average level of the sea surface. Superposed on the reflections predicted by this $D\rho$ are many reflections that, when added vectorially, can be represented by a small quantity ηe^{-ir} which fluctuates randomly in both space and time. Then we can write

$$F = |1 + D\rho\, e^{-i\alpha} + \eta e^{-ir}|. \tag{40}$$

Because $D\rho\ e^{-i\alpha}$ is nearly unity in a maximum and η is assumed small, $F_{max} \approx 2$. Near a minimum, η becomes important, as $F_{min} \approx \eta$. Equation (39) then becomes

$$D\rho \approx \frac{2 - \eta}{2 + \eta} \approx 1 - \frac{\eta}{2}. \tag{41}$$

Thus we see that values of $D\rho$ obtained by the use of Eq. (39) and a rapidly moving receiving terminal are always at the mercy of the rapid time variations in reflecting properties of the sea and will give values of $D\rho$ less than the theoretical values when these are very nearly unity, as they are for horizontal polarization.

When $D\rho$ is somewhat less than unity, Eq. (39) can be rewritten to show the range of variation likely to be encountered,[1]

$$D\rho \approx \frac{|1 + D\rho \pm \eta| - |1 - D\rho \pm \eta|}{|1 + D\rho \pm \eta| + |1 - D\rho \pm \eta|}. \tag{42}$$

This equation indicates that the measured values of $D\rho$ should scatter about the average value of $D\rho$ when $D\rho$ is not near unity; and as η is equally likely to have a given value in either a maximum or minimum, the scatter should be nearly symmetrical about the average value. The data in Figs. 5 16 and 5·17 are not sufficiently plentiful either to confirm or to refute this point, although no marked disagreement is apparent.

This discussion makes no pretense of rigor or completeness, but it makes plausible a connection between time-varying surface-roughness effects and the values of reflection coefficient deduced from measurements of the interference pattern. During the course of these experiments attempts to correlate the minor day-to-day variations of measured reflection coefficient for horizontal polarization with sea roughness and direction of wave travel were unsuccessful. A wide range of roughness could not be investigated in this case, because navigational difficulties limited flights to days when wind speeds were low, and measurements could be made only for a calm sea.

[1] This is not a true representation of Eq. (39), as F_{max} is not necessarily $|1 + D\rho \pm \eta|$ but rather, for the purpose of these measurements, the maximum value of Eq. (40) encountered as the receiving terminal traversed the region of a maximum; a similar statement is true for F_{min}. Also note that η does not necessarily have the same value at any two positions or times.

It is reasonable to expect that the observed value of ρ would depend upon the surface area illuminated if surface roughness and inhomogeneity are important in the reflection process. Thus the ρ appropriate for a transmitter height of 30 ft might not be appropriate for a height of 1000 ft. In the former case only a small surface area of approximately homogeneous surface roughness might be involved, but in the latter many such areas with different roughnesses might be involved. During the 10-cm measurements the fixed-terminal height was varied from 25 to 125 ft and the aircraft altitude from 500 to 5000 ft; at 3 cm the fixed terminal was at 20 ft and flights were made at 500 to 1000 ft. No dependence of ρ upon the area illuminated was observed within the narrow limits allowed by this variation of terminal heights.

This type of experiment is susceptible to several sources of experimental error, which may be divided roughly into three classes. The first consists of instrumental errors. At very short ranges and high angles the interference pattern was traversed so rapidly that the recorder could not indicate the full depth of the minima, and it was possible to use only the measurements made at grazing angles less than about 5° on 10 cm and 3° or 4° on 3 cm.

The second possible source of error arises from impure polarization of the radiation from the antennas, but in the present instance this error is at most about 2 per cent in the product $D\rho$ and can be neglected.[1] It is worth noting that measurements based on Eq. (39) are most subject to error when $D\rho$ is nearly unity, as weak stray fields tend to "fill up" the minima in the interference pattern and decrease the calculated value of $D\rho$.

The third and most important sources of experimental error arise from physical factors outside the equipment and constitute the principal limitations of the method. When one terminal is on the shore, the properties of that part of the ocean surface where most of the measured reflection occurs are likely not to be representative of the open sea, because of the effects of shore line, shallow bottom, underwater obstacles, and so forth, on the water-wave structure. To prevent extraneous "filling up" of the interference minima, care must be taken to ensure that no radiation strikes objects other than the sea surface; this precaution will

[1] The upper limit of this error may be estimated in the following manner. Assume that the ratio of amplitude of field radiated with the wrong polarization to that with correct polarization is ϵ for both transmitting and receiving antennas. When the receiving antenna is in a minimum of the interference pattern and $D\rho$ is near unity, the received field is obtained primarily on the wrong polarization and is at most ϵ^2 times the correctly polarized free-space field. In the maximum of the interference pattern, however, the former is outweighed by the correctly polarized field, which is nearly twice the free-space value. Then the value for $D\rho$ obtained from Eq. (39) is approximately $(2 - \epsilon^2)/(2 + \epsilon^2)$, for $D\rho$ near unity. Thus impure polarization reduces the apparent value of $D\rho$.

frequently require the use of narrow beams, a fact that complicates the aircraft-navigation problem. Nonstandard refraction may distort the interference pattern, rendering it useless as a quantitative measure of reflection from the surface (see Sec. 4·21).[1] Nevertheless the principal source of difficulty lies with navigation of the aircraft. Ideally one would use radar or, better still, radar-beacon navigational techniques, but these were not available; consequently excessive deviations from correct position and aircraft heading frequently necessitated discarding considerable sections of record.

5·7. Interference Measurements at Short Ranges.[2]—The Sperry Gyroscope Company[3] has measured reflection coefficients by varying the terminal heights on a transmission path of a few hundred feet in length. Over such a short path the effects of surface roughness may not be pronounced, and the observed values of ρ are frequently in better agreement with the theory for smooth surfaces. Measurements made over extremely calm sea water with ripples of 2 in. or less are in good agreement with theory (using $\varepsilon_1 = 69$ and $\sigma = 6.5$ mhos/m) for grazing angles up to 20°. On occasion the observed values for vertical polarization are slightly below the theoretical values for grazing angles between 12° and 28°. Figure 5·18 shows the observed values of ρ for a tidal canal with ripples estimated to be about 1 in. high. Measured values of ρ for very calm fresh water are in good agreement with the theoretical values for a perfect dielectric with $\varepsilon_1 = 80$, for angles up to 18°. At greater grazing angles the values are somewhat below the theoretical values for both polarizations.

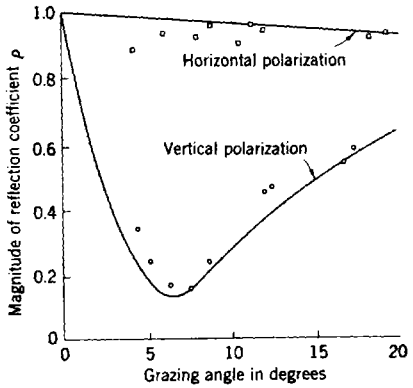


FIG. 5·18.—Values of the magnitude of the reflection coefficient observed by the Sperry Gyroscope Company at 10 cm over a tidal canal with 1-in. ripples. The circles give measured points and the curve represents the theoretical values for smooth water.

[1] The measurements reported here were for the most part accompanied by meteorological soundings; except for possible shallow evaporation ducts (the effects of which were not determined), the *M*-profiles were required to be standard for the field-strength measurements to be used for measurement of ρ. For the measurements to be considered reliable, the field strength at the maxima in the interference pattern was required to vary inversely with range and the observed ranges of maxima and minima were required to check closely with the ranges calculated for standard refraction.

[2] By William T. Fishback and Donald E. Kerr.

[3] See E. M. Sherwood, "S-band (10-cm) Measurements of Reflection Coefficients for Various Types of Earth," Sperry Gyroscope Co. Report No. 5220–129, Oct. 29, 1943.

Measurements were made by the Navy Electronics Laboratory[1] in sheltered harbor waters near San Diego at a fixed range of 4900 yd and grazing angle of approximately 1° 39′, using radar systems and specially constructed targets instead of the customary one-way circuit.[2] The height of the target was varied through the interference pattern and the resulting variation in echo strength was used to obtain ρ by a procedure that may be illustrated in principle in the following way: If ϕ is assumed to be 180°, the equation for F can be written

$$F^2 = 1 + (D\rho)^2 - 2D\rho \cos(k\,\Delta R);$$

F^2 varies between $(1 - D\rho)^2$ and $(1 + D\rho)^2$ as $k\,\Delta R$ varies between $2n\pi$ and $(2n + 1)\,\pi$ (where n is an integer) and is a linear function of $\cos(k\,\Delta R)$. A plot of F^2 as a function of $\cos(k\,\Delta R)$, derived from the echo strength,[3] then gives a straight line of slope $-2D\rho$. If the slope rather than the value of F^2 in a minimum is used as a measure of $D\rho$, the difficulties from rapid fluctuations in the minimum are avoided, and the accuracy then depends upon fitting a straight line through a series of points that are also subject to fluctuation, although perhaps relatively less than those in the minimum.

After corrections were made for finite vertical target extent, the effective values of ρ for horizontal polarization were so near unity that the departure from unity could not be detected on wavelengths between 2 m and 44 cm. Similar measurements on vertical polarization gave values between 0.6 and 0.4, slightly lower than the theoretical values for a smooth surface.

Well-defined specular reflection of vertically polarized 1.25-cm waves from the sea was observed by the Radiation Laboratory on a one-way transmission path across the harbor at Gloucester, Mass. The terminals were about 40 ft above mean sea level. The variations of tide level of 8 to 10 ft changed ΔR by about three wavelengths, producing a periodic variation of received field strength as the interference pattern moved past the receiving antenna. Because of the fluctuations of the signal caused by the rough surface the ratio of field strength at the maxima and minima of the interference pattern could not be determined accurately, but it appeared to be in the neighborhood of 3 or 4, corresponding to values for ρ_s of 0.5 to 0.6. The grazing angle ψ_2 was about 0.2°. It is interesting to note

[1] L. Anderson, J. B. Smyth, F. R. Abbott, and R. Revelle, "Radar Wave Propagation," NRSL Report No. WP-2, Nov. 30, 1942, pp. 10–12. Further results were presented at a conference at the Radiation Laboratory in July 1943.

[2] This method is susceptible to error because of the finite vertical dimensions of the target. At short wavelengths it may intercept several lobes; even if this is not true, it is likely to have sufficient vertical extent to cover a very sharp minimum. In the present instance this effect was corrected approximately by integrating the field over the target to obtain the echo strength. The principle of this method is discussed in Sec. 6·5.

[3] The echo voltage output of a linear *radar* receiver is proportional to F^2.

that in a maximum the signal fluctuated about 3 db but in a minimum 8 to 10 db. This is in keeping with the discussion in Sec. 5·6.

Some of the most interesting evidence of specular reflection from the sea has been obtained by the General Electric Company of England.[1] One-way transmission was carried out with vertical polarization on wavelengths of 5.81 and 6.35 mm between a fixed shore transmitter at heights of 15 and 48 ft and a shipborne receiving terminal at a height of 63 ft. The range varied between about 1 and 10 km, but the grazing angle was always less than 1°. The ship followed a radial course from the shore station, and the resulting signal trace, obtained on recording milliammeters operating from the receivers, showed well-developed and fairly regular interference patterns. The positions of the maxima and minima agreed well with theory, and there was no evidence of superrefraction or other extraneous effects.[2] The reflection coefficient ρ_v was calculated from relative field strengths at the maxima and minima of the interference pattern and was found to vary between 0.25 and 0.8 on different runs; theory for a smooth sea and the same range of grazing angles predicted ρ_v to vary between 0.87 and 0.97. The sea was reported to be fairly calm in all cases, having ripples of less than 12 in. and low swell. On the basis of the Rayleigh criterion discussed in Sec. 5·4 the heights of the surface irregularities would be required to be much less than 12 in. (more nearly 1 in.) over most of the range if ρ were to approach the theoretical value. Further experiments of this type, with the addition of objective measurements of sea state, are very much needed.

5·8. Interpretation of Measurements.[3]—The measurements described in Secs. 5·6 and 5·7 indicate that for wavelengths of about 10 cm or less the amount of specular reflection is markedly affected by surface roughness. When the effective part of the reflecting surface is very smooth, which is usually true only over short paths, the measured values of ρ are in good agreement with the theoretical values for a smooth surface calculated from the known values of conductivity and dielectric constant. When the surface is rough, there is enough scattering of the incident radiation, particularly if it is horizontally polarized, to diminish the average intensity of the reflected wave appreciably. This can be equally true either at short or long ranges, but it is particularly noticeable over long transmission paths, where it is improbable that the whole reflection area will be sufficiently smooth for efficient specular reflection. The effects of surface roughness are manifested not by lowering ρ to a fixed

[1] H. R. L. Lamont and A. G. D. Watson, *Nature*, **158,** 943 (1946).

[2] The wavelengths were chosen to place the 5.81-mm signal in the oxygen absorption band and the 6.35-mm signal in a region where absorption should be negligible (see Chap. 8). The 5.81-mm signal showed, in addition to the interference pattern, an exponential range attenuation, from which a value of attenuation coefficient of 1.5 db/km was estimated, a value considerably larger than that given in Chap. 8.

[3] By Donald E. Kerr and William T. Fishback.

value but by causing it to vary over a wide range during short time intervals.

Because of the uncertainty in the field strength resulting from these fluctuations there is no unique value of field strength above the ocean at a given point in the interference region; the field strength varies continually and rapidly. In terms of coverage, this means that a unique coverage pattern does not exist, even for the overidealized case of a point target and an isotropic antenna radiation pattern. Prediction of radar coverage is then a statistical problem involving the instantaneous value of ρ for the sea surface and orientation and position of the target, in addition to the transmitted power, antenna gain, receiver sensitivity, etc. In the absence of more detailed information than that presented in Fig. 5·16 we suggest that a reasonable procedure is to calculate an "average" coverage diagram, using the theoretical value of ρ for vertical polarization. No universally applicable average value of ρ for horizontal polarization can be determined from Figs. 5·16 and 5·17. It should be noted, however, that the observed values appear to be, on the average for any angle, slightly higher than those for vertical polarization, and a conservative procedure would be to use the values for vertical polarization.

Very few measurements are available for wavelengths longer than 10 cm, but practical evidence with meter-wave radar (using horizontal polarization almost without exception) indicates that ρ for the sea is very nearly unity for grazing angles up to several degrees. The Rayleigh criterion does not seem always to be reliable, as large roughness effects have occasionally been observed even when the critical roughness was not exceeded. This is probably to be expected, for, as mentioned earlier, the Rayleigh criterion cannot be considered adequate for the present problem. Much more experimental information is needed, therefore, on the prevalence and magnitude of the fluctuations, and a more comprehensive theory of the scattering from rough surfaces is also required. The practical importance of these questions requires no underlining. The answers are needed to determine to what extent free-space coverage is modified in the interference region by reflection from the sea.[1]

REFLECTION COEFFICIENT OF LAND

5·9. Measurements over Long Ranges.[2]—Reflection-coefficient measurements over land were made by the Radiation Laboratory[3] at a wave-

[1] The reader should not forget that for very small grazing angles ρ is effectively unity even for fairly rough surfaces. This fact is undoubtedly of importance for long-distance transmission caused by surface ducts, in which the elementary waves strike the surface at very small grazing angles.

[2] By William T. Fishback.

[3] P. J. Rubenstein and W. T. Fishback, "Preliminary Measurements of 10-cm Reflection Coefficients of Land and Sea at Small Grazing Angles," RL Report No. 478, Dec. 11, 1943.

length of 10 cm using a technique identical with that described in Sec. 5·6. Measurements were made over flat, sandy land, with little vegetation, near Orlando, Fla.; over very flat farm land with low vegetation on Long Island; and over hilly, rocky land, thickly settled or wooded, near Boston. Flights were made between 600 and 10,000 ft, with the ground station sited at heights of 100 ft or less above the immediately surrounding terrain. For all three locations the signal recordings showed minor variations but no regular interference patterns up to grazing angles of about 5°.

In general the received power followed the inverse-square law of attenuation within the horizon, indicating that free-space propagation was occurring. Figure 5·19 shows the results of such a flight made at Boston

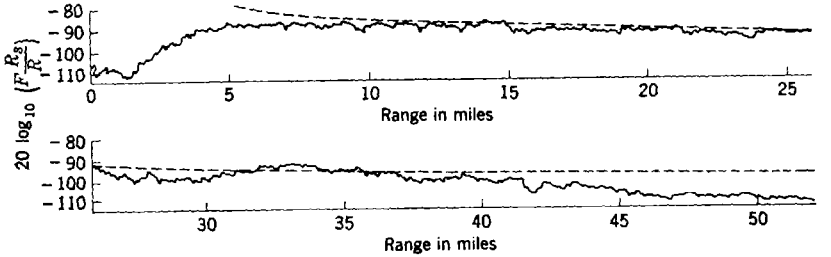


FIG. 5·19.—Reflection-coefficient field-strength flight of July 16, 1943, using horizontally polarized 10-cm radiation with terminal heights of 100 and 3000 ft over rough land. The dashed line indicates the free-space field.

with horizontal polarization. The observed signal obeys the inverse-square law well at ranges of less than 43 miles and has variations in strength of up to 5 db, but it does not exhibit any regular interference pattern.

The results obtained at the Florida site were similar and were substantiated by calibration flights on a near-by 10-cm radar system, which showed no extension of detection range and no observable lobe structure. Calibration flights for a 30-cm radar system at the same site did show noticeable lobe structure, however.

The locales selected for these measurments represent a broad range of surface conditions, from the very rough terrain near Boston to as flat land as will ordinarily be encountered. The results indicate that specular reflection over long ranges at wavelengths of 10 cm or less is negligible for most land surfaces. (Exceptional cases are cited below.) On the basis of the Rayleigh criterion, described in Sec. 5·4, for each degree of surface roughness and grazing angle there should be a wavelength region in which the reflection changes in character from specular to diffuse; longer wavelengths undergo appreciable specular reflection, but the shorter do not. The experiments indicate that this transition wavelength is greater than 10 cm for most land surfaces and grazing angles down to about $\frac{1}{4}°$. The radar evidence at the Florida site suggests that it is less than 30 cm for a

fairly smooth surface, although the measurements were not sufficiently detailed to determine the magnitude of specular reflection.

That specular reflection of microwaves over land is possible, given a smooth enough surface, is evident from observations made by the U.S. Army Signal Corps in California.[1] One-way transmission measurements were made on a wavelength of 6 cm on an overland path 170 miles in length between Blue Ridge (5860 ft) and Mt. Hamilton (4300 ft). The grazing angle for this path is 0.19°. Fading of as much as 30 db was observed at times when soundings made near the point of reflection in San Joaquin Valley revealed essentially standard M-profiles.[2] Installation of a second receiver below the first by an amount equivalent to one-half lobe of the expected interference pattern revealed that fading on the two circuits was complementary, that is, when the signal strength increased on one, it decreased on the other. The most obvious explanation of this phenomenon is that large specular reflection occurred, producing a lobe structure that shifted slightly in the vertical direction because of minor variations in the shape of the M-profile.[3] The geometry of the region of reflection lends plausibility to this suggested mechanism, because for many miles about the reflection point the land is semiarid and extremely flat, with almost no vegetation and an average slope of less than 5 ft per mile. It is reasonable to expect specular reflection over land at 6-cm wavelengths with as smooth a reflecting surface as this and the small grazing angle involved. This path is probably not typical of those to be encountered in microwave communications or radar.

A second interesting example of efficient reflection from very flat ground is offered by some results observed on the GCA (Ground-controlled Approach) radar instrument landing system. This system employs a very narrow 3-cm beam that scans vertically to measure the altitude of an approaching aircraft. It is a common occurrence for a strong and well-defined mirror image of the aircraft to appear as a result of reflection from the near-by ground, usually a very flat section of an airport.

Measurements of the reflection coefficient at 3 cm were made by the Bell Telephone Laboratories[4] on a 12.6-mile transmission path between Beer's Hill and Deal, N.J. (This is the path used for the experiments described in Sec. 4·23). The profile of the path was irregular enough to preclude exact calculation of the grazing angle, but its maximum value was considerably less than 1°. The interference pattern was investigated

[1] T. J. Carroll, "Complementary Diversity Reception on Microwaves," OCSigO Report No. RPS-4, Jan. 3, 1946; R. Bateman, *Proc. IRE*, **34**, 662 (1946).

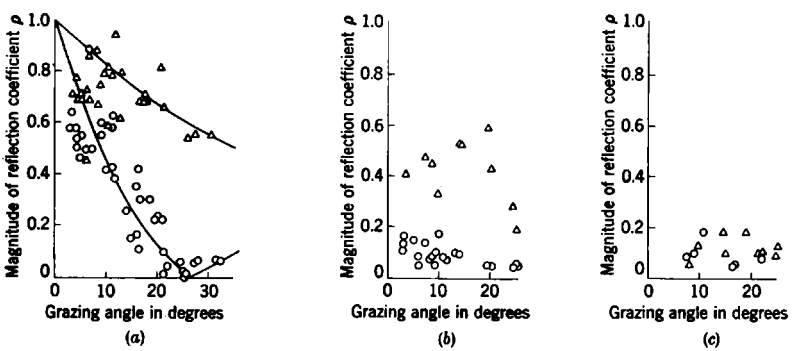
[2] Ground-based equipment was used, and irregularities above the sounding height may have occurred.

[3] Note the similarity to the results obtained in a somewhat different manner over water, described in Sec. 4·8.

[4] W. M. Sharpless, "Ground Reflection Coefficient Experiments at X-band (3 cm)," BTL Memorandum No. MM-44-160-250, Dec. 15, 1944.

by varying the transmitter height at such times when obviously standard refractive conditions prevailed. Such conditions occurred on Sept. 22 and Dec. 4, 1944. The results obtained on these two days show no marked difference in spite of the seasonal difference in foliage. The over-all space variation in signal was observed to be less than 4 db, but it was so regular and repeated so exactly throughout all the observations that undoubtedly a slight amount of specular reflection was taking place. The values of ρ determined from the experiment are 0.17 for horizontal and 0.20 for vertical polarization. Attempts to determine the height of the effective ground-reflection surface have been carried out for two different assumed reflection points. Both show the height to be above the true surface by an appreciable amount. This is not surprising in view of the irregular profile of the path, which undoubtedly renders indeterminate the area or areas at which specular reflection occurs.

5·10. Measurements at Short Ranges.[1]—The Sperry Gyroscope Company[2] also made measurements at wavelengths of 10 cm over various types



(*a*) Over small, dry sand hillocks.

(*b*) Over dry, slightly rolling land with grass 4 to 18 in. high.

(*c*) Over a beet field with weeds.

FIG. 5·20.—Values of the magnitude of the reflection coefficient observed by the Sperry Gyroscope Company at 10 cm.

of land. These were similar to their overwater measurements (see Sec. 5·7). High values of ρ were observed over level, dry, or wet sand; a finely plowed clay-sand field; tidal flats; and an airport. Moderate values were found over a slightly rolling field with grass 4 to 18 in. tall. Virtually no reflection was found when more thickly vegetated areas were investigated. Figure 5·20 shows typical values of ρ observed. Although an experiment carried out at very short ranges might be expected to be less affected by surface roughness, even this experiment confirms the fact that a very

[1] By William T. Fishback.

[2] E. M. Sherwood, "S-band (10-cm) Measurements of Reflection Coefficients for Various Types of Earth," Sperry Gyroscope Co. Report No. 5220–129, Oct. 29, 1943.

smooth surface is required to prevent diffuse reflection of microwave radiation. The types of surface over which this experiment reveals good reflection are usually limited in area and would be likely to affect only a small portion of a radar system coverage diagram or a very fortunately sited communications link, such as the one used in the U.S. Army's California test.

As was mentioned in Sec. 5·1, measurements have been made[1] of the reflecting and absorbing properties of various samples of earth at 9 cm to determine their permittivities and conductivities. The reflecting properties were determined by measurements over a 50-ft path with terminal heights of 10, 18.5, and 26.5 ft (corresponding to grazing angles of 20°, $36\frac{1}{2}$°, and $46\frac{1}{2}$°). Narrow antenna beams were used, and separate measurements made of the direct and reflected radiation. The surface consisted of a plot 20 by 30 ft which was raked and leveled with a plank until it was "optically" smooth. The reflecting properties of this sample for different degrees of wetness were found to be consistent according to Fresnel reflection theory for the various grazing angles available and were used to obtain the values of ground constants listed in Sec. 5·1.

Measurements were made over this and some other natural samples to determine the effect of surface roughness on the value of ρ. The trend predicted by the Rayleigh criterion (Sec. 5·4) was verified; it was found that if the product of the depth of the irregularity and sin ψ_2 was equal to $\lambda/5$, the values of ρ were reduced by one-half; if the product was $\lambda/2$, the observed values of ρ were 0.1 or less. Even a growth of short grass on the very smooth sample of earth was sufficient to reduce the value of ρ for horizontal polarization at these high grazing angles; the decrease was not so pronounced on vertical polarization, and on occasion an increase in the value of ρ was observed.

Attempts were made to simulate conditions over the sea by raking the sample into a series of ridges having wavelengths of 60 to 120 cm and amplitudes of 5 to 16 cm. The ridges were placed at angles of 0°, 45°, and 90° with the direction of transmission. In all three cases the observed values of ρ were materially decreased, so much so that evaluation of the effect of the various angles of inclination was not possible. Considerable spatial variation was present in the reflected radiation. The signal level measured over the reflected path was observed to vary with small displacements of the antennas, and the observed value of ρ was found to vary as the geometrical point of reflection moved from a wave trough to a wave crest.

5·11. Measurements of Time Variations.[2]—It has been shown that the reflection coefficient of even dry ground, as measured over short ranges on

[1] L. H. Ford and R. Oliver, "An Experimental Investigation of the Reflection and Absorption of Radiation of 9-cm Wavelength," RRB/C107, Oct. 27, 1944; also *Proc. Phys. Soc.*, **58**. 265 (1946).

[2] By Herbert Goldstein.

carefully leveled plots, can be high; it is the common experience, however, that specular ground reflection is rarely found on microwaves for long ranges and appreciable grazing angles. This is not surprising because the irregularity of most terrain is such that the Rayleigh roughness criterion is exceeded. As a result the intensity of the wave reflected from the ground surface is small compared with the direct wave.

It also seems reasonable that this small scattered wave would fluctuate strongly under certain circumstances. Thus, if the reflection area is heavily vegetated, one would expect that the nature of the surface would be constantly changing with time—because of the wind. Again, if the reflection takes place over city areas and includes roads and railroad yards, similar effects seem reasonable.

The transmission over a number of short paths, about 10 miles long, across the city of Boston was briefly observed simultaneously on wavelengths of 9.2 and 3.2 cm. At the longer wavelengths the fluctuations were about $\pm$ 0.2 db, with an average period of about 5 sec. On 3.2 cm the amplitude of fluctuation was about $\pm$ 0.5 db, with a period nearer to 2 sec. The dependence of the signal characteristics on wavelength is in agreement with the hypothesis that the fluctuation arises from the ground-scattered wave, and not from atmospheric turbulences (see Sec. 6·14).

Additional evidence for this view comes from the results of crossing the polarizations of transmitter and receiver. From the normal depolarization of the paraboloid antennas a decrease of 20 to 30 db is to be expected. Actually decreases of 10 to 25 db were observed, and the signal fluctuated violently between these limits. It seems therefore that over these paths there is an appreciable component scattered in such a manner as to rotate the direction of polarization.

As in the case of fluctuations in the reflection from water, the problem of ground reflection has been too little explored.

5·12. Interpretation of Measurements.[1]—The experiments described in Secs. 5·9 to 5·11 indicate that wavelengths of about 10 cm or less suffer specular reflection according to the Fresnel equations if the surface is sufficiently smooth but that they are sufficiently scattered by most types of earth surface so that little or no specular reflection will be observed except at almost zero grazing angle. The critical factor in determining specular reflection is the roughness of the surface, not the conductivity and dielectric constant of the surface. Reflection can be observed at either long or short ranges, as indicated by the California and Sperry measurements, but it is probable that it will be observed oftener at short ranges. Terrain similar to that over which Sperry found high reflection coefficients is fairly common, but it is usually not extended in area (desert regions are a possible important exception). A microwave system operating over long ranges may illuminate a smooth plowed field, but it may also illuminate trees,

[1] By William T. Fishback.

fences, buildings, and other large objects in the first few Fresnel zones to such an extent that the reflection will become largely diffuse. The specular reflection observed at long ranges on microwaves in California is understandable, in view of the small grazing angle and the presence of a very large and very flat area in the region of the point of reflection, but frequent occurrence of such a combination on overland microwave long-range communications links seems improbable. It seems even less probable that reflections and lobe structure would be pronounced on a 10-cm search radar system, for in this case the flat reflecting surface is likely to be limited in azimuth. (Specular reflection over ground has been observed on microwave radar systems, but only when the system was at a low site near an airport runway or similar flat surface.) Such systems operating over land should not be expected a priori to provide other than free-space coverage.

No quantitative data are available for long wavelengths; but judging from the meager qualitative information available for 30 cm and above, it would appear that specular reflection over ordinary terrain improves rapidly with increasing wavelength, but whether or not in accordance with the Rayleigh criterion one cannot say. Accurate determination of this trend is a problem for the future.

ERRORS IN RADAR HEIGHT MEASUREMENTS

By William T. Fishback

Although the range and azimuth of a radar target can be easily determined with considerable accuracy, the problem of measuring height is more involved and frequently susceptible to considerable error. To find the height of a radar target, it is convenient to measure the angular elevation of the target. Two common methods consist of varying the elevation angle of the radar antenna to find the direction from which maximum echo is received on single beam or varying the elevation angle to equalize the echo on two overlapping beams symmetrically displaced relative to the mechanical axis of the antenna.[1] Once the elevation angle is determined, graphs, tables, or automatic computers may be used to determine the target height over a spherical earth as a function of elevation angle and range, assuming standard refraction. This assumption may lead to errors at very low angles under nonstandard refractive conditions; and although the magnitude of the angular error will be small, the resulting error in height may be excessive. Much greater errors will be introduced at low elevation angles when appreciable surface reflection occurs. When this is true, maximizing signal strength by varying tilt of the antenna will not

[1] The latter scheme has been referred to in the past as "pip-matching" or "lobe-switching."

guarantee that the antenna is pointed directly at the target, nor do equal echo strengths on overlapping beams guarantee that the angle of elevation of the antenna is the true angle of elevation of the target. This phenomenon has not been analyzed completely, but a qualitative summary is possible of the effects that surface reflection can have on radar height measurements. In the following discussion standard refraction will be assumed.

5·13. Qualitative Discussion.—We shall first consider the case of height measurement by maximizing echo strength on a single beam. In terms of the transmission formulas, we must find the value of antenna elevation angles ξ that makes the pattern-propagation factor F a maximum for a given ψ_1. In the case of free-space propagation we have, from Eq. (2·453),

$$F = f(\psi_1 - \xi),$$

which has a maximum value of unity when $\psi_1 = \xi$. In other words, there will be no height-finding error introduced in this case, as one would expect. When surface reflection takes place, errors are introduced. When the target is in a lobe maximum, we must maximize

$$F = |f(\psi_1 - \xi) + D\rho f(-\psi_2 - \eta - \xi)|,$$

but for a target in a minimum we must maximize

$$F = |f(\psi_1 - \xi) - D\rho f(-\psi_2 - \eta - \xi)|,$$

by variation of ξ alone. In the former case F has a maximum value for $\xi < \psi_1$; in the latter case the maximum occurs for $\xi > \psi_1$, where ξ is a positive angle. There will also be negative maximum values of ξ where $\xi > (-\psi_2 - \eta)$ in a lobe maximum and $\xi < (-\psi_2 - \eta)$ in a lobe minimum. As a result the indicated elevation angle will oscillate about the true elevation angle and its mirror image as the target moves through the interference pattern. We can see the physical reason for this variation of ξ by considering a radar with its antenna pointed directly at a target in a lobe maximum. If the antenna is lowered, the return along the direct ray will not be greatly reduced, but the return along the reflected ray will be materially increased. Thus the total returned signal will be increased by decreasing the antenna elevation angle. In a like manner, if the target is in a minimum and the antenna is raised, there will be less destructive interference from the reflected ray and the signal return will be increased. A similar argument holds for oscillations about the mirror image of the true elevation angle.

An example of the errors produced is shown in Fig. 5·21 in which F^2 (a measure of the received field at the radar or of the energy density incident on the target) has been plotted as a function of ξ for targets at

elevation angles of 1°, 2°, and 3°. An "error-curve" antenna pattern with a 3° beamwidth was used, and it was assumed that $\psi_1 \approx \psi_2 + \eta$, which is a good approximation for a radar at a very low altitude looking at an aircraft. An effective reflection $D\rho$ of 0.8 was used. The upper and lower solid curves give the value of F^2 when the target is in a maximum and minimum respectively; the dotted curve shows the value of F^2 that would be observed in free space, which is just f^2. It can be seen that for high enough angles accurate height measurements will be possible. This

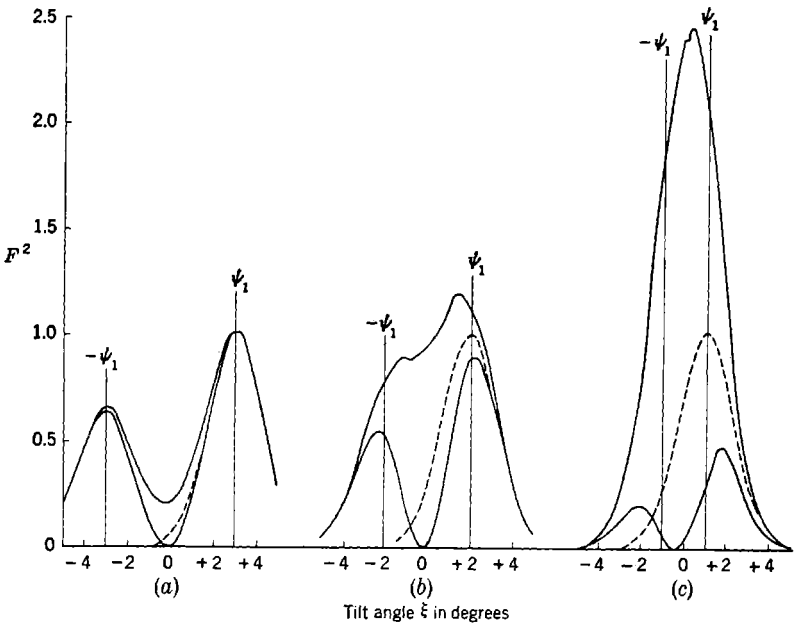


FIG. 5·21.—Relative intensity of illumination of a target at elevation angles of (*a*) 3°, (*b*) 2°, (*c*) 1°, as a function of tilt angle. Half-power beamwidth = 3°; $D\rho = 0.8$.

region is between 2° and 3° in this case. As the elevation angle of the target decreases, the field strength in a maximum will increase and that in a minimum will decrease. At the same time the indicated elevation angles in a maximum will approach zero; indeed as indicated in Fig. 5·21*c* the system will scarcely be able to resolve the target and its image. In a minimum, however, the value of $|\xi|$ for maximum F^2 will diverge more and more from zero as the true elevation of the target decreases.

The figure would seem to indicate that at low elevation angles the indicated height would oscillate about the true height, varying in both directions. If the low elevation angle occurs at a great range, however, the value of F^2, even when at its maximum, may be very small in an interference minimum (see Fig. 5·21). As a result, the indicated heights

may all be too small, as the echo may be strong enough to be detected and maximized only when the value of F^2 is sufficiently large, near an interference maximum. It has been observed in practice that indicated heights for low elevation angles and great ranges are all too low because of this phenomenon.[1]

In the "lobe-switching" scheme of height-finding, we do not maximize F but require that the F associated with the upper beam be equal to the F associated with the lower beam. Let ξ be the elevation angle of the cross-over point of the two beams, and let 2ζ be the angular separation between the beam maxima, as illustrated in Fig. 5·22. Now when the target is in a lobe maximum, we must have

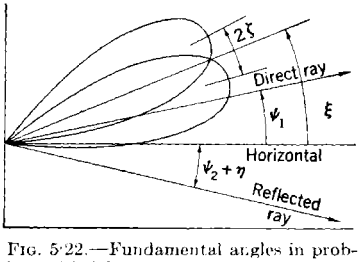


FIG. 5·22.—Fundamental angles in problem of height-finding by lobe-switching.

$$\begin{aligned} F &= |f(\psi_1 - \xi - \zeta) + D\rho\, f(-\psi_2 - \eta - \xi - \zeta)| \\ &= |f(\psi_1 - \xi + \zeta) + D\rho\, f(-\psi_2 - \eta - \xi + \zeta)|, \end{aligned} \tag{43}$$

whereas in a lobe minimum we require

$$\begin{aligned} F &= |f(\psi_1 - \xi - \zeta) - D\rho\, f(-\psi_2 - \eta - \xi - \zeta)| \\ &= |f(\psi_1 - \xi + \zeta) - D\rho\, f(-\psi_2 - \eta - \xi + \zeta)|. \end{aligned} \tag{44}$$

Though for a single beam it seemed reasonable to expect two regions of oscillation, one about the true elevation and the other about the mirror image, in lobe-switching it would seem reasonable to expect a third region of oscillation at the point where the antenna was pointed horizontally. Because the return along the reflected ray is always less than along the direct ray ($D\rho < 1$), this third branch should occur in the region of very slightly negative tilt angles. This hypothesis is confirmed in Fig. 5·23, which is a plot of indicated height against true height at a constant range where the dotted lines A and B are solutions of Eq. (43) and C, D, and E are solutions of Eq. (44). The solid lines are the graph of indicated height that the radar system could observe. The figure shows that two negative branches do indeed exist with a lobe-switching scheme but that in the region ED, these negative branches merge and accurate height measurement is possible only when the target is near an interference minimum. In this region the oscillations become discrete "islands," which are the only part of the negative branch of practical importance. Most errors will occur during the oscillations in AC with occasional excursions into DE.

[1] It should also be pointed out that side lobes have been neglected in Fig. 5·21. They will cause very small variations of indicated elevation angle in the region where good height measurement is to be expected (Fig. 5·21*a*).

At greater heights (and elevation angles) where the allowed regions are AC, BE, and BD, the system will remain in AC and will give fairly accurate height data.

No precise quantitative analysis of the limiting angles for good height measurement has been carried out, but some plane-earth calculations for 10 cm,[1] assuming an error-curve antenna pattern and the theoretical values of ρ for a smooth sea, give an estimate of the limiting angle of elevation below which height measurements are likely to be seriously in error. The angular error will be 0.1° or less for horizontal polarization if the antenna is elevated by at least its half-power beamwidth and for vertical polarization by 75 per cent of its beamwidth. The calculations showed that the advantage of single-beam maximizing over lobe-switching was negligible, eventually disappearing as beamwidths were narrowed. Operational evidence indicates that these values are somewhat pessimistic, at least for conical scanning, and that height-finding can be made successfully at somewhat lower angles. This is probably attributable primarily to the fact that the average reflection coefficient for horizontal polarization and sea water tends to be lower than the theoretical values for a smooth sea (see Secs. 5·5 to 5·8).

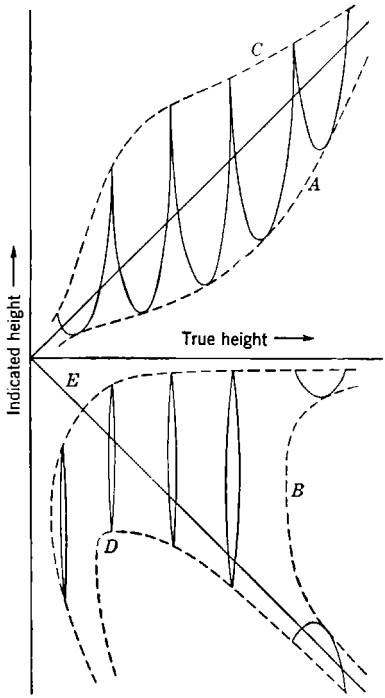


FIG. 5·23.—Height indicated by a lobe-switching radar as a function of true height at constant range.

Although this fact may indicate the reason for the increased region of good height measurements, it also shows that height cannot be measured down to zero elevation angle over water. As long as substantial specular reflection occurs, oscillations in the indicated elevation angle will occur with an amplitude depending upon ρ, the antenna beamwidth, and the elevation angle. Because reasonably large values of ρ occur near grazing incidence over water for both polarizations down to wavelengths at least as low as 3 cm, the only obvious practical way to increase the region of accurate height measure-

[1] R. A. Hutner *et al.*, "Radar Height-finding," RL Report No. 21, April 6, 1943. The analysis referred to was essentially empirical; it is probably not suited to general application in a strictly quantitative manner.

ments using the present schemes is to decrease the vertical beamwidth of the antenna by increasing the vertical aperture or by shortening the wavelength.

Overland specular reflection of microwaves is likely to be sufficiently small to permit accurate height measurements at small elevation angles, but care must be taken in siting a radar so that it will not illuminate any large area sufficiently flat to give appreciable specular reflection.

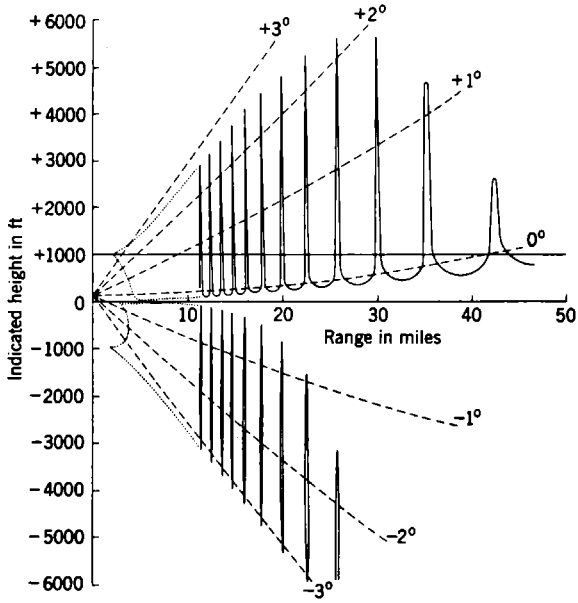


FIG. 5·24.—Theoretical values of indicated height of a target at 1000 ft viewed by a lobe-switching, horizontally polarized 10-cm radar sited at 125 ft and having a beamwidth of 4.4°.

5·14. Illustrative Examples.—An example of the oscillations in indicated height is shown in Fig. 5·24, in which indicated height is plotted as a function of range for a target height of 1000 ft. The curves are drawn for a horizontally polarized, conically scanning 10-cm radar system having a half-power beamwidth of 4.4° and sited at a height of 125 ft. Theoretical values of ρ for a smooth sea have been used in computing the curves. At short ranges only the envelopes of maxima and minima are shown. It will be seen that accurate height measurement will be possible in this case only at ranges of about 2 miles or less. At greater ranges large oscillations will occur, and on occasion the system may even indicate negative heights if it becomes locked in one of the "islands" of the negative branch. Figure 5·25 shows the envelopes of the oscillations for the geometry of Fig. 5·24 with horizontal and vertical polarizations. The difference between

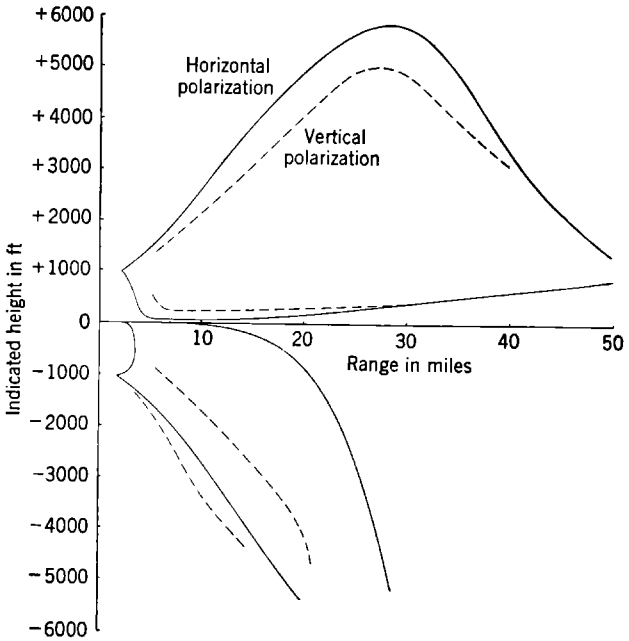


FIG. 5·25.—Effect of polarization on amplitudes of indicated height oscillations shown in Fig. 5·24.

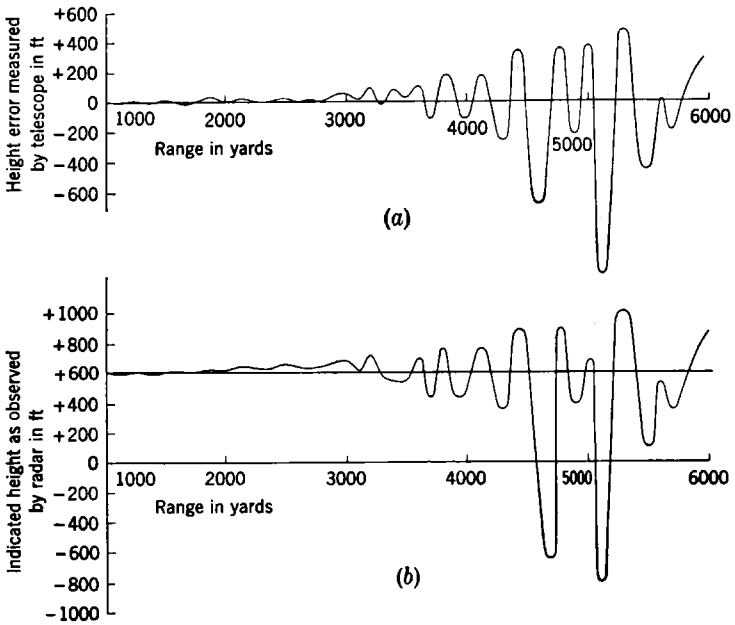


FIG. 5·26.—Height-finding errors observed by a laboratory model of SCR 584 tracking a plane at 600 ft. (*a*) Height error as measured by telescope; (*b*) height of target as indicated by radar.

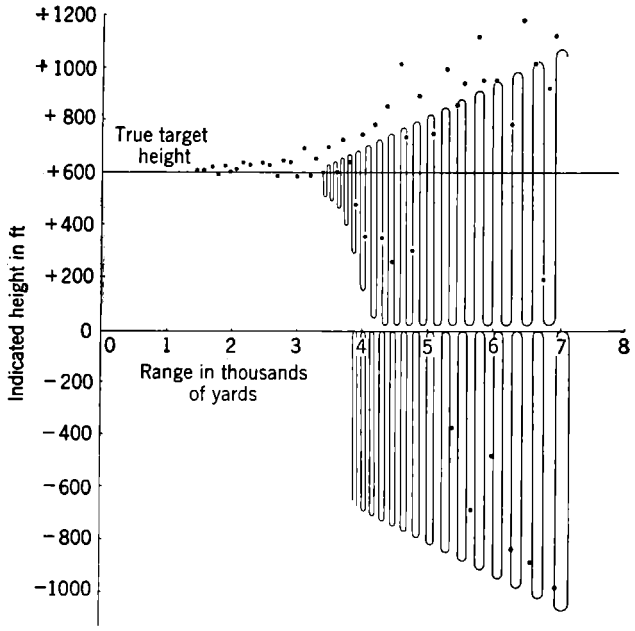


FIG. 5·27.—Comparison of theoretical and observed indicated heights for a target at 600 ft tracked by a laboratory model of SCR 584 at 100 ft over the sea.

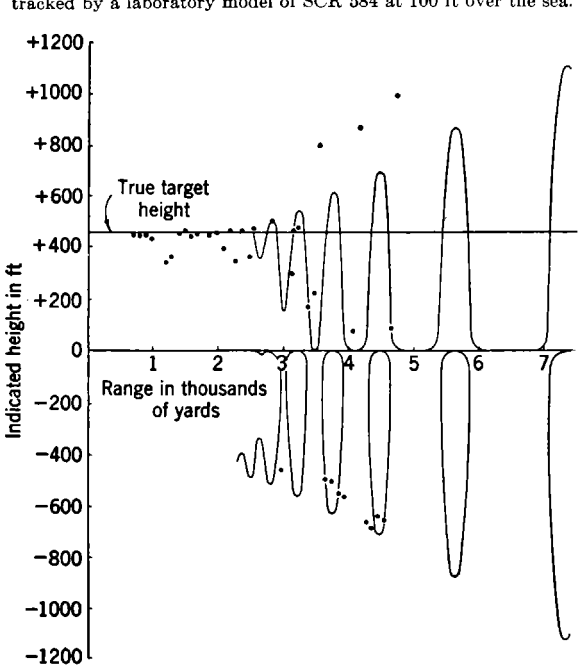


FIG. 5·28.—Comparison of theoretical and observed indicated heights for a target at 450 ft tracked by a laboratory model of SCR 584 at 25 ft over the sea.

polarizations is slight in the positive branch and would be even less had experimental values of ρ been used (see Fig. 5·16).

Experimental confirmation of height-finding errors is contained in some experiments carried out with an early laboratory model of the SCR-584 radar set in which the system was used to track an airplane flying a radial course at a constant height. During the flights the height indicated by the radar was recorded and checked by telescopic measurements of the angular displacement between the target and the antenna axis. Figure 5·26 shows the indicated height as observed by the radar and the height error as measured by the telescope as a function of range for a flight at 600 ft. In Fig. 5·27 points from several such flights are shown along with the appropriate theoretical curves. Figure 5·28 shows a similar plot for a target height of 450 ft and a radar height of 25 ft. In both cases the positions of maxima and minima check surprisingly well, although there is a tendency for the amplitude of the oscillations to exceed that predicted by the theoretical curves. This phenomenon is probably caused by variations in the instantaneous value of ρ and by mechanical overshoot of the tracking system. In all, about 30 sets of data were taken with this system for antenna heights of 25 and 100 ft and for target heights from 20 to 600 ft. On all flights large oscillations of indicated height were observed and frequent large "negative heights" were encountered.

CHAPTER 6

RADAR TARGETS AND ECHOES

By Donald E. Kerr and Herbert Goldstein

In the earlier parts of this book it has been assumed that radar targets could be characterized by the radar cross section σ, which has been treated primarily from the phenomenological point of view. In this chapter we shall examine the properties of σ and its relation to the properties of the target. It will be found that in only a negligible number of extremely simple cases is it feasible to calculate σ from the geometry; the remaining cases, which for the most part are those of greatest practical interest, are beyond the scope of existing mathematical methods.

THE RADAR CROSS SECTION OF ISOLATED TARGETS

By Donald E. Kerr

If the target subtends a sufficiently small angle, as viewed from the energy source, the incident wavefronts may be considered plane. The problem of calculating σ of such a target in free space is thus reduced to the problem of scattering of a plane wave. The problem can be solved exactly for a sphere and approximately for a few other simple shapes. The solution for the sphere will be given in some detail because the results are needed in later sections, and because the general method serves as a pattern that can be used in approximate form for other cases. In all cases the medium external to the target is assumed to be free space.

6·1. Scattering from a Sphere.—The diffraction of a plane wave by a sphere was first given by Mie[1] and has been elaborated upon by Stratton,[2] whose method of treatment will be outlined very briefly here. For details Stratton's analysis should be consulted.

To describe the scattering of plane electromagnetic waves from a sphere, solutions of the vector wave equation must be found that add up to give a plane wave at large distances from the sphere but are of the proper form to satisfy boundary conditions at the surface of the sphere. As is well known, in a charge-free homogeneous and isotropic medium **E**, **B**, **D**, and **H** satisfy the vector wave equation

$$\nabla^2\mathbf{C} + k^2\mathbf{C} = 0. \tag{1}$$

[1] Mie, *Ann. Physik*, **25.** 377 (1908).

[2] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, 1941, p. 563*ff*.

In spherical coordinates R, θ, ϕ, it is possible to construct solutions to Eq. (1) from the vectors $\mathbf{M}$ and $\mathbf{N}$ derived from the solutions of the *scalar* wave equation $\nabla^2\psi + k^2\psi = 0$ in the following manner:

$$\mathbf{M} = \nabla\psi \times \mathbf{R} = \frac{1}{k}\nabla \times \mathbf{N}, \tag{2}$$

$$\mathbf{N} = \frac{1}{k}\nabla \times \mathbf{M}, \tag{3}$$

where $\mathbf{R}$, the position vector, is $\mathbf{i}_r R$. It is clear that $\mathbf{M}$ and $\mathbf{N}$ are proper functions to represent $\mathbf{E}$ and $\mathbf{H}$, because each is proportional to the curl of the other, and it is also easily shown that $\nabla \cdot \mathbf{M} = \nabla \cdot \mathbf{N} = 0$. We shall express $\mathbf{E}$ and $\mathbf{H}$ in the form

$$\mathbf{E} = -\sum_n (a_n\mathbf{M}_n + b_n\mathbf{N}_n), \qquad \mathbf{H} = \frac{k}{i\omega\mu}\sum_n (a_n\mathbf{N}_n + b_n\mathbf{M}_n), \tag{4}$$

where the coefficients a_n and b_n are to be determined.[1]

The elementary solutions of the scalar wave equation are

$$\psi_{\substack{e\\o}mn} = \begin{matrix}\cos\\ \sin\end{matrix} m\phi\, P_n^m(\cos\theta)\, z_n(kR), \tag{5}$$

in which $P_n^m(\cos\theta)$ are the associated Legendre polynomials and $z_n(kR)$ are spherical Bessel functions appropriate to the range of R under consideration. The subscripts e and o describe solutions that are even or odd in ϕ. Inside the sphere of radius a, where the field remains finite, we use the Bessel function of the first kind:

$$z_n(kR) = j_n(kR) = \sqrt{\frac{\pi}{2kR}}\, J_{n+\frac{1}{2}}(kR), \qquad R \leqslant a. \tag{6}$$

Outside the sphere where the field is a traveling wave we use

$$z_n(kR) = h_n^{(2)}(kR) = \sqrt{\frac{\pi}{2kR}}\, H_{n+\frac{1}{2}}^{(2)}(kR), \qquad R \geqslant a, \tag{7}$$

where $H_{n+\frac{1}{2}}^{(2)}(kR)$ is the Hankel function of the second kind.[2]

The incident and scattered waves and the field inside the sphere must now be expressed in terms of $\mathbf{M}$ and $\mathbf{N}$. If Eq. (5) is inserted into Eqs. (2) and (3), and if the time factor is removed by writing $\mathbf{M} = \mathbf{m}e^{i\omega t}$ and $\mathbf{N} = \mathbf{n}e^{i\omega t}$, we find

$$\mathbf{m}_{\substack{e\\o}mn} = \mp \mathbf{i}_\theta \frac{m}{\sin\theta} z_n(kR)\, P_n^m(\cos\theta) \begin{matrix}\sin\\ \cos\end{matrix} m\phi - \mathbf{i}_\phi z_n(kR) \frac{dP_n^m(\cos\theta)}{d\theta} \begin{matrix}\cos\\ \sin\end{matrix} m\phi, \tag{8}$$

[1] The vectors $\mathbf{M}$ and $\mathbf{N}$ are assumed to form a complete set in the absence of space charge.

[2] The Hankel function of the second kind is required here because the time variation is assumed to be $e^{+i\omega t}$.

$$\mathbf{n}_{\substack{e_{mn}\\o_{mn}}} = \mathbf{i}_r \frac{n(n+1)}{kR} z_n(kR)\, P_n^m(\cos\theta) \begin{matrix}\cos\\ \sin\end{matrix} m\phi + \mathbf{i}_\theta \frac{1}{kR}\frac{d}{dR}[Rz_n(kR)] \frac{d P_n^m(\cos\theta)}{d\theta} \begin{matrix}\cos\\ \sin\end{matrix} m\phi \mp \mathbf{i}_\phi \frac{m}{kR\sin\theta}\frac{d}{dR}[Rz_n(kR)]\, P_n^m(\cos\theta) \begin{matrix}\sin\\ \cos\end{matrix} m\phi. \quad (9)$$

If the positive direction of propagation is the z-axis, the incident plane wave to be represented is of the form

$$\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} = \mathbf{a}\begin{pmatrix}E_0\\ H_0\end{pmatrix} e^{-ikR\cos\theta} \qquad (z = R\cos\theta),$$

where $\mathbf{a}$ is a unit vector with components along the x- and y-axes. It can be shown that a plane wave traveling in the z direction in free space and polarized on the x-axis can be expressed in terms of the vector functions by

$$\left.\begin{aligned}\mathbf{E}^i &= \mathbf{i}_x E_0 e^{i(\omega t - kz)} = E_0 e^{i\omega t}\sum_{n=1}^{\infty}(-i)^n \frac{2n+1}{n(n+1)}(\mathbf{m}_{o1n} + i\mathbf{n}_{e1n}),\\ \mathbf{H}^i &= \mathbf{i}_y \frac{E_0}{\eta_0} e^{i(\omega t - kz)} = -\frac{E_0}{\eta_0} e^{i\omega t}\sum_{n=1}^{\infty}(-i)^n \frac{2n+1}{n(n+1)}(\mathbf{m}_{e1n} - i\mathbf{n}_{o1n});\end{aligned}\right\} \quad (10)$$

where η_0 is the intrinsic impedance of free space, 120π ohms, and where

$$\mathbf{m}_{\substack{e_{on}\\o_{en}}} = \pm\,\mathbf{i}_\theta \frac{1}{\sin\theta} j_n(kR)\, P_n^1(\cos\theta) \begin{matrix}\cos\\ \sin\end{matrix}\phi - \mathbf{i}_\phi j_n(kR) \frac{dP_n^1(\cos\theta)}{d\theta} \begin{matrix}\sin\\ \cos\end{matrix}\phi, \quad (11)$$

$$\mathbf{n}_{\substack{o_{en}\\e_{on}}} = \mathbf{i}_r \frac{n(n+1)}{kR} j_n(kR)\, P_n^1(\cos\theta) \begin{matrix}\sin\\ \cos\end{matrix}\phi + \mathbf{i}_\theta \frac{[kRj_n(kR)]'}{kR}\frac{dP_n^1(\cos\theta)}{d\theta} \begin{matrix}\sin\\ \cos\end{matrix}\phi \pm \mathbf{i}_\phi \frac{[kRj_n(kR)]'}{kR}\frac{P_n^1(\cos\theta)}{\sin\theta} \begin{matrix}\cos\\ \sin\end{matrix}\phi. \quad (12)$$

The primes indicate differentiation with respect to the argument. It is of interest to note that m in Eqs. (8) and (9) has become unity, because of the physical requirement that the field must depend upon ϕ only as $\cos\phi$ or $\sin\phi$. The subscript for m is dropped to simplify notation. Equations (10) to (12) express the unperturbed plane wave incident on the sphere.

The induced secondary field consists of two parts, one scattered by the sphere and the other inside the sphere. They will be referred to as the scattered and transmitted fields, denoted by superscripts s and t. The transmitted field will involve the j_n radial functions, but the scattered field behaves as a spherical wave for large R and requires the $h_n^{(2)}$ of Eq. (7) in place of j_n. This modification in Eqs. (11) and (12) will be indicated by $\mathbf{m}^{(2)}_{\substack{e_{on}\\ o_{en}}}$ and $\mathbf{n}^{(2)}_{\substack{o_{en}\\ e_{on}}}$. If the propagation constant of the sphere is k_1, where

$k_1^2 = \omega^2\mu_1\epsilon_1 - i\omega\mu_1\sigma_1$, we can now write the scattered and transmitted fields

$$\left.\begin{aligned} \mathbf{E}^s &= E_0 e^{i\omega t} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} (a_n^s \mathbf{m}_{on}^{(2)} + i b_n^s \mathbf{n}_{en}^{(2)}), && (13) \\ \mathbf{H}^s &= -\frac{E_0}{\eta_0} e^{i\omega t} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} (b_n^s \mathbf{m}_{en}^{(2)} - i a_n^s \mathbf{n}_{on}^{(2)}), && (14) \end{aligned}\right\} R \geqslant a,$$

$$\left.\begin{aligned} \mathbf{E}^t &= E_0 e^{i\omega t} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} (a_n^t \mathbf{m}_{on} + i b_n^t \mathbf{n}_{en}), && (15) \\ \mathbf{H}^t &= -\frac{k}{\omega\mu_1} E_0 e^{i\omega t} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} (b_n^t \mathbf{m}_{en} - i a_n^t \mathbf{n}_{on}), && (16) \end{aligned}\right\} R \leqslant a.$$

The boundary conditions for continuity of tangential components of $\mathbf{E}$ and $\mathbf{H}$ at $R = a$ are

$$\left.\begin{aligned} \mathbf{i}_r \times (\mathbf{E}^i + \mathbf{E}^s) &= \mathbf{i}_r \times \mathbf{E}^t, \\ \mathbf{i}_r \times (\mathbf{H}^i + \mathbf{H}^s) &= \mathbf{i}_r \times \mathbf{H}^t, \end{aligned}\right\} \tag{17}$$

where the incident field is given by Eq. (10). Carrying through this operation yields four equations for the coefficients a_n^s, b_n^s, a_n^t, b_n^t, of which only the first two are of interest for the present purpose. If we assume that $\mu_1 = \mu_0$, and if we introduce the complex refractive index n_c and a convenient measure of the size of the sphere ρ,

$$n_c = \sqrt{\epsilon_c} = \frac{k_1}{k} = \sqrt{\epsilon_1 - i\sigma_1/\omega\epsilon_0},$$

$$\rho = ka = \frac{2\pi a}{\lambda},$$

$$\left.\begin{aligned} a_n^s &= -\frac{j_n(\rho)\,[n_c\rho j_n(n_c\rho)]' - j_n(n_c\rho)\,[\rho j_n(\rho)]'}{h_n^{(2)}(\rho)\,[n_c\rho j_n(n_c\rho)]' - j_n(n_c\rho)\,[\rho\, h_n^{(2)}(\rho)]'}, \\ b_n^s &= -\frac{n_c^2 j_n(n_c\rho)\,[\rho j_n(\rho)]' - j_n(\rho)\,[n_c\rho j_n(n_c\rho)]'}{n_c^2 j_n(n_c\rho)\,[\rho h_n^{(2)}(\rho)]' - h_n^{(2)}(\rho)\,[n_c\rho j_n(n_c\rho)]'}. \end{aligned}\right\} \tag{18}$$

The complexity of Eq. (18) and the lack of tables of j_n and h_n for complex argument place severe limitations on calculations for spheres of arbitrary size and complex refractive index. The progress that has been made thus far will be described in Chap. 7; for the moment, however, we shall continue with the development of general expressions for the power scattered and absorbed by the sphere.

The effect of the sphere on the incident wave may be most conveniently determined by integrating the Poynting vector over the surface of a spherical surface concentric with the scattering sphere. The time average

of the radial component of the Poynting vector is

$$\bar{S}_R = \tfrac{1}{2}\mathrm{Re}(\mathbf{E} \times \mathbf{H}^*) = \tfrac{1}{2}\mathrm{Re}(E_\theta H_\phi^* - E_\phi H_\theta^*),$$

where $\mathbf{E}$ and $\mathbf{H}$ are the sum of the incident and scattered fields. With appropriate grouping of terms, this becomes

$$\begin{aligned} S_R = \tfrac{1}{2}\mathrm{Re}[&(E_\theta^i H_\phi^{i*} - E_\phi^i H_\theta^{i*}) + (E_\theta^s H_\phi^{s*} - E_\phi^s H_\theta^{s*}) \\ &+ (E_\theta^i H_\phi^{s*} - E_\phi^i H_\theta^{s*} + E_\theta^s H_\phi^{i*} - E_\phi^s H_\theta^{i*})]. \end{aligned} \tag{19}$$

The integral of S_R over the external spherical surface gives the net *outflow* of energy from the sphere. As the sphere is a sink rather than a source of energy, this quantity is negative; it will be denoted by $-P_a$, where P_a is the power loss in the dielectric of the sphere. The integral of the first term on the right of Eq. (19) is zero, as it gives the net outflow of energy in the incident plane wave. The second term, when integrated, yields the total power scattered out of the incident wave and will be denoted by P_s, a positive quantity. If we denote the integral of the third term of Eq. (19) by $-P_t$, we see that

$$P_t = P_a + P_s;$$

that is, P_t gives the total power removed from the beam by scattering and dielectric loss. We must now evaluate

$$P_t = -\tfrac{1}{2}\,\mathrm{Re}\int_0^{2\pi}\int_0^{\pi}(E_\theta^i H_\phi^{s*} - E_\phi^i H_\theta^{s*} + E_\theta^s H_\phi^{i*} - E_\phi^s H_\theta^{i*})\,R^2 \sin\theta\, d\theta\, d\phi. \tag{20}$$

The integration is performed at a very large value of R, and the asymptotic expansions for $z_n(kR)$ are employed:

$$j_n(kR) \approx \frac{1}{kR}\cos\left(kR - \frac{n+1}{2}\pi\right); \tag{21}$$

$$h_n^{(2)}(kR) \approx \frac{i^{n+1}}{kR}\,e^{-ikR}. \tag{22}$$

The integration of Eq. (20) may now be performed, using Eqs. (21), (22), (8), and (9) and certain orthogonality relations of the Legendre functions. This is a lengthy and complicated task and will not be given here. The final result is

$$P_t = \frac{-\pi E_0^2}{\eta_0 k^2}\,\mathrm{Re}\sum_{n=1}^{\infty}(2n+1)(a_n^s + b_n^s).$$

The *total* scattering cross section of the sphere Q_t is the ratio of the total scattered power P_t to the incident Poynting vector $\bar{S}_z^i = E_0^2/2\eta_0$;

$$Q_t = -\frac{2\pi}{k^2}\,\mathrm{Re}\sum_{n=1}^{\infty}(2n+1)(a_n^s + b_n^s). \tag{23}$$

We shall employ Q_t in the following chapters in connection with attenuation by spherical raindrops.

The *radar* cross section may be obtained from the radial component of the Poynting vector of the scattered wave, given by the second term of Eq. (19) and computed at a large distance for the backward direction $\theta = \pi$.

$$\overline{S}_R^s = \tfrac{1}{2} \operatorname{Re} (E_\theta^s H_\phi^{s*} - E_\phi^s H_\theta^{s*}). \tag{24}$$

If Eq. (22) is employed in Eqs. (13) and (14), it is readily found that for very large R,

$$E_\theta^s = \eta_0 H_\phi^s = \frac{-E_0 e^{-ikR}}{ikR} \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left(a_n^s \frac{P_n^1(\cos\theta)}{\sin\theta} + b_n^s \frac{dP_n^1(\cos\theta)}{d\theta} \right) \cos\phi, \tag{25}$$

$$E_\phi^s = -\eta_0 H_\theta^s = \frac{E_0 e^{-ikR}}{ikR} \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left(a_n^s \frac{dP_n^1(\cos\theta)}{d\theta} + b_n^s \frac{P_n^1(\cos\theta)}{\sin\theta} \right) \sin\phi. \tag{26}$$

Then Eq. (24) becomes

$$\overline{S}_R^s = \frac{E_0^2}{2\eta_0 (kR)^2} \operatorname{Re} \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{2n+1}{n(n+1)} \frac{2m+1}{m(m+1)} \left\{ \frac{P_n^1 P_m^1}{\sin^2\theta} \left(a_n^s a_m^{s*} \cos^2\phi + b_n^s b_m^{s*} \sin^2\phi \right) + \frac{dP_n^1}{d\theta} \frac{dP_m^1}{d\theta} \left(a_n^s a_m^{s*} \sin^2\phi + b_n^s b_m^{s*} \cos^2\phi \right) + \frac{P_m^1}{\sin\theta} \frac{dP_n^1}{d\theta} \left(a_m^{s*} b_n^s \cos^2\phi + a_n^s b_m^{s*} \sin^2\phi \right) + \frac{P_n^1}{\sin\theta} \frac{dP_m^1}{d\theta} \left(a_n^s b_m^{s*} \cos^2\phi + a_m^{s*} b_n^s \sin^2\phi \right) \right\}. \tag{27}$$

This equation gives the power scattered in the direction θ, ϕ. For back scattering it must be evaluated for $\theta = \pi$. For this purpose we employ the relations

$$-\left[\frac{P_n^1(\cos\theta)}{\sin\theta} \right]_{\theta=\pi} = \left[\frac{dP_n^1(\cos\theta)}{d\theta} \right]_{\theta=\pi} = (-1)^n \frac{n(n+1)}{2}$$

and after some manipulation obtain

$$S_R^s(\pi) = \frac{E_0^2}{2\eta_0 (kR)^2} \left| \sum_{n=1}^{\infty} (-1)^n \left(n + \frac{1}{2} \right) (a_n^s - b_n^s) \right|^2. \tag{28}$$

Recalling the definition of σ, in the present notation,

$$\sigma = 4\pi R^2 \frac{\overline{S}_R^s(\pi)}{\overline{S}_z^i},$$

where $\overline{S}_z^i$ is the magnitude of the Poynting vector incident on the sphere, $E_0^2/2\eta_0$. It is particularly convenient to employ the ratio of σ to the geometrical cross section of the sphere,

$$\frac{\sigma}{\pi a^2} = \frac{1}{\rho^2}\left|\sum_{n=1}^{\infty}(-1)^n(2n+1)(a_n^s - b_n^s)\right|^2 \tag{29}$$

Although Eq. (29) is simple in appearance, numerical calculations for a wide range of sphere radii and dielectric properties present serious difficulties. We consider at first the case for $\rho \ll 1$. If we expand Eq. (18), discarding powers of ρ higher than the sixth and writing $\varepsilon_c = n_c^2$, we have[1]

$$a_1^s = -\frac{i}{45}(\varepsilon_c - 1)\rho^5, \tag{30a}$$

$$b_1^s = -i\frac{2}{3}\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\rho^3\left(1 + \frac{3}{5}\frac{\varepsilon_c - 2}{\varepsilon_c + 2}\rho^2 - i\frac{2}{3}\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\rho^3\right), \tag{30b}$$

$$b_2^s = -\frac{i}{15}\frac{\varepsilon_c - 1}{2\varepsilon_c + 3}\rho^5. \tag{30c}$$

The physical interpretation of the a_n^s and b_n^s coefficients aids in visualizing the mechanism of the scattering process. It can be shown that b_1^s is the amplitude of a field distribution on the sphere corresponding to a simple electric dipole oriented along the x-axis. Also a_1^s is directly related to a magnetic dipole field distribution, and a_2^s to a magnetic quadrupole. In general, the a_n's and b_n's are proportional to the amplitudes of magnetic and electric multipoles, which are induced by the incident wave.[2]

Returning to Eq. (30), we observe that for sufficiently small ρ and finite ε_c only the first term of Eq. (30b) is important,

$$b_1^s = -i\frac{2}{3}\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\rho^3;$$

that is, the sphere behaves like an electric dipole. Its dipole moment is easily shown to be

$$p = 4\pi\epsilon_0\frac{\varepsilon_c - 1}{\varepsilon_c + 2}a^3E_0.$$

[1] Values for b_1^s and b_2^s are taken from L. Goldstein, in *Radio Wave Propagation*, Academic Press, New York, 1949, p. 271*ff*. Goldstein reports that b_1^s and b_2^s as given by Stratton are incorrect, that Eq. (35), p. 571, should have $2n + 1$ rather than $2n + 2$, and that the right side of Eq. (36) should be negative.

[2] For a detailed discussion of the multipole fields and their magnitudes, see M. Born, *Optik*, Springer, Berlin, 1933, pp. 274*ff*.

Equation (29) yields, for this case,

$$\frac{\sigma}{\pi a^2} = 4\left|\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\right|^2 \rho^4.$$

This is precisely the well-known Rayleigh scattering law (for the backward direction), which occurs in all problems of scattering of electromagnetic radiation by very small dielectric spheres. Two cases that are of some interest in special applications are

$$\text{Finite } |\varepsilon_c| \gg 1, \qquad \frac{\sigma}{\pi a^2} = 4\rho^4 = 6.234\left(\frac{a}{\lambda}\right)^4 \times 10^3,$$

$$|\varepsilon_c| \ll 1, \qquad \frac{\sigma}{\pi a^2} = \rho^4 = 1.559\left(\frac{a}{\lambda}\right)^4 \times 10^3.$$

The scattering coefficients for spheres with infinite ε_c (either infinite dielectric constant or conductivity or both) may best be obtained directly from Eq. (18):

$$a_n^s = -\frac{j_n(\rho)}{h_n^{(2)}(\rho)}, \qquad b_n^s = -\frac{[\rho j_n(\rho)]'}{[\rho h_n^{(2)}(\rho)]'}. \tag{30d}$$

If, as before, these expressions are evaluated for small ρ, we discover that *both* a_1^s and b_1^s now contain terms in ρ^3 but b_2^s begins with ρ^5. If we retain only the ρ^3 terms,

$$a_1^s = -\frac{i}{3}\rho^3, \tag{30e}$$

$$b_1^s = -i\frac{2}{3}\rho^3. \tag{30f}$$

[Equation (30*f*) can be obtained directly from Eq. (30*b*), but this is not so simple for Eqs. (30*e*) and (30*a*)]. We now observe the interesting fact that the magnetic dipole term is half as large as the electric dipole term and in phase with it. The effect is to increase σ by a factor $\frac{9}{4}$ over the value for large but finite ε_c:

$$|\varepsilon_c| = \infty, \qquad \frac{\sigma}{\pi a^2} = 9\rho^4 = 1.403\left(\frac{a}{\lambda}\right)^4 \times 10^4.$$

Figure 6·1 shows the radar cross section for a sphere with infinite conductivity. For small a/λ the Rayleigh law is followed closely,[1] but above about $a/\lambda = 0.1$ the higher-order multipole moments become larger and fluctuate in phase and magnitude in a complicated way, causing the curve to exhibit marked oscillations. The oscillations diminish slowly in amplitude,[2] and σ approaches asymptotically the geometrical cross section as a/λ

[1] Unfortunately the dashed line labeled "Rayleigh Law" is displaced to the right of its correct position. It should follow $\sigma/\pi a^2 = 1.403\ (a/\lambda)^4 \times 10^4$.

[2] Presently available tables permit calculations only up to $\rho = 10$.

becomes very large. It will be shown in Sec. 6·3 that this is just the value to be expected on the basis of geometrical optics.

The analysis of scattering by a sphere has been sketched in sufficient detail to show the principles required for an exact solution. It is clear that fulfillment of boundary conditions equivalent to Eq. (17), in order that coefficients of the scattered waves can be obtained, automatically limits the shape of the target to a handful of the simplest geometrical forms.

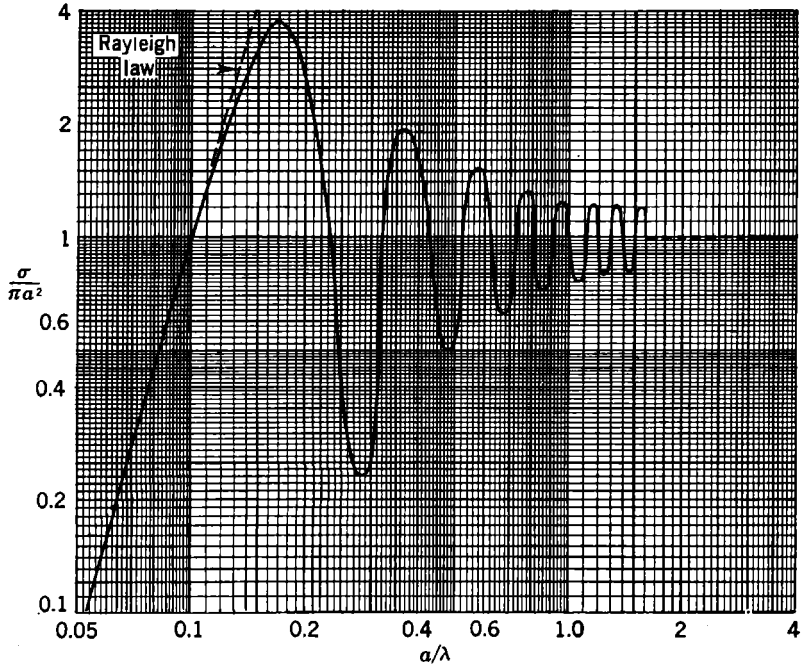


FIG. 6·1.—Ratio of back-scattering cross section to geometrical cross section as a function of radius to wavelength for a sphere with infinite conductivity. The line labeled "Rayleigh Law" is misplaced. See footnote on page 452.

This geometry must correspond to a coordinate system in which the wave equation is separable, and which gives rise to ordinary differential equations with solutions of known properties. These properties must be sufficiently simple that the necessary manipulation of these functions does not require unreasonable effort. As an illustration of how difficult these operations can be with a seemingly simple target, the analysis of scattering from an ellipsoid has been completed only for limited range of shapes because the general solution offers such great practical computational difficulties.[1] The remainder of this chapter will be devoted to approximate methods of calculation of cross section and to measurements of cross section of targets that

[1] F. V. Schultz, "Scattering by a Prolate Spheroid," Aeronautical Research Center, University of Michigan, Mar. 1, 1950.

become progressively so complicated that the methods become essentially empirical in nature.

6·2. Vector Form of Huygens' Principle.—As the rigorous method of the preceding section cannot be applied to most problems, we shall consider a second method that forms the basis for solution of essentially all diffraction problems that can be solved. It is the familiar Kirchhoff-Huygens principle, which states that if the value of a field quantity is known at every point on any closed surface surrounding a source-free region, each elementary unit of surface can be considered as a radiating source, and the total field at any interior point is given by integrating the contributions of all the individual elements over the surface. This principle is generally given

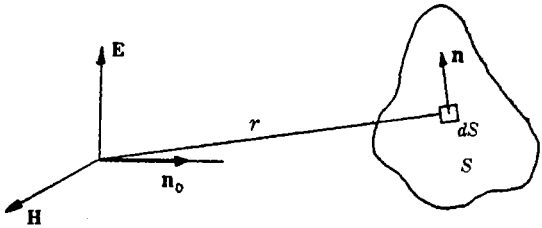


FIG. 6·2.—Plane wave incident upon a surface S of arbitrary shape. The direction of propagation is $\mathbf{n}_0$, and the normal to the surface at any point is $\mathbf{n}$.

in a form appropriate for scalar fields, such as sound. It is frequently applied in this form to vector fields, such as those of electromagnetic waves, despite the fundamental differences between the two kinds of field. Such procedures cannot yield information on polarization, however, and consequently are not suitable for the present purpose. Recently Stratton and Chu have given a particularly elegant and compact statement of the Kirchhoff-Huygens principle for electromagnetic waves,[1] which will be used here to calculate scattering cross sections. (The same method could have been applied to the sphere.)

If a plane wave falls on an object of arbitrary surface S, shown in Fig. 6·2, the object scatters the incident wave, and we may consider the currents and charges within the object as the source of the scattered wave. We shall specialize the present analysis in two ways: (1) Only scattering back along the direction of the incident wave will be calculated and (2) the scattering object will be considered to have infinite conductivity. The second restriction requires all fields inside the object and the tangential component of $\mathbf{E}$ and the normal component of $\mathbf{H}$ on the surface to be zero. Then the scattered field is given by

$$\mathbf{H}^s = -\frac{1}{4\pi}\int_S \left[(\mathbf{n} \times \mathbf{H}_i) \times \nabla \frac{e^{-ikr}}{r}\right] dS, \tag{31}$$

[1] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, Sec. 8·14.

where the surface of integration is the surface of the object and a closed surface at infinity (the integral over this surface is zero), $\mathbf{n}$ is a unit normal to the surface, $\mathbf{H}_t$ is the *total* magnetic field on the surface and is entirely tangential to the surface.[1]

Because we consider only targets at distances large compared with dimensions of the target and only back scattering,

$$\nabla\left(\frac{e^{-ikr}}{r}\right) \approx -\mathbf{n}_0\left(ik\,\frac{e^{-ikr}}{r}\right),$$

where r is the distance from any element of area dS to the point of observation (in this case the source of the wave). The r in the denominator may be replaced by R, the distance to the center of the target. In the exponent, however, we must use better approximations. Then Eq. (31) becomes

$$\mathbf{H}^s = -\frac{ik}{4\pi R}\int_S \mathbf{n}_0 \times (\mathbf{n} \times \mathbf{H}_t)e^{-ikr}\,dS. \tag{32}$$

The term $\mathbf{n} \times \mathbf{H}_t$ can be recognized as the surface current $\mathbf{K}$ on S, which sometimes aids in visualizing the scattering pattern.

As the system is linear, $\mathbf{H}_t$ will be proportional to the incident magnetic field, the magnitude of which is H_0. Then we may write

$$\mathbf{H}_t = \mathbf{i}_t H_0 \zeta(x,y,z), \tag{33}$$

where $\zeta(x,y,z)$ is a complex function describing the variation of field over the surface and $\mathbf{i}_t$ is a vector (not in general of unit length) tangent to the surface and giving the direction of the total surface field.

The ratio of the scattered to the incident Poynting vector is proportional to the square of the ratio of scattered to incident fields; so the scattering cross section becomes

$$\sigma = 4\pi R^2 \frac{S^s}{S^i} = 4\pi R^2 \left|\frac{H^s}{H_0}\right|^2$$

or

$$\sigma = \frac{\pi}{\lambda^2}\left|\int_S \mathbf{n}_0 \times (\mathbf{n} \times \mathbf{i}_t)\,\zeta(x,y,z)e^{-ikr}\,dS\right|^2. \tag{34}$$

Thus far the only approximation has been that for large distances. It has been assumed that $\mathbf{H}_t$ is known, but $\mathbf{H}_t$ will be known exactly only if the boundary value problem can be solved. It was pointed out in the pre-

[1] Equation (31) is Stratton's Eq. (20), *ibid.*, p. 466, adapted to the present case. It might be mentioned here that a method of approach more frequently used involves calculation of the surface currents, the resulting vector potential, and then the field from the vector potential. This method is equivalent to the one to be given here but involves more mathematical manipulation without increasing the clarity of the procedure in the present case. The derivation of Eq. (31) requires that the fields and their first derivatives be continuous on S, but we shall apply it to cases in which this condition is not strictly fulfilled. Because we consider only objects larger than the wavelength, the necessary correction terms are small and will not be given here.

ceding section that exact solutions can be obtained in only a very few cases. We shall next consider applications of Eq. (34) to cases in which it is possible to obtain approximate formulas for $\mathbf{H}_t$ that will allow Eq. (34) to be integrated. The approximation to be used is stated formally in Eqs. (53) and (54); physically, it is that $\mathbf{H}_t$ is twice the tangential component of the magnetic field of the incident wave.

6·3. Scattering from Planes and Curved Surfaces. *Large Planes.*—A plane linearly polarized wave is incident upon a plane surface, as shown in

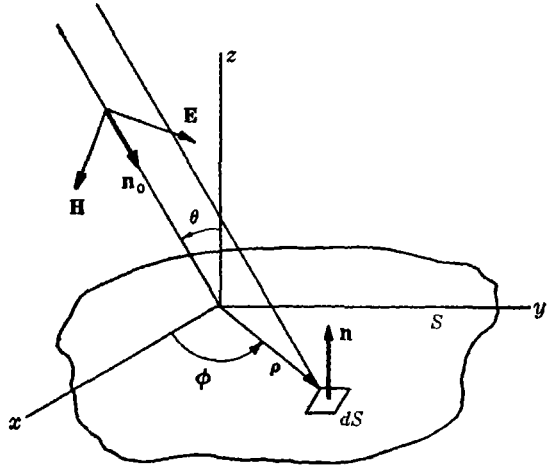


FIG. 6·3.—Plane linearly polarized electromagnetic wave incident upon a plane surface S.

Fig. 6·3. The plane of incidence is the xz-plane, and the incident magnetic field is given by

$$\mathbf{H}^i = \mathbf{a} H_0 e^{-ik(R - x\sin\theta)}, \tag{35}$$

where $\mathbf{a}$ is a unit vector in a plane perpendicular to $\mathbf{n}_0$. If α is the azimuth of $\mathbf{a}$ in this plane, measured from the plane of incidence,

$$\mathbf{a} = \mathbf{i}_x \cos\alpha \cos\theta + \mathbf{i}_y \sin\alpha + \mathbf{i}_z \cos\alpha \sin\theta.$$

In order to determine $\mathbf{H}_t$ we now introduce the assumption that the surface S is so large in terms of wavelength that $\mathbf{H}_t$ has the same value that it would have if the surface were infinite in extent. On the back side the field is then zero, and on the front side $\mathbf{H}_t$ is twice the tangential component of the incident wave:[1]

[1] One occasionally encounters the incautious statement that the surface magnetic field is twice the incident field. Clearly this is true only if the incident magnetic vector is perpendicular to the plane of incidence and thus $\alpha = \pi/2$. The correct statement of the boundary condition is given in more general terms in Eqs. (53) and (54) later in this section.

$$\mathbf{H}_t = \mathbf{i}_t\, 2H_0\, e^{-ik(R - x\sin\theta)}, \tag{36}$$

where

$$\mathbf{i}_t = \mathbf{i}_x \cos\alpha\cos\theta + \mathbf{i}_y \sin\alpha.$$

One then finds without difficulty that

$$\mathbf{n}_0 \times (\mathbf{n} \times \mathbf{i}_t) = \mathbf{a}\cos\theta.$$

Equation (32) then becomes

$$\mathbf{H}^s = -\mathbf{a}\,\frac{ikH_0\cos\theta}{2\pi R}\, e^{-i2kR} \int_S e^{-i2kx\sin\theta}\, dS. \tag{37}$$

This equation shows, as might be expected, that maximum back scattering occurs for normal incidence ($\theta = 0$) and that there is no depolarization of the scattered wave. The last conclusion is not strictly true because the surface fields near the edges are not actually as we have assumed them; but as long as the smallest dimension of the surface is several wavelengths long, the present procedure can be expected to give satisfactory accuracy.[1]

The radar cross section σ may now be found from Eqs. (33), (34), and (36).

$$\sigma = \frac{4\pi}{\lambda^2}\cos^2\theta \left| \int_S e^{-i2kx\sin\theta}\, dS \right|^2, \tag{38}$$

where the integration is performed over the front, or illuminated surface. For *normal incidence* Eq. (38) reduces to the important formula

$$\sigma_{\max} = \frac{4\pi A^2}{\lambda^2}, \tag{39}$$

where A is the area of the surface, which may have any shape as long as it is not so complicated that the initial assumptions regarding the surface fields are seriously in error.

If the surface is rectangular, with total dimensions of a and b parallel to the x- and y-axes, respectively, Eq. (38) becomes

$$\sigma(\theta) = \frac{4\pi A^2}{\lambda^2}\left[\frac{\sin(ka\sin\theta)}{ka\sin\theta}\right]^2 \cos^2\theta. \tag{40}$$

If the incident wave had come from the direction θ, ϕ, where ϕ is the azimuth angle measured from the x-axis, an analysis similar to that above would show that

$$\sigma(\theta,\phi) = \frac{4\pi A^2}{\lambda^2}\left[\frac{\sin(ka\sin\theta\cos\phi)}{ka\sin\theta\cos\phi}\;\frac{\sin(kb\sin\theta\sin\phi)}{kb\sin\theta\sin\phi}\right]^2 \cos^2\theta. \tag{41}$$

[1] For scattering from objects such as flat strips or ribbons having widths comparable to the wavelength see P. M. Morse and P. J. Rubenstein, *Phys. Rev.*, **54**, 895 (1938).

The factors in Eqs. (40) and (41) having the form $(\sin x)/x$ occur in all diffraction problems in which a uniform field distribution occurs across a slit or rectangular area. There is one difference between the expressions above and those ordinarily encountered, however; in Eq. (41), for instance, we have $ka \sin\theta \cos\phi$, whereas in cases in which the primary source of radiation is in the surface S one finds $\frac{1}{2}\, ka \sin\theta \cos\phi$. The difference occurs, of course, because in the radar case the phase differences governing the interference between the elementary current elements are double those for the one-way case. The result is that the main lobe of the diffraction pattern of a given surface acting as a radar target is only approximately half as wide as that from an obstacle or opening of the same dimensions scattering in the forward direction.

Comparison of Eqs. (39) and (40) or (41) shows that at normal incidence σ is inversely proportional to λ^2, whereas for any appreciable value of θ, σ is a rapidly oscillating function of λ and approaches zero for very small λ. This is to be expected; for as ka becomes large, most of the incident energy experiences specular reflection and little is scattered backward. The diffraction pattern was indicated qualitatively in Fig. 1·3.

If the surface S is a circular disk the integration in Eq. (34) must be performed in polar coordinates ρ, ϕ in the xy-plane. Then

$$2kx \sin\theta = 2k\rho \cos\phi \sin\theta \text{ and } dS = \rho\, d\rho\, d\phi.$$

Using well-known integral definitions of Bessel functions[1] Eq. (38) becomes

$$\sigma(\theta) = \frac{4\pi A^2}{\lambda^2}\left[2\,\frac{J_1(2ka\sin\theta)}{2ka\sin\theta}\right]^2 \cos^2\theta, \tag{42}$$

where J_1 designates the first-order Bessel function and a is the radius of the disk. The factor in square brackets has the same general oscillatory nature as the corresponding $(\sin x)/x$ terms in Eq. (40). Its first three zeros are at 3.83, 7.02, and 10.17, and it is unity for $\theta = 0$. The factor 2 is included within the brackets because $J_1(x)/x = \frac{1}{2}$ when $x = 0$.

Cylinders. Scattering from cylinders may also be calculated with the same kind of approximation that was used for plane surfaces. Two cases will be considered first: that of longitudinal polarization, in which the cylinder is oriented with its axis parallel to the electric field of the incident wave, and that of transverse polarization, in which the cylinder is perpendicular to the incident electric field and to the direction of propagation. The procedure will be to find the surface fields from a solution of the boundary value problem on the assumption that the fields on a finite length of cylinder do not differ from those on an infinitely conducting cylinder of infinite length and then to calculate the cross section by Eq. (40).

For the case of longitudinal polarization let the cylinder be oriented

[1] Jahnke and Emde, *Tables of Functions*, Dover reprint, 1943, p. 149.

along the z-axis of the coordinate system, as shown in Fig. 6·4. The incident wave traveling along the x-axis is obtained from the expansion of a plane wave in cylindrical coordinates:

$$\mathbf{E}^i = \mathbf{i}_z E_0 e^{-ikx} = \mathbf{i}_z E_0 \sum_{n=0}^{\infty} (2 - \delta_{0n})(-i)^n \cos n\phi J_n(kr)$$

$$\mathbf{H}^i = \frac{-1}{i\omega\mu_0} \nabla \times \mathbf{E}^i = \frac{E_0}{ik\eta_0} \sum_{n=0}^{\infty} (2 - \delta_{0n})(-i)^n \left[\mathbf{i}_r \frac{n}{r} \sin n\phi J_n(kr) + \mathbf{i}_\phi k \cos n\phi J_n'(kr) \right],$$

where δ_{0n} is the Kronecker delta, equal to unity for $n = 0$ and zero for all other values of n. If the cylinder is infinite in length, the scattered wave

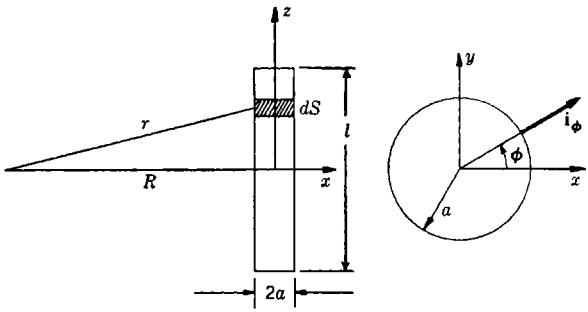


FIG. 6·4.—Geometry for calculation of scattering from a cylinder.

will be given by a similar series, except that in order to represent outward traveling waves for large r the Bessel functions of the first kind must be replaced by the Hankel functions $H_n^{(2)}(kr)$, each with a coefficient a_n, which must be determined from the boundary conditions. Recalling that the magnitude of the incident magnetic wave is $H_0 = E_0/\eta_0$, we may now write the total field as the sum of the incident and scattered fields:

$$\mathbf{H} = -iH_0 \sum_{n=0}^{\infty} (2 - \delta_{0n})(-i)^n \left\{ \mathbf{i}_r \frac{n}{kr} \sin n\phi \left[J_n(kr) + a_n H_n^{(2)}(kr) \right] + \mathbf{i}_\phi \cos n\phi \left[J_n'(kr) + a_n H_n^{(2)\prime}(kr) \right] \right\}. \quad (43)$$

At the cylinder surface, where $r = a$, the radial component of the magnetic field is zero, which requires that

$$a_n = -\frac{J_n(\rho)}{H_n^{(2)}(\rho)}, \qquad \rho = ka. \quad (44)$$

Referring to Eq. (33), we see that $\mathbf{H}_t$ is obtained from Eqs. (43) and (44) by setting $r = a$. We then have the longitudinal surface current $\mathbf{K}$, given by

$$\mathbf{i}_r \times \mathbf{i}_t H_0 \zeta(x,y,z) =$$

$$- \mathbf{i}_z i H_0 \sum_{n=0}^{\infty} (2 - \delta_{0n})(-i)^n \cos n\phi \left[\frac{J_n'(\rho)H_n^{(2)}(\rho) - J_n(\rho)H_n^{(2)\prime}(\rho)}{H_n^{(2)}(\rho)} \right]. \quad (45)$$

The numerator of the expression in brackets is the Wronskian of J_n and $H_n^{(2)}$, given by $2i/\pi\rho$. The vector product of Eq. (45) with $\mathbf{n}_0 = \mathbf{i}_x$ changes the direction to that of the y-axis, and Eq. (34) becomes

$$\sigma = \frac{\pi}{\lambda^2} \left| \frac{2}{\pi\rho} \sum_{n=0}^{\infty} \frac{(2 - \delta_{0n})(-i)^n}{H_n^{(2)}(\rho)} \int_S e^{-ikr} \cos n\phi \, dS \right|^2. \quad (46)$$

If R is the distance from the point of observation to the center of the cylinder, $r \approx R + a \cos \phi$. The surface element $dS = a \, d\phi \, dz$, and the integration is performed over the angular range from zero to 2π and over the length l. The ends are omitted, as the currents and fields are assumed to be zero there.

Integration of Eq. (46) requires the integral representation of Bessel functions mentioned earlier.[1] After integration Eq. (46) becomes

$$\sigma_l = \frac{4l^2}{\pi} \left| \sum_{n=0}^{\infty} (-1)^n (2 - \delta_{0n}) \frac{J_n(\rho)}{H_n^{(2)}(\rho)} \right|^2, \quad (47)$$

which can be written in a form more convenient for numerical computations as

$$\sigma_l = \frac{4l^2}{\pi} \left| \sum_{n=0}^{\infty} (-1)^n (2 - \delta_{0n}) e^{i\delta_n} \sin \delta_n \right|^2, \quad (47a)$$

where

$$\tan \delta_n = - \frac{J_n(\rho)}{N_n(\rho)}.$$

Analysis of the transverse-polarization case proceeds in similar fashion and leads to

$$\sigma_t = \frac{4l^2}{\pi} \left| \sum_{n=0}^{\infty} (-1)^n (2 - \delta_{0n}) \frac{J_n'(\rho)}{H_n^{(2)\prime}(\rho)} \right|^2, \quad (48)$$

[1] *Ibid.*

[2] The angles δ_n (and the δ_n' for the case of transverse polarization) have been tabulated by P. M. Morse *et al.*, in "Scattering and Radiation from Circular Cylinders and Spheres," MIT Underwater Sound Laboratory, Applied Mathematics Panel Report No. 62.1R, February, 1945.

which can be rewritten in a manner similar to Eq. (47*a*) as

$$\sigma_t \approx \frac{4l^2}{\pi}\left|\sum_{n=0}^{\infty}(-1)^n(2-\delta_{0n})e^{i\delta_n'}\sin\delta_n'\right|^2, \tag{48a}$$

where

$$\tan\delta_n' = -\frac{J_n'(\rho)}{N_n'(\rho)},$$

the primes denoting differentiation with respect to the argument.

The dependence of σ on cylinder radius is important. For very small ρ, Eqs. (47) and (48) become

$$\sigma_t \approx \pi l^2\left(\frac{1}{\ln\rho}\right)^2, \tag{49}$$

$$\sigma_t \approx \frac{9\pi}{4}l^2\rho^4. \tag{50}$$

The fourth-power law in σ_t is reminiscent of the Rayleigh scattering law for the sphere discussed in Sec. 6·1; in fact, closer scrutiny reveals that there is a marked analogy between scattering by spheres and cylinders. The coefficients a_n^s and b_n^s of Eq. (18) were identified with multipole moments from the surface currents induced by the wave incident on the sphere. The a_n of Eq. (44) is analogous to a_n^s representing magnetic multipoles for the sphere and similarly the analysis of the transverse-polarization case for the cylinder involves a scattering coefficient b_n representing electric multipoles, which is analogous to a_n^s for the sphere. In the spherical case both a_n^s and b_n^s varied with ρ in the same manner for very small ρ but this is not true for the cylinder, with the result that the cross section shows a marked dependence on polarization. Eq. (49) shows that σ_l is considerably larger than σ_t for ρ less than about 0.2. For larger values of ρ the complete expressions must be employed. They show that σ_l decreases monotonically with increasing ρ but that σ_t departs markedly from the fourth-power law and oscillates in a manner very similar to that of the curve in Fig. 6 1. The square of the series in Eqs. (47) and (48) approaches the limiting value $\pi\rho/4$ for ρ greater than about 6, and both cross sections approach the limiting value

$$\sigma_t \approx \sigma_l \xrightarrow[\rho\to\infty]{} \rho l^2 = \frac{2\pi a l^2}{\lambda}, \tag{51}$$

independent of polarization. This is also the value that would be predicted by geometrical optics. This point will be discussed further later in the section.

Equations (47) and (48) were derived for the direction of propagation perpendicular to the axis of the cylinder. If the incident electric field lies in the xz-plane of Fig. 6·4 with the direction of propagation $\mathbf{n}_0$ forming an

angle β with the x-axis, the induced currents along the cylinder cause a diffraction pattern similar to that from the plane reflector, and the effective value of ρ is multiplied by the factor $\cos \beta$. Consequently for this case the cross section becomes

$$\sigma = \sigma_1 \left[\frac{\sin (kl \sin \beta)}{kl \sin \beta} \right]^2 \tag{52}$$

where σ_1 is obtained from Eq. (47) after replacing ρ by $\rho \cos \beta$. For very small $\rho \cos \beta$, σ_1 is given by Eq. (49) with the same substitution; and for large $\rho \cos \beta$, σ_1 is obtained by using Eq. (51) in the same way.

The degree of approximation involved in neglecting the effect of finite length of the cylinder cannot be evaluated directly, but some useful qualitative statements can be made. If ρ is small, the assumption of constant-current distribution is likely to be seriously in error because of resonance effects and standing waves unless the wire is very long. If the diameter of the cylinder is roughly $\lambda/4$ or more, however, and if the cylinder is several wavelengths long, the current is probably near the predicted value a fraction of a wavelength from the end. The correct current distribution, if it were known, would produce a dependence on l different from the factors l^2 and $\left[\frac{\sin (kl \cos \beta)}{kl \cos \beta} \right]^2$ occurring in the preceding analysis.

An important conclusion to be drawn from the difference in reflecting properties of a thin cylinder for the two polarizations is that a wave polarized linearly at an angle with the cylinder axis will be scattered with elliptical polarization. [The difference in amplitudes of the two components of the scattered wave is indicated by Eqs. (49) and (50), but the differences in phase must be obtained by performing the integration of Eq. (32).] If the radius of the cylinder is greater than about one wavelength, the depolarizing effect essentially disappears. This fact forms the basis for some important relations that will now be developed to extend the methods given above to other cases of practical interest and to illustrate the relation between these results and those of geometrical optics.

Simplified Formula for Cross Section.—If a plane wave is incident upon a perfectly conducting surface having undulations or convolutions with radii of curvature greater than about one wavelength, the induced currents and the fields radiated from any infinitesimal unit of area are very nearly those which would be obtained from the same area if it were part of an infinite plane, tangent to the surface at the location of the element of area. The currents and fields on the surface are determined by the boundary condition that the surface magnetic field is entirely tangential and is twice the tangential component of the magnetic field of the incident wave. For convenience we assume the incident wave traveling along the z-axis, with its magnetic vector of amplitude H_0 in the

direction **a**, perpendicular to the z-axis. On the surface the magnetic field is given by

$$\mathbf{H}_t = \mathbf{i}_t 2H_0 e^{-ikz}, \tag{53}$$

where $\mathbf{i}_t$ is given by

$$\mathbf{i}_t = \mathbf{a} - (\mathbf{a} \cdot \mathbf{n})\mathbf{n} \tag{54}$$

and $\mathbf{n}$ is the outward normal to the surface.[1]

To calculate the scattered field we must evaluate the triple vector product of Eq. (32), using Eq. (54). Using standard vector identities we find

$$\mathbf{n}_0 \times (\mathbf{n} \times \mathbf{i}_t) = -\mathbf{a}2(\mathbf{n}_0 \cdot \mathbf{n}) \cdot$$

If dS is the element of area on the scattering surface, the projection of dS onto the xy-plane will be denoted by dA. The relation between these areas is

$$dA = -(\mathbf{n}_0 \cdot \mathbf{n})\, dS$$

The scattered field is, then, from Eq. (32),

$$\mathbf{H}^s = \mathbf{a}\frac{ikH_0}{2\pi R}\int_A e^{-i2kR}\, dA, \tag{55}$$

and the cross section given by Eq. (34) reduces to

$$\sigma = \frac{4\pi}{\lambda^2}\left|\int_A e^{-i2kz}\, dA\right|^2, \tag{56}$$

where the integration on dA is now to be performed in the xy-plane, using the known shape of the surface to express z as a function of the coordinates in the xy-plane.

It is convenient to employ polar coordinates ρ, ϕ in the xy-plane. Then

$$dA = \rho\, d\rho\, d\phi = \frac{d\rho}{dz}\, dz\, d\phi,$$

and Eq. (56) becomes

$$\sigma = \frac{4\pi}{\lambda^2}\left|\int_0^{2\pi}\int_0^{z(\phi)} \frac{d\rho}{dz}\, e^{-i2kz}\, dz\, d\phi\right|^2. \tag{57}$$

Scattering by Quadric Surfaces.—Integration of Eq. (57) is not difficult for quadric surfaces but becomes cumbersome for other shapes. We shall consider an elliptic paraboloid and an ellipsoid in order to obtain a scattering formula of some practical importance and in order to illustrate the relative importance of scattering from the "nose" and from the remaining parts of the surface.

[1] Equation (54) can be derived from the fact that the tangential component of **a** is obtained by subtracting from **a** its component normal to the surface.

From analytic geometry, the equation for an elliptic paraboloid opening outward along the positive z-axis is

$$\frac{z}{c} = \frac{x^2}{b_1^2} + \frac{y^2}{b_2^2}, \tag{58}$$

where b_1, b_2, and c are constants. The quantities of interest in the present case are the radii of curvature of the surface in the xz- and yz-planes at the origin. These radii are easily shown to be

$$a_1 = \frac{b_1^2}{2c}, \qquad a_2 = \frac{b_2^2}{2c}.$$

Transforming to polar coordinates ρ, ϕ in the xy-plane, Eq. (58) becomes[1]

$$z = \frac{\rho^2}{2a_1}\left[1 - \left(1 - \frac{a_1}{a_2}\right)\sin^2\phi\right]. \tag{59}$$

If Eq. (59) is employed in Eq. (57) and the integration on z is carried out to a height h, independent of ϕ,

$$\sigma = \pi a_1 a_2 |e^{-i2kh} - 1|^2 = 4\pi a_1 a_2 \sin^2 kh. \tag{60}$$

The -1 of the complex expression represents the contribution of the tip of the paraboloid, whereas the exponential expresses the interference between successive Fresnel zones on the paraboloid.[2] If the height of the paraboloid were not well defined, that is, if the top edge were cut very irregularly, the contributions of the last zone (or zones) would tend to cancel, causing the exponential term to disappear, then

$$\sigma = \pi a_1 a_2. \tag{61}$$

If the scattering surface is a closed surface, such as an ellipsoid, the contribution of the last Fresnel zone disappears, leaving Eq. (61). This point may be illustrated briefly by the same method used for the paraboloid. The equation for an ellipsoid with semiaxis h_0 along the axis and with radii of curvature a_1 and a_2 at the lower tip is, again in polar coordinates,

$$z^2 = h_0^2 - \rho^2 \frac{h_0}{a_1}\left[1 - \left(1 - \frac{a_1}{a_2}\right)\sin^2\phi\right].$$

Substituting into Eq. (57) and again integrating to a height $h < h_0$,

$$\sigma = \pi a_1 a_2 \left[\left(\frac{h}{h_0}\right)^2 + 4\left(1 - \frac{h}{h_0}\right)\sin^2 kh\right], \tag{62}$$

where terms in $(2kh_0)^{-1}$ have been dropped. In this case the successive Fresnel zones again cause σ to depend upon the height of the section of

[1] For $a_1 = a_2$ the surface is a paraboloid of revolution.

[2] Each zone corresponds to an increment $\Delta h = \lambda/4$ rather than the $\lambda/2$ used in optics because of the doubled path length in the present case.

the ellipsoid, but in a rather different manner from that found for the preceding case. If $h = h_0$, or if h is very irregular around the periphery, the argument used above shows that again $\sigma = \pi a_1 a_2$. In the important special case of a sphere $a_1 = a_2$ and $\sigma = \pi a^2$, in agreement with the results of Sec. 6·1 for a sphere with large a/λ.

In the present approximation the back half of the surface is omitted from the integration of Eq. (57), as the fields are assumed to be zero there. The cross section of the sphere is, of course, independent of aspect, but this is not true for the other surfaces. The scattering surface has been oriented tip foremost with its symmetry axis along the direction of propagation of the incident wave. We have seen that if the back edge is sufficiently irregular, or if the surface is closed,[1] the scattering can be accounted for by the shape of the tip. The proof has been given for a surface that can be represented by an equation of the second degree. Surfaces of different shapes require other methods of integration of Eq. (57) than that given here.

Small Objects. — Before proceeding to geometrical optics we quote without proof some useful values for σ for targets that are comparable in dimensions to the wavelength.

For a thin perfectly conducting prolate spheroid, $\lambda/2$ long (essentially a "half-wave" dipole), the maximum value of σ is approximately $0.88\lambda^2$, and the average value over all orientations and polarizations is approximately $0.11\lambda^2$.

A small perfectly conducting disk shows markedly a phenomenon discussed in Sec. 6·1 in connection with the sphere; the induced surface currents give rise to both electric and magnetic dipole moments of comparable magnitude, and the disk acts like both a straight wire and a loop radiator.[2] As a consequence the maximum echo is not obtained when the disk is oriented normal to the incident radiation, but rather when both the incident electric vector and the direction of incidence are parallel to the surface of the disk. The cross section at normal incidence and the average cross section over all orientations and directions of polarization are

$$\frac{\sigma}{\pi a^2} = \frac{64}{9\pi^2}\left(\frac{2\pi a}{\lambda}\right)^4 = 1125\left(\frac{a}{\lambda}\right)^4, \qquad \frac{2\pi a}{\lambda} < 1.$$

The fourth-power dependence is again an illustration of the Rayleigh scattering law. It is interesting to note that the numerical coefficient for the sphere is about 12.5 times that for the disk.

Geometrical Optics. — Thus far the analysis of scattering has been carried out in terms of solutions of the wave equation, in which the ratio

[1] More correctly, for the purpose of the present approximation the surface needs only to be "half closed" in the forward direction, as only the illuminated portion is considered.

[2] H. A. Bethe, "Theory of Diffraction by Small Holes," RL Report No. 128, Jan. 23, 1942; *Phys. Rev.*, **66**, 163 (1944).

of target dimensions to wavelength plays a prominent role. In the limit as the wavelength approaches zero, however, it is possible to make further approximations that lead to geometrical optics, where polarization and phases are neglected and energy flow follows straight-line ray paths determined from the target and antenna geometry (in a homogeneous medium).

It is convenient to generalize temporarily the back-scattering cross section used up to this point by introducing the *differential* scattering cross section σ', defined for arbitrary scattering angles θ, ϕ, where θ is the acute angle between the surface normal and the direction of propagation of the incident wave (assumed to be the z-axis) and ϕ is the azimuth measured in the xy-plane.[1] We denote σ' as the cross section of an isotropic scatterer that would scatter the same amount of power in any given direction as the given target. This definition is expressed by

$$dP^s = \overline{S}^i \sigma' \frac{d\Omega}{4\pi}, \tag{63}$$

where dP^s is the power scattered into the solid angle $d\Omega$ in the direction (θ,ϕ) and $\overline{S}^i$ is the magnitude of the incident Poynting vector. When $\theta = 0$, σ' is the back-scattering cross section σ used elsewhere in this volume.

At the reflection point the angle of incidence is equal to the angle of reflection. The angle between the direction of incident and reflected waves is then 2θ. If we denote the solid angle enclosing the normal to the surface by $d\Omega_0$,

$$d\Omega_0 = \sin\theta\, d\theta\, d\phi.$$

Similarly, the solid angle enclosing the direction of reflection is $d\Omega$;

$$d\Omega = \sin 2\theta\, d(2\theta) d\phi = 4\cos\theta\, d\Omega_0.$$

But the scattered power is

$$dP^s = \overline{S}^i \cos\theta\, dS,$$

where dS is an element of area on the target and is not to be confused with the Poynting vector. Combining the foregoing expressions in Eq. (63), we have

$$\sigma' = \pi \frac{dS}{d\Omega_0}, \tag{64}$$

which is independent of θ and is a function only of the shape of the surface at any point. As σ' is independent of θ, we now drop the prime. The power scattered per unit solid angle is

$$\frac{dP^s}{d\Omega} = \frac{S^i}{4}\frac{dS}{d\Omega_0}. \tag{65}$$

[1] We use here the method of R. C. Spencer, "Reflections from Smooth Curved Surfaces," RL Report No. 661, Jan. 26, 1945.

Application of Eq. (64) to the case of a second-degree surface with principal radii of curvature a_1 and a_2 gives $dS = a_1 a_2 \, d\Omega_0$, and Eq. (64) reduces to Eq. (61);

$$\sigma = \pi a_1 a_2. \tag{61}$$

Marcus has analyzed this case in greater detail[1] and reports that for a radar antenna of circular shape the correct result is

$$\sigma = \frac{\pi}{2} \sqrt{a_1 a_2} \, (a_1 + a_2). \tag{66}$$

Neither Eq. (61) nor (66) is valid if $a_2/a_1 \gg 1$ or for flat or cylindrical surfaces.[2]

The mean value of σ averaged over all orientations of a smooth convex surface may be obtained from Eq. (64).

$$\bar{\sigma} = \frac{\pi \int \frac{dS}{d\Omega_0} \, d\Omega_0}{\int d\Omega_0} = \frac{\pi S}{4\pi} = \frac{S}{4};$$

thus, σ is one-fourth of the total surface area. (The simplest example is a sphere.) The requirement for convexity of the surface can perhaps be well illustrated by the extreme case of the corner reflector, which gives an enormously greater value of σ than $S/4$.

For a flat plate normal to the beam and large enough that the reflected cone of rays fills the entire receiving antenna area the effective cross section may be shown from Eq. (64) to be

$$\sigma = \pi R^2, \tag{67}$$

where R, as customary, is the distance from radar to target. This odd-appearing result causes the echo power in the radar transmission equation to vary as R^{-2} rather than as R^{-4}; this is the usual result when a target is very large instead of appearing essentially as a point as viewed from the radar.

The absence of wavelength in Eq. (61) to (66) is characteristic of geometrical optics. These formulas are the limits approached by those of physical optics as the wavelength approaches zero. Few actual radar targets are so smooth or symmetrical that these formulas can be applied to a large fraction of the target. If, however, the formulas are used with

[1] P. M. Marcus, "Reflection of Radiation from Curved Surfaces," RL Report No. 1029, Jan. 16, 1946.

[2] For these formulas to be valid the target must have sufficient curvature that at least the first Fresnel zone is included on the surface. For a sphere of radius a this requires the diameter d of the spherical segment to be

$$d \geqslant \sqrt{2a\lambda}.$$

some care (in particular, with due regard for Fresnel zone effects), they can often explain the major characteristics of echoes from prominent structural features of targets such as ships, aircraft, water towers, and similar objects.[1]

Polarization.—All of the discussion and formulas up to this point have assumed the incident radiation to be a linearly polarized plane wave, and the definition of σ has referred to the echo energy carried by that part of the scattered field which is polarized in the same direction as the field incident on the target. Most radar targets produce some depolarization of the scattered wave; and if the receiving antenna system is arranged to receive other types of polarization than that transmitted, the received echo may vary over a wide range as the characteristics of the receiving antenna are varied. Some very interesting preliminary experiments indicate that targets such as wooded hills or houses produce a large degree of depolarization of the radar echo whereas ships or water towers produce very much less depolarization.[2] Insufficient data are given to draw any general conclusions, but the experiments are highly thought-provoking and suggest numerous other investigations in a field that has received practically no attention. Certainly the idealizations employed here of infinite conductivity and simple geometrical shapes do not begin to indicate the possibilities of interesting (and, we hope, useful) properties of radar echoes, and further research along these lines is urged.

One further result is worthy of discussion. If the outgoing wave from the radar antenna is elliptically rather than linearly polarized, the effective back-scattering cross section of the target depends upon the degree of ellipticity of the incident radiation as well as upon the target geometry and wavelength. In particular, if the radiated wave is circularly polarized, and if the target produces no depolarization of the back-scattered wave, its effective cross section is zero![3] This peculiar result is demonstrated in Appendix A, along with an application of the powerful Lorentz reciprocity theorem to radar scattering.

An interesting application of this property of the back scattering of circularly polarized waves was suggested by E. M. Purcell to reduce the intensity of echoes from rain.[4] As individual raindrops do not de-

[1] For further discussion of limitations and applications of the formulas from geometrical optics see J. F. Carlson and S. A. Goudsmit, "Microwave Radar Reflections," RL Report No. 195, Feb. 20, 1943; S. A. Goudsmit and P. R. Weiss, "Possible Measurement of Radar Echoes by Use of Model Targets," RL Report No. 196, Mar. 4, 1943; also R. C. Spencer, *op. cit.*, and P. M. Marcus, *op. cit.*

[2] M. Kessler, C. E. Mandeville, and E. L. Hudspeth, "The Depolarization of Microwaves," RL Report No. 458, Nov. 1, 1943.

[3] This assumes the usual radar circuit arrangement in which the echo is measured in the transmission line (or waveguide) along which the radiated power flows to the antenna system.

[4] See Sec. 3·10 of Vol. 1 of the Radiation Laboratory Series.

polarize the back-scattered wave, and as multiple scattering is too small to be important, a rainstorm fulfills the necessary conditions on the target. Tests revealed a reduction of rain echo of about 25 db. No effort was made to fabricate the antenna with great accuracy, and it seems not unlikely that further reduction could be obtained if it were desired. This scheme has the disadvantage that echoes from most other targets suffer some attenuation also. For some purposes, however, this fact might be of secondary importance. It appears that further generalization of the Lorentz reciprocity theorem given in Appendix A, coupled with further antenna developments, could lead to useful applications of the polarization properties of radar echoes.

COMPLEX TARGETS

By Donald E. Kerr

Thus far we have considered only targets composed of single kinds of surfaces, isolated in free space. It was shown that subject to certain simplifying assumptions σ could be determined uniquely in terms of target geometry, orientation, and wavelength. Such highly idealized situations occur rarely in practice, however. Most targets are complicated combinations of surfaces of such irregular shape that the scattering pattern is a maze of lobes which may or may not have a well-defined peak in one general direction.

An airplane is an example of a combination of fairly simple smooth surfaces from which portions of characteristic scattering patterns can sometimes be identified but which is still too complicated to permit accurate analysis. A ship presents an even more difficult problem, not only because of its geometry but because of the nonuniform illumination of its area resulting from the interference pattern formed by reflections from the sea surface. A hillside with trees moved by the wind presents a still different problem as the respective scattering elements move independently of each other. The echoes from such targets usually fluctuate rapidly over a wide intensity range. Because measurement of such echoes is very difficult, the concept of radar cross section encounters difficulties from the practical standpoint.

We should distinguish between two aspects of the problem, however. It is still true that for every possible type of incident radiation and configuration of the elements of a complex target a unique scattering pattern exists, and from this pattern an effective value of σ could in principle be determined uniquely for any particular direction. When the target (or some of its parts) is in motion, however, σ varies so rapidly and over such a wide range that the final outcome of any measurement depends markedly on the observer and his instruments. In the present state of the radar art, the latter factors play a prominent role in determining the available data on σ for aircraft and to a lesser degree for ships.

Because of the need for some means of describing complex targets certain desperate artifices have been evolved to preserve the concept of cross section for such targets. When used with proper precautions the results so obtained serve a useful purpose, although to be used effectively they must be supplemented by further detailed information on the nature of the scattering. The remainder of this chapter summarizes both types of information.

6·4. Radar Cross Section of Aircraft.—Because of the multilobed scattering pattern of aircraft, definitions of cross section have meaning only in relation to the specific method employed to observe the echo. During the war many measurements of cross section over a wide wavelength range were made, but unfortunately almost none were made with sufficient accuracy or uniformity of measurement technique that they can be considered meaningful.[1] Probably the most definitive approach (at least in the wavelength range of 10 cm and below) is that of Linford and his collaborators.[2] Linford defines the cross section as that value of σ obtained from the measured quantities and the radar equation

$$\sigma = \frac{(4\pi)^3 R^4}{G^2 \lambda^2} \frac{P_r}{P_t},$$

which is exceeded during one-half of a series of measuring-time intervals. In this way the essential feature of the *probability* of detecting the echo is introduced into the definition of σ. (Whether or not this degree of probability is satisfactory for field operations depends upon the specific problem; if necessary the required degree of probability could be modified, and the resulting value of σ would be changed accordingly.)

In his measurements, Linford employed a measuring-time interval of 5 sec, and the echo was considered as "seen" if the echo was distinguishable in the noise on an A-scope at least once during the interval. The strongest echo occurring during the interval was measured by a signal generator and was plotted as a function of range as shown in Fig. 6·5. The line is drawn through the points in such a way that half the measured points are above and half are below it (except for the closest ranges, where data were diffi-

[1] Disagreement between different observers' results by a factor of 10 in cross section on a given type of aircraft have occurred not infrequently. This factor corresponds to a difference in signal level of 10 db, which at first glance seems very large. Further examination reveals, however, that there are numerous understandable reasons for disagreement, such as errors in determination of antenna gains, r-f system losses, absolute power levels, ground-reflection effects, and method and duration of echo display and observation and differences in criteria as to what type of averaging or peak measurements of echo amplitude are most appropriate.

[2] L. B. Linford, D. Williams, V. Josephson, and W. Woodcock, "A Definition of Maximum Range on Aircraft and Its Quantitative Determination," RL Report No. 353, Nov. 12, 1942.

cult to obtain). The fact that the slope corresponds to the inverse fourth-power range law indicates that ground reflections were negligible.[1] The intercept on the horizontal axis gives the free-space range R_0, where we now mean R_0 to be the range at which the echo is visible in the noise on an A-scope during 50 per cent of a series of 5-sec intervals.[2]

Two more important factors must be specified before the figures above have a unique meaning. The measurements were made with the radar trained directly on the target and with the operator watching for the echo at a known spot on an A-scope. The measurement intervals were so long that the highest possible performance was observed (variation of the interval from 2 to 32 sec produced about 5 per cent variation in σ). Extrapolation of these results to situations in which the observation time may be 0.01 sec or less because of rapid scanning and narrow beamwidths and to the use of an intensity-modulated indicator such as a PPI may lead to appreciable error in prediction of performance. In general, the effective values of σ are likely to be noticeably lower than those obtained during observation times of the order of seconds. A discussion of the many factors affecting detection of echoes near noise level is beyond the scope of this volume.[3] The problem has been mentioned here only to point out the need for use of the concept of radar cross section with appropriate regard for its relation to the properties of the radar system.

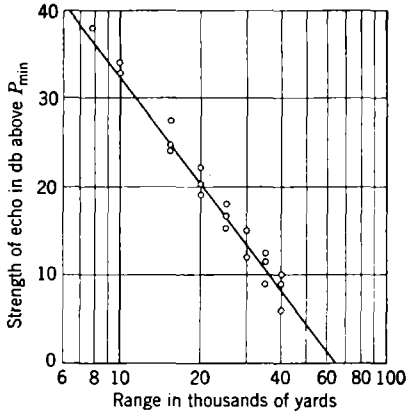


FIG. 6·5.—Echo strength from an aircraft as a function of range, as obtained by a 10-cm radar system. The target is viewed over rough terrain, with head-on aspect. The heavy line, drawn as explained in the text, gives a free-space range R_0 of 63,000 yd.

Some of the results obtained by Linford[4] are given in Table 6·1.

Simultaneous comparisons on 10 and 3 cm (and a few rough checks on 1.25 cm) indicate that for most aircraft σ is not noticeably dependent upon wavelength in this range of wavelengths. Polarization is also of negligible importance. (This result would be expected in view of the remarks of the

[1] Figure 5·19 shows measurements made over the same flight path to determine the amount of specular reflection present. Both one-way and radar measurements showed the propagation to correspond essentially to that in free space.

[2] Because of screening by near-by hills the measurements could not be carried beyond 40,000 yd in the case illustrated.

[3] Consult Vols. 1 and 24 of the Radiation Laboratory Series. See also A. V. Haeff, *Proc. IRE*, **34**, 857 (1946).

[4] *Op. cit.*, also other unpublished data.

preceding section concerning targets with radii of curvature much greater than the wavelength.) All measurements were made with either head- or tail-on aspect. Some aircraft show noticeably larger cross sections head-on than tail-on, but many others show no difference. Usually an aircraft viewed from the side gives very strong echoes in a series of sharp lobes, but these can hardly be used as a basis for a useful definition of σ.

TABLE 6 1.—AIRCRAFT CROSS SECTIONS

Aircraft type	σ, ft^2
OS2U	170
Curtiss-Wright 15D	410
J2F Grumman amphibian	440
B-17	800
B-18	640
AT-11	200
PBY	560
Taylorcraft	170

For the longer radar wavelengths, greater than about 1 m, a large amount of work has been done to calculate σ by treating sections of the aircraft as cylinders, ellipsoids, or various combinations of curved surfaces. Some of these methods have achieved a moderate degree of success, but we shall not discuss them here. For the longer wavelengths the scanning period is so long that it presents no difficulties, and in general the whole detection problem is less difficult than it is in the microwave range. It is important to observe that for the longer wavelengths the effective cross section is noticeably dependent upon both wavelength and polarization. Resonance effects in the various parts of the aircraft become important. Many model studies have been made, and they probably represent the most fruitful method for future study of both microwave and meter wave cross sections.[1] A large amount of work remains to be done both in the field of measurements of scattering and in the application of the results to radar operating techniques in a way to minimize uncertainties and confusion.

6·5. Radar Cross Section of Ships.—Most of the difficulties described in the preceding section and, in addition, others more serious are encountered when one attempts to define and measure σ for ships and other surface targets. Although an aircraft can usually be considered to intercept and scatter a plane wave,[2] the wave incident on a ship is far from plane, and its characteristics are strongly dependent upon range, wave-

[1] For example, see the following report: R. B. Jacques, "B-17E Bomber at 100 Mc. Reflection Patterns," Ohio State University Research Foundation, NDRC Report No. 15-759-22, March 18, 1944; G. Sinclair, *Proc. IRE*, **36**, 1364 (1948).

[2] Unless, of course, the vertical extent of the aircraft is comparable to the thickness of the lobes of the interference pattern. The scattering cross section of any target is independent of its position only if its dimensions are small compared with the space interference pattern.

length, radar height, and reflecting properties of the sea surface. Very often at the range of threshold detection the bulk of a ship lies below the lowest lobe in the interference pattern (in the intermediate region discussed in Sec. 2·15), and the tops of the rigging or stacks may supply the echo first detected.[1] If range is decreased or radar height increased, the deck structures and hull receive increasing illumination, until finally they usually supply the major part of the echo (although the orientation affects the contribution of the hull).

Analysis of scattering under these circumstances requires two major simplifications. (1) The target must be replaced by a presumably equivalent model of simple shape and (2) assumptions must be made regarding the distribution of field over the target. All procedures adopted thus far for ships represent extreme oversimplification of the true situation but have the advantage of obtaining, with a reasonable amount of effort, results that bear sufficient resemblance to experimental facts to be useful. No attempt will be made here to analyze or compare in detail the various proposed methods. None rests on firm enough theoretical grounds or has enough unambiguous supporting experimental evidence that it can claim superiority over all the others. Here again, as in the case of aircraft, only certain major essentials of the scattering pattern can be obtained, and the mass of "fine structure" of the lobe patterns cannot be predicted.

Once the incident field and geometry of the target model have been decided upon, surface currents can be estimated and the back-scattered field can be computed. The simplest analysis is that in which the effective area of the ship is assumed to be concentrated at a single point above the surface and the incident wave is assumed to be that obtained over a flat, perfectly conducting surface. If the heights of the radar and target are z_1 and z_2, respectively, and the ship cross section is σ, the pattern-propagation factor F is $2\sin(2\pi z_1 z_2/\lambda R)$, and the radar equation (Sec. 2·2) becomes

$$\frac{P_r}{P_t} = \frac{G^2\lambda^2}{(4\pi)^3 R^4}\left[2\sin\left(\frac{2\pi z_1 z_2}{\lambda R}\right)\right]^4 \sigma, \tag{68}$$

where the antenna pattern is omitted because of the small angles involved. This will be recognized as the equation for the plane-earth version of the curve of Fig. 5·8. It predicts large fluctuations in echo strength at ranges less than that of the most remote maximum of the interference pattern $4z_1z_2/\lambda$; whereas for ranges sufficiently greater that the argument of the sine function is small, Eq. (68) becomes

$$\frac{P_r}{P_t} = 4\pi G^2\lambda^2\left(\frac{z_1 z_2}{\lambda}\right)^4 \frac{\sigma}{R^8}. \tag{69}$$

[1] This situation tends to be reversed in the presence of surface ducts, which may illuminate the hull strongly; sufficiently deep ducts extend the detection range far beyond the horizon.

Although these two equations are not correct in detail because of the crudity of the underlying assumptions, the second contains two important features that are qualitatively correct. If radar sets on two different wavelengths are to be compared and $G^2\lambda^2$ and σ are held constant, Eq. (69) predicts an advantage in low coverage for the shorter wavelength, the dependence having the form λ^{-4}. (If the effective antenna area is held constant instead of $G^2\lambda^2$ the dependence has the form λ^{-6}.) A striking

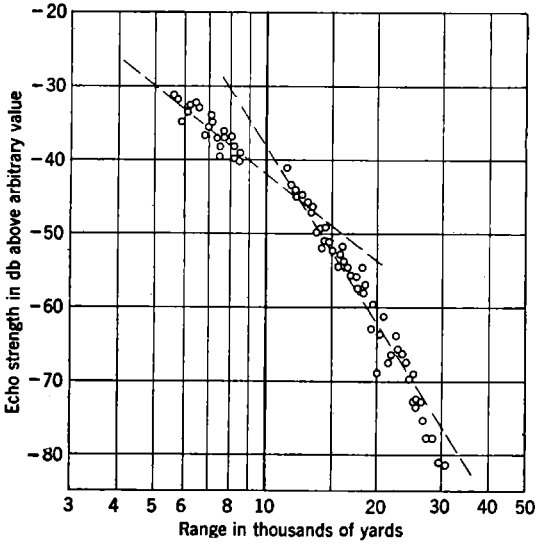


FIG. 6·6.—Typical plot of ship echo as a function of range. Radar height is 125 ft; $\lambda = 30$ cm. The straight lines correspond to variation of echo strength with R^{-4} and R^{-8}. (*NRL data.*)

advantage of the shorter wavelengths for detection of small surface objects and ships on a smooth ocean surface is actually found in practice.

The second important feature of Eq. (69) is the R^{-8} range dependence. The true dependence is actually more complicated than this, of course, but within the limits of possible measurement accuracy it represents many actual cases reasonably well.[1] If the threshold signal is reached in the region in which Eq. (69) is valid,

$$R_{\max} = \left(\frac{4\pi G^2\lambda^2\sigma P_t}{P_{\min}}\right)^{1/4} \left(\frac{z_1 z_2}{\lambda}\right)^{1/2}. \tag{70}$$

A plot of echo power from a medium-sized ship as a function of range usually contains a section for short ranges that varies as R^{-4} as predicted by Eq. (68), but the fluctuations of the sine factor are absent or are poorly defined because the echo comes from a large number of scattering regions

[1] It is far from correct if superrefraction is present (*cf.* Fig. 1.5).

rather than from a single point. At a range that is sometimes fairly well defined, the R^{-4} dependence changes to something approximating R^{-8}. The latter variation may persist to the maximum detection range or may become even more rapid, depending upon how soon the earth's curvature begins to be effective. Figure 6·6 shows a plot of echo strength from a ship, illustrating the features just described. (Marked deviations from this example are not infrequent, however.) As in the aircraft case discussed in the preceding section, obtaining useful numerical data of the type shown in Fig. 6·6 requires a standardized technique for accurate signal measurement in order to minimize the ambiguities caused by the large and rapid echo fluctuations that invariably occur. Usually the information plotted is the maximum echo measured by a signal generator in a fixed time interval. (For the data in Fig. 6·6 the interval was 15 sec.)

Although a ship is not a point target at a fixed height, there is sufficient similarity between curves like that of Fig. 6·6 and the trend predicted by Eqs. (68) and (69), omitting the sinusoidal fluctuations, that useful information can be obtained on this hypothesis. Assuming that the transition between the R^{-4} and R^{-8} sections of the echo-vs.-range curves occurs in the vicinity of the most remote maximum of the interference pattern, where $R = 4z_1z_2/\lambda$, the height z_2 may be estimated. Having z_2, σ may be obtained from Eq. (69). Measurements made in this way show a wide range of values of σ, depending on the type of vessel. Some of these measurements are summarized in Table 6·2. The radar height is 110 ft.[1]

TABLE 6·2.—SHIP CROSS SECTIONS

Ship	$\sigma \times 10^{-3}$ ft²	
	$\lambda = 3$ cm	$\lambda = 10$ cm
Tanker	24	24
Cruiser	150	150
Small freighter	1.5	1.5
Medium freighter	80	80
Large freighter	160	160
Small surfaced submarine	1.5	0.4

As would be expected, there was no noticeable dependence of cross section on wavelength or polarization in this range (except for one case), as most of the reflecting portions are much larger than the wavelength. There is, however, a marked dependence on orientation of the vessel at any wavelength. Broadside aspect gives very strong echoes, presumably because of the large areas of hull with large radius of curvature.

[1] Balzer, Counter, Fairbank, Gordy, and Hudspeth, "Overwater Observations at X and S Frequencies," RL Report No. 401, July 26, 1943.

It is clear at this juncture that we have departed very far from the original concept of cross section defined in Chap. 2. Cross section is no longer a function of the target properties alone but depends upon height of the radar, the range, wavelength, and the over-all performance of the radar set. Consequently, to be useful it must be stated in conjunction with all of the other quantities affecting it. This point will become even more apparent as other methods of approach are considered.

Instead of representing a ship by a point target at a fixed height, some investigators choose for a model a flat, perfectly conducting plate projecting vertically from the ocean surface, viewed along the normal to the plate.[1] The echo from the plate may be calculated by the methods given earlier if the effects of reflection at the surface are taken into account. We shall consider that the reflection coefficient at the ocean surface is -1. The field incident upon the target is in effect the sum of two waves, one directly from the source and one from the image of the source. The wave from the image is 180° out of phase with the direct wave because of the phase reversal upon reflection at the surface.[2] The surface current K on the surface of the sheet is then twice the sum of the direct and reflected magnetic field strengths. If R_1 is the distance from the source at the height z_1 to the point z_2 on the sheet, if R_2 is the distance from the image to the point z_2, and if only small angles are considered,

$$K = 2H_0(e^{-ikR_1} - e^{-ikR_2}) = 2H_0e^{-ikR_1}(1 - e^{-ik\,\Delta R}),$$

where $\Delta R = R_2 - R_1$. Equation (32) in Sec. 6·3 gives a formula for the back-scattered field in terms of the surface current.[3] We now have two fields to add at the source, however; one from the current sheet on the plate and one from its image. The integrand of Eq. (32) will then contain $K(e^{-ikR_1} - e^{-ikR_2})$ rather than Ke^{-ikR_1} alone. The back-scattered field from a plate of width b and total height h is then

$$H^s = \frac{iH_0b}{\lambda R}\int_0^h e^{-i2kR_1}(1 - e^{-ik\,\Delta R})^2\,dz_2, \tag{71}$$

[1] M. V. Wilkes, J. A. Ramsay, and P. B. Blow, "Theory of the Performance of Radar on Ship Targets," ADRDE and CAEE Joint Report, CAEE Ref. 69/C/149, ARDE Ref. R04/2/CR252, July 1944. See also a series of reports from the Naval Research Laboratory entitled "Radar Cross Section of Ship Targets," Parts I-VI, by Katzin, Ament, McDonald, and Barr. These reports give measurements for the wavelength range from 1.5 m to 10 cm. See also M. V. Wilkes and J. A. Ramsey, *Proc. Camb. Phil. Soc.*, **43**, 220 (1947).

[2] It will be recalled from Sec. 5·1 that this is a good approximation for either polarization at very small grazing angles because of the finite conductivity of the water.

[3] The vector notation is omitted here because the scattering sheet is assumed much larger than the wavelength and there is no depolarization by the plate or by the ocean surface.

where R is the surface range of the target. Introducing the customary flat-earth approximations from Sec. 2·2,

$$R_1 \approx R\left[1 + \frac{1}{2}\left(\frac{z_2 - z_1}{R}\right)^2\right], \qquad \Delta R \approx \frac{2z_1z_2}{R},$$

and rewriting the factor $(1 - e^{-ik\,\Delta R})^2$ in terms of the exponential definition of the sine, we have for Eq. (71)

$$H^s \approx -i\,\frac{4H_0b}{\lambda R}\,e^{-i2kR(1+z^2/2R^2)}\int_0^h e^{\frac{-i2\pi z_2^2}{\lambda R}}\sin^2\left(\frac{2\pi z_1z_2}{\lambda R}\right)dz_2. \tag{72}$$

An alternative way of writing the intergral in Eq. (72) is to retain the exponential form of Eq. (71) and to expand the square in the usual way. The integral then becomes the sum of three Fresnel integrals of the form

$$\frac{\sqrt{\lambda R}}{2}\int_{u_1}^{u_2} e^{-iu^2}\,du, \tag{73}$$

where u_1 and u_2 are different for each of the three integrals and one integral has an additional external phase factor.[1]

In the work of the Naval Research Laboratory group (*loc. cit.*) the phase factor in the integrand of Eq. (72) has been ignored; that is, it is assumed that all parts of the surface are excited in phase and the total variation of the field on the surface of the target is expressed by the $\sin^2$ term. This is hardly an acceptable approximation in a general analysis of the flat plate[2], although from a practical point of view this fact is not likely to be important because of the inadequacy of the flat-plate flat-earth model in representing in detail an actual ship.

If Eq. (72) is integrated omitting the phase factor in the integrand, one finds

$$H^s = \frac{2H_0bh}{\lambda R}\left(1 - \frac{\sin\Delta}{\Delta}\right), \tag{74}$$

where

$$\Delta = \frac{4\pi z_1h}{\lambda R}, \tag{75}$$

[1] This procedure was followed in the analysis of echoes from a standpipe viewed over water (see Sec. 6·15). It is particularly advantageous if the effects of the tilt of the target are to be considered.

[2] At the maximum of the lowest lobe of the interference pattern the exponent in the phase factor of the integrand is $\pi\lambda R/8z_1^2$, which not infrequently may be too large to neglect.

and irrelevant external phase factors have been omitted. The cross section is then

$$\sigma = 4\pi R^2 \left|\frac{H^s}{H_0}\right|^2 = 4\sigma_0 \left(1 - \frac{\sin \Delta}{\Delta}\right)^2, \tag{76}$$

where σ_0 is the cross section of the plate in free space at normal incidence, namely,

$$\sigma_0 = 4\pi \left(\frac{bh}{\lambda}\right)^2. \tag{77}$$

M. Katzin points out that the factor $\left(1 - \frac{\sin \Delta}{\Delta}\right)$ can be approximated, without incurring an error of more than 2.6 db at any range, by $\Delta^2/6$ for $\Delta < 0.78\pi$ and by unity for $\Delta > 0.78\pi$. This fact provides a convenient means of distinguishing between the behavior of σ for small and large ranges, called the "near" and "far" zones, respectively.[1] For the two zones we have

$$\text{Near zone, } \Delta < 0.78\pi, \qquad \sigma \approx 4\sigma_0, \tag{78}$$

$$\text{Far zone, } \Delta > 0.78\pi, \qquad \sigma \approx \left(\frac{z_1}{R}\right)^4 \left(\frac{4\pi h}{\lambda}\right)^4 \frac{\sigma_0}{9}. \tag{78a}$$

Katzin defines two kinds of cross section on the basis of these expressions, the second of which differs from our usage in this volume. The near-zone cross section is defined as

$$\sigma_n = 4\sigma_0, \tag{79}$$

and the far-zone cross section is

$$\sigma_f = \frac{1}{9}\left(\frac{4\pi h}{\lambda}\right)^4 \sigma_0. \tag{80}$$

This method of defining σ_f splits off the $(z_1/R)^4$ factor pertaining to a given set of measurement conditions in order that a quantity dependent only on the target properties may be obtained. Application of a numerical value of σ_f measured for one geometry to a different geometry can be expected to give reliable results only in so far as the plane-reflector model represents the actual target and the earth curvature is unimportant.

A further consequence of the definition of σ_f is that its numerical value is likely to be extremely large (10^{12}ft^2 or larger for medium-sized vessels) but the reduction to a magnitude comparable with values determined by other methods such as those of Table 6·2 is obtained when the factor $(z_1/R)^4$ is applied to reduce σ_f to our form of σ [Eq. (78)]. This

[1] The dividing line between the near and far zones, $\Delta = 0.78\pi$, corresponds to a range $R \approx 5z_1h/\lambda$, which is 25 per cent greater than the range to the most distant maximum of the interference pattern at the height h over a plane earth.

step also emphasizes that the effective value of cross section is in reality a function of range and radar height.

The wavelength trend inherent in the definition of σ_f has been verified roughly in a few cases in the wavelength range from 1.5 m to 10 cm but in general is not so simple as indicated.[1]

The transmission equations for the two zones are[2]

$$\left(\frac{P_r}{P_t}\right)_{\text{near}} = \frac{G^2\lambda^2}{(4\pi)^3R^4}\frac{\sigma_0}{4}, \tag{81}$$

$$\left(\frac{P_r}{P_t}\right)_{\text{far}} = 4\pi G^2\lambda^2\left(\frac{z_1h}{\lambda}\right)^4\frac{\sigma_0}{9R^8}, \tag{82}$$

or

$$\left(\frac{P_r}{P_t}\right)_{\text{far}} = \frac{G^2\lambda^2}{(4\pi)^3R^4}\left(\frac{z_1}{R}\right)^4\sigma_f. \tag{83}$$

With these equations and the necessary absolute calibrations of the radar set σ_n and σ_f may be determined from the type of data presented in Fig. 6·6[3].

The relation of the actual dimensions of the ship to b and h or to σ_0 is not clear, because of the artificiality of the flat-plate model. The fact that σ_0 is proportional to $(bh)^2$ does not necessarily imply that σ_0 should be proportional to the square of the geometrical cross section of the ship.[4]

On the other hand, the ADRDE-CAEE group, assuming incoherent superposition of the scattered energy, predicts that the effective cross section is proportional to bh and reports that its data are not inconsistent with the view that the effective area is proportional to the silhouette area of the ship and is independent of wavelength (presumably in the vicinity of 10 cm, as no data are reported for other wavelengths).

[1] The predicted wavelength dependence arises from the fact that the model is a flat plate viewed at normal incidence. If, on the other hand, the target were assumed to consist of a combination of plane and curved surfaces with random orientations, the wavelength dependence would be expected to be much less marked. This would appear to be a matter worthy of further investigation.

[2] Compare with Eqs. (68) and (69).

[3] In the past the lack of reliable means of absolute measurements of echo strength and radar set characteristics has often necessitated comparison with echoes from "standard targets," the free-space cross section and appropriate value of F for these targets supposedly being known accurately. This procedure is difficult in practice and is beset with numerous pitfalls. It is to be expected that technical advances will soon make possible absolute measurements of sufficient accuracy that calibration by external means will be unnecessary.

[4] The British Army Operational Research Group, however, reports that the effective cross section for $\lambda = 10$ cm of ships, viewed at aspects other than broadside, is proportional to the *square* of the broadside silhouette area. See "The Performance of 10 cm Radar on Surface Craft," AORG Report No. 155, Jan. 3, 1944. The approach is essentially empirical.

The ADRDE-CAEE approach also differs from that of the Naval Research Laboratory in that incoherent scattering is assumed from elementary units of area. The received echo power is then assumed to be given by

$$P_r \propto \frac{G^2\lambda^2 b}{R^4} \int_0^h \sin^4\left(\frac{2\pi z_1 z_2}{\lambda R}\right) dz_2. \tag{84}$$

(A further correction is also made for the actual variation of the incident field strength at very low altitudes as the ship moves into the diffraction region.) An elaborate set of charts has been made, predicting echo strength on various types of ships as a function of radar performance, wavelength, and set height. The effect of earth curvature becomes fairly important for high-performance sets with high sites.

Two other methods have been proposed that consider the ship as an array of isotropically scattering elements with random relative phases. Although expressed in slightly different ways, the two methods are essentially identical. The Radio Research Laboratory suggests that in the far zone the cross section be obtained by a formula of the type

$$\sigma \propto \frac{1}{R^4}\left(\frac{z_1}{\lambda}\right)^4 \sum_i \sigma_i z_i^4, \tag{85}$$

where σ_i is the free-space cross section of the ith horizontal strip of the ship and z_i is its height above the surface.[1]

The Telecommunications Research Establishment suggests[2] instead an "echoing parameter" T, defined by

$$T^2 = \int z_2^4 \, dS, \tag{86}$$

where dS is an element of area at the height z_2. This parameter is derived from an analysis that assumes the elementary scattering elements to scatter isotropically with random phases. For a perfectly conducting plane earth, the back-scattered Poynting vector is then

$$S^s = \frac{P_t G}{\pi^2 R^4} \int \sin^4\left(\frac{2\pi z_1 z_2}{\lambda R}\right) dS. \tag{87}$$

Restricting the analysis to the region well below the maximum of the first lobe in order that the sine function may be replaced by its argument, Eq. (87) becomes

$$S^s = \frac{P_t G}{\pi^2 R^4}\left(\frac{2\pi z_1}{\lambda R}\right)^4 T^2,$$

[1] Kuhn and Sutro, "Theory of Ship Echoes as Applied to Naval RCM Operations," RRL Report No. 411-93, July 14, 1944.

[2] W. Walkinshaw and J. E. Curran, "Screened and Unscreened Radar Coverage for Surface Targets," TRE Report No. T1666, 1944.

where T^2 is given by Eq. (86). Although T has the dimensions of volume, there is no relation to the volume of the ship. Estimates of T have been made from ship silhouettes. In terms of σ as used in this volume,

$$\sigma = \left(\frac{4\pi z_1}{\lambda R}\right)^4 T^2. \tag{88}$$

The foregoing discussion has attempted to trace only the fundamental ideas underlying current methods of estimating ship cross sections theoretically, without reviewing the experimental data that have been presented in support of the several points of view. (Security regulations do not permit such a review.) In a detailed analysis of all the available evidence, the ADRDE-CAEE group remarks (*loc. cit.*), "In view of the uncertainties in the definition of σ and its dependence on the situation of the target we are of the opinion that far from simplifying the problem one adds to its difficulties by endeavoring to define and measure the radar cross section of ships." Even if one does not agree entirely with this point of view, it must be admitted that there is a large measure of truth in it. At the same time, however, it emphasizes the need for further development of a practical means of representing a ship in the radar transmission equation.

SEA ECHO

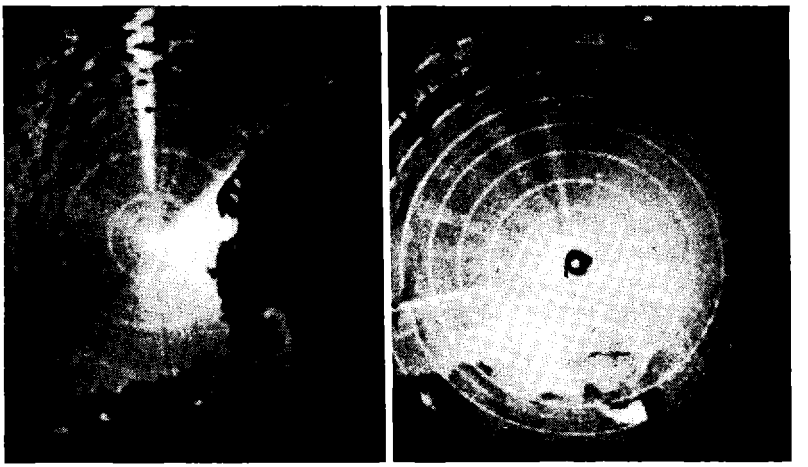
By Herbert Goldstein

6·6. Nature of the Problem.—"Sea echo" is a term used to denote a type of radar reflection associated in some manner with the surface of the sea. This definition has been deliberately worded in a vague fashion because the precise connection between the echo and the sea surface is not yet understood. The purpose of the next few sections is to present the existing information on the characteristics of sea echo and to discuss, in the light of this information, which phenomena of the sea surface may be responsible for the echo.

The appearance of the echo is well illustrated in Fig. 6·7, which shows two PPI photographs taken with a high-power 10-cm airborne system on two occasions. The sweep lengths differ slightly but are about 100 miles in both cases. Figure 6·7*a* was taken with the system slightly inland, near Boston, Mass., on a day when sea echo was negligible. The distinctive geographical features of the region—Cape Cod, Nantucket, Martha's Vineyard, etc.—are clearly recognizable. Figure 6·7*b* was taken with the same system in roughly the same location but on a day when sea echo was present out to ranges of more than 50 miles. The practical importance of the sea-echo problem is strikingly demonstrated by this photograph (which is more representative of normal conditions during winter than is Fig. 6·7*a*), for the signal acts as a "built-in jammer," almost completely obscuring

Cape Cod and the Massachusetts Bay area. Sea echo may thus often constitute a serious limitation on the effectiveness of radar systems.[1]

Whatever its nature may be, the target responsible for sea echo can always be resolved into a number of individual scatterers that can be treated as independent of each other. These scatterers may range in size from small spray droplets to large-scale surface waves or even a train of waves. Let us consider the radar echo received at one given instant, that is, at some



(*a*) (*b*)

FIG. 6·7.—Two photographs of the PPI display on a 10-cm airborne radar system. In both photographs the system was near Boston, Mass. (*a*) Sea echo was negligible on the day the photograph was taken. (*b*) Strong sea echo is present out to 50 miles.

given range on the indicators. Then the contribution to this signal by the reflection from the jth scatterer is given by

$$P_{rj} = \frac{G^2\lambda^2}{(4\pi)^3 R_j^4} f^4(\theta, \phi_j) p(R_j)\ \sigma_j. \tag{89}$$

In this formula the jth scatterer is identified by a pair of coordinates ϕ_j, R_j, where ϕ is the azimuth measured from the electrical axis of the antenna and R is the slant range from the radar system. It will be assumed in the following discussion that the depression angle θ is sufficiently small so that slant range and ground range can be used interchangeably.[2] The function $f(\theta, \phi)$ is the antenna pattern function. The

[1] Circuits have been devised that considerably improve the appearance of the PPI for navigational purposes (see Vol. 24, Chap. 11, Radiation Laboratory Series); not, however, without some sacrifice of weaker signals.

[2] In Secs. 6·6 to 6·12, θ will be used interchangeably for depression angle and grazing angle. There is a slight difference between the two angles because of the curvature of the earth, but it is completely negligible in the following discussions.

function $p(R)$ is proportional to the magnitude of the Poynting vector of the incident wave at R at such a time that the echo reflected from the scatterer returns to the radar at the instant t_0 under consideration and is related to the pulse shape in the following manner. If $g(R\text{-}ct)$ gives the instantaneous distribution of the transmitted power in range, then

$$p(R) = g(2R\text{-}ct_0) \tag{90}$$

The quantity σ_j is the radar cross section of the scatterer, as defined in Chap. 2, except that it includes the pattern-propagation factor F^4. Hence σ_j, as used here, may possibly depend on the height (if any) of the scatterer above the sea[1] and on the depression angle θ.

The total signal received will be the sum of the contributions from all scatterers. As the individual scatterers are considered independent, the received power,[2] averaged over many pulse periods, is just the sum of the powers received from each scatterer separately,

$$\overline{P}_r = \frac{G^2\lambda^2}{(4\pi)^3} \sum_j \frac{f^4(\theta_j, \phi_j)}{R_j^4}\, p(R_j)\, \bar{\sigma}_j, \tag{91}$$

where σ_j is the time average of the scatterers' cross sections.

It will now be assumed that the sum involves so large a number of scatterers that the summation may be replaced by an integral. For this purpose a density function $N(R, \phi, \bar{\sigma})R\, dR\, d\phi\, d\bar{\sigma}$ is introduced, which gives the number of scatterers in an area element $R\, dR\, d\phi$ for which the radar cross section lies between $\bar{\sigma}$ and $\bar{\sigma} + d\bar{\sigma}$. It will be further assumed that the scatterers are distributed uniformly and homogeneously over the surface of the sea, so that the function N depends only on $\bar{\sigma}$. With these assumptions, Eq. (91) becomes

$$\overline{P}_r = \frac{G^2\lambda^2}{(4\pi)^3} \int_0^\infty \int_{-\pi}^{+\pi} f^4(\theta, \phi)\, d\phi\, \frac{p(R)}{R^3}\, dR \int_0^\infty \bar{\sigma}\, N(\bar{\sigma})\, d\bar{\sigma}. \tag{92}$$

The integral over $\bar{\sigma}$ defines a dimensionless quanity, which for simplicity of reference will be denoted by σ^0 and which will be called the *average radar cross section of sea echo per unit area of the sea surface.* Similarly, the integral of f^4 over the azimuth ϕ defines a beamwidth Φ, a function of θ, given by

$$\int_{-\pi}^{+\pi} f^4(\theta, \phi)\, d\phi = \Phi f^4(\theta, 0). \tag{93}$$

[1] It will not depend on the antenna gain because the height of the scatterers must be so low that there is no difference in angle between the direct and reflected waves.

[2] By received power is meant the power in the r-f received signal as obtained by averaging over a small interval of range, small compared with the pulse length but covering many cycles of the radio frequency. By average received power is meant the average of these values over many sweeps.

With these definitions Eq. (91) simplifies to

$$\overline{P}_r = \frac{G^2\lambda^2\Phi}{(4\pi)^3}\int_0^\infty f^4(\theta, 0)\,\sigma^0(\theta)\,\frac{p(R)\,dR}{R^3}. \tag{94}$$

The function $p(R)$ is different from zero only over a distance given roughly by the pulse length. In that distance, under the geometrical conditions assumed, f^4, σ^0, and R^{-3} are essentially constant and may be taken out from under the integral sign. The resulting integral defines the product $P_t\tau$ in terms of the pulse shape from Eq. (90),

$$\int_0^\infty p(R)\,dR = \frac{P_t\tau}{2}\,c, \tag{95}$$

where P_t and τ are closely proportional to the peak power and pulse duration as conventionally defined. The final expression for the time average of the power received at one given range on an indicator is, therefore,

$$\overline{P}_r = P_t\frac{G^2\lambda^2}{(4\pi R)^3}\,\Phi f^4(\theta)\,\frac{\tau c}{2}\,\sigma^0(\theta) \qquad (\theta \text{ small}), \tag{96}$$

where $f^4(\theta, 0)$ has been abbreviated by $f^4(\theta)$.

As has been noted, the quantity σ^0 is dimensionless, being the ratio of a cross section to unit area. It is entirely correct therefore to speak of $\log \sigma^0$, and in later discussions values of σ^0 will almost always be given, for convenience in decibel notation, as $10 \log_{10} \sigma^0$. It is often not feasible to obtain σ^0 in absolute measure, the values being indeterminate to the extent of an unknown multiplicative factor. When the decibel notation is used in such cases, the values will be labeled "referred to an arbitrary datum." At all other times it will be assumed that σ^0 is given in absolute measure. The value of σ^0 need not necessarily be less than unity, for the radar cross section of a surface may be much greater than the geometrical area.

The azimuthal beamwidth Φ used in Eq. (96) is not identical with the value as defined conventionally in terms of half-power points on the antenna pattern. For the usual antenna patterns conversion between the two values is obtained by a nearly constant multiplicative factor of the order of unity. The difference is of no significance in this work.

Equation (96) is limited to small grazing angles only as a consequence of the geometrical conditions assumed. A more general expression can be derived which is not subject to these geometrical limitations and which therefore holds for all θ. The final result is

$$\overline{P}_r = \frac{G^2\lambda^2}{(4\pi)^3 z^2}\int_0^{\pi/2} f^4(\theta)\Phi(\theta)\sigma^0(\theta)\,\sin\theta\; p\left(\frac{z}{\sin\theta}\right)d\theta, \tag{97}$$

where z is the height of the radar above the sea surface. In addition to the usual assumption of a large number of scatterers with uniform distribution over the sea surface, Eq. (97) assumes only that the field over

any individual scatterer is uniform and that the azimuthal beamwidth Φ is small compared with unity.

A useful limiting case that can be obtained from Eq. (97) occurs with "pencil" beams at such high angles that the pulse packet covers a much larger region of the sea surface than is defined by the beamwidths. The function $f^4(\theta)$ then varies much more rapidly than the other quantities in the integral in Eq. (96), and they may therefore be taken outside the integral. The remaining integral defines a vertical beamwidth Θ in analogy to Eq. (93), and the formula for $\overline{P}_r$ reduces to

$$\overline{P}_r = \frac{P_t G^2 \lambda^2 \Theta \Phi}{(4\pi)^3 2^2} \sin \theta \sigma^0(\theta) = \frac{P_t G^2 \lambda^2 \Theta \Phi}{(4\pi)^3 R^2 \sin \theta} \sigma^0(\theta) \qquad (\theta \text{ large}), \qquad (98)$$

where θ is not the variable of Eq. (97), but is now the depression angle of the axis of the beam. Equation (98) is frequently referred to as the "large-angle" formula, in contradistinction to Eq. (96), although it holds only for pencil beams and therefore breaks down even at high angles if a "fan" beam, such as the cosecant-squared beam, is used. The transition between Eqs. (96) and (98) may be considered to occur at the angle for which the distance on the sea surface illuminated by the vertical section of the beam is equal to the pulse length on the sea surface, that is, an angle θ such that

$$\tan \theta = \frac{\Theta R}{\frac{\tau c}{2}}. \qquad (99)$$

Equation (97) and companion Eqs. (96) and (98) are essentially phenomenological. They require no assumptions about the specific nature of the scatterers responsible for the echo. As a consequence, the information that can be obtained from them is rather limited. Thus the important quantity σ^0 is independent of many of the characteristics of the radar set but is a function of at least five parameters:

1. The angle that the incident ray makes with the horizontal (θ).
2. Wavelength.
3. Direction of polarization.
4. State of the sea.
5. Azimuth relative to the wave pattern on the surface of the sea (that is, upwind or downwind, etc.).

Equations (96) to (98) tell us nothing about the dependence of σ^0 on any of these parameters; this information can be obtained only by experiment or, by assuming a particular scattering mechanism, from theory.

The succeeding sections are to a large extent devoted to the presentation of the experimental and theoretical information about the properties of σ^0. However, as sea echo is so intimately connected with the sea surface, the pertinent oceanographic knowledge of the characteristics of the sea surface is first summarized in Sec. 6·7 and the validity then con-

sidered (Sec. 6·8) of the fundamental assumption that lies behind Eqs. (96) to (98), that is, that the distribution of the sea-echo scatterers is homogeneous. The conditions under which it is possible to speak of a quantity σ^0 having thus been established, experiments carried out at the Radiation Laboratory on the wavelength dependence of σ^0 are described in Sec. 6·9, which goes into some detail in order to introduce the peculiarities of, and the difficulties involved in, the experimental measurement of sea-echo characteristics. The results of all known and reliable measurements of the properties of σ^0 are then given briefly (see Sec. 6·10). The scattering cross section σ^0 naturally refers to the *time average signal*; however, the *fluctuation* of sea echo is one of its most striking characteristics. Although the mechanism of this fluctuation properly belongs in one of the later sections of this chapter, Sec. 6·11 tells what information about the scattering process may be obtained from a study of the fluctuations. Finally, in Sec. 6·12 the theories of sea echo that have been proposed are examined in the light of all the experimental evidence, and suggestions are made concerning the direction that future investigations should take.

6·7. Nature of the Sea Surface.—It is obvious that the phenomenon of sea echo is in some manner related to the surface of the sea. For a proper understanding of the phenomenon the present state of knowledge about the characteristics of the sea surface must be briefly reviewed. Most of this knowledge was gained in recent years, some of it from wartime research.[1]

The outstanding feature of the sea's appearance is the presence of periodic or near-periodic waves on the surface. These waves are obviously caused by the action of the wind on the sea,[2] but they may outlast, or outrun, the wind system responsible for their growth. Wave structures still driven by their generating winds are commonly referred to as "sea";

[1] H. U. Sverdrup, *Oceanography for Meteorologists*, Prentice-Hall, New York, 1942. See pp. 133-137 for a summary of the information available before the war. For a longer treatment see H. B. Bigelow and W. T. Edmondson, *Wind Waves at Sea Breakers and Surf*, U.S. Govt. Printing Office, 1947. V. Cornish, *Ocean Waves and Kindred Geophysical Phenomena*, Macmillan, London, 1934, is a classic, notable especially for its mathematical appendix by H. Jeffreys. American research during the war is to be found chiefly in three publications of the Hydrographic Office, prepared by the Scripps Institution of Oceanography: "Wind Waves and Swell, Principles in Forecasting," *HO Misc.* 11,275; "Breakers and Surf, Principles in Forecasting," *HO* 234, and "Wind, Sea and Swell, Theory of Relations in Forecasting," H. U. Sverdrup and W. H. Munk, *HO* 601. Two papers on the subject have also been published by H. U. Sverdrup and W. H. Munk: *Trans. Am. Geophys. Union*, **27**, 823 and 828 (1947). Some of the similar British work is to be found in *Trans. Roy. Soc.* (*London*), A, **240**, No. 824 (1948) by N. F. Barber and F. Ursell. A summary of the results of classical hydrodynamics is given in M. P. O'Brien, "A Summary of the Theory of Oscillatory Waves," Beach Erosion Board, Corps of Engineers, U.S. Army, Technical Report No. 2, 1942.

[2] Isolated waves such as "tidal" waves are of no importance in the radar problem.

waves that are no longer driven by the wind but are in the process of decaying are termed "swell."

When the wind is high the waves of a sea often exhibit little similarity to the conventional picture of a regular succession of sinusoidal waves whose crests and troughs extend indefinitely. The sea surface more often resembles heaps of water irregularly piled and of irregular dimensions, occasionally breaking, that is, showing "white water." Such a sea is called "short-crested," as the crest length is of the same order of magnitude as the distance between crests.

"Swell" tends to be much more regular, with near-sinusoidal waves and long crests. As the waves of swell are decaying, their amplitudes are likely to be small, and they may often be completely obscured by a superimposed "sea" caused by local winds.

Even the more regular "swell" is not characterized by a particular wavelength or amplitude; a whole range of values will be present in any given region. One can speak only in terms of "spectra" or probability distributions of these quantities. This variability, an essential characteristic of the sea surface, is even further emphasized in rough seas, where it is often doubtful whether "wavelength" or "wave height" can be defined. The shape of the waves differs radically from the sinusoid and resembles more a trochoid,[1] presenting a peaked appearance. Under the influence of the wind these peaks become still sharper until they break, forming whitecaps. Theoretically the limit of stability occurs for a peak angle of 120°, at which angle the ratio of wave height (from trough to crest) to wavelength, denoted by H/L, exceeds 1/7. These limiting values have been reasonably well verified experimentally. (In actual seas the steepness ratio H/L is found to vary from 1/7 to about 1/20. In swell, of course, the ratio may be much lower, of the order of 1/100.) The formation of whitecaps and spray by the peaking action of the wind (which can occur even when the wave height is small) is to be clearly distinguished from breakers on a beach. Here the action is caused not by the wind but by the modification of *swell* waves as they advance into shallow water. The wavelength decreases and the wave height increases until the peaks break over.

The preceding description of the sea surface has been confined to the large-scale structure, that is, to dimensions of the order of feet. Superimposed upon this structure there are always much smaller waves, or ripples, presumably caused by very local eddies in the wind. In fact, the wavelength spectrum probably exhibits a continuous range from waves of many feet down to ripples of only a few inches. These ripples are comparable in size to the microwave wavelengths, and there is no reason to believe they are not as pertinent, if not much more so, for the

[1] The trochoid is the wave shape of finite amplitude predicted by classical hydrodynamic theory; *cf.* M. P. O'Brien, *op. cit.*, p. 24.

radar problem than are the large and obvious sea waves. Yet nothing is known about these ripples, and it is almost impossible to make observations about their nature in an actual sea. In seeking some empirical correlation between sea-echo intensity and sea state it is therefore necessary to deal with the large-scale wave structure, keeping in mind that any such correlation may only be symptomatic of a more fundamental connection between sea echo and the small ripples.

The great difficulties facing a theory that attempts to account for even the large-scale structure are obvious. By its neglect of viscosity and other dissipative processes, classical hydrodynamics can say nothing about the generation of waves and very little about the dimensions of waves. It does furnish a number of useful relations dealing with the propagation and shape of waves. Thus, for a sinusoidal wave of negligible amplitude the wavelength, period, and velocity are connected by the relations

$$L = \frac{g}{2\pi} T^2 = \frac{2\pi}{g} U^2, \tag{100}$$

where U is the wave velocity and the other symbols have obvious meanings. These equations assume infinite depth, but they may be easily modified to include the effect of finite depth. Although they hold strictly only for very small wave amplitudes, the changes for finite waves (such as the trochoid) are not important.

Jeffreys[1] was the first to construct a satisfactory theory of the generation of waves by wind, involving the turbulent nature of the wind in the lee of an obstacle and the viscosity of the water. On the basis of this theory the Scripps group has been able to build up a more or less complete picture of the growth and decay of waves. A differential equation for the history of the waves has been constructed, essentially from consideration of the "energy budget," that is, all the processes whereby a wave can gain or lose energy. The constants in the solutions obtained from the differential equation are evaluated by means of an experimental graph of wave steepness H/L as a function of the "wave age," the ratio of wave velocity to wind velocity. The results of this theory may be summarized as follows.

Any given wind structure is, of course, not uniform over an unlimited region of the sea. The length of the path to the observer over which a substantially constant wind blows is known as the "fetch." At any given fetch the wave amplitude increases with the length of time that the wind has been blowing. After a certain time a steady state is reached at which the wave height remains constant. The time that it takes to reach this state of equilibrium, known as the "minimum duration," increases with the length of the fetch and the wind velocity. The characteristics of the waves in this equilibrium state are also a function of the fetch and wind velocity. If the wind blows for a time shorter than the minimum duration the wave height does not depend upon the fetch but only upon the time

[1] See V. Cornish, *op. cit.*

and wind velocity.[1] For as long as the wind is building or maintaining the waves, the wave velocity is equal to or slightly less than the wind velocity.

In a similar fashion the transition from a "sea" to "swell" has been worked out. Swell waves usually arise only from the highest, long-crested waves of the sea, as these waves have the most energy and therefore take the longest time to die down.

It has been emphasized that at any given time a whole spectrum of wave heights and lengths will be found in the sea. The term wave height as used above must therefore refer only to some particular characteristic or moment of the spectrum. There is both theoretical and experimental evidence that the particular characteristic involved is the average of the highest third of the spectrum. This is stated by the Scripps group only as a tentative conclusion; a more definite statement must await more detailed experimental data.

The preceding discussion makes obvious the futility of attempting to describe completely the sea state by a single number. Nevertheless, for lack of anything better, a number of scales of sea state have been described and widely used.[2] The use of such scales, especially for the radar problem, can be characterized only as unfortunate. Even experienced seamen on shipboard will differ in their estimates by one or more numbers in the sea scale. Yet sea-scale numbers have often been assigned solely by visual observation from a high-flying airplane. Particularly pernicious is the practice of associating certain sea states with wind force in the Beaufort scale.[3] This "scale" apparently arose from an attempt to give aviators flying over water a method of estimating wind speed, but it is often used in the reverse fashion. It has already been emphasized that the wave structure depends not only on wind speed but also on the duration and fetch of the wind. The sea state cannot be determined simply from a statement of the wind speed.

Admittedly, however, no convenient objective method of specifying sea state is yet available to replace the sea scales.[4] Ideally, one should

[1] See *HO Misc.* 11,275 for graphs for wave height and period as a function of fetch, wind velocity, and duration.

[2] A summary of the more important of these scales will be found in *HO Misc.* 11,275, p. 35.

[3] A brief discussion of a scale of this sort is given in Bureau of Aeronautics, U.S. Navy, *Aerology for Pilots*, McGraw-Hill, New York, 1943, p. 49.

[4] Whenever terms are used in subsequent sections to denote sea states, reference is to the Douglas scale as given in "Instructions to Marine Meteorological Observers," *U.S. Weather Bur. Circ. M*, 7th ed., 1941. An abbreviated form of this scale is given here:

No.	Term	Wave height, ft
0	Calm	0
1	Smooth	<1
2	Slight	1–3
3	Moderate	3–5
4	Rough	5–8
5	Very rough	8–12
6	High	12–20

know the spectra of the various wave characteristics, such as wave height and wavelength, that could be obtained from records of the instantaneous sea height. During the war, pressure recorders were developed for this purpose,[1] and they will probably be widely used in the future. Stereophotography from shipboard has been used in the past, but the analysis is extremely lengthy.[2] The method of taking stereophotographs from low-flying planes with special "strip" cameras shows considerable promise.[3] All the methods require considerable analysis, however, before the desired quantities are obtained. The radar problem is further complicated in that none of the methods so far proposed yield information about quantities likely to be of the most primary concern for sea echo, namely, small ripples or spray.

6·8. Validity of the Fundamental Assumptions.—The form of the "phenomenological" Eqs. (96) to (98) for the average sea-echo intensity depends, aside from purely geometrical factors, upon the assumption that the scatterers are uniformly distributed over the surface of the sea. Before discussing the dependence σ^0 on the various parameters, it is well first to examine the conditions under which this assumption is valid, for only then can one truly speak of a "cross section per unit area." This examination can take two possible forms. One can try to verify those predictions of Eqs. (96) to (98) which are independent of the scattering mechanism, namely, that the echo is proportional to the beamwidth and pulse length and varies inversely as the cube of the range (or as $1/R^2$ for high angles) for constant angle of incidence. Or the distribution of the scatterers can be studied directly by examining the appearance of the echo on a high-resolution radar system. Both methods have been used to some extent.

Measurement of the dependence on range at constant angle of incidence has obvious technical difficulties and has been attempted only once.[4] A 3-cm airborne system was used at several grazing angles from 5° to 90°. At normal incidence the expected $1/R^2$ relation was obtained. For the more interesting lower angles the results were interpreted as $1/R^3$ at large distances but $1/R^2$ for small distances. Such behavior is to be expected qualitatively (*cf.* Sec. 6·6), but the apparent "crossover" points all occur at ranges far greater than those calculated from Eq. (99). The

[1] Brief descriptions of these recorders are to be found in N. J. Holter, *Electronics*, **19,** 94 (1946), in a note in *Nature*, **157,** 165 (1946), and in "Ocean Surface Waves," *Ann. N.Y. Acad. Sci.*, **51,** 343 (1949).

[2] See Bigelow and Edmondson, *op. cit.*, for some contour maps obtained in this fashion by Schumacher.

[3] Recently a radio altimeter has been used at Telecommunications Research Establishment in England to measure wave heights with an accuracy of ± 1 ft. See TRE Report No. T2044, "A Preliminary Report on an Airborne Sea Swell Recorder."

[4] E. W. Cowan, "X-band Sea Return Measurements," RL Report No. 870, Jan. 10, 1946.

accuracy of the measurements is believed not to be great enough to justify the interpretation put on them. One can say only that the results, to within experimental error, are not inconsistant with Eqs. (96) to (98).

Measurement of the variation of echo intensity with beamwidth also involves many technical difficulties because it is almost impossible to change beamwidth without affecting other properties of the system, such as antenna gain. In the course of the experiments on the frequency dependence of σ^0 (Sec. 6·9), measurements were occasionally made on the 3-cm system with both 48- and 18-in. paraboloids, corresponding to a ratio of 2.5 in beamwidth. After correction for the change in antenna gains, the observed differences in received power were between 4 and 7 db, compared with the expected 4 db. The discrepancies are within experimental error and are probably due to uncertainties in the vertical polar diagram of the antennas.

The dependence of echo power on pulse length has been more frequently studied. In the experiments at 3 cm cited previously,[1] measurements were made with pulse duration of 2, 1, and 0.5 μsec. As expected, the return at lower angles was roughly proportional to pulse length and became independent of it at higher angles. Similar results were obtained at 10 cm for $\tau = 1$ and 0.5 μsec in the course of the experiments on frequency dependence.

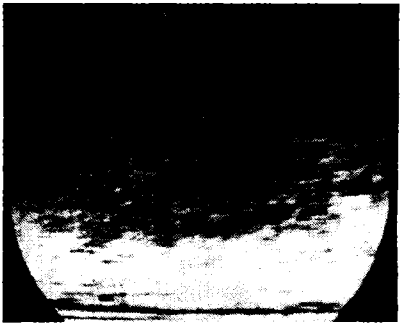


FIG. 6·8.—Photograph of a delayed-sweep B-scope on a 1.25-cm radar, showing sea echo. Dimensions are roughly 20° by 2000 yd. The pulse length is 0.15 μsec.

Measurements have also been made at the Telecommunications Research Establishment[2] with a narrow-beam system operating on 1.25 cm with pulse lengths of 0.5 and 0.2 μsec. The average change obtained was 4.7 db, compared with the 4 db expected. The use of such high-resolution systems, however, has at other times revealed striking examples of breakdown of the "uniformity" assumptions. These deviations are best illustrated by means of the second method, *i.e.*, study of the appearance of the echo.

Figure 6·8 is a photograph of the delayed-sweep B-scope on a system operating at 1.25 cm with a pulse length of 0.15 μsec and 0.8° beamwidth. The dimensions of the picture are roughly 20° by 2000 yd, with the center of the pattern at 2500 yd. The picture was made toward the end of a

[1] *Ibid.*

[2] A. L. Cockroft, H. Davies, and R. A. Smith, "A Quantitative Study of Sea Returns at Centimetre Wavelengths for Moderately Small Angles of Elevation," TRE Report No. T1933, September 1945; also *Proc. Phys. Soc.*, **58**, 717 (1946).

"northeast" storm, with the system looking out over Boston Bay. Instead of the customarily uniform grainy appearance, as in Fig. 6·7*b*, the echo takes on the form of parallel straight rows. Figure 6.9 shows a photograph of the 2000-yd expanded-sweep A-scope taken at the same time. On this indicator the echo is concentrated in the form of "spikes" with intervening

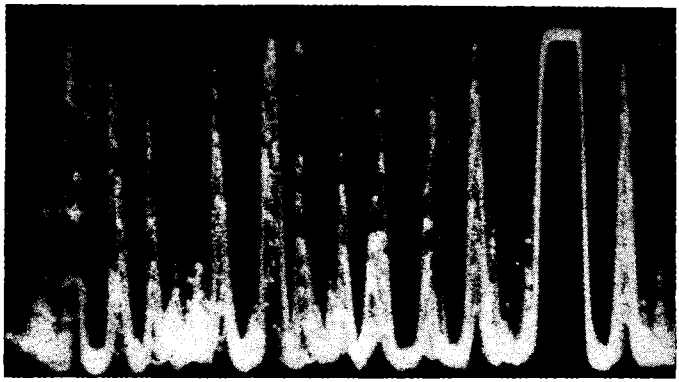


FIG. 6·9.—Photograph of the expanded-sweep A-scope on the system of Fig. 6·8, showing "spiky" appearance of sea echo. The sweep length was 2000 yd. The square-topped pulse is an artificial echo from a signal generator.

regions of little or no signal.[1] (This photograph may be contrasted with Fig. 6·10 which shows the more normal appearance of the echo on the A-scope.)

The echo appearance in Figs. 6·8 and 6·9 is strongly suggestive of the pattern of the waves on the sea surface. This inference is supported by

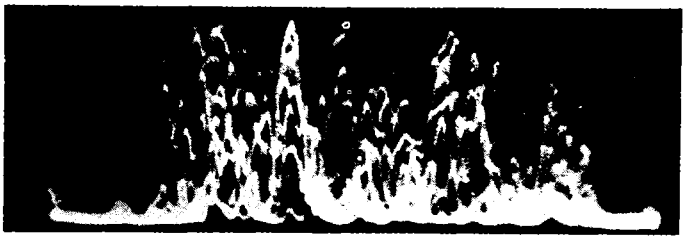


FIG. 6·10.—Normal appearance of sea echo on the A-scope; wavelength 3 cm, sweep length 4000 yd, pulse length 1 μsec.

the following measurements. Assuming that the radar is in effect "seeing" the waves, the distance between rows, or between "spikes," should correspond to the wavelength. The value obtained, when corrected for finite

[1] A "spiky" appearance of the A-scope under similar circumstances has also been reported from Telecommunications Research Establishment (*cf.* TRE Report No. T1933).

depth of the bottom, is 250 ft, which is in itself not unreasonable and agrees fairly well with predictions based on *HO Misc.* 11,275. Furthermore, the wave velocity could be measured, for the "spikes" traveled along the A-scope and could be tracked in range for 10 to 15 sec. The result of this measurement, 25 mph, is again consonant with the wind speed at the time, viz., 32 mph. Furthermore, an internal check can be made on these measurements, because, as stated in Sec. 6·7, classical hydrodynamics furnishes relations between wavelength, wave velocity, and period. The wavelength calculated by Eq. (100) from the measured wave velocity is 260 ft, in excellent and probably somewhat fortuitous agreement with the observed value.

One must therefore conclude that in this case the scatterers are, in effect, not uniformly distributed but appear concentrated in definite regions of the waves.[1] That this concentration may be only apparent is clear from a consideration of the length-to-height ratio of the waves. As stated in Sec. 6·7, this ratio varies between 7/1 and 20/1 for waves in a "sea." The crest of the wave will thus obscure the trough, diffraction being neglected, unless the grazing angle is greater than 6° even for the flatter waves. As the grazing angle was about 1°, it is obvious that even with diffraction a large portion of the possible scattering region must not have been illuminated by the incident beam. This "shadow" effect will obviously enter only at low angles and whenever the pulse length is less than the water wavelength and width of the beam is less than the crest length. These conditions are almost never to be expected when an airborne system is being used, when the pulse length is 1 sec or more and the beamwidth several degrees, or when the sea is short-crested. For such cases the concept of a cross section per unit area is therefore meaningful and will be freely used. The cross section per unit area can, however, be a function of position over the sea surface because of local variations in winds and currents: A striking example of such variation is shown in Fig. 6·11 which is a portion of a photograph of a 5-mile PPI on a 1.25-cm system operating with a 0.75-μsec pulse. The system was located on Fishers Island, off the Connecticut coast, and the sea surface at the time was calm, with no more than slight chop. Almost all the signal is sea echo, expect for Fishers Island itself and near-by islands in the upper half of the oscilloscope. The striated appearance is fairly marked and is possibly due to freak currents. It is definitely not due to a shadow effect, for there were no large waves present. The regions making up the striations were much larger than the pulse packet, and within each region one can still speak of a scattering cross section per unit area. Such marked local variations were found only for relatively calm sea states.

[1] It should be pointed out that the waves observed were too far out to be breaking. In fact these waves, which actually came from the storm center off shore, were not very obvious by visual observation from shore, being masked by a cross-wind "chop."

6·9. Frequency Dependence of Sea Echo.—Two possible mechanisms for sea echo have been frequently suggested. According to one view, back scattering by the droplets of spray present in whitecaps and other breaking water constitutes the major portion of the echo. The other view proposes that the observed signal is the result of diffraction from the corrugated surface of the sea or even specular reflection if a portion of a wave happens to be favorably oriented. These two hypotheses differ widely in their predictions for the frequency dependence of the cross section.

FIG. 6·11.—Photograph of the PPI on a 1.25-cm system, showing the striated appearance of the sea echo. Range markers are 1 mile apart. The bright line is the north heading.

By the droplet theory, if the radii of the droplets are much less than $\lambda/2\pi$, only the dipole terms in the back-scattering cross section are important, and the frequency dependence is given by the familiar Rayleigh law, namely, as λ^{-4}. Because of the interference lobes of the field above the sea surface, the frequency variation of the received echo should be even faster, possibly as much as λ^{-8}. The alternative mechanism predicts a much slower variation. If only a geometrical reflection is involved, then the cross section is, of course, independent of frequency. Even on the basis of a diffraction phenomenon, if the dimensions of the surfaces involved are comparable to or greater than λ, the dependence should not be faster

than λ^{-2}. In either case the differences with the droplet theory should be sufficiently marked to be measurable experimentally.

A series of measurements was carried out at the Radiation Laboratory to make just such a comparison.[1] The experiment will be discussed in detail in order to show some of the difficulties involved in the study of sea echo and the precautions that must be taken to obtain significant results.

Three truckborne systems were used, operating on wavelengths of 9.2, 3.2, and 1.25 cm, respectively. The systems, whose characteristics are summarized in Table 6·3, were experimental ones, specifically designed for accurate measurements of signal strength. (The majority of the measurements were made with the 3- and 9-cm systems, as difficulties with the 1.25-cm system prevented it from being used as often or with the same accuracy as the other systems.) A few details of the systems are worthy of

TABLE 6·3.—CHARACTERISTICS OF RADAR SYSTEMS

System No.	Wavelength, cm	Antenna dimensions, in.	Beamwidths, degrees		Pulse length, μsec	Peak power, kw.
			E-plane	*H*-plane		
XT-3	9.2	48 (paraboloid)	6.8	8.9	2,1,0.5	35
XT-2	3.2	(*a*) 48 (paraboloid)	1.9	1.9	1,0.33,0.15	20
		(*b*) 18 (paraboloid)	4.8	4.8	1,0.33,0.15	20
T-14	1.25	48 × 6 (parabolic cylinder)	0.8	6.1	0.75,0.35,0.15	10

notice. The direction of the polarization in each system could be changed rapidly, in the particular case of the 1.25-cm set, the change was effected by inserting a metal-strip half-wave plate in front of the antenna.

It is a desirable precautionary measure to have all systems measure the echo from the same area of the sea surface in case the assumption of homogeneous distribution of scatterers is not satisfied. Most of the measurements on 3.2 cm were therefore made with an 18-in.-diameter antenna, whose beamwidth nearly matched that on 9.2 cm. In addition, the pulse length used was in all cases either 1.0 or 0.75 μsec. It should be noted, however, that the pulse length need not be accurately known in practice, as Eq. (96) for the received power can be rewritten in the form

$$\overline{P}_r = \frac{p_t G^2 \lambda^2}{(4\pi R)^3 \Omega} \Phi f^4(\theta) \frac{c}{2} \sigma^0(\theta), \tag{101}$$

where Ω is the pulse recurrence frequency and p_t is the average transmitted power. (It is just this average power, and not the peak power P_t, that is measured directly by the usual test equipment.) The pulse-recurrence frequency was crystal-controlled and hence known very accurately.

[1] For a more comprehensive discussion see H. Goldstein, *Phys. Rev.*, **70** 938 (1946).

The instantaneous sea-echo intensity fluctuates rapidly and over wide limits. In order to meaure σ^0, the time average of the received echo must be obtained. A number of methods have been devised (see Sec. 6·10), but an electrical averaging technique was chosen as the most direct and objective procedure. A narrow section of the video signal (shorter than the pulse length) is selected by a gate and passed through an amplifier having a long time constant. To smooth out all fluctuations it was found that the averaging time had to be as long as 30 sec. The reading of the output meter on the amplifier was converted to r-f input level by means of a calibrated r-f signal generator.[1]

The systems were set up at a number of locations on the New England coast at elevations of 70 and 120 ft above mean sea level. Because of the relatively low elevation the grazing angle of the beam with the horizontal was restricted to less than 2° and usually had to be below 1° to avoid shallow water close to shore. Unfortunately sea echo at these elevations was not very extensive, never extending farther than 8000 yd. Therefore, for seas calmer than "moderate," measurements could often be made only at angles greater than 1° and may not be representative of open-sea conditions.

Possible sources of error in the measurements are of two types. (1) There may be errors in the calibration and constants of the system. On the 9.2- and 3.2-cm systems, calibration was made frequently and by several well-known methods. The limits of error in the values of received power, transmitted power, and antenna gain were ± 0.5 db for each. In addition there were a number of minor corrections to be applied in determining σ^0: loss in the radome, correction for the vertical polar diagram, TR switch recovery losses (at close ranges only), etc. The uncertainty in each of these corrections was small, but the cumulative effect is not negligible. The probable systematic error in the absolute measurement of σ^0 on 3 or 9 cm is, therefore, estimated to be 2 db. Because of difficulties in obtaining consistent absolute measurements of received and transmitted power and owing to additional losses from water vapor attenuation, probable systematic error on 1.25 cm is estimated to be 4 db. (2) In addition to such systematic errors there are random errors arising from the inherent scatter in the data. The value of the cross section is not strictly constant with time but fluctuates as the sea surface changes and the number and size of the scatterers varies. Some of this fluctuation is smoothed out in the electrical averaging process, but the slower variations (over periods of minutes) are not affected. Furthermore there is generally little correlation

[1] Strictly speaking, this procedure yields an average of the *video* signal, which differs from the average of the signal at the input because of the nonlinearity of the receiver. A correction can be calculated from the measured receiver law and the probability distribution of the instantaneous sea-echo amplitudes. Conditions were always so adjusted that this correction was never greater than 1 db and was roughly the same in all systems (see H. Goldstein, *op. cit.*).

between the scatter on different frequencies. These random errors were reduced somewhat by repeated measurements made in rapid succession. The scatter in the values of such a sequence was usually 1 to 2 db.

The variation of the cross section from 9.2 to 3.2 cm is expressed by the ratio

$$\frac{\sigma^0(3.2)}{\sigma^0(9.2)} = \beta,$$

usually given in decibels, that is, $10 \log_{10}\beta$. Measured values of the ratio varied from -8 to $+$ 18 db. A definite difference was found between the two polarizations, the ratio being consistently smaller on vertical polarization. The results also show a correlation with the roughness of the sea. Unfortunately, it was not possible to measure the sea state beyond some rough and unsatisfactory visual observations (along with surface wind data). The cross section at any given angle, wavelength, and polarization, however, should be a function solely of the parameters lumped under the term *sea roughness*. This cross section can therefore serve as a dependent variable, or scale, to illustrate any correlation of the frequency dependence with sea roughness.

The 110-odd observations made are spread rather thinly over a variety of sea states and grazing angles. The observed ratios were divided into three broad ranges of grazing angle and were then further separated into groups according to the measured value of the cross section on 9.2 cm for horizontal polarization. The average ratio for each group is shown in Table 6·4. For the middle group an attempt has been made to translate the cross-section scale into the conventional sea-state scale using visual observations and predictions based upon the observed winds and fetch. The very rough correspondence so obtained between the scales is shown in the table. This correspondence cannot necessarily be carried over to the other groups because the cross section varies with angle in a way that is not well known and that probably changes with sea state.

The values of β listed in the table are to be compared with the figure of $+18.5$ db, corresponding to a variation as $1/\lambda^4$, or with $+37$ db, corresponding to $1/\lambda^8$. With the exception of two points, the experimental values are much smaller, well outside the expected error. The results thus do not support the spray-droplet theory.

The data on the ratio $\sigma^0(1.25)/\sigma^0(3.2)$ are too few to permit any detailed analysis. They cover angles from 0.45° to 0.90° and have been measured only for moderate seas on horizontal polarization. The average value of the ratio was found to be about $+5$ db for these conditions, compared with $+16.5$ db predicted by a $1/\lambda^4$ variation.

An important by-product of this investigation is information on the dependence of the cross section on the polarization. This information is presented as the ratio of the cross section for vertical polarization to the

TABLE 6·4.—EXPERIMENTAL VALUES OF THE FREQUENCY DEPENDENCE OF SEA ECHO

$\theta \geqslant 1.0°$

Interval of $\sigma_\lambda^0(9.2)$, db	No. in interval		β, db	
	Hor.	Vert.	Hor.	Vert.
−40 to −50	2	0	+ 8.5	
−50 to −60	0	0		
−60 to −70	5	4	+13	+ 2
−70 to −80	1	2	+18	+11

$0.65° \leqslant \theta < 1.0°$

Interval of $\sigma_\lambda^0(9.2)$, db	No. in interval		β, db		Approx. sea state
	Hor.	Vert.	Hor.	Vert.	
−30 to −40 db	6	4	− 2	− 5	Rough
−40 to −50	15	11	+ 3	− 1	Moderate
−50 to −60	5	7	+ 9	+ 3	Slight
−60 to −70	2	0	+18	....	Smooth

$\theta < 0.65°$

Interval of $\sigma_\lambda^0(9.2)$, db	No. in interval		β, db	
	Hor.	Vert.	Hor.	Vert.
−30 to −40 db	6	6	0.0	−5
−40 to −50	11	11	+4	0.0
−50 to −60	5	4	+6	+3

cross section for horizontal polarization $(\sigma_v^0/\sigma_h^0)_\lambda$ usually measured simultaneously on 9.2 and 3.2 cm. The data should be almost completely free from systematic error, as the ratio is independent of the absolute calibration of the system. Almost the only sources of systematic error are changes in the antenna characteristics and faulty calibration of the signal-generator attenuator, both of which are deemed negligible. There is still a random error, of course, arising from the fluctuations in the values of σ^0. If anything, this random scatter tends to be larger than that for the frequency-dependence data because of the unavoidable time difference between measurements on the two polarizations.

The observed values of the ratio vary between −8 and +22 db. A definite correlation exists between the values of the ratio on 9.2 and 3.2

cm, as illustrated by Fig. 6·22 in Sec. 6-12. There the ratio for $\lambda = 9.2$ cm is plotted as ordinate, and the ratio for $\lambda = 3.2$ cm as abscissa. The circles represent experimental values for grazing angles between 0.65° and 1.35°. Although the points show considerable scatter, the presence of a correlation is clearly indicated. High values of the ratio on 9.2 cm are almost always accompanied by high values on 3.2 cm. Similarly, low values on 9.2 cm are paired with low values on 3.2 cm. The ratio on 3.2 cm, it will be noted, is almost always lower than the simultaneous ratio on 9.2 cm. In fact, although vertical polarization almost always returns the stronger echo on 9.2 cm, the reverse is frequently the case on 3.2 cm.

An analysis of the data also indicates a correlation of the "polarization ratio" with sea state. Large values of the ratio are found only when the sea is calm; in fact, it is then one of the most striking characteristics of the echo. As the sea becomes rougher, the ratio steadily decreases, there being little difference between the two polarizations for very rough seas on 9.2 cm. Under the same conditions the echo on 3.2 cm, however, is considerably stronger on horizontal polarization.

Very few observations of polarization dependence are available on 1.25 cm. The values obtained indicate that the echo is about the same for both polarizations, with a scatter of roughly ± 3 db.

A word might be said about the absolute magnitude of the cross section. From Table 6·4 the low-angle cross section on 9.2 cm for horizontal polarization is seen to vary between -30 and -70 db. These figures mean that the echo from each square foot of the sea's surface can be accounted for by a cross section of between 10^{-3} and 10^{-7} sq ft. The values seem low, and this in itself mitigates against specular reflection playing an important role. The radar beam usually covers a considerable number of square feet, however, and the total cross section can be rather large. Thus, in a system with a pulse length of 1 μsec and a beamwidth of 3°, the beam covers an area of 3×10^6 sq ft at a range of only 40,000 yd (23 miles). The total sea-echo cross section under these conditions then varies between 3000 and 0.3 sq ft. The larger figure is quite sufficient to mask the presence of a large aircraft or a surfaced submarine.

6·10. Measurements of the Properties of Sea-echo Cross Section.— Section 6·8 has shown that except for certain rare conditions, the concept of a cross section per unit area for sea echo σ^0 is meaningful. The present section attempts to summarize the available experimental information on the properties of σ^0 and its dependence on various parameters. Table 6·5 lists the attendant circumstances for nine investigations on sea echo. No attempt is made to list all the known researches. A good many investigations, especially older reports and those in the nature of system service trials, have been omitted. It is believed, however, that the table includes all reported investigations making a significant contribution to our knowledge concerning the sea-echo cross section.

The items listed in Columns 1 to 4 of Table 6·5 need no explanation. Column 5 indicates the parameters as a function of which σ^0 has been measured, that is, grazing angle of incidence θ, wavelength λ, polarization p, and state of the sea S. In Column 6 is stated the method of measuring $\overline{P}_r$. All investigations except II, in which only the maximum range of sea echo was measured, used some kind of signal generator. The measurement of sea-echo intensity presents some peculiar difficulties because of the rapid fluctuations of the signal, and various procedures have been devised in order to obtain some definite quantity to measure. In one investigation the highest, or peak, signal occurring in a given time interval was observed. This is not believed to be a good choice; Secs. 6·11 and 6·18 will show the probability distribution for the instantaneous signal amplitude to be such that there is a finite probability for *any* amplitude, no matter how high. A more precisely defined level is the amplitude exceeded by a specified fraction of the pulses, usually in the neighborhood of 2 to 5 per cent. In practice it is found that this level can be reproducibly estimated by visual observation of the A-scope. This method is denoted as "mean peak signal." Sea-echo intensity has also been measured by finding the r-f level of a signal-generator pulse that can just be detected in the sea echo, analogous to the determination of minimum discernible signal in noise. The level so obtained is reproducible for one observer and system but, as is well known, depends upon pulse length, pulse-recurrence frequency, method and size of display and upon the observer. In addition, the level, unlike the "mean peak signal," is not connected in a known way with the average intensity. By far the best method of measurement, and the one that is least subjective, involves averaging by electrical means, as described in Sec. 6·9. This method has the additional advantage that it yields exactly the quantity most suited for theoretical discussion, viz, $\overline{P}_r$.

Most of the investigations indicate the state of the sea or the height of the waves. Column 7 indicates the type of observation upon which such statements are based.

Variation of Maximum Range with Height. The way in which the maximum range of sea echo[1] on the indicator varies with the height of the system is of obvious practical importance. The range enters into the basic formula [Eq. 96] for the received intensity at low angles in two places: in the R^3 term in the denominator and as a parameter of σ^0 in the form $z/R = \theta$. Thus, Eq. (96) can be rewritten

$$\overline{P}_r = \frac{P_t G^2 \lambda^2 \Phi f^4(\theta)}{R^3} \frac{\tau c}{2} \sigma^0\left(\frac{z}{R}\right). \tag{102}$$

Until the specific function $\sigma^0(z/R)$ is known, the variation of maximum range with height cannot be predicted. For discussion purposes, it is con-

[1] Defined as the range at which the sea echo is lost in noise.

TABLE 6·5.—A SUMMARY OF THE SIGNIFICANT INVESTIGATIONS ON SEA ECHO

Number (1)	System height, ft (2)	λ (3)	θ, deg. (4)	Properties of σ^0 studied (5)	Method of measurement (6)	Type of observations on sea state (7)	Organization and reference (8)
I	60 and 20	9,3	< 2	θ,p,S	Signal generator, r-f; peak signal	Wind speed, visual estimate	RL: O. J. Baltzer, W. M. Fairbank, and J. D. Fairbank, RL Report No. 536, Mar. 14, 1944
II	250–13,000 (airborne)	10.3	< 5	θ	Maximum range	Wind speed, visual estimate	RL: Unpublished Memoranda
III	Airborne	3	5–90	θ,p,S	Signal generator, r-f; mean peak signal	Visual estimate	RL: E. W. Cowan, RL Report No. 870, Jan. 10, 1946
IV	70 and 120	9,3,1	< 1.5	p,λ,S	Signal generator, r-f; electrical averaging	Wind speed, visual estimate	RL: H. Goldstein, *Phys. Rev.*, **70**, 938 (1946); see also Sec. 6·9
V	Airborne	10	< 5	θ,S	Signal generator, r-f; minimum discernible	Wind speed, visual estimate	RL: Unpublished
VI	500	10,3,1	0.5–5	θ,λ,S	Signal generator, i-f; mean peak signal	Measurements from submarine, wind speed	TRE: A. L. Cockroft, H. Davies, and R. A. Smith, TRE Report No. T1933, September 1945*
VII	3000–10,000 (airborne)	3	10–90	θ,S	Signal generator, i-f; (indirectly); mean peak signal	Visual estimate	TRE: G. S. R. Maclusky and H. Davies, TRE Report No. T1956, Nov. 20, 1945*
VIII	3500 (Mt. Snowden)	10	0.8–2	θ,S	Signal generator, r-f; electrical averaging	Not stated, probably visual estimate	ASE: H. Bondi, ASEE Witley, Interim Report No. Mk38/45 Feb. 16, 1945
IX†	150 and less	10,3.3	< 1.2	θ,λ,S	Electrical averaging	Not stated, probably visual estimate	ASE: T. Gold and W. Renwick, ASEE Witley, Report No. XRC, Sept. 8, 1946

* Also see H. Davies and G. G. Mcfarlane, *Proc. Phys. Soc.*, **58**, 717 (1946).

† Information on this investigation was unfortunately received too late for inclusion in the discussion in the text.

venient to assume that σ^0 has the form $(z/R)^n$. Equation (102) then can be solved for the range;

$$R_m \propto \overline{z^{n+3}}. \tag{103}$$

As long as n is positive, the maximum range will increase more slowly than the height.

This procedure can obviously be reversed; from measurements of the variation of range with height the functional dependence of σ^0 on θ can be deduced; it will, in fact, be proportional to the cube of the maximum range. As the measurements do not require any additional test equipment beyond the normal indicators, the method is tempting. Unfortunately, what con-

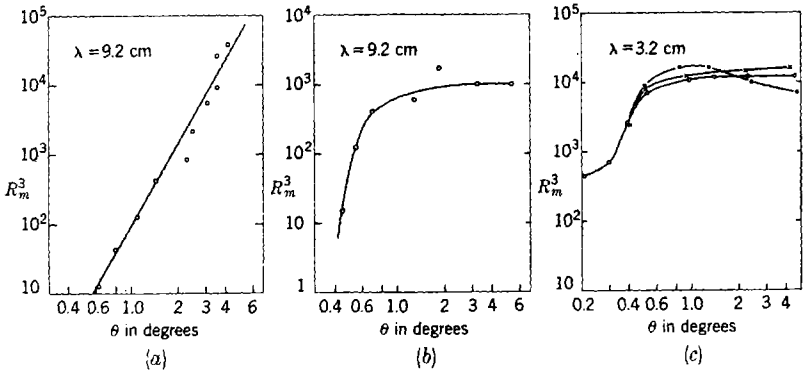


FIG. 6·12.—Three typical curves of maximum range of sea echo vs. grazing angle, replotted from Reference II, Table 6·5. The quantity R_m is in statute miles.

stitutes maximum range is not clear, and the readings depend upon the observer and the type of indicator. Furthermore, since $\sigma^0(\theta) \propto R_m^3$, small errors in R_m result in large errors in σ^0; or to put it another way, the variation of maximum range with height is not a sensitive function of σ^0.

Nevertheless, because of its simplicity, the method has been applied frequently, usually in the course of field trials of new systems. The results obtained in Reference II of Table 6·5 are typical and probably represent the most careful measurements made in this way.[1] Two commercial 9- and 3-cm systems, the ASE and ASD respectively, were used, mounted in separate light airplanes. The only modification of the systems consisted in adding a 5-in. A-scope. Unfortunately, almost all the measurements were made for relatively calm seas and all for horizontal polarization only.

Three typical curves are shown in Fig. 6·12 where the logarithm of R_m^3 is plotted against the logarithm of the angle θ. In all cases, $\sigma^0(\theta)$ rises very

[1] Other references are G. A. Garrett and K. L. Mealey, "Photographic Polarization Tests," RL Report No. 382, May 1943; "Trials of Mk XI ASV (3 cm)," RNAS Report No. 707/1/2, March 1945; "ASD in a Ventura," TRE Report No. T1651, September 1943.

rapidly at low angles, approximately as θ^n where n is about 3 or 4. Occasionally, as in Fig. 6·12*a*, this rapid rise continues to the highest observed angles. More often, as in Fig. 6·12*b*, the initial rate of increase is not maintained, and there appears a definite tendency for saturation; hence the cross section eventually becomes substantially independent of angle. This tendency is even more pronounced in the curves for $\lambda = 3$ cm.

The difficulties of this method are well illustrated in the curves of Fig. 6·12*c* for 3 cm, where the solid and open circles represent simultaneous measurements by two observers in the same airplane and the crosses represent measurements made by a third observer from oscilloscope photographs taken at the same time. According to one observer, $\sigma^0(\theta)$ decreased by a factor of 2 in going from $\theta = 1°$ to 5°, whereas, according to another, it increased by a factor of 1.4 in the same interval.

Obviously, the method of maximum range can furnish only a general idea of the dependence of σ^0 on θ; more painstaking procedures must be used to obtain accurate information.

Variation of σ^0 with θ. Eight of the nine investigations listed in Table 6·5 give data on the variation of σ^0 with grazing angle. As has already been seen, however, all are not equally reliable, nor are many of them at all comparable. Thus, in none of the American work with airborne systems was attention paid to obtaining favorable oceanographic conditions. In most cases the winds were not high, and the fetch was almost always 50 miles or less. On the other hand, because the prevailing winds are on-shore in Britain, the English work was probably done with sea states rougher than any encountered in the airborne work in this country; this may explain some of the apparent contradictions found. Again, the elevation of land-based systems is usually far less than those attained in airborne work, and the normal range is thus far smaller. Therefore, in order to obtain sufficient sea echo to work with, the land-based measurements are usually made in weather unsuited to aircraft operation. In any case, it is risky to try to find the variation of σ^0 with θ from the land; too often one merely measures the change in the sea roughness as shore is approached. Trustworthy data on the angle dependence of σ^0 can probably be obtained only from aircraft or from ships in rough seas, well away from shore.[1]

The investigations are rather sharply divided into two classes: low angle, in which the grazing angle is less than about 5°, and high angle, where θ varies from 5° up to normal incidence. These two classes will be treated separately, especially as it seems likely that two different mechanisms are involved.

Two of the researches listed in Table 6 5, III and VII, fall into the latter class, both on 3·2 cm. In III, a night-fighter system (AIA) was used, to

[1] Difficulties arising from nonuniformity in the sea surface were also encountered in airborne measurements as a result of the airplane's ability to cover large distances rapidly, but with proper planning of flight procedure such difficulties can be eliminated.

which considerable calibrated r-f test equipment had been added so that absolute measurements could be made. Data were taken for both polarizations, employing the "mean peak signal" criterion. A noteworthy feature of this investigation was that the range was kept constant as the angle was varied. In Fig. 6·13 the data for the roughest sea measured (yet for which the wind speed was below 12 mph) have been replotted to show the cross section in absolute measure. The range of variation is extremely large, almost 60 db from 15° to normal incidence.

In the corresponding English investigation a similar radar system and criterion for signal level were used, but only horizontal polarization was available. Also, the measurements were made by calibrating the gain control with an i-f signal generator; thus absolute figures could not be obtained. The range of sea states covered was much greater, however, than in Reference III. Figure 6·14 is a replot of only two of the many curves given in the original report. Figure 6·14*a* was obtained for a slight sea (Douglas 2), while Fig. 6·14*b* is listed as very rough (Douglas 5). The rate of increase of σ^0 with θ and hence the total range of values are much less than in Reference III. The dotted lines show the function $\sin \theta$, which represents the θ dependence of σ^0 fairly well up to grazing angles of 40° to 60°. The general level of σ^0 changes only slightly with sea state, in fact, the cross section at normal incidence is constant to within $\pm$ 3 db. Below 45°, σ^0 increases by only 10 db from the calmest to the roughest seas. In comparing these two investigations it should be emphasized again that the description "calm seas" in Reference VII probably means a rougher sea than any in III.

Of the five investigations at low angles (I, V, VI, VIII, and IX), the least trustworthy is probably I, because of the unsatisfactory criterion for signal level (peak signal in 10 sec) and because measurements were made very close to shore. The spread in the data is considerable, but the results can be roughly represented by $\sigma^0 \propto \theta^n$, for $\theta < 1°$, with the exponent n varying[1] between 1.3 and 4.0. There is some indication that the exponent tends to be less on 3 cm than on 9 cm. No significant variation of the exponent with sea state is noticeable.

Land-based systems were also used in Reference VI, but elevation was much higher, 500 ft, and the location (Llandudno in North Wales) such that the fetch was large. As a result, the sea-echo intensity was sufficiently high for measurements to be made at much greater distances from shore. Data were obtained from two systems, both on 1.25 cm, but with different beamwidths and pulse lengths, and from two others on 3.2 and 9 cm, respectively, all with horizontal polarization. Signal measurements

[1] Similar results were obtained at an almost identical location using the radar systems described in Sec. 6·9 but with the averaging of $\overline{P}_r$ obtained by long-exposure photographs of the A-scope.

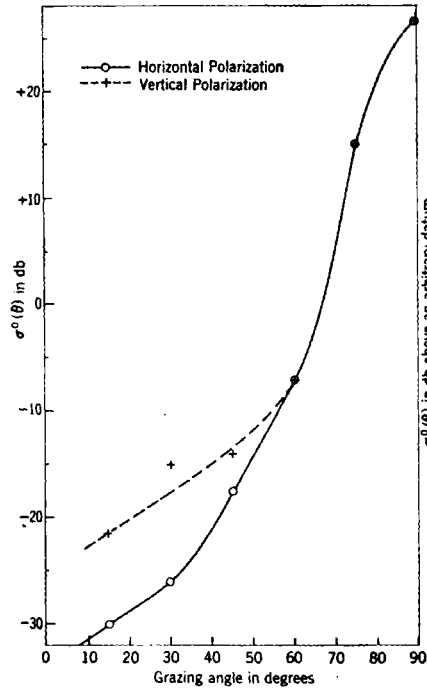


FIG. 6·13.—Sea-echo cross section per unit area on 3.2 cm for high grazing angles as replotted from Reference III of Table 6.5.

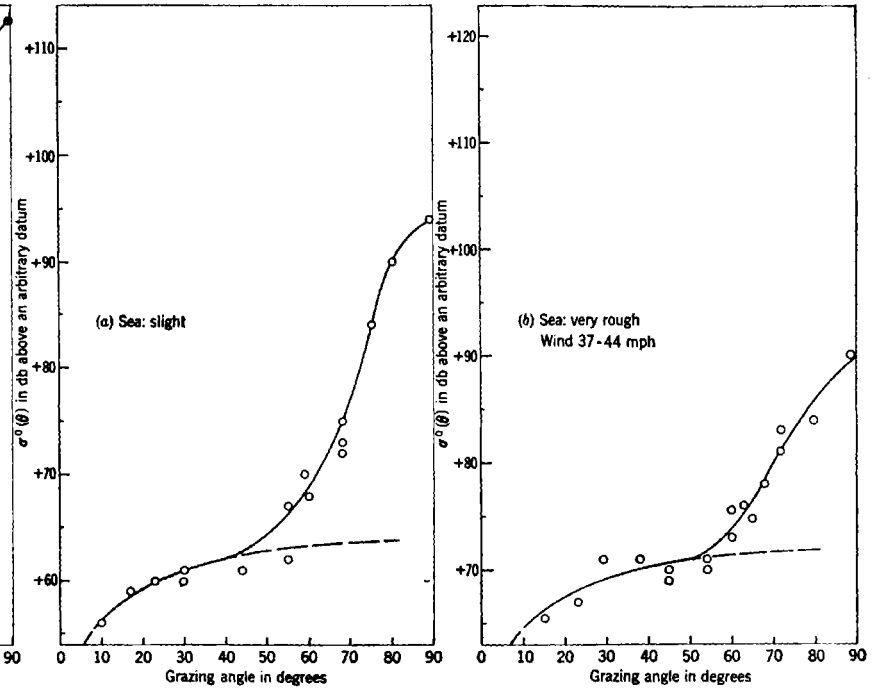


FIG. 6·14.—Two typical curves of sea echo cross section, in arbitrary units, on 3.2 cm for high grazing angles as replotted from Reference VII of Table 6.5. The solid curve and circles are for horizontal polarization. The broken curve is $\sigma^0 \propto \sin\theta$.

were made with an i-f signal generator and calibrated gain controls, but an absolute calibration for 1.25 and 3.2 cm was obtained by measuring the echo from metal spheres. Values of σ^0 on the two 1.25-cm systems differed consistently by 4 db, a fact that may reflect the errors involved in the calibration. This is the only sea-echo investigation for which actual measurements of wave height were obtained. These were made on a submarine sufficiently submerged to be unaffected by the waves, with the wave height read through the periscope by reference to a scale lashed to the superstructure.

All the data on 3 and 1 cm are very well represented by a linear dependence of σ^0 on θ. Figure 6·15 reproduces two of the graphs shown in the original report, in which $\overline{P}_r$ is plotted against the logarithm of the range. The straight line indicates the $1/R^4$ variation that corresponds to $\sigma^0 \propto \theta$. All the curves given in the report can be fitted by a power-law variation of $\overline{P}$, with R, in which the exponent lies between 3.5 and 4.5 at the extremes. The scatter of the points is much smaller than is usually encountered in investigations of this nature. It may be noted here that the ratio σ^0/θ was remarkably independent of wave height for waves above 2 ft high. The absolute values of σ^0 accord well with corresponding measurements of Reference IV.

Because of poor functioning of the system, the data on 9 cm were much less reliable and no absolute calibration was obtained. The one curve reproduced in the report, at a wave height of 6 ft, definitely indicates, however that σ^0 rises faster than θ for small angles.

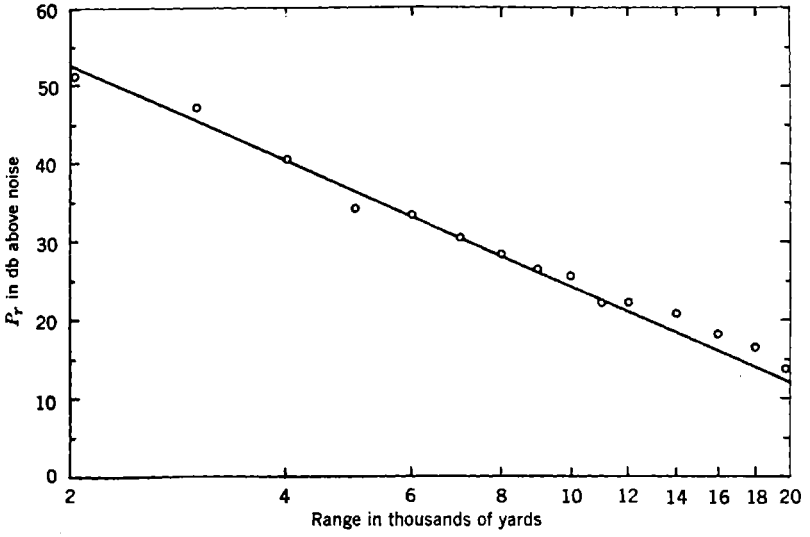
An extreme case of land-based measurements is represented by Reference VIII, as the radar was located atop Mt. Snowden, North Wales, at an elevation of 3500 ft, and the data should be free of some of the uncertainties characteristic of land-based work. Only one system, on 10 cm and with horizontal polarization, was used. Averaging of the signal was performed electrically as in Reference IV.

It is stated in this report that the results are well fitted in all cases by the following form of σ^0

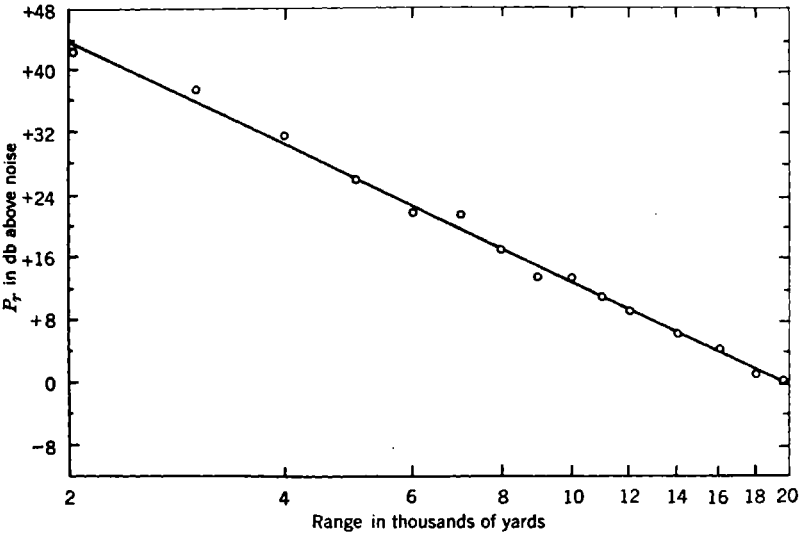
$$\sigma^0 \propto e^{-\frac{\theta}{\theta_0}}, \qquad 0.8° \leqslant \theta \leqslant 1.9°, \tag{104}$$

where θ_0 is a constant whose value is about $3.8° \pm 1°$. Unfortunately, only an interim report has been issued, and actual curves are not shown. Too much emphasis should not be put on the various analytic forms in which the θ-dependence of σ^0 is expressed. Thus, the above exponential dependence, over the limited range of angles, can be replaced by the power-law formula $\sigma^0 \propto \theta^{3.3}$ with an error of less than $\pm$ 1 db, which is undoubtedly well within the experimental scatter.

The most complete set of data for 10 cm is given by Reference V. The measurements were made in the course of design flight trials for a high-



(*a*) Wavelength 3.2 cm Wave height 6 ft



(*b*) Wavelength 1.25 cm Wave height 4 ft

FIG. 6·15.—Two typical plots of sea-echo intensity as a function of range. The circles are experimental points and the straight lines indicate $1/R^4$. Taken from Reference VI, Table 6.5.

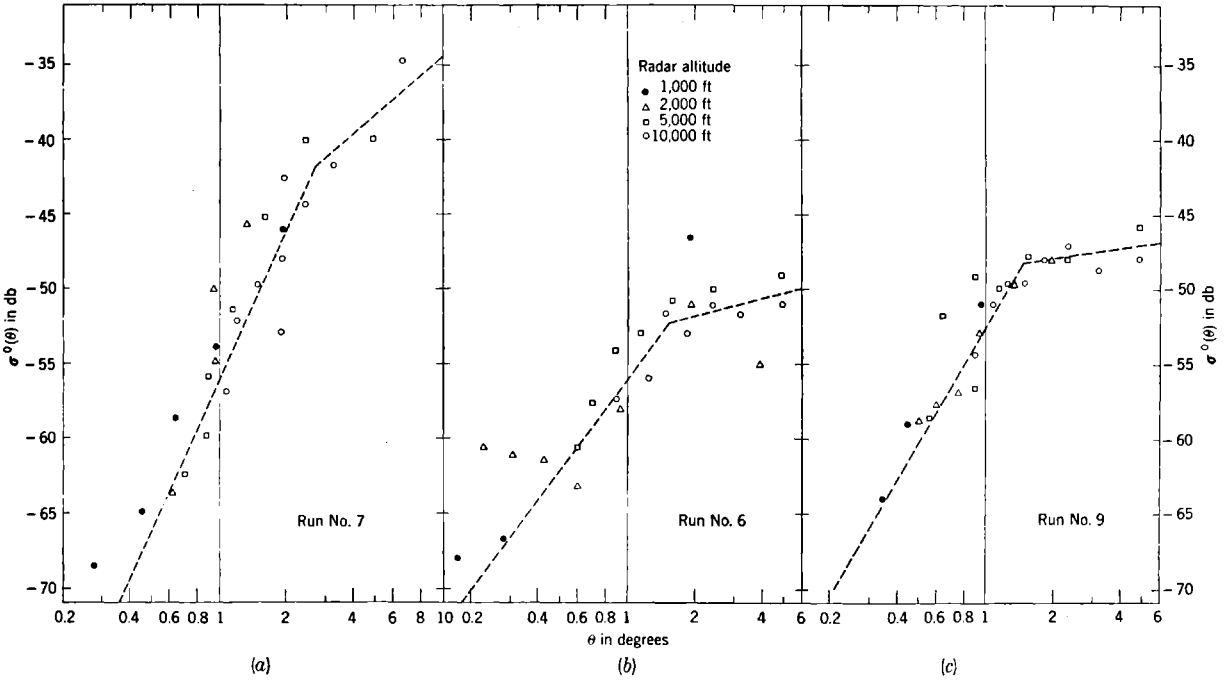


FIG. 6·16.—Typical plots of σ^0 (approximately in absolute measure) vs. grazing angle θ. Calculated from the data of Reference V, Table 6.5.

power 10-cm airborne system operating with horizontal polarization. In a typical run the sea-echo intensity was measured as a function of range by the "minimum-discernible-signal-generator" technique, usually for several altitudes. About ten such runs were obtained over a period of several months.

The work is somewhat marred by the almost complete lack of attention to the proper oceanographic conditions (admittedly often enforced by the limited range of the aircraft) and by the large number of observers with varying degrees of training (with the result that separate runs are often not comparable).

TABLE 6·6.—PROPERTIES OF THE CURVES OF σ^0 VS. θ AS OBTAINED FROM REFERENCE V, TABLE 6·5

Run No.	$\sigma^0(1°)$, db	Power-law exponent		θ_0, degrees
		$\theta < \theta_0$	$\theta > \theta_0$	
1	−74	3.6	0	4
2	−63	2.4	0	4
3	−62	2.5	0	4
4	−60	2.8	0.7	2.1
5	−59	3.1	1.0	2.0
6	−56	2.0	0.3	2.0
7	−54	3.6	1.3	2.7
8	−52.5	2.1	0.0	2.0
9	−52.5	2.6	0.2	1.5
10	−42	3.8	0.4	1.1

The results have been recalculated and plotted as a function of the grazing angle θ. In this form the curves obtained at several altitudes should coincide, ideally. Figure 6·16 reproduces three of the plots. The condition of the sea surface cannot be determined from the information available, but it seems likely that Fig. 6·16*a* refers to a slight sea and Fig. 6·16*b* and *c* to a moderate sea (no sea state rougher than moderate was encountered.) The separate altitudes at which the measurements were made are indicated by different symbols. The scatter of the points is considerable, as might be expected, but the data for the various altitudes seem well intermingled, and the deviations appear random. Exceptions occur only at low altitudes and angles, and can probably be attributed to the low signal-to-noise ratio under these conditions.

There is obviously considerable latitude in fitting the points with a smooth curve, but for convenience the general behavior may be represented by segments of two power laws, which in the coordinates used become straight lines and are shown by the dotted lines in Fig. 6·16. This type

of representation is also suggested by some theoretical considerations. All the ten plots have been fitted in the above manner,[1] and the exponents of the two power laws are listed in Table 6·6. Also listed are the angles at which the two segments intersect and the value of the fitted curve at $\theta = 1.0°$. This last is given in absolute measure and was obtained in a crude fashion from assumed noise figures and constants of the system. The values should therefore be depended on only for orders of magnitude. It should also be remembered that the calculation did not include the unknown and variable difference between the minimum discernible and average signal level. The order of increasing values of $\sigma^0(1°)$ as given in the table should correspond approximately to the order of increasing ea roughness. All values of the cross section are considerably smaller than those measured for rough seas in the experiments discussed in Sec. 6·9. This fact is another indication that the waves did not exceed moderate heights.[2]

Almost all the investigations described refer to somewhat different external conditions and cannot therefore be compared with each other. It is tempting, though not altogether legitimate, to try to summarize the various results and build a composite picture of the curve of σ^0 as a function of θ. Certainly it seems clear that at 10 cm the curve rises very steeply at low angles, being approximated by a power law whose exponent lies between 2 and 4. At somewhat higher angles σ^0 increases much more slowly; the curve lies between a constant value and a linear dependence on θ. The angle corresponding to the "knee" of the curve decreases as the sea becomes rougher (*cf.* Table 6·6), ranging from 5° for calm seas to less than 1° at very rough seas.

The initial rapid rise is presumably present at 3 cm also (as evidenced by Fig. 6·12*c*), but the angle at which the curve begins flattening off is apparently much smaller at 3 cm than at 9 cm for the same sea state. The guess might be hazarded that the trend is continued in the same direction at 1.25 cm, but the data for this wavelength are very fragmentary.

The cross section does not continue to rise slowly indefinitely. At some higher angle a new scattering mechanism appears to become important and σ^0 begins to rise rapidly again, reaching a maximum at normal incidence. The onset of this second rapid rise, at least for 3 cm, occurs at about 45° to 60° in rough seas and at 15° to 30° in calmer seas. The maximum values of σ^0 can become fairly large, greatly exceeding unity, indicating that some specular reflection is taking place.

[1] The fitting was at times more ambiguous than in the case of the three examples shown.

[2] *Note added in proof:* The data of Reference IX, limited to angles less than 1° and wave heights from 2 to 4 ft, are fitted well by a cross section σ^0 varying as $\theta^2 e^{-k/\theta}$, where k is 0.02 radians for $\lambda = 10$ cm and 0.01 for $\lambda = 3.3$ cm.

Variation of σ^0 with λ. The manner in which the cross section changes with wavelength is of great importance for theoretical considerations, but measurements of this quantity have been rare because of the difficulties involved in making absolute measurements.

Reference VI of Table 6·5 presents some data on the change of cross section between 3.2 and 1.25 cm for horizontal polarization. Two systems at 1.25 cm and one system at 3.2 cm were calibrated absolutely by means of a standard target consisting of a sphere suspended from a balloon. The ratio of the cross section at the two wavelengths $\sigma^0(1.25)/\sigma^0(3.2)$ was independent of θ in the range measured and independent of waveheight above 2 ft. The value obtained for the ratio under these conditions were +3 db using one of the 1-cm systems and +7 db using the other. As has been remarked, the difference of the values is an indication, perhaps, of the accuracy of the absolute calibration. The smaller value is stated to be the more reliable and is considered to be indicative of a variation as $1/\lambda$. Unfortunately, dependable absolute values on 9 cm could not be obtained, but there was some evidence that the ratio of the cross section on 9 cm to that on 3 cm decreased somewhat as the sea became rougher.

The results obtained in Reference IV have been discussed in detail in Sec. 6 9 and need be summarized here only briefly. With horizontal polarization it was found that the increase in the cross section in going from 9.2 to 3.2 cm varied between +18 db with relatively calm seas to −2 db when the sea was rough. When the polarization was vertical, over the same wavelength region, the cross section increased by 11 db for calm sea and *decreased* slightly (5 db) for rough sea. A few rough measurements were also made at 1.25 cm, horizontal polarization, and the cross section was found to be approximately 5 db larger than that at 3.2 cm.

The two investigations discussed barely overlap, but their results are not mutually inconsistent.[1]

References in the literature to sea echo at wavelengths greater than 10 cm are very meager. It would appear that the echo is not very strong outside the microwave region, although it is risky to generalize from such qualitative information, or rather lack of information. But if the cross section does not vary much with wavelength, one might expect sea echo to be as prominent on the large CHL (low-frequency) radar installations on the English coast as on the near-by microwave sets. As this is not the case, it seems probable that σ^0 decreases very rapidly with wavelength above 10 cm, at least for horizontal polarization.

[1] *Note added in proof:* The results of Reference IV on horizontal polarization also bear on this question. Strictly interpreted, they would indicate a wavelength dependence between 10 and 3 cm that is a function of angle, varying from a dependence as λ^1 around 1° to λ^4 at about 0.2° grazing angle. The range of sea states was too limited to indicate any variation with sea roughness.

Dependence of σ^0 on Polarization. The earlier qualitative studies of polarization effects all agreed that the cross section of sea echo on 10 cm was much greater for vertical polarization than for horizontal.[1] In the investigation of Reference I it was similarly found that for 9.1 cm and *calm seas,* the echo could be as much as 20 to 30 db greater for vertical than for horizontal polarization. The difference steadily decreased, however, as the sea became rougher; thus even for moderate seas the echo was about the same on the two polarizations. With a wavelength of 3.2 cm and for *calm seas,* the echo was greater for vertical polarization but the difference was not so great as for 9.1 cm. As the sea became rougher, the behavior on 3.2 cm paralleled that on 10 cm, and it was found that the ratio of intensities even reversed at moderate seas, with the echo becoming stronger on horizontal polarization.

This general picture has been confirmed by the investigation described in Sec. 6·9 (Reference IV), and the correlation between the behaviors on 9 and 3 cm has been worked out in greater detail (*cf.* Fig. 6·22 in Sec. 6·12). A few rough observations were also made on 1.25 cm, but no significant difference between the polarizations could be found. The only other known quantitative study of polarization effects is contained in Reference III. It is there stated that for 3.2 cm at high angles the echo was invariably stronger on vertical than on horizontal polarization, the ratio at times exceeding 10 db. This result is not inconsistent with the observations previously mentioned, as the sea condition during the investigation was never rougher than "slight."

Variation of σ^0 with State of the Sea. It is clear that the intensity of sea echo increases as the sea becomes rougher. Because of the difficulties of measuring objectively the sea state, as discussed in Sec. 6·7, quantitative information is, however, extremely rare. Even many obvious questions of a qualitative nature still remain unanswered. Thus, it is of great importance for theoretical discussions to know how the presence or absence of whitecaps at a given wave height affects sea echo, but no clear answer has yet been given.[2] In general it must be stated that very little progress has been made toward determining what parameters of the sea state, if any, are clearly correlated with sea-echo intensity.

The most reliable measurements of sea echo as a function of wave height are contained in Reference VI of Table 6·5, as the wave heights were actually measured from a submerged submarine. It will be remembered that the cross section was observed to be a linear function

[1] *Cf.* G. A. Garrett and K. L. Mealey, *op. cit.,* and E. M. Lyman and J. J. Hibbert, "Radar Target Contrast," RL Report No. 375, June 1942.

[2] In "Report on an Investigation into the Nature of Sea Echo," TRE Report No. T1497, May 1943, a single instance is cited in which the sea echo increased considerably when whitecaps developed because of a cross wind that did not influence the height of the waves. It would be highly desirable to determine if this fact is true in general.

of θ between 0.5° and 5° on $\lambda = 3.2$ and 1.25 cm. The ratio σ^0/θ is therefore considered to be an empirical constant. Figure 6·17 reproduces from this reference the plot of σ^0/θ vs. wave height for the three systems used. Above wave heights of 2 ft the cross section is substantially constant or increases only slightly with frequency. Below 2 ft there is good evidence for a rapid variation of σ^0/θ with wave height. (The existence of a "saturation" effect at large wave heights is in accord with general qualitative impressions obtained in the course of the Radiation Laboratory

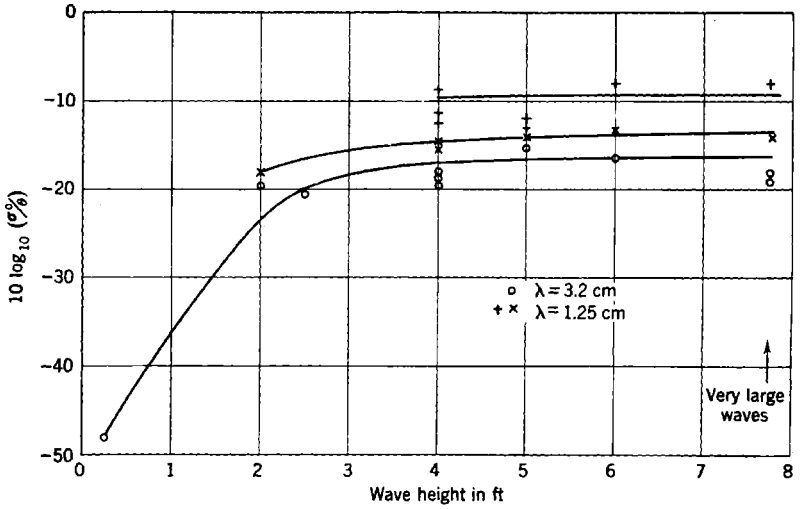


FIG. 6·17.—Variation of the ratio σ^0/θ at low grazing angles with wave height, as measured on three systems. Taken from Reference VI, Table 6.5.

investigations.) It is stated that the maximum value of σ^0/θ depended on wavelength as

$$\left(\frac{\sigma^0}{\theta}\right)_{\max} = \frac{\lambda_0}{\lambda}, \tag{105}$$

where $\lambda_0 \approx 0.08$ cm. As the measurements were of the "mean peak signal" this value is two to three times larger than for the true σ^0.

The behavior at 10 cm is even less well known. It is believed on the basis of qualitative data that a saturation effect is still present but that it occurs for much larger wave heights, in the neighborhood of 10 ft. The only quantitative data are given in Reference VIII, where it is stated that the echo increases about 10 db for every foot of wave height up to 6 ft and at roughly 8 db per ft from then on to a height of 14 ft, which was the maximum observed. It is not certain how these measurements were obtained.

Mention should be made of the variation of sea echo with the azimuth of the beam relative to the wave structure. All the investigations that deal with this variation (II, III, VII, VIII, and others[1]) agree with rare unanimity that the echo is strongest when the set is looking into the wind. The relative increase over the downwind value varies but is generally about 5 db. The phenomenon can be marked and has at times been proposed as a means of determining wind direction from aircraft, though with questionable reliability. Unfortunately, this universally observed phenomenon is amenable to explanation by almost all the theories of sea echo yet proposed and consequently sheds little light upon the nature of the scattering process. The lee side of a wave driven by the wind tends to be steeper than the windward side, and diffraction or specular reflection is expected to be greater when looking into the wind. Correspondingly, breaking waves will break over on the lee side, and the scattering from the increased quantities of spray and foam equally well account for the observations.

There is not such complete agreement about the intensity in the cross-wind direction. Generally it has been found to be intermediate between the values for upwind and downwind. Occasionally it is stated to be even smaller than for downwind. In Reference III both types of behavior are reported.

6·11. The Fluctuation of Sea Echo.—One of the most striking characteristics of sea echo as viewed on the A-scope is its rapid and irregular fluctuation, which gives the signal a family resemblance to ordinary receiver noise. This fluctuation is well illustrated in the photograph of the A-scope shown in Fig. 6·10 and in Fig. 6·18 which is a plot of the instantaneous intensity of sea-echo signal on 9.2 cm for a period of about 1 sec. Later sections of this chapter discuss the fluctuations of sea echo in detail as an example of the general subject of fluctuations of clutter signals. This section, however, will be concerned primarily with that information about the nature of the sea-echo scatterers which is furnished by a study of the fluctuations. As will be seen, this information is disappointingly meager.

The range or extent of the fluctuation is specified in detail by the probability distribution for the instantaneous echo intensity, that is, the fraction of signals whose intensity lies within a given interval. Such a distribution, which tells nothing about the frequency of fluctuation, is called the *first probability distribution*. The origin of the majority of sea-echo fluctuation is thought to be the varying interference between the echoes from scatters that lie within the region illuminated by the radar beam. As the scatterers move about, the relative phases of their echoes change and the total echo (which is the vector sum of the component echoes) varies correspondingly. It can then be shown (*cf* Sec. 6·19) that when the average

[1] See, also, RL Report No. 382, RNAS Report No. 707/1/2, TRE Report No. T1551.

intensity of the component echoes is constant in time and the number of independent scatterers is large, the probability of receiving an echo of intensity P in range dP is

$$W_1(P)\,dP = e^{-\frac{P}{P_0}}\frac{dP}{P_0}, \tag{106}$$

where P_0 is the average intensity. Strictly speaking, this distribution is only the limiting form approached as the number of scatterers becomes in-

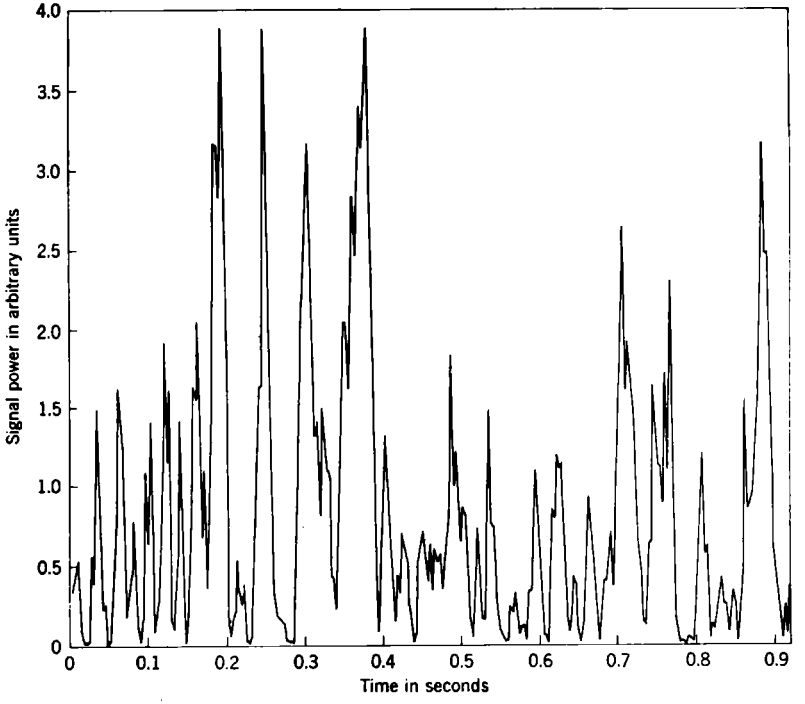


FIG. 6·18.—Plot of the r-f intensity of sea echo on 9.2 cm as obtained from pulse-to-pulse photographs of the A-scope.

finite. It is conceivable, therefore, that the number of scatterers present could be found by studying the deviations of the actual distribution from the simple exponential law. Unfortunately, the limiting process is a rapid one. Although the difference between two and three scatterers is fairly large, with higher numbers the distinction decreases rapidly. If there are as few as five comparable scatterers present, then the experimental distribution should be indistinguishable from the limiting form.

Whenever the average intensity is constant in time, the distribution obtained experimentally agrees with the exponential form within the

limits of error. Figure 6·19 shows the distribution obtained from the analysis of 2000 pulses on $\lambda = 3.2$ cm. The fraction of signals in unit power interval (that is, the smallest power interval used in the histogram) is plotted on a logarithmic scale against P/P_0 as the abscissa. With these coordinates, Eq. (106) becomes a straight line as shown. Equally good agreement with theory has been obtained even with the highest resolution available. The most that can be deduced from the first probability

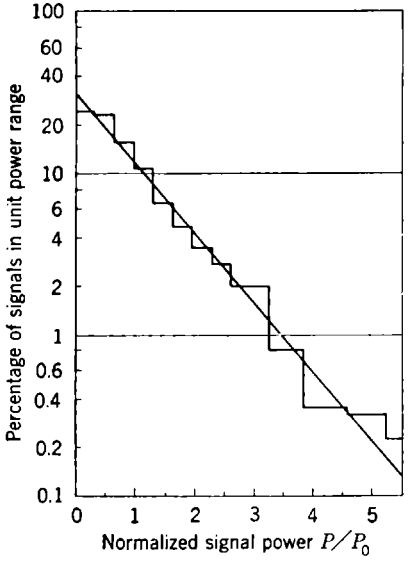


FIG. 6·19.—Probability distribution of the instantaneous signal powers for sea echo. The histogram shows the analysis of 2000 pulses, 0.002 sec apart, on 3.2 cm. The straight line is the theoretical prediction, assuming random scatterers.

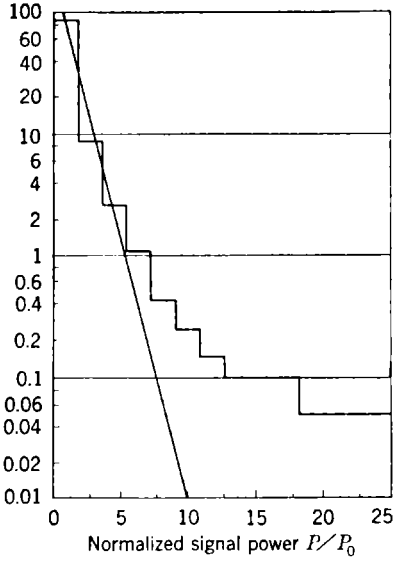


FIG. 6·20.—Probability distribution of instantaneous signal powers for sea echo. The histogram shows the analysis of 2000 pulses, 0.024 sec apart, on 9.2 cm. The straight line is the theoretical prediction, assuming random scatterers.

distribution, therefore, is that within an area of the sea 100 by 20 yd there are at least four or five independent scatterers present. This conclusion seems reasonable on the basis of almost any theory of sea echo. It does exclude, however, the possibility that specular reflection from an occasional steep surface of a wave plays an important role.

The fast fluctuations of sea echo may be described as the Doppler beats between the various scatterers. The rate of fluctuation, therefore, depends on the velocities of the scatterers relative to each other. Actually there is no single rate of fluctuation, but instead there is a continuous spectrum, as might be expected from the irregular nature of the fluctuation. Typical power spectra, which are proportional to the fraction of the energy of the incoming signal for which the fluctuation frequency lies between ν and

$\nu + d\nu$, are shown in Figs. 6·55 and 6·56 in Sec. 6·21. The maximum of the spectrum is always at zero frequency; the curve decreases monotonically with increasing frequency, roughly similar to a Gaussian curve.

For sea echo on 9.2 cm the width of the spectrum, defined as the frequency at which the power spectrum reaches half of its maximum value, has been measured to be between 25 and 40 cps. The corresponding relative velocities of the scatterers are a few miles per hour. It should be remembered that the fluctuation frequency depends, not upon their mass motion or average velocity, but rather upon the fluctuations of the individual velocities about the average. The scatter in velocities thus

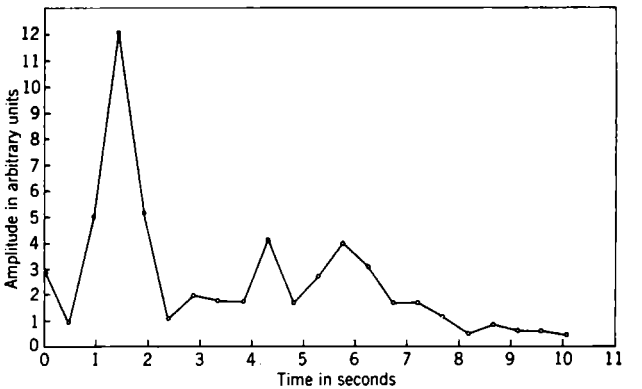


FIG. 6·21.—The amplitude of sea echo as a function of time on 9.2 cm, averaged for a period slightly less than 0.5 sec. (This represents part of the data used in obtaining Fig. 6·20.)

obtained from the frequency spectrum seems numerically reasonable, whether the fluctuation has its origin in the slightly different velocities of ripples driven by the wind or refers to the turbulences in clouds of spray drops. Very little information about the fundamental scattering mechanism can be obtained from the fast fluctuations of sea echo.

It should be mentioned that if the fluctuations are principally Doppler beats, the width of the fluctuation spectrum should be proportional to the signal frequency. Experimental verification of this point has not been easy to obtain because of the difficulty of making measurements simultaneously in time and space. Satisfactory agreement, however, was obtained with the one pair of measured spectra that referred to sea echo from the same patch of sea (see Fig. 6·56).

In addition to the fast fluctuations of sea echo, averaging about 30 cps at 9 cm, a slower fluctuation, measured in cycles per second or fractions of a cycle per second, is almost always present. Figure 6·21 is a graph of the amplitude of sea echo, averaged over intervals of about 0.5 sec, plotted as a function of time up to 10 sec. The fluctuations shown are

far too large to be only statistical fluctuations attributable to the finite averaging interval. No detailed statistical analysis of this slow fluctuation has been made, but the maximum of the spectrum is probably in the neighborhood of 1 to 2 cps. The spectrum, however, often has components with periods as long as 30 sec. The extent of this fluctuation spectrum points up the experimental difficulties in measuring the "average" intensity of sea echo. Because of these slow fluctuations such measurements must be expected to have an inherent scatter of 1 or 2 db with any reasonable averaging time.

The origin of these slow fluctuations must clearly be sought in the constantly changing nature of the sea surface. Ripples and other possible scattering surfaces grow and decay. Waves become peaked and break over; clouds of spray are formed, remaining suspended for a short time before settling on the surface. All these changes occur in times not incompatible with the observed periods of the slow fluctuations. Once again, the phenomenon offers little help in distinguishing between the fundamental scattering processes.

The presence of the slow fluctuations has a considerable influence on the first probability distributions obtained from an analysis of the echo in a finite time interval. During this interval the "average" intensity is not constant and the random process is no longer a stationary one. Deviations from the theoretical distributions are therefore to be expected. Figure 6·20 shows the first probability distribution obtained from 2000 measurements covering a time interval of 48 sec. (Figure 6·21 was prepared from the same data.) At very high signal intensities the number of observed pulses is much higher than predicted by the exponential law. It can in fact be shown (*cf.* Sec. 6·20) that in the first approximation the effect of the slow fluctuations is always to increase the standard deviation of the distribution above the theoretical value for a stationary process.

6·12. Theories of Sea Echo.—In formulating any theory of sea echo care must be taken to distinguish between high and low grazing angles of incidence, for the origin of the phenomenon is likely to be different in the two regions. From the fact that σ^0 at normal incidence is considerably greater than unity, at least for calm seas, it seems quite clear that specular reflection plays the dominant role at high angles. The size of the reflecting surfaces can be estimated from values of the cross section at normal incidence. Suppose that in unit area of the sea surface there are n flat, horizontal reflecting surfaces. Then, at normal incidence,

$$\sigma^0(90°) = \frac{4\pi nA^2}{\lambda^2} = \frac{4\pi A}{\lambda^2}\,\epsilon, \tag{107}$$

where $\epsilon = nA$, the fraction of the sea surface occupied by the flat reflecting surfaces. From Reference III, Sec. 6·10, it is found that $\sigma^0(90°) = -25$ db

on $\lambda = 3.2$ cm for slight seas. Assuming arbitrarily a value for ϵ of 0.1 for the sake of discussion, we then find that A must be about 3 ft², with appropriate changes for other values of ϵ.

The objection has been raised that the ocean surface would hardly be sufficiently smooth ($\lambda/4$ being only 0.3 in.) to constitute a specular reflector. Although there is considerable force to the objection, it is seen that the observed echo can be accounted for with only a small fraction of the total surface contributing, yet with surfaces of modest dimensions. It is to be expected, of course, that with rougher seas the extent of the reflection would decrease sharply, in accord with the decreased cross section at normal incidence found in Reference VII for rough seas.

The phenomena at high angles are thus qualitatively explicable on the basis of specular reflection. As quantitative data at high angles are meager, the rest of the section will be devoted to discussion of the phenomena at low grazing angles, especially below 5°.

The failure of the cross section to exhibit the characteristic $1/\lambda^4$ frequency dependence, as shown in Sec. 6·9, is sufficient to eliminate small spray droplets, of circumference small compared with λ, as the principal sea-echo scatterers. Attention is therefore turned to diffraction from the corrugated sea surface as the next most likely mechanism.[1]

The scattering from a sinusoidal surface has often been calculated,[2] with varying degrees of rigor. Most of these attempts assume uniform wavelength and wave height and infinite crest lengths. Such a model of the sea is extremely artificial and leads to some phenomena that undoubtedly have no reality, namely, scattering peaks at angles corresponding to discrete spectra. An attempt to take into account the intrinsic irregularity of the sea surface has been made by the Mathematics Group at the Telecommunications Research Establishment.[3] The surface was still assumed to consist of sinusoidal waves, but successive waves were considered to have different amplitudes and wavelengths distributed according to a Gaussian law. It was assumed also that the wave height was small compared with the wavelength. The calculation should, therefore, best apply to calm seas in the absence of steep and breaking waves.

The computation of the scattered field was carried out by a modified Kirchhoff-principle method, essentially as used earlier in this chapter in obtaining the scattering from plates and cylinders. Currents flowing on the surface of the sinusoidal corrugations were calculated from the incident

[1] Specular reflection is very improbable, as the occurrence of sufficiently steep surfaces is extremely unlikely.

[2] A partial list includes Lord Rayleigh, *Theory of Sound*, 2d ed., Vol. II, Macmillan, London, 1896, p. 89; T. L. Eckersley, BRL Report No. 506, January 1945, and unpublished work by A. J. F. Siegert and by L. I. Schiff and H. Feshbach.

[3] G. G. Macfarlane, "Sea Returns and the Detection of Schnorkel," TRE Report No. T1787, February 1945.

field as if the scattering surface were not present, and the reradiated field in turn was calculated from these currents. Some difficulty was encountered because of the sharp edges at the junctions between successive waves, but it is believed that "edge effects" have been eliminated from the final answer. As would be expected, the discrete spectra disappear in the averaging process.

The final formula for the cross section per unit area is

$$\sigma^0(\theta,\lambda) = \frac{L\lambda}{A^2(4\pi)^2} e^{-z^2} I_m(z^2), \tag{108}$$

where

$$z = \frac{4\pi A}{\lambda} \sin\theta \approx \frac{4\pi A\theta}{\lambda} \tag{109}$$

and

$$m = \frac{2L}{\lambda} \cos\theta \approx \frac{2L}{\lambda}, \tag{110}$$

for small θ. The wavelength of the sea waves is denoted by L; the wave amplitude by A (half the crest-to-trough height H); and I_m represents the modified Bessel function of order m.

This result can be simplified if θ is restricted to certain ranges. If the argument of the modified Bessel function is small compared with the order, more particularly if

$$z^4 \ll 4m, \tag{111}$$

then the function is well approximated by the first term in its series expansion. Therefore, for angles much less than

$$\theta_1 = \left(\frac{8L}{\lambda}\right)^{1/4} \frac{\lambda}{4\pi A}, \tag{112}$$

the cross section has the form

$$\sigma^0(\theta) = \frac{L\lambda}{(4\pi)^2 A^2} \left(\frac{z^2}{2}\right)^{m+1} \frac{e^{-z^2}}{\Gamma(m+1)}, \tag{113}$$

where, as usual, Γ denotes the gamma function. If, furthermore, z is much less than 1, that is, $\theta \ll \lambda/4\pi A$, then the exponential term is unimportant and the θ dependence is determined by $\frac{1}{2} z^2$ and hence

$$\sigma^2(\theta) \propto \theta^n, \tag{114}$$

where

$$n = 2m + 2 = \frac{4L}{\lambda} + 2. \tag{114a}$$

At the other extreme, when the argument of the modified Bessel function is greater than the order, that is, the angle is greater than

$$\theta_2 = \frac{L}{2\pi A}, \tag{115}$$

then the asymptotic expansion for the Bessel function may be used. The formula for the cross section then reduces to

$$\sigma^0(\theta) = \frac{\sqrt{2L}}{(4\pi)^{3/2}A} \sin \theta. \tag{116}$$

Thus, at large angles, the cross section increases only slowly with angle. For angles lying between θ_1 and θ_2 an intermediate behavior prevails.

The qualitative agreement of the picture thus arrived at with the experimental results presented in Sec. 6·10 is striking. It will be recalled that the cross section, especially in calm seas, was found to increase rapidly with angle at first, following a power law, and then to flatten off above some critical angle, with a linear or slower dependence on angle. However, if the accepted values for wavelengths and wave steepness are substituted in the formulas, a number of quantitative contradictions appear.

Consider first the critical angle θ_1. If one assumes the steepest possible waves, with $L = 14A$, then the formula for the critical angle can be written

$$\theta_1 = 0.26 \left(\frac{\lambda}{A}\right)^{3/4} \tag{117}$$

At $\lambda = 10$ cm, with a wave height of as much as 7 ft (beyond the limits of applicability of the theory), it is found that θ_1 is about 2.5° and is larger for less steep waves. Hence at grazing angles of 1° the approximate forms for σ^0 given by Eqs. (113) and (114) should always hold. The value of σ^0 depends upon the parameter $m + 1$ which is both the argument of the factorial function in the denominator and the exponent of a quantity less than unity in the numerator. This parameter in turn depends on the ratio of the sea wavelength to the radar wavelength. It is seen that if ordinary waves are assumed, the value of m is quite large and σ^0 is extremely small. In fact, if the steepest possible waves are assumed and A taken to be as small as λ (and therefore $L = 14\lambda$), then $\sigma^0 = -760$ db at 1°! Furthermore, the cross section then varies as the angle raised to the fifty-eighth power!

Approaching the matter from the other direction and taking the measured exponent of θ (which is never larger than 4), then one finds that the maximum sea wavelength which contributes to the echo is only $\lambda/2$ or about 2 in. at $\lambda = 10$ cm. These waves can thus only be small irregularities on the surface of the larger waves. In fact, they correspond

to ripples smaller than have been as yet observed. Again assuming maximum possible steepness ratio, one finds that the computed cross section at 1° is still too small, about −100 db, but it is much closer to actual values.

It may be objected that these quantitative contradictions result from stretching the applicability of the model too far and that the general picture is not thereby invalidated. Calculations for trochoidal waves have, in fact, been attempted by the same group at the Telecommunications Research Establishment; and when they are published, it will be interesting to see whether or not this more realistic model furnishes closer agreement with experiment. But there are some fundamental difficulties with this mechanism which will not be affected by changes in the details of the model.

For example, no provision has been made for the "shadow effect," that is, that much of the wave lies in the geometrical shadow of the wave crest at low angles. The importance of this effect is evident from the fact that with a steepness ratio of 1/20, the troughs of waves are visible only at angles of 6° or greater. The reality of the effect is also clearly demonstrated in the B-scope photograph in Fig. 6·8. Although there is some illumination in these shadowed portions by diffraction, nevertheless at low angles of 1° or less it is expected that the shadow effect would greatly influence the calculations.

More important is the total inability of this mechanism to account for the striking polarization changes that are observed. The TRE calculations assume perfect reflection and obtain the same back scattering for both polarizations. It was shown in Sec. 6·3 that on the basis of this method of computing the induced currents, the *back scattering* is the same on both polarizations as long as the convolutions of the surface have radii of curvature greater than the radar wavelength. Even a more rigorous calculation, such as the Sommerfeld solution for a half plane or wedge, yields the same result for back scattering. The assumption of perfect reflection is not strictly valid, of course. The reflection coefficient on horizontal polarization is close to unity, but that for vertical polarization is somewhat smaller. Presumably this would have the effect of diminishing the scattering on vertical polarization, but only in proportion to the decrease of the reflection coefficient. Yet for calm seas (where this mechanism should be most applicable) the observed difference between the two polarizations is in the other direction and may be as great as a factor of 1000!

In the light of these grave difficulties in the surface-scattering theory it does not seem pointless to examine once again in detail the droplet theory. For on this theory the polarization dependence admits at least of a qualitative explanation. The total field incident upon the spray droplets consists not only of the direct beam but also of the field reflected

from the sea surface. On horizontal polarization the reflection coefficient is very nearly unity and the phase change is π. As a result, the total field is practically zero just above the water surface but then rises to a maximum at the height corresponding to the first interference lobe, etc. On vertical polarization the phase shift is also π, but the magnitude of the reflection coefficient is sensibly different from unity even at low angles. Although the maxima of the lobes are not greatly affected, the minima of the pattern are not so deep as on horizontal polarization. On calm seas the conditions should be most favorable for setting up this interference pattern, and it would also be expected that the spray droplets, if present, would be close to the surface. Under these conditions, the incident field for horizontal polarization would be small and the back scattering should be far greater on vertical polarization.

As the radar wavelength is shortened, the height of the first lobe decreases. Hence, for the same height of spray droplets, the average incident field on horizontal polarization increases as the wavelength decreases. The advantage on vertical polarization should therefore decrease, as is actually observed.

As the sea becomes rougher, the interference pattern, especially at the minima, tends to be destroyed. Furthermore, the spray droplets are thrown to greater heights and cover larger portions of the interference lobe. The polarization dependence should therefore decrease as the sea becomes rougher. For very rough seas, especially on the shorter radar wavelengths, the height of the spray is enough to cover several lobes of the interference pattern. The average incident field on horizontal polarization is then the greater, because the magnitude of the reflection coefficient is larger, and the echo should then become slightly weaker on vertical polarization.

In order to illustrate these changes more concretely the calculations have been carried through explicitly on the basis of a simple model. It is assumed that the sea surface is perfectly smooth and horizontal and that the reflection coefficient on both 9.2 and 3.2 cm is the same as that calculated from the constants of sea water at 10 cm (see Sec. 5·1). The cross section per unit area is then determined by the expression

$$\sigma^0(\theta) = \int_0^\infty \int_0^\infty \sigma(r)\, n(r,h)\, F^4(z,\theta,\lambda)\, dz\, dr, \tag{118}$$

where $F(z,\theta,\lambda)$ is the pattern-propagation factor given at low angles by

$$F = 1 + \exp\left(-\frac{4\pi i z\theta}{\lambda} - i\phi\right),$$

$\sigma(r)$ is the cross section of a single drop of radius r, and $n(r, z)$ is the distribution function of the number of drops per unit area according to radius and height z above the sea surface.

It will also be assumed that the distribution of drops in height is independent of their size. This is undoubtedly not true; the very small drops would presumably be found at greater heights. Only the largest drops, however, would contribute to the echo, and the range of drop sizes to be considered is therefore fairly limited. The distribution function can then be split into two functions

$$n(r, z) = n_1(r)n_2(z), \tag{119}$$

and the cross section is given by

$$\sigma^0(\theta) = N\bar{\sigma} \int_0^\infty n_2(z)\, F^4(z,\theta,\lambda)\, dz, \tag{120}$$

where $\bar{\sigma}$ is the average cross section per drop and N is the total number of drops per unit area of the sea.

Finally, it is necessary to assume an explicit function for the distribution in height. The barometric law

$$n_2(z) = \frac{1}{z_0} e^{-\frac{z}{z_0}} \tag{121}$$

must be qualitatively correct and will be used. The quantity z_0 is an adjustable parameter which is a measure of the average drop height.

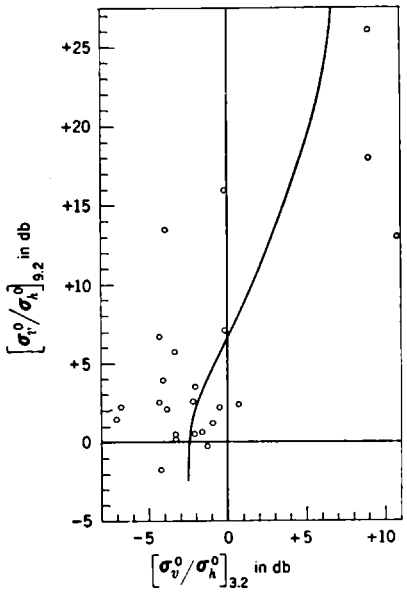


FIG. 6·22.—Correlation diagram for the polarization ratio of sea echo on 9.2 and 3.2 cm.

With these assumptions, the ratio of σ_v^0/σ_h^0 has been calculated for $\theta = 1°$ at $\lambda = 9.2$ and 3.2 cm as a function of the parameter z_0. The results are plotted on Fig. 6·22 as the continuous line. The curve comes to a definite terminus in the lower left-hand corner at a value of -4 db on both wavelengths. Also plotted are the experimental points, obtained as described in Sec. 6·9, for a range of θ between 0.65° and 1.35°. Considering the artificiality of the model and the inherent scatter of the experimental points, the qualitative agreement is satisfactory.

The dependence of σ^0 on the grazing angle θ is also in agreement with the drop theory, at least on horizontal polarization where most of the measurements have been made. The inital rapid rise of σ^0 with θ occurs at angles for which the height of the first lobe is considerably greater than

the drop height, and therefore $\sigma^0 \propto \theta^4$. As the angle increases, the lobe height decreases until the drops cover several lobes and σ^0 should become independent of θ, except for the shadow effect.

On the basis of the model used above, a simple explicit formula for $\sigma^0(\theta)$ can be obtained for horizontal polarization, where the reflection coefficient is practically unity and the phase shift almost π:

$$\sigma^0(\theta) = N\bar{\sigma}(2\rho - 1)\frac{24\left(\frac{4\pi z_0\theta}{\lambda}\right)^4}{1 + 5\left(\frac{4\pi z_0\theta}{\lambda}\right)^2 + 4\left(\frac{4\pi z_0\theta}{\lambda}\right)^4}. \qquad (122)$$

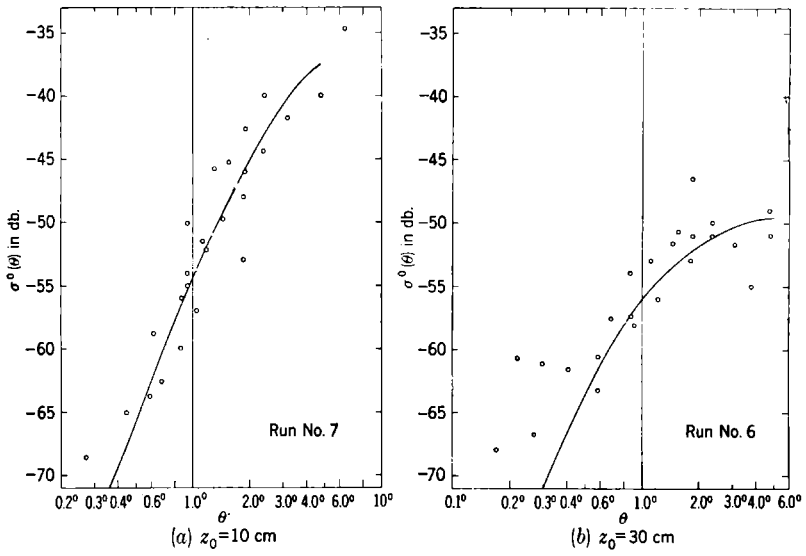


FIG. 6·23.—Two experimental plots of σ^0 as a function of θ on 10 cm (taken from Fig. 6·16) and fitted by theoretical curves predicted by the drop theory.

This formula has been fitted to the curves of σ^0 vs. θ in Reference V of Table 6·5 by suitable choice of z_0. Figure 6·23 reproduces two of the plots in which the fit was better than average. The values of z_0 ranged from less than 1 in. up to slightly more than 1 ft. The scale of increasing z_0 agreed approximately with the scale of increasing sea roughness. Thus the predictions of the drop theory are not inconsistent with the present experimental knowledge about the dependence of σ^0 on θ.

The data on the frequency dependence of σ^0 show that the drops cannot be of circumference small compared with λ, for then σ^0 should vary as λ^{-4} in rough weather and as λ^{-8} in calm weather. Actually, as seen from the results of Sec. 6·9, the dependence is more like λ^0 in rough weather and λ^{-4}

in very calm weather. To explain this frequency dependence with the drop model it is necessary to assume that the drops are such that their individual cross sections are practically independent of λ. This condition would be satisfied with drops of dimensions comparable to λ, i.e., in the neighborhood of the first resonance maximum (Fig. 6·1). The minimum diameter would be in the neighborhood of 1 cm and the cross section not far from πr^2, or roughly 1 cm^2.

Under these assumptions the frequency dependence of σ^0 stems entirely from the effects of the interference pattern. For calm seas the variation of σ^0 with λ should, therefore, go as λ^{-4}. The observed decrease of the frequency dependence with increasing sea roughness can then be interpreted as the result of two factors. (1) The height of the drops increases, covering greater portions of the interference lobes. (2) The increasing roughness tends to "wash out" the whole interference pattern.

Some idea about the magnitude of $N\bar{\sigma}$ required may be obtained from the values of σ_h(9·2). The data of Table 6·4, Sec. 6·9, require values of $N\bar{\sigma}$ ranging from -30 db in very rough weather to -50 db and possibly -60 db in very calm weather. With the foregoing drop size, the corresponding variation in N is 10 drops per square foot to one drop every 10 or every 100 sq ft.

The improbabilities of the drop theory in this form are fully recognized. The greatest polarization effects have been observed with seas so calm that no spray, not to speak of blobs of water $\frac{1}{2}$ in. in diameter, could be perceived. Another strong objection is that the reflection coefficient for any actual sea, as is well known, is considerably less than the theoretical value for horizontal polarization. In rough seas the effect of the surface irregularities should almost completely destroy any interference pattern. Nevertheless the effect of surface roughness should be less marked here than in the usual setup for measuring reflection coefficient. Because the height of the drops is small, the size of the flat area necessary to ensure almost perfect reflection is correspondingly small (*cf.* Sec. 5·4). It might even be furnished by the surface of the crest of a wave. Finally, drops of this size are unstable and should break up. They would also be expected to fall to the surface rather rapidly.

It is therefore not likely that the drop mechanism represents the actual state of affairs. However, it remains the only way at present to account for the polarization dependence. There is thus no really satisfactory explanation available at this time for the origin of sea echo.

Any theoretical discussion of sea echo must at present be based on experimental data that are, at best, inadequate and often contradictory. Future progress in the sea-echo problem requires a systematic and accurate study of the characteristics of the echo. As the polarization dependence appears to be the most distinguishing feature of the present theories, it is especially important to measure σ^0 as a function of θ for both polariza-

tions. Both the form of the curve and the absolute level should be determined. These measurements are best made from the air. Especial attention should be paid to oceanographic factors, and a careful study made of the correlation of the radar data with the sea state. All possible means of obtaining quantitative information about the sea surface should be exploited, as, for example, submerged pressure recorders on expendable radio buoys and "strip" stereophotographic cameras.[1]

On the theoretical side, a rigorous computation of the scattering from an irregular surface is the most pressing desideratum. In order to yield realistic results, the boundary conditions imposed should be "statistical" in nature; that is, the surface should be specified, not by an artificially simple model, but by probability functions for the height and slope of the surface at any given point. It would also be necessary to use more rigorous methods than have hitherto been customary, as the surface irregularities of importance must be of the dimensions of the wavelength or less. The possibility of obtaining a polarization dependence through the interference between multiple reflections from the sea surface also deserves investigation.

Some of the details of the drop theory also seem to warrant further development. The frequency dependence of single drops should be worked out in detail to beyond the first resonance maximum. Attention should be paid to attenuation and multiple scattering in dense layers of drops. It might be possible to account for the frequency dependence of σ^0 in that fashion without postulating large drops.

THE ORIGINS OF ECHO FLUCTUATIONS

By Herbert Goldstein

A portion of this volume is devoted to a description of the effects of varying atmospheric refraction on the amplitude of radar echoes from targets near or beyond the horizon. It is a common observation, however, that the echoes from less distant targets will often also show large fluctuations. It is not at all obvious that the same mechanism is responsible in both cases, and the following sections discuss in detail the possible sources of the fluctuations of echoes from near-by targets.

6·13. The Limitations of System Stability.—The radar system itself can be a source of variations in the amplitude of the received signal. Apparent fluctuations in the echo may be caused by changes in the amplitude of the transmitted signal or by variations in the transmitter and local-oscillator frequency and in the gain of the receiver and video amplifiers. It is outside the scope of this book to treat in detail the many sources of system instability and the methods by which such instability

[1] A program of this sort had been projected in the Radiation Laboratory but was abandoned at the cessation of hostilities.

may be reduced. A few typical figures will be presented, however, to give an idea of the magnitude of echo variations that can be ascribed to system fluctuations. These data were taken on an experimental 9-cm mobile radar system.[1] They are characteristic of the stability that may be obtained in the laboratory without too great an expenditure of effort but by no means represent the limits that can be achieved if sufficient care is taken. On the other hand, they are not applicable to field systems, especially airborne sets, which are designed for maximum sensitivity and compactness and in which such refinements as carefully regulated power supplies cannot often be included.

Stability of a system or of its components cannot be specified without a statement of the time interval involved. The source responsible for fluctuations over fractions of a second may have no connection with the source that gives rise to variations over minutes or hours. Thus the pulse-to-pulse amplitude fluctuation of the magnetron pulse because of inherent "jitter" could be easily reduced to less than $\pm$ 0.06 db, the limit of observation. But careful voltage regulation is necessary in order to keep the variation over minutes down to $\pm$ 0.10 db; otherwise even ordinary line voltage changes of 1 or 2 per cent will cause pulse variations several times as large. In general, the parameters of a radar system cannot be maintained constant over a period of days or longer without repeated adjustments and the frequent use of test equipment.

Ripple in the receiver and indicators is a common cause of apparent fluctuation. This is especially true in airborne sets where jitter arising from microphonic action will often make the indicators unusable. With adequate filtering the ripple can be greatly reduced. Thus the apparent fluctuation of a signal generator was reduced to a point where the standard deviation of the amplitude was 0.12 db, which is close to the limit of observation. The high-gain i-f amplifiers in receivers are naturally difficult to stabilize over long periods of time, and they cannot be depended on to have a constant gain even for a few minutes. They are often especially sensitive to changes in heater voltage. These difficulties can be avoided by measuring echo amplitudes with reference to an r-f signal generator pulse that passes through the same receiver as the echo.

Frequency stability of the system depends upon the type of transmitter used—magnetron, klystron, etc.—and even on the individual magnetron. No exact measurements of frequency fluctuations have been made, but an examination of a properly operating 10-cm magnetron with a spectrum analyzer shows that the pulse-to-pulse changes must be less than 200 kc/sec. Likewise, the frequency modulation within the pulse can be determined from the shape of the spectrum and is usually less than 200 kc/sec/μ sec. These figures, expecially for frequency modulation, must be increased somewhat at higher frequencies.

[1] The system was of the ordinary type, not designed for MTI operation.

In most radar systems fluctuations in the magnetron frequency are compensated by an AFC which keeps the local oscillator in tune with the magnetron. Some types of AFC, however, if not adjusted precisely, will "hunt" slightly and thus cause a certain amount of jitter in the final signal. In addition, frequency changes can cause variations in the echo amplitude if the cross section is a sensitive function of the frequency. Thus, let us consider a target consisting of two equal scatterers spaced so that they are barely unresolved, that is, a distance $\tau c/2$ apart, where τ is the pulse duration and c the speed of light. The number of wavelengths in the round-trip distance between them is $n = \tau\nu$, where ν is the frequency. If the number n changes by $\frac{1}{2}$, the interference between the echoes from the two scatterers can change from destructive to constructive. The frequency shift required is $1/2\tau$ or, for a 1-μsec pulse, 500 kc/sec. This is larger than the limits of pulse-to-pulse frequency instability, at least for 10 cm, but frequency shifts as large or larger are often caused by slight supply-voltage changes.

The question arises whether or not it is possible to find a complex target that is more "frequency sensitive" than just two equal scatterers. Although no general answer can yet be given to this question, it seems likely that two equal scatterers is in fact the most extreme case. Thus, let us suppose that instead of two targets, there are N targets, equally spaced, with the distance between the first and last equal to $\tau c/2$. Then it can be shown that the frequency shift necessary to go from maximum amplitude to minimum amplitude is the same as with just two targets.[1]

It is instructive to consider a more general class of targets, namely, those consisting of a large number of uniformly dense scatterers, with randomly distributed phases. The individual scattered amplitudes do not necessarily have to be equal. As will be shown in the following sections, most types of "clutter" targets fall in this category, and even a ship may be well approximated by such a model.

The instantaneous amplitude received by the radar is given by

$$V = \sum_j v_j s(x_j) e^{2ikx_j}, \tag{123}$$

where v_j is the amplitude of the jth scatterer alone, $s(x_j)$ is a real function giving the variation of the incident field in space and depends solely on the pulse shape, k is the wave number $2\pi/\lambda$, and x is the radial coordinate from the radar. The instantaneous signal power is given essentially by

$$P \propto VV^* = \sum_{j,l} v_j v_l^* e^{2ik(x_j - x_l)} s(x_j) s(x_l). \tag{124}$$

[1] Equivalent to this statement is the well-known fact that the apparent width of a monochromatic line in a diffraction grating spectrum depends solely on the width of the whole grating and not on the spacing between the slits.

Let us suppose that an imaginary experiment is now performed, in which two radar systems located at the same spot and looking at exactly the same target are identical except that one operates on a wavelength λ_1 and the other on $\lambda_2 = \lambda_1 + \Delta\lambda$. If P_1 is the instantaneous signal power on one set and P_2 the corresponding simultaneous value on the second set, then the quantity $\overline{(P_2 - P_1)^2}$ (where the average is over all possible random configurations of the individual scatterers) will be a function of $\Delta\lambda$ and is a measure of the frequency sensitivity of the target cross section. From Eq. (124),

$$P_2 - P_1 \propto \sum_{j,l} v_j v_l^* s(x_j) s(x_l)\, e^{2ik(x_j - x_l)}\, \psi_{jl}, \tag{125}$$

where

$$\psi_{jl} = e^{2i\,\Delta k\,(x_j - x_l)} - 1. \tag{126}$$

Squaring Eq. (125), one obtains

$$(P_2 - P_1)^2 \propto \sum_{j,l,p,q} v_j v_l^* v_p v_q^* s(x_j) s(x_l) s(x_p) s(x_q)\, e^{2ik(x_j - x_l - x_p + x_q)}\, \psi_{jl}\, \psi_{pq}^*. \tag{127}$$

When the average over all possible configurations is computed, only those terms will remain for which the exponent vanishes, as the phases are assumed randomly distributed.[1] The exponent vanishes only when $j = l$ and $p = q$ or when $j = p$ and $l = q$. The average is therefore given by

$$\overline{(P_2 - P_1)^2} \propto \sum_{j,p} v_j^2 v_p^2 s^2(x_j) s^2(x_p) \psi_{jj} \psi_{pp}^* + \sum_{j,l} v_j^2 v_l^2 s^2(x_j) s^2(x_l) \psi_{jl}^2. \tag{128}$$

From Eq. (125), ψ_{jj} and ψ_{pp}^* vanish identically and the average reduces to

$$\overline{(P_2 - P_1)^2} \propto \sum_j v_j^2 s^2(x_j) \sum_l v_l^2 s^2(x_l) \psi_{jl}^2. \tag{129}$$

To evaluate this quantity further it will be assumed that the number of scatterers is sufficiently large that the sums may be replaced by integrals. Let N be the number of scatterers in the beam per unit length and $f(v^2)$ the distribution function for v^2, that is, $f(v^2)\, dv^2$ is the fraction of scatterers for which v^2 lies between v^2 and $v^2 + dv^2$. This distribution is assumed independent of x. Then the average signal power is given by

$$\overline{P} \propto \sum_j v_l^2 s^2(x_j) = N \int_0^\infty s^2(x)\, dx \int_0^\infty v^2 f(v^2)\, dv^2, \tag{130}$$

and Eq. (128) becomes

$$\overline{(P_2 - P_1)^2} \propto N^2 \int_0^\infty \int v^2 f(v^2) s^2(x) \int_0^\infty \int v'^2 f(v'^2) s^2(x') \psi^2(x,x') dx'\, dv'^2 dx\, dv^2, \tag{131}$$

[1] This operation is really a time averaging if the scatterers are moving about randomly, as in the case of storm echoes, etc.

where $\psi^2(x,x') = 4 \sin^2 [\Delta k(x-x')]$. If there is now introduced a normalized function

$$S^2(x) = \frac{s^2(x)}{\int_0^\infty s^2(x)\, dx}, \tag{132}$$

then Eq. (131) can be simply written as

$$\overline{(P_2 - P_1)^2} = 4\overline{P}^2 \int_0^\infty \int_0^\infty S^2(x)\, S^2(x') \sin^2 [\Delta k(x - x')]\, dx'\, dx. \tag{133}$$

These integrals cannot be evaluated without a knowledge of the function, $S^2(x)$, which depends solely on the pulse shape. The rise and fall time of the pulse is usually small compared with the period of oscillation of $\sin^2[\Delta k(x-x')]$, and there is therefore little error in assuming the pulse to be rectangular. The function $S^2(x)$ is then

$$S^2(x) = \frac{1}{\delta}, \qquad R \leqq x \leqq R + \delta$$

$$= 0 \text{ elsewhere,}$$

where R is the range and δ is the radar pulse length, equal to $\tau c/2$. With this choice for the pulse shape, Eq. (133) reduces to

$$\overline{(P_2 - P_1)^2} = 2\overline{P}^2 \left[1 - \left(\frac{\sin \pi\tau\, \Delta\nu}{\pi\tau\, \Delta\nu} \right) \right]^2. \tag{134}$$

The corresponding mean-square change can be easily computed for the case of two equal scatterers a distance δ apart and turns out to be

$$\overline{(P_2 - P_1)^2} = 2\overline{P}^2 \sin^2 (\pi\tau\, \Delta\nu). \tag{135}$$

The first maximum of Eq. (135) occurs at $\Delta\nu = 1/2\tau$, in agreement with the rough conclusions given earlier. But in Eq. (134), the first maximum occurs at $\Delta\nu = 1/\tau$. Hence this type of target, consisting of many random scatterers, is less frequency sensitive than the simple case of two equal scatterers. It should be remarked again that a large fraction of radar targets fall within this classification and that the assumptions made in the derivation for the sake of mathematical simplicity could be removed without materially affecting the results.

It seems safe to conclude that with a carefully regulated system, fluctuations caused by frequency shifts of the transmitter will not be important.

6·14. Atmospheric Variations.—It has been shown in the preceding chapters that varying atmospheric refraction may have a profound influence on the strength of signals coming from beyond the horizon. The question immediately occurs whether or not this same phenomenon

might not also be of importance in the variations of echoes from near-by targets. The problem is most simply attacked by considering one-way transmission paths rather than the actual radar situation. Unfortunately experimental data of this type are rather meager, most attention having been paid to the effects in the diffraction region.

Anderson and others[1] have measured transmission over a 13.5-mile land path in which the angle of the ray with the horizontal was 2.5°. The results obtained have already been discussed in Secs. 4·12 and 4·13; it will be recalled that the maximum variations were $\pm \frac{1}{2}$ db and always of a long period. The signal strengths over longer nonoptical paths measured at the same time showed much larger fluctuations.

Pertinent theoretical work is even more meager. Freehafer has studied the case of a rather pronounced S-shaped duct by means of ray theory calculated with the aid of the MIT differential analyzer. The transmitter was assumed located at a height of 150 ft, in the middle of the duct. At a distance of 7 miles the calculated field differed from the standard by not more than 2 db. At 12 miles the deviation was not much greater.

Despite this paucity of information, it seems safe to conclude that near-by echoes may be somewhat affected by nonstandard refraction but that the effects are many orders of magnitude smaller than those commonly observed in the diffraction region. Furthermore, the changes that may occur are of a fairly long period; hence one should speak of *variations* and not of *fluctuations*. It should be remarked that these conclusions are based solely on one-way transmission data. There is no known unambiguous case in which changes of radar echoes from clearly optical targets have been correlated with the presence of nonstandard atmospheric refraction.

There might be other mechanisms whereby changes in refraction would cause echo fluctuations. Thus, the beam might actually be bent "off" the target. It is easy to calculate, however, that even in the extreme case of grazing incidence on the boundary between two greatly dissimilar air masses, the beam deviation is not greater than a small fraction of a degree. Sharpless and Crawford[2] have measured the deviations occurring over somewhat longer paths than are of interest here and found that they never exceeded 0.1° in azimuth or 0.70° vertically. Only with exceptionally large antennas would this amount of bending be at all important. (See Secs. 4·23 and 4·24 for an account of these experiments.)

Another interesting possibility is the existence of a phenomenon analogous to the familiar twinkling of stars. The image of stars close to

[1] P. A. Anderson *et al.*, NDRC Project PDRC-647, Washington State College, Report No. 4, Oct. 26, 1943.

[2] A. B. Crawford and W. M. Sharpless, *Proc. IRE*, **34**, 845 (1946). See also A. W. Straiton, *Proc. IRE.*, **37**, 808 (1949).

the horizon will often show large intensity fluctuations with standard deviations as high as 60 per cent. The rate of fluctuation varies from 25 to 150 per second. Since the time of Kepler this twinkling has been known to be associated with atmospheric refraction. More recently it has been shown[1] that the effect is produced by small inhomogeneities in the refractive index which act as weak lenses focusing and defocusing the rays. The wavefront thus becomes somewhat corrugated. The light and dark shadow bands seen on the surface of the earth at the instant before a total eclipse of the sun are evidence of these corrugations. Similar shadow bands exist for stars except that they are too weak to be perceived directly; but as the bands pass across our vision, the apparent intensity of the star image fluctuates. The size of the inhomogeneities has been measured to be from 4 to 60 in. These striae have been found to move at speeds comparable to that of the wind.

It would seem, a priori, that similar effects should occur at microwave frequencies. There is no conclusive evidence, however, either theoretical or experimental, that they are of practical importance for transmissions over paths not extending into the diffraction region.

Some calculations on the possible effect of these inhomogeneities have been made by Scott,[2] but the most detailed theory so far is that of Bergmann.[3] He has considered the problem solely as one in geometrical optics, neglecting diffraction effects and assuming that the variations of the refractive index are isotropic in space. The rms intensity fluctuation in db on a one-way path is given by

$$\sqrt{\overline{L^2}} = 0.0795R^{3/2}\sqrt{\alpha\overline{(\nabla^2 n)^2}}, \tag{136}$$

where R is the range, α is the "patch" size defined in a certain manner, $\nabla^2 n$ is the Laplacian of the refractive index, and the averaging is to be taken over all space.

Meteorological data that can furnish us with values for α and $\nabla^2 n$ are practically nonexistent. A. H. Woodcock, of the Woods Hole Oceanographic Institution, has made some measurements with wet- and dry-bulb thermometers in a situation of unstable cold air moving over water at a higher temperature.[4] The meteorological conditions were thus favorable for the production of inhomogeneities. The maximum deviation of the

[1] For a description of the experimental investigations and further references see R. W. Wood, *Physical Optics*, 3d ed., Macmillan, New York, 1934, p. 91; J. A. Anderson, *Jour. Optical Soc. Am.*, **25**, 152 (1935); and M. Minnaert and J. Hontgast, *Z. Astrophysik*, **10**, 86 (1935).

[2] J. M. C. Scott, ADRDE Memorandum No. 42, March 8, 1944.

[3] P. G. Bergmann, *Phys. Rev.*, **70**, 486 (1946).

[4] Unpublished data. These measurements are not the same as those referred to in Sec. 3·26 and have been analyzed in a different fashion, as the above section deals with the purely meteorological aspects of the situation.

index from the average was as high as 1×10^{-5}, but the rms deviation was about 2.5×10^{-6}. These calculated fluctuations are probably too small, as a result of the large lag of the thermometers. The data have been analyzed statistically to furnish the quantities needed in Eq. (136). The patch size α was about 40 yd. At a range of 20,000 yd, an extreme case, the computed rms echo fluctuation is 0.15 db, which is entirely too small to be of importance.

The maximum fluctuation of the index of refraction is probably higher than the data indicate, but not enough to affect the result greatly. If, however, the size of the inhomogeneities were much smaller, let us say comparable to those which are of importance for star twinkling, that is, much less than 1 yd, then for the same absolute deviation of the index the Laplacian would be much greater and hence the fluctuation would become large. But if the inhomogeneities are small in extent, then the corrugations in the wavefront are also small; hence the average antenna aperture would include many of the "shadow bands." Now it can be shown that the fluctuations in phase will always be small. Bergmann has derived the following expression for the rms fluctuation of the optical path length:

$$s = \sqrt{\gamma(n - n)^2 R}, \tag{137}$$

where γ is another patch size defined in a slightly different manner from α. Using the same data as for Eq. (137) with a range of 20,000 yd, s turns out to be 0.3 cm. If the patch size were only a few inches, the phase fluctuations would be even smaller. Hence the various shadow bands across the antenna aperture are in phase, and the total signal measured would tend to be an average over the several fluctuations. The corresponding phenomenon has been observed optically, that is, the intensity fluctuations of a star image decrease as the telescope aperture is increased.

There is no experimental evidence of fluctuations on short paths that can be indisputably ascribed to atmospheric inhomogeneities. Whenever the data have been subjected to close examination, other mechanisms have always been found that fit the observations as well or better. Simultaneous transmission of 9- and 3-cm waves was briefly studied on an overland path of about 10 miles, and fluctuations observed that were slightly larger than those due to purely instrumental causes. The rms fluctuation on 3 cm, 0.27 db, however, was roughly three times the fluctuation on 9 cm. A statistical analysis showed that the rates of fluctuation were also inversely proportional to the wavelength. Finally, the cross correlation between the simultaneous signal intensities on the two wavelengths was completely negligible. These facts agree perfectly with the assumption of a small amount of ground reflected signal that is variable, chiefly in phase. On the other hand it has been seen that only such inhomogeneities as are large compared with the wavelength are likely to have much effect practically. But in this case one would expect, in con-

tradiction to the experimental data, that the fluctuations would be independent of wavelength.

It seems likely, therefore, that inhomogeneities of the refractive index will have a negligible effect on microwave transmission over short distances. On longer paths, extending into the diffraction region, the effects are greater. It is obvious, however, that here, as in the influence of nonstandard refraction on near-by echoes, the investigations have been initiated and much further work remains to be done.

6·15. Fluctuations in the Space Interference Pattern.—If a microwave transmission path is over water and to a smaller extent if it is over land, the field incident on the target consists of both a direct and reflected component. Changes in the amplitude and phase of the reflected component relative to the direct one are a fruitful source of variations in the radar signal. Two types of phenomena are to be distinguished:

1. The reflecting surface is smooth but moves up and down relative to transmitter and target. A space interference pattern is still present, but the positions of the maxima and minima relative to the target are changed.
2. The reflecting surface is not smooth but irregular; hence there may be many reflecting and scattering points whose contributions will be changing in time. The whole nature of the interference pattern is changed and is often almost destroyed.

A striking illustration of the first type of effect is the frequently observed variation of signal strength with height of the tide. As a simple case let us consider a point target, and assume a wide beam vertically and a reflection coefficient of -1. The variation in echo power is contained in the factor F^4 in the transmission formula. With the foregoing assumptions this factor can be written as

$$F^4 = 16 \sin^4 \left[\frac{2\pi}{\lambda R} (h_1 - h_0)(h_2 - h_0) \right], \tag{138}$$

where h_1 and h_2 are the transmitter and target heights above a given datum level and h_0 is the tide height referred to the same level. Obviously if the tides are of sufficient range, F^4 can vary from a minimum of zero to its maximum value of 16.

Figure 6·24 shows a plot of the received echo at 9 cm from a 40-in. square corner reflector at a range of 6000 yd. The transmitter was 210 ft and the target height about 50 ft above mean sea level. The circles represent the echo power on horizontal polarization, as measured at half-hour intervals in a 72-hr period (covering some 12 tidal cycles), and show a range of more than 25 db. The continuous curve was calculated on the basis of Eq. (137) and was shifted vertically on the decibel scale to agree with the maximum values. Good agreement is obtained except in the neighborhood of the minimum. There are a number of

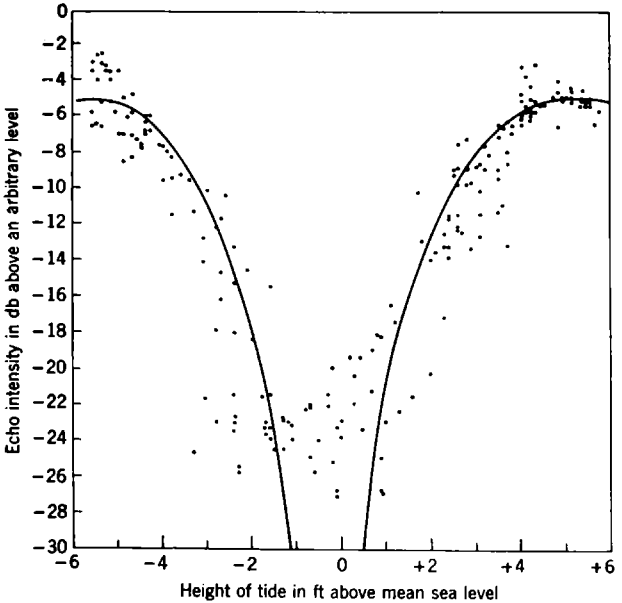


FIG. 6·24.—Tidal cycle in the echo from a 40-in. corner reflector on 9.2 cm. Transmitter height 210 ft, target height 50 ft, range 6000 yd. The dots are experimental points; the solid line is the theoretical curve for a point target, assuming the reflection coefficient is −1.

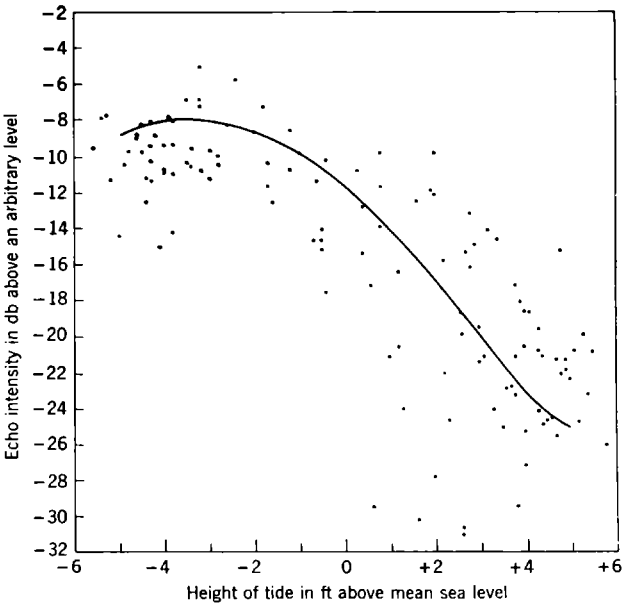


FIG. 6·25.—Tidal cycle in the echo from the Strawberry Hill standpipes on 9.2 cm. Transmitter height 120 ft; the targets extend in height from 100 to 160 ft, range 10,000 yd. The dots are experimental points; the solid line shows the predicted variation on the basis of the theoretical model.

reasons for this exception. The reflection coefficient, even for a smooth surface, is not exactly -1 at the angles involved, and the water, of course, was not smooth. The fluctuations caused by surface roughness are greatest at the minimum where the direct and reflected rays are out of phase (*cf.* Secs. 5·6 and 5·8). In addition, the corner reflector is not exactly a point target but is sufficiently large to cover an appreciable portion of one of the interference lobes; this should have the effect of filling in the minimum without appreciably changing the signal at its maximum.

In fact, if the vertical extent of the target is sufficient to cover several lobes of the interference pattern, it might be thought that the "tidal effect" would be small if not completely negligible. In practice, however, it is found that echoes from extended targets often do show a large variation with tide.

Figure 6·25 is a plot of the signal strength from Strawberry Hill on Nantasket, Mass., over a 4-day period. The radar, operating on 9 cm, horizontal polarization, was located about 10,000 yd away on a hill 120 ft above mean sea level. The target consists of a large steel standpipe 57 ft tall on a hill 100 ft high, along with a much smaller neighboring water tank. As before, the circles represent experimental points measured at half-hour intervals over a period of 4 days. Although the scatter is large, there is a definite variation of some 18 db with tide.

The standpipe covers one and one-half lobes of the interference pattern,[1] including two minima, a maximum, and most of another maximum. Nevertheless the effect of these lobes is not "averaged out," essentially because the component parts of the target scatter coherently. The nature of the tidal variations of echoes from extended targets is best understood by examining the various components of the received signal. Referring to Fig. 6·26 one sees that the echo can be divided into three portions:

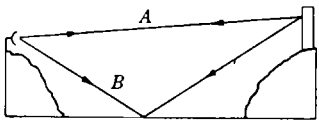


FIG. 6·26.—Illustrating tidal variations with extended targets.

1. The incident and reflected waves travel along path A.
2. The incident wave travels along path A, but the received wave along path B; or equivalently, the field is incident on path B and reflected on path A.
3. Both the incident and reflected waves travel along path B[2].

The first component is obviously uninfluenced by the tide, but for the other two the variation in the height of the sea causes a corresponding

[1] The water tank can probably be neglected as a target compared with the standpipe.

[2] An equivalent analysis was given in Sec. 6·5 for the flat-plate model of a ship target.

change in the angle of incidence on the target or in the angle of scattering.

The reradiation pattern of a geometrically simple object, such as the standpipe, as obtained from Fresnel diffraction theory, has a sharp peak in the direction corresponding to specular reflection. The width of the peak corresponds roughly to λ/l, where l is the vertical extent of the target. Away from the peak the amplitude of the scattered signal is small and varies only slowly with angle. The phase of the diffracted wave, on the other hand, is almost stationary in the neighborhood of the specular reflection peak but varies rapidly outside this region. If now the angles are such that all three components of the received signal correspond to diffraction angles away from the main peak, so that their amplitudes are of the same order of magnitude, then the change in the relative phases with tide may be great enough to cause large variations in the total signal. Obviously, it is possible that at one stage of the tide the amplitudes and phases may be such as to result in almost complete cancellation of the signal.

In the case of the Strawberry Hill standpipe the first component (direct wave) comes almost exactly in the direction of specular reflection, providing the target is exactly vertical; it would then be much larger than the other components. Little tidal variation is then to be expected. If the target is not exactly vertical, however, but leans at an angle of the order of λ/l, the specular reflection beam is thrown away from the receiver and destructive interference may easily take place. The angle involved is small, 20′ of arc or less. The continuous curve in Fig. 6·25 was calculated on the assumption of an angle of about 8′, chosen to give the best fit with the experimental points. The corresponding deviation of the tower from the vertical would be not more than 1 in., which seems reasonable.

The true state of affairs is probably much more complicated, because of the presence of the water tank and the hill itself. This simplified model has been considered in detail, however, in order to demonstrate that large tidal variations of signal strength may occur even with extended targets.[1]

The phenomena resulting from the roughness of the sea have already been described in Sec. 5·5. Targets viewed over water often exhibit characteristically violent and rapid fluctuations because of the irregular nature of the reflection (especially in the minimum of a tidal cycle; *cf.* Figs. 6·24 and 6·25). Although the fundamental origin of these fluctuations seems clear, many questions about the magnitude of the effects and the conditions favorable for their occurrence remain to be answered by further investigations.

The discussion has so far been restricted to reflection from water surfaces. In Sec. 5·10 it has been shown that the reflection coefficient of even

[1] Similar effects for both point and extended targets have been reported by G. C. Southworth, A. P. King, and S. D. Robertson, "K-band Radar Transmission," BTL Memorandum No. MM44-160-115, May 1944.

dry ground, as measured at short ranges on carefully leveled plots, can be high. It is common experience, however, that ground reflection is rarely found for long ranges. This situation is quite understandable, since the irregularities of most terrain are such that the Rayleigh roughness criterion [Eq. (5 19)], is greatly exceeded. As a result, the intensity of the wave reflected from the ground surface is small compared with that of the direct ray.

It also seems reasonable that this small scattered wave will fluctuate violently under certain circumstances. Thus, if the reflection area is heavily vegetated, one would expect that the nature of the surface would constantly change with time because of the wind. Again, if the reflection takes place over city areas including roads and railroad yards, similar effects seem reasonable.

The transmission over a number of short paths, about 10 miles long, across the city of Boston, Mass., was briefly observed simultaneously on 9.2 and 3.2 cm. At the longer wavelength the rms fluctuation was 0.1 db and the autocorrelation was down to 0.5 at 4 sec. On 3.2 cm, the rms fluctuation was about 0.27 db with the corresponding width of the correlation function near 1.6 sec. Vanishing cross correlation was found between the two records. As has already been pointed out in Sec. 6·14, the dependence of the signal characteristics on wavelength is in agreement with the hypothesis that the fluctuation arises from the ground-scattered wave and not from atmospheric turbulences.

Additional evidence for this view comes from the behavior when the polarization of transmitter and receiver were crossed. From the normal depolarization of the paraboloid antennas a decrease of 20 to 30 db is to be expected. Actually decreases of 10 to 25 db were observed, and the signal fluctuated violently between these limits. It seems, therefore, that over these paths there is an appreciable component scattered in such a manner as to rotate the direction of polarization.

As in the case of fluctuations in the reflection from water, the question with regard to ground reflections has been too little explored. Much more work must be done before the magnitude and importance of these fluctuations are known.

6·16. Isolated Moving Targets.—On coherent types of radar systems the echo from a moving target will show a periodic fluctuation because of the target's radial velocity relative to the radar. With noncoherent systems this Doppler beat, so-called, is absent but echoes from moving planes or ships still show large fluctuations.

The most striking characteristic of aircraft echoes is an almost periodic variation of the signal strength. It has been shown conclusively that this variation arises from a periodic modulation of the plane's radar cross section as the propellers rotate. The nature and properties of "propeller

modulation" have been studied intensively and are reported in detail elsewhere.[1,2]

Figure 6·27 shows the signal from a B-26 bomber on a 10-cm system for an interval of about 0.1 sec. The periodicity is strikingly evident. The fundamental frequency is given, not by the shaft speed itself, but by the shaft speed times the number of blades per propeller and is about 50 to 60 cps. It is obvious from the figure that the signal must be rich in harmonics of this fundamental frequency. Figure 6·28 shows the

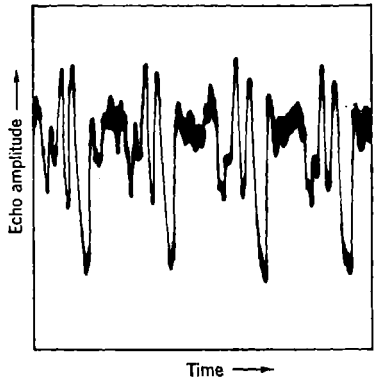


FIG. 6·27.—Modulation from propeller rotation on the 10-cm echo from a B-26.

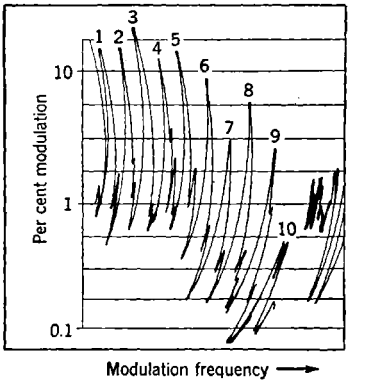


FIG. 6·28.—Frequency spectrum of the modulation shown in Fig. 6·27.

frequency spectrum of the signal in Fig. 6·27. The peaks indicate the intensity of the individual harmonics in terms of per cent modulation. It is to be noted that the harmonic at three times the fundamental frequency is stronger than the fundamental. There is an appreciable contribution from as high as the tenth harmonic.

As is to be expected, the propeller modulation percentage is a function of the plane's aspect. It is naturally greatest head-on, slightly less for the tail aspect, and considerably less broadside, where the reflection from the fuselage predominates. Figure 6·29, which is a record[3] by Ashby and Martin of the propeller modulation of the signal from a B-26 as a function of aspect, illustrates these differences. The plot also shows that the pattern has a very detailed lobe structure. The same general dependence

[1] J. L. Lawson, Ed., "Detection of Propeller and Sambo Modulations," RL Report No. S10, May 16, 1944; R. M. Ashby and F. Martin, RL Report No. 914, March 28, 1946. Also see Vol. 24, Chap. 10, of the Radiation Laboratory Series.

[2] British work in the same field is reported by G. H. Beeching and N. Corcoran, "The Characteristics of S-band Aircraft Echoes," ADRDE Research Report No. 253, May 8, 1944.

[3] Ashby and Martin, *op. cit.*

of propeller modulation on aspect is reported by Beeching and Corcoran[1] who also state that the greatest modulation, head-on, was found in a Mosquito plane (65 per cent) and the least for a B-17 bomber (25 per cent).

If the propeller modulation is smoothed out by averaging the signal over a time long compared with the fundamental period, a much slower variation

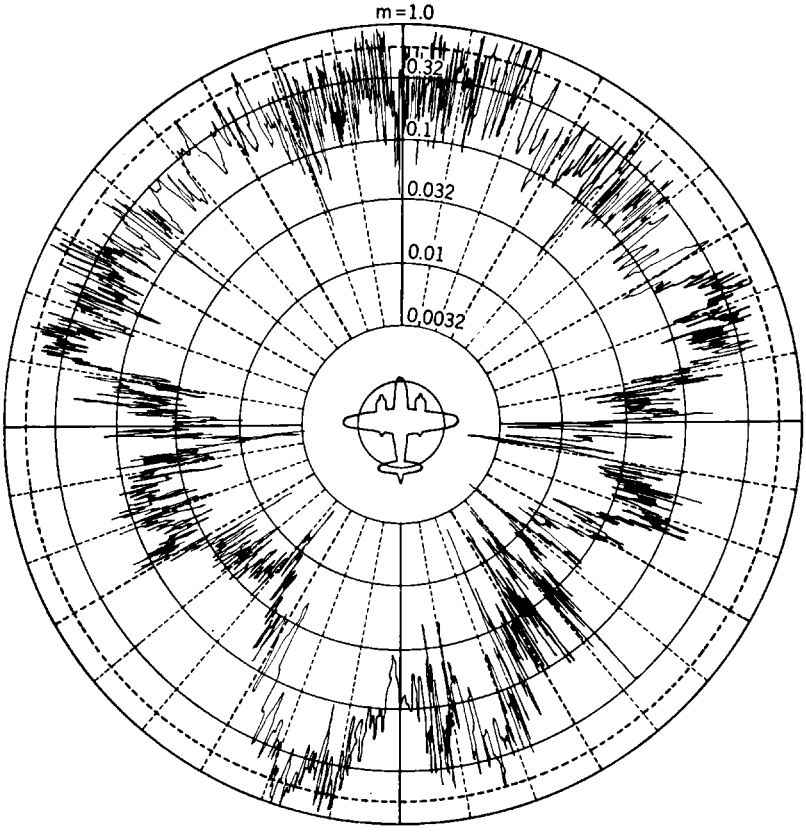


FIG. 6·29.—Per cent propeller modulation of the echo on 10 cm from a B-26 bomber, as a function of azimuth.

in signal strength is found to be present in addition to the propeller modulation. Figure 6·30 shows the echo intensity from an AT-11 aircraft, head-on, averaged over intervals of 0.03 sec and plotted for 3.2 sec. These measurements were made from pulse-to-pulse A-scope photographs and show a fading exceeding 10 db. A frequency analysis of the data indicates

[1] *Op. cit.*

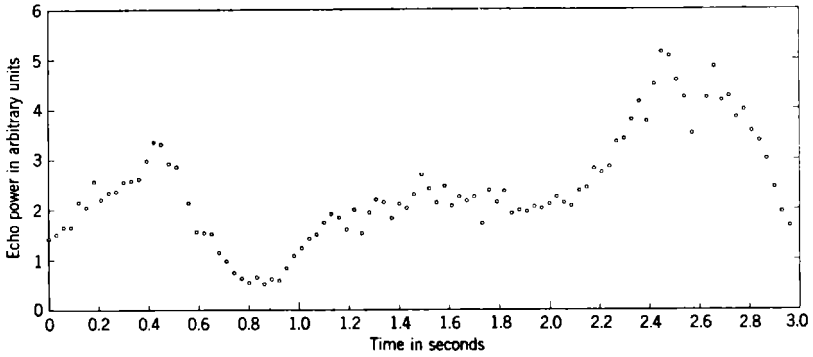


FIG. 6·30.—The echo from an AT-11 airplane on 9 cm as a function of time. Each point represents an average over a 30-msec period.

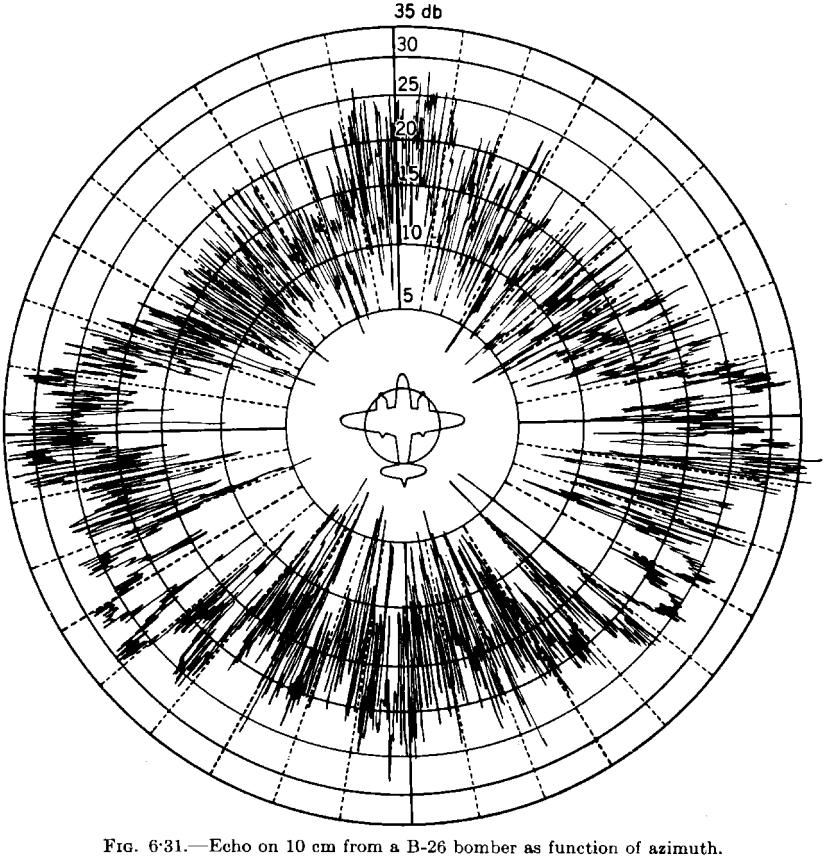


FIG. 6·31.—Echo on 10 cm from a B-26 bomber as function of azimuth.

a maximum in the spectrum of about 0.6 cps. Similar results have been obtained by recording the AGC voltage in the radar receiver.[1]

This fluctuation, slow compared with propeller modulation, is almost entirely attributable to the fine lobe structure of the radar cross section as a function of aspect. One would expect the width of the lobes to be given roughly by $\lambda/2l$, where l is some characteristic length of the airplane. Thus at 10 cm, with $l = 30$ ft, the width of the lobes should be of the order of magnitude of 0.3°. An airplane in flight will change aspect by many times this figure, yawing being the most important motion for these considerations. These small changes in aspect due to yaw give rise, therefore, to large changes in the echo. Figure 6·31 shows a record taken by Ashby and Martin of the signal strength from a B-26 bomber as a function of azimuth. The detailed lobe structure is evident, although quantitative conclusions are difficult to draw because the recorder could not follow the rapid changes satisfactorily.

These results provide an adequate explanation for the "spotty" nature of aircraft echoes on microwave scanning systems. Normally the beam is on the target for a time long compared with the propeller modulation period but short compared with the slower fading due to aspect changes. Consequently it becomes a matter of chance whether the beam hits the target when it is oriented to give a large echo or a small echo. If the general signal level is not too high above noise, the echo will apparently disappear for one or more scans, as has been observed in many cases. The fluctuating nature of reflection from the rough surface of the sea also often contributes to the fading and "spotty" signals.

Ship echoes, of course, do not have propeller-modulation effects but nevertheless show deep and often rapid fading.[2] These fluctuations are the result of changing aspect or variable reflection from the sea surface. The reradiation pattern of a ship should have an even finer lobe structure than that of aircraft, as ship dimensions are usually much greater.

Several attempts have been made to correlate the position of a ship with the instantaneous intensity of the echo. Cooper[3] obtained simultaneous measurements of a trawler's roll and slew along with the echo on a 10-cm system but failed to find any correlation. From a battleship, Koller[4] searched on 10-cm radar for regular periods in the echo which might be attributed to roll, also without success. At these high frequencies,

[1] Lawson, *op. cit.*

[2] There are several reports, however, that echoes from PT-boats show a characteristic rapid flutter that has been ascribed to vibration caused by the engines. See "Use of Shore Based 10 cm Radar for Watching Ships and Aircraft," AORG Publication N3, September 1944.

[3] R. I. B. Cooper, "The Fading of S-band Echoes from Ships in the Optical Zone," RRDE Report No. 265, December 1944.

[4] L. R. Koller, "Identification of Naval Spoofs," RRL Report No. 411–129, Nov. 29, 1944.

however, the lobes of the diffraction pattern may be so extremely fine that the aspect would have to be measured with extreme precision before any significant correlation could be attempted. Also, the slight differences in the ship's position from one roll to the next would mask any regular period. [In confirmation, Koller was able to find a regular repetition of fluctuations corresponding to the roll period on a much lower frequency (700 Mc/sec).]

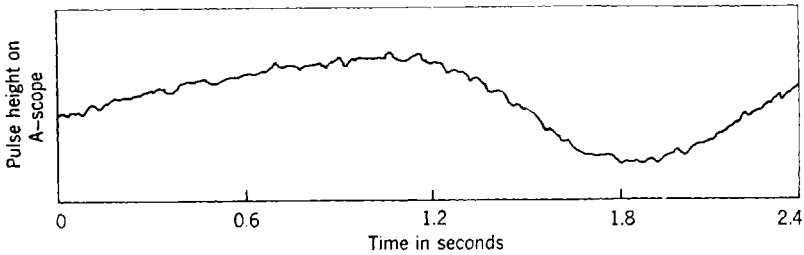


FIG. 6·32.—Echo from a battleship on 10 cm. (*Taken from RRL Report No.* 411-129.)

In addition, the varying nature of the sea reflection tends to mask any correlation with ship aspect.

The rates of fluctuation of ship echoes cover an extremely wide range. Figure 6·32, taken from Koller's report, shows a slow fade of several seconds in the echo from a battleship on 10 cm. The ordinate is height of signal on the A-scope, and the abscissa is time, up to 2.4 sec.[1] At the other ex-

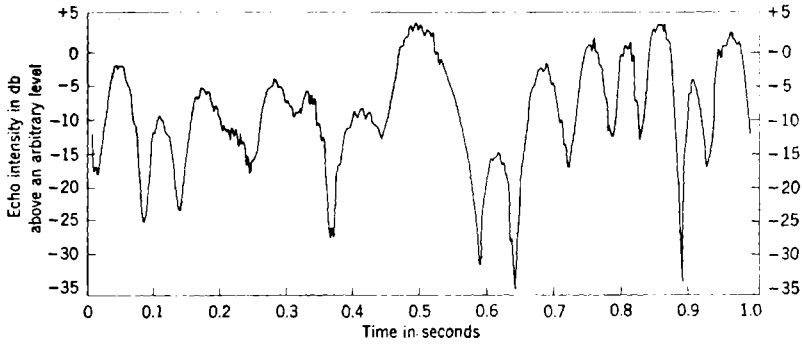


FIG. 6·33.—Plot of the r-f intensity of the echo from a battleship on 9.2 cm at a range of 5000 yd.

treme, Fig. 6·33 is a plot taken from a pulse-to-pulse film of a battleship echo on a 9-cm system with a pulse-recurrence frequency of 333 pps. The signal intensity is given in decibels referred to an arbitrary datum, and the time scale extends to 1.0 sec. Frequencies up to 20 cps are present in the

[1] No receiver calibration is given in the original report, but apparently the bottom line corresponds to the A-scope baseline and the top line to the saturation level of the receiver. The dynamic range is probably around 20 db.

record. It seems likely that ship echoes show such high rates of fluctuation only when the ship is turning relative to the radius vector from the radar. The fine lobes of the reradiation pattern then pass rapidly over the receiving antenna. Such is the situation in Fig. 6·33. Coales and Hopkins[1] have found fluctuations at rates of about 10 to 20 cps in the 10-cm echo from a destroyer turning at the rate of 0.8°/ sec. This is in rough agreement with the expected lobe structure from a target about 200 ft long.

The rate of fluctuation would be expected to increase linearly with r–f frequency. The only data on this point are those given by Koller (*op. cit.*), who found no significant difference between the rates at $\lambda = 10$ cm and at $\lambda = 40$ cm.[2] The lack of agreement was interpreted as indicating that the

FIG. 6·34.—Section of the pulse-to-pulse record of the echo from a battleship on 9.2 cm. Pulse-recurrence frequency 333 pps, pulse length 1 μsec.

effective reflecting surfaces are different at the two wavelengths, masts and rigging playing a proportionately larger role at 40 cm than at 10 cm.

In general, ships are sufficiently large structures so that the various reflecting surfaces are at significantly different ranges. If the receiver bandwidth is wide for the pulse length (i.e. much greater than $1/\tau$), then the echo pulse will have a complex shape, with different portions of the pulse often fluctuating independently of each other. This phenomenon is clearly illustrated in Fig. 6·34, which shows a strip of the pulse-to-pulse film from which Fig. 6·33 was obtained. The pulse length is 1 μsec, and the receiver bandwidth, about 3 Mc/sec. The complexity of the echo is one of the most convenient ways of identifying the signal as coming from ships.

As the reflecting elements on a vessel are so numerous and diverse, it has often been suggested that a ship may be considered simply as a group of independent scatterers whose echoes have random phases (*cf.* Sec. 6·5). The theory of the return from such a group is presented in detail in Sec. 6·19, where it is stated that the probability of obtaining an echo intensity between P and $P + dP$ is given by

$$W_1(P)\,dP = e^{-\frac{P}{P_0}}\frac{dP}{P_0}, \tag{139}$$

where P_0 is the average intensity. There are several a priori objections to this picture. As all reflecting objects are more or less rigidly connected,

[1] J. R. Coales and M. Hopkins, "Report on Trials to Determine the Variation of the Apparent Reflecting Point of Plane 10 cm Waves from a Destroyer," ASE Report No. M627, July 1944.

[2] The measurements, however, were not made simultaneously on the two frequencies.

the phases cannot change independently. Also there are likely to be a few surfaces whose contributions are far greater than the others. It seems reasonable, however, that this model should be a fair approximation to the actual case.

Cooper[1], Koller[2], and Wilkes, Ramsay, and Blow[3], and others have tried to verify the probability distribution [Eq. 139] or rather its integral. The published reports in general show fairly large deviations from the exponential law. With the exception of Cooper, the investigators have

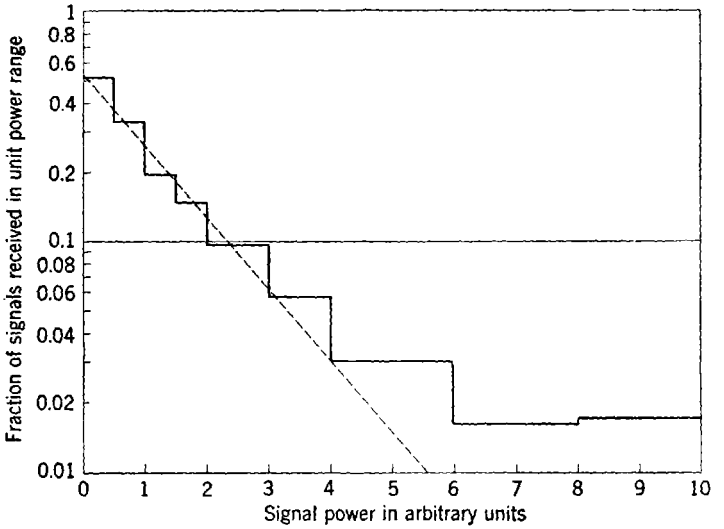


FIG. 6·35.—Probability distribution of the levels of 333 echoes from a battleship on 9.2 cm. The dotted line indicates the exponential formula predicted theoretically for random scatterers.

assumed linearity of the receivers and indicators, an assumption well known to be frequently erroneous. It is therefore difficult to tell if the deviations are significant. Cooper, whose work is the most extensive, used a signal generator for calibration and in general found good agreement with Eq. (139), although occasionally he found an excess of high-intensity signals. This difference is in line with the idea that a few large surfaces predominate. Corroborative evidence is to be found in the probability distribution of a battleship echo obtained from the data shown in 6·33. The results are plotted in Fig. 6·35 in histogram form, showing the number of pulses whose intensities are in a given range. The probability is plotted on a logarith-

[1] *Op. cit.*

[2] *Op. cit.*

[3] M. V. Wilkes, J. A. Ramsay, and P. B. Blow, "The Theory of the Performance of Radar on Ship Targets." Joint ADRDE and CAEE Report, July 27, 1944. CAEE Ref. 69/C/149, ADRDE Ref. RO4/2/CR252.

mic scale, as this is the most sensitive method for detecting differences from the exponential law. The points for low intensities follow a straight line fairly well, but the excess at high levels is plainly evident.

6·17. Interference Phenomena in Complex Targets.—The echo fluctuations predicted on the basis of the several mechanisms discussed in the previous sections are almost always small, so much so that they are usually second-order effects. Only in the case of targets viewed over water or isolated moving targets such as ships and aircraft has it been possible so far to account for the observed large fluctuations. The echoes in the general class of "ground signals" should, according to the above discussions, be fairly steady, and occasionally they are. The Provincetown, Mass., standpipe, as viewed from Race Point on Cape Cod, returns a very steady echo. The standard deviation of the pulse-to-pulse fluctuation on 9.2 cm was measured to be 0.27 db, which is only slightly larger than the fluctuation inherent in the signal-generator pulse as given in Sec. 6·13. Such cases are the exception, however. Most ground targets at one time or another will show large and often rapid variations.

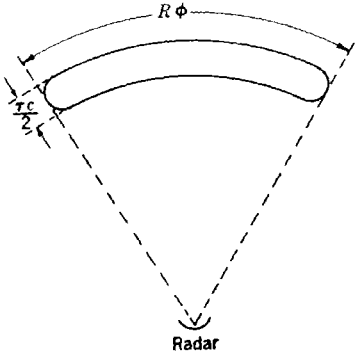


FIG. 6·36.—Illustration of the term "target area."

The explanation lies in the realization that a radar echo is almost never the reflection from a single target but rather from all the objects within a "target area." This area is defined by the beamwidth and radar pulse length (see Fig. 6·36); its size is given roughly by

$$\frac{R\phi r_c}{2}, \tag{140}$$

where R is the range, ϕ the azimuth beamwidth, and r the pulse duration. All objects in this region (and within the vertical extent of the beam) will contribute to the signal received at the radar at one given range on the indicators. Because the various portions of the beam are coherent, the contributions from each of the component targets must be added vectorially, taking into account the phase of each echo. The total echo is thus the result of interference between echoes from a number of targets that may be independent and unconnected. The phase of these echoes provides a sensitive means for changing the interference, for in the microwave region motions of only a few inches are sufficient to change the relative phase by π or more radians. *The swaying of trees in the wind, the motion of vehicles on a road, even the spinning of rooftop ventilators all affect the phase pattern and cause the echo to fluctuate.*

Let us examine in detail some simple examples of this interference mechanism. The Provincetown standpipe was mentioned above as a target with an extremely steady echo. Figure 6·37 shows the profile of the path from the radar to the target. The series of hills (really sand dunes) along the path acts as a screen between the radar and all but the top portion of the standpipe.[1] Effectively then, the target area contains only one target, this being a very rigid and solid structure. The echo consequently shows negligible fluctuation. Similar cases have been observed where intervening hills screen all but one isolated target, and the echo has invariably been steady. On the other hand, in cases where the target area contains many trees, houses, etc., in addition to a prominent object such as a standpipe or water tank, the echo will often fluctuate badly.

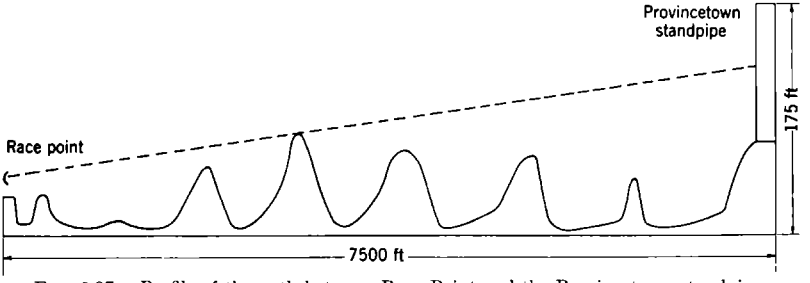


FIG. 6·37.—Profile of the path between Race Point and the Provincetown standpipe.

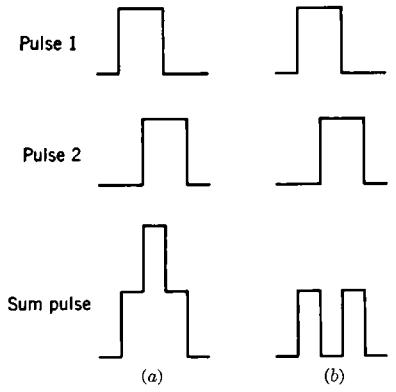


FIG. 6·38.—Interference of two equal point scatterers. (*a*) Constructive interference; (*b*) destructive interference.

The next most simple situation involves not one but two targets in the target area. Suppose the area contains two equal point scatterers separated d in range. Ideally, the echo from the nearer scatterer can be represented by the top pulse of Fig. 6·38*a*; the echo from the second one, just below it, would arrive at a time $2d/c$ later. The total signal received would be the sum of these two, and its appearance would depend on the relative phase of the two echoes. If they were in phase, then the echoes would interfere constructively in the overlapping region to produce an echo as shown at the bottom of Fig. 6·38*a*. If they were out of phase by π radians, there would

[1] The portion of the standpipe illuminated because of diffraction can be shown to be fairly small

be destructive interference and the echoes would cancel as in Fig. 6·38*b*.

An actual target closely resembling this idealized picture has been found in the antenna towers of radio station WBZ at Hull, Mass., as viewed from Deer Island, 6 miles away (see Figure 4·1). The target consists of two 500-ft towers supported by guy wires up to 400 ft and located about 60 yd apart in range. A 100-ft hill in the line of sight and 2 miles away ensures that only the towers are seen by the radar. The right-hand signal in the 16-frame-per-second strip reproduced in Fig. 6·39 shows the appearance of the echo on a 9-cm system with a pulse length of 1 μsec. Both constructive and destructive interference are well illustrated, along with the evolution of the intermediate stages.[1] The differences between these photographs and Fig. 6·38 are all attributable to the finite receiver bandwidth and to the fact that the towers are not exactly equal scatterers.

The phase between the two component echoes can be changed in several ways to produce the various degrees of interference observed. Thus the index of refraction in the region between the towers and, hence, the optical path length might change. For a phase shift of π radians the index would be required to vary by $\Delta n = \lambda/4d$. In this case, at $\lambda = 9$ cm, Δn would be 4×10^{-4}, which is even larger than the average value of $(n - 1)$. A change in the transmitter frequency could also produce the necessary phase shift, but we have already seen in Sec. 6·13 that the required frequency change is outside the limits of the transmitter instability. The most plausible explanation, then, is that there is a physical motion of the towers relative to each other and a consequent variation of the phase. If the wavelength is 10 cm, the relative motion that would shift

[1] The left-hand signal was the output of a narrow-band receiver on a separate antenna 25 ft. away from the radar. The striking diversity effect thus shown arises from the slightly different aspects of the towers at the two locations.

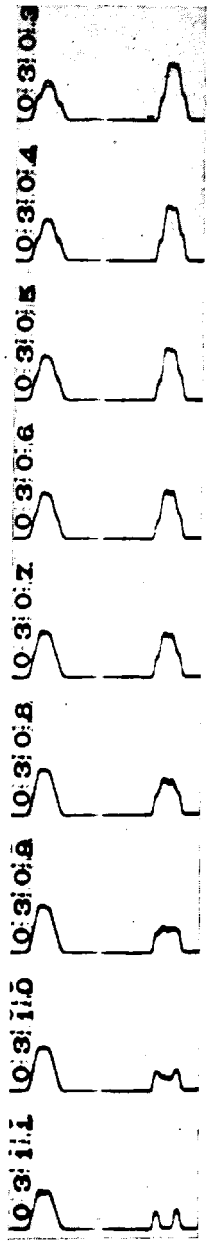


FIG. 6·39.—The right-hand signal is the echo from WBZ radio towers photographed at intervals of $\frac{1}{16}$ sec. (Left-hand echo is from the same target on a separate diversity receiver 25 ft away.)

the phase by π radians is only 1 in. It is plausible that the tall towers would sway to an even greater extent even in fairly gentle winds.

To test this conclusion, a 4-day continuous watch was kept on this target with a 9-cm system. At half-hour intervals the maximum variation in signal strength within a 1-min period was noted. These measurements were compared with the surface wind speed at the time (as recorded at a spot 5 miles from the target). A definite correlation was found between the amount of fluctuation and the wind speed. By means of the chi-square test[1] the probability of obtaining this distribution of measurements by chance, if the fluctuation were really independent of wind speed, was calculated to be 1×10^{-7}.

The examples discussed so far have been simple ones and are rather artificial and met with infrequently. Of greater practical importance is the case where the target area contains a large number of scatterers, not merely one or two. The entire class of so-called "clutter" echoes—"chaff," sea echo, precipitation echoes, and ground clutter—falls within this group of targets. Because the area contains a large number of component targets, echoes of this type can be treated only statistically. The remaining sections of this chapter are devoted to a discussion of theoretical and experimental investigations of these echoes.

THE FLUCTUATIONS OF CLUTTER ECHOES

By HERBERT GOLDSTEIN

6·18. The Nature of Clutter Echoes.—"Clutter" is a term used to denote the radar echo from an extended target covering a region much larger than the target area. The target consists of many individual targets more or less independent of each other and sufficiently numerous that any given target area usually contains a large number of them. This class of echoes includes chaff, sea, and precipitation echoes and ground clutter.

"Chaff" is a code name for a confusion type of interference scheme that uses large numbers of resonant dipoles.[2] The dipoles are usually narrow strips of tin foil, cut to resonate at the proper wavelength and dropped in huge quantities from aircraft. When properly distributed they return an echo covering a large volume of space. At microwave frequencies a single package may contain hundreds of thousands of dipoles.

It is obvious that the target area will include a large number of the elementary targets and that the echo phases, as determined by the radial distance from the radar, will be random. That is to say, the number of dipoles with any given phase will be the same as those with any other phase. In

[1] See T. C. Fry, *Probability and Its Engineering Uses*, Van Nostrand, New York, 1928.

[2] The term *window*, often used in the same connection, is a generic term of British origin.

general the chaff cloud will move with the wind, but in a noncoherent system such mass motion, which changes all phases alike, has no effect on the returned signal. In addition, the dipoles will move relative to each other because of the way they fall and because of local turbulences. This relative motion will change the relative phases of the elementary echoes, causing the resultant signal to fluctuate. The faster the motion, the more rapid the fluctuations. An alternative and very fruitful way of looking at the rates of fluctuation is to consider them as the differences between Doppler frequencies from the various targets. All targets having the same characteristics as a uniform cloud of chaff will be called here an *assembly of independent random scatterers* and often abbreviated to *random scatterers*.

The nature of *precipitation echo*, or *storm echo*, is discussed in Chap. 7, where it is shown that the signal is caused by scattering either from water drops or from water in a solid form, that is, ice, snow, etc. Here, too, the target area contains a large number of elementary scatterers with randomly distributed phases. Furthermore, because of turbulences within a precipitation area there will be relative motion of the scatterers, causing fluctuations in the total signal. This type of target therefore also belongs in the class of assemblies of independent random scatterers.

It has already been pointed out that there is still much debate over the specific mechanism responsible for *sea echo*. Present evidence supports the view that the echo is caused by scattering either from small ripples on the sea surface or from suspended spray drops. In either case the target area contains many such ripples or drops, and it is not implausible that the conditions for random scatterers would again be satisfied. This will be assumed to be true, and the consequences of the assumptions examined. Any deviations from the behavior expected on this basis will be used to obtain further information about the fundamental scattering mechanism involved.

A bewildering variety of targets other than man-made objects—rocks, earth, tree trunks, branches, bushes, leaves, grass, for example—may give rise to *ground clutter*. Some of these targets return steady echoes, and some move in the wind, changing the interference between the component echoes and causing the total signal to fluctuate. It is a common observation that the harder the wind blows the more rapid are the fluctuations of ground clutter. A quantitative demonstration of this phenomenon was obtained with the following experiment.

At intervals in a 24-hr period, motion pictures at the rate of 16 frames per second were taken of the A-scope on a 9-cm system, showing the echo from a portion of the Blue Hills Reservation near Boston, Mass. The target area chosen is heavily wooded and almost free from man-made objects. During this 24-hr period the surface wind speed, as measured about 2 miles from the target, varied from over 30 mph to dead calm. Two hundred frames were chosen in each run, and the echo height in each frame measured and converted into relative r–f intensity in decibels. As a

measure of the fluctuation, the difference in the decibel values for two consecutive frames was then found and averaged over the 200 frames. The values so obtained are plotted against wind speed in Fig. 6·40 indicated by the circles.

At low wind speeds the fluctuation is only about 0.2 db, the same order of magnitude as for the steadiest targets. At the other extreme, with wind speeds of 30 mph or more, the fluctuation is fifteen times as great. It should be noticed that the change in the signal from one frame to the next

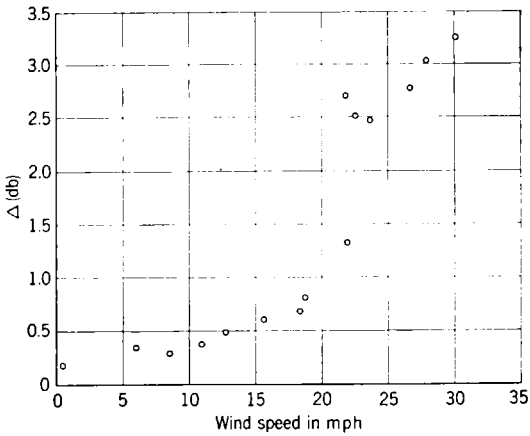


FIG. 6·40.—Average fluctuation of a ground clutter echo during 0.06 sec, plotted against wind speed.

fails to be a good measure of the fluctuation when the period of fluctuation becomes comparable to the exposure time for a single frame ($\frac{1}{32}$ sec). Such a failure occurs at the higher wind speeds, and the actual fluctuations are larger than indicated on the figure.

The fluctuation seems to increase rapidly in the neighborhood of 20 mph. It may possibly be significant that this corresponds to the wind speed at which small trees and large branches begin to move as a whole, according to the Beaufort wind scale. The jump would then indicate that there is no substantial change in the phase relationships until this large motion takes place.

There is no sharp dividing line, of course, between stationary and moving targets. Almost all objects making up ground targets will move to some extent. Nevertheless it seems feasible to separate the targets into two classes, those which move less than $\lambda/4$ and those which move in the wind by much more than $\lambda/4$. The former class of targets gives rise to a steady echo; the latter fulfills fairly well the conditions for an assembly of independent random scatterers. The total signal is thus the sum of these two types of echo. The membership of the classes will depend on the wind

speed and the wavelength. At the high wind speeds only rocks and heavy tree trunks could be classified as stationary. A target that moves only $\frac{1}{2}$ in. would be considered at rest for 9 cm but would certainly be in the assembly of random scatterers for a wavelength of 1 cm. While data at wavelengths other than 9 cm are scarce, it has been noted that ground clutter at 1.25 cm shows large and rapid fluctuations at wind speeds far smaller than those corresponding to the "jump" at 9 cm.

6·19. The Theory of Clutter Fluctuations.—The preceding section has shown that the targets responsible for clutter echoes generally fall in one of two classes: either an assembly of independent random scatterers or such an assembly plus a number of stationary targets. The description of the echo from these types of targets forms a branch of the general theory of random processes dating back to Rayleigh[1] and which has undergone considerable development in recent years. Here only the results of the theory will be given; derivations will be found in Chap. 6 of Vol. 24 of the Radiation Laboratory Series.[2] The two types of targets will be treated separately.

Assemblies of Independent Random Scatterers.—The received field strength will be of the form

$$E = V(t) \cos [\omega_0 t + \phi(t)], \tag{141}$$

where ω_0 is the transmitter frequency. The functions $V(t)$ and $\phi(t)$ will depend on time chiefly because of the variation of the signal characteristics with range and will be roughly periodic with the pulse-recurrence frequency. Only the echo from a given target area, that is, at a given range, is of interest, however. If a small interval of time is chosen centered at a given range and short compared with the pulse duration but long compared with the r-f period, then within that interval V and ϕ will be constant, but they will still vary from one sweep to the next. It is with this variation in time and not with the variation from the changes with range that this section will be concerned. Although the quantities $V(t)$ and $\phi(t)$ are thus defined only at discrete times that are multiples of the repetition period, they will still be considered as continuous functions of time.

From Eq. (141) it is seen that the signal can be completely specified by a vector $\mathbf{R}(t)$ whose magnitude is equal to V and which makes an angle with some reference direction equal to the phase ϕ. The Cartesian coordinates of this vector will be called X and Y. The term $\mathbf{R}(t)$ can be written as the vector sum

$$\mathbf{R} = \sum_j \mathbf{r}_j, \tag{142}$$

[1] *Cf.* the two-dimensional "random walk" problem in Lord Rayleigh, *Theory of Sound*, 2d ed., Vol. I, Macmillan, London, 1894, p. 35.

[2] See also, S. O. Rice, *Bell System Tech. Jour.*, **23,** 282 (1944), and **24,** 46 (1945); Ming Chen Wang and G. E. Uhlenbeck, *Rev. Mod. Phys.*, **17,** 323 (1945); and S. Chandreskhar, *ibid.*, **15,** 1 (1943).

where $\mathbf{r}_j$ is the contribution to the total signal of one of the individual targets making up the target area. The components of $\mathbf{r}_j$ will be denoted by x_j and y_j.

It is now possible to describe in more mathematical terms the conditions that must be fulfilled by an assembly of independent random scatterers:

1. There must be a large number of scatterers in the target area. What constitutes a large number will be discussed below.
2. The phases of the component echoes must be random; that is,

$$\overline{x_j} = \overline{y_j} = 0, \tag{143}$$

where the bar denotes averaging over time.

3. The probability of receiving an echo $\mathbf{r}_j$ from the jth target $w_j(\mathbf{r}_j)$ must be independent of the signal from the kth target. That is, if $w(\mathbf{r}_j, \mathbf{r}_k)$ is the probability of receiving echoes $\mathbf{r}_j$ and $\mathbf{r}_k$ from the jth and kth scatters respectively, then

$$w(\mathbf{r}_j, \mathbf{r}_k) = w_j(\mathbf{r}_j)w_k(\mathbf{r}_k). \tag{144}$$

Under these conditions the probability that any given measurement of $\mathbf{R}$ lies between $\mathbf{R}$ and $\mathbf{R} + d\mathbf{R}$ is given by[1]

$$W_1(\mathbf{R})\, d\mathbf{R} = \frac{1}{\pi P_0} e^{-\frac{X^2+Y^2}{P_0}} dX\, dY, \tag{145}$$

where P_0 is the time average of $\mathbf{R}^2$. Following Uhlenbeck, W_1 will be referred to as the *first probability distribution* because it deals with the value of $\mathbf{R}$ at only one time. The shape of this distribution is *independent of the distribution of the component amplitudes*, providing only that there are a large number of them, strictly speaking, an infinite number. Thus there is no requirement that the individual $\mathbf{r}$'s be equal or even that the amplitudes be constant with time but only that the distribution of $\mathbf{r}$ be stationary in time so that $\overline{\mathbf{R}^2}$ $(= \Sigma\overline{\mathbf{r}^2})$ shall be constant. It is, in fact, a direct consequence of the central-limit theorem of probability theory that W_1 approaches the form given in Eq. (144) as the number of scatterers increases toward infinity, independent of the distribution of the $\mathbf{r}$'s.

The rate at which the limiting form is approached does depend on the distribution. Thus Eq. (145) holds exactly for a finite number of targets if the distribution for $\mathbf{r}$ is Gaussian. Goudsmit[2] and Schiff[3] have studied the case of finite number of scatterers for certain special distributions. If all the $\mathbf{r}$'s are equal, then Eq. (145) is a good approximation, even if $n = 4$ or 5. When the number is 10 or more, the differences are completely negligible.

[1] Henceforth this type of expression will be shortened to "$\mathbf{R}$ in range $d\mathbf{R}$."

[2] S. A. Goudsmit, RL Report No. 196, March 4, 1943.

[3] L. I. Schiff, unpublished memorandum, June 5, 1943.

Equation (145) defines a bivariate Gaussian distribution centered about $X = 0$, $Y = 0$. *The most probable value of* $\mathbf{R}$ *is therefore zero.* The first probability distribution in amplitude $\mathbf{R}$ is given by

$$W_1(R)\,dR = \frac{2}{P_0}\,e^{-\frac{R^2}{P_0}}\,R\,dR. \tag{146}$$

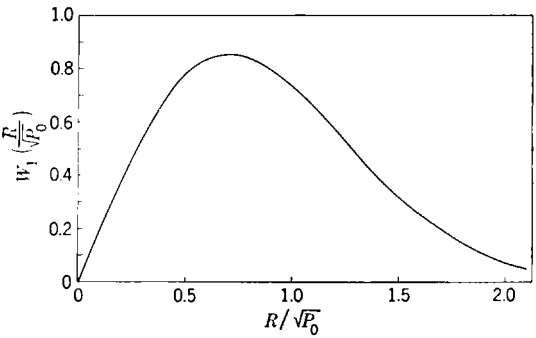


FIG. 6·41.—First probability distribution in amplitude for ramdon scatterers.

Figure 6·41 is a plot of this distribution. The most probable value of the amplitude is equal to 0.707 of the rms value, $\sqrt{P_0}$. The average value of R is $\sqrt{\pi/2}$ times the rms value. Let P denote $\mathbf{R}^2$, which is proportional to the signal energy. Then the distribution in P is

$$W_1(P)\,dP = e^{-\frac{P}{P_0}}\frac{dP}{P_0}$$

and is shown in Fig. 6·42. It should be noticed that the most probable value of P is zero.[1]

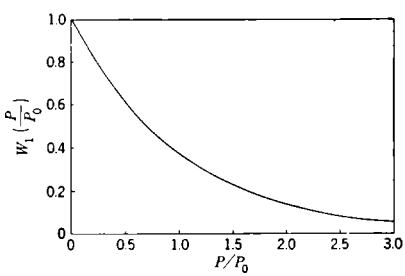


FIG. 6·42.—First probability distribution in power for random scatterers.

From these distributions the fluctuations in the signal are seen to be fairly large. The signal power will be less than one-half the average level 29 per cent of the time but greater than twice the average 14 per cent of the time and above five times the average 1 per cent of the time. In fact, although the most probable value of P is zero, there is a finite probability for any value of P, no matter how large. A commonly used measure of the amount of fluctuation is the standard deviation

$$\frac{\sqrt{(P - P_0)^2}}{P_0} = \frac{\sqrt{P^2 - P_0^2}}{P_0}. \tag{148}$$

[1] A similar situation occurs in the kinetic theory of gases. With the Maxwellian distribution the most probable *velocity* is zero, but the most probable *speed* is different from zero.

From Eq. (146) one obtains

$$\overline{P^2} = 2P_0^2, \tag{149}$$

and hence the standard deviation is 100 per cent. It should be emphasized again that because of the randomness of the component vectors the amplitude of fluctuation does not depend at all upon the number of scatterers, provided that there are many of them.

The first probability distribution tells us about the magnitude of the fluctuation. Information about the *rate* of fluctuation is obtained from the second probability distribution, which gives the probability of measuring $\mathbf{R}_1$ in range $d\mathbf{R}_1$ at time t_1 and measuring $\mathbf{R}_2$ in range $d\mathbf{R}_2$ at time t_2. As we are assuming stationary processes, time enters in only as the difference $\tau = t_2 - t_1$. If the scatterers are moving independently, it can be shown[1] that the second probability distribution is

$$W_2(\mathbf{R}_1, \mathbf{R}_2, \tau)\, d\mathbf{R}_1\, d\mathbf{R}_2 = \frac{dX_1\, dX_2\, dY_1\, dY_2}{\pi^2 P_0^2 (1 - g^2)} \times \exp\left[- \frac{X_1^2 + Y_1^2 + X_2^2 + Y_2^2 - 2g(X_1X_2 + Y_1Y_2)}{P_0(1 - g^2)} \right] \tag{150}$$

where g is the correlation coefficient for the Cartesian coordinates of $\mathbf{R}$ and is defined by

$$g = \frac{\overline{X_1X_2}}{\overline{X_1^2}} = \frac{\overline{Y_1Y_2}}{\overline{Y_1^2}}. \tag{151}$$

The correlation coefficient is obviously a function of the time interval τ and is 1 if $\tau = 0$ and approaches zero as τ increases to infinity and $\mathbf{R}_1$ and $\mathbf{R}_2$ become independent of each other.

It has already been pointed out that the rate of fluctuation depends upon the relative motion of the contributing scatterers or, more precisely, on the component of the relative velocities in the direction of the radar system. Siegert[2] has derived the relation connecting the correlation function g and the distribution of relative velocities

$$g(\tau) = \frac{\left| \int_0^\infty r^2\, dr \int_{-\infty}^{\infty} w(r,v_z)\, e^{4\pi i v_z \tau/\lambda}\, dv_z \right|}{\int_0^\infty r^2\, dr \int_{-\infty}^{\infty} w(r,v_z)\, dv_z}, \tag{152}$$

where v_z is the component of the velocity of the scatterer in the direction of the radar system and $w(r,v_z)\, dr\, dv_z$ is the probability of finding a scatterer that returns a signal of amplitude r in range dr and with a velocity component v_z in range dv_z. In addition to the previous conditions Eq. (152) assumes (1) that the amplitudes of the individual echoes are constant in time, that is, the fluctuation is caused only by variations in phase and (2) that phase changes arise solely from shifts in positions of the scatterers.

[1] See Vol. 24, Chap. 6.

[2] A. J. F. Siegert, RL Report No. 773, January 1946.

Equation (152) can be written more simply as

$$g(\tau) = \frac{\overline{|r^2 e^{4\pi i v_2 \tau/\lambda}|}}{\overline{r^2}}. \tag{152a}$$

The time interval τ occurs only in the ratio τ/λ. Hence if simultaneous measurements are made on the same target but at different wavelengths, the correlation functions plotted as a function of τ/λ should be identical. Another way of saying it is that the rate of fluctuation should be proportional to the transmitter frequency, which is obvious from the interpretation of the fluctuation as Doppler frequencies. The fluctuation frequency spectra on several wavelengths should coincide when plotted against $\nu\lambda$.

Equation (150) can be conveniently rewritten in terms of the amplitudes and phases of the two vectors

$$W_2(\mathbf{R}_1,\mathbf{R}_2,\tau)d\mathbf{R}_1\, d\mathbf{R}_2 = \frac{1}{\pi^2 P_0^2(1-g^2)} \exp\left[-\frac{R_1^2 + R_2^2 - 2gR_1R_2\cos\theta}{P_0(1-g^2)}\right] R_1\, dR_1\, R_2\, dR_2\, d\theta\, d\phi_1, \tag{153}$$

where θ is the angle between $\mathbf{R}_1$ and $\mathbf{R}_2$. The second probability distribution for P_1 and P_2 is obtained by integration over θ and ϕ_1,

$$W_2(P_1,P_2,\tau)\, dP_1\, dP_2 = e^{-\frac{P_1+P_2}{P_0(1-g^2)}} J_0\left(\frac{2ig\sqrt{P_1P_2}}{P_0(1-g^2)}\right)\frac{dP_1\, dP_2}{P_0^2(1-g^2)}, \tag{154}$$

where J_0 is the zeroth-order Bessel function. From this distribution one can in turn calculate the correlation function for P, defined by

$$\rho(P,\tau) = \frac{\overline{(P_1 - P_0)(P_2 - P_0)}}{\overline{(P_1 - P_0)^2}} = \frac{\overline{P_1P_2} - P_0^2}{\overline{P_1^2} - P_0^2} = \frac{\overline{P_1P_2} - P_0^2}{P_0^2}; \tag{155}$$

and it is found that

$$\rho(P,\tau) = g^2. \tag{156}$$

In a similar fashion the distributions and correlation functions of other functions of $\mathbf{R}_1$ and $\mathbf{R}_2$ may be calculated. Thus the correlation function for R has been found by Uhlenbeck[1] to be

$$\rho(R,\tau) = \frac{2E(g) - (1-g^2)K(g) - \frac{\pi}{2}}{2 - \frac{\pi}{2}}, \tag{157}$$

where $E(g)$ and $K(g)$ are the complete elliptic integrals of the first and second kind, respectively. Sollfrey and Barlow[2] have calculated the

[1] G. E. Uhlenbeck, RL Report No. 454, Oct. 15, 1943.

[2] Unpublished memorandum by W. Sollfrey and E. Barlow.

correlation coefficient for $\ln R$:

$$\rho(\ln R,\tau) = \frac{6}{\pi^2}\sum_{k=1}^{\infty}\frac{g^{2k}}{k^2}. \tag{158}$$

A quantity that is often of practical interest is the ratio P_2/P_1, denoted by q. The distribution of q can be found from

$$W_2(q,\tau)\,dq = \int_0^{\infty} W_2(q,P_1,\tau)P_1\,dP_1\,dq. \tag{159}$$

Using Eq. (153) and the integral

$$\int_0^{\infty} xe^{-ax}J_0(bx)\,dx = \frac{a}{(a^2+b^2)^{\frac{3}{2}}},$$

the distribution is found to be

$$W_2(q,\tau)\,dq = (1-g^2)\frac{(1+q)\,dq}{[(1+q)^2-4g^2q]^{\frac{3}{2}}}. \tag{160}$$

It is more convenient to use the distribution with q expressed in decibels. Let $y = 10\log_{10} q$. Then

$$W_2(y,\tau)\,dy = \frac{0.230\,(1-g^2)\,e^{0.230y}\,(1+e^{0.230y})\,dy}{[e^{0.460y}+2(1-2g^2)\,e^{0.230y}+1]^{\frac{3}{2}}}. \tag{161}$$

Equation (161) will be used as a convenient method of verifying the second probability distribution experimentally.

The first and second probability distributions describe completely the fluctuations of the echo. The correlation function g by itself contains all needed information on the rates of fluctuation. It is often desirable, however, to present this information in the form of a frequency spectrum. A relation exists, known as the Wiener-Khintchine theorem[1], that links the correlation function with a suitably defined frequency spectrum.

Let $z(t)$ be any continuous function arising from a stationary random process, and let $f(\omega)$ be its complex frequency spectrum as obtained from the Fourier integral theorem. Then the normalized power spectrum of $z(t)$ is defined as

$$F(\omega) \equiv \frac{|f(\omega)|^2}{\int_0^{\infty}|f(\omega)|^2\,d\omega}. \tag{162}$$

The correlation function of $z(t)$ is given by

$$\rho(z,\tau) = \frac{\overline{[z(t_1)-\bar{z}]\,[z(t_1+\tau)-\bar{z}]}}{\overline{[z(t_1)-\bar{z}]^2}}. \tag{163}$$

[1] See Ming Chen Wang and G. E. Uhlenbeck, *op. cit.*, p. 326.

Then it can be shown that $F(\omega)$ and $\rho(z,\tau)$ are cosine transforms of each other, that is,

$$F(\omega) = \frac{2}{\pi}\int_0^\infty \rho(z,\tau)\cos\omega\tau\, d\tau, \tag{164}$$

and

$$\rho(z,\tau) = \int_0^\infty F(\omega)\cos\omega\tau\, d\omega. \tag{165}$$

Hence the frequency spectrum can always be obtained from the correlation function, and vice versa.

In all practical situations the signal in the receivers is the sum of the radar echo and a certain amount of receiver noise. In such cases the total signal can be described by the sum

$$\mathbf{R} = \sum_j \mathbf{r}_j + \mathbf{N},$$

where $\mathbf{N}$ is a vector representing the noise envelope.[1] It is well known that the noise distributions are identical with those already given[2] for $\mathbf{R}$. However, because the smallest time interval of interest here is the time between sweeps, the correlation function of the noise vector can be assumed to be zero. The first probability distribution for the combined vector is given by

$$W_1(\mathbf{R})\, d\mathbf{R} = \int W_1(\mathbf{N})\, W_1(\mathbf{R} - \mathbf{N})\, d\mathbf{N}\, d\mathbf{R}. \tag{166}$$

Similarly, the second probability distribution may be found from

$$W_2(\mathbf{R}_1,\mathbf{R}_2,\tau)\, d\mathbf{R}_1\, d\mathbf{R}_2 = \int\int W_1(\mathbf{N}_1)\, W_1(\mathbf{N}_2)\, W_2(\mathbf{R}_1 - \mathbf{N}_1,\, \mathbf{R}_2 - \mathbf{N}_2,\tau)\, d\mathbf{N}_1\, d\mathbf{N}_2\, d\mathbf{R}_1\, d\mathbf{R}_2. \tag{167}$$

On carrying out the integrations it is found that the introduction of noise has not affected the shape of the distributions but that P_0 and g are transformed as follows:

$$\begin{aligned} P_0 &\to P_0 + N_0^2 = P_0(1+s), \\ g &\to \frac{g}{1 + \dfrac{N_0^2}{P_0}} = \frac{g}{1+s}, \end{aligned} \tag{168}$$

where N_0 is the rms noise voltage and s is the ratio of noise power to echo power. By means of these relations the measured values of g can be corrected for the presence of noise in order to obtain the true correlation function of the echo alone.

[1] It is assumed that the noise is added in one of the linear stages of the receiver.

[2] See S. O. Rice, *op. cit.*

Random Scatterers plus a Steady Signal.—If the echo from an assembly of independent random scatterers is added to a constant echo, then the total signal will be represented by a vector

$$\mathbf{R} = \mathbf{S} + \sum_j \mathbf{r}_j. \tag{169}$$

The sum $\sum \mathbf{r}_j$ represents the contributions from the random scatterers, and $\mathbf{S}$ the signal that is assumed constant in time. The first and second probability distributions for $\mathbf{R}$ obviously differ from those given above only in that they are now functions of $\mathbf{R} - \mathbf{S}$. Thus the analogue of Eq. (145) is now

$$W_1(\mathbf{R})\, d\mathbf{R} = \frac{1}{\pi P_0} e^{-\frac{(X-S)^2+Y^2}{P_0}} dX\, dY, \tag{170}$$

where, without loss of generality, the x-axis has been taken in the direction of $\mathbf{S}$. The quantity P_0 is no longer the mean square value of $\mathbf{R}$ but only the mean square of the contribution from the random scatterers. Instead, we have

$$\overline{R^2} = \overline{P} = P_0 + S^2. \tag{171}$$

The first probability distribution for P becomes[1]

$$W_1(P)\, dP = e^{-\frac{P+S^2}{P_0}} J_0\left(2i\,\frac{S}{P_0}\sqrt{P}\right)\frac{dP}{P_0}. \tag{172}$$

If we denote by m^2 the ratio S^2/P_0, which can be called the ratio of the steady to random average power, then the distribution can be rewritten as

$$W_1(P)\, dP = (1+m^2)\, e^{-m^2} e^{-\frac{P}{\overline{P}}(1+m^2)} J_0\left(2im\sqrt{1+m^2}\sqrt{\frac{P}{\overline{P}}}\right)\frac{dP}{\overline{P}}. \tag{173}$$

In Fig. 6·43 this distribution is plotted against $P/\overline{P}$ for several values of m^2. For $m^2 = 1$ or less there is not much difference from the simple exponential that holds for random scatterers. The most probable value of P is still zero, and the range of fluctuation still very large. As m^2 is increased above 1, a maximum appears in the distribution and the fractional fluctuation decreases. When m^2 becomes very much greater than 1, the curve has the form of a sharp peak about $P = \overline{P}$ and approaches asymptotically the Gaussian distribution

$$W_1(P)\, dP \approx (1+m^2)\,\frac{e^{-\frac{(R-S)^2}{\overline{P}}(1+m^2)}}{m\sqrt{4\pi}}\,\frac{dP}{\overline{P}}. \tag{174}$$

[1] This equation is identical with the well-known formula for the distribution in power of noise plus signal.

The standard deviation of $W_1(P)$ for all values of m is

$$\frac{\sqrt{\overline{(P-\bar{P})^2}}}{\bar{P}} = \frac{\sqrt{1+2m^2}}{1+m^2} \tag{175}$$

and decreases as $\sqrt{2}/m$ when m^2 is much greater than unity.

The second probability distribution in **R** can similarly be immediately written down:

$$W_2(\mathbf{R}_1, \mathbf{R}_2, \tau)\, d\mathbf{R}_1\, d\mathbf{R}_2 = \frac{dX_1\, dX_2\, dY_1\, dY_2}{\pi^2 P_0^2(1-g^2)} \times$$

$$\exp\left\{-\frac{(X_1-S)^2+Y_1^2+(X_2-S)^2+Y_2^2-2g[(X_1-S)(X_2-S)+Y_1Y_2]}{P_0(1-g^2)}\right\} \tag{176}$$

The quantity g is now the correlation function for Y and for X-S. The relation between g and the velocity distribution of the random scatterers remains unchanged, of course.

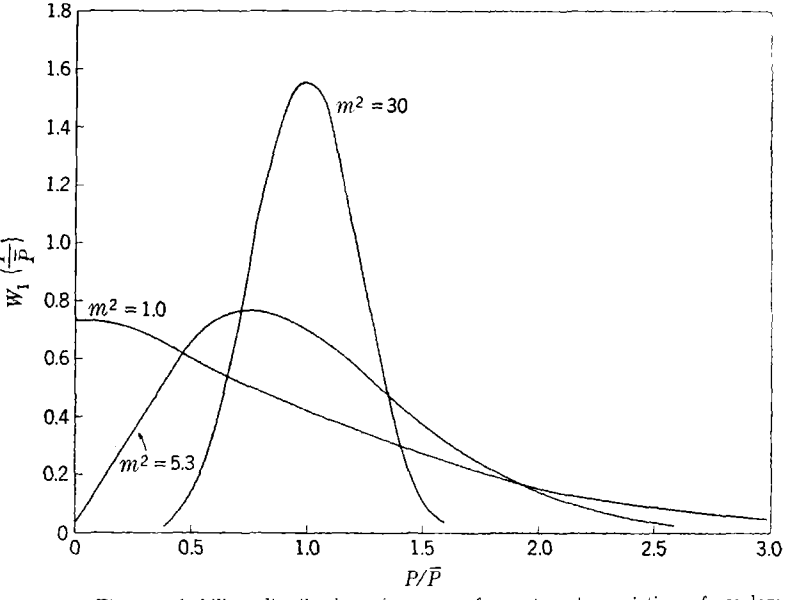


FIG. 6·43.—First probability distributions in power for a target consisting of random scatterers plus a fixed target, for several values of m^2.

It has not yet been possible to carry out the integrations needed to obtain $W_2(P_1, P_2, \tau)$. Some of the moments of the distribution can be found, however. Thus the average value of $\overline{P_1P_2}$ has been determined by Barlow to be

$$\overline{P_1P_2} = P_0^2\,[1 + g^2 + 2m^2(1+g^2) + m^4]. \tag{177}$$

The correlation coefficient of P is therefore

$$\rho(P,\tau) = \frac{g^2 + 2m^2 g}{1 + 2m^2}. \tag{178}$$

Many of the relations derived for random scatterers remain valid, as for example, the Wiener-Khintchine theorem and the effect of noise on P_0 and g.

6·20. Experimental Techniques in the Study of Clutter Fluctuations.— The preceding sections have presented a theoretical picture of the nature of clutter fluctuations. This general picture has been confirmed and further details obtained in an extensive experimental investigation of clutter fluctuations. Measurements were made on three experimental truckborne systems, operating on wavelengths of 9.2, 3.2, and 1.25 cm, respectively. Special care was taken to reduce the intrinsic fluctuations of the radar systems to negligible levels, as discussed in detail in Sec. 6·13.

To cover the entire spectrum of echo fluctuations, measurements of the signal level must be made as often as possible, namely, each time an echo is received. Of the several techniques available for making such measurements,[1] photography of the individual sweeps on the A-scope was chosen as furnishing the largest amount of information. The methods developed for this type of photography are described in detail elsewhere.[2] A blue-screen short-persistence cathode-ray tube running at potentials slightly higher than customary was used in conjunction with a high-speed camera in which the film moved continuously to prevent overlapping of successive traces. It was possible to record traces with writing speeds as high as 70 cm/μsec and at rates up to 4000 traces per second.

The basic pulse-recurrence rate was about 1000 pps, accurately maintained by a quartz-crystal-controlled oscillator. By "counting down" from this frequency the A-scope could be triggered at a number of rates down to 42 pps. If the echo fluctuation was slow enough to permit even more infrequent measurements, conventional motion-picture cameras were employed at various speeds down to four frames per second.

Figure 6·44 reproduces a strip from a typical film, showing sea echo on the 9.2-cm system. The interval betweem traces is 3 msec, and the sweep length is 1500 yd, the most frequently used sweep. For convenience in reference, each frame area is provided with an identification number, obtained by photographing a synchronized counter prior to photographing the A-scope. The prominent pulse on the right is an artificial echo pulse,

[1] Specialized methods have been developed for obtaining specific types of information, such as the superheterodyne audio analyzer for frequency spectrum (*cf.* RL Report No. 914) and the "fading analyzer" discussed by R.I.B. Cooper, "The Fading of S-band Echoes from Ships in the Optical Zone," RRDE Report No. 265, December 1944.

[2] H. Goldstein and P. D. Bales, "High Speed Photography of the Cathode-ray Tube," *Rev. Sci. Instruments*, **17**, 89 (1946).

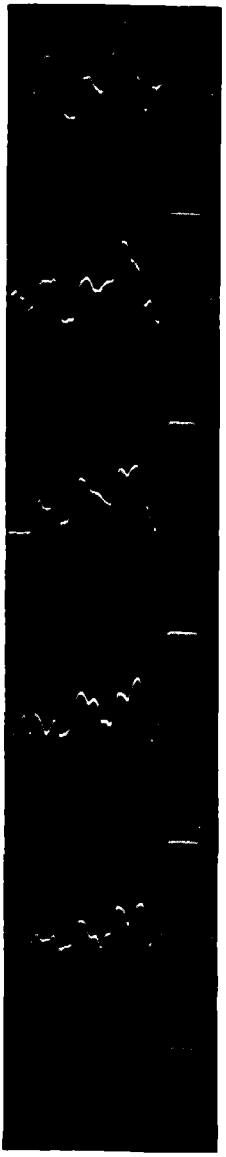


FIG. 6·44

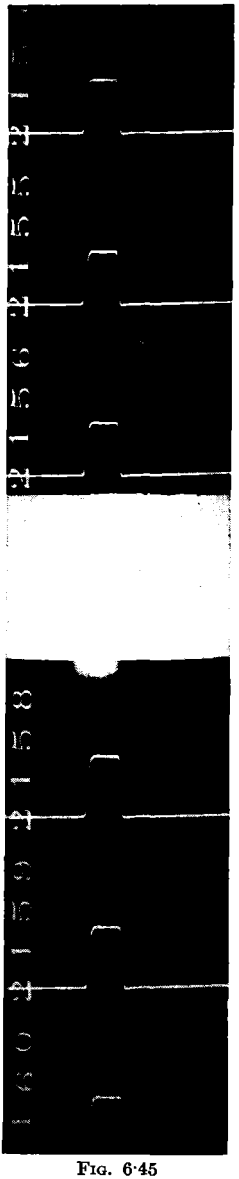


FIG. 6·45

FIG. 6·44.—Section of a pulse-to-pulse film of sea echo on 9.2 cm. Pulse length 1 μsec, pulse-recurrence frequency 333 pps, sweep length 1500 yd.

FIG. 6·45.—Section of a calibration film showing a signal generator pulse for two calibration levels 2 db apart.

1 μsec wide, obtained from a signal generator. The r-f level of the pulse is deliberately made so high that it saturates the first video stages of the receiver.[1] Such a saturated pulse performs several functions. It may serve as a range marker. More important, the top of the pulse provides a convenient reference level down from which the echo amplitude is measured. The baseline is not always suitable as a datum, especially as it tends to be obscured in any extended clutter signal.

The negative film was projected in a microfilm reader and the deflection at some given range read off to the nearest 0.05 cm, with the maximum deflection about 8 cm.[2] Between 2000 and 4000 sweeps were recorded on a single 100-ft roll of 16-mm film, depending upon the film speed. Of these, about 1000 consecutive sweeps were usually measured, although on occasion up to 3000 sweeps have been measured.[3]

Up to this point the data are in the form of deflections on the A-scope and must first be converted into some unit proportional to r-f intensity before being analyzed statistically. The necessary curve of deflection vs. r-f level is known as the "receiver calibration curve." It is obtained by photographing separately the pulse from an accurately calibrated r-f signal generator. The r-f level of the pulse is varied in steps of about 2 db until the entire dynamic range of the receiver is covered (usually about 20 to 30 db). Figure 6·45 is a strip from such a calibration film taken at 16 frames per second and shows the pulses for two levels 2 db apart.

Some typical calibration curves for various receiver gains are plotted in Fig. 6·46 in a *normalized* form, that is, deflection, as per cent of saturation, plotted against r-f level in decibels. (Necessary corrections have been made for the width of the trace.) The horizontal scale for each value of receiver gain has been shifted so as to make the points coincide at 60 per cent of saturation. It will be noticed that for a range of gain of about 30 db (as determined from the horizontal shifts) the calibration curve plotted in this form is substantially independent of receiver gain. Outside this range, significant changes in the shape of the curve begin to appear. Therefore when the receiver is being calibrated, an effort is made to set the receiver gain to approximately the same level as when the echo was originally photographed, but obviously no great accuracy is required.

The calibration curve cannot be represented by some simple law, as is often assumed. Below about half saturation, the deflection is usually a power-law function of r-f voltage, with the exponent varying between 1.0 and 2.5 depending on the receiver. Above 60 per cent saturation, the

[1] The over-all bandwidth of the video stages was from 4 to 10 Mc/sec, depending on the system.

[2] The graininess of the negative determined the optimum magnification.

[3] Much of the success of these measurements was due to the efforts of Margaret Harwood, whose careful attention to detail and unflagging interest resulted in data of the necessary high degree of precision.

curve flattens off and approaches the saturation level asymptotically. A detailed discussion of the sources of the nonlinearities lies outside the scope of this section,[1] but such nonlinearities appear to be present in all receivers commonly employed in microwave radar systems.

From the calibration curve a conversion table is constructed, listing the r-f intensity corresponding to A-scope deflection in steps of 0.05 cm. With this table the measured deflections can be converted rapidly to r-f intensity. Depending upon the eventual use of the data, similar conversion tables

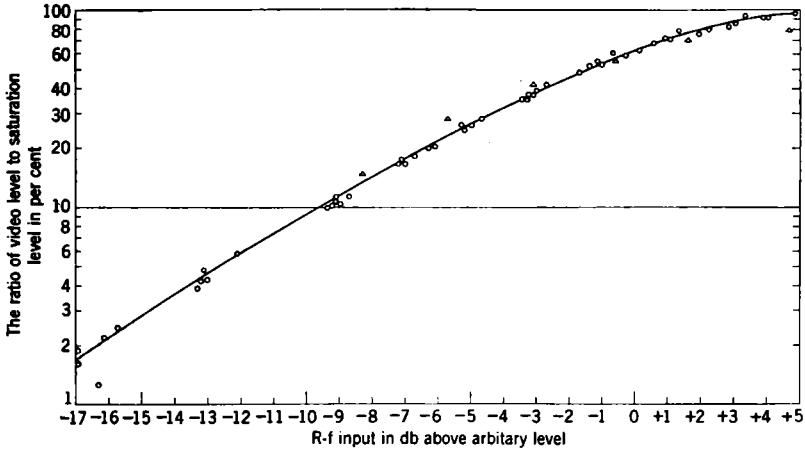


FIG. 6·46.—Typical curve of the receiver law. The circles represent data for six gains covering a range of 30 db, each shifted horizontally to coincide at 60 per cent saturation. The triangles show the data for a gain of 5 db below this range.

may be constructed for r-f voltage, r-f level in decibels, or any other desired function of r-f intensity.

The method of calibration described here has proved satisfactory if noise is not appreciable. No entirely adequate procedure for calibration has been found, however, when the gain is so high that the noise level is a sensible fraction of the maximum deflection. The noise modulates the calibration pulse with the result that its height varies randomly from sweep to sweep. In practice the average height of this modulated pulse is measured by averaging the lowest and highest amplitudes. It is then assumed that to a first approximation the deflection is linear in r-f power, so that the average height therefore corresponds to the signal-generator r-f power plus noise power. The calibration curve is than plotted, adding a constant noise power to the signal-generator attenuator reading In principle the form of the calibration so found could be used to calculate

[1] The chief source appears to be the second detector and succeeding video stages; but, as evidenced by the dependence on gain, some nonlinearity occurs in the early i-f stages.

a correction (due to noise) for various signal generator powers and the procedure repeated in a method of successive approximations, but this process is too tedious to be of much use. Hence an effort was made to work with high signal-to-noise ratios in order to avoid these difficulties.

From the converted data the first probability distribution is obtained in the obvious fashion by dividing the intensity scale into a number of adjoining intervals and counting the number of measurements falling in each interval.[1] The rates of fluctuation are first obtained in the form of correlation functions. This quantity has been defined for an infinite sample as

$$\rho(P,\tau) = \frac{\overline{[P(t) - P_0]\,[P(t+\tau) - P_0]}}{\overline{[P(t) - P_0]^2}} = \frac{\overline{P(t)\,P(t+\tau)} - P_0^2}{\overline{P(t)^2} - P_0^2}. \quad (179)$$

For a finite sample it is possible to give several definitions for ρ, all of which approach the limit [Eq. (178)] as the size of the sample becomes infinite. Solely on the basis of convenience in computing, the following definition has been used. Let T be the interval between successive measurements, and $n = \tau/T$. Then the correlation function is obtained from

$$\rho(P,\tau) = \rho(P,n) = \frac{\frac{1}{N}\sum_{j=1}^{N} P_j P_{j+n} - \left(\frac{1}{N}\sum_{j=1}^{N} P_j\right)^2}{\frac{1}{N}\sum_{j=1}^{N} P_j^2 - \left(\frac{1}{N}\sum_{j=1}^{N} P_j\right)^2}. \quad (180)$$

The first sum in the numerator requires n measurements of P in addition to the sample of N values. The number of such "excess" measurements never exceeded 4 or 5 per cent of N and was usually much smaller. Under these conditions all definitions of ρ for finite samples yield substantially the same value. The sums required in Eq. (180) are most efficiently and rapidly obtained with punched-card computing machines. The correlation function was computed in this manner in integral steps of n usually up to $n = 20$ or 40 or, in rare cases, up to $n = 80$.

Figure 6·47 shows a typical correlation function plotted against τ or, equivalently, against n. The correlation function should be, of course, identically equal to 1.0 for $\tau = 0$. If the curve in Fig. 6·47 is extrapolated back to $\tau = 0$, it approaches a value less than 1.0. As has been shown in Sec. 6·19 this discontinuity at the origin is caused by the presence of receiver noise, random from one sweep to the next, in addition to the echo under study. The magnitude of the discontinuity is determined by the signal-to-noise ratio and may be calculated from Eq. (167).

[1] The counting may be done manually or by the use of punched-card machines.

When τ is so large that successive values of P are uncorrelated, then $\rho(P,\tau)$ approaches zero. In Fig. 6·47, however, it is seen that instead $\rho(P,\tau)$ appears to approach asymptotically a curve above the axis, which in turn is only very slowly tending toward zero. This asymptotic behavior is indicative of a slow secular variation of the average power P_0. Such a slow fluctuation might perhaps be explained by the change in orientation of chaff dipoles and the accompanying change in cross section, or, in the case of sea echo, it might be caused by the rise and fall of a sea wave, producing variation in the number and size of the scatterers. Whatever the source, this variation is always very slow compared with the fast "Doppler beat"

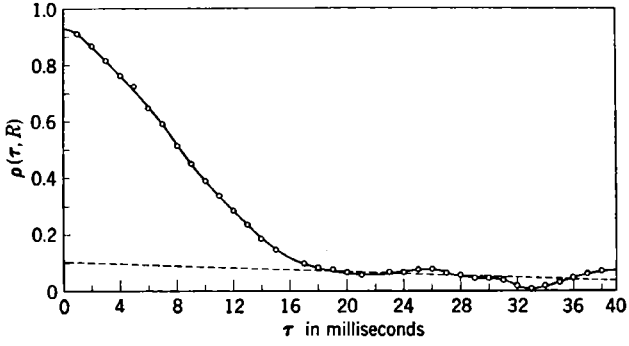


FIG. 6·47.—Voltage correlation function for chaff on 3 cm.

fluctuation. A rigorous treatment of this case would appear to be difficult, as we are no longer dealing with a stationary process. A first approximation, however, may be obtained in the following way. The sample of N values is divided into l groups, each containing m measurements separated τ in time. We assume that $m\tau$ is long compared with the periods of "Doppler beat" fluctuation but small compared with the period of the variation of P_0 so that within each group P_0 can be assumed constant. The correlation function [Eq. (180)] can now be written (for random scatterers) as

$$\rho(P,\tau) = \frac{(1+g^2)\frac{1}{l}\sum_j P_{0j}^2 - \left(\frac{1}{l}\sum_j P_{0j}\right)^2}{\frac{2}{l}\sum_j P_{0j}^2 - \left(\frac{1}{l}\sum_j P_{0j}\right)^2}. \tag{181}$$

The sums can be written as averages

$$\rho(P,\tau) = \frac{(1+g^2)\,\widetilde{P_0^2} - \tilde{P}_0^2}{2\widetilde{P_0^2} - \tilde{P}_0^2}, \tag{182}$$

where the tilde denotes averaging over the distribution of the P_0's. As $g^2 \to 0$, $\rho(P,\tau)$ then approaches a value different from zero, which can be written as

$$\lim_{\tau \to \infty} \rho(P,\tau) = \frac{\delta^2}{2\delta^2 + 1}, \tag{183}$$

where δ is the standard deviation of the distribution for P_0. It should be noted that the asympototic value of ρ is always positive. The first probability distribution will also be affected by the secular variation of P_0. The standard deviation for P is no longer 1 but $\sqrt{1 + 2\delta^2}$, indicating an excess of large values of P in the distribution. Values of δ^2 obtained separately from the measured standard deviation of the first probability distribution and from the measured asymptote of ρ agree satisfactorily.

Equation (182) suggests that the intrinsic correlation function arising from "Doppler beat" fluctuation may be obtained by subtracting from the measured ρ the asymptotic curve, indicated by the dashed line in Fig. 6·47. All the measured correlation functions have been corrected in this way and then multiplied by a factor that brings the extrapolated value of $\rho(0)$ up to unity, thus also correcting for the noise discontinuity.

It will be noticed that ρ does not approach the asymptotic curve in Fig. 6·47 smoothly but oscillates about it. In the corrected correlation function this behavior appears as oscillation about the horizontal axis. It is believed that these oscillations are due entirely to statistical fluctuations arising from the finite size of the sample, since for infinite samples g^2 is always positive. In practice, therefore, these oscillations have been smoothed out, and the curve is drawn to approach the axis monotonically.

The power frequency spectrum is, as stated by the Wiener-Khintchine theorem, the cosine transform of the correlation function. Often the corrected correlation function can be accurately approximated by a function whose cosine transform is well known, for example, the Gaussian error curve. If not, then the integration is performed numerically, using Simpson's rule. By test with a known function, the numerical integration has been found to be sufficiently accurate up to about half the pulse-recurrence rate.

The preceding discussion has been based on the implicit assumption that the data are given in the form of intensity or power, as is usually preferable. As a consequence, however, of the large range of values encountered in a given sample (often greater than 1000 to 1), the punched-card computations often prove to be unduly lengthy. If the deflections are converted into r-f voltages or amplitudes instead of intensities, the range is greatly reduced; therefore this conversion was often made. Equation (157) supplies the theoretical relation between the correlation functions for amplitude and intensity, assuming random scatterers.

The correlation function ρ $(\mathbf{R},\tau)$ is never very different from g^2, the maximum deviation being 0.027. Usually this small difference may be neglected, and the two correlation functions considered equivalent. Whenever it is desired to use Eq. (157) to obtain g^2 from $\rho(R,\tau)$, however, correction must first be made for any slow fluctuation present, and the statistical oscillations smoothed out, as described above. The correlation function shown in Fig. 6·47 is an example of an amplitude correlation function. In Fig. 6·48 the same curve is shown after the various corrections have been made and after it has been converted to an intensity correlation function.

Significant deviations between experimental first probability distributions and theoretically predicted functions yield information about the size and number of scatterers contributing to the given clutter signal. To determine if observed deviations are significant, it is important to know the magnitude of the experimental errors in the first probability distribution. Three types of errors may be distinguished:

1. Statistical fluctuations because of finite size of the sample.
2. Errors in reading the film.
3. Errors in the calibration.

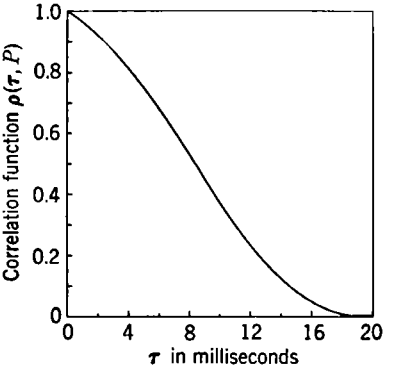


FIG. 6·48.—The correlation function for chaff on 3.2 cm, converted to an intensity basis and corrected as outlined in the text.

The fluctuations in the number of measurements falling within a given intensity range are distributed according to the well-known Poisson distribution. The rms fractional deviation is $1/\sqrt{n}$, where n is the expectation value of the number of measurements, provided n is small compared with the size of the sample N. It must be remembered that independent trials are assumed. Normally there is correlation between successive measurements, and n should be replaced by an n', which is the number of measurements out of n that may be considered independent (that is, that have little correlation between them).

Errors in the reading of the film may be either systematic or random. If the trace had a large slope and was faint, the tendency was to read a value consistently too high. It is difficult to estimate the magnitude of such an error; fortunately, it occurred only infrequently. A source of random error was almost always present, however; depending upon the computer and the nature of the trace, there is a certain minimum measurable difference in deflection, ranging from 0.05 cm to several millimeters, but usually about 1 mm. As stated above, the first probability distribution is obtained by dividing the range of intensities into several groups and count-

ing the number of measurements falling in each group. Let us consider now an interval equal to the minimum measurable difference, centered about the deflection corresponding to the boundary between two groups. Of the pulses in this interval zone, some will by chance be assigned to one group and some to the other. The rms fluctuations of the number in a group, caused by such an "uncertainty zone" at one boundary, is given by $\frac{1}{2}\sqrt{m}$, where m is the total number of measurements falling in the

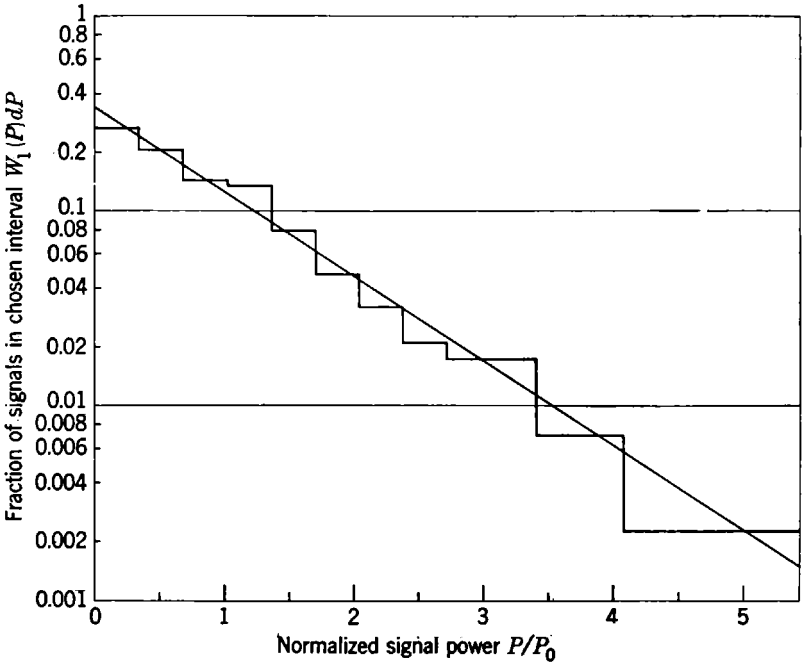


FIG. 6·49.—Typical first probability distribution for chaff echo on 9.2 cm. The histogram shows the experimental distribution of 1000 pulses; the straight line is the theoretical formula, using the measured average power.

uncertainty zone.[1] It should be noted that m includes all measurements falling in the interval, as all trials are independent. This type of error will be referred to as "boundary error."

Errors caused by improper calibration are, of course, difficult to correct. They are best minimized by proper choice of receiver gain when the film is taken. An attempt was made always to set the gain in such a way that most of the pulses were well below the saturation level, as the errors would be greatest in the neighborhood of this level. Any effect

[1] The reasoning employed is as follows: It is equally likely that a measurement will be assigned to either group; the average number going to any one group is $m/2$; the rms fluctuation in this number is $\sqrt{(m/2)\ (1 - \frac{1}{2})} = \frac{1}{2}\sqrt{m}$.

attributable to incorrect calibration should appear as a systematic deviation from the theoretical distribution, whereas most of those discussed above would be random.

6·21. Experimental Results.—The previous section has described the experimental techniques involved in recording the fluctuations of clutter echoes and in analyzing these records in accordance with the theory developed in the previous sections. This present section summarizes the results obtained in this manner at the Radiation Laboratory. Those targets which may considered as assemblies of random scatterers, that is, chaff, precipitation, and sea echo, are discussed first. Ground clutter, in which the target contains fixed as well as moving scatterers, is considered later.

Chaff. It has been pointed out in Sec. 6·18 that the large numbers of randomly moving dipoles which make up the chaff cloud should ideally satisfy the conditions for random scatterers prescribed in Sec. 6·16. In fact the experimental analysis of chaff echo is in full agreement with the theory. Thus Fig. 6·49 compares a typical experimental first probability distribution, on 9.2 cm, with the theoretical exponential curve. The histogram represents the analysis of 1000 pulses, and the straight line, corresponding to the theoretical curve, uses the value of the measured average intensity of the sample analyzed. The deviations between the two plots are random and do not indicate any significant differences. Table 6·7 compares the magnitude of the deviations in each interval with the expected rms fluctuations arising from both the finite size of the sample and the "boundary error" discussed in Sec. 6·20. In the

TABLE 6·7.—DEVIATIONS BETWEEN THEORETICAL AND EXPERIMENTAL FIRST PROBABILITY DISTRIBUTIONS FOR CHAFF ECHO; λ = 9.2 CM, 1000 PULSES ANALYZED

Interval of $\frac{P}{P_0}$	No. of pulses in interval		Difference	Expected rms fluctuation in number	
	Experimental	Theoretical		Due to finite sample	Due to boundary error
0 –0.34	275	292	−17	17	3
0.34–0.68	205	202	+ 3	14	4
0.68–1.01	143	144	− 1	12	4
1.01–1.35	136	102	+34	10	3
1.35–1.69	79	72	+ 7	8	3
1.69–2.03	47	52	− 5	7	3
2.03–2.36	32	37	− 5	6	2
2.36–2.70	21	26	− 5	5	2
2.70–3.38	35	31	+ 4	5	2
3.38–4.05	14	16	− 2	4	1
4.05–5.40	9	12	+ 3	3	1
5.40–8.43	4	4	0	2	1

fifth column it has been assumed that all the measured pulses are independent of each other. As there is actually considerable correlation between neighboring pulses, the figures in this column represent minimum values. The "boundary error" calculation assumes that differences in pulse height of less than 1 mm were not measurable, which is likewise a minimum figure. It is seen that the actual deviations are, with one exception, well inside the expected statistical fluctuations. Equally good results have been ontained in all measurements on 9.2 and 3.2 cm of chaff cut to resonate at 10 cm. The first probability distribution of the

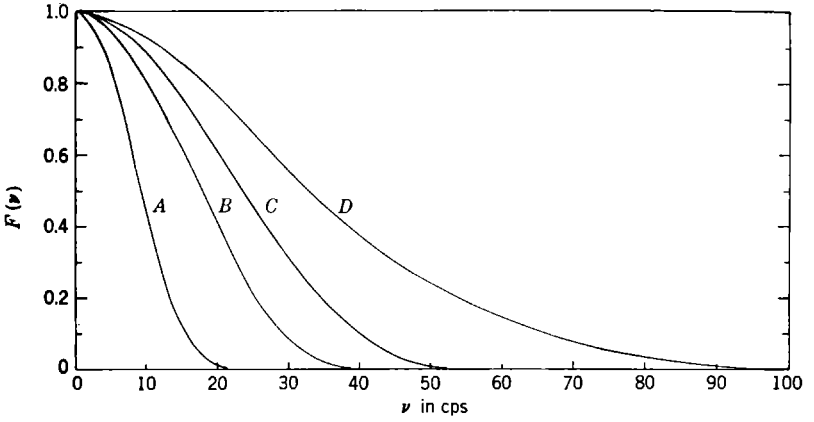


FIG. 6·50.—Power frequency spectrum for fluctuations of chaff echo, as measured on four occasions.

echo on a frequency of 515 Mc/sec from chaff cut for the "Wurzburg" band ($\lambda \approx 50$ cm) has also been measured, and good agreement was likewise obtained with theory.[1]

All the measured frequency spectra of the fluctuations of chaff echo have had roughly the same shape, resembling error curves centered at the origin. The width of the spectra, however, is quite variable even at one wavelength. Thus Fig. 6·50 shows the spectra for the echo of chaff cut for 10 cm, as measured on 9.2 cm on four occasions. It is significant that the widest spectrum *D* was obtained with gusty winds up to 25 mph whereas the wind speed was 10 mph or less for the other cases. The width of the spectrum depends upon the relative velocity of the chaff dipoles, that is, the so-called "horizontal-dispersal rate," and it is to be expected that this rate will depend on the speed and gustiness of the wind.

The narrowest spectrum, Curve *A*, was obtained with chaff of the same electrical properties as used for the other curves but having slightly

[1] G. P. Kuiper, "A Study of Chaff Echoes at 515 Mc," RRL Report No. 411-73, December 1943.

different mechanical and aerodynamical properties. In addition, the chaff was dispensed from a slowly moving blimp instead of being thrown out into the turbulent slip stream of an airplane.

The measured spectra (including a few not shown in Fig. 6·50) do not show any significant dependence on the "age" of the chaff, that is, the length of time between the dropping and the measurement. The data are, however, too limited to allow a definite conclusion. As a matter of interest, the ages for the four curves *A*, *B*, *C*, and *D* of Fig. 6·50 are 3 min, 20 sec, 6 min, and 10 min, respectively.

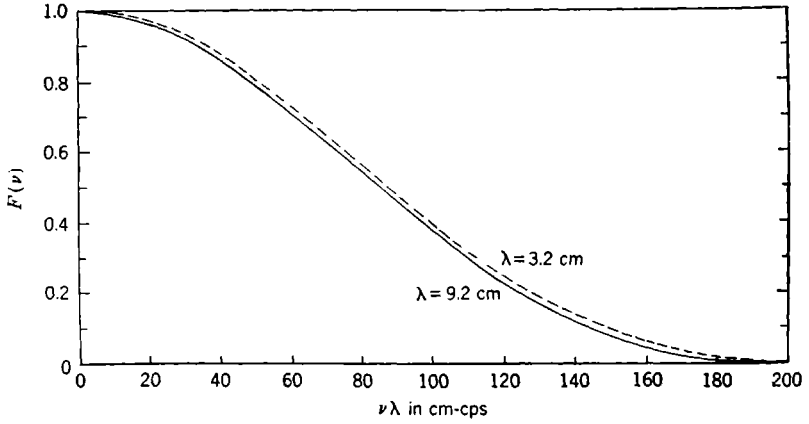


FIG. 6·51.—Power frequency spectrum for the fluctuations of chaff echo on 9.2 and 3.2 cm, plotted as a function of the product $\nu\lambda$.

If the fluctuation arises solely from the Doppler beats of the moving chaff dipoles, then it is seen from Sec. 6·18 that the correlation function is a function of the product $v_z\tau/\lambda$. Hence the width of the frequency spectrum of the fluctuation should be proportional to the radar frequency for the same velocity distribution. More exactly, if simultaneous measurements of the spectra are made on several wavelengths, the curves should coincide when plotted as functions of the product of the fluctuation frequency and radar wavelength. Accordingly, in Fig. 6·51 the experimental spectra for chaff measured simultaneously on 3.2 and 9.2 cm are plotted against $\nu\lambda$. The small discrepancy between the two curves is well within experimental error.

The spectrum of chaff has also been measured[1] at 515 Mc/sec, and the maximum frequency present to any appreciable extent was found to be 4 cps on horizontal polarization. When "scaled" to 9 cm, this value roughly corresponds to Curve *A* of Fig. 6 50. Considering that the chaff size was different, the wavelength dependence is at least qualitatively verified. The frequency of fluctuation on vertical polarization was found

[1] G. P. Kuiper, *op. cit.*

to be somewhat slower, with a maximum of 3 cps. (All the microwave measurements were made with horizontal polarization.)

Mention must also be made of some results on the frequency of fluctuation, obtained in England, that are not in agreement with the above conclusions.[1] Pulse-to-pulse photographs of chaff echoes were made on 212 and 3000 Mc/sec. A statistical analysis was not carried out, but it is stated that there was qualitative evidence of fluctuation frequencies in the range 10 to 25 cps on both frequencies.

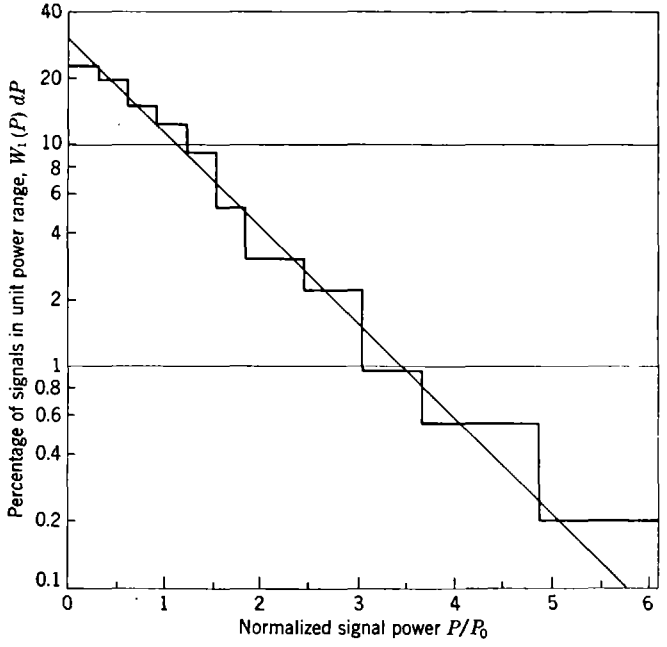


FIG. 6·52.—First probability distribution of the echo from a rainstorm on 3.2 cm. The histogram shows the analysis of 1000 pulses; the straight line represents the theoretical behavior for random scatterers.

From the measured spectra one might expect to obtain information about the distribution of the relative velocities. A detailed analysis of this sort has not been carried out; indeed, there is no unique velocity distribution corresponding to a given spectrum. If, however, the spectrum has a Gaussian shape, some information about average speeds can be easily obtained. Let $\bar{v}$ be defined such that one-half of the scatterers have relative velocities in the direction of the radar set lying between $-\bar{v}$ and $+\bar{v}$. Then it has been shown by Siegert[2] that with a Gaussian

[1] "Final Report on ADRDE Window Trials," ADRDE Report No. 250, April 1944.
[2] A. J. F. Siegert, RL Report No. 773, January 1946.

spectrum, $\bar{v}$ is given by

$$\bar{v} = 0.00650\lambda\nu_{1/2} \text{ ft/sec}, \tag{184}$$

where λ is in centimeters and $\nu_{1/2}$ is the frequency at which the power spectrum is down to one-half of its original value. Applying this formula to the curves of Fig. 6·50 (though they are not exactly Gaussian in shape) one finds values of $\bar{v}$ ranging between 0.6 and 2.0 ft/sec. These results are of the same order of magnitude as the horizontal-dispersal rate as actually measured from motion pictures of chaff dipoles.

The correlation functions obtained for chaff show that in addition to the fast Doppler-beat flunctuations, a much slower variation is almost always present to some extent. (This slow variation has been subtracted out in obtaining the curves of Fig. 6·50. The origin of this type of fluctuation, which extends to a few cycles per second, is not yet clear. Possibly it may be due to the rotation of the dipoles (about an axis perpendicular to their length, which would cause variations in the amplitude scattered by each dipole.)

Precipitation Echo. The scatterers responsible for precipitation echoes are almost undoubtedly either raindrops or water particles in solid form. One would therefore expect that the conditions for treating the target as an assembly of random scatterers are as well satisfied as with chaff. This expectation is borne out by experiment. Figure 6·52 shows a typical experimental first probability distribution as obtained from the analysis of 1000 pulses on 3.2 cm of the echo from a rain shower. The deviations of the histogram from the theoretical curve (plotted as a straight line on this scale) are random. An analysis similar to that made for chaff echo shows that the deviations are within the expected statistical fluctuations.

Unfortunately, only a small amount of data is available on the spectrum of the fluctuations of precipitation echo. Three films were measured on 9.2 cm, and one on 3.2 cm; the echo in each case came from a shower or thunderstorm. The fluctuations are always very rapid, being several times greater than those for chaff. Figure 6·53 is a plot of the power spectrum for the three films measured on 9.2 cm. The spectra have roughly Gaussian shapes except in one case where there is a pronounced tail above 80 cps. It is doubtful if this tail is significant, as it occurs at frequencies so close to the pulse-recurrence rate that the whole method may be expected to break down. The other curves were obtained with much higher recurrence rates.

The width of the three spectra, given by $\nu_{1/2}$, varies by a factor of 2. As in the case of chaff, it is therefore not possible to speak of *the* fluctuation spectrum, but one must expect to find the spectrum depending on the particular storm, probably varying even within the storm region, and changing with time. Additional evidence for this variability was furnished by an attempted experiment on the wavelength dependence of the fluctua-

tion. Two films were taken of the echo from a rain shower on 9.2 and 3.2 cm, respectively, separated by a time interval of a few minutes and a range interval of a few thousand yards. The fluctuation spectra, instead of depending linearly on wavelength, were practically identical both in shape and width. As all other evidence points to the validity of the fundamental mechanism assumed for the fluctuations, it must be concluded that the rate of fluctuations can change by a factor of 3 even over short intervals of time and space.

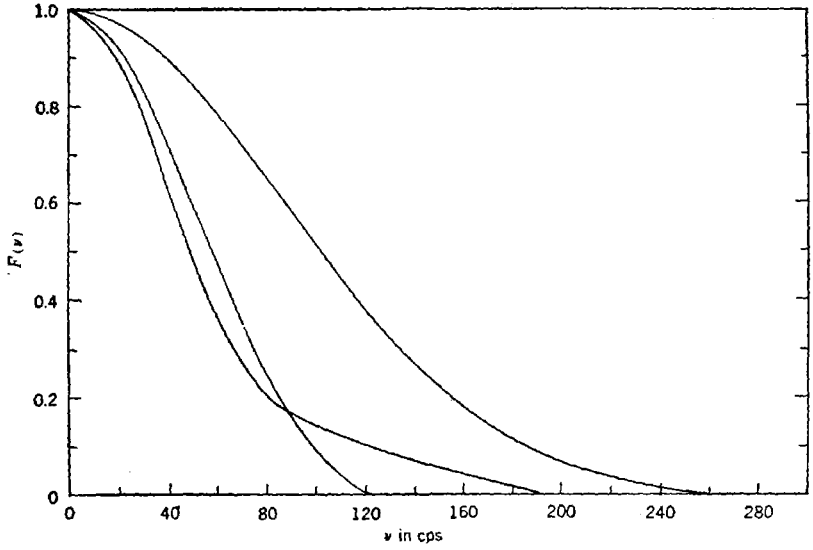


FIG. 6·53.—Power frequency spectrum of the fluctuations of precipitation echo on 9.2 cm as measured on three occasions.

A few measurements of the frequency spectra for precipitation echoes have been made by an entirely different method, using a superheterodyne-type audio analyzer.[1] Although far quicker than the photographic analysis, the results obtained by this method are in a much cruder form and comparison is difficult. The frequency at which the spectrum is one-tenth of the value at the origin, $\nu_{0.1}$, varied between 200 and 300 cps, which is somewhat higher than for the curves in Fig. 6·53. Some evidence was found for an increase of fluctuation frequency with range. This is attributed to the larger area illuminated by the beam at great distances, thus including regions of more diverse wind velocities.

The foregoing method is capable of considerable development and, in a more perfected state, would be preferable to the photographic technique

[1] RL Report No. 773. See also R. M. Ashby, F. W. Martin, and J. L. Lawson, "Modulation of Radar Signals from Airplanes," RL Report No. 914, March 28, 1946.

because of the immense saving in time. It would be interesting to use the method for a detailed study of the variation of the fluctuation spectra from storm to storm and the variation with range, pulse length, and wavelength.[1]

An average figure for the speeds of the drops relative to each other may be obtained from Eq. (184), but even greater caution must be exercised than for the case of chaff because of possible deviations from a Gaussian spectrum. Figure 6·53 yields values of v between 3 and 5 ft/sec. These relative velocities should be connected with the turbulence existing in the regions of precipitation, and in this light the values seem low. It must be remembered, however, that the turbulence in a storm is mostly vertical and the fluctuation depends on the *horizontal* velocities. Futhermore, the velocity distribution is very broad and extends far beyond v. A detailed study of the connection between the fluctuation spectra and the degree of turbulence is of considerable practical importance. If some correlation exists it might be possible from a study of the radar signal to inform aviators of dangerous precipitation areas.

It has been suggested that the fluctuation arises not so much from the interference between randomly moving drops as between blocks of drops, the members of which move as a whole. Although this is undoubtedly true to some extent, if there are many such blocks moving randomly relative to each other, the spectrum is indistinguishable from that due to completely randomly moving drops. If there is such an "order" in the distribution of the velocities, it might be possible, however, to make the illuminated area so small that only a few blocks are present, and then differences would appear.

The correlation functions for precipitation echo are particularly marked by an almost complete absence of the secular variation found for chaff and sea echo. Actually there is a slow variation, from a change of the number and size of the drops as the character of the precipitation changes in space and time. The periods involved, seconds or minutes, are too long, however, to appear on the correlation functions.

Measurement of the correlation function or the spectrum does not of itself constitute a verification of the form of the second probability distribution obtained in Sec. 6·19. In fact, the previous discussion has implicitly assumed this distribution. It has already been pointed out that some degree of verification may be attained by an examination of distributions derived from the second probability distribution. Thus in Eq. (160), a distribution is given for the ratio of the echo intensities at two instants τ apart. Figure 6·54 shows an experimental distribution of this sort (actually in terms of $\Delta(\text{db}) = 10 \log_{10}[P(t+\tau)/P(t)]$) obtained from the analysis of 1000 pulses on 9.2 cm for the echo from a rain shower.

[1] Such a program is being carried out in the Weather Radar Project of the Department of Meteorology, Massachusetts Institute of Technology.

The smooth curve is the theoretical formula [Eq. (161).] using the measured correlation function for the time interval in question (0.003 sec.). The differences between the theoretical curve and the experimental histogram are not completely random. However, this was one of the earliest films measured, and the technique was still so crude that the errors in measurement were large.

Sea Echo. Unlike the previous targets, the nature of the scatterers responsible for sea echo is not well known and it is not possible to predict

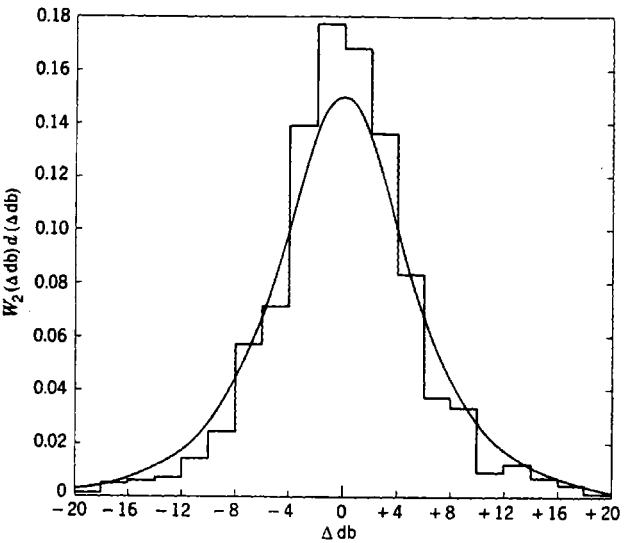


FIG. 6·54.—Probability distribution of the logarithm of the intensity ratio of two pulses, 3 milliseconds apart, for a typical precipitation echo on 9.2 cm. The histogram shows the analysis of 1000 pulses, and the continuous line is the theoretical curve based on the measured correlation function.

with surety the properties of the fluctuation; rather, the experimental data are analyzed for what information can be inferred from them about the scattering mechanism. In Sec. 6·11 this question is discussed in detail in the light of current theories of sea echo. Here it may be stated that whether it is assumed the echo arises from spray droplets or from surface irregularities, there should be a sufficiently large number of independent scatterers present to satisfy the conditions for random scatterers. The interpretation of the spectrum in terms of the velocity distribution, however, would be open to question.

A striking feature of the sea-echo signal is the presence of the slow or secular variation to a much more marked extent than with the other echoes of this class. The period of this variation, which must be associated with the lifetime of the scatterers, is of the order of seconds or fractions

of a second. Within a given record of the echo covering several seconds, the average intensity will not in general be truly constant. As has been pointed out in Sec. 6·20, this has the effect of increasing the standard deviation of the first probability distribution, and one would therefore expect to find an excess of high-intensity echoes. A rather extreme example of such behavior is shown in Fig. 6·20, Sec. 6·11, where there is plotted the experimental first probability distribution as obtained from an analysis of 2000 pulses on 9.2 cm covering a time interval of 48 sec. The straight line is the theoretical exponential curve for $W_1(P)\,\Delta P$, using the intensity averaged over the entire time interval. As explained in Sec. 6·20, the standard deviation of the distribution of "average" intensities may be obtained both from the first probability distribution and from the finite asymptote of the correlation function. The values computed from the data of Fig. 6·20 and from the correlation function are large, as would be expected, $\delta = 0.33$ and 0.48 respectively and are in rough agreement with each other.

If a time interval is chosen in which the "average" intensity is sensibly constant, then the first probability distribution has always been found to agree with the theoretical prediction. Figure 6·19 in Sec. 6·11 illustrates such a situation. Here the first probability distribution was obtained from the analysis of 2000 pulses in a much shorter time interval, 4 sec, during which there happened to be little variation of the average intensity. The deviations are well within statistical error.[1] Thus the rms fluctuation of the echo relative to the average intensity is 0.93, compared with the theoretical value 1.0 derived in Sec. 6·19. The asymptote of the correlation function is correspondingly low. Equally good agreement with theory has been obtained with the highest resolution available, that is, 0.15 μsec pulse and 0.8° beamwidth. It must be concluded that even in areas as small as 20 by 90 yd the number of sea-echo scatterers is at least greater than five.

The power spectrum of the fluctuations of sea echo is again roughly Gaussian in shape and has about the same width as the spectra for chaff. The range of widths encountered is much smaller, however, the values for $\nu_{\frac{1}{2}}$ lying between 25 cps and 35 cps for 9.2 cm.[2]

No dependence of the spectrum on pulse length was found in one comparison between 0.15 and 1 μsec on 3.2 cm. Figure 6·55 presents the two spectra as measured about 1 min apart on the two pulse lengths and illustrates the close agreement. Although this is the only example measured, it is in accord with general qualitative observations.

[1] It should be pointed out that even in the extreme example of Fig. 6·20 the number of excess pulses is less than 2 per cent of the total.

[2] Somewhat larger values have been reported by T. Gold. See ASEE Witley Report No. XRD/46/3, July 10, 1946. The differences are perhaps attributable to rougher sea conditions than those encountered in our measurements.

Several attempts have been made to compare the fluctuation spectra on two wavelengths. Two such comparisons, between 3.2 and 9.2 cm, were made, but unfortunately not simultaneously either in time or in range. The ratio of the widths in these two cases was 2.0 and 2.5, respectively, compared with 2.88 for the ratio of wavelengths. One comparison was made between 9.2 and 1.25 cm which was simultaneous in space, however, and only a few minutes apart in time. The two power frequency spectra are plotted in Fig. 6.56 as a function of $\nu\lambda$. The widths of the spectra are seen to be in about the right proportion. The slight difference in shapes is not significant, especially in light of the

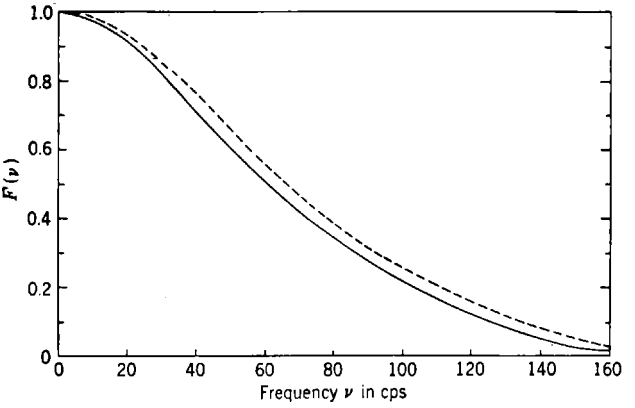


FIG. 6·55.—Power frequency spectrum of the fluctuations of sea echo on 3.2 cm for two pulse lengths. The continuous curve is for a pulse length of 1 μsec; the dashed curve is for 0.15 μsec.

large corrections made on the correlation function because of the slow "secular" variation. On the basis of the wavelength dependence, it seems fair to conclude that the fast fluctuations of sea echo are attributable, at least in major part, to the Doppler beats between moving scatterers and that the median relative velocity v is of the order of 1 or 2 ft/sec. The growth and decay of the individual scatterers also gives rise to fluctuations in the echo; but as already noted, the frequency is much slower.

The shape of the second probability distribution has been checked in the same manner as for precipitation echo, that is, by means of the distribution of the ratio of two intensities τ apart in time. Figure 6·57 shows the plot of the distribution of the ratio, expressed in decibels as usual, as obtained by analysis of 1000 pulses 0.003 sec apart on 9.2 cm. The smooth curve is the theoretical prediction calculated from the measured correlation function. It is seen that the deviations are random and the over-all agreement is very good.

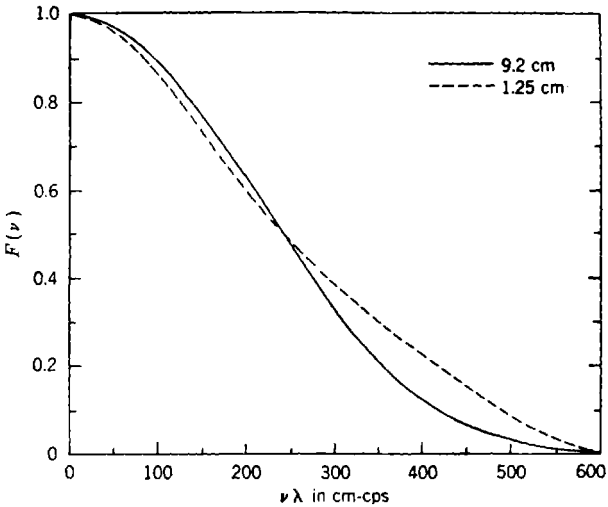


FIG. 6·56.—Power frequency spectrum for the fluctuations of sea echo on 9.2 and 1.25 cm plotted as a function of the product $\nu\lambda$.

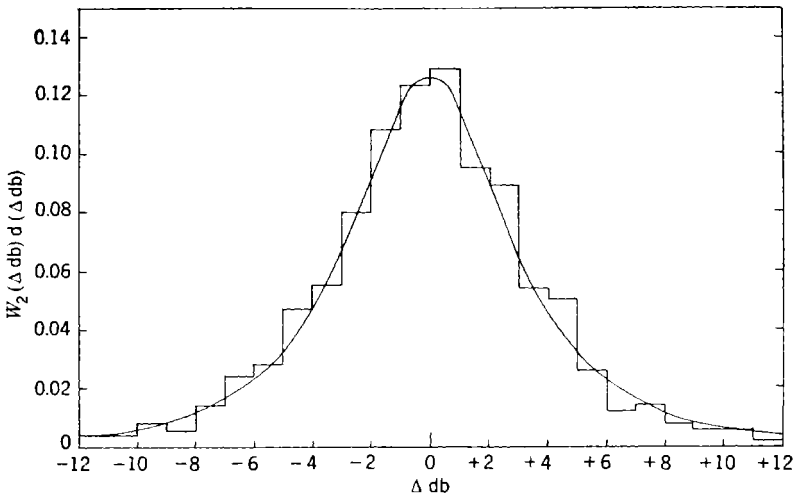


FIG. 6·57.—Plot of the probability distribution of the logarithm of the intensity ratio of two pulses, 3 milliseconds apart, for sea-echo. The histogram shows the analysis of 1000 pulses; the continuous curve is the theoretical prediction using the measured correlation function.

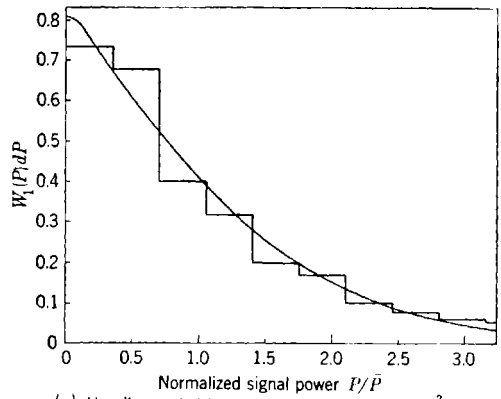
Ground Clutter. The previous targets could all be closely approximated by assemblies of random independently moving scatterers. In Sec. 6·18 it was pointed out that the targets responsible for ground clutter also include assemblies, such as leaves, branches, etc., that move in the wind. In addition there are scatterers with random relative phases that are, however, fixed, for example, tree trunks, rocks, etc. The total echo is the sum of the echoes from both classes of targets.

The first probability distribution for the echo has been derived in Eq. (173) as a function of m^2, the ratio of the intensity of the steady echo to the average power of the echo from the moving scatterers, that is, S^2/P_0. This ratio will be called the steady-to-random ratio for convenience in discussion. For $m^2 = 0$, the distribution reduces to the familiar exponential for random scatterers alone. When m^2 is small, that is, less than 1, there is little difference between Eq. (173) and the exponential. As m^2 increases, a maximum appears in the function. When $m^2 >> 1$, the distribution can be approximated by a narrow Gaussian curve centered about the intensity of the steady echo.

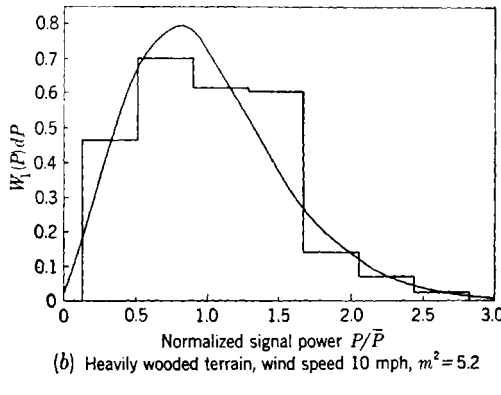
All the steps in this sequence in the development of $W_1(P)\,dP$ can be illustrated with observed ground-clutter distributions. At very high wind speeds almost everything moves in the wind and there are very few "steady" targets. As the wind velocity decreases, for the same average intensity of ground clutter, the steady-to-random ratio would be expected to increase.

The echo from a heavily wooded hill on 9.2 cm has been measured for gusty winds of about 50 mph, and the first probability distribution is indistinguishable from the simple exponential curve. A less extreme case is illustrated in Fig. 6·58*a*, which shows the first probability distribution of the echo from a densely wooded section at a wind speed of 25 mph. The flattening of the distribution for low values of signal intensity is evident, and the histogram has been fitted with a theoretical curve for $m^2 = 0.8$. In Figure 6·58*b* is shown the distribution of the echo from similar terrain but at a wind speed of 10 mph. There is now a maximum in the distribution at slightly less than the average value of the echo intensity, and the histogram is best fitted by a theoretical curve with $m^2 = 5.2$. The final stage in the sequence is illustrated in Fig. 6·58*c*. The distribution shown there is likewise for a wind speed of 10 mph, but the target consisted chiefly of rocks with sparse vegetation. All the echoes are closely clustered around the average value, and the theoretical curve best fitting the experimental distribution corresponds to $m^2 = 30$.

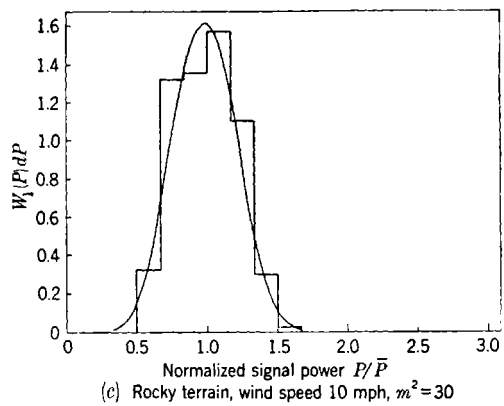
As demonstrated by the last two illustrations, the relative contribution from the steady targets is greater when the vegetation is sparse than when it is dense. A detailed study of the variation of m^2 with type of vegetation has not yet been made, nor is much known about how the



(*a*) Heavily wooded terrain, wind speed 25 mph, $m^2 = 0.8$



(*b*) Heavily wooded terrain, wind speed 10 mph, $m^2 = 5.2$



(*c*) Rocky terrain, wind speed 10 mph, $m^2 = 30$

FIG. 6·58.—Several curves of the first probability distribution for ground clutter on 9.2 cm. Experimental data are shown by histograms; theoretical curves, fitted to the data, by continuous curves.

steady-to-random ratio varies with frequency. One would expect considerable change, as a target can be classified as "steady" only in relation to the wavelength. A tree branch moving back and forth 1 cm would be steady at 10 cm but would go through all possible phase angles at 1.25 cm. The value of m^2 should therefore decrease with wavelength, other factors being constant, and there is some qualitative evidence in support of this prediction. The question obviously depends on the relative contribution of leaves and grass to the echo, compared with that from branches and tree trunks. At 10 cm, at least, it would seem that

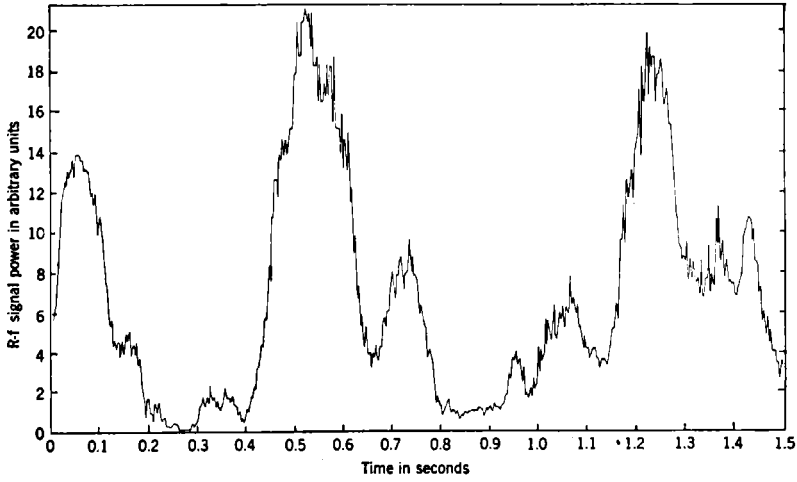


FIG. 6·59.—Plot of the echo intensity on 9.2 cm from heavily wooded terrain, at a wind speed of 22 mph.

the contribution of the latter can be considerable, as no striking variation of the average intensity with season has been observed. The situation may be different at the shorter wavelengths; there the foliage may play a more important role. Thus ground clutter on 1.25 cm always shows large and rapid fluctuations even when perfectly steady on 3.2 cm. Unfortunately, quantitative comparisons on several frequencies have so far been made at such high wind speeds that m^2 is very small at all microwave frequencies.

The fluctuation frequencies of ground clutter are considerably smaller, by an order of magnitude, than the fluctuations of chaff or sea echo, with a correspondingly smaller median speed v. The relatively slow rate is illustrated in Fig. 6·59 which is a plot of the echo intensity on 9.2 cm of ground clutter at a wind speed of 22 mph. The time scale covers 1.5 sec, during which only a few oscillations of the signal occur, in comparison with the rapid fluctuations shown in the similar record of sea echo (Fig. 6·18).

The shape of the spectrum is again roughly similar to those of the other clutter echoes, though the differences from the Gaussian shape are more pronounced. A considerable number of spectra have been measured, mostly at 9.2 cm, for various terrains and at several wind speeds. Pulse-to-pulse records of the signal intensity have been used most of the time, but occasionally the fluctuation is so slow that this procedure is no longer efficient, and A-scope motion pictures at 16 frames per second have been employed instead.

The widths of the spectra naturally increase with wind speed and, in addition, depend on the terrain to some extent. It would be pointless to reproduce all the spectra in detail. Instead, three points of a number of representative spectra for wooded terrain at 9.2 cm are listed in Table 6·8

TABLE 6·8.—EXAMPLES OF THE FLUCTUATION SPECTRA FOR WOODED GROUND CLUTTER; $\lambda = 9.2$ CM

Wind, mph	m^2	$\nu_{0.8}$, cps	$\nu_{0.5}$, cps	$\nu_{0.1}$, cps
17	1.3	1.44	0.77	1.67
10	5.2	0.65	1.10	2.07
23	1.0	0.72	1.27	2.7
22	1.0	1.9	3.1	5.9
23	0.8	1.9	3.3	8.6
30	0.2	3.3	5.9	10.1
50	0	2.1	3.9	14.8

in order to furnish an estimate of the width and shape. The three points chosen are the frequencies at which the power frequency spectrum is down to 80 per cent (−1 db), 50 per cent (−3 db), and 10 per cent (−10 db), respectively, of its maximum value. In addition, the wind speed and the steady-to-random ratio m^2 are listed when available.

Ground clutter most closely resembles the echo from randomly moving scatterers at high wind speeds, and it would therefore be expected that the widths of the spectra at these wind speeds would be proportional to the radar frequency. A comparison among several frequencies was made twice for heavily wooded terrain at gale winds or higher. One comparison, between 3.2 and 9.2 cm, gave a ratio of less than 2 for the widths of the spectra, compared with the wavelength ratio of 2.88. The second one, between wavelengths of 3.2 and 1.25 cm, gave a corresponding ratio of 2.3, compared with the wavelength ratio of 2.56. These last two spectra are plotted in Figure 6·60 as a function of $\nu\lambda$. It is difficult to know how to interpret the results, especially as there are the usual uncertainties about the simultaneity of the measurements in time. It would seem, however, that the width of the spectrum is not quite proportional to the radio frequency.

The spectrum at 9.2 cm was measured at the same time as the spectra at 3.2 and 1.25 cm mentioned above. However, the shape of the spectrum

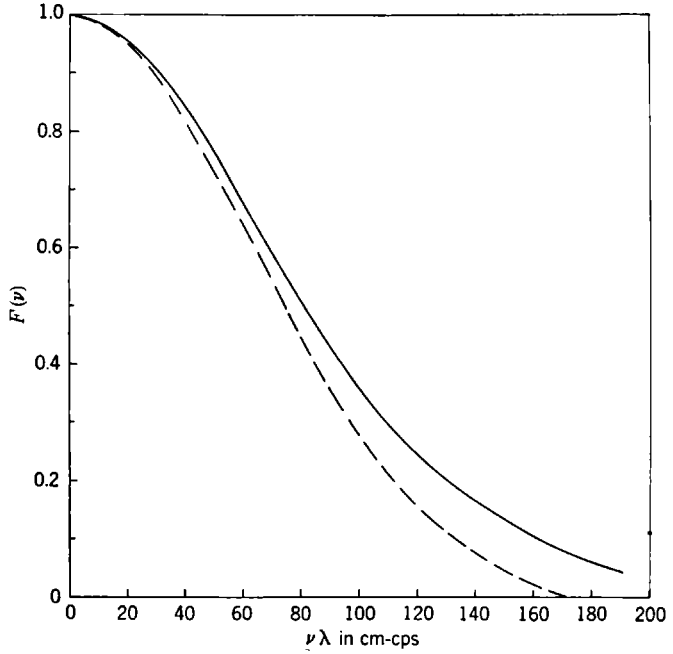


FIG. 6·60.—Power frequency spectra of the fluctuations for typical ground clutter in gale winds, plotted as a function of the product of fluctuation frequency and radar wavelength. The solid line is the curve for 3.2 cm, and the dashed line for 1.25 cm.

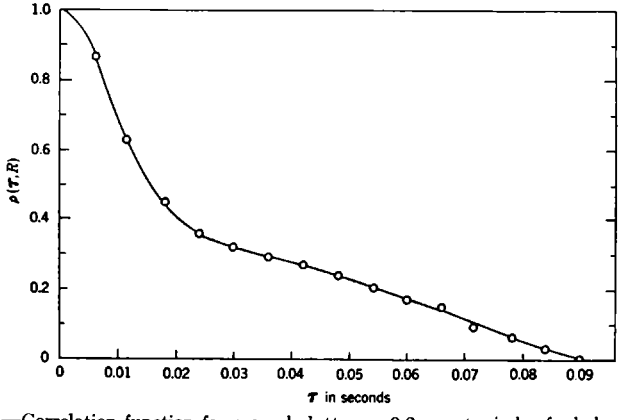


FIG. 6·61.—Correlation function for ground clutter on 9.2 cm at winds of whole gale force, showing anomalous appearance.

(listed in the last line of Table 6·8) was so markedly out of line with the others that no comparison was possible. The strange shape was the result of a kink in the correlation function, as shown[1] in Fig. 6·61, perhaps indicative of two sources of fluctuation. No explanation is offered for this strange behavior, which was not found elsewhere. This instance is cited to emphasize that much is still unknown about the fluctuations of ground clutter. It is believed the general outlines of the picture have been established, but many of the details, especially with regard to frequency dependence, await further investigation.

[1] The correlation function in the figure is already corrected for a small positive asymptote because of a very slow "secular" variation.

CHAPTER 7

METEOROLOGICAL ECHOES

By Herbert Goldstein, Donald E. Kerr, and Arthur E. Bent

ORIGIN OF THE ECHO

By Herbert Goldstein

Under certain meteorological conditions characteristic echoes will appear on microwave radar systems, often covering considerable area and characterized by irregular, diffuse boundaries and rapidly fluctuating intensity. A large body of evidence, presented in the following sections, supports the belief that all such echoes are caused by incoherent scattering from condensed water, either in the form of raindrops or in some solid form such as ice crystals or snow. Section 7·1 treats the phenomenon of incoherent scattering from any type of scatterer randomly distributed in volume. This section will provide a common language in terms of which one can express the results of the several theories and make comparison with experiment. In Sec. 7·2 it is shown that wherever careful investigations have been made, precipitation has with few exceptions been found at the points from which the echoes appear to arise. Section 7·3 presents some elementary and crude calculations of the magnitude of rain echoes to be expected on the basis of the raindrop theory and shows that the experimental values are not inconsistent with these results. Some of the alternative theories not involving condensed water are discussed in Sec. 7·4, and it is shown that these theories cannot account for the observed magnitude of the echo. Various modifications of the hypothesis that rain echoes are produced as a result of incoherent scattering from drops have been proposed at one time or another. Thus, it has been suggested that coherent scattering may occur at times. In Sec. 7·5 reasons are given for believing these suggested modifications in the hypothesis to be without practical significance.

The rapid fluctuation of meteorological echoes is so striking as to provide one of the easiest ways to identify them. On the assumption that the echoes originate from precipitation, the fluctuation is naturally expected because of the changing interference effects among the randomly moving scatterers. A detailed discussion of the fluctuations, as an example of clutter-echo fluctuations in general, was presented in Chap. 6, especially Secs. 6·18 and 6·21, and need not be repeated here.

7·1. The Echo from Incoherent Scatterers Distributed in Volume.— The amplitude of the r-f signal received by a radar is modulated by the range variation of the specific targets. The frequency of this modulation, however, is always very slow compared with the radio frequency. It is always possible, therefore, to set up a small interval of time, large compared with the r-f period, during which the amplitude of the echo is constant. The echo r-f power during such an interval, whose center lies at a given time t_0 after the transmitted pulse, will be called the *instantaneous echo power* at a range $R_0 = ct_0/2$. This instantaneous echo power will, of course, vary from one sweep to the next, producing the *fluctuation* of the echo. By *average echo power* at a given range will be meant the average of the instantaneous echo power over many pulses.

Let R_j, θ_j, ϕ_j be the range, elevation, and azimuth coordinates of some particular (point) scatterer in the target. If $p(R - ct)$ gives the instantaneous r-f power of the transmitted pulse as a function of range, then the power incident on the scatterer at such an instant of time that the returned echo is received at a time t_0 is proportional to $p(2R_j - ct_0)$. The contribution of this scatterer to the total received power (in free space) is then

$$P_{rj} = \frac{p(2R_j - ct_0)\, G^2\lambda^2}{(4\pi)^3 R_j^4} f^4(\theta_j, \phi_j)\, \sigma_j, \tag{1}$$

where σ_j is the radar cross section of the scatterer and $f(\theta,\phi)$ is the antenna pattern. In Sec. 7·5 it will be proved in detail that the total back scattering can be treated as incoherent provided (1) the distribution of scatterers is random and (2) the rise and fall time of the pulse is long compared with the r-f period. The total average received power is, therefore, just the sum of the echo powers from each of the scatterers:

$$\overline{P}_r = \frac{G^2\lambda^2}{(4\pi)^3} \sum_{\text{all } j} \frac{p(2R_j - ct_0)}{R_j^4} f^4(\theta_j, \phi_j)\, \sigma_j. \tag{2}$$

It will now be assumed that the number of contributing scatterers is very large and that they are uniformly distributed in space, completely filling the radar beam, and that the distribution of cross sections is independent of position in space. The sum in Eq. (2) can be then replaced by an integral by introducing a density function $n(\sigma)d\sigma$, which gives the number of scatterers per unit volume with cross section lying between σ and $\sigma + d\sigma$. If an element of volume in space is written as $R^2\, dR\, d\Omega$, where Ω represents the solid angle, then Eq. (2) reduces to

$$\overline{P}_r = \frac{G^2\lambda^2}{(4\pi)^3} \int_0^\infty \frac{p(2R - ct_0)}{R^2}\, dR \int_{(4\pi)} f^4(\theta, \phi)\, d\Omega \int_0^\infty n(\sigma)\, \sigma\, d\sigma. \tag{3}$$

When the range is large compared with the pulse length, R is essentially constant and equal to R_0 throughout the region where $p(2R - ct_0)$ is differ-

ent from zero and hence may be taken outside the first integral. If the pulse were exactly rectangular in shape, this integral would be given by

$$\int_0^\infty p(2R - ct_0)\, dR = P_t \frac{\tau c}{2}, \tag{4}$$

where P_t is the peak transmitted power and τ the pulse duration. The pulse is never exactly rectangular, but it is obvious that Eq. (4) still holds to a good degree of approximation if the conventionally defined values are substituted for P_t and τ.

It is clear that the second integral of Eq. (3) defines approximately the solid angle of the beam. If the beamwidths are not too large (1 radian), then the integral can be said to define two beamwidths Θ, Φ such that

$$\int f^4(\theta, \phi)\, d\Omega = \Theta\, \Phi. \tag{5}$$

It should be pointed out that although Θ and Φ are proportional to the vertical and horizontal beamwidths defined in terms of the half-power points, they are not necessarily identical with them. Thus, if the antenna pattern is approximated by a Gaussian curve, Θ or Φ is about 80 per cent of the angle between half-power points. Usually this difference is of no importance.

The last integral is, of course, simply the definition for the average cross section and can be written

$$\int_0^\infty \sigma n(\sigma)\, d\sigma = N\bar{\sigma}, \tag{6}$$

where N is the total number of scatterers per unit volume and σ is the average cross section. It is convenient to replace the product $N\bar{\sigma}$ by a single symbol η, which may be called the *radar cross section per unit volume.* The dimensions of η are those of reciprocal length; and whenever numerical values are given in later sections, the units will be reciprocal meters. Occasionally η will be expressed in decibels referred to a level of 1 m^{-1}, that is, 10 $\log_{10}(\eta/1\mathrm{m}^{-1})$.

By combining these results, we find the final phenomenological formula for the average received power to be

$$\overline{P_r} = \frac{P_t G^2 \lambda^2}{(4\pi)^3 R^2}\, \Theta\, \Phi\, \frac{\tau c}{2}\, \eta. \tag{7}$$

It should be noted that $\frac{1}{2}\Theta\Phi\tau cR^2$ is just the volume illuminated by the beam. In fact, Eq. (7) could have been derived, though not as rigorously as above, by the following brief argument. If the scattering is incoherent, the total echo power is proportional to the number of scatterers or, with uniform density, proportional to the volume. The total received power is

then given by the echo from a *single* scatterer of average cross section $\bar{\sigma}$ times the number of scatterers per unit volume times the volume illuminated by the beam. Equation (7) then follows immediately, with $\eta = N\bar{\sigma}$.

There are a number of instructive variant forms into which Eq. (7) can be put. Thus, the product $P_t\tau$ is proportional to the *average* transmitted power p_t, and one can write

$$\overline{P_r} = \frac{p_t G^2 \lambda^2}{(4\pi)^3 \omega_r R^2} \Theta \Phi \frac{c}{2} \eta, \tag{8}$$

where ω_r is the pulse-recurrence frequency. Also, $\Theta\Phi$ is approximately equal to $4\pi/G$, where G is the antenna gain. The antenna gain is given by $4\pi A/\lambda^2$, where A is the effective area of the antenna; hence Eq. (7) can be rewritten as

$$\overline{P_r} = \frac{P_t A}{4\pi R^2} \frac{c\tau}{2} \eta \tag{9}$$

or

$$\overline{P_r} = \frac{p_t A}{4\pi \omega_r R^2} \frac{c\eta}{2} \tag{10}$$

Thus at constant λ the echo will be directly proportional to the *area* of the antenna, whereas for a point target the echo is proportional to the *square of the area.*

The preceding discussion has not specified the nature of the scatterers beyond requiring that they be distributed randomly with uniform density. The results therefore are equally applicable in a wide variety of theories about the origins of precipitation echoes.

7·2. Evidence of Direct Correlation between Meteorological Echoes and Precipitation.—The most direct way of determining the origin of meteorological echoes is to have an observer with appropriate instruments present at the point where the scatterers appear to originate. This method has obvious technical difficulties and has rarely been attempted on a sufficiently large scale or with adequate precautions. Where it has been employed, it has been found almost invariably that the echo is connected with precipitation of some form. In a few early experiments carried out at the Radiation Laboratory, an airplane was directed to the region of the echo and precipitation was invariably observed. During the summer, echoes were often observed to move in the direction of the radar system. The passage of these echoes directly over the station was always accompanied by rain from showers or thunderstorms. Occasionally it happened that echoes were received directly overhead from thunderstorm cumulus clouds several minutes before the onset of rain. It seems obvious that in such cases rain was present but was kept aloft by updrafts.

The most extensive investigation of this sort to date has been reported by the Canadian Army Operational Research Group.[1] During the summer months an almost continuous watch was kept for meteorological echoes on a high-power radar system operating on 10 cm. The recurrence frequency was sufficiently low that a sweep length of 160 miles was available on the PPI. The progress of the storm echoes was carefully watched visually and continuously recorded with a motion-picture camera. By these arrangements the possibility of confusion because of "second-sweep echoes," that is, echoes of range greater than what corresponded to the interpulse time, was eliminated. (Thunderstorm echoes have occasionally been received at ranges of 250 miles, and the presence of such second-sweep echoes on high-recurrence-frequency systems has often led to erroneous conclusions.)

To provide the meteorological information a network of ground observers was set up, the network covering the region visible to the radar. When an echo was observed near or directly over one of the observers, a telephonic report of the current weather was obtained. *In the 48 cases in which an observer was located in the echo area rain was invariably reported.* Apparently, and somewhat unexpectedly, there was no instance of rain suspended aloft. The converse of this correlation was *not* found, that is, light to moderate rain was occasionally reported where no echo was found. This is not unexpected, as is shown by a calculation of the expected intensity carried out in the next section.

A similar experiment was carried out in winter by the same group.[2] Some instances were reported of echoes obtained in the absence of rain, but these were definitely attributable to snowstorms.

The experiments cited here are the only known investigations in which a deliberate and careful attempt was made to ascertain the meteorological nature of the target. But there is a very large body of observations of a more casual nature supporting the thesis that the echoes arise only from precipitation. Thus storm echoes were used extensively during the war in several theaters of operation for routing of airplane flights to avoid dangerous precipitation areas. Although admittedly not a scientific proof, this universal procedure for such practical purposes lends considerable support to the hypothesis that connects the echoes with precipitation areas.

[1] J. S. Marshall, R. C. Langille, W. M. Palmer, R. A. Rodgers, G. P. Adamson, and F. F. Knowles, "Summer Storm Echoes on Radar MEW," CAORG Report No. 18, November, 1944. Also J. S. Marshall, L. G. Eon, and L. G. Tibbler, "An Analysis of Storm Echoes in Height Using MHF," CAORG Report No. 30, June 25, 1945; and J. S. Marshall, R. C. Langille, and W. M. Palmer, "Measurement of Rainfall by Radar," *Jour. Meteor.*, **4**, 186 (1947).

[2] J. S. Marshall, R. C. Langille, W. M. Palmer, and L. G. Tibbler, "S-band Radar Echoes from Snow," CAORG Report No. 26, June, 1945.

Mention must also be made of reports of meteorological echoes under circumstances that appear to exclude the presence of precipitation. Most of these are in the nature of casual observations, but a number have been reported in sufficient detail to merit attention. In general the echoes observed fall in two classes. Of the first type are extended echoes closely resembling normal precipitation echoes. One of the most authentic cases of this type appears to be an instance[1] in which an airplane was directed to an area corresponding to an echo observed on a radar operating at 10 cm. Not only was an absence of precipitation noted, but no clouds were present. There were, however, several columns of smoke from brush fires, and the correlation of the echoes with the location of the smoke was said to be so positive as to exclude the possibility of second-sweep echoes.

The second class of echoes has a distinctive appearance, different from the normal precipitation echo. Known as "angels," these echoes have been observed by an increasing number of investigators. The characteristic experimental arrangement apparently favored by most observers has been to point a radar set vertically upward, recording the echoes that appear on the A-scope. With such setups "angel" echoes have been observed in clear weather at ranges up to several thousand yards.[2] They appear most frequently as isolated signals, lasting a fraction of a minute and lacking the usual rapid fluctuation of precipitation echoes. No definite connection with weather conditions has been found, but it is noted they are more numerous on summer nights under calm conditions. "Angels" have been detected on a number of wavelengths from 10 cm to 1 cm. It is worthy of mention that similar echoes have been observed with vertical-looking sonars, using sound wavelengths of the same order of magnitude as microwaves.[3]

With a vertical radar no information is furnished about the horizontal extent of the echoes. Baldwin,[4] however, has observed them with an automatic-tracking radar operating on 3 cm and has found that the echoes appear to be as sharply defined in space as the pulse length and beam width will allow. The "angels" were observed to move in space with speeds the order of magnitude of wind speeds, and Baldwin was able to

[1] Unpublished information from Major J. O. Fletcher, Signal Corps Ground Agency, 1943.

[2] H. T. Friis, "Radar Reflections from the Lower Atmosphere," *Proc. IRE* **35,** 494 (1947); W. B. Gould, "Radar Reflections from the Lower Atmosphere," *Proc. IRE*, **35,** 1105L (1947); *cf.* also A. W. Friend, *Proc. IRE*, **37,** 116 (1949). In addition there seems to be much work as yet unpublished.

[3] G. W. Gilman, A. B. Coxhead, and F. H. Willis, "Reflection of Sound Signals in the Troposphere," *Jour. Acoust. Soc.*, **18,** 274 (1946).

[4] M. W. Baldwin, Jr., BTL Memoranda Nos. MM44–150–2, 3, and 4, July 18, 1944, to Aug. 11, 1944, paper presented at the joint URSI-IRE meeting, May, 1946. A summary will be found in the *BTL Record*, **25,** 75 (1947). See also a recent letter in *Proc. IRE*, **36,** 363 (1948).

track a single echo for a few minutes at a time. Ranges varied between 800 and 2600 yards and heights from ground level to 2000 yards. Though weak, the echoes were clearly above noise, and from the characteristics of the system one can estimate that the radar cross section of the object responsible for the echo lay between 10^{-6} and 10^{-4} m^2.

Similar "angel" echoes have been seen in unpublished observations made at the Radiation Laboratory, and later at the Cambridge Field Station of the Watson Laboratories, using a high-power early-warning system on 10 cm. Confusing ground echo was eliminated by means of MTI. The observations tallied closely with those of Baldwin, the echoes appearing as isolated signals moving irregularly with speeds up to 20 mph. Occasionally such echoes were seen as far as 20 miles, indicating (from the constants of the set) that the cross section could be as large at 10^{-2} m^2. The "angels" seemed especially prominent after sunset, and at times the echoes would almost completely cover the PPI in a stippled dot pattern out to a distance of 10 to 15 miles.

By themselves the experimental data available are not sufficient to determine the source of the "angels" echo. We can exclude precipitation, for all the observations were made in clear weather. The tendency at first seemed to be to ascribe the echoes to inhomogeneities in the index of refraction. Thus, to quote one out of many, Baldwin states (*loc. cit.*): "The most attractive possibility is that the echoes are returned from clumps of water vapor." This possibility is analyzed quantitatively in Sec. 7·4, and it is there shown that the gradients of index of refraction believed to exist in the atmosphere are much too small to account for the observed echoes on microwaves. It cannot be said by any means that this explanation is ruled out, for there may exist a microstructure in the atmosphere too fine to be detected by present instruments. It does serve, however, to make the possibility much less attractive. Much larger signal strength can be computed by assuming that the echo arises by reflection from a horizontal layer,[1] but such an assumption seems unrealistic on two counts. From the theoretical side the notion of a horizontal reflecting layer assumes that the reflection from all parts of the layer is coherent, requiring the layer to be uniform horizontally to within a radar wavelength over the width of the beam. This seems highly unlikely even at 10 cm, much less at 3 or 1 cm. Experimentally, the idea of a horizontal layer seems to have been inspired by the prevalent use of vertical-looking radars to observe the "angels." The experiments of Baldwin and others indicate, however, that the volume of space giving rise to the echo is most likely smaller than can be resolved by the radar and hence cannot be in the form of horizontal layers.

[1] W. E. Gordon, "A Theory of Radar Reflections from the Lower Atmosphere," *Proc. IRE*, **37**, 41 (1949), and A. W. Friend, *loc. cit.*

One is tempted, therefore, to look elsewhere for an explanation. It is known that echoes can be observed even from small birds, and British investigators have paid considerable attention to such echoes.[1] It is stated that the cross section of a large sea bird on 10 cm can be as high as 0.2 m^2. Some of the echoes may possibly be accounted for in this manner, although Baldwin failed to see any in a telescope mounted on his antenna. Recently, Crawford[2] has produced convincing evidence that the "angels" seen on a vertical-looking radar were due to the presence of insects. Many of the observed properties of the "angels" fit in with such a source—their isolated nature, motion in space, and lack of rapid fluctuation—and it seems probable that a large proportion of the echoes arise in this fashion. Conceivably, too, such extraneous phenomena as "second-trip echoes" or antenna side lobes might contribute some of the apparent echoes.

Many of the observed "angels," perhaps most of them, can be explained in one of these ways. However, there seems at the present time to be a residuum of observations that cannot be accounted for by such explanations. Thus "angels" have been seen in winter when the ground was below freezing and the presence of insect life highly unlikely.[3] For these echoes one would have to fall back on the atmospheric-inhomogeneity theory. It deserves repeating, however, that to obtain anything like the right order of magnitude for the echo, the spatial dimensions of the inhomogeneities must be far smaller than any yet observed. It is a matter of inches, not of yards or hundreds of yards. Recent investigations have indeed revealed an atmospheric microstructure much finer than previously conceived,[4] but even so the observations are too coarse for the inhomogeneities of interest. There is a great need for meteorological instruments capable of detecting and measuring the extremely fine microstructure that is needed to account for the echoes.

On the radar side there is need for more controlled observations capable of producing quantitative data. Once we have a body of numerical information on the magnitude of the echoes, their variation with range and radar frequency, and their fluctuation rates, it may be possible to progress beyond the field of speculation to more solidly established reasoning. It is conceivable that such radar research may provide a powerful tool for studying the structure of the lower atmosphere.

[1] D. Lack, "Radar Echoes from Birds," British AORG Report No. **257**, February, 1945.

[2] A. B. Crawford, "Radar Reflections in the Lower Atmosphere," *Proc. IRE*, **37**, 404 (1949).

[3] Private communication from R. Wexler.

[4] J. R. Gerhardt and W. E. Gordon, "Microtemperature Fluctuations," *Jour. Meteorol.*, **5**, 197 (1948).

7·3· The Approximate Magnitude of Rain Echoes on the Drop Theory. The cross section per unit volume η is defined in Sec. 7·1 in terms of the distribution of the scatterers in radar cross section. If the scatterers consist of spherical water drops, differences in cross section are due solely to differences in radii, and η may be given in terms of the distribution of the number of drops per unit volume according to their radii:

$$\eta = \int_0^\infty \sigma n(a)\, da. \tag{11}$$

When the circumference of the drops is very small compared with λ, then the scattering is entirely dipole in character; that is, only the first term in the multipole expansion for the scattered intensity need be retained (*cf.* Sec. 6·1). The dimensions of raindrops are usually small enough to satisfy this condition; but as will be seen in Sec. 7·6, the higher-order scattering cannot always be entirely neglected, especially at the shorter wavelengths. However, here it is desired only to establish the order of magnitude of the scattering, and for this purpose it seems sufficient to consider dipole scattering alone, at least for the longer wavelengths.

Under these conditions the cross section for a drop of radius a is given by

$$\sigma(a) = 4\pi a^2 \left(\frac{2\pi a}{\lambda}\right)^4 \left(\frac{\varepsilon - 1}{\varepsilon + 2}\right)^2, \tag{12}$$

where ε is the dielectric constant. In the microwave region ε is very large compared with unity; thus the last factor in Eq. (12) can be replaced by 1. The cross section per unit volume is then given by

$$\eta = 4\pi \int_0^\infty a^2 \left(\frac{2\pi a}{\lambda}\right)^4 n(a)\, da. \tag{13}$$

As only a rough value of η is desired, it will also be assumed that only drops of a single radius a are present. Since the total number of drops per unit volume N times the volume of each drop, $4\pi a^3/3$, is numerically equal to the mass of water per unit volume W, the formula for η can now be written

$$\eta = 6\pi \times 10^{-4} \frac{W}{\lambda} \left(\frac{2\pi \tilde{a}}{\lambda}\right)^3, \tag{14}$$

where W is expressed, as is conventional, in grams per cubic meter, λ and $\tilde{a}$ in centimeters. The radius $\tilde{a}$ used in Eq. (14), it should be emphasized, corresponds to a much larger radius than the median of the drop-size distribution, as the cross section varies as the sixth power of the radius and the larger drops are proportionately more important.[1] Actually no data

[1] Equation (14) follows rigorously from Eq. (13) if $\tilde{a}^3$ is defined as the mean sixth power of the radius divided by the mean cubed of the radius. A more detailed discussion of drop-size distributions and the calculation of η from Eqs. (11) and (13) is given in Secs. 7·6 and 7·7.

for W and $\bar{a}$ are available for the precipitation from which echoes have been measured. The comparison of the experimentally determined intensities and the expected values will therefore be carried out by first assuming a reasonable value for W. Using the observed values of η the corresponding magnitudes of $\bar{a}$ will then be calculated from Eq. (14). It will be shown that the drop radii thus arrived at are reasonable.

The intensities of a number of precipitation echoes have been measured using the same radar systems and electrical technique for averaging the received power that have been described in Sec. 6·9. The cross sections per unit volume η of several summer thunderstorms and cloudbursts at a wavelength of 9.2 cm were found to lie between -60 and -70 db, with -65 db being an average figure. (It will be remembered that η expressed in db's is referred to a level of one reciprocal meter.) Because the indications are that the rate of rainfall in all cases was in excess of 1 in./hr, it seems reasonable to assign W a value of 1 g/m^3. With the figure of -65 db for the calculated radius, $\bar{a}$ is 0.18 cm, which is not out of line with accepted values for very heavy rain.

In one instance the cross section per unit volume for 9.2 cm for general precipitation accompanying a "northeaster" was about -85 db with a variation of $\pm$ 7 db. The rate of precipitation was a little less than 0.1 in./hr (light to moderate rain according to Humphreys' classification) for which a value of W about 0.1 g/m^3 is usually quoted. The corresponding figure for $\bar{a}$ is 0.08 cm, a result that again seems quite reasonable, as drop radii up to 0.15 cm are reported in the literature for light rains.

Some rough measurements of the intensity of echoes from tropical showers have been reported,[1] using an airborne system on $\lambda = 3.2$ cm, ASD-1. Thus, as a typical example, the "peak echo" from a shower 5 miles away was measured to be 85 db below 1 watt. The fluctuations of precipitation echoes are such that there is a finite probability of receiving any echo no matter how high. Hence the term "peak echo" is somewhat ambiguous, and the relation between this figure and the average power $\overline{P}_r$ cannot be stated with any degree of certainty (*cf.* the similar discussion for sea echo, Sec. 6·10). However, it is probable that the difference is about 5 db, plus or minus a few db's, so that the quoted figure for "peak echo" corresponds to $\overline{P}_r = -90$ db. With the constants of the system usually quoted, the calculated value of η is -56 db. Assuming then that $W = 1$ g/m^3, the drop radius found from Eq. (14) is then $\bar{a} = 0.08$ cm. The result is certainly reasonable; in fact it looks somewhat smaller than might be expected. It should be remembered, however, that the absolute calibration of the system was extremely crude, and this could well account for the apparent discrepancy.

[1] A. E. Bent, "Echoes from Tropical Rain on X-band Airborne Radar," RL Report No. 728, June 15, 1945.

Direct attempts at verifying the wavelength dependence predicted by the Rayleigh law have been made by simultaneous measurements of η on 9.2 and 3.2 cm using the radar systems described in Sec. 6·9. (The absolute values of η quoted in the first part of the present section were obtained from this experiment.) The expected ratio of the cross section per unit volume for the two frequencies is +18 db. Actually, the measured ratio for thunderstorms and cloudbursts averaged between +12 and +15 db. Of this difference several db's may be attributable to inaccurate calibration of the systems, but it is not believed that the entire difference could be accounted for in this manner. It had been thought that the remainder of the difference results from the beginning of the breakdown of the Rayleigh law at the larger drop sizes and the shorter wavelengths. Recent rigorous calculations show, however, that the cross section at 3 cm is actually above the Rayleigh law for drops smaller than 0.9 cm in diameter. We are therefore without an explanation for this discrepancy.

In any case the differences between the experimental and predicted ratios should not be unduly stressed. Even at worst the figures correspond to a wavelength dependence of λ^{-3}. This is sufficiently close to λ^{-4} as predicted by the Rayleigh law to favor decisively the drop theory over some other nonfrequency sensitive mechanism such as ions or refractive index fluctuations. These mechanisms were proposed at a time when accurate measurements of η were impossible and it was incorrectly believed that the drop theory was qualitatively inadequate to explain the magnitude of the echoes. Other consequences of these mechanisms are examined in detail in Sec. 7·4, where it is shown that their contributions to the echo are negligible.

7·4. Possible Alternative Theories to Scattering by Drops.[1]—Several theories have been proposed to account for meteorological echoes that do not involve scattering from drops. Some of these theories will be analyzed in this section, and it will be shown that they predict echo intensities far below what is observed or computed from the drop theory.

It has been pointed out that raindrops become electrically charged in the process of formation and that consequently there exists in space a neutralizing cloud of ions. Under the influence of the incident electric field these ions will oscillate and in so doing will reradiate energy in all directions, some of which will be received by the radar. A quantitative expression for this energy is easy to calculate by considering the ion as a dipole performing forced oscillations in the incident field. If $E_0e^{+i\omega t}$ represents the incident electric field at the ion, then the equation of motion of the ion is

$$m\frac{d^2x}{dt^2} = eE_0e^{+i\omega t}, \tag{15}$$

[1] This section is based on an unpublished memorandum by A. J. F. Siegert.

neglecting damping forces (about which very little can be said quantitatively). The amplitude of the forced oscillation of the ion is then given by

$$x_0 = \frac{eE_0}{\omega^2 m}. \tag{16}$$

The Poynting vector for the field radiated in the direction of the radar by an oscillating dipole of moment $p = ex_0e^{+i\omega t}$ is given by[1]

$$S = \frac{\omega^4 e^2 x_0^2}{2(4\pi)^2 c^3 \epsilon_0} \frac{1}{R^2} = \frac{e^4 E_0^2}{2(4\pi m)^2 c^3 \epsilon_0} \frac{1}{R^2}, \tag{17}$$

where R is the distance between the ion and the radar. The power received by the radar is then

$$P_r = \frac{SG\lambda^2}{4\pi} = \frac{(e^2E_0)^2 G\lambda^2}{2(4\pi)^3 m^2 c^3 \epsilon_0 R^2}. \tag{18}$$

The amplitude of the incident field E_0 is related to the incident Poynting vector and to the (peak) transmitted power by

$$S_i = \frac{E_0^2}{2}\sqrt{\frac{\epsilon_0}{\mu_0}} = \frac{P_t G}{4\pi R^2}. \tag{19}$$

By means of Eq. (19), E_0 in Eq. (18) can be expressed in terms of the transmitted power, and after some simplification the formula for the received power reduces to

$$P_r = \frac{P_t G^2 \lambda^2}{(4\pi)^3 R^4} \frac{1}{4\pi} \left(\frac{e^2}{\epsilon_0 mc^2}\right)^2. \tag{20}$$

By comparison with the standard formula for the power received by the radar it is clear that the radar cross section of each ion is

$$\sigma = \frac{1}{4\pi}\left(\frac{e^2}{\epsilon_0 mc^2}\right)^2; \tag{21}$$

That is, the cross section per ion is of the order of the classic electron radius squared (assuming one electronic charge per ion). If the ions radiate incoherently, then the cross section per unit volume η is just

$$\eta = \frac{1}{4\pi} \frac{Qe^3}{\epsilon_0^2 m^2 c^4}, \tag{22}$$

where Q is the total charge per unit volume. According to a figure quoted in Humphreys[2] the charge per unit volume in a rainstorm rarely exceeds

[1] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, pp. 435 and 284.

[2] W. J. Humphreys, *Physics of the Air*, McGraw-Hill, New York, 1940. More recent data by R. Gunn, *Phys. Rev.*, **71**, 181 (1947), would seem to indicate that the free-charge density even in a thunderstorm is not much more than 10^{-6} coulomb/m^3.

2×10^{-3} coulomb/m^3. If the ions are then assumed to be all electrons, the value of η turns out to be -140 db above 1 meter. This figure is far below the experimentally measured values of -60 to -85 db quoted in the preceding section.

It does not seem possible to introduce modifications into this theory that would increase the expected cross section substantially. On the contrary most of the factors neglected in the preceding discussion would tend to decrease the cross section, as, for example, the presence of damping forces. If the ions were of atomic mass rather than electrons, as would seem more plausible, the cross section would be reduced by a factor of at least 4×10^6. Some increase in the cross section would be expected if all the ions released by one raindrop in its formation stayed sufficiently close together to radiate coherently. The total cross section per unit volume is then obtained by multiplying Eq. (21) by n^2N, where n is the number of ions per raindrop and N is the number of drops per unit volume. As Q is nNe, this results in increasing η by a factor of n, which would have to be excessively large in order to bring the cross section up to observable levels.

Finally it should be pointed out that the predicted cross section of Eq. (22) is independent of wavelength, which is in violent contradiction with experiment. Scattering from atmospheric ions is thus completely inadequate to explain meteorological echoes. (Scattering by ions does enter in such cases as the echo from an exploding shell, where the charge density is much greater.)

As has been mentioned in Sec. 7·2, it has often been suggested that the meteorological echoes occasionally observed under conditions apparently excluding the presence of precipitation are caused by scattering from inhomogeneities in the atmospheric dielectric constant. There are two possible techniques that can be used to investigate the magnitude of the echoes arising from such inhomogeneities. One can obtain the appropriate solutions of the wave equation for a medium in which the velocity of propagation varies with position in space and then calculate the Poynting vector back in the direction of the radar. Alternatively one can go back to the microscopic source of the index of refraction, the polarization of the gas molecules under the influence of the incident electric field. Each of these molecules will then radiate like a dipole, and the radiated energy received by the radar from the assembly of illuminated molecules can then be calculated. It is this latter method which will be followed, as the framework for dealing with scattering from large aggregations of scatterers has already been set up.

A strictly uniform and stationary distribution of molecules will not exhibit back scattering, for the distribution of the phases of the scatterers will be the same as for points uniformly distributed along the incident wave train, and the scattered radiation will therefore interfere destructively except in the forward direction. Thus a perfect crystal or a uniform

medium such as glass transmits light without scattering. Scattering will occur only when there are deviations from uniformity. Such deviations can arise in two ways. Even if the time average of the distribution is uniform, at any given instant the density of scatterers will exhibit statistical fluctuations about the average. These statistical density fluctuations result in a net amount of energy being scattered which is proportional to the number of scatterers and which is therefore designated as the "incoherent" scattering. It is just this type of mechanism which has been discussed in connection with the scattering by raindrops and which in the case of visible light and air molecules gives rise to the blue of the sky. It is highly unlikely that the analogous molecular scattering at microwaves could be observed with detection sensitivities that can be visualized at present, as the scattering per molecule is so small.

In addition scattering will also occur if the time-average distribution is not uniform,[1] as in the case of atmospheric inhomogeneities. The scattering in this case is coherent; that is, it depends upon the square of the number of scatterers present. Now, the space variation of particle density giving rise to the scattering can be real or can be a fictitious variation due to the pulsed nature of the incident field. Presumably this latter effect could result in a certain amount of coherent scattering from even a uniform distribution of raindrops. The calculation of the scattering by atmospheric inhomogeneities of the index of refraction and of the coherent scattering from raindrops therefore belongs to the same problem and can be treated together. Appendix B presents a treatment of the problem due to A. J. F. Siegert that is applicable to both cases. It is there shown [Eq. (B·17)] that the radar cross section for the coherent scattering is

$$\sigma_c = |p|^2 \left| \int_0^\infty \bar{n}(R) e^{-\frac{4\pi i R}{\lambda}} \, dR \right|^2, \tag{23}$$

where $|p|^2/4\pi$ is the cross section per scatterer and $\bar{n}(R)$ is the time average of the number of scatterers per unit distance. In terms of the density of molecules $\rho(R)$ this can be written

$$\sigma_c = |p|^2 \left| \int \rho(R) e^{-\frac{4\pi i R}{\lambda}} \, dV \right|^2, \tag{24}$$

the volume integration to be carried over the volume illuminated by the radar beam. Under the influence of the incident field the gas molecule acts like a radiating dipole with p given by[2]

$$p = \left(\frac{2\pi}{\lambda}\right)^2 \frac{\alpha}{\sqrt{4\pi\,\epsilon_0}}. \tag{25}$$

[1] "Time average" means over a time long compared with the time between molecular collisions.

[2] *Cf.* Stratton, *loc. cit.*

The quantity α appearing in Eq. (25) is the molecular polarizability which in a gas is related to the index of refraction n by the formula

$$\rho\alpha = \epsilon_0(n^2 - 1) \approx 2\epsilon_0(n - 1), \tag{26}$$

as n is very close to unity. Substituting Eqs. (25) and (26) into Eq. (24), the cross section can be written

$$\sigma_c = \frac{1}{\pi}\left(\frac{2\pi}{\lambda}\right)^4 \left| \int (n - 1)e^{-\frac{4\pi i R}{\lambda}} dV \right|^2. \tag{27}$$

This expression may be put in a frequently useful form by performing an integration by parts:

$$\int_0^\infty (n - 1)e^{-\frac{4\pi i R}{\lambda}} dR = \frac{-\lambda}{4\pi i}(n - 1)e^{-\frac{4\pi i R}{\lambda}}\Big|_0^\infty + \frac{\lambda}{4\pi i}\int_0^\infty \frac{d(n - 1)}{dR} e^{-\frac{4\pi i R}{\lambda}} dR.$$

The first term here always vanishes, if for no other reason than that the pulsed nature of the radiation makes $(n - 1)$ effectively zero at the origin and at infinity. Hence Eq. (27) can also be written as

$$\sigma_c = \frac{1}{4\pi}\left(\frac{2\pi}{\lambda}\right)^2 \left| \int \frac{d(n - 1)}{dR} e^{-\frac{4\pi i R}{\lambda}} dV \right|^2. \tag{28}$$

Perhaps the simplest type of inhomogeneity that might be considered is a uniform gradient of the index, starting at a distance R from the radar, ending at $R + \Delta R$, and extending over an area A perpendicular to the direction of the radar beam. The resulting cross section is [cf. Eq. (A·19)]

$$\sigma_c = \frac{A^2}{4\pi}\left[\frac{d(n - 1)}{dR}\right]^2 \sin^2\left(\frac{2\pi\,\Delta R}{\lambda}\right). \tag{29}$$

It is seen that the maximum cross section is independent of wavelength. Information on the magnitude of the inhomogeneities existing in the atmosphere is very scanty, chiefly because sufficiently accurate and rapid meteorological instruments are not available. In the investigation mentioned in Sec. 6·14 the largest observed gradient of the index of refraction was 2×10^{-7}/m. Such a gradient would provide a radar cross section of 10^{-6}m^2 (sufficient to account for some of the "atmospheric" echoes mentioned in Sec. 7·2) if it extended over an area 150 yd on a side. Although at first sight this figure may seem reasonable, it is actually entirely too large an area to be practicable. Our model requires that all portions of the volume of inhomogeneity start and stop at the same distance from the radar, as if we were dealing with the reflection from a plane sheet normal to the radar beam. This condition cannot be met in actuality with an area whose dimensions are so large compared with the wavelength. Instead it is likely that the area must be broken up into a number of small patches.

Within each patch the radiators scatter coherently, but the patches must be considered as radiating incoherently relative to each other for the time-average echo.[1] From Eq. (29) the time-average cross section is then

$$\sigma_c = \frac{}{8\pi}\overline{\left(\frac{d(n-1)}{dR}\right)^2}\left(\frac{A_p}{A}\right), \tag{30}$$

with A the normal area illuminated by the beam and A_p the projected area of a "coherent" patch. Clearly any reasonable guess for such a patch area would result in a cross section many powers of 10 smaller than would be apparently given by Eq. (29).

Note added in proof: The recent work of J. R. Gerhardt and W. E. Gordon, "Micro-temperature Fluctuations," *Jour. Meteorol.*, **5**, 197 (1948) has revealed an atmospheric microstructure finer than that used in the illustrations above, but the qualitative conclusions remain the same.

A thoroughgoing treatment of the echo arising from atmospheric inhomogeneities would require considering the index of refraction as a random function of space and time with characteristic frequency spectra and then applying the statistical methods of Sec. 6·19 to the evaluation from the cross section as given by Eq. (27).[2] To consider a simpler method, let us suppose that the inhomogeneities consist of irregularly spaced "clouds" of molecules, such that in each cloud the distribution of the index of refraction is

$$(n - n_0) = n_1 e^{-\frac{\pi}{a^2}(x^2+y^2+z^2)}, \tag{31}$$

the origin being at the center of the inhomogeneity and a being a measure of the size of the cloud. The cross section per inhomogeneity may be easily evaluated from Eq. (27) and turns out to be

$$\sigma_c = \frac{1}{\pi}\left(\frac{2\pi}{\lambda}\right)^4 n_1^2 a^6 e^{-\frac{8\pi a^2}{\lambda^2}}. \tag{32}$$

The separate inhomogeneities scatter incoherently with respect to each other, and their contributions to the total cross section are therefore simply additive. Hence the cross section per unit volume is

$$\frac{\sigma_c}{V} = \eta = \frac{1}{\pi}\left(\frac{2\pi}{\lambda}\right)^4 n_1^2 a^6 e^{-\frac{8\pi a^2}{\lambda^2}} N,$$

where N is the number of inhomogeneities per unit volume. We are certainly overestimating the magnitude of the scattering if we take the inho-

[1] The time average now being over a time long compared with that required for the inhomogeneities to move $\lambda/2$ with respect to each other.

[2] For recent applications of such techniques see a series of reports by N. G. Parke, "Microwaves in an Irregular Atmosphere," under Air Force Contract No. AF 19(122) 91; also H. G. Booker and W. E. Gordon, *Proc. IRE*, **38**, 401 (1950).

mogeneities to be spaced a distance $2a$ apart, and on this basis the cross section per unit volume is

$$\eta = \frac{1}{8\pi a}\left(\frac{2\pi a}{\lambda}\right)^4 n_1^2 e^{-\frac{8\pi a^2}{\lambda^2}}. \tag{33}$$

From the data presented in Sec. 6·14 one can estimate a value for a around 40 yd. At a wavelength of 10 cm the value of η would then be about $n_1^2 \times 10^{-1.7\times 10^6}$. It may well be argued, however, that the inhomogeneities present are much smaller in size than present meteorological instruments can detect. The maximum value of η [Eq. (33)] occurs for a value of a very close to $\lambda/4$. Even then, for a wavelength of 10 cm, a value of η as large as 100 db below 1 m would require fluctuations in n of 10^{-5} within the space of 1 in.

The conclusion one is led to by all these examples seems inescapable: Inhomogeneities in the index of refraction will not reflect an appreciable radar signal in the microwave region unless the fluctuations in the index are more fine grained and larger in magnitude than we have reason to expect at present. It cannot be too strongly emphasized that further progress in this field depends vitally on investigations into the microstructure of the atmosphere, a field at present practically untouched.

7·5 Modifications of the Drop Theory.—Early measurements of the intensities of precipitation echoes were of a semiquantitative nature and often quite inaccurate Many of them, however, seemed to indicate that the echo was much larger than could be accounted for by reasonable figures for drop size and concentration. The more accurate (though still relatively crude) measurements detailed in Sec. 7·2 have now dispelled such doubts. At the time, a number of modifications of the picture of incoherent scattering from drops were proposed, all tending to increase the magnitude of the echo.

Some of them can be disposed of quickly. Thus it was suggested that there were aerodynamic forces present between two falling raindrops tending to draw them together. If the forces were sufficiently strong, the two drops might come close enough so that they radiated coherently. Without going into the magnitude of the forces involved, it is clear that at best such a violation of the statistical independence of the drop positions would increase the amplitude of scattering by a factor of 2 and the intensity by a factor of 4. With present techniques, such a factor lies within the uncertainties of measurement. Again, it was suggested that aerodynamic and electrical forces might distort the raindrop from a spherical shape. However, any reasonable distortion, that is, short of producing an ellipsoid whose major axis corresponds to the length of a resonant dipole, would change the scattering by only a small factor, and it is only orders of magnitude that would be significant practically. In any case, such distortions are negligible, as it has been shown to a high

degree of accuracy that the intensity of meteorological echoes is independent of the direction of polarization of the incident radar wave (cf. Vol. 1 of this series, Sec. 3·10).

Requiring more attention is the suggestion that the scattering is not entirely incoherent, but that part is coherent. This is an especially attractive hypothesis, as incoherent scattering is proportional only to the number of scatterers, but coherent scattering is proportional to the *square* of the number, thus providing the possibility of an enormous increase in echo. The relative proportion of incoherent and coherent back scattering from an assembly of particles has already been discussed in Sec. 7·4 in connection with the echoes from index of refraction gradients and examined in detail in Appendix B. One has only to change the name of the scatterers from gas molecules to raindrops. It was pointed out in Sec. 7·4 that a perfectly uniform spatial distribution of like particles would not scatter backward, for the distribution of phases of the individual scattered amplitudes would then be uniform, and the total sum would cancel to zero. An echo will be obtained only if there is some inhomogeneity in the distribution, some lack of uniformity. Thus even a distribution that in the average is uniform will deviate instantaneously from the average if the scatterers are moving independently, resulting in a net scattered amplitude at any given instant. This instantaneous amplitude will vary with time over large limits and will average[1] to zero. There will, however, be a time-average scattered intensity different from zero, proportional to the number of scatterers, hence representing the incoherent scattering.

In addition to this source of nonuniformity the time-average distribution itself might be nonuniform, either because of actual inhomogeneities in the spatial distribution of the raindrops or because of the pulsed nature of the radar which causes the amplitude of the incident field to vary rapidly in space at any given instant. This latter source is likely to cause the larger effect, as a pulse rise time of 0.1 μsec (16 yd) is not uncommon and much shorter rise times are technically possible.[2] On the other hand it seems highly unlikely that the edges of a cloud of raindrops would be as sharply defined. We therefore confine our attention to the apparent inhomogeneities caused by the rise and fall times of the radar pulse.

The scattering from such time-average nonuniformities is proportional to the square of the number of scatterers involved and therefore represents a coherent contribution [cf. Appendix B, Eq. (B·7)]. Clearly the "sharper" the pulse the larger will be the magnitude of the effect.

[1] Time average here denotes averaging over a time long compared with that required for a scatterer to move $\lambda/4$.

[2] It should be pointed out that it is the dimensions of the received pulse in the i-f section of the receiver that are important, not those of the transmitted pulse.

In fact, for a pulse with discontinuous sides the coherent scattering can be of the same order of magnitude as the incoherent scattering or larger. This is, however, an unrealistic picture of a pulse, and a less artificial model might be a trapezoidal pulse, i.e., one with a flat top and uniformly sloping sides. It is shown in Appendix B [Eq. (B·21)] that the ratio of the coherent to incoherent scattering with such a pulse is always less than

$$\frac{\sigma_c}{\sigma_i} \leqslant N\left(\frac{\lambda}{2\pi a}\right)^4\left(\frac{a}{b}\right)^2, \tag{34}$$

where b is the pulse length between 50 per cent points in amplitude, a is the width of the sloping sides of the pulse, and N is the total number of scatterers in the beam, given by

$$N \approx nR^2\Theta\Phi b \tag{35}$$

(a uniform cloud of like scatterers is assumed for simplicity). In Eq. (35) n is the density of scatterers, R the range and Θ and Φ the beamwidths of the radar.

The factor $(a/b)^2$ in Eq. (34) arises simply from the proportion of scatterers in the sloping portions of the pulse length to the total number and is thus a geometrical factor. On the other hand the term involving $\lambda/2\pi a$ is of great physical importance; it tells us that the "unit" in terms of which we are to measure the dimensions of the inhomogeneity is $\lambda/2\pi$. Stated otherwise, coherent contribution will be the larger the sharper the pulse edges are in relation to $\lambda/2\pi$. Suppose one were deliberately to design a radar to obtain the maximum coherent scattering. It would have a short, fast-rising pulse and at the same time operate on the longest wavelength feasible and have a broad beam so as to include a large number of scatterers. These last properties are not usually associated with short pulses and consequent high resolution in range. A practicable microwave system approaching these characteristics might have the following specifications:

$$\begin{aligned} \lambda &= 10 \text{ cm}, \\ a/b &= 1, \\ a &= 0.1\ \mu\text{sec} = 16 \text{ m}, \\ \Theta &= \Phi = 0.1 \text{ radians}. \end{aligned}$$

A figure of 2000 drops per cubic meter is reasonable for a thunderstorm, and under these conditions and at a range of 50 miles the maximum possible ratio of coherent to incoherent scattering is about

$$\frac{\sigma_c}{\sigma_i} = 0.16,$$

which is undetectable. For more normal radar specifications and especially at shorter wavelengths, the ratio would be even smaller.

Furthermore, a trapezoidal pulse is still not a good picture of an actual pulse, as it is discontinuous in slope. A real pulse will be continuous in all derivatives. One would expect therefore that a pulse made up of error integral curves, in which all derivatives are continuous, would be closer to actuality. The differentiated pulse would then consist of two Gauss error curves (with opposite signs). If these curves are written in the form of

$$Ae^{-\frac{\pi R^2}{a^2}}$$

where A is a constant determined by normalization, then the ratio of coherent to incoherent scattering [Eq. (B·23)] turns out to be always less than

$$\frac{\sigma_c}{\sigma_i} \leqslant N\left(\frac{a}{b}\right)^2\left(\frac{\lambda}{2\pi a}\right)^2 e^{-\frac{\pi}{2}\left(\frac{2\pi a}{\lambda}\right)^2}, \tag{36}$$

where b has the same significance as in Eq. (34). The quantity a is roughly the distance in which the pulse rises from 0.03 to 97 per cent of maximum power. Equation (36) depends upon the quantities involved in roughly the same fashion as in the previous case, but the magnitude clearly is much smaller. Using the foregoing figures, $2\pi a/\lambda \approx 10^3$ and the coherent scattering on this basis is completely negligible.

With actual pulses the coherent scattering probably lies between the two models used, and it is likely to be closer to the second one. In any case it seems certain that the effect is completely unimportant with present and foreseeable techniques for short pulses and wideband receivers.[1]

THE INTENSITY OF METEOROLOGICAL ECHOES

By Herbert Goldstein

The preceding sections have presented the evidence upon which is based the belief that meteorological echoes arise as incoherent scattering from raindrops or other condensed forms of water. In the course of the discussion an approximate formula, valid for small drops, was given for the magnitude of the scattering cross section per unit volume to be expected on this hypothesis [Eq. (7·14)]. We have now to show how the accurate cross section for any drop size is to be calculated, a task requiring knowledge of both the radar cross section of a single drop as a function of radius and the drop-size distribution. Obtaining the first is almost entirely a problem in electromagnetic theory and can be carried out to any desired degree of accuracy. The second, however, requires purely meteorological measurements that are extremely difficult to make and for which no meaningful values are at present available. Hence, accurate calculation of the echo to be expected from a precipitation area

[1] *Note added in proof:* This problem has also been considered by C. Domb [*Proc. Camb. Phil. Soc.*, **43**, 587 (1947)], and similar conclusions were reached.

must at the present time remain an academic problem. We shall, however, outline the procedure to be followed when more reliable drop-size distributions are obtained in the future.

Section 7·6 reviews briefly the defining formulas for the cross section per unit volume and discusses the factors entering into it besides the drop-size distribution. In particular, methods of computing the cross section for a single drop are described in detail. In the following section (7·7) the difficulties attendant upon meaningful measurements of the drop-size distribution are discussed briefly. Finally, in Sec. 7·8 an indication is given of the method of calculating the echo from precipitation other than liquid water such as snow, ice crystals, and hailstones.

7·6. The Radar Cross Section of Single Drops.—In Sec. 7·1 it was shown that the average echo, at a given range, from a "cloud" of water drops was proportional, among other things, to the volume of the cloud illuminated by the beam and to a cross section per unit volume η defined by Eq. (6):

$$\eta = \int_0^\infty \sigma n(\sigma)\, d\sigma = \int_0^\infty \sigma(a) n(a)\, da. \tag{37}$$

Here $\sigma(a)$ is the radar cross section for a drop of radius a, and $n(a)\, da$ gives the number of drops per unit volume with radius between a and $a + da$. The rigorous procedure for obtaining $\sigma(a)$ has already been outlined in Sec. 6·1. Briefly, one sets up approximate solutions of the wave equation to represent the incident field and the induced field inside and outside the spherical drop. The conditions that must then be met in order to satisfy the boundary conditions serve to fix the coefficients in the expansion of the scattered wave in spherical harmonics. One can then compute the Poynting vector of the wave scattered back in the direction of the radar and obtain the following cross section:

$$\frac{\sigma(a)}{\pi a^2} = \frac{1}{\rho^2}\left|\sum_{n=1}^{\infty}(-1)^n(2n+1)(a_n^s - b_n^s)\right|^2. \tag{38}$$

Here the quantities a_n^s are coefficients of the terms in the expansion of the scattered field arising from the induced *magnetic* dipole, quadrupole, etc., and b_n^s refers to the corresponding *electric* multipole terms. Explicit formulas for these coefficients have been given in Sec. 6·1, Eq. (6·18):

$$\begin{aligned} a_n^s &= -\frac{j_n(\rho)\,[n_c\rho j_n(n_c\rho)]' - j_n(n_c\rho)\,[\rho j_n(\rho)]'}{h_n^{(2)}(\rho)\,[n_c\rho j_n(n_c\rho)]' - j_n(n_c\rho)\,[\rho h_n^{(2)}(\rho)]'}; \\ b_n^s &= -\frac{n_c^2 j_n(n_c\rho)\,[\rho j_n(\rho)]' - j_n(\rho)\,[n_c\rho j_n(n_c\rho)]'}{n_c^2 j_n(n_c\rho)\,[\rho h_n^{(2)}(\rho)]' - h_n^{(2)}(\rho)\,[n_c\rho j_n(n_c\rho)]'}. \end{aligned} \tag{39}$$

The quantity ρ is $2\pi a/\lambda$; n_c is the complex index of refraction of the water; j_n and $h_n^{(2)}$ are the spherical Bessel function and the spherical Hankel

function of the second kind, respectively. The primes indicate differentiation with respect to the argument.

The complex index of refraction of water required in Eq. (39) has been measured accurately over a wide range of temperatures and in the wavelength region from 1.25 to 10 cm. A summary of these measurements, based on a critical survey by Saxton, is given in detail in Sec. 8·6. The results agree sufficiently well with the Debye theory of anomalous dispersion of polar liquids[1] that the theory may be used with confidence to extrapolate to shorter wavelengths.

The complicated expressions [Eqs. (38) and (39)] can be simplified considerably if the approximation is made that ρ and $|n_c\rho|$ are very much less than unity, that is, that we are dealing with drops small compared with the wavelength. The Bessel and Hankel functions in Eq. (39) can then be expanded in ascending powers of their arguments. It turns out that for finite conductivity, the first term in the expansion of the coefficient b_n^s, corresponding to the electric $2n$-pole, is of the order of ρ^{2n+1} whereas the first term for the magnetic $2n$-pole is ρ^{2n+3}. Hence, if in expanding the coefficients a_n^s and b_n^s we neglect all terms in ρ^6 or higher, then the only nonvanishing terms in Eq. (38) correspond to electric dipole, quadrupole, and magnetic dipole oscillations. The expansions of these coefficients through terms in ρ^5 are, from page 451,

$$a_1^s = \frac{-i}{45}(\varepsilon_c - 1)\rho^5, \tag{40}$$

$$b_1^s = -\frac{2i}{3}\left(\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\right)\rho^3\left[1 + \frac{3}{5}\left(\frac{\varepsilon_c - 2}{\varepsilon_c + 2}\right)\rho^2\right], \tag{41}$$

$$b_2^s = \frac{-i}{15}\left(\frac{\varepsilon_c - 1}{2\varepsilon_c + 3}\right)\rho^5. \tag{42}$$

These formulas suffice to expand σ correctly up to terms in ρ^9, and such expansions have been given by Ryde[2] and L. Goldstein[3] in terms of the real and imaginary parts of the index of refraction or dielectric constant. The results will not be given here as they are very complicated, and may

[1] *Cf.* P. Debye, *Polar Molecules*, Chemical Catalog Co., New York, 1929, Sec. 8·1 and Chap. V.

[2] J. W. Ryde, "Echo Intensities and Attenuation Due to Clouds, Rain, Hail, Sand and Dust Storms at Centimetre Wavelengths," GEC Report No. 7831, October 1941. There are a number of algebraic mistakes in Appendix 1 of this paper which are corrected in GEC Report No. 8516, Aug. 3, 1944, by J. W. Ryde and D. Ryde.

[3] L. Goldstein, "Absorption and Scattering of Microwaves by the Atmosphere," Columbia University Division of War Research Propagation Group, Report No. WPG-11, New York, May 1945; also, a chapter in *Radio Wave Propagation*, Academic Press, New York, 1949.

be obtained from Eqs. (38) and (40) to (42) by straightforward, albeit tedious, algebra.

A further approximation of the coefficients a_n^s and b_n^s yields a simple result having an important physical interpretation. Suppose the drops are so small that only the first nonvanishing power in ρ can be neglected. In this approximation only the electric dipole term is retained:

$$b_1^s = -\frac{2i}{3}\left(\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\right)\rho^3, \tag{43}$$

and the cross section becomes, from page 452,

$$\sigma(a) = 4\pi a^2 \left(\frac{2\pi a}{\lambda}\right)^4 \left|\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\right|^2. \tag{44}$$

This is the famous Rayleigh scattering formula, used previously in Sec. 7·3, and is the dominant term in the scattering from raindrops throughout most of the microwave region. It is completely adequate for computing the echo from very small drops such as those encountered in clouds and fog.

Rigorously speaking, the cross section, as given by the Rayleigh approximation, depends upon wavelength not only because of the $1/\lambda^4$ factor but also because of the wavelength dependence of ε_c. Practically, however, this latter effect is very small. Thus at 18°C, ε_c varies from $78.5 - i\ 12.3$ for $\lambda = 10$ cm to $34.2 - i\ 35.9$ at $\lambda = 1.24$ cm. The corresponding variation $\left|\frac{\varepsilon_c - 1}{\varepsilon_c + 2}\right|^2$ is from 0.9286 to 0.9206! Clearly as long as the magnitude of ε_c is large we may safely replace the quantity in absolute signs in Eq. (44) by unity.

The procedure for obtaining the rigorous cross section is straightforward, though it rapidly becomes involved. The first step is to put the coefficients a_n^s and b_n^s in a slightly different form by introducing variations of the spherical Bessel functions defined by

$$\begin{aligned} S_n(z) &= zj_n(z), \\ C_n(z) &= zn_n(z), \\ E_n(z) &= zh_n^{(2)}(z). \end{aligned}$$

The coefficients [Eq. (39)] can then be written

$$a_n^s = -\frac{\dfrac{S_n'(\rho)}{E_n'(\rho)}\dfrac{\dfrac{S_n(\rho)}{S_n'(\rho)} - \dfrac{1}{n_c}\dfrac{S_n(n_c\rho)}{S_n'(n_c\rho)}}{\dfrac{E_n(\rho)}{E_n'(\rho)} - \dfrac{1}{n_c}\dfrac{S_n(n_c\rho)}{S_n'(n_c\rho)}},$$

$$b_n^s = -\frac{\dfrac{S_n'(\rho)}{E_n'(\rho)}\dfrac{\dfrac{S_n(\rho)}{S_n'(\rho)} - n_c\dfrac{S_n(n_c\rho)}{S_n'(n_c\rho)}}{\dfrac{E_n(\rho)}{E_n'(\rho)} - n_c\dfrac{S_n(n_c\rho)}{S_n'(n_c\rho)}},$$

where the prime denotes differentiation with respect to the argument. The advantage of this formulation is that the arguments of the functions are complex only in the combination

$$Q = \frac{S_n(n_c\rho)}{S_n'(n_c\rho)},$$

and all the other functions may be obtained from well-known tables. Unfortunately tables of the spherical Bessel functions of complex argument involved are not available, and in effect one must construct them for the region of interest. The remainder of the calculation then consists in separating the real and imaginary parts of the coefficients, a process without complications though extremely tedious.

Calculations of this nature have been carried out both in England and in this country. In England J. W. Ryde and his collaborators have made extensive computations, but unfortunately only a preliminary report is available at present.[1] The coefficients a_n^s and b_n^s have also been calculated by the Applied Mathematics Group at the National Bureau of Standards[2] using Saxton's data for index of refraction. From these results F. T. Haddock has computed the radar cross section as a function of drop size and wavelength from 3.0 cm to 0.3 cm. Some of the values for the cross section obtained in this manner are reproduced in Table 7·1 by the kind permission of Mr. Haddock. The cross sections are presented in the form of the ratio $\sigma/\pi a^2$ for the three wavelengths 3.0 cm, 1.25 cm, and 0.3 cm, and for a range of the parameter ρ from 0.15 to 3.0.

An examination of these results for values of ρ equal to 1.0 and smaller reveals an unexpected phenomenon. With the example of the infinitely conducting sphere in mind (cf. Fig. 6·1), it might be thought that the first deviations from the Rayleigh law would be to lower the values of the cross section. Indeed, the three-term expansion represented by Eqs. (40–42) predicts that at 3.0 cm the cross section should be 0.8 db below the Rayleigh value at $\rho = 0.20$, and 6.2 db smaller at $\rho = 0.52$. The rigorously calculated cross sections given in Table 7·1 show that on the contrary the deviations are such as to *increase* the value of σ above that given by the Rayleigh law for $\rho < 0.90$ at the wavelengths of 3.0 cm and 1.25 cm. This behavior is strikingly illustrated in Fig. 7·1, where the ratio of the correct cross section to that predicted by the Rayleigh law is plotted against ρ for $\lambda = 3.0$ and 1.25 cm. It is seen that the cross section at 3 cm is at times more than twice as large as that predicted by the Rayleigh law! The physics underlying this phenomenon is not very

[1] *Jour. IEE*, IIA, **93**, 101 (1946), and also a chapter in *Meteorological Factors in Radio-Wave Propagation*, The Physical Society, London, 1946.

[2] Cf. *Tables of Scattering Functions for Spherical Particles*, Applied Mathematics Series, No. 4, National Bureau of Standards, February, 1949.

TABLE 7·1.—CROSS SECTIONS FOR RAINDROPS OF VARIOUS SIZES AND WAVELENGTHS

ρ	$\sigma/\pi a^2$			$\sigma/\pi a^2$	
	λ = 3.0 cm	λ = 1.25 cm	λ = 0.3 cm	ρ	λ = 0.3 cm
0.150	1.6×10^{-3}	1.8×10^{-3}	1.6×10^{-3}	1.30	0.9830
0.200	4.8×10^{-3}	5.7×10^{-3}	5.3×10^{-3}	1.4	0.5820
0.250	0.0114	0.0137	0.013	1.5	0.2669
0.300	0.0251	0.0298	0.0275	1.6	0.09360
0.350	0.0801	0.0620	0.0509	1.7	0.08659
0.400	0.205	0.126	0.0874	1.8	0.2227
0.450	0.357	0.237	0.143	1.9	0.4426
0.500	0.512	0.395	0.223	2.0	0.6705
0.550	0.6635	0.6017	0.3364	2.1	0.8444
0.600	0.8417	0.8458	0.4902	2.2	0.9219
0.650	1.077	1.123	0.6911	2.3	0.8876
0.700	1.383	1.417	0.9348	2.4	0.7575
0.750	1.748	1.716	1.208	2.5	0.5713
0.800	2.103	1.994	1.523	2.6	0.3803
0.850	2.367	2.229	1.710	2.7	0.2388
0.900	2.507	2.394	1.875	2.8	0.1792
0.950	2.545	2.482	1.957	2.9	0.2082
1.00	2.524	2.486	1.956	3.0	0.3098
1.10			1.755		
1.20			1.398		

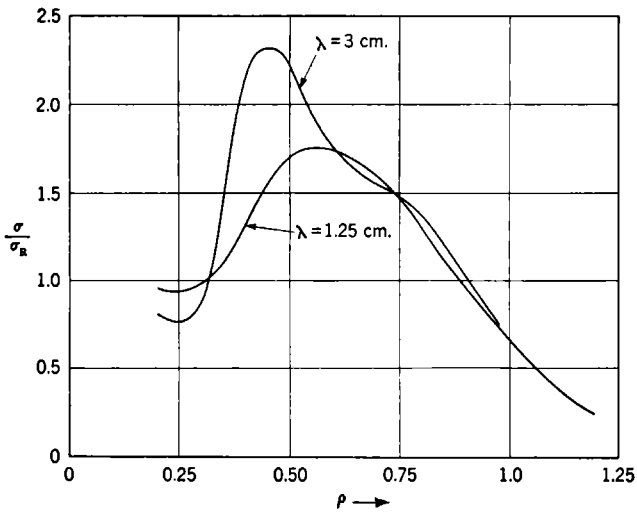


FIG. 7·1.—The ratio of the back scattering cross section of a raindrop to the cross section predicted by Eq. (44), with $\rho = 2\pi a/\lambda$ and with the expression involving ϵ_c equal to 0.93. (Adapted from calculations by F. T. Haddock.)

clear.[1] Presumably the values of the index of refraction are such that the series expansion converges very slowly even for values of ρ around 0.5 or less, and the behavior is the result of contributions from many orders of multipoles.

A drop diameter of 0.5 cm, the largest usually encountered, corresponds to $\rho = 0.52$ at a wavelength of 3.0 cm. Since a deviation of a few db is not greatly significant in the present stage of the radar art, the data of Fig. 7·1 show that the Rayleigh law is a good approximation for all raindrops at $\lambda = 3.0$ cm. It can certainly be used with safety at larger wavelengths. Even for a wavelength of 1.25 cm the Rayleigh law remains valid for all but the largest drops, and should be sufficient for order-of-magnitude calculations.

Eventually, as ρ increases, the cross sections do drop below the Rayleigh prediction, exhibiting oscillations characteristic of the various resonances of the spherical drop, as would be expected. The qualitative picture is therefore closely similar to that for the infinitely conducting sphere, as depicted in Fig. 6·1. One would not expect complete quantitative agreement, of course. For one thing, the slopes of the Rayleigh law portions are different because of the presence of the added magnetic dipole term when the conductivity is infinite. Also, the resonance maxima and minima for the waterdrop tend to oscillate about lower average values as compared to the metal sphere. Nevertheless, the agreement between the two sets of curves is remarkably good. The positions of the maxima and minima agree almost exactly, as might be expected since these positions are dictated mainly by the geometrical conditions for resonance. If the wavelengths and drop sizes are such that the Rayleigh law is no longer valid, and if errors of a few db may be tolerated, then the results for the infinitely conducting sphere provide a satisfactory "universal curve" on which to base calculations for the return from raindrops. Such a curve, joined on properly to the Rayleigh law for large λ and small ρ, should be quite adequate for the task of predicting η measured experimentally for drop-size distributions.

It should be pointed out at this time that there are other factors affecting the echo magnitude, besides drop cross section and size distributions, which may introduce appreciable uncertainties. Thus, at wavelengths around 1.25 cm the echo will be attenuated appreciably by water vapor and by rain areas in the atmospheric region between the the radar and the target. These, of course, affect the echo from all targets at this wavelength. More peculiar to precipitation echoes is an effect that may be termed "self-absorption" (in analogy to the corresponding spectroscopic phenomenon)—attenuation by the target itself.

[1] There is little question about the accuracy of the calculations; Ryde's graphs, obtained quite independently, appear to present the same phenomenon.

Following Ryde,[1] we may consider this effect by first treating a somewhat indealized example. Let the radar set transmit a square pulse starting at the time $t = 0$ and ending at the time $t = \tau$. The radar pulse length in space is then $b = \tau c/2$. Suppose the target be a uniform precipitation area starting at a range R_0. The radar echo measured at a range R, that is, at a time $2R/c$, will include contributions from a strip $R - b$ to R in range. For the case of zero attenuation the received power at time $2R/c$ is proportional to

$$P_r \propto \int_{R-b}^{R} \eta(x)\, dx,$$

where

$$\begin{aligned} \eta(x) &= 0, && x < R_0 \\ &= \eta, \text{ a constant} && x > R_0. \end{aligned}$$

Hence the variation with range will be

$$\begin{aligned} P_r &\propto \eta(R - R_0), && R < R_0 + b, \\ P_r &\propto \eta b, && R > R_0 + b. \end{aligned}$$

The radar echo from a sharp-edged cloud (neglecting coherent radiation) thus would appear to have a sloping edge on the radar indicator, for a distance of a pulse length. If now we introduce a uniform attenuation, the received power would be proportional to

$$P_r \propto \int_{R-b}^{R} e^{-2\alpha(x-R_0)}\, \eta(x)\, dx.$$

With the same uniform drop distribution in the cloud the power received by the radar varies in range as

$$P_r \propto \eta e^{-\alpha(R-R_0)} \frac{\sinh [\alpha(R - R_0)]}{\alpha}, \qquad R < R_0 + b,$$

$$P_r \propto \eta e^{-2\alpha\left(R-R_0-\frac{b}{2}\right)} \frac{\sinh \alpha b}{\alpha}, \qquad R > R_0 + b.$$

The echo will still appear to have a sloping front edge, rising to a maximum of $(\eta b\, e^{-\alpha b} \sinh \alpha b)/\alpha b$ and then decreasing uniformly with a logarithmic attenuation of 2α. The point to be observed is that the echo is attenuated in even the first pulse length, so that choosing only the closest parts of a precipitation echo may not be sufficient to eliminate the effects of attenuation.

Consideration of some reasonable figures, however, shows that the effect is not likely to be troublesome. The maximum reduction within the first pulse length is by a factor $(e^{-\alpha b} \sinh \alpha b)/\alpha b$. At $\lambda = 3$ cm, the longest wavelength at which attenuation is likely to be significant,

[1] *Loc. cit.*

measured values of α range up to 2.5 db/km for a rainfall of 4 in./hr (heaviest tropical shower). With a pulse length of 1 μsec the reduction in the first pulse length is 0.4 db, which is negligible. In practice the attenuation even through a mile of rainfall characterized as heavy (1 in./hr) is not noticeable.

With shorter wavelengths the self-absorption may be much more pronounced. Thus at $\lambda = 1.25$ cm, a 1 μsec pulse, and a value of α of 19 db/km, which has been reported for rainfall of 4 in./hr, the reduction in the first pulse is 2.6 db and the echo drops rapidly with increasing penetration into the rain. Although the effect will be proportionately smaller when the rains are not as heavy as in this example, the possibility of attenuation self-absorption should be kept in mind at such short wavelengths.

It is fortunate that the self-absorption is rarely important for it would be difficult to know just what the attenuation is. Theoretically there is no simple relation between the attenuation α and the cross section per unit volume η. The attenuation is to a large extent an absorption effect and is not very sensitive to the drop-size distribution. The back-scattering, on the other hand, does depend in a drastic fashion on slight accidental features of the drop-size distribution. Nor does there seem to be a feasible method for simultaneously measuring the attenuation along with the radar echo. Conceivably it could be done by investigating the pulse-length dependence of the echo [which should vary with $(\sinh \alpha b)/\alpha$]. Closer examination shows, however, that the effect is not sufficiently striking to provide accurate results even if special measuring techniques were used.

7·7. Drop-size Distribution.—To obtain the cross section per unit volume from a given region of precipitation, one must know, among other factors, the distribution of the drops in size. As the dominant term in the cross section is proportional to a^6, the very large drops, even though few in number, account for a large proportion of the scattering. Slight changes in the high end of the distribution will therefore cause a disproportionate change in the echo power, though hardly affecting the median drop diameter or total rainfall. It is therefore necessary to know the drop-size distribution to great accuracy to obtain even correct order of magnitude estimates of the scattering. It is believed that the distributions so far obtained are neither accurate nor meaningful enough to satisfy these conditions.

Three methods have been used in the past to obtain drop sizes: The photographic, the filter paper or blotter, and the flour method. The first, involving photography of the drops as they fall or after they are collected, has been frequently used for fog and cloud droplets, very rarely for raindrops. In the second method an absorbent material such as filter paper or large blotter is dusted with a dry but water-

soluble organic dye and is then exposed for a measured interval in the rain. At each point where a drop has struck the paper, a colored circular area appears whose diameter is a unique function of the original drop diameter (as long as the drops are not so large that "splashing" is serious). Calibration is obtained by using drops of known size (usually measured by weighing a definite number of drops obtained from capillary tubes with known pressures). The third method, devised by Laws and Parsons,[1] is in reality a variant of the second in which the filter paper is replaced by a pan of dry sifted flour. A drop striking the flour forms a small ball of dough, which may be hardened by drying and then examined at leisure. Again, the relation between ball and drop diameter is found by separate calibration.

These three methods suffer from a number of common disadvantages:

1. Calibration errors, particularly in the filter paper and flour methods are quite appreciable. Thus one investigator using the filter paper method estimated calibration accuracy of around ± 10 per cent.[2]
2. In all the methods the diameter of each drop must be measured individually and recorded by the observer. As a result the labor involved in measuring a large sample is prohibitive. Hence the measurements so far reported have been made with samples far too small to provide sufficient statistical accuracy. For the same reason the number of samples measured has been small so that no information is yet available as to the reproducibility of the distributions or on possible correlations with more conveniently observable meteorological parameters.
3. Perhaps the most important disadvantage is that the measurements do not supply the number of drops per unit volume of the target region with diameters in a given interval but rather the number of drops in that interval striking the ground per unit time. From a knowledge of the terminal velocities of the drops one can indeed convert this information into drop-size distributions in the volume immediately above the ground. The relation between drop diameter and terminal velocity must be obtained empirically, and the several measurements that have been made do not agree very closely.[3] Fortunately all the results agree in that the velocity is practically uniform for the large drops that are of interest here, so that the uncertainties in the actual terminal velocity are not

[1] J. O. Laws and O. A. Parsons, *Trans. Am. Geophys. Union*, p. 452 (1943).

[2] G. T. Rado, "Measurement of the Attenuation of K-band Waves by Rain," RL Report No. 603, March 7, 1945. This report contains a detailed discussion of the precautions necessary to minimize systematic errors.

[3] For the figures given by A. C. Best see Table 8·10. Other values will be found in Laws, *Trans. Am. Geophys. Union*, p. 709 (1941), where references to earlier work are given.

very important. (This does not preclude the possibility that strong up or down drafts may cause great changes in the apparent terminal velocity. There is considerable evidence that strong vertical drafts are not common.)

Even if one can obtain the correct distribution just above the ground, it is doubtful that such a distribution reflects faithfully the situation in the target region, which is usually far above the ground. It is obvious that processes of evaporation and condensation may cause considerable changes in the diameter of a given drop in its long journey to the ground. Indeed radar observations with sharp beams often reveal considerable vertical structure in the precipitation regions (see Sec. 7·14). To be meaningful, drop-size distributions must be obtained at the location from which the echo arises and at the time when the radar observations are made.

It therefore does not seem worth while to reproduce here any of the detailed distributions given in the literature. A general idea of the shape of the distributions can be obtained from the results of Laws and Parsons[1] shown in Table 8.12 below. It must be emphasized again that caution must be used in basing any general conclusions on so restricted a sampling of rain distributions.

A method that appears to obviate many of the objections raised above is at present under development at the Weather Radar Project being conducted now (1949) by the Meteorology Department of the Massachusetts Institute of Technology for the U.S. Army Signal Corps. Briefly, the raindrops interrupt a narrow beam of light focused on a photoelectric cell, the characteristics of the resulting pulse depending upon the drop diameter. The pulses will be automatically sorted into ten or more channels depending on drop diameter and will be counted electrically. Thus, the number of drops in a given interval will appear as a reading on a counter in the appropriate channel. This device, dubbed "Disdrometer," is mounted on an airplane, with the direction of the light beam normal to the line of flight. As the forward velocity will therefore be much greater than the downward velocity of the drops, uncertainties in the terminal-velocity relation are unimportant. The counters of the various channels are photographed at the same time that measurements of the radar echo intensity are made of the immediate vicinity of the plane. Schemes have been carefully worked out to ensure exact synchronization of the measurements in both time and space. Other meteorological quantities such as total water content will be measured at the same time.

A project of this kind requires elaborate and painstaking preparation. Once smoothly under way, however, it should provide rapidly much

[1] *Loc. cit.* References to earlier measurements will be found in this paper.

meaningful data with which to test quantitatively the theory presented here. In addition, it should be possible to investigate extensively empirical correlation between echo intensities and meteorological quantities that may be conveniently measured. It might thus be possible to devise precedures either for estimating echo intensities from easily available data or for using measured intensities to furnish information about the characteristics of the precipitation involved.[1]

7·8. Echoes from Solid Precipitation.—Liquid drops are not the only form in which condensed water may appear in the atmosphere; it may also appear as ice crystals, hailstones, snow, etc. These solid forms are of great importance meteorologically. At least one theory of the formation of rain requires the prior presence of ice crystals. In one respect the calculation of the echo to be expected from solid precipitation forms is simpler than for raindrops, as the imaginary part of the index of refraction is almost completely negligible for ice (*cf.* Table 8·16). Absorption effects can then be neglected and the dielectric constant treated as real, with the value $\varepsilon_1 = 3.06$ from 0° to −50°C. On the other hand we are no longer dealing with simple spherical shapes; the corresponding boundary value problem for the actual shapes (needles, plates, etc.) becomes complicated and has not yet been developed fully. As in the case of raindrops, we are indebted to J. W. Ryde and his collaborators[2] for almost all of the calculations made thus far of the echoes from solid precipitation forms.

Ice Crystals.—Crystals of ice, such as are to be found in clouds, have very complicated geometries. Fortunately, however, their dimensions are always very small compared with the wavelength so that only the electric dipole excitation is of importance. To a first approximation this is purely a volume phenomenon independent of the shape of the scatterer. In effect, the radar beam cannot "see" the shape of an object much smaller than the wavelength. We can thus take over in this case the Rayleigh formula given previously [(Eq. 7·14)] with slight modifications because of the density of ice being slightly different from unity:

$$\eta = 0.11 \frac{Wm\kappa}{\rho^2\lambda^4}\left(\frac{\varepsilon - 1}{\varepsilon + 1}\right)^2, \tag{45}$$

where W is the mass of ice per cubic meter in grams, m is the average mass of a single crystal in grams, ρ is the density, and κ is a dimensionless factor, of the order of unity, arising from the deviation from spherical shape. The quantity depending upon the dielectric constant can no longer be replaced by unity; because of the decreased dielectric constant com-

[1] *Cf.* H. R. Byers *et al.*, "The Use of Radar in Determining the Amount of Rain Falling over a Small Area," *Trans. Am. Geophys. Union*, **29**, 187 (1948).

[2] J. W. Ryde, "Echo Intensities and Attenuation Due to Clouds, Rain, Hail. Sand and Dust Storms at Centimetre Wavelengths," General Electric Co., Ltd., Report No. 7831, Oct. 13, 1941; also unpublished memorandum.

pared with liquid water it is here much smaller, about 0.165. Consequently the echo from a cloud of ice crystals will always be considerably less than from a cloud of rain droplets of the same size and concentration. As ice crystals are always smaller than usual raindrops, it is to be expected that the echo from ice crystals will be negligible except in unusual circumstances. Substituting in the appropriate values of ϵ and ρ (0.916 for ice), Eq. (45) can be written

$$\eta = 0.022 \frac{Wm}{\lambda^4} \kappa \text{ m}^{-1}. \tag{46}$$

Hailstones.—Deviations from spherical shape are not important for hailstones; on the other hand they frequently have dimensions comparable with the wavelength. It is therefore necessary to use the rigorous series expansion [Eqs. (38) and (39)], evaluating as many terms as needed to obtain the desired accuracy. This has been done by Ryde, who has computed the echo per unit volume from a storm of uniform hailstones. In this case η is equal to $n\sigma$, where n is the number of stones per unit volume and σ the radar cross section of each stone. One can also express η as

$$\eta = \frac{\sigma}{mv} p, \tag{47}$$

where p is the precipitation rate, the mass striking the ground per unit area in unit time, v the terminal velocity and m the mass of each stone. Ryde has given a table, reproduced in Table 7·2, of η/p (with p given as an equivalent liquid precipitation rate in millimeters per hour) for various stone diameters and wavelengths. The values of the terminal velocity used were obtained by extrapolation from raindrop data. It will be noticed that η/p exhibits maxima and minima for $\lambda = 1$ cm, indicating

TABLE 7·2.—CROSS SECTION PER UNIT VOLUME FOR HAILSTORMS*

Stone diameter, cm	$\frac{\eta}{p}$, m^{-1} per mm per hr.		
	$\lambda = 10$ cm	$\lambda = 3$ cm	$\lambda = 1$ cm
0.1	6.75×10^{-13}	67.73×10^{-11}	7.57×10^{-9}
0.2	3.32×10^{-12}	3.31×10^{-10}	2.58×10^{-8}
0.3	7.93	1.01×10^{-9}	4.94
0.4	1.69×10^{-11}	1.69	3.53
0.5	2.89	3.31	0.69
0.75	5.23	7.11	2.23
1.0	1.42×10^{-10}	9.15	6.35
1.5	3.05	1.15×10^{-8}	
2.0	5.85	4.96×10^{-9}	

* After Ryde.

the presence of resonance peaks as in the case of perfectly conducting spheres. In general, Ryde finds that hailstorms may be expected to give considerably less echo than a rainstorm of the same precipitation rate with drop distributions such as given by Laws and Parsons.

Snow Crystals.—The individual crystals of snow are usually hexagonal plates with diameters ranging from 0.5 to 3 mm. For the smaller snow crystals Eq. (45) or (46) would be used, but larger crystals would require a more rigorous formula, complicated here by the platelike shape of the scatterers. Ryde has compared the echo from a snowstorm of uniform crystal size with that from a rainstorm of equivalent precipitation rate, again using Laws' and Parsons' drop-size distributions. He comes to the conclusion that for snow crystals weighing between 1 and 2 mg the echoes from the two storms would be about the same. This is in rough agreement with the observations by J. S. Marshall *et al.*, mentioned in Sec. 7·2.

Snow often occurs in more complex forms; individual snow crystals aggregate to form flakes. Such flakes can attain dimensions comparable to the wavelength, 2 cm in diameter or larger, and their complicated shape makes a rigorous computation extremely difficult. A somewhat similar phenomenon is thought to occur with ice crystals in one of the stages leading to the formation of raindrops; that is, aggregates of many ice crystals are formed. Here the dimensions are still small compared with the wavelength, and two equivalent procedures may be used to approximate the echo. Ryde assumes a spherical shape for the aggregate and computes the echo from Eq. (45) and in substituting the density ρ dielectric constant ϵ takes into account the air spaces between the crystals. Effectively this means using a sphere whose volume is that of the sum of crystals composing the aggregate. Alternatively one can say that the N crystals forming the cluster are so close to each other, compared with the wavelength, that they radiate coherently and the total echo is N^2 that of a single crystal. That this view is equivalent to the one used by Ryde is seen from the fact that for electric dipole excitation the cross section is proportional to the square of the total volume, or N^2 times the volume squared of each crystal.

A cloud of ice crystal aggregates should therefore have an echo several orders of magnitude larger than the same cloud of dispersed ice crystals. If these crystals melted to form raindrops, the echo would further increase, chiefly because of the increase in the dielectric constant. In the process of melting, the clusters still have a terminal velocity characteristic of the ice crystal aggregates. As they fall, however, they attain rapidly the much larger terminal velocity of raindrops. As a consequence the concentration of drops decreases and with it the intensity of the echo. Thus one might expect to find a maximum in the echo intensity roughly at the freezing level where the falling clusters of ice crystals begin to melt. Such "bright bands" at the freezing isotherm have indeed been observed by several groups (see Sec. 7·14, for example).

Finally it should be mentioned that suspended solid matter other than water could conceivably give rise to an echo. Sandstorms, dust clouds, and smoke should also scatter electromagnetic radiation. Here the particle sizes are so small compared with the wavelength that the Rayleigh formula [Eq. (45) or equivalent] may be used with complete confidence. The diameters of sand particles are comparable to those of raindrops, and appreciable echoes may be expected from sandstorms. As pointed out by Ryde, however, the particles in a true sandstorm rarely rise above 6 ft. Dust particles, on the other hand, are usually much finer. According to W. F. Gibbs,[1] to remain suspended in the air for appreciable times, dust particles must be 10 microns or less in diameter. For a given mass concentration the Rayleigh formula predicts that the echo varies as the volume, or cube of the dimensions, of each scatterer. The echo from the fine particles of dust should therefore be extremely small. Thus for the extreme case, quoted by Ryde, of grain elevator dust with a concentration of 1 kg/m^3 and particle diameter 15 microns the value of η at 10 cm is about 10^{-22}m^{-1}: It is just conceivable that the echo from dust storms could be detected at sufficiently short wavelengths (as short as 1 cm). The situation is even less favorable for smoke particles, whose average diameter is about 1 micron, with mass concentrations usually less than 1 g/m^3. Under such conditions, even at a wavelength of 1 cm, η is not bigger than 10^{-12}m^{-1}. Such calculations tend to make one highly skeptical about reports of echoes from "smoke."

GENERAL PROPERTIES OF PRECIPITATION ECHOES

BY DONALD E. KERR AND ARTHUR E. BENT

The preceding sections have discussed evidence in favor of the hypothesis that most meteorological echoes arise from precipitation in either liquid or solid form. We shall now present observational information that not only lends further support to this reasoning but summarizes the general properties of these echoes and their relation to known meteorological facts. Assuming acceptance of the hypothesis, we shall employ the term *precipitation echoes*, which is now in general use by workers in the field.[2]

7·9. Identifying Characteristics.—When precipitation echoes are observed on radar systems, they are likely to be mingled with the other echoes (from land targets, ships, aircraft, etc.) from which they must be distinguished. The ease with which precipitation echoes can be distinguished

[1] *Clouds and Smokes*, London, 1936.

[2] From the historical point of view, it is of interest to point out that when these echoes were first observed in the early days of microwave radar, they were dubbed "cloud echoes." This is an unfortunate choice of terminology, as it conveys the wrong impression to uninitiated radar personnel. It appears to be losing ground as fuller knowledge of the phenomenon is disseminated. There still remain the "angel" echoes, however, which may sometimes be correctly represented by the earlier term.

from others varies widely and depends upon numerous factors within the radar system and without. There are several characteristics of precipitation echoes that aid in varying degrees in identifying them. These characteristics are discussed briefly below and are illustrated by examples in later sections.

Motion.—In general, precipitation echoes show horizontal motion[1] characteristic of the local air currents surrounding the rainstorm causing the echoes. Speeds up to 100 mph have been observed on occasion, and at the other extreme the rain at times appears to be practically stationary for appreciable intervals. On numerous occasions radar echo-velocity measurements have been found to agree with upper-air wind-velocity measurements made by other means and with surface observations of the motion of rainstorms. Valuable meteorological information concerning winds aloft in overcast weather and the time of arrival of precipitation at a given point may often be obtained by purely radar means. The characteristic motion is usually of considerable aid in distinguishing precipitation echo from echoes from fixed land targets.

Altitude.—Precipitation echoes have been received from areas at levels indicated to be from close to the ground up to 40,000 ft. (The maximum altitude has probably been exaggerated by the beamwidth, but at most this error is a few thousand feet.) Many precipitation echoes are seen on systems with the beam directed horizontally, but in most cases echo amplitude is increased by elevating the beam above the horizon. When an echo is received from an area indicated to be above the surface, precipitation is suggested unless the size is such as to involve confusion with aircraft. Experience with airborne radar shows that the elevation test is not of much advantage at ranges beyond 25 to 30 miles. At close range, however, surface targets are lost upon elevation of the antenna beam but precipitation echoes remain. Precipitation echoes have been observed with the beam depressed below the horizon by reflection from near-by water or ice surfaces.

Distance.—Because of the vertical development of the phenomena involved, precipitation echoes may be observed at large ranges. Thunderstorms have been seen at distances up to 250 miles. Echoes thus appear on ground-based systems at distances far beyond normal land signals. (Echoes from distant land targets are at times seen as a result of atmospheric refraction, but the appearance of an echo at great range usually suggests a precipitation echo.)

Size.—The size of these echoes as observed on radar systems varies within wide limits. In general, precipitation echoes are distinguished by the large size of the echo area as compared with usual targets. Whole quadrants of the PPI may be filled with echo in extreme cases, even when using sweeps of 50 miles or more. Where the size of the echo area is great,

[1] Vertical motions are discussed at a later point.

it is easy to distinguish from other types, but cases are observed where it is small and the possibility of confusion is greater. This is particularly true when using the long sweeps that make it impossible to examine the radial extent of the echo in detail.

The shape of precipitation echoes is also quite variable, normally being irregular. The smaller echoes, such as those of high amplitude received from thunderstorms and heavy showers, may be of a circular or oval shape with fairly even boundaries. The weaker echoes, such as those usually received from precipitation associated with stratiform-type clouds, are of irregular shape with indistinct boundaries where the amplitude of the echo merges into the noise level. Shape of precipitation echoes changes with time, sometimes very rapidly.

Average Intensity.—The preceding sections have indicated the echo intensities to be expected from various types of precipitation. In practical cases the intensity may vary from an indefinitely small lower limit to magnitudes sufficiently strong to obliterate strong echoes from near-by land targets.[1]

Rapid Fluctuations of Intensity.—It was shown in Chap. 6 that precipitation echoes fluctuate in intensity in a manner best described by a frequency spectrum. Visual observation of the radar indicator is not suited to obtaining information on frequency spectra, but under proper circumstances it can lead to positive identification of precipitation echoes on the basis of their fluctuations in intensity. The PPI and B-type indicators are not suited to observation of fluctuations, but the expanded A-scope is the ideal instrument for this purpose.[2] The outstanding property of precipitation echoes viewed on an A-scope is the rapid dancing or fluttering of the echo, extending over a range interval of a few hundred yards to miles in some cases. The effect is frequently described by saying that it is "fuzzy," "furry," or lacelike in structure, depending upon the details of the indicator and nature of the storm. If the echo is very weak, it often appears as a small "bump" in the thermal noise normally present with the gain control well advanced.

Figure 7·2 shows a photograph of precipitation echoes on an A-scope with a 10 mile sweep. As this is a time exposure, it represents the superposition of the echoes from many pulses, each echo different from all the others. The detail visible to the eye upon direct observation of the indicator is thus lost, but enough of the general appearance is preserved to suggest the actual fuzzy appearance. This echo is to be contrasted to echoes from

[1] Assuming special anticlutter circuit arrangements are not employed to reduce the effective intensity of the precipitation echoes. Such circuits are outside the scope of this volume.

[2] We assume familiarity with the properties of radar indicators. See Vol. 22 of this series.

most fixed targets, which, with proper adjustment of the gain control, break the baseline of the indicator trace cleanly, giving a series of sharp, discrete spikes that also fluctuate but in general tend to retain their identity to a far greater extent than do the components of precipitation echo.

FIG. 7·2.—Photograph of precipitation on an A-scope with a 10-mile sweep. Most of the detail is lost, but the characteristic "fuzzy" appearance is still partly visible.

If the rain area is small or if the sweep length is many miles, the fine detail of the echo is not visible. *The most satisfactory indicator for identification and visual study of precipitation echoes is the expanded A-scope*, which allows a small range interval of 2000 to 4000 yd at any distance to be spread across the entire indicator. Figure 7·3 shows precipitation echo (not the same storm as in Fig. 7·2) on an expanded 4000-yd sweep. The improvement in detail over Fig. 7·2 is obvious. The lacy structure of the echo can be observed fairly clearly, but the detail is still in no way comparable to that observable visually. This detail, revealed on such an indicator, is the best means of identification of precipitation echoes; in fact, we know of no instance in which examination of an echo of unknown origin in the foregoing manner has failed to make possible a definite decision as to whether or not the echo originated from precipitation.[1] Unfortunately relatively few radar systems are equipped with this type of indicator.

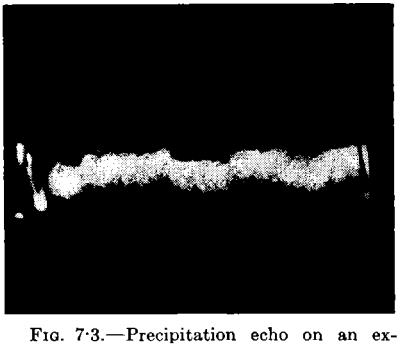


FIG. 7·3.—Precipitation echo on an expanded A-scope 4000 yd in length. The rectangular notch in the center is a "range gate" and has no connection with the present discussion.

Brightness.—The PPI presentation is controlled by limiting circuits that operate to cut off signals of extreme amplitude so that echoes above a certain level appear on the scope as of equal brilliance. Precipitation echoes may easily be above limiting levels. Some, particularly those from light to moderate general rain, have a faint, cloudlike appearance that is very typical. This appearance might also

[1] A possible exception is sea echo, which on an A-scope may be indistinguishable from precipitation echo. In this case auxiliary information from other indicators and from operations such as varying antenna tilt may be necessary to permit positive identification.

be described as foggy or transparent. Brightness is largely a matter of adjustment of receiver gain and limiting circuits and hence must be used with caution for purposes of identification.

Stratification.—Many precipitation areas show a layer structure when viewed from below. This is most typical of situations of general rain. As many as four reflecting layers have been observed at one time. These layers are often of wide horizontal extent. There is much variation in the characteristics of these layers, which are discussed more fully at a later point.

7·10. Confusion and Masking of Other Echoes.—A practical problem often confronting radar operators is that of positive identification of a new echo in a very short time and without recourse to special indicators, system adjustments, or operating procedure. Under some circumstances, general operating experience and familiarity with the characteristics described in the preceding section are sufficient, but there are times when this is not true. For instance, fixed radar systems operating in temperate latitudes show thunderstorms and widespread precipitation areas as large bright masses on PPI or B indicators, and even a casual glance at the indicator reveals the origin of the echo. (Examples are shown in Secs. 7·11 to 7·13.) By contrast, however, airborne radar in the tropics continually encounters small isolated thundershowers that are frequently very deceiving. They may be much less than a mile in diameter, appear to have fairly sharply defined boundaries; and when viewed on an indicator having a sweep length of more than about 15 miles, they are easily mistaken for isolated land masses or ships. Even if an expanded A-scope were available, there would seldom be an opportunity in high-speed aircraft to stop the antenna, align it accurately, and examine each echo in detail.[1] This problem will probably be reduced to a minor annoyance as operator training and experience improve and as circuit developments become available that permit anticlutter action at the flip of a switch.

In Sec. 6·17 it was pointed out that all scattering elements within a volume determined by the radar beamwidth and the pulse length contribute to the echo. At large distances this volume may become very large because of the effect of the finite beamwidth, and consequently precipitation echoes may interfere with echoes from other targets at the same range but in a slightly different direction from the radar. This undesirable effect may be reduced by decreasing beamwidth and pulse length and by employing certain types of anticlutter circuits. These matters will not be discussed here, however.

[1] Some radar systems have expanded PPI or B-scopes that can often be very effective for this purpose and are much easier to use than the A-scope. The rapid fluctuation of the echo cannot be observed, but the hazy appearance and irregular shape are usually fairly satisfactory symptoms.

Figure 7·4 shows an example of the type of masking effect that may be encountered with airborne radar. The precipitation echo, easily identified by its irregular, diffuse appearance, obliterates a large fraction of the landscape echoes normally visible. The degree of seriousness of this masking phenomenon depends very much upon the details of the situation. The

FIG. 7·4.—A PPI photograph from a 1.25-cm airborne radar system operating at an altitude of 3000 ft near a thunderstorm over a city. Full-scale range is 10 miles. The landscape targets, faint because the receiver gain has been reduced to improve contrast, are heavily masked by the thunderstorm echoes. The masking effect would be even more marked with a higher gain control setting.

quantity of most interest in this connection is the excess of echo strength that a desired echo must have in order to be detected or tracked in the midst of precipitation echo. This is a complicated matter, depending upon numerous factors both within and without the radar system. The reader is referred to Chap. 11 in Vol. 24 for a discussion of threshold signals in clutter.

PRECIPITATION ECHO PROPERTIES AND METEOROLOGICAL STRUCTURE

BY DONALD E. KERR AND ARTHUR E. BENT

7·11. Classification of Echo Types.—For purposes of description, meteorological conditions that result in radar echoes may be divided broadly into two groups, those in which precipitation is localized and those in which

precipitation is widespread.[1] The first group includes precipitation in unstable air masses and frontal systems where the air is completely or conditionally unstable, whereas the second group includes precipitation in stable air masses and in frontal systems in which the air participating in the upglide movement is stable.

The first group includes thunderstorms, showers, and what are termed for convenience in classification as *active convective situations*. This last subdivision is necessary to include cases of radar echoes from precipitation in *cumulus congestus* and some other cloud types, where precipitation is present in the cloud but does not reach the ground. (Cases have been described in Sec. 7·2.) Thunderstorms and showers of the tropics are a typical example of precipitation in unstable air masses. The cold-front thunderstorm or shower is a typical illustration of precipitation in frontal systems in which the air is completely or conditionally unstable.

The second group includes widespread rains such as those resulting from the mechanical lifting of stable air over mountain ranges or from the upglide movement in a warm front.

This classification is convenient for purposes of description, but many variations and combinations occur. Type cases are used here for purposes of illustration, but in practice such large variations in degree occur that identification is not always easy. The same problems are encountered in radar observation of precipitation that exist in conventional meteorological observation and forecasting. Often echoes are observed that do not fit well into this simplified classification. This is another indication of the variability of meteorological phenomena and calls for experience and skill in interpretation.

7·12. Thunderstorms.—Figure 7·5 shows a typical isolated convective thunderstorm of a hot summer afternoon in middle latitudes. The small, sharply defined echoes in the first 15 miles are from fixed land targets in the vicinity of Boston. The three bright, hazy echoes are from a well-developed thundershower at about azimuth 200°, a heavy shower at about azimuth 130° that is probably either an early or late stage of a thundershower, and a smaller and much less well-developed shower at azimuth 70°. Examination on the A-scope showed extremely strong echoes from the shower at 200° and weaker echoes from the other two.

The development of the large shower is further revealed by the study of its vertical structure shown in Figs. 7·6 and 7·7. In obtaining Fig. 7·6 an expanded elevation indicator (EEI) was used to scan the shower in the vertical plane at a fixed azimuth of 190°. By means of a 10/1 synchro drive the vertical angle scale is expanded by a factor of 10. Appreciable echo is obtained from elevation angles of $-10°$ to $+32°$. Because of the beam-

[1] A. E. Bent, "Radar Echoes from Atmospheric Phenomena," RL Report No. 173, March 13, 1943; "Radar Detection of Precipitation," *Jour. Meteorol.*, **2**, 78 (1946).

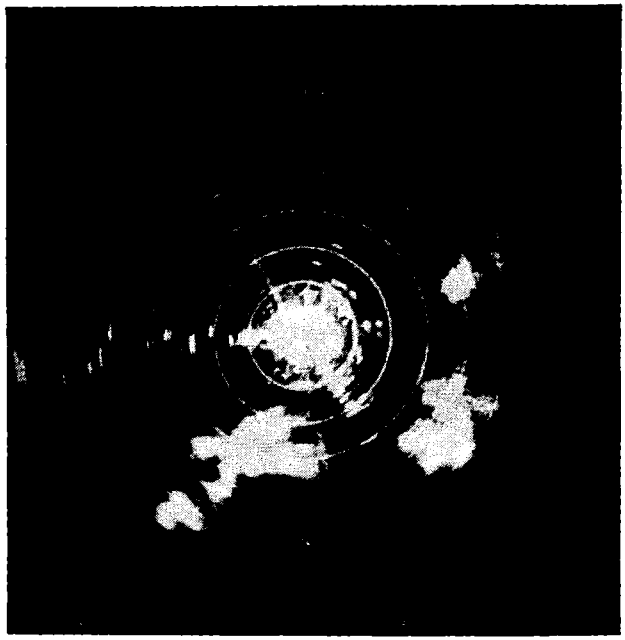


FIG. 7·5.—A PPI photograph of a typical summer-afternoon convective thunderstorm in middle latitudes, as seen on a 10-cm radar system at the Radiation Laboratory on July 14, 1942. Range marks 5 miles, pulse length 1 μsec, beamwidth 6°. Zero azimuth (north) is at the top of the figure.

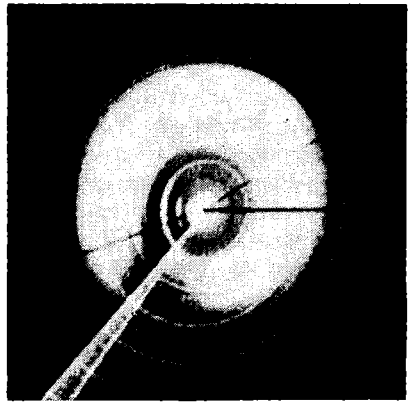


FIG. 7·6.—Photograph of an EEI showing the echo from the thundershower of Fig. 7·5 at an azimuth of 190°. The angular scale indicates ten times the true angle of elevation, with the black horizontal line indicating the true horizontal. Five-mile range marks.

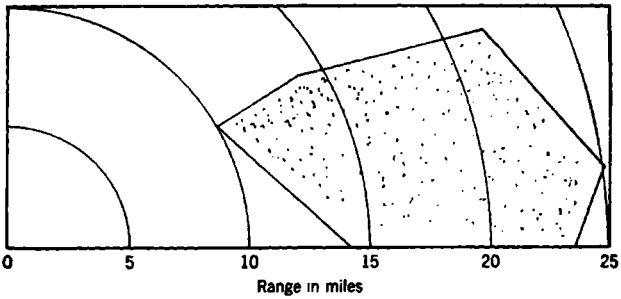


FIG. 7·7.—Rough replot of the EEI presentation of Fig. 7·6 in rectangular coordinates in the vertical plane, showing the vertical development of the thundershower of Fig. 7·5 at an azimuth of 190°. As no correction has been made for the beamwidth, the true echoing region is somewhat smaller than this replot indicates.

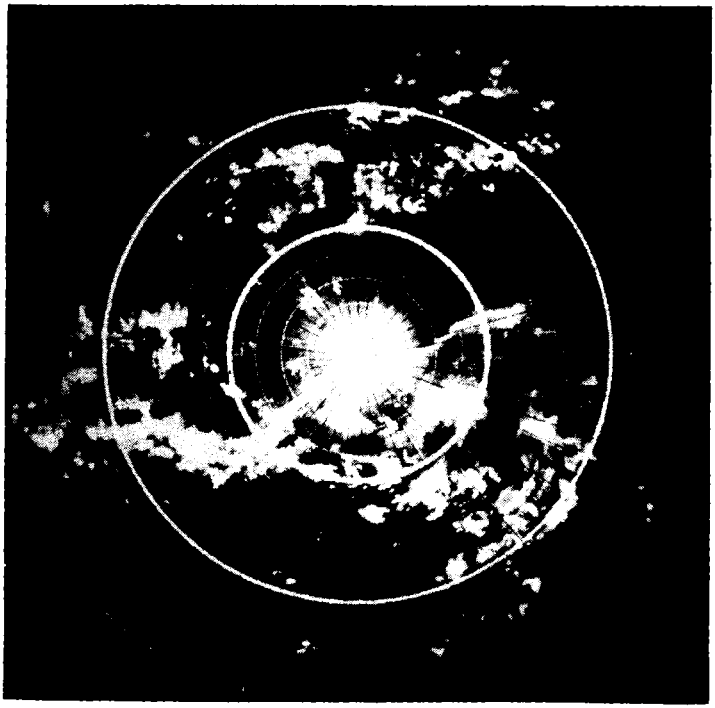


FIG. 7 8.—Photograph of the PPI of a high-power 10-cm radar system in Florida, showing intense summer thunderstorms. In this case a special feedback circuit is operating to suppress all but the most intense echoes from the centers of the columns of precipitation. Full-scale range is 150 miles.

width of about 6° the echo region is actually somewhat smaller than indicated by the EEI. If the echoing region is replotted from the EEI in rectangular coordinates, it appears as shown in Fig. 7·7.[1]

Figure 7·8 shows a PPI photograph of a strong convective thunderstorm in Florida, where such storms are a predominant feature of the summer climate. In this case special circuits have been arranged to give a high

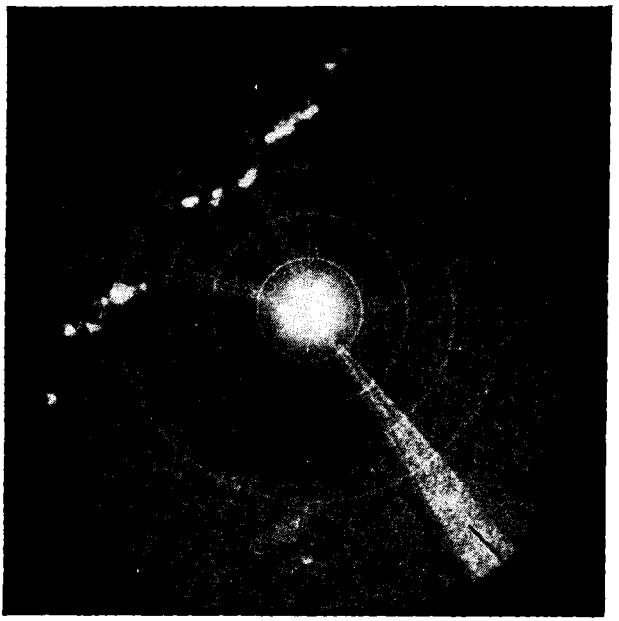


FIG. 7·9.—A line of thunderstorms in a cold front as shown on the PPI of a 3.2-cm radar system at the Radiation Laboratory on July 22, 1943. Beamwidth is 1.5°, pulse length 1 μsec; and the range marks are 20 miles apart.

degree of suppression of the rapidly fluctuating precipitation echo, with the result that only the most intense echoes from the centers of the columns of precipitation are visible.[2] No echoes from fixed land targets are visible in this photograph because the set is operating at a low altitude over flat terrain and the horizon range is only a few miles.[3]

[1] The EEI presentation is not very useful in practice because of the tediousness of translating from polar to rectangular coordinates. The range-height indicator (RHI) is preferable (see Sec. 7·14 for RHI photographs).

[2] Without this circuit in operation the precipitation echoes were much more intense and widespread. Unfortunately we have no comparable photographs showing the echoes without the circuit in operation.

[3] Further study is being made of thunderstorm structure by radar means; for example, see H. R. Byers and R. R. Braham, "Thunderstorm Structure and Circulation," *Jour. Meteorol.* **5**, 71 (1948).

The arrangement of thunderstorms along the line of a cold front results in heavy precipitation and radar echoes having a similar appearance on the indicator. The long straight line of echoes is so characteristic that it is usually readily identified. (Occasionally intense precipitation areas may happen to be arranged in an elongated form that has a superficial similarity to the cold-front thunderstorm echo. Known meteorological conditions

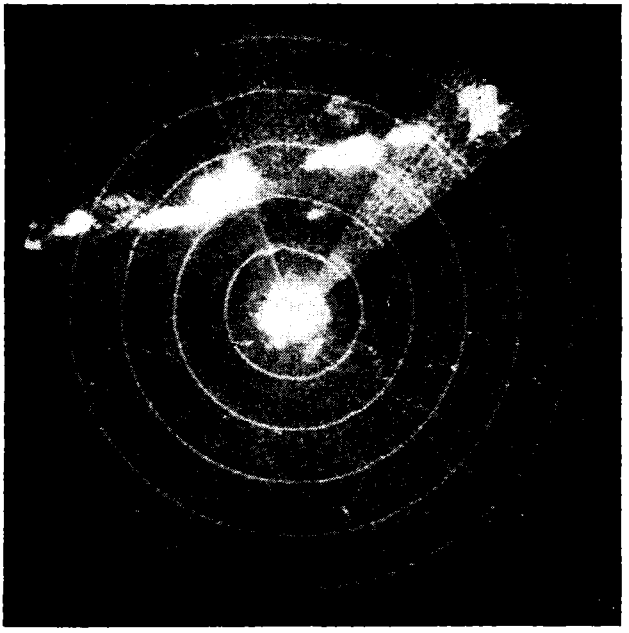


FIG. 7·10.—Cold-front storms on the PPI of a 3.2-cm radar system at the Radiation Laboratory on June 29, 1943. Beamwidth is 1.5°, pulse length 1 μsec; and range marks are 20 miles apart.

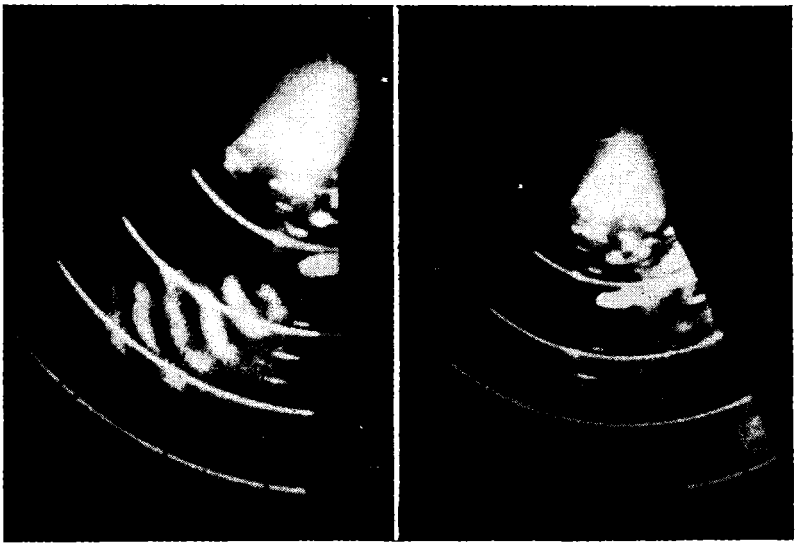
at the time often give a clue as to what might reasonably be expected, however.) As much variation exists in radar cold-front echoes as would be expected (from a meteorological point of view) in the precipitation itself. Many cold fronts, of course, do not have thunderstorms at all.

Figure 7·9 shows the typical line of thunderstorms accompanying an advancing cold front. A series of similar photographs revealed rapid movement of the front for a time, followed by an almost complete halt. While the front was moving, the echoes did not change appreciably in size or intensity, but when the front stopped, the echoes became stronger and covered a larger area.

Figure 7·10 shows another cold front at closer range. In this case it is not known definitely that the echoes shown came from a thunderstorm, although it appears probable. Vertical scanning revealed echoes at 18,000

ft, they were very strong, and light thunder was heard near the echo area. This is a case in which positive differentiation between cold-front showers and thunderstorms could not be made without further information, although there was no doubt about existence of the cold front, as it was confirmed by other observations.

7·13. Other Forms of Localized Precipitation. *Showers in Unstable Air Masses.*—These are the typical showers of hot summer afternoons in tem-



(*a*) (*b*)

FIG. 7·11.—Instability showers on the PPI of a 10-cm radar system at the Radiation Laboratory on Oct. 30, 1943. Range marks 4 nautical miles apart, pulse length 0.8 μsec, and beamwidth 2°. (*a*) Made at 2:05 P.M.; (*b*) made at 2:10 P.M. North is at the top of the picture.

perate latitudes or in the tropics. (There are also many other types of instability showers.) An example was shown in Fig. 7·5, in which two showers were found near a fully developed thunderstorm. In that case the difference between the thunderstorm and the showers appeared to be largely a matter of degree of development (the dividing line is not sharply defined). The appearance on the PPI frequently supplies insufficient information to permit estimation of the intensity of the storm, and examination of the vertical extent by vertical scanning is necessary.

Another example of afternoon showers that is interesting because of its odd configuration and short duration is shown in Fig. 7·11. Meteorological observations from Blue Hill Observatory[1] showed the showers to be highly

[1] The vicinity of the observatory is shown by the intense oval-shaped echo about 9 miles south of the radar.

localized and of brief duration, concluding with a row of towering cumulus clouds to the east of the echo region.

Showers in Frontal Systems with Complete or Conditional Instability.—Cold-front showers are an excellent example of this type of precipitation. The resulting echoes have the linear grouping characteristic of a cold front. As indicated earlier, both cold-front showers and thundershowers are likely to have a similar appearance on the PPI, and vertical scanning is often of considerable aid in deducing the structure of the precipitation source.

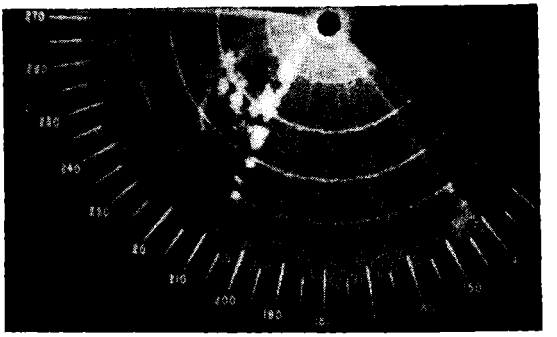


FIG. 7·12.—Cold-front showers on the PPI of a 10-cm radar system at the Radiation Laboratory on Dec. 2, 1942. Range marks are 10 miles apart; beamwidth 2°.

Figure 7·12 shows echoes from precipitation identified as showers forming part of a cold front that appeared from the west, moved across the field of the PPI, and disappeared to the southeast. The heavy shower at the center passed over the laboratory.

7·14. Widespread Precipitation.—By definition this group of echoes includes those from precipitation in stable air masses and frontal systems where the air participating in the upglide movement is stable. A typical example is the lifting of air in the upglide of a warm front. The classification cannot always be strictly maintained, as instability often plays a part in the cases considered here. From the radar point of view the principal distinction is in the widespread nature of the precipitation as contrasted to the extremely localized extent of precipitation in thunderstorms and showers resulting from convection.

The rate of rainfall of this type may vary through wide limits, and the amplitude of the radar echo shows corresponding variation. In general, however, the amplitude of the echoes is at a lower level than is found with convective showers and thunderstorms, resulting in a faint, foggy appearance on intensity-modulated indicators.

The area from which the echoes is received is variable but is characteristically large in contrast to that of the echoes already considered. More or less complete coverage of the PPI on a 50-mile sweep is not uncommon. Echo areas may be continuous or broken up into irregular shapes. Bound-

aries of areas are irregular and taper gradually in intensity to the noise level. Figure 7·13 shows typical echo from this type of precipitation, which was associated with the warm sector of a low-pressure system.

In some cases the rate of rainfall may be relatively steady over a large area; but if convective instability occurs, this relative uniformity tends to be interrupted by areas of alternately heavy and light rainfall, resulting in roughly corresponding variations in echo intensity. The nearly uniform

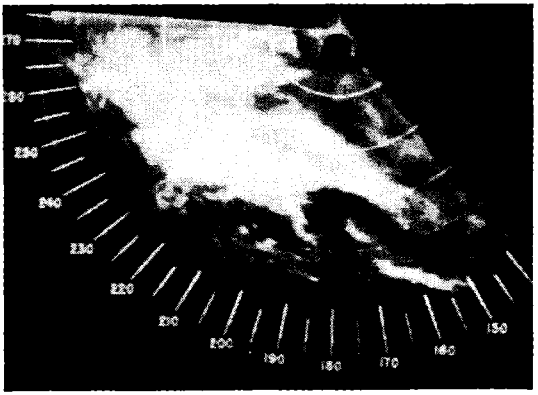


FIG. 7·13.—Echo from widespread rain associated with the warm sector of a low-pressure system on Dec. 1, 1942, as viewed on a 10-cm radar system at the Radiation Laboratory. Horizontal beamwidth about 0.8°, pulse length 1 μsec, range markers spaced 10 miles. This photograph illustrates very well the hazy, nearly "transparent" appearance of precipitation echoes.

echo over a wide area in Fig. 7·13 suggests lack of appreciable convective instability.

Layer-type Echoes.[1]—One of the most interesting properties of the echoes observed under conditions of widespread precipitation is pronounced stratification. Sometimes well-defined, nearly continuous horizontal layers are indicated for a considerable distance, whereas at other times multiple layers or "clumps" of limited extent are observed, often accompanied by considerable motion or buildup and decay of the echoes.

If a radar with a PPI is pointed directly upward at an echoing layer and the antenna and indicator yoke continue to rotate the echo appears as a circular ring. Figure 7·14 shows an example of a layer-type echo observed in this manner with warm tropical air overrunning cold surface air. The boundary between the air masses was estimated from Weather Bureau information to be at about 10,000 ft, and the freezing isotherm at about 8,000 ft. If the antenna is tilted downward, the diameter of the ring increases in the expected manner with angle, confirming the existence of a

[1] A. E. Bent, "Radar Echoes from Precipitation Layers," RL Report No. 689, Aug. 20, 1945.

layer.[1] As many as four layers have been observed simultaneously for limited periods at heights between ½ and 6 miles.

The explanation of the origin and behavior of these layers in terms of meteorological structure has not been found. Some property—large drop-size or high drop concentration or both—causes η (the radar cross section per unit volume) of the precipitation in a limited region to exceed greatly that of surrounding regions. The observation of vertical motion coupled with the possibility of strong updrafts has led to suggestions that updrafts hold large drops suspended. Another plausible suggestion was outlined in Sec. 7·8. It appears that obtaining the correct explanation for these echoes is likely to bring to light much new knowledge of the physical processes of precipitation.

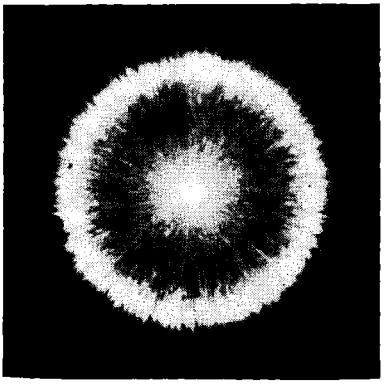


FIG. 7·14.—Photograph of the PPI showing the typical ringlike structure obtained when the antenna is pointed directly upward at a horizontal layer. Pulse length is 1 μsec, and wavelength 3.2 cm. This layer is about 2000 ft thick and 2 miles high at the top.

The best indicator thus far available for observing the vertical structure of precipitation echoes is the range-height indicator (RHI), which presents a cross section in the vertical plane at a fixed azimuth, with independently adjustable magnification on the horizontal and vertical range scales. The RHI is particularly useful when the stratification is not limited to a single well-developed layer, as in Fig. 7·14, but occurs in broken patches or in curved or warped surfaces. Figure 7·15 shows an RHI photograph of later developments of the situation of Fig. 7·14. The three photographs are part of a series showing extremely rapid changes in density and shape in periods as short as 1 min. The antenna was pointed at a fixed azimuth slightly south of west. The surface wind while the pictures were being made was northeast, about 20 mph, and the winds at 10,000 ft were estimated by the Weather Bureau to be from the southwest. The local rate of rainfall was about 0.04 in./hr. The radial velocity of parts of the echoing areas of the bottom photograph was about 60 mph. The amount of information available in a detailed study of the kind suggested by these photographs is enormous, and it presents both a challenge and a valuable research tool to the meteorologist which, it is to be hoped, will be explored fully in the near future.[2]

[1] Antenna side lobes must be reduced to a very low level if observations of layer echoes are to be free from confusion.

[2] For a summary of later work, consult the series of reports now being issued by the Department of Meteorology, Massachusetts Institute of Technology (Army Contract W-36-039-SC-32038).

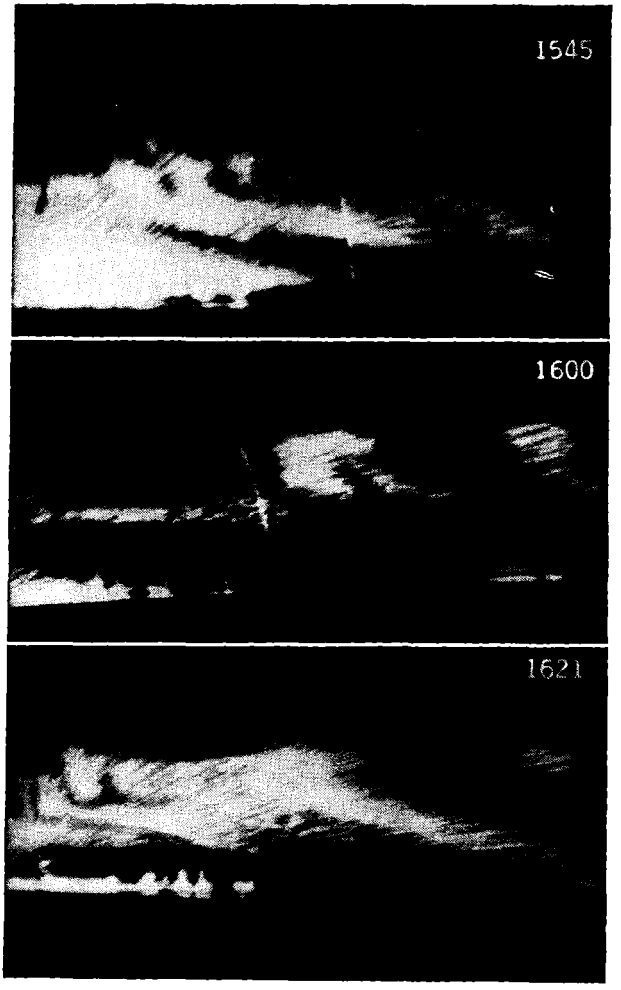


FIG. 7·15.—RHI photograph of the meteorological structure about 2 hr after that of Fig. 7·14, using the same radar system. The time of making each photograph is indicated on the right. A 5-mile range mark shows faintly.

7·15. Cyclonic Storms of Tropical Origin.—Tropical storms, with their associated heavy precipitation in characteristic patterns, make ideal subjects for radar study.[1] They are known as hurricanes or typhoons, depending primarily upon their origin. We present here a limited sample of the currently available information on these storms.

[1] The term "characteristic" is perhaps anticipatory, as not enough different storms of this type have been studied thus far to predict what patterns are characteristic, or typical.

Hurricanes.—The hurricane of Sept. 14–15, 1944 passed close to Cambridge, Mass., where it was observed on a number of radar systems. Rain began in the Boston area some time before the approach of the severe part of the storm, resulting in the usual echoes associated with widespread precipitation. The center of the storm passed near Boston at about midnight on Sept. 14.

Figure 7·16 shows the general heavy rain associated with the forward gradient of the storm about 3 hr before the center arrived. During the

FIG. 7·16.—Echoes from the hurricane of Sept. 14–15, 1944, observed on a high-power 10-cm radar system at the Radiation Laboratory, at 2118 hours, EST. The "bands" of precipitation were characteristic of the echoes before arrival of the center of the storm.

hour centered around the time when this picture was taken, about half an inch of rain fell. Successive pictures made later clearly showed the approach of the central part of the storm until midnight, when no further rain was seen to the south.

The most curious feature of these echoes was the "bands" of precipitation that advanced northward in a nearly straight line toward Boston prior to the arrival of the center of the storm. They were correlated with the arrival of bursts of excessively heavy rainfall at the radar location. Shortly after Fig. 7·16 was made, the storm was scanned in the vertical plane on an EEI, which revealed a thin layer (as in Fig. 7·14) at a height of about 9000 ft. (The duration of this layer is not known.)

The early stages of the hurricane were also observed by the 3-cm radar at Provincetown, Mass. (see map in Fig. 4·1). Figure 7·17 shows a PPI photograph of the intense localized precipitation preceding the storm. The individual echoing areas appear to fall in the classification of Sec. 7·13, but collectively they form part of a system of clearly recognizable bands of precipitation.

To summarize the radar information secured in this one case: The maximum range on the general precipitation associated with the forward quadrant of this storm was of the order of 60 to 80 miles; the approach of the clear central eye of the storm was tracked in from about 25 miles; there were concentrated bands of precipitation extending from east to west, moving north; in the general rain ahead of the central part of the storm there were

FIG. 7·17.—Localized precipitation preceding the hurricane of Sept. 14–15, 1944, as seen on the PPI of a 3-cm radar system at Provincetown, Mass., at 1015 hours, EST, Sept. 14. The small echo areas are part of larger "band" structures characteristic of this hurricane. Beamwidth 2°, pulse length 1 μsec, 10-mile range marks. The white line points north.

many small scattered areas of intense precipitation; after the eye of the storm had approached sufficiently close, no rain could be observed in the southerly quadrants; and in the early part of the storm there was a well-defined echoing layer at about 9000 ft at one time.

An excellent detailed radar study of this hurricane was made by Army Air Forces units as it passed through Florida and New Jersey. High-powered radar systems were employed in a number of cases, and their much greater coverage revealed that the "bands" described above were part of a large circulation system that rotated about the center of the storm as it traveled northward.[1] By the time the storm reached Boston, much of the

[1] Unfortunately we do not have the details available. The complete radar record is worth serious study by anyone interested in hurricanes.

"band" structure had been destroyed, but it is still sufficiently recognizable in Fig. 7·16 that there can be little doubt that it is the same phenomenon that was prominent farther south.

Typhoons.—On Dec. 18, 1944, a typhoon was observed by radar east of the Philippine Islands by a U. S. Naval vessel.[1] A PPI photograph of a high-powered 10-cm system is shown in Fig. 7·18. The approach at long

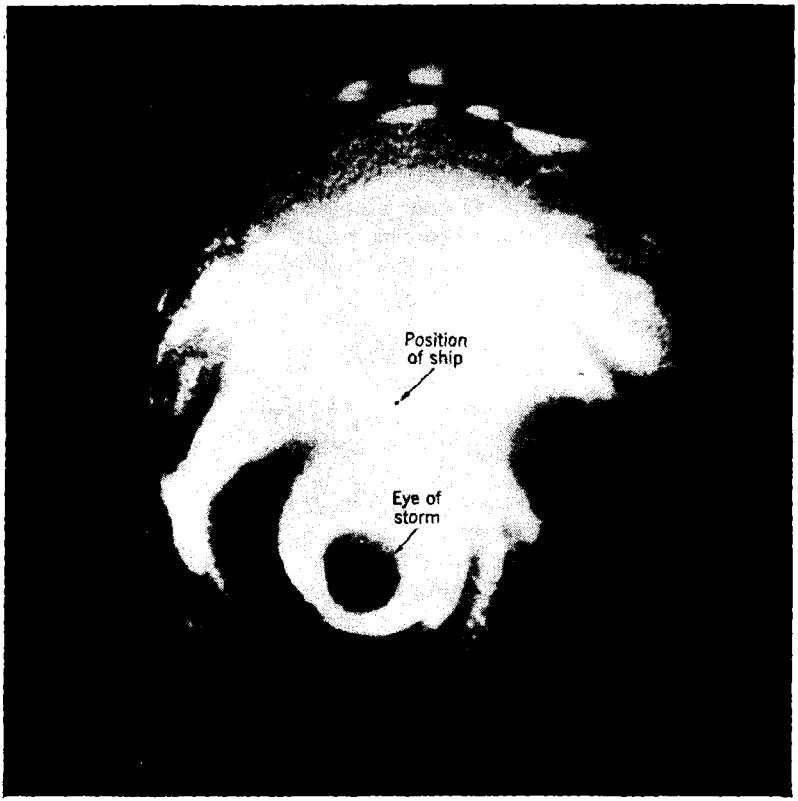


FIG. 7·18.—Typhoon near the Philippine Islands in December 1944, as observed on the PPI of a U. S. Navy 10-cm ship radar.

range was followed by standard meteorological techniques. The storm was tracked for about 3 hr by radar at a range of about 30 miles while the ship followed a course roughly parallel to the storm track. In the photograph the center was 35 miles distant from the ship, bearing 010° true. Motion of the storm was toward the northwest. Note that the distribution of heavy precipitation is clearly shown. The clear central eye of the storm

[1] G. F. Koscoe, "Highlights of the December 1944 Typhoon Including Photographic Radar Observations," Fleet Weather Central Paper No. 10, Feb. 10, 1945.

is readily observed, surrounded by an inner heavy circular rain area. Lying farther from the center and separated from the outer edge of the inner rain area by a relatively clear space is found a crescent-shaped precipitation area.

The distribution of precipitation observed by radar around this storm was not in agreement with expected maximum rainfall in the right front quadrant, and the meteorologists who observed this case concluded that although maximum convergence may have occurred in the right semicircle, it was largely dissipated in accelerating the air stream and thus produced relatively little rising air and precipitation. The clear space between the inner and outer rain areas was explained by a large vertical eddy cell built up over the left quadrant of the storm, the descending currents of which produced the clear sector through compressional heating. Although the explanations in this particular case are speculative, they suggest future applications in meteorological research.

CHAPTER 8

ATMOSPHERIC ATTENUATION

By J. H. Van Vleck, E. M. Purcell, and Herbert Goldstein

THE RELATION BETWEEN ABSORPTION AND DISPERSION

By J. H. Van Vleck

In studying the propagation of microwaves through the atmosphere one is interested both in the absorption and in the dispersion, that is, how the index of refraction varies with frequency. A very general relation exists between the refractive index and the absorption coefficient that enables one to determine the absorption if the dependence of the refractive index on the frequency is known throughout the spectrum from $\nu = 0$ to $\nu = \infty$. Vice versa, the refraction can be computed (apart from an additive constant ε_∞) if the absorption is known for all wavelengths. In other words, assumptions cannot be made regarding how the refractive index varies with frequency without implications as to the amount of absorption. This fact has a bearing on the theory of nonstandard refraction. For instance, various investigators have noted that a variation of 1 per cent in the refractive index between $\lambda = 1$ and $\lambda = 10$ cm would lead to interesting differences in the trapping or duct phenomena at the two wavelengths. Such a variation, however, would necessarily lead to an unreasonably high absorption, in contradiction with experiment.

8·1. Properties of the Complex Dielectric Constant.[1]—The general relations between the refractive index and the absorption arise out of a correlation of these two quantities, respectively, with the real and imaginary parts of a complex dielectric constant. The relations in question were first noted by Kramers in 1927.[2] Analogous relations between the real and imaginary parts of the impedance rather than of the dielectric constant have been well known in circuit theory since about 1932.[3] The

[1] The material in section 8·1 is condensed from J. H. Van Vleck, RL Report No. 735, May 28, 1945.

[2] H. A. Kramers, *Atti cong. intern. fisici*, Como, **2,** 545 (1927). One of the relations is also implicitly contained in slightly earlier papers by R. de L. Kronig, *Jour. Optical Soc. Am.*, **12,** 547 (1926), and by H. Kallman and H. Mark, *Ann. Physik*, **82,** 585 (1927).

[3] *Cf.* Y. W. Lee, *Jour. Math. Phys.*, **11,** 83 (1932); also O. Brune, *ibid.*, **10,** 191 (1931); C. M. Gewertz, *ibid.*, **12,** 91 (1933); W. Kigenhuis and F. L. Stumpers, *Physica*, **8,** 289 (1941); E. A. Guillemin, Circuit Theory Lectures Nos. 11 and 14, M.I.T.

proof is practically identical in the two cases and is closely related to what the mathematicians call the "Hilbert transform."

If the complex dielectric constant $\varepsilon = \varepsilon_1 - i\varepsilon_2$ is known, then the refractive index n and absorption γ in decibels per kilometer can immediately be calculated, for

$$n = \sqrt{\varepsilon_1}, \qquad \gamma = \frac{2\pi}{\lambda} 10^6 \varepsilon_2 \log_{10} e, \tag{1}$$

provided $n^2 - 1$ is small compared with unity.

The general relations between absorption and refraction are most simply expressed in terms of ε_1, ε_2 rather than of n, γ and are

$$\varepsilon_1(\nu) - \varepsilon_\infty = \frac{2}{\pi} \int_0^\infty \frac{\nu' \varepsilon_2(\nu')}{\nu'^2 - \nu^2} d\nu', \tag{2}$$

$$\varepsilon_2(\nu) = -\frac{2}{\pi} \nu \int_0^\infty \frac{\varepsilon_1(\nu') - \varepsilon_\infty}{\nu'^2 - \nu^2} d\nu', \tag{3}$$

where $\varepsilon(\nu)$ and ε_∞ mean, respectively, the dielectric constant for frequency ν and for infinite frequency. By means of Eq. (2), we compute ε_1 if we know ε_2 for all ν; conversely Eq. (3) determines ε_2 if we know ε_1. The integrals in Eqs. (2) and (3) diverge at the resonance singularity $\nu = \nu'$ if taken literally, and here we are to use the so-called Cauchy principal value, defined by

$$\int_0^\infty f \, d\nu' = \lim_{\delta \to 0} \left(\int_0^{\nu - \delta} f \, d\nu' + \int_{\nu + \delta}^\infty f \, d\nu' \right).$$

It would be well to stress the generality of the relations Eqs. (2) and (3). They require only that the complex dielectric constant $\varepsilon(\nu)$ be an analytic function of ν throughout the lower half of the complex plane for ν and that $\varepsilon(-\nu)$ be the conjugate of $\varepsilon(\nu)$ along the real axis in this plane. These conditions are really no assumptions at all, for they are necessarily fulfilled by virtue of the fact that the polarization of a particle cannot antecede the arrival of the disturbing electric field that produces it. The validity of Eqs. (2) and (3) is general and quite irrespective of the model, which can be either classical or quantum-mechanical.

The "particle" or structural unit can be an atom, electron, molecule, water drop, ice crystal, or anything else provided only that a considerable number of them be included in a volume unit whose dimensions are small compared with the path length.

The proof of the relations Eqs. (2) and (3) will be sketched only briefly. The validity of the law of cause and effect requires that $\varepsilon(\nu)$ have no singularities in the lower half of the complex ν-plane, as it can be shown that otherwise the response to a pulsed wave would antecede arrival of the latter. Hence by the residue theorem, one has

$$\int \frac{[\varepsilon(\nu') - \varepsilon_\infty]}{\nu' - \nu} d\nu' = 0 \tag{4}$$

if the path of integration is a closed curve formed by an infinite semicircle in the lower half plane and the real axis, except that the singularity at $\nu = \nu'$ is avoided by replacing a small segment of the real axis by a semicircle of radius δ with origin at $\nu = \nu'$. If δ be made to approach zero, the contribution of this small semicircle to the left side of Eq. (4) is $\pi i\,[\varepsilon\,(\nu) - \varepsilon_\infty]$. The contribution of the infinite arc is zero, and the linear portion of the contour involves the same path as implied by the Cauchy principal value. By virtue of these various observations and the further fact that $\varepsilon_1(\nu') = \varepsilon_1(-\nu')$, $\varepsilon_2(\nu') = -\varepsilon_2(-\nu')$, Eqs. (2) and (3) follow from Eq. (4). There are two relations, Eqs. (2) and (3), because in the complex equation (4), the real and imaginary parts of the left-hand side must vanish separately.

From Eqs. (2) and (3) it follows that any observable change in refractive index in the microwave region would imply a fantastically high absorption. For estimating orders of magnitude it will suffice to assume that ε_1 falls off linearly from $\nu = 0$ to $\nu = \nu_1$ and is constant thereafter. Let us suppose that $\varepsilon_1 - 1$ is one-tenth of 1 per cent lower for $\lambda_1 = c/\nu_1 = 1$ cm than the value 0.000558 for $\lambda = \infty$. This change is probably far too small to have any importance for trapping or duct phenomena. On evaluation of the integral it is found that the absorption would be at least 0.5 db/km for all wavelengths shorter than about 2 cm. This attenuation would be so high as to cause a great deal of difficulty and lead to contradiction with experiment. Thus we can safely conclude that *for all frequencies in the microwave region, the static value of the dielectric constant can be used without appreciable error.*

In actual materials, ε_1 varies with ν in a more complicated way than in the linear fashion assumed in the preceding paragraph in order to estimate orders of magnitude. We may distinguish between two cases, viz., the nonresonant and the resonant. In the first of these the real part of the dielectric constant decreases monotonically as the frequency is increased from zero to infinity. The standard theory of nonresonant absorption is that of Debye in classical theory and transcribed into quantum mechanics by Kronig and Gorter.[1] It assumes a molecule whose free rotations are resisted by a viscous force and leads to the formula

$$\varepsilon_1 - i\varepsilon_2 = \frac{\varepsilon_0 - \varepsilon_\infty}{1 + \dfrac{i\nu}{\Delta\nu}} + \varepsilon_\infty, \tag{5}$$

where $\Delta\nu$ is the line-breadth constant and $\varepsilon_0 - \varepsilon_\infty$ has the significance

$$\varepsilon_0 - \varepsilon_\infty = \frac{4\pi N\mu^2}{3kT}, \tag{6}$$

in terms of the dipole moment μ of the molecule and the number N of molecules per unit volume. The refraction and absorption predicted by Eq. (5)

[1] P. Debye, *Polar Molecules*, Chemical Catalog Co., New York, 1929, Chap. 5; C. J. Gorter and R. de L. Kronig, *Physica*, **3**, 1009 (1936); R. de L. Kronig, *ibid.*, **5**, 65 (1938).

are shown in Fig. 8·1.[1] The absorption implied by a given change $\epsilon_0 - \epsilon$ in dielectric constant is of the same order of magnitude as in the example discussed in the preceding paragraph.

If the absorption is of the resonant rather than nonresonant type, there is still no escape from the conclusion that to avoid fantastically large absorption, the refraction must be appreciably constant over the microwave

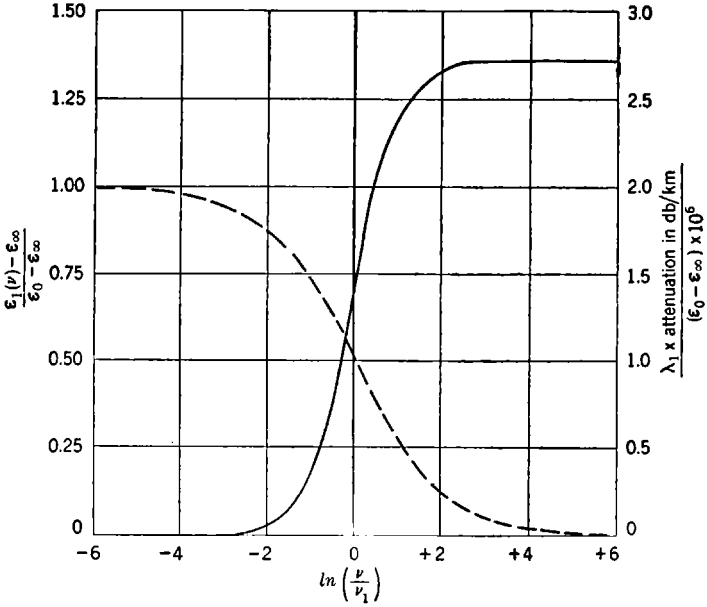


FIG. 8·1.—Dependence of nonresonant refraction and absorption on frequency. The curves are based on the theory of Debye. The dashed and solid curves relate to the refraction and absorption. The notation ν, λ, is used for $\Delta\nu$, $c/\Delta\nu$ respectively.

domain. With a resonant frequency ν_0, the quantum-mechanical formula for the complex dielectric constant is

$$\epsilon - 1 = \frac{A\nu_0^2}{(\nu_0^2 - \nu^2)}\left\{1 - \frac{\Delta\nu}{2}\frac{\left(\frac{\nu}{\nu_0}+1\right)\frac{\nu}{\nu_0}}{[-i(\nu_0-\nu)+\Delta\nu]} - \frac{\Delta\nu}{2}\frac{\left(\frac{\nu}{\nu_0}-1\right)\frac{\nu}{\nu_0}}{[i(\nu_0+\nu)+\Delta\nu]}\right\}, \tag{7}$$

[1] In many tests it is customary to plot the absorption per wavelength rather than the absorption per fixed unit length. The absorption curve then, because of an extra factor λ, shows a maximum at a particular frequency rather than increasing monotonically with frequency as in Fig. 8·1.

where A is a dimensionless constant. In Eq. (7), the line-broadening is supposed due to collisions, and throughout the chapter we neglect the influences of the Doppler effect or of radiation damping, as their effects in the microwave region are negligible. In case the line-breadth constant $\Delta\nu$ is small compared with the resonance frequency, by far the most important part of the absorption and dispersion is that in the immediate vicinity of the resonance center. In this region, the formulas for ε_1 and ε_2 become substantially

$$\varepsilon_1 - 1 = \frac{A\nu_0}{2}\left[\frac{(\nu_0 - \nu)}{\Delta\nu^2 + (\nu - \nu_0)^2}\right] \tag{8}$$

and

$$\varepsilon_2 = \frac{A\nu_0}{2}\left[\frac{\Delta\nu}{\Delta\nu^2 + (\nu - \nu_0)^2}\right]. \tag{9}$$

These formulas show the situation to be different from the nonresonant case, where ε_1 decreased monotonically with increasing ν. The maximum

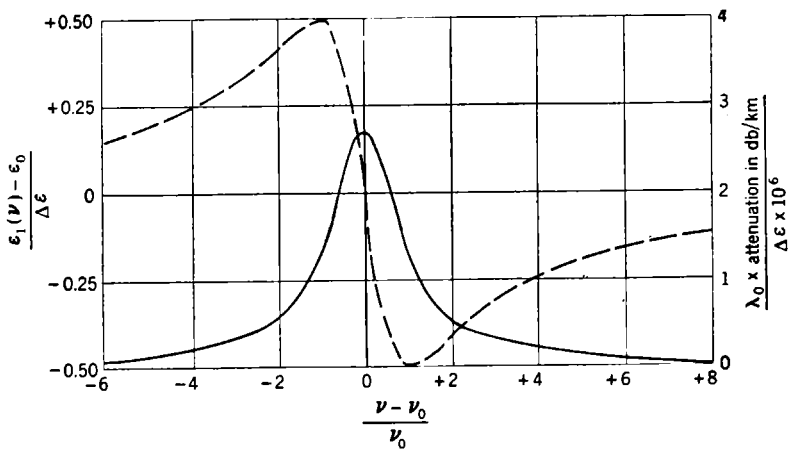


FIG. 8·2.—Dependence of resonant absorption and refraction on frequency in the vicinity of resonance.

and minimum values of $\varepsilon_1 - 1$, viz., $\pm A\nu_0/4\Delta\nu$, are achieved at $\nu = \nu_0 \mp \Delta\nu$. As $\nu_0 \gg \Delta\nu$, the difference $\Delta\varepsilon = A\nu_0/2\Delta\nu$ between these extremes is much larger than $\varepsilon_\infty - \varepsilon_0$ which by Eq. (7) equals A. The situation in the immediate vicinity of the absorption line is sketched in Fig. 8 2. The maximum value of the absorption (assuming $\nu_0 \gg \Delta\nu$) is achieved at $\nu = \nu_0$ and by Eqs. (1) and (9) is $0.434(2\pi) \times \Delta\varepsilon \times 10^6(1/\lambda_0)$ where $\lambda_0 = c/\nu_0$. Thus if there is a resonant line that changes the value of $n^2 - 1$ for air by 1 part in 1000 (that is, $\Delta\varepsilon = 5.8 \times 10^{-7}$), the absorption at the center of the line will amount to $1.6/\lambda_0$ db/km. Again an appreciable change in the refraction cannot occur without being manifest in absorption.

Sections 8·2 and 8·3 treat the atmospheric attenuation by lines in O_2 and H_2O located at about $\frac{1}{2}$ cm and 1.35 cm respectively. These resonances are of importance from the standpoint of absorption, but from the preceding formulas the concomitant changes in refraction are found to be inconsequential; namely, the two extremes of the kink in the values of $n^2 - 1$ which is caused by the O_2 absorption near $\frac{1}{2}$ cm should differ from each other by only about 1 part in 300. This estimate may be too large because the absorption by oxygen is complicated by the fine structure from molecular rotation and so cannot be treated as a single line. In the case of the water-vapor line at 1.35 cm, the absorption is much weaker than that for O_2 and gives a change in $n^2 - 1$ amounting only to about one part in several thousand, too small an effect to have any appreciable bearing on nonstandard refraction. In heavy rains, the absorption arises more from drops than from uncondensed vapor; it is then essentially nonresonant in character and of a type to be studied in Sec. 8·7. In extreme torrential rains, occurring mostly in the tropics, the attenuation in the 1-cm region can be as large as several decibels per kilometer, and then $n^2 - 1$ might differ from the static value by one part in a few hundred, but such behavior will be relatively rare.

THEORY OF ABSORPTION BY UNCONDENSED GASES[1]

By J. H. Van Vleck

The absorption of microwaves in the atmosphere arises from both free molecules and suspended solid particles, that is, dust grains and especially the water drops condensed in fogs or rains. The present section is concerned only with the molecular absorption. The influence of particles will be treated in Sec. 8·7. The attenuation brought about by the gaseous and particle mechanisms enters additively; in other words, the decay that we now compute is always present as a "background" superposed on any effects produced by particles. (The size of the two effects will be compared in Sec. 8·7.) When it rains, they are often of the same rough order of magnitude. In a noncondensed atmosphere the ingredients that absorb microwaves are (1) oxygen, which has a magnetic interaction with the incident radiation because the O_2 molecule is normally in a paramagnetic $^3\Sigma$ state, and (2) water vapor, which contributes because of the electric polarity of the H_2O molecule. In both cases, there are certain frequency regions where the absorption is abnormally large because of resonance. In the case of oxygen, this region occurs at wavelengths in the vicinity of $\frac{1}{2}$ cm (and also $\frac{1}{4}$ cm), and at the peak of the resonance the absorption is over 15 db/km. In the case of water, the resonance center is at 1.35 cm,

[1] The material in Secs. 8·2 and 8·3 was for the most part given in RL Reports Nos. 175 and 664, which present a more complete discussion. Further improvements in numerical values found here are given by J. H. Van Vleck, *Phys. Rev.*, **71**, 413*ff* (1947).

but with a much weaker absorption. Naturally the amount of absorption depends on the humidity—at saturation at 20°C the absorption at maximum resonance amounts to about 0.4 db/km—enough to have some bearing on radar performance over long path lengths.

The following tables give a bird's-eye view of the absorption due to the two causes.

TABLE 8·1.—THEORETICAL VALUES OF ATMOSPHERIC ABSORPTION BY OXYGEN

λ, cm	γ, db/km	λ, cm	γ, db/km	λ, cm	γ, db/km
100	0.0014	1.00	0.014	0.465	5.0
30	0.0050	0.667	0.077	0.435	0.51
10	0.0066	0.588	0.32	0.400	0.19
3	0.0072	0.556	1.99	0.200	0.03
1.5	0.0089	0.50	14		

TABLE 8·2.—THEORETICAL VALUES OF THE RATIO OF ATMOSPHERIC ABSORPTION BY WATER VAPOR TO NUMBER OF GRAMS OF H_2O PER CUBIC METER

λ, cm	γ/ρ, db/km per g per m³	λ, cm	γ/ρ, db/km per g per m³	λ, cm	γ/ρ, db/km per g per m³
10	0.000025	1.43	0.015	1.0	0.0059
3	0.00039	1.35	0.020	0.667	0.0041
2	0.0016	1.25	0.018	0.5	0.0058
1.5	0.011	1.11	0.009	0.3	0.014

The entries in Table 8·1 give the decay constant γ in decibels per kilometer from the absorption by oxygen. In other words, besides the decrease due to the inverse square law, the *intensity* of the beam diminishes at the rate $10^{-0.1\gamma l}$, where l is the effective path length in kilometers.

The absorption by water vapor is, of course, directly proportional to the vapor content in Table 8·2, and consequently we give the absorption divided by the density ρ of water vapor in grams per cubic meter. For temperate latitudes (20°C) in summer, there are on the average about 7.5 g of water per cubic meter. On the other hand, at saturation at 20°C, sea level, the value is $\rho = 17$ g/m³. Under tropical conditions the content can be even higher.

The calculations for the preceding table are made on the assumption that the air traversed is at 76-cm pressure and 20°C. When waves are being used at a substantial altitude, allowance must be made for the fact that the quotient of density by temperature, to which absorption is proportional, diminishes at high altitudes. Approximate allowance for this fact can be made in the case of oxygen by taking the effective path length l

to be $l_0\left(1-\frac{h+h'}{20}\right)$, where l_0 is the actual path length and h, h' are, respectively, the elevation of the initial and terminal points in kilometers. In considering the effect of water, allowance must be made for the variation of the fractional vapor content of the atmosphere with altitude, and an appropriate mean density must be used. The corrections to be made because the temperature may not be 20°C are relatively unimportant except, of course, for the influence on the amount of water vapor. The absorption by oxygen is roughly proportional to $1/T^2$ as long as the pressure is constant. The temperature corrections for water will be discussed more fully in Sec. 8·3.

At wavelengths less than about 1 mm, the absorption by water vapor is so high as to make such wavelengths useless for transmission except over exceedingly short distances. This statement does not apply, of course, if the wavelength is reduced indefinitely. Obviously, for instance, the atmosphere is not opaque in the optical region—otherwise we could not live. The opening up of the atmosphere to transmission, however, occurs only when the frequency becomes high compared with the rotational frequencies of the water molecule, and then we are in the infrared rather than in the very short microwave region.[1] Values of the absorption in the millimeter region will be given in Sec. 8·3, Fig. 8·7.

8·2. Oxygen.—To have appreciable absorption, it is necessary that the incident frequency not differ too much from an atomic or molecular resonance frequency, which by quantum theory is proportional to the difference between two molecular energy levels. Even the shortest microwave is very long by ordinary optical standards; hence appreciable absorption is to be expected only if there are closely spaced energy levels between which transitions are permitted by the selection rules of quantum mechanics. The nitrogen and oxygen molecules, which are the important constituents of dry air, are both nonpolar or, in other words, devoid of a permanent electric moment active in the infrared or microwave region. As a result, only permitted transitions in electric dipole radiation connect different electronic states and give rise to absorption in the visible or ultraviolet rather than the microwave region.

It must not, however, be overlooked that the oxygen molecule is paramagnetic or, in other words, has a permanent magnetic moment and can consequently absorb microwaves. Absorption is usually thought of as arising only from electric dipoles, but it can arise from magnetic polarity also, as Maxwell's equations are symmetric in **E** and **H**. At first sight, it might appear that any effect of this kind would be negligible, for molecular magnetic moments are much smaller than electric ones, corresponding to the fact that one Bohr magneton is about $\frac{1}{100}$ of a Debye

[1] For a calculation of the absorption by water vapor in the infrared, see W. M. Elsasser, *Astrophys. Jour.*, **87**, 497 (1938).

unit of electric polarity. Actually, however, the resulting attenuation is not negligible, because of the great path lengths in the atmosphere and because there is resonance to the energy intervals in the oxygen triplets if the incident wavelength is about $\frac{1}{2}$ cm. The magnetic moment has diagonal matrix elements also involving no energy change and hence zero frequency; or in classical language, there exists a constant term in its Fourier expansion. A corresponding term is not found for electric dipoles. A completely undisturbed constant moment does not give absorption; but when allowance is made for interruption by collisions, the constant or diagonal elements are capable of absorbing energy and give rise to a feeble attenuation in the conventional microwave region.

From band spectra the oxygen molecule is known to be in a $^3\Sigma$ ground state and to carry a magnetic dipole moment of 2 Bohr magnetons (spin quantum number $S = 1$; gyromagnetic ratio $g = 2$). This magnetic moment interacts with the "end-over-end" rotation of the oxygen molecule to form what is sometimes called a "rho-type triplet."[1] If K be the rotational quantum number exclusive of spin, then K is compounded vectorially with the spin S to yield a resultant inner or rotational quantum J for the molecule, which can take on the values $J = K - 1$, K and $K + 1$. The illustration in Fig. 8·3 is not to scale, as the spacing between the different rotational states K is large compared with the intervals within the triplet and involves energy intervals appropriate to the infrared rather than the microwave region. Because the oxygen nucleus is devoid of spin, the Pauli exclusion principle allows K to take on only odd values, but K can range from 1 to ∞, and only the rotational states 7, 9 are shown in Fig. 8·3. The components $J = K - 1$ and $J = K + 1$ nearly coincide and differ from $J = K$ by intervals that depend on K but which, with one exception, are about 2 cm^{-1}, corresponding to a wavelength of $\frac{1}{2}$ cm. The one exception is furnished by the components $J = 0$, 1 of $K = 1$ for which the separation is approximately 4 cm^{-1}, implying absorption in the vicinity of $\frac{1}{4}$ cm. At first sight it might seem that the small separation of $J = K - 1$ and $J = K + 1$ gives us just what we want for a low-frequency resonance, of the order $\lambda = 10$ cm, in the

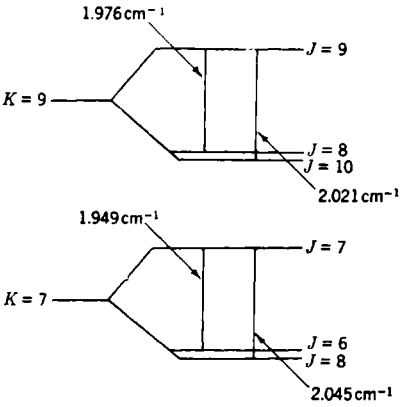


FIG. 8·3.—Spin triplets in the molecular spectrum of oxygen.

[1] For a detailed discussion of this triplet and further references, see R. Schlapp, *Phys. Rev.*, **51**, 343 (1937).

conventional rather than the very short microwave region. This is not really the case, however, as there are no matrix elements of the magnetic moment vector connecting $J = K - 1$ with $J = K + 1$, because of the selection rule that J cannot change by more than one unit.

The general quantum-mechanical formula for the absorption coefficient γ for microwaves of frequency ν is

$$\gamma = 10^6(\log_{10} e)\left(\frac{8\pi^3\nu N}{3hc}\right)\frac{\sum_j \sum_i \left\{|\mu_{ij}|^2 f(\nu_{ij}, \nu)\right\} e^{\frac{-E_j}{kT}}}{\sum_j e^{\frac{-E_j}{kT}}}. \tag{10}$$

Here μ_{ij} is the matrix element of the dipole moment connecting two stationary states i, j, of energy E_i, E_j respectively; ν_{ij} is the frequency of the corresponding spectral line, given by the Bohr frequency condition $h\nu_{ij} = E_i - E_j$; and N is the number of molecules per cubic centimeter. The frequency of the incident radiation is denoted by ν. The factor $10^6\log_{10}e$ is inserted in order that the absorption be expressed in decibels per kilometer. The factor $f(\nu_{ij}, \nu)$ is the so-called "structure factor" determining the shape of the absorption line and is given by the expression

$$f(\nu_{ij}, \nu) = \frac{\nu}{\pi\nu_{ij}}\left[\frac{\Delta\nu}{(\nu_{ij} - \nu)^2 + \Delta\nu^2} + \frac{\Delta\nu}{(\nu_{ij} + \nu)^2 + \Delta\nu^2}\right]. \tag{11}$$

Near resonance, Eq. (11) is practically identical with

$$f(\nu_{ij}, \nu) \approx \frac{1}{\pi}\left[\frac{\Delta\nu}{(\nu_{ij} - \nu)^2 + \Delta\nu^2}\right]. \tag{12}$$

Equation (12) is the form of structure factor commonly given in the literature and is adequate for most purposes, especially in the optical region. This approximation, however, is not adequate in all applications in microwaves, as we shall sometimes be interested in cases where $(\nu_{ij} - \nu)$ and even $\Delta\nu$ may be of the same order of magnitude as ν. It is then necessary to use the more refined formula[1] of Eq. (11). For the theory underlying Eq. (11), which involves inclusion of a term omitted in the standard Lorentz theory of collision broadening, the reader is referred to a paper by Van Vleck and Weisskopf.[1] The constant $\Delta\nu$ is the so-called line-breadth constant. It is proportional to the number of collisions per unit time but cannot be calculated theoretically, as the collision cross section is not known. Uncertainty as to the value of $\Delta\nu$ has been the greatest stumbling block in the theoretical calculation of absorption. This point will be discussed more fully later, but it can be mentioned here that the microwave measurements themselves have contributed a great deal of information on the magnitude of $\Delta\nu$. In fact,

[1] J. H. Van Vleck and V. Weisskopf, *Rev. Mod. Phys.*, **17**, 227 (1945).

from a theoretical standpoint, one of the interesting features of microwave absorption is that it covers a region where $\Delta\nu$ may be comparable to ν, and the distinction between Eqs. (11) and (12) consequently cease to be trivial. In other words, questions of line profile can be studied to greater advantage in the microwave than in the optical region.

As $h\nu_{ij} = E_i - E_j = -h\nu_{ji}$, we have $f(\nu_{ij}, \nu) = -f(\nu_{ji}, \nu)$ and the contribution of a typical term i, j in Eq. (10) to the absorption is positive or negative (induced emission) according as the energy of the state i is greater or less than that of j. To any given term i,j in the double sum in Eq. (10) there is also a corresponding term j,i whose contribution is of opposite sign and which would cancel the effect of i,j completely were it not for the difference in the Boltzmann factors for the states i and j. However, in case the frequencies ν_{ij} are small compared with kT/h, this difficulty of nearly compensating terms may be avoided by using the approximation

$$\nu_{ij}e^{\frac{-E_j}{kT}} + \nu_{ji}e^{\frac{-E_i}{kT}} = \frac{1}{2}\frac{h\nu_{ij}^2}{kT}\left(e^{\frac{-E_j}{kT}} + e^{\frac{-E_i}{kT}}\right), \tag{13}$$

which shows that the positive and negative terms nearly cancel if $h\nu_{ij} \ll kT$. Use of the simplification (13) is fully warranted in the present calculation on oxygen, as the width of the rho-type triplet is only about $\frac{1}{100}$ of kT/h. When Eq. (13) is employed, Eq. (10) becomes

$$\gamma = 10^6(\log_{10} e)\,\frac{8\pi^3\nu N}{6ckT}\,\frac{\sum_j \sum_i \left[|\mu_{ij}|^2 \nu_{ij} f(\nu_{ij}, \nu)\right] e^{\frac{-E_j}{kT}}}{\sum_j e^{\frac{-E_j}{kT}}}. \tag{14}$$

The advantage of using Eq. (14) rather than Eq. (10) is that all the terms of the numerator of Eq. (14) are positive.

In the present specific application to oxygen, the index i (or j) signifies a trio of quantum numbers J, K, M, (or J', K', M'), where we have $J = K - 1$, K, $K + 1$ as previously explained, and where M is the equatorial quantum number associated with the space quantization of J. Various simplifications are possible, however. In the first place, the energy does not depend on M and only slightly on J; hence, in the Boltzmann factors, but not, of course, in the frequency differences ν_{ij} we can take $E(K, J) = E_K$. The matrix elements $\mu(KJM; K'J'M')$ of the magnetic moment are of the type $K' = K$, $J' = J$, $J - 1$, $J + 1$. Consequently it is found that Eq. (14) can be written

$$\gamma = 10^6(\log_{10} e)\,\frac{4\pi^3\nu N}{3ckT}\,\frac{\sum_K \left[G_{K+}\mu_{K+}^2 + G_{K-}\mu_{K-}^2 + F(\nu)\mu_{K0}^2\right] e^{\frac{-E_K}{kT}}}{\sum_K 3(2K+1)\, e^{\frac{-E_K}{kT}}}. \tag{15}$$

Here we have used the abbreviations

$$\mu_{K0}^2 = \sum_{J=K-1}^{J=K+1} \sum_{M=-J}^{M=J} \sum_{M'=-J}^{M'=J} |\mu(KJM;\ KJM')|^2 ,$$

$$\mu_{K\pm}^2 = \sum_{M=-K}^{M=K} \sum_{M'=-K\mp 1}^{M'=K\pm 1} |\mu(KKM;\ KK \pm 1M')|^2 ,$$

$$\nu_{K\pm} = |\ \nu(KK;\ KK \pm 1)\ |,$$

$$G_{K\pm} = 2\ |\ \nu_{K\pm} f(\nu_{K\pm},\ \nu)\ |,$$

$$F(\nu) = \lim_{\nu_{ij} \to 0} [\nu_{ij} f(\nu_{ij},\ \nu)] = \frac{2\nu\,\Delta\nu}{\pi(\nu^2 + \Delta\nu^2)}.$$

The intensity formulas of quantum mechanics[1] appropriate to the type of vector addition involved in a rho-type triplet (Hund's Case *b*) show that

$$\left.\begin{aligned} \mu_{K+}^2 &= \frac{4\beta^2 K(2K+3)}{K+1}, \qquad \mu_{K-}^2 = \frac{4\beta^2(K+1)(2K-1)}{K}, \\ \mu_{K0}^2 &= \frac{8\beta^2(K^2+K+1)(2K+1)}{K(K+1)}, \end{aligned}\right\} \tag{16}$$

where β is the Bohr magneton $eh/4\pi mc$.

The third member of Eq. (15), that is, the term with factor $F(\nu)$, represents the nonresonant part of the absorption from the diagonal matrix elements that are the quantum-mechanical equivalent of a constant term in the Fourier expansion. Without appreciable error the approximation $\mu_{K0}^2 = 8\beta^2\ (2K+1)$ can be made in Eq. (16), as most rotational states have $K \gg 1$. Then the sum cancels from numerator and denominator in the third member of Eq. (15); consequently the nonresonant part of the absorption is given by the very simple formula

$$\gamma_{NR} = 10^6(\log_{10} e)\,\frac{64\pi^2\beta^2 N}{9ckT}\,\frac{\nu^2\,\Delta\nu}{\nu^2 + \Delta\nu^2} = \frac{0.34\nu^2\,\Delta\nu}{\nu^2 + \Delta\nu^2}. \tag{17}$$

This expression has precisely the same structure as the well-known classical formula of Debye for absorption at radio frequencies in dielectric liquids, which is obtained by substituting Eqs. (5) and (6) in Eq. (1). The square of the dipole moment in this connection is $8\beta^2/3 = g^2 S(S+1)\,\beta^2/3$, where g is the Landé factor and S is the spin quantum number; the factor $\frac{1}{3}$ appears because one-third of the total mean square moment is of the diagonal variety, the other two-thirds being consumed by the nondiagonal type $J' = J \pm 1$.

[1] *Cf.* E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra*, Cambridge, London, 1935, p. 66, or J. H. Van Vleck, *Electric and Magnetic Susceptibilities*, Oxford, New York, 1932, p. 167, Eq. (88).

The nonresonant effect represents the most important part of the absorption when the incident wavelength is very long compared with the resonance wavelength $\frac{1}{2}$ cm but nevertheless does not give enough absorption to be important for most practical purposes, as can be seen from the entries in Table 8·1 for $\lambda \geq 3$ cm. The distinction between the resonant and nonresonant absorption is, in a certain sense, a purely artificial one, as the diagonal part of the absorption is merely a special case of the general resonant variety in which the resonance frequency happens to be zero. Prior to the revision, already mentioned, of the Lorentz theory, the diagonal and nondiagonal contributions to the absorption had to be treated in different ways, as the one type did not properly reduce to the other when the resonant frequency was set equal to zero, or in other words, the general formula for absorption with an arbitrary absorption frequency did not include the Debye-Kronig expression as a special case. However, we now have a unified theory in which this difficulty is avoided.

The first and second members of Eq. (15) give the contribution of the nondiagonal elements to the absorption, which is particularly important in the vicinity of $\lambda = \frac{1}{2}$ cm. This resonant absorption is not given by so simple an expression as Eq. (17). In the Boltzmann factors we can take $W_K = BK(K+1)$ with $B = 1.44$ cm^{-1} with the result that at 293°K we have

$$e^{-\frac{E_K}{kT}} = 10^{-0.00301K(K+1)}.$$

Because of the exclusion principles only odd value of K are involved in the summation. Without appreciable error, the sum in the denominator can be replaced by the integral

$$3\int_0^\infty Ke^{-BK^2/kT}\,dK = \frac{3kT}{2B}.$$

On the other hand, the first and second members of the numerator must be summed in detail, and it is necessary to know the energy levels fairly accurately. Here the values shown in Table 8·3 have been adopted. They are based largely on the spectrosopic measurements of Dieke and Babcock,[1] except that (1) the value of ν_{K-} for $K = 1$ cannot be observed directly and is obtained from a theoretical formula of Schlapp[2] and (2)

[1] G. H. Dieke and H. D. Babcock, *Proc. Nat. Acad. Sci.*, **13**, 670 (1927). Recently an improved determination of the triplet intervals has been made. [H. D. Babcock and L. Herzberg, *Astrophys. Jour.*, **108**, 167 (1948).] The new values of these intervals were not available at the time this chapter was written and differ only very slightly from those given in Table 8·3. Hence, use of the revised intervals would not affect the calculated absorption appreciably except at very low pressures, where the contributions of the individual resonance lines become resolved.

[2] R. Schlapp, *loc. cit.*

in some cases the measured values have been smoothed out where they seem likely to be unreliable.

TABLE 8·3.—INTERVALS IN THE OXYGEN TRIPLETS

K	$\frac{\nu_{K-}}{c}$	$\frac{\nu_{K+}}{c}$	K	$\frac{\nu_{K-}}{c}$	$\frac{\nu_{K+}}{c}$
1	3.962	1.872	15	1.877 cm^{-1}	2.10 cm^{-1}
3	2.084	1.950	17	1.864	2.12
5	2.011	1.981	19	1.838	2.14
7	1.976	2.021	21	1.815	2.16
9	1.949	2.045	23	1.81	2.18
11	1.919	2.06	$\geqslant 25$	1.80	2.20
13	1.899	2.08		$-0.01 \times (K - 25)$	$+0.01 \times (K - 25)$

In regions well removed from the absorption band (that is, for $\lambda > 1$ or $< \frac{1}{3}$ cm), one can make the centroid approximation of considering all the frequency denominators to be the same and use the approximate formula

$$\gamma = 0.34 \left(\frac{\nu^2}{c^2}\right) \left[\frac{\frac{\Delta\nu}{c}}{\left(2 - \frac{\nu}{c}\right)^2 + \left(\frac{\Delta\nu}{c}\right)^2} + \frac{\frac{\Delta\nu}{c}}{\left(2 + \frac{\nu}{c}\right)^2 + \left(\frac{\Delta\nu}{c}\right)^2} + \frac{\frac{\Delta\nu}{c}}{\left(\frac{\nu}{c}\right)^2 + \left(\frac{\Delta\nu}{c}\right)^2} \right].$$

In comparing the theory with experiment, one must first of all select some value of $\Delta\nu$. No direct measurements on the line breadth $\Delta\nu$ are possible in the infrared spectrum, since oxygen is electrically nonpolar. By extrapolating from CO and other gases, Van Vleck originally (in 1942) considered 0.1 cm^{-1} as the most reasonable choice for $\Delta\nu/c$, but the subsequent experimental work of Beringer, who measured the absorption of microwaves in a waveguide filled with oxygen, shows that 0.1 cm^{-1} is definitely too high a value for $\Delta\nu/c$ and that $\Delta\nu/c$ is probably between the limits 0.02 and 0.05 cm^{-1}. This conclusion is in agreement with the fact that improved infrared technique practically always seems to indicate that the conventional values of $\Delta\nu/c$ are too large. Besides Beringer's experiments, which will be discussed in Sec. 8·4, we have the measurements on the absorption of O_2 made by Mueller,[1] who determined how

[1] G. E. Mueller, BTL Report No. MM-44-160-150, July 3, 1944; also *Proc. IRE*, **34**, 181 *P* (1946). *Note added in proof.* During the printing of the present volume further determinations of the absorption of oxygen by transmission measurements in wave guide have been made. See M. W. P. Strandberg, C. Y. Meng, and J. G. Ingersoll, *Phys. Rev.*, **75**, 1524 (1949).

much the transmission of air over a long path deviated from the inverse square law at a wavelength of 0.625 cm. At this wavelength, practically all the absorption is due to the oxygen in the air. Mueller concludes that $\Delta\nu/c$ must be less than 0.09 db/km. As the theoretical value of $\Delta\nu/c$ at 0.625 cm is 6.5 ($\Delta\nu/c$), Mueller's measurements seem to demand that the line-breadth constant $\Delta\nu/c$ not exceed 0.014 cm^{-1}, which is lower than the range of values yielded by Beringer's investigations. Mueller's experiment requires extreme precision. The value 0.09 db/km corresponds to a reduction of only 0.8 per cent in the transmitted signal over the quarter-mile path that he used. Mueller informs the author that it is allowable to raise his upper limit to 0.13 db/km. This corresponds to a line breadth 0.02 cm^{-1}, the lowest value of $\Delta\nu/c$, which seems reasonable in the light of Beringer's experiments, although a sharp definition of his least admissible $\Delta\nu/c$ is not possible. In addition, by the radiometer method to be described in Sec. 8·5, measurements have been made on the total absorption by oxygen in the earth's atmospheric layer.[1] The wavelengths (1, $1\frac{1}{4}$, $1\frac{1}{2}$ cm) employed in these experiments are so far from resonance that the effect is weak and hard to determine with precision. The indications are that the best fit is obtained if $\Delta\nu/c$ is of the order of 0.02 cm^{-1}. In the combined light of the Beringer, Mueller,[2] and Dicke data, 0.02 cm^{-1} is perhaps the most reasonable value for $\Delta\nu/c$ and is that used for Table 8·1.

The values of the computed absorption by the oxygen of the atmosphere are shown in Fig. 8·4 as a function of wavelength. The various varieties of points in Fig. 8·4 give Beringer's experimental points. The measurements on pure or 50 per cent O_2 are reduced to "air equivalent" by assuming that at given total pressure the absorption is directly proportional to the concentration of oxygen. This procedure appears legitimate, since Beringer's data show the collision cross section to be substantially the same for $O_2 - N_2$ and $O_2 - O_2$ impacts. Figure 8·4 shows clearly that the early choice $\Delta\nu/c = 0.1$ cm^{-1} for the line breadth does not fit the data well and that Beringer's measurements probably fix $\Delta\nu/c$ between the limits 0.02 and 0.05 cm^{-1}.

The wobbles in the curve for $\Delta\nu/c = 0.02$ cm^{-1} are explained by the fact that at certain wavelengths there is an extremely close resonance to a particular rotational transition. This effect does not show up at 0.05 or 0.1 cm^{-1}, as the lines are so broad that they overlap, whereas at 0.02 one has an incipient resolution of the contributions of the different rotational states. By working with low pressures and well-stabilized frequencies, it should ultimately be possible to resolve this fine structure. In fact, micro-

[1] R. H. Dicke, R. Beringer, R. L. Kyhl, and A. B. Vane, *Phys. Rev.*, **70**, 340 (1946).

[2] Measurements similar to those of Mueller have also been made by H. R. L. Lamont [*Proc. Phys. Soc.*, **61**, 562 (1948)].

wave spectroscopy should ultimately prove the method par excellence of accurately determining the low energy levels of molecules as well as of line profiles. Beringer's measurements already include some data at low pressures, but the scatter of the individual runs is considerable. The theoretical curves for the absorption as a function of total pressure, with attempts to correlate with Beringer's data, are published elsewhere.[1]

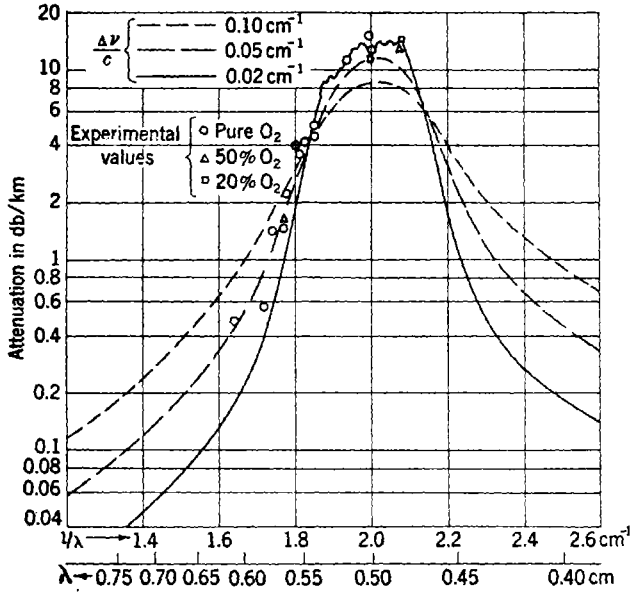


FIG. 8·4.—Atmospheric attenuation by oxygen as a function of wavelength. The measurements, when made on other than 20% O_2, are reduced to "air equivalent values" by assuming that at given total pressure the absorption is proportional to the partial pressure of oxygen.

8·3. Uncondensed Water Vapor.[2]—Actually, of course, the atmosphere is never completely dry, and we should therefore consider the absorption by water vapor in addition to that by oxygen. It is essential that the states of the water molecule be treated quantum-mechanically, for the precise spacing of the rotational energy levels cannot be deduced classically. The polarity to be considered is now electric rather than magnetic, of course. The important fact to note is that there is no absorption of the nonresonant type, or, in the language of quantum mechanics, the electric moment matrix of the molecule contains no diagonal elements. This situation is the result of the selection rules connecting transitions

[1] See RL Reports Nos. 175 and 664, which present a more complete discussion; also J. H. Van Vleck, *Phys. Rev.*, **71**, 413 (1947).

[2] See also J. H. Van Vleck, *Phys. Rev.*, **71**, 425 (1947).

between the energy levels of different symmetry types. It can be shown by group theory or otherwise that the rotational energy levels of the water molecule are of four different symmetry types, which are usually designated as $++$, $--$, $-+$, and $+-$. The selection rules permit the rotational quantum number J to change only by not more than one unit and allow $++$ to combine only with $--$ and $-+$ to combine only with $+-$. Nonresonant absorption would require diagonal matrix

TABLE 8·4.—LOW ROTATIONAL ENERGY LEVELS OF H_2O MOLECULES IN CM^{-1}

J_τ	$\frac{E}{hc}$	J_τ	$\frac{E}{hc}$	J_τ	$\frac{E}{hc}$
$++\ 0_0$	0	$++\ 4_4$	488.19	$++\ 5_{-2}$	416.00
		$-+\ 4_3$	488.19	$+-\ 5_{-3}$	399.44
$+-\ 1_1$	42.36	$--\ 4_2$	383.93	$--\ 5_{-4}$	326.50
$--\ 1_0$	37.14	$+-\ 4_1$	382.47	$-+\ 5_{-5}$	325.27
$-+\ 1_{-1}$	23.76	$++\ 4_0$	315.70		
		$-+\ 4_{-1}$	300.33	$++\ 6_6$	1045.14
$++\ 2_2$	136.15	$--\ 4_{-2}$	275.23	$-+\ 6_5$	1045.14
$-+\ 2_1$	134.88	$+-\ 4_{-3}$	224.71	$--\ 6_4$	888.74
$--\ 2_0$	95.19	$++\ 4_{-4}$	221.94	$+-\ 6_3$	888.70
$+-\ 2_{-1}$	79.47			$++\ 6_2$	757.72
$++\ 2_{-2}$	70.08	$+-\ 5_5$	742.18	$-+\ 6_1$	756.78
		$--\ 5_4$	742.18	$--\ 6_0$	661.54
$+-\ 3_3$	285.46	$-+\ 5_3$	610.34	$+-\ 6_{-1}$	649.03
$--\ 3_2$	285.26	$++\ 5_2$	610.16	$++\ 6_{-2}$	602.71
$-+\ 3_1$	212.12	$+-\ 5_1$	508.80	$-+\ 6_{-3}$	552.94
$++\ 3_0$	206.35	$--\ 5_0$	503.90	$--\ 6_{-4}$	542.74
$+-\ 3_{-1}$	173.38	$-+\ 5_{-1}$	446.39	$+-\ 6_{-5}$	447.17
$--\ 3_{-2}$	142.30			$++\ 6_{-6}$	446.66
$-+\ 3_{-3}$	136.74				

elements and so would demand that $++$ combine with $++$, etc.; this is not allowed. (The corresponding selection rule for the magnetic dipole moment of oxygen is that $+$ combines with $+$ and $-$ with $-$; consequently, there was a nonresonant absorption in O_2 which at very long wavelengths was more important than the resonant.) To have an appreciable contribution to the absorption it is necessary, therefore, that there be some small resonance frequencies, or, in other words, two closely spaced combining energy levels. Table 8·4 shows the energy levels of the water molecule, according to Dennison,[1] up to those of total rotational quantum number 6. The subscript τ specifies a second quantum number whose details need not be described here.

The table shows that there are many closely spaced energy levels. For instance, the energy difference between the states 3_3 and 3_2 is such as to correspond to an absorption maximum at a wavelength of 5 cm. Com-

[1] D. M. Dennison, *Rev. Mod. Phys.*, **12**, 189 (1940).

bination of these two levels is not allowed by the selection rules, however; this is also true, with one exception, for all the small intervals in the table. The exception is the pair 5_{-1}, 6_{-5}, which can combine. According to the table, the corresponding absorption maximum is at $\nu/c = 0.78$ cm^{-1} or $\lambda = 1.28$ cm. Unlike the situation in the optical region, however, where the wavelength is a quantity known with great accuracy, the wavelength of the transition in question is not determined with much accuracy from the infrared data. The trouble is, of course, that small differences between relatively large quantities (the excitation energies above the very lowest state) can be determined only if the energy levels themselves are known with extreme accuracy. According to an earlier table of energy levels by Randall, Dennison, Ginsburg, and Weber,[1] the estimated wavelength of the line in question is 1.46 cm. The difference 0.18 between this value and Dennison's estimate 1.28 is probably comparable to the error inherent in trying to work out the levels from the infrared data; for one thing, it is commensurate with the breadth of the line, as we shall see later. After the development of microwaves and radar had aroused interest in knowing the wavelengths of absorption resonances accurately, measurements of the absorption of water vapor as a function of frequency in the vicinity of 1 cm were made at Columbia Radiation Laboratory by a method to be described in Sec. 8·5; the wavelength of the transition 5_{-1}, -6_{-5} was found to be 1.344 cm.[2]

The intensity and breadth of the line must also be known. The intensity can be calculated from the known dimensions of the water molecule and from the quantum mechanics of the unsymmetrical top. The so-called "line strength" (square of the dipole moment summed over the various Zeeman components and hence inclusive of the statistical weight) is found to be 0.17 $(1.84 \times 10^{-18})^2$ in esu. The line-breadth constant $\Delta\nu$ has steadily diminished as time has progressed and the measurements have become more accurate. Early infrared work indicated that $\Delta\nu/c$ was in the vicinity of 0.3 cm^{-1}. In view of the importance of this quantity for microwave physics, improved measurements[3] of the breadth of water-vapor lines were made in the infrared laboratory at the University of Michigan. These experiments gave the result $\Delta\nu/c = 0.17$ and later 0.11 cm^{-1}. The breadth of the line can be measured far more accurately, however, with microwave than with infrared technique, inasmuch as in the microwave region the breadth is comparable to the

[1] H. M. Randall, D. M. Dennison, N. Ginsburg, and L. R. Weber, *Phys. Rev.*, **52,** 160 (1937).

[2] More recent work places this line at 1.3481 cm. See C. H. Townes and F. R. Merritt, *Phys. Rev.*, **70,** 558 (1946).

[3] A. Adel, NDRC Report No. 14—320, University of Michigan, Oct. 10, 1944, and especially the supplement thereto. His "half breadth" is the same as twice our quantity $\Delta\nu/c$.

frequency of the line itself. The recent microwave measurements at Columbia Radiation Laboratory (see Sec. 8·5 and Fig. 8·5) yield the shape of the line in detail and indicate that for low densities of water vapor, the best choice of the parameter $\Delta\nu/c$ is 0.087 cm^{-1}. The agreement with the Michigan data is within the latter's limit of error, which is estimated to be 0.02 cm^{-1}. In the calculation of the values of the ab-

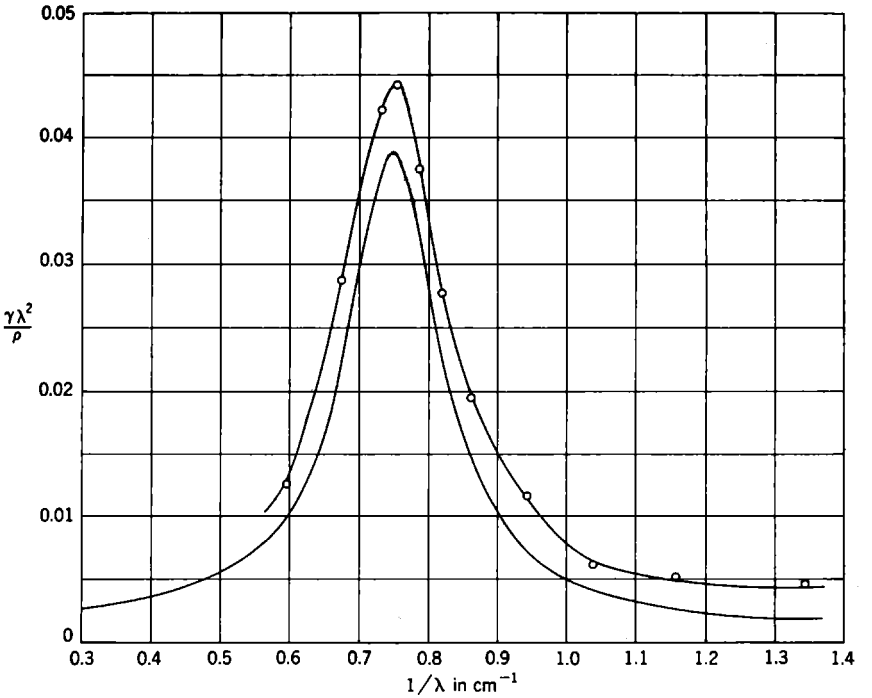


FIG. 8·5.—Theoretical and experimental results on attenuation by water vapor in the 1— cm region at a temperature of 45°C. The lower curve gives theoretical values based on $\Delta\nu/c = 0.087$ cm^{-1}. The upper curve represents measurements made at the Columbia Radiation Laboratory. The units used are: γ in decibels per kilometer, λ in cm, and ρ in grams per cubic meter.

sorption shown in Fig. 8·5, the round value $\Delta\nu/c = 0.1$ cm^{-1} will be used as a sort of mean of the Columbia and Michigan determinations. An interesting result of the Columbia experiments is that at atmospheric pressure the line breadth $\Delta\nu/c$ depends somewhat on the partial pressure of the water vapor. For instance, at a content of 50 g/m^3 (which could be obtained without condensation in the Columbia experiments because a temperature of 318°K was used) the value of $\Delta\nu/c$ is about 0.107, instead of 0.087 cm^{-1}. The dependence on the H_2O vapor pressure is such as to indicate that $H_2O - H_2O$ collisions have about five times as great a cross

section as the $H_2O - O_2$ or $H_2O - N_2$. This observation is another illustration of the wealth of information on collision broadening furnished by microwave spectroscopy. In applications to radar, the value of $\Delta\nu/c$ obtained by extrapolation to zero vapor pressure can be employed without appreciable error, as at ordinary temperatures the atmosphere is saturated at comparatively low densities (about 20 g/m³). In Fig. 8·5 the experimental values of the absorption are those per gram of water per cubic meter but extrapolated to zero vapor pressure so as to make the collision effects entirely those of water with air rather than with other water molecules.

The formula for the absorption in decibels per kilometer from the $5_{-1} - 6_{-5}$ line in water vapor at $T = 293°K$ is

$$\gamma = \frac{0.00350\rho}{\lambda^2}\left[\frac{\frac{\Delta\nu}{c}}{\left(\frac{1}{\lambda} - \frac{1}{1.35}\right)^2 + \left(\frac{\Delta\nu}{c}\right)^2} + \frac{\frac{\Delta\nu}{c}}{\left(\frac{1}{\lambda} + \frac{1}{1.35}\right)^2 + \left(\frac{\Delta\nu}{c}\right)^2}\right], \qquad (18)$$

where ρ is the number of grams of water per cubic meter and λ is in centimeters. Equation (18) is essentially a special case of Eq. (14), with the moment now electric rather than magnetic, of course. As we are dealing with only one line, there is only one pair of terms [combinations i,j and j,i in the summation over i,j in the numerator of Eq. (14)]. The calculation of the summation in the denominator involves the evaluation of the partition function which, in arriving at Eq. (18), has been computed numerically by summing over the various stationary states.

If we set $\Delta\nu/c = 0.1$ cm^{-1}, we obtain the values of the absorption given in Table 8·5.

TABLE 8·5.—ATTENUATION FROM THE LINE OF WATER VAPOR AT $\lambda = 1.35$ CM, IN DB/KM PER G OF H_2O/M³

λ, cm	γ/ρ	λ, cm	γ/ρ	λ, cm	γ/ρ
10	0.0000133	1.43	0.0147	1.0	0.0047
3.0	0.000256	1.35	0.0193	0.667	0.00150
2.0	0.00136	1.25	0.0168	0.5	0.00107
1.5	0.0101	1.11	0.0082	0.33	0.00085

Besides the absorption from the $5_{-1} - 6_{-5}$ transition, the absorption from the other lines should be taken into consideration. To evaluate their effect it is necessary to know the line strengths of the various transitions. These have been obtained from some theoretical calculation of King, Hainer, and Cross[1] for the lines of low rotational quantum number,

[1] G. W. King, R. M. Hainer, and P. C. Cross, *Phys. Rev.*, **71**, 433 (1947).

which are the most important ones, and by rough approximation for the higher levels. The resulting formula is

$$\gamma = \frac{0.012\rho \frac{\Delta\nu}{c}}{\lambda^2}. \tag{19}$$

The corresponding values of the absorption are given in Table 8·6 on the assumption that $\Delta\nu/c = 0.1$ cm^{-1}.

TABLE 8·6.—ATTENUATION FROM LINES OF WATER VAPOR OTHER THAN THAT AT 1.35 CM IN DB/KM PER G OF H_2O/M^3

λ, cm	γ/ρ	λ, cm	γ/ρ	λ, cm	γ/ρ
10	0.000012	1.35	0.00064	0.667	0.0026
3	0.00013	1.25	0.00074	0.5	0.0047
2	0.00029	1.0	0.0012	0.3	0.013
1.5	0.00052				

Comparison of Tables 8·5 and 8·6 shows that in the ordinary microwave region, the absorption from resonance with the $5_{-1} - 6_{-5}$ line far overshadows the combined effect of all the other lines.

The absorption maximum in the vicinity of $\lambda = 1.33$ cm from resonance with the $5_{-1} - 6_{-5}$ line is confirmed in measurements made at the Columbia Radiation Laboratory by a technique to be described in Sec. 8·5. The maximum absorption is somewhat larger than that given by theory, amounting at 1.35 cm to 0.03 db/km per g of H_2O per cubic meter instead of the theoretical value 0.0199. Instead of graphing the absorption itself, it is more instructive to graph the absorption multiplied by the square of the wavelength, as then the contribution of the nonresonant lines given by Eq. (19) becomes merely an additive constant independent of frequency. The comparison of the experimental curve with the theoretical one based on $\Delta\nu/c = 0.087$ is shown in Fig. 8·5.[1] The theoretical curve can be brought into fairly good agreement with the experimental one by raising the former by a constant amount. This

[1] The ordinates used in drawing the theoretical curve are about 10 per cent lower than the corresponding values in the tables. This change is to allow for the fact that in order to obtain higher vapor pressures, the experiments at the Columbia Radiation Laboratory were made at a temperature of 318°K rather than at room temperature. It can be shown that theoretically, for given vapor density, the absorption from the line at 1.33 cm should be proportional to $e^{\frac{-280}{T}}(1/T^2)$ in the vicinity of resonance, whereas the residual absorption in Eq. (19) should be proportional to $1/T^{3/2}$. These corrections allow for the fact that at constant total pressure the line-breadth constant $\Delta\nu$ should be proportional to $1/T^{1/2}$. The value $\Delta\nu/c = 0.087$ used in drawing Fig. 8·5 applies to 318°. The corresponding value at 293° is 0.091 cm^{-1}.

corresponds to assuming empirically a nonresonant absorption about four or five times the theoretical value. The explanation of this anomaly may possibly be that different lines have different line-breadth constants. This might be the case, for instance, if, because of resonance effects in energy exchange, the most potent collisions are those in which the colliding molecules merely exchange rotational states and hence both have nearly the same Boltzmann factors, inasmuch as the energy transfer is small compared with kT. Then molecules in the populous states have larger line breadths than those in less occupied ones, as they are more likely to collide with molecules of comparable energy. Thus the effective $\Delta\nu/c$ to be used in Eq. (19) might be larger than that involved in Eq. (18), the important lines that contribute to Eq. (19) involving more heavily populated states than the pair involved in the resonant absorption (18). It is doubtful, however, if the line-breadth constant varies enough from line to line to alter the calculated nonresonant absorption by a factor 4 or 5. Existing experimental evidence on line widths in general indicates that the variation with rotational quantum number is not this large.

More likely the reason that the observed nonresonant or background absorption is so much larger than that computed is that the calculation utilizes the Lorentz theory of line-broadening, which assumes that the collisions are of very short duration. This assumption works well in the center of a line, but on theoretical grounds may be expected to fail in the wings. In the wing region a more appropriate theory is probably the statistical one of Margenau and others. Hence, it may well be that the infrared absorption lines tail off less rapidly than one would expect on the basis of the Lorentz structure factor (11). The background absorption contributed by the infrared lines in the microwave region would then be greater than that yielded by the theory which was utilized in drawing Fig. 8·5.

In any case, the agreement between theory and experiment is good when it is considered that the measurements are made on absolute intensities, without any adjustable constant except, to a certain extent, the line-breadth parameter $\Delta\nu/c$. Absolute intensities in the infrared region have often been in error experimentally by a factor 10 or so, whereas the disagreement between theory and experiment as to the magnitude of the height of the resonance peak in Fig. 8·5 amounts to only a few per cent. The size of the theoretical matrix element for the $5_{-1} - 6_{-5}$ transition may be in error by about a fraction of a per cent, as a result of not allowing for centrifugal expansion. This correction is not one of consequence, however.

Figure 8·6 summarizes the results by showing the absorption by both oxygen and water as a function of the reciprocal of the wavelength. In order to have a basis of comparison, it is assumed that 1 per cent of the

molecules in the atmosphere are H_2O, corresponding to a content of 7.5 g of H_2O per cubic meter. In a temperate climate this is approximately the amount of water in an average day. When there is saturation, the content can be two or three times higher and even greater in the tropics. Figure 8·6 shows that oxygen has a resonance near $\frac{1}{4}$ cm, which has not as yet been investigated experimentally, as well as that at $\frac{1}{2}$ cm, which has been studied in detail. The resonance at $\frac{1}{4}$ cm is explained by the

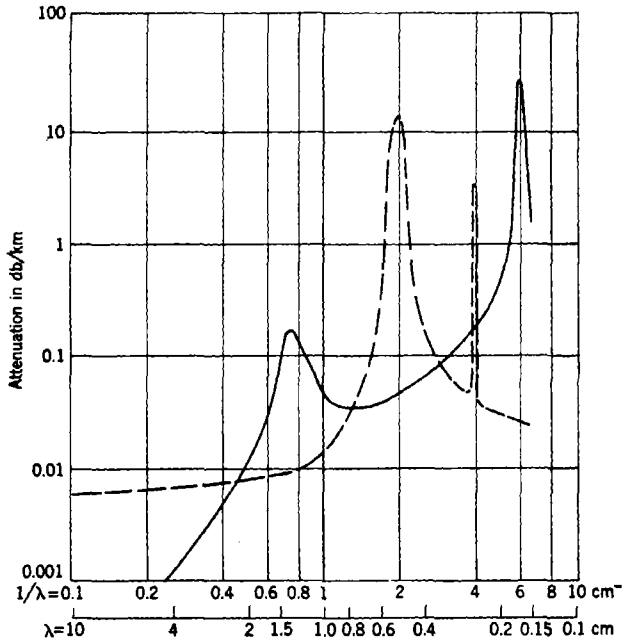


FIG. 8·6.—Theoretical values of atmospheric attenuation by oxygen and uncondensed water vapor at sea level for a temperature of 20°C. The solid curve gives the attenuation by water in an atmosphere containing 1 per cent water molecules ($\rho = 7.5$ g/m^3) for $\Delta\nu/c = 0.10$ cm^{-1}. The dashed curve is the attenuation by oxygen for $\Delta\nu/c = 0.02$ cm^{-1}.

fact that one of the doublets for $K = 1$ is radically different from that for other rotational states (*cf.* Table 8·3).

Figure 8·7 gives the extension of Fig. 8·6 to the millimeter region, where the effect of oxygen is negligible compared with that of H_2O and is consequently omitted. In Fig. 8·7 it is assumed that the line breadth $\Delta\nu/c$ is 0.11 cm^{-1}, the value indicated by the University of Michigan data. If, actually, the line breadth varies from line to line, the absorption may be larger than indicated by Fig. 8·7. This figure then loses quantitative significance but at least shows where the important resonances are located. It is to be emphasized that in drawing Fig. 8·7, Eq. (19) cannot be used,

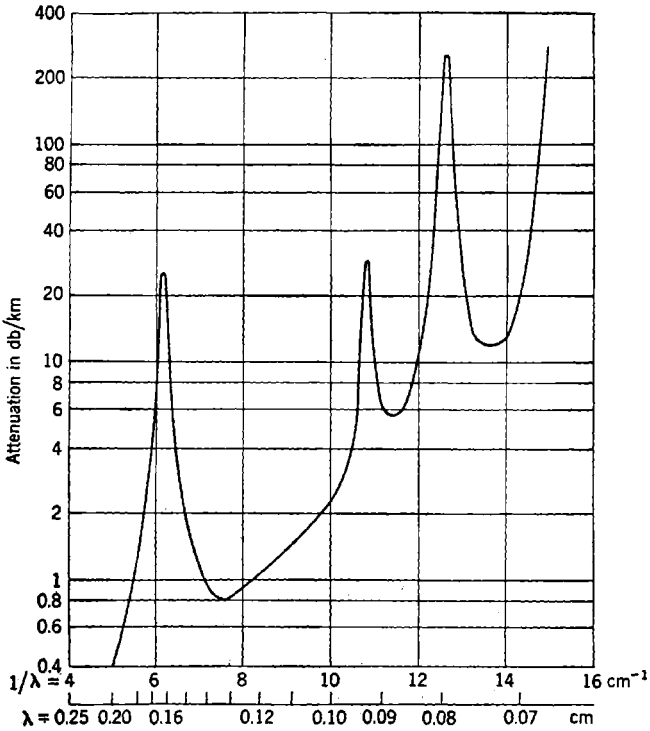


FIG. 8·7.—Theoretical values of atmospheric attenuation by uncondensed water vapor in the millimeter region for a temperature of 20°C. The curve is for an atmosphere containing 1 per cent water molecules ($\rho = 7.5$ g/m^3) for $\Delta\nu/c = 0.11$ cm^{-1}.

as one is too close to resonance with many of the lines to permit the assumptions $\nu \ll \nu_{ij}$ basic to Eq. (19). It will be noted that there are some very strong resonances in the millimeter region—notably the line at 0.164 cm.

MEASUREMENT OF ATMOSPHERIC ABSORPTION

By E. M. PURCELL

The theory of absorption of microwave energy by oxygen and by water vapor has been outlined in the preceding sections. It permits prediction of the attenuation in decibels per kilometer path length at any frequency, providing certain parameters in the theoretical formulas are known. The quantities required are the frequencies of the spectral lines responsible for the absorption and the collision widths of the lines under the prevailing conditions. The information derived from previous spectroscopic observations was neither sufficiently complete nor sufficiently accurate for the

present purpose. Moreover, until recently few direct measurements of microwave absorption had been made, and even in the optical and infrared region the opportunity for a direct comparison of theory and experiment as to the *absolute* intensity of an absorption line has been rare. For these reasons and because of the urgent need for quantitative information on the water-vapor and oxygen effects, direct measurements of microwave absorption have been undertaken by a number of methods, some of which will be described in this section.

The experiments are difficult because the absorption effects with which one is concerned are extremely weak, being in most cases measured in tenths or even hundredths of a decibel per kilometer. Consequently either one must work with very long paths, handicapped by field conditions and uncertain knowledge of conditions along the path; or if one uses a short "laboratory" path, he must employ a technique of the utmost refinement for measuring a small effect with precision.

8·4. Direct Measurement of Absorption by Oxygen.—The absorption by oxygen in the neighborhood of 0.5-cm wavelength was determined by measuring the attenuation in an oxygen-filled waveguide some 6 m long.[1] The attenuation was compared with that of the same waveguide evacuated, and the difference, corrected for the effect of the refractive index of the gas upon the guide wavelength and hence upon the guide attenuation, was ascribed to the oxygen absorption. The attenuation produced by the resistance of the waveguide walls was many times larger than the attenuation caused by the gas. Also, at the time of the experiments, no 0.5-cm oscillator was in existence, and the measurements had to be carried out using the low power available as the second harmonic from a crystal rectifier driven at 1 cm. The general scheme of measurement[2] is shown in Fig. 8·8.

Most of the apparatus appearing in the block diagram is associated with the 1000-cps modulation that was applied to the signal to facilitate amplification after detection. In the absence of gas the 1000-cps component of the detector output was balanced out, the introduction of gas then producing an unbalance signal. The waveguide was provided with gastight nonreflecting windows, and the system could be flushed or evacuated quickly. Methods of this type have the important advantage that the composition and pressure of the small amount of gas required can be easily controlled. The attenuation was measured as a function of wavelength and total gas pressure for various mixtures of oxygen and nitrogen, as well as for pure oxygen. A comparison of the experimental data with the theory is shown in Fig. 8·4. The interpretation of the

[1] E. R. Beringer, "The Absorption of One-half Centimeter Electromagnetic Waves in Oxygen," RL Report No. 684, Jan. 26, 1945.

[2] *Ibid.*

experimental results is complicated because the absorption arises not from a single line but from many closely spaced lines that are incompletely resolved at the gas pressures used. The absorption cross section per oxygen molecule, in mixtures of oxygen and nitrogen, was found to depend on the *total* gas pressure only. This important observation

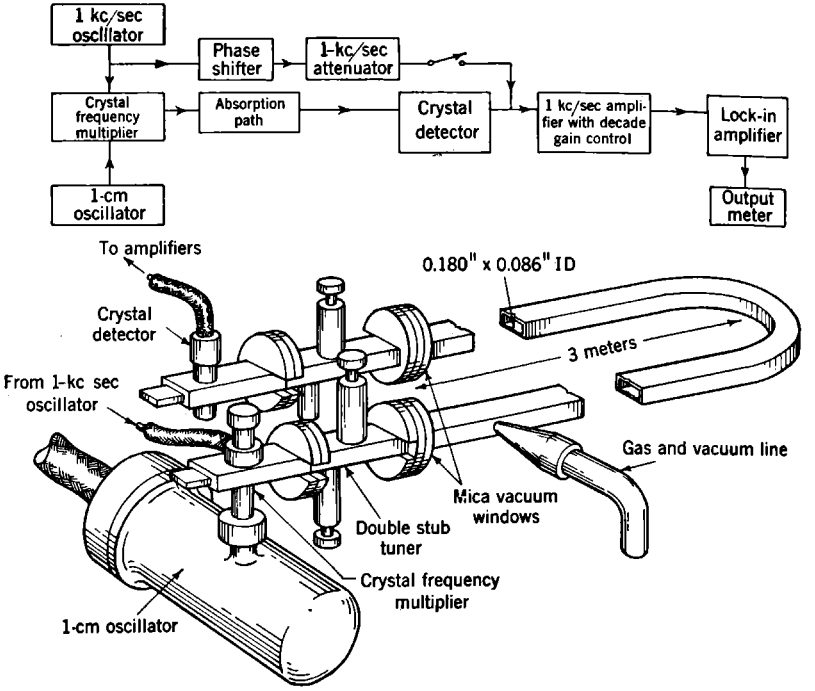


FIG. 8·8.—Arrangement of apparatus for the direct measurement of absorption by oxygen.

illustrates the advantage of the laboratory type of experiment in which the parameters can be varied systematically.

8·5. Measurements of Water-vapor Absorption.—The measurement of the attenuation by water vapor in the 1.0- to 1.5-cm range is still more difficult because the absorption is even smaller. It will be recalled that the attenuation anticipated is only a few tenths of decibel per kilometer, even at high absolute humidities. The attenuation in silver waveguide, at 1.25 cm, is, by comparison, some 500 db/km. Thus in a length of waveguide in which the resistive attenuation is not prohibitively high, the additional attenuation from water vapor may amount to a few thousandths of a decibel. If the water is introduced at high density as supersaturated steam, this disadvantage can be lessened. Such experiments have been performed, using a small resonant cavity rather than a

waveguide.[1] The conditions prevailing, however, were far from the conditions in the normal atmosphere—in particular the collision width is determined by $H_2O - H_2O$ collisions rather than $H_2O - O_2$ and $H_2O - N_2$ collisions—and the application of the result to microwave propagation is not direct.

The measurement of the attenuation under conditions typical of the atmosphere can be carried out in several ways if one is provided with a long path over which the attenuation from other effects is not too large. Three methods, vastly different in conception and technique, will be described.

The first method, though the least accurate, can from the point of view of microwave propagation be said to be the most direct, as it consists of determining the effect of atmospheric attenuation upon a radar signal.

Radar Methods.—An airborne radar set on a wavelength of 1.25 cm was furnished with means for introducing at will a fixed known amount of atten-

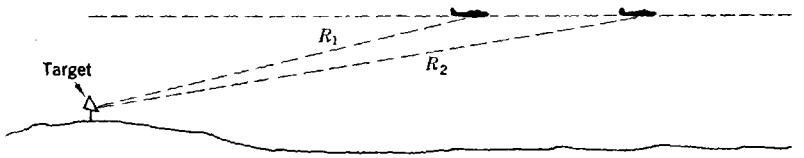


FIG. 8·9.—Use of a standard target for the measurement of attenuation in the atmosphere

uation (approximately 20 db) in the receiving system. A radar target of large and constant cross section was provided in the form of a corner reflector constructed with great attention to accuracy and solidity. The reflector was mounted on a site free from other conspicuous targets and was observed by the radar, which was flown back and forth on a course extending radially from the target (see Fig. 8·9). The maximum range on this target was determined in repeated flights, both with and without the additional attenuation in the radar receiver. If these ranges are R_1 and R_2 respectively, it is easy to show that the one-way atmospheric attenuation α averaged over the path is

$$\alpha = \frac{\frac{A}{2} - 20 \log_{10}\left(\frac{R_2}{R_1}\right)}{(R_2 - R_1)} \text{ db/km}, \tag{20}$$

where A is the value, in decibels, of the fixed attenuator inserted in the receiving channel. This conclusion is justified only if the signal comes entirely from the target, with no reflections from other surfaces, and if all

[1] J. A. Saxton, "Dielectric Constant and Absorption Coefficient of Water Vapor for Radiation of Wavelength of 1.6 cm (Frequency 18,800 Mc/sec)," RRBS No. 17, Apr. 22, 1944; also a chapter in *Meteorological Factors in Radio-wave Propagation*, The Physical Society, London, 1946.

other parameters affecting signal strength and signal discernibility are held constant. The radar-wise reader will at once think of the numerous precautions that are necessary (*cf.* Chaps. 4, 5, and 6). The nature of the target site and of the intervening terrain had to be considered, and the set itself was maintained in laboratory condition and provided with power monitors and special test instruments.

The absolute humidity over the path was obtained from soundings made either by radiosonde at a neighboring meteorological station or by means of a psychrometer in the plane itself (see Secs. 3·27 and 3·29). The attenuation measured includes the effect of oxygen, of course, as well as water vapor, but the latter was much the larger effect at this wavelength. In the hands of skilled observers the method gave convincingly consistent results which proved to be in satisfactory agreement with the values obtained later by the other methods to be described.

Microwave Radiometer.—A method that had no counterpart in previous r-f or microwave measurements was based on the microwave radiometer.[1] This device is in effect an extremely sensitive radiation thermometer, sensitive to thermal radiation in the microwave range and within a band of frequencies defined by the pass band of the microwave receiver which is the central part of the apparatus. By an ingenious automatic comparison scheme, very small changes in the noise input to this receiver are measured accurately; and if the receiver input is connected to a microwave antenna, the effective noise temperature of the antenna can be measured. This "antenna temperature" is determined by the amount of thermal radiation received by the antenna and hence by the temperature of any radiating medium toward which the antenna is directed. Because of the basic connection at any frequency between absorption and emissivity, if such an antenna is directed upward, the thermal radiation incident upon it will be that originating in the atmosphere itself, to the extent that the atmosphere is an *absorbing* medium; otherwise it will have come from some region outside the atmosphere. The following more precise statement is easily derived: If the atmosphere be assumed isothermal at absolute temperature T_a, if the effective radiation temperature of space beyond the atmosphere is T_s, and if the total absorption coefficient of the atmosphere is χ, the radiation temperature of the antenna will be

$$T_r = \chi T_a + (1 - \chi)T_s. \tag{21}$$

Now χ will depend on the length of path through the atmosphere and, for a horizontally homogeneous atmosphere, will thus depend in a calculable way upon the secant of the angle θ which this path makes with the vertical (see Fig. 8·10). Measurements of T_r at two values of θ will suffice to determine χ and T_s if T_a is known. The analysis is complicated, of course, if

[1] R. H. Dicke, "The Measurement of Thermal Radiation at Microwave Frequencies," RL Report No. 787, Aug. 22, 1945.

T_a varies with height, as it does, but the principle is the same. Actually measurements were made over a series of values of θ from 0° to 60°, and the expected variation of T_r with θ was accurately verified.

The temperature T_s is of considerable interest. If the radiation temperature of space were nearly the same as T_a, the determination of χ would, of course, be difficult. It was found, however, that T_s, at 1 cm, is at most a *few degrees* Kelvin. In addition to facilitating the determination of χ, we thus have the interesting information that the radiation coming from outer space in the 1-cm range is extremely small, in contrast to effects observed at much lower frequencies. Incidentally the experiment demonstrates the transparency of the ionosphere at microwave frequencies.

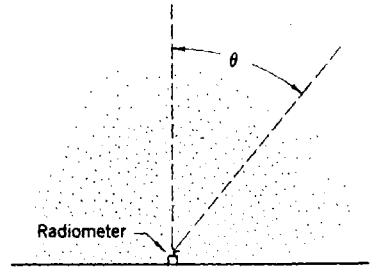


FIG. 8·10.—Radiometer arrangement to measure atmospheric absorption.

It would, of course, have been possible, in determining χ, to use the sun as a *hot* source rather than using the rest of space as a *cold* source. But the latter method is actually far superior, as it allows a series of measurements of temperature as a function of angle to be repeated in a short time, at any time, and in any azimuth plane.

The value of χ, once determined, has to be correlated with the absolute humidity, measured at various heights in the atmosphere. In an extensive series of measurements in Florida the humidity information was obtained from radiosondes and occasionally from aircraft soundings (see Secs. 3·27 and 3·29). These soundings and the allowances that had to be made for the variation with pressure and temperature of the water-vapor absorption are the principal sources of uncertainty in the final results. The radiometer measurement itself is capable of remarkable precision, temperature differences being measured with a sensitivity of a few tenths of a degree centigrade. It is to be expected that the method can be applied to measure the oxygen absorption, a case that is largely free from the above-mentioned difficulties. It should be possible to obtain an accurate measure of the absorption in the wings of the oxygen line. It is perhaps well to point out that conditions in the microwave region are highly favorable for this type of measurement because of the ideal *cold source* provided by space and because neither χ nor $(1 - \chi)$ is very small, for the magnitudes of attenuation of interest.

Nonresonant Echo Box.—A third method was developed and applied to the water-vapor problem at the Columbia Radiation Laboratory.[1] This method makes use of a long path also, but the path is obtained by multi-

[1] W. E. Lamb, Jr., *Phys. Rev.*, **70**, 308 (1946); G. E. Becker and S. H. Autler, *ibid.* **70**, 300 (1946).

ple reflections within a large copper-lined chamber, and the atmosphere traversed is therefore controllable. On the other hand the ratio of the power lost in the walls to that absorbed in the gas is reduced over the value typical of a waveguide method by a factor that is roughly λ/d, where d is the length of one side of the chamber. In the Columbia experiments this factor is of the order of 10^{-3}. Thus the disadvantages of both the waveguide methods and the free-space methods are to a large extent avoided. It had often been suggested that attenuation could be measured using such an echo chamber by determining the rate of decay of energy in the chamber after initial excitation. (This suggests some similar experiments in acoustics.) The success of the Columbia method, however, can be traced to certain important new techniques.

The Columbia "box" was a cube approximately 2.5 m on a side. Such a large resonator displays an enormous number of resonance modes, which in general overlap in frequency owing to the broadening of the resonances by the wall losses. The object is to excite a large number of these modes by a short pulse of microwave energy and then to measure the rate of subsequent decay of energy; the energy density averaged over many modes should decrease, after removal of the excitation according to a time factor $e^{-\gamma t}$. The quantity γ should depend both on the losses in the wall and on the attenuation in the gas within the chamber. A determination of γ with the gas present and absent would enable the attenuation in the gas, in decibels per kilometer, to be inferred. A direct determination of γ, however, is fraught with serious difficulties. Even though a large number of modes are excited, the observed decay does not follow a simple exponential curve but oscillates violently and irregularly above and below it, owing to the interference of the different modes at the location of any single pickup probe. This difficulty was avoided by measuring instead a quantity proportional to the time average of the energy density at a point in the chamber averaged in turn over many such points. This was done by distributing through the chamber chains of thermocouples, whose "hot" junction had been coated with material absorbent at microwave frequencies. All thermocouples were connected in series, the total voltage developed being measured using a potentiometer and a sensitive galvanometer. The chamber was excited by a pulsed magnetron at a convenient repetition rate. During each pulse the microwave energy in the box increases rapidly to some maximum value and is subsequently dissipated, at a much slower rate, in the walls and in any other absorbing medium within the box. The rate of dissipation is determined by the Q of the box, as in the case of a simple resonant circuit. It is not hard to see that the thermocouple voltage e_t observed is then related to Q by an as yet unknown factor of proportionality which will, however, remain constant as long as the intensity of excitation remains constant. In other words $e_t = kQ$, where Q is defined in the usual way as ω times the ratio of energy stored to power dissipated. The

relative changes in Q produced by the introduction of an absorbing gas can thus be determined, but this determination alone does not suffice to give the attenuation in absolute terms; additional information is needed from which the Q of the chamber can be found, that is, from which k can be determined. This additional information was obtained by introducing a known loss in the form of a window in one wall. The chamber was so large that it was possible to have a window that was at once small compared to the extent of the wall but large compared with λ; in this case it was assumed and verified by experiment that the power lost through such a window is just proportional to its area, and therefore it can be taken to be equal simply to the total power incident on this area from all directions. This allows the chamber to be "calibrated." The absolute humidity was determined by wet-dry bulb readings and was controllable over a wide range.

Using a series of specially constructed magnetrons, the Columbia group extended their measurements over the range 0.75 to 1.7 cm. The results, shown in Fig. 8·5, undoubtedly represent the most accurate measurements to date on the water-vapor line at 1.3 cm. The usefulness of the method, in general, is limited mainly by the requirement of a high-power pulsed source, a rather large quantity of the gas to be investigated, and the restriction on the temperatures and pressures that the chamber can withstand.

ATTENUATION BY CONDENSED WATER

By Herbert Goldstein

8·6. Phenomenology of Attenuation by Precipitation.— In passing through regions containing precipitation, whether in liquid or solid form, microwave radiation will be attenuated through two mechanisms: (1) energy will be scattered out of the beam; (2) energy will be absorbed and converted into heat. The total energy removed from the beam per unit time by a single particle, say a spherical drop of radius a, is given by

$$W = SQ(a,\lambda), \tag{22}$$

where S is the magnitude of the incident Poynting vector and $Q(a,\lambda)$ is a quantity called the total cross section of the drop (*cf.* Sec. 6·1 and Stratton[1]). The total cross section can be interpreted as the sum of a scattering cross section Q_s and an absorption cross section Q_a. From Eq. (22) the decrease in S in passing through a layer of precipitation dl in thickness is

$$-dS = S\,dl \int_0^\infty n(a)\,Q(a,\lambda)\,da, \tag{23}$$

where $n(a)\,da$ gives the number of drops per unit volume (cubic meter) with

[1] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 569.

radius a in range da.[1] Equation (23) can be integrated immediately, leading to the attenuation formula

$$S = S_0 e^{-\int \alpha \, dl}, \tag{24}$$

where

$$\alpha = \int_0^\infty n(a) \, Q(a,\lambda) \, da. \tag{25}$$

In Eq. (25) and elsewhere in the following sections, n is given in number per cubic meter, Q in square centimeters, and a in centimeters. It is therefore more convenient to give the attenuation[2] in terms of a quantity γ in units of decibels per kilometer, related to α by

$$\gamma = 434\alpha \qquad \text{db/km}. \tag{26}$$

Common experience indicates that precipitation is never uniform over an extended region. Detailed experiments, described in Sec. 8·9, have revealed large differences in rainfall at points only a few hundred feet apart. In general, therefore, $n(a)$ and hence α and γ are functions of the distance l along the path. Only in exceptional conditions of uniform precipitation is it allowable to write the exponent in Eq. (24) simply as $-\alpha l$.

It is seen that the attenuation can be calculated rigorously only if the drop-size distribution is known. The techniques for determining this distribution are at present crude and cumbersome (Sec. 7·7), and the published data are open to question. As a result both of the marked spatial nonuniformity of the precipitation and of the difficulty in measuring the drop-size distribution, it is clear that exact calculations of attenuation to be expected for any given practical path are not feasible. Although the purely electromagnetic side of the theory is adequately known, the necessary meteorological data would be prohibitively detailed.

It would therefore be highly desirable to relate the attenuation to some meteorological parameter more conveniently measurable than the drop-size distribution. When the drops are very small, as they are in clouds and fog, it turns out that the attenuation is directly proportional to the total water content per unit volume, independent of drop-size distribution. For with small drops, the scattering cross section is proportional to

$$Q_s \propto \lambda^2 \left(\frac{2\pi a}{\lambda}\right)^6, \tag{27}$$

[1] Although this equation and subsequent discussion are specifically for drops, it is obvious that the same general formulas hold for other precipitation forms such as ice, snow, etc., with a as a generalized parameter of the particle dimensions.

[2] All attenuations given here are for "one-way" transmission.

that is, Rayleigh scattering. The absorption cross section on the other hand is proportional to

$$Q_a \propto \lambda^2 \left(\frac{2\pi a}{\lambda}\right)^3. \tag{28}$$

For sufficiently small drops, therefore, the attenuation is caused almost entirely by absorption and is directly proportional to a^3, or the total volume of the drops, per unit volume of the atmosphere. These conditions certainly hold for clouds or fogs at all wavelengths considered here. Even for raindrops this conclusion is valid for all wavelengths above the microwave region and to some extent for wavelengths of a few centimeters.

However, the water content per unit volume is still not a conveniently measurable parameter of precipitation. In fact the only such parameter is the rate of precipitation, that is, the volume of water reaching the ground per unit time. It has become customary to assume a linear relation between attenuation and precipitation rate and to report the attenuation in decibels per kilometer per millimeter per hour. This practice is to some extent supported by experimental results (see Sec. 8·9). It should be emphasized, though, that a priori there is no theoretical reason for expecting a linear relation between the two quantities. This is clearly shown by comparing Eq. (25) with the expression for the total precipitation rate

$$p = 15.1 \int_0^\infty n(a)\, v(a)\, a^3\, da, \tag{29}$$

where $v(a)$ is the terminal velocity of the drops in meters per second and p is in millimeters per hour. In general α is not necessarily proportional to p. Unless a given precipitation rate corresponds always to a definite drop-size distribution and vice versa, α is not even a single-valued function of p. All that can be said is that the results should cover a region in the α-p plane.

It has been pointed out by L. Goldstein[1] that it is occasionally possible to delimit the region of the α-p plane covered by the results. Let us define a function

$$g(a,\lambda) = \frac{Q(a,\lambda)}{15.1\, v(a)\, Q^3}. \tag{30}$$

At constant wavelength g is a function of drop size only and at most wavelengths has clearly defined maximum and minimum values, $g_2(\lambda)$ and $g_1(\lambda)$ respectively, within the drop-size range involved. Then

$$15.1 g_1 \int_0^\infty n(a)\, v(a)\, a^3\, da \leqslant \alpha \leqslant 15.1 g_2 \int_0^\infty n(a)\, v(a)\, a^3\, da, \tag{31}$$

or

$$g_1 p \leqslant \alpha \leqslant g_2 p. \tag{32}$$

[1] L. Goldstein, "Absorption and Scattering of Microwaves by the Atmosphere," Wave Propagation Group, Columbia University Division of War Research, Report No. WPG-11, May 1945.

That is, for given wavelength, the values of α should lie in the α-p plane in the region bounded by two straight lines $\alpha = g_1 p$ and $\alpha = g_2 p$. Theoretical values of g_1 and g_2 for a number of wavelengths will be given in Sec. 8·7. It turns out that at 1.25 and 3 cm, g varies little over the region of drop radius expected to cover the largest number of drops, thus explaining in part the frequently observed linear dependence of α on p.

8·7. Calculation of Attenuation by Water Drops.—The fundamental calculation of the scattering and absorption of electromagnetic waves by a dielectric sphere is due to Mie[1]. In Sec. 6·1 the method of calculation and the most important results have been outlined briefly, following the presentation by Stratton. Extensive calculations of attenuation, based on Mie's results, have been carried out by Ryde and Ryde.[2] Almost all the quantitative data presented in this section have been taken from their report.

The total cross section has been given in Sec. 6·1 as[3, 4]

$$Q = -\frac{\lambda^2}{2\pi} \operatorname{Re} \sum_{n=1}^{\infty} (2n+1)(a_n^s + b_n^s), \tag{33}$$

where λ is the free-space wavelength and a_n^s and b_n^s are given by Eq. (6·18) in Sec. 6·1.

To evaluate these expressions, the complex index of refraction of the water must be known for microwave frequencies. This quantity n_c may be given in terms of the real part of the index of refraction n and the imaginary part $n\chi$ by

$$n_c = n(1 - i\chi), \tag{34}$$

or in terms of the complex dielectric constant

$$n_c^2 = \varepsilon_c = \varepsilon_1 - i\varepsilon_2. \tag{35}$$

The real and imaginary parts of n_c and ε_c are related by

$$\varepsilon_1 = n^2(1 - \chi^2), \quad \varepsilon_2 = 2n^2\chi. \tag{36}$$

The values used by Ryde and Ryde are those given by Saxton,[5] who com-

[1] G. Mie, *Ann. Physik*, **25**, 377 (1908).

[2] J. W. Ryde and D. Ryde, "Attenuation of Centimetre and Millimetre Waves by Rain, Hail, Fogs and Clouds," GEC Report No. 8670, May 1945. See also the earlier reports GEC No. 7831, October 1941 (J. W. Ryde), and GEC No. 8516, August 1944 (Ryde and Ryde), where less accurate values of the refractive index were used.

[3] Also J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 569, Eq. (29), as corrected by L. Goldstein, *op. cit.*

[4] *Note added in proof:* Since the foregoing material was written, the total cross section has also been computed at the National Bureau of Standards. See *Tables of Scattering Functions for Spherical Particles*, Applied Mathematics Series, No. 4, Jan. 25, 1949.

[5] See papers by J. A. Saxton and J. A. Lane in *Meteorological Factors in Radio-wave Propagation*, The Physical Society, London, 1946.

bined his measurements in the neighborhood of 1 cm with the results of previous investigators at longer wavelengths. It was found that the data could be well fitted by curves plotted from the Debye formula [*cf.* Eq. (8·5)]:

$$\epsilon_c = \frac{\epsilon_0 - \epsilon_\infty}{1 + \frac{i\Delta\lambda}{\lambda}} + \epsilon_\infty, \tag{37}$$

where ϵ_0, ϵ_∞, and $\Delta\lambda$ are constants adjusted to fit the experimental points.

Table 8·7 lists the index of refraction obtained from these curves for several wavelengths and for the temperature range from 0° to 40°C. The constants to be used in the Debye formula are presented in Table 8·8. These curves represent the experimental data in the region from 1 to 10 cm so well that it is believed that they may be used for extrapolation into the millimeter region.

TABLE 8·7.—COMPLEX INDEX OF REFRACTION OF WATER*

T, °C	λ = 1.24 cm		λ = 3.2 cm		λ = 10 cm	
	n	$n\chi$	n	$n\chi$	n	$n\chi$
0	4.68	2.73	7.10	2.89	8.99	1.47
10	5.74	2.92	8.00	2.33	9.02	0.90
18	6.4	2.8	8.30	1.90	8.9	0.69
20	6.53	2.77	8.33	1.72	8.88	0.63
30	7.10	2.48	8.39	1.31	8.71	0.45
40	7.47	2.11	8.35	1.02	8.33	0.36

* Ryde and Ryde, after Saxton.

TABLE 8·8.—CONSTANTS FOR THE DEBYE FORMULA*

T, °C	ϵ_0	ϵ_∞	$\Delta\lambda$, cm
0	88	5.5	3.59
10	84	5.5	2.24
18	81	5.5	1.66
20	80	5.5	1.53
30	76.4	5.5	0.112
40	73	5.5	0.0859

* Ryde and Ryde, after Saxton.

The complicated expressions for the coefficients in Eq. (33) admit of some simplification if the circumference of the drop is small compared with the wavelength. The coefficients b_n^s may be identified formally with excitation of *electric* dipole, quadrupole, and higher-order multipole oscillations, and in a similar manner the a_n^s can be identified with *magnetic* dipole, quad-

rupole, and higher-order oscillations. When the dimensions of the drop are small compared with the exciting wavelength, it is reasonable to expect that the higher multipole oscillations will be only feebly excited. As a first approximation it therefore seems reasonable to neglect all coefficients except b_1^s and b_2^s (electric dipole and quadrupole respectively) and a_1^s for the magnetic dipole. Furthermore, the Bessel functions appearing in these coefficients have as argument either $\rho(=2\pi a/\lambda)$ or $n_c\rho$. As it has been assumed that ρ is small compared with unity, these Bessel functions may be expanded in ascending powers of ρ. If terms in ρ^7 or higher are neglected, the resultant expression for Q is

$$Q = \frac{\lambda^2}{2\pi}\rho^3(c_1 + c_2\rho^2 + c_3\rho^3), \qquad \rho \ll 1, \tag{38}$$

where

$$c_1 = \frac{6\varepsilon_2}{(\varepsilon_1+2)^2+\varepsilon_2^2},$$

$$c_2 = \frac{\varepsilon_2}{15}\left[\frac{3(7\varepsilon_1^2+4\varepsilon_1-20+7\varepsilon_2^2)}{[(\varepsilon_1+2)^2+\varepsilon_2^2]^2} + \frac{25}{(2\varepsilon_1+3)^2+4\varepsilon_2^2} + 1\right],$$

$$c_3 = \frac{4}{3}\frac{(\varepsilon_1-1)^2(\varepsilon_1+2)^2+\varepsilon_2^2[2(\varepsilon_1-1)(\varepsilon_1+2)-9]+\varepsilon_2^4}{[(\varepsilon_1+2)^2+\varepsilon_2^2]^2}.$$

It will be noticed that both c_1 and c_2 vanish if $\varepsilon_2 = 0$, indicating that these terms are caused essentially by absorption. On the other hand if $\varepsilon_2 = 0$, c_3 reduces to the familiar Rayleigh scattering form,

$$c_3 = \frac{4}{3}\frac{(\varepsilon_1-1)^2}{(\varepsilon_1+2)^2}. \tag{39}$$

For water ε_2 is not zero; in fact it is the same order of magnitude as ε_1. Hence for sufficiently small ρ the absorption term c_1 must predominate and Q is given simply by

$$Q = \frac{4\pi^2}{\lambda}c_1 a^3 = 4.093\frac{M}{\lambda}c_1, \tag{40}$$

where M is the mass of condensed water per unit volume of air, in grams per cubic meter, and λ, as always in these sections, is in centimeters. From Eq. (40) and with the above values of index of refraction at 18°C, Ryde and Ryde have computed the attenuation γ, for $M = 1$ g/m^3. Some of the results are given in Table 8·9.

Empirically it is found that in the region from $\lambda = 0.5$ cm to $\lambda = 10$ cm c_1 varies as $1/\lambda$ and that γ can be written as

$$\gamma = \frac{0.438M}{\lambda^2}\ \text{db/km} \tag{41}$$

to an accuracy of 5 per cent.

TABLE 8·9.—ATTENUATION BY SMALL DROPLETS OF WATER FOR A TEMPERATURE OF 18°C*

λ, cm	γ/M, db/km per g per m³
0.2	7.14
0.5	1.65
0.7	0.876
1.0	0.438
1.25	0.280
2.0	0.112
3.0	0.050
5.0	0.0178
10.0	0.0045

* After Ryde and Ryde.

Table 8·9 and Eq. (41) are based on the values of the refractive index at 18°C. The correction for other temperatures is not always negligible, as is indicated by Table 8·10, which lists the correction factor $\phi_1(T)$ defined such that

$$c_1(T) = \phi_1(T)\, c_1(18°).$$

TABLE 8·10.—CORRECTION FACTOR $\phi_1(T)$ FOR COEFFICIENT c_1*

λ, cm	0°C	10°C	18°C	20°C	30°C	40°C
0.5	1.59	1.20	1.0	0.95	0.73	0.59
1.25	1.93	1.29	1.0	0.95	0.73	0.57
3.2	1.98	1.30	1.0	0.95	0.70	0.56
10.0	2.0	1.25	1.0	0.95	0.67	0.59

* After Ryde and Ryde.

It is seen that from 0° to 40°C the attenuation decreases by more than a factor of 3. This large temperature effect is caused by the large change of the line width ($\Delta\lambda$ in the Debye formula) with temperature, which affects chiefly the absorption ε_2. When terms other than c_1 enter into the attenuation, the temperature effect is not nearly so great.

The approximation of the attenuation solely by the absorption term c_1 of the electric dipole contribution is valid only for very small drop radii. Although the upper bound on ρ for which the approximation is valid is not easy to determine, the formulas certainly hold for fog and for clouds in which the droplets are extremely small, with diameters of the order of 0.001 to 0.005 cm. Tables 8·9 and 8·10 and Eqs. (40) and (41) may safely be used in these cases. It should be mentioned that a value of $M = 1$ g/m³ represents about the maximum water content of fogs, with the possible exception of heavy sea fogs. In most cases M is much less than 1. Ryde

and Ryde give an empirical relation between an average $\overline{M}$ and optical visibility in the fog:

$$\overline{M} = 1660D^{-1.43},$$

where D is the optical visibility in feet and $\overline{M}$ is such that in 95 per cent of the cases M lies between $\frac{1}{2}M$ and $2M$. Such a relation may prove helpful when more definite information on M is lacking.

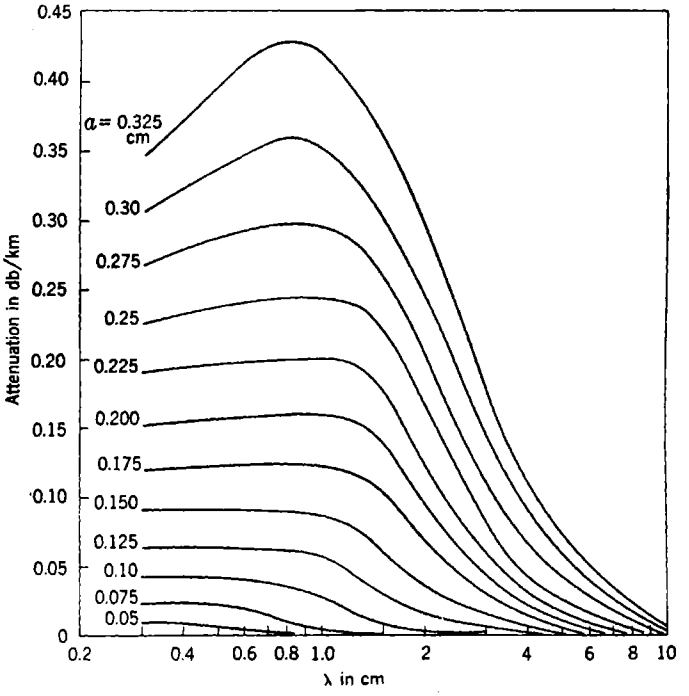


FIG. 8·11.—Theoretical values of attenuation by raindrops for various drop radii, expressed in decibels per kilometer per drop per cubic meter. (*From Ryde and Ryde.*)

For droplets larger than those encountered in fog or clouds, the c_1 term fails to be a good approximation to the attenuation. The whole three-term expansion [Eq. (38)] fails to hold for $\rho > 0.1$. Hence for drop sizes such as encountered in actual rain it is necessary to return to the original expression for Q [Eq. (33)] and use the rigorous form of the coefficients. Ryde and Ryde have developed a computation scheme for evaluating the real and imaginary parts of the coefficients that is discussed briefly in Sec. 7·6. For a 1 per cent accuracy it is stated that terms up to $n = 10$ must be retained for $\rho = 6$ and up to $n = 15$ for $\rho = 10$.

The results of these laborious computations are summarized in Table 8·11, where $Q(a,\lambda)$ is listed for λ from 0.3 to 10 cm and a from 0.025 to 0.325 cm. Actually what is listed is 434 times Q so that the table gives the atten-

uation in decibels per kilometer per drop per cubic meter. The data are also presented in Fig. 8·11. Although the attenuation for all drop radii rises sharply with increasing frequency at longer wavelengths, the behavior is quite different for wavelengths of 1.0 cm and shorter. For the larger-drop radii the attenuation curves even exhibit maxima in the neighborhood of $\lambda = 0.8$ cm, with decreasing attenuation for shorter wavelengths. This variation indicates the great degree to which the simple $1/\lambda^2$ dependence for clouds and fog fails to hold with the larger raindrops.

TABLE 8·11.—ATTENUATION IN DECIBELS PER KILOMETER PER DROP PER CUBIC METER AT A TEMPERATURE OF 18°C*

a_{cm} \ λ_{cm}	$434Q(a,\lambda)$								
	0.3	0.4	0.5	0.6	1.0	1.25	3.0	3.2	10.0
0.025	$8.34\cdot10^{-4}$	$5.80\cdot10^{-4}$	$3.46\cdot10^{-4}$	$1.74\cdot10^{-4}$	$5.62\cdot10^{-5}$	$2.71\cdot10^{-5}$	$4.21\cdot10^{-6}$	$3.55\cdot10^{-6}$	$2.94\cdot10^{-7}$
0.050	$9.02\cdot10^{-3}$	$6.80\cdot10^{-3}$	$4.67\cdot10^{-3}$	$3.10\cdot10^{-3}$	$1.00\cdot10^{-3}$	$6.25\cdot10^{-4}$	$5.05\cdot10^{-5}$	$4.77\cdot10^{-6}$	$2.53\ 10^{-6}$
0.075	$2.18\cdot10^{-2}$	$2.12\cdot10^{-2}$	$2.04\cdot10^{-2}$	$1.85\cdot10^{-2}$	6.91	$4.19\cdot10^{-3}$	$4.16\cdot10^{-4}$	$3.24\cdot10^{-4}$	9.39
0.100	4.07	4.03	3.97	3.85	$2.76\cdot10^{-2}$	$1.37\cdot10^{-2}$	$2.21\cdot10^{-3}$	$1.70.10^{-3}$	$2.49\cdot10^{-5}$
0.125	6.35	6.31	6.22	6.15	5.53	4.00	6.98	5.75	5.61
0.150	8.90	8.87	8.81	8.79	8.50	7.60	$1.56\cdot10^{-2}$	$1.40\cdot10^{-2}$	$1.13\cdot10^{-4}$
0.175	$1.18\cdot10^{-1}$	$1.19\cdot10^{-1}$	$1.21\cdot10^{-1}$	$1.22\cdot10^{-1}$	$1.21\cdot10^{-1}$	$1.15\cdot10^{-1}$	2.76	2.54	2.09
0.200	1.52	1.53	1.56	1.57	1.58	1.55	4.35	3.98	3.65
0.225	1.90	1.92	1.94	1.95	1.98	1.95	6.35	5.75	6.10
0.250	2.26	2.32	2.35	2.39	2.42	2.38	8.68	7.94	9.76
0.275	2.68	2.77	2.85	2.91	2.94	2.84	$1.13\cdot10^{-1}$	$1.05\cdot10^{-1}$	$1.51\cdot10^{-2}$
0.300	3.07	3.22	3.37	3.48	3.53	3.30	1.43	1.38	2.24
0.325	3.46	3.75	3.97	4.15	4.20	3.89	1.74	1.73	3.16

* After Ryde and Ryde.

The results given in Table 8·11 or Fig. 8·11 are still not in a very useful form. To facilitate comparisons with the usual type of experimental results definite statements must be made about the drop velocities and size distributions. Table 8·12 gives the terminal velocity relation obtained by Best,[1] as quoted by Ryde and Ryde. The values differ somewhat from

TABLE 8·12.—RAINDROP TERMINAL VELOCITY*

Radius, cm	Velocity, m/sec	Radius, cm	Velocity, m/sec
0.025	2.1	0.175	8.35
0.05	3.9	0.020	8.70
0.075	5.3	0.225	9.0
0.10	6.4	0.25	9.2
0.125	7.3	0.275	9.35
0.15	7.9	0.30	9.5
		0.325	9.6

* After Best.

[1] A. C. Best, "Water in the Atmosphere," Part I of the Interim Report of the Ultra Short Wave Panel Working Committee, July 18, 1944.

the figures given by Humphreys,[1] which are not so extensive. By dividing the attenuation per drop by the terminal velocity one can obtain the function g defined by Eq. (30) and set upper and lower bounds on the attenuation, as explained in Sec. 8·6. Table 8·13 lists these bounds, in decibels per kilometer per millimeter per hour at 18°C as obtained from Ryde and Ryde's data for two ranges of drop sizes. Also listed are the radii corresponding to these bounds. The concept of such limits is most useful only when the function g possesses well-defined extrema in the region of drop

TABLE 8·13.—UPPER AND LOWER LIMITS ON ATTENUATION, IN DECIBELS PER KILOMETER PER MILLIMETER PER HOUR, AT 18°C

λ, cm	For all rains with drop radii 0.025 cm $\leq a \leq$ 0.325 cm				For all rains with drop radii 0.075 cm $\leq a \leq$ 0.25 cm			
	$434g_1$	a_1, cm	$434g_2$	a_2, cm	$434g_1$	a_1, cm	$434g_2$	a_2, cm
10.0	0.000254	0.10	0.000635	0.325	0.000254	0.10	0.000450	0.25
3.0	0.00687	0.05	0.0415	0.20	0.0123	0.075	0.0415	0.20
1.25	0.0550	0.025	0.188	0.15	0.124	0.075	0.188	0.15
1.0	0.0845	0.325	0.285	0.10	0.112	0.25	0.285	0.10
0.6	0.0835	0.325	0.547	0.075	0.110	0.25	0.547	0.075
0.5	0.0798	0.325	.....	.....	0.108	0.25	0.604	0.075
0.3	0.0695	0.325	.....	.....	0.104	0.25	0.645	0.075

radii of interest. For wavelengths shorter than 0.6 cm, however, g appears to increase monotonically as the drop radius decreases, and therefore no upper bound has been listed. Also, for many wavelengths the lower bound corresponds to the extremely large drop radii which occur only infrequently in actual rains. In the second part of the table the bounds have been recalculated for a more restricted range of drop radius, from 0.075 to 0.25 cm, which probably includes most of the drops found in heavier rains (see Table 8·14). It is noteworthy that with this range, at a wavelength of 1.25 cm the upper and lower bounds are quite close together. At this wavelength the attenuation should therefore be substantially independent of drop-size distribution, at least for heavier rains. To a somewhat lesser extent the same behavior is to be expected at 3.0 and 10 cm.

The lack of sensitivity of the attenuation at these wavelengths to the drop size is fortunate, as reliable and extensive information on drop-size distributions is not available at present. What is generally conceded to be the best data so far have been obtained by Laws and Parsons.[2] The results are experimentally obtained, not in the form of the distribution $n(a)$, but as a related distribution $m(a)$, the fraction of the total volume of

[1] W. J. Humphreys, *Physics of the Air*, McGraw-Hill, New York, 1940.

[2] J. O. Laws and D. A. Parsons, "The Relation of Drop Size to Intensity," *Trans. Am. Geophys. Union*, p. 452, (1943).

water striking the ground due to drops of radius a. The two distributions are connected by

$$m(a)\,da = \frac{n(a)\,\frac{4\pi a^3}{3}\,v(a)\,da}{\int_0^\infty n(a)\,\frac{4\pi a^3}{3}\,v(a)\,da} = \frac{15.1n(a)v(a)\,da}{p}, \tag{42}$$

from Eq. (29). With the definition of g as given in Eq. (30) it is seen that the attenuation can be calculated as

$$\gamma = 434p \int_0^\infty m(a)g(a,\lambda)\,da, \text{ db/km}. \tag{43}$$

Table 8·14 presents some of the $m(a)$ distributions obtained by Laws and Parsons, as quoted by Ryde and Ryde. The interval of radius da is 0.025 cm except for the lowest interval, which extends from 0.0 to 0.0375 cm. In all other cases the a listed in the table is the mid-point of the interval. It appears from the data that the mean drop radius increases with the total rate of precipitation, at least for the group of rains measured by Laws and Parsons. In the light of the behavior of the function $g(a,\lambda)$ as obtained from Table 8·9 one can therefore predict that the attenuation per millimeter per hour for 3 cm and less would be greater for lighter rains than for heavy rains.

From these distributions and the cross section $Q(a,\lambda)$ or the derived function $g(a,\lambda)$, the total attenuation may be computed. (In practice the

TABLE 8·14.—FRACTION OF TOTAL VOLUME REACHING GROUND CONTRIBUTED BY DROPS OF VARIOUS SIZES*
(Drop radius interval, da = 0.025 cm)

p, mm/hr	0.25	1.25	2.5	12.5	25	50	100	150
a, cm	$m(a)\,da$							
0.025	0.28	0.11	0.07	0.03	0.02	0.01	0.01	0.01
0.050	0.50	0.37	0.28	0.12	0.08	0.05	0.04	0.04
0.075	0.18	0.31	0.33	0.25	0.18	0.12	0.09	0.07
0.100	0.03	0.14	0.19	0.25	0.24	0.20	0.14	0.12
0.125	0.01	0.05	0.08	0.17	0.20	0.21	0.17	0.14
0.150	.....	0.02	0.03	0.10	0.13	0.16	0.18	0.18
0.175	.....	0.01	0.01	0.04	0.08	0.11	0.15	0.16
0.200	.....	.....	0.01	0.02	0.03	0.07	0.09	0.12
0.225	.....	.....	.....	0.01	0.02	0.03	0.06	0.08
0.250	.....	.....	.....	0.01	0.01	0.02	0.03	0.04
0.275	.....	.....	.....	.....	0.01	0.01	0.02	0.02
0.300	.....	.....	.....	.....	.....	0.01	0.01	0.01
0.325	.....	.....	.....	.....	.....	.....	0.01	0.01

* After Laws and Parsons.

integration of the distribution is, of course, replaced by a summation.) Table 8·15 lists the result of these calculations at 18°C for the rains of Table 8·14, as given by Ryde and Ryde. The main features of the results are also shown in Figs. 8·12 and 8·13.

TABLE 8·15.—ATTENUATION AT 18°C FOR A NUMBER OF RAINS*

λ, cm	0.3	0.5	0.6	1.0	1.25	3.0	3.2	10
p, mm/hr	Attenuation, db/km							
0.25	0.305	0.160	0.106	0.037	0.0215	0.00224	0.0019	0.00001
1.25	1.15	0.720	0.549	0.228	0.136	0.0161	0.0117	0.00042
2.5	1.98	1.34	1.08	0.492	0.298	0.0388	0.0317	0.00079
12.5	6.72	5.36	4.72	2.73	1.77	0.285	0.238	0.00364
25	11.3	9.49	8.59	5.47	3.72	0.656	0.555	0.00728
50	19.2	16.6	15.3	10.7	7.67	1.46	1.26	0.0149
100	33.3	29.0	27.0	20.0	15.3	3.24	2.80	0.0311
150	46.0	40.5	37.9	28.8	22.8	4.97	4.39	0.0481

* From Ryde and Ryde.

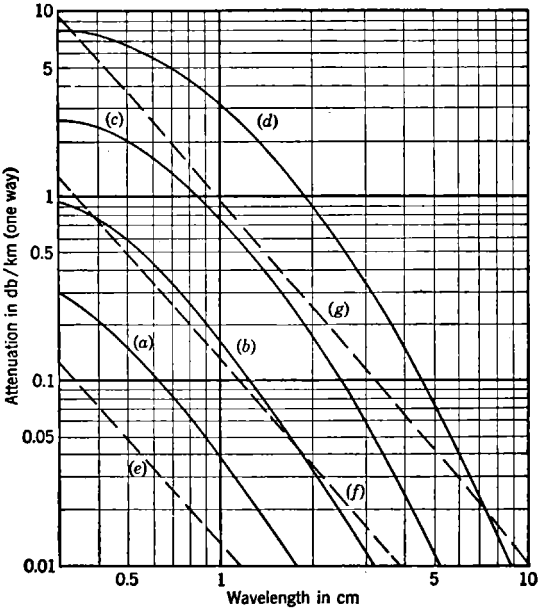


FIG. 8·12.—Theoretical values of attenuation by rain and fog. Solid curves show attenuation in rain of intensity. (*a*), 0.25 mm/hr (drizzle); (*b*), 1 mm/hr (light rain); (*c*), 4 mm/hr (moderate rain); (*d*), 16 mm/hr (heavy rain). Dashed curves show attenuation in fog or cloud. (*e*), 0.032 g/m³ (visibility about 2000 ft); (*f*), 0.32 g/m³ (visibility about 400 ft); (*g*), 2.3 g/m³ (visibility about 100 ft).

When the attenuation is plotted against rate of precipitation at constant wavelength, it is found that smooth curves can be accurately drawn through the data. These curves are shown in Figure 8·13. The considerations given above have indicated that such a behavior is to be expected for wavelengths from 1.0 to 3.0 cm; for the other wavelengths it must be considered to some extent as fortuitous. The curves of Fig. 8·13 indicate that it is reasonable to interpolate between the values of Table 8·15 to obtain figures

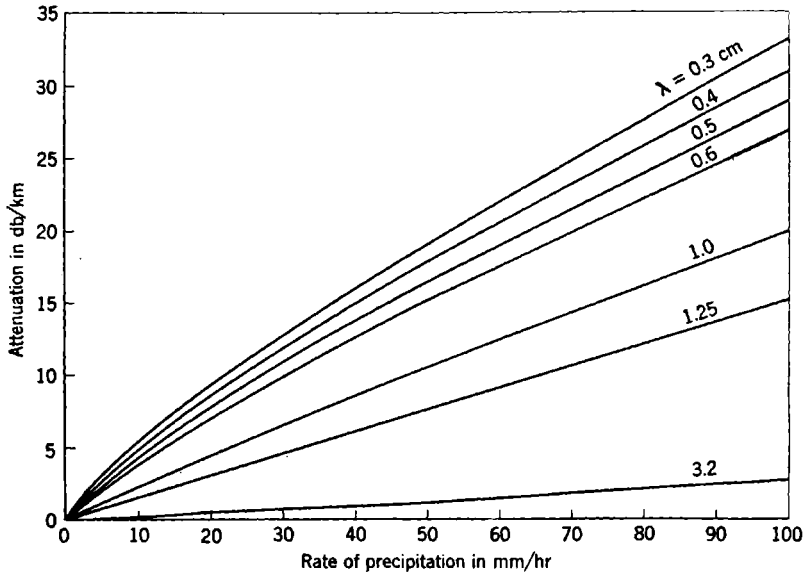


FIG. 8·13.—Theoretical values of attenuation as a function of rate of precipitation using raindrop distributions of Laws and Parsons. (*From Ryde and Ryde.*)

for other rates of precipitation. In Fig. 8·12 the solid lines are curves of the attenuation vs. wavelength for precipitation rates of 0.25, 1, 4, and 16 mm/hr. On Humphreys' classification[1] these correspond respectively to drizzle, light rain, moderate rain, and heavy rain. The dashed lines in the same figure represent the attenuation for clouds and fog as given by Eq. (41) for water contents of 0.032, 0.32 an 2.3 g/m^3 respectively. It is seen that the attenuation due to a heavy sea fog exceeds that due to moderate rain, but it must be remembered that although the drops are much larger in the rain, the total water content of the heavy fog greatly exceeds that of the rain, and to a first approximation the attenuation depends only on water content.

For wavelengths from 3.2 to 1.0 cm the attenuation is roughly proportional to the rate of precipitation. At shorter wavelengths the slope of the

[1] W. J. Humphreys, *op. cit.*

curve is not constant but decreases as precipitation rate increases for light rains, as was predicted above. Even at wavelengths of 1.25 and 1.0 cm this deviation from linearity is present for rains less than 2.5 mm/hr, although it is not apparent on Figure 8·13 because of the scale. Table 8·16 lists the slopes of the straight-line approximation for wavelengths from 1.0 to 3.2 cm, along with the accuracy to which they represent the data of Table 8·15 above 2.5 mm/hr.

TABLE 8·16.—SLOPE OF APPROXIMATE LINEAR RELATION BETWEEN ATTENUATION AND PRECIPITATION RATE AT 18°C

(For the rains of Table 8·14 above 2.5 mm/hr)

Wavelength, cm	Slope, db/km per mm per hr
1.0	0.200 ± 0.008
1.25	0.150 ± 0.007
3.0	0.028 ± 0.005
3.2	0.024 ± 0.004

It must be emphasized again that Tables 8·15 and 8·16 and Figs. 8·12 and 8·13 apply only to the rains whose drop-size distributions are given in Table 8·14. How representative these distributions are cannot be said until much more extensive data are available. Nevertheless, certain general qualitative conclusions seem warranted. For any applications thus far proposed, *the attenuation by rain at wavelengths of about 10 cm and above is completely unimportant.*

For wavelengths down to about 3 cm rain attenuation is seldom likely to be serious, except possibly in cases of the heaviest tropical downpours. It should be kept in mind in considering rain attenuation that although the rate of rainfall at a given location may be high for a brief period, it must be high at one time over a considerable distance in order to be important. For this reason brief intense showers are less significant than very large areas of fairly steady rain; fortunately in the latter case the rainfall intensity is usually much lower than in the brief, violent shower. According to Humphreys' classification a precipitation rate of 40 mm/hr represents "excessive rain." At 3.2 cm a shower of this intensity and 1 mile in diameter would attenuate the radar echo from a target behind it by about 3 db. Although this decrease might be visible on the PPI for a weak signal (that is, one that does not saturate the video amplifiers) it would not affect the appearance of strong echoes from ships or land.

Clearly the degree of applicability of the preceding remarks depends upon the wavelength under consideration. The irregular nature of rainfall is such that a mere statement of attenuation in terms of rate of rainfall is likely to give little information about the actual attenuation over an

extended path; consequently the use of such figures should be tempered by experience.

In the region below about 3 cm the importance of rain attenuation increases rapidly with decreasing wavelength, until *at wavelengths of about* $1\frac{1}{4}$ *cm or less the attenuation even from moderate rain far exceeds that from uncondensed water vapor or oxygen.* In this case even small rainfall areas may be of considerable importance if they contain intense precipitation (a not unlikely situation).

It is tempting to try to estimate the relative frequencies and geographical distribution of rains sufficiently heavy to be important for various wavelengths, and in fact some noteworthy attempts have been made by several investigators. The existing climatological data are entirely too sparse and lacking in detail for the conclusions to be considered very significant, however, and consequently they will not be considered here.

It should be pointed out that other effects of rain have frequently been attributed by uncritical observers to attenuation by absorption and scattering, particularly in connection with radar. One of the most frequent situations arises when a thin film of water forms over a radome housing an antenna or over a window of a supposedly weatherproof antenna feed device. Even at wavelengths of 10 cm or more the water film not only may absorb a considerable fraction of the energy incident upon it but may provide a serious mismatch for the transmission system, with the usual attendant undesirable consequences. A second phenomenon often incorrectly ascribed to rain attenuation is the apparent shadow immediately behind rain areas, as seen on PPI or B-scopes. Such shadows can, of course, be caused by attenuation, but they are more often likely to be caused by receiver saturation or various types of incorrect circuit adjustments. A third phenomenon is the masking of radar echoes by echoes from the rain at or near the target—something that sounds obvious but is not always so obvious on the radar indicator.

The calculation of attenuation has so far been for a temperature of 18°C. Ryde and Ryde have investigated the temperature dependence in the form of a multiplication factor $\phi(T)$ which corrects the values given for 18°C. Table 8·17 lists $\phi(T)$ for several rains and for temperatures from 0° to 40°C. For wavelengths of 1.25 cm and less, which are the only ones of importance, it is seen that the effect of temperature is quite small, less than 20 per cent over the entire range, and can usually be neglected.

8·8. Calculation of Attenuation by Precipitation in Solid Form.—The refractive index of ice has been measured by Dunsmuir and Lamb[1] between wavelengths of 3 and 9 cm. Within these limits it appears that the real and imaginary parts of the index of refraction are independent of wavelength. It therefore seems reasonable to assume that the same values hold

[1] R. Dunsmuir and J. Lamb, "The Dielectric Properties of Ice at Wavelengths of 3 and 9 cm," Manchester University Report No. 61, March 1945.

TABLE 8·17.—TEMPERATURE-CORRECTION FACTOR $\phi(T)$*

Precipitation rate, mm/hr	λ, cm	Correction factor $\phi(T)$				
		0°C	10°C	18°C	30°C	40°C
0.25	0.5	0.85	0.95	1.0	1.02	0.99
	1.25	0.95	1.0	1.0	0.90	0.81
	3.2	1.21	1.10	1.0	0.79	0.55
	10.0	2.01	1.40	1.0	0.70	0.59
2.5	0.5	0.87	0.95	1.0	1.03	1.01
	1.25	0.85	0.99	1.0	0.92	0.80
	3.2	0.82	1.01	1.0	0.82	0.64
	10.0	2.02	1.40	1.0	0.70	0.59
12.5	0.5	0.90	0.96	1.0	1.02	1.00
	1.25	0.83	0.96	1.0	0.93	0.81
	3.2	0.64	0.88	1.0	0.90	0.70
	10.0	2.03	1.40	1.0	0.70	0.59
50	0.5	0.94	0.98	1.0	1.01	1.00
	1.25	0.84	0.95	1.0	0.95	0.83
	3.2	0.62	0.87	1.0	0.99	0.81
	10.0	2.01	1.40	1.0	0.70	0.58
150	0.5	0.96	0.98	1.0	1.01	1.00
	1.25	0.86	0.96	1.0	0.97	0.87
	3.2	0.66	0.88	1.0	1.03	0.89
	10.0	2.00	1.40	1.0	0.70	0.58

* After Ryde and Ryde.

for even shorter wavelengths. Table 8·18 lists n and $n\chi$ for several temperatures, as given by Ryde and Ryde.[1] The most noteworthy difference between the properties of water in liquid or solid form is the smallness of the absorption in the latter case. Besides greatly facilitating the calculations, the small value of $n\chi$ means that the attenuation for very small particle size will be small compared with the attenuation from droplets of equal water content.

TABLE 8·18.—REFRACTIVE INDEX OF ICE FOR λ = 3.0 to λ = 9.0 CM*

T, °C	n	$n\chi$
0	1.75	0.00105
−10°	1.75	0.000285
−30°	1.75	0.000145
−50°	1.75	0.000110

* Ryde and Ryde, from Dunsmuir and Lamb.

[1] The values differ slightly, but not significantly, from the figures given by E. L. Younker, "Dielectric Properties of Water and Ice," RL Report No. 644, December 1944.

Calculations for the attenuation by spherical hailstones can be carried out analogously to the case of drops, though with considerably more ease, as $n\chi$ is small. It turns out that the three-term expansion [Eq. (38)] is valid up to $\rho = 0.5$, above which the more exact series must be used. Table 8·19 is a condensation of a similar table by Ryde and Ryde giving $Q(a,\lambda)$, or rather $434Q(a,\lambda)$, the attenuation in decibels per kilometer for a concentration of one hailstone per cubic meter, at 0°C. Ryde and Ryde also give the terminal velocity relation as obtained by interpolation from Humphreys and calculate the function $g(a,\lambda)$. Except for wavelengths in the millimeter region the attenuation by dry hailstones is small compared with that from rains of equivalent water content or precipitation rate.

TABLE 8·19.—ATTENUATION BY DRY HAILSTONES AT 0°C, DB/KM PER STONE PER M³

a, cm	$434Q(a,\lambda)$			
	$\lambda = 0.5$ cm	$\lambda = 1.25$ cm	$\lambda = 3.2$ cm	$\lambda = 10$ cm
0.025	$3.75 \cdot 10^{-6}$	$2.81 \cdot 10^{-7}$	$7.43 \cdot 10^{-8}$	$2.29 \cdot 10^{-8}$
0.050	$3.16 \cdot 10^{-4}$	$7.57 \cdot 10^{-6}$	$7.15 \cdot 10^{-7}$	$1.85 \cdot 10^{-7}$
0.10	$1.41 \cdot 10^{-2}$	$4.00 \cdot 10^{-4}$	$1.37 \cdot 10^{-5}$	$1.57 \cdot 10^{-6}$
0.15	8.91	$3.31 \cdot 10^{-3}$	$1.19 \cdot 10^{-4}$	5.99
0.20	$1.79 \cdot 10^{-1}$	$2.53 \cdot 10^{-2}$	6.15	$1.80 \cdot 10^{-5}$
0.25	2.63	9.55	$2.27 \cdot 10^{-3}$	4.64
0.375	4.07	$6.31 \cdot 10^{-1}$	$2.63 \cdot 10^{-2}$	$3.24 \cdot 10^{-4}$
0.50	4.08	$1.59 \cdot 10^{0}$	$1.59 \cdot 10^{-1}$	$1.67 \cdot 10^{-3}$
0.75	.....	2.69	$1.18 \cdot 10^{0}$	$1.70 \cdot 10^{-2}$
1.00	.....	.....	3.98	9.77
1.50	.....	.....	14.5	$1.10 \cdot 10^{0}$
2.00	.....	.....	17.9	5.37

For ice and snow particles that are not spherical it is no longer possible to use the Mie formulas. Although the calculations have been extended to ellipsoidal particles,[1] the computations are extremely complicated. However, when the size of the particles is very small compared with the wavelength, it is found that the attenuation again reduces to a term proportional to the volume of the particles similar to the c_1 term for spheres [*cf.* Eq. (38)]. For spherical ice crystals, where χ^2 terms may be neglected, c_1 reduces to

$$c_1 = \chi \frac{12n^2}{(n^2 + 2)^2}. \tag{44}$$

Similarly, for long needles

$$c_1 = \chi \frac{4n^2}{9}\left[1 + \frac{8}{(n^2 + 1)^2}\right], \tag{45}$$

[1] See F. Moglich, *Ann. Physik*, **83**, 609 (1927), where references to the earlier literature are given.

and for flat plates

$$c_2 = \chi \frac{4n^2}{9}\left(2 + \frac{1}{n^4}\right). \tag{46}$$

With the values of n and χ from Table 8·18 the attenuation γ can be easily calculated. The resultant formulas are summarized in Table 8·20, where M is the water content in grams per cubic meter. A comparison with Eq. (41) for drops indicates that at 0°C and for λ about 1 cm the attenuation by ice clouds is two magnitudes smaller than the attenuation from a cloud of liquid droplets with the same water content. At −40°C the attenuation is still less. It seems safe to conclude that the attenuation by ice crystals can almost always be neglected, as M rarely goes above 0.5 g/m³.

TABLE 8·20.—ATTENUATION BY ICE CRYSTAL CLOUDS*

Type of particle	c_1	Attenuation γ, db/km	
		$T = -40$°C	$T = 0$°C
Spheres	1.43	$0.00044M/\lambda$	$0.0035M/\lambda$
Needles	2.02	$0.00062M/\lambda$	$0.0050M/\lambda$
Disks	2.86	$0.00087M/\lambda$	$0.0070M/\lambda$

* After Ryde and Ryde.

Although detailed calculations of the attenuation by dry snow are faced with obvious difficulties, it seems reasonable that it should be less than that for hailstones of comparable dimensions and that because of the small value of χ, it would be less than the attenuation by rains of equal precipitation rates.

8·9. Measurements of Attenuation by Rain.—In 1935 Wolff and Linder[1] attempted to measure the attenuation of 9-cm waves over a 2-mile path. Although measurements were made at precipitation rates up to 75 mm/hr, no detectable attenuation could be observed. It was concluded that the attenuation at 9 cm must be less than 8×10^{-4} db/km per mm per hr, which is consistent with the prediction of Table 8·9. The first positive measurements were obtained by the Clarendon Laboratory group at Oxford early in 1942, for a wavelength of about 1 cm.[2] Some time later the Bell Telephone Laboratories published the results of similar measurements at 1.09 and 3.2 cm.[3] Subsequently, work has been done at 0.6 cm, also by Bell

[1] I. Wolff and E. G. Linder, "Transmission of 9 cm Electromagnetic Waves," *Broadcast News*, 18 10, (1935), reprinted in *Radio at Ultra-high Frequencies*, RCA Institutes Technical Press, 1940.

[2] M. G. Adam, R. A. Hull and C. Hurst, Committee on Valve Development, Clarendon Laboratory, *Misc.* 3, June 1942.

[3] S. D. Robertson, BTL Memorandum No. MM42-160-87, Aug. 1, 1942; A. P. King and S. D. Robertson, *ibid.*, No. MM42-160-93, Aug. 26, 1942; S. D. Robertson, *ibid.*, No. MM43-160-2, Jan. 5, 1943. Also S. D. Robertson, "The Effect of Rain upon the Propagation of Waves in the 1 and 3 cm Region," *Proc. IRE*, **34**, 178 *P* (1946).

Laboratories,[1] and at 1.25 cm by the Radiation Laboratory[2] and by the Naval Electronics Laboratory.[3] All the experiments used the same basic procedure. A simple transmitter and receiver were set up at opposite ends of a short path. Changes in signal strength were usually measured by comparison with a calibrated attenuator in the r-f line of the receiver, although the recorded output appeared in the form of a crystal, a-f amplifier, or bolometer output. In all cases rates of precipitation were measured simultaneously with one or more rain gauges. Sometimes the rate of rainfall was measured by a funnel and graduate read by an observer, a method that is accurate only for high precipitation rates; at other times commercial recording rain gauges were used. In several of the experiments measurements were also made of the raindrop radii by the "filter paper" method (*cf.* Sec. 7·7). However, in no case were significant distributions obtained, and no use was made of the results. Table 8·21 summarizes the experimental details of each investigation.

The possible sources of error are roughly the same for each experiment. Generally speaking the purely electrical measurements were adequately accurate with the possible exception of the pioneer Clarendon work. The greatest uncertainties arose in the meteorological measurements. Most important of these was the question of nonuniformity in the spatial distribution of the rain intensity. Wherever several rain gauges were used, there was considerable evidence of nonuniform rainfall even over distances of several hundred meters, and there is every reason to believe similar conditions held for all the paths. The BTL measurements minimized this source of error by using a short path, while in the NRSL investigation rain gauges were spaced 220 m apart. It seems likely that even closer spacings would be helpful. (By the same token there must inevitably be considerable uncertainty whenever the measurements or theory are used to predict the attenuation for a specific radar or transmission application. The rainfall will never be uniform along any practical path, and hence it will be possible to calculate only the order of magnitude of the attenuation to be expected.)

Another source of error is the method of measuring the rates of precipitation. Rado[4] has discussed the accuracies of the various types of gauges and has pointed out that none of the gauges is designed specifically for measuring high rates of rainfall. Two gauges placed close together occasionally differ by as much as 20 per cent, apparently as a result of aerodynamic effects of local updrafts.

[1] G. E. Mueller, "Propagation of K/2 Band Waves," BTL Memorandum No. MM44-160-150, July 3, 1944. Also, G. E. Mueller, "Propagation of 6-millimeter Waves," *Proc. IRE*, **34**, 181 *P* (1946).

[2] G. T. Rado, "Measurements of the Attenuation of K-band," RL Report No. 603, Mar. 7, 1945.

[3] L. J. Anderson, J. P. Day, C. H. Freres, J. B. Smyth, A. P. D. Stokes, and L. G. Trolese, "K-band Attenuation Due to Rainfall," NRSL Report No. WP-20, June 1945; and *Proc. IRE*, **35**, 351 (1947).

[4] G. T. Rado, *op. cit.*

TABLE 8·21.—DETAILS OF ATTENUATION EXPERIMENTS

Organization	λ, cm	Type of transmitter	Type of receiver	Path length, m	No. of rain gauges	Type of rain gauges	Range of rainfall, mm/hr
Clarendon	0.96	Square-wave modulated klystron	Crystal and galvanometer	2000	1	Dine's tilting syphon	0–20
BTL	1.09 3.20	C-w klystron	Superheterodyne and calibrated output meter	380 ($\lambda = 1.09$) 270 ($\lambda = 3.20$)	1	Funnel and graduate	0–100
BTL	0.62	Square-wave modulated klystron and second-harmonic crystal generator	Crystal and audio amplifier	380	1	Funnel and graduate	0–100
RL	1.25	Pulsed magnetron	Bolometer and audio amplifier	4000	3	Several types of recording gauges	0–30
NEL (NRSL)	1.25	A-f modulated klystron	Superheterodyne and a-f amplifier	1950	9	Four funnel and graduate; five tipping bucket	0–80

Precautions must also be taken to keep the rain out of the transmitting or receiving r-f line. Usually this was done by placing large enough shelters over the transmitter and receiver. In the Radiation Laboratory investigation the more positive step was taken of blowing a stream of hot air out through the feed of the two antennas.

Other possible, though not important, sources of error are reflections from ground or other objects and lack of time synchronism between the records of attenuation and rain intensity.

The experimental points obtained in each investigation when plotted on the γ-p plane usually show considerable random scatter. Although this behavior is to be expected in any case on fundamental theoretical

TABLE 8·22.—SUMMARY OF ATTENUATION MEASUREMENTS

Organization	λ, cm	Attenuation, db/km per mm per hr		
		Upper bound	Lower bound	Average
BTL	3.2	0.090	0.012	0.019
RL	1.25	0.40	0.09	0.17
NRSL	1.25	0.34	0.23	0.25
BTL	1.09	0.27	0.15	0.18
Clarendon	0.96	0.25	0.10	0.15
BTL	0.62	0.37	0.27	0.31

grounds, some of the scatter arises from lack of uniformity of the rainfall over the path. Thus in the NEL (NRSL) experiment, where enough gauges were used to provide an adequate picture of the rainfall distribution, the scatter of the points is quite small. For some of the points the rainfall was substantially uniform over the path, and for the others the records could be analysed using the points already obtained.

Table 8·22 summarizes the results of the measurements by listing the slope of the "average" straight line best fitting the experimental data, along with the bounds that include most of the points. These figures are to be compared with the theoretical bounds listed in Table 8·13 and the approximate linear relations given in Table 8·16. For the higher wavelengths the experimental upper bounds exceed the theoretical limit, but the large range of values measured in most cases indicates that the difference is caused by experimental uncertainty. The average values, in all but one case, lie well within the theoretical limits, with the Clarendon figure appearing to be slightly low. The one exception is the NEL result, where even the lowest experimental bound exceeds the theoretical maximum. It is difficult to see what may be the source of the discrepancy, as in many respects the work seems to be the most careful of the entire group. The suggestion put forward by the authors, that the theoretical assumption of incoherent scat-

tering is incorrect, is not at all tenable. A good deal of the attenuation is caused by absorption, and here there can be no question of coherence or incoherence as the absorption must be proportional to the number of drops without regard to their distribution. With the scattering portion of the attenuation the same considerations hold as for back scattering from rain. In Sec. 7·5 it was shown that the coherent portion of the scattering is completely negligible unless the concentration of drops changes rapidly over a

TABLE 8·23.—COMPARISON OF EXPERIMENTAL AND THEORETICAL ATTENUATION
(λ = 0.6 cm)

p, mm/hr	Attenuation, db/km	
	BTL	Ryde and Ryde
10	3.7	3.8
20	7.4	7.0
30	10.6	10.0
40	13.6	12.7
50	16.0	15.4
70	20.4	20.0
100	26.5	27.0

strip about one wavelength wide. It seems highly unlikely that the boundary of the rain could be so sharply defined. It would appear much more likely that the origin of the difference is to be found on the meteorological side of the experiment, perhaps in the terminal velocity relation, or in the dielectric constant of the water.

The measurements at 0.62 cm by Bell Telephone Laboratories show very little scatter and clearly exhibit the increased slope for low precipitation rates. Their average curve is in very good agreement with the curve calculated by Ryde and Ryde for the rains given in Table 8·14 as shown by Table 8·23, where these two curves are compared. The close agreement is perhaps to some extent fortuitous, considering the uncertainties on both sides, but is nevertheless satisfying.

APPENDIX A

APPLICATION OF THE LORENTZ RECIPROCITY THEOREM TO SCATTERING

By Donald E. Kerr

There are in current use several reciprocity theorems, all of which are often loosely referred to as "the" reciprocity theorem. We apply here the version developed by H. A. Lorentz.[1]

We begin with a purely formal manipulation of Maxwell's equations. Let us assume that by some as yet unspecified means two different distributions of electromagnetic fields can be established in the region of interest. We denote these field distributions by the subscripts 1 and 2 on the corresponding field vectors. Each set of field vectors, assumed to vary harmonically in time with angular frequency ω, satisfies the Maxwell equations[2]

$$\nabla \times \mathbf{H} = (\sigma + i\omega\epsilon)\mathbf{E}, \tag{1}$$

$$\nabla \times \mathbf{E} = -i\omega\mu\mathbf{H}. \tag{2}$$

By forming the four following scalar products and employing Eqs. (1) and (2) one finds easily that

$$\mathbf{H}_2 \cdot \nabla \times \mathbf{E}_1 - \mathbf{E}_1 \cdot \nabla \times \mathbf{H}_2 - (\mathbf{H}_1 \cdot \nabla \times \mathbf{E}_2 - \mathbf{E}_2 \cdot \nabla \times \mathbf{H}_1) = 0, \tag{3}$$

which may be recognized as

$$\nabla \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) = 0. \tag{4}$$

By the divergence theorem, the integral of the divergence of a vector throughout a volume is equal to the integral of the normal component of the vector over a surface enclosing the volume; hence,

$$\int_S \mathbf{n} \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)\, dS = 0, \tag{5}$$

where $\mathbf{n}$ is the outward normal to the surface S enclosing the region of interest. This is the Lorentz reciprocity theorem.

It should be pointed out that Eq. (5) is a purely formal relation between the two sets of field vectors. The vector product $\mathbf{E} \times \mathbf{H}$ might tempt one to seek to interpret it as a Poynting vector, but it has no such simple meaning. Here $\mathbf{E}$ and $\mathbf{H}$ are, in general, three-dimensional vectors and are complex as well.

[1] *Amst. Ak. van Wetenschappen*, **4**, 176 (1895-96).

[2] Regions containing batteries or generators are excluded.

We now define the meaning of the subscripts, with reference to Fig. A·1. The fields $\mathbf{E}_1$ and $\mathbf{H}_1$ are those obtained *with the target removed* and with a driving voltage V_1 from the radar transmitter across the transmission line[1] in the plane AA'. The fields $\mathbf{E}_2$ and $\mathbf{H}_2$ are those *with the target in place and with the radar transmitter in operation, delivering the same* V_1 *that it delivered in the absence of the target.*[2] It should be observed that

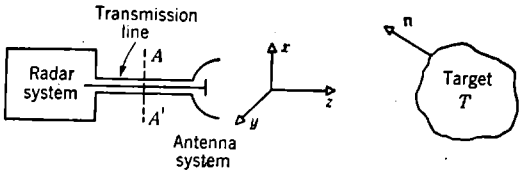


FIG. A·1.—Schematic representation of radar system and target in free space. The radar system is enclosed by a perfectly conducting shield and the transmission line and antenna surfaces also have infinite conductivity. The plane AA' in the transmission line is the plane across which the transmitted and received voltages are measured.

both fields 1 and 2 are *total* fields. They both satisfy Maxwell's equations; field 2 satisfies the boundary conditions on the several surfaces that form S, and field 1 satisfies the boundary conditions on all the surfaces except the target. Using these two fields, we now proceed to derive relations between them and the power radiated by the transmitter and the echo power from the target.

We now apply Eq. (5), integrating over an outer surface at infinity, the target, the outer surfaces of the radar and antenna system, the inner metal surfaces down to the plane AA', and finally over the plane AA' inside the transmission line.

The triple scalar product in Eq. (5) can be written

$$\mathbf{n} \cdot \mathbf{E} \times \mathbf{H} = \mathbf{H} \cdot \mathbf{n} \times \mathbf{E} = -\mathbf{E} \cdot \mathbf{n} \times \mathbf{H}; \tag{6}$$

thus only the field components tangential to the surface S contribute to the integral. The integral vanishes on all parts of the radar shielding, transmission line, and antenna system as a result of the fact that $\mathbf{n} \times \mathbf{E}_1 = \mathbf{n} \times \mathbf{E}_2 = 0$.

It can be shown that the integration of Eq. (5) over the surface at infinity vanishes, but the details will not be given here.[3]

The remaining surfaces are the plane AA' and the surface of the target. In the transmission line the fields and the current and

[1] A coaxial line is shown, but the results to be derived are valid for any nonradiating transmission system.

[2] It is assumed that the impedances terminating the transmission line on both ends are matched to its real characteristic impedance, Z_0, and that the echo signal does not disturb the impedance match.

[3] See J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York, 1941, p. 487.

voltage are related by

$$E_r = \frac{V}{r \ln\left(\frac{b}{a}\right)}, \qquad H_\phi = \frac{I}{2\pi r}, \tag{7}$$

where r and ϕ are the usual cylindrical coordinates in the line, a and b are the inner and outer radii of the line, and V and I are the line voltage and current. If Eq. (7) is substituted into Eq. (5), we have[1]

$$\int_{AA'} \mathbf{n} \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)\, dS = V_1 I_2 - V_2 I_1. \tag{8}$$

This expression may be simplified further. If the positive direction for current is along the positive z-axis, and if we recall that impedances are matched on both ends to Z_0,

$$I_1 = \frac{V_1}{Z_0}.$$

With the target in place the total voltage V_2 at AA' is the sum of V_1 and the echo voltage v. The total line current I_2 consists of the sum of the original I_1, flowing to the right in Fig. A·1 under the influence of V_1, and the echo current $-v/Z_0$, negative because it flows to the left under the influence of v. Hence

$$V_2 = V_1 + v,$$

$$I_2 = \frac{V_1 - v}{Z_0},$$

and

$$V_1 I_2 - V_2 I_1 = -2v \frac{V_1}{Z_0}. \tag{9}$$

Employing Eqs. (9) and (8) in Eq. (5), we have

$$\frac{V_1 v}{Z_0} = \frac{1}{2} \int_T \mathbf{n} \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)\, dS. \tag{10}$$

This expression may be transformed to include the transmitted power and the echo power if we employ the fact that the average power flow in either direction in the line is given by a formula of the type $P = VV^*/2Z_0$.

If we now multiply Eq. (10) by its complex conjugate, and if in conformity with earlier notation we denote the transmitted power P_1 by P_t and the received power P_2 by P_r, Eq. (10) becomes

$$P_r = \frac{1}{4P_t} \left| \frac{1}{2} \int_T \mathbf{n} \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)\, dS \right|^2. \tag{11}$$

[1] Equation (8) holds in its present form for c-w radar but is zero for pulsed operation. We shall not introduce the additional complexity required to establish the results for the pulse case.

We restrict the analysis from here on to that for a large perfectly conducting target having radii of curvature greater than about one wavelength (a "smooth" target in the sense of Sec. 6·3). For such a target the second term of the integrand of Eq. (11) becomes

$$\mathbf{n} \cdot \mathbf{E}_2 \times \mathbf{H}_1 = \mathbf{H}_1 \cdot \mathbf{n} \times \mathbf{E}_2 = 0$$

because of the infinite conductivity.

The first term of the integrand may be obtained from Eqs. (6·53) and (6·54) of Sec. 6·3 which yield, in the present case,

$$\mathbf{H}_2 \approx 2[\mathbf{H}_1 - (\mathbf{n} \cdot \mathbf{H}_1)\mathbf{n}]. \tag{12}$$

The integrand of Eq. (11) then becomes

$$\mathbf{n} \cdot [\mathbf{E}_1 \times \mathbf{H}_1 - (\mathbf{n} \cdot \mathbf{H}_1)\mathbf{E}_1 \times \mathbf{n}].$$

The triple scalar product in the second term is

$$\mathbf{n} \cdot \mathbf{E}_1 \times \mathbf{n} = \mathbf{E}_1 \cdot \mathbf{n} \times \mathbf{n} = 0;$$

thus for this case the integral reduces to a simple expression in terms of the incident fields. We now impose the customary restriction that the distance from radar to target is sufficiently large that the surfaces of constant phase in the incident wavefronts are substantially plane, and we assume propagation along the z-axis; after some manipulation the integrand then becomes

$$\mathbf{n} \cdot \mathbf{E}_1 \times \mathbf{H}_1 = \mathbf{n} \cdot \mathbf{E}_1 \times (\mathbf{i}_z \times \mathbf{E}_1) \frac{1}{\eta_0} = \frac{1}{\eta_0} (\mathbf{E}_1 \cdot \mathbf{E}_1)(\mathbf{n} \cdot \mathbf{i}_z),$$

where $\eta_0 = 120\pi$ ohms, the intrinsic impedance of free space.

As in Sec. 6·3, we replace $\mathbf{n} \cdot \mathbf{i}_z \, dS$ by dA, the projection of dS of the target on to the xy-plane, at right angles to the direction of propagation of the incident wave. Introducing these modifications, Eq. (11) becomes

$$P_r = \frac{1}{4P_t} \left| \frac{1}{\eta_0} \int \mathbf{E}_1 \cdot \mathbf{E}_1 \, dA \right|^2. \tag{13}$$

Most radar systems in use at the present time employ antenna systems that radiate essentially linearly polarized waves in the direction of maximum transmission; and in general, data on target cross sections relate to that part of the echo energy which is linearly polarized in the same direction as is the radiated wave. The effects of other types of polarization may be examined by the present analysis, however. For example, if the antenna radiates an elliptically polarized wave, we may represent it by

$$\mathbf{E}_1 = E_0 e^{-ikz} [\mathbf{i}_x + \mathbf{i}_y g e^{-i\beta}], \tag{14}$$

where g is a real scalar quantity and β is the time phase between the

two field components at right angles in space. The accompanying magnetic field is

$$\mathbf{H}_1 = \frac{E_0}{\eta_0} e^{-ikz} [-\mathbf{i}_x g e^{-i\beta} + \mathbf{i}_y]. \tag{15}$$

The Poynting vector giving the incident average power flow is

$$\bar{S}_z^i = \frac{1}{2} \operatorname{Re}(\mathbf{E}_1 \times \mathbf{H}_1^*) = \frac{E_0^2}{2\eta_0} (1 + g^2). \tag{16}$$

Using Eq. (14), the integrand of Eq. (13) becomes[1]

$$\mathbf{E}_1 \cdot \mathbf{E}_1 = E_0^2 (1 + g^2 e^{-i2\beta}) e^{-i2kz}. \tag{17}$$

Using Eqs. (16) and (17) and manipulating the complex quantities, we find that Eq. (13) becomes

$$P_r = \frac{1}{P_t} (\bar{S}_z^i)^2 \left[1 - \frac{4g^2 \sin^2\beta}{(1 + g^2)^2} \right] \left| \int e^{-i2kz} \, dA \right|^2. \tag{18}$$

This equation may be recast in a familiar form if one recalls that the Poynting vector at a distance R in the direction of maximum transmission from an antenna of gain G is[2]

$$\bar{S}_z^i = \frac{P_t G}{4\pi R^2}.$$

With this modification and slight rearrangement Eq. (18) now becomes the radar transmission equation

$$\frac{P_r}{P_t} = \frac{G^2 \lambda^2}{(4\pi)^3 R^4} \left[1 - \frac{4g^2 \sin^2\beta}{(1 + g^2)^2} \right] \left(\frac{4\pi}{\lambda^2} \left| \int e^{-i2kz} \, dA \right|^2 \right). \tag{19}$$

The two brackets contain the factor heretofore called the back-scattering cross section σ, but now modified in a way to emphasize its meaning in relation to polarization and antenna properties. The polarization effect is given by the factor in the first brackets, and the second bracketed quantity may be recognized as the cross section given by Eq. (6·56) in Sec. 6·3 for linear polarization and a smooth target.

For the case of linear polarization $\beta = 0$, and Eq. (19) reduces to the simple form discussed in Sec. 6·3. For circular polarization, however, $g = 1$, $\beta = \pm\pi/2$, and Eq. (19) predicts that the echo power is zero.[3]

[1] The earlier remark about the $\mathbf{E} \times \mathbf{H}$ products not being interpretable as Poynting vectors is well illustrated by the $\mathbf{E}_1 \cdot \mathbf{E}_1$ of Eq. (17) which developed from them. It is a complex number having no direct connection with average energy flow as it stands.

[2] The gain must be defined appropriately for elliptical polarization. It is the ratio of S given by Eq. (16) to S at the same distance from an isotropic antenna radiating the same power as the antenna system under investigation.

[3] Apparently this fact has been generally known for several years, but the present author is not aware of its origin. It first came to his attention in P. D. Crout and F. E. Bothwell, "A Theoretical Treatment of Radar Target Return," RL Report No. 719.

The skeptical reader is invited to consider again the system shown in Fig. A·1, with particular regard to the antenna system. Let the transmission system be split into two branches on the right of AA'. Each branch leads to an antenna that radiates a linearly polarized wave, but the antennas are oriented so that their polarizations are at right angles. The antennas are identical and are supplied with equal amounts of power, but the line from the branch to one antenna is $\lambda/4$ longer than the other line. This combination of antennas produces, at large distances, a circularly polarized wave described by Eq. (14), in which $g = 1$ and $\beta = \pi/2$. The x- and y-components of the wave incident upon the target are scattered independently and in an identical manner (by the *smooth* target) and upon return to the antenna system are received independently by the same antenna that transmitted them. Because of the additional $\lambda/4$ in one side of the transmission line, the total excess phase delay in this side is π for the two-way journey. The two signals at AA' are equal in magnitude and are in phase opposition, hence zero total signal.[1] To say it another way, the reflection from the target changes a right-circularly polarized wave radiated by the antenna system into a left-circularly polarized wave to which the antenna system will not respond.

The foregoing result emphasizes the fact that the earlier definition of σ in terms of the incident and back-scattered Poynting vectors, namely,

$$\sigma = 4\pi R^2 \frac{S^s}{S^i},$$

is useful only for linear polarization of the incident wave and the component of the back-scattered wave that is similarly polarized. In the case discussed above the back-scattered Poynting vector is certainly not zero, although the reciprocal relation between transmitting and receiving properties of the antenna system causes it to appear to be zero.

The reciprocity theorem is well known in antenna theory, where it is useful for deriving general properties of antenna and transmission systems.[2] In assuming a smooth target we have deliberately chosen a situation that emphasizes the antenna properties. The general case in which the target or the transmission path depolarizes the incident wave will not be treated here. The present development has been given primarily to bring to light some of the fundamentals underlying the concept of radar cross section.

[1] This would not occur if the echo signal were measured in either branch of the antenna feed line. The conventional arrangement, however, is to detect the echo in the transmission line at some location corresponding to AA'.

[2] See Vol. 12 of this series, or Frank and Von Mises, *Die Differential und Integralgleichungen der Mechanik und Physik*, Bd. 2, Vieweg, 1935, pp. 953–963; also John R. Carson, *Bell System Tech. Jour.*, **3**, 393 (1924); **9**, 325 (1930); *Proc. IRE*, **17**, 925 (1929); J. C. Slater, *Microwave Transmission*, McGraw-Hill, New York, 1942, p. 251; S. A. Schelkunoff, *Electromagnetic Waves*, Van Nostrand, New York, 1943, p. 477.

APPENDIX B

COHERENT AND INCOHERENT SCATTERING FROM ASSEMBLIES OF SCATTERERS

By A. J. F. Siegert and Herbert Goldstein

It will be assumed for simplicity that all the scattering particles return a signal of equal magnitude. The assumption is not a necessary one, but the introduction of a distribution in the magnitude of the scattering would only complicate the mathematics without affecting the conclusions. Let the distance of a convenient reference point on a given scatterer, denoted by "center of the scatterer," from the radar be r. No matter what the nature of the scatterer is, the scattered amplitude will be proportional to the incident field amplitude and inversely proportional to the distance r. Hence the scattered field at the receiver due to one scatterer can be represented as

$$E^s = \frac{p}{\sqrt{4\pi}} \frac{E_0(r)}{r} e^{-4\pi i(r/\lambda)}, \tag{1}$$

where $E_0(r)$ is the incident field amplitude and p is a constant of proportionality, by assumption the same for all the scatterers. The total scattered amplitude from all the scatterers is then

$$E^s = \frac{p}{\sqrt{4\pi}} \int_0^\infty n(r) \frac{E_0(r)}{r} e^{-4\pi i(r/\lambda)}\, dr, \tag{2}$$

where $n(r)\ dr$ is the number of scatterers with centers lying between r and $r + dr$. In the case of pulsed transmission the incident field $E_0(r)$ not only will vary with r as $1/r$ but will also exhibit the characteristic pulse shape. It is convenient, however, to treat the incident field as a continuous wave and lump the effects of the pulse shape in with the spatial variation of the density of scatterers by assuming a fictitious distribution of the scatterers corresponding to the pulse shape. Further, the variation of r over the distance of the pulse length is usually small so that Eq. (2) can be approximated to a high degree of accuracy by

$$E^s = \frac{p}{4\pi} \frac{E_0(R)}{R} \int_0^\infty n(r)\, e^{-4\pi i(r/\lambda)}\, dr, \tag{3}$$

where R is some appropriately defined range, say, to the center of the pulse. The scattered Poynting vector at the receiver will then be

$$S^s = |p|^2 \frac{S_0(R)}{4\pi R^2} \int_0^\infty \int n(r) n(r') e^{-\frac{4\pi i}{\lambda}(r - r')}\, dr\, dr'. \tag{4}$$

Actually the distribution of scatterers is rarely static; the scatterers are moving, and the number of scatterers present in any given volume changes with time. The distribution function should therefore more

properly be written $n(r,t)$. As a consequence, the scattered signal will vary with time; we shall be concerned, however, with the time average of the signal:

$$\overline{S^s} = |p|^2 \frac{S_0(R)}{4\pi R^2} \int_0^\infty \int \overline{n(r,t)n(r',t)}\, e^{-\frac{4\pi i}{\lambda}(r-r')}\, dr\, dr', \tag{5}$$

the bar denoting a time average. The incident Poynting vector is given in terms of the radar characteristics by[1]

$$S_0(R) = \frac{P_t G}{4\pi R^2},$$

while the average received power is connected with $\overline{S^s}$ by

$$\overline{P_r} = \frac{\overline{S^s} G \lambda^2}{4\pi}.$$

Hence the received power is

$$\overline{P_r} = \frac{P_t G^2 \lambda^2}{(4\pi)^3 R^4}\, \sigma,$$

where the radar cross section σ is given by

$$\sigma = |p^2| \int_0^\infty \int \overline{n(r,t)\, n(r',t)}\, e^{-\frac{4\pi i}{\lambda}(r-r')}\, dr\, dr'. \tag{6}$$

It is instructive to break σ up into terms by writing

$$\overline{n(r,t)\, n(r',t)} = \bar{n}(r)\, \bar{n}(r') + [\overline{n(r,t)\, n(r',t)} - \bar{n}(r)\, \bar{n}(r')],$$

where $\bar{n}(r)$ is the time average of the distribution function. The cross section can now be written

$$\sigma = |p^2| \left| \int_0^\infty \bar{n}(r)\, e^{-\frac{4\pi i r}{\lambda}}\, dr \right|^2 + |p^2| \int \int [\overline{n(r,t)\, n(r',t)} - \bar{n}(r)\, \bar{n}(r')]\, e^{-\frac{4\pi i}{\lambda}(r-r')}\, dr\, dr'. \tag{7}$$

The first term will obviously be proportional to the square of the total number of scatterers; it is therefore the "coherent" part of the scattering; that is, the individual amplitudes rather than intensities are added together. We shall show that this term vanishes if $\bar{n}(r)$ is constant. The nature of the second term is more clearly shown by observing that

$$\overline{n(r,t)\, n(r',t)} - \bar{n}(r)\, \bar{n}(r') = [\overline{n(r,t) - \bar{n}(r)}][\overline{n(r',t) - \bar{n}(r')}]. \tag{8}$$

Thus this term arises solely from the fluctuations in the density of the scatterers about the time average distribution. With a purely static

[1] For simplicity we assume free-space transmission.

distribution this contribution must vanish, leaving only the first term. Note that for $r = r'$ the right-hand side of Eq. (8) is just the mean square fluctuation of density, and it has been shown in many places that for independent scatterers this mean square fluctuation is just $\bar{n}(r)$.[1]

In general a time average such as is involved in the second term in σ can be obtained as follows. Let it be required to find

$$\overline{\int_0^\infty n(r,t)\, n(r',t)\, F(r')\, dr'} = \int_0^\infty \overline{n(r,t)\, n(r',t)}\, F(r')\, dr', \tag{9}$$

where $F(r)$ is any function of r alone (independent of time). In the course of time the assembly of scatterers will pass through an infinity of configurations, each configuration being specified by saying that scatterer number 1 is at r_1, scatterer number 2 is at r_2, and so on. The time average involved in Eq. (9) can be thought of as an average over all configurations obtained by multiplying the integrand by the probability for a given configuration $W(r_1, r_2 \cdots r_N)$ (that is, the relative frequency of the configuration) and integrating over all configurations. The density function $n(r,t)$ can be expressed in terms of the specification of the given configuration by using the well-known Dirac δ function.[2] For a configuration in which the centers of the N scatterers are located at $r_1, r_2, \cdots r_N$, the number of such centers located between r and $r + \Delta r$ is

$$n(r,t)\, \Delta r = n(r;r_1,r_2, \cdots r_N)\, \Delta r = \sum_{l=1}^{N} \int_r^{r+\Delta r} \delta(s - r_l)\, ds; \tag{10}$$

that is, each position r_l has a unit contribution to the number involved, providing r_l lies inside the region, zero otherwise. From the properties of the δ function it then follows that

$$\int_0^\infty F(r')\, n(r',t)\, dr' = \int_0^\infty F(r')\, n(r',r_1 \cdots r_N)\, dr' = \sum_{k=1}^{N} F(r_k). \tag{11}$$

The time average required in Eq. (9) can now be expressed as a configuration average:

$$\Delta r \int_0^\infty \overline{n(r,t)\, n(r',t)}\, F(r')\, dr' =$$

$$\int_0^\infty \cdots \int_0^\infty W(r_1 \cdots r_N)\, dr_1 \cdots dr_N \sum_{l=1}^{N} \int_r^{r+\Delta r} \delta(s - r_l)\, ds \sum_{m=1}^{N} F(r_m). \tag{12}$$

[1] *Cf.* the derivation given in a similar connection in M. Born, *Optik*, Springer, Berlin, 1933, p. 374.

[2] See, for example H. Margenau and G. M. Murphy, *The Mathematics of Physics and Chemistry*, Van Nostrand, New York, p. 325.

We now introduce the assumption that scatterers are independent, at least in the sense that the presence of the kth scatterer in any given interval does not prejudice the presence of other scatterers.[1] This assumption naturally requires that the distance between scatterers be large compared to their dimensions, a necessary condition that is usually amply satisfied. With this assumption the probability of any configuration $r_1 \cdots r_N$ is just the product of the individual probabilities of finding a given scatterer at a given position;

$$W(r_1 \cdots r_N) = \prod_{k=1}^{N} W(r_k),$$

where $W(r_k)\, dr_k$ is the probability of finding the kth scatterer in the interval between r_k and $r_k + dr_k$, and is given simply by

$$W(r)\, dr = \frac{\bar{n}(r)}{N}\, dr. \tag{13}$$

Equation (12) thus can be written as

$$\Delta r \int_0^\infty \overline{n(r,t)\, n(r',t)}\, F(r')\, dr' = \sum_{l=1}^{N} \sum_{m=1}^{N} \int \cdots \int \prod_{k=1}^{N} \frac{\bar{n}(r_k)\, dr_k}{N} \int_r^{r+\Delta r} \delta(s - r_l)\, dr_l\, F(r_m). \tag{14}$$

For values of the index k not equal to l or m the integrations over r_k can be performed immediately and lead to unity, for the integrand in each case consists solely of $n(r_k)/N$. For the remaining terms we must distinguish between the cases $l = m$ and $l \neq m$. In the former case the integrations over r_k can be carried out immediately for all k except $k = l$, whereas in the latter case two integrations are left: $k = l$, and $k = m$:

$$\Delta r \int_0^\infty \overline{n(r,t)\, n(r',t)}\, F(r')\, dr' = \sum_l \int_0^\infty dr_l\, \frac{\bar{n}(r_l)}{N} \int_r^{r+\Delta r} \delta(s - r_l)\, ds\, F(r_l) + \sum_{\substack{l \\ (l \neq m)}} \sum_m \int\int_0^\infty dr_l\, dr_m\, \frac{\bar{n}(r_l)\bar{n}(r_m)}{N^2} \int_r^{r+\Delta r} \delta(s - r_l)\, ds\, F(r_m). \tag{15}$$

By carrying out the integral over r_l first this reduces to

$$\frac{1}{N} \sum_l \int_r^{r+\Delta r} \bar{n}(s) F(s)\, ds + \frac{1}{N^2} \sum_{(l \neq m)} \sum \int_0^\infty dr_m \bar{n}(r_m) F(r_m) \int_r^{r+\Delta r} n(s)\, ds.$$

[1] It is exactly this assumption which determines the "Gaussian" nature of the statistics we use.

Each term in the summation over l and m is identical; hence the sums reduce to N times each term, or

$$\int_r^{r+\Delta r} \bar{n}(s)F(s)\,ds + \int_0^\infty \bar{n}(r')F(r')\,dr' \int_{r'}^{r+\Delta r} \bar{n}(s)\,ds.$$

Finally we can let $\Delta r \to 0$ so that the integral over s becomes just the integrand times Δr:

$$\Delta r \int \overline{n(r,t)n(r',t)}\,F(r')\,dr' = \Delta r \left[\bar{n}(r)F(r) + \int_0^\infty \bar{n}(r')\bar{n}(r)F(r')\,dr'\right].$$

Hence it follows that

$$\int_0^\infty\int \left[\overline{n(r,t)n(r',t)} - \bar{n}(r)\bar{n}(r')\right] F(r')G(r)\,dr\,dr' = \int_0^\infty \bar{n}(r)F(r)G(r)\,dr. \quad (16)$$

Applying this result to σ [Eq. (7)] with $F(r') = e^{4\pi i r'/\lambda}$ and $G(r) = e^{-4\pi i r/\lambda}$ we obtain

$$\sigma = |p^2| \int_0^\infty \bar{n}(r)\,dr + |p^2| \left| \int_0^\infty \bar{n}(r) e^{-\frac{4\pi i r}{\lambda}}\,dr \right|^2. \quad (17)$$

The first term is just $N|p^2|$, the cross section per scatterer times the number of scatterers, and therefore represents the usual incoherent scattering, and the second term, as has been remarked, represents the coherent scattering.

The integral involved in the coherent scattering obviously vanishes if $\bar{n}(r)$ is essentially constant, for it represents the average of $\bar{n}(r)$ over an oscillatory function, which by itself averages to zero. Clearly the magnitude of the coherent scattering will depend upon how much the density function $\bar{n}(r)$ varies over a period of the oscillatory function, that is, in a distance λ. Thus, qualitatively it can already be seen that the magnitude of the coherent scattering will depend upon the gradient of the density of particles. In fact, because the density of particles can be taken to be zero at the origin and infinity, one can integrate by parts and obtain such a dependence explicitly.

$$\int_0^\infty \bar{n}(r) e^{-\frac{4\pi i r}{\lambda}}\,dr = +\frac{\lambda}{4\pi i} \int_0^\infty \frac{d\bar{n}}{dr} e^{-\frac{4\pi i r}{\lambda}}\,dr. \quad (18)$$

As an example of these results let us calculate the coherent scattering from a uniform density gradient extending over a distance Δr less than the pulse length. (It will be assumed that the edges of the pulse have sufficiently gradual "sides" that they do not contribute to the coherent scattering.) We then have

$$\sigma_c = |p|^2 \left(\frac{\lambda}{4\pi}\right)^2 \left(\frac{d\bar{n}}{dr}\right)^2 \left| \int_{\frac{-\Delta r}{2}}^{\frac{+\Delta r}{2}} e^{-\frac{4\pi i}{\lambda} r} \right|^2$$

or

$$\sigma_c = 4|p|^2 \left(\frac{\lambda}{4\pi}\right)^4 \left(\frac{d\bar{n}}{dr}\right)^2 \sin^2\left(\frac{2\pi\,\Delta r}{\lambda}\right). \tag{19}$$

Thus the maximum signal is already obtained when the gradient extends over a strip no wider than $\lambda/2\pi$.

It has already been pointed out that the effect of pulsed transmission is the same as if there were a corresponding pulselike spatial distribution of scatterers, so that a certain amount of coherent scattering can be expected because of the fictitious gradient of n corresponding to the

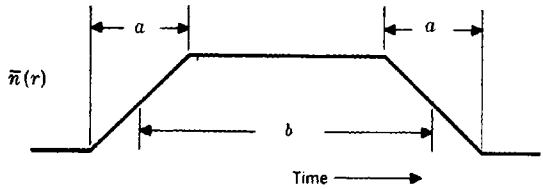


FIG. B·1.—Trapezoidal pulse.

pulse edges. Suppose the pulse sides sloped uniformly for a distance a as in Fig. B·1. The slope of n is related to the total number of particles N by

$$\frac{d\bar{n}}{dr} = \frac{N}{ba},$$

where b is a pulse length defined as the distance between the points where n is 50 per cent of its maximum value. Then we have

$$\left|\int_0 \frac{d\bar{n}}{dr}\, e^{-\frac{4\pi i r}{\lambda}}\, dr\right| = 4\,\frac{N}{ba}\,\frac{\lambda}{4\pi}\cos\left(\frac{2\pi b}{\lambda}\right)\sin\left(\frac{2\pi a}{\lambda}\right),$$

and the cross section for coherent scattering is

$$\sigma_c = |p^2|\left(\frac{N}{ba}\right)^2\left(\frac{\lambda}{4\pi}\right)^4\left[4\cos\left(\frac{2\pi b}{\lambda}\right)\sin\left(\frac{2\pi a}{\lambda}\right)\right]^2. \tag{20}$$

The ratio of coherent to incoherent scattering is then

$$\frac{\sigma_c}{\sigma_i} \leqslant N\left(\frac{\lambda}{2\pi b}\right)^2\left(\frac{\lambda}{2\pi a}\right)^2. \tag{21}$$

Even such a trapezoidal pulse is not a faithful reproduction of an actual pulse, as it possesses discontinuities in the first derivative $d\bar{n}/dr$. It can be seen from a comparison of Eqs. (17) and (18) that a discontinuity in $d\bar{n}/dr$ produces the same sort of effect as a discontinuity in $\bar{n}$

itself, although diminished by a factor proportional to λ. Such discontinuities are artificial. A realistic model of a pulse should be smooth in all its derivatives. A convenient form satisfying this characteristic is a pulse formed by an error integral curve reflected about R, as shown in Fig. B·2.

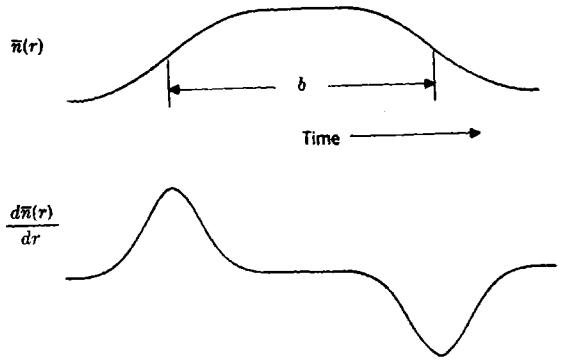


FIG. B·2.—Error-integral pulse.

For sufficiently sharp pulses,

$$\begin{aligned} \frac{d\bar{n}}{dr} &\approx \frac{N}{ba} \exp\left\{-\frac{\pi}{a^2}\left[r-\left(R-\frac{b}{2}\right)\right]^2\right\}, \qquad r < R, \\ \frac{d\bar{n}}{dr} &\approx -\frac{N}{ba} \exp\left\{-\frac{\pi}{a^2}\left[r-\left(R+\frac{b}{2}\right)\right]^2\right\}, \qquad r > R, \end{aligned} \tag{22}$$

with b again approximately the pulse width between the half-density points. Providing $b \gg a$ we can write

$$\int_0^\infty \frac{d\bar{n}}{dr} e^{-\frac{4\pi i r}{\lambda}}\, dr \approx \frac{N}{ba}\left[e^{\frac{4\pi i}{\lambda}\left(R-\frac{b}{2}\right)} - e^{-\frac{4\pi i b}{\lambda}\left(R+\frac{b}{2}\right)}\right] \int_{-\infty}^{+\infty} e^{-\frac{\pi r^2}{a^2}} e^{-\frac{4\pi i r}{\lambda}}\, dr.$$

From Eqs. (18) and (17) the coherent scattering cross section is then given by

$$\sigma_c = 4|p|^2 \left(\frac{N}{ba}\right)^2 \left(\frac{\lambda}{4\pi}\right)^2 \sin^2\left(\frac{2\pi b}{\lambda}\right) \left| \int_{-\infty}^{+\infty} e^{-\frac{\pi r^2}{a^2}} e^{-\frac{4\pi i r}{\lambda}}\, dr \right|^2.$$

The integration is easily effected by completing the square in the exponent, and the final result is

$$\sigma_c = N^2 |p|^2 \left(\frac{\lambda}{2\pi b}\right)^2 \sin^2\left(\frac{2\pi b}{\lambda}\right) e^{-\frac{8\pi a^2}{\lambda^2}}.$$

The ratio of the coherent to incoherent scattering then has an upper bound given by

$$\frac{\sigma_c}{\sigma_t} \leqslant N\left(\frac{\lambda}{2\pi b}\right)^2 e^{-\frac{8\pi a^2}{\lambda^2}} \tag{23}$$

In comparison with a trapezoidal pulse of approximately the same "sharpness" it is seen that the smooth pulse has a much smaller coherent scattering.

Name Index

A

Abbott, F. R., 361, 428
Adam, M. G., 688
Adamson, G. P., 592
Adel, A., 658
Alexander, F. E. S., 265, 370
Ament, S., 177, 476
Anderson, J. A., 533
Anderson, L. J., 285, 316, 328, 329, 334, 341, 361, 382, 428, 689
Anderson, P. A., 226, 283, 284, 286, 335, 341, 532
Arakawa, 201
Archer-Thompson, H., 419
Ashby, R. M., 540, 576
Atkinson, J. R., 363
Austin, L. W., 2
Autler, S. H., 669

B

Babcock, H. D., 653
Baker, B. B., 4
Baldwin, M. W., Jr., 593
Bales, P. D., 562
Baltzer, O. J., 475, 501
Barber, N. F., 486
Barker, C. B., 406
Barker, C. L., 283
Barlow, E., 557, 561
Barr, 476
Bateman, R., 341, 432
Bauchman, R. W., 226, 374
Becker, G. E., 669
Becker, J. A., 272
Beeching, G. H., 540, 541
Bent, Arthur E., 597, 621–640
Bergmann, P. G., 533
Beringer, R., 655, 656, 665
Best, A. C., 616, 679
Bethe, H. A., 465
Bigelow, H. B., 486, 490
Binnian, W., 226, 374
Blow, P. B., 476, 546
Boltzmann, L., 190
Bondi, H., 501
Booker, H. G., 175, 318, 370–373, 603
Born, M., 4, 451, 701
Börnstein, R., 190
Bothwell, F. E., 697
Braaten, A. M., 341
Braham, R. R., 630
Bremmer, H., 22, 402, 404
Brissman, D. N., 292
Brooke, N., 419
Broxon, J. W., 190
Brune, O., 641
Brunt, D., 198, 211, 259
Burgoyne, R. H., 223, 226, 249, 300
Burke, C. J., 228
Burrows, C. R., 40, 41, 53, 112, 125, 336, 374, 398
Byers, H. R., 181, 255, 618, 630

C

Carlson, J. F., 468
Carroll, T. J., 341, 432
Carson, John R., 698
Carter, R. H. A., 9
Chandreskhar, S., 553
Coales, J. R., 545
Cockroft, A. L., 491, 501
Condon, E. U., 652
Cooper, R. I. B., 543, 546, 562
Copson, E. T., 4
Corcoran, N., 540, 541
Cornish, V., 486, 488
Counter, V. A., 475
Cowan, E. W., 490, 501
Coxhead, A. B., 593
Craig, Richard A., 194–198, 220–264, 265, 266–271, 346
Crain, C. M., 191
Crawford, A. B., 336, 339, 374, 386, 388, 532, 595
Cross, P. C., 660
Crout, P. D., 697
Curran, J. E., 480

D

Dane, F. P., 316
Davies, H., 491, 501
Day, J. P., 316, 342, 350, 689
Debye, P., 13, 189, 609, 643
Decino, A., 336
Dennison, D. M., 657, 658
Dicke, R. H., 655, 668
Dieke, G. H., 653
Doble, S. M., 290
Dodson, H. W., 98, 113, 136
Domb, C., 98, 113, 125, 607
Dunmore, F. W., 284
Dunsmuir, R., 685, 686
Durkee, A. L., 341

E

Eccles, W. H., 2
Eckersley, T. L., 71, 151, 519
Edmondson, W. T., 486, 490
Elsasser, W. M., 271, 272, 346, 648
Emde, F., 458
Emmons, Gardner, 230, 236, 248
Englund, C. R., 336, 339, 374
Engstrom, J. A., 378
Eon, L. G., 592
Eyres, N., 175

F

Fairbank, J. D., 501
Fairbank, W. M., 475, 501
Ferrell, E. B., 53, 336, 374
Feshbach, H., 519
Fishback, William T., 112–140, 373–385, 396, 421–444
Fitzsimmons, K. E., 283, 335
Fleming, J. A., 2
Fletcher, J. O., 593
Ford, L. H., 398, 434
Frank, N. H., 417
Frank, P., 698
Freehafer, J. E., 1–3, 9–22, 41–112, 174–180
Freres, C. H., 689
Friend, A. W., 13, 339, 593, 594
Friis, H. T., 31, 593
Fry, T. C., 550
Furry, Wendell H., 71, 140–174

G

Gächter, K., 13, 191
Gamow, G., 71
Garrett, G. A., 502, 512
Gerdel, R. W., 268
Gerhardt, J. R., 378, 385, 595, 603
Gewertz, C. M., 641
Gibbs, W. F., 621
Gill, J. R., 98, 113, 136
Gilman, G. W., 593
Ginnings, D. C., 188
Ginsburg, N., 658
Gold, T., 419, 501, 579
Goldstein, Herbert, 419–421, 434, 445, 481–621, 671–692, 699–706
Goldstein, L., 451, 609, 673, 674
Goldstein, S., 208
Gordon, W. E., 272, 318, 387, 389, 594, 595, 603
Gordy, W. O., 475
Gorter, C. J., 643
Goudsmit, S. A., 468, 554
Gould, W. B., 593
Gray, M. C., 112, 125
Green, C. B., 272
Guillemin, E. A., 641
Gunn, R., 599

H

Haddock, F. T., 611, 612
Haeff, A. V., 471
Hainer, R. M., 660
Hamlin, E. W., 389, 390
Harrison, L. P., 183
Hartree, D. R., 175, 177
Harwood, Margaret, 564
Hatcher, R. W., 341
Haurwitz, B., 232, 255, 271, 279, 346
Hayes, W. A., 363
Heaviside, O., 2
Hector, L. G., 190
Hershberger, W. D., 9
Herzberg, L., 653
Hibbert, J. J., 512
Holter, N. J., 490
Hontgast, J., 533
Hopkins, M., 545
Howard, B. E., 98, 113, 136

Howlett, J., 175
Hoyle, F., 328, 412, 419
Hudspeth, E. L., 468, 475
Hull, Ross A., 339, 342
Hull, R. A., 688
Humphreys, W. J., 599, 680, 683, 684
Hunt, L. E., 336
Hurst, C., 688
Hutner, R. A., 113, 137, 440

I

Ingersoll, J. G., 654
Ingham, J., 363
Inglefield, S., 290

J

Jacques, R. B., 472
Jahnke, E., 458
James, R. W., 363
Jarkowski, H., 323
Jeffreys, H., 486
Jenkins, F. A., 4, 417
Jordan, A. R., 190
Josephson, V., 470

K

Kallman, H., 641
Katz, Isadore, 198–199, 202–207, 264–265, 272, 295, 336
Katzin, M., 226, 374, 476, 478
Kempton, A. E., 367
Kennelly, A. E., 2
Kerr, Donald E., 3–8, 22–41, 328–353, 361–373, 385–430, 445–481, 621–640, 693–698
Kessler, M.. 468
Kigenhuis, W., 641
King, A. P., 538, 688
King, G. W., 660
Knowles, F. F., 592
Kock, W. E., 386
Koller, L. R., 543, 546
Koscoe, G. F., 639
Kramers, H. A., 641
Kronig, R. de L., 641, 643
Kuhn, T., 480
Kuiper, G. P., 572, 573
Kyhl, R. L., 655

L

Lack, D., 368, 595
LaGrone, A. H., 389
Lamb, H., 346
Lamb, J., 685, 686
Lamb, W. E., Jr., 669
Lamont, H. R. L., 429, 655
Landolt, H., 190
Lane, J. A., 674
Langille, R. C., 592
Laws, J. O., 616, 617, 620, 680, 681, 683
Lawson, J. L., 540, 543, 576
Lee, Y. W., 641
Lettau, H., 208, 214
Lewis, W. D., 31
Linder, E. G., 688
Linford, L. B., 470, 471
Lorentz, H. A., 693
Lukes, George D., 199
Lyman, E. M., 512

M

McDonald, A. W., 476
MacFarlane, G. G., 177, 412, 501, 519
Maclusky, G. S. R., 501
McPetrie, J. S., 323
Mandeville, C. E., 468
Marconi, G., 1
Marcus, P. M., 467, 468
Margenau, H., 701
Mark, H., 641
Marshall, J. S., 592, 620
Martin, F. W., 540, 576
Mealey, K. L., 502, 512
Megaw, E. C. S., 323, 325–327, 342, 373
Meng, C. Y., 654
Merritt, F. R., 26, 658
Middleton, W. E. K., 182, 272, 280
Mie, G., 674, 687
Miller, J. C. P., 95
Mimno, H. R.. 3
Minnaert, M., 533
Mitchell, R. M., 295, 336
Moglich, F., 687
Montgomery, R. B., 181–189, 200–202, 208–223, 226, 230, 231, 249, 268. 292, 300
Morse, P. M., 457, 460
Mueller, G. E., 654, 655, 689
Mumford, W. W., 336, 339, 374

Munk, W. H., 486
Murphy, G. M., 701

N

Namias, J., 260
Nash, J. P., 373
Neelands, L. J., 295, 336
Nicholson, J. W., 411
Nicholson, P., 175
Noether, F., 3
Norton, K. A., 40, 112, 350, 416

O

O'Brien, M. P., 486, 487
Oliver, R., 398, 434
Omberg, A. C., 416
Osborne, N. S., 188

P

Palmer, W. M., 592
Parke, N. G., 318, 603
Parker, F. D., 113
Parsons, O. A., 616, 617, 620, 680, 681, 683
Pearcey, T., 175, 178, 179, 316, 334
Pearson, G. L., 272
Pekeris, C. L., 12, 111, 175, 177, 180, 317, 374, 378, 385
Petterssen, S., 263
Pryce, M. H. L., 51, 98, 113, 125, 412
Purcell, E. M., 468, 664–671

R

Rado, G. T., 616, 689
Ramsay, J. A., 367, 476, 546
Ramspeck, G. D., 378
Randall, H. M., 658
Rayleigh, Lord, 519, 553
Renwick, W., 501
Revelle, R., 361, 428
Riblet, H. J., 406
Rice, S. O., 412, 553, 559
Robertson, S. D., 538, 688
Rodgers, R. A., 592
Rossby, C.-G., 214, 218
Rubenstein, P. J., 34, 112, 189–193, 294–322, 341, 354–361, 385–395, 421, 430, 457
Ryde, D., 609, 674–683, 685–688, 692
Ryde, J. W., 609, 611, 613, 614, 618–620, 674–683, 685–688, 692

S

Sänger, R., 13, 191
Saxton, J. A., 610, 667, 674, 675
Schelkunoff, S. A., 28, 31, 698
Schelleng, J. C., 53, 336, 374
Schiff, L. I., 519, 554
Schlapp, R., 649, 653
Schultz, H. L., 190
Schuster, A., 411
Scott, J. M. C., 533
Seay, P. A., 272
Sharpless, W. M., 386, 388, 432. 532
Sherwood, E. M., 427, 433
Shortley, G. H., 652
Sicinski, L., 323
Siegert, A. J. F., 519, 556, 574, 598, 699–706
Sinclair, G., 472
Slater, J. C., 417, 698
Smith, R. A., 491, 501
Smith-Rose, R. L., 325, 341, 347
Smyth, J. B., 315, 334, 346, 350. 361, 428, 689
Sollfrey, W., 557
Sommerfeld, A., 3, 40
Southworth, G. C., 538
Spencer, R. C., 466, 468
Spilhaus, A. F., 278
Starnecki, B. J., 312, 323, 327
Steiger, O., 13, 191
Stephenson, S. T., 283, 285, 329, 335
Stickland, A. C., 325, 341
Stimson, H. F., 188
Stokes, A. P. D., 285, 329, 689
Straiton, A. W., 385, 390, 398, 532
Stranathan, J. D., 13, 191
Strandberg, M. W. P., 654
Stratton, J. A., 4, 61, 396, 445, 454 455, 599, 601, 671, 674, 694
Stumpers, F. L., 641
Sutro, 480
Sverdrup, H. U., 212, 219, 486

T

Tangl, K., 190
Terman, F. E., 398
Tibbler, L. G., 592
Tolbert, C. W., 398
Tomlin, M., 316
Townes, C. H., 26, 658
Tregidga, A. C., 191

Trevor, B., 9
Trolese, L. G., 334, 342, 346, 350, 361, 689

U

Uhlenbeck, G. E., 553, 557, 558
Ursell, F., 486

V

van der Pol, B., 22, 402, 404
Van Vleck, J. H., 189, 641–664
Vane, A. B., 655
Varley, G. C., 369
von Mises, R., 698

W

Walkinshaw, W., 480
Walz, F. C., 190
Wang, Ming Chen, 553, 558
Waterman, A. T., Jr., 387
Watson, A. G. D., 429
Watson, G. N., 3, 22
Weber, L. R., 658
Weiss, P. R., 468
Weisskopf, V., 650
Westwater, F. L., 326, 363
Wexler, R., 595
White, H. E., 4, 417
Wickizer, G. S., 341
Wilkes, M. V., 476, 546
Williams, D., 470
Willis, F. H., 593
Wolff, I., 688
Wood, R. W., 533
Woodcock, A. H., 269, 533
Woodcock, W., 470
Woodward, A. M. W., 95
Woodward, P. M., 95

Y

Younker, E. L., 686

Z

Zahn, C. P., 191
Zenneck, J., 2, 3

Subject Index

A

Absorption, atmospheric, by oxygen, 26, 646–656, 663
  by water vapor, 26, 646, 647, 656–664
  dielectric, 24, 676–678
  nonresonant, 643, 652–653
  by particles, 23, 676–678
  relation of, to index of refraction, 641–646
  resonant, 644, 646, 653
Absorption coefficient, 24, 647, 650
  for oxygen, 647, 651–656, 663
    measurements of, 654–656, 665
    theory of, 648–656
  for water vapor, measurements of, 659, 666–671
    theory of, 656–664
Additivity rule, 190
Adiabatic lapse rate, 194
Adiabatic process, 186, 193
ADRDE-CAEE, 479–481
Advection fog, 230
Aerometeorographs, 347
Aeropsychrograph, 272, 297, 298
  on PBY aircraft, 276
Aeropsychrograph housing, 274, 275
Air, dry, density of, 292
  homogeneous, 193, 207
  molecular weight of, 292
  specific heat of, 293
  unstable, modification of, 226
  warm, over cold water, 237–250
    probable modification processes for, 249–250
Air Defense Research and Development Establishment (ADRDE), 175, 479–481
Air-mass boundary, 339
Aircraft echoes, azimuth variation of, 543
  propeller modulation of, 539–541
  slow fading of, 543
Aircraft psychrometer, 287, 288
Airy integral, 95, 149, 150
  asymptotic expansion of, 150
"Angels," 593–595
Angle, of arrival, affecting radar echoes, 532
    measurements of, 386–391
    by ray theory, 391
  of penetration, 17
Antenna gain, definition of, 28–29
Antenna pattern, cosecant-squared, 38
  symmetrical, 38
Antenna pattern function, 27–31, 116, 437–440, 482, 589, 629
  definition of, 27
Antenna pattern symmetry in pattern-propagation factor, 38
Antennas, metal-lens, 386
Anticyclone, Pacific, 262
  subtropical, 261
Anticyclonic conditions, 347
Antigua, B.W.I., 226
  experiment in, 374
Anti-Stokes lines, 73
  for Eckersley modes, 84
Assembly of random scatterers, 551, 554
  average echo from, 589–591, 699–700
  coherence of scattering from, 601, 699–706
  frequency sensitivity of echo from, 529
  theory of fluctuation from, 553
Asymptotic expansion of Airy integral, 150
Asymptotic representations of solution of wave equation, 70, 75–77, 92–95, 98, 101, 105, 110
Atmospheric absorption (*see* Absorption, atmospheric)
Atmospheric inhomogeneities, as cause of echo fluctuations, 531–535
  meteorological evidence for, 267–271, 533, 595, 603
  as producing radar echo, 594, 600–604, 700, 703
  theory of effect on echo fluctuations, 533–535
  wavelength dependence of resulting transmission instability, 534

Atmospheric waves, 346
Attenuation, by atmospheric gases (*see* Absorption)
by fog and clouds, 24, 682
by hailstones, 25
within meteorological echoes, 613–615
by precipitation, measurements of, 688–692
by rain, connection of, with back scattering, 615
for small drops, 676–677, 682
wavelength dependence of, 682, 684
by solid precipitation, 685–688
Attenuation constant, in clouds or fog, 672–673, 677
for hailstones, 687
for ice and snow, 687–688
for precipitation, 672
approximate formulas for, 676
limits on, 673–674, 680, 691
relation of, to precipitation rate, 673, 680, 683–684, 691
rigorous values for, 678–680, 682–684
temperature dependence of, 677, 685
for very small drops, 676–677
for rain, anomolous values for, at 1.25 cm, 691–692
experimental values for, 691
for various rains, 682–684
Attenuation function, 122
for first mode, 126, 161, 168–174
Audio-frequency spectrometers, 562, 576
Austin-Cohen formula, 2, 3

B

Back-scattering cross section, definition of, 33
Barrier, 76
nonstandard layer as, 168
Beam width, affecting sea echo, 483–485, 491
Bell Telephone Laboratories, 336, 348, 386, 432
Bilinear M-profile and angle of arrival, 393
comparison of, with theory and experiment, 307–313
fitting of, to data, 308
Bilinear modified-index profile, analytical statement of, 142
Bilinear N-profiles, examples of, 143
Birds, echoes from, 595
Black-body radiation, by clouds, 254
by earth, 254
Blue Hill, 296, 551
Boundary conditions, at earth's surface, 58, 62–64, 111–112
for meteorological quantities, 201, 255
simplified, 140
at infinity, 71, 76
statistical nature of, 430, 527
for targets, 445, 448, 454, 456, 458, 462, 476
for trapped modes, 78
Boundary error in correlation function, 570–572
Boundary friction, 221
Boundary layer, turbulent (*see* Turbulent boundary layer)
Branch cuts, 66, 109
Branch points, 66, 109
Brewster's angle, 399
Bright band from snow echoes, 620
British Meteorological Office, 288, 289, 341

C

Calibration of receiver output, 564–566
Cambridge Field Station of Watson Laboratories, 594
Cambridge University, 175, 177
Canadian Army Operational Research Group, 592
Cardigan Bay, 323
Caustic, 19, 58, 391, 393
Central Radio Propagation Laboratory, 350
Chaff, 550
first probability distributions for echo from, 571–575
fluctuations in echo from, 551, 572–575
frequency spectra of echoes from, 572
mean spread in velocity of, 574
voltage correlation function of, 567
Characteristic, 20
of ray, 49
Characteristic curve, 200–202
modification of, by varying water temperature, 252

Characteristic diagram, 184, 189, 191, 200–203, 236, 300
Characteristic functions, 176
for first mode, 161
Characteristic values, 21, 68, 70, 71, 176
for bilinear N-profile, 146, 162
for first mode, 161
for large values of q and $s = +\sqrt{\frac{1}{4}}$, 171
methods for calculating, 146
for s nearly unity, 163
of trapped modes, 75
Clapeyron's equation, 183, 279
Cloud cover, 255
Clutter echoes, definition of, 550
measuring techniques for, 562–571
nature of fluctuations in, 550–553
theory of fluctuations from, 553–562
Coherent scattering, from assemblies of random scatterers, 601–607, 699–706
magnitude of, from drops, 600, 605–607
Cold fronts, 263
echoes from, 23, 630–632
Conduction (thermal), 213, 219
Conductivity (thermal), 208, 209, 293
Connection formulas, 71, 75
Conservative property, 186, 200, 206, 207
Contour diagram, 32
free-space, 30
Contours, of constant field strength, 130
of field strength in surface duct, 178, 179
Convection, 193, 220, 268, 283
Convective mixing, 225 253, 258
Cooling, from below, 220, 228–231, 251, 254, 300
over land, 234
over water, 232
of initially stratified air over water, 251
nocturnal (*see* Nocturnal cooling)
Correlation function, in amplitude, 557, 567–569
connection of, with frequency spectrum, 559
definition of, 556
effect of steady signal on, 562
errors in measurement of, 569
for finite sample, 566, 568
measurement of, 566
for nonstationary signals, 567–569
of ln R, 558
in power, 568
in presence of noise, 559, 566
Correlation function, statistical fluctuations in, 568, 569
and velocity distribution of scatterers, 556
Correlations, between angle of arrival and M-profile, 388
between one-way transmission and refraction, California coast, 328–335
Flathead Lake, 335
Irish Sea, 322–327
Massachusetts Bay, 301–322
overland circuits, 336–353, 341 (table)
South Wales to Mt. Snowdon, 328
between radar echoes and M-profiles, 355–361, 365–368
between radar echoes and one-way transmission, 10, 361, 367
between radar echoes and precipitation, 591–592
between radar echoes and superrefraction, 363, 367–373
Coverage, distortion of, by refraction, 9–11, 366, 369, 370, 374–385, 427, 473
and specular reflection from earth, 406
statistical properties of, 430
wavelength dependence of, 409
Coverage diagram, 32
coarse, 130, 139
approximate calculation of, 138–140
free-space, 32
generalized equation for, 36
for plane earth, 39
radar, free-space, 34
Cross section, effective, of antenna, 31
of raindrops, 608–612
approximate expressions for, 609
tables of, 612
per unit area for sea echo, 483
per unit volume for echoes from ions, 599
for meteorological echoes, 590, 618–621
Cusps, 393

D

Debye formula for dielectric constant, 609, 643, 652, 675
Deer Island, 296, 297
Depolarization, 389, 457, 468
by rough surfaces, 539
by scattering cylinder, 462

Depth-loss, 169
Depth-loss effect, 168, 170
Dew point, 182, 186, 207
  potential, 198, 200
Dew-point hygrometer, 182, 186
Dielectric constant, complex, 397, 448
    properties of, 641
  table of, 398
    for water, 674
  of air, experimental information on, 190
  of earth, 111, 398
  general properties of, 641–646
  properties of real and imaginary part of, 642
  relation of, to absorption, 641–646
  resonant and nonresonant, 643
  of sea water, 111, 398
  for water, 398, 609–613, 674–675
  (*See also* Index, of refraction)
Diffraction by earth's surface, 6, 34
Diffraction region, 5, 6, 109
Diffusion, 213, 248, 252
Diffusion theory, 175, 238
Diffusivity, 209, 293
Diplane, 11
Dipole moment, 13, 190, 451, 461, 599, 609, 646, 649, 650, 652, 658, 675, 677
Direct ray, 5, 36, 37, 38, 99, 100, 104, 107 116
Disdrometer, 617
Diurnal cycle, 253–260
Diurnal effects, 337–339, 343, 351, 368
  on California coast, 331
  in Irish Sea, 321, 324
  in Massachusetts Bay, 321
Divergence factor, 40, 99, 113–115, 137, 138, 401, 404
Diversity effects, and average signal level, 306
  frequency, 304, 306
  spatial, 304, 306, 345, 420, 432
Diversity reception, 136
Dominant terms, 73, 103
Drop-size distribution, 672, 680–681
  method of measuring, 615–618
Droplet theory of sea echo, 494, 497, 519, 522–527
  difficulties of, 526
  polarization dependence of, 523–525
Drops, attenuation by, 615, 678–685
  cross sections of, 608–613
Dry-adiabatic lapse rate, 194, 280
Duct, 9, 14, 17–22, 143
  associated with M-inversions, 15
  conditions for, over water, 9, 223–226, 228–234, 254–260
  elevated, 14, 18
  surface (*see* Surface duct)
  trapping in, 21
  (*See also* M-inversion)
Duct height, 21, 225, 238, 243
  definition of, 224
  and height of temperature inversion, 244
  ratio of, to M-deficit, 225, 241, 243
Duxbury, Mass., 298

E

Earth, effective radius of, 6, 53, 115, 339, 392, 406
Earth curvature, 40, 48, 50
Earth-flattening method, Maxwell's equations for, 51
Eastern Point, Gloucester, Mass., 295–297
Echo box, nonresonant, 669–671
Eckersley modes, 71, 84
  generalized discussion of, 85–87
  phase-integral method for, 87, 151
Eddy diffusion, 208, 219, 232, 238, 245
Eddy diffusivity, 208, 211, 213, 218, 238
  height distribution of, 219, 239
  ratio of, to eddy viscosity, 211–212
Eddy stress, 210, 216, 218
Eddy transfer of heat and water vapor, 211, 212
Eddy viscosity, 208, 210, 213, 214, 217, 218
  dynamic, 210
  ratio of, to eddy diffusivity, 211–212
  Rossby's distribution of, 215
Edge effects in scattering, 457, 462, 464, 465
Eikonal, 41
Ekman spiral, 213, 214
Electrolytic hygrometer, 182, 284, 286
Ellipsoid, scattering by, 464
Elliptic paraboloid, scattering by, 464
Emagram, 188
English Channel, 367

Equilibrium, neutral, 213, 223, 226, 238
definition of, 195
stable, 220, 228, 234, 238
definition of, 196
fluctuations in, 270–271
shear in, 234–237
unstable, 220, 225, 235, 268
definition of, 196
fluctuations in, 268
Euler's equation, 45
Evaporation coefficient, 224
definition of, 219
Evaporation duct, 329, 336, 364, 374
Expanded elevation indicator, 627
Extra-optical path, 6

F

Fading, 302, 324, 335, 339, 340
amplitude of, 307, 336, 343, 346, 348, 350
early studies of, over land, 337
effect of path geometry in, 343
on Mt. Snowdon, 328
on 1.25 cm, 312
period of, 304, 350
proposed mechanisms for, over land, 339
of radar echoes, 354, 363, 367
as random process, 313–315, 328
roller (*see* Roller fading)
of ship echoes, 543
synchronism of, 307, 328, 336, 345, 346
wavelength dependence of, 307, 324, 343, 348
Federal Communications Commission, 350
Fermat's principle, 6, 45, 99
Fetch, 488
Field attenuation, 30, 31
Field strength, constant, contours of, 130
contours of, in surface duct, 178, 179
from dipole, 21, 58
integral for, 95, 101
free-space, 29
from $\lambda/2$ doublet, 30
level surfaces of, 42
standard, definition of, 14
with standard refraction, 112
Field-strength calculation, for bilinear profile, 168
sample of, in interference region, 117–122
at total horizon range, 125
Figure of merit, 135
Filter-paper method for drop-size distribution, 615
Finite number of scatterers, fluctuations from, 515, 554
First-hop length of ray, 83
First probability distribution, in amplitude, 554
for chaff echoes, 571–572
definition of, 514, 554
for finite number of scatterers, 554
for ground clutter, 582–583
in power, 555
for precipitation echo, 575
in presence of secular variation, 518
for random scatterers plus steady signal, 560–561
for sea echo, 515–516, 579
Flat-earth formulas, region of validity of, 121–122
Flathead Lake, 335
Flour method for drop-size distribution, 616
Fluctuation frequency spectrum, as Doppler beats, 516, 575
for sea echo, 516–518, 579–581
Fluctuations, in M, 268–271
of radar echoes, from aircraft, 539–543
due to atmospheric variations, 531–535
from chaff, 551, 572–575
from complex targets, 547–550
experimental techniques for, 562–571
origin of, 527
from precipitation, 551
from ships, 543–547
from two interfering simple targets, 549
due to varying ground reflection, 539
due to varying water reflection, 535–538
of sea echo, 486, 514–518, 551
in stable equilibrium, 270–271
in unstable air, 269, 270, 375
in unstable equilibrium, 268
Focus, 19, 58, 391, 393
Fog, 230, 326
advection, 230
and fading, 346
vs. poor visibility and substandard transmission, 324
Foliage, 584

Forecasting of propagation conditions, 368
Free-space range, 31, 33
  for one-way transmission, 32
Free-space transmission, 27–34
Frequency dependence of sea echo, 494–499, 511, 524
Frequency sensitivity of radar targets, 529–531
Frequency spectrum, for chaff echoes, 572–575
  connection of, with correlation function, 559
  definition of, 558
  dependence of, on pulse length, 579
  for ground clutter, 582–587
    anomalous shapes of, 587
    effect of wind speed on, 585
  for sea echo, 579–582
Frequency stability of radar systems, 528
Fresnel diffraction, 412
Fresnel equations for specular reflection, 396
Fresnel integrals, 477
Fresnel zones, 417
  on targets, 464, 468
Frictional influence, layer of, 213, 231
Frontal inversion, 263
Frontal passage, 264
Frontal surface, 263
Fronts, 263, 319, 325, 331, 347

G

Gain of antenna, 29, 31
Gamow mode, 83, 309
Gamow phase-integral method, 151, 167
Gaussian distribution of clutter echoes, 555
Geometrical optics, 4, 6, 41, 50
  in back scattering, 461, 465–468
Geometrical parameters for standard propagation in interference region, 116
Geostrophic wind, 213
Gradients, horizontal, 267
  of refractive modulus, 192, 198
  of temperature, 192
  of vapor pressure, 192
  vertical, 193
Gravitational waves, 346
Grazing angle, 108, 114, 115
Grazing path, 6
Geogory humidiometer, 289
Ground-controlled approach (GCA) radar, 432
Ground clutter, definition of, 551
  division of, into moving and stationary targets, 552, 582
  first probability distributions for, 582–583
  fluctuations of, 551–553, 582–587
    effect of wind speed on, 551, 584
  frequency spectra for, 584–587
Ground reflections, effect of, on echo fluctuations, 539

H

Hailstones, 25
  attenuation by, 25, 687
  echoes from, 619
Hair hygrometer, 182
Half-period zones, 413
Hankel functions, modified, of order one-third, 94
Harvard University, 339
Haze layers and signal fluctuations, 346
Heat, latent, 183
  of vaporization of water, 188, 293
Heat capacity of air, 293
Heating from below, 193, 220, 221, 227, 228, 300, 364
  over land in daytime, 222
  over tropical oceans, 222, 223
Height, natural unit of, 96
    as function of wavelength, 123
Height-gain effect, 10, 304, 307, 350, 375, 378
Height-gain function, 22, 112, 122, 168
  differential equation for, 145
  for first mode, 128
  of first mode for bilinear N-profile, 164–166
  normalization of, 65, 78, 86, 141–142
  of second mode for bilinear N-profile, 167
  unnormalized, 78
Height-loss effect, 10
Height measurements by radar, errors in, from refractive effects, 385, 436
    from specular reflection, 436
Hertz vector, 58–60

Hertzian dipoles, free-space transmission equation for, 32
High-resolution appearance of sea echo, 491
High wind conditions, 347
Homogeneous air, definition of, 193
Homogeneous layer, 221, 222, 226
Horizon, 5
Horizon ray, 17
Horizontal velocities of raindrops, 577
Humidiometer, 289
Humidity, 182, 206
  absolute, 182
  relative, 182, 184, 206
  specific, 185, 187, 188, 206
    in turbulent boundary layer, 218
Humidity deficit, 224, 229
  definition of, 222
Hurricanes, 23, 636–640
Huygens' principle, 412, 454
Hydrostatic equilibrium, stable, 220
  unstable, 220
Hygrometer, dew-point, 182, 186
  electrolytic, 284, 286

I

Ice, index of refraction for, 686
Ice-bulb thermometer, 283
Ice crystals, attenuation by, 687–688
  echoes from, 618
Ice point, 292
Incoherent scatterers, 588–591, 699–706
Index, modified (*see* Modified index)
  of refraction, of air, 6, 189
    atmospheric, constancy of, in microwave region, 641, 643
    dependence of, on meteorological parameters, 13, 191
    for ice, 686
    for water, 609–613, 674–676
  (*See also* Dielectric constant)
Inhomogeneity of rain, 24, 672, 684, 689
Insects, echoes from, 595
Instability of radar systems, causing echo fluctuations, 527–529
  limits of, 528
Interference effects, as source of echo fluctuations, 535–538, 543–545, 547–550
Interference pattern, 5, 38–40, 113–122, 407–410, 421–444
  fluctuations of, 535
Interference region, 5, 6, 113–122, 408
Intermediate region, 125, 409
Ions, echo from, 598
Irish Sea experiment, 322
  results of, summary of, 323
Isobaric cooling, 186
Isotropic antenna, definition of, 28

J

Jordan's lemma, 110

K

Kinematic viscosity, 208, 214, 293
Kite ascents, 262
KXS trials, 327, 363

L

Lag coefficient, 280
  of ML-313, 287
  for psychrograph, 280–281
Laminar sublayer, 209, 213, 219
Laplace integral, 88
Lapse rate, 193
  adiabatic, 194
  of dew point in homogeneous air, 196
  dry-adiabatic, 194, 280
  moist-adiabatic, 194, 279
  saturation-adiabatic, 194
  in subsidence, 260
  of vapor pressure, 196
  of wet-bulb temperature in homogeneous air, 197
Latent heat, 183
  of vaporization of water, 188, 293
Layer, elevated, 14
    (*See also* Duct, elevated)
  of frictional influence, 213
    outer, 214
  homogeneous, 193, 199, 231, 235
  nonstandard, as barrier, 75, 76, 83, 168
  saturated, 207
  standard, 14, 235
  substandard (*see* Substandard layer)
  superstandard (*see* Superstandard layer)
  surface (*see* Surface layer)
Leaking modes, 22, 71, 84, 333, 378, 382

Length, natural unit of, as function of wavelength, 124
Line of sight, 5
Line-breadth constant, 650, 654
  for oxygen, 654–656
  for water vapor, 658–664
Line broadening, 645, 650
Linear-exponential M-profile, contours of field strength for, 178–180
Linear-exponential profile, 169, 174
Linear modified-index profile, 14–16, 87
Lloyd's mirror, 4
Lobe-switching, 436
Lobes, 39, 40, 407–410
Loci of $A_{(m)}^{(\infty)}(s)$ and $A_{(m)}^{(\infty)}(s)$ in complex plane, 158
Locus of maxima, 137
Logarithmic distribution, 223
  of specific humidity, 218
  of temperature, 223
  in turbulent boundary layer, 215
  of velocity, 217
  of water vapor, 223
Long-wave transmission, proposed mechanisms for, 1
Lorentz reciprocity theorem (*see* Reciprocity theorem)

M

M, fluctuations in, 268–271
M-deficit, 222, 229, 230, 234, 237–239, 244, 245
M-gradient, standard, definition of, 14
M-inversion, 14
  definition of, 15
  elevated, 18, 223, 249, 261–266, 302–304, 329–335, 339, 359, 361, 362, 364, 366, 372, 373, 382
    overhanging, 304, 332
    and 256 Mc/s transmission, 315
  shallow, 224
  with subsidence, 261
M-profile, bilinear (*see* Bilinear M-profile)
  characteristic values for, 146–163
  definition of, 14
  linear-exponential, 312
  power-law, 312
  standard, definition of, 14
Manchester, University of, 175, 177
Mapping, method of, for obtaining characteristic values, 146
Mark VI aircraft psychrometer, 288, 289
Massachusetts Bay, 226, 295, 296
  meteorological measurements and analysis of, 297
  radio measurements of, 296
M.I.T. Weather Radar Project, 577, 617
Maxwell's equations in earth-flattening method, 51
Mechanical integration by differential analyzers, 177
Mechanical mixing, 238, 252
  and temperature excess, 231
  and wind speed, 231
Metal-lens antennas, 386
Meteor expedition, 262
Meteorograph, 337
Meteorological conditions, nonuniform, 315, 318–319, 368
Meteorological constants, 292
Meteorological echoes, average echo from, 589–591
  from causes other than precipitation, 593–595
  and coherent scattering, 605–607
  correlation of, with precipitation, 591–592
  cross section of, per unit volume, 590
    absolute measurements of, 597–598
    approximate calculation of, 596–597
  dependence of, on antenna area, 591
    on beam width, 590
    on drop size, 596
    on pulse length, 590
  magnitude of, from atmospheric inhomogeneities, 600–604
  nondrop theories of, 598
  from solid precipitation, 618
  validity of Rayleigh scattering law for, 598
  variants of drop theories of, 604–608
  (*See also* Precipitation echoes)
Meteorological Instruments, 272
Microwave spectroscopy, 25
Migratory highs, 261
Millimeter waves, absorption of, by water vapor, 664
  raindrop radar cross sections for, 612
Mixing, 200, 201
  convective, 225
  mechanical (*see* Mechanical mixing)
Mixing ratio, 185, 187, 188, 200, 206

ML-313 psychrometer, 287
lag coefficient of, 287
Mode, single, in duct, 19
Modes, elementary, 18
Modification, of air over water, for long trajectories, 248–250
height of, 238, 241, 242
Modification cross section, 239–241
Modified index, 12, 50, 53, 191, 192
definition of, 12
inadequacy of, over land, 352
Modified-index gradient, 192
Modified-index profile, bilinear, analytical statement of, 142
idealizations and definitions of types of, 14
nonlinear, 174
Modulus, potential, 198–200, 207
definition of, 199
Modulus deficit, 224
definition of, 222
Moist-adiabatic lapse rate, 194, 297
Monsoons, 371, 372
Mt. Snowdon, 328
Multiple moments, of cylinder, 461
of raindrops, 608–610
of sphere, 7, 451–452
Multiple-ray paths, 388, 390, 395
Multipoles on raindrops, 676

N

N-inversion, definition of, 15
N-profile, 14
bilinear, characteristic values for, 162
examples of, 143
National Physical Laboratory, 325, 341
Natural units, advantages of, in computation, 136
Naval Research Laboratory, 226, 374, 477, 480
Navy Electronics Laboratory, 262, 328, 350, 382
Navy Radio and Sound Laboratory, 328, 341
Negative heights, 444
Newton-Raphson method for obtaining characteristic values, 146
Nocturnal cooling, 221, 253–257
Nocturnal inversion, 346, 348
Nocturnal radiation, 220, 232, 234, 254, 359
and elevated M-inversions, 316
Nocturnal surface inversions, 339, 348
Noise, detection of radar echoes in, 471
effect of, on first and second probability distributions, 559
on receiver output calibration, 565
Nonlinear receiver characteristics, effects of, on measurement accuracy, 564–571
Normalization, of height-gain function, 141–142
Numerical integration of Riccati equation, 177

O

Occlusions, 347
Office of Chief Signal Officer, 341
Optical path, 6
Optical path length, 47, 50, 99–100, 104, 107
Optics, geometrical (*see* Geometrical optics)
physical, 4, 58
Over-water transmission, general characteristics of, for Massachusetts Bay, 301–307
Overland transmission, general features of, 340
Overland transmission paths, 341
Oxygen, 26
absorption by, 646–656, 663
magnetic dipole transitions in, 649
rotational spectrum of, 649, 653–654
selection rules for, 649–650
Oxygen absorption, measurements on, 655, 656, 665–666

P

Paraboloid, elliptic, 464
of revolution, 464
Path length, optical (*see* Optical path length)
Path-length difference, 37, 108, 114, 115
Paths of integration, in normalization, 141, 177
in κ-plane, 68
in t-plane, 89

Pattern-propagation factor, 34, 117
definition of, 35
in diffraction region, 109, 122, 141
in interference region, 38, 99, 115–117
in natural coordinates, 98
sample calculation of, for diffraction region, 124
extending through intermediate region, 133–135
simple properties of, 36–41
Perturbation calculations, 177
Phase constant of first mode, 162
Phase front measurements, 389
Phase-integral methods, 70, 177
for Eckersley modes, validity of, 87
relation of, to ray tracing, 82
Phenomenological description of attenuation by precipitation, 671–674
of average signal strength, from meteorological echoes, 589–591
from sea echo, 482–486
Photography of A-scope for study of echo fluctuations, 562
Physical optics, 4, 58
Pilot balloon observations, 298
Pip-matching, 436
Point of reflection, 113
Point Loma, San Diego, 329
Polarization, 140
as affecting σ, 33, 461, 462, 468, 471, 475, 696–699
affecting sea echo, 498, 522, 527
and boundary conditions at earth's surface, 58, 63, 111, 112, 124, 140
circular, 697–699
in scattering from smooth targets, 468
effect of, on reflection coefficient, 40, 98, 397–401, 408, 423–424, 427–429, 434
in refraction studies, 297, 337, 343, 351, 389
elliptical, 696–697
purity in, relation of, to reflection coefficient, 426
of radiated wave, 29
Polybutylmethacrylate, 273, 281
Potential temperature, 194, 195, 200, 202
definition of, 194
Power-law M-profile, 312
Power-law profiles, 174
Precipitation, attenuation by, 671–692
as causing meteorological echoes, 581–598, 604–615, 618–621
effects of, on transmission, 339
onset of, 339
Precipitation echoes, 22, 621
band structure of, 632, 637
confusion and masking of, 625–626
first probability distribution for, 575
fluctuation of, 551, 575–578, 588
frequency spectra for, 575–578
from hurricanes and typhoons, 636–640
identification of, 621–625
from layers, 634–636
magnitude of, 596–598
meteorological structure of, 626
probability distribution for ratio of successive amplitudes, 578
self-absorption of, 613
variability in frequency spectra of, 575, 576
from widespread rain, 633
Precipitation rate and attenuation by precipitation, 673, 683, 691
Pressure gradient, 193
Pressure recorders, 490
Principle of similarity, 201, 232, 233, 238, 240, 243, 268
Probability distribution for ratio of amplitudes, 558, 577–578
(*See also* First probability distribution; Second probability distribution)
Propagation, nonstandard, geographical distribution of, 369
Propeller modulation, 539–543
Provincetown Standpipe, 547, 548
Pseudoadiabatic chart, 188
Psychrograph, 272
circuit diagram of, 277
lag coefficient for, 280–281
Psychrograph housing for ground-based soundings, 283
Psychrometer, 182
aircraft, 287
ML-313, 287
Psychrometric formula, 187
Pulse integrator, 496
Pulse length affecting sea echo, 491
Pulse shape, and magnitude of coherent scattering, 605–607, 704–706
in sea echo, 484

R

Race Point, Cape Cod, 295, 296, 298
Radar, ground-controlled approach (GCA), 432
Radar cross section, 33, 445
of aircraft, 470–472
method of measurement of, 470
polarization dependence of, 471–472
table of, 472
wavelength dependence of, 471–472
of circular disk, 458
of complex targets, 469
of cylinders, 458–462
definition of, 33
dependence of, on polarization, 698
differential, 466
of flat plate, 457
from flat-plate model, 476–479
general formula for, 455
of sea echo (*see* Sea echo)
of ships, 472
aspect dependence of, 479
far-zone, 478
incoherent scattering of, 479–481
near-zone, 478
polarization dependence of, 475
table of, 475
wavelength dependence of, 475, 479
simplified formula for, 462
of small objects, 465
of smooth target, 697
of sphere, 450, 452–453
statistical nature of, 470
per unit volume, for precipitation echoes, 596
definition of, 590
Radar detection near surface, range dependence of, 474
wavelength dependence of, 409, 474
Radar detection ranges and duct height, 358
Radar echoes, fluctuations of (*see* Fluctuations, of radar echoes)
long-range, 9, 358, 367–373
from precipitation, 22
(*See also* Precipitation echoes)
vs. range, for ships, 474
steady, examples of, 547
Radar equation for ships, 473
Radar measurements program, Massachusetts Bay, 354
Radar range, free-space, 33
Radar set, for measurements of sea echo, 495
SCR-584, 444
Radar transmission equation, for elliptical polarization and smooth target, 697
for flat-plate ship model, 479
free-space, 33
for sea echo, 484
Radiation, 232, 234, 248
nocturnal (*see* Nocturnal radiation)
Radiation cooling, 343
Radiation temperature of space, used in water-vapor absorption measurements, 669
RCA Laboratories, 338, 344
Radio meteorological forecasting, 175
Radio-optical path, definition of, 6
Radio Research Laboratory, 480
Radiometer, microwave, 668-669
Radiosonde, 207, 255, 272, 298, 347, 348
Radiosonde ascents, 261, 348
Rain, attenuation in, 673, 678–685
incorrect ideas about, 685, 691
measurements of, 688–692
wavelength dependence of, 684
echoes from, 591–592
magnitude of, 596–598
(*See also* Meteorological echoes; Precipitation echoes)
nonuniform spatial distribution of, 672, 684, 689
Rain gauges, 689
Raindrops, attenuation constant for, 678–685
multipole moments of, 608–610, 676
and radar cross section, 608–613
size distributions of, 615–618, 681
terminal velocity of, 679
Random processes, theory of, 553
Random scatterers, 551, 553
assembly of (*see* Assembly of random scatterers)
plus steady signal, first and second probability distributions for, 560–562
Random-walk problem, 553
Range, natural unit of, 97
Range-height indicator, 630, 635–636
Ray, 4, 12, 54
characteristic of, 15, 49
curvature of, in relation to index gradient, 44

Ray, definition of, 6
differential equation for, 42–50
direct (*see* Direct ray)
reflected (*see* Reflected ray)
in surface duct, 19
tangent, 5, 6, 116
turning point of (*see* Turning point, of ray)
in waveguide, penetration of ducts in, 18
Ray curvature, relative, 12
relative to earth, 50
Ray family, 15, 47
linear N-profile, 16
simple surface duct, 17
Ray theory, 332, 391
Ray tracing, 17, 18
errors of, 7, 18
relation of, to phase-integral methods, 82
Ray-tracing conditions, as criterion of validity of ray methods, 57
Ray-tracing formulas, 41–50
Ray-tracing methods, limitations of, 53–58
Rayleigh roughness criterion, 411, 430, 431, 434, 539
Rayleigh scattering law, 23, 465
for cylinder, 461
deviations from, 598, 612
for flat plate, 8, 465
for meteorological echoes, 597–598, 610, 618
for sea echo, 494
for sphere, 7, 452
for suspended matter, 621
Received power, definition of, for pulses, 483
minimum useful, 32
Receiver calibrations, 564
Reciprocity theorem, 97, 114, 468, 693–698
Reflected ray, 5, 19, 36, 37, 38, 99, 100, 101, 108, 116
variations of, 424, 539
Reflection, from earth's surface, 5, 34
and radar height measurements, 436
from elevated M-inversion, 334
geometry of, from plane earth, 36
ionospheric, 3
from rough terrain, 340, 351, 471
specular, from terrain, 344, 391
wavelength dependence of, 436
time variation and rough surface, 418
Reflection coefficient, 37, 97–99, 102, 105, 396
dependence of, on grazing angle, 401–403
of ocean, 537
and sea echo, 418
for smooth plane, 396
for smooth sea, theory of, 402–403
and surface roughness, 411
for very smooth land, 432
Reflection coefficient measurements, interpretation of, 429, 435
long-range, over land, 430
over sea, 421
rapid variations in, 419, 538
short-range, over land, 433
over water, 427
Refraction, atmospheric, 34
standard, 14, 22, 53, 95
in diffraction region, 122
Refraction studies by radar, 353–373
Refractive index, of air, 189
complex, 448
variable gradients of, 9, 192
(*See also* Index, of refraction)
Refractive modulus, 192, 203, 207
approximate formula for, 203
definition of, 12, 191
errors of, 203
gradient of, 192, 198
errors in, 204
potential, 199, 201
approximate formula for, 205
Residues, sum of, of field integral, 110
Resistance coefficient, 216, 217
Resistance thermometer, 289
Riccati equation, 147
numerical integration of, 177
Ripple theory of sea echo, 494
Ripples, 487
Roller fading, 302, 304, 306, 324, 328, 335, 337–343
Rossby distribution of eddy viscosity, 215
Rossby equivalent potential temperature diagram, 188, 200, 201
Rotational spectrum, of oxygen, 649, 653–654
of water vapor, 656–658
Rough surfaces, and reflection coefficient, 411, 423–425, 434
(*See also* Reflection coefficient)

Rough surfaces, and time variation of reflecting properties, 434
Roughness length, 217

S

Saddle points, 90, 92
Sandstorms, echoes from, 621
Saturated vapor, 182, 183
Saturation-adiabatic lapse rate, 194, 279
Saturation curves, 184
Saturation mixing ratio, 185, 188
Saturation specific humidity, 185, 188
Saturation vapor pressure, 183
Scalar potential, diffraction region, 65
Scanning, conical, 440
Scanning loss, effect of, on σ, 471
Scattering, and absorption by particles, 23
  approximate formula for, 463
  by atmospheric irregularities, 317, 600
  in backward direction, 455
  coherent (*see* Coherent scattering)
  by earth's surface (*see* Reflection coefficient)
  by flat plate, 8
  general formula for, 454
  geometrical optics in, 465–468
  by quadric surfaces, 463
  by sphere, 7, 445
  treated by reciprocity theorem, 693–698
Scattering coefficients, for cylinder, 459
  for sphere, 448
Scattering cross section (*see* Radar cross section)
Scattering functions for spherical particles, 611
Scintillation, 302, 305, 306, 324, 335, 339, 350, 389
SCR-584 radar set, 444
Scripps Oceanographic Institution, 488
Sea, definition of, 486
Sea breeze, 264–266, 370
Sea clutter (*see* Sea echo)
Sea echo, σ° vs. θ, 503–510, 521, 524, 525
  appearance of, on airborne system, 481–482
    with high resolution, 491–492
  azimuth variation of, 514
  beam width affecting, 483–485, 491
  dependence of, on beam width and pulse length, 491
    on range, 490
Sea echo, σ° vs. θ, difficulties in measurement of, 496, 500, 503
  errors in measurement of, 496
  experiments with, table of, 501
  first probability distribution of, 514–516, 579
  fluctuations of, 514–518, 578
  dependence of, on wavelength, 517, 580
  formula for average signal, 484, 495, 500
  frequency spectra for, 516–518, 579–581
  at large depression angles, 485, 490, 503–510, 518–519
  maximum range of, vs. height, 500–503
  methods of measuring, 496, 500
  normal appearance of, on A-scope, 492
  polarization affecting, 498, 522, 527
  probability distribution for ratio of successive amplitudes, 580
  radar cross section of, 483
  radar cross section per unit area, absolute values of, 499, 506
    definition of, 483
    factors involved in, 485
    polarization dependence of, 497–499, 512, 522, 526
    sea state dependence of, 506, 512–514
    variation of, over sea surface, 493
    wavelength dependence of, 494–499, 511, 524
    when meaningful, 490–493
  radar transmission equation for, 484
  range dependence of, 490
  saturation of, in rough seas, 513
  as scattering from surface waves, 519–522
  shadow effect in, 493, 522
  slow variation of, 517–518, 578–579
  "spiky" appearance of, on A-scope, 492
  theories of, 494, 518–527
  velocity distribution of scatterers, 517
Sea-echo studies, future progress in, 526
Sea scales, 489
Seasonal effects, 338, 339, 343, 345, 351, 368, 370
  in Irish Sea, 324
  in Massachusetts Bay, 321
Second probability distribution, 556
  in power, 557
  for random scatterers plus steady signal, 561
Secular variation, effect of, on correlation function, 567

Secular variation, effect of, of ground clutter, 587
of sea echo, 517–518, 578–581
Selection rules, for oxygen, 649–650
for water vapor, 656–658
Self-absorption in precipitation echoes, 613
Shadow effect in sea echo, 493, 522
Shadow region, 17
Shear, 232, 234, 248, 250, 271
in stable equilibrium, 234–237
Shearing stratification, 235, 248, 300
variables governing, 235
Ship echoes, first probability distributions for, 545–546
fluctuations of, 543–546
frequency of fluctuation of, 543–544
Showers, in frontal systems, 633
in unstable air masses, 632
Sign convention for angles, 38, 116
Signal level, average, 304
over desert, 350
vs. duct height, 309–312
for Flathead Lake, 335
in Irish Sea, 323
over land, range of variation in, 340, 343, 349
for nonoptical vs. optical paths, 351
and substandard layers, 302, 303
and duct height, 306, 309–312
wavelength comparison, experimental, 310–312
maximum, 323, 352
midday, seasonal drifts in, 338
Signal types, over land, 337, 342
Massachusetts Bay, 301–307
over water, sudden changes in, 304
Similarity, principle of (*see* Principle of similarity)
Skip effect, 359, 360, 370–371
Smoke, echoes from, 593, 621
Snell's law, 6, 15, 19, 45, 46, 391
Snow, attenuation by, 688
echoes from, 620
Solar heating, 193, 220–222, 232, 259
Solar radiation, 220, 254
Sommerfeld integral, 65
Sonde, wired (*see* Wired sonde)
Sounding instruments, future requirements for, 292
Sounding methods, 282
Soundings, 207
problems in execution of, 291
by radiosonde, 185
(*See also* Radiosonde)
representation of, 206–207
synthetic, 298, 300, 309
Space interference pattern, changes in, as cause of echo fluctuations, 535–538
Specific heat of dry air, 188, 293
Speed correction, 276
for aeropsychrograph, 277–280
measurement technique for, 280
wet-bulb, 279
Sperry Gyroscope Company, 427
Spheres, scattering from, 445–453, 608–613
Spherical wave in terms of cylindrical waves, 61
Stability, importance of, in radar measurements, 527–529
of radar system, as affecting echo fluctuations, 527–529
limits of, 528
Standard deviation of fluctuation from random scatterers, 555
Standard pressure, 195
Stars, twinkling of, 532–533
Stationary phase, 103
principle of, 99
Statistical distribution of signal strength, Massachusetts Bay, 320
Steady-to-random ratio for ground clutter, 582
Steepest descent, method of, 90, 98, 105, 106
path of, 92
Stereophotography of sea echo, 490, 527
Stokes lines, 73, 77
Stokes phenomenon, 71, 72, 91
Storm structure, 23, 623, 625, 627–640
Strawberry Hill Standpipe, 537–538
Structure factor for microwave absorption, 650, 662
Sturm differential equation, 69
Subdominant terms, 73
Subsidence, 260
Subsidence inversion, 260–262
Substandard layer, 143, 168, 301, 302, 315
analytical complications of, 172
definition of, 14
(*See also* Surface layer)

Superstandard layer, 237
definition of, 15
elevated, 304
(*See also* Surface layer)
Surface of sea, 486–490
methods of observing, 490
Surface current, 455, 460, 473, 476
Surface duct, 14, 18, 75, 176, 304, 335, 364
deep, 302, 306, 308, 378
formation of, over land, 352
shallow, 302, 306, 308, 374
Surface layer, 14, 220–226, 228–234, 238–260, 348
standard, 230, 258
substandard, 14, 230, 258, 303, 306
over land, 258
and observed signal layer over water, 316
and scintillations, 305
superadiabatic, 221, 222
superstandard, 14, 229, 258
Surface wave, 2, 3, 40, 66, 103, 109
Swell, definition of, 487

T

T-14, experimental radar, 495
Tangent plane, 116
Tangent ray, 5, 6, 116
Target cross sections under nonstandard refraction, 354, 473
Target properties by radar measurements, 369
Targets, complex, fluctuations in echoes from, 547
standard, 479
Techniques in measurement of echo fluctuations, 562–570
Telecommunications Research Establishment (TRE), 175
Temperature, potential (*see* Potential temperature)
wet-bulb (*see* Wet-bulb temperature)
Temperature excess, 229, 231, 236, 237, 245–247
definition of, 222
negative and positive, 220
Temperature inversion, 192, 238
elevated, 231, 330
(*See also M*-inversion, elevated)
height of, 241–244
and *M*-gradient, 13, 192, 259
Temperature inversion, nocturnal, 258
in summer, 259
in winter, 259
10-per cent rule, 227, 228, 300
Tephigram, 188
Terminal velocity for raindrops, 616, 679
Terrestrial radiation, 254
Texas, University of, 389
Theorem, fundamental, 65
Thermal conductivity, 209
of earth, 254–255
Thermal stability, effect of, on shear, 234
Thermistor, 272
temperature-resistance curve for, 273
Thermocouples, 290, 323
Thermometric conductivity, 209, 293
Thunderstorms, 23, 627
Tidal effects, 321, 420
as cause of echo fluctuations, 535–539
for extended targets, 537–538
for point targets, 535
Time-average signal, from meteorological echoes, 589–591
from sea echo, 482–485
Trade-wind inversion, 375
Transitional layers, 172
Transmission in free space, 27–34
Transmission equation, free-space, one-way, 31
one-way, generalized, 35
radar, generalized, 35
Trapped modes, 22, 71, 75, 82, 149
boundary condition for, 78
characteristic values of (*see* Characteristic values)
Trapping, criteria for, 21, 84
by elevated inversions, 332
strong, 167
Tree-motion and radar echo fluctuation, 547, 552, 584
Triplets of oxygen spectrum, 649, 653, 654
Troposphere, 9
Turbulence, 23, 211, 250
mechanical, 221, 225, 268
and precipitation fluctuations, 577
thermal, 220, 221
Turbulence constant, universal, 216
Turbulence inversion, 231
Turbulent boundary layer, 214, 215, 217–219
logarithmic distributions in, 215
thickness of, 218

Turning point, for mode of propagation in duct, 20
of ray, 16, 17, 19, 46, 83, 100
Twinkling of stars, 532
Two equal scatterers as radar target, 529, 548–550
Typhoons, 23, 636–640

U

Ultra Short Wave Propagation Panel, 322
U.S. Army Signal Corps, 432
U.S. Navy Electronics Laboratory (*see* Navy Electronics Laboratory)
U.S. Weather Bureau, 339
Units, system of, 28
Universal gas constant, 292
Universal turbulence constant, 216

V

Values, characteristic (*see* Characteristic values)
Vapor pressure, 182, 187, 192, 206
lapse rate of, 196
potential, 198, 200, 202
over sea water, 200
Vapor-pressure curve for sea water, 184
Vaporization, latent heat of, 293
Variational methods, 177
Vector plane wave, expansion of, in spherical vector wave functions, 447
Vector potential method for scattering calculations, 455
Vector wave equation, 445
Viscosity, dynamic, 208
eddy (*see* Eddy viscosity)
kinematic, 208, 214, 293

W

Warm front, 263
Washington State College, 226, 284, 285, 335, 341, 347, 348
Water, index of refraction for, 675
molecular weight of, 293
supercooled, 183
Water temperature, 300, 310
measurements of, 299
variation of, 250–254
Water vapor, 26
equation of state for, 182
rotational spectrum of, 656–658
selection rules for, 656–658
Water-vapor absorption, 312, 363, 646, 647, 656–671
measurements of, 658, 661, 666–671
nonresonant part of, 662
Water-vapor content of atmosphere, seasonal variation and effect of, on M-profile, 259
Water-vapor lapse and M-gradient, 13, 192
Wave equation, formal expression for, 60
solution of, for diffraction region, 109
for interference region, 99
vector form, 59, 445
Wavefront, 42
of elementary wave, 63
Waves, elementary, 18
on sea, 486–493
origin of, 488
scattering from, 492, 519–527
WBZ radio towers, 296, 549
Weather Radar Project, 577, 617
Wet-bulb depression, 187
coefficient of, 188
Wet-bulb temperature, 182, 186, 187, 200, 206
potential, 198, 200
White caps, sea echo in presence of, 512
Wiener-Khintchine theorem, 558, 568
Wind, effect of, on ground-clutter fluctuation, 547, 551–553, 582–584
Wind direction and strength of sea echo, 514
Wind speed, and superrefraction, 375
variation in, along trajectory, 251
Window, 550
in infrared region, 26
(*See also* Chaff)
Winds aloft in Massachusetts Bay, 299
Wired sonde, 269, 283, 284, 329, 348
ground unit of, 286
housing for, 285
WKB approximation, 72
Wronskian, 58, 69, 79, 94

X

XT-2, experimental radar, 354, 495
XT-3, experimental radar, 10, 354, 495