Radio Occultations Using Earth Satellites: A Wave Theory Treatment

DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES

Issued by the Deep Space Communications and Navigation Systems
Center of Excellence
Jet Propulsion Laboratory
California Institute of Technology

Joseph H. Yuen, Editor-in-Chief

Previously Published Monographs in this Series

1. *Radiometric Tracking Techniques for Deep-Space Navigation*
C. L. Thornton and J. S. Border

2. *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation*
Theodore D. Moyer

3. *Bandwidth-Efficient Digital Modulation with Application to Deep-Space Communications*
Marvin K. Simon

4. *Large Antennas of the Deep Space Network*
William A. Imbriale

5. *Antenna Arraying techniques in the Deep Space Network*
David H. Rogstad, Alexander Mileant, and Timothy T. Pham

Radio Occultations Using Earth Satellites: A Wave Theory Treatment

William G. Melbourne

Jet Propulsion Laboratory
California Institute of Technology

MONOGRAPH 6
DEEP SPACE COMMUNICATIONS AND NAVIGATION SERIES

Radio Occultations Using Earth Satellites: A Wave Theory Treatment

April 2004

The research described in this publication was carried out at the
Jet Propulsion Laboratory, California Institute of Technology, under
a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service
by trade name, trademark, manufacturer, or otherwise, does not constitute
or imply its endorsement by the United States Government or the
Jet Propulsion Laboratory, California Institute of Technology.

Table of Contents

Foreword ..... *xvii*

Preface ..... *xix*

Acknowledgments ..... *xxiii*

Chapter 1: Radio Occultation Using Earth Satellites Background and Overview 1

1.1 Introduction 1

1.1.1 History of the Occultation Technique 2

1.1.2 Occultations from Earth Satellites 3

1.1.3 The Global Positioning System 8

1.1.4 Timing 11

1.1.5 Ephemerides 12

1.2 Information Content in GPS Occultation Observations 13

1.2.1 Connected Phase 13

1.2.2 Sample Rate Versus Vertical Resolution 14

1.2.3 Inverting Radio Occultation Data 14

1.2.4 Assimilating Radio Occultation Data 17

1.2.5 Rays and Stationarity 18

1.2.6 Excess Doppler 21

1.3 Scientific Applications of GPS Occultation Observations 23

1.3.1 Weather 24

1.3.2 Climate 27

1.4 Problems from Multipath and Some Remedies 28

1.4.1 Spectral/Holographic Techniques 32

1.4.2 Back Propagation 34

1.4.3 The Canonical Transform Technique 41
1.4.4 The Impact Parameter Space Curve 41
1.4.5 A Full-Spectrum Wave Theory 43
1.5 Overview 47
1.5.1 Chapter 2 47
1.5.2 Chapter 3 48
1.5.3 Chapter 4 49
1.5.4 Chapter 5 50
1.5.5 Chapter 6 52
1.5.6 Appendices 60
1.6 Limitations and Simplifications 61
1.6.1 Ionosphere 61
1.6.2 Placing the Occulted GPS Satellite at Infinity 61
1.6.3 Time 63
1.6.4 Spherical Symmetry 63
1.6.5 Coplanarity 63
1.6.6 Circularity 64
1.6.7 Treating the GPS Signal as a Harmonic Wave 64
1.7 Recommendations for the Next Chapters 65
References 66
Chapter 2: Scattering of Electromagnetic Waves from a Spherical Boundary Using a Thin Phase Screen Model and Scalar Diffraction Theory 77
2.1 Introduction 77
2.1.1 Multipath Scenarios 79
2.1.2 An Overview of Chapter 2 82
2.2 Geometric Optics in a Spherical Medium 83
2.2.1 Multipath, Shadow Zones, and Caustics According to Geometric Optics 86
2.2.2 Thin-Atmosphere Conditions 90

2.3 Thin Phase Screen Models 92

2.3.1 The Helmholtz–Kirchoff Integral Theorem 93

2.3.2 The Space Curve for Impact Parameter.....94

2.3.3 The Fresnel Phase Screen 100

2.3.4 Suitability of the Thin-Screen Model for Diffraction Analysis.....101

2.3.5 A Phase Profile for the Thin Screen 102

2.3.6 Bending-Angle Perturbations 103

2.3.7 Case A: A Discontinuity in Refractivity 105

2.3.8 Case B: A Discontinuity in Scale Height 107

2.3.9 Case C: A Discontinuity in Lapse Rate.....109

2.4 Multipath Using a Thin Phase Screen Model 109

2.5 Scalar Diffraction: The Rayleigh–Sommerfeld Integral 114

2.6 The Stationary-Phase Technique 116

2.6.1 Necessary Conditions for Validity of the Geometric Optics Approach 121

2.7 Numerical Results Using Thin-Screen/Scalar Diffraction..... 122

2.7.1 Fresnel Response to a Discontinuity in Lapse Rate 125

2.7.2 Fresnel Response to a Discontinuity in Refractivity.....128

2.7.3 A Boundary Layer.....131

2.8 Sensing a Boundary in the Ionosphere..... 133

2.8.1 Fresnel Effects at a Boundary.....136

2.8.2 Amplitude Effects at a Boundary 139

2.8.3 Ray Splitting.....140

2.8.4 Doppler Information at a Boundary.....146

2.8.5 Fresnel Effects Using an Improved Electron Density Model 148

2.9 The Error in the Recovered Refractivity Resulting from Fresnel Phase Perturbations 152

2.10 Fresnel Transform Techniques 156

2.10.1 Adjoining the Constraint $A(u) \equiv 1$ 158

2.10.2 Diffraction Integral with Multipath 159

2.10.3 A Numerical Example.....159

2.10.4 Fresnel Aliasing.....161

2.10.5 Numerical Results 162

References 167

Chapter 3: Scattering from a Large Transparent Sphere Based on Maxwell's Equations: Mie Scattering Theory 171

3.1 Introduction..... 171

3.2 Scalar Potentials..... 174

3.2.1 Series Expansions for Scalar Potentials 176

3.3 Multiple Internal Reflections..... 179

3.4 Fresnel Formulas for Reflection and Transmission Amplitudes..... 181

3.4.1 Conservation Principle.....184

3.4.2 Scattering Angles and Intensities 185

3.4.3 Caustics.....189

3.5 Mie Scattering Theory: Obtaining the Scattering Coefficients at a Boundary 192

3.5.1 Transmission and Reflection Coefficients for Incident External Waves 192

3.5.2 Transmission and Reflection Coefficients for Incident Interior Waves.....196

3.5.3 Aggregate Scattering.....198

3.6 The Problem of Slow Convergence..... 200

3.7 The Sommerfeld–Watson Transformation 201

3.8 Evaluating Scattering Coefficients with Asymptotic Expansions..... 203

3.9 Expressing Scattering Coefficients in Terms of Phasors 209

3.10 Asymptotic Forms for the Hankel and Legendre Functions Evaluated at the LEO 214

3.11 Geometric Optics Interpretation of Mie Scattering Theory 216

3.12 Evaluating Mie Scattering by Integration of the Scattering Phasor 218

3.13 Interpreting Scattering Using the Stationary-Phase Technique 220

3.13.1 Stationary Phase 220

3.13.2 Comparison with Diffraction from a Knife-Edge 223

3.13.3 Stationary-Phase Points for the Scattering Integrals 224

3.13.4 Vestigial Rainbow Effects 229

3.13.5 Asymptotic Solutions for a Transparent Refracting Sphere 232

3.13.6 Correspondence Between the Stationary-Phase Value for Spectral Number and Snell's Law 236

3.13.7 Stationary Value for the Scattering Phase 237

3.13.8 Evaluating the Scattering Integrals Using the Stationary-Phase Technique 240

3.14 Duality Between Stationary-Phase Concepts in Electrodynamics and in Geometric Optics 243

3.15 Diffraction from a Large, Transparent, Refracting Sphere Using Mie Scattering Theory 243

3.15.1 Numerical Integrations 243

3.15.2 Comparison with Thin-Screen/Scalar Diffraction Results 246

3.16 Looking for Rainbows 249

3.16.1 Rainbow Effects on the Refracting Sphere 249

3.16.2 Rainbow Effects at LEO Distances 251

3.16.3 Assessing Rainbow Effects Using the Third-Order Stationary-Phase Technique 253

3.17 Limiting Cases 257

3.17.1 The Perfectly Reflecting Sphere 257

3.17.2 Geometric Optics Interpretation 263

3.17.3 Limiting Cases: The Strongly Absorbing Sphere 265

3.17.4 Limiting Cases: Reflections from a Sphere of Infinite Radius 269

References 269

Chapter 4: Wave Propagation in a Stratified Medium: The Thin-Film Approach 271

4.1 Introduction 271

4.2 Thin-Film Concepts 272

4.2.1 Cartesian Stratification 272

4.3 The Characteristic Matrix 276

4.4 The Stratified Medium as a Stack of Discrete Layers 277

4.4.1 The Characteristic Matrix when $n(x)$ = constant 278

4.4.2 A Stack of Homogeneous Layers when $n(x)$ is Piecewise Constant 279

4.4.3 Range of Validity 282

4.4.4 The TM Case 283

4.5 The Characteristic Matrix for an Airy Layer 284

4.6 Incoming and Outgoing Waves and Their Turning Points 288

4.6.1 Eikonal and Cophasal Normal Paths 293

4.6.2 Defocusing 296

4.7 Concatenated Airy Layers 298

4.8 Osculating Parameters 303

4.8.1 At a Turning Point....309
4.9 Airy Functions as Basis Functions 310
4.10 Wave Propagation in a Cylindrical Stratified Medium 311
4.11 Wave Propagation in a Spherical Stratified Medium 318
4.12 Correspondence between Characteristic Matrices for Cartesian and Spherical Stratified Airy Layers 324
References 327

Chapter 5: Propagation and Scattering in a Spherical-Stratified Refracting Medium 329
5.1 Introduction 329
5.2 Maxwell's Equations in a Stratified Linear Medium 334
5.2.1 Scalar Potential Functions....335
5.3 Modified Spherical Bessel Functions.... 337
5.4 Asymptotic Forms.... 341
5.5 Modified Mie Scattering in a Spherical Stratified Medium 343
5.5.1 Incoming Waves....347
5.5.2 Evaluating the Spectral Coefficients in a Stratified Medium....352
5.5.3 Outgoing Waves....357
5.5.4 Correspondence between Cartesian and Spherical-Stratified Phase Quantities....358
5.5.5 Absorption....359
5.6 More Geometric Optics: Cumulative Bending Angle, Bouguer's Law, and Defocusing.... 360
5.6.1 Defocusing....363
5.7 More Asymptotic Forms 366

5.7.1 Equating $dG\left[\rho^{\dagger}(\nu),\nu\right]/d\nu$ to $\tilde{\alpha}(\nu,\nu)$ and $d^2G\left[\rho^{\dagger}(\nu),\nu\right]/d\nu^2$ to $d\tilde{\alpha}(\nu,\nu)/d\nu$ 373

5.7.2 Fixing the Form for $g(\hat{y})$ when $\hat{y}>0$ by Asymptotic Matching.....374

5.7.3 Behavior of the WKB Solution at a Turning Point.....377

5.7.4 Setting $G[\rho,\nu]$ for $\rho \le \rho^{\dagger}(\nu)$ 379

5.8 Spectral Representation of an Electromagnetic Wave in a Spherical Stratified Medium 379

5.8.1 Behavior of $\partial G/\partial\nu$ 381

5.8.2 Accuracy of the Osculating Parameter Technique 385

5.8.3 Numerical Comparisons.....389

5.8.4 Comparison of Phase Delays in an Airy Layer from Wave Theory and Geometric Optics 398

5.8.5 Asymptotic Matching the Spectral Coefficients for Incoming and Outgoing Waves 399

5.9 Interpreting Wave Theory in a Refracting Medium Using the Stationary Phase Technique..... 403

5.9.1 Geometric Interpretation of the Phasors.....403

5.9.2 Stationary Phase Conditions 405

5.9.3 Stationarity of $\Psi(\pm,\pm)$ 406

5.9.4 Plane Waves 407

5.9.5 The Electric Field for an Incoming Wave.....409

5.9.6 Evaluating the Electric Field Vector Using the Stationary Phase Technique.....412

5.9.7 An Outgoing Electric Field.....416

5.9.8 The Electric Field at the LEO.....418

5.10 Comparison of Geometric Optics and Wave Theory 418

5.10.1 Comparison of Wave Theory with Geometric Optics 419

5.10.2 Duality between Systems 423

5.10.3 Amendments to Account for Wave-Front Curvature from the Finite Distance of the Emitting GPS Satellite 425

5.11 The Electric Field at a Turning Point 426

5.11.1 Fourier Transform of the Airy Function 429

5.11.2 Fresnel Transform of the Airy Function 431

5.12 Caustics and Multipath 432

5.12.1 A Numerical Example of Multipath and Caustics 433

5.12.2 A Geometric Optics Interpretation of Multipath 440

5.12.3 Amplitude Variability from Geometric Optics 443

5.12.4 Third-Order Theory 443

5.12.5 Reduction of Multipath Spectral Width by Defocusing 452

5.12.6 Combined Water Vapor and Dry Air Refractivity Model 453

5.13 Spectral Coefficients in a Spherical Refracting Medium with an Embedded Discontinuity 458

5.13.1 A Medium with a Discontinuity in its Refractive Gradient 459

5.13.2 A Transparent Sphere Embedded in a Refracting Medium 468

5.14 The Scattered Field from a Perfectly Reflecting Sphere Embedded in a Refracting Medium 469

5.14.1 Stationary Phase Analysis 475

5.14.2 Results from Wave Theory 483

5.14.3 Law of Reflection 486

References 488

Chapter 6: The Inverse Problem: Using Spectral Theory to Recover the Atmospheric Refractivity Profile 491

6.1 Introduction 491

6.2 GPS Receiver Operations 492

6.2.1 Adverse Signal Conditions 496

6.2.2 Flywheeling 502

6.2.3 Refractivity Error from a Single-Ray Paradigm 504

6.3 Spectral Representation of the Field at the LEO 506

6.3.1 Stopped Field at the LEO 509

6.3.2 The Obliquity Factor 509

6.3.3 Doppler Variability 511

6.4 Refractivity Recovery 514

6.4.1 Super-Refractivity 521

6.4.2 Improving the Accuracy of $G^{\dagger}(\nu)$ 526

6.4.3 Resolution Issues 528

6.5 Summary 529

References 531

Appendix A: Miscellaneous Derivations 533

A.1 Bending Angle and Bouguer's Law, Eq. (2.2-2) 533

A.2 Defocusing 535

A.3 Excess Doppler 536

A.4 Scalar Diffraction Theory 542

A.5 The First Fresnel Zone 544

A.6 The Abel Transform 545

A.7 Sensitivity of the Recovered Refractivity to an Error Source 546

A.8 Perturbation of the Bending-Angle Profile Near a Boundary 546

A.9 Bending-Angle Perturbation 550

A.10 Bending-Angle Perturbation by a Discontinuous Refractivity, Case A 551
A.11 The Fresnel Transform Pair 552
A.12 Ray Path Phase Delay 553
References 555
Appendix B: Caustic Surfaces 557
References 561
Appendix C: Multiple Ray Path Separation Altitudes 563
C.1 Case C: A Discontinuous Lapse Rate 563
C.2 Case B: A Discontinuous Scale Height 565
C.3 Case A: A Discontinuous Refractivity 566
Appendix D: Third-Order Stationary Phase Theory 571
D.1 Maximum Flaring 571
D.2 Minimum Signal Amplitude in a Shadow Zone 574
D.3 Accuracy of the Stationary Phase Technique 579
References 580
Appendix E: Bending by a Gaussian Electron Density Distribution 581
Appendix F: The Effect of Cycle Slips on Recovered Refractivity 583
Appendix G: Using the Sommerfeld–Watson Transformation 587
G.1 Introduction 587
G.2 Application to a Perfectly Reflecting Sphere 591
G.3 Application to a Refracting Sphere 593
G.4 Aggregate Scattering 596
Reference 598
Appendix H: Characteristic Matrix in a Stack of Airy Layers 599

Appendix I: Field Equations in a Stratified Medium 605

Appendix J: Conditions for Near-Equivalence between $dG^{\dagger}(\nu)/d\nu$ and $\tilde{\alpha}(\nu,\nu)$, and between $d^2G^{\dagger}(\nu)/d\nu^2$ and $d\tilde{\alpha}(\nu,\nu)/d\nu$ 607

Glossary of Terms 611

Acronyms 619

Foreword

The Deep Space Communications and Navigation Systems Center of Excellence (DESCANSO) was established in 1998 by the National Aeronautics and Space Administration (NASA) at the California Institute of Technology's Jet Propulsion Laboratory (JPL). DESCANSO is chartered to harness and promote excellence and innovation to meet the communications and navigation needs of future deep-space exploration.

DESCANSO's vision is to achieve continuous communications and precise navigation—any time, anywhere. In support of that vision, DESCANSO aims to seek out and advocate new concepts, systems, and technologies; foster key technical talents; and sponsor seminars, workshops, and symposia to facilitate interaction and idea exchange.

The Deep Space Communications and Navigation Series, authored by scientists and engineers with many years of experience in their respective fields, lays a foundation for innovation by communicating state-of-the-art knowledge in key technologies. The series also captures fundamental principles and practices developed during decades of deep-space exploration at JPL. In addition, it celebrates successes and imparts lessons learned. Finally, the series will serve to guide a new generation of scientists and engineers.

Joseph H. Yuen
DESCANSO Leader

Preface

Providing an accurate quantitative description of wave propagation through a complex medium remains one of the more difficult mathematical problems of our age. In electrodynamics, many alternative-solution techniques for Maxwell's equations have emerged over the past century; all involve approximations in some form or another. These techniques range from approximations to the equations themselves, e.g., WKB, perturbation theory, parabolic equations, finite element, dynamic programming, scalar diffraction-based multiple phase-screen propagation using the Helmholtz–Kirkhoff theorem, etc., to straight ray theory techniques, with a good measure of hybrid wave/optics and spectral techniques included. Moreover, particularly when the wavelength of the wave is small compared to the scale of the refracting medium, one is forced to analytic and geometric simplifications to render such problems tractable. When one compares wave propagation literature from diverse fields, such as acoustics, electrodynamics, and seismology, one finds an almost bewildering variety of analysis techniques; one also finds, from time to time, essentially the same technique applied in different fields, but with different labels.

The radio occultation technique using spacecraft to sound a refracting medium is a relative newcomer in wave propagation processes and in the inverse problem of inferring certain physical properties of the medium. The first planetary occultation with spacecraft was accomplished about 40 years ago using the NASA/JPL Mariner 4 spacecraft as it passed behind Mars. The Mariner radio signal received on the Earth first transected the Martian ionosphere and atmosphere during its immersion before being eclipsed by Mars, and it passed through these media again after its egress from behind Mars. Using ray theory techniques, partly borrowed from seismology, these

initial experiments successfully recovered accurate vertical refractivity profiles for the Martian atmosphere, and related density and pressure information.

These early ray theory techniques work well in a thin medium such as the Martian atmosphere and even in the upper reaches of other planetary atmospheres, such as the Earth's. But, as anyone who seriously studies sunrises and sunsets could predict, dense atmospheres with even simple mesocale layered structures, not to mention more complex structures, can lead to difficulty in the inversion process using basic ray theory techniques. Multiple rays from the transmitter following different paths through the atmosphere to the receiver, and atmosphere-induced broadening of the spectral lines of the individual rays, lead to ambiguity problems and analysis difficulties in a purely ray-theoretic approach. Moreover, ray theory, without recourse to at least an assist from scalar diffraction theory, imposes a certain resolution limit in the spatial dimensions perpendicular to the line of sight, the so-called first Fresnel zone, which is analogous to the Airy disk used in optical instruments to set resolution performance. Because of these limitations in basic ray theory, radio occultation investigators turned early on to wave/optics techniques to unravel multipath ambiguities and to sharpen the resolution using synthetic aperture concepts that approach the limit offered by wave theory. The thick atmospheres of the Jovian planets and the desire to sharpen the resolution of observations of the rings of these planets led to innovations in wave/optics techniques applied to the occultation experiments carried out with the Voyager spacecraft beginning about 25 years ago. The relatively recent advent of low Earth orbiters carrying occultation-capable GPS receivers (a half-dozen in the past decade and probably twice that many in the next decade) has resulted in a marked expansion in radio occultation science, in research and development in analysis techniques, and in technology development of new instrumentation and data information systems. The Earth's atmosphere is far more accessible than those of our solar system neighbors.

These modern inversion techniques, which on the whole have worked moderately well in adverse signal conditions such as those in the Earth's lower troposphere, may be classified, broadly speaking, as either spectral/holographic-based or as scalar diffraction-based. They all use some wave-related concept to recover the refractive bending angle(s) of the "received" ray(s), and through a certain ray-theoretic inversion process, they recover the vertical refractivity profile of the atmosphere. This hybrid wave-ray methodology is often referred to as a "wave/optic" analysis process.

This monograph develops, as its central theme, a purely wave-theoretic approach. The amplitude and phase of a harmonic wave at a given point are formally expressed in terms of a complex spectral series, the spectral coefficients of which depend on the refractive gradient of the medium in addition to the spectral number and the initial conditions for the wave. Therefore, the spectral coefficients vary with position in the medium because of

the non-zero refractive gradient there, whereas in the homogeneous medium they are invariant. The total contribution to the electromagnetic field at a given point is obtained by summing this spectral series over all spectral numbers. Stationary phase points in spectral number signify the major neighborhoods contributing to the evaluation of the field. The goal here is to recover the spectral coefficients from the time series of observed amplitude and phase of the received wave. The magnitude of a spectral coefficient relates to the absorption by the atmosphere; its phase shift relative to the constant phase that would hold in a homogeneous medium relates to the cumulative effect of the refractive gradient of the atmosphere on that spectral component of the wave. The phase shift for the spectral coefficient is given by a spectral density function for refractive gradient-induced phase delay. From the recovery of this spectral density function, the profile for the refractivity gradient is obtained from the formal integral relationship between these quantities. Although full-spectrum wave theory approaches for these kinds of problems are notorious for their complexity and cumbersomeness, it nevertheless is possible to render the mathematics into a reasonably tractable form through judicious use of asymptotic techniques, stationary phase concepts, and so on. Special topics familiar in ray theory, such as rays, caustics, reflections, defocusing, multipath, and so on, are treated here in a wave-theoretic framework.

William G. Melbourne
April 2004

Acknowledgments

Most of this monograph was written during 2000–2002 while I was on assignment from JPL at GeoForschungsZentrum-Potsdam (GFZ), now a leading Earth science institution in Germany, and in the late 19 century a principal European center in the nascent field of mathematical geodesy. I wish to thank my colleagues at GFZ-Potsdam and DRL-Neustrelitz, Georg Beyerle, Jens Wickert, and Norbert Jakowski, for many stimulating discussions, and Professor Christoph Reigber, for creating an extraordinary scientific *Atmosphäre* at GFZ for visitors and staff alike. I also wish to thank the JPL senior management for granting me this "sabbatical." I especially wish to acknowledge support by the Solid Earth Program in NASA Code Y under Dr. John LaBrecque, who played a key programmatic role in the success of GPS applications in space. To Cami Melbourne, my friend and wife, I owe a great debt for her forbearance and technical support with the preparation of text and figures. Some of the work here sprang from concepts originating in the JPL groups involved in research and development of high-accuracy GPS receiver technology for flight applications, and in the radio occultation groups that use this technology. In this regard, particular thanks to George Hajj, Tom Meehan, Larry Romans, and Tom Yunck for many discussions over the years.

There is one person that I would like to especially acknowledge for the inspiration he has left me through his pioneering work in radio occultation. That is Gunnar Lindal, who retired from JPL nearly 15 years ago. Although I headed a relatively large organization within JPL for a long time, in which Gunnar worked, I never really worked with him; our few interactions were usually of a management or programmatic nature. However, eventually I

studied most of the papers that he wrote from the 1960s through the early 1990s. I regret not spending more time with him when I had the chance.

William G. Melbourne
April 2004

Chapter 1
Radio Occultation Using Earth Satellites Background and Overview

1.1 Introduction

This monograph is concerned with the phase and amplitude of an electromagnetic wave during a radio occultation episode. This wave has passed through an intervening medium from a distant emitter and arrives at a receiver. The receiver measures the phase and amplitude of the wave over the duration of the occultation episode. These measurement sequences can be used to infer physical properties about the intervening medium.

Radio occultation refers to a sounding technique in which a radio wave from an emitting spacecraft passes through an intervening planetary atmosphere before arriving at the receiver. The words "occultation," "occulted," or "occulting" imply that the geometry involving the emitter, the planet and its atmosphere, and the receiver changes with time. Although an occulting or eclipsing planet (or moon) usually is involved, the word has come in recent times to also include non-occulting events, for example, satellite-to-satellite sounding through the ionosphere or receiving a reflected wave from a reflecting surface. From the perspective of the receiver for the strictly occulting case, the emitter is seen to be either rising or setting with respect to the limb of the occulting planet. As the radio wave from the emitter passes through the intervening atmosphere, its velocity and direction of propagation are altered by the refracting medium. The phase and amplitude of the wave at the receiver consequently are altered relative to their values that would hold without the intervening medium or the occulting planet. As time evolves, profiles of the phase variation and the amplitude variation at the receiver are generated and

recorded by the receiver. These profiles provide information about the refractive properties of the intervening medium.

In seismology, an array of seismometers spread over some geographical extent is used to study the various types of seismic waves arriving at each seismometer from a remote earthquake. With the array, one can measure the differential arrival times at the different stations in the array and also the spectral properties of the various waves. The different paths followed by these various waves and certain physical properties of the medium through which they pass can be inferred from these observations. In a radio occultation, it is the kinematics of the emitter/receiver pair over the duration of an occultation episode that provides analogous information.

1.1.1 History of the Occultation Technique

An astrometric observation in the eighteenth century measuring the times of ingress and egress of a lunar occultation of a star was probably among the first scientific applications of the technique. The method of lunar distances, the common seaman's alternative to the relatively expensive chronometer in the eighteenth and nineteenth centuries for keeping accurate Greenwich time for longitude determination at sea, and even for calibrating marine chronometers over very long voyages, depended crucially on an accurate lunar ephemeris. This in turn depended on accurate astrometric observations of the moon relative to the background stars and a good dynamical theory for the lunar orbit. Much later, the limb of the moon has been used as a knife-edge to obtain the microwave diffraction pattern from certain quasars [1]. The fringe spacing and amplitude of this diffraction pattern provide information about the angular distribution of radiant intensity from these very compact radio sources. Also, planetary atmospheres have been studied by analyzing stellar refraction and scintillation effects that occur during the ingress and egress periods when the star is occulted by the planet [2–4].

The radio occultation technique to sound planetary atmospheres using spacecraft began almost at the dawn of the era of planetary exploration. The first spacecraft to Mars in 1964, Mariner 4, flew along a trajectory that passed behind Mars as viewed from Earth [5,6]. The extra carrier phase delay and amplitude variation observed on the radio link between Mariner 4 and the Earth-based radio telescopes as Mariner 4 passed behind Mars and emerged from the other side provided valuable density information about its very tenuous atmosphere and also about its ionosphere [7]. Since then a score of experiments involving planetary missions have been undertaken to study the atmospheres of almost all of the planets in the solar system, including several moons and the rings of Saturn [8–12].

1.1.2 Occultations from Earth Satellites

Sounding of the Earth's atmosphere and ionosphere using the occultation technique theoretically can be performed with any two cooperating satellites. Prior to the Global Positioning System (GPS) becoming operational, a few early radio occultation experiments from a satellite-to-satellite tracking link had been conducted. These included the radio link between the Mir station and a geostationary satellite [13] and between GEOS-3 and ATS-6 [14]. However, this monograph focuses primarily on the carrier phase and amplitude measured by a GPS receiver onboard a low Earth orbiting spacecraft (LEO) while tracking the navigation signals emitted by a GPS satellite during its occultation by the Earth's limb [15]. GPS/MET (Global Positioning System/Meteorology), an occultation experiment that flew on MicroLab-1 and launched in 1995, was the first occultation experiment using the GPS [16–19]. Although experimental, over 11,000 occultations were used from GPS/MET to recover refractivity, density, pressure, temperature, and water vapor profiles [19]. GPS/MET provided a definitive engineering proof-of-concept of the occultation technique, and its data set became an experimental platform for implementing improved tracking and data processing schemes on subsequent Earth satellites with GPS occultation capability. GPS/MET also provided a basis for assessing the scientific and societal value of the technique in such diverse applications as meteorology, boundary layer studies, numerical weather prediction (NWP), and global climate change. Since then the Challenging Minisatellite Payload (CHAMP) (2001) [20], Satellite de Aplicaciones Cientificas-C (SAC-C) (2001), and Gravity Recovery and Climate Experiment (GRACE) (2002) satellites have been launched, from which radio occultation observations are now more or less continually made [21]. These missions alone could return nearly 1000 occultations per day. Future operational missions are planned, such as the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) configuration of occultation-dedicated LEOs to be launched in 2005 [22–24]. This system will provide near-real-time sounding information from about 4000 globally distributed occultations per day, which will be assimilated into NWP programs. Also, other global navigation satellite systems (GNSSs), such as the Russian Global Navigation Satellite System (GLONASS), and future systems, such as the planned European system, Galileo, will broaden the opportunities for dedicated satellite-to-satellite occultation missions [25,26].

Figure 1-1 depicts in exaggerated form a typical occultation scenario involving a LEO and the GPS satellite constellation. For a setting occultation, the about-to-be occulted GPS satellite will be seen from the LEO to be setting with respect to the limb of the Earth. The duration of a typical analyzed occultation through the neutral atmosphere (from roughly 100-km altitude to sea level) is less than 100 s. A ray from this GPS satellite passes through the

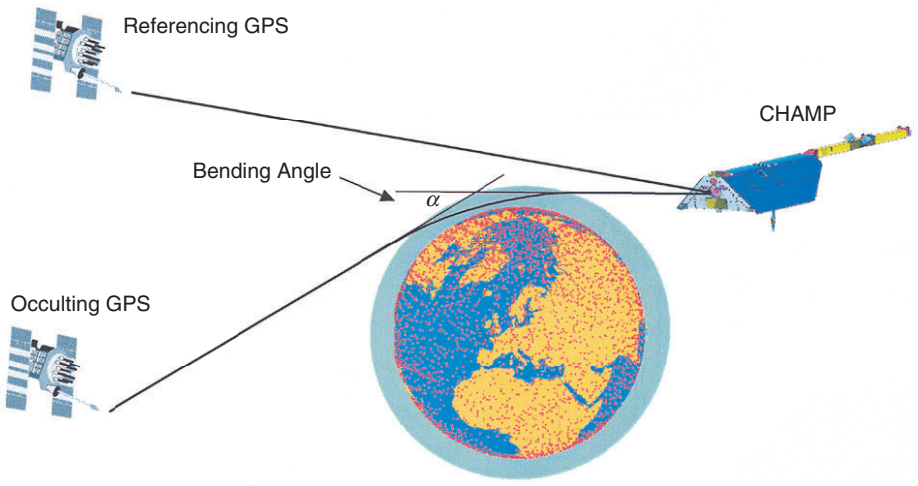

Fig. 1-1. Occultation scenario for a low Earth orbiter. From [91].

upper layers of the Earth's atmosphere at a near-horizontal rate of descent, skimming a deepest layer at its tangency point. A ray in this context may be defined by the normal to the cophasal surface of the carrier wave emitted by the GPS satellite. After the tangency point, the ray begins its near-horizontal rate of ascent through the upper layers, exits the atmosphere, and continues on to the LEO. The maximum total refractive bending of the ray is very small, about 1 deg from dry air. Additional refraction from water vapor, especially for tangency points in the lower troposphere, can double or triple that bending angle. A small bending-angle profile is a hallmark of a "thin" atmosphere. The index of refraction of dry air at microwave frequencies is about the same for light. At sunset the apparent sun for an observer on the ground is refracted through its own diameter, about 1/2 deg. The secular trend in the refractive gradient of dry air with altitude (which is near-exponential) is evident in the oblate shape of the apparent sun as its lower limb touches the horizon. Rays from the bottom limb usually are bent more than rays from the top limb. The fractured shape of the solar disk at some sunsets is caused by abrupt departures of the refractivity profile at low altitudes from the secular trend.

The ray arriving at the LEO from the occulted GPS satellite may not be unique, and indeed may not even exist in a geometric optics context. But for our purpose, we assume in Figure 1-1 that it does exist and that it is unique at the epoch of the observation. The inclination of the ray to the local geopotential surface is very slight; at 100 km from the tangency point it is about 1 deg in dry air, increasing linearly with ray path distance from the tangency point. At 400 km from the tangency point, the ray is about 10 km higher in altitude. Because of the near-exponential decrease in dry air density with height, it follows that most of the information about the atmosphere at a given epoch is

contained in a relatively narrow section of the ray about its tangency point, a few hundred kilometers in length [27].

As time evolves, the tangency point of the ray arriving at the LEO drifts deeper on average into the atmosphere for a setting occultation. Typically for the occultations selected for analysis, the initial cross-track angle of descent of the tangency point relative to the local vertical is less than 30 deg. The excess phase delay observed at the LEO, which is simply the extra phase induced by the refracting medium, will continue to increase on average over the course of the occultation because of the increasing air density with depth. This continues until the tangency point of the ray nears or contacts the planetary surface. The total excess delay can exceed 2 km with refractive bending angles up to 4 deg near the surface when water vapor is in abundance and the vertical gradient of its density is large. Defocusing and multipath, which tend to become strong in the lower troposphere, may reduce the signal amplitude to below a detection threshold before the ray contacts the surface, terminating the occultation episode several seconds or sometimes tens of seconds prematurely. However, sometimes the signal returns several seconds later, and sometimes in very smooth refractivity conditions even a knife-edge diffraction pattern from the limb or interference fringes from an ocean reflection are observed in the phase and amplitude before the direct signal from the GPS satellite is completely eclipsed [28,29]. On average, over the entire globe 80 percent of the CHAMP occultations reach to within 1 km of the surface, and 60 percent reach to within 1/2 km [21]. These encouraging statistics should improve in the future as new signal-tracking algorithms are implemented in the GPS receivers onboard the LEOs.

Figure 1-2 shows results from an early occultation from the GPS/MET experiment [18]. Here the excess phase delay of the L1 carrier in meters and its time derivative, excess Doppler in hertz, are shown over the last 90 seconds of the occultation. The lower abscissa shows coordinated Universal time (UTC), and its scale is linear. The upper abscissa shows the altitude of the ray path tangency point, and its scale is non-linear. The refractive gradient of the atmosphere increases markedly with depth, which effectively slows the average rate of descent of the tangency point because the refractive bending angle of the ray increases on average with depth. Therefore, the LEO must travel increasingly farther along its orbit to intercept these progressively deeper penetrating and more refracted rays. Near the Earth's surface, the average rate of descent of the tangency point typically is an order of magnitude smaller than it is in the upper atmosphere.

Figure 1-3 shows the amplitude of the L1 carrier for the same occultation [18]. The ordinate is the signal-to-noise ratio (SNR) in voltage, SNR_V, that would apply if the individual L1 carrier amplitude measurements were averaged over 1 s. The actual sample rate in this figure is 50 Hz. The averaging

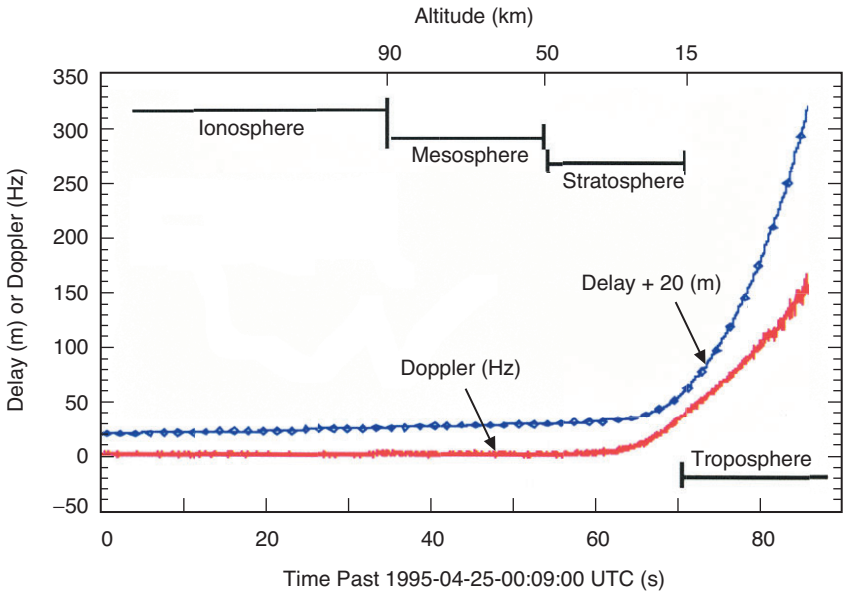

Fig. 1-2. Profile of excess Doppler and phase for a particular occultation of GPS PRN no. 28 observed by the GPS/MET experiment on MicroLab-1 on April 25, 1995, near Pago Pago. Redrawn from [18].

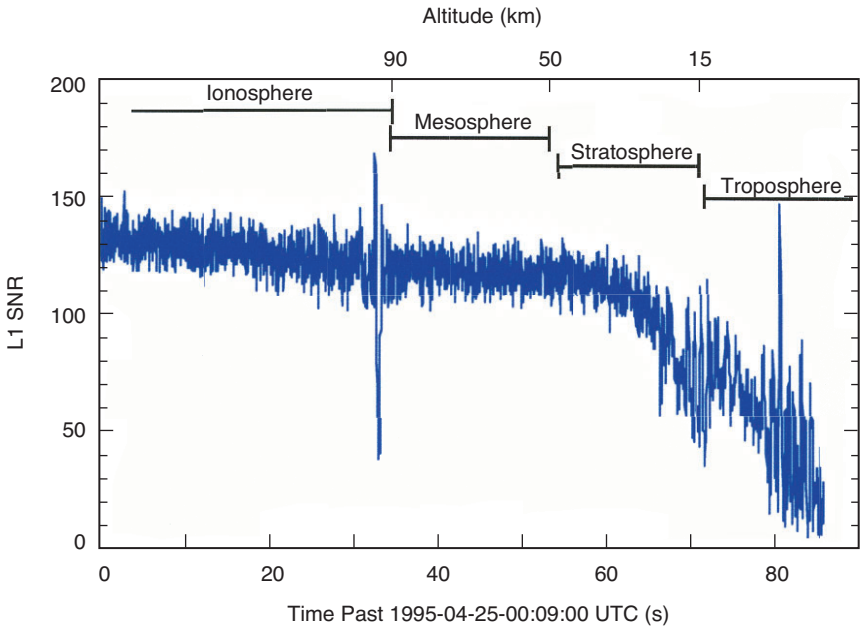

Fig. 1-3. Observed voltage SNR of the L1 carrier versus time for the same occultation shown in Fig. 1-2. Redrawn from [18].

time is 20 ms. Therefore, the thermal noise ($\sim 1/\mathrm{SNR_V}$) for the points in this figure is $\sqrt{50}$ larger than the $\mathrm{SNR_V}$ values would imply. This figure clearly shows the secular defocusing caused by the increasing refractive gradient with depth. The refractive gradient disperses the directions of progressively deeper rays after exiting the atmosphere; the gradient effectively "de-collimates" the rays, thereby diluting their power at the LEO by spreading it over a larger area. The ratio of a small "collimated" area proportional to Δa, shown in Fig. 1-4(a), to the resulting de-collimated area proportional to $\Delta\sigma$ at the LEO is the defocusing factor. Appendix A derives a simple form for this factor:

$$\frac{\Delta a}{\Delta\sigma} = \zeta = \left(1 - D\frac{d\alpha}{da}\right)^{-1} \tag{1.1-1}$$

where D is effectively the distance of the LEO from the Earth's limb and $d\alpha/da$ is the radial gradient of the refractive bending angle α. Figure 1-4(b) shows a relatively mild multipath scenario, including a shadow zone where $d\alpha/da$ is temporarily larger. Both the shadow zone and the interference from multipath waves cause the variability in SNR. The secular trend of $\mathrm{SNR_V}$ in Figure 1-3 is effectively given by $\zeta^{1/2}(\mathrm{SNR_V})_o$, where $\zeta^{1/2}$ is the defocusing from air and $(\mathrm{SNR_V})_o$ is the voltage SNR of the GPS signal that would be received at the LEO without the planet and its intervening atmosphere, the so-called "free-space" value. In addition to secular defocusing effects, this figure shows the strong transients in signal amplitude as the point of tangency of the ray passes through certain narrowly defined horizontal layers of the lower ionosphere and when it crosses the tropopause. These transients signal the presence of multipath-induced interference between different rays arriving concurrently at the LEO from different levels in the ionosphere and atmosphere. As the tangency point of the main ray cuts further down through successive layers of the atmosphere and into the middle and lower troposphere, the prevailing interference evident in this figure likely is induced by variable water vapor concentrations.

These transients observed in phase and amplitude raise a number of issues, some of which have been better dealt with than others. Significant progress on dealing with multipath has been made using back propagation and spectral techniques. Deep troughs in amplitude, shadow zones, and super-refractivity episodes, as well as caustics, are difficult for ray theory. How well these techniques work when the validity of ray theory itself is being strained is still to be established. The low SNR associated with many of these transients also makes it difficult to maintain connection in the LEO-observed phase measurements across them. We return to these topics later.

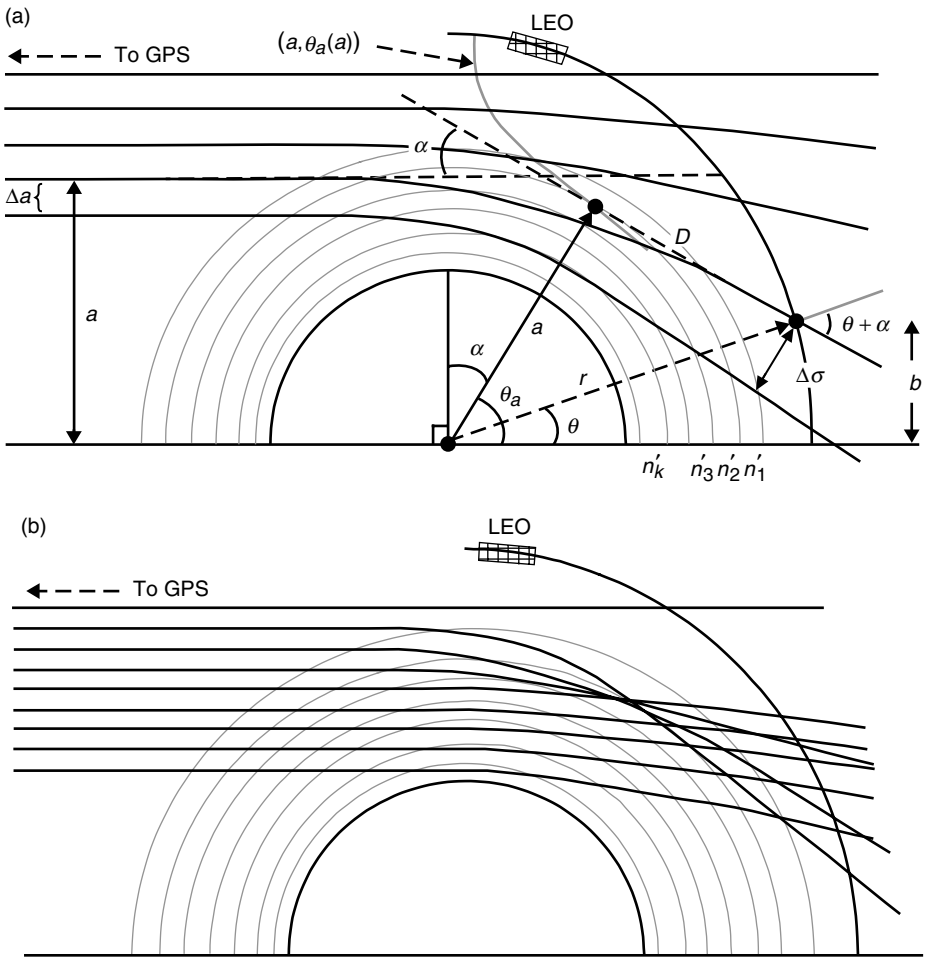

Fig. 1-4. Dispersive bending from the refractive gradient for an occultation observed from a LEO: (a) an ordered set of rays from an increasing refractive gradient with depth; no multipath, and (b) a non-monotonic refractive gradient results in a shadow zone and multipath.

1.1.3 The Global Positioning System

The GPS is operated by the U.S. Air Force (USAF). A system description of the GPS and the signal structure of its broadcast navigation signals are found in [30,31] and in many internet web sites. It suffices here to note a few details.

The GPS constellation is comprised of 24 satellites plus some on-orbit spares, more or less globally distributed up to 55-deg latitude. Their orbits are near-circular with a semi-major axis of about 4.1 Earth radii and with an inclination to the equator of 55 deg. A GPS satellite will be observed from a LEO to rise or set with respect to the Earth's limb on average once per 2 to 3

minutes, or at a rate of several hundred per day. The geographical distribution of the occultation tangency points is more or less global, but the actual distribution depends somewhat on the inclination of the LEO orbit plane and its altitude. For example, a LEO in a polar orbit returns fewer tangency points in the equatorial zone; fewer occulted GPS satellites in polar directions are viewed from that LEO because the GPS orbit inclination is only 55 deg. Figure 1-5 shows the global distribution of tangency points obtained over almost one month in 2001 for analyzed occultations from the CHAMP satellite. The lower density of points in the equatorial zone reveals CHAMP's near-polar orbital inclination. The geographical distribution also reveals vague clustering and striations resulting from commensurabilities between the satellite orbit periods. On roughly one-third of these occultations, a reflected ray from the Earth's surface also was detected [29].

The waveforms of the navigation signals broadcast by a GPS satellite are directional; they bathe the entire Earth with essentially full power. The 3-dB point of their radiant power distribution lobe is about 1400 km above the Earth's surface. Thus, a GPS receiver onboard a LEO, with an orbit radius typically well below this 3-dB threshold altitude, achieves about the same performance as a receiver on the ground.

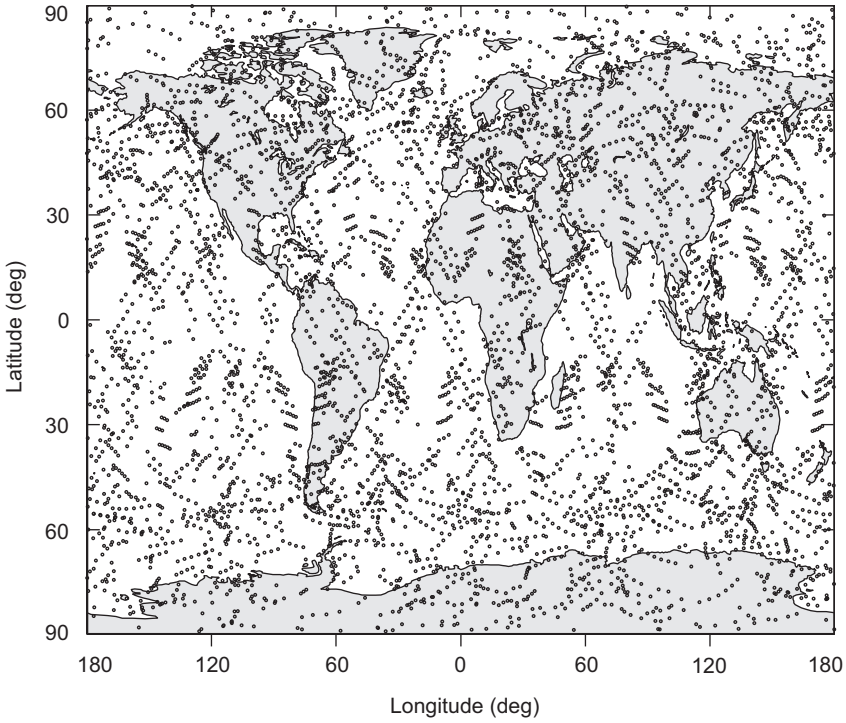

Fig. 1-5. Geographical distribution of nearly 4000 occultations obtained from CHAMP between May 14 and June 10, 2001. Redrawn from [29].

Each GPS satellite continuously broadcasts a set of square-wave codes by coherently modulating the phase of the carrier of the transmitted radio signal. These codes, which are unique to the broadcasting satellite, are pseudorandom and mutually orthogonal. They are used for ranging and for transmitting almanac and timing information. The mutual orthogonality property of the codes enables the receiver by cross-correlation techniques to isolate the received signals broadcast by a given satellite from all others, and to process in parallel the signals from all satellites in view of the receiver. The GPS satellites broadcast ranging codes on a pair of phase coherent L-band carriers, the L1 carrier at a frequency of 1575.4 MHz and L2 at 1227.6 MHz. These include an encrypted precision (P) code with a chip rate of 10.23 MHz on both carriers and the clear access or coarse acquisition (C/A) code at 1.023 MHz on the L1 carrier. The dual carriers are needed primarily to eliminate (or determine) the refraction effect from the ionosphere. For a microwave in the ionosphere, the refractivity is very nearly proportional to the local electron density and inversely proportional to the square of the carrier frequency. Therefore, the range and phase information received separately from the two carriers can be applied in concert to nearly completely decouple the ionospheric refraction effect by using this dispersive property of the ionospheric plasma. Newer versions of the GPS satellites planned for launch this decade will have an additional carrier at 1176.45 MHz (L5) and a C/A-like code also on L2. This will significantly improve receiver tracking operations using clear access ranging codes and increase the accuracy of the ionosphere calibration.

The signal structure of the ranging codes on a GPS signal is designed for near-real-time point positioning. By concurrently tracking four or more GPS satellites in diverse directions, the ranging code measurements can yield within a few seconds absolute point positions with an accuracy of roughly 10 m, and also one's time relative to the GPS clocks. One can obtain near-real-time relative positions (by concurrent tracking with two or more GPS receivers) with sub-meter or even sub-decimeter accuracy.

For occultation applications, however, one needs not the ranging information, but only very accurate measurements of the phase and amplitude of the L1 and L2 carriers, which are by-products of the range code tracking. For occultation applications, we may consider the radio signals arriving at a LEO from a GPS satellite as being a pair of spherical monochromatic waves from a distant point source at frequencies of 1575 MHz and 1228 MHz, respectively, plus Doppler shifts from kinematics and refraction of up to a few tens of kilohertz. Therefore, a high-performance GPS receiver used for space geodetic applications with millimeter-level accuracy requirements is a natural choice for occultation applications because it is designed to measure the phase of each carrier with sub-millimeter accuracy. Chapter 6 discusses certain additional aspects of such a space-rated receiver adapted for operations onboard a LEO.

1.1.4 Timing

A key factor for an accurate phase delay measurement is the epoch of the measurement. The measured phase at the receiver depends on the true phase accumulation between the emitter and the receiver, and the difference between the clock epochs of the emitter and receiver. Knowing the offset in time between the two clocks at a given instant is essential. More precisely, knowing the variability of this offset with time is essential. A constant offset is of no theoretical consequence (but it can be an operational problem) because the refraction information in the phase measurements is contained in their change with time. Each GPS satellite carries onboard up to four very precise cesium and/or rubidium frequency standards for controlling time and time intervals.

In 1999, the U.S. national policy was modified regarding certain operational aspects of the GPS. This change of policy led to the discontinuation of Selective Availability (SA) in May 2000, which had limited the accuracy of near-real-time point positioning to potential adversaries and to civilian users without access to decryption capability. By the late 1990s, SA became increasingly viewed as a cost and productivity issue for many GPS applications, military and civilian. As new technology and alternate means became available to the Department of Defense (DoD) for limiting access to the GPS, it became clear that, even though continuing SA provided a marginal defense benefit, it incurred an economic liability. SA deliberately degrades the near-real-time point-positioning accuracy of the GPS by at least an order of magnitude by causing the clock epoch errors in the GPS satellites to pseudo-randomly wander, nominally by the light-time equivalent of roughly 100 m over several minutes. Although the maximum deviation of the error is bounded and it can be averaged down substantially over 10 minutes or more, the short-term variability of SA poses a significant problem for clock epoch interpolation. SA dithers the onboard GPS master clock oscillator frequency at 10.23 MHz, which is derived from the atomic frequency standards and from which the chip rates of the codes and the frequencies of the L1 (154×10.23) and L2 (120×10.23) carriers are generated. The clock epochs depend on the integral over time of the oscillator frequency. The magnitude of the SA dithering has several possible levels of severity, which were set by certain alert or defense conditions. Over the years prior to 2000, SA had been set at a relatively low level but not at zero. Today SA is set to zero, but it has not been eliminated.

For carrier phase applications requiring high accuracy with SA turned on, multiple ground stations concurrently tracking all the GPS satellites in view at a relatively high sample rate of 1 Hz were needed [32,33]. This high-rate tracking was required to limit interpolation and/or extrapolation errors in rendering measured phases from different GPS satellites to a common epoch. Using these tracking data and applying a "double-differencing" or equivalent scheme among the tracked phase measurements referenced to common transmit epochs,

one can eliminate clock offset errors among the GPS satellites [34]. With these differencing operations, one incurs a penalty in increased effective thermal noise in the phase measurements, which becomes a concern in threshold situations where thermal noise becomes a limiting accuracy factor.

With SA off, the full inherent accuracy of the atomic frequency standards on the GPS satellites (with an effective 10-s frequency stability of a few parts in 10^{12}) can be realized with less costly ground tracking operations [35] and with more leisurely sample rates of 0.1 to 0.03 Hz. The interpolation error in the measured phase change over 10 s from GPS clock instability (with SA off) usually can be kept to well below 1 mm.

In addition to the GPS clock errors, the clock error in the LEO receiver also must be eliminated. However, the LEO usually carries an inferior frequency standard, which requires another strategy for eliminating this error source. This is depicted in Fig. 1-1, which shows the GPS receiver onboard the LEO observing the setting GPS satellite and simultaneously a second GPS satellite clear of any intervening medium, a so-called "clock" or reference satellite. (We assume that the dual-frequency phase measurements eliminate phase effects from the ionosphere.) Differencing the phase measurements from concurrent tracking of these two (now synchronized) GPS satellites at the same reception epoch eliminates the LEO clock error. The cost of this strategy in measurement precision is essentially a $\sqrt{2}$ increase in data noise, usually not a significant limitation except in threshold detection situations.

1.1.5 Ephemerides

For occultation applications, the knowledge error in the relative velocity between satellites must be controlled to a few tenths of a millimeter per second over the roughly 100-s duration of an occultation episode for the neutral atmosphere. This translates into a precision orbit determination (POD) requirement on the LEO of about 30 cm in accuracy, readily achievable these days with a GPS space geodetic receiver onboard. For the GPS constellation, the orbit information for each GPS satellite is included in its navigational signal and is maintained and operated by the USAF, usually with an accuracy of a few meters, or to a few hundred parts per billion. But for many scientific applications of the GPS, particularly space geodetic applications, one needs an accuracy of at most a few parts per billion. Pursuant to this goal, the International GPS Service (IGS) was inaugurated under the auspices of the International Union of Geodesy and Geophysics (IUGG) about a decade ago. The IGS consists of a globally distributed network of over 350 GPS ground tracking stations operated by almost as many different organizations, several communication and data information centers, data analysis centers, and a central bureau for oversight and user interfacing [36,37]. The IGS is comprised of over 200 organizations from over 80 countries collaborating on maintaining

and operating its various system elements. The scientific products from the IGS include very accurate and reliable ephemerides for the GPS satellites accurate to about 1 decimeter, plus a realization of a terrestrial reference frame for ground-based GPS receivers accurate to 1 to 2 parts per billion. These IGS-provided accuracies exceed the requirements for occultation applications.

1.2 Information Content in GPS Occultation Observations

A radio occultation observation profile consists of sequences of amplitude and phase measurements of the L1 and L2 carriers obtained by the LEO receiver over the course of an occultation episode. For the neutral atmosphere, a typical occultation contains a few thousand data points, depending on the sample rate. The measurement error on each point is statistically independent. Clock epoch errors in the transmitter and receiver have been removed using the differential tracking techniques described earlier. Using the POD information about the LEO and the GPS satellites, one can accurately calculate the phase accumulation from satellite kinematics. Subtracting this from the phase measurement leaves as a remainder the excess phase from refraction. Thus, the information content about the refracting medium is contained in the amplitude and excess phase measurement sequences, such as those shown in Figs. 1-2 and 1-3.

1.2.1 Connected Phase

An important aspect of the excess phase measurement sequence is its connectedness. To maintain an accurate phase profile over time, successive phase measurements must (or at least should) be connected, which means that the integer number of cycles accumulated between successive measurement epochs must be exactly accounted for in addition to the fractional cycle determinations made by the receiver at each epoch. In benign signal conditions, the receiver itself can produce connected phase measurements using an internal phase model based on previous phase measurements to extrapolate forward to the next measurement epoch. But in adverse conditions, the raw measurements must be supplemented with the application of more realistic models of the excess phase between epochs. Based on statistical studies from actual occultation profiles, the uncertainty in empirical excess Doppler models for the Earth's atmosphere is about 10 Hz, but with some outliers, which tend to become exacerbated deeper in the moist lower troposphere [34]. Thus, the maximum uncertainty in accumulated cycles between measurement epochs usually is about $10\Delta t$, where Δt is the separation time between successive epochs. For $\Delta t = 0.02$ s, the uncertainty in the integer cycle count usually is 0.2 or less. Fixing cycle breaks can be a significant off-line data-editing task, and sometimes flags are required to denote unfixable cycle breaks.

1.2.2 Sample Rate Versus Vertical Resolution

The canonical sample rate for the phase and amplitude measurements is 50 Hz, which is convenient for the receiver because that is the chip rate of the GPS header code. This is another phase-modulating, very low-rate, square-wave code on the L1 and L2 carriers that carries almanac, timing, health, and other information for the tracked satellite. The individual phase and amplitude samples can be averaged and reported at a coarser rate to reduce thermal noise effects, i.e., to increase the SNR, but at a cost of potentially poorer vertical resolution in the refracting medium. The average vertical velocity of a ray path tangency point is 2 to 3 km/s in the upper atmosphere and an order of magnitude smaller in the lower troposphere, where the refractive gradient is much stronger, and even slower where very large refractive gradients from water vapor occur. Thus, averaging the measurements over a time span broader than where thermal noise is no longer the limiting error source will further limit the vertical resolution with little benefit in measurement accuracy.

Resolution topics, such as the first Fresnel zone, and related topics are discussed further in Chapter 2 and Appendix A. The first Fresnel zone defines the resolution perpendicular to the ray path of a single occultation observation, much like the Airy disk provides a resolution threshold for a circular lens in an optical instrument. As with the Airy disk, the rays passing through this zone are more or less in phase and interfere "constructively"; rays outside this zone interfere destructively. In a refracting medium, the first Fresnel zone is elliptically shaped. The vertical radius is given by

$$\mathscr{F} = \sqrt{\lambda D \,|\zeta|} \tag{1.2-1}$$

where λ is the wavelength of the GPS carrier, about 20 cm; D is the limb distance of the LEO, roughly 3000 km; and ζ is the defocusing factor. The vertical width of the first Fresnel zone is on average about 1.5 km in the upper atmosphere and roughly 1/2 km in the lower troposphere from dry air alone. But near the Earth's surface, it can be an order of magnitude smaller when strong refractive gradients from water vapor are present. It suffices here to note that the wave properties of the GPS signal, phase, and amplitude can be used with a sequence of observations to achieve a vertical resolution that is a small fraction of the width of the local first Fresnel zone, perhaps as small as 10 percent. A discussion of data-smoothing strategies involving the trade-off between resolution and measurement accuracy can be found in [34].

1.2.3 Inverting Radio Occultation Data

Although a wave-theoretic approach to recovering the refraction properties of the atmosphere is the principal theme of this monograph, it often is more

convenient and useful to cast propagation processes in terms of geometric optics. The following discussion is set in a ray-theoretic context.

The radio occultation profiles of phase and amplitude are analyzed, broadly speaking, in either a stand-alone mode or in an assimilation mode. In the stand-alone mode, the observation sequences are used to determine a profile for the index of refraction $n(\boldsymbol{r})$ from some sort of inversion process without preponderant reliance on a priori information. Usually some kind of symmetry or stratification geometry must be invoked for the medium to unambiguously determine $n(\boldsymbol{r})$. Effectively, the number of degrees of freedom in defining $n(\boldsymbol{r})$ must be no more than the number of independent observations in an occultation profile in order to achieve an unambiguous determination without using a priori information. For the Earth, local spherical symmetry that accounts for the Earth flattening usually is assumed. Local inhomogeneity in the refractivity along an equipotential surface usually is a small error source for the middle troposphere and higher, but water-vapor-induced horizontal gradients in refractivity near sea level can become a significant error source [38].

Figure 1-4 shows a schematic of one "onion layer" approach. Within each layer, the gradient of the index of refraction is assumed constant. Thus, we have a sequence of unknown parameters, $n_1', n_2', \cdots, n_M'$, to be determined from M or more observations. At a specific epoch within the occultation episode, the ray traverses the atmosphere down to a minimum depth, the kth layer, but no deeper. Therefore, the observations for this ray will depend only on the k parameters $n_1', \cdots, n_k'$. If an unambiguous relationship holds between the observation at a given epoch and the ray—and this is a big if—then it is straightforward to form a linearized system of equations involving a triangular information matrix relating the observation sequence to these refraction parameters. The schematic in Fig. 1-4(a) depicts (with exaggerated bending) an ordered set of rays, i.e., no multipath, with a one-to-one relationship between the ray arriving at the LEO and the value of a, the impact parameter of the ray. With each successive measurement epoch in this figure, the ray at its tangency point passes through a new layer that is lower in altitude and that was not "sounded" before. In Fig. 1-4(b), we have multipath for some periods when more than one ray arrives at the LEO at the same time; the impact parameter of each ray can be moving either up or down with time, depending on the type of ray. Therefore, an analysis scheme must first be applied to the observations to discriminate among concurrently arriving rays.

In the linearized version relating observations to refraction parameters, an observation at a given epoch is the actual observation minus the predicted observation based on an initial estimate for parameter set $\hat{n}_1', \hat{n}_2', \cdots, \hat{n}_M'$, and based on a provisional ray path passing through the medium that follows Snell's law. Inversion of this information matrix, or inverting a weighted least-squares matrix version if over-determined, yields a determination of the

corrections $\Delta\hat{n}_1', \Delta\hat{n}_2', \cdots, \Delta\hat{n}_M'$. One then updates $\hat{n}_1', \hat{n}_2', \cdots, \hat{n}_M'$ by adding the corrections to these predictions and generates a new ray path, forms new observation differences, and iterates until convergence is achieved. Assuming that $n_0 = 1, n_0' = 0,$ one can integrate this converged sequence $\hat{n}_1', \hat{n}_2', \cdots, \hat{n}_M'$ downward to obtain a recovered profile $\hat{n}(\boldsymbol{r})$ for the index of refraction. As a practical matter, at very high altitudes, above 50 to 60 km, the noise in the occultation observable from measurement error or ionospheric calibration error becomes larger than the signal from the atmosphere. In these high altitudes, to initiate the downward integration process, one can statistically combine in a maximum-likelihood sense the noisy occultation observable with an a priori value furnished by a model atmosphere, along with the appropriate covariance matrices [39]. Because refractivity varies approximately exponentially with altitude at these heights, an error in this matching and handing-over process (from strong reliance on a priori information at higher altitudes to strong reliance on the data at lower altitudes) is damped out exponentially with decreasing altitude and usually is nearly gone within 2 scale heights.

When spherical symmetry applies and only a single ray arrives at the LEO over time, or when multiple rays can be discriminated according to their excess Doppler signature, then both the bending-angle and impact parameter values α and a are obtained from the derived excess Doppler information for each ray. Excess Doppler for a given ray is the time derivative of its excess phase. In this case, the Abel integral transform [see Eq. (1.2-5)] can be applied [40,41]. This integral transform directly recovers without iteration the profile $\hat{n}(r)$ from the observed bending-angle and impact parameter sequences for each evolving ray arriving at the LEO. At very high altitudes where the atmospheric signal is weak, the same kind of statistical hand-over using predicted excess Doppler from a model is still used to initiate the integral downward.

The accuracy of the recovery of the index of refraction can be determined theoretically based on an assessment of the error sources within the occultation system. It also can be inferred from statistical comparisons with NWP model values from, for example, the European Centre for Medium-Range Weather Forecasts (ECMWF) [19,42]. Fractional accuracy for the refractivity recovery of 0.1 to 1.0 percent is achievable from about 5 km up to 30 to 40 km. Here the assumption of local spherical symmetry usually is valid and the signal from the atmosphere is relatively strong compared to the measurement noise. But in the lower troposphere, horizontal variations in refractivity can become a significant error source when laminar symmetry is assumed, and tracking becomes more difficult and the measurements are noisier. These factors can limit further improvement with depth of the accuracy of the refractivity recovery.

If the chemical composition of the medium is known, one knows precisely the relationship between the density of the medium and its refractivity. In this case, one obtains the density profile $\rho(r)$ of the refracting medium from the

recovered profile for $n(r)-1$. Invoking hydrostatic equilibrium for the medium in a gravity field $g(r)$, one has the hydrostatic equation for the pressure $dp/dr = -g\rho$. Integrating this yields the pressure profile $p(r)$. Invoking thermodynamic equilibrium allows one to apply the ideal gas law to the recovered density and pressure profiles to obtain the temperature profile.

For the Earth, the gaseous constituents of the neutral atmosphere affecting the refractivity at microwave frequencies are dry air and water vapor. For an air/water system in thermodynamic equilibrium, the Clausius–Clapeyron equation gives the partial pressure of water vapor, which varies exponentially with $-1/T$ [43]. Reducing the temperature by 10 percent reduces the saturated water vapor content by an order of magnitude. Below a threshold temperature of about 250 K, it is too cold for any residual water vapor to significantly contribute to the refractivity at L-band frequencies. The altitude corresponding to a temperature of 250 K typically is in the range 6 to 8 km. For altitudes above this threshold and below roughly 40 km, the potential accuracy of the temperature recovery is less than 1 K, probably less than 0.5 K for the 10- to 25-km range [21]. For higher atmospheric temperatures at lower altitudes, water vapor becomes an additional component in the refractivity. We speak of the "dry delay," which is the phase delay caused by the refractivity from air, and the "wet delay," which is the delay solely from water vapor. The latter arises from the permanent electric dipole moment of the water molecule, which is a significant contributor to the refractivity at microwave frequencies. The problems with the wet delay generally are that the density of water vapor is highly variable in space and time, its relative abundance or specific humidity is uncertain, and its refractivity per mole at L-band frequencies is much larger, about 16 to 18 times larger, than the refractivity per mole of dry air. Especially in the lower troposphere, this variability in specific humidity results in large refractive gradients that can double or even triple the refractive bending from dry air alone. In the stand-alone mode, therefore, the radio occultation technique requires ancillary information for those regions where the temperature is above 250 K to convert the recovered refractivity profile into unambiguous density, pressure, temperature, and specific humidity profiles. In temperate and tropical regions where the water vapor content is a major uncertainty, the temperature profile usually is relatively better known from NWP models, and it is often used to aid the occultation recovery of water vapor [19,21].

1.2.4 Assimilating Radio Occultation Data

In the assimilation mode, the occultation observable profiles are assimilated through statistical inference processes into a global atmospheric model where the basic meteorological quantities of the refracting medium, density, pressure, temperature, specific humidity, cloud, aerosol and precipitation distributions,

wind fields, and so on are defined. Therefore, the index of refraction $n(\boldsymbol{r})$ is already constrained to varying degrees in space and time by a usually much larger and more global data set in the atmospheric model. Covariance matrices characterizing the statistical aspects of the global data set and the atmospheric models controlling thermodynamic and dynamic processes are also part of this global system. In contrast to the stand-alone mode, no symmetry or stratification assumptions about the refracting medium are theoretically required here. Moreover, the granularity of the model for the refracting medium for numerical computation or, equivalently, the four-dimensional cell size (three in space and one in time) is already established and controlled by the continuity equation. Usually the number of degrees of freedom, even when constrained by the physical equations in the model, is vastly more than found in a typical stand-alone occultation model.

The four-dimensional variational analysis (4DVAR) technique in NWP programs is a prime example of the assimilation mode for using the occultation data [44–47]. Here the occultation profiles are assimilated competitively in a statistical treatment of the data. The radio signal from the position of the occulted GPS satellite is propagated through this global model atmosphere with its free parameters provisionally fixed. At each observation epoch in the occultation sequence of M observations, the difference between an actual observable of the GPS signal (e.g., excess phase, excess Doppler or bending angle, signal amplitude) and its propagated value based on the provisional parameter set is computed. This difference becomes a component of an additional M-dimensional observation vector in the global data set. One adjusts the free parameters of the global model to minimize a cost function characterizing the goodness of fit of all the data weighted by their covariance matrix inverses, including the occultation data. The efficacy of the occultation information in this environment to modify the values of the free parameters obviously depends on the information content in the global system and on the assigned covariance matrices.

1.2.5 Rays and Stationarity

The Poynting vector, which defines the direction and magnitude of the radiant power in an electromagnetic wave at a given point in terms of the cross product of the electric and magnetic field vectors, is perpendicular to the cophasal surface of the wave at that point. The limiting form that the Poynting vector takes as the wavelength of the wave is shrunk to zero defines the tangent vector of the ray in geometric optics. The second-order differential equation that results from this limiting process gives the curvature of the ray at any point in terms of the gradient of the refractivity there. This limiting process also gives the amplitude of the ray based on the Poynting vector, its reflection and transmission properties that would apply to an electromagnetic wave of very

short wavelength, and so on. What is missing in ray theory, by definition, is diffraction. Also, ray theory (as geometric optics is usually defined) cannot deal adequately with caustic rays or with trapped rays in a super-refracting medium where the curvature of a ray, if it could exist in this environment, would temporarily be greater than the curvature of the local equipotential surface. A caustic ray is one that first comes in contact with the envelope to a family of rays, for example, a spectral component of a rainbow. Born and Wolf provide a comprehensive discussion of the foundations of geometric optics following this limiting approach in [48], which cites many references giving the historical development of geometric optics.

The same differential equation for the ray (the Euler equation) comes from the calculus of variations upon applying Fermat's principle (see Appendix A). The phase accumulation φ along a path through a three-dimensional refracting medium from point A to point B may be written in parametric form as

$$\varphi = k\int_A^B n(\boldsymbol{r})ds = k\int_A^B n(\boldsymbol{r})\left(\dot{x}^2+\dot{y}^2+\dot{z}^2\right)^{1/2}dt \tag{1.2-2}$$

where $k = 2\pi/\lambda$ is the wave number of the wave, $n(\boldsymbol{r})$ is the index of refraction at the position $\boldsymbol{r} = \hat{\boldsymbol{x}}x + \hat{\boldsymbol{y}}y + \hat{\boldsymbol{z}}z$, $ds = (\dot{x}^2+\dot{y}^2+\dot{z}^2)^{1/2}dt$ is the incremental arc length along the path, t is an arbitrary parameter denoting position along the path, $\dot{(\ )} = d(\ )/dt$, and the triad $(\dot{x},\dot{y},\dot{z})$ defines the slope of the path. The path is described by the functions $x(t)$, $y(t)$, and $z(t)$, and we now stipulate that they describe a ray path. Fermat's principle requires that this path integral along a ray be stationary. In other words, the phase accumulation from the same initial point to the same end point along any other physical path neighboring the ray path would differ from the phase accumulation along the ray path in a second-order manner. Let the neighboring path be described by the functions $x(t)+\varepsilon\xi_x(t)$, $y(t)+\varepsilon\xi_y(t)$, and $z(t)+\varepsilon\xi_z(t)$, where ε is a small parameter. The functions $\xi_x(t)$, $\xi_y(t)$, and $\xi_z(t)$ are completely arbitrary other than that they must be physical and they must satisfy the boundary conditions, i.e., $\xi_x(t_A) = \xi_x(t_B) = 0$, etc. Then $d\varphi/d\varepsilon$ must be zero when evaluated along the ray, on which $\varepsilon = 0$. This stationary-phase condition requires that the ray path at every point satisfy the Euler differential equation or its integral equivalent where a discontinuity in $n(\boldsymbol{r})$ occurs. Usually the stationary value of φ along a ray is a local minimum, i.e., $d^2\varphi/d\varepsilon^2 > 0$, but not always. Anomalous rays provide a local maximum, and caustics involve a breakdown of the assumption of a non-zero second-order variation, which is embedded in the foundations of geometric optics.

The calculus of variations is a powerful technique for developing basic ray theory and many of its attributes. It may be used to develop the eikonal equation for obtaining phase delay along a ray, the transversality condition for

relating the change in end-point phase delay to small changes in end-point position (including relating excess Doppler to refractive bending angle), the Hamilton–Jacobi theory for obtaining the end-point phase delay expressed as a field variable in terms of the end-point position variables for a family of rays, and finally, for obtaining the canonical variable representation of a ray. We discuss these aspects in Chapter 2 and in Appendices A and B.

The stationary-phase property of rays affects one's strategy for processing occultation data using ray theory, depending on one's objective. In a sense, stationarity is both a blessing and a curse. Regarding the latter, consider the process of finding the ray path passing through the end points A and B for a given trial index of refraction profile $\hat{n}(\boldsymbol{r})$. One can forward propagate from A or backward propagate from B using Snell's law and a ray-tracing method. By adjusting the slope parameters of the ray at these end points, one theoretically can converge on a complete ray path that matches the boundary conditions, if such a ray exists. The problem is that the calculated phase delay obtained by following this provisional path also must match the observed value φ for each observation epoch. This in turn requires adjustments in $\hat{n}(\boldsymbol{r})$ to reach the observed values for φ; for the stand-alone mode, this would require an inversion process such as that described earlier. But along the ray, the phase delay is stationary with respect to small variations in the slope parameters or to variations in any other parameter, for example, the impact parameter, that characterizes an alternative nearby path that satisfies the boundary conditions. There is a singularity here. In effect, the "density" of admissible paths around the actual ray with a phase delay within the limits $(\varphi, \varphi + \delta\varphi)$, all satisfying the boundary conditions, becomes infinite along the actual ray where $\delta\varphi = 0$. This singularity in path density forms the basis of Fermat's principle. The idea is that all physical paths in the vicinity of the actual ray have the same phase delay to first order. Therefore, the phasor, $\exp(i\varphi)$, for each path neighboring the ray is the same, to first order, and when the phasors from all of these possible nearby paths are summed (Huygen's principle) and averaged to obtain the total field at point B, they sum constructively, reinforcing the amplitude. For paths with larger deviations from the ray, their first-order phase variation with respect to path deviation is not zero. Therefore, their phase delays for different paths in their neighborhood are distributed more or less randomly over 2π radians, and their phasors destructively combine. No average field results from those paths with a non-zero first-order variation, only fluctuations. This concept of constructive reinforcement for nearby paths about a ray and destructive reinforcement for paths away from the ray forms the basis for the idea of a ray, the basis for Fermat's principle, and the basis for the first Fresnel zone. It also is the central concept in the Feynman sum-over-histories technique used in quantum electrodynamics to calculate the quantum wave function for the probability amplitude of a quantum event [49].

This singularity in path density results in numerical instability when the measured phase sequence per se is used as a criterion to reconstruct the most probable ray path for each observational epoch. On the actual ray, $d\varphi / d\varepsilon = 0$, where ε is a parameter describing the departure of a trial path from the actual ray path. By tinkering with ε, one tries to converge to the actual ray path. One would use an iterative differential correction scheme of the linear form $\delta\varepsilon = -(d\varphi / d\varepsilon)^{-1}\delta\hat{\varphi}$, where $\delta\hat{\varphi}$ is the observed value of φ minus the computed value $\hat{\varphi}$ based on the provisional profile $\hat{n}(\boldsymbol{r})$ and the provisional path through this medium between the end points. Here one attempts to determine iteratively the maximum-likelihood estimates of $\delta\varepsilon$, $\hat{n}(\boldsymbol{r})$, and the path. This scheme will run into computational trouble after successive iterations as the iterated path approaches the actual ray path because $d\varphi / d\varepsilon \to 0$ as $\varepsilon \to 0$.

1.2.6 Excess Doppler

Stationarity also brings a blessing, the transversality condition. If one slightly changes the end-point values of the ray, a new ray will pass through the new end points. The phase delay along the new ray between the new end points minus the phase delay along the old ray between the old points depends only on local conditions at the end points, not on the path in between. This is rather remarkable considering that the phase delay itself is computed from a path integral along a new path, which explicitly accounts for the variability in $\hat{n}(\boldsymbol{r})$ along the entire ray path. This local rather than global dependency on change is a unique property of a stationary-phase path.

Using the transversality condition, it is shown in Appendix A that the rate of change in excess phase resulting from the satellite velocities is given by

$$\lambda f_D = n_{\mathrm{G}}\boldsymbol{T}_{\mathrm{G}} \cdot \boldsymbol{V}_{\mathrm{G}} - n_{\mathrm{L}}\boldsymbol{T}_{\mathrm{L}} \cdot \boldsymbol{V}_{\mathrm{L}} - \dot{r}_{\mathrm{GL}} \tag{1.2-3}$$

where $f_D = (d\varphi / dt) / 2\pi$ is the rate of change of the excess phase in cycles, or excess Doppler shift of the ray, and λ is the wavelength of the harmonic wave. Here G denotes the position of the end point of the ray at the transmitting GPS satellite, and L denotes the end point position at the LEO. Also, $\boldsymbol{V}$ is the velocity vector of the satellite at an end point of the ray, $\boldsymbol{T}$ is the unit tangent vector of the ray at an end point, and $\dot{r}_{\mathrm{GL}}$ is the radial velocity between the end points G and L (see Fig. 1-6). POD information provides a determination of $\boldsymbol{r}_{\mathrm{G}}$, $\boldsymbol{V}_{\mathrm{G}}$, $\boldsymbol{r}_{\mathrm{L}}$, and $\boldsymbol{V}_{\mathrm{L}}$, and therefore $\dot{r}_{\mathrm{GL}}$, at each observational epoch. It follows that the excess Doppler shift for a given ray, a quantity directly obtained from the observed excess phase profile for that ray, imposes a constraint on the unknown unit ray path tangent vectors $\boldsymbol{T}_{\mathrm{G}}$ and $\boldsymbol{T}_{\mathrm{L}}$. If we assume coplanar propagation, usually a good assumption, then $\boldsymbol{T}$ is defined by a single deflection angle

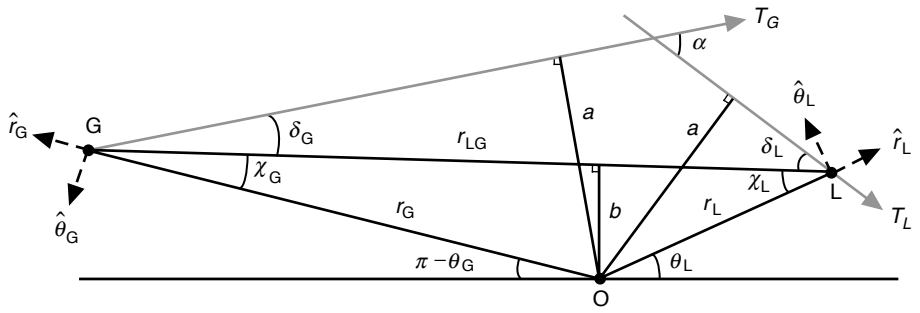

Fig. 1-6. Ray path geometry from point G to point L in the plane of propagation for a spherical symmetric medium.

relative to the vector $\boldsymbol{r}_{GL}$ between end points G and L (see Fig. 1-6). Ray tracing and Snell's law provide a second constraint on these two deflection angles. If we assume the satellites are out of the refracting medium, then $n_G = n_L = 1$. Thus, one determines the refractive bending angle, the angle α between $\boldsymbol{T}_G$ and $\boldsymbol{T}_L$, provided that the ray is unique. No singularity arises in use of the excess Doppler to converge to the actual ray path from point G to point L.

The catch is that Eq. (1.2-3) requires that the ray arriving at point L from point G be unique. Equation (1.2-3) applies only to a single ray, not to a composite. One must be sure a priori that the observed phase profile results uniquely from a single ray or at least very nearly so, not from some composite of full-fledged rays arising from multiple stationary-phase paths through the atmosphere, all starting from point G and arriving at point L. These multipath rays arrive concurrently at L, but each has its own unique excess Doppler signature because their $\boldsymbol{T}$ vectors differ. We return to this multipath issue later.

Assuming that one has recovered the bending angle without ambiguity for a given ray, one obtains the impact parameter

$$a = n\,|\,\boldsymbol{r} \times \boldsymbol{T}\,| = nr\sin\gamma \tag{1.2-4}$$

for that ray, which is a constant along the ray when spherical symmetry applies. This is Bouguer's law, essentially Snell's law for a spherical symmetric medium. Here γ is the angle between the radius vector and the tangent vector of the ray. In ray theory, the constancy of this impact parameter a is the analog of the conservation of angular momentum in classical mechanics. This is shown in Fig. 1-4. The impact parameter is related to the tangency point (r_*, θ_*) of the ray by the condition $a = r_* n(r_*)$. Then, as mentioned earlier, with the ordered series of values for the bending-angle and impact parameter pair (α_k, a_k), $k = 1, 2, \cdots, M$, one can form the function $\alpha(\rho) = \alpha_k$, $\rho = a_k$, $k = 1, 2, \cdots, M$, and one can use the Abel transform

$$\log n(a) = \frac{1}{\pi}\int_a^\infty \frac{\alpha(\rho)}{\sqrt{\rho^2 - a^2}}\, d\rho \tag{1.2-5}$$

to recover the refractivity profile in terms of the impact parameter a. The Abel transform is derivable from the basic integral equation for the bending angle in a spherical symmetric medium (see Appendix A):

$$\alpha(a) = -2a\int_a^\infty \frac{d\log n}{d\rho}\frac{d\rho}{\sqrt{\rho^2 - a^2}} \tag{1.2-6}$$

By substituting this form for $\alpha(a)$ into the integral operation on the right-hand side (RHS) of Eq. (1.2-5), one indeed recovers $\log n(a)$. If significant multipath occurs within a section of the observation series, then it should be noted that $a_k,\ k = 1,2,\cdots,M$ will not be time-ordered during that section. But, as long as one can successfully recover a sequence (α_k, a_k) that is ordered, then the Abel transform can be applied.

If spherical symmetry does not apply, then some ray tracing technique combined with the appropriate stratification model for the refractivity can be used to recover the refractivity profile, or one can use the assimilation mode.

Note from Eq. (1.2-6) that the existence of a ray between given end points with an impact parameter value of a requires that $\rho \geq a$ at all points along the ray. In a super-refracting layer, $d\rho / dr = (n + rn') < 0$, and this condition $\rho \geq a$ can be violated for a certain range of tangency points r_*. Within this range, no rays exist. Chapter 6 briefly discusses super-refractivity in a spherical shell.

1.3 Scientific Applications of GPS Occultation Observations

We briefly review the major uses of occultation observations, which fall broadly into the categories of meteorology, weather prediction, and global climate change. Accuracy, resolution, and the global distribution of the GPS occultations are key factors. Regarding accuracy, we have already noted the accuracy of the refractivity recovery from the stand-alone mode, 0.1 to 1 percent fractional accuracy above the lower troposphere, where the assumption of local spherical symmetry is likely to be satisfactory, and below 40 km. This translates into the same fractional accuracy range for the density determination for temperatures below 250 K. It also implies dekameter accuracy for a geopotential height determination of a point on a constant pressure surface, for example, the difference in the potential for the 300-mbar surface minus the potential of the ocean geoid, with the difference divided by the gravity coefficient [50]. It also translates for the same altitude range into sub-kelvin accuracy in temperature recovery.

Regarding these accuracy figures, the caveats about the applicability of spherical symmetry and the temperature being colder than 250 K have more to do with the model used and the analysis approach, and less to do with the inherent accuracy of the measurements. Good, clean phase and amplitude measurements down to the Earth's surface could yield 0.1-K accuracy for temperature recovery if we knew the water vapor content and the stratification. So, these caveats are somewhat misleading and tend to understate the inherent strength of the occultation data if our strategy is instead to assimilate them into a 4DVAR process.

Radio occultations bring new capabilities to the study of processes in meteorological models. With occultation data, one can study the fine structure of various refractive boundaries, such as the tropopause, a marine boundary, and a sporadic E-layer in the ionosphere. Temperature and pressure recovery from the middle troposphere through the middle stratosphere can lead to better understanding of energy transport and exchange processes across the tropopause and of how the atmosphere radiates dynamical energy and heat through acoustic gravity waves. Gravity waves are quite evident in occultation-derived temperature profiles in the stratosphere [51–54]. Phase and amplitude measurements made down to or near the Earth's surface provide powerful constraints on water vapor, which is such an important greenhouse gas and so relevant to energy transfer and balance processes in the atmosphere and to aerosol growth and cloud formation. Occultation-derived specific humidity distributions will provide sharper information about cyclogenesis processes, baroclinic wave and wind field development, and so on. More detail about many promising new capabilities is provided in [22,55].

1.3.1 Weather

The utility of occultation data when assimilated into 4DVAR programs depends greatly on the degree of complementarity of the information content in the occultation data set compared to the 4DVAR data set. The utility will be highest when other data are relatively sparse or when certain physical parameters of the 4DVAR model are less well constrained. This includes both data-sparse geographical locations and certain locations in altitude, even in data-rich geographical areas. Occultation profiles are in their best form for tangency points in the upper troposphere up to the lower stratosphere. Here the technique can provide unprecedented accuracy in density, pressure, and temperature, and it offers very sharp (perhaps dekameter-level) vertical resolution.

The utility of the occultation data also depends on their latency. The ephemeral aspects of the weather force very short latency requirements. That is why planned operational LEO occultation constellations, such as COSMIC,

have a 2- to 3-hour latency goal: occultation data to be completely assimilated within 3 hours of the actual occultation event.

Regarding geographical complementarity, consider again the planned COSMIC constellation, which will return on average about 10 occultations within a 3-hour period over a geographical area the size of the continental U.S. Obviously, 10 occultations from southern ocean or polar locations where weather data are sparse will have a bigger impact in weather models for those areas than will 10 occultations obtained over the continental U.S. and southern Canada in weather models for that area. In North America, the myriad weather sensors and observation programs is constraining.

Nonetheless, the water vapor information provided by the radio occultation data has a global utility. These data are powerful because they highly constrain the water vapor uncertainty. Accurate knowledge of specific humidity distributions, particularly in the lower troposphere, is key to more accurate weather prediction, for example, the ongoing development of wind fields, cloud formation, and precipitation in cyclogenesis processes. It is the water vapor information that makes reaching the Earth's surface with a high percentage of the radio occultations such an important goal for operational programs [56].

Pursuing this complementarity concept further, let us see how the water vapor information content in an occultation might fold into a 4DVAR process. The refractivity at L-band frequencies in the neutral atmosphere is given by [57,58]

$$N = n - 1 = \left(77.6 \frac{p}{T} + 3.73 \times 10^5 \frac{p_{\mathrm{w}}}{T^2} \right) \times 10^{-6} + \text{neglible terms} \quad \text{(1.3-1a)}$$

where p is the pressure of the moist air in mbars, p_{w} is the partial pressure of the water vapor, and T is temperature in kelvins. Here and throughout this monograph, we use an unconventional definition for refractivity, $N = n - 1$, instead of the standard $N = (n-1) \times 10^6$. Let us rewrite the refractivity equation in terms of the molecular number densities, n_{a} and n_{w}, in moles / m^3, where n_{a} and n_{w} are the dry air and water vapor number densities, respectively. We use the ideal gas law $p_{\mathrm{x}} = n_{\mathrm{x}} RT$, where R is the universal gas constant. Then Eq. (1.3-1a) becomes

$$\left.\begin{aligned} N &= c\left(n_{\mathrm{a}} + b(T) n_{\mathrm{w}}\right) + \text{neglible terms} \\ c &= 6.45 \times 10^{-6} \ \ \mathrm{m}^3 \cdot \mathrm{mole}^{-1}, \quad b(T) = 16.0\left(\frac{300}{T}\right) + 1 \end{aligned}\right\} \quad \text{(1.3-1b)}$$

The relatively high sensitivity of the refractivity to the density of water vapor should be noted; the value of $b(T)$ typically is 16 to 18. Because the mean

molecular mass of water is 18 g/mole and dry air is 29 g/mole, the sensitivity of the specific refractivity is even greater.

Suppose now that we have inverted the occultation data to obtain an estimate $\hat{N}(r)$ of the refractivity. We could make a similar case for complementarity using a basic occultation observable, such as excess phase, Doppler, or amplitude. But it is easier to show using the recovered refractivity. Assume that at some epoch during the occultation episode the tangency point of the ray is located at a particular altitude where water vapor is a significant factor. Suppose that the error in the recovered refractivity there, $\delta\hat{N}$, is a random variable that is Gaussian distributed with a mean value of zero and a standard deviation of $\sigma_{\hat{N}}$. Then, from Eq. (1.3-1b), the probability density distribution $P[\delta\hat{N}]$ for $\delta\hat{N}$ written in terms of corresponding Gaussian errors in $\hat{n}_{\mathrm{a}}$ and $\hat{n}_{\mathrm{w}}$ is given by

$$P[\delta\hat{N}] = \frac{c^2 b}{2\pi\sigma_{\hat{N}}^2} \exp\left[-\frac{c^2 b^2}{2\sigma_{\hat{N}}^2} \left(\frac{\delta\hat{n}_{\mathrm{a}}^2}{b^2} + \frac{\delta\hat{n}_{\mathrm{w}}^2}{1^2} \right) \right] \tag{1.3-2}$$

Thus, a contour of constant probability density in $\delta\hat{N}$ is an ellipse with the semi-major axis along (or nearly so) the n_{a} axis and the semi-minor axis along the n_{w} axis. The coefficient $b(T)$ gives the ratio of these axes. The left-hand error ellipse, System A, in Figs. 1-7(a) and 1-7(b) gives qualitatively the contour for a constant probability density for $\delta\hat{N}$ from an occultation expressed in terms of the corresponding errors, $\delta\hat{n}_{\mathrm{a}}$ and $\delta\hat{n}_{\mathrm{w}}$, in the statistical estimates of n_{a} and n_{w}. For example, the ellipse might be the contour corresponding to the $1-\sigma$ values for $\delta\hat{n}_{\mathrm{a}}$ and $\delta\hat{n}_{\mathrm{w}}$; in this case, the probability of finding $\delta\hat{N}$ within the area inside of this ellipse is 40 percent.

Several comments can be made about Figs. 1-7(a) and 1-7(b). At this altitude of the tangency point and at this observation epoch, the error ellipse from the occultation information, System A, highly constrains the range of probable values for $\hat{n}_{\mathrm{w}}$ compared to the range of probable values for $\hat{n}_{\mathrm{a}}$. System B in Fig. 1-7(a) shows another error ellipse for $\delta\hat{n}_{\mathrm{a}}$ and $\delta\hat{n}_{\mathrm{w}}$ from another information system. It is derived from a presumably higher-dimensional covariance matrix describing the uncertainties in the recovered values for the global parameter set in System B mapped to the same epoch and altitude and projected into our two-dimensional $(\delta\hat{n}_{\mathrm{a}}, \delta\hat{n}_{\mathrm{w}})$ space. The information in System B is statistically independent from the information in System A. In Fig. 1-7(a), the information in System B is more or less "parallel" with the information in System A, nearly the same strength and the same weakness. In Fig. 1-7(b), System B is complementary to System A, more or less orthogonal, as characterized by the crossed error ellipses; the strength in

one system is a weakness in the other. The error ellipses for the two systems, A and B, may or may not be of comparable scale or eccentricity, but one can make an argument that they are often comparable. With the 2×2 covariance matrices for System A and System B, one can statistically combine the information in these two systems to obtain minimum variance estimates for n_a and n_w. The combined covariance matrix yields the System C minimum variance error ellipse. In Fig. 1-7(a), the combined System C, when the information in the two systems is comparable in accuracy, shows a roughly $1/\sqrt{2}$ improvement in probable errors all around, but it inherits the same strengths and weaknesses from the two systems. In the orthogonal case in Fig. 1-7(b), System C inherits the strengths from each system and none of the weaknesses; its error ellipse is dramatically improved. Water vapor content in NWP models is very difficult to know accurately in place and time, and it is generally less well-known in these models than are the total pressure and temperature. The near-orthogonal error ellipse in Fig. 1-7(b) for System B probably is more typical for NWP models.

In summary, in weather applications, complementarity from radio occultations, if they are timely, comes mainly from locations with relatively sparse meteorological data or from regions with more loosely constrained parameters. The other major contribution to NWP from the radio occultations is their highly constrained water vapor information, as just discussed. The long-perceived bête noir of the radio occultation technique for the Earth—water vapor—is in fact its strength.

1.3.2 Climate

In global climate change studies, detecting weak signatures in a noisy Earth system environment over decadal periods and longer is a major challenge. Here,

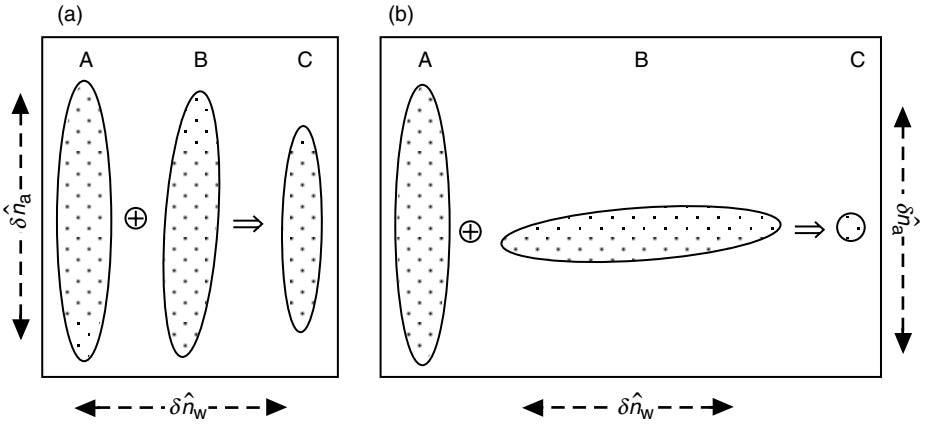

Fig. 1-7. Error elipses for independent information systems A and B, and their minimum variance combination C: (a) parallel information and (b) complementary information.

the stability, high accuracy, and fine vertical resolution of the GPS occultations are key assets for use in climate studies. Except for the ionosphere, the GPS observations are essentially calibration-free, with no instrumental biases or calibrations that vary slowly. Even with the ionosphere, it would be a decadal variation in the usually small refraction effects from third- and higher-order frequency terms that could possibly "leak into" long-term observations in the stratosphere. This seems an unlikely scenario because modeling the ionosphere to aid the dual-frequency corrections significantly reduces errors from higher-order terms on the recovered refractivity in the stratosphere [59,60].

In monitoring possible secular trends, one can spatially average the recovered temperatures at a given altitude, or geopotential heights at a given pressure, to reduce by perhaps an order of magnitude the effect of random errors and variations. Thus, all the points lying within a specified geographical area, say a region 500 × 500 km in size, could be lumped and then averaged over time, weeks or months, enhancing the probability of detecting signatures with spectral power at very long periods.

Secular variations in the geopotential height of the 300-mbar surface, for example, can be a sensitive indicator of secular changes in the average temperature, $\overline{T}$, below. The sensitivity coefficient, $\partial h / \partial \overline{T}$, is about 30 m/K. A geopotential height change of 30 m at 300 mbar almost surely would be detectable over a long series of occultation observations, especially if it were regionally averaged, but ascertaining that it was a temperature signal rather than some other possible physical signal would remain a challenge.

A predicted secular effect associated with the global warming scenario is a warming troposphere but a cooling stratosphere. Therefore, monitoring averaged changes in recovered temperature profiles across the tropopause could be a very sensitive indicator of this warming trend. Other signatures associated with global warming to which averaged radio occultations might be applied are increased water vapor content and variability in the tropics, changes in the residual water vapor in the upper troposphere, and warming in the lower troposphere in polar regions. More examples and details on climate applications are given in [22,26].

1.4 Problems from Multipath and Some Remedies

The main problem from multipath, other than stressing the signal-tracking operations by the receiver, is in the ambiguity that it introduces regarding the appropriate excess Doppler value to use for each arriving ray. We have already noted in Fig. 1-3 the interference in amplitude from multipath rays for a particular occultation from GPS/MET. Figures 1-8(a) through 1-8(d) from CHAMP show similar interference in amplitude for a selected group of occultation profiles. The de-trended phase measurements for these occultations show similar interference. One should note from Fig. 1-8 the generally

improved SNR on CHAMP compared with GPS/MET. The GPS antenna on CHAMP has about 4 dB more gain, which on average yields about a factor of two increase in free space $\mathrm{SNR_V}$. These profiles reveal striking and highly transient features in signal amplitude as the ray path tangency point descends through the atmosphere. Occultations passing through narrowly defined layers of the ionosphere also reveal similar transients in SNR and phase as a result of sharp gradients in electron density. Examples of passing through a sporadic E-layer in the lower ionosphere are seen in Fig. 1-8(b) and will be found later in Chapter 2 in Figs. 2-15 and 2-16 from GPS/MET.

A characteristic of many of these transients is their brevity, which, when the downward velocity of the ray is considered, translates into an altitude range that is usually narrow compared to the vertical width of the first Fresnel zone. Moreover, deep troughs in the SNR of the received signal combined with contiguous and abrupt flaring (and concomitant transients in phase) are often observed. These are almost sure signs of multipath and/or shadow zones, and even of ducting on occasion through a super-refracting layer [Fig. 1-8(a)]. In some instances, caustic rays are evident and also diffraction fringes when the physical refractivity feature is sharp enough and the neighboring environment is smooth enough.

Multipath is essentially the rule in occultations, not an exception. There are benign periods, of course, but even those usually show some small interference or scintillation that exceeds the thermal noise; it is a matter of degree. For example, the SNR in Fig. 1-8(c) at $t = 0$ predicts a 1 percent thermal noise level $\left(\sqrt{50} / \mathrm{SNR_V}\right)$, but the actual noise is 2 to 3 times larger. One can average the samples down from a 50-Hz rate to a smaller rate, for example 1 Hz. Because the typical frequency difference, Δf, between interfering rays is several hertz, averaging dramatically improves the multipath "noise" compared to thermal noise. Over an averaging time $\Delta t >> (\Delta f)^{-1}$, the amplitude of the fringes from multipath averages down as $1/\Delta t$, whereas thermal noise averages down as $(1/\Delta t)^{1/2}$. On the other hand, averaging can degrade the vertical resolution potential. For a LEO, the vertical motion of the ray path tangency point spans the vertical diameter of the local first Fresnel zone in a time interval of $\Delta T \approx 0.6\zeta^{-1/2}$ s, where ζ is the defocusing factor, unity at the top of the atmosphere and usually 0.1 to 0.01 at sea level. Therefore, averaging observations over a time interval longer than ΔT exceeds the Fresnel limit, which is the resolution from only a single observation. Sub-Fresnel vertical resolution, obtained from an ensemble of observations by means analogous to the resolution obtained from synthetic aperture radar systems, and a holy grail for boundary studies, can be considerably hampered by time-averaging the observations. Ultimately good SNR is the key.

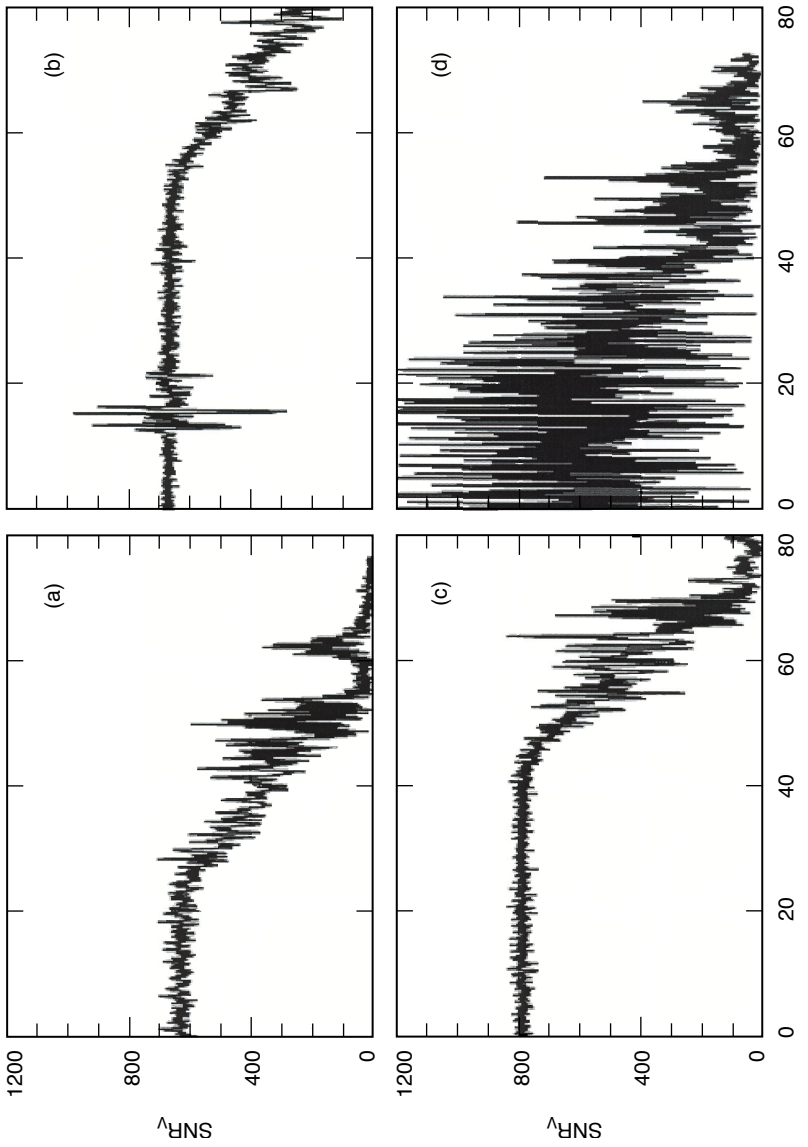

Fig. 1-8. Voltage SNR profiles from four CHAMP occultations in October 2001: (a) probable ducting in the lower troposphere (at ~54 to 62 s), (b) sporadic E-layer (at ~15 s), (c) strong, clean signal above the tropopause (at ~55 s), and (d) troublesome.

Figure 1-9 shows a phasor diagram in the complex plane containing a snapshot at a given instant of the phasors for three interfering rays, the "main" ray m, and the additional rays a and b. This is one of the simplest multipath scenarios, which usually comes in odd-numbered packets, 3, 5, 7,⋯, except at caustic contact points, or when a reflected ray is present. The vector sum

$$\hat{\boldsymbol{E}}(t) = \boldsymbol{E}_{\mathrm{m}}(t) + \boldsymbol{E}_{\mathrm{a}}(t) + \boldsymbol{E}_{\mathrm{b}}(t) \tag{1.4-1}$$

gives the observed amplitude and excess phase. The excess phase and amplitude of the individual components are not directly observable. The amplitudes of these rays, E_{m}, E_{a}, and E_{b}, usually differ and they also vary with time, but usually at a much slower rate than $\hat{E}(t)$ can vary because of interference. Their amplitudes are determined by their respective defocusing factors, which usually vary relatively slowly except near a caustic contact point. The individual excess phase of each contributing ray, φ_{m}, φ_{a}, or φ_{b}, depends on its individual ray path through the medium, which is changing with time as the ray path tangency point migrates downward or upward through the medium as a result of the orbital motions of the satellites. Because there are about 10^8 wavelengths along the ray path, the phases of these rays change rapidly with time. The relative phases, $\varphi_{\mathrm{a}} - \varphi_{\mathrm{m}}$ and $\varphi_{\mathrm{b}} - \varphi_{\mathrm{m}}$, vary at a rate that depends on the size of the refractivity perturbation that caused the multipath and on the local defocusing by smooth dry air. Typically this difference in rate is below 10 to 15 Hz. The relative values of the amplitudes E_{a} and E_{b}, compared to the amplitude of the main ray E_{m}, determine whether we have deep scintillation or light interference from these particular rays. Each of these three rays has its own excess Doppler signature because the directions of their ray path tangent vectors leaving the transmitting GPS satellite and arriving at the LEO are different. The excess Doppler relationship given in Eq. (1.2-3) is valid for each of these contributing rays, but not to the Doppler for $\hat{\boldsymbol{E}}(t)$. In

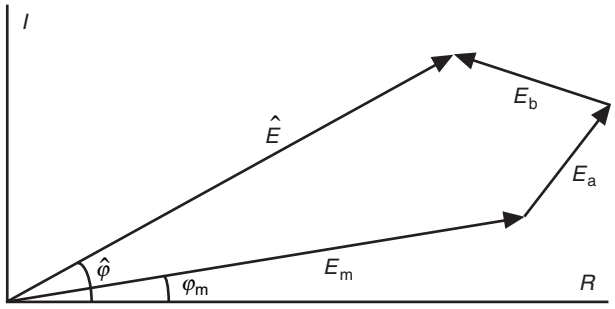

Fig. 1-9. Complex diagram showing composite amplitude $\hat{E}$ and phase $\hat{\varphi}$ from three multipath rays.

other words, $d\hat{\varphi}/dt$ does not yield per se a bending angle from Eq. (1.2-3) because $\hat{\boldsymbol{E}}(t)$ does not represent an actual ray, only a composite.

The interesting point about these multiple rays is that, although they arrive at the LEO at the same time, they all have different excess Doppler values because their tangent vectors at the LEO differ. Thus, from Eq. (1.2-3) it follows that when we have multiple rays arriving concurrently at the LEO, they arrive with distinct bending angles.

Figure 1-10 shows a schematic diagram for a hypothetical multipath scenario in which up to 5 rays concurrently arrive at the LEO. The ordinate is the excess Doppler shift for individual rays, and the abscissa is the epoch of the observation at the LEO. For this scenario, the phase and amplitude contributions from possibly 1, 3, or 5 rays will simultaneously be registered in the total amplitude and phase measured at the LEO. The actual number of contributing rays depends on the observation epoch t_k, $k = 1,2,\cdots$, as shown in Fig. 1-10(a). The measured composite phase will be an unknown combination of the phases of the individual rays weighted by their respective amplitudes per the vector diagram in Fig. 1-9.

1.4.1 Spectral/Holographic Techniques

What this multipath problem, described in Fig. 1-10, needs is a transformation that converts the multi-valued time series of observations (in Doppler or bending angle) into a single-valued series. There are several ways of accomplishing this. One is to convert the multi-valued time series into a single-valued spectral series, as suggested in Fig. 1-10(b). Here ω is a spectral variable, for example, a Fourier variable from a fast Fourier transform (FFT)

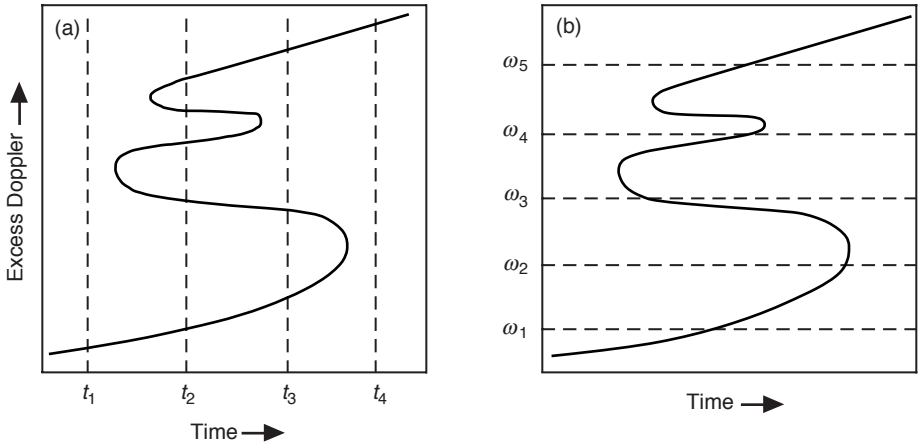

Fig. 1-10. Hypothetical multipath scenario for a setting occultation: (a) the phase at the LEO is measured at the epochs t_k, $k = 1,2,\cdots$, resulting in multiple Doppler contributions at some epochs, and (b) the time series of observations is transformed into a single-valued spectral series.

that can be made proportional to or at least related in a one-to-one way to the excess Doppler. In Fig. 1-10(b), we have converted through a Fourier transform the time series of phase and amplitude measurements of the field, in which the bending angle may not be a unique function of time, into a spectral series in which the bending angle is a unique function of the spectral variable. The peak power in the Fourier spectrum locates the excess Doppler from an individual ray and, therefore, the bending angle per Eq. (1.2-3). This approach was applied to the open-loop phase and amplitude measurements observed from the Earth from the occultation of Voyager 2 by Uranus in 1986 [10]. Figure 1-11 shows an open-loop power spectrum over time, composed of contiguous strips of 10-s temporal width and an excess Doppler breadth of about 15 Hz.

The word "radio-holography" is often used to describe this class of "wave/optic" analysis techniques. Radio-holographic techniques first "stop" the phase rate of the LEO observations by subtracting a time-dependent phase predicted from a realistic model that includes both geometric delay from satellite kinematics and the refractive delay [62–69]. This narrows the effective bandwidth around the spectral peak of the reference ray used in the reference model. The variability of the Doppler shift in the model typically ranges from several hertz to at most a few tens of hertz. Then a complex spectral algorithm,

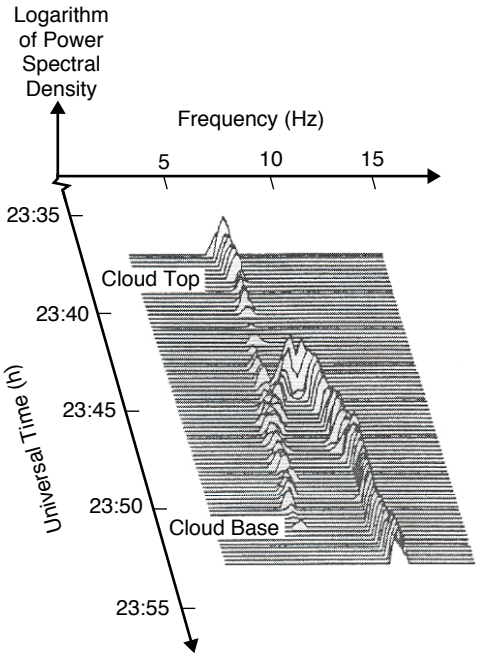

Fig. 1-11. The power spectrum of the radio signal from Voyager 2 while occulting behind Uranus' lower troposphere in 1986. Redrawn from [10].

for example, an FFT, is applied to obtain the power spectrum over the relatively narrow remaining bandwidth of the stopped phase profile within a prescribed time window or time width. Maximum power yields the Doppler tone, or tones when atmospheric multipath is present. From Eq. (1.2-3), one then obtains the bending angle for each tone. The temporal evolution of the tones is obtained by sliding the sample time window used in the FFT. Under the assumption of spherical symmetry, the one-to-one relationship between the recovered excess Doppler/bending angle and the impact parameter of the ray obtained from Bouguer's law yields the value of the impact parameter. From this approach, one recovers a unique bending-angle profile versus impact parameter for a given ray member, but corrupted as always by measurement noise and modeling errors. From this point, the Abel transform yields the refractivity profile versus impact parameter, and then from the relationship $a = r_* n(r_*)$ at a tangency point, one obtains $n(r)$.

Figure 1-12 shows two snapshots of the bending-angle spectra recovered from a GPS/MET occultation over the Sea of Okhotsk in the Russian Far East using a holographic technique [29,54,67,68]. Figure 1-12(a) shows a narrow (half-power width ~20 μrad) single tone in the mesosphere at 56-km altitude. This translates into a vertical resolution of about 60 m. Figure 1-12(b) shows the complex tone structure about 1/2 minute later for the same occultation. Here the tangency point of the direct ray lies deep in the lower troposphere between 1 and 2 km above sea level. This figure also includes a weaker tone corresponding to a near-specular reflection from the ocean.

1.4.2 Back Propagation

A second, fundamentally different approach to achieving an equivalent single-valued time series and to improving resolution is to map the field measurements recorded by the LEO onto another surface much closer to the refracting medium. This technique is based on the Helmholtz–Kirchoff integral theorem from classical electrodynamics. This integral theorem expresses the amplitude and phase of an electromagnetic wave at a given point in terms of an integral involving the distribution of the amplitude and phase of the wave over an enclosing surface. It explicitly accounts for the retardation time between any point on the radiating surface and the interior point, that is, the travel time between these points resulting from the finite speed of light. The theorem is valid when the scale of the radiating surface is very much larger than the wavelength of the wave. Both the Rayleigh–Sommerfeld and Fresnel–Kirchoff scalar diffraction theories follow directly from this asymptotic theorem [48,70]. These scalar diffraction theories enable both the forward-propagation techniques using a phase screen model to mimic the observations [71–74] and the backward-propagation techniques to map the observations to a back plane [75–77].

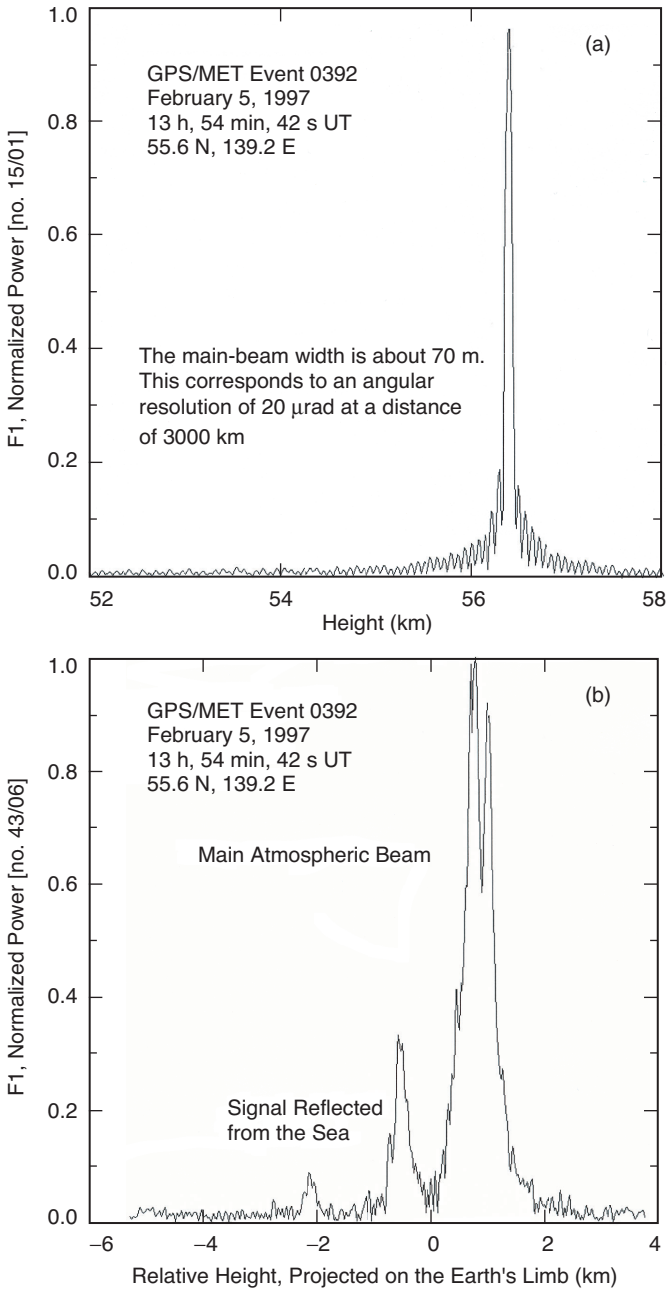

Fig. 1-12. Bending-angle spectra from a GPS/MET occultation using the radio holographic technique [67]: (a) narrow spectral distribution from a single ray in the mesosphere and (b) complex ray structure low in the troposphere, including an ocean surface reflection.

The integral theorem also can be used to propagate a wave through a succession of surfaces. For example, one form of the multiple phase screen technique involves placing successive parallel planes spaced along the paraxial direction of a wave through an inhomogeneous medium. The planes are mounted perpendicular to the paraxial direction, essentially the preferred direction of propagation. The medium between the screens is taken as homogeneous. Therefore, to compensate for the actual inhomogeneous medium between any two adjacent screens, a phase offset is added to the wave in the subsequent screen. This technique can be more tractable for propagating a wave than solving Maxwell's equations. In two-dimensional problems, the multiple phase screen technique involves a one-dimensional scalar diffraction integral applied successively, whereas Maxwell's equations form a second-order system that in general requires simultaneous integration over two dimensions [78–82].

In the multipath problem, the mapped or back-propagated field (toward the emitter) can provide an equivalent virtual time series of phase and amplitude values in another surface that has the favorable property of no or at least substantially fewer multi-valued virtual Doppler points. The one-dimensional back-propagation diffraction integral in a vacuum (see Eq. (A-22) in Appendix A) is given by

$$E(\boldsymbol{r}_1) = \sqrt{\frac{i}{\lambda}} \int_c \left(\frac{E(\boldsymbol{r}_2)}{\sqrt{r_{12}}} \exp(-ikr_{12}) \left(\hat{\boldsymbol{r}}_{12} \cdot \hat{\boldsymbol{n}}(\boldsymbol{r}_2) \right) \right) ds_2 \qquad (1.4\text{-}2)$$

where $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$; $\boldsymbol{r}_2$ denotes a point at the LEO; ds_2 denotes an incremental arc length along the path C, defined by the trajectory of the LEO over which observations were made during the occultation; and $\hat{\boldsymbol{n}}(\boldsymbol{r}_2)$ is the outward unit normal vector to C. The intervening medium is a vacuum. We apply this path integral using the LEO observations of phase and amplitude to obtain the mapped field at the position $\boldsymbol{r}_1$. Because maintaining phase coherence in this diffraction integral is so important, the three-dimensional relative motion of the transmitting GPS satellite and the LEO over the integration span of the occultation observations must be accurately modeled. Also, phase connection in the mapped phase along the back surface must be maintained. In addition, the stationary-phase technique can be used to set practical integration limits in the diffraction integral for the LEO phase and amplitude measurements as a function of the position $\boldsymbol{r}_1$. These and other details are found in [76,77,83].

Figure 1-13 provides a one-dimensional schematic for the concept. Here the LEO travels vertically downward in the LEO plane and vertical distance in this plane is proportional to elapsed time, 2 to 3 km/s. Multiple rays from different altitudes in the atmosphere arrive concurrently at the LEO plane. We use the scalar diffraction integral in Eq. (1.4-2) to propagate the field backward from

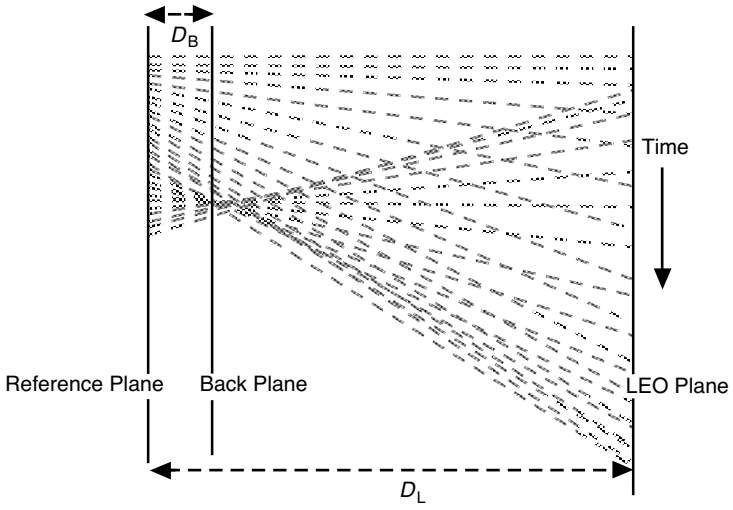

Fig. 1-13. Back propagation geometry. The measured field in the LEO plane is back propagated to the back plane to reduce multipath.

the LEO plane toward the emitter to the position of the back plane mounted perpendicular to the LEO/GPS line. Figure 1-13 suggests that a significant reduction in multiple rays can be achieved from transforming the observations to an equivalent "more focused" set in a back plane [76,77,83], particularly when the actual observation distance is large compared to the vertical scale in the atmosphere over which multipath occurs. The limb distance of a LEO is about 3000 km, but the altitude range in the troposphere from which most multiple rays arrive at the LEO is 10 km or less. Reducing this at least 300 to 1 ratio to 30 to 1 or even 10 to 1 can improve the multipath problem and reduce the Fresnel zone.

Here is a heuristic way of looking at the back-propagation concept, which is further discussed in Chapter 2. To simplify the math, let us place the emitting GPS satellite at infinity, as indicated in Fig. 1-4, and we assume a circular orbit for the LEO. Consider a point (r,θ) well out of the atmosphere ($n \equiv 1$) through which a ray passes after traversing the atmosphere. Suppose we displace that point by a small angular increment to $(r,\theta+\Delta\theta)$, holding the radius r fixed. A new ray arrives at that new point with a new impact parameter value, $a+\Delta a$. Then the difference between the excess phase on the new ray and the excess phase on the old one through the point (r,θ) is given by $\Delta\varphi = (\partial\varphi/\partial\theta)\Delta\theta$. From Eq. (1.2-3), we have

$$\frac{\partial\varphi}{\partial\theta} = k\left(\boldsymbol{T}\cdot\frac{\partial\boldsymbol{r}}{\partial\theta} - \frac{\partial r_{GL}}{\partial\theta}\right) = -k(a-b) = -k\left(D\alpha + a\frac{\alpha^2}{2} + \cdots\right) \qquad (1.4\text{-}3)$$

where $b = |\boldsymbol{r} \times \boldsymbol{r}_{GL}| / r_{GL} \to r\sin\theta$, the impact parameter of the straight line between points G (now placed at infinity along the direction $\theta_G = \pi$ in Fig. 1-6) and L. Also, from Eq. (1.2-4) we have Bouguer's law for the impact parameter of the ray, $a = |\boldsymbol{T} \times \boldsymbol{r}| = r\sin(\theta + \alpha)$. For collimated incident rays, the angle γ between the ray and the radius vector is simply $\gamma = \theta + \alpha$. These quantities are indicated in Fig. 1-4. The distance D is given by

$$D = r\cos(\theta + \alpha) \tag{1.4-4}$$

D is essentially the limb distance minus the small increment $a\alpha$. It is the distance between the point (r, θ) and the point (a, θ_a) with $\theta_a = \pi/2 - \alpha(a)$, as shown in Fig. 1-4. The latter point is on the impact parameter space curve associated with the rays after their encounter with the atmosphere. The excess Doppler shown in Fig. 1-10 is proportional to $\partial\varphi/\partial\theta$ in Eq. (1.4-3). Multiplying Eq. (1.4-3) by $\lambda\dot{\theta}_L / 2\pi$ yields the relationship between excess Doppler (assuming that $\dot{r}_L = 0$) and the bending angle

$$\lambda f_D = -D\dot{\theta}_L\alpha + O[\alpha^2] \tag{1.4-5}$$

where $\dot{\theta}_L$ is the projection of the LEO orbital rate in the plane of propagation, typically in the range 0.7 to 1.0 mrad/s.

We now construct $\partial^2\varphi/\partial\theta^2$. When θ is varied and r is held fixed, $\Delta a = r\cos(\theta + \alpha)(\Delta\theta + \alpha'\Delta a)$, where $\alpha' = d\alpha/da$. Therefore,[1]

$$\left.\begin{aligned} \frac{\partial a}{\partial\theta} &= \frac{r\cos(\theta+\alpha)}{1 - \alpha' r\cos(\theta+\alpha)} \\ \frac{\partial b}{\partial\theta} &= r\cos\theta \end{aligned}\right\} \tag{1.4-6}$$

It follows from Eqs. (1.4-3) and (1.4-6) that

$$\frac{\partial^2\varphi}{\partial\theta^2} = -k\left(\frac{D^2\alpha'}{1 - D\alpha'} + a\alpha + O[\alpha^2]\right), \quad D = r\cos(\theta + \alpha) \tag{1.4-7}$$

---

[1] Incidentally, Eq. (1.4-6) provides an expression for estimating the duration of an occultation in the neutral atmosphere. Noting that $\Delta\theta = \int (\partial a/\partial\theta)^{-1} da$, we have $\Delta T = \dot{\theta}_L^{-1}\int_{a_1}^{a_2} (\partial a/\partial\theta)^{-1} da \doteq \left(a_2 - a_1 + D(\alpha_1 - \alpha_2)\right)\left(\dot{\theta}_L D\right)^{-1} \approx 50 - 100$ s. The duration strongly depends on the magnitudes of $\dot{\theta}_L$ and α near sea level.

The first point to notice is that, on a highly defocused ray with $|D\alpha'|>>1$, it follows that $\partial^2\varphi/\partial\theta^2 \to kr\cos\theta$. Therefore, the LEO excess Doppler rate approaches essentially a constant in strong defocusing, that is, $df_D/dt \doteq (\partial^2\varphi/\partial\theta^2)\dot{\theta}_{\mathrm{L}}^2/2\pi \to \dot{\theta}_{\mathrm{L}}^2 r\cos\theta_{\mathrm{L}}/\lambda$, or 10 to 15 Hz/s, depending primarily on the value of $\dot{\theta}_{\mathrm{L}}$. This is an important coherence issue for holographic techniques applied to the lower troposphere. Holographic techniques subtract a model phase from the observations to "stop" their phase variability, i.e., to greatly narrow their spectral bandwidth, making an FFT or a similar complex spectral analysis practicable. The near-constancy of the excess phase acceleration in highly defocused areas is an important factor in the accuracy of the modeled phase profile.

The second point about Eq. (1.4-7) concerns choosing a value for the back-plane distance D_{B}. Referring to the excess Doppler profile shown in Fig. 1-10 evaluated along the LEO orbit, the undulations in the phase rate curve in the back plane will be greatly diminished if we choose $D_{\mathrm{B}} << D$. Basically we want $\partial\varphi/\partial\theta$ to be monotonic, no turning points with respect to θ, or with respect to time, or with respect to any other displacement metric, for example, arc length along the path defined by the intersection of the back plane with the plane of propagation. It follows from Eq. (1.4-7) that this is equivalent to requiring that the defocusing on the back plane be such that $\zeta_{\mathrm{B}}^{-1} = 1 - D_{\mathrm{B}}(d\alpha/da) > 0$ for all values of a spanned by the occultation episode. This condition on the value of back-plane distance D_{B}, i.e., that $D_{\mathrm{B}}(d\alpha/da) < 1$ for any point along the back plane, ensures monotonicity in excess phase rate versus displacement along the back plane if we can place the back plane close enough to the impact parameter space curve defined by (a,θ_a). This may not be achievable with a fixed back plane.

Why not set $D_{\mathrm{B}} = 0$? No multipath and no defocusing there. It turns out that that is exactly the right choice, but the problem is that one knows a priori the value of neither α nor a. From Eq. (1.4-4), it follows that setting $D_{\mathrm{B}} = 0$ is equivalent to placing the back plane at the point (a,θ_a) with $\theta_a = \pi/2 - \alpha(a)$. Setting $D_{\mathrm{B}} = 0$ is impossible to uniformly accomplish for all values of a with a fixed back plane because the impact parameter space curve as a function of a is in general not straight, and it is markedly non-linear in multipath zones.

So, with the back-plane methodology, some compromise in choosing D_{B} must be made. The difference in phase between the point (a,θ_a) on the impact parameter space curve and a nearby point $(r_{\mathrm{B}},\theta_{\mathrm{B}})$ on the back plane is given to first order by

$$\Delta\varphi = -a\Delta\theta + 0\Delta r \tag{1.4-8}$$

where $\Delta\theta = \theta_B - \theta_a$ and $\Delta r = r_B - a$. The first-order phase change in the radial direction is zero because the ray through the point (a,θ_a) is orthogonal to $\boldsymbol{r}_a$ at that point. We select $r = r_B$ and $\theta = \theta_B = \pi/2 - \alpha_B$ as the fixed point to position the back plane perpendicular to the LEO/GPS line, with α_B given a fixed value. Then the phase difference $\Delta\varphi$ for a given value of a between the point (a,θ_a) on the impact parameter space curve and a nearby point on the back plane is given to first order by $\Delta\varphi = -a(\theta_B - \theta_a) + 0(r_B - r_a) = -a(\alpha(a) - \alpha_B)$. We now have a new caustic possibility in the back plane, $d\varphi/da = \alpha_B - \alpha - ad\alpha/da = 0$. This arises on the back plane whenever it is away from the impact parameter space curve $r = a,\ \theta_a = \pi/2 - \alpha(a)$, generated by varying a. Since most caustics occur where refractivity gradients are large, and therefore where α is large, a good rule of thumb to reduce the probability of caustics occurring on the back plane has been to set $\alpha_B = \alpha_{Max}$, where α_{Max} is the largest typical bending angle encountered in the lower troposphere, where major multipath is likely to occur, 30 to 40 mrad. This choice places the back plane close to the impact parameter space curve in that vicinity [65,83]. Even though this plane is in the Earth's atmosphere, it is treated as though it were in a vacuum. It serves only as a platform for recovery of the refractivity profile.

We designate the mapped phase using the diffraction integral in Eq. (1.4-2) from the LEO to the back plane as $\varphi_B(s_B)$, where s_B is path length along the back plane perpendicular to the LEO/GPS line. From Eq. (1.2-3) or Eq. (1.4-5), it follows that, when $d\varphi_B/da$ is monotonic in the back plane, then the bending angle of a ray intersecting the back plane can be inferred unambiguously from the directional derivative of $\varphi_B(s_B)$. For collimated incident waves, it follows that

$$\frac{d\varphi_B}{ds_B} = -k\sin\alpha \tag{1.4-9}$$

By assuming spherical symmetry, one obtains the impact parameter for each bending angle using Bouguer's law. From the profile of bending-angle and impact parameter pairs, one can use the standard geometric optics Abel transform technique to recover the refractivity profile. In severe multipath situations, the caustic condition on the back plane can be violated, but experience with actual observations has established that the technique significantly mitigates multipath problems [83].

The back-plane technique gains two benefits, reduced multipath and also enhanced vertical resolution. The Fresnel zone is smaller at the back plane by roughly the factor $\sqrt{|D_B\zeta_B / D_L\zeta_L|}$.

1.4.3 The Canonical Transform Technique

We have noted that rays are stationary-phase paths. Therefore, their space coordinates can be transformed with Hamilton–Jacobi theory into canonical coordinates or ray coordinates. The canonical transform technique [84] essentially uses Hamilton–Jacobi theory to transform the two-dimensional space coordinates (z, x) and their slopes dx/dz for a ray system, that is, for a family of rays generated by varying, for example, the impact parameter a, into three-dimensional phase–space coordinates (z, x, a). Here x is vertical directed in Fig. 1-4 and z is horizontal directed. In these transformed coordinates, a becomes the canonical conjugate momentum to the variable x. One then applies a Fourier integral transform similar to the scalar diffraction integral in Eq. (1.4-2) to map the field from the phase and amplitude measurements along the LEO path in space coordinates onto the a-plane in phase–space coordinates, from which the phase $\varphi(a)$ is obtained. On this a-plane, $k\alpha = -d\varphi(a)/da$. The technique yields an unambiguous determination of α and a, provided that spherical symmetry holds and super-refractive zones are avoided, and subject, of course, to the quality of the LEO observations and the modeling. The technique fails in super-refractivity zones because it is in part ray theory. The technique avoids the potential back-plane caustic and multipath problems of the general back-propagation technique, and its vertical resolution potential is comparable to wave theory.

1.4.4 The Impact Parameter Space Curve

Another less elegant but conceptually straightforward method for avoiding caustic and multipath problems maps the phase and amplitude measurements made by the LEO along its path to the impact parameter space curve associated with the rays after their encounter with the atmosphere. As Fig. 1-4 indicates, the rays are ordered through a transverse plane located perpendicular to the incoming rays at a position prior to their entering the atmosphere. This pre-encounter impact parameter space curve is a simple straight line perpendicular to the incoming collimated rays, or a circular arc of radius $\rho_G/2$ if the incoming rays are from a spherical wave. This is equivalent to the proposition that the rays are single-valued as a function of impact parameter a before encountering the atmosphere; a one-to-one relationship holds between the value of the impact parameter of the ray intersecting the transverse plane (or arc) and the position of that intersection point on the plane (or arc). If spherical symmetry holds, the rays will continue to be single-valued after passing through the atmosphere, that is, each ray will continue to have a unique value of the impact parameter. Therefore, after transecting a spherical symmetric atmosphere, there will remain a one-to-one relationship between the impact parameter of a ray and the angle through which it was refracted. The point

$r_a = a,\ \theta_a = \pi/2 - \alpha(a)$, located on an outgoing ray, also preserves this one-to-one relationship with that ray, if spherical symmetry holds. Bouguer's law in Eq. (1.2-4) requires it. The curve traced out by the point $(r_a,\ \theta_a)$ as a is varied is the space curve for the impact parameter *after* atmospheric encounter. As mentioned earlier, there will be no caustics on this post-encounter curve, but it is no longer a simple curve because of the dispersive refractive gradient.

It is important to keep in mind that we are not so much interested in ray reconstruction or mapping the electromagnetic field through space as we are in recovering the refractivity profile. Any construction, however non-physical, that facilitates refractivity recovery becomes a valid candidate for this application. In this regard, a fictitious curve placed after a ray's atmospheric encounter that denumerates the impact parameter uniquely along its path should suffice for spherical symmetric geometry because the rays through points in this curve will be unique [77]. This holds even though the impact parameter space curve actually lies within the atmosphere.

On the post-encounter impact parameter space curve, the phase of the intersecting ray with an impact value a can be obtained from the defining integral for phase delay given in Eq. (1.2-2) using Bouguer's law. Let this phase be defined by $\varphi_a(a,\theta_a)$ [with $\theta_a = \pi/2 - \alpha(a)$]. It can be shown from Eq. (1.2-2) (with the GPS transmitting satellite at infinity and the phase referred to the line $\theta = \pi/2$) that

$$\varphi_a(a,\theta_a) = k\left(a\alpha(a) + \int_a^\infty \alpha(\omega)d\omega\right) \tag{1.4-10}$$

The space curve itself is generated by the tip of the vector $\boldsymbol{r}_a(a)$, with $r_a = a$ and $\boldsymbol{r}_a \cdot \boldsymbol{T}_a = 0$, where $\boldsymbol{T}_a$ is the unit tangent vector of the ray passing through the tip of $\boldsymbol{r}_a$. The angle ψ_a between the tangent vector $d\boldsymbol{s}_a$ to the impact parameter space curve and $\boldsymbol{r}_a$ is given by

$$\left.\begin{aligned} \cos\psi_a &= \frac{\boldsymbol{T}_a \cdot d\boldsymbol{s}_a}{ds_a} = \frac{1}{\sqrt{1+(a\alpha')^2}} \\ \sin\psi_a &= \frac{-a\alpha'}{\sqrt{1+(a\alpha')^2}} \end{aligned}\right\} \tag{1.4-11}$$

with

$$\frac{ds_a}{da} = \sqrt{1+(a\alpha')^2} \tag{1.4-12}$$

From Eq. (1.4-10), it follows that

$$ka\frac{d\alpha}{da} = \frac{d\varphi_a}{da} \tag{1.4-13a}$$

or

$$\alpha(a) = -\frac{1}{k}\int_a^\infty \frac{1}{a}\frac{d\varphi_a}{da}da \tag{1.4-13b}$$

By replacing φ_a in this integral with the mapped phase $\varphi_a(\boldsymbol{r}_a)$ obtained from the back-propagating diffraction integral of the LEO observations given in Eq. (1.4-2), we recover the profile for $\alpha(a)$. Note that, from Eq. (1.4-10), $\varphi_a \to 0$ as $a \to \infty$ if $\alpha \to 0$.

Chapter 2 discusses one method for recovering the impact parameter space curve $(a, \theta_a(a))$ starting from an initial known point $(a_1, \theta_a(a_1))$ on the curve. It uses the back-propagation diffraction integral to map by iteration the LEO observations to a converged point $(a_1 + \Delta a, \theta_a(a_1) + \Delta\theta_a)$ near the initial point. Each successive converged point on that curve yields the recovered values for $\varphi_a(a_k)$ and $\alpha(a_k)$, $k = 1, 2, \cdots$. For a spherical symmetric atmosphere, the recovered impact parameter space curve can be used with the Abel transform to recover $n(a)$. This method also fails in super-refractive layers. Also, Eq. (1.4-13) shows that the actual shape of the (a, θ_a) curve, however complicated it might be, is of no theoretical consequence in recovering $\alpha(a)$, although a complicated shape could slow the iteration process.

The difference between the phase $\varphi(a)$ on the a-plane in phase space, which uses the canonical transform technique, and $\varphi_a(a)$ on the post-encounter curve in space coordinates (a, θ_a) is simply

$$\varphi(a) = \varphi_a(a) - ka\alpha(a) = k\int_a^\infty \alpha da \tag{1.4-14}$$

1.4.5 A Full-Spectrum Wave Theory

The various spectral or holographic approaches and the scalar diffraction approaches, both the forward- and backward-propagation techniques, involve a hybrid of wave and ray theory concepts. In the end, profiles of bending angle versus impact parameter are obtained from which the refractivity is recovered using the Abel integral transform.

In the quest to deal with multipath problems and related resolution issues, there is yet another approach that seems not to have been fully considered in radio occultations. This involves a full-spectrum wave-theoretic treatment of the electromagnetic propagation through a refracting medium. The technique presented here uses a modified Mie scattering approach adapted to an

inhomogeneous medium. It is applicable to the more general case of a scattering surface embedded in a transparent, spherical refracting medium. An interesting feature of this approach is its full-spectral series representation of the wave.

The spectral series is a solution to the Helmholtz equation for the time-independent part of the wave. Each spectral component of integer number l consists of a variable spectral coefficient multiplied by the basis functions, which in their native form are the spherical Hankel functions of integer order l and the spherical harmonic functions. The effect of the inhomogeneity in the propagation medium is contained in the spectral coefficients, leaving the basis functions unchanged from their functional form in a homogeneous medium.

To derive the variation of the spectral coefficients across a spherical surface, one invokes the continuity conditions from electrodynamics that apply to the radial and tangential components of the incident, reflected, and transmitted electromagnetic field vectors, just as is done in Mie scattering theory. A limiting process yields the differential equations describing the variation of the spectral coefficients in a stratified medium. The accuracy of the technique deteriorates when the truncation assumption about the smallness of certain curvature terms in the field equations becomes invalid, for example, near a turning point. The related Wentzel–Kramers–Brillouin (WKB) method has similar problems at a turning point. The technique is not applicable to a medium with significant back scattering, but it is adaptable to an absorbing medium.

In this wave theory, each spectral coefficient contains a spectral density function for the phase delay induced by the refractivity gradient of the atmosphere. Basically the spectral density function gives the cumulative phase delay induced by the atmospheric refractivity gradient on the lth spectral component of a wave that has traveled from outside the atmosphere down to a given radial distance. This spectral density for phase delay, which is a function of the radial coordinate and the spectral number, can be formally written in terms of an integral over the radial coordinate that explicitly contains the refractivity gradient. Without a refractivity gradient, there is no phase delay in the spectral coefficient.

This series representation of the wave can be directly linked to the Fourier transform of the time series of LEO-observed amplitude and phase. The spectral density function can be recovered by taking the Fourier transform of the stopped version of the formal spectral series from wave theory and the Fourier transform of the stopped phase and amplitude observations made by the LEO over time. Equating these two transforms enables recovery of the spectral density function for the phase delay. The refractivity gradient follows from an inversion of the defining integral for the derivative of the spectral density function.

In a wave theory, the fundamental observations are the phase and amplitude measurements of the field, not Doppler observations or the resolving of possible ambiguities contained therein from multipath. There is no need to propagate the field or map it from one surface to another. The spectral series gives the field at an arbitrary point, even inside of the refracting medium, provided turning points are avoided and the curvature truncation assumption is valid. Multipath and caustic rays become almost irrelevant in wave theories. This applies to parabolic equation techniques with multiple phase screens and to the full-spectrum technique described here. A wave theory calculates the phase and amplitude of a caustic ray as accurately as any other ray. The predicted amplitude of the field at the LEO when it is located on a caustic ray is proportional to $|d^2a/d\theta^2|^{1/3}$ in wave theory, unlike the infinite value predicted in (second-order) geometric optics.

In a full-spectrum wave theory, there will be stationary-phase points in spectral number from where the principal contributions to the summation of the spectral series originate. These stationary-phase points in spectral number closely correspond to the impact parameter values (in phase units) of the multiple rays, including caustic contact points. But in the full-spectrum technique, all points in spectral number are swept up together into a spectral density function that holds for all relevant spectral numbers, not just stationary-phase points.

As mentioned above, we are more interested in the refractivity recovery than in calculating the electromagnetic field vector at the LEO. In this regard, the main emphasis in this full-spectrum technique is on computing the spectral density function for the phase delay, a somewhat simpler task than computing the electromagnetic wave from the spectral series itself. Moreover, being fundamentally a result from wave theory, the recovered refractivity profile has the potential for a much finer radial resolution than dictated by the local first Fresnel zone. This also applies to certain holographic and scalar diffraction techniques mentioned above. Theoretical vertical resolution for wave techniques is proportional to the carrier wavelength divided by the angular synthetic aperture of the observations $\dot{\theta}\Delta T$, or roughly $150(\Delta T)^{-1}$ m for a LEO. Here ΔT is in seconds; it is the minimum of either the time span of the selected observation sequence or the coherence limit of the observations. This would suggest a dekameter-level resolution potential, but the actual vertical resolution will be limited by atmospheric vicissitude, such as horizontal variations in refractivity caused by advection and turbulence, and also by other GPS signal structure and data processing limitations [72]. Perhaps 10 percent of the vertical axis of the local first Fresnel zone is a practicable goal, although the evidence for this conjecture is still inchoate at this time.

Full-spectrum wave-theoretic approaches are powerful for a certain class of geometries, but they are notoriously cumbersome. That is why so many

alternate schemes for propagating waves through complex media have evolved over the years since Mie scattering theory was first formulated in the early 20th century. The use of the Helmholtz–Kirkchoff integral theorem for multiple phase screen propagation is one example. Other parabolic wave equation techniques for propagating waves through a medium can be found in [81]. Spherical symmetry or some similar symmetry is almost de rigueur by definition in a full-spectrum wave-theoretic approach. When the wavelength is small compared to the scale of the refracting medium, the spectral series are slow to converge. The basis functions in the spectral representation of the wave, the spherical Bessel functions, and the spherical harmonic functions become difficult to evaluate in their native forms for large values of spectral number.

Nevertheless, accurate asymptotic forms for the Bessel functions in terms of the Airy functions exist that are quite comprehensive and easy to use [85,86]. Most computer mathematics programs carry a full library of the Airy functions of the first and second kind and their derivatives. Also, the asymptotic forms for the spherical harmonic functions for large spectral number are given by sinusoid functions. While asymptotic functions are used throughout the series, the full-spectral character of the wave equation solution is retained. On the issue of convergence, the stationary-phase technique can be used to identify neighborhoods in spectral number that contribute to the summation of the series representation of the wave, thereby eliminating enormous tracts of non-contributing spectral numbers. Moreover, the stationary-phase technique can be used to establish a duality between certain wave-theoretic quantities evaluated at the stationary-phase points in spectral number and the corresponding quantities in geometric optics. For example, when a stationary-phase point exists, and when super-refractivity situations are avoided, then there is a close correspondence between the spectral density function for phase delay evaluated at a stationary-phase point in spectral number and the eikonal function in geometric optics for the corresponding ray. This is the ray with an impact parameter (in phase units) equal to the stationary-phase point in spectral number.

The essential difference between a full-spectrum wave theory and most holographic techniques [62–69] is in the treatment of the wave itself. Loosely speaking, a full-spectrum wave theory involves a pre-stationary-phase process; holographic techniques involve a post-stationary-phase process. In the full-spectrum technique, the time-independent part of a harmonic wave is described by a spectral distribution within an infinite series summation over spectral number, here an equivalent spectral integral. Holographic treatments describe the wave as a finite sum of the complex amplitudes from the multiple rays involved at an observation epoch, as the schematic in Fig. 1-10 suggests. In the recovery process, both techniques apply a spectral analysis to the phase and amplitude measurements, for example, a fast Fourier transform. The full-spectrum technique recovers the spectral distribution for the phase delay

induced by the refractivity gradient, and it does so with a limit in spectral resolution that is imposed by the uncertainty principle in the Fourier analysis, $\sim 1/\Delta T$, where ΔT is the time interval used in the observations. From the spectral derivative of the recovered spectral distribution for the phase delay, the gradient of the refractivity profile is recovered from inverting the formal integral equation that expresses this spectral derivative in terms of the refractivity gradient. From this one obtains the profile of the change in refractivity over the span of impact parameter values set by the time interval ΔT. One could place, if needed, the recovered spectral density for phase delay into the spectral integral for the wave to compute the stationary-phase points of this integral for successive increments in observation epochs. This gives profiles of the impact parameter values of the multipath rays, including caustic rays if they occur there, the bending angles of the rays, and their angular power spectra.

In comparison, a holographic technique obtains the maximum-likelihood estimate for the bending angle and impact parameter associated with each of the finite number of rays detected in the Fourier analysis of the observations within a given time increment, as well as their angular power spectra. And it then applies the Abel integral transform from ray theory to the recovered bending-angle and impact parameter sequences for each ray to obtain the profile of the change in refractivity over the integration limits. The integration limits in impact parameter for each recovered ray are set to obtain the refractivity profile over the impact parameter range applicable to that ray. Appropriate hand-over at each caustic contact point of the integration limit in the Abel integral from the one nascent ray to the other provides the connections between integral segments, and it yields the entire refractivity profile.

If for certain positions of the LEO there are no stationary-phase values in spectral number in the full-spectrum technique, then in ray theory there will be no rays or eikonal functions to give phase delays. But in a wave theory, there still will be a field predicted for such locations, diminished in amplitude probably, but not zero.

The overviews of Chapters 5 and 6 in Section 1.5 give additional details on this wave-theory approach.

1.5 Overview

1.5.1 Chapter 2

Chapter 2 is largely a revision of an earlier Jet Propulsion Laboratory (JPL) publication [87]. Chapter 2 discusses the changes in the phase and signal amplitude profiles that are observed by a LEO during an occultation of a GPS satellite as a result of a sharp change across a spherical surface in a refraction-related property. This would include (1) the refractivity itself, (2) its scale

height, and (3) the lapse rate of the temperature profile. These sharp changes are embedded in an otherwise smoothly varying refractivity profile. Across the tropopause the lapse rate can change abruptly. Strong refractivity gradients often are associated with a water vapor boundary in the lower troposphere or an electron density change in a sporadic E-layer of the ionosphere. A discontinuous refractivity profile could be a useful limiting case.

Chapter 2 uses a combination of geometric optics and scalar diffraction theory in a thin-screen model to interpret these transients in terms of multipath interference, shadow zones, diffraction, and caustics that are likely to be observed by the LEO. In other words, we primarily address the forward-propagation effects of certain precipitous changes in the vertical profile of the atmospheric or ionospheric refractivity, such as a discontinuous change in refractivity or in one of its derivatives. The applicability of the thin-screen approach in a medium with strong refractive gradients is discussed. The stationary-phase technique for interpretation and for aiding the computation of diffraction integrals is discussed.

This hybrid thin-screen/scalar diffraction approach is used to calculate the phase and amplitude perturbations that would be observed by the LEO from these perturbations in refractivity. Although one customarily thinks of diffraction as producing high-frequency fringe effects, which is true, it also can soften or mitigate the harsh effects predicted by geometric optics, partially filling in the troughs in shadow zones and rounding the peaks in flaring regions where caustics are a factor.

Use of an inverse transform technique to sharpen the resolution of localized features is briefly covered, and it is applied to a specific occultation where the ray tangency point crosses a sporadic E-layer.

1.5.2 Chapter 3

Chapter 3 also is a revision of an earlier JPL publication [88]. To prepare the development of a wave theory in a refracting medium, we review basic Mie scattering theory. Chapter 3 discusses the scattered electromagnetic field from a spherical scattering surface in a homogeneous medium. The field is expressed as series solutions to the Helmholtz equation involving spherical Bessel functions and spherical harmonic functions summed over integer spectral number l. In Mie theory, the scattering coefficients in the spectral series solutions are determined by applying the continuity conditions from electrodynamics that must hold across a boundary bearing a discontinuity ΔN in refractivity. The total scattering from the sphere includes all scattering modes, which arise from external reflection, refraction, and internal reflections.

It is well-known that these spectral series converge slowly with spectral number when the radius r_o of the refracting sphere is very large compared to the wavelength λ of the incident wave. Asymptotic forms for the Bessel and

Legendre polynomial basis functions in terms of the Airy functions and the complex exponential functions, which are applicable when $r_o / \lambda >> 1$, can be used to greatly facilitate the computation of the scattered field. The asymptotic scattering series is evaluated by converting it into an integration of a phasor multiplied by a slowly varying function of l. Neighborhoods around spectral number points yielding stationary-phase values are the principal contributors to the scattering integrals.

Certain attributes of the scattered field, such as phase, amplitude, reflection, bending angle, defocusing, and caustics, are evaluated at the stationary points in spectral number. A close correspondence is established between these attributes in electrodynamics and their analogs in geometric optics when the observation point is sufficiently away from the shadow boundary.

Rainbow caustics arising from one or more internal reflections within the refracting sphere are evaluated and compared to the predictions from the thin-screen model. For small discontinuities in refractivity, rainbow effects are shown to be prominent only within relatively narrow directions. Third-order stationary-phase theory is used to determine their characteristics in terms of LEO position, the value of ΔN, the ratio r_o / λ, and the number of internal reflections prior to being observed.

Special limiting cases, such as scattering by a perfectly reflecting sphere, by a strongly absorbing sphere, and by a sphere with infinite radius also are discussed in terms of stationary-phase concepts with numerical examples.

When the magnitude of the discontinuity in refractivity at the scattering surface is sufficiently small, $|\Delta N| <\sim (2\pi\lambda / r_o)^{2/3}$, about 30×10^{-6} for the Earth's radius or equivalently about 20 mrad of refractive bending at the surface, and when the observation point is sufficiently far from the limb, e.g., at the LEO, then one obtains good agreement in amplitude between the computations based on Mie scattering theory and those based on the thin-screen model combined with scalar diffraction theory.

1.5.3 Chapter 4

In Chapter 4, we review a technique that uses the unitary state transition matrix for the system of first-order electromagnetic wave equations in a refracting medium that is stratified and thin, and whose electromagnetic properties are linear [48,89,90]. This approach has been useful for calculating the propagation of an electromagnetic wave through a thin film. Although not essential to the development of the modified Mie scattering approach, several key propagation concepts are introduced here that are used in the Mie formulation. These include the representation of incoming and outgoing standing waves, the use of osculating parameters, their asymptotic forms, asymptotic matching of incoming and outgoing solutions, and evaluation of the accuracy of the osculating parameter technique.

1.5.4 Chapter 5

In wave theory, we are concerned with solutions to Maxwell's equations. For a harmonic wave in a homogeneous medium, the time-independent parts of the electric and magnetic field vectors are solutions to the Helmholtz equation, which become the components in a spectral series representation of the wave. Each component consists of a constant spectral coefficient multiplied by the basis functions used in the spectral series, which in a spherical medium are the spherical Hankel functions and spherical harmonic functions. Chapter 5 develops a spectral representation for a harmonic electromagnetic wave in a spherical refracting medium. The resulting spectral series for the field in the refracting medium also are expressed in terms of the spherical Hankel and spherical harmonic functions as basis functions. The spectral coefficients for these basis functions are derived for a wave propagating through this refracting medium. The spectral coefficients, which in a homogeneous medium are functions only of spectral number, also vary with radial distance in a refracting medium. Through a modified Mie scattering theory, these spectral coefficients also account for, if applicable, scattering from a spherical surface embedded in the refracting medium. The scattering surface could be defined by a discontinuity in refractivity, or in its gradient or a higher derivative, or by a reflecting surface, or by an absorbing surface.

Chapter 5 uses an osculating parameter technique to calculate the spectral coefficients describing wave propagation in a refracting medium. The method for obtaining the spectral coefficients is similar to a certain parabolic equation technique for wave propagation [81]. In a typical parabolic equation technique, one forms the reduced solution $u(\boldsymbol{r}) = \exp(-i\boldsymbol{k}\cdot\boldsymbol{r})\psi(k\boldsymbol{r})$ and then eliminates less relevant terms. Here $\psi(k\boldsymbol{r})$ is a solution to the Helmholtz equation and $\boldsymbol{k}$ is the wave number vector in the paraxial direction. In our formulation, the form is, for a given spectral number l, $u(r) = \exp(-iG[knr])\psi(knr)$, where $-G[knr]$ is the phase delay incurred by the lth spectral component of the wave *only* from the refractive gradient. The "optical" delay is retained in the original Helmholtz solution $\psi(knr)$. In the continuous medium overlying the scattering surface, the solution is applicable when the medium induces negligible back scattering of the wave, i.e., the medium is transparent. The technique can be adapted to an absorbing medium. It is shown for the limiting case, where the spherical stratification approaches a Cartesian stratification, that the osculating parameter solution becomes the WKB solution for a variable index of refraction. The accuracy of this technique and its range of applicability are addressed. The technique fails at or below a "turning point" (which is spectral-number-dependent in wave theory) because the truncation assumption regarding the smallness of certain curvature terms in the defining differential equations for the electromagnetic field is not valid there. The WKB method also fails for the same reason without appropriate connection formulas between

regimes above a turning point and below, the tunneling regime. On the other hand, one must match the incoming and outgoing spectral coefficients at a turning point to ensure the absence of singularities at the origin from the Hankel functions, and to obtain the correct form for the outgoing spectral coefficients. The outgoing coefficients are the only coefficients applicable at the LEO. Individual Hankel functions are singular at the origin, but the sum of the first and second kinds equals the spherical Bessel function of the first kind, which is well-behaved at the origin. This matching of incoming and outgoing coefficients is accomplished by an asymptotic technique, even though the individual forms for the spectral coefficients fail at and below the turning point.

Chapter 5 is relevant to the applicability of the single thin phase screen/scalar diffraction model, discussed in Chapter 2, to a spherical atmosphere with a significant gradient in refractivity. The thin-screen model serves as a surrogate for the actual atmosphere. Both the thin-screen/scalar diffraction approach and the wave-theory approach lead to predictions of the observed phase and amplitude of the wave having passed through an intervening atmosphere and perhaps having encountered an embedded scattering surface. The wave-theory approach explicitly accounts for scattering and for the effects of the refractive gradient in the surrounding medium on the phase and amplitude of the electromagnetic wave propagating through the medium. In comparison, the thin-screen model lumps these effects into a phase delay profile embedded in the thin screen. This profile mimics the extra phase delay experienced by the wave due to the atmospheric refractivity profile. In accounting for the scattering effects from and through the spherical surface of discontinuity, wave theory also includes the possibility that the scattered wave has undergone one or more reflections inside this surface. Interference, shadowing, diffraction, and caustics can be evaluated using both the thin phase screen/scalar diffraction approach and the full-wave theory approach. The level of agreement between these two approaches and how that level depends on the adversity of the wave propagation conditions in the atmosphere are addressed here. Good agreement with thin-screen results is obtained when "thin atmosphere" conditions hold in a large sphere, i.e., no caustic points in the screen, when rainbow caustic directions are avoided, and when the LEO is some distance from the refracting medium.

Stationary-phase concepts are applied to key observables, such as phase delay, propagation direction, and wave amplitude, the stationary-phase values of which are established in spectral number using a phasor representation in the wave-theory approach. Stationary-phase values for these observables also are obtained in impact parameter space in geometric optics. Chapter 5 establishes a duality between stationary-phase variables in spectral number, when they exist, and their counterparts in geometric optics.

Chapter 5 introduces a general spectral density function $G[\rho,\nu]$ for the phase delay in the spectral coefficient induced by the refractive gradient in the

medium. This function of radial coordinate $\rho = krn(r)$ and spectral number ν accounts for the extra phase delay at the radial position ρ on each spectral component induced by the refractive gradient on an incident wave that asymptotically is either planar or spherical at large approaching distances relative to the refracting sphere. Correspondence between this spectral density function for phase delay and key concepts from geometric optics, cumulative bending angle and the eikonal function giving path delay, is established.

Chapter 5 also describes the properties of turning points, caustics, shadow zones, and multipath from a wave-theoretic point of view in a spectral framework. It also discusses these quantities in a second-order geometric optics framework, including its shortcomings near caustics or in dealing with nascent ray pairs with nearly merged impact parameters. Third-order stationary-phase theory is introduced to develop a more accurate ray-theoretic approach near caustics.

Chapter 5 develops a phasor-based approach for evaluating the spectral series using numerical integration combined with the stationary-phase technique. The numerical integration of the spectral representation has been aided by the stationary-phase technique to identify contributing neighborhoods in spectral number, greatly improving their efficiency.

Special topics are addressed—for example, calculating the field when a reflecting surface is embedded in the refracting medium and dealing with turning point computational difficulties where the osculating parameter technique degrades in accuracy. Numerical solutions for the spectral representation of the field at the LEO are presented for various refractivity profiles and scattering surfaces. Numerical examples also include the field at the LEO from refractivity profiles giving multipath, shadow zones, super-refractivity, and caustics, and the field at the LEO from a perfectly reflecting sphere embedded in an overlying refracting medium.

1.5.5 Chapter 6

Here our goal is use the wave-theory concepts in Chapter 5 to develop a technique to recover the refractivity profile from occultation observations of phase and amplitude. The chapter begins with a discussion of GPS receiver operations and ends with a discussion of a recovery technique using the spectral representation developed in Chapter 5. We provide a brief summary here of the recovery sections.

In spherical coordinates, the coplanar components of the time-independent solutions to the Helmholtz equation for the electric field evaluated at the point (ρ, θ) in the plane of propagation may be expressed in spectral series form:

$$\left.\begin{aligned} E_r(r,\theta) &= \frac{E_o}{2\rho^2}\sum_{l=1}^{\infty} l(l+1)\left(a_l^-(\rho)\xi_l^-(\rho)+a_l^+(\rho)\xi_l^+(\rho)\right)P_l^1(\cos\theta) \\ E_\theta(r,\theta) &= \frac{E_o}{2\rho}\sum_{l=1}^{\infty}\left(a_l^-(\rho)\xi_l'^-(\rho)+a_l^+(\rho)\xi_l'^+(\rho)\right)\frac{\partial}{\partial\theta}P_l^1(\cos\theta) \end{aligned}\right\} \quad (1.5\text{-}1)$$

where $a_l^{\pm}(\rho)$ are the spectral coefficients, $\xi_l^{\pm}(\rho)$ are the spherical Hankel functions of the first (+) and second (–) kinds, $\rho = krn(r)$, and $P_l^m(\cos\theta)$ are the associated Legendre polynomials with $m=1$. In this coordinate frame, the emitting source is located in the direction $\theta=\pi$. Here E_o is an arbitrary constant when the emitting source is infinitely far, but in the finite case it also accounts for space loss from the wave-front curvature. Wave theory allows for the possibility of two polarization modes for the field vectors, the so-called transverse electric (TE) and transverse magnetic (TM) modes. Here we assume a TM wave, i.e., the magnetic field vector is oriented perpendicular to the plane of propagation. The electric field calculated from Eq. (1.5-1) includes both incoming (–) and outgoing (+) waves, and it takes into account the possibility that radiant energy has come to the point (r,θ) via any sector in the refracting sphere, for example, from the bottom limb and from the top limb. But for an occultation geometry from a distant emitting source and when the LEO is located at a point $(r_{\mathrm{L}},\theta_{\mathrm{L}})$ well into the top and outgoing sectors, $\pi/2-\alpha_{\mathrm{L}} >> \theta_{\mathrm{L}} >> 0$, only the terms with plus signs contribute to the evaluation of these series at the LEO when $r_{\mathrm{L}}/\lambda >> 1$.

These spectral series converge only after the value of the spectral number becomes comparable to ρ_o, the radial phase distance of the refracting source, about 2×10^8 for GPS wavelengths. To compute the actual electromagnetic field from these series, one uses special techniques, for example, contour integration in the complex spectral number plane, saddle point integration, and numerical integration aided by stationary-phase theory. One also needs a number of asymptotic forms for the Hankel functions and the Legendre polynomials generally valid for large values of the spectral number l and the radial coordinate ρ. For $l,\rho >> 1$, the Bessel functions, or spherical Hankel functions in this case, can be written accurately in terms of the Airy functions. Similarly, the spherical harmonic functions can be written in terms of sinusoids.

Hardly any of that computational firepower is necessary when one is interested only in recovery of the refractivity profile. The important aspect for refractivity recovery is the variation in phase of each spectral coefficient $a_l^+(\rho)$, which varies with radial coordinate because of the refractive gradient in the medium. For a refracting medium with negligible back scattering and absorption, these coefficients evaluated at the LEO can be written in the form

$$a_l^+(\rho_{\mathrm{L}}) = f(l)\exp\left(-i2G^\dagger(\nu)\right),\ \ \nu = l + \frac{1}{2} \tag{1.5-2}$$

where $f(l)$ is a function of spectral number. The specific functional form of $f(l)$ depends on the asymptotic boundary conditions for the incoming wave from the emitting GPS satellite. Chapter 5 mainly uses a collimated incident wave for the spectral coefficients, but the adjustments to account for a spherical incident wave are noted there, and they are used in Chapter 6. The quantity $-2G^\dagger(\nu)$ is the spectral density function of the phase delay for the lth spectral component evaluated at the position of the LEO for a wave that has originated from a distant source; both the LEO and the emitter are assumed to be out of the refracting medium. For a homogeneous medium, $G^\dagger(\nu) \equiv 0$. Otherwise, $G^\dagger(\nu)$ is obtained from the general spectral density function $G[\rho,\nu]$, which is given to sufficient accuracy by the integral

$$\left.\begin{aligned} G[\rho,\nu] &= \int_\rho^\infty \left(\frac{d\log n}{d\rho}\right) g(\hat{y})d\rho, \quad \rho \geq \rho^\dagger(\nu) = \nu - \hat{y}^\dagger K_\nu,\ \ g(\hat{y}^\dagger) = 0 \\ g(\hat{y}) &\doteq \pi K_\nu^2\left(\mathrm{Ai}'[\hat{y}]^2 + \mathrm{Bi}'[\hat{y}]^2 - \hat{y}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)\right), \\ \hat{y} = \hat{y}\left[\frac{\nu}{\rho}\right] &\doteq \frac{1}{4K_\nu^4}\left(\nu^2 - \rho^2\right),\ \ K_\nu = \left(\frac{\nu}{2}\right)^{1/3},\ \ \nu = l + \frac{1}{2} \end{aligned}\right\} \tag{1.5-3}$$

Here $\mathrm{Ai}[\hat{y}]$ and $\mathrm{Bi}[\hat{y}]$ are the Airy functions of the first and second kind, and $\hat{y}$ is their argument; its exact functional form is given in terms of the ratio ν/ρ [85]. The quantity $\hat{y}^\dagger = 0.441\cdots$ is the zero point of $g(\hat{y})$, which is unique. The spatial difference between ν and $\rho^\dagger$ is $k^{-1}\hat{y}^\dagger K_{\rho_o} \approx 7$ m, almost negligible. For $\rho < \rho^\dagger(\nu)$, the form for $g(\hat{y})$ in Eq. (1.5-3) does not apply; the correct form rapidly approaches zero for increasing $\hat{y} > 0$.

The interpretation of the general form $G[\rho,\nu]$ is as follows: $-G[\rho,\nu]$ is the phase delay accumulated by the spectral coefficient $a_l^-(\rho)$ for a wave that has traveled from infinity (where $dn/d\rho \to 0$) down through a refracting medium to a radial distance r with $\rho = krn(r)$. For a given spectral number ν, the rate of phase accumulation by the corresponding spectral coefficient $a_l^-(\rho)$ rapidly approaches zero for decreasing ρ values less than ν. In this regime, these spectral coefficients are unaffected by the refracting medium overhead or by a scattering surface contained therein. Therefore, $G[\rho,\nu] \to$ constant for decreasing $\rho < \nu$.

The integration variable $\rho = krn(r)$ is used for convenience. Crossing a super-refractivity boundary where $d\rho / dr$ reverses sign requires special treatment, which is discussed in Chapter 6.

$G^{\dagger}(\nu)$ applies to an outgoing wave evaluated at the LEO outside of the refracting medium. The general form for $G^{\dagger}(\nu)$ when $d\rho / dr > 0$ is given by

$$G^{\dagger}(\nu) = G[\rho^{\dagger}, \nu], \;\; \rho^{\dagger} \doteq \nu - \hat{y}^{\dagger} K_{\nu}, \;\; K_{\nu} = \left(\frac{\nu}{2}\right)^{1/3}, \;\; \hat{y}^{\dagger} = 0.441\cdots \qquad (1.5\text{-}4)$$

This connection between the outgoing and general forms for the spectral density functions is accomplished by asymptotically matching the incoming and outgoing spectral coefficients at a turning point to satisfy the general boundary condition that the spectral series shall have no singularity at the origin. It uses the property that $G[\rho, \nu] \rightarrow$ constant for decreasing $\rho < \rho^{\dagger}(\nu)$ to match the incoming and outgoing coefficients. The accuracy of this asymptotic matching technique, which involves an approximation, is discussed in Chapters 5 and 6.

For $\rho > \rho^{\dagger}(\nu)$ and provided that super-refractivity situations are avoided, it can be shown that $\partial G[\rho, \nu] / \partial \nu \doteq \tilde{\alpha}(\rho, \nu)$. Here $\tilde{\alpha}(\rho, \nu)$ is the cumulative one-way refractive bending angle evaluated at ρ for an incoming ray with an impact parameter equal to ν. Thus, $2\tilde{\alpha}(\nu, \nu) = \alpha_{\mathrm{L}}(\nu)$, a virtual bending angle of a ray with an impact parameter value of ν that would be observed at the LEO. If we can recover the profile of $2G^{\dagger}(\nu)$ versus spectral number, then it is clear from Eqs. (1.5.3) and (1.5-4), and from the close association of $2dG^{\dagger}(\nu) / d\nu$ with the total virtual bending angle $\alpha_{\mathrm{L}}(\nu)$ under the conditions just cited, that we can recover the profile of $d \log n / d\rho$ versus ρ, and thence $N(\rho)$. Note that scattering and reflecting surfaces that are embedded in the refracting medium, including a discontinuity in $n(r)$, can be calculated using the defining integral for $G[\rho, \nu]$ in Eq. (1.5-3). For the case of a reflecting surface, the complete spectral series breaks conveniently into two series, one for the direct field and one for the reflected field.

One technique for accomplishing the recovery of $2G^{\dagger}(\nu)$ is to convert the spectral series in Eq. (1.5-1) into an integral representation in phasor form using the asymptotic forms for the Hankel functions and the spherical harmonic functions. This technique also is applicable for evaluating the electromagnetic field at the LEO using numerical integration aided by the stationary-phase technique. This converted series in scalar form is given to sufficient accuracy (for $\nu < \rho_{\mathrm{L}}$) by

$$\left.\begin{aligned} E(r_{\rm L},\theta_{\rm L}) &= \frac{E_o}{\sqrt{2\pi\rho_{\rm L}\sin\theta_{\rm L}}}\int_0^{\infty}\left(\frac{\sin\theta_\nu^{\rm L}}{\cos\theta_\nu^{\rm G}\cos\theta_\nu^{\rm L}}\right)^{1/2}\exp(i\Psi)d\nu, \\ \Psi &= \sqrt{\rho_{\rm G}^2-\nu^2}+\sqrt{\rho_{\rm L}^2-\nu^2}+\nu(\theta_\nu^{\rm G}+\theta_\nu^{\rm L}-\theta_{\rm L})-2G^{\dagger}(\nu)-\frac{\pi}{4}, \\ \theta_\nu^{\rm G} &= \sin^{-1}\left(\frac{\nu}{\rho_{\rm G}}\right),\ \theta_\nu^{\rm L}=\sin^{-1}\left(\frac{\nu}{\rho_{\rm L}}\right),\ \nu<\rho_{\rm L}<\rho_{\rm G} \end{aligned}\right\} \quad (1.5\text{-}5)$$

Here $\exp(i\Psi)$ is the phasor. $\Psi=\Psi(\rho_{\rm G},\rho_{\rm L},\theta_{\rm L},\nu)$ gives the spectral density of the complete phase delay at the LEO position $(r_{\rm L},\theta_{\rm L})$ relative to the emitting GPS satellite, located at $(r_{\rm G},\theta_{\rm G})$ with $\theta_{\rm G}=\pi$. $\Psi(\rho_{\rm G},\rho_{\rm L},\theta_{\rm L},\nu)$ includes the geometric delay terms and the term $-2G^{\dagger}(\nu)$ for the delay from the refractive gradient. The diagram in Fig. 1-14 shows the total geometric delay from the emitter to the receiver expressed in spectral number space. The term, $(\rho_{\rm L}^2-\nu^2)^{1/2}$, gives the geometric delay in phase along a straight line between the LEO and the tangency point of the line on a sphere of spectral number radius $\nu<\rho_{\rm L}$ centered at the origin. The term $\nu(\theta_{\rm L}-\theta_\nu^{\rm L})$ is an arc length along this sphere of radius ν, and it is subtracted from $(\rho_{\rm L}^2-\nu^2)^{1/2}$ to correct it to the intersection of the sphere with the line $\theta=\pi/2$, which is the fixed reference line for computing phase delays at the LEO. Similarly, the term $(\rho_{\rm G}^2-\nu^2)^{1/2}+\nu\theta_\nu^{\rm G}$ is the geometric phase delay from the emitting GPS satellite along the straight line to the tangency point on the sphere of radius ν and thence along the sphere to the line $\theta=\pi/2$.

Although we have set the upper bound in the spectral integral to ∞, as a practical matter the stationary-phase contributions to the integral come from

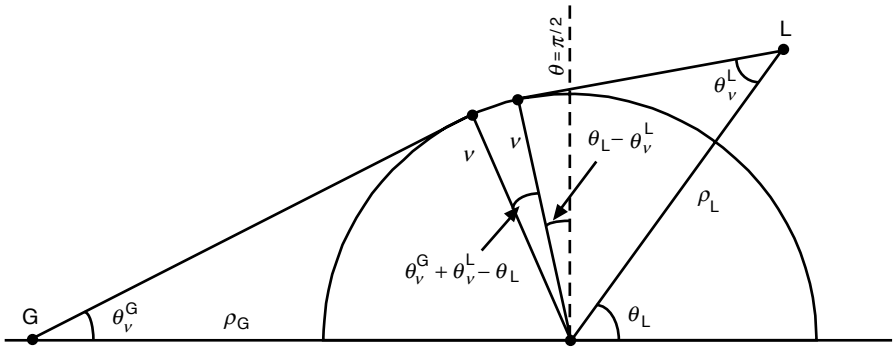

Fig. 1-14. Geometric phase delay diagram in spectral number space.

spectral numbers that are smaller than ρ_L, about 10 percent smaller for a LEO. Therefore, the asymptotic forms given in Eq. (1.5-5) for $\nu < \rho_L$ are valid.

For given satellite positions $(\rho_G, \rho_L, \theta_L)$, the stationary values of Ψ with respect to ν define the neighborhoods in spectral number that contribute to the spectral integral. When a stationary value of Ψ exists, the spectral number providing that stationary value closely equals the impact parameter (in phase units) of the corresponding ray, if near-super-refractivity conditions are avoided. In this case, the stationary value of Ψ essentially equals the value of the eikonal function for the corresponding ray, which is the phase delay φ given in Eq. (1.2-2) for the path integral along the ray (see Eq. (1.5-9) and also Appendix A for its derivation). The stationary values of Ψ with respect to spectral number involve a trade-off between geometric delay and atmospheric delay.

Next, one counter-rotates or "stops" the phase rate of $E(r_L, \theta_L)$ using a realistic model $\varphi_m(\rho_G, \rho_L, \theta_L)$ for the phase, which includes the Doppler shift from the LEO and GPS satellite velocities and the rate of change of the atmospheric phase delay based on geometric optics. Then the phasor in the spectral integral for $E(r_L, \theta_L)$ in Eq. (1.5-5) becomes $\exp[i(\Psi - \varphi_m)]$. This general phase-stopping operation is sometimes referred to as forming a hologram, the coherent mixing of two or more waves. The same holographic operation is performed on the LEO observations, and then a fast Fourier transform (or a similar discrete spectral transform) is performed. Equating these two series, the Fourier transform of the stopped LEO observations to the Fourier transform of the stopped version for $E(r_L, \theta_L)$ in Eq. (1.5-5), leads to an explicit evaluation of $2G^{\dagger}(\nu)$ versus spectral number. The resolution of this recovery is limited by the uncertainty principle in the discrete spectral analysis technique used on the data. The refractivity profile can be obtained by inverting the integral equation for $2G^{\dagger}(\nu)$ defined in Eqs. (1.5-3) and (1.5-4).

It is more efficient to recover the refractivity from $2dG^{\dagger}(\nu)/d\nu$.[2] It can be shown from Eq. (1.5-5) that

$$2\frac{dG^{\dagger}(\nu)}{d\nu} = (\theta_\nu^G + \theta_\nu^L - \theta_{LO}) + i\dot{\theta}_L \left.\frac{d\log\hat{E}[\omega]}{d\omega}\right|_{\omega=\omega_\nu}, \quad \omega_\nu = \omega_m + \dot{\theta}_L \nu \qquad (1.5\text{-}6)$$

where $d\hat{E}[\omega]/d\omega$ is the spectral derivative of the Fourier transform of the stopped $E(r_L, \theta_L)$, θ_{LO} is the orbit angle of the LEO in the plane of propagation

[2] In wave theory, the reason for using $dG^{\dagger}(\nu)/d\nu$ instead of $G^{\dagger}(\nu)$ to recover the refractivity is similar to the reason in geometric optics for using excess Doppler instead of phase. It turns out that $\partial G[\rho,\nu]/\partial\rho = 0$ at $\rho = \rho^{\dagger}$, but not $\partial G[\rho,\nu]/\partial\nu$.

at the temporal center of the data interval, and $\dot{\theta}_{\mathrm{L}}$ is the component of the LEO orbital angular velocity relative to the occulted GPS satellite in the plane of propagation. Note in Eq. (1.5-6) that $\omega_m \approx \pm 2 \times 10^5$ rad / s, or ±(30 to 35) kHz; the sign depends on whether the occultation is rising or setting. On the other hand, within the time interval ΔT over which the Fourier transform is applied, $\omega_\nu / 2\pi$ varies over only a few tens of hertz at most. Equating $d\hat{E}[\omega_\nu] / d\omega_\nu$ in Eq. (1.5-6) to the first derivative of the Fourier transform of the LEO amplitude and stopped phase data yields a determination of the profile for $2dG^{\dagger}(\nu) / d\nu$. This recovered profile for $2dG^{\dagger}(\nu) / d\nu$ is expressed in terms of the refractivity profile through an integral equation derivable from Eqs. (1.5-3) and (1.5-4). This is given by

$$\left.\begin{aligned} 2\frac{dG^{\dagger}}{d\nu} &\doteq 2\pi K_\nu \int_{\rho^{\dagger}}^{\infty} \left(\frac{d\log n}{d\rho}\right)\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right) d\rho, \\ \hat{y} &\doteq \frac{1}{4K_\nu^4}\left(\nu^2 - \rho^2\right),\ K_\nu = \left(\frac{\nu}{2}\right)^{1/3},\ \rho^{\dagger} = \nu - \hat{y}^{\dagger} K_\nu \end{aligned}\right\} \qquad (1.5\text{-}7)$$

The left-hand side (LHS) of this integral equation is evaluated from the spectral derivative of the Fourier transform of the stopped LEO observations. The RHS is a convolution integral that must be inverted to obtain $\log n(\rho)$ in terms of the recovered sequence of values for $2dG^{\dagger} / d\nu$. By replacing the Airy functions with their asymptotic forms applicable for negative values of $\hat{y}$, $\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \to \pi^{-1}(-\hat{y})^{-1/2}$, one can show that the asymptotic version of the integral in Eq. (1.5-7) equals $\alpha_{\mathrm{L}}(\nu)$, which is given in Eq. (1.2-6).

One can apply a wave-theory analog of the Abel integral transform to recover $\log n(\rho)$, but in wave theory the kernel in the integral equation giving the weighting distribution of the contributions to $\log n(\rho)$ is not a Dirac delta function, as it is with the Abel transform, but rather it is spread over a characteristic width ΔL. This characteristic width corresponds to $\sim -2 \le \hat{y} \le 0$. Over this interval, the Airy functions make the transition from their tunneling forms to their negative argument asymptotic forms (sinusoids) corresponding to ray theory. ΔL is given by

$$\Delta L = 2\frac{\lambda K_{\rho_o}}{2\pi} = \left(\frac{n_o r_o \lambda^2}{\pi^2}\right)^{1/3} \qquad (1.5\text{-}8)$$

where r_o is the radius of curvature for the refracting surface. For GPS wavelengths at sea level, $\Delta L \approx 30$ m. Away from super-refractive areas, shadow zones, and caustics, the principal differences in results from a full-

spectrum wave theory versus a wave/optics approach originate from within this relatively narrow altitude band.

Chapter 5 establishes the correspondence between a stationary value with respect to spectral number for the spectral density function of the total delay, $\Psi^* = \Psi(\rho_G, \rho_L, \theta_L, \nu^*)$, when a stationary value exists, and the phase delay or eikonal function $\mathscr{S}(\rho_G, \rho_L, \theta_L, \rho_*)$ for the corresponding ray. This is the ray with an impact parameter value $\rho_* = ka = kr_* n(r_*) \doteq \nu^*$, where ν^* is the spectral number providing a stationary value for $\Psi(\rho_G, \rho_L, \theta_L, \nu)$. From geometric optics, one can show [see Appendix A, Eqs. (A-55) and (A-56)] for an emitting GPS satellite located at the point (r_G, θ_G) with $\theta_G = \pi$, that the phase delay at the LEO position (r_L, θ_L) along a ray with an impact parameter value of ρ_* is given by

$$\mathscr{S} = \rho_G \cos(\chi_G + \delta_G) + \rho_L \cos(\chi_L + \delta_L) + \rho_* \alpha(\rho_*) + \int_{\rho_*}^{\infty} \alpha(\omega) d\omega \qquad (1.5\text{-}9)$$

where δ_G and δ_L are the ray path deflection angles with $\alpha = \delta_G + \delta_L$, and χ_G and χ_L are orbit-related internal angles of the triangle OLG shown in Fig. 1-6. It follows from Eq. (1.5-5) that

$$\Psi(\rho_G, \rho_L, \theta_L, \nu^*) \Leftrightarrow \mathscr{S}(\rho_G, \rho_L, \theta_L, \rho_*), \;\; \nu^* \doteq \rho_* \qquad (1.5\text{-}10)$$

At a stationary point in spectral number, $\nu^* \doteq \rho_*$, and it follows from Fig. 1-14 that $\theta^G_{\nu^*} + \theta^L_{\nu^*} - \theta_L \doteq \alpha_L(\rho_*)$. But if for certain positions $(\rho_G, \rho_L, \theta_L)$ there are no stationary-phase values in spectral number for $\Psi(\rho_G, \rho_L, \theta_L, \nu)$, then in geometric optics there are no rays, no bending angle, and no eikonal function. This would apply to super-refractivity conditions and to strict shadow zones resulting from a discontinuity in the gradient of the refractivity (not necessarily super-refracting) or from an eclipsing limb. But even in these severe situations, $\Psi(\rho_G, \rho_L, \theta_L, \nu)$ still exists as a spectral density function, and it still provides a value for the field at (r_L, θ_L) according to Eq. (1.5-5). The field won't be zero there, just greatly diminished with damping fluctuations. Chapter 2 discusses wave-theory predictions based on scalar diffraction theory of the amplitude and phase at a LEO from various perturbations in the refractivity profile that compromise ray-theory accuracy in the transition regions (see Figs. 2-10 and 2-11). These include a strict shadow zone in amplitude caused by a discontinuity in the gradient of the refractivity, which is not super-refracting, and also by a discontinuity in refractivity, which is super-refracting.

As already mentioned, the potential radial resolution of this wave-theory technique is essentially proportional to λ / L, where L is the component of the distance traveled by the LEO in the propagation plane perpendicular to the limb

direction over ΔT, the time interval. ΔT may be limited by coherence issues; an error in φ_m used to stop the phase rate results in a loss of coherence in the Fourier integral transform if ΔT is too large. Horizontal variations in refractivity probably pose a larger limitation to resolution in the lower troposphere. A horizontal error δD in the assumed value of the limb distance (effectively from an unknown mesoscale horizontal refractivity gradient) translates into a vertical resolution limit that is equal to at least $500(\delta D / D)^{1/2}$ m [72,73]. A 1 percent error in D translates into a 50-m vertical resolution limit.

The formal integral for $G[\rho, \nu]$ in Eq. (1.5-3) enables one to consider a variety of features embedded in the refracting medium. In addition to a discontinuity in $n(\rho)$ or one of its derivatives, or a specular reflecting surface, or an absorbing medium by adding a complex term to $n(r)$, there are other possibilities. Spectra of recovered bending-angle profiles show broadening to varying degrees, which arise from measurement noise and atmospheric "noise" from turbulence, scintillation, and so on [29,68]. With regard to atmospheric noise, one could treat $n(\rho)$ as a stochastic variable, for example, a first-order Markoff process, with a mean value at any altitude equal to the recovered value, but with statistical parameters that are also recovered from the observations.

1.5.6 Appendices

Miscellaneous topics in geometric optics and wave theory are covered by a number of appendices. Appendix A deals principally with concepts from geometric optics and scalar diffraction theory, rays and refractive bending angle, defocusing, the Fresnel zone, the bending-angle perturbation profile from a perturbation in the refractivity profile, phase delay, and so on. Appendices A through F are intended to accompany Chapter 2. Appendix B briefly discusses a caustic surface for a ray family in terms of envelope theory. Appendix C estimates the separation altitudes of the tangency points of multipath rays as a function of the refractivity perturbation across the boundary. This includes a discussion of the effects of a discontinuity in refractivity or in one of its gradients, and the resulting caustics that can follow. Appendix D deals with third-order stationary-phase theory, an important adjunct when dealing with caustic rays in a ray-theoretic context. An index is developed for setting the accuracy of second-order geometric optics (using third-order theory) in the vicinity of a caustic contact point. Appendix E gives the bending-angle profile versus altitude for a Gaussian refractivity distribution, useful in discussing multipath, quasi-shadow zones, and caustic rays. Appendix F comments on the effect of cycle slips from either receiver operations or recovery analyses. Appendix G gives a short summary of the contour integration technique using the complex spectral number plane for summing spectral series over real integers. Appendix H develops the characteristic matrix for a stack of Airy

layers, which is discussed in Chapter 4. Appendix I summarizes the electromagnetic field equations in a stratified medium and the use of modified scalar potentials to derive the field equations. Appendix J addresses the conditions for near-equivalence between certain wave-theory phase delay quantities and their counterparts in geometric optics.

1.6 Limitations and Simplifications

This book is surely not oriented toward working with actual data. Numerous simplifying assumptions have been made to minimize the impact on an already arcane mathematical framework. Regarding the treatment of actual data from the GPS observations, regretfully we have ignored several aspects essential to use of the occultation technique for science. Not the least of our omissions is a discussion of rendering raw data streams into validated data with minimal cycle breaks, obtaining corrections to clock epoch offsets in transmitters and receivers, dealing with the attendant light-time problem to ensure that the differencing of the phase measurements to eliminate clock offsets occurs at common transmitter and/or receiver epochs, strategies for smoothing and sampling noisy data, applying estimation theory, using precision orbit determination, and, very importantly, using data information systems. The references [34,51,91,92] address many of these aspects. In addition, we have simplified the model in which GPS occultations are assumed to occur, again to minimize the impact on the mathematics. Notable examples are the following.

1.6.1 Ionosphere

Except in Chapter 2, where the thin-screen/scalar diffraction technique has been used on the sporadic E-layer, phase effects from the ionosphere have been ignored. For a comprehensive overview of ionospheric applications, see [93]. Linearly combining the dual-frequency signals from the GPS eliminates most of the ionospheric effect [93–95], but small terms involving f^{-3} and higher degrees remain that can be significant, particularly for tangency points in the upper stratosphere. Modeling the ionosphere to capture the higher-degree terms has had some success. Also, extrapolation of the dual-frequency correction from ray tangency points in the mid-troposphere to the lower troposphere (where the increased noise, particularly on the L2 phase from defocusing and interference, can result in errors in the local dual-frequency correction that exceed the correction) has been useful [59,76].

1.6.2 Placing the Occulted GPS Satellite at Infinity

Placing the emitting GPS satellite at infinity simplifies certain equations, mostly in wave theory. In geometric optics and scalar diffraction theory, this approximation can be compensated for to a certain extent by also replacing the

limb distance of the LEO, $D_{\rm L}$, with the reduced limb distance, $D^{-1} = D_{\rm L}^{-1} + D_{\rm G}^{-1}$, where $D_{\rm G}$ is the actual limb distance of the occulted GPS satellite; D is about 10 percent smaller than $D_{\rm L}$. With this simplification, the fundamental relationship between excess Doppler f_D and bending angle α of a ray is given by Eq. (1.4-5). $D\dot{\theta}_{\rm L}$ is essentially the component of the orbital velocity of the LEO in the plane of propagation perpendicular to $\boldsymbol{r}_{\rm LG}$, the direction of the GPS/LEO straight line, or, equivalently, $D\dot{\theta}_{\rm L}$ is the vertical rate of descent or ascent of the GPS/LEO straight line through the atmosphere, 2 to 3 km/s. For the Earth's atmosphere, $\alpha <\sim 0.05$ rad for dry air and water vapor combined, and above the lower troposphere α is less than 0.01 rad. Thus, we have this near-linear relationship between excess Doppler and bending angle given in Eq. (1.4-5).

In contrast, for the case where the GPS satellite is at a finite distance, the appropriate form when the satellites are outside of the refracting medium is given by Eq. (1.2-3). This form can be reduced to geometric quantities in the form

$$\lambda f_D = \boldsymbol{T}_\perp \cdot (\boldsymbol{V}_{\rm L}\delta_{\rm L} - \boldsymbol{V}_{\rm G}\delta_{\rm G}) + {\rm O}\left[\delta_{\rm L}^2, \delta_{\rm G}^2\right] \tag{1.6-1}$$

Here, $\boldsymbol{T}_\perp$ is a unit vector in the propagation plane perpendicular to $\boldsymbol{r}_{\rm LG}$, and $\delta_{\rm L}$ and $\delta_{\rm G}$ are the (small) deflection angles between the vector $\boldsymbol{r}_{\rm LG}$ and $\boldsymbol{T}_{\rm L}$ and $\boldsymbol{T}_{\rm G}$, respectively (see Fig. 1-6). Thus, $\alpha = \delta_{\rm L} + \delta_{\rm G}$, and $\delta_{\rm L}$ is about 8 to 9 times larger than $\delta_{\rm G}$. Nevertheless, dealing with two ray path deflection angles instead of one (even though mutually constrained by Snell's law) and accounting for curvature effects in the approaching wave front from the occulted GPS satellite would further complicate the matter.

As described above, in wave theory a spectral integral representation is used to describe the electromagnetic wave. The main impact of adopting a finite value for $\boldsymbol{r}_{\rm G}$ would fall on the asymptotic form that the lth incoming spectral coefficient $a_l^-(\rho)$ would assume for large values of $\rho = knr$ out of the atmosphere. Here we have assumed that the asymptotic form for $a_l^-(\rho)$ corresponds to a collimated wave, a very simple form involving only the spectral number, $i^{l-1}(2l+1)/l(l+1)$. But with $\boldsymbol{r}_{\rm G}$ finite, the asymptotic form for $a_l^-(\rho)$ would correspond to a spherical wave $\exp(i\rho_{\rm LG})/\rho_{\rm LG}$ with its center at $(r_{\rm G}, \theta_{\rm G})$ with $\theta_{\rm G} = \pi$ (see Fig. 1-6). Therefore, it is an explicit function of spectral number and $\rho_{\rm G}$ through a spherical Hankel function $\xi_l^+(\rho_{\rm G})$—not intractable, but certainly more complicated [96]. The asymptotic form of $a_l^-(\rho)$ for the case of a spherical incident wave is briefly discussed in Chapter 5, Eq. (5.5-3b). This asymptotic form involves an additional phase offset,

denoting a shift of the reference point for phase calculation from the line $\theta = \pi / 2$ used in the collimated case to the position of the emitter. Also, there is an amplitude term to account for the space loss from a spherical wave. But almost all of the subsequent discussion in Chapter 5 assumes a collimated incident wave. The adjustments to account for a finite GPS satellite distance are briefly presented without derivation in Section 5.10, where ray-theory correspondence is addressed, and in Section 6.4 in connection with refractivity recovery.

1.6.3 Time

We have assumed Newtonian time throughout this monograph. But with actual data when the predicted phase must be computed for a wave moving at the speed of light, several different times and time intervals are involved: proper and coordinate times from special and general relativity, asynchronous timekeeping among all of the receiver and transmitter clocks in the system, transmit time, receive time, topocentric time, atomic time, and coordinated Universal time [34,97]. Keeping track of all these aspects of representing and synchronizing time, essential in actual data processing, unnecessarily complicates our presentation.

1.6.4 Spherical Symmetry

For the terrestrial planets with mild oblateness, the assumption of spherical symmetry is moderately accurate, but for the Jovian planets it was demonstrated about 25 years ago that oblateness must be carefully accounted for [98–100]. The usual approach for the Earth, which has a 20-km difference between its equatorial and polar radii, or 0.3 percent, is to apply the oblateness correction by adjusting the location of the geocenter using a local geoid. For a given geographical location of the tangency point of the occultation, the local radius of curvature vector for the sphere can be adjusted to fit the radius of curvature vector for the local geoid. Thus, a local spherical symmetry is assumed that more or less fits the curvature of the geopotential surface at that geographical location. The radius vector of the local sphere is equated to and aligned with the radius of curvature vector for the geopotential surface. We have assumed spherical symmetry here.

Other aspects of departures from spherical symmetry, for example, horizontal variability in the refractivity, are not considered here [38].

1.6.5 Coplanarity

The assumption of coplanarity between the propagation plane and the satellite orbit planes runs into difficulty mainly when kinematics arise. In Chapter 6, we briefly address this by identifying two different orbit angles that

locate the position of the LEO relative to the occulted GPS satellite, θ in the LEO orbit plane and $\tilde{\theta}$ in the plane of propagation. The obliquity between these two planes is readily expressed in terms of orbit elements, and it is nearly constant over an occultation episode for the neutral atmosphere. Depending on the actual orbits, the obliquity factor $d\tilde{\theta}/d\theta$ ranges between 0.6 and 1.0 for most occultations used for analysis. As discussed above, the term $d\tilde{\theta}/dt$ appears in the expressions for the excess Doppler, but it can be expressed in a near-linear form in terms of $d\theta/dt$ through spherical trigonometry when the obliquity factor is given. Also, departures from planarity from cross-track horizontal variations in refractivity are not considered here.

1.6.6 Circularity

The assumption that the LEO orbit is circular only arises here in connection with the representation of the excess Doppler in terms of the atmospheric bending and the satellite velocity. It is a fairly valid assumption, but with actual data no such approximation is made. The orbits of the LEO and the GPS satellite constellation and their clock epoch offsets are determined independently from POD information based on continuous GPS tracking data from the LEO and from a network of ground stations including the IGS.

1.6.7 Treating the GPS Signal as a Harmonic Wave

The C/A and P ranging codes on the GPS carriers are pseudorandom, phase-modulating square waves that fully suppress the carrier tone. Each transition of a code, occurring at a frequency or chip rate of 1.023 MHz for the C/A code and at 10.23 MHz for the P code, involves a change in phase of the carrier of either zero or 180 deg in accordance with the pseudorandom recipe for that particular code. Therefore, the radio frequency (RF) power spectra of the GPS navigation signals have somewhat complicated shapes. The L1 signal carries both the C/A and P codes, but on the present GPS satellites the L2 carries only the P code; however, all are phase coherent at the transmitter. Both the C/A and P codes produce $(\sin x/x)^2$-like spread spectra with single-side bandwidths of about 1 and 10 MHz, respectively. The side lobes of the P-code spectrum at the higher harmonics of 10 MHz are attenuated at a higher rate than the $(\sin x/x)^2$ envelope decays to minimize spillover power in nearby user bands. The power in the C/A code is 3 dB greater than the power in the P code. The GPS receiver applies an appropriately time-delayed and Doppler-shifted replica of the transmitted C/A code to the received signal, and carries out a number of cross-correlation and digital signal processing tasks. Through these operations, the receiver isolates the navigation signal emitted by a specific GPS satellite from all others using the orthogonality property of the ranging codes. The receiver in essence collapses the spread-spectrum signal for each carrier

from the selected GPS satellite into a single tone, which is the equivalent of a harmonic wave with a frequency equal to that of the received carrier. Then the receiver measures the phase and amplitude of this equivalent harmonic wave in the presence of noise. Some of the operational aspects of extracting the phase and amplitude measurements from noisy digital data streams are discussed briefly in Chapter 6. For details on digital signal processing by a particular space geodetic GPS receiver, the TurboRogue, and on dealing with the encrypted P code—the so-called anti-spoofing (AS)—for recovery of the L2 phase, see [101,102]. We assume here that the received signal is indeed from harmonic waves with the L1 and L2 frequencies plus the Doppler shifts from satellite motion and from refraction. We also assume that the receiver is reporting at a suitable sample rate, nominally at 50 Hz, perhaps higher if needed, the measured phase and amplitude for this harmonic wave. Also, we assume that phase connection between successive measurements has been successful—a task not always completely achievable with actual data in adverse signal conditions.

1.7 Recommendations for the Next Chapters

Which chapters to pursue depends on one's interest. Chapter 2 and Appendices A through E have a great deal of basic material pertaining to geometric optics, scalar diffraction theory, multipath, shadow zones, caustics, stationary-phase theory, and third-order stationary-phase theory for dealing with caustics.

Chapter 3 discusses basic Mie scattering theory, but with an emphasis on a phasor-based spectral representation for the scattered wave and associated stationary-phase concepts. It can be used as a reference as needed in the later chapters for asymptotic forms, phasor representations, correspondence with ray theory concepts, and so on.

Chapter 4 stands alone. It addresses thin-film theory, which uses a unitary state transition matrix to describe electromagnetic wave propagation in a laminar medium. Several useful concepts found also in Chapters 3 and 5 are developed in Chapter 4, which offers an easier Cartesian framework for their introduction: incoming and outgoing standing waves, osculating parameters, asymptotic matching methodology, turning points, and problems therefrom. It also introduces the Airy layer wherein the refractivity gradient is constant. It is called an Airy layer because the wave equation solutions in this medium are Airy functions. It has important uses in Chapter 5 for asymptotic matching of incoming and outgoing spectral coefficients and for dealing with a turning point.

Chapter 5 provides the basis for the modified Mie scattering theory applicable to a refracting medium with or without an embedded scattering surface.

However, if one's interest is primarily in recovery issues using the full-spectrum wave-theory approach, then the main effort should be on Chapter 6 with occasional reference to Chapter 5 for the basic wave-theory fundamentals regarding the full-spectrum treatment in a spherical symmetric refracting medium.

References

[1] C. Hazard, "The Method of Lunar Occultations and Its Application to a Survey of the Radio Source 3C212," *Monthly Notices of the Royal Astronomical Society*, vol. 124, pp. 343–357, 1962.

[2] S. Chandrasekhar, "A Statistical Basis for the Theory of Stellar Scintillation," *Monthly Notices of the Royal Astronomical Society*, vol. 112, pp. 475–483, 1952.

[3] W. Hubbard, J. Jokopii, and B. Wilking, "Stellar Occultations by Turbulent Planetary Atmospheres: A Wave-Optical Theory Including a Finite Scale Height," *Icarus*, vol. 34, pp. 374–395, 1978.

[4] E. Saltpeter, "Interplanetary Scintillations. I. Theory," *The Astrophysical Journal*, vol. 147, pp. 433–448, 1967.

[5] A. Kliore, T. Hamilton, and D. Cain, *Determination of Some Physical Properties of the Atmosphere of Mars from Changes in the Doppler Signal of a Spacecraft on an Earth Occultation Trajectory*, Technical Report 32-674, Jet Propulsion Laboratory, Pasadena, California, 1964.

[6] G. Fjeldbo, *Bistatic Radar Methods for Studying Planetary Ionospheres and Surfaces*, Science Report No. 2, NsG-377, SU-SEL-64-025, Stanford Electronics Laboratories, Stanford, California, 1964.

[7] A. Kliore, D. Cain, G. Levy, V. Eshleman, G. Fjeldbo, and F. Drake, "Occultation Experiment: Results of the First Direct Measurement of Mars' Atmosphere," *Science*, vol. 149, pp. 1243–1248, 1965.

[8] G. Fjeldbo, A. J. Kliore, and V. R. Eshleman, "The Neutral Atmosphere of Venus as Studied by Mariner V Radio Occultation Experiments," *The Astronomical Journal*, vol. 76, no. 2, pp. 123–140, 1971.

[9] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120–166, 1986.

[10] G. Lindal, J. Lyons, D. Sweetnam, V. Eshleman, D. Hinson, and G. Tyler, "The Atmosphere of Uranus: Results of Radio Occultation Measurements with Voyager 2," *Journal of Geophysical Research,* vol. 92, pp. 14,987–15,001, 1987.

[11] G. Tyler, "Radio Propagation Experiments in the Outer Solar System with Voyager," *Proceedings of the IEEE*, vol. 75, pp. 1404–1431, 1987.

[12] G. Lindal, "The Atmosphere of Neptune: An Analysis of Radio Occultation Data Acquired by Voyager 2," *The Astronomical Journal,* vol. 103, pp. 967–982, 1991.

[13] O. Yakovlev, I. Matyugov, and I. A. Vilkov, "Radio-Wave Phase and Frequency Fluctuations as Observed in Radio Occultation Experiments on the Satellite-to-Satellite Link," *Journal of Communications Technology and Electronics*, vol. 41, no. 11, pp. 993–998, 1996.

[14] A. S. Liu, *On the Determination and Investigation of the Terrestrial Ionospheric Refractive Indices Using GEOS-3/ATS-6 Satellite-to-Satellite Tracking Data; Final Report*, NASA-CR-156848, November 1, 1978.

[15] T. Yunck, G. Lindal, and C. H. Liu, "The Role of GPS in Precise Earth Observations," *Proceedings of the IEEE: Position Location and Navigation Symposium*, Orlando, Florida, 1988.

[16] R. Ware, M. Exner, D. Feng, M. Gorbunov, K. Hardy, B. Herman, Y.-H. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Solheim, X. Zou, R. Anthes, S. Businger, and K. Trenberth, "GPS Sounding of the Atmosphere from Low Earth Orbit: Preliminary Results," *Bulletin of the American Meteorological Society*, vol. 77, pp. 19–40, 1996.

[17] E. Kursinski, G. Hajj, W. Bertiger, S. Leroy, T. Meehan, L. Romans, J. Schofield, D. McCleese, W. Melbourne, C. Thornton, T. Yunck, J. Eyre, and R. Nagatani, "Initial Results of Radio Occultation Observations of Earth's Atmosphere Using the Global Positioning System," *Science*, vol. 271, pp. 1107–1110, 1996.

[18] G. A. Hajj, E. R. Kursinski, W. I. Bertiger, S. S. Leroy, T. Meehan, L. J. Romans, and J. T. Schofield, "Initial Results of GPS–LEO Occultation Measurements of Earth's Atmosphere Obtained with the GPS/MET Experiment," *Proceedings of the IAG Symposium*, G1, GPS Trends in Precise Terrestrial, Airborne, and Spaceborne Applications, IUGG XXI General Assembly, Boulder, Colorado, July 2–14, 1995, Springer-Verlay, pp. 144–154, 1996.

[19] C. Rocken, R. Anthes, M. Exner, D. Hunt, S. Sokolovskiy, R. Ware, M. Gorbunov, W. Schreiner, D. Feng, B. Herman, Y.-H. Kuo, and X. Zou, "Analysis and Validation of GPS/MET Data in the Neutral Atmosphere," *Journal of Geophysical Research,* vol. 102, pp. 29,849–29,866, 1997.

[20] J. Wickert, T. Schmidt, C. Marquardt, Ch. Reigber, K.-H. Neumayer, G. Beyerle, R. Galas, and L. Grunwaldt, "GPS Radio Occultation with CHAMP: First Results and Status of the Experiment," *Proceedings of IAG Scientific Assembly*, *Vistas for Geodesy in the New Millenium*, September 2–7, 2001, Budapest, Hungary, Springer Series, vol. 125, pp. 273–278, 2002.

[21] G. Hajj, C. Ao, B. Iijima, D. Kuang, E. Kursinski, A. Mannucci, T. Meehan, L. Romans, M. de la Torre Juarez, and T. Yunck, "CHAMP and SAC-C Atmospheric Occultation Results and Intercomparisons," submitted to *Journal of Geophysical Research, Atmosphere* August 2, 2002.

[22] R. Anthes, C. Rocken, and Y. H. Kuo, "Applications of COSMIC to Meteorology and Climate," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 115–156, 2000.

[23] Y.-H. Kuo, X. Zou, F. Vandenberghe, W. Huang, and R. Anthes, "Assimilation of GPS/MET Data for Numerical Weather Prediction," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 157–186, 2000.

[24] C. Rocken, Y.-H. Kuo, W. Schreiner, D. Hunt, S. Sokolovskiy, and C. McCormick, "COSMIC System Description," *Terrestrial, Atmospheric and Oceanic Science*, vol. 11, pp. 21–52, 2000.

[25] T. Yunck, C. H. Liu, and R. Ware, "A History of GPS Sounding," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 1–20, 2000.

[26] T. P. Yunck, G. A. Hajj, E. R. Kursinski, J. A. LaBrecque, S. T. Lowe, M. M. Watkins, and C. McCormick, "AMORE: An Autonomous Constellation Concept for Atmospheric and Oceanic Observation," *Acta Astronautica*, vol. 46, no. 2-6, pp. 355–364, 2000.

[27] B. Ahmad and G. Tyler, "The Two-Dimensional Resolution Kernel Associated with Retrieval of Ionospheric and Atmospheric Refractivity Profiles Using the Abelian Inversion of Radio Occultation Ohase Data," *Radio Science*, vol. 33, pp. 469–478, 1997.

[28] A. Pavelyev, A. Zakharov, A. Kucherjavenkov, E. Molotov, I. Siderenko, I. Kucherjavenkova and D. Pavelyev, "Propagation of Radio Waves Reflected from Earth's Surface at Grazing Angles between a Low-Orbit Space Station and a Geostationary Satellite," *Journal of Communications Technology and Electronics*, vol. 42, no. 1, pp. 51–57, 1997.

[29] G. Beyerle, K. Hocke, J. Wickert, T. Schmidt, C. Marquardt, and Ch. Reigber, "GPS Radio Occultations with CHAMP: A Radio Holographic Analysis of GPS Signal Propagation in the Troposphere and Surface Reflections," *Journal of Geophysical Research*, vol. 107, no. D24, 4802, doi: 10.1029/2001JD001402, 2002.

[30] *Global Positioning System*, ed. P. M. Janiczek, vols. 1–3, Washington, DC: Institute of Navigation, 1980.

[31] B. Parkinson and J. Spilker, eds., *Global Positioning System: Theory and Application*, American Institute of Aeronautics and Astronautics, Washington, DC, 1996.

[32] G. Gendt, P. Fang, and J. Zumberge, "Moving IGS Products Towards Real-Time," ftp://igscb.jpl.nasa.gov/IGSCB/resource/Pubs/99_tec/99tr_AC_WS_7.pdf 1999.

[33] R. Galas, J. Wickert, and W. Burghardt, "High Rate Low Latency GPS Ground Tracking Network for CHAMP," *Physics and Chemistry of the Earth*, part A, vol. 26, issues 6–8, pp. 649–652, 2001.

[34] G. Hajj, E. Kursinski, L. Romans, W. Bertiger, and S. Leroy, "A Technical Description of Atmospheric Sounding by GPS occultation," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, pp. 451–469, 2002.

[35] J. Zumberge, and G. Gendt, "The Demise of Selective Availability and Implications for the International GPS Service," *Physics and Chemistry of the Earth,* part A, vol. 26, issues 6–8, pp. 637–644, 2001.

[36] G. Beutler, M. Rothacher, S. Schaer, T. A. Springer, J. Kouba, and R. E. Neilan, "The International GPS Service (IGS): An Interdisciplinary Service for Earth Sciences," *Advances in Space Research*, vol. 23, no. 4, pp. 631–653, 1999.

[37] R. Neilan, J. Zumberge, G. Beutler, and J. Kouba, "The International GPS Service: A Global Resource for GPS Applications and Research," *Proc. ION-GPS-97*, The Institute of Navigation, pp. 883–889, 1997.

[38] B. Ahmad and G. Tyler, "Systematic Errors in Atmospheric Profiles Obtained from Abelian Inversion of Radio Occultation Data: Effects of Large Scale Horizontal Gradients," *Journal of Geophysical Research*, vol. 104, no. D4, pp. 3971–3992, 1999.

[39] S. Sokolovskiy and D. Hunt, "Statistical Optimization Approach for GPS/MET Data Inversions," *URSI GPS/MET Workshop*, Tucson, Arizona, February 1996.

[40] R. Phinney and D. Anderson, "On the Radio Occultation Method for Studying Planetary Atmospheres," *Journal of Geophysical Research*, vol. 73, no. 5, pp. 1819–1827, 1968.

[41] G. Fjeldbo and V. Eshelman, "The Atmosphere of Mars Analyzed by Integral Inversion of the Mariner IV Occultation Data," *Planetary Space Science*, vol. 16, pp. 1035–1059, 1968.

[42] E. Kursinski, G. Hajj, J. Schofield, R. Linfield, and K. Hardy, "Observing Earth's Atmosphere with Radio Occultation Measurements Using the Global Positioning System," *Journal of Geophysical Research,* vol. 102, pp. 23,429–23,465, 1997.

[43] O. Tetens, "Über Einige Meteorologische Begriffe," *Zeitschrift für Geophysik*, vol. 6, pp. 297–309, 1930.

[44] J. Eyre, "Assimilation of Radio Occultation Measurements into a Numerical Weather Prediction System," ECMWF Technical Memorandum 199, European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom, 1994.

[45] X. Zou, Y.-H. Kuo, and Y.-R. Guo, "Assimilation of Atmospheric Radio Refractivity Using a Nonhydrostatic Adjoint Model," *Monthly Weather Review*, vol. 123, pp. 2229–2249, 1995.

[46] X. Zou, F. Vandenberghe, B. Huang, M. Gorbunov, Y.-H. Kuo, S. Sokolovskiy, J. Chang, J. Sela, and R. Anthes, "A Ray-Tracing Operator and Its Adjoint for the Use of GPS/MET Refraction Angle Measurements," *Journal of Geophysical Research, Atmospheres*, vol. 104, no. D18, pp. 22,301–22,318, 1999.

[47] Y.-H. Kuo, X. Zou, F. Vandenberghe, W. Huang, and R. Anthes, "Assimilation of GPS/MET Data for Numerical Weather Prediction," *Terrestrial, Atmospheric and Oceanic Sciences*, 1999.

[48] M. Born and E. Wolf, *Principles of Optics,* 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.

[49] R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, New York: McGraw-Hill, 1965.

[50] S. Leroy, "Measurement of Geopotential Heights by GPS Radio Occultation," *Journal of Geophysical Research*, vol. 102, pp. 6971–6986, 1997.

[51] K. Hocke, "Inversion of GPS Meteorology Data," *Annales Geophysicae*, vol. 15, pp. 443–450, 1997.

[52] K. Hocke, T. Tsuda, and A. de la Torre, "A Study of Stratospheric GW Fluctuations and Sporadic E at Midlatitudes with Focus on Possible Orographic Effect of Andes*," Journal of Geophysical Research*, vol. 107, no. D20, 4428, doi: 10.1029/2001JD001330, October 2002.

[53] K. Igarashi, A. Pavelyev, K. Hocke, D. Pavelyev, and J. Wickert, "Observation of Wave Structures in the Upper Atmosphere by Means of Radio Holographic Analysis of the Radio Occultation Data," *Advances in Space Research*, vol. 27, no. 6–7, pp. 1321–1326, 2001.

[54] A. Pavelyev, K. Igarashi, Ch. Reigber, K. Hocke, J. Wickert, G. Beyerle, S. Matyugov, A. Kucherjavenkov, D. Pavelyev, and O. I. Yakovlev, "First Application of the Radio Holographic Method to Wave Observations in the Upper Atmosphere," *Radio Science*, vol. 37, no. 3, pp. 15-1–15-11, 2002.

[55] *Atmospheric Profiling Mission*, ESA Report SP-1196, no. 7, European Space Agency, April 1996.

[56] Y.-H. Kuo, X. Zou, and W. Huang, "The Impact of GPS Data on the Prediction of an Extratropical Cyclone: An Observing System Simulation Experiment," *Dynamics of Atmospheres and Oceans*, vol. 27, pp. 439–470, 1997.

[57] E. K. Smith, Jr., and S. Weintraub, "The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies," *Proceedings of the IRE*, vol. 41, pp. 1035–1037, 1953.

[58] D. Thayer, "An Improved Equation for the Radio Refractive Index in Air," *Radio Science*, vol. 9, pp. 803–807, 1974.

[59] G. A. Hajj and L. J. Romans, "Ionosphere Electron Density Profiles Obtained with the Global Positioning System: Results from the GPS/MET Experiment," *Radio Science,* vol. 33, no. 1, pp. 175–190, 1998.

[60] M. E. Gorbunov, S. V. Sokolovsky, and L. Bengtsson, *Space Refractive Tomography of the Atmosphere: Modeling of Direct and Inverse Problems*, Report 210, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.

[61] T. Meehan, personal communication, Jet Propulsion Laboratory, Pasadena, California, 1995.

[62] A. G. Pavelyev, A. V. Volkov, A. I. Zakharov, S. A. Krutikh, and A. I. Kucherjavenkov, "Bistatic Radar as a Tool for Earth Investigation Using Small Satellites," *Acta Astronautica*, vol. 39, no. 9–12, pp. 721–730, 1996.

[63] K. Hocke, A. Pavelyev, O. Yakovlev, L. Barthes, and N. Jakowski, "Radio Occultation Data Analysis by the Radioholographic Method," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 61, pp. 1169–1177, 1999.

[64] K. Igarashi, A. Pavelyev, K. Hocke, D. Pavelyev, I. Kucherjavenkov, S. Matyugov, A. Zakharov, and O. Yakovlev, "Radio Holographic Principle for Observing Natural Processes in the Atmosphere and Retrieving Meteorological Parameters from Radio Occultation Data," *Earth, Planets and Space*, vol. 52, no. 11, pp. 893–899, 2000.

[65] M. E. Gorbunov, A. S. Gurvich, and L. Kornblueh, "Comparative Analysis of Radioholographic Methods of Processing Radio Occultation Data," *Radio Science*, vol. 35, no. 4, pp. 1025–1034, 2000.

[66] S. Sokolovsky, "Modeling and Inverting Radio Occultation Signals in the Moist Troposphere," *Radio Science*, vol. 36, no. 3, pp. 441–458, 2001.

[67] A. Pavelyev, Y.-A. Liou, Ch. Reigber, J. Wickert, K. Igarashi, K. Hocke, and C.-Y. Huang, "GPS Radio Holography as a Tool for Remote Sensing of the Atmosphere and Mesosphere from Space," *GPS Solutions*, vol. 6, pp. 100–108, 2002.

[68] K. Igarashi, K. Pavelyev, J. Wickert, K. Hocke, and D. Pavelyev, "Application of Radio Holographic Method for Observation of Altitude Variations of the Electron Density in the Mesosphere/Lower Thermosphere Using GPS/MET Radio Occultation Data," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, no. 8–11, pp. 959–969, 2002.

[69] O. I. Yakovlev, A. G. Pavelyev, S. S. Matyugov, D. A. Pavelyev, and V. A. Anufriev, *Radioholographic Analysis of Radiooccultation Data*, Scientific Technical Report 02/05, GeoForschungsZentrum-Potsdam, Potsdam, Germany, 2002.

[70] J. Goodman, *Introduction to Fourier Optics*, New York: McGraw-Hill, 1968.

[71] E. N. Bramley, "The Accuracy of Computing Ionospheric Radiowave Scintillation by the Thin-Phase Screen Approximation," *Journal of Atmospheric and Solar-Terrestrial Physic*s, vol. 39, pp. 367–373, 1977.

[72] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120–166, 1986.

[73] W. G. Melbourne, E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, E. R. Kursinski, T. K. Meehan, L. E. Young, and T. P. Yunck, *The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring*, JPL Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, April 1994.

[74] M. Meincke, *Inversion Methods for Atmospheric Profiling with GPS Occultation,* Scientific Report No. 99-11, Danish Meteorological Institute, Copenhagen, Denmark, 1999.

[75] A. J. Devaney, "A Filtered Backpropagation Algorithm for Diffraction Tomography," *Ultrasonic Imaging*, vol. 4, pp. 336–350, 1982.

[76] M. E. Gorbunov, A. S. Gurvich, and L. Bengtsson, *Advanced Algorithms of Inversion of GPS/MET Satellite Data and Their Application to Reconstruction of Temperature and Humidity*, Report 211, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.

[77] E. Karayel and D. Hinson, "Sub-Fresnel Scale Vertical Resolution in Atmospheric Profiles from Radio Occultation," *Radio Science*, vol. 32, no. 2, pp. 411–423, 1997.

[78] R. H. Hardin and F. D. Tappert, "Applications of the Split-Step Fourier Method to the Numerical Solution of Nonlinear and Variable Coefficient Wave Equations," *SIAM Review,* vol. 15, p. 423, 1973.

[79] D. Knepp, "Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves," *Proceedings of the IEEE*, vol. 71, no. 6, pp. 722–737, 1983.

[80] J. Martin and S. Flatté, "Intensity Images and Statistics from Numerical Simulation of Wave Propagation in 3-D Random Media," *Applied Optics*, vol. 27, no. 11, pp. 2111–2126, 1988.

[81] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, United Kingdom: The Institution for Electrical Engineers, 2000.

[82] G. Beyerle, M. Gorbunov, and C. Ao, "Simulation Studies of GPS Radio Occultation Measurements," *Radio Science*, in press, 2003.

[83] C. Ao, G. Hajj, T. Meehan, S. Leroy, E. Kursinski, M. de la Torre Juárez, B. IiJima, and A. Mannucci, "Backpropagation Processing of GPS Radio Occultation Data," in *CHAMP Science Meeting*, Ch. Reigber, H. Lühr, and P. Schwintzer, eds., January 21–24, 2002, Springer-Verlag, Berlin, Germany, pp. 415–422, 2003.

[84] M. E. Gorbunov, "Canonical Transform Method for Processing Radio Occultation Data in the Lower Troposphere," *Radio Science,* vol. 37, no. 5, 1076, doi:10.1029/2000RS002592, 2002.

[85] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards Series 55, Washington, DC, 1964.

[86] F. Olver, *Introduction to Asymptotics and Special Functions*, New York: Academic Press, 1974.

[87] W. G. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occultation Observations from a Low Earth Orbiter, Part 1*, JPL Publication 99-5, Jet Propulsion Laboratory, Pasadena, California, July 1998.

[88] W. G. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occultation Observations from a Low Earth Orbiter, Part 2*, JPL Publication 99-5, Jet Propulsion Laboratory, Pasadena, California, April 1999.

[89] F. Abelès, "Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Applications to Thin Films," *Annales de Physique*, vol. 5, pp. 596–640, September–October 1950 (in French).

[90] F. Abelès, "Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Application to Thin Films. II. Thin Films," *Annales de Physique*, vol. 5, pp. 706–782, November–December 1950 (in French).

[91] J. Wickert, G. Beyerle, T. Schmidt, C. Marquardt, R. König, L. Grunwald, and C. Reigber, "GPS Radio Occultation with CHAMP," *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*, eds. C. Reigber, H. Luhr, and P. Schwintzer, proceedings from the First CHAMP Science Meetiing, Potsdam, Germany, January 21–24, 2002, Springer-Verlag, 2002.

[92] http://genesis.jpl.nasa.gov/

[93] G. Hajj, L. Lee, X. Pi, L. Romans, W. Schreiner, P. Strauss, and C. Wang, "COSMIC GPS Ionosphere Sensing and Space Weather," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 235–272, 2000.

[94] S. Bassiri and G. Hajj, "Higher-Order Ionospheric Effects on the Global Positioning System Observables and Means of Modeling Them," *Manuscripta Geodaetica*, vol. 18, pp. 280–289, 1993.

[95] V. V. Vorob'ev and T. G. Krasil'nikova, "Estimation of the Accuracy of the Atmospheric Refractive Index Recovery from Doppler Shift Measurements at Frequencies Used in the NAVSTAR System," *Physics of the Atmosphere and Ocean*, vol. 29, no. 5, pp. 602–609, 1994.

[96] J. Jackson, *Classical Electrodynamics,* 2nd ed., New York: John Wiley & Sons, Inc., 1975.

[97] T. D. Moyer, *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation*, Monograph 2, Deep Space Communications and Navigation Series, Jet Propulsion Laboratory, Pasadena, California, October 2000.

[98] W. Hubbard, "Ray Propagation in Oblate Atmospheres," *Icarus*, vol. 27, pp. 387–388, 1976.

[99] A. J. Kliore, P. M. Woiceshyn, and W. B. Hubbard, "Temperature of the Atmosphere of Jupiter from Pioneer 10/11 Radio Occultations," *Geophysical Research Letters*, vol. 3, pp. 113–116, 1976.

[100] G. F. Lindal, D. N. Sweetnam, and V. R. Eshleman, "The Atmosphere of Saturn: An Analysis of the Voyager Radio Occultation Measurements," *The Astronomical Journal,* vol. 90, pp. 1136–1146, 1985.

[101] J. B. Thomas, *Signal-Processing Theory for the TurboRogue Receiver*, JPL Publication 95-6, Jet Propulsion Laboratory, Pasadena, California, 1995.

[102] C. Ao, T. Meehan, G. Hajj, A. Mannucci, and G. Beyerle, "Lower Troposphere Refractivity Bias in GPS Occultation Retrievals," submitted to *Journal of Geophysical Research*, 2003.

Chapter 2
Scattering of Electromagnetic Waves from a Spherical Boundary Using a Thin Phase Screen Model and Scalar Diffraction Theory

2.1 Introduction

This monograph focuses principally on the calculation of the electromagnetic field vector observed by a low Earth orbiting (LEO) spacecraft during an occultation of a Global Positioning System (GPS) satellite that is broadcasting a navigation signal. Of specific interest is the change in the observed signal amplitude and phase of the electric field that results from a sharp change across a spherical surface of a refraction-related property of the atmosphere. This would include a discontinuity in the atmospheric refractivity itself, or in some other quantity, such as the gradient of the refractivity caused, for example, by a change in scale height or the lapse rate of the temperature profile. Discontinuities such as these, and also milder dislocations, can result in multiple rays arriving concurrently at the LEO or, conversely, in shadow zones where no rays (according to geometric optics) arrive. Diffraction also is a by-product of these types of refractivity changes, but the amplitudes and frequencies of the diffraction fringes very much depend on the sharpness of the change and its type. Also, sometimes, caustics are present for certain LEO orbital positions.

The calculation of the observed electromagnetic field for these situations can be very difficult. Even with certain geometric assumptions in place, such as spherical symmetry, GPS/LEO orbital coplanarity, and a planar wave front for the incident wave, the exact wave-theoretic solution to this problem based on

Maxwell's equations is not readily tractable. Mie scattering theory is an example of a wave-theoretic approach to this spherical problem. Chapter 3 uses Mie scattering theory to study scattering from a spherical surface in an otherwise homogeneous medium. Chapter 5 uses a modified version of Mie scattering to deal with a scattering surface that is embedded in a refracting medium. Both of these chapters involve a full-wave theory approach, which results in spectral series representations of the electromagnetic field that are solutions to the Helmholtz equation.

Here in Chapter 2, however, we use a combination of geometric optics and scalar diffraction theory applied to a thin phase screen model to develop the phase and amplitude profile that would be observed by a LEO as a result of a change in a refractivity-related quantity. The thin screen becomes a proxy for the actual atmosphere. A vertical profile for the atmosphere-induced phase delay is embedded in the thin-screen model in such a way that the phase and amplitude profile observed by the LEO over time matches (according to geometric optics) the actual profile. This model in a ray-theoretic framework greatly simplifies the calculations. But, when diffraction effects are included, one must also deal with a convolution integral over the vertical profile in the thin phase screen. In wave theory, the spectral series representations of the electromagnetic field are essentially integrals over spectral number space. In the thin-screen approach, the integral representation of scalar diffraction effects is effectively over impact parameter space.

In Fig. 2-1 we see a ray with its point of tangency below the surface r_o, that is, $r_* < r_o$. This surface at $r = r_o$ marks a boundary separating the "+" and the "–" regimes. Across this surface a single discontinuity is assumed to occur (in an otherwise benign medium) in one of the parameters that define the relationship between the refractivity and the radial distance r. We will consider the effects of small discontinuities at r_o in three parameters: N_o, the refractivity at r_o; H_{p_o}, the pressure scale height; and $\gamma = T^{-1} dT / dr$, the normalized lapse rate of the temperature profile. The profile of the observed bending angle of the signal from an occulted GPS satellite will change as a result of a discontinuity at r_o in any one of these quantities.

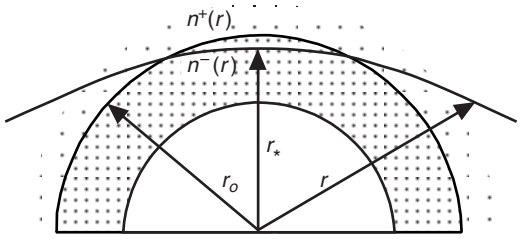

Fig. 2-1. Ray path in a stratified medium with a spherical boundary separating two regimes.

2.1.1 Multipath Scenarios

Figure 2-2 shows different multipath scenarios. Figure 2-2(a) is a schematic ray diagram for a collimated beam of plane waves encountering from the left a sphere of lower refractivity than that in the surrounding medium, a so-called

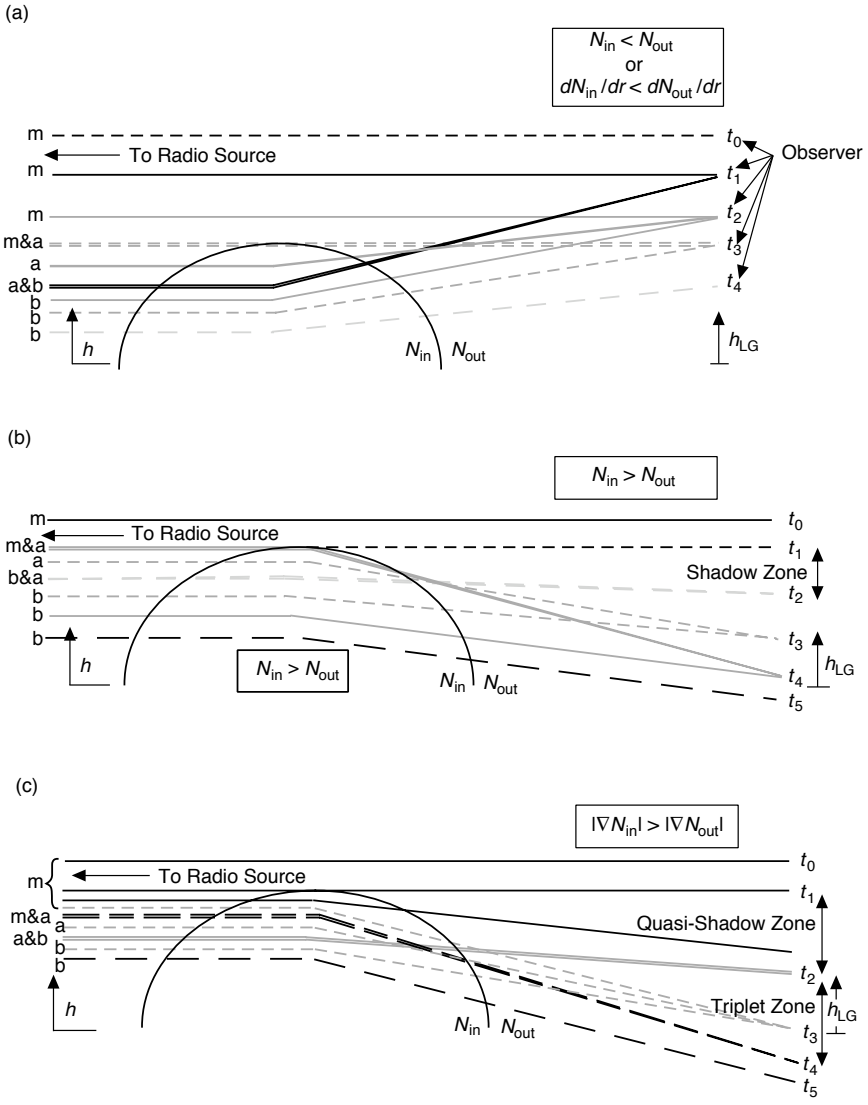

Fig. 2-2. Multipath scenarios caused by a refracting sphere according to geometric optics. Rays arriving on the right at the same point have a common observation epoch denoted by $t_0,t_1,t_2,\cdots$: (a) $N(r)$ or its gradient is less inside than outside, (b) $N(r)$ is greater inside with super-refractivity, and (c) the gradient of $N(r)$ is greater inside without super-refractivity. The altitude of the observer is denoted by h_{LG}, and the altitude of the impact parameter of the ray is denoted by h.

sub-refraction scenario. For a setting occultation, the observer may be thought of here as traveling vertically downward on the right-hand side of the panels along the points labeled $t_0, t_1, t_2, \cdots$, which denote the epochs at which one observes the rays arriving at those corresponding points. The altitude of the observation point is denoted in the figure (relative, say, to the bottom of the frame) by the variable h_{LG}. There is a one-to-one and almost linear relation between these points in time and the values of h_{LG}, $\dot{h}_{\mathrm{LG}} \doteq D\dot{\theta}_{\mathrm{L}} = 2-3$ km/s, where D is the LEO limb distance and $\dot{\theta}_{\mathrm{L}}$ is its orbit angle rate relative to the direction of the GPS satellite and projected into the plane of propagation. The altitude of the corresponding observed ray before passing through the sphere (shown on the left edge of the figure) is denoted by h, which is closely related to the impact parameter of the ray minus the reference. For observations taken before t_1, only a single ray is encountered, and the measured phase and amplitude of the signal can unambiguously be assigned to this singlet. The epoch t_1 marks the first contact by the observer with a triplet ray system, the rays of which have been labeled in Fig. 2-2 as "m" for the main ray, "a" for the anomalous ray, and "b" for the branching ray. Rays observed at the same epoch arrive at the same time point on the right. At the position of the observer at t_1 in Fig. 2-2(a), the newly created a and b rays are collinear, but thereafter the points of tangency of the observed a and b rays, that is, their h or impact parameter values, migrate in opposite directions: for the a ray upward and for the b ray downward. Thus, at the epoch t_1, the altitude h_{LG} of the observer is stationary with respect to the altitude h of the a,b ray system. This corresponds to dh_{LG}/dh being zero at t_1 for either the a or the b ray, and having opposite polarities on these two rays for later times as they drift apart. This results in a separation in observed bending angles and/or excess Doppler, which are both offset from the bending angle and excess Doppler associated with the main ray m, the third ray. Between epochs t_1 and t_3, the observer must deal with a triplet ray system for this case. Worse scenarios with quintuplets and higher numbers can readily be constructed. As time is shown progressing in Fig. 2-2(a), the observer finally reaches the point, at epoch t_3, where the m and a rays have merged—their respective points of tangency, or h values, have met at the boundary of the sphere; dh_{LG}/dh also is zero here. This marks the end of the line for these two rays, and from this epoch onward (for example, at t_4) one has only the singlet b ray remaining; it now becomes m, the new main ray. We show later that caustic rays occur where $dh_{\mathrm{LG}}/dh = 0$, in this figure at t_1 and at t_3. The scalar diffraction version of this scenario in Fig. 2-2(a), which is discussed later, is shown in Figs. 2-10(a) and 2-10(c) and in Figs. 2-11(c) and 2-11(d).

Figure 2-2(b) shows a schematic ray diagram for the converse scenario where the refractivity is abruptly larger within the sphere than outside. This

creates a super-refractivity zone. Here one observes a blackout (according to geometric optics) for the time interval from t_1 to t_2; the observer moves into a shadow zone at t_1 and exits it at t_2. This is accompanied by a flaring at t_2; at this point $dh_{\mathrm{LG}} / dh = 0$. Thereafter, the observer sees a doublet ray system. This doublet is a result of the condition $rn(r) - r_* n^-(r_*) \geq 0, \forall\, r \geq r_*$ being violated for tangency points nearing the boundary, a hard discontinuity in refractivity in this example, which causes the main ray to terminate abruptly at the boundary and no rays to arrive in the shadow zone. The scalar diffraction version of this scenario is discussed in Section 2.7 (see Figs. 2-11(a) and 2-11(b).

Figure 2-2(c) relaxes the discontinuity in refractivity shown in Fig. 2-2(b). It imposes a continuous refraction profile, but it invokes a large but finite gradient in refractivity between the boundary and a short distance below it. Below this level, the gradient resumes its reference profile. However, the magnitude of the gradient in the transition region is constrained to ensure that the ray existence condition $rn(r) - r_* n^-(r_*) \geq 0, \forall\, r \geq r_*$ holds for all values of the tangency point r_* near the boundary, thereby ensuring that rays exist for all values of the impact parameter. The radius of curvature of the rays here is greater than the spherical radius. This scenario results in a triplet ray system (for $t_2 \leq t \leq t_4$), although in Fig. 2-2(c) the m ray becomes severely defocused in this multipath region. Figures 2-10 (b) and 2-10(d) show a diffraction version of this scenario. This scenario also is further discussed in Section 2.8 in connection with an actual occultation taken by Global Positioning System/Meteorology (GPS/MET) (see Figs. 2-15 and 2-16). Evidently, on this particular occasion the tangency point of the ray(s) descended through a sharp positive radial gradient in electron density at the bottom of a sporadic E-layer.

Multipath with caustics has been seen in radio occultation observations of Uranus and Neptune [1,2]. When convective mixing of atmospheric gases plays a minor role, layering is a plausible result. To a first approximation, the multipath scenario shown in Fig. 2-2(c) probably occurred in the Uranus occultation displayed in Fig. 1-11. It shows an open-loop power spectrum over time, observed in 1986 from the Earth, of the Voyager 2 radio signal as the spacecraft followed an occulting trajectory behind Uranus [1]. The power spectrum over time is composed of contiguous power spectrum strips 10 s wide. Over roughly 1 hour the signal from Voyager passed through Uranus' atmosphere. During the 10- to 20-minute interval shown in Fig. 1-11, the signal in its descent into the atmosphere encountered a layer composed of gases of higher refractivity.

For an Earth limb sounder, one might expect to encounter this kind of feature upon first contact of the tangency point of the ray with a marine layer in the lower troposphere. However, for the Earth, the time interval for completion of the transient would be measured in a few seconds rather than the several

minutes shown in Fig. 1-11 for Uranus, but the Doppler spread between tones would be comparable. We note that in Fig. 1-11, just prior to and just after the epoch of first contact with the caustic, the signal power of the main ray is very low. This suggests that it passed through a narrow transition region at the boundary with a large gradient in refractivity and, consequently, is causing severe defocusing.

A closed-loop tracking receiver is unlikely to deal adequately with the kind of multipath scenarios shown here when its signal correlation and tracking-loop feedback logic are predicated on encountering only singlet tones. Even the phase and amplitude data from a receiver that tracks in the so-called "flywheel" mode will complicate significantly the recovery process unless the trifurcation and subsequent annihilation of these tones are properly accounted for or effectively eliminated by preprocessing the data. In the flywheel mode, the receiver uses its last several reliable phase measurements to extrapolate forward in time to attempt to keep the radio frequency (RF) sample window centered in Doppler space at the most likely location of the tone(s). If the receiver had been using its last good estimate of the Doppler tone from the m ray for its flywheeling forward, what does it do at the epoch t_3, which is the end of the line? To continue onward along the evolution of the b ray with time, does one account for the integer cycle difference that has accumulated between the m and b rays during the interval from t_1 to t_3? Appendix F discusses the bias in recovered refractivity that can result from missing cycles.

2.1.2 An Overview of Chapter 2

To apply a thin phase screen model for calculating signal amplitude and phase observed by a LEO, we need a few concepts from geometric optics for the actual atmosphere, which we assume here is spherical symmetric. This includes the integral expression for refractive bending angle in terms of the radial profile for the index of refraction, the Abel transform for recovering the refractivity profile, and other related quantities. Geometric optics is used to obtain the observed change in the refractive bending angle and phase delay as a function of impact parameter due to any one of the discontinuity scenarios described earlier. Then, we introduce two thin phase screen models and discuss their use and their correspondence to the actual atmosphere. Multipath, caustics, and shadow zones are discussed in the framework of a thin-screen model, first in a geometric optics context and later in a wave-theory context. Next, to account for diffraction effects in carrier phase and amplitude, we introduce into the thin-screen model a particular scalar diffraction theory in the form of the Rayleigh–Sommerfeld integral. To numerically evaluate this diffraction integral, we introduce the stationary-phase concept and its role, not so much as a stand-alone computation technique in terms of Fresnel integrals, but as an aid for other computation techniques, and for assigning the appropriate phase

profile to the thin-screen model. Numerical results for Fresnel diffraction from a discontinuous refractivity profile then are presented. Also presented are the diffraction effects from a continuous refractivity profile but with a discontinuous lapse rate.

We then use thin-screen/scalar diffraction techniques developed here to treat certain caustic, multipath, and shadow effects that can be observed by the LEO as the ray path tangency point crosses an ionosphere layer; for examples, see Figs. 2-15 and 2-16 in Section 2.8. We discuss the errors in the retrieved refractivity and temperature profile that can result from use of a straight Abel transform recovery algorithm without accounting for the deficiency in ray theory for this case. Finally, in later sections, we briefly discuss the potential of the Fresnel transform technique for enhanced resolution in thin atmosphere conditions and sketch how it might be used in multipath situations. Use of a scalar diffraction technique to sharpen the resolution of localized features is briefly covered.

2.2 Geometric Optics in a Spherical Medium

To use a thin-screen model to calculate these effects, we need some concepts from geometric optics applicable to the actual atmosphere. We need to express the refractive bending angle as a function of the index of refraction, and we need a few other related concepts.

The refractive bending angle observed by a LEO during an occultation is obtained directly from the Doppler observations made by the LEO and the precision orbit determination (POD) information about the satellites. Appendix A, see Figs. A-2 and A-3 and Eqs. (A-8) through (A-18), derives the relationship between the excess Doppler and the atmosphere-induced refractive bending angle α of a ray from the occulting GPS satellite. The excess Doppler is the observed Doppler minus the Doppler due to the relative motion between the LEO and the observed GPS satellite. Both the case where the GPS satellite is located at a finite distance from the Earth's limb and the limiting case where it is infinitely afar are given. The latter case[1] yields a particularly simple near-linear relationship between α and the excess Doppler f_D, which is given by

$$\lambda f_D = V_\perp \alpha + \mathrm{O}[\alpha^2] \tag{2.2-1}$$

[1] Placing the occulted GPS satellite at infinity is adequate for our purpose, but not for treating actual data. For the finite case, the bending angle equals the sum of two deflection angles, δ_{G} and δ_{L}. δ_{G} is the deflection angle of the ray asymptote at the GPS satellite (see Fig. A-3) relative to the straight line passing through the positions of the GPS and LEO satellites; δ_{L} is the deflection angle at the LEO. In practice δ_{L} is roughly an order of magnitude larger than δ_{G}. These deflection angles are mutually constrained by ray tracing or by Bouguer's law when spherical symmetry applies.

Here λ is the wavelength of the carrier of the observed signal, and $V_\perp$ is the component of the cross-velocity of the straight-line path between the GPS and LEO satellites lying in the propagation plane, the plane containing the GPS satellite, the LEO, and the ray. In Figs. 2-2 and 2-3, $V_\perp = -dh_{LG}/dt$ for a setting occultation. When the LEO orbital altitude is about 10 percent of the Earth's radius, $|V_\perp| \cong 2.5$ km/s $\pm 20\%$. The 20 percent spread results from the typical range of obliquity values between the propagation and orbit planes in the set of occultations used for data analysis. $V_\perp$ is essentially constant over the relatively short duration of an occultation. For the Earth's atmosphere at sea level, α is about 20 mrad for the dry air component; thus, the range of f_D is a few hundred hertz for dry air. Water vapor in the lower troposphere can double or triple this range.

Geometric optics as a ray theory can be developed from different approaches. One approach is to start from a wave representation based on Maxwell's equations for a harmonic wave in a refracting medium, and then take its limiting form as the wavelength of the wave is made to approach zero. The general properties of rays, such as ray path curvature as a function of refractive gradient, reflection coefficients, and so on, can be deduced. An account of this approach is given in [3]. Another approach is to start from Fermat's stationary-phase principle for rays in general and apply the Calculus of Variations to the path integral for the phase delay along the ray. Each of the rays in Fig. 2-3 satisfies Fermat's principle, which requires that each must be a path of stationary phase. In other words, the total phase accumulation or delay along the path from beginning to end would undergo only a second-order variation as

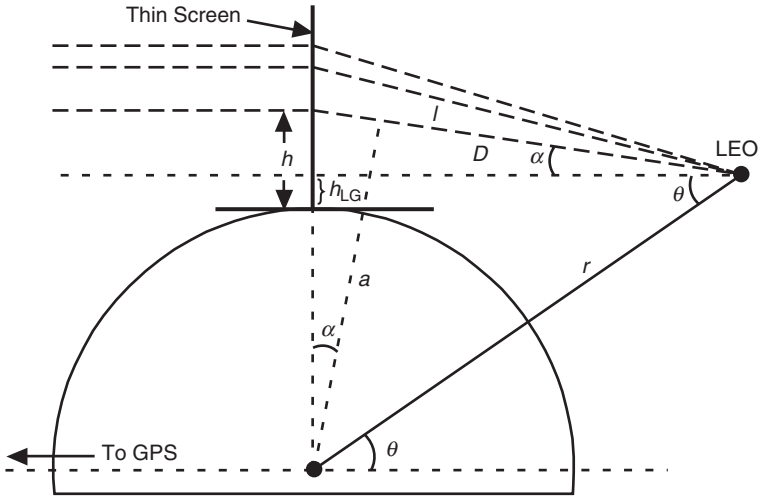

Fig. 2-3. Thin-screen geometry for a LEO. The figure shows multiple rays, with different bending angle and impact parameter values, all converging at the LEO.

a function of a nearby trial deviation of the path from the actual path, while still satisfying the boundary conditions. Usually the stationary value of the phase delay along a ray is a local minimum. However, for the scenarios shown in Fig. 2-2 the a ray, unlike the m and b rays, is a path that provides a local maximum in the phase delay. We know from the Calculus of Variations that this can happen when a ray at some interior point along its path comes in contact with the envelope associated with the family of rays whose members are generated by varying some parameter, such as the impact parameter of a ray. In optics this envelope is known as the caustic, or burning curve. Appendix A provides a brief introduction to ray theory using the stationary-phase concept. Appendix B discusses the properties of a caustic.

From Eqs. A-1 through A-5 in Appendix A, it follows that the bending angle $\alpha(r_*)$ and the impact parameter a in a locally spherical symmetric atmosphere are given by

$$\alpha(r_*) = -2a\int_{r_*}^{\infty} \frac{1}{n}\frac{dn}{dr}\frac{1}{\sqrt{n^2r^2 - a^2}}dr \tag{2.2-2a}$$

$$a = n(r_*)r_* = n(r)r\sin\gamma = \text{constant} \tag{2.2-2b}$$

where γ is the angle between the radius vector and the tangent vector of the ray. When we place the GPS satellite at an infinite distance, then $\gamma = \theta + \alpha$. The quantity r_* is the radial distance of the turning point of the ray or its point of tangency with the Earth's limb (Fig. 2-1). The quantity $n(r) = 1 + N(r)$ is the index of refraction, and $N(r)$ is the refractivity. (For convenience, we define $N = n - 1$, not $N = 10^6(n-1)$; the latter is the customary definition of refractivity.) The quantity a in Eq. (2.2-2b) is the so-called impact parameter, and it is a ray-specific quantity. When spherical symmetry applies, it has a constant value when evaluated at any point (r,θ) on a specific ray. The relationship given in Eq. (2.2-2b) for a is known as Bouguer's law. This is analogous to the conservation of angular momentum in a central force field. Bouguer's law is the spherical equivalent of Snell's law.

For the spherical symmetry assumed here, and when at a specific observational epoch there is a unique ray producing the excess Doppler, then Eqs. (2.2-1) and (2.2-2) show that the Doppler observations between the LEO and occulted GPS satellite, plus the POD information about the satellites, yield both the bending angle α for the ray and its impact parameter a. Hence, one can form a data sequence $(\alpha_\kappa, a_\kappa)$, $\kappa = 1,2,\cdots,M$, from the excess Doppler sequence obtained from the M observations made over the occultation episode. This sequence $(\alpha_\kappa, a_\kappa)$ describes the evolution over time of the bending angle and impact parameter for a specific family of locally unique rays that are

generated by the orbital motions of the satellites. When the geometry does not admit spherical symmetry, then the equations and procedures for recovering bending angle α and ray path tangency point r_* involve ray-tracing techniques, but they lead to the same general result, a data sequence $(\alpha_\kappa, r_{*_\kappa})$. We avoid this complication here.

When spherical symmetry does apply, and there is a unique ray from the GPS satellite to the LEO at each epoch in the observational sequence, one can invoke the Abel transform to recover the radial profile of the refractivity from the sequence $(\alpha_\kappa, a_\kappa)$. This is given by (see Appendix A)

$$N(a) \doteq \log n(a) = \frac{1}{\pi}\int_a^\infty \frac{\alpha(\xi)}{\sqrt{\xi^2 - a^2}}\, d\xi \tag{2.2-3}$$

Here $\alpha(\xi)$ is the functional form of the bending angle versus impact parameter ξ obtained from the data sequence $(\alpha_\kappa, a_\kappa)$, $\kappa = 1, 2, \cdots, M$. Implicit in the form for the Abel transform given in Eq. (2.2-3) is the vanishing of the bending angle for large values of the impact parameter, that is, $\alpha(\xi) \to 0$ as $\xi \to \infty$.

The gradient of $n(r)$ in Eq. (2.2-2a) is given by dn/dr. But, note that this integral readily allows a change of variable. For example, we can define another variable $\rho = kn(r)r$; Eq. (2.2-2) becomes

$$\alpha(\rho_*) = -2\rho_* \int_{\rho_*}^\infty \frac{1}{n}\frac{dn}{d\rho}\frac{1}{\sqrt{\rho^2 - \rho_*^2}}\, d\rho, \quad \rho_* \equiv ka = \rho \sin\gamma \tag{2.2-2′}$$

The functional form for the bending angle is transformed from $\alpha(r_*)$ to $\alpha(\rho_*)$, which is acceptable provided we know or can recover the functional form for $n(\rho)$, and provided no super-refractive zones exist. The integrals in Eqs. (2.2-2′) and (2.2-3) form an Abel transform pair.

A problem arises when multiple rays from the same GPS satellite meet at the LEO, which is shown in Fig. 2-3. In this case, the signal arriving at the LEO is a composite of these multiple rays. The superposition of these rays causes interference in the phase and amplitude of the signal. Spectral techniques are useful for unraveling the excess Doppler frequency and amplitude for each interfering ray. We next address multipath in a ray theoretic context.

2.2.1 Multipath, Shadow Zones, and Caustics According to Geometric Optics

A polarity change in the gradient of the bending angle is an almost sure sign that multipath and/or shadow effects will be seen, particularly if the observer is far from the location of these reversals in the gradient. Diffraction effects also are likely to be present in at least part of the observations if the

change is sharp enough. Figure 2-3 provides the basic geometry for a particular thin-screen model, which we discuss later. Here h is the altitude (relative to an arbitrary but defined reference altitude R, say, sea level) of a point in a thin screen through which the ray from the occulted GPS satellite passes on its way to the LEO. The screen in this figure is oriented perpendicularly to the straight line between the LEO and the GPS satellite; h_{LG} is the altitude in the thin screen of the intersection of the LEO–GPS line. It varies nearly linearly with time during an occultation episode, about 1 minute in duration for the neutral atmosphere. Figure 2-3 shows a collimated beam from the GPS satellite to the left of the screen, which is tantamount to setting the GPS distance D_{G} from the Earth's limb to infinity. It is appropriate in this case instead to use for LEO limb distance the quantity D, which is the reduced distance between the LEO and the thin screen minus a small quantity $a\alpha$. It is given by

$$D^{-1} = D_{\mathrm{L}}^{-1} + D_{\mathrm{G}}^{-1} \tag{2.2-4}$$

where D_{L} is essentially the distance of the LEO from the limb of the Earth, and similarly for D_{G}. In calculating the phase at the LEO, use of the reduced distance D instead of D_{L} accounts for the extra phase from the wave front curvature resulting from the finite distance D_{G} of the GPS satellite from the Earth's limb. D_{G} is about 4 Earth radii, and D_{L} is about 1/2 an Earth radius; therefore, D is about 10 percent smaller than D_{L}.

From the geometry in Fig. 2-3, we see that h denotes the height at which the asymptote of the ray intersects the thin screen. From Bouguer's law in Eq. (2.2-2), we have the relationship for this height h in the thin screen:

$$\left.\begin{aligned} & h + R = a\sec\alpha = r\sin(\theta+\alpha)\sec\alpha = \\ & r(\sin\theta + \cos\theta\tan\alpha) = h_{\mathrm{LG}} + R + r\cos\theta\tan\alpha \\ & \text{or} \\ & h = h_{\mathrm{LG}} + D\alpha(h) + \mathrm{O}\left[\alpha^2\right],\ \ D = r\cos(\theta+\alpha) \end{aligned}\right\} \tag{2.2-5}$$

Here the bending angle has been expressed as a function of thin-screen altitude $\alpha(h)$. We show later that when $d(\alpha^2)/da$ is less than a critical value, there is a one-to-one relationship between h and the impact parameter a. In this case, it is straightforward to transfer between $\alpha(h)$ and $\alpha(a)$.

Figure 2-3 shows an example where three altitudes result in bending angles so that the rays from these particular tangent points with the Earth's limb all simultaneously converge at the LEO. Figure 2-4 shows a hypothetical curve for h_{LG} versus h in the neighborhood of this multipath zone that might correspond to the multiple rays shown in Fig. 2-3. This depicts a classic multipath situation.

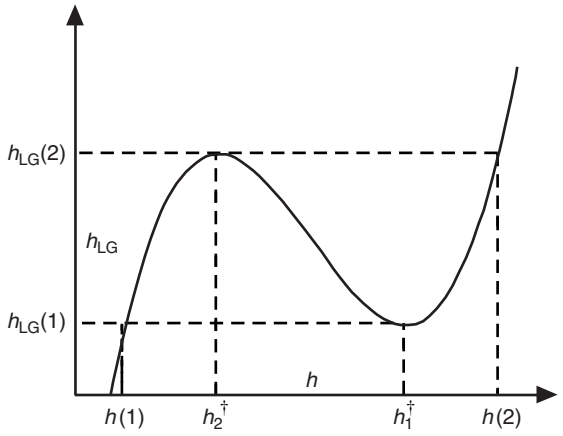

Fig. 2-4. Hypothetical curve showing thin-screen altitudes h versus h_{LG} in the neighborhood of a multipath zone. Multipath occurs within the range $h_{LG}(1) \le h_{LG} \le h_{LG}(2)$, $h(1) \le h \le h(2)$, and $h_1^\dagger$ and $h_2^\dagger$ mark the caustic contact altitudes, where $dh_{LG} / dh = 0$.

As discussed earlier in connection with Fig. 2-2, the condition $dh_{LG} / dh = 0$ will be found to hold at two or more altitudes in the neighborhood defined by the turning-point altitudes for these three rays. These altitudes mark the contact points with the envelope surface to the ray family, the vertical continuum of ray paths generated by the evolution of h_{LG} over time. In the immediate neighborhood about a point $\left(h^\dagger, h_{LG}^\dagger\right)$ where $dh_{LG} / dh = 0$, h_{LG} has a locally quadratic (or higher) dependence on h. This results in two rays from altitudes slightly above and below $h^\dagger$ that concurrently meet at the LEO when h_{LG} is on the concave side of the h_{LG} versus h curve, and none from these two ray families when h_{LG} is on the convex side. This is depicted in Fig. 2-4 at each of the two caustic contact points. This is multipath in its simplest form.

Taking the reciprocal of dh_{LG} / dh, we have $dh / dh_{LG} \to \infty$ at $\left(h^\dagger, h_{LG}^\dagger\right)$; so, here there is a singularity in the density of ray paths arriving at the LEO. Caustics occur near here. Therefore, whenever the gradient of the bending angle $d\alpha / dh$, which is nominally negative for the Earth's atmosphere, takes the value $d\alpha / dh = D^{-1}$, about 0.3 mrad/km, then dh_{LG} / dh will be zero.

In Appendix A it is shown that the defocusing function for the atmosphere is given by

$$\frac{1}{\zeta} \doteq 1 - D\frac{d\alpha}{da} \tag{2.2-6}$$

The defocusing function provides a measure of the dispersal of the rays caused by the refractive gradient. Therefore, the defocusing function gives the relative amplitude of a ray at the LEO as a result of the ray having passed through the atmosphere compared to the amplitude that would have been obtained without the atmosphere. From Eq. (2.2-5) we see that in the screen $\zeta \doteq dh / dh_{\mathrm{LG}}$.

Appendix A also derives the vertical radius of the first Fresnel zone $\mathscr{F}(h)$, which is given by

$$\mathscr{F} = \sqrt{\lambda D \,|\zeta|} \tag{2.2-7}$$

The condition $\zeta^{-1} = 0$ marks the first contact with a caustic in geometric optics; rainbows[2] occur at such points. It also marks the first (or last) encounter of the LEO with the multiple-ray system. This condition also corresponds to the vertical diameter of the first Fresnel zone growing infinite. Therefore, geometric optics predicts an infinite amplitude at this point, but no such event is observed there, partially because of the failure of geometric optics in the neighborhood of this first contact point. The first caustic contact point corresponds to a point where not only is the phase delay along each of the multiple rays stationary with respect to a path variation, for example, the impact parameter of the path, but its second-order variation also is zero.

Geometric optics is based on a second-order stationary-phase theory. A non-zero second-order variation is required in this theory. A third-order theory is needed to handle caustics. Here, the use of the words "ray" or "path of stationary phase" implicitly assumes that a geometric optics treatment is valid. Thus, necessary conditions for validity of a geometric optics treatment are that the rays exist (see Footnote 3). Also, the altitude differences of the multipath rays at their tangency points should significantly exceed the diameter of the first Fresnel zone, and caustics are to be avoided.

A necessary condition for the existence of a caustic is that the defocusing factor $\zeta \to \infty$ at some altitude. A caustic surface will be generated by the

---

[2] The rainbow is a caustic phenomenon. The altitude h of the first ray making caustic contact in Figure 2-2(b) (at t_4) corresponds for a raindrop to the value of the impact parameter of the incident ray from the Sun that gives rise to a rainbow. The scattering angle of a ray, after being refracted upon entering the raindrop, internally reflected, and refracted again upon exiting, depends on the impact parameter of the ray. This scattering angle becomes stationary when the impact parameter is located about 85% (for the primary rainbow) of the raindrop radius out from the center. The impact parameters of the Sun's rays are uniformly distributed prior to hitting the raindrop. But the exiting rays with impact parameter values in the nearby neighborhood about this 85% value "pile up" at the stationary scattering angle, about 138 deg. Therefore, the amplitudes of these rays add up constructively because they all have incurred essentially the same phase accumulation at the observer. A "caustic" ray is the result.

continuum of ray paths whose points of tangency lie in some neighborhood around this critical altitude. When a caustic does occur, multipath situations also will arise where two or more ray paths arrive at the LEO from different altitudes. Appendix B gives a short discussion of caustics in a geometric optics context.

Caustics also are predicted in a wave theory framework, but there their form is softened and, although brightening is predicted, the infinite amplitude predicted by (second-order) geometric optics does not appear. Caustics in a wave theory framework are discussed in Chapters 3 and 5.

2.2.2 Thin-Atmosphere Conditions

To calculate the effects of the discontinuities related to refractivity, we use certain approximations that are applicable to a thin atmosphere. By a "thin atmosphere" we mean that the atmosphere is thin geometrically and optically. Specifically, the characteristic length of the ray path through the atmosphere $\bar{L}$ defined by

$$\bar{L} = \frac{2\int_0^\infty sN[r(s)]ds}{\int_0^\infty N[r(s)]ds} \doteq 2\sqrt{\frac{2r_o H_{p_o}}{\pi(1-\beta)}} \tag{2.2-8}$$

satisfies the condition that $\bar{L}/r_o << 1$. Here s is arc length along the ray path measured from the point of tangency. The second expression on the right-hand side (RHS) of Eq. (2.2-8) applies to an atmosphere with an exponentially decreasing refractivity profile and with a scale height H_{p_o}. See Appendix A for more detail.

The parameter β is defined by

$$\beta = -\frac{r}{n}\frac{dn}{dr} \doteq \frac{rN(r)}{H_p} \tag{2.2-9}$$

where the second expression applies to an exponential model for refractivity. This parameter β is essentially the ratio of the impact parameter of a ray to its radius of curvature. It must satisfy the condition $\beta << 1$ or, equivalently, $d(rn)/dr >> 0$. In effect, the radius of curvature of the ray at any point must be much greater than the local radius of curvature of the "iso-refractivity" or equipotential surface there. Super-refractivity situations where $\beta > 1$ over

substantive[3] path lengths are specifically excluded here for the reference refractivity profile. In the case of the Earth, we have the happy circumstance that for dry air $\bar{L} / r_o \approx 0.05$. At the Earth's surface the dry air component of β is about 0.2, and at the tropopause it typically is around 0.05. On the other hand, across a sharp marine layer boundary, β can exceed unity.

Equation (2.2-8) accounts for first-order ray path curvature effects through the term $(1-\beta)^{1/2}$ in the denominator. One can calculate ray path curvature effects once the refractivity profile is specified [4]. If s is arc length along the ray from its tangency point, then to first order the radial coordinate of a point on the ray is given by

$$r \doteq r_* + (1-\beta)\frac{s^2}{2r_*} \tag{2.2-10}$$

With $\beta = 0$, this gives the straight-line relationship between chord length and radial coordinate. With $\beta > 0$, the ray is bent radially inward. Both Eqs. (2.2-8) and (2.2-10) show that the thin atmosphere test fails when β nears or exceeds unity.

Atmospheric refractivity for the Earth may be classified in increasing difficulty as (1) locally spherical symmetric without multipath at the LEO, (2) locally spherical symmetric with multipath at the LEO, (3) not locally spherical symmetric because of horizontal variations, and (4) time variable through turbulence, advection, and so on. Our reference refractivity profile will satisfy the thin atmosphere conditions, and it will be spherical symmetric without multipath. We will explicitly avoid locally harsh refractivity conditions, such as critical refractivity conditions found at times in the lower troposphere, other than the discontinuities under study, which actually do include super-refractivity cases. Our discontinuities or sharp transitions are assumed to be sufficiently mild that we need not account via the Fresnel formulas for reflections from the boundary, nor for Love or Rayleigh-like waves propagating along or near the boundary surface, nor for evanescent waves. We also assume that our reference atmosphere is temporally smooth and

---

[3] The validity of Eq. (2.2-2) requires that $rn(r) \geq a$ along the ray path. This condition is required to obtain real solutions to the Euler differential equation (see Appendix A), which all ray paths must satisfy in a spherical symmetric medium. This condition $rn(r) \geq a$ must hold at all points along the ray path to obtain a stationary-phase path satisfying the boundary conditions and passing through the turning point at r_*. In other words, if this condition is not satisfied, then there is no ray with the impact parameter value a satisfying the boundary conditions. This global condition translates into a local necessary condition (but not sufficient; see Section 6.4.2) that requires that $\beta < 1$ in some neighborhood about a turning point.

does not cause appreciable scintillation (although scintillation is almost always present to some degree in sounding data). Even a cursory review of wave and/or ray theory literature in such disparate fields as electrodynamics and seismology will reveal that propagation across a boundary is potentially a very complicated problem without simplifying assumptions, such as those in the foregoing discussion.

We note from Eq. (2.2-8) that $\bar{L}$ is about 400 km for dry air, but the along-track resolution δL is closer to

$$\delta L = 2\sqrt{\frac{2r_o\delta r}{1-\beta}} \approx 225\sqrt{\frac{\delta r}{1-\beta}} \text{ km} \tag{2.2-11}$$

where δr is the vertical resolution. This is the path length in a shell of thickness δr, essentially the chord length corrected to first order for ray path curvature $(1-\beta)^{-1/2}$. Using the Fresnel vertical diameter for δr, we obtain about 180 km at sea level for dry air and 280 km in the stratosphere. Because both the defocusing factor ζ and the ray path curvature parameter β depend strongly on the refractivity and its gradient, δL for the water-vapor-laden lower troposphere can vary widely. Near-super-refractivity situations where $\beta \to 1$ render the above definition of δL less meaningful.

2.3 Thin Phase Screen Models

Diffraction effects occur to a varying degree whenever a sharp change in refractivity or in one of its derivatives occurs. Sommerfeld defined the word "diffraction" more than a century ago to cover those effects not predicted by ray theory. Because diffraction effects for a spherical geometry are difficult to calculate, one often attempts to substitute a simpler model that yields basic ray theory results plus diffraction effects, and which has acceptable agreement with rigorous wave theory results. In a thin-screen model, one substitutes for the atmosphere a thin-screen proxy, nominally transparent, through which any passing ray experiences a position-dependent phase delay upon emerging from the screen. The relationship between position in a thin screen and altitude of the turning point of the actual ray in the atmosphere depends on the choice of screen and on the index of refraction profile. The thin-screen model is a surrogate for the transmission effects of the real atmosphere on the traversing GPS signal. One chooses a phase profile for the thin screen that attempts to match the LEO-observed phase delay profile. This is accomplished by assigning a phase delay, embedded in the thin screen at a given thin-screen position, that corresponds to the actual atmosphere-induced phase delay predicted by geometric optics. If a certain thin atmosphere condition holds, the assigned phase delay values in the screen can be made to be a single-valued function of the thin-screen position. In this case, the assignment leads to the

proper total refractive bending angle $\alpha(a)$ when a path of stationary phase is followed from the GPS satellite along the incoming ray asymptote through the thin screen and along the outgoing ray asymptote to the LEO. When this thin atmosphere condition is satisfied, the profile of this phase delay in the thin-screen yields the LEO-observed phase and Doppler of the primary rays (excluding, for example, reflections) predicted by geometric optics for the actual atmosphere, including multipath and shadow regions where appropriate. This thin-screen model then is used to calculate the observed diffraction effects using a scalar diffraction theory such as the Rayleigh–Sommerfeld integral [5].

The thin-screen model has been discussed in many references over the years, where it has been applied to scintillation and diffraction studies [6–11]. For phase scintillation applications, [9] also discusses the wave propagation coherency conditions that an atmosphere with a finite scale height must satisfy in order for the thin-screen model to remain valid. These coherency conditions can be related in part to the "thin atmosphere" conditions given earlier in Section 2.2, namely, that β and $\bar{L} / r_o$ should be small.

Chapter 1 points out that the thin-screen concept can be extended to include multiple thin screens in tandem; see, for example, [10,11]. Here one replaces propagation through the actual atmosphere with field values holding on m thin screens, which are separated along the axis of propagation by some assigned distance D_k, $k = 1,2,\cdots, m$. The propagation of the wave between successive screens is accomplished using the Fresnel–Kirchoff scalar diffraction integral that applies asymptotically when the wavelength of the propagated wave becomes a very small fraction of the scale of the screens. The actual inhomogeneous medium in between screens is replaced by a homogeneous medium, but the phase accumulation that would have resulted is lumped into an extra position-dependent phase increment that is added to the wave at each screen. In this way the field values for amplitude and phase on one plane are mapped to the field values on the next plane, and so on. This approach has been shown to work well in a "forward" propagation process through an inhomogeneous medium.

2.3.1 The Helmholtz–Kirchoff Integral Theorem

This theorem provides the basis for scalar diffraction theory [3,12]. It may be used in either a "backward" or "forward" propagation mode, i.e., toward an emitter or away from it. In the backward mode, the scalar diffraction integral is used to map the phase and amplitude sequences measured by the LEO through the vacuum to an equivalent surface much closer to the Earth's limb. Appendix A derives both the forward and backward modes from the Helmholtz–Kirchoff theorem. From Fig. 2-2 it is clear that, if we could move the surface containing the LEO and its vertical motion to another surface much closer to the limb, fewer rays would cross each other between the limb and the

closer surface. A virtual LEO moving in the closer surface would experience reduced multipath (see Fig. 1-13). The refractivity profile of the real atmosphere is recovered from the phase and amplitude profiles mapped from the LEO trajectory to the new closer surface. The so-called "back plane" propagation method is one example of this technique [13–15]. Here the chosen surface for convenience is planar, actually a line in coplanar propagation, and it is oriented perpendicular to the LEO–GPS line, or perhaps canted slightly. In this closer "plane" with hopefully far fewer multipath episodes, one derives bending-angle and impact parameter profiles. From Eq. (2.2-1) it follows that $d\varphi_B = k\alpha ds_B$, where φ_B is the mapped phase in the back plane mounted perpendicular to the LEO–GPS line, and ds_B is an incremental arc length along the plane. Therefore, the gradient of the mapped phase along this plane yields the bending angle in that back plane, from Bouguer's law the impact parameter. The Abel transform then yields the refractivity profile.

2.3.2 The Space Curve for Impact Parameter

There is nothing in the Helmholtz–Kirchoff integral theorem (see Appendix A), the basis for scalar diffraction theory, that requires planarity for the phase screen; it could be a curved surface. Or, if it is planar, it need not be mounted perpendicular to the LEO–GPS line; these are matters of convenience. We could as well adopt as a canting angle $\alpha_m = \alpha(a)$ for the clockwise rotation. This is suggested in [14]. In this case, the "height" h in Eq. (2.2-5) becomes $h = a - R$. It is strictly the impact parameter minus the reference radius, and it becomes an arc length along a curve defined by the intersection of the plane of propagation with the impact parameter surface. This space curve for the impact parameter in general is not straight. Since there is a one-to-one relationship between impact parameter and bending angle when spherical symmetry holds, then we are assured that $h(\alpha)$ will be monotonic. Of course, we don't know a priori the value of $\alpha_m = \alpha(a)$ and, therefore, its shape without first processing the observations.

We show one way to recover the impact parameter curve in space coordinates from the LEO phase and amplitude observations. To simplify the geometry here we assume that the occulted GPS satellite is infinitely far in the $\theta = \pi$ direction; therefore, the angle γ in Bouguer's law in Eq. (2.2-2) between the radius vector and the tangent vector of the ray becomes $\gamma = \theta + \alpha$. The finite case is a straightforward extension of the following, but it requires more care in handling the relative motion of the satellites during the occultation.

The phase delay φ along the ray $\mathscr{R}$ from its turning point at (r_*, θ_*) to a point (r, θ) can be obtained from its defining integral:

$$\varphi = k\int_{\mathscr{R}} nds = k\int_{r_*}^{r} n\sqrt{1+(r\theta')^2}\,dr = k\int_{r_*}^{r}\frac{n^2 r}{\sqrt{n^2r^2-a^2}}dr \qquad (2.3\text{-}1)$$

where in the last integral we have used Bouguer's law in Eq. (2.2-2) to rewrite $d\theta/dr$ in terms of the impact parameter. We assume that the point (r,θ) is out of the atmosphere where $n \equiv 1$. By integrating by parts and using Eq. (2.2-2), we obtain for the phase at (r,θ)

$$\left.\begin{aligned}\varphi &= k\left(r\cos(\theta+\alpha)+a\alpha+\int_a^\infty \alpha(\rho)d\rho\right),\\ a &= r\sin(\theta+\alpha),\ \ \alpha=\alpha(a)\end{aligned}\right\} \qquad (2.3\text{-}2)$$

Setting $\alpha \equiv 0$, $\varphi \to kr\cos\theta$, which is the cumulative phase along a planar wave traveling in the $\hat{z}$ direction (see Fig. 2-5) from the line defined by $\theta = \pi/2$. Therefore, $\varphi - kr\cos\theta$ is the excess phase caused by the refractivity. If we hold r and θ fixed, then we note from Eq. (2.3-2) that

$$\frac{\partial\varphi}{\partial a} = k\left(-a\frac{d\alpha}{da}+\alpha+a\frac{d\alpha}{da}-\alpha\right)\equiv 0 \qquad (2.3\text{-}3)$$

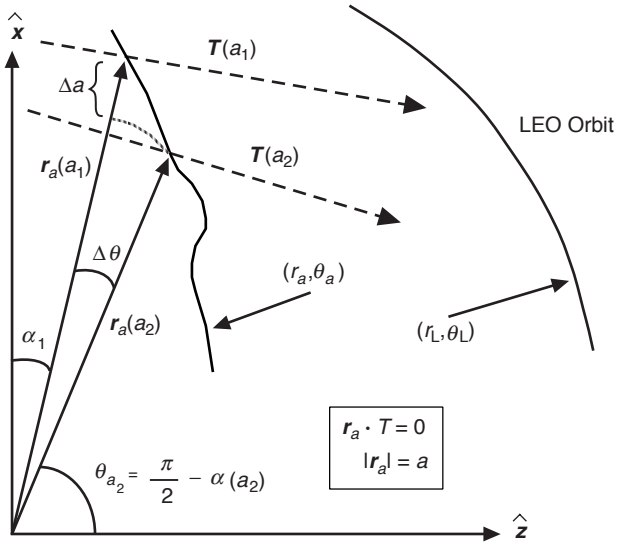

Fig. 2-5. Schematic of the impact parameter space curve (r_a, θ_a), which is traced out by the tip of vector $\boldsymbol{r}_a(a)$ as a is varied. Bouguer's law requires that $|\boldsymbol{r}_a| = a$ and that $\boldsymbol{r}_a \cdot \boldsymbol{T} = 0$, where $\boldsymbol{T}$ is the unit tangent vector of the ray passing through the end point of $\boldsymbol{r}_a(a)$.

which merely underscores the stationary property of the phase on a ray, and the difficulty in using phase directly in ray theory to recover the impact parameter of the ray.

If we vary r and θ in Eq. (2.3-2), holding a fixed, we obtain

$$d\varphi = k\left(-ad\theta + \cos(\theta+\alpha)dr\right) \tag{2.3-4}$$

We now set $r = a$ and $\theta = \theta_a = \pi/2 - \alpha(a)$. By doing so, we have positioned our point (r,θ) at the tip of the impact parameter vector $\boldsymbol{r}_a(a)$. This is indicated in Fig. 2-5, which shows the geometry (but definitely not to scale) of the impact parameter space curve (r_a,θ_a) with $r_a = a$. The form for $\boldsymbol{r}_a(a)$ in terms of unit Cartesian vectors $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{z}}$ is given by

$$\boldsymbol{r}_a(a) = a\left(\hat{\boldsymbol{z}}\sin\alpha(a) + \hat{\boldsymbol{x}}\cos\alpha(a)\right) \tag{2.3-5}$$

Bouguer's law requires that $|\boldsymbol{r}_a(a)| = r_a = a$ and that $\boldsymbol{r}_a(a)$ is perpendicular to $\boldsymbol{T}(a) = \hat{\boldsymbol{z}}\cos\alpha(a) - \hat{\boldsymbol{x}}\sin\alpha(a)$, the unit tangent vector of the ray passing through the tip of $\boldsymbol{r}_a(a)$. At this point (a,θ_a), a small change in position alters the phase by an amount $d\varphi$. According to Eq. (2.3-4), to first order this is given by

$$d\varphi\big|_{\theta=\pi/2-\alpha} = -akd\theta + 0dr \tag{2.3-6}$$

There is no first-order variation in phase in the radial direction because the ray through the tip of $\boldsymbol{r}_a(a)$ is perpendicular to $\boldsymbol{r}_a(a)$.

This impact parameter vector $\boldsymbol{r}_a(a)$ traces out the impact parameter space curve (r_a,θ_a), as indicated by the schematic example in Fig. 2-5. The incremental arc length vector $d\boldsymbol{s}_a$ along this space curve and the angle ψ_a between $d\boldsymbol{s}_a$ and $\boldsymbol{r}_a(a)$ are given by

$$\left.\begin{aligned} d\boldsymbol{s}_a &= \frac{d\boldsymbol{r}_a}{da}da = \left(\hat{\boldsymbol{r}}_a - \hat{\boldsymbol{\theta}}_a a\alpha'\right)da,\ \ \boldsymbol{r}_a = \hat{\boldsymbol{r}}_a a \\ \cos\psi_a &= \boldsymbol{r}_a \cdot \frac{d\boldsymbol{s}_a}{ds_a} = \frac{1}{\sqrt{1+(a\alpha')^2}},\ \ \sin\psi_a = \frac{-a\alpha'}{\sqrt{1+(a\alpha')^2}} \end{aligned}\right\} \tag{2.3-7}$$

where $\alpha' = d\alpha/da$; also, $\hat{\boldsymbol{r}}_a$ and $\hat{\boldsymbol{\theta}}_a$ are unit orthogonal vectors in the rotating frame. It follows that, except at super-refractivity points, the impact parameter values are denumerated uniquely along the curve traced out by $\boldsymbol{r}_a(a)$ as a is varied. With spherical symmetry, a unique relationship holds between a and α, except at a super-refractivity point, which is discussed later.

To generate the impact parameter curve, we start from a known point on the curve where no multipath exists and, assuming that it is a high point, we

continue downward. Let $\alpha(a_1,\theta_1)$, $\theta_1 = \pi/2 - \alpha(a_1)$ denote such a known point, as shown in Fig. 2-5. We assume that, in addition to knowing unambiguously the value of α at (a_1,θ_1), we also know the value of α'. The back propagation diffraction integral, for coplanar propagating in a vacuum the phase and amplitude measurements made by the LEO to another point at $\boldsymbol{r}_1$ (see Eq. (A-22) in Appendix A), is given by

$$E(\boldsymbol{r}_1) = \sqrt{\frac{i}{\lambda}} \int_C \left(\frac{E(\boldsymbol{r}_\mathrm{L})}{r_{1\mathrm{L}}^{1/2}} \exp(-ikr_{1\mathrm{L}}) \left(\hat{\boldsymbol{r}}_{1\mathrm{L}} \cdot \hat{\boldsymbol{n}}(\boldsymbol{r}_\mathrm{L}) \right) \right) ds_\mathrm{L} \tag{2.3-8}$$

where $r_{1\mathrm{L}} = |\boldsymbol{r}_1 - \boldsymbol{r}_\mathrm{L}|$, $\boldsymbol{r}_\mathrm{L}$ denotes a point at the LEO, ds_L denotes an incremental arc length along the curve C defined by the trajectory of the LEO over which observations were made during the occultation, and $\hat{\boldsymbol{n}}(\boldsymbol{r}_\mathrm{L})$ is the outward unit normal vector to the curve C. The intervening medium between points $\boldsymbol{r}_\mathrm{L}$ and $\boldsymbol{r}_1$ is taken as a vacuum. We apply this path integral using the LEO observations to obtain the mapped phase and amplitude at the impact parameter point $\boldsymbol{r}_1 = \boldsymbol{r}_a(a_1)$. (There are several important details required in practice related to establishing integration limits, preserving phase coherency, i.e., $|\delta r_{1L}| << \lambda$, where $\delta r_{1\mathrm{L}}$ is the error in $r_{1\mathrm{L}}$, over the span of the integral by accounting for the relative satellite motion, and maintaining phase connection. We omit those details here [13–15].) Even though this point $\boldsymbol{r}_1 = \boldsymbol{r}_a(a_1)$ is well inside the refracting atmosphere, we treat it as a proxy curve in a vacuum, just like a thin screen. Let the phase of the mapped field $E(\boldsymbol{r}_a(a_1))$ from the diffraction integral in Eq. (2.3-8) at this impact parameter point be designated as $\hat{\varphi}_1 = \hat{\varphi}(a_1)$. Next, we alter $\boldsymbol{r}_a(a_1)$ by a small amount to $\boldsymbol{r}_a(a_2)$ with $a_2 = a_1 + \Delta a$, as shown in Fig. 2-5 with $\Delta a < 0$. It follows that a first approximation for $\boldsymbol{r}_a(a_2)$ is given by

$$\boldsymbol{r}_a(a_2) \doteq \boldsymbol{r}_a(a_1) + \hat{\boldsymbol{r}}_{a_1} \Delta a - \boldsymbol{T}_1 a_1 \Delta\theta \tag{2.3-9}$$

where $\boldsymbol{T}_1 = \boldsymbol{T}(a_1)$ is the unit tangent vector in the direction of the ray passing through the point (a_1,θ_1), $\theta_1 = \pi/2 - \alpha(a_1)$. As a zeroth approximation for $\Delta\theta$, we have $\Delta\theta^{(0)} = -\alpha'(a_1)\Delta a$. To this provisional point for $\boldsymbol{r}_a(a_2)$ at $(a_1 + \Delta a, \theta_1 + \Delta\theta^{(0)})$, we use the diffraction integral again to obtain the phase $\hat{\varphi}_2^{(0)}$ from the mapped complex field $E(a_1 + \Delta a, \theta_1 + \Delta\theta^{(0)})$. Taking the difference in the mapped phases at (a_1,θ_1) and at $(a_1 + \Delta a, \theta_1 + \Delta\theta^{(0)})$,

$\Delta\hat{\varphi}^{(0)} = \hat{\varphi}_2^{(0)} - \hat{\varphi}(a_1)$, it follows from Eq. (2.3-5) that an updated estimate of $\Delta\theta$ is given by

$$\Delta\theta^{(1)} = -\frac{\Delta\hat{\varphi}_2^{(0)}}{\bar{a}k}, \quad \bar{a} = \frac{a_1 + a_2}{2} = a_1 + \frac{\Delta a}{2} \tag{2.3-10}$$

The updated estimate of θ_2 is given by $\theta_2^{(1)} = \theta_1 + \Delta\theta^{(1)}$. Mapping the LEO field again using the diffraction integral to the updated provisional point on the impact parameter curve $\left(a_2, \theta_2^{(1)}\right)$ gives us $\hat{\varphi}_2^{(1)}$. We again obtain from the difference in mapped phases $\Delta\hat{\varphi}^{(1)} = \hat{\varphi}_2^{(1)} - \hat{\varphi}_2^{(0)}$, and use Eq. (2.3-10) to obtain an updated value $\Delta\theta^{(2)}$. We iterate until convergence is achieved. This gives the point (a_2, θ_2) with $\alpha(a_2) = \pi/2 - \theta_2$. Note that at this new point (a_2, θ_2) we now have the converged phase $\hat{\varphi}(a_2) = \hat{\varphi}_2$, the bending angle $\alpha(a_2) = \alpha_2 = \pi/2 - \theta_2$, and an updated value for $\alpha_2' = -(\hat{\varphi}_2 - \hat{\varphi}_1)/a_2$ to start the next step at $a_3 = a_2 + \Delta a$. By this way of succession we generate the impact parameter curve $\boldsymbol{r}_a(a)$ over the range of impact parameter values relevant to the occultation.

The most important aspect of this approach is that it facilitates direct recovery of the refractivity profile. Since we now have recovered an unambiguous bending-angle profile $(a, \alpha(a))$, we can use the Abel transform in Eq. (2.2-2') to recover $\log n(a)$, and thence $\log n(r_*)$ from $a = r_* n(r_*)$.

If a super-refracting layer occurs at a given altitude range, then ray theory imposes restrictions on the location of tangency points in and below the layer. In a super-refracting layer located in the range $r_d \le r \le r_u$, the condition $rdn/dr + n \le 0$ holds. Here the radius of curvature of a ray would be shorter than the local radius of curvature of the equipotential surface. For tangency points in the layer and below it down to a certain critical altitude denoted by $r^c < r_d$, a ray could not escape in a spherical stratified medium. The ray with its tangency point at the lower critical altitude $r_* = r^c$ below the layer just manages to pass through the layer and escape from its upper boundary at $r = r_u$ parallel to it. This is analogous to the parallel path direction required by Snell's law of a refracted ray exiting a planar surface at the critical internal angle of incidence. It follows from Bouguer's law that the ray through this lower critical tangency point below the layer has the same impact parameter value as the ray just grazing the top of the layer where super-refractivity first sets in, i.e., $n(r^c) r^c = n(r_u) r_u$. (See Section 6.5.) These two rays have in general different bending angles. Therefore, the impact parameter space curve has a discontinuity in θ_a, i.e., two different bending angles, but with the same value for the impact parameter. Hence, $ds_a/da \to \infty$, $dx_a/dz_a \to -\tan\alpha(a)$. The

break is initially parallel to the top ray just grazing the top of the super-refractivity layer, and finally it is parallel to the emerging ray from the lower critical tangency point below the bottom of the layer.

Ray theory alone cannot tell us about $n(r)$ between the upper boundary of a super-refracting layer at $r = r_u$ and the critical altitude below the lower boundary at $r = r^c$ because no ray with its tangency point within these bounds can exist in a spherical stratified medium. There is a hiatus within the bounds $r^c \le r_* \le r_u$. However, in practice a super-refractive layer tends to be relatively thin, usually caused by a marine layer. (In radio transmission research, super-refractivity is called ducting.) Therefore, one can use the recovered profiles for $n(r)$ above and below the super-refractivity zone $r^c \le r_* \le r_u$, plus some atmospheric physics, to estimate $n(r)$ in between. In wave theory, this super-refractive zone corresponds to the attenuation or tunneling part of the Airy function of the first kind. Wave theory (Section 5.8) predicts a very weak but non-zero field for this zone, analogous to the Fresnel decay in amplitude observed behind a shadow boundary from a knife-edge. Section 2.7 includes a discussion of a thin-screen/scalar diffraction model for the field observed at the LEO from a super-refractive boundary, a limiting case where $dn/dr \to -\infty$ with Δn finite. Geometric optics already becomes inaccurate near these boundaries at $r_* \le r^c$ and at $r_* \ge r_u$ (see Fig. 2-12).

The vertical resolution potential of the impact parameter space curve approach is essentially that of wave theory. This is discussed in Chapters 1 and 6, but vertical resolution is more limited by departures from spherical symmetry and other data processing properties [4,5].

Regarding caustic contact points, they have of course disappeared on the $\boldsymbol{r}_a(a)$ space curve for the impact parameter. The existence of a caustic contact point is observer-position dependent, as with rainbows. From Bouguer's law in Eq. (2.2-2), if we alter θ and take into account that a new ray with a new impact parameter value will pass through that altered point $(r, \theta + d\theta)$, we have

$$\frac{da}{d\theta} = \frac{r\cos(\theta+\alpha)}{1 - r\cos(\theta+\alpha)\alpha'} \tag{2.3-11}$$

which is zero when we set $r = a$ and $\theta = \pi/2 - \alpha$.

On the other hand, if we set $\theta = \theta_{\mathrm{L}}$ and $r = r_{\mathrm{L}}$, then we have the defocusing-dependent relationship between a and θ_{L}, i.e., $da/d\theta_{\mathrm{L}} = \zeta_{\mathrm{L}} r_{\mathrm{L}} \cos(\theta_{\mathrm{L}} + \alpha)$, where $\zeta_{\mathrm{L}}^{-1} = 1 - r_{\mathrm{L}} \cos(\theta_{\mathrm{L}} + \alpha)\alpha'$. $da/d\theta_{\mathrm{L}}$ can be positive or negative depending on the sign of the defocusing factor ζ_{L}. If we set $\theta_{\mathrm{L}} = \dot{\theta}_{\mathrm{L}} t + \theta_{\mathrm{L}o}$, then we have the relationship between a and time, $da/dt = \zeta_{\mathrm{L}} \dot{\theta}_{\mathrm{L}} r_{\mathrm{L}} \cos(\theta_{\mathrm{L}} + \alpha)$. For anomalous rays, da/dt is positive for a

setting occultation; the impact parameter migrates upward with time for these anomalous rays and at the same time downward for normal rays.

It follows that, for the purpose of forward propagating to the LEO the perturbations in phase and amplitude caused by the atmospheric refractivity perturbations, we could assign a phase profile $\varphi(a)$ to the impact parameter space curve that is given from Eq. (2.3-2) by

$$\varphi(a) = k\left(a\alpha(a) + \int_a^{\infty} \alpha(\rho) d\rho \right) \tag{2.3-12}$$

where $\alpha(a)$ is the reference plus perturbed bending-angle profile calculated for a spherical symmetric atmosphere from geometric optics. Moreover, given $\alpha(a)$, we can generate the impact parameter space curve $\boldsymbol{r}_a(a)$ from its definition in Eq. (2.3-6). We show in Chapter 5 using a full-spectrum wave theory (see Table 5-1) that the stationary values with respect to spectral number of the spectral density function for phase delay in wave theory, when they exist, are very closely equal to $\varphi(a)$, given in Eq. (2.3-12).

On the other hand, the impact parameter space curve generated by $\boldsymbol{r}_a(a)$ is likely to have a complicated shape in multipath situations. Computing the distance from the $\boldsymbol{r}_a(a)$ curve to the LEO is somewhat less convenient than simply using the Fresnel approximation with a planar screen. This alternate approach works well when the caustic avoidance condition is met, which relates to the thin atmosphere conditions cited above in Eqs. (2.2-8) and (2.2-9).

2.3.3 The Fresnel Phase Screen

Here we use a single planar thin-screen model to study the multipath and diffraction processes from a single boundary embedded in a laminar atmosphere possessing an otherwise smoothly varying radial gradient in refractivity. The screen is nominally oriented orthogonal to the LEO–GPS line, and the rays impinging on and emerging from the screen are the straight-line asymptotes from the LEO and the GPS satellite (see Fig. 2-3). The relationship between the altitudes h and h_{LG} is given in Eq. (2.2-5). The single thin-screen model also could be applied to multiple but spherical symmetric boundaries, given the bending-angle profile according to Snell's law that results from a ray transecting these multiple layers (see Appendix A). As an analysis technique for recovery of the refractivity profile, the usefulness of the planar thin-screen model is limited to thin atmospheres, conditions not always found, particularly in the lower troposphere. Also, if atmospheric refractivity has significant horizontal gradients, then the multiple-screen approach would be a better choice.

To maintain a one-to-one relationship between h and a, the condition $dh/da > 0$ must hold. It follows from Eq. (2.2-5) that this condition is equivalent to the condition

$$\frac{dh}{da} = \sec\alpha\left(1 + a\tan\alpha\frac{d\alpha}{da}\right) \doteq 1 + a\alpha\frac{d\alpha}{da} > 0 \qquad (2.3\text{-}13)$$

which is equivalent to proscribing any caustics from occurring in the phase screen (where $dh/da = 0$). If α is given by an exponential refractivity model, then it can be shown [see Eq. (2.3-18)] that the condition in Eq. (2.3-13) is equivalent to the condition $\beta < (2\pi)^{-1/2} \approx 0.4$. This is basically double the refractive bending from dry air at sea level. In the lower troposphere, this condition can readily be violated at a marine layer boundary, but as α decreases secularly with altitude, Eq. (2.3-13) is more readily satisfied, even with larger values of $d\alpha/da$. For h lying within the range of values where the condition in Eq. (2.3-13) is violated, i.e., $dh/da = 1 + a\alpha(d\alpha/da) < 0$, there are in general at least three values of bending angle that apply. The assignment of a unique phase function to the screen [based on Eq. (2.3-7)] is not possible for this range of h values.

On the other hand, we could cant the screen clockwise by a small angle α_m. Then, in that screen the altitude relationship would be $h = a\sec(\alpha - \alpha_m) - R$, and the caustic avoidance condition becomes $1 + a(\alpha - \alpha_m)\alpha' > 0$, which might be easier to satisfy. The thin-screen phase function to be given in Eq. (2.5-1c) would have an extra factor of $\cos\alpha_m$.

2.3.4 Suitability of the Thin-Screen Model for Diffraction Analysis

The value of the thin-screen approach for qualitative study of Fresnel processes in the Earth's atmosphere partially depends on the refractivity being fairly stratified along equipotential surfaces as well as the thin screen being viewed from afar. Also, the thin atmosphere condition should hold. If the along-track inhomogeneity in refractivity is significant, then better results likely would require a multi-screen approach. We know that the fidelity of the diffraction results from the single thin-screen model is fairly good for a grazing occultation in a uniform medium from an opaque surface—e.g., the Earth's limb [16]. When $r_o >> \lambda$, the diffraction effects from a three-dimensional opaque object can be calculated by replacing the shadow zone on the surface of that object with a disk perpendicular to the incident ray path and bounded by the silhouette of the object [17]. However, a curved and nearly transparent surface, across which a discontinuity in refractivity or in one of its derivatives occurs, can support scattered rays that arise from a combination of external and internal reflections and surface propagation (evanescent waves) yielding, for

example, rainbow and waveguide effects. These would be more difficult to model with the thin screen without first solving Maxwell's equations or some other wave propagation technique for the actual surface and surrounding medium and then adopting a phase and amplitude profile for the thin screen that corresponds to the actual observed phase and amplitude.

There are three-dimensional diffraction techniques discussed in the literature that use both geometric and physical optics approaches to evaluate scattering from a variety of relatively simple geometrical objects. Helmholtz's equation in a stratified medium, as a boundary value problem in Potential Theory, can formally be solved for simple surfaces, such as spheres and cylinders, and with relatively simple asymptotic boundary conditions, such as incident planar electromagnetic waves. Solutions for scattering from a conducting or a dielectric sphere in a uniform medium were developed starting about 95 years ago by G. Mie, P. Debye, G. Watson, and others. Mie scattering theory arose from the study of scattered light from droplets. Asymptotic solutions to Helmholtz's equation for $r_o >> \lambda$ also were worked out during the early decades of the 20th century, and parabolic equation techniques for wave propagation have evolved greatly in the last 20 years [18].

Chapter 5 presents a wave theory approach that deals more rigorously with the diffraction problem discussed here. There a discontinuity in a refraction-related property is embedded in a spherical atmosphere possessing a significant refractive gradient. A modified Mie scattering theory technique is developed to deal with wave propagation through the refracting medium and across the discontinuity. Good agreement holds for a single scattering surface between the thin-screen results here and the more rigorous wave-theoretic results in Chapter 5, under the caustic avoidance assumption in Eq. (2.3-13) and for positions away from external reflections and rainbow caustics. The latter are caustics arising from multiple internal reflections and refraction within the sphere.

2.3.5 A Phase Profile for the Thin Screen

For the purpose of describing diffraction and multipath processes, we use here the perpendicular mounted phase screen described above, keeping in mind the possibility that phase screen caustics might occur. Let $\varphi(h)$ be the nominal thin-screen phase delay at a thin-screen altitude h. We must now set this profile using this model so that the predicted phase observed at the LEO matches the actual atmosphere-induced phase delay observed by the LEO. Let $\psi\left(h_{\mathrm{LG}}, \tilde{h}\right)$ be the cumulative phase observed by the LEO for a ray starting from the thin screen at an arbitrary altitude $\tilde{h}$. From Fig. 2-3, we see that

$$\psi = kl + \varphi = k\sqrt{\left(\tilde{h} - h_{\mathrm{LG}}\right)^2 + D^2} + \varphi(\tilde{h}) \qquad (2.3\text{-}14)$$

where k is the wave number of the ray, $k = 2\pi / \lambda$. In geometric optics, a ray path is defined by the condition that the observed phase obtained from that path is stationary. The stationary-phase condition on the ray path, from the GPS satellite (infinitely afar) through the thin screen to the LEO, requires that $\partial\psi / \partial\tilde{h} = 0$; that is,

$$\frac{\partial\psi}{\partial\tilde{h}} = k\frac{\tilde{h} - h_{\mathrm{LG}}}{l} + \frac{d\varphi}{d\tilde{h}} = k\sin\alpha + \frac{d\varphi}{d\tilde{h}} = 0 \tag{2.3-15}$$

It follows for small bending angles that

$$\varphi(h) = k\int_h^\infty \sin(\alpha(h'))dh' \doteq k\int_h^\infty \alpha(h')dh' \tag{2.3-16}$$

Here it is assumed that $\alpha(h) \to 0$ as $h \to \infty$. The LEO-observed stationary phase is given by $\psi(h_{\mathrm{LG}}) = \psi[h_{\mathrm{LG}}, h(h_{\mathrm{LG}})]$, where h is the altitude in the thin screen providing the stationary-phase value for the LEO located at h_{LG}, and it is given in terms of the impact parameter by Eq. (2.2-5). It may not be unique if multipath is present. Comparing the phase $\varphi(h)$ in Eq. (2.3-16) for the planar thin screen with the phase $\varphi(r,\theta)$ given in Eq. (2.3-2) for a ray at the point (r,θ), the difference is the term $r\cos(\theta+\alpha) + a\alpha$. Through second order, this term is just $r\cos\theta\sec\alpha$, the slant distance to the point h in the screen, which from Eq. (2.3-14) is $\left(\left(\tilde{h} - h_{\mathrm{LG}}\right)^2 + D^2\right)^{1/2}$.

2.3.6 Bending-Angle Perturbations

To apply the appropriate phase profile $\varphi(h)$ for the thin screen according to Eq. (2.3-16), we need the appropriate form for $\alpha(r_*)$. It is convenient to break $\alpha(r_*)$ into two parts: the reference bending-angle profile $\alpha(r_*)$ from a reference refractivity profile, and the perturbed component of the bending angle $\delta\alpha(r_*)$ due to the refractivity-related discontinuity. For the reference profile, we use the bending angle from an exponential profile for the atmospheric refractivity. This is given by

$$n = 1 + N_o \exp\left(-\frac{r - r_o}{H_{p_o}}\right) \tag{2.3-17}$$

Appendix A [Eq. (A-30)] derives the form for $\alpha(r_*)$ that this refractivity profile generates. It is given by

$$\alpha(r_*) = \sqrt{\frac{2\pi r_*}{H_{p_*}}} N(r_*)\left(1 + (\sqrt{2}-1)\beta_* + 0.28\beta_*^2 + \cdots\right) \qquad (2.3\text{-}18)$$

Here $N(r) = n(r) - 1$ and $\beta = \mid n'r / n \mid = Nr / H_p$. This expression is very accurate for thin atmospheres, i.e., $\beta <\approx 1/2$.

We now need expressions for $\delta\alpha(r_*)$ for the different cases. A ray with its point of tangency below the discontinuity travels through both the "+" and the "–" regimes when its radius of curvature is greater than r_o, as indicated in Fig. 2-1. It follows that the bending angle observed by the LEO is given from Eq. (2.2-2) by

$$\alpha(r_*) = \begin{cases} -2a^- \left(\int_{r_*}^{r_o} \left(\frac{n'}{n}\right)^- \frac{dr}{\sqrt{(n^- r)^2 - (a^-)^2}} + \int_{r_o}^{\infty} \left(\frac{n'}{n}\right)^+ \frac{dr}{\sqrt{(n^+ r)^2 - (a^-)^2}} \right), & r_* \le r_o \\ \alpha^+(r_*), & r_* > r_o \end{cases} \qquad (2.3\text{-}19)$$

where $a^- = r_* n^-(r_*)$, and where $n^+(r)$ and $n^-(r)$ are functions describing the index of refraction in the + and – regimes, respectively. Here $n' = dn/dr$. In Eq. (2.3-10), $\alpha^+(r_*)$ is the bending angle for points of tangency of the ray lying in the + regime where $n^+(r)$ applies, and, therefore, it is given by Eq. (2.2-2) with $n = n^+(r)$ or by Eq. (2.3-18) when the refractivity profile is given by Eq. (2.3-17). Equations (2.2-2) and (2.3-19) are valid only over integration intervals where $n(r)$ is differentiable. These equations have to be appended with discrete Snell's law terms at points where $n(r)$ is discontinuous, and when $n^-(r_o) > n^+(r_o)$, the inequality $r_* \le r_o$ in Eq. (2.3-19) also must be amended to account for critical refraction. For the case where the gradient of $n(r)$ is at least piecewise continuous, Eq. (2.3-19) can be recast into the form

$$\alpha(r_*) - \alpha^-(r_*) = \begin{cases} -2a^- \int_{r_o}^{\infty} \left[\left(\frac{n'}{n}\right)^+ \frac{1}{\sqrt{(n^+ r)^2 - (a^-)^2}} - \left(\frac{n'}{n}\right)^- \frac{1}{\sqrt{(n^- r)^2 - (a^-)^2}} \right] dr, & r_* \le r_o \\ \alpha^+(r_*) - \alpha^-(r_*), & r_* > r_o \end{cases} \qquad (2.3\text{-}20)$$

Here, $\alpha^-(r_*)$ has a connotation similar to $\alpha^+(r_*)$; i.e., it is the bending angle that would be obtained from Eq. (2.2-2) if $n^-(r)$ applied to the entire atmosphere. Equation (2.3-20) may be generalized from a single boundary geometry to the multiple-layer case (see Appendix A).

The evaluation of the integral in Eq. (2.3-20) requires differential techniques, which will be found in Appendix A. Equation (2.3-20) involves terms of the form $\alpha^-(r_*) - \alpha^+(r_*)$ plus additional terms to account for the differences between a^-, a^+, and $a_o^\pm = r_o n^\pm(r_o)$. For example, if we write $\alpha^\pm(r_*) = \alpha\left[p^\pm, r_*\right]$, where p is the parameter to be varied, then we can linearize $\alpha^+(r_*) - \alpha^-(r_*)$ by the form

$$\alpha^+(r_*) - \alpha^-(r_*) \doteq \frac{\partial \alpha}{\partial p} \Delta p \tag{2.3-21}$$

It is shown in Appendix A, for example, that when p is the scale height,

$$\frac{\partial \alpha}{\partial H_{p_o}} \hat{=} \frac{\alpha}{H_{p_o}} K_H = \frac{\alpha}{H_{p_o}} \left(-\frac{1}{2} + \mathrm{O}\left[\beta, (\gamma H_{p_o})\right] \right) \tag{2.3-22}$$

Also, when p is the normalized lapse rate $\gamma = (dT/dr)T^{-1}$, which is a constant, one obtains

$$\frac{\partial \alpha}{\partial \gamma} \hat{=} \alpha H_{p_o} K_\gamma = \alpha H_{p_o} \left(\frac{3}{8} + \mathrm{O}\left[\beta, (\gamma H_{p_o})\right] \right) \tag{2.3-23}$$

In Appendix A, the integrals in Eq. (2.3-20) are computed to zeroth order in β (i.e., no ray path curvature corrections are applied, but for Case A, β is included), and to first order in the discontinuous parameter of interest for the three cases cited above: discontinuous refractivity, discontinuous scale height, and discontinuous lapse rate. The reference refractivity and refractive bending-angle profiles are given for a constant lapse rate by Eqs. (A-37) through (A-39).

2.3.7 Case A: A Discontinuity in Refractivity

Case A may be considered as a limiting case for a ray that crosses a very sharp boundary, such as a well-delineated marine layer in the lower troposphere or a sporadic E-layer in the ionosphere. Let $n(r)$ have a discontinuity across the surface $r = r_o$. Then $n(r)$ has the form

$$n(r) = n^-(r) + \mathscr{H}(r - r_o)\left[n^+(r) - n^-(r)\right] \tag{2.3-24}$$

where $\mathscr{H}(r-r_o)$ is the Heaviside function and $n^+(r)$ and $n^-(r)$ are well-behaved reference functions, as obtained, for example, from Eqs. (A-37) and (A-38). Therefore, the gradient of $n(r)$ has the form

$$\frac{dn}{dr} = \frac{dn^-}{dr} + \mathscr{H}(r-r_o)\left[\frac{dn^+}{dr} - \frac{dn^-}{dr}\right] + \delta(r-r_o)\left[n^+(r_o) - n^-(r_o)\right] \qquad (2.3\text{-}25)$$

where $\delta(r-r_o)$ is the Dirac delta function. Let σ be defined by

$$\sigma^2 = \frac{(r_o - r)}{H_{p_o}}, \quad r \le r_o \qquad (2.3\text{-}26)$$

where H_{p_o} is the pressure scale height at $r = r_o$. It can be shown (Appendix A) that an expression for Eq. (2.3-20) valid to first order in ΔN but without ray path curvature corrections is given by

$$\begin{aligned}\alpha(r_*) - \alpha^-(r_*) \doteq \alpha^-(r_o)&\left[\left(1 + \frac{\Delta N_o}{N_o}\right)F(\sigma_+) - F(\sigma)\right] \\ &+ 2\left(\sin^{-1}\left(\frac{r_* n_*^-}{r_o n_o^+}\right) - \sin^{-1}\left(\frac{r_*}{r_o}\right)\right), \; r_* \le \tilde{r} \le r_o, \\ &\alpha^+(r_*) - \alpha^-(r_*), \quad r_* > r_o\end{aligned} \qquad (2.3\text{-}27)$$

where $\Delta N_o = n^+(r_o) - n^-(r_o)$ and where $F(x)$ is given by

$$F(x) = (1 - \mathrm{erf}(x))\exp(x^2) \qquad (2.3\text{-}28)$$

The quantity σ_+ is given by

$$\sigma_+^2 = \sigma^2 + \left(\frac{\Delta N_o}{N_o}\right)\left(\frac{N_o r_o}{H_{p_o}}\right) \qquad (2.3\text{-}29)$$

which is essentially equal to σ, except near $\sigma = 0$. The second term in Eq. (2.3-29) is just the fractional change in refractivity times the ratio of r_o to the radius of curvature of the ray, both of which are nominally small quantities. When $\Delta N_o < 0$, $\sigma_+ = 0$ corresponds to the critical internal reflection condition and $\tilde{r}$, the maximum value of r_* at this critical point, is given by $\tilde{r} = r_o n^+(r_o)/n^-(r_o) \doteq r_o(1 + \Delta N_o)$. In geometric optics, no rays exist from the GPS satellite to the LEO with turning points located in the interval $\tilde{r} < r_* \le r_o$.

One can improve the accuracy of Eq. (2.3-27) by including bending along the ray path of integration in Eq. (2.3-20); a power series in β would result. Equation (A-49) in Appendix A accounts for most of the first-order bending for Case A.

Figure 2-6(a) shows $\alpha(r)-\alpha^{-}(r)$ near $r=r_o$ based on a numerical integration of Eq. (2.2-2a) plus the Snell term. The exponential refractivity profile given in Eq. (2.3-17) is used for the reference. The values of the reference parameters used for this figure are $N_o=200\times10^{-6}$, $r_o-R_E\approx 3\,\text{km}$, $H_{p_o}=7\,\text{km}$, $\gamma=0$, and $\Delta N_o/N_o=-1/20$. Most of the jump in α arises from the Snell term: $\Delta\alpha\doteq 2\left(-2\Delta N_o\right)^{1/2}$.

2.3.8 Case B: A Discontinuity in Scale Height

Let ΔH_{p_o} denote a discontinuity in a piecewise constant pressure scale height, which occurs across the surface $r=r_o$. In this case, the gradient of $n(r)$ is discontinuous across this surface, but $n(r)$ is differentiable at interior points of the integration intervals defined by Eq. (2.3-20). From Appendix A, Eqs. (A-44) and (A-45), it can be shown that to first order in ΔH_{p_o}

$$\alpha(r_*)-\alpha^{-}(r_*)\doteq$$

$$\begin{Bmatrix} \dfrac{\Delta H_{p_o}}{H_{p_o}}\alpha^{-}(r_o)\left[\left(K_H-\sigma^2\right)F(\sigma)+\dfrac{\sigma}{\sqrt{\pi}}\right]+\mathrm{O}\left[\left(\dfrac{\Delta H_{p_o}}{H_{p_o}}\right)^2\right],\ r_*\le r_o \\ \alpha^{+}(r_*)-\alpha^{-}(r_*),\ \ r_*\ge r_o \end{Bmatrix} \tag{2.3-30}$$

where K_H is given by Eq. (2.3-22) for $\gamma=0$, and $\sigma^2=\left(r_o-r_*\right)/H_{p_o}$. Appendix A also discusses the $\gamma\neq 0$ case. Depending on the physical circumstances present at a boundary layer in the real atmosphere, one could use a linear combination of Eqs. (2.3-27) and (2.3-30) to model the effect of a discontinuity in temperature or a boundary demarcating a lower regime containing water vapor. The values of the parameters used in Fig. 2-6(b) are $\alpha^{-}\left(r_o\right)=5.44$ mrad, $\Delta H_p/H_p=-1/7$, $r_o-R_E=10$ km, $H_{p_o}=7$ km, and $\gamma=0$.

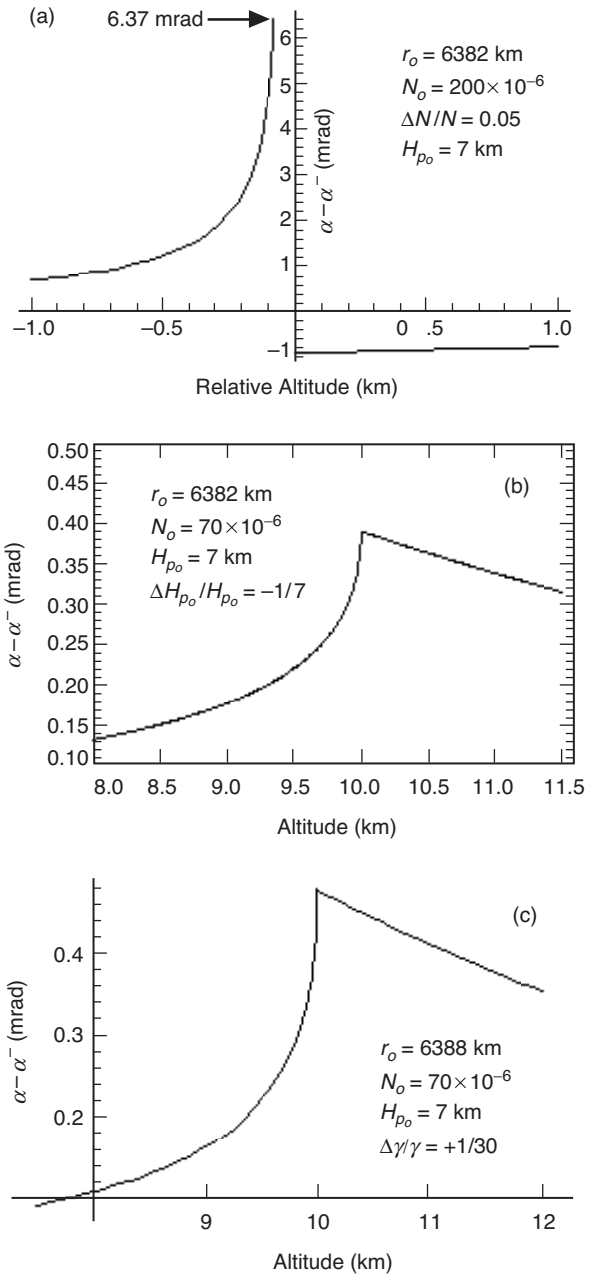

Fig. 2-6. Perturbations in bending angle profile for different refractivity scenarios: (a) perturbation in bending angle for Case A; no rays exist in the super-refractivity region where r_o – 79 m < r_* < r_o; (b) perturbation for Case B; and (c) perturbation for Case C. All figures use an exponential reference refractivity profile.

2.3.9 Case C: A Discontinuity in Lapse Rate

Let $\Delta\gamma$ denote a discontinuity in the piecewise constant normalized lapse rate across the surface $r = r_o$. In this case, the gradient of $n(r)$ is discontinuous across this surface, but $n(r)$ is differentiable at all interior points of the integration. From Appendix A, Eqs. (A-44) and (A-45), it follows that Eq. (2.3-20) becomes to first order in $\left(\Delta\gamma H_{p_o}\right)$

$$\alpha(r_*) - \alpha^-(r_*) \doteq$$

$$\begin{cases} \left(\Delta\gamma H_{p_o}\right)\alpha^-(r_o)\left[\left(K_\gamma + \dfrac{3\sigma^2}{2} + \dfrac{\sigma^4}{2}\right)F(\sigma) - \dfrac{1}{\sqrt{\pi}}\left(\dfrac{5\sigma}{4} + \dfrac{\sigma^3}{2}\right)\right], & r_* \le r_o \\ \alpha^+(r_*) - \alpha^-(r_*), & r_* \le r_o, \end{cases} \tag{2.3-31}$$

where K_γ is given by Eq. (2.3-23) (or by Eq. (A-40) for the $\gamma \neq 0$ case), and $\sigma^2 = \left(r_o - r_*\right)/H_{p_o}$. Figure 2-6(c) shows the behavior of $\alpha(r) - \alpha^-(r)$ near $r = r_o$. The values of the reference parameters used to obtain this figure are $N_o = 70 \times 10^{-6}$, $r_o - R_E = 10$ km, $H_{p_o} = 7$ km, $\gamma = 0$, and $\Delta\gamma = +1/30$, or about a 7 K/km discontinuity in lapse rate.

We note in all three cases that $\alpha(r_*) = \alpha^+(r_*)$ when $r_* > r_o$; also, for Cases B and C, $\alpha(r_*)$ is continuous across the surface $r = r_o$. It follows for Cases B and C that the difference in altitudes of the points of tangency of the two rays about the surface $r = r_o$, which is given by

$$\left\{r_*^+ - r_*^- \doteq D\left[\alpha^+\left(r_o^+\right) - \alpha\left(r_o^-\right)\right], \quad D^{-1} = D_{\text{LEO}}^{-1} + D_{\text{GPS}}^{-1}\right\} \tag{2.3-32}$$

approaches zero. We will see below that when a caustic surface is generated, one has to be more careful about the definition of the condition $r_* \to r_o$.

2.4 Multipath Using a Thin Phase Screen Model

We show now some examples of the effect of these discontinuity surfaces on the relationship between the altitude of the ray path turning point and the altitude of the LEO–GPS line. Figures 2-7(a) through 2-7(d) show the relationship between h and h_{LG}, based on the thin-screen relationship given in Eq. (2.2-5), in the vicinity of the discontinuity at r_o for Cases A, B, and C. The same parameter values hold here that were used in Fig. 2-6.

Figure 2–7(a), which corresponds to the scenario shown in Fig. 2–2(b), shows a super-refractivity situation. This results from a hard transition in

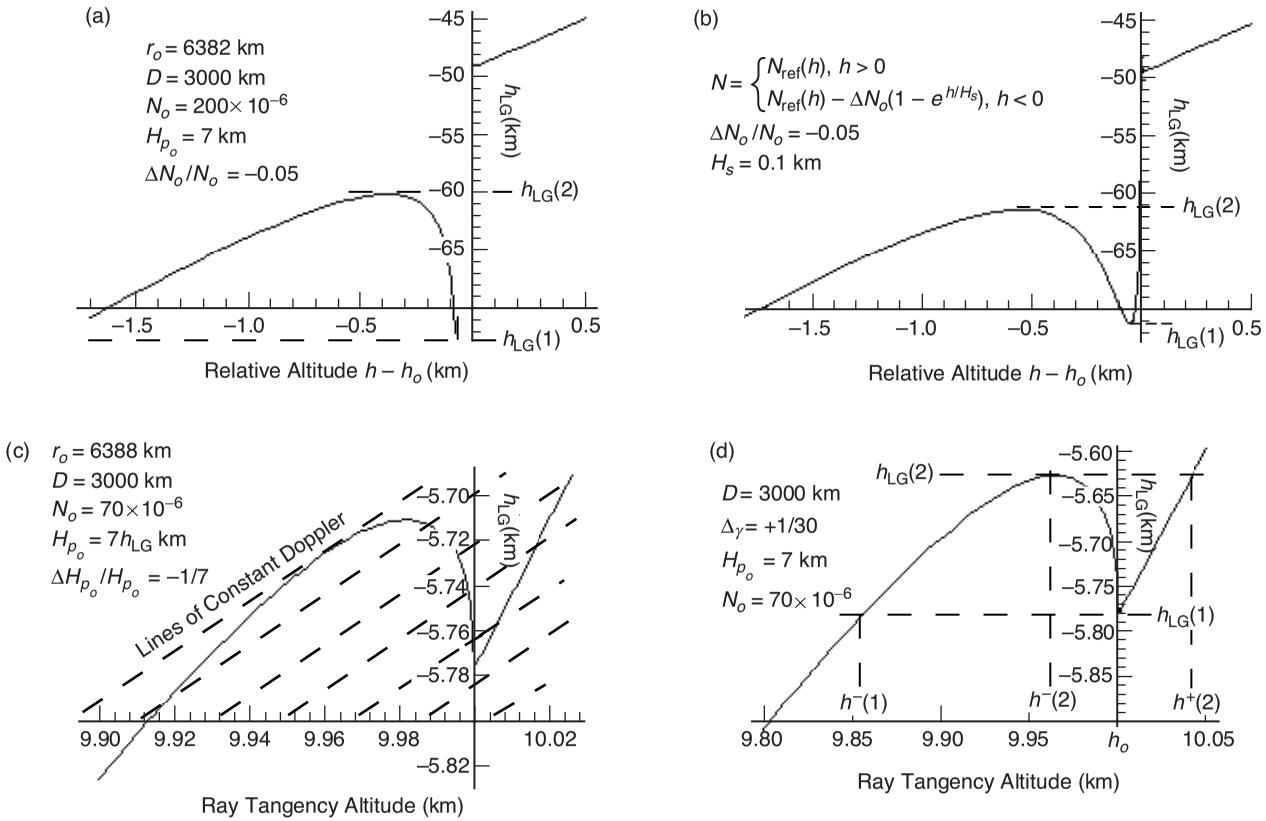

Fig. 2-7. Thin-screen relationship between altitude h of the impact parameter and altitude h_{LG} of the LEO–GPS straight line for different refractivity perturbation scenarios: (a) Case A, a super-refractivity scenario with $N_o^- > N_o^+$, (b) the same as (a) except that the profile of N has been softened to the profile given in the panel, (c) Case B, a scale height discontinuity, and (d) Case C, a lapse rate discontinuity.

refractivity (infinite gradient) at the boundary, a Case A scenario. Here a multipath doublet configuration results. The doublet begins at an altitude in h–space of about 79 m below the boundary (corresponding to the critical reflection point $\tilde{r}$) and continues down to an altitude $h(1)$, or about 1.6 km below the boundary. No rays exist for tangency points lying between the boundary and 79 m below. Figure 2–7(a) also shows the shadow zone in h_{LG}-space (~–60.5 ≤ h_{LG} ≤ ~–49.3 km or, equivalently, 4 to 5 s of elapsed time) within which no signal at all (according to geometric optics) will be received by the LEO. Below the first contact point with the caustic in h_{LG}–space, the altitude difference between the doublet rays grows to well in excess of the Fresnel diameter.

Figure 2-7(b) uses the same conditions that were used in Fig. 2-7(a) except that the discontinuity in refractivity has been replaced by a sharp continuous change over a narrow transition zone just below the boundary. Although the magnitude of the refractivity gradient is large in this zone, it is bounded so that the ray existence condition $\beta < 1$ holds for all values of the impact parameter. By softening the discontinuity, one obtains a triplet ray system and two caustics. This yields a quasi-shadow zone where the highly defocused main ray continues, and also a triplet ray system between the contact points with the two caustics, one at $\left(h(1), h_{\mathrm{LG}}(1)\right)$ and the other at $\left(h(2), h_{\mathrm{LG}}(2)\right)$. The caustic contact points mark the upper and lower boundaries of the multipath zone.

Figure 2-7(c) shows for Case B the range of ray path altitudes and h_{LG} values over which three (when $\Delta H_p < 0$) mutually interfering ray paths will be received by the LEO. Contours of constant excess Doppler also are overlaid on this figure. The excess Doppler value for the dashed contour through the point at $h_{\mathrm{LG}} = -5.78$, $h = 10$ km, is about 80 Hz. The spacing between contours is about 100 mHz.

Figure 2-7(d) shows for Case C the range of ray path altitudes and h_{LG} values over which three mutually interfering ray paths will be received by the LEO when $\Delta\gamma > 0$. The point where $dh_{\mathrm{LG}} / dh = 0$, at a ray path altitude of $h = h^{-}(2) \cong 9.96\,\mathrm{km}$ in Fig. 2-7(d), is a singularity point where the vertical radius of the first Fresnel zone approaches infinity, which is the condition for the existence of a caustic. For Case C, we see from Fig. 2-7(d) that the altitude differences between the tangency points of the three mutually interfering ray paths depend non-linearly on the location of h_{LG} within the multipath zone $h_{\mathrm{LG}}(1) \le h_{\mathrm{LG}} \le h_{\mathrm{LG}}(2)$. These multiple rays arise as a result of an abrupt downward increase in refractivity or decrease (in magnitude) of its gradient. Figure 2-7(d) shows that the maximum separation, $h_o - h^{-}(1)$, is about 170 m for a temperature lapse rate discontinuity of +7 K/km at $h_o = 10$ km, which is small compared to the local mean Fresnel zone but still a significant source of

interference. The half-width in Doppler space of the local first Fresnel zone is given by

$$f_{\mathcal{F}} = \dot{h}_{LG} \frac{|\zeta(h)|}{\mathcal{F}(h)} \tag{2.4-1}$$

where $\dot{h}_{\mathrm{LG}}$ is about $2.5\ \mathrm{km/s} \pm 20\%$. For the assumed ambient medium at an altitude of 10 km, $f_{\mathcal{F}}$ is roughly 2 Hz. The maximum separation in ray path tangency points in Fig. 2-7(d), which is $(h(1) - h_o)$, corresponds to a tonal separation of about 1 Hz, well within the half-width of the first Fresnel zone and well within the 25–Hz bandwidth of the high sample rate mode in a TurboRogue GPS flight receiver. Figures 2–7(a) through 2-7(d) also may be used to obtain the time and Doppler intervals between caustic events. The abscissa should be multiplied by approximately 5 to yield a Doppler scale in hertz, and the vertical scale should be divided by approximately 2.5 to obtain a time scale in seconds. Appendix C provides further details on the separation geometry at the tangency points of the rays. It also shows the near-quadratic dependence of the separation scale (and, therefore, the difference in frequency of the tones in Doppler space) on ΔH_p for Case B and $\Delta\gamma$ for Case C. Caustics produce inherently nonlinear effects, and their existence depends on the polarities of $\Delta\gamma$ and ΔH_p.

Because the Doppler spreads in Figs. 2–7(b) and 2-7(c) are small compared to the Doppler spread of the first Fresnel zone, use of geometric optics to predict received signal amplitude and phase would be inaccurate. A diffraction treatment should be used.

When $\Delta\gamma < 0$, the left-hand portion of the curve in Fig. 2-7(c) that is less than $h = h_o = 10$ km will be flipped about the horizontal axis (at $h_{\mathrm{LG}} = h_{\mathrm{LG}}(1) = -5.78$); no caustic occurs and no multipath arises due to this kind of discontinuity. However, severe defocusing will occur at the boundary; a quasi-shadow zone results when the refractivity gradient is steep. Appendix C provides further details and an example of this converse case.

For Case B, no caustic surface is generated when $\Delta H_p / H_p > 0$ at $h = h_o$, but shadowing may be severe if the gradient is large. Diffraction effects will soften the shadowing.

We also can estimate the relative intensity (in a geometric optics context) and the relative phases of the multipath rays observed by the LEO. To obtain the relative intensity, we use the defocusing function $\zeta(h)$, which is given by Eq. (2.2–6). For Case C, for example, $|\zeta(h)|^{1/2}$ is shown in Figs. 2–8(a) and 2–8 (b) for $\Delta\gamma < 0$ and $\Delta\gamma > 0$, respectively. It should be noted, however, that the altitude separation of the multipath rays for this value of $\Delta\gamma$ is well within

the first Fresnel diameter, and, therefore, the geometric optics approach will break down. Here $\zeta(h)$ will not provide a realistic measure of relative intensities; a wave theory approach is required. Figure 2-8(b) exhibits the flaring of intensity based on $\zeta(h)$ in the vicinity of the singularity point where $dh_{\mathrm{LG}} / dh \to 0$, which marks the onset of a caustic condition. Here $|\zeta|^{1/2} \to \infty$ in the vicinity of the caustic contact point. The scalar diffraction versions of these cases are given in Figs. 2–10(c) and 2–10(d). Irrespective of whether or not the altitudes of the tangency points of the multiple rays lie within the Fresnel zone, a ray optics approach based on stationary-phase theory to only second order breaks down at the point of contact with a caustic. A third-order stationary-phase theory is given in Appendix D.

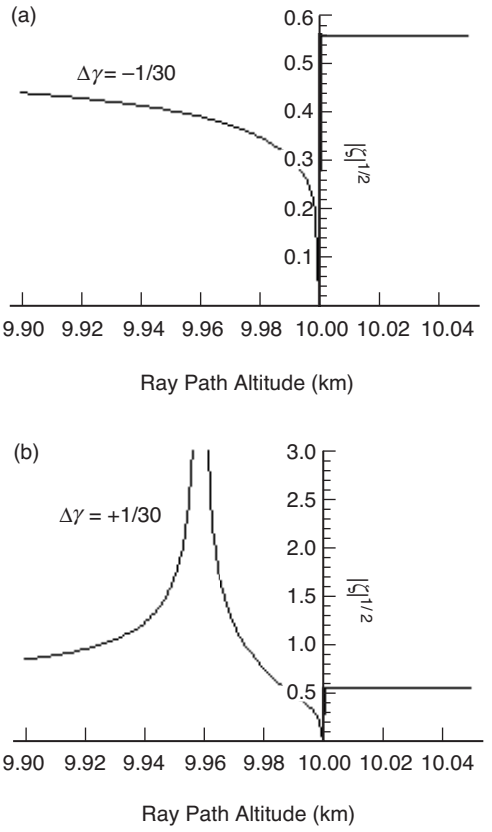

Fig. 2-8. Defocusing function $|\zeta|^{1/2}$ for Case C: (a) $\Delta\gamma < 0$; a shadow zone results; and (b) $\Delta\gamma > 0$; $\zeta \to \infty$ near the caustic contact. The scalar diffraction versions are shown in Figs. 2-10(c) and 2-10(d).

The phase difference between multipath rays observed by the LEO is given by the difference in values of the Fresnel phase function [see Eq. (2.5-1)] for the two rays. We show later that this is given by

$$\Delta = \Phi^- - \Phi^+ =$$

$$\frac{2\pi}{\lambda D}\left(\frac{1}{2\zeta^+\left(r_*^+\right)}(r_*^- - r_*^+)^2 + D\int_{r_*^-}^{\infty}\left(\alpha^- - \alpha^+\right)dr_*\right) + \mathrm{O}\left[\frac{\left(r_*^- - r_*^+\right)^3}{\lambda D^2}\right] \qquad (2.4\text{-}2)$$

where $r_* - R$ is the altitude of a ray at its point of tangency with the Earth's limb. The first term on the RHS of Eq. (2.4-2) results from the increase in path length due to the additional bending $\delta\alpha$. The second term results from the additional delay through the atmosphere. For cases with a caustic, the observed phase Φ^- must be further delineated, for example, by Φ^a and Φ^b and r_*^a and r_*^b, to indicate the multipath branch to which one is referring. Again, geometric optics breaks down for the small values of $\Delta\gamma$ used in these examples and also near the point of first contact with the caustic. The actual observed phase differences will differ significantly from those predicted in Eq. (2.4-2) because of diffraction effects.

Finally, one can carry out the vector addition of the multipath signals

$$E\exp(i\Psi) = E^+\exp\left(i\Psi^+\right) + E^a\exp\left(i\Psi^a\right) + E^b\exp\left(i\Psi^b\right) \qquad (2.4\text{-}3)$$

using Eqs. (2.2-6) and (2.4-2) for each ray to obtain a total electric field vector observed by the LEO as predicted by geometric optics. When the altitude differences of the tangency points of the multiple signals are significantly larger than $2\mathscr{F}(h)$ and well away from caustics, this approach usually should be valid.

2.5 Scalar Diffraction: The Rayleigh–Sommerfeld Integral

To calculate wave theory effects from these perturbations in refractivity on bending angle, we use scalar diffraction theory applied to a thin-screen phase profile $\varphi(h)$, which is defined in Eq. (2.3-16) in terms of the bending-angle profile. Scalar diffraction theory is derived from the Helmholtz–Kirchoff integral theorem from classical electrodynamics to relate the amplitude and phase distributions of an electromagnetic wave over a surface to the amplitude and phase at a point interior to the surface [3,12]. This theorem is valid when the wavelength of the wave is small compared to the scale of the radiating surface. Appendix A presents a version of this surface integral; also, its contraction to a two-dimensional coplanar geometry is derived there. It

provides the basis for scalar diffraction theory. As was pointed out in Section 2.3, Eq. (2.3-8), one can use this integral to map the observed amplitude and phase measurements made by the LEO backwards to an equivalent set of observations on a surface closer to the Earth's limb. This is the basis for the back-plane propagation technique used in radio occultations to reduce multipath [13–15].

Similarly, the Helmholtz–Kirchoff surface integral also can be used to map the phase and amplitude forward from an emitting surface to the LEO [4,5]. The forward one-dimensional version of this integral [see Appendix A, Eq. (A-22)] leads directly to the Rayleigh–Sommerfeld integral for scalar diffraction. We adopt a thin-screen model to provide the phase and amplitude of the emitting source in the screen. For the purpose of forward propagating to the LEO, we use a simple phase screen mounted perpendicular to the LEO–GPS line. This should work well provided that the perturbation in refractivity does not violate the uniqueness condition $dh/da > 0$ or, equivalently, $1 + a\tan\alpha(d\alpha/da) > 0$. This ensures a one-to-one relationship between h and a. We also use in this integral the Fresnel approximation. Here the ray path length from a height h in the screen to the LEO is given by $l \doteq D + (h - h_{\mathrm{LG}})^2 / 2D$, (Fig. 2–3). For typical LEO orbits, $D >> |h - h_{\mathrm{LG}}|$, so this approximation is adequate. For a LEO orbit altitude of 600 km, D is roughly 3000 km, whereas the range of $h - h_{\mathrm{LG}}$ is only a few tens of kilometers.

Let $E(h_{\mathrm{LG}})$ and $\psi(h_{\mathrm{LG}})$ be the normalized amplitude and phase that are observed by the LEO at an altitude h_{LG} as a result of the Earth's intervening atmosphere. If the Earth's atmosphere were absent, the amplitude of the signal at the LEO would be unity and the phase would be kD. Diffraction effects from the Earth's opaque limb are ignored here. It can be shown (see Appendix A) that $E(h_{\mathrm{LG}})$ and $\psi(h_{\mathrm{LG}})$ are given through the diffraction integral formulation by

$$E(h_{\mathrm{LG}})\exp\left(i\psi(h_{\mathrm{LG}})\right) = \frac{1}{1+i}\exp(ikD)\sqrt{\frac{2}{\lambda D}}\int_{-\infty}^{\infty} A(h)e^{i\Phi(h,h_{\mathrm{LG}})}dh, \quad \text{(2.5-1a)}$$

$$\Phi(h,h_{\mathrm{LG}}) \doteq \varphi(h) + \frac{\pi}{\lambda D}(h - h_{\mathrm{LG}})^2, \quad \text{(2.5-1b)}$$

$$\varphi(h) = k\int_{h}^{\infty}\alpha(h')dh' \quad \text{(2.5-1c)}$$

We define $\Phi(h,h_{\mathrm{LG}})$ to be the Fresnel phase function for the screen. Equation (2.5-1b) shows that it equals the thin-screen phase function $\varphi(h)$ plus the extra geometric delay from the bending angle, $k(l - D)$. For a transparent

phase screen, $A \equiv 1$. In actuality, the integration limits in this integral are finite, but one can for convenience set these limits to infinity. If one is interested in edge phenomena, such as knife-edge diffraction effects from the Earth's limb, one can set $A = 0$ below a certain altitude. We will show using the stationary-phase technique that, when the stationary-phase points are located well away from questionable boundaries, the resulting error from using infinite integration limits is negligible because the significant contributions to the integral come only from neighborhoods where the phase is stationary, or nearly so.

As a check, let us set $\varphi(h) \equiv 0$ and $A \equiv 1$. Then the Fresnel phase function becomes simply the extra geometric phase delay from any point in the screen to the LEO, $\Phi \doteq \pi(h - h_{\mathrm{LG}})^2 / \lambda D$. The integral in Eq. (2.5-1) reduces to a complete complex Fresnel integral with a value of $(1+i)(\lambda D / 2)^{1/2}$. Thus, $E(h_{\mathrm{LG}}) = 1$ and $\psi(h_{\mathrm{LG}}) = kD$, as predicted.

There are alternative methods for evaluating the convolution integral in Eq. (2.5-1). One approach much favored prior to the advent of modern computers was the saddle-point method. Here we discuss the stationary-phase technique, which expresses the results in terms of Fresnel integrals. However, the stationary-phase technique as a computational technique is limited to a few ideal situations with isolated stationary-phase points. When complexities of a real atmosphere are introduced, this approach becomes cumbersome. Even handling caustics becomes a problem unless a higher-order Taylor series expansion is used in Eq. (2.5-1). The technique does provide some insights, however, and also asymptotic forms that are useful in determining the limiting value of a numerical integration. It is in that spirit that the discussion in the following section is presented.

2.6 The Stationary-Phase Technique

The stationary-phase technique uses the fact that the principal contributions to an integral with a rapidly oscillating kernel come from those neighborhoods where it has the least change. Traditionally the technique mainly has been applied to neighborhoods around well-isolated stationary-phase points or near end points associated with a boundary of some sort. In this technique, the integral in Eq. (2.5-1a) is evaluated in terms of a Fresnel integral by expanding $\Phi(h, h_{\mathrm{LG}})$ in a Taylor series in h about the thin-screen altitude $h = h^*(h_{\mathrm{LG}})$ providing the least change in $\Phi(h, h_{\mathrm{LG}})$ with h, while retaining only terms up to the second degree. This technique fails in the case of multiple stationary points when they are not sufficiently isolated to ensure adequate phase windup in the neighborhood between them. It also fails when the expansion of the Fresnel phase to only the quadratic term is insufficient, for example, when the quadratic term is zero or near zero, which occurs near a caustic. A third-order treatment is needed for caustics.

We show an example where the technique is sound. In this case, we assume that there is a single stationary-phase point, well away from any boundaries that the geometry might include. If we expand the Fresnel phase $\Phi(h, h_{\mathrm{LG}})$ in Eq. (2.5-1a) in a Taylor series, we obtain

$$\left.\begin{aligned} \Phi(h, h_{\mathrm{LG}}) &= \Phi(h^*, h_{\mathrm{LG}}) + \left.\frac{\partial \Phi}{\partial h}\right|_{h^*} (h - h^*) + \frac{1}{2}\left.\frac{\partial^2 \Phi}{\partial h^2}\right|_{h^*} (h - h^*)^2 + \cdots, \\ \Phi(h^*, h_{\mathrm{LG}}) &= k\int_{h^*}^{\infty} \alpha(h')\,dh' + \frac{\pi}{\lambda D}(h^* - h_{\mathrm{LG}})^2, \\ \left.\frac{\partial \Phi}{\partial h}\right|_{h^*} &= k\left(\frac{1}{D}(h - h_{\mathrm{LG}}) - \alpha(h)\right)^*, \quad \left.\frac{\partial^2 \Phi}{\partial h^2}\right|_{h^*} = \frac{\pi}{\lambda D}\left(1 - D\frac{d\alpha}{dh}\right)^* \end{aligned}\right\} \tag{2.6-1}$$

Now we set $h^* = h_{\mathrm{LG}} + D\alpha(h^*)$ to null the first-order term. This condition is identical with the thin-screen relationship in Eq. (2.2-5). Then the diffraction integral in Eq. (2.5-1) becomes

$$\left.\begin{aligned} E(h_{\mathrm{LG}})e^{i\psi(h_{\mathrm{LG}})} &= e^{i\left(kD + \Phi(h^*, h_{\mathrm{LG}})\right)} \frac{1}{1+i}\sqrt{\frac{2}{\lambda D}} \int_{-\infty}^{\infty} e^{\left(i\frac{\pi}{\lambda D \zeta^*}(h - h^*)^2\right)} dh, \\ \zeta^{-1} &= \left(1 - D\frac{d\alpha}{dh}\right)^* \end{aligned}\right\} \tag{2.6-2}$$

Here ζ is the defocusing function [Eq. (2.2-6)] evaluated at the stationary-phase altitude h^*. We assume that it is positive in this example, but on anomalous rays $\zeta < 0$. Now we make a change of integration variable $v = (2/\lambda D\zeta)^{1/2}(h - h^*)$ to obtain

$$\left.\begin{aligned} E(h_{\mathrm{LG}})e^{i\psi(h_{LG})} &\doteq \sqrt{\zeta}\, e^{i\left(kD + \Phi(h^*, h_{\mathrm{LG}})\right)} \frac{1}{1+i}\int_{-\infty}^{\infty} e^{i\frac{\pi}{2}v^2} dv = \sqrt{\zeta}\, e^{i\left(kD + \Phi(h^*, h_{\mathrm{LG}})\right)}, \\ kD + \Phi(h^*, h_{\mathrm{LG}}) &= kD\left(1 + \frac{1}{2}\alpha^2(h^*)\right) + k\int_{h^*}^{\infty} \alpha(h)\,dh, \\ h^* &= h_{\mathrm{LG}} + D\alpha(h^*) \end{aligned}\right\} \tag{2.6-3}$$

Implicit in the integration here is that the infinite limits on the integral are valid. Otherwise, an incomplete Fresnel integral would result, yielding diffraction fringes. Moreover, we have assumed here that h^*, the stationary value for h, is

unique. In the Fresnel phase expression, the term $D(1+\alpha^2/2) \doteq D\sec\alpha$ is the slant range from the LEO to the stationary-phase point in the thin screen. The second term, the integral term, is the total refractivity-induced phase delay, that is, it is the extra phase delay incurred by the ray after passing completely through the refracting atmosphere. We note the defocusing term $\zeta^{1/2}$ in the amplitude term in Eq. (2.6-3), in accordance with geometric optics.

Neither of the integration assumptions used to obtain Eq. (2.6-3) is likely to hold when refractive boundaries have to be considered, such as one of the scenarios described by Cases A, B, or C. We need to develop the perturbation to the nominal thin-screen phase-delay profile that results from a discontinuity on the surface $r = r_o$. For a ray descending through the + regime of the atmosphere, the nominal phase delay in the thin-screen model is taken to be that corresponding to the bending-angle profile $\alpha^+(r)$, which is given by Eq. (2.3-18) (see also Appendix A). Because the turning point of this ray is above the boundary, it is unfettered by the effects of the discontinuity below. The thin-screen phase function may be broken into two parts:

$$\varphi(h) = \left\{ \begin{array}{l} \varphi_{\text{Ref}}(h), \;\; h \geq h_o, \\ \varphi_{\text{Ref}}(h) + \delta\varphi(h), \;\; h \leq h_o \end{array} \right\} \tag{2.6-4}$$

where $\varphi_{\text{Ref}}(h)$ is the reference thin-screen phase function derived from the integral in Eq. (2.3-16) in terms of the reference form for the bending angle for $\alpha^+(r_*)$, nominally that given in Eq. (2.3-18). Similarly, the perturbed thin-screen phase function $\delta\varphi(h)$ is obtained from Eq. (2.3-16) and the form for the perturbed bending angle, for example, Eq. (2.3-31) for Case C. Thus,

$$\delta\varphi(h) = \left\{ \begin{array}{l} 0, \;\; h \geq h_o \\ k\int_h^{h_o} \delta\alpha(h')\,dh', \;\; h \leq h_o \end{array} \right\} \tag{2.6-5}$$

Here $h = n(r_*)r_*\sec\alpha - R$. The quantity $\delta\alpha(h) = \alpha(r_*) - \alpha^+(r_*)$ is the perturbation in refractive bending angle. Thus, there is a break in the diffraction integral in Eq. (2.5-1). We can use any one of Eqs. (2.3-27), (2.3-30), or (2.3-31) combined with the appropriate expression for $\alpha^-(r_*) - \alpha^+(r_*)$ to develop explicit expressions for $\delta\alpha(r)$ $\delta\varphi(h)$ and for $\delta\varphi(h)$, which then can be used with $\varphi_{\text{Ref}}(h)$ in the diffraction integral to calculate the LEO-observed phase and amplitude perturbations. Appendix A gives explicit forms for $\delta\varphi(h)$ for the different discontinuity scenarios discussed here. Here h denotes a height in the thin screen, but r_* denotes a geocentric radial position in the atmosphere of the ray path tangency point with the Earth's limb. Also, $r_* - R$ denotes the

height in the atmosphere of the tangency point; h is essentially equal to $r_* n(r_*) \sec\alpha - R$, as has already been discussed.

As the point of tangency of the ray traversing the atmosphere approaches the boundary, a perturbation in phase is observed by the LEO because of the discontinuity. This results from changes in both the atmospheric delay and the refractive bending angle. Figures 2–9(a) through 2–9(c) show the Fresnel phase function $\Phi(h, h_{\mathrm{LG}})$ versus h and the incomplete diffraction integral $\left|\int_h^\infty \exp\left[i\Phi(h', h_{\mathrm{LG}})\right] dh'\right| / (\lambda D)^{1/2}$ for three different h_{LG} values. A Case C scenario has been used. These figures should be compared to the thin-screen Case C scenario shown in Fig. 2–7(d), except that there the discontinuity in lapse rate is 1/7 the magnitude used to generate Figs. 2–9(a) through 2-9(c). The integral is obtained from a numerical integration. The phasor $\exp\left[i\Phi(h, h_{\mathrm{LG}})\right]$ in this integral rapidly oscillates except near stationary-phase points. Therefore, care must be exercised in evaluating the limiting value of the integral. Figures 2-9(a) and 2-9(b) show two cases where h_{LG} lies outside of the multipath zone; in Fig. 2-9(a) h_{LG} is well above the zone, and in Fig. 2-9(b) it is just below. In Fig. 2-9(c), h_{LG} lies in the multipah zone, that is, where $h_{\mathrm{LG}}(1) < h_{\mathrm{LG}} < h_{\mathrm{LG}}(2)$. Here $h_{\mathrm{LG}}(1)$ and $h_{\mathrm{LG}}(2)$ mark the lower and upper boundaries of the multipath zone, respectively. From Fig. 2-7(d), it follows that the lower limit, $h_{\mathrm{LG}}(1)$, is independent of the magnitude of the discontinuity, such as $\Delta\gamma$, and depends only on the altitude of the discontinuity and the nominal value of $\alpha^+(r_o)$. However, $h_{\mathrm{LG}}(2)$ corresponds to a nominal ray path altitude *above* h_o (when $\Delta\gamma > 0$); that is, it is the altitude of tangency point for the main ray [the m ray in Fig. 2–2(a)]. It depends on the magnitude of the discontinuity.

When h_{LG} does not lie within a multipath zone, as shown in Figs. 2-9(a) and 2-9(b), and the convexity of the Fresnel phase function $\Phi(h, h_{\mathrm{LG}})$ is such that it rapidly increases as h moves away from its stationary-phase value, then the stationary-phase approximation for the diffraction integral may be accurate. This would hold for Fig. 2-9(a), even though there are reversals in polarity of the convexity of $\Phi(h, h_{\mathrm{LG}})$ (around $h = 10\,\mathrm{km}$), because $|\partial\Phi/\partial h| >> 0$ in that neighborhood. There is a unique and isolated stationary-phase point.

However, when the convexity of $\Phi(h, h_{\mathrm{LG}})$ not only reverses sign but also induces a significant slowdown in the growth rate of $\Phi(h, h_{\mathrm{LG}})$ to near zero, which is shown in Fig. 2-9(b) (around $h = 10\,\mathrm{km}$), care in using the stationary-phase approximation must be exercised. This figure typifies the case where a quasi-stationary-phase altitude makes a small but significant contribution to the overall integral in Eq. (2.5-1), which affects both the amplitude and phase of the observed signal. A geometric optics approach to situations like this would

be compromised *even though there is a unique stationary-phase altitude* for the example in Fig. 2-9(b) (at about 7.4 km).

Finally, Fig. 2-9(c) shows $\Phi(h, h_{LG})$ when h_{LG} falls within the multipath zone between the two caustic contact points, which results in a triplet of

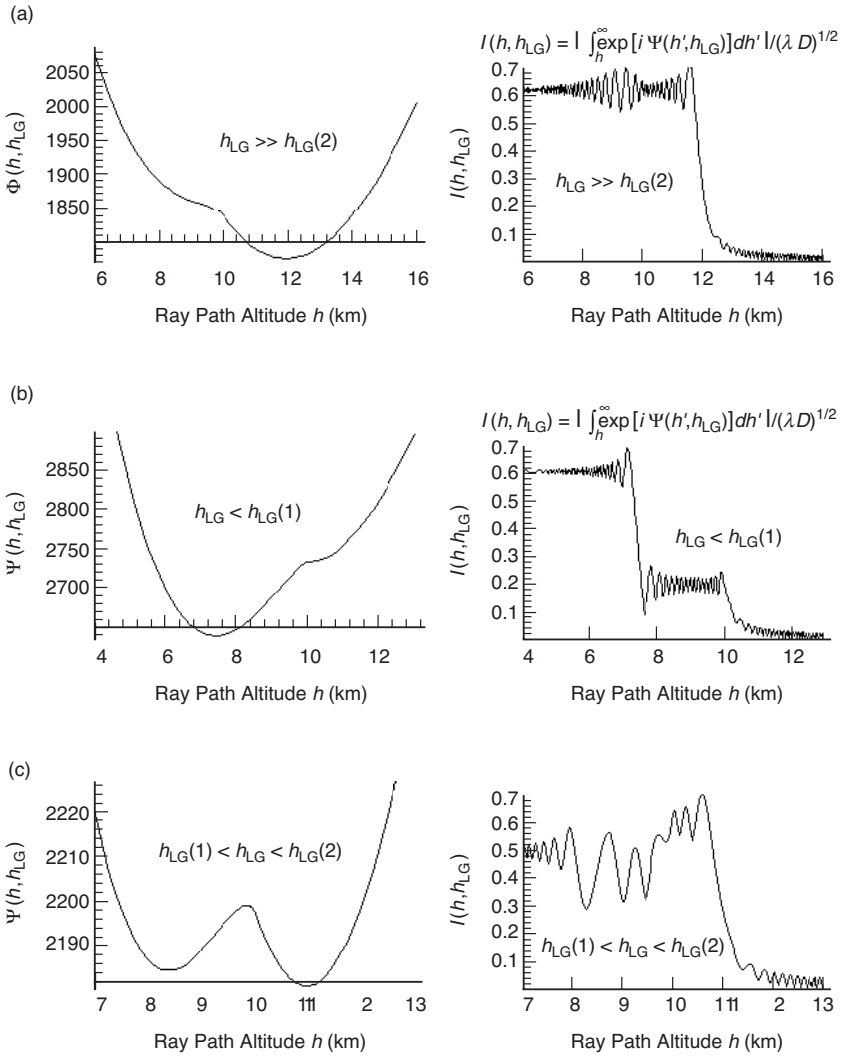

Fig. 2-9. Pairs of figures showing Fresnel phase Φ (left) and the incomplete scalar diffraction integral $I(h, h_{LG})$ (right) for three different fixed altitudes of the LEO–GPS line h_{LG}. A Case C discontinuity in lapse rate at 10 km has been used with $\Delta\gamma = +7/30$. In (a), h_{LG} lies well above the upper boundary of the multipath zone at $h_{LG}(2)$, (b) h_{LG} lies below the lower boundary at $h_{LG}(1)$ but near it, and (c) h_{LG} lies in the multipath zone.

stationary-phase points. Here the stationary-phase approximation is likely to be seriously compromised. Multipath rays with narrow separations at their tangency points comparable to or smaller than the radius of the first Fresnel zone are difficult for the stationary-phase approximation. This is because the technique in these circumstances tends to yield a biased phase error in the individual E vector computation for each stationary-phase point, i.e., each ray, which can significantly alter the resultant E vector after vector addition of the contributions from the multiple ray segments. Near a caustic contact point at $\left(h^{\dagger}, h_{\mathrm{LG}}^{\dagger}\right)$, $\Phi\left(h, h_{\mathrm{LG}}\right)$ must be expanded in a power series in $\left(h - h^{\dagger}\right)$ that includes third-order terms to obtain sufficient accuracy.

Similarly, regarding Fig. 2–9, geometric optics should be accurate for situations depicted by Fig. 2–9(a), less accurate for Fig. 2–9(b), and inadequate for Fig. 2–9(c) because of co-mingling of the stationary-phase contributions from overlapping neighborhoods in impact parameter.

One can generalize Fig. 2–9(c) to situations where $\Phi\left(h, h_{\mathrm{LG}}\right)$ supports multiple multipath zones. It should be clear that for n multipath zones one generally will have up to $2n + 1$ stationary-phase points, i.e., $2n + 1$ rays, and also at least $n + 1$ caustics. Whenever a situation arises where the stationary-phase technique degrades in accuracy, so also will geometric optics as an approximate second-order description of electromagnetic wave processes. It should be pointed out, however, that in multipath situations where the separation altitudes between the tangency points of the rays are large compared to their respective first Fresnel zones, then a geometric optics approach can be sound. However, in the vector addition of these multiple rays to compute the complete field vector, their relative phases must be maintained with high accuracy.

2.6.1 Necessary Conditions for Validity of the Geometric Optics Approach

Let h^a, h^b, and h^c denote successive ray path altitudes at which the Fresnel phase has stationary values for a given value of the LEO–GPS straight line altitude above the Earth's limb, h_{LG} (see Fig. 2–3). Then necessary conditions for ensuring accuracy of the stationary-phase approximation technique with multiple stationary-phase points would be

$$\left|\Phi\left(h^a, h_{\mathrm{LG}}\right)() - \Phi\left(h^b, h_{\mathrm{LG}}\right)\right| >> \pi, \quad \left|\Phi\left(h^b, h_{\mathrm{LG}}\right) - \Phi\left(h^c, h_{\mathrm{LG}}\right)\right| >> \pi\Big\} \quad (2.6\text{-}6)$$

These conditions are tantamount to requiring the altitudes of the stationary-phase points to be separated by a distance much greater than the first Fresnel zone. However, for cases with a single stationary-phase point, as demonstrated by Fig. 2–9(b), or even for multiple stationary-phase points where the

conditions in Eq. (2.5-8) hold, we also must ensure that $\Phi(h, h_{LG})$ has no near-stationary points elsewhere. A strong condition would require that there be no polarity changes in the convexity of $\Phi(h, h_{LG})$ over the h–space. A somewhat relaxed condition would require that $|\partial\Phi/\partial h| >> 0$ within any neighborhood bounded by reversals in polarity of the convexity of $\Phi(h, h_{LG})$. Appendix D, which is principally concerned with third-order stationary-phase theory to deal with caustics, provides a more explicit bound on $|\partial\Phi/\partial h|$ for zero convexity points that are isolated ($|\Delta\Phi| >> \pi$) from end points and stationary-phase points.

2.7 Numerical Results Using Thin-Screen/Scalar Diffraction

Figures 2–10 and 2–11 show for several discontinuities, both positive and negative, the resulting phase and amplitude perturbations, $\psi(h_{LG})$ and $E(h_{LG})$. We refer to these as "Fresnel effects," but in fact they are a combination of both diffraction and interference. Away from caustic rays and deep shadow zones, the interference in amplitude and phase is largely predictable by complex addition of the rays from geometric optics. These figures are based on a numerical integration of Eq. (2.5-1). Figure 2–10 shows a Case C scenario, a lapse rate discontinuity. Figure 2–11 shows a Case A scenario, a discontinuity in N. In both the figures, the exponential refractivity model given in Eq. (2.3-17) was used to generate the reference phase profile for the Fresnel phase function given in Eq. (2.5-1) for the thin screen. The perturbed Fresnel phase was obtained from Eq. (2.6-5) using the perturbed bending-angle profile given in Eq. (2.3-27) for Case A and Eq. (2.3-31) for Case C. To simplify the calculations, the atmospheric altitude difference $r_* - r_o$ was replaced directly by the thin-screen altitude difference $h - h_o$ without converting to the impact parameter value.

These numerical integrations were aided by the stationary-phase technique, which was used to isolate the neighborhoods in h-space that contribute to the integral. Numerical integration is aided by using the asymptotic forms for the incomplete Fresnel integrals, which these diffraction integrals assume in their limit. These asymptotic forms also provide the characteristic wavelengths and magnitudes of the asymptotic modulations in signal phase and amplitude. The asymptotic expansions for the Fresnel integrals can be written in the form

$$\int_0^x \exp\left(i\frac{\pi}{2}y^2\right)dy = \frac{1+i}{2} - \left(g(x) + if(x)\right)\exp\left(i\frac{\pi}{2}x^2\right) \qquad (2.7\text{-}1)$$

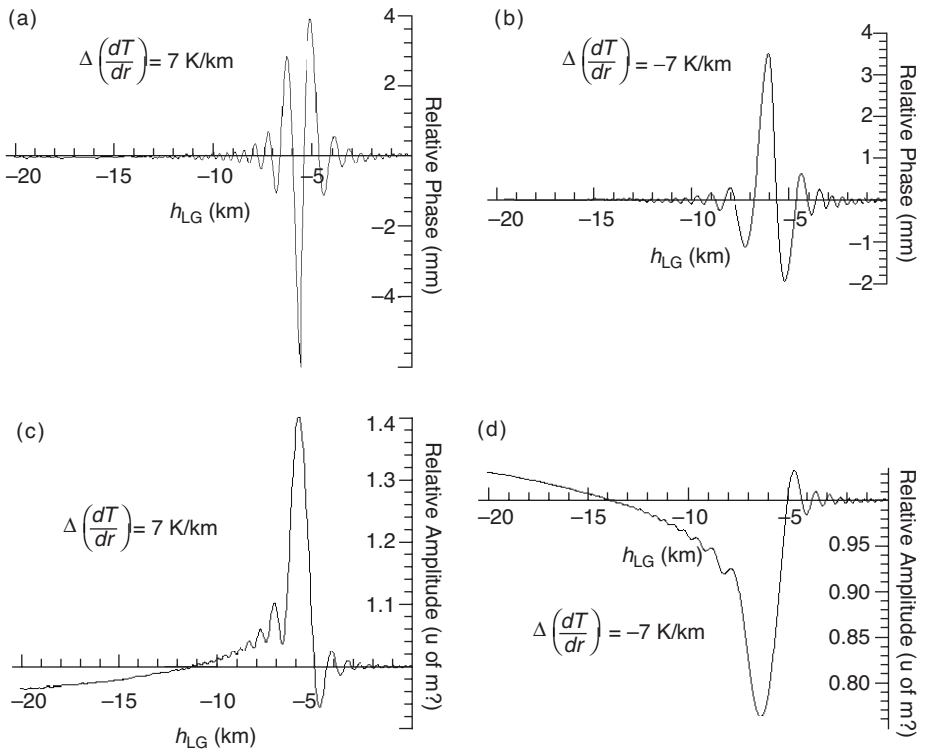

Fig. 2-10. Fresnel perturbations in received signal due to a discontinuity $\Delta\gamma = \pm 1/30$ in lapse rate at an altitude of 10 km. The phase variations are shown in millimeters for (a) $\Delta\gamma = 1/30 > 0$ and (b) $\Delta\gamma < 0$. The geometric optics terms for the reference and perturbed phase have been removed to stop phase windup. The amplitude variations are shown for (c) $\Delta\gamma = 1/30 > 0$ and (d) $\Delta\gamma < 0$.

where $f(x)$ and $g(x)$ are given by

$$f(x) = \frac{1}{\pi x} - \frac{3}{\pi^3 x^5} + \cdots, \quad g(x) = \frac{1}{\pi^2 x^3} - \frac{5}{\pi^4 x^7} + \cdots \Bigg\} \tag{2.7-2}$$

By fitting the free parameters (scale factor and mean) associated with these functions in Eq. (2.7-1) to the numerical integration of the diffraction integral in its asymptotic regime, one can calculate the limiting value of the integral without actually numerical integrating to the limit, thereby reducing computations.

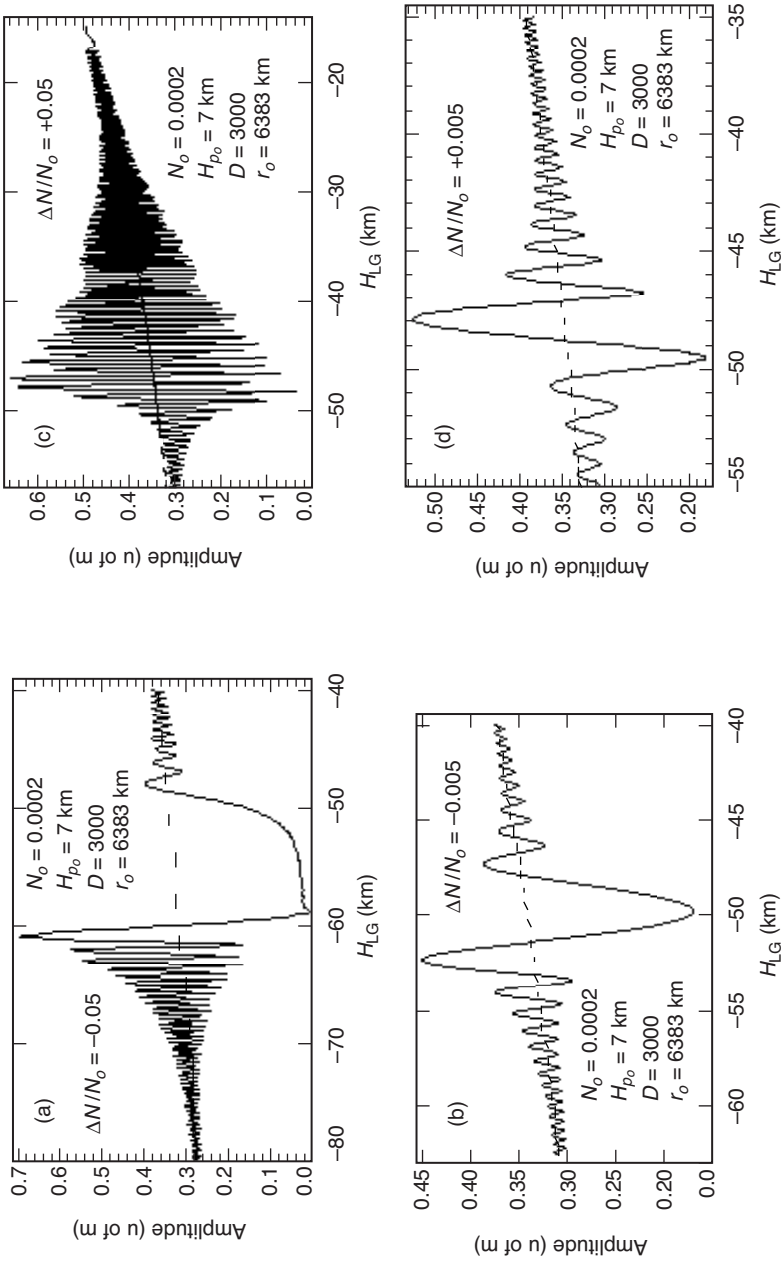

Fig. 2-11. Fresnel amplitude and phase perturbations at the LEO caused by a discontinuity in refractivity ($\Delta N = N_o^+ - N_o^-$) at about 4-km altitude: amplitude perturbations with (a) $\Delta N/N_o = -0.05$, (b) $\Delta N/N_o = -0.005$, (c) $\Delta N/N_o = +0.05$, and (d) $\Delta N/N_o = +0.005$, and phase perturbations minus geometric optics terms with (e) $\Delta N/N_o = -0.05$ and (f) $\Delta N/N_o = +0.05$. Super-refractivity occurs when $\Delta N < 0$. The dashed curves in (a) through (d) are from the defocusing factor for the dry-air exponential model used in the reference refractivity profile.

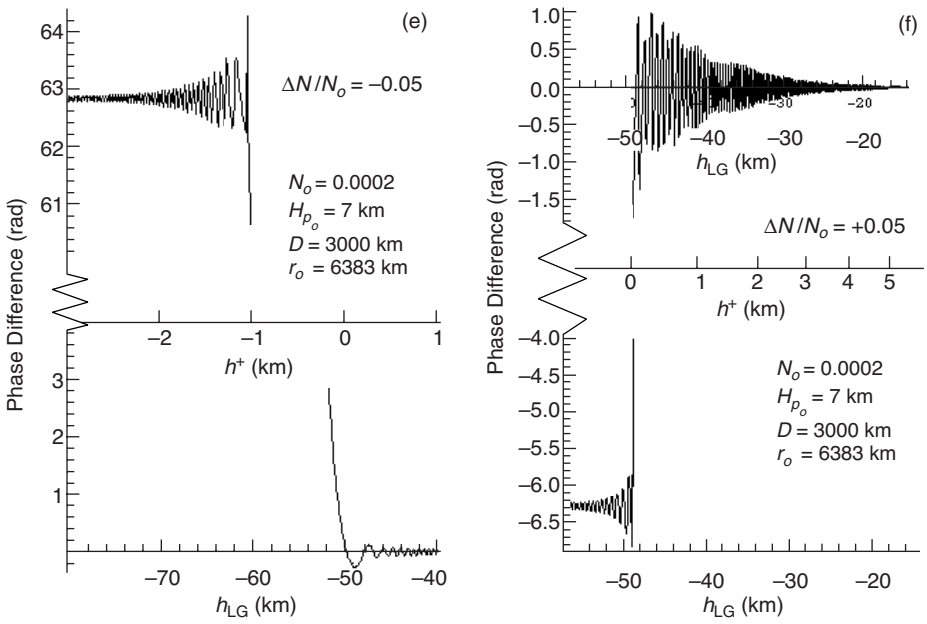

Fig. 2-11 (cont'd.).

2.7.1 Fresnel Response to a Discontinuity in Lapse Rate

Figures 2–10(a) through 2–10(d) show the Fresnel effects in phase and amplitude. The abscissa h_{LG} has the favorable property of varying nearly linearly with time during an occultation (within $\pm 0.5\%$ for the spans in these figures). The equivalent elapsed time in these figures is 5 to 7 s. The ray path altitude in Fig. 2–10 ranges from ~6 km to ~12 km and its scale, if shown, would be non-linear, and it would be compressed by defocusing compared to the h_{LG} scale. The defocusing here is about 1/3. The corresponding altitudes of the points of tangency in the actual atmosphere would be 400 to 500 m lower (essentially $(n-1)r_o$ lower). The boundary marking the discontinuity in lapse rate, at a thin-screen altitude h_o of 10 km above sea level, is crossed when the altitude of the LEO–GPS straight line reaches the altitude $h_{LG} = h_{LG}(1) = -5.78$ km. The dry air exponential refractivity model was used to generate the reference bending angle from Eq. (2.3-17) and the reference phase delay; it yields at $h_o = 10$ km a reference bending angle of 5.44 mrad. The pressure scale height is 7 km, and the temperature is 210 K. At a LEO orbital altitude of 700 km, the limb distance is $D = 3000$ km. The principal Fresnel effects in phase and amplitude are completed within a time interval of 2 to 3 s.

The LEO-observed phase delay from the geometric optics phase delay (including the perturbed part below the boundary) has been suppressed in

Figs. 2-10(a) and 2-10(b) to stop the phase windup. This leaves only the Fresnel effects. For ray path altitudes below h_o, the question arises when a caustic is present (whenever $\Delta\gamma > 0$) as to which of the new rays should be used for phase-stopping. The geometric optics phase delay offset Δ^- from the phase delay of the main ray (m), or + ray, is shown versus h_{LG} in Fig. 2–12. This is the difference in phase of a– regime ray at the LEO relative to the phase of the + regime ray that either exists for $h_{\mathrm{LG}} \geq h_{\mathrm{LG}}(1)$ or would exist if the + regime were continued below the discontinuity. Figure 2–12 shows the phase offset Δ^- for the two rays in the multipath zone: the anomalous a ray and the branching b ray. The quantity Δ^{a} gives the phase offset for the a ray. This corresponds to the a ray in Fig. 2–2(a), which as a function of h_{LG} begins at the cusp at an

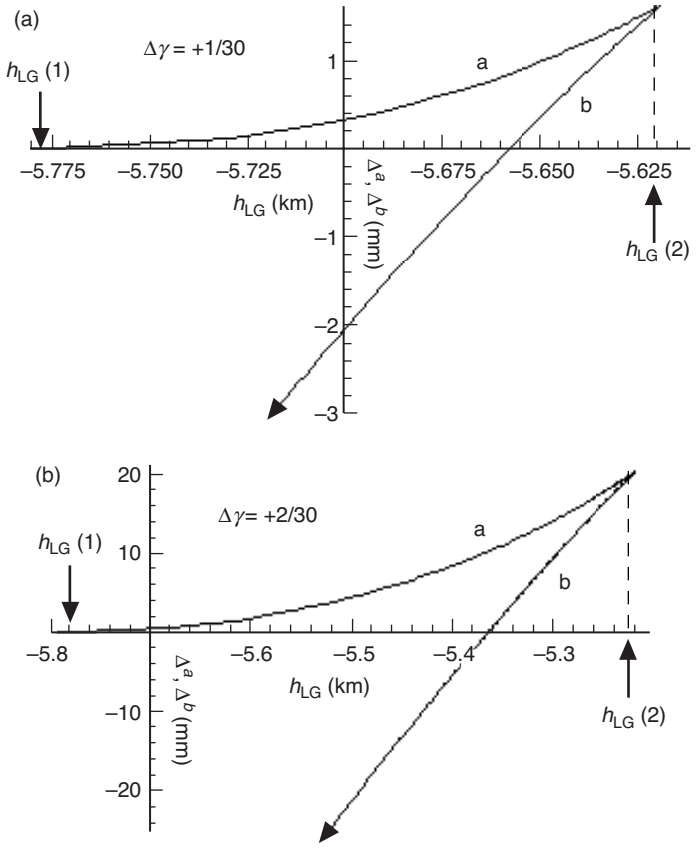

Fig. 2-12. LEO phase delay offsets from the main ray for Case C: (a) $\Delta\gamma$ = 1/30 and (b) $\Delta\gamma$ = 2/30. Phase offsets are for the ordinary branching ray, b, and also for the anomalous ray, a. Cusp marks the contact point with the caustic and its offset scales roughly as $(H_{p_o}\Delta\gamma)^4$. Delay is in range units for L1.

altitude of $h^a[h_{\mathrm{LG}}(2)]$ and terminates at the higher altitude $h_o = h^a[h_{\mathrm{LG}}(1)]$. This point $h_o = h^a[h_{\mathrm{LG}}(1)]$ also marks the termination of the lower descent of the + ray. Δ^{b} gives the phase offset for the b ray. It continues downward indefinitely into the – regime. The phase from the b ray has been applied in Fig. 2–10 to stop the phase windup. Consequently, there is a discontinuity in phase in Figs. 2-10(a) and 2-10(b) at $h_{\mathrm{LG}} = h_{\mathrm{LG}}(2)$. The magnitude of the discontinuity is about 1.6 mm for $\Delta\gamma = +1/30$ and about 20 mm for $\Delta\gamma = +2/30$.

Figures 2–10(a) and 2–10(c) show the Fresnel response for $\Delta\gamma = 1/30 > 0$. The flaring in amplitude in Fig. 2-10(a) at $h_{\mathrm{LG}} \cong -6\,\mathrm{km}$ and $h^+ \cong 9.8\,\mathrm{km}$ corresponds to the first contact with the caustic surface. In a descending occultation, the + ray at the altitude $h^+[h_{\mathrm{LG}}(2)]$, which is above h_o, abruptly splits into three rays, two of which, the a and b rays in the – regime, arrive at the LEO essentially mutually in phase initially from an altitude slightly below h_o ($h^-[h_{\mathrm{LG}}(2)] - h_o \approx -40\,\mathrm{m}$). The third, the + ray, which is the main ray or the m ray in Fig. 2–2, arrives from the original altitude above h_o, ($h^+[h_{\mathrm{LG}}(2)] - h_o \approx +50\,\mathrm{m}$) with an initial phase offset from the other two that is given by the cusp offset in Fig. 2–12(a).

The cusps in Fig. 2–12 mark the initial contact with the caustic surface for a descending occultation and also the point of nascence of the a and b rays. The altitude of the ray at the cusp is determined in terms of the Fresnel function $\Phi(h, h_{\mathrm{LG}})$ by the conditions that both $\partial\Phi/\partial h$ and $\partial^2\Phi/\partial h^2$ are zero there. The magnitude of the initial phase offset at the cusp is very sensitive to $\Delta\gamma$. It can be shown from Appendix C that this phase offset generated by a discontinuous lapse rate can be evaluated in terms of a power series in $\Delta\gamma H_{p_o}$ and that the leading term is $(\Delta\gamma H_{p_o})^4$. The relative magnitude of the flaring at the caustic contact point will depend on how flat the Fresnel phase function is in that neighborhood. This is discussed further in Appendix D.

If the phase offset at the cusp from the phase of the + ray approaches a radian, as is nearly the case for Fig. 2–12(b) (which corresponds to an unrealistically large lapse rate discontinuity of +14K/km), then a tracking-loop-driven GPS receiver would begin to encounter difficulty in such a neighborhood. The concomitant flaring at the caustic contact would likely lure the receiver to begin tracking one or the other of the nascent rays, a or b, there. These figures (Figs. 2-10(a), 2-10(b), and 2-12) also suggest that significant Fresnel ringing in phase and amplitude will persist well below the site of the discontinuity but that only faint Fresnel perturbations are evident for ray path altitudes higher than a Fresnel radius above the discontinuity when $\Delta\gamma > 0$. Asymptotic Fresnel effects typically attenuate slowly in their continuation past a discontinuity.

A lapse rate change of $\Delta\gamma = +2/30$ is too large for the tropopause and also one does not find a strict discontinuity there. Moreover, the tracking statistics through the tropopause from more recent LEOs is good. On the other hand, sharp changes in refractive gradient due to a water vapor layer in the lower troposphere or due to sharp electron density changes in the lower ionosphere (for example, across the boundary of a sporadic E-layer) can and do cause tracking difficulty for grazing ray paths. Multipath and diffraction effects for these situations can be quite severe, as they are extremely sensitive to the magnitude of the change in refractive gradient. The severity of the Fresnel phase perturbations in the vicinity of a discontinuity in scale height is, as in Case C, very sensitive to the magnitude of ΔH_{p_o}, roughly the phase perturbations scale as $(\Delta H_{p_o})^4$. Moreover, from Figs. 2–6(b) and 2-6(c) and from Appendix C, one concludes that the profile of $\delta\alpha(r)$ for Case B will be similar to that for Case C when $\Delta H_{p_o} / H_{p_o} \approx -\Delta\gamma H_{p_o}$. For example, it is shown from Appendix A that a value $\Delta H_{p_o} / H_{p_o} = -(\Delta\gamma H_{p_o})(H_{\rho_o} / H_{p_o})^2$ gives for Cases B and C the same value for the discontinuity in the refractive gradient at r_o; only the subsequent bending-angle profiles differ, but only slightly. When this correspondence between ΔH_{p_o} and $\Delta\gamma$ holds, one would expect to see Fresnel perturbations for Case B that are similar to those for Case C.

Figures 2–10(b) and 2-10(d) show examples of Fresnel effects when $\Delta\gamma < 0$. Here the caustic-induced flaring is absent. The Fresnel effects are smaller and their onset is lowered in altitude to h_o. Also, the severe darkening at $h = h_o$ predicted by Eq. (2.2–6) and shown in Fig. 2–8(a) is very prominent but modulated by diffraction.

2.7.2 Fresnel Response to a Discontinuity in Refractivity

Figure 2–11 shows the thin-screen/scalar diffraction prediction for a discontinuity in N itself at $r = r_o$, a Case A scenario. Here r_o corresponds to a lower troposphere altitude of about 1/2 scale height above the Earth's surface. The discontinuities are $\Delta N_o / N_o = \pm 0.05$ and $\Delta N_o / N_o = \pm 0.005$. The same exponential refractivity model given in Eq. (2.3–17) is used here for the reference bending angle and Fresnel phase function with $N_o = 0.0002$ and $H = 7$ km. The dashed curves in these figures are the defocusing from the bending-angle profile for the exponential reference refractivity based on geometric optics. The amplitude of the incident wave just before encountering the atmosphere is normalized to unity.

Figures 2-11(a) and 2-11(b) show the Fresnel response in amplitude to a negative discontinuous change $\Delta N = N^+ - N^- < 0$. These can provide limiting

forms for the amplitude perturbations experienced by a ray encountering from above a sharp water vapor layer at an altitude of about 4 km. A 5 percent discontinuity in refractivity for these physical conditions corresponds roughly to only a 10 to 20 percent change in water vapor density. In Appendix C, the sensitivity of the corresponding Doppler spread between multiple rays to the magnitude of $\Delta N_o / N_o$ for $\Delta N_o < 0$ is discussed (see Fig. C–1).

Figures 2-11(c) and 2-11(d), which show the Fresnel response to a positive discontinuity in N, provide examples of what might be observed as the tangency point of the ray drops below the bottom of such a water vapor layer and enters a regime of drier and speedier air. This corresponds to the ray diagram in Fig. 2–2(a).

A relatively small negative change in refractivity, 5 percent for Fig. 2-11(a), leads to a very deep shadow zone lasting roughly 4 s. In Appendix C, it is shown that the shadow zone duration is closely proportional to $(-\Delta N)^{1/3}$. Thus, even very small negative discontinuities can result in significant shadow zones. In Fig. 2-11(b), the magnitude of ΔN is a factor of ten smaller than that used in Fig. 2-11 (a).

On the other hand, a relatively small positive ΔN results in significant interference well above the boundary. It is shown in Appendix C that the duration of the multipath region for a positive discontinuity is closely proportional to $(\Delta N)^{1/2}$. Therefore, the maximum altitude separation of the rays and also the Doppler spread for a positive discontinuity is approximately proportional to $(\Delta N)^{1/2}$.

The super-refractivity condition from a discontinuity in N creates errors in our thin-screen/scalar diffraction prediction for observed phase. The phase, unlike the amplitude, is extremely sensitive to model approximations. The uniqueness condition, $1 + a \tan\alpha (d\alpha / da) > 0$, is effectively violated across the discontinuity in the bending angle (the Snell term). Chapter 3, which uses Mie scattering theory to predict phase and amplitude at the LEO, also compares those wave theory results with the thin-screen results for similar values of ΔN (see Figs. 3–23 through 3–26).

One example of the phase perturbations from a discontinuity in N is shown in Figs. 2–11(e) and 2–11(f) in order to point out that the phase effects across a discontinuity depend strongly on the polarity of ΔN. This also is borne out in the Mie scattering treatment in Chapter 3. In Figs. 2–11(e) and 2–11(f), the phase windup has been stopped by subtracting from the predicted phase observed at the LEO the geometric optics phase applicable to each regime, the "+" and "–" regimes, as already explained earlier for Fig. 2–10. In Fig. 2-11(e), where $\Delta N / N = -0.05$, we have subtracted the geometric optics phase profile using the reference bending angle model at all altitudes, including across the shadow zone. However, no rays exist in the shadow zone [see Fig. 2–7(a)] because of the discontinuity in refractivity. For Fig. 2-11(e), the shadow zone is

defined by $h_{LG}(2) \le h_{LG} \le h_{LG}(1)$; $h_{LG}(2) \approx -60.5$ km, and $h_{LG}(1) \approx -49.3$ km, a temporal separation of 3 to 4 s. The shadow zone marks a transition region where the Fresnel phase migrates from the reference Fresnel phase profile in the + regime to the perturbed Fresnel phase profile in the – regime. The latter involves (in a geometric optics context) the weighted (by their respective defocusing factors) vector addition of the phase contributions from both the b and a rays. The point $h_{LG}(2) \approx -60.5$ km marks first contact with the caustic. Below this point, we have subtracted the phase from the branching ray b. Therefore, across this transition region we see cycle slipping relative to the prediction provided by the reference Fresnel phase profile, which here is the extrapolated reference profile from the main ray that would apply in the shadow zone without the discontinuity. The cycle slip essentially equals the difference between the phase of the nascent rays at the caustic contact point at the end of the shadow zone at $h_{LG} = h_{LG}(2)$, minus the phase of the main ray extended to that caustic point, even though the main ray doesn't exist there. For the example in Fig. 2-11(e), this amounts to about 10 cycles. For the opposite polarity of the same magnitude, i.e., $\Delta N / N = +0.05$ shown in Fig. 2-11(f), which involves a positive ΔN of the same magnitude, the cycle slip is an order of magnitude smaller.

The ramification of this for a statistical net loss of cycles by a receiver attempting to track a GPS signal that passes through a series of sharp positive and negative transitions in refractivity should be noted. A GPS receiver with an ordinary low-gain antenna is likely to have more difficulty tracking through a deep trough in the shadow zone because of poorer signal-to-noise ratio (SNR), as shown in Fig. 2–11(a), than it would in a strong (but messy) signal condition, as shown in Fig. 2–11(c). Note from Figs. 2–11(e) and Fig. 2–11(f) the difference in the rate of unaccounted-for phase windup between these positive and negative ΔN scenarios. Statistically this could result in a net loss of cycles incurred by the receiver in attempting to track the phase during successive episodes of sharp transitions of positive and negative ΔN. A net loss of cycles leads to an underestimate of N.

The frequency spectrum of the fringes in Fig. 2–11(e) is biased significantly higher than that for Fig. 2–11(f), even though the magnitude of the discontinuity in refractivity is the same in the two sets of figures. This is because of the on-average wider separation of the multipath tones when $\Delta N > 0$ (see Appendix C, Fig. C–2). The multipath separation altitudes in the panels for $\Delta N > 0$ are sufficiently wide compared to the vertical width of the first Fresnel zone that a geometric optics approach to construct these figures using Eqs. (2.2–5), (2.2–6), and (2.4–1) should be fairly accurate. Also, no caustics occur in the neighborhood about these altitudes in this case. Use of geometric optics is, of course, much easier when corrections for diffraction can be neglected. On the other hand, for the $\Delta N < 0$ case shown in Figs. 2–11(a) and 2–11(b), the dire

predictions of geometric optics of infinite signal power at the contact point with the caustic surface and zero power in the shadow zone are far from accurate. Ray theory based upon second-order theory fails in these regions, or at least it is compromised. In the former region, it fails because $dh_{\mathrm{LG}} / dh \to 0$ at the caustic contact point and, therefore, $\partial^2 \Phi\left(h, h_{\mathrm{LG}}\right) / \partial h^2 = 0$ at that point, rendering the quadratic form of the stationary-phase technique useless. In the latter region, it fails because $\partial \Phi\left(h, h_{\mathrm{LG}}\right) / \partial h \neq 0$ at any altitude; there are no stationary-phase points there. The scalar diffraction integral does a good job with these two situations.

2.7.3 A Boundary Layer

An example of a two-sided boundary is shown in Figs. 2–13(a) through 2-13(d). This example might be found in a marine layer in the lower troposphere with very sharp boundaries. Here the refractivity, shown in Fig. 2–13(a), abruptly within 50 m rises by 5 percent over a very short range compared to the first Fresnel zone from the dry air reference refractivity profile used in Fig. 2–11. For this transition, $r_o n_o' + n_o = -0.2$; therefore, n_o' exceeds the super-refractivity limit of $-n_o / r_o \approx -1 / 6400\ \mathrm{km}^{-1}$. Then at an altitude of 1/2 kilometer lower, the refractivity abruptly drops by 5 percent. The bending-angle profile for this example is shown in Fig. 2–13(b). It is obtained from Eq. (2.2–2), taking care to avoid integrating across a very narrow super-refractivity interval near the upper boundary where $nr_* < a$. Figure 2-13(c) shows the $\left(h, h_{\mathrm{LG}}\right)$ relationship using the thin-screen model. In this figure, the curve for h_{LG} versus h from the reference refractivity profile is dashed below the upper boundary. Figure 2-13(d) shows the Fresnel response in amplitude for this model. The dashed curve in this figure is the defocusing from the dry air reference refractivity profile. Qualitatively, this figure is largely a composite of Figs. 2–11(a) and 2–11(c). The shadow zone due to the strong refractivity gradient over the transition at the upper boundary (where $dN / dh < 0$) is largely filled in by the "throw-back" rays from the reverse transition ($dN / dh > 0$) at the lower boundary. It goes almost without saying that the refractivity transitions across boundary layers in the real atmosphere can and do produce very complicated interference/diffraction patterns in amplitude and phase, such as those shown in Figs. 1–3 and 1–8.

Higher SNR will be called for in future missions if the fine structure in the narrow refracting layers is to be fully studied to their sub-Fresnel resolution limit. This means, for example, the use of an antenna array that is effectively several decimeters in size along the vertical dimension, in-phase and quadrature dual-band coarse acquisition (C/A) recovery of carrier phase, and so on. Sample rates somewhat higher than the canonical 50-Hz GPS data transmission chip rate may prove beneficial.

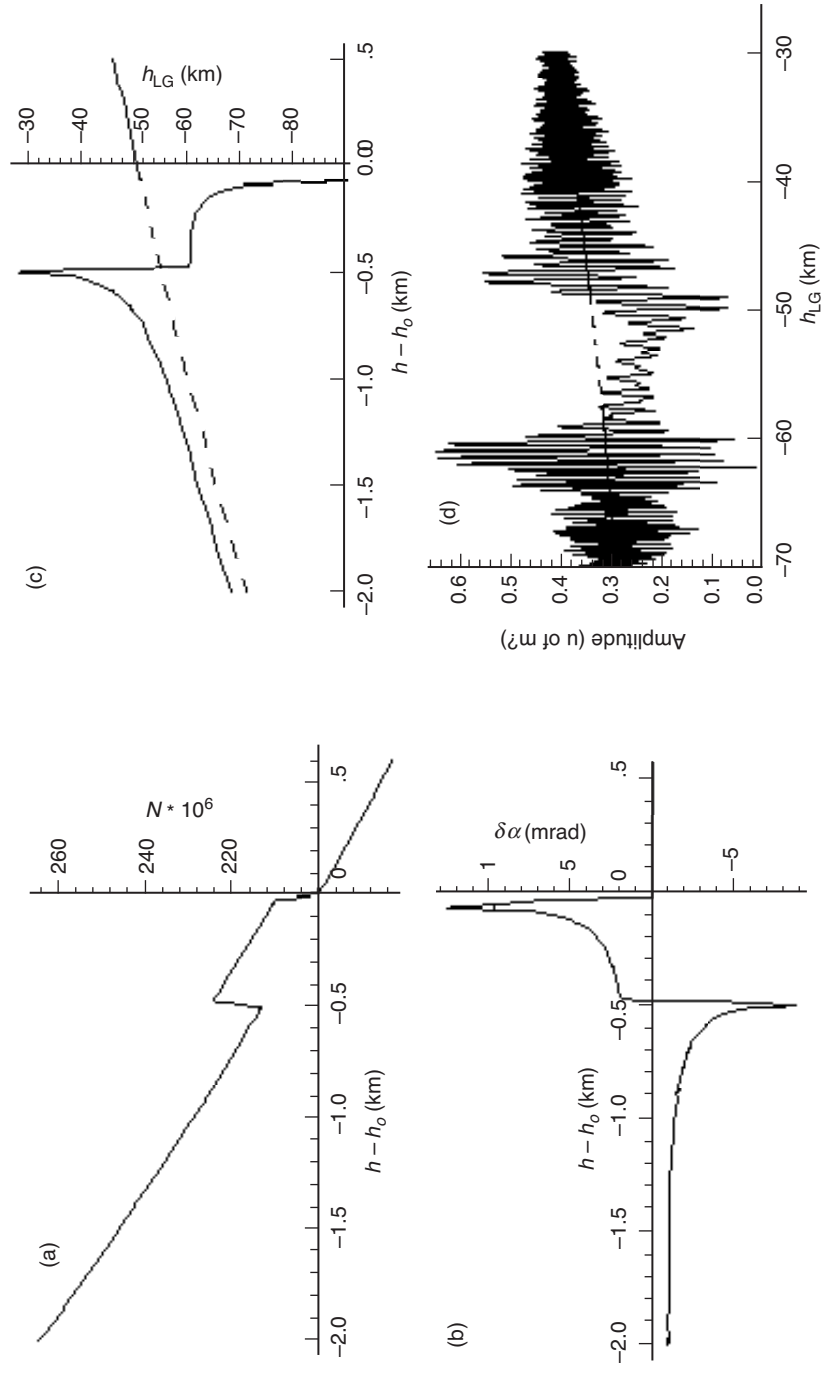

Fig. 2-13. Boundary layer at 4-km altitude with a sharp transition in refractivity of 5 percent at the upper and lower boundaries: (a) the refractivity profile, (b) the resulting bending angle perturbation, (c) the thin-screen relationship, and (d) the amplitude perturbations. The same reference refractivity model was used in Fig. 2-11. The changes in *N* occur over a transition interval of 50 m at each boundary. The upper boundary is super-refractive, $r_o n'_o = -1.3$.

2.8 Sensing a Boundary in the Ionosphere

In the ionosphere, the plasma frequency is typically about 3 MHz. Hence, the ionosphere, except on occasions of severe storms resulting in deep scintillation, is essentially transparent for radio signals at L–band frequencies, with typical lower ionosphere refractivity in the $10^{-5} - 10^{-6}$ range. On the other hand, the Doppler and SNR signatures in observed GPS signals that traverse the ionosphere at grazing angles can reflect a high degree of spatial and temporal variability in electron density. An example is provided in Fig. 2–14 [19], which shows the vertical profiles of the L1 and L2 bending angles obtained from GPS/MET for a particular occultation on April 24, 1995. As the tangency points of the two rays descend through the lower ionosphere (~50 to 250 km) during this occultation, sharp changes in the vertical gradient of the electron density in this neighborhood can be inferred. Around 100 km these profiles reveal a layer with near-discontinuities in electron density across their upper and lower boundaries. Figure 1–3 also shows a sharp transient in SNR_V for this same occultation around 100-km altitude.

The GPS/MET observations for certain occultation profiles show a near blackout or fadeout in SNR in the GPS signal when the point of tangency of the ray at the Earth's limb drops through this kind of boundary in the ionosphere. Figure 2–15 [23] captures in fine detail an example of this fading (associated with a sporadic E-layer) in the same occultation taken on April 25, 1995. It charts at a 50–Hz sample rate the variation of the voltage signal-to-noise ratio

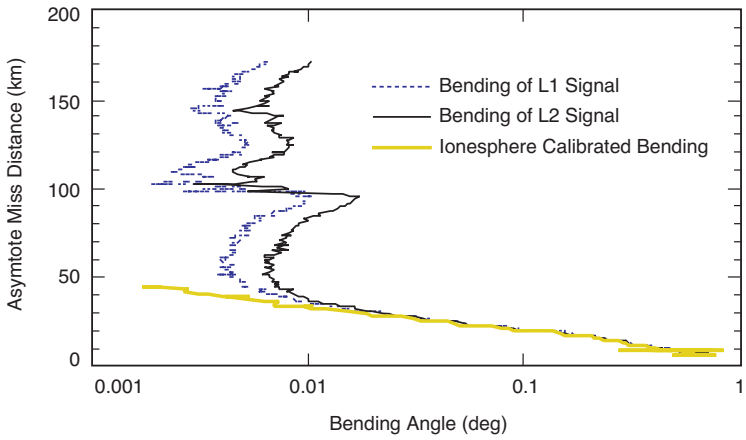

Fig. 2-14. Example of GPS/MET-observed bending angle profiles for the L1 and L2 signals, and also the "ionosphere-free" profile. Redrawn from [18].

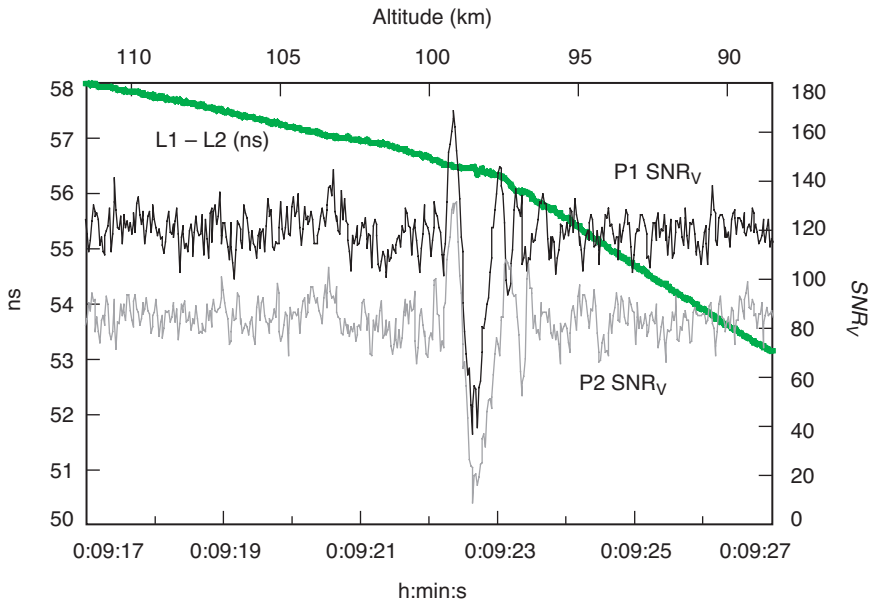

Fig. 2-15. Signal amplitude profiles over a 10-s interval for the L1 and L2 carriers observed by GPS/MET for an occultation on April 25, 1995, sampled at 50 Hz. Amplitude is voltage signal-to-noise ratio SNR_V at 1 s. The phase difference of the L1 and L2 carriers is in nanoseconds. The 100-km altitude of the transient indicates the crossing of a sporadic E-layer. Redrawn from [23].

SNR_V, rated 1-s average,[4] for the L1 and L2 carriers over a 10–s time interval; the variability in the two SNR_V profiles is highly correlated, reflecting true variability in electron density. This figure also shows the change in the difference of the observed L1 and L2 carrier phase [L1–L2] with time (each carrier phase in nanosecond units).

In Fig. 2–16, the excess Doppler for the L1 and L2 carriers for this occultation is shown at 50–Hz and 5–Hz sample rates [24]. The latter is obtained by applying a sliding box filter of 200-ms width to the 50–Hz series. A value of 100 for SNR_V corresponds to a thermal noise in an individual L1 phase measurement over a 20-ms average of about 2 mm. The thermal noise on most 50–Hz Doppler points in Fig. 2–16 is about 0.7 Hz on L1 and 1.0 Hz on L2. Below $SNR_V \cong 30$ the receiver is increasingly likely to encounter difficulty tracking the phase, particularly when significant phase accelerations are also present. On this particular occasion, the receiver maintained lock.

[4] Here SNR_v is expressed as the ratio of the signal amplitude to noise amplitude in voltage that would be obtained by averaging over 1 s, even though the actual averaging time in Figs. 2-15 and 2-16 is 20 ms, the reciprocal of the sample rate.

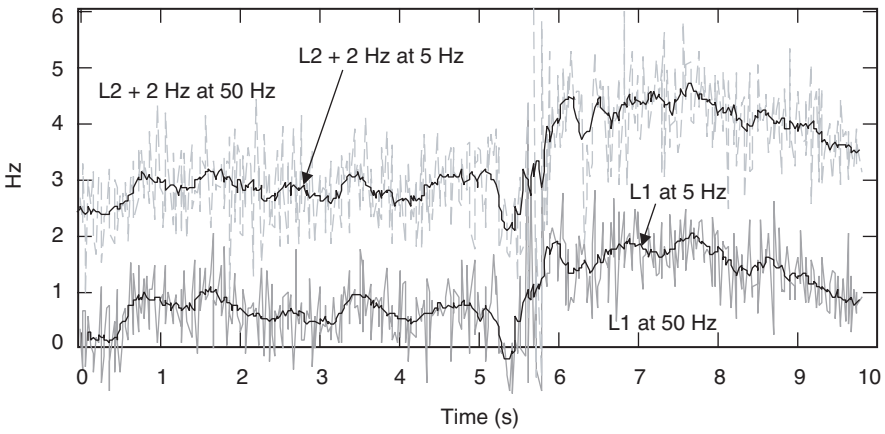

Fig. 2-16. The excess Doppler profile for the same time interval shown in Fig. 2-15. Thermal noise sampled at 50 Hz is about 0.7 Hz (1 σ) on the L1 Doppler and 1.0 Hz for L2. The 5-Hz series are obtained by averaging the 50-Hz points over 200 ms. Redrawn from [23,24].

The Doppler change is greater for L2 after crossing the "boundary" at approximately $t = 5.7$ s. The dip in Doppler just prior to impacting the boundary indicates a negative radial gradient in the electron density transient above the boundary. The abrupt rise in Doppler just afterwards indicates a large positive gradient below.

The precipitous fade and recovery of SNR_V in Fig. 2–15 and the abrupt change in slope of [L1–L2], all occurring within about 1/3 s about coordinated Universal time (UTC) = 00:09:22.7, mark a lower boundary of an ionospheric layer of higher electron density. For this occultation, a very abrupt roll-off of electron density occurs with descending altitude. The scenario shown in Fig. 2–2(c) best matches this situation.

The measured SNR_V and Doppler profiles on L1 and L2 each provide a number of key quantities: the change in Doppler across the fadeout zone (and hence the change in bending angle), the epoch of onset (and hence altitude from POD information and ray tracing) of the fadeout (which may differ on L1 from that on L2 because of ray splitting), the widths of the fadeout zones, and their depths. For the example in Fig. 2–15, the difference in onset of the L1 and L2 rays is about 0.1 s or ~200 m. Using these observed quantities and a model, one can recover, in addition to the altitude of the transition boundary, the change in electron density at the boundary, its gradients on each side of the boundary (with an additional global ionosphere model), the mean distance D of the layer from the LEO, and the separation of the L1 and L2 rays at the boundary.

2.8.1 Fresnel Effects at a Boundary

An example of this approach and a discussion of some Fresnel features are now given. Let the boundary be located at $r = r_o$. We present here the response of the ray in terms of the LEO-observed phase and amplitude using a simple *locally* spherical symmetric model in which the electron density undergoes a sharp change at the boundary but is otherwise smooth. For calculating the change in bending angle due to this roll-off in electron density, the assumption of local spherical symmetry should be useful because of the abrupt nature of the roll-off and the narrowness of its extent in the radial direction. (For the ionosphere at large, the assumption of spherical symmetry has been shown to be unreliable [19].) We will obtain an estimate of the change in electron density profile by assuming that over a short altitude range around $r = r_o$ (roughly 1 to 2 km) the variations in bending angle due to the ionosphere at large can be ignored. We also assume here that the radius of curvature of the ray is very much larger than the local radius of curvature of the boundary, and that the latter equals r_o and is radial directed.

The refractivity ($\times 10^{-6}$) for carrier phase in the ionosphere is given by the principal term

$$N = -\kappa \frac{n_e}{f^2} \tag{2.8-1}$$

where n_e is the electron density, f is the carrier frequency, and $\kappa = 40.3\,\mathrm{m}^3 \cdot s^{-2} \cdot e^{-1}$. A fuller description of ionospheric refractivity, including the higher-order terms, is given in [4,20,21]. Let Δn_e be the change in the electron density in the neighborhood of r_o relative to a reference electron density profile for the ionosphere at large that is assumed to change more gradually with radial distance. For one example of a global reference model, see Appendix E. We assume that the spatial variability of the reference profile is small compared to the magnitude of the local electron density gradient at the roll-off. Inasmuch as the entire fadeout zone in Fig. 2–15 spans only about 1/3 s of time, the boundary width is comparable to or less than the first Fresnel zone for this example; a Fresnel treatment will be needed to predict the observed phase and amplitude. Under these conditions, the bending-angle profile resulting from a discontinuous change in refractivity, as provided by Case A, should be a useful initial approach. The change $\Delta\alpha_S$ below the boundary $r = r_o$ for this case is given by the discrete Snell term from Eq. (2.3–27) by

$$\Delta\alpha_S(r) = \begin{cases} 2\left(\sin^{-1}\left(\frac{n^- r}{n^+ r_o}\right) - \sin^{-1}\left(\frac{r}{r_o}\right)\right) \doteq 2\sqrt{2}\left(\sqrt{v} - \sqrt{\Delta N + v}\right), \; r \le \tilde{r} < r_o \\ 0, \;\; r > r_o \end{cases} \qquad (2.8\text{-}2)$$

where[5] $v = (r_o - r)/r_o = (h_o - h)/r_o$ and $n^+ = n^- + \Delta N$. Here ΔN is obtained from Eq. (2.8–1) and it is a negative quantity when $\Delta n_e > 0$; i.e., when the electron density increases upward across the boundary. The quantity $\tilde{r}$ marks the upper limit of the ray due to the internal critical refraction limit when $\Delta N < 0$; it is defined by the condition $n^- \tilde{r} = n^+ r_o$. When $r \le \tilde{r}$, the signal while traveling through the lower region experiences a slowdown in phase velocity due to the increased number of electron "holes" with depth in that region. For $r > r_o$, $\Delta\alpha(r) = 0$.

Using the thin-screen model and the stationary-phase condition, it follows that the phase delay $\varepsilon(h) = \delta\varphi(h)$ embedded in the thin screen is given by

$$\varepsilon(h) = \begin{cases} k\int_r^{\tilde{r}} \Delta\alpha_S(r')dr' \doteq \frac{2^{5/2}}{3} k r_o \left(v^{3/2} - (-\Delta N)^{3/2} - (\Delta N + v)^{3/2}\right), r < \tilde{r} \\ 0, \;\; r > \tilde{r}, \quad v = (r_o - r)/r_o = (h_o - h)/r_o \end{cases} \qquad (2.8\text{-}3)$$

Figures 2–17 through 2-19 show the change in observed phase and amplitude that result from the diffraction integral in Eq. (2.5–1) using this Case A model. For the L2 carrier, a discontinuous electron density of $\Delta n_e = +4.48\,x10^{10}\ \mathrm{e/m^3}$ corresponds to $\Delta N_{\mathrm{L}2} = -1.2\times10^{-6}$. The abscissa $r_o - r_{\mathrm{LG}}$ is the difference in altitudes of the boundary and the tangent point of the straight line between GPS/MET and the occulted GPS satellite; $r_o - r_{\mathrm{LG}}$ varies nearly linearly with time. Depending on the averaging time used to construct the phase measurements (and hence Doppler), Fig. 2–17 shows that the Fresnel perturbations[6] can be a significant error source for sharp transitions in electron density if unaccounted for in the inversion process that converts Doppler into an electron density.

---

[5] For notational convenience in this section we have dropped the subscript "$_*$" from r_*, the turning point of the ray, i.e., $r \equiv r_*$. Also, in the ionosphere, we can set $r - r_o = h - h_o$ because the nominal defocusing is minimal.

[6] The Fresnel perturbation is the difference between the phase profile shown in Fig. 2–17 minus the profile from geometric optics, which is given by $\Delta = 1/2kD\left(\Delta\alpha_S(v)\right)^2 + \varepsilon(v), \;\; r_o v = r_b + D\Delta\alpha_S(v)$.

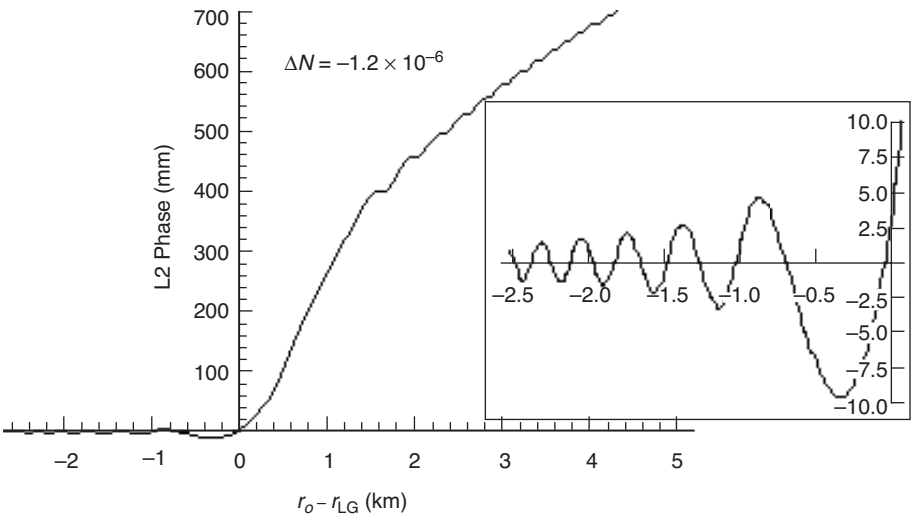

Fig. 2-17. Fresnel perturbations in phase at the LEO from a discontinuity $\Delta N_{L2} = N_{L2}^{+} - N_{L2}^{-} = -1.2 \times 10^{-6}$ at 100-km altitude. For L2, this is a change in electron density of $\Delta n_e = +4.48 \times 10^{10}$ e/m^3. Inset shows perturbations for h_{LG} just above the boundary.

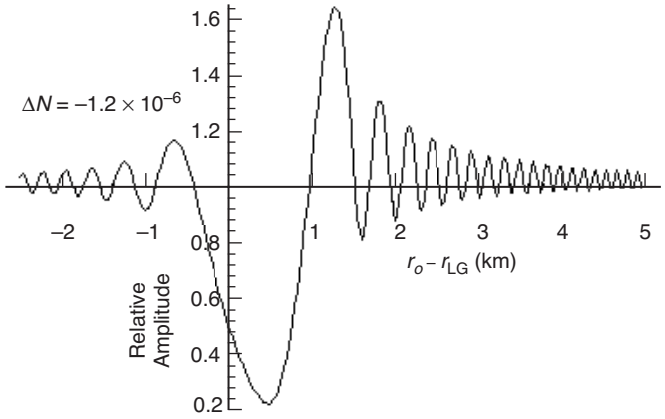

Fig. 2-18. Fresnel perturbation in L2 signal amplitude due to an upward discontinuity $\Delta n_e = +4.48 \times 10^{10}$ e/m^3 at 100-km altitude.

Figure 2–19 shows the Fresnel difference $\mathrm{L1} - \mathrm{L2}$, each carrier measured in range units (ns) using the Case A model with $\Delta n_e = +4.48\,x\,10^{10}\,\mathrm{e/m^3}$. Figures 2–17 through 2–19 span about 3 s of time. For the specific occultation shown in Figs. 2–15 and 2–16, $\dot{r}_{\mathrm{LG}} = -2.3\,\mathrm{km/s}$.

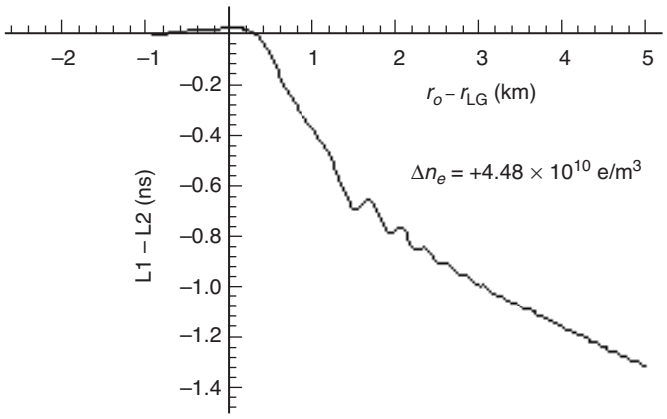

Fig. 2-19. Differenced L1–L2 phase perturbation due to a discontinuity Δn_e at 100-km altitude.

Above the boundary, the L1–L2 profile for the GPS/MET occultation shown in Fig. 2–15 is sloped downward at an average rate of about –0.28 ns/s (which reflects the presence of a more global electron density distribution above the boundary, and which yields average L1 and L2 bending angles of about +0.05 and +0.08 mrad, respectively). Below the boundary, the average slope is about –0.85 ns/s. Thus, the change in average slope as the ray cuts down across the boundary and into the lower region is about –0.57 ns/s. For the small angles involved here, the scale of the observed change in phase delay is nearly proportional to the change in average slope of L1–L2 in Fig. 2–15. One can obtain an estimate of Δn_e by adjusting its value used to generate Fig. 2–19 to obtain a match in the average slope differences of the L1–L2 profiles across the boundary in Figs. 2–15 and 2–19.

2.8.2 Amplitude Effects at a Boundary

The widths of the fadeout zones in SNR_V, shown in Fig. 2–15 for L1 and L2, provide additional information. Let B be a measure of the width in r_{LG}-space across the SNR fadeout zone. A simple measure is the distance between the boundary and the altitude of the first contact with the caustic, near which the flaring would be at maximum. Figure 2–18 shows the Fresnel perturbation in observed signal amplitude using the Case A model with $\Delta N_{L2} = -1.2 \times 10^{-6}$. However, we note from this figure that the Fresnel perturbations on the SNR tend to soften the response of the observed signal to the discontinuity. This somewhat fills in the trough in the SNR profile and widens the distance between the point of maximum Fresnel overshoot just above the boundary and the point of maximum flaring, due to the caustic, just below the boundary. See Appendix D for a discussion of Fresnel effects on

location and magnitude of the flaring. For the value of Δn_e used in Fig. 2–18, the altitude in r_{LG}-space of the first contact point with the caustic based on geometric optics is +1.05 km (see Appendix C). The actual maximum in Fig. 2–18 occurs at about +1.3 km, while the maximum Fresnel overshoot occurs at –0.7 km. Thus, this particular measure for B (from boundary to contact) will underestimate the width, more like a half-width. A correction to B for Fresnel effects can be developed. Asymptotic expressions for the Fresnel integrals can be used to isolate the point of maximum overshoot, and Appendix D provides Fresnel corrections to the position of maximum flaring and its maximum value.

From the thin-screen model using Eqs. (2.2–5) and (2.5–6), and using a power series expansion of Eq. (2.8–2) and the condition that $\partial h / \partial h_{\mathrm{LG}} \rightarrow \infty$ at first caustic contact, one obtains an expression for the "ray" offset $\nu^{\dagger} = \left(r_o - r^{\dagger}\right) / r_o = \Delta r_o / r_o$:

$$\nu^{\dagger} \doteq \frac{1}{2}\left(-\frac{2D\Delta N}{r_o}\right)^{2/3}, \quad \nu^{\dagger} >> |\Delta N| \tag{2.8-4}$$

For B, using the thin-screen model, one obtains

$$B = r_o - r_{\mathrm{LG}}^{\dagger} = r_o \nu^{\dagger} + D\Delta\alpha\left(r^{\dagger}\right) \doteq \frac{3}{2} r_o \left(-2D\Delta N / r_o\right)^{2/3} = 3 r_o \nu^{\dagger} \tag{2.8-5}$$

Let $B_{\mathrm{L}1}$ and $B_{\mathrm{L}2}$ denote the widths for L1 and L2, respectively. Then it follows from Eqs. (2.8-1) and (2.8–5) that

$$\frac{B_{\mathrm{L}1}}{B_{\mathrm{L}2}} \doteq \left(\frac{f_{\mathrm{L}2}}{f_{\mathrm{L}1}}\right)^{4/3} \tag{2.8-6}$$

which appears to hold roughly for Fig. 2-15. From Eqs. (2.8-4) and (2.8-5), one obtains for Δr and B the values $\Delta r = 350\,\mathrm{m}$ and $B_{\mathrm{L}1} \approx 1.0\,\mathrm{km}$ for $\Delta N_{\mathrm{L}1} = -1.2 \times 10^{-6}$. For $\dot{h}_{\mathrm{LG}} = -2.3\,\mathrm{km/s}$, this yields a temporal width of about 1/3 s, which is about the observed value in Fig. 2-15 for the half-width.

Figure 2-18 also shows a small asymptotic bias in amplitude above unity for the lower altitude range, which is a lense effect caused by the positive bending angles below the boundary.

2.8.3 Ray Splitting

The information about the change in slopes of the L1 and L2 carrier phase profiles and about the fadeout width provides redundant determinations of Δn_e; one can use this redundancy to recover an estimate for D. Let p denote a

parameter such as the RF carrier frequency; in this case $p = f^{-2}$. Let $\psi(r_p, r_{\mathrm{LG}})$ be the excess carrier phase observed by the LEO—that is, the difference of the observed phase minus the predicted phase in the absence of the atmosphere and ionosphere. Here $r_p - R_E$ is the altitude of the point of tangency with the Earth's limb for the ray path corresponding to the parameter value *p*; similarly, r_{LG} is the position of the straight line between the LEO and the observed GPS satellite at its point of tangency. We use the subscript *p* on the ray path tangency point to denote the dependence of r_p at a fixed epoch (and, therefore, for a given value of r_{LG} on the parameter *p*, that is, $r_p = r_p[p, r_{LG}]$). If one varies *p* while holding the end points fixed and hence r_{LG} fixed, then r_p also must vary in such a way that the new ray path satisfies the stationary-phase condition. The observable α_{p_k} is given to first order in α_{p_k} (Appendix A) from the excess Doppler by

$$\alpha_{p_k} = \alpha[p_k, r_{p_k}] \doteq -\frac{\lambda_k}{2\pi \dot{r}_{\mathrm{LG}}} \dot{\psi}(r_{p_k}, r_{\mathrm{LG}}) = -\frac{c}{\dot{r}_{\mathrm{LG}}} \dot{L}k \tag{2.8-7}$$

where $\mathrm{L}k$ is the *k*th carrier phase in time units and *c* is the velocity of light. We have written $\alpha_p = \alpha[p, r_p]$ to reflect the direct and indirect dependency of the bending angle on the parameter *p*. The quantities $\dot{\psi}$ and $\dot{r}_{\mathrm{LG}}$ are known from the observations, and the POD information is known for the occulted GPS satellite and for GPS/MET. Therefore, bending-angle profiles, such as those shown in Fig. 2-14, follow directly from the Doppler measurements and the POD information. The quantities $\dot{\psi}_{\mathrm{L1}}$ and $\dot{\psi}_{\mathrm{L2}}$ in hertz are shown in Fig. 2-16.

Because of ionosphere-induced ray splitting, the tangent points for the L1 and L2 rays will not quite be the same for a given epoch. For example, there is ~0.1 s difference in the epochs of onset of the L1 and L2 fadeout zones in Fig. 2-15. This reflects an altitude difference due to ray splitting from the ionosphere at large that is of the order of 100 m. Figure 2-20 shows the profile of the ray splitting term $D(\alpha_{\mathrm{L1}} - \alpha_{\mathrm{L2}})$ using Eq. (2.8-7) for the same occultation shown in Figs. 2-15 and 2-16. Figure 2-20 shows the profile for a 50-Hz sample rate, and it shows the profile for a 5-Hz sample rate, which is derived by applying a sliding box filter of 200-ms width to the 50-Hz time series. The magnitude of the thermal noise reduces as the square root of the ratio of the sample rates, but the magnitude of the Fresnel fringes, which are essentially sinusoidal over sampling widths used in Fig. 2-20, reduces linearly. After contacting this particular ionospheric layer, the ray splitting essentially triples in magnitude. We should account for this (usually) small difference in altitude or, equivalently, a bending-angle difference for this example of the order of a tenth of a milliradian.

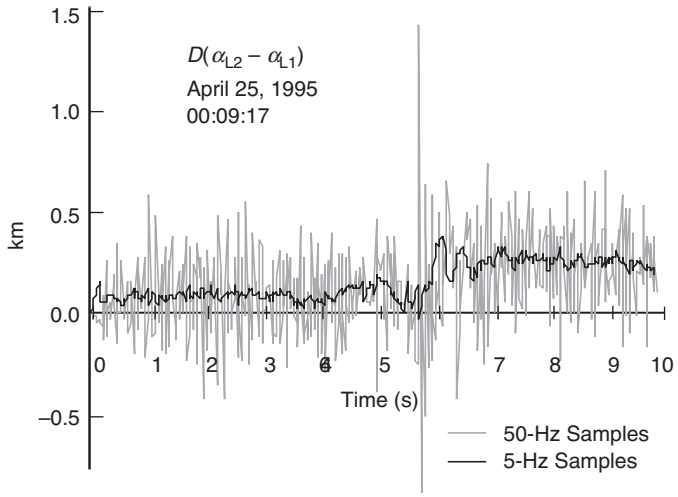

Fig. 2-20. Ray path splitting of the L1 and L2 signals for the same occultation and epoch shown in Figs. 2-15 and 2-16.

Reference [22] uses a linear combination of bending angles to eliminate ionosphere effects to second degree in frequency from the observations. We use a similar approach here to account for ray splitting. To a good approximation, the total bending can be decomposed into

$$\alpha_p = \alpha\left[p, r_p\right] = \alpha^A\left(r_p\right) + \alpha^I\left(p, r_p\right) \tag{2.8-8}$$

where α^A and α^I are, respectively, the atmospheric and ionospheric components of the bending. However, when the tangent point of the ray lies in the ionosphere, $\alpha^A \equiv 0$. Note that α^I depends on p directly [through the refractivity given by Eq. (2.8-1)] and also indirectly through r_p, which for a fixed epoch also depends on p. That is,

$$\alpha^I\left(p, r_p\right) \doteq pI\left(r_p\right) \tag{2.8-9}$$

where $I\left(r_p\right)$ is the path integral $\kappa\int \nabla n_e \times d\mathbf{s}$ along the ray through the ionosphere from the LEO to the GPS satellite with an impact parameter value r_p. For example, when the ray path tangency point lies in the ionosphere and if spherical symmetry were to apply,

$$I\left(r_p\right) \doteq \kappa r_p\left(\int_{r_p}^{r_{\mathrm{LEO}}} \frac{n_e'}{\sqrt{r^2 - r_p^2}}dr + \int_{r_p}^{\infty} \frac{n_e'}{\sqrt{r^2 - r_p^2}}dr\right) \tag{2.8-10}$$

Equation (2.8-10) has been simplified somewhat relative to Eq. (2.2-2) by setting the index of refraction terms in the integral to unity; in the upper stratosphere and the ionosphere, the differences in these terms from unity is very small. Also, the third- and fourth-order ionosphere terms have been ignored (except indirectly through ray splitting).

At a fixed epoch where r_{LG} is fixed, that is, when the end points of the ray are fixed, varying p will cause r_p also to vary so as to maintain the stationary-phase condition on the resulting ray path. Therefore,

$$\frac{d\alpha}{dp} = \frac{\partial\alpha}{\partial p} + \frac{\partial\alpha}{\partial r_p}\frac{dr_p}{dp}, \quad r_{\mathrm{LG}} \text{ fixed} \tag{2.8-11}$$

But using the thin-screen approximation [Eq. (2.2-5)],

$$\frac{dr_p}{dp} = D\left(\frac{\partial\alpha}{\partial p} + \frac{\partial\alpha}{\partial r_p}\frac{dr_p}{dp}\right) \tag{2.8-12}$$

or

$$\frac{dr_p}{dp} = \zeta_p D\frac{\partial\alpha}{\partial p} \tag{2.8-13}$$

where ζ_p is the defocusing factor

$$\zeta_p^{-1} = 1 - D\frac{\partial\alpha}{\partial r_p} = 1 - D\left(\frac{d\alpha^A}{dr_p} + p\frac{dI}{dr_p}\right) \tag{2.8-14}$$

The differential expression in Eq. (2.8-13) should be valid provided $\left|\zeta^{-1}\right| > 0$, but its validity will progressively worsen as the ray nears a caustic surface where $\zeta^{-1} \to 0$. Combining Eqs. (2.8-11) and (2.8-13) yields

$$\frac{d\alpha}{dp} = \zeta_p\frac{\partial\alpha}{\partial p} = \zeta_p\frac{\partial\alpha^I}{\partial p} = \zeta_p I\left(r_p\right) \tag{2.8-15}$$

Let $\bar{r}$ be the position of the tangent point of the L1 and L2 rays in the absence of an ionosphere, or, equivalently, when $p = 0$. We expand $\alpha = \alpha\left[p, r_p\right]$ about this position to obtain

$$\alpha_p = \alpha\left[p, r_p\right] = \alpha[0, \bar{r}] + p\left(\frac{d\alpha_p}{dp}\right)_{\bar{r}} + \frac{1}{2}p^2\left(\frac{d^2\alpha_p}{dp^2}\right)_{\bar{r}} + \cdots \tag{2.8-16}$$

where $\bar{r}$ is given to first order in p by

$$\bar{r} = \frac{p_1 r_{p_2} - p_2 r_{p_1}}{p_1 - p_2} = \frac{f_1^2 r_{p_1} - f_2^2 r_{p_2}}{f_1^2 - f_2^2} \doteq 2.546 r_{p_1} - 1.546 r_{p_2} \qquad (2.8\text{-}17)$$

Evaluating Eq. (2.8-16) at p_1 and p_2 and linearly combining to eliminate the first-order term in p, we obtain

$$\alpha_3 = \frac{p_1 \alpha_{p_2} - p_2 \alpha_{p_1}}{p_1 - p_2} = \alpha[0, \bar{r}] - \frac{1}{2} p_1 p_2 \left(\frac{d^2 \alpha_p}{dp^2} \right)_{\bar{r}} + \cdots \qquad (2.8\text{-}18)$$

or, upon using Eq. (2.8-15),

$$\alpha_3 \doteq \alpha[0, \bar{r}] - \frac{1}{2} D \left[\zeta_{p_1} \zeta_{p_2} \frac{\partial}{\partial r_p} \left(\alpha_{p_1}^I \alpha_{p_2}^I \right) \right]_{\bar{r}} \qquad (2.8\text{-}19)$$

Figure 2-21 shows this linear combination of bending angles using the excess L1 and L2 Doppler data from the occultation shown in Fig. 2-16. At these altitudes (~100 km), α_3 should be zero. It is not quite zero because of ray splitting and the presence of other higher-order effects. Inasmuch as the magnitude of the bending angle itself from the ionosphere is only 0.1 to 0.2 mrad, the error in the α_3 linear combination is as large as the correction itself, which suggests that an alternative technique for correcting for ray splitting and higher-order effects might be advisable [19].

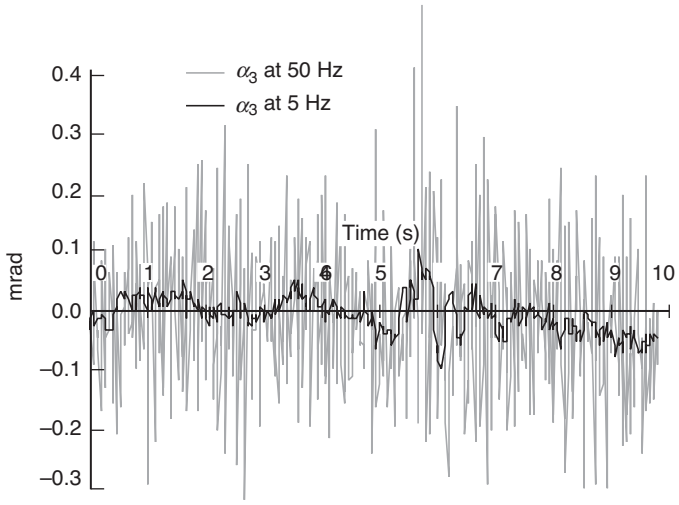

Fig. 2-21. Ionosphere-free bending angle α_3 for the same occultation and epoch shown in Figs. 2-15, 2-16, and 2-20.

The ray-splitting term involves the gradient of the square of the ionospheric bending, which is a consequence of the stationary-phase condition on the ray. When $\bar{r}$ lies above the atmosphere, $\alpha[0,\bar{r}] \equiv 0$, and for this case we have

$$\left.\begin{aligned} \alpha_3 &\doteq -Dp_1p_2\left[\frac{I(r)I^{'}(r)}{\left(1-Dp_1I^{'}(r)\right)\left(1-Dp_2I^{'}(r)\right)}\right]_{r=\bar{r}} \\ \alpha^A(\bar{r}) &\equiv 0 \end{aligned}\right\} \tag{2.8-20}$$

Thus, in the ionosphere the linear combination of the observables α_{p_1} and α_{p_2} that forms α_3 also provides a measure of ray-splitting effects. For the case of spherical symmetry and where n_e denotes the local transient in electron density, the integral $I'\left(r_p\right)$ is given by

$$I'\left(r_p\right) \doteq 2\kappa r_p\int_{r_p}^{\infty}\frac{\sqrt{r^2-r_p^2}}{r^2}n_e'dr + 2\kappa r_p^2\int_{r_p}^{\infty}\frac{n_e''}{r\sqrt{r^2-r_p^2}}dr \doteq 2\kappa r_p\int_{r_p}^{\infty}\frac{n_e''}{\sqrt{r^2-r_p^2}}dr \tag{2.8-21}$$

which normally is dominated by n_e'' for r near r_p when r_p lies in the ionosphere. Also, from Eq. (2.8-16) we can obtain determinations of $I\left(r_p\right)$ and $I'\left(r_p\right)$ from the observables α_{p_1} and α_{p_2} when $\bar{r}$ lies in the ionosphere. One obtains

$$I(\bar{r}) = \frac{\alpha_{p_1}\alpha_{p_2}\left(p_2-p_1\right)}{p_1p_2\left(\alpha_{p_2}-\alpha_{p_1}\right)},\ \alpha[0,\bar{r}] \equiv 0 \tag{2.8-22}$$

and

$$I^{'}(\bar{r}) = \frac{1}{D}\frac{\alpha_{p_2}p_1-\alpha_{p_1}p_2}{p_1p_2\left(\alpha_{p_2}-\alpha_{p_1}\right)},\ \alpha[0,\bar{r}] \equiv 0 \tag{2.8-23}$$

We note that the ray splitting itself is given by

$$r_{p_2}-r_{p_1} = D\left(\alpha_{p_2}-\alpha_{p_1}\right) \doteq D\left(p_2-p_1\right)\left(\zeta_p I(r)\right)_{\bar{r}} \tag{2.8-24}$$

A suggestion that has been made independently by several investigators is to offset the epochs of the Doppler measurements to eliminate (to first order) the ionosphere effects on derived bending angles—in effect, to let the L1 ray

"catch up" with the L2 ray. Let Δt_p be the offset of the observational epoch on the "p_kth" carrier for which the change in derived bending angle at that offset epoch nulls (to first order) the dispersive term in Eq. (2.8-16). That is, we require that $\Delta t_p \, d\alpha_p / dt = -\zeta_p \, \alpha_p^I$. From Eqs. (2.2-5) and (2.2-6), we have

$$\frac{d\alpha_p}{dt} = \dot{r}_b \zeta_p \frac{\partial \alpha_p}{\partial r_p} \tag{2.8-25}$$

and, from Eq. (2.8-16), we obtain by expanding through the first derivative in time

$$\begin{aligned} \alpha_p\Big|_{t=t_o+\Delta t_p} &= \alpha_p\Big|_{t_o} + \Delta\alpha_p = \alpha[0,\bar{r}] + \zeta_p \alpha_p^I \\ &+ \left(\dot{r}_b \zeta_p \frac{\partial}{\partial \bar{r}} \alpha[0,\bar{r}] + \dot{r}_b \zeta_p \frac{d}{d\bar{r}} \left(\zeta_p \alpha_p^I \right) \right) \Delta t_p \end{aligned} \tag{2.8-26}$$

from which it follows that to first order

$$\Delta t_p = -\left[\frac{\alpha_p^I}{\dot{r}_b} \left(\frac{\partial}{\partial r_p} \alpha[0, r_p] \right)^{-1} \right]_{r_p = \bar{r}} \tag{2.8-27}$$

Forming the "α_3" linear combination again, but with an offset epoch applied to each carrier according to Eq. (2.8-27), one obtains, when $D|\partial\alpha/\partial r| << 1$,

$$\begin{aligned} \tilde{\alpha}_3 = \frac{p_1 \alpha_{p_2}\Big|_{t_o+\Delta t_2} - p_2 \alpha_{p_1}\Big|_{t_o+\Delta t_1}}{p_1 - p_2} &= \alpha[0,\bar{r}] - D\Bigg[\left(\zeta_{p_1} \zeta_{p_2} \right)^2 \\ \times \left\{ \left(\frac{\partial \alpha}{\partial r} \right) \frac{\partial}{\partial r} \left(\alpha_{p_1}^I \alpha_{p_2}^I \right) - \zeta_{p_1} \zeta_{p_2} \alpha_{p_1}^I \alpha_{p_2}^I \frac{\partial^2 \alpha}{\partial r^2} \right\} &\left(\frac{\partial \alpha}{\partial r} \right)^{-1} \Bigg]_{r=\bar{r}} + \cdots \end{aligned} \tag{2.8-28}$$

When $D|\partial\alpha/\partial r| >> 1$, for example, when $\bar{r}$ is located low in the troposphere, then the quadratic error term in Eq. (2.8-28) is more complicated, but it still will carry the quartic power of the defocusing factors ζ_{p_1} and ζ_{p_2} as the leading term. This factor will be small, so the error term usually should be small.

2.8.4 Doppler Information at a Boundary

In the following, the subscript "p_k" is replaced by "Lk", $k = 1,2$. Returning to the problem of determining Δn_e across a boundary, two alternate approaches

can be followed. One is to de-trend the L1 and L2 Doppler profiles using the POD-provided profile of ρ, where ρ is the LEO–to–GPS satellite range; the bending angles $\alpha_{\mathrm{L}1}$ and $\alpha_{\mathrm{L}2}$ follow directly from the excess Doppler. This is the approach that was used in [19] to obtain the results shown in Fig. 2-14. The changes in $\alpha_{\mathrm{L}1}$ and $\alpha_{\mathrm{L}2}$ across the boundary form an Abel transform with Δn_e if local spherical symmetry is assumed, which might be a good assumption given the limited altitude range over which Δn_e applies ($\Delta h = \sim 1\,\mathrm{km}$ in Fig. 2-15, and an equivalent horizontal span of $2(\Delta h r_o)^{1/2} = \sim 200\,\mathrm{km}$). Applying the Abel transform to these differenced observable streams would yield Δn_e; however, if the transient is particularly sharp compared to the size of the first Fresnel zone, then we would expect imperfect recovery due to the limitations in geometric optics.

The other approach is simply to use the profile of L1–L2, which does not require POD information. For this approach, we have

$$\left[\alpha_{\mathrm{L}1}(r_{\mathrm{L}1}) - \alpha_{\mathrm{L}2}(r_{\mathrm{L}2})\right] = -\frac{c}{\dot{r}_{\mathrm{LG}}}\left[\dot{\mathrm{L}}1 - \dot{\mathrm{L}}2\right] + \mathrm{O}\left[\alpha^2\right] \tag{2.8-29}$$

When spherical symmetry applies, $\Delta\left[\alpha_{\mathrm{L}1}(r) - \alpha_{\mathrm{L}2}(r)\right]$ and $\Delta\left[N_{\mathrm{L}1}(r) - N_{\mathrm{L}2}(r)\right]$ form an Abel transform pair. However, the quantity [L1–L2] in Fig. 2-15 is formed from the measured phases of the two carriers at a common epoch and is so formed to eliminate the geometric range term between the LEO and the occulted GPS satellite. As Eq. (2.8-29) shows, [L1–L2] is affected by ray splitting, and it should be corrected if $\left[r_{L1} - r_{L2}\right]$ significantly differs from zero. Following the earlier discussion and using Eq. (2.8-16), one obtains for the change in the [L1–L2] profile in the neighborhood of the boundary

$$\Delta\left[\dot{\mathrm{L}}1 - \dot{\mathrm{L}}2\right] \doteq$$
$$\left(\frac{\dot{r}_b(p_1 - p_2)\Delta I}{c\left(1 - Dp_1 I'\right)\left(1 - Dp_2 I'\right)()}\right) = \frac{\dot{r}_b}{c}\left[\zeta_{p_1}\zeta_{p_2}\Delta\left[\alpha_{L1} - \alpha_{L2}\right]\right]_{r_p = \bar{r}} \tag{2.8-30}$$

If $\zeta_p^{-1} = 1 - DpI'$ is near unity, one obtains $\Delta I(\bar{r})$ directly from Eq. (2.8-30); otherwise, Eq. (2.8-30) can be evaluated by iteration, first obtaining the profile for $\left[\alpha_{\mathrm{L}1}(r_{\mathrm{L}1}) - \alpha_{\mathrm{L}2}(r_{\mathrm{L}2})\right]$ from Eq. (2.8-29) and then using this profile over time to obtain an estimate of *I*. Alternatively, one can obtain an estimate of $r_{\mathrm{L}1} - r_{\mathrm{L}2}$ from the difference in epochs of fadeout onset; use of Eq. (2.8-24) and an assumed value for D yields an estimate of *I*.

Finally, correcting Eq. (2.8-29) for ray splitting and applying an Abel transform (when spherical symmetry holds) to Eq. (2.8-30) yields Δn_e in the vicinity of r_o:

$$\begin{aligned}\Delta n_e(\bar{r}) &\doteq \frac{f_1 f_2}{\kappa\pi}\frac{f_1 f}{f_1^2 - f_2^2{}_2}\int_{\bar{r}}^{\infty}\frac{\Delta[\alpha_{L1}(r)-\alpha_{L2}(r)]}{\sqrt{r^2-\bar{r}^2}}dr \\ &= 3.03x10^{16}\int_{\bar{r}}^{\infty}\frac{\Delta[\alpha_{L1}(r)-\alpha_{L2}(r)]}{\sqrt{r^2-\bar{r}^2}}dr\end{aligned} \tag{2.8-31}$$

With the value for Δn_e obtained from Eq. (2.8-31) and using the L1–L2 profile, one obtains D from Eq. (2.8-5), which is

$$D = r_o\frac{f_{Lk}^2}{2\kappa\Delta n_e(r_o)}\left(\frac{2B_{Lk}}{3r_o}\right) \tag{2.8-32}$$

The value for *D* so obtained is around 2800 km for the occultation in Fig. 2-15.

2.8.5 Fresnel Effects Using an Improved Electron Density Model

One can improve on the Case A model somewhat. Comparing Figs. 2-15 and 2-18, one notes that a simple model given by a discontinuous electron density at a boundary is not sufficient to explain all of the features in Fig. 2-15 in the vicinity of r_o. Prior to onset of the fadeout zone, Fig. 2-15 shows an SNR_V flaring relative to the base SNR_V that is approximately 33 percent in L1 and nearly 50 percent in L2, which are far in excess of the Fresnel overshoot shown in Fig. 2-18. This suggests that another caustic geometry applies just above the boundary, in addition to the caustic below. In this case, it follows that, in addition to the large positive gradient in electron density just below the boundary, the gradient must become negative just above the boundary (to obtain a positive gradient in refractivity in this region and, therefore, a negative increment in bending angle). Moreover, the dip in the excess Doppler profiles shown in Fig. 2-16 just prior to onset of the fadeout followed by an abrupt rise just afterwards also suggests a sharp negative gradient in electron density above the boundary and a positive gradient below. Apparently, the electrodynamic and electrochemical environment in this region can result in a concentration of free electrons confined to a narrow range of altitudes at roughly 100 km just above the recombination zone. This is depicted in Fig. 2-22. A sporadic E-layer roughly fits this description.

We use a simpler model that will generate caustics above and below the boundary as well as the sharp negative and positive changes in excess Doppler. It is given by

$$\Delta n_e = \begin{Bmatrix} \dfrac{\Delta n_e(r_o)}{1+\gamma}\left[1+\gamma\exp\left(-\beta(r-r_o)\right)\right], & r \geq r_o \\ 0, & r < r_o \end{Bmatrix} \tag{2.8-33}$$

Following the same methodology as given for Case A (with $\alpha^+ = 0$), three regimes will apply: the region above the boundary, the narrow region between the critical refraction altitude up to the boundary, and the region below the critical refraction altitude. One obtains for $\Delta\alpha(r)$ in the two regimes

$$\Delta\alpha = \begin{Bmatrix} \dfrac{\Delta N_o \gamma}{1+\gamma}\sqrt{2\pi\beta r}\,\exp\left(-\beta(r-r_o)\right), & r \geq r_o \\ \dfrac{\Delta N_o \gamma}{1+\gamma}\sqrt{2\pi\beta r}\,\mathrm{F}[\sigma] + \Delta\alpha_S, & r \leq \tilde{r} < r_o \end{Bmatrix} \tag{2.8-34}$$

where $\Delta\alpha_S$ is the discrete Snell bending-angle term given in Eq. (2.8-2), $\Delta N_o = -\kappa\Delta n_e(r_o)/f^2$, $\mathrm{F}[\sigma]$ is defined by Eq. (2.3-28), $\sigma = \sqrt{\beta(r_o - r)}$, and $\tilde{r} = (1+\Delta N_o)r_o$, which denotes the critical refraction point for $\Delta n_e(r_o) > 0$. The corresponding refractivity and bending-angle profiles are shown in Figs. 2-23 and 2-24. This would produce a dip in excess Doppler prior to onset of fadeout and an abrupt rise just after onset, such as shown in Fig. 2-16.

The free parameters in this model are ΔN_o, γ, and β, and we can adjust them using the relationships developed in this section to best fit the observed profiles in Figs. 2-15 and 2-16. Using the stationary-phase condition for the

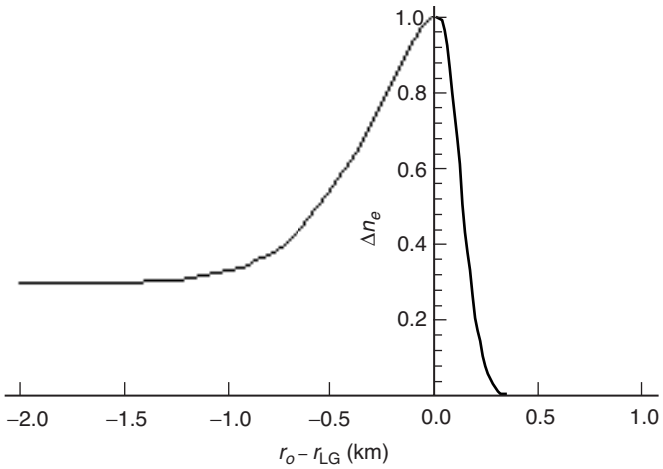

Fig. 2-22. Trial electron density profile near the boundary, $r = r_o$.

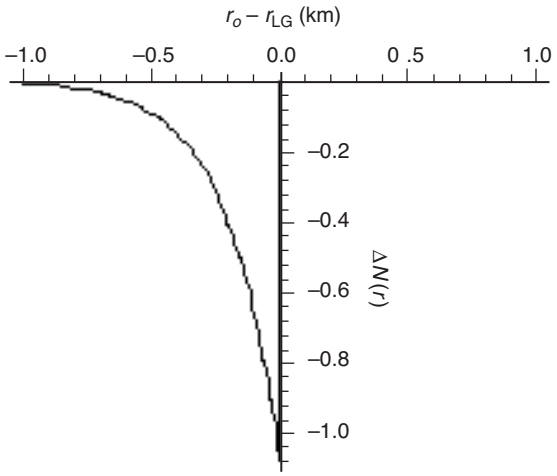

Fig. 2-23. Trial refractivity profile according to Eq. (2.8-33).

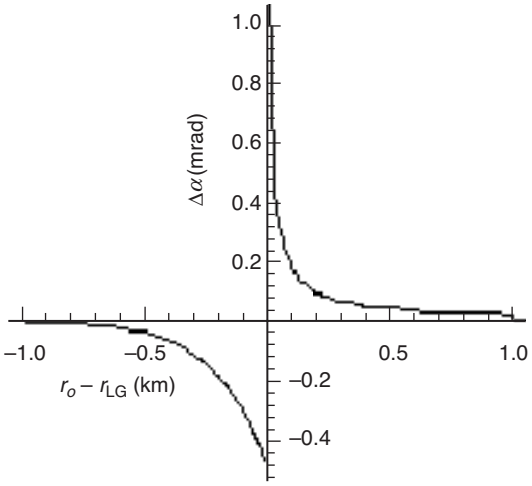

Fig. 2-24. Refractive bending angle from Eq. (2.8-34).

thin-screen model, one integrates Eq. (2.8-34) to obtain the phase delay $\varepsilon(h)$, which is to be embedded in the thin screen. It is given by

$$\varepsilon(h) = \left\{ \begin{array}{l} \dfrac{2\pi}{\lambda\beta}\alpha(r), \quad r > \tilde{r} \\ \dfrac{2\pi}{\lambda\beta}\left[\alpha(r) - \Delta\alpha_S(r) + 2\sigma\right] + \varepsilon_S(r), \quad r \le \tilde{r} < r_o \end{array} \right\} \tag{2.8-35}$$

where $\Delta\alpha_S$ and ε_S are the Snell terms given by Eqs. (2.8-2) and (2.8-3), respectively, and $h = r - R$. From the Rayleigh–Sommerfeld integration in Eq. (2.5-1), one obtains the observed Fresnel response to this kind of electron density distribution.

Figure 2-25 shows the Fresnel perturbations in observed signal amplitude in the vicinity of r_o for $\Delta N_o = -1.2 \times 10^{-6}$, $\gamma = 10$, and $\beta = 5\,\text{km}^{-1}$. The similarity of Fig. 2-25 in the fadeout zone, including the neighboring crests due to flaring, to those in Fig. 2-15 should be noted. Figure 2-26 shows the excess Doppler, including the Fresnel perturbations that result from the assumed electron density model given in Eq. (2.8-33). The black curve is a smoothed version obtained by applying a 1/4-s-wide box filter to the 40-Hz samples. The qualitative agreement between Figs. 2-16 and 2-26 should be noted in the vicinity of $r_{\text{LG}} = r_o$. The abscissa range shown in Figs. 2-25 and 2-26 is equivalent to about 3 s of observations.

Fresnel expressions for estimating minimum signal amplitude in a trough or maximum signal amplitude at the crest of a flaring region, and for calculating the positions of these features in h_{LG}-space, will be found in Appendix D. However, from geometric optics we can estimate the altitude of the flaring using the condition that $dr_{\text{LG}} / dr = 0$ at the first contact point with a caustic. Using the thin-screen relationship in Eqs. (2.2-5) and (2.8-34) for $\Delta\alpha$, one obtains for the position $r_{\text{LG}}^{\dagger}$ of the first contact point at the upper caustic

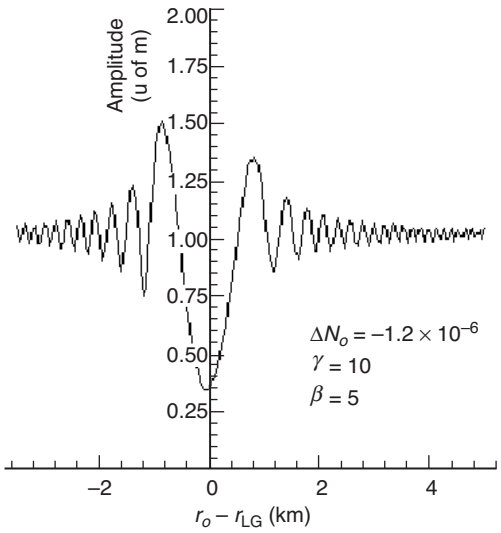

Fig. 2-25. Fresnel perturbation in signal amplitude using the thin-screen phase profile from Eq. (2.8-35). This figure should be compared with Fig. 2-15.

$$r_{\mathrm{LG}}^{\dagger} = r_o + r^{\dagger} + \beta^{-1} = r_o + \frac{1}{\beta}\left(\log\left[-\Delta N_o D\gamma \frac{\left(2\pi r_o \beta^3\right)^{1/2}}{1+\gamma} \right] + 1 \right) \qquad (2.8\text{-}36)$$

which predicts the contact point (where flaring should be near maximum) at $r_o - r_{\mathrm{LG}}^{\dagger} = -0.6\,\mathrm{km}$, or about 0.2 s prior to the time of impact of the ray with the boundary. The exact location of the point of maximum amplitude in Fig. 2-25, which can be accurately predicted using the third-order theory in Appendix D, is $r_o - r_{\mathrm{LG}}^{\dagger} = -0.84\,\mathrm{km}$.

Finally, we can apply the Abel transform to the L1 Doppler series shown in Fig. 2-26, which is based on the thin-screen model using the refractivity model given in Eq. (2.8-33) and shown in Fig. 2-23. This will give us one qualitative estimate of the fidelity of the Abel transform for two transient situations: for $r_{\mathrm{LG}} > r_o$ the transition scale in refractivity is comparable to the scale of the first Fresnel zone, but at $r_{\mathrm{LG}} = r_o$ there is a sharp transition, a discontinuity $\Delta N_o = -1.2 \times 10^{-6}$, in this case. The recovery of the refractivity profile based on the Abel transform and the original model are shown in Fig. 2-27, as too is the truth model $N(r)$ given by Eq. (2.8-33). Here $N(-\infty) = \Delta N_o / (1+\gamma)$. Even the smoothed profile (at 4 Hz) shows a bias as a result of the sharp transition in true refractivity at the boundary.

2.9 The Error in the Recovered Refractivity Resulting from Fresnel Phase Perturbations

Since the Abel transform is a linear operator, we can use it, if spherical symmetry applies, to assess the error in the recovered refractivity that results

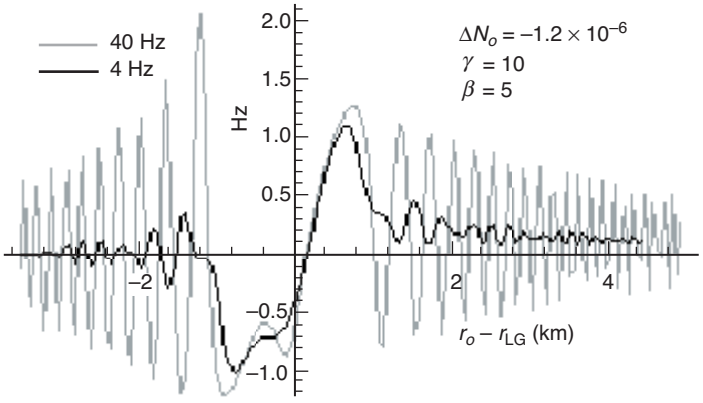

Fig. 2-26. Residual excess L1 Doppler including Fresnel perturbations for the electron density model given in Eq. (2.8-33). This figure should be compared with Fig. 2-16.

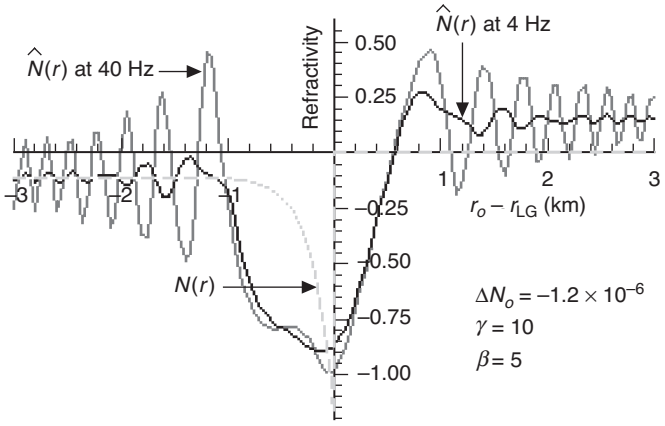

Fig. 2-27. Recovered refractivity profile $\hat{N}(r)$ obtained from the Abel transform versus the truth model $N(r)$ given by Eq. (2.8-33).

from using a geometric optics approach when Fresnel phase perturbations are present (see Appendix F). Let $\delta\hat{N}$ ($\times 10^{-6}$) be the error in recovered refractivity, which is obtained from the Abel transform by the expression

$$\delta\hat{N}(a) = \frac{1}{\pi}\int_a^\infty \frac{\delta\alpha(\mu)}{\sqrt{\mu^2 - a^2}}d\mu \tag{2.9-1}$$

where a is the impact parameter for the ray path. For error assessment purposes, it can be taken as $R_E + h_o$, the altitude in the thin screen corresponding to the real altitude of the tangency point of the ray to the Earth's limb, plus the Earth's radius R_E. Now the error in the inferred bending angle is given to a good approximation by

$$\delta\alpha = \frac{\lambda}{2\pi}\frac{\delta\dot{\varphi}}{\dot{h}_{\mathrm{LG}}} \doteq \frac{\lambda}{2\pi}\frac{d(\delta\varphi)}{dh_{\mathrm{LG}}} \tag{2.9-2}$$

where $\delta\varphi$ is the phase perturbation caused by some error source. Let $\delta\varphi_F(h_{\mathrm{LG}})$ be the phase perturbation due to Fresnel effects, such as that shown in Fig. 2-10. Then Eq. (2.9-1) becomes

$$\delta N_F(h) = \frac{\lambda}{2\pi^2}\int_h^\infty \frac{d(\delta\varphi_F)}{dh_{\mathrm{LG}}}\frac{1}{\sqrt{\mu^2 - r^2}}d\mu \tag{2.9-3}$$

where the dummy variable μ is given by $\mu = R_E + h'$ and r is given by $r = R_E + h$. To avoid differentiating $\delta\varphi_F(h_b)$, Eq. (2.9-3) may be written in an alternate form:

$$-\frac{\lambda}{2r\pi^2}\frac{d}{dr}\int_h^\infty \left\{\delta\varphi_F[h_{\mathrm{LG}}(h')]\left[\frac{d\zeta}{dh'}\sqrt{\mu^2-r^2}+\frac{\mu\zeta(h')}{\sqrt{\mu^2-r^2}}\right]\right\}d\mu \qquad (2.9\text{-}4)$$

where ζ is the defocusing factor, $\zeta = dh / dh_{\mathrm{LG}}$.

For an atmosphere in local thermodynamic equilibrium and of known constituents, the consequent error in the retrieved temperature profile due to unmodeled Fresnel perturbations follows from the relation

$$\left.\frac{\delta T}{T}\right|_h = -\left.\frac{\delta N}{N}\right|_h + \frac{\int_h^\infty \delta N dh}{\int_h^\infty N dh} \qquad (2.9\text{-}5)$$

When the spectrum of $\delta N(h)$ has most of its power at short wavelengths (i.e., <1 km), which is the case for Fresnel-induced variations, then the first term on the RHS of Eq. (2.9-5) will be dominant. In any case, the second term is given by

$$\int_h^\infty \delta N_F(h')dh' = \frac{\lambda}{2\pi^2}\int_h^\infty \frac{d(\delta\varphi_F)}{dh_{\mathrm{LG}}}\cos^{-1}\left(\frac{r}{\mu}\right)d\mu, \quad \mu = R_E + h' \qquad (2.9\text{-}6)$$

Figures 2-28(a) through 2-28(d) provide an evaluation of the Fresnel effects on the recovery of $N(h)$ and $T(h)$ for the Case C example given in Fig. 2-10(b), with a lapse rate discontinuity of 14 K/km. Here, as in Fig. 2-10(b), the geometric optics component due to the lapse rate discontinuity has been suppressed because the Abel transform would recover this component. A lapse rate discontinuity of 14 K/km results in an error of a few tenths of a kelvin in retrieved temperature in the vicinity of the discontinuity. As expected, the maximum error occurs near first contact with the caustic, which for the + ray (in Figs. 2-28(c) and 2-28(d) the abscissa is the turning point altitude h^+ of the + ray) is located at $h^+(2) - h_o = +0.16\,\mathrm{km}$, that is, it is *above* the discontinuity. The – rays are located at $h^{a,b}(2) - h_o = -0.16\,\mathrm{km}$. This example clearly shows that even though the altitude differences of the turning points of these multipath rays are well within the first Fresnel zone, a Fresnel diffraction treatment is required to properly account for the effects of the caustic. On the other hand, SNR limitations, at least on GPS/MET, constrain the utility of high-resolution Doppler sampling. A typical averaging interval used by investigators

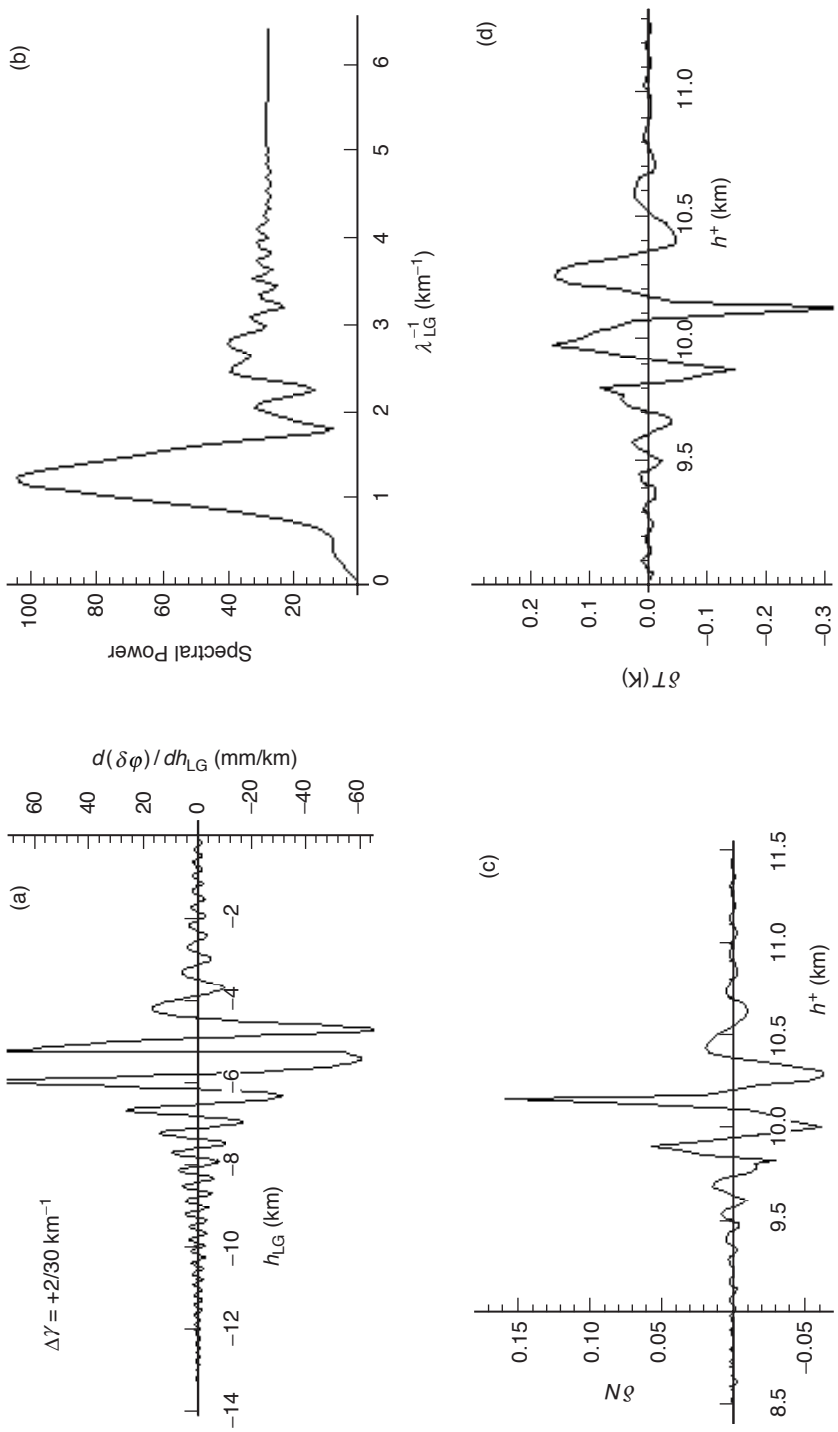

Fig. 2-28. Fresnel perturbations and errors for Case C with a lapse rate discontinuity of 14 K/km at an altitude of 10 km: (a) Doppler perturbations, (b) Fourier spectrum of the Doppler perturbation, (c) refractivity error, and (d) temperature error.

in analyzing GPS/MET data has been about 1/3 s, or about 1 km in h_{LG}, which largely washes out Fresnel effects [see Figs. 2-28(c) and 2-28(d)].

2.10 Fresnel Transform Techniques

Because of their highly localized character, transients such as those observed across a sporadic E-layer would appear to be good candidates for treating by inverse diffraction techniques to secure sub-Fresnel vertical resolutions. Limitations in resolution using inverse diffraction techniques can arise from along-track inhomogeneity, from an uncertainty in the cross-track gradient of the TEC (total along-track electron content per unit columnar area, 1 TEC = 10^{16} electrons/m^2) between the LEO and the layer, and also from the uncertainty in the effective or mean along-track distance of the layer from the LEO. Any model error that causes a loss of coherence as the integration limits are broadened between the integrands ($E(h_{\mathrm{LG}})\exp\left[i\Psi(h_{\mathrm{LG}})\right]$ in the diffraction integral [Eq. (2.5-1)] and $A(h)\exp\left[i\Phi(h,h_{\mathrm{LG}})\right]$) in its Fresnel transform will degrade the resolution.

For a spherical shell with a 1-km vertical thickness, the along-track extent of the layer would be roughly 200 km. It can be shown [4,5] that δh, the vertical resolution limit due to an error δD in the *assumed* along-track distance, is given by

$$\delta h \geq 1.25\sqrt{\frac{\lambda D}{2}\frac{\delta D}{D}} \tag{2.10-1}$$

Thus, an uncertainty of 5 percent in the along-track distance (about 150 km) translates into a lower resolution limit of about 20 percent of the radius of the first Fresnel zone; the limit scales as the square root of the error in the knowledge of the distance D of the layer.

Uncertainty in the cross-track gradient in TEC between the LEO and the layer averaged over the vertical extent associated with the diffraction integration interval (a few kilometers) is another problem. Although a time sequence of dual-band phase measurements allows a determination of the overall (LEO-to-GPS) cross-track gradient in TEC, the component of cross-track TEC gradient arising from the electron density spatial variability *between* the LEO and the boundary layer is partially unknown. This component can "leak into" the diffraction and thereby corrupt the resolution because of the 13 percent difference in Fresnel zone radii between the L1 and L2 carriers. It can be shown [4] that an error in the knowledge of this cross-track gradient degrades the resolution by

$$\delta h \geq D \frac{1.75 \times 10^{18}}{f^2} \delta\left(\frac{d(TEC)}{dh_{\mathrm{LG}}} \right) \text{ m} \tag{2.10-2}$$

It follows that a near-field error of 0.1 TEC/km maps into a vertical resolution lower limit of about 30 percent of the radius of the first Fresnel zone; note that this limit scales linearly with the error in the near-field component of cross-track TEC gradient.

With these caveats in mind[7] (and others; see, for example, [5]), we briefly discuss enhanced resolution using the Fresnel transform. The phase perturbation in the thin screen due to $\delta\alpha$ theoretically can be recovered with enhanced resolution from the observed phase and amplitude variations by using Fresnel transform techniques. We may generalize Eq. (2.5-1) to the form

$$\left.\begin{aligned} E(u_{\mathrm{LG}}) \exp\left(i\psi(u_{\mathrm{LG}})\right) &= \frac{1}{1+i} \int_{-\infty}^{\infty} A(u) \exp\left(i\left(\varphi(u) + \frac{\pi}{2}(u - u_{\mathrm{LG}})^2 \right) \right) du \\ u = h\left(\frac{2}{\lambda D} \right)^{\frac{1}{2}}, \quad u_{\mathrm{LG}} &= h_{\mathrm{LG}} \left(\frac{2}{\lambda D} \right)^{1/2} \end{aligned}\right\} \tag{2.10-3}$$

Here u is the thin-screen altitude h expressed in Fresnel units, and similarly for u_{LG}, $\varphi(u)$ is the thin-screen phase profile, and $A(u)$ is the normalized amplitude of the radiation emitted by the thin screen ($A(u)$ would be unity throughout the screen when no attenuation exists). Then $A(u)\exp[i\varphi(u)]$ and $E(u_{\mathrm{LG}})\exp\left[i\psi(u_{\mathrm{LG}})\right]$ form a Fresnel transform pair. Multiplying the integral in Eq. (2.10-3) by $\exp\left[-i(v - u_{\mathrm{LG}})^2\right]$, integrating on u_{LG}, and using the Dirac delta function (see Appendix A) yields

$$A(u)\exp\left(i\varphi(u)\right) = \frac{1}{1-i} \int_{-\infty}^{\infty} E(u_{\mathrm{LG}}) \exp\left(i\left(\psi(u_{\mathrm{LG}}) - \frac{\pi}{2}(u - u_{\mathrm{LG}})^2 \right) \right) du_{\mathrm{LG}} \tag{2.10-4}$$

The perturbation in atmospheric refractivity would be obtained from

$$\delta N(h) \doteq 10^6 \frac{\lambda}{2\pi^2 a} \frac{d}{da} \int_a^{\infty} \frac{\xi \varepsilon(h)}{\sqrt{\xi^2 - a^2}} d\xi \tag{2.10-5}$$

[7] The effect of along-track inhomogeneities on resolution can be assessed using the multiple-screen approach [10,11] combined with a two-dimensional random walk model to account for variations in refractivity in the along-track and radial directions.

where $\xi = h + R_E$, $h = a\sec\alpha - R_E$. Here $\varepsilon(h)$ is the perturbation in the thin-screen phase profile obtained by subtracting the reference profile $\varphi_{\text{Ref}}(h)$ from the profile $\hat{\varphi}(h)$ recovered from the inverse Fresnel transform in Eq. (2.10-4).

Strategies for using Fresnel transform techniques as outlined by Eqs. (2.10-4) and (2.10-5) to achieve resolutions past the Fresnel limit are discussed in [4,5]. Sub-Fresnel-zone resolution can be obtained in certain physical situations. This was clearly demonstrated by [5] in recovering certain physical parameters of Saturn's rings from radio occultation observations. The extent to which these techniques are applicable to the Earth's atmosphere and ionosphere will depend in part on how laminar the atmosphere is and how benign the ionosphere is.

2.10.1 Adjoining the Constraint $A(u) \equiv 1$

One could strengthen the recovery of the thin-screen phase profile by adjoining a constraining condition to Eq. (2.10-3) requiring that $A(u) \equiv 1$ and applying some least-squares criterion to the resulting linear operators. One of the shortcomings of the formal application of the inverse Fresnel transform in Eq. (2.10-4) is that this approach does not take into account the presence of errors in the phase and amplitude observations and one is free to adjust the recovered values of $A(u)$ accordingly without a priori information about $A(u)$ being imposed.

A straightforward alternative approach is to linearize the forward Fresnel transform in Eq. (2.10-3) in terms of the variables φ, E, and ψ by varying them about their nominal values and then to convert the resulting variational equations, both the in-phase and quadrature components, into a discrete, over-determined, linear least-squares estimation system. The state vector is the thin-screen phase sequence $\{\delta\varphi\} = \{\varepsilon\}$. These vector elements are to be adjusted and iterated according to some least-squares criterion that minimizes the weighted root-mean-square (rms) of the "O–C" vector residuals $\{\delta E\}$ and $\{\delta\psi\}$. It can be shown with $A(u) \equiv 1$ that the linearized version of Eq. (2.10-3) is given by

$$\left.\begin{aligned} \begin{Bmatrix} \delta E(u_{\text{LG}}) \\ E(u_{\text{LG}})\delta\psi(u_{\text{LG}}) \end{Bmatrix} &= \frac{1}{2}\int_{-\infty}^{\infty} \begin{Bmatrix} \cos\Omega + \sin\Omega \\ \cos\Omega - \sin\Omega \end{Bmatrix} \varepsilon(u)\,du \\ \Omega(u, u_{\text{LG}}) &= \psi(u_{\text{LG}}) - \varphi(u) - \pi(u - u_{\text{LG}})^2 / 2 \end{aligned}\right\} \tag{2.10-6}$$

Converting the integrals in Eq. (2.10-6) into discrete forms to handle discrete sampling should be done with caution because it can introduce subtle effects related to aliasing, sampling bandwidth versus the spectral breadth of the phase and amplitude observations, and SNR levels.

2.10.2 Diffraction Integral with Multipath

One can generalize Eqs. (2.10-4) and (2.10-5) further to account for multipath, following the lines suggested in Chapter 3, footnote 2. Suppose the LEO GPS receiver has successfully "tracked" the amplitude and phase profiles of the individual tones, or has at least operated in an open-loop mode so as to enable this recovery using Fourier transform techniques. Let $E^{(j)}(u_{\mathrm{LG}})$ and $\psi^{(j)}(u_{\mathrm{LG}})$ be the amplitude and phase of the j th tone "tracked" by the receiver on the LEO. Then the stationary-phase condition must apply to each tone, and it follows that

$$E^{(j)}(u_{\mathrm{LG}})e^{i\psi^{(j)}(u_{LG})} = \frac{1}{1+i}\int_{-\infty}^{\infty} A^{(j)}(u)e^{i\left(\varphi^{(j)}(u)+\frac{\pi}{2}(u-u_{\mathrm{LG}})^2\right)}du, \quad j=1,2,\cdots \tag{2.10-7}$$

In this case, $A^{(j)}$ would not be unity because the incident signal power would have to be shared among the tones so that $\sum A^{(j)} = 1$. If geometric optics is applicable, the local defocusing factor for each tone could be used to obtain an a priori estimate of the relative values of $A^{(j)}$. The Fresnel transform of Eq. (2.10-7) is given by

$$A^{(j)}(u)e^{i\varphi^{(j)}(u)} = \frac{1}{1-i}\int_{-\infty}^{\infty} E^{(j)}(u_{\mathrm{LG}})e^{i\left(\psi^{(j)}(u_{LG})-\frac{\pi}{2}(u-u_{\mathrm{LG}})^2\right)}du_{\mathrm{LG}}, \quad j=1,2,\cdots \tag{2.10-8}$$

2.10.3 A Numerical Example

With the 50-Hz L1 $\mathrm{SNR_V}$ and phase data from the occultation profile shown in Figs. 2-15 and 2-16 as input into Eq. (2.10-4), one can generate estimates of $\varepsilon(u)$ for the layer. We assume that over the few seconds of interest here the phase contribution from the ionosphere at large is essentially constant. One can improve on this procedure by modeling the contribution from the ionosphere at large over this short time interval; for example, it is easily shown that adding a constant Doppler term Δf_d to the excess Doppler series in Fig. 2-16 (to null its mean value) results in an offset in the recovered profile of $\varepsilon(u)$ [from Eq. (2.10-4)] in h-space that is equal to $D\Delta\alpha^I$, or about 100 m; $\Delta\alpha^I$ is the bending-angle increment resulting from the Doppler offset Δf_d. More specifically, we perturb the observed phase: $\psi \to \psi^T + 2\pi\Delta f_d t$, where t is a linear function of h_{LG} and $t=0$ when $h_{\mathrm{LG}}=0$. Let $\tilde{A}(u)\mathrm{Exp}[i\tilde{\varphi}(u)]$ be the recovered thin-screen amplitude and phase including the effect of this linear perturbation in observed phase. Then from Eq. (2.10-4) it can be shown that

$$\left.\begin{aligned}\tilde{A}(u)e^{i\tilde{\varphi}(u)} &= e^{i\frac{\pi}{2}\left((u+\Delta u)^2-u^2\right)}\frac{1}{1-i}\int_{-\infty}^{\infty}E\left(u_{\mathrm{LG}}\right)e^{i\left(\Psi^T\left(u_{\mathrm{LG}}\right)-\frac{\pi}{2}\left(\Delta u+u-u_{\mathrm{LG}}\right)^2\right)}du_{\mathrm{LG}},\\ \Delta u &= \sqrt{\frac{2D}{\lambda}}\Delta\alpha^I,\ \Delta\alpha^I=\lambda\frac{\Delta f_d}{\left|\dot{r}_{\mathrm{LG}}\right|},\ u=h\left(\frac{2}{\lambda D}\right)^{1/2},\ u_{\mathrm{LG}}=h_{\mathrm{LG}}\left(\frac{2}{\lambda D}\right)^{1/2}\end{aligned}\right\} \tag{2.10-9}$$

Thus, the recovered profiles of $\tilde{A}(u)$ and $\tilde{\varphi}(u)$ are offset[8] in the thin screen by Δu. We neglect this offset Δu here because we are interested in the resolution potential of Eq. (2.10-4).

We renormalize the excess L1 Doppler series shown in Fig. 2-16 so that its average value for the first 5 seconds in this figure is zero. Then we integrate this Doppler series with time to obtain a renormalized L1 phase profile (see Fig. 2-29), which along with the normalized $\mathrm{SNR_V}$ series (its average value over the first 5 seconds is set to unity) are used in Eq. (2.10-4) to obtain the profiles for the thin-screen amplitude and phase in the vicinity of the transient, which from Fig. 2-16 is located around $t = 5.7\,\mathrm{s}$.

As discussed earlier, the integration limits in Eq. (2.10-4) affect the resolution that can be achieved when modeling errors are present. Thus, the optimal integration limit in Eq. (2.10-4) involves a trade-off between accuracy and resolution. For an uncertainty in D of 5 percent, the single-sided integration limit giving the best resolution is about 3 to 4 Fresnel radii, or about 1 to 2 s of elapsed time; the best resolution itself would be given by Eq. (2.10-1), and would be about 20 percent of a Fresnel radius in this case.

---

[8] In Fresnel transform theory, the pair (u_{LG}, u) form a conjugate variable set that is analogous to that formed by spectral number and coordinate position, or by frequency and time in Fourier transform theory, or, equivalently, by position and momentum, or by energy and time in quantum mechanics. The reader can verify that there is an analogous "Heisenberg uncertainty" principle that holds between Δu_{LG} and Δu. In diffraction integral processes, this translates into the following proposition. Suppose that a transient occurs in the thin-screen phase profile of characteristic width Δu (for example, a Gaussian-shaped transient with a 1-σ half-width of Δu). This transient causes a corresponding transient in the observations of characteristic width $\Delta u_{\mathrm{LG}} = \Delta T \dot{r}_{\mathrm{LG}}(2/\lambda D)^{1/2}$ (the envelope of which for this example also would be a Gaussian with a 1-σ half-width of Δu_{LG}). Then the uncertainty principle states that the product $\Delta u \cdot \Delta u_{\mathrm{LG}} \geq 1/\pi$. This inequality is of course a hallmark of classical diffraction: the narrower the aperture through which waves must pass, the broader their overall pattern at reception.

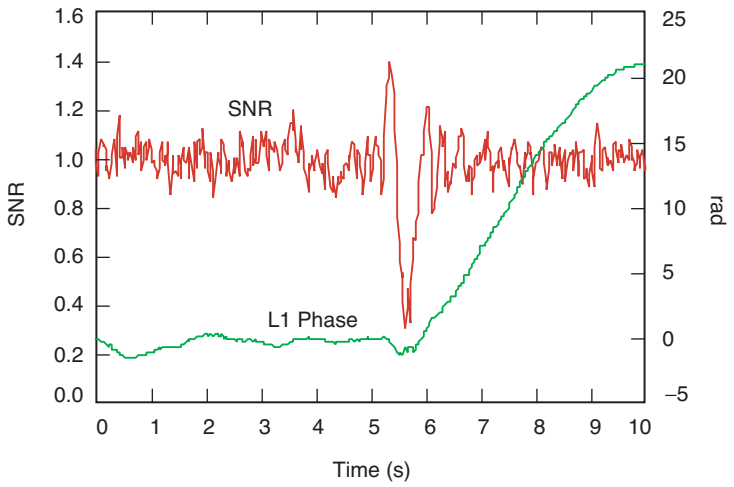

Fig. 2-29. Renormalized L1 SNR and excess phase for the same occultation and epoch shown in Figs. 2-15 and 2-16.

2.10.4 Fresnel Aliasing

The finite temporal sample width (20 ms) of the L1 Doppler series shown in Fig. 2-16 can result in a kind of "Fresnel aliasing" in the convolution; so, one must use integration limits that are well below a certain threshold at which the effects of this aliasing become prominent. Let k be an integer that denotes the position in h_{LG}-space of the LEO–GPS line at the kth observation epoch, and similarly, let j denote a thin-screen position in h-space. When $h_{\mathrm{LG}} = h$, $k = j$; thus, $h_{\mathrm{LG}} = (\dot{r}_{\mathrm{LG}}\Delta T)k + c$ and $h = (\dot{r}_{\mathrm{LG}}\Delta T)j + c$, where c is a constant (we have assumed here that the defocusing is negligible). Here ΔT is the temporal sample interval, 20 ms. In this case, the aliasing threshold for the Fresnel phase in Eq. (2.10-4) is given by the discrete form of the stationary-phase condition

$$\psi(k+1) - B(j-k-1)^2 - \psi(k) + B(j-k)^2 = \pm 2n\pi \qquad (2.10\text{-}10)$$

where n is an integer and where

$$B = \pi \frac{(\dot{r}_{\mathrm{LG}}\Delta T)^2}{\lambda D} \qquad (2.10\text{-}11)$$

The principal contributions to the continuous version of the inverse diffraction integral in Eq. (2.10-4) come from the neighborhood(s) around the point(s) where k to the nearest integer satisfies the condition given in Eq. (2.10-10) with $n = 0$. However, in the discrete version of Eq. (2.10-4), pseudo-stationary-phase contributions also arise with $n = \pm 1, \pm 2, \cdots$. Let $M^* = |k^* - j|$ be the

aliasing threshold, which must significantly exceed the single-sided range of k values about a given value of j that can be used in the deconvolution to avoid Fresnel aliasing effects. Expanding Eq. (2.10-10) yields

$$M^* = \left| \frac{\lambda D}{(\dot{r}_b \Delta T)^2} \left(f_d \Delta T \pm n \right) - 1/2 \right| \tag{2.10-12}$$

where f_d is the excess Doppler in hertz. For $n = 1$, M^* is about 240 for this occultation, which exceeds the integration limit for optimal resolution when $\delta D / D \geq 0.3\%$ and when the uncertainty in the near-field cross-track gradient in TEC is greater than ~0. 1 TEC/km.

2.10.5 Numerical Results

The results of this inverse diffraction integral are shown in Figs. 2-29 and 2-30(a) through 2-30(c). Figure 2-30 shows the renormalized excess L1 phase and SNR for the occultation shown in Figs. 2-15 and 2-16. Figure 2-30(a) shows the recovered thin-screen amplitude and phase using Eq. (2.10-4) without applying the a priori constraint $A(u) \equiv 1$. Figure 2-30(b) shows the gradient of the recovered thin-screen phase delay, which is proportional to bending angle. Note the change in time scale in Figs. 2-30(a) through 2-30(c) versus Fig. 2-29. Although the recovered profile for the thin-screen amplitude is roughly unity over most of its time series, it deviates significantly from unity at the transient; it is in this neighborhood where adjoining the a priori constraint might strengthen the recovery of the thin-screen phase profile. On the other hand, it is well known from the mathematical concept of observability in estimation theory that in the presence of mismodeled effects or certain other error sources, one might obtain better least-squares determinations of the parameters of interest by allowing certain other parameters (not well-observed in the first place) a wider latitude in the least-squares process to absorb the effects of such error sources. A good strategy here is not clear without first investigating the spectral nature of the error sources and constructing the sensitivity matrix of the estimated parameter set.

To obtain the electron density profile across the transient, we relate the phase profile in the thin-screen model to the change in refractivity across a thin, spherical ionospheric shell. In thin-screen methodology involving a single screen, the actual along-track phase accumulation through the shell is replaced with a thin-screen phase that is given by the stationary-phase condition through Eq. (2.6-5), which relates the thin-screen phase profile to the refractive bending-angle profile based on geometric optics. For the very small bending angles involved here and assuming local spherical symmetry, we can use the approximate relationship for the thin-screen phase perturbation:

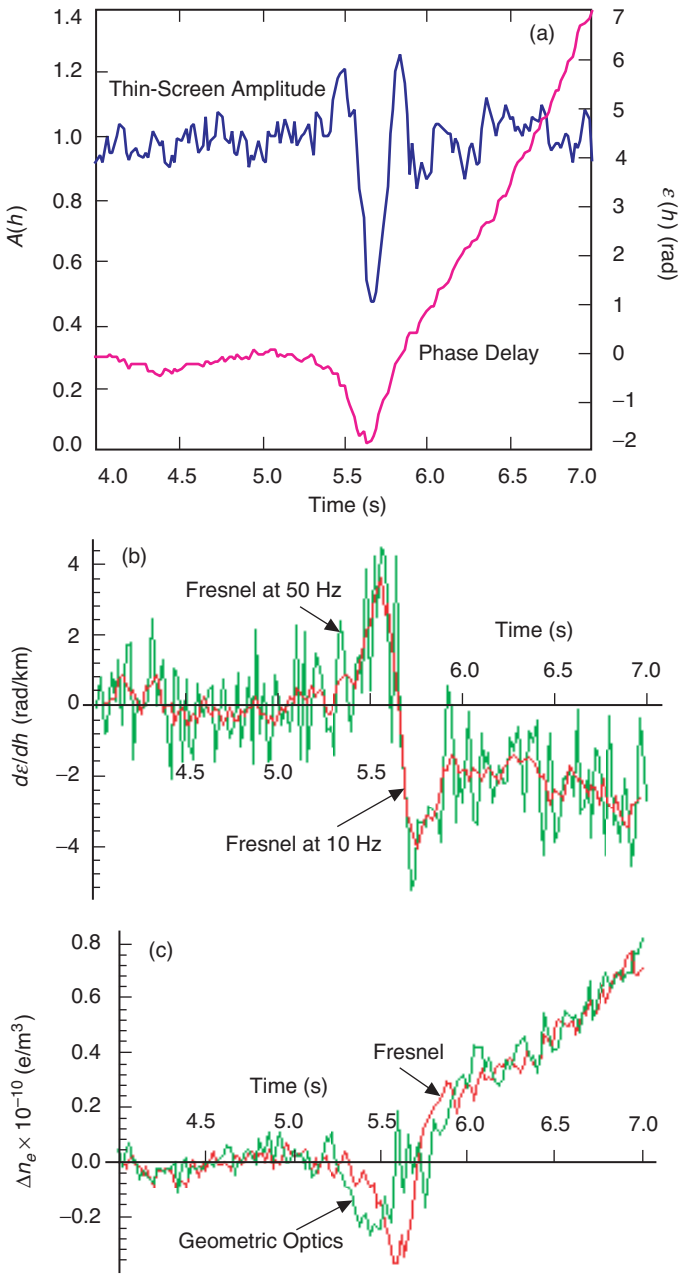

Fig. 2-30. Recovered thin-screen properties from the Fresnel transform without the adjoining a priori constraint $A(u) = 1$: (a) recovered thin-screen phase and amplitude, (b) recovered gradient in phase (bending angle), and (c) change in recovered refractivity using the Fresnel transform versus the Abel transform. The data are the same as in Fig. 2-29. Note the change in time scale.

$$\varepsilon(r) \doteq 2k\int_0^\infty \Delta N(r')ds$$
$$= 2k\int_r^\infty \Delta N \frac{r'dr'}{\sqrt{r'^2 - r^2}} = -2k\int_r^\infty \frac{d(\Delta N)}{dr'}\sqrt{r'^2 - r^2}\,dr' \qquad (2.10\text{-}13)$$

As Fig. 2-31 shows, even a delta-function-like impulse in refractivity across a thin ionospheric shell results in a lingering effect in the thin-screen phase profile because the rays continue to traverse the shell twice at lower depths. Upon differentiating ε with respect to r and using the Abel transform (and setting $\Delta N(\infty) = 0$), it follows that

$$\Delta N(r) \doteq -\frac{1}{\pi k}\int_r^\infty \frac{d\varepsilon}{d\xi}\frac{d\xi}{\sqrt{\xi^2 - r_o^2}} \qquad (2.10\text{-}14)$$

which is essentially the same form as given by Eq. (2.9-3). The result of this integration is shown by the curve labeled "Fresnel" in Fig. 2-30(c). The half-width of the bulge in electron density is about 300 m. The curve labeled "Geometric Optics" is the geometric optics version, that is, the profile obtained from applying the Abel transform directly to the L1 Doppler series in Fig. 2-16, renormalized so that the average excess Doppler over the first 5 seconds is zero. As has already been suggested by Fig. 2-27, the geometric optics version "tracks" the medium-to-long wavelength variability (compared to the Fresnel scale) in refractivity fairly well. But just prior to and just after encountering the sharp transition in refractivity at $t = 5.6$s in Fig. 2-30(c), the fidelity of this algorithm appears to degrade significantly.

We also can compare the constrained and unconstrained versions of the

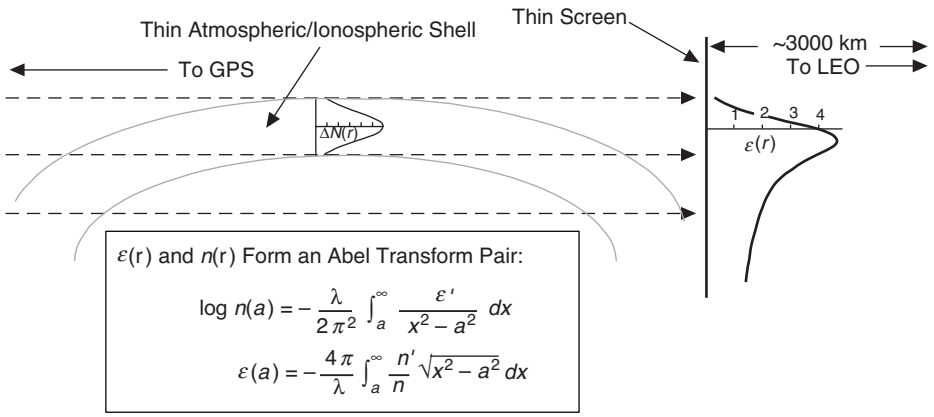

Fig. 2-31. The relationship between refractivity and the thin-screen phase for a spherical shell.

Fresnel recovery of the electron density profile around the transient. Following the discussion around Eq. (2.10-6) in which the constraint $A(u) \equiv 1$ is formally imposed on the linearized version of the forward Fresnel transform, we first convert it into a discrete form and then invert it using a least-squares algorithm. In this case, one obtains a formal solution for the least-squares estimate of the offsets in the thin-screen phase relative to the nominal profile, which is given by

$$\delta\hat{\varepsilon} = \Gamma \cdot \mathbf{L} \tag{2.10-15}$$

where $\delta\hat{\varepsilon}$ is the $m \times 1$ matrix giving the least-squares estimate of the thin-screen phase offset with $m \le \text{NMAX}/2$. Γ is the $m \times m$ covariance matrix, and $\mathbf{L}$ is the $m \times 1$ information matrix, which for the simple discrete version used here and assuming a stationary white noise process for the observation set (stationarity manifestly does not hold near the transient!), is given by

$$\left.\begin{aligned} \mathbf{L}_j &= \frac{A}{2}\sum_{k=1}^{\text{NMAX}} \beta_k \begin{Bmatrix} \cos\Omega_{jk} + \sin\Omega_{jk} \\ \cos\Omega_{jk} - \sin\Omega_{jk} \end{Bmatrix}^{\mathrm{T}} \begin{Bmatrix} \delta E_k \\ E_k \delta\psi_k \end{Bmatrix} \\ \Omega_{jk} &= \psi_k - \varepsilon_j - B(j-k)^2, \\ \beta_k &= 1, \ \forall\, k \ne 1 \ \text{ or NMAX}; \ \ \beta_1 = \beta_{\text{NMAX}} = 1/2 \\ A &= \dot{r}_b \Delta T \left(\frac{2}{\lambda D}\right)^{1/2}, \ B = \frac{\pi}{2} A^2, \end{aligned}\right\} \tag{2.10-16}$$

Here NMAX is the number of discrete pairs of amplitude and phase observations selected for the inversion process and NMAX should be at least as great as the Nyquist limit for the selected number of thin-screen points m (but significantly less than the Fresnel aliasing limit $2M^*$). For large NMAX and for a stationary white noise process, $\mathbf{\Gamma}^{-1}$ asymptotically approaches the particularly simple form

$$\left.\begin{aligned} \left[\mathbf{\Gamma}^{-1}\right]_{ij} &\to \frac{A^2}{4} \frac{\sin\left[\varepsilon_i - \varepsilon_j + B(i-j)(i+j-1)\right]}{\sin\left[B(i-j)\right]}, \ \ i \ne j \\ \left[\mathbf{\Gamma}^{-1}\right]_{ii} &\to \frac{A^2}{2}\text{NMAX} \end{aligned}\right\} \tag{2.10-17}$$

which is nearly diagonal when $1 << \text{NMAX} < 2M^*$. A simple trapezoidal rule has been used to convert the continuous forms to discrete forms used in Eqs. (2.10-16) and (2.10-17). This should suffice provided the spatial sampling interval is small (it is about 1/11 of a Fresnel scale for this occultation) and provided the sampled observation vectors are low-pass filtered to eliminate the

possibility of certain high-frequency components in their continuous power spectrum leaking in through aliasing. These discrete linearized systems can be applied iteratively using any one of a variety of estimation algorithms (e.g., Householder orthogonal transformation, singular-value decomposition, square-root information filter) until convergence is obtained.

When the converged profile for the thin-screen phase is passed through Eq. (2.10-14), one recovers the refractivity with the a priori constraint $A(u) \equiv 1$ imposed. A comparison of electron density profiles [using Eq. (2.8-1)] with and without this constraint is shown in Fig. 2-32. Steeper gradients in electron density near the transient and a higher peak are significant features of the constrained case. These result in greater defocusing in the constrained case than in the unconstrained case to achieve the observed SNR fadeout and recovery instead of allowing $A(u)$ also to vary to achieve the same effect.

The question of the fidelity of the different Fresnel versions presented here for recovering electron density is moot without first performing a number of simulations with realistic error sources including along-track inhomogeneity and sampling effects. Chapters 3 and 5 of this monograph deal indirectly with the question of the fidelity of the thin-screen model by studying scattering effects based on Maxwell's equations applied to a spherical surface. Fresnel transform theory and the thin-screen model appear to work well in thin atmosphere conditions, which makes certain narrow features observed in the ionosphere good candidates for this approach. Also, tropopause features should be well modeled. But lower troposphere features often violate the thin atmosphere condition for the thin screen, $1 + a\alpha(d\alpha / da) > 0$ or $\beta < (2\pi)^{-1/2} \approx 0.4$, and they even exceed the super-refractivity threshold $\beta = 1$ on occasion. The impact parameter space curve provides a better platform for these lower troposphere features, working exceptionally well except for super-refractive layers.

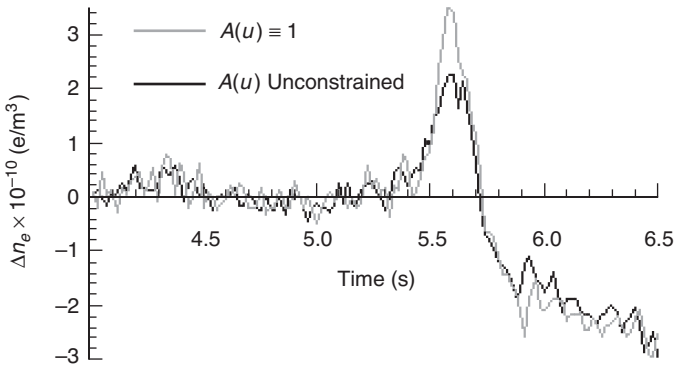

Fig. 2-32. Fresnel recovery of the transient in electron density distribution for the occultation shown in Fig. 2-15 around $t = 5.6$ **s, with and without adjoining the a priori constraint** $A(u) \equiv 1$.

For observation epochs greater than $t = 6.5\text{s}$ shown in Fig. 2-32, the recovered profile for Δn_e gradually levels off as the tangency point of the ray descends into the upper mesosphere and into a more neutral medium. This regime where Δn_e levels off could provide a zero point from which the absolute value of the electron density near the transient can be estimated. However, variability from the ionosphere at large may corrupt this estimate. Figure 2-14 provides one example of the trend in bending angles for the upper mesosphere.

References

[1] G. Lindal, J. Lyons, D. Sweetnam, V. Eshleman, D. Hinson, and G. Tyler, "The Atmosphere of Uranus: Results of Radio Occultation Measurements with Voyager 2," *Journal of Geophysical Research,* vol. 92, pp. 14,987–15,001, 1987.

[2] G. Lindal, "The Atmosphere of Neptune: An Analysis of Radio Occultation Data Acquired by Voyager 2," *The Astronomical Journal,* vol. 103, pp. 967–982, 1991.

[3] M. Born and E. Wolf, *Principles of Optics,* 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.

[4] W. G. Melbourne, E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, E. R. Kursinski, T. K. Meehan, L. E. Young, and T. P. Yunck, *The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring*, JPL Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, April 1994.

[5] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," Icarus, vol. 68, pp. 120–166, 1986.

[6] E. Bramley, "The Diffraction of Waves by an Irregular Medium," *Proceedings of the Royal Society*, vol. A225, pp. 515–518, 1954.

[7] E. Saltpeter, "Interplanetary Scintillations. I. Theory," *The Astrophysical Journal*, vol. 147, pp. 433–448, 1967.

[8] E. N. Bramley, "The Accuracy of Computing Ionospheric Radiowave Scintillation by the Thin-Phase Screen Approximation," *Journal of Atmospheric and Solar-Terrestrial Physic*s, vol. 39, pp. 367–373, 1977.

[9] W. Hubbard, J. Jokopii, and B. Wilking, "Stellar Occultations by Turbulent Planetary Atmospheres: A Wave-Optical Theory Including a Finite Scale Height," *Icarus*, vol. 34, pp. 374–395, 1978.

[10] D. Knepp, "Multiple Phase-Screen Calculation of the Temporal Behavior of Stochastic Waves," *Proceedings of the IEEE*, vol. 71, no. 6, pp. 722–737, 1983.

[11] J. Martin and S. Flatté, "Intensity Images and Statistics from Numerical Simulation of Wave Propagation in 3-D Random Media," *Applied Optics*, vol. 27, no. 11, pp. 2111–2126, 1988.

[12] J. Goodman, *Introduction to Fourier Optics*, New York: McGraw-Hill, 1968.

[13] A. J. Devaney, "A Filtered Backpropagation Algorithm for Diffraction Tomography," *Ultrasonic Imaging*, vol. 4, pp. 336–350, 1982.

[14] E. Karayel and D. Hinson, "Sub-Fresnel Scale Vertical Resolution in Atmospheric Profiles from Radio Occultation," *Radio Science*, vol. 32, no. 2, pp. 411–423, 1997.

[15] M. E. Gorbunov, A. S. Gurvich, and L. Bengtsson, *Advanced Algorithms of Inversion of GPS/MET Satellite Data and Their Application to Reconstruction of Temperature and Humidity*, Report 211, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.

[16] G. Watson, "The Diffraction of Electric Waves by the Earth," *Proceedings of the Royal Society*, vol. A95, pp. 83–99, London, United Kingdom, 1918.

[17] J. Jackson, *Classical Electrodynamics,* 2nd ed., New York: John Wiley & Sons, Inc., 1975.

[18] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, United Kingdom: The Institution for Electrical Engineers, 2000.

[19] G. A. Hajj, E. R. Kursinski, W. I. Bertiger, S. S. Leroy, T. Meehan, L. J. Romans, and J. T. Schofield, "Initial Results of GPS–LEO Occultation Measurements of Earth's Atmosphere Obtained with the GPS/MET Experiment," *Proceedings of the IAG Symposium*, G1, GPS Trends in Precise Terrestrial, Airborne, and Spaceborne Applications, IUGG XXI General Assembly, Boulder, Colorado, July 2–14, 1995, Springer-Verlay, pp. 144–154, 1996.

[20] M. Gu and F. Brunner, "Theory of the Two Frequency Dispersive Range Correction," *Manuscripta Geodætica*, vol. 15, pp. 357–361, 1990.

[21] S. Bassiri and G. Hajj, "Higher-Order Ionospheric Effects on the Global Positioning System Observables and Means of Modeling Them," *Manuscripta Geodætica*, vol. 18, pp. 280–289, 1993.

[22] V. V. Vorob'ev and T. G. Krasil'nikova, "Estimation of the Accuracy of the Atmospheric Refractive Index Recovery from Doppler Shift Measurements at Frequencies Used in the NAVSTAR System," *Physics of the Atmosphere and Ocean*, vol. 29, no. 5, pp. 602–609, 1994.

[23] T. Meehan, personal communication, Jet Propulsion Laboratory, Pasadena, California, 1995.

[24] W. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occulatation Observations From a Low Earth Orbiter, Part 1*, JPL Publication 96-5, Jet Propulsion Laboratory, Pasadena, California, July 1998.

Chapter 3
Scattering from a Large Transparent Sphere Based on Maxwell's Equations: Mie Scattering Theory

3.1 Introduction

We consider here Mie scattering theory. Incident electromagnetic waves are reflected from and refracted by a large, transparent sphere of uniform refractivity N, which is slightly different from the uniform refractivity of the surrounding medium. The ratio of the radius of the sphere to the wavelength of the incident waves, r_o / λ, is assumed to be sufficiently large so that certain asymptotic solutions to the Helmholtz equation are applicable. The refractivity difference across the boundary of the sphere is assumed to be sufficiently small (for our case, $N << 0.001$). Therefore, the difference in the reflection coefficients from the Fresnel formulas for incident waves with in-plane polarization and with cross-plane polarization is negligible; for grazing conditions, this difference is of the order N.

In Fig. 3-1, an electromagnetic wave is characterized by its Poynting vector, which in Gaussian units is given by $\boldsymbol{S} = c(\boldsymbol{E} \times \boldsymbol{H}) / 4\pi$, where $\boldsymbol{E}$ and $\boldsymbol{H}$ are, respectively, the electric and magnetic field vectors of the wave. Here c is the velocity of light in a vacuum. The incoming plane waves and the scattered waves are characterized by $\boldsymbol{S}^{(i)}$ and $\boldsymbol{S}^{(s)}$, respectively. $\boldsymbol{S}$ provides the power of the wave passing through a perpendicular unit area. The scattered waves are assumed to include both reflected and refracted waves. The incoming wave is assumed to be planar and harmonic. In this case, the electric field may be written as $\boldsymbol{E} = \hat{\boldsymbol{x}} E_o \exp[i(nkz - \omega t)]$, where $\hat{\boldsymbol{x}}$ is a unit vector directed in-plane

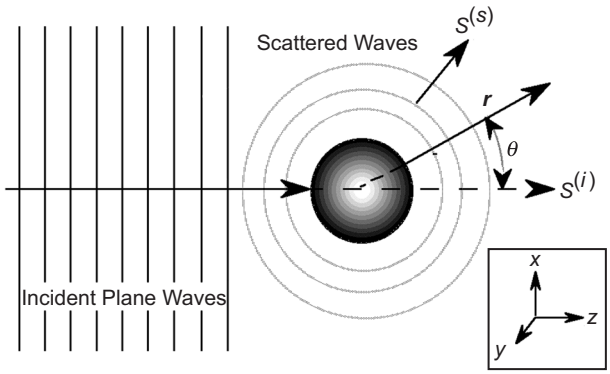

Fig. 3-1. Scattering of incident plane waves from a large refracting sphere.

and perpendicular to $\boldsymbol{S}^{(i)}$ (see inset in Fig. 3-1). Here $k = 2\pi / \lambda = \omega / c$ is the wavenumber in a vacuum, where ω is the angular frequency of the wave. The index of refraction in Gaussian units is given by $n = \sqrt{\varepsilon\mu}$, where ε is the permittivity (or dielectric coefficient) of the medium (assumed to be linear) and μ is its magnetic permeability. In a vacuum, $\varepsilon = \mu = 1$. We will assume a purely dielectric medium with zero charge and no current density throughout the medium. Also, we will assume zero magnetization, so that $\mu \equiv 1$ throughout the medium. Only ε will change discontinuously by a small amount across the boundary of the sphere, and here it will be assumed to be otherwise invariant with $n = 1$ in the surrounding medium and $n = 1 + N$ inside. We will be interested in the amplitude and phase of the total electric field vector $\boldsymbol{E}^{(i)} + \boldsymbol{E}^{(S)}$ in the vicinity of the low Earth orbiting (LEO) spacecraft.

Scattering from a conducting or a dielectric sphere is widely discussed in the classical electrodynamics literature [1–8]. Much of the material presented here is almost a hundred years old and has been discussed extensively in the past. We mostly follow the Mie formulation given in [8] in the ensuing discussion. This latter reference does not, however, deal with refracted and internally reflected components of scattered waves including rainbow caustics, nor with computational techniques and asymptotic forms. We have supplemented from other sources, mostly cited above, or from our own invention from time to time.

If much of the material in this chapter does involve old ground plowed over many times, why discuss it all? The mathematics of Mie scattering is surely ponderous, and its many asymptotic expansions are complicated, although computationally essential. Several considerations favor including the material. A principal goal here is to establish the mathematical infrastructure that is used in Chapter 5, which deals with discontinuities in the refraction model across a

spherical surface that is embedded in a refracting medium possessing a significant gradient. There, all of the Mie scattering apparatus is needed. Concepts from the "simpler," pure Mie scattering problem, such as the spectral representation in basic wave theory, the formalism for representing incoming and outgoing spherical waves, the reflection and transmission spectral coefficients, asymptotic expansions, phasor representation, stationary-phase concepts, caustics, etc., all apply to the more difficult problem of an embedded discontinuity in a refracting medium.

A second motivation for the material presented in this chapter stems from the particular emphasis throughout this monograph on stationary-phase concepts. The stationary-phase technique is well known. It is often used to evaluate an integral along the real axis with an integrand comprised of a variable phasor multiplied by a relatively slowly varying real factor. It also is very useful as an aid to other integration techniques by identifying stationary or near-stationary-phase neighborhoods of the integration variable that provide the principal contributions to the integral. But stationary-phase concepts have a much broader relevance. Geometric optics is based on a second-order stationary-phase methodology applied to the phase delay integral along the ray path. The impact parameter associated with a given ray in fact provides a stationary value of the phase delay integral for that ray. In wave theory, the spectral integral representation of a wave involves a certain phasor generated from the scattering coefficients and related geometric delays. The spectral number providing a stationary value for that phasor shows a close correspondence to impact parameter in geometric optics. In fact, one can establish a duality between certain wave theoretic quantities evaluated at stationary-phase points in spectral number and their analogous concepts in geometric optics. Hence, Chapter 3 devotes considerable attention to developing a phasor-based spectral representation for wave theory that is applicable to real spectral numbers, and to developing this stationary-phase duality with geometric optics that was just mentioned. All of these concepts carry over to Chapter 5. Lest the classical approaches for evaluating scattering integrals be completely overlooked, this chapter and also Appendix G discuss briefly the use of the complex spectral number technique. There, a summation over spectral number on the real axis is transformed into a summation of residues from a related contour integral in the complex spectral number plane, which may result in a more rapid convergence of the summation.

A third motivation for the material in this chapter stems from the desire to assess the efficacy of the technique of using scalar diffraction theory applied to the thin phase screen model. The focus here is on the calculation of the LEO-observed scattered field from rays at nearly grazing conditions with the sphere. One of our tasks is to compare the scattering at nearly grazing angles from this sphere based on a solution to Maxwell's equations with that predicted by the thin-screen technique. Specifically, we will compare Mie scattering results in

this chapter and modified Mie scattering results in Chapter 5 with Fresnel-diffracted wave forms obtained from applying scalar diffraction theory to a thin phase screen model. The thin-screen model serves as a proxy for the refracting/reflecting sphere, and it is defined in Chapter 2, Eqs. (2.5-1), (2.8-2), and (2.8-3). The thin-screen methodology is less complex than Mie scattering theory, and it offers a considerable computational advantage if its soundness can be established. When thin atmosphere conditions apply or, more specifically, when thin-screen caustics are avoided, the agreement between these two systems in LEO-observed phase and amplitude is good.

A fourth motivation is actually related to the first. Chapter 6 discusses the inverse problem of recovering the profile for the index of refraction from the phase and amplitude observations made by the LEO. For that purpose, we use the phasor-based spectral representation of the wave developed in Chapter 5.

To develop a wave theory approach to scattering, we use the general spectral series solutions to the Helmholtz equation for an electromagnetic field to develop the appropriate representations for incoming and outgoing waves and for reflected and transmitted waves. These are addressed in the next sections. We then consider the reflection and transmission coefficients across a planar dielectric surface based on geometric optics and the Fresnel formulas. Next we develop the analogous wave theory spectral coefficients for incident and scattered waves from a refracting sphere and the associated spectral series representations for waves transmitted through the sphere and reflected by it, which is Mie scattering theory. The problem of slow convergence of these spectral series for the scattered and transmitted waves from a large sphere ($r_o >> \lambda$) is reviewed. The asymptotic expressions for large spectral number and radial coordinate values needed for numerical evaluation are reviewed. Then we develop a phasor-based approach for solving these asymptotic spectral series aided by the stationary-phase technique. Along the way, we compare various stationary-phase concepts from wave theory with those from geometric optics and scalar diffraction theory. Numerical results based on Mie scattering theory for the phase and amplitude of a wave scattered from a large refracting sphere are presented and compared with results obtained from applying scalar diffraction theory to the thin phase screen model. Rainbows and caustics are addressed, and special limiting cases, such as scattering from a perfectly reflecting sphere and from an absorbing sphere, are discussed.

3.2 Scalar Potentials

For a harmonic wave in a homogeneous medium without charge and current densities, Maxwell's vector field equations for the time-independent part of the wave are given by

$$\left.\begin{aligned} \nabla \times \boldsymbol{E} &= ik\mu\boldsymbol{H} \\ \nabla \cdot \boldsymbol{E} &= 0 \\ \nabla \times \boldsymbol{H} &= -ik\varepsilon\boldsymbol{E} \\ \nabla \cdot \boldsymbol{H} &= 0 \end{aligned}\right\} \tag{3.2-1a}$$

By forming $\nabla \times (\nabla \times \boldsymbol{E})$ and $\nabla \times (\nabla \times \boldsymbol{H})$, and noting the identity $\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2 \boldsymbol{A}$, Maxwell's equations can be converted into a pair of Helmholtz equations:

$$\left.\begin{aligned} \nabla^2 \boldsymbol{E} + \mu\varepsilon k^2 \boldsymbol{E} &= 0 \\ \nabla^2 \boldsymbol{H} + \mu\varepsilon k^2 \boldsymbol{H} &= 0 \end{aligned}\right\} \tag{3.2-1b}$$

The vector solutions to these electromagnetic field equations also can be obtained from a pair of linearly independent scalar solutions to the Helmholtz equation,

$$\nabla^2 \Pi + \mu\varepsilon k^2 \Pi = 0 \tag{3.2-1c}$$

These scalar potential functions are denoted by ${}^{\mathrm{e}}\Pi(r,\theta,\phi)$ and ${}^{\mathrm{m}}\Pi(r,\theta,\phi)$. Here ${}^{\mathrm{e}}\Pi$ and ${}^{\mathrm{m}}\Pi$ are the so-called "electric" and "magnetic" scalar potentials, respectively. The time-independent components, $\boldsymbol{E}$ and $\boldsymbol{H}$, are obtained from vector calculus operations on a vector form $(\Pi \boldsymbol{r})$ of these scalar potentials (due to Hertz). The field vectors are given by [8]

$$\left.\begin{aligned} \boldsymbol{E} &= \nabla \times \nabla \times ({}^{\mathrm{e}}\Pi \boldsymbol{r}) + ik\mu \nabla \times ({}^{\mathrm{m}}\Pi \boldsymbol{r}) \\ \boldsymbol{H} &= \nabla \times \nabla \times ({}^{\mathrm{m}}\Pi \boldsymbol{r}) - ik\varepsilon \nabla \times ({}^{\mathrm{e}}\Pi \boldsymbol{r}) \end{aligned}\right\} \tag{3.2-2}$$

If one forms the curl and divergence of $\boldsymbol{E}$ and $\boldsymbol{H}$ as these fields are defined by Eq. (3.2-2), it may be shown using Eq. (3.2-1c) that these scalar potentials do indeed generate vector fields that satisfy Maxwell's equations for a harmonic wave, that is, Eq. (3.2-1a).

In our problem, these two scalar potentials essentially account for the two possible linear polarization modes of the electromagnetic field. Note that the vector $\nabla \times (\Pi \boldsymbol{r})$ is orthogonal to $\boldsymbol{r}$, whereas in general $\nabla \times \nabla \times (\Pi \boldsymbol{r})$ is not. Thus, the term $-ik\varepsilon \nabla \times ({}^{\mathrm{e}}\Pi \boldsymbol{r})$ in Eq. (3.2-2) generates a magnetic field without a radial component, that is, a transverse magnetic field. Similarly, $ik\mu \nabla \times ({}^{\mathrm{m}}\Pi \boldsymbol{r})$ generates a transverse electric field. When the electromagnetic field is invariant in the cross-plane direction (along the y-axis in Fig. 3-1), that is, when $\nabla \Pi \cdot \hat{\boldsymbol{y}} \equiv 0$, the curl operation $\nabla \times ({}^{\mathrm{m}}\Pi \boldsymbol{r})$ generates the cross-plane

component of the field $\boldsymbol{E}_{\perp}$, and $\nabla\times\nabla\times({}^{e}\boldsymbol{\Pi}\boldsymbol{r})$ generates the in-plane component $\boldsymbol{E}_{\parallel}$.

The symmetry in Eqs. (3.2-1) and (3.2-2) for $\boldsymbol{E}$ and $\boldsymbol{H}$ should be noted. It is rooted in the space–time four-vector covariant form that Maxwell's equations take in the framework of Special Relativity. We have chosen Gaussian units to emphasize this point. Note that Maxwell's equations remain invariant when $\boldsymbol{E}$ and $-\boldsymbol{H}$, and simultaneously μ and ε, are exchanged. With this symmetry we need deal with only one polarization mode. For example, if we know $\boldsymbol{E}$ and $\boldsymbol{H}$ for one mode, their forms for the other mode follow directly.

For plane waves in a homogeneous medium, $H\sqrt{\mu}=E\sqrt{\varepsilon}$. Also, $\boldsymbol{H}$ and $\boldsymbol{E}$ are orthogonal and in-phase. We take the polar axis of the sphere $\hat{z}$ to be directed parallel to the direction of propagation of the incident plane wave, that is, parallel to $\boldsymbol{S}^{(i)}$ in Fig. 3-1. We choose the azimuthal orientation of our coordinate frame about the z-axis so that the in-plane polarization of the incident wave lies in the plane defined by $\phi=0$.

3.2.1 Series Expansions for Scalar Potentials

The scalar potentials for incident plane waves of arbitrary polarization can be represented by a pair of series expansion solutions to the Helmholtz equation in spherical coordinates. The basis functions are the spherical Bessel functions for the radial coordinate r and the spherical harmonic functions for the angular coordinates. The harmonic functions are given by the Legendre polynomials for the angular coordinate θ and the sinusoids for the azimuthal angle ϕ. Here θ is the angle between the z-axis, the direction of $\boldsymbol{S}^{(i)}$, and the radius vector $\boldsymbol{r}$, which points to the position at which we wish to evaluate the field. The azimuthal angle ϕ lies in the xy-plane in Fig. 3-1 and is measured clockwise about the z-axis; $\phi=0$ along the x-axis.

The time-independent scalar component of a planar and harmonic wave of unit amplitude, $\exp(ikz)$, traveling in a non-conducting, homogeneous medium is given by a series expansion in terms of these basis functions by the form

$$\exp(ikz)=\frac{1}{kr}\sum_{l=0}^{\infty}i^{l}(2l+1)\psi_{l}(kr)\,P_{l}(\cos\theta),\quad z=r\cos\theta \tag{3.2-3}$$

This is Bauer's identity [9], which is derivable from a multipole series expansion of $\exp(ikr\cos\theta)$ and by use of the addition theorem for spherical harmonic functions. This series is the analog of a Fourier decomposition; the spherical Bessel function, $\psi_{l}(kr)$, and the Legendre polynomial, $P_{l}(\cos\theta)$, provide a complete basis in r and θ space for the series expansion. An individual term in the series multiplied by $\cos\phi$, that is,

$\psi_l(kr)P_l(\cos\theta)\cos\phi/kr$, provides the functional form of the lth in-plane polarized spherical wavelet and is itself a solution to the Helmholtz equation.

Scalar potential solutions to the Helmholtz equation also can be written in terms of these basis functions. A general form for a scalar potential solution to the Helmholtz equation that is well behaved at the origin is given by the series

$$\Pi = \sum_{l=0}^{\infty}\sum_{m=0}^{l}\left(a_{lm}\cos(m\phi)+b_{lm}\sin(m\phi)\right)\frac{\psi_l(kr)}{kr}P_l^m(\cos\theta) \qquad (3.2\text{-}4)$$

Here a_{lm} and b_{lm} are constant coefficients to be determined, and $P_l^m(x)$ is the associated Legendre polynomial of degree m and order l. Using this general series, one forms the two vector curl operations in Eq. (3.2-2), $\nabla\times(\Pi\boldsymbol{r})$ and $\nabla\times\nabla\times(\Pi\boldsymbol{r})$, and equates these expressions respectively (with $n\equiv 1$) to the vector equivalent of Eq. (3.2-3) for a planar wave with in-plane polarization (with $\boldsymbol{E}_o$ directed along the x-axis in Fig. 3-1). By equating the two expressions on a term-by-term basis, one can show that the coefficients a_{lm} and b_{lm} are forced to assume the values

$$\left.\begin{array}{l} a_{lm}=b_{lm}=0,\ \ m\neq 1,\ \text{or}\ l=0 \\ \left\{\begin{array}{l} a_{l1}=k^{-1}E_o i^{l-1}\dfrac{2l+1}{l(l+1)} \\ b_{l1}=0 \end{array}\right\}\text{electric},\quad \left\{\begin{array}{l} a_{l1}=0 \\ b_{l1}=k^{-1}E_o i^{l-1}\dfrac{2l+1}{l(l+1)} \end{array}\right\}\text{magnetic} \end{array}\right\} \qquad (3.2\text{-}5)$$

The identity $\nabla\times\nabla\times(\Pi\boldsymbol{r})=(2+\boldsymbol{r}\cdot\nabla)\nabla\Pi-\boldsymbol{r}\nabla^2\Pi$, the Helmholtz equation, and the differential equation defining the spherical Bessel function, $d^2\psi_l/dx^2+\left(1-l(l+1)/x^2\right)\psi_l=0$, are used to prove Eq. (3.2-5). It follows that the electric and magnetic scalar potentials for a plane wave in a uniform medium ($n\equiv 1$) with an in-plane polarization of constant amplitude E_o are given by the series expansions

$$\left.\begin{array}{l} {}^{\mathrm{e}}\Pi^{(i)}=\dfrac{E_o}{rk^2}\displaystyle\sum_{l=1}^{\infty} i^{l-1}\frac{2l+1}{l(l+1)}\psi_l(kr)P_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\Pi^{(i)}=\dfrac{E_o}{rk^2}\displaystyle\sum_{l=1}^{\infty} i^{l-1}\frac{2l+1}{l(l+1)}\psi_l(kr)P_l^1(\cos\theta)\sin\phi \end{array}\right\} \qquad (3.2\text{-}6)$$

From Eq. (3.2-2), ${}^{\mathrm{e}}\Pi$ generates the transverse magnetic, or TM, field, that is, the component of the $\boldsymbol{H}$ field perpendicular to $\boldsymbol{r}$; ${}^{\mathrm{m}}\Pi$ generates the transverse

electric, or TE, field and, therefore, the component of the $\boldsymbol{E}$ field perpendicular to $\boldsymbol{r}$. Here $P_l^1 = -\partial P_l / \partial\theta$. In the azimuthal coordinate ϕ, the vector components in the vector version of Bauer's identity only involve $\sin\phi$ and $\cos\phi$; thus, m in Eq. (3.2-5) can have only the values ± 1. The function $\psi_l(x)$ is related to the Bessel function of the first kind, $J_{l+1/2}(x)$, which is of fractional order $l+1/2$, through the expression[1]

$$\psi_l(x) = \sqrt{\frac{\pi x}{2}} J_{l+1/2}(x) \tag{3.2-7}$$

The scattered waves also can be represented by scalar potentials expressed in a series expansion solution to the Helmholtz equation, except in this case the appropriate cylindrical functions to use in the series involve Hankel functions of the first and second kinds, $H^+_{l+1/2}(x)$ and $H^-_{l+1/2}(x)$. Hankel functions behave asymptotically as $\exp(\pm ikr)/\sqrt{kr}$ for large r, where the plus sign applies to the first kind and the negative sign to the second kind. We will show that the scalar potentials that use $H^+_{l+1/2}(x)$ result in scattered waves that behave as $\exp(ikr)/kr$, as $r \to \infty$, which matches the spherical form that the scattered wave front must assume at large distances from the scattering source. We will need the spherical Hankel functions $\xi_l^{\pm}(x)$, which are defined in terms of $H^{\pm}_{l+1/2}(x)$ by

$$\left.\begin{aligned} \xi_l^+(x) &= \sqrt{\frac{\pi x}{2}} H^+_{l+1/2}(x) = \sqrt{\frac{\pi x}{2}}\left(J_{l+1/2}(x) + iY_{l+1/2}(x)\right) \\ \xi_l^-(x) &= \sqrt{\frac{\pi x}{2}} H^-_{l+1/2}(x) = \sqrt{\frac{\pi x}{2}}\left(J_{l+1/2}(x) - iY_{l+1/2}(x)\right) \end{aligned}\right\} \tag{3.2-8}$$

where $Y_{l+1/2}(x)$ is the Neumann function of order $l+1/2$ and it is the second solution to the Bessel equation, linearly independent of $J_{l+1/2}(x)$. These outgoing and incoming wave forms are discussed further in Chapters 4 and 5.

[1] We define the spherical Bessel functions ψ_l and χ_l in terms of the integer Bessel functions by $\psi_l(x) = (\pi x/2)^{1/2} J_{l+1/2}(x)$, $\chi_l(x) = (\pi x/2)^{1/2} Y_{l+1/2}(x)$, where $J_\nu(x)$ and $Y_\nu(x)$ are the νth-order Bessel functions of the first and second kind; [8] uses these forms for ψ_l and χ_l, but in most references they are divided by x.

3.3 Multiple Internal Reflections

In the case of a sphere that is not strongly absorbing, scattered waves arise from external reflections from the surface of the sphere, and also from refracted waves that pass through the interior of the sphere. We also have to account for refracted rays that exit the sphere after one or more internal reflections; this mechanism can give rise to rainbow effects. This has been shown in a geometric optics context in Fig. 3-2.

Internal to the sphere, standing waves are established that include both incoming and outgoing waves. The outgoing waves must involve Hankel functions of the first kind because of the asymptotic boundary conditions with large r cited above. However, the Hankel functions involve the Neumann function, which is unbounded at the origin. Hence, the appropriate cylindrical function for the aggregate incoming and outgoing series for internal standing waves must be a purely spherical Bessel function, i.e., of the form provided by the $\psi_l(x)$ functions given in Eq. (3.2-7), which are well-behaved at the origin. This can be accomplished by noting from Eq. (3.2-8) that

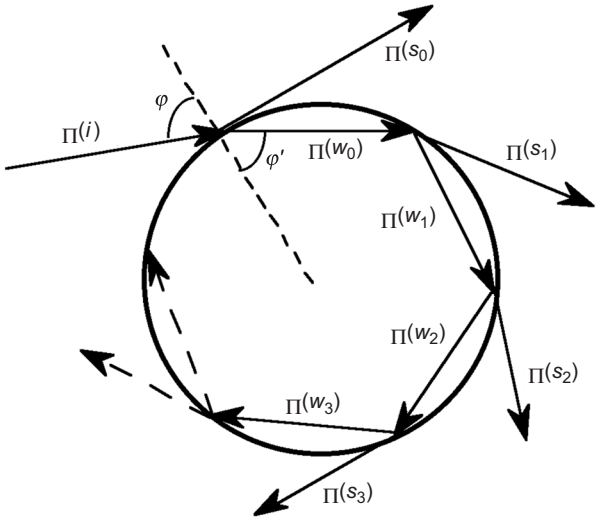

Fig. 3-2. Scattering schematic from a weakly absorbing sphere.

$$\psi_l(x) = \frac{1}{2}\xi_l^+(x) + \frac{1}{2}\xi_l^-(x) \tag{3.3-1}$$

Therefore, we express the scalar potential series for the outgoing internal waves in terms of spherical Hankel functions of the first kind using $\xi_l^+(x)$, and the incoming internal waves by spherical Hankel functions of the second kind using $\xi_l^-(x)$. If these two series are equally weighted, their sum gives an equivalent series in $\psi_l(x)$.

We define the scattered waves in total as comprising those waves that are reflected from the external surface of the sphere as a result of the incident plane wave and also all waves that have passed into the sphere and subsequently escaped from it after zero or more internal reflections. The scalar potential series for the different degrees of scattered waves are defined by

$$\left\{\begin{array}{l} {}^{\mathrm{e}}\Pi^{(s_j)} = \dfrac{E_o}{rk^2}\displaystyle\sum_{l=1}^{\infty} {}^{\mathrm{e}}b_l^{(j)}\xi_l^+(kr)P_l^1(\cos\theta)\,\cos\phi \\ {}^{\mathrm{m}}\Pi^{(s_j)} = \dfrac{E_o}{rk^2}\displaystyle\sum_{l=1}^{\infty} {}^{\mathrm{m}}b_l^{(j)}\xi_l^+(kr)P_l^1(\cos\theta)\,\sin\phi \end{array}\right\}, j = 0,1,2,\cdots; r \geq r_o \tag{3.3-2}$$

The expansion in degree is limited to $m = 1$ because of the form for the incident wave in Eq. (3.2-6) and because of the continuity conditions in electrodynamics [see Eq. (3.5-1)]. The scattered waves being outgoing, their scalar potentials must use $\xi_l^+(x)$. The scattering coefficients $\left({}^{\mathrm{e}}b_l^{(j)}, {}^{\mathrm{m}}b_l^{(j)}\right)$, $j = 0,1,2,\cdots$, are to be determined from the boundary conditions at the surface of the sphere, which will be discussed later. The scattering coefficients $\left({}^{\mathrm{e}}b_l^{(0)}, {}^{\mathrm{m}}b_l^{(0)}\right)$ apply to the zeroth-degree scattered waves, which are those components of the incident plane wave that undergo only an external reflection from the surface of the sphere. For an opaque or an absorbing sphere, only these scattering coefficients would be relevant to our scattering problem. The scattering coefficients $\left({}^{\mathrm{e}}b_l^{(j)}, {}^{\mathrm{m}}b_l^{(j)}\right)$, $j = 1,2,\cdots$, correspond to those components of the incident plane wave that have been transmitted inward and refracted across the surface of the sphere and then have undergone $j-1$ internal reflections within the sphere before escaping from it after being refracted again.

What about the standing waves inside the sphere? Their scalar potentials satisfy Eq. (3.2-1c) inside where $n \neq 1$; they are of the form

$$\left\{\begin{array}{l} {}^{\mathrm{e}}\Pi^{(w_j)} = \dfrac{E_o}{rn^2k^2}\displaystyle\sum_{l=1}^{\infty} {}^{\mathrm{e}}a_l^{(j)}\psi_l(nkr)P_l^1(\cos\theta)\,\cos\phi \\ {}^{\mathrm{m}}\Pi^{(w_j)} = \dfrac{E_o}{rnk^2}\displaystyle\sum_{l=1}^{\infty} {}^{\mathrm{m}}a_l^{(j)}\psi_l(nkr)P_l^1(\cos\theta)\,\sin\phi \end{array}\right\}, j = 0,1,2,\cdots; r \le r_o \qquad (3.3\text{-}3)$$

The $\psi_l(x)$ functions are used to avoid a singularity at the origin. The coefficients $\left({}^{\mathrm{e}}a_l^{(j)}, {}^{\mathrm{m}}a_l^{(j)}\right)$ apply to those waves internal to the sphere that have undergone j internal reflections. These coefficients also are to be determined from the boundary conditions at the surface of the sphere. However, we must distinguish between incoming waves and outgoing waves. For incoming internal waves that have been transmitted across the boundary surface, or reflected from its inner side, we use in the scalar potential series the form $a_l^{(j)}\xi_l^-(nkr)/2$. For outgoing internal waves that are about to be either reflected from the inner side of the boundary surface or transmitted across it, we use the form $a_l^{(j)}\xi_l^+(nkr)/2$. We show that with this formalism in place for the internal waves, the incident external plane wave also will partition naturally into its incoming component where the Hankel functions of the second kind will apply. The other incident component, the "ongoing" waves, so to speak, that effectively do not interact with the sphere at all, must be represented in terms of Hankel functions of the first kind in order to match the asymptotic boundary conditions as $r \to \infty$.

3.4 Fresnel Formulas for Reflection and Transmission Amplitudes

How many internal reflections do we have to take into account for a given accuracy? We should answer this question before taking up the determination of the transmission and scattering coefficients. Guidance can be obtained from the Fresnel formulas in geometric optics for the transmission across and reflection from a dielectric surface. Maxwell's equations require that both tangential components of $\boldsymbol{E}$ and of $\boldsymbol{H}$ must be continuous across any surface that is devoid of free charges and surface currents. Also, the normal components of $\varepsilon\boldsymbol{E}$ and $\mu\boldsymbol{H}$ must be continuous. When these conditions are applied to waves impinging on a plane surface of infinite extent, one obtains the Fresnel formulas [8], which provide the ratios of the transmission and reflection amplitudes relative to the incident amplitude at the boundary. For an incident wave with in-plane polarization of the $\boldsymbol{E}$ field, these ratios are denoted by $T_{\|}/I_{\|}$ and $R_{\|}/I_{\|}$, where $I_{\|}$ is the amplitude of the incident ray. For a cross-plane polarization, they are denoted by $T_{\perp}/I_{\perp}$ and $R_{\perp}/I_{\perp}$. These ratios are

$$\left.\begin{aligned} T_{\|} &= \frac{2n_1\cos\varphi_1}{n_2\cos\varphi_1 + n_1\cos\varphi_2} I_{\|} = \frac{2\sin\varphi_2\cos\varphi_1}{\cos(\varphi_1-\varphi_2)\cdot\sin(\varphi_1+\varphi_2)} I_{\|} \\ R_{\|} &= \frac{n_2\cos\varphi_1 - n_1\cos\varphi_2}{n_2\cos\varphi_1 + n_1\cos\varphi_2} I_{\|} = \frac{\tan(\varphi_1-\varphi_2)}{\tan(\varphi_1+\varphi_2)} I_{\|} \end{aligned}\right\} \quad (3.4\text{-}1a)$$

and

$$\left.\begin{aligned} T_{\perp} &= \frac{2n_1\cos\varphi_1}{n_1\cos\varphi_1 + n_2\cos\varphi_2} I_{\perp} = \frac{2\sin\varphi_2\cos\varphi_1}{\sin(\varphi_1+\varphi_2)} I_{\perp} \\ R_{\perp} &= \frac{n_1\cos\varphi_1 - n_2\cos\varphi_2}{n_1\cos\varphi_1 + n_2\cos\varphi_2} I_{\perp} = -\frac{\sin(\varphi_1-\varphi_2)}{\sin(\varphi_1+\varphi_2)} I_{\perp} \end{aligned}\right\} \quad (3.4\text{-}1b)$$

where the subscript "1" denotes the medium on the (left) incident side of the boundary and the subscript "2" denotes the transmitted side (see Fig. 3-3). The angle of incidence φ_1 is related to the refracted angle φ_2 through Snell's law, $n_1\sin\varphi_1 = n_2\sin\varphi_2$, and the second set of equalities in Eqs. (3.3-1a) and (3.3-1b) are obtained by applying Snell's law. Here the magnetic permeability of the sphere is assumed to be the same as that of the surrounding medium, which is assumed to be a vacuum.

One can easily show from Eqs. (3.4-1a) and (3.4-1b) that $T_{\perp}/T_{\|} = \cos(\varphi_1-\varphi_2) = 1 + \mathrm{O}[n_2 - n_1]$ (with $I_{\perp} = I_{\|} = 1$), and similarly for $R_{\perp}/R_{\|}$, one obtains $R_{\perp}/R_{\|} = -\cos(\varphi_1+\varphi_2)/\cos(\varphi_1-\varphi_2)$. For near-grazing conditions and $N << 1$, $R_{\perp}/R_{\|}$ also is essentially unity. Therefore, we consider in the sequel only the in-plane Fresnel coefficients and we set $I_{\|} \equiv 1$.

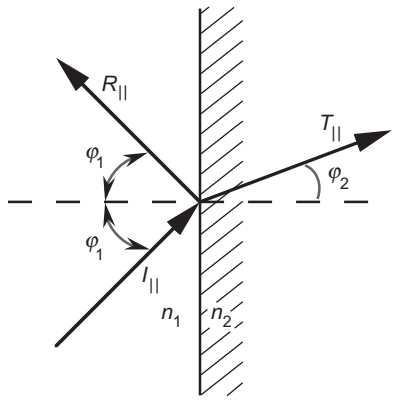

Fig. 3-3. Incident, reflected, and transmitted waves across a planar boundary.

The reflection coefficient $R_{||}$ will be negative if $n_2 > n_1$ when φ_1 exceeds the Brewster angle, that is, when $\varphi_1 + \varphi_2 > \pi/2$. This is to be interpreted as a reversal in phase of the reflected ray of π radians relative to the phase of the incident ray. In our radio occultation problem, a ray at a near-grazing angle at a boundary ensures that $\varphi_1 + \varphi_2 \approx \pi$ and, therefore, that a phase reversal always occurs for reflections from the incident side when $n_1 < n_2$, that is, the external reflection.

Figure 3-4 is a plot of the in-plane coefficients as a function of the impact parameter or impact distance of the ray, which in this case is the penetration depth h of the undeflected ray into the surface of the sphere (for a sphere, $\varphi \equiv \pi - \theta$) and given by

$$h = r_o(1 - \sin\theta), \quad \varphi = \pi - \theta, \quad \pi/2 \le \theta \le \pi \tag{3.4-2}$$

One concludes from Fig. 3-4 that the transition from total external reflection to nearly total transmission is completed within a very short impact distance when Nr_o is small. For $N = 1.2 \times 10^{-6}$ and $r_o = 6400$ km, the equivalent impact distance for this figure is less than 10 m, which is far smaller than the first Fresnel zone, $(\lambda D)^{1/2} \approx 1$km. Figure 3-4 and, in fact, the concept of "impact distance" apply to the LEO observables only where $Nr_o >> (\lambda D)^{1/2}$, or for a LEO using Global Positioning System (GPS) signals, when $N >> 10^{-4}$. In our examples to follow, $N << 10^{-4}$, so the first Fresnel zone acts like a

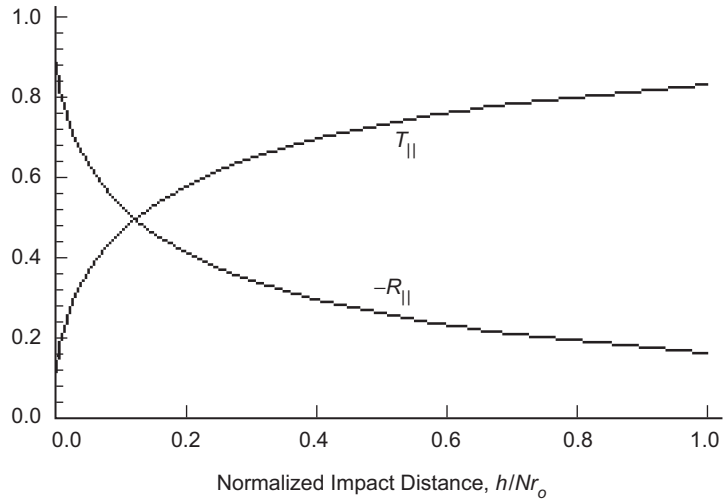

Fig. 3-4. Transmission and reflection ratios for an external wave impinging on a sphere of radius r_o and uniform refractivity N.

"smoothing" filter, blurring the variability implied by Fig. 3-4 that otherwise would be observed by the LEO.

We note from Eq. (3.4-1a) that

$$n_2 T_{\|} = n_1(1 + R_{\|}) \tag{3.4-3}$$

which will be useful later in evaluating the coefficients in Mie scattering theory.

3.4.1 Conservation Principle

Let $R[\varphi, n]$ and $T[\varphi, n]$ denote the reflection and transmission ratios, respectively, for an external wave with either in-plane or cross-plane polarization impinging on the outer surface at an angle of incidence φ (see Fig 3-3) and with $n = n_2 / n_1$. Similarly, $R\left[\varphi', n^{-1}\right]$ and $T\left[\varphi', n^{-1}\right]$ are the reflection and transmission ratios for an internal wave impinging on the inner side of the boundary surface at an angle of incidence φ', which is given from Snell's law by $\varphi' = \arcsin\left(n^{-1} \sin\varphi\right)$.

Let $A^{(j)}$ be the amplitude of the scattered ray of jth degree at the surface of the sphere (disregarding space losses and using a geometric optics approach). Then, it follows for a perfectly transparent sphere that $A^{(j)}$ is given by

$$\left.\begin{aligned} A^{(0)} &= R[\varphi, n] \\ A^{(j)} &= T[\varphi, n]\left(R\left[\varphi', n^{-1}\right]\right)^{j-1} T\left[\varphi', n^{-1}\right], \quad j = 1, 2, \cdots \end{aligned}\right\} \tag{3.4-4}$$

From Poynting's vector, the radiant power per unit area carried by an incident wave striking a boundary is given by $-\hat{\boldsymbol{r}} \cdot \boldsymbol{S}^{(i)}$, where $\hat{\boldsymbol{r}}$ is the outward unit normal vector to the tangent plane at the point of impact of the ray with the boundary surface. This quantity must be conserved in a non-absorbing medium. Therefore, the total radiant power carried away by the reflected and transmitted waves at the boundary must equal the incident power; that is,

$$\hat{\boldsymbol{r}} \cdot \boldsymbol{S}^{(i)} + \hat{\boldsymbol{r}} \cdot \boldsymbol{S}^{(R)} - \hat{\boldsymbol{r}} \cdot \boldsymbol{S}^{(T)} = 0 \tag{3.4-5}$$

$\boldsymbol{S}^{(R)}$ and $\boldsymbol{S}^{(T)}$ are the Poynting vectors for the reflected and transmitted waves. From Snell's law and the Poynting vector for plane waves ($|\boldsymbol{S}| = c\sqrt{\varepsilon / \mu} E^2 / 4\pi$), it can be shown for an incident ray at both the external and internal sides of the boundary that the conservation constraint for power given in Eq. (3.4-5) results in the conditions

$$\left.\begin{aligned}\left(R[\varphi,n]\right)^2+\left(T[\varphi,n]\right)^2\frac{\tan\varphi}{\tan\varphi'}&=1\\ \left(R\left[\varphi',n^{-1}\right]\right)^2+\left(T\left[\varphi',n^{-1}\right]\right)^2\frac{\tan\varphi'}{\tan\varphi}&=1\end{aligned}\right\} \tag{3.4-6}$$

From Eqs. (3.4-4) and (3.4-6), it follows that

$$\sum_{j=0}^{\infty}\left(A^{(j)}\right)^2=1 \tag{3.4-7}$$

Thus, $\left(A^{(j)}\right)^2$ can be considered as the relative intensity of the jth scattered component.

3.4.2 Scattering Angles and Intensities

The scattering angle $\Psi^{(j)}(\varphi,n)$ that the jth degree ray undergoes upon departing from the sphere after $j-1$ internal reflections is easily obtained from Snell's law and ray tracing through the sphere (see Fig. 3-2); this angle is given by

$$\left.\begin{aligned}\Psi^{(j)}&=(\pi-2\varphi)(j-1)+2j(\varphi-\varphi')\\ \varphi'&=\sin^{-1}\left(n^{-1}\sin\varphi\right),\ j=0,1,2,\cdots\end{aligned}\right\} \tag{3.4-8}$$

where $\Psi^{(j)}(\varphi,n)$ is measured clockwise from the positive z-axis (see Fig. 3-1). Thus, $\Psi^{(0)}$ is the scattering angle for the external reflected wave (and, therefore, negative); $\Psi^{(1)}$ is the angle for the "primary" ray, which passes through the sphere without reflection and is doubly refracted. The scattering angle $\Psi^{(2)}$ is for a ray that is internally reflected once and refracted twice. When $\Psi^{(2)}$ assumes its stationary value with respect to the incident angle φ of the incoming ray, that is, when a caustic surface is formed (see Appendix B), that value of $\Psi^{(2)}$ marks the scattering angle of the primary rainbow from a raindrop. Similarly, the stationary value of $\Psi^{(3)}$ marks the direction of the secondary rainbow.

We can evaluate the relative importance of the higher-degree scattered rays to our diffraction problem by plotting both $\Psi^{(j)}$ and $A^{(j)}$ versus incident angle φ or, equivalently, the impact distance of the ray [Eq. (3.3-2)], which is measured positive downward from the surface at $r=r_o$.

Figure 3-5 displays the normalized scattering angle $\Psi^{(j)}(h)/(2N)^{1/2}$ versus $h/\left(Nr_o\right)$ for various scattering degrees. In Fig. 3-5(a), the $j=0$ curve is

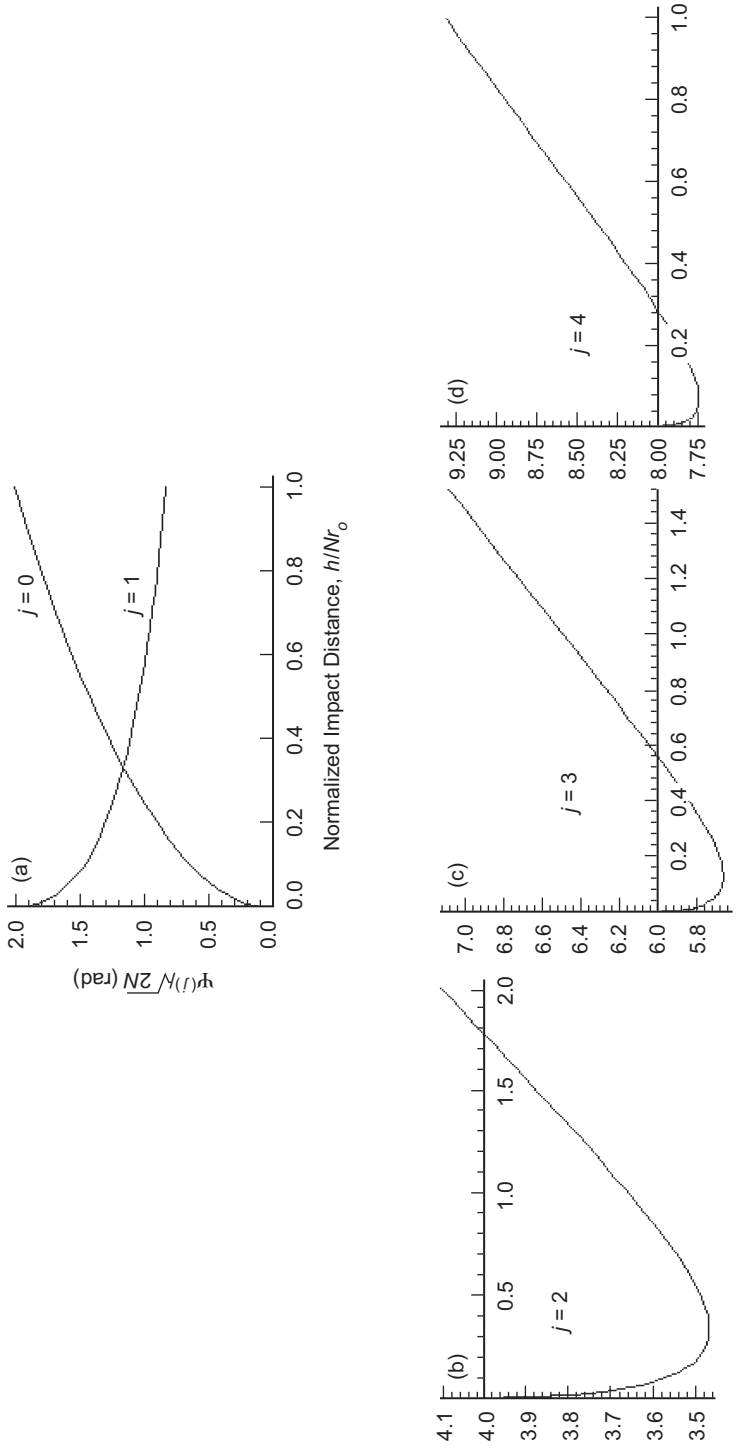

Fig. 3-5. Normalized scattering angle profiles for different j values: (a) external reflected wave and refracted wave, no internal reflection, and scattering angle for waves with (b) one, (c) two, and (d) three internal reflections.

for $-\Psi^{(0)}(2N)^{-1/2}$. In Fig. 3-5(b), rainbow caustics occur where $\partial\Psi^{(j)}/\partial h = 0$; the impact distance giving this stationary value is defined by $h = h_*^{(j)}$. We will show in Mie scattering theory (see Section 3.16 and Fig. 3-29) that when N exceeds a threshold limit, $N >\sim \left(\lambda / r_o\right)^{2/3}$, $\approx 10^{-5}$, the scattering intensity will become prominent at a scattering angle near $2(6N)^{1/2}$ for $j = 2$. Below this threshold, the rainbow effects, particularly when seen at LEO distances, will be largely washed out. We are dealing with a single sphere [in contrast to an ensemble of raindrops, all refracting and reflecting coherently, and an ensemble whose extent subtends an angle at the observer that is much larger than that subtended by the first Fresnel zone $(\lambda / D)^{1/2}$)]. Consequently, the rainbow effects from the single sphere that are seen by the LEO will be mostly washed out if $h_* << (\lambda D)^{1/2}$.

For near-grazing ray path conditions ($\varphi \approx \pi / 2$), the stationary value of the scattering angle with respect to impact distance is given by

$$\Psi_{\min}^{(j)} \doteq \frac{2}{n}\sqrt{\frac{n^2-1}{j^2-1}}\left(j^2 - n\right) = 2\left[2N\left(j^2-1\right)\right]^{1/2} + \mathrm{O}\left[N^2\right] \tag{3.4-9}$$

Moreover, the impact distance providing that stationary value is given by

$$h_*^{(j)} \doteq r_o\left(1 - \sqrt{\frac{j^2-n^2}{j^2-1}}\right) = \frac{Nr_o}{j^2-1} + \mathrm{O}\left[N^2\right] \tag{3.4-10}$$

Figure 3-6 shows the relative amplitudes of the scattered waves $A^{(j)}(h)$. It can be shown that the value $A_*^{(j)}$ at the stationary scattering angle is given by

$$A_*^{(j)} = \frac{4j}{(j+1)^2}\left(\frac{j-1}{j+1}\right)^{j-1} \tag{3.4-11}$$

which approaches $4/(j\mathrm{e}^2)$ for large values of j. Appreciable intensities occur in scattering for degrees $j \geq 2$ because the near-grazing angles of incidence result in significant internal reflections.

Figure 3-7 shows that as the impact distance $h \to 0^+$, an infinite number of quasi-evanescent waves are generated (see Fig. 3-2). These waves follow chords in the sphere; as the impact distance $h \to 0^+$, the angle that these chords make at their vertices with respect to the normal vector to the surface approaches the critical reflection angle $\arcsin(1/n)$. These boundary waves spin off scattered waves like a fireworks pinwheel. The Fresnel formulas can be used to figure out the damping rate. When the refractivity is such that a

rotational commensurability exists, that is, when there is an exact integer number of chords spanning the sphere after one revolution ($\pi(2N)^{-1/2}$ = an integer), a kind of resonance occurs at these scattering points. This has an analog in complex angular momentum theory of sub-atomic particles in high-energy physics involving scattering resonances for selected impact parameter/energy values of the bombarding particles.

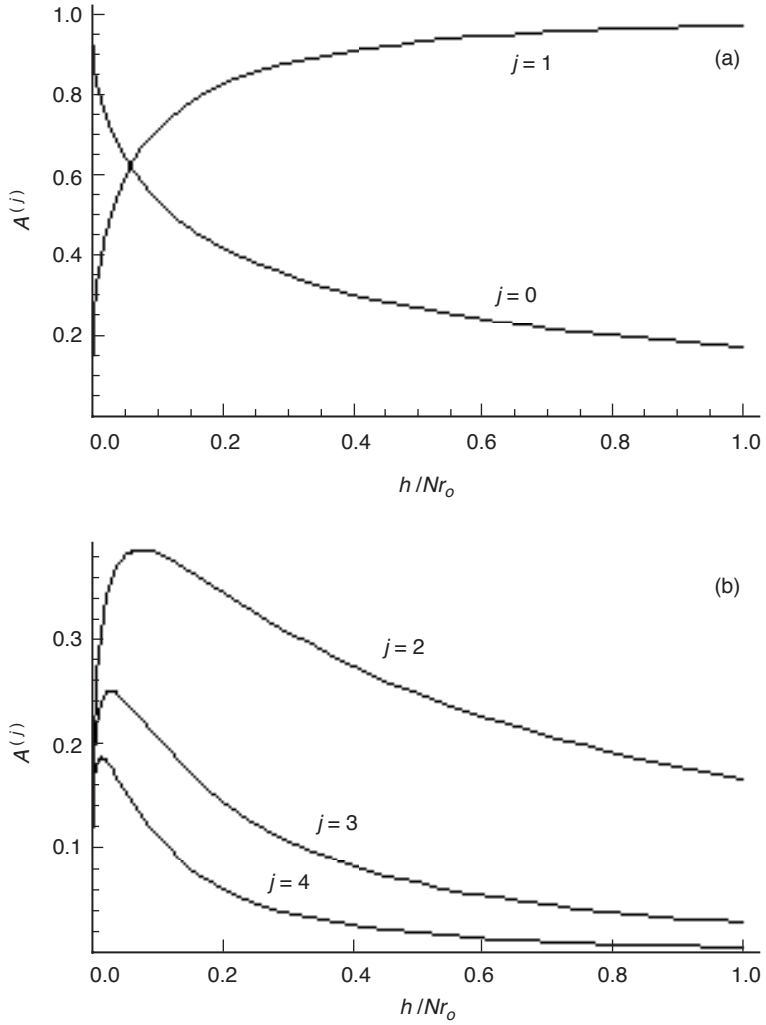

Fig. 3-6. Relative intensities of scattered waves: (a) external reflected wave, and refracted wave with no internal reflection, and (b) refracted waves with one, two, and three internal reflections.

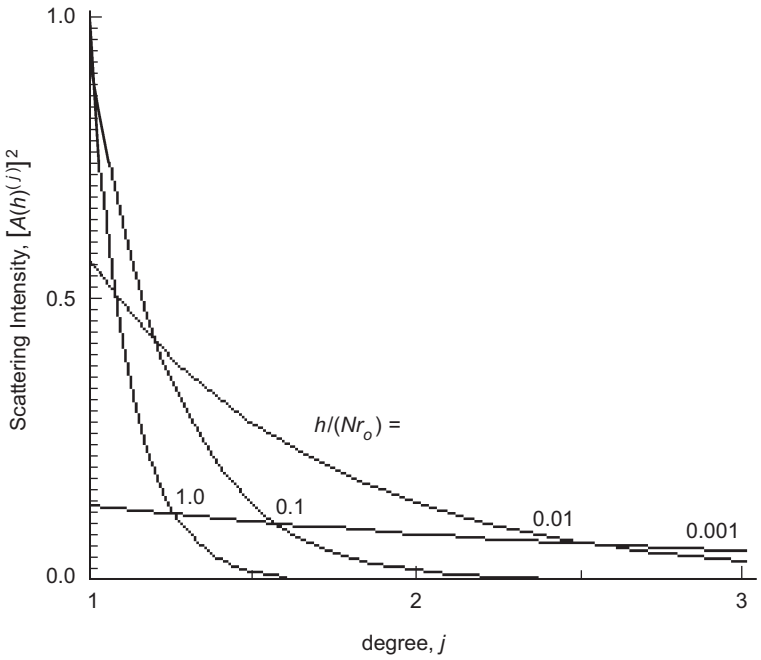

Fig. 3-7. Scattering intensity versus scattering degree shown parametrically with impact distance.

3.4.3 Caustics

Another measure of the contributions of the higher-degree scattered waves is to examine their "intensity density." The impact distances of the ray paths of the incident plane waves will be equally distributed. Therefore, the scattered waves will "pile up" at a scattering angle that has a stationary value with respect to the impact distance of the impinging external ray, that is, at an impact distance where $\partial\Psi^{(j)}/\partial h = 0$. This defines the neighborhood where a caustic surface is formed by the ray system inside the sphere; one example of this is manifested by the rainbow in scattering from a raindrop. We define "intensity density" as $\left(A^{(j)}\right)^2\left(\partial\Psi^{(j)}/\partial h\right)^{-1}$, which has been shown in Fig. 3-8 for the low-degree scattering. Caustic resonances occur for all scattering degrees when $j \geq 2$, but with rapidly diminishing effect with increasing j [see Eq. (3.4-11)].

We note that geometric optics predicts an infinite intensity density at the caustic resonances. Geometric optics is based on second-order stationary-phase theory, which breaks down at caustics. A third-order treatment leads to more realistic predictions. However, Mie scattering theory yields the correct predictions for these difficult points. Even without Mie scattering theory, we

can see early on from this purely geometric optics-based discussion that the thin-screen model will not account for the scattering at degrees higher than 1 without further enhancement.[2] We will show that Mie scattering theory yields resonances at these scattering angles near $2\left(2N\left(j^2-1\right)\right)^{1/2}$. We should also expect to see interference fringes at a scattering angle slightly above the stationary point that arise from incident plane waves arriving at the sphere with their ray paths at slightly offset impact distances above and below the point providing the stationary scattering angle. Therefore, scattered rays that arrive at the LEO from these separate impact points will have followed slightly different paths, which will have slightly different optical lengths. These unequal optical lengths result in phase offsets in the arriving rays at the LEO and, hence, phase interference will result. This interference has its analog in supernumerary arcs seen sometimes on the bright side of the primary rainbow, which were first explained by Thomas Young in the early 19th century in terms of his wave theory for light. For scattering higher than degree $j = 2$, these effects are diminished [see Eq. (3.4-11)] and also the scattering angles are larger, increasing essentially linearly with increasing degree [see Eq. (3.4-9)].

For the LEO, which is at a distance D from the limb, this resonance occurs when $\partial\Psi^{(j)}/\partial h = D^{-1} \approx 0$. For the degree $j = 2$ caustic, this occurs at an epoch $\sim 2\times 10^3 (2N)^{1/2}$ sec after first passing into the geometric shadow zone. Thus,

---

[2] We could, however, embed the thin-screen model with a modified phase and amplitude profile that attempts to account for the multi-degree scattering. Let $E^{(j)}\left(h_{\mathrm{LG}}\right)\exp\left[i\vartheta^{(j)}\left(h_{\mathrm{LG}}\right)\right]$ denote the amplitude and phase of the *j*th degree scattered ray that is observed by the LEO, which is located at a perpendicular distance D from the thin screen and at an altitude h_{LG} relative to the altitude of the zero impact point, i.e., at $r = r_o, \theta = \pi/2$. Using the thin-screen model, and expressing distances in phase units, the Rayleigh–Sommerfeld integral becomes

$$E^{(j)}\exp\left[i\vartheta^{(j)}\right] = \sqrt{\frac{1}{\pi D}}\left(\frac{1}{\pi D}\right)\frac{1}{1+i}\int_0^\infty A^{(j)}\left(h_S\right)\exp\left[i\frac{\left(h_S-h_{\mathrm{LG}}\right)^2}{D}+i\varepsilon^{(j)}\left(h_S\right)\right]dh_S,\ j=0,1,\cdots$$

where $\varepsilon^{(j)}\left(h_S\right)$ is the phase embedded in the thin screen at an impact distance h_S due to the *j*th scatterer, and $A^{(j)}\left(h_S\right)$ is its relative amplitude. For each value of *j*, the scattered ray must satisfy the stationary-phase condition; therefore, $d\varepsilon^{(j)}/dh_S = -k\Psi^{(j)}$, where $\Psi^{(j)}(h_S)$ is given from Eqs. (3.4-2) and (3.4-8). The relative amplitude $A^{(j)}\left(h_S\right)$ is given from Eq. (3.4-4). Fresnel diffraction effects from the *j*th-degree scattering principally would be evident at scattering angles that are in the neighborhood of the *j*th rainbow caustic, i.e., where $d\Psi^{(j)}(h_S)/dh_S = 1/D \approx 0$. In this modified form, the combined electric field observed by the LEO would be given by summing the weighted phasors for all degrees, $j = 0,1,2\ldots$.

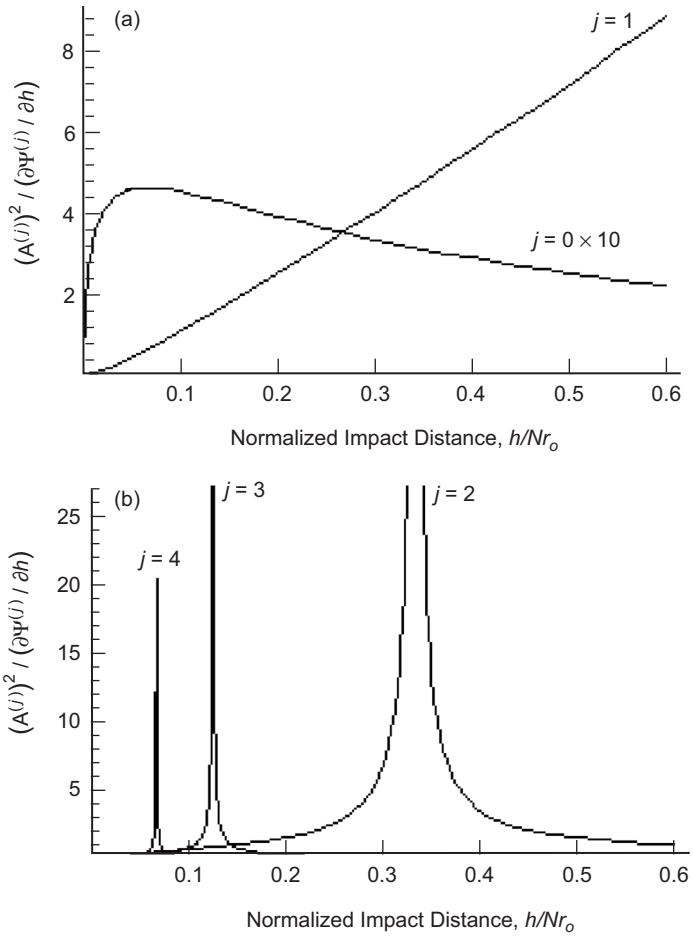

Fig. 3-8. Intensity densities for scattered waves: (a) external reflected wave, and refracted wave with no internal reflection, and (b) refracted waves with one, two, and three internal reflections.

from a super-refracting marine layer with a near hard discontinuity of N across its upper boundary, we might expect to "see" these rainbow effects of degree $j = 2$ as the LEO descends deeper into the refracted shadow zone. In our case, the orbital velocity of a LEO at an orbital radius of $1.1r_o$ is of the order of 1 mrad/ms. So, even when the deceleration of the rate of decent of the impact parameter of the ray from defocusing in the lower troposphere is taken into account, transient rainbow boundary phenomena are completed within seconds (see Fig. 3-29 for $j = 2$ and $j = 3$ effects).

To fully account for the scattering exactly at the shadow zone boundary, one needs to keep a large number of scattering terms. Even Mie theory may have computational difficulty on such boundaries when N is large. The boundary (or more precisely, the chords at the critical internal reflection angle) of our spherical model is a caustic surface, which presents a singular situation.

It is ironic that our seemingly simple, transparent sphere would present such complications. An opaque sphere is easier.

3.5 Mie Scattering Theory: Obtaining the Scattering Coefficients at a Boundary

To obtain the relationships between the scattering and transmission coefficients, we again use the conditions from Maxwell's equations that the tangential components of $\boldsymbol{E}$ and $\boldsymbol{H}$ must be continuous across a charge and current-free boundary that bears a discontinuity in ε and/or μ: $\boldsymbol{E}_{\tan}^{(i)} + \boldsymbol{E}_{\tan}^{(S)} = \boldsymbol{E}_{\tan}^{(w)}$; $\boldsymbol{H}_{\tan}^{(i)} + \boldsymbol{H}_{\tan}^{(S)} = \boldsymbol{H}_{\tan}^{(w)}$. In addition, Maxwell's equations also require that the radial components of $\varepsilon\boldsymbol{E}$ and $\mu\boldsymbol{H}$ be continuous across the boundary. These conditions were used earlier to obtain the Fresnel formulas. It follows that a correspondence between these coefficients and the Fresnel ratios should exist, which will be established.

3.5.1 Transmission and Reflection Coefficients for Incident External Waves

First we will deal with the incident external wave and its reflection and transmission at the boundary surface. As the incident external wave is an incoming one at the boundary, its scalar potential series in Eq. (3.2-6) will naturally partition into a series where $\xi_l^-(x)$ is used for the incident component, and into a series for the "ongoing" component where $\xi_l^+(x)$ is used. Why is this plausible? There are two reasons. In the limit as $N \to 0$ the incoming external waves must match the incoming internal waves at the boundary, which involves only $\xi_l^-(x)$. The second reason, as we will show shortly, is that when we calculate the scattering coefficients $({}^{\mathrm{e}}b_l^{(0)}, {}^{\mathrm{m}}b_l^{(0)})$ that arise from the incident external waves, they cancel completely the $\xi_l^+(x)$ component of that incident field.

The continuity conditions on the components of $\boldsymbol{E}$ and $\boldsymbol{H}$, which were mentioned above, are given in terms of an equivalent set of continuity conditions on the scalar potentials for these waves, ${}^{\mathrm{e}}\Pi r$ and ${}^{\mathrm{m}}\Pi r$, and also on the radial components of their gradients across the surface of the sphere [8]. These are given by

$$\left.\begin{aligned}
\frac{\partial}{\partial r}\left[r\left({}^{\mathrm{e}}\Pi^{(i)}+{}^{\mathrm{e}}\Pi^{(s_o)}\right)\right]_{r=r_o^+} &= \frac{\partial}{\partial r}\left[r\left({}^{\mathrm{e}}\Pi^{-(w_o)}\right)\right]_{r=r_o^-} \\
\frac{\partial}{\partial r}\left[r\left({}^{\mathrm{m}}\Pi^{(i)}+{}^{\mathrm{m}}\Pi^{(s_o)}\right)\right]_{r=r_o^+} &= \frac{\partial}{\partial r}\left[r\left({}^{\mathrm{m}}\Pi^{-(w_o)}\right)\right]_{r=r_o^-} \\
\left[r\left({}^{\mathrm{e}}\Pi^{(i)}+{}^{\mathrm{e}}\Pi^{(s_o)}\right)\right]_{r=r_o^+} &= n^2\left[r\left({}^{\mathrm{e}}\Pi^{-(w_o)}\right)\right]_{r=r_o^-} \\
\left[r\left({}^{\mathrm{m}}\Pi^{(i)}+{}^{\mathrm{m}}\Pi^{(s_o)}\right)\right]_{r=r_o^+} &= \left[r\left({}^{\mathrm{m}}\Pi^{-(w_o)}\right)\right]_{r=r_o^-}
\end{aligned}\right\} \tag{3.5-1}$$

Here $\Pi^{-(w_0)}$ denotes the partitioned component of the scalar potential for the interior wave that has just crossed the boundary, and, therefore, it corresponds to an incoming wave. In general, the partitioned scalar potential series for interior incoming (–) and outgoing (+) waves are given from Eqs. (3.3-1) and (3.3-3) by

$$\left.\begin{aligned}
{}^{\mathrm{e}}\Pi^{\pm(w_j)} &= \frac{E_o}{2rn^2k^2}\sum_{l=1}^{\infty}{}^{\mathrm{e}}a_l^{(j)}\xi_l^{\pm}(nkr)P_l^1(\cos\theta)\cos\phi \\
{}^{\mathrm{m}}\Pi^{\pm(w_j)} &= \frac{E_o}{2rnk^2}\sum_{l=1}^{\infty}{}^{\mathrm{m}}a_l^{(j)}\xi_l^{\pm}(nkr)P_l^1(\cos\theta)\sin\phi
\end{aligned}\right\}, j = 0,1,\cdots; r \le r_o \tag{3.5-2}$$

We apply the continuity conditions in Eq. (3.5-1) at all points on the surface of the sphere to the series expansions to the scalar potentials for the incident plane wave in Eq. (3.2-6) and to the zeroth-degree scattering and transmitted components in Eqs. (3.3-2) and (3.5-2). Using the orthogonality properties of the spherical harmonics, we obtain for each integer value of spectral number the transfer equations required by the continuity conditions:

$$\left.\begin{aligned}
{}^{\mathrm{e}}b_l^{(0)}\xi_l^{+'}(x_o) + i^{l-1}\frac{2l+1}{l(l+1)}\psi_l'(x_o) &= \frac{1}{2n}{}^{\mathrm{e}}a_l^{(0)}\,\xi_l^{-'}(nx_o) \\
{}^{\mathrm{m}}b_l^{(0)}\,\xi_l^{+'}(x_o) + i^{l-1}\frac{2l+1}{l(l+1)}\psi_l'(x_o) &= \frac{1}{2}{}^{\mathrm{m}}a_l^{(0)}\,\xi_l^{-'}(nx_o) \\
{}^{\mathrm{e}}b_l^{(0)}\,\xi_l^{+}(x_o) + i^{l-1}\frac{2l+1}{l(l+1)}\psi_l(x_o) &= \frac{1}{2}{}^{\mathrm{e}}a_l^{(0)}\,\xi_l^{-}(nx_o) \\
{}^{\mathrm{m}}b_l^{(0)}\,\xi_l^{+}(x_o) + i^{l-1}\frac{2l+1}{l(l+1)}\psi_l(x_o) &= \frac{1}{2n}{}^{\mathrm{m}}a_l^{(0)}\,\xi_l^{-}(nx_o)
\end{aligned}\right\} l = 1,2,\cdots \tag{3.5-3}$$

where $\xi_l^{\pm'}(x)$ is the derivative of $\xi_l^{\pm}$ with respect to x. We define the dimensionless radial coordinate

$$x = kr, \;\; x_o = kr_o \tag{3.5-4}$$

Thus, $x_o = 2.1 \times 10^8$ for the L1 GPS carrier when $r_o = 6400\,\text{km}$; "x" here should not be confused with the "up" spatial coordinate shown in Fig. 3-1. Unless otherwise noted, the definition given in Eq. (3.5-4) will apply henceforth.

Solving the system of linear equations in Eq. (3.5-3) and using the relationship between spherical Bessel functions and spherical Hankel functions given in Eq. (3.3-1), we obtain for the scattering and transmission coefficients of zeroth degree

$$\left.\begin{aligned}
{}^{\mathrm{e}}b_l^{(0)} &= -i^{l-1}\frac{2l+1}{l(l+1)}\left[\frac{1}{2}+\frac{1}{2}\frac{\xi_l^-(x_o)\xi_l^{-\prime}(nx_o)-n\xi_l^{-\prime}(x_o)\xi_l^-(nx_o)}{\xi_l^+(x_o)\xi_l^{-\prime}(nx_o)-n\xi_l^{+\prime}(x_o)\xi_l^-(nx_o)}\right] \\
{}^{\mathrm{m}}b_l^{(0)} &= -i^{l-1}\frac{2l+1}{l(l+1)}\left[\frac{1}{2}+\frac{1}{2}\frac{n\xi_l^-(x_o)\xi_l^{-\prime}(nx_o)-\xi_l^{-\prime}(x_o)\xi_l^-(nx_o)}{n\xi_l^+(x_o)\xi_l^{-\prime}(nx_o)-\xi_l^{+\prime}(x_o)\xi_l^-(nx_o)}\right] \\
{}^{\mathrm{e}}a_l^{(0)} &= i^{l-1}\frac{2l+1}{l(l+1)}\left[\frac{-2i}{\xi_l^+(x_o)\xi_l^{-\prime}(nx_o)-n\xi_l^{+\prime}(x_o)\xi_l^-(nx_o)}\right]n \\
{}^{\mathrm{m}}a_l^{(0)} &= i^{l-1}\frac{2l+1}{l(l+1)}\left[\frac{-2i}{n\xi_l^+(x_o)\xi_l^{-\prime}(nx_o)-\xi_l^{+\prime}(x_o)\xi_l^-(nx_o)}\right]n
\end{aligned}\right\} \tag{3.5-5}$$

Here, the Wronskian of the spherical Hankel functions,

$$\mathcal{W}[\xi_l^+(x),\xi_l^-(x)] = \xi_l^+(x)\xi_l^{-\prime}(x) - \xi_l^{+\prime}(x)\xi_l^-(x) = -2i \tag{3.5-6}$$

has been used to obtain the transmission coefficients. If we insert these expressions for the external reflection coefficients $({}^{\mathrm{e}}b_l^{(0)}, {}^{\mathrm{m}}b_l^{(0)})$ into Eq. (3.3-2) for the zeroth-degree scattered wave, we note that the term $-i^{l-1}(2l+1)/(2l(l+1))$ in that series just cancels exactly the $\xi_l^+(x)$ component of the scalar potential series for the incident external plane wave given in Eq. (3.2-6). In other words, the $\xi_l^+(x)$ component in the spectral representation of the incident plane wave never makes it across the boundary of the sphere. Therefore, we can ignore it in calculating all transmission and reflection coefficients, if for $({}^{\mathrm{e}}b_l^{(0)}, {}^{\mathrm{m}}b_l^{(0)})$ we also eliminate the $-i^{l-1}(2l+1)/2l(l+1)$ term in Eq. (3.5-5), the so-called "–1/2" term.

Also, the "electric" coefficients $({}^{\mathrm{e}}a_l^{(0)}, {}^{\mathrm{e}}b_l^{(0)})$ and the "magnetic" coefficients $({}^{\mathrm{m}}a_l^{(0)}, {}^{\mathrm{m}}b_l^{(0)})$ differ from each other's counterpart by a small quantity of order N when near-grazing conditions are assumed. We will focus primarily on the electric potential and in-plane ($\phi = 0$) observations in future sections.

We define

$$\begin{Bmatrix} \mathrm{e} \\ \mathrm{m} \end{Bmatrix} \tilde{b}_l^{(0)} = \begin{Bmatrix} \mathrm{e} \\ \mathrm{m} \end{Bmatrix} b_l^{(0)} + i^{l-1} \frac{2l+1}{2l(l+1)} \tag{3.5-7}$$

Thus, $({}^{\mathrm{e}}\tilde{b}_l^{(0)}, {}^{\mathrm{m}}\tilde{b}_l^{(0)})$ are the zeroth-degree scattering terms that result from the partitioned scalar potential series for incoming external waves that is comprised of $\xi_l^-(x)$ terms only. We note from Eq. (3.5-5) that

$$\begin{Bmatrix} \mathrm{e} \\ \mathrm{m} \end{Bmatrix} \tilde{b}_l^{(0)} 2\frac{\xi_l^+(x_o)}{\xi_l^-(x_o)} + 1 = \begin{Bmatrix} \mathrm{e} \\ \mathrm{m} \end{Bmatrix} a_l^{(0)} \frac{\xi_l^-(nx_o)}{\xi_l^-(x_o)} \tag{3.5-8}$$

The similarity to Eq. (3.4-3) for the relationship between the Fresnel ratios should be noted. If we establish the following correspondence:

$$2\frac{\xi_l^+(x_o)}{\xi_l^-(x_o)} \tilde{b}_l^{(0)} \leftrightarrow R[\varphi, n], \quad n^{-1}\frac{\xi_l^-(nx_o)}{\xi_l^-(x_o)} a_l^{(0)} \leftrightarrow T[\varphi, n] \tag{3.5-9}$$

we would expect the reflection and transmission coefficients, $a_l^{(j)}$ and $\tilde{b}_l^{(j)}, j \geq 1$, for the interior standing waves and escaping waves to adhere to the form given for the amplitude ratios $A^{(j)}$ given in Eq. (3.4-4). In this case, the Fresnel ratios are complex and dependent on the spectral number l. The ratio $\xi_l^+(x_o)/\xi_l^-(x_o)$ provides phase information. Note that $\xi_l^-(x)$ is the complex conjugate of $\xi_l^+(x)$ when x is real. Let $2\Omega^{(0)}$ be the phase of this ratio; it follows that $\mathrm{Exp}[i2\Omega^{(0)}] = \xi_l^+(x_o)/\xi_l^-(x_o)$.

3.5.2 Transmission and Reflection Coefficients for Incident Interior Waves

For interior waves reflecting from the inner boundary or escaping across it, the continuity conditions from electrodynamics require

$$\left.\begin{aligned}
\frac{\partial}{\partial r}\left[r\left({}^{\mathrm{e}}\Pi^{+(w_{j-1})}+{}^{\mathrm{e}}\Pi^{-(w_j)}\right)\right]_{r=r_o^-} &= \frac{\partial}{\partial r}\left[r\left({}^{\mathrm{e}}\Pi^{(s_j)}\right)\right]_{r=r_o^+} \\
n^2\left[r\left({}^{\mathrm{e}}\Pi^{+(w_{j-1})}+{}^{\mathrm{e}}\Pi^{-(w_j)}\right)\right]_{r=r_o^-} &= \left[r\left({}^{\mathrm{e}}\Pi^{(s_j)}\right)\right]_{r=r_o^+} \\
\frac{\partial}{\partial r}\left[r\left({}^{\mathrm{m}}\Pi^{+(w_{j-1})}+{}^{\mathrm{m}}\Pi^{-(w_j)}\right)\right]_{r=r_o^-} &= \frac{\partial}{\partial r}\left[r\left({}^{\mathrm{m}}\Pi^{(s_j)}\right)\right]_{r=r_o^+} \\
\left[r\left({}^{\mathrm{m}}\Pi^{+(w_{j-1})}+{}^{\mathrm{m}}\Pi^{-(w_j)}\right)\right]_{r=r_o^-} &= \left[r\left({}^{m}\Pi^{(s_j)}\right)\right]_{r=r_o^+} \quad j=1,2,\cdots
\end{aligned}\right\} \tag{3.5-10}$$

Here ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}\Pi^{+(w_{j-1})}$ represents the outgoing component of the internal standing wave that impinges on the inner surface of the sphere after having undergone $j-1$ prior internal reflections. Thus, ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}a_l^{(j-1)}$, which has been determined already from the $j-1$st encounter with the boundary where it was a reflection coefficient (for $j \geq 2$), assumes the "incident wave" role in the jth encounter. On the other hand, ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}\Pi^{-(w_j)}$ represents the reflected and, therefore, incoming wave just after its jth internal reflection. The series forms for ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}\Pi^{+(w_{j-1})}$ and ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}\Pi^{-(w_j)}$ are the partitioned forms of Eq. (3.3-3) and are given in Eq. (3.5-2). The scalar potential in the right-hand side (RHS) of Eq. (3.5-10) represents the outgoing transmitted waves, which are the scattered waves of jth degree, and is given by Eq. (3.3-2). Here ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}b_l^{(j)}$ acts as a transmission coefficient at this jth internal encounter with the boundary. Applying the continuity conditions in Eq. (3.5-10) to all interior points on the surface of the sphere, one obtains the recurrence relations for ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}a_l^{(j)}$ and ${}^{\left\{\mathrm{e}\atop\mathrm{m}\right\}}b_l^{(j)}$:

$$
\left.
\begin{aligned}
&\begin{Bmatrix} e \\ m \end{Bmatrix} b_l^{(0)} = -\frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)} \left(1 + \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^-}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right), \\
&\begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(0)} = i^{l-1} \frac{2l+1}{l(l+1)} \left(\frac{-2i}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right) n, \\
&\left.
\begin{aligned}
&\begin{Bmatrix} e \\ m \end{Bmatrix} b_l^{(j)} = -i \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(j-1)}}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} = -i \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(j-1)}}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \left(- \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right)^{j-1} \\
&= \frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)} n \left[\left(\frac{-2i}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right) \left(- \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right)^{j-1} \left(\frac{-2i}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right) \right], \\
&\begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(j)} = -\begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(j-1)} \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} = \begin{Bmatrix} e \\ m \end{Bmatrix} a_l^{(0)} \left(- \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right)^j \\
&\qquad = i^{l-1} \frac{2l+1}{l(l+1)} n \left[\left(\frac{-2i}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right) \left(- \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right)^j \right],
\end{aligned}
\right\} j = 1, 2, \cdots \\
&{}^{e}\mathcal{W}_l^{\pm} = \xi_l^{\pm}(x_o) \xi_l^{\pm\,\prime}(nx_o) - n \xi_l^{\pm\,\prime}(x_o) \xi_l^{\pm}(nx_o), \\
&{}^{m}\mathcal{W}_l^{\pm} = n \xi_l^{\pm}(x_o) \xi_l^{\pm\,\prime}(nx_o) - \xi_l^{\pm\,\prime}(x_o) \xi_l^{\pm}(nx_o), \\
&{}^{e}\mathcal{W}_l = \xi_l^{+}(x_o) \xi_l^{-\,\prime}(nx_o) - n \xi_l^{+\,\prime}(x_o) \xi_l^{-}(nx_o), \\
&{}^{m}\mathcal{W}_l = n \xi_l^{+}(x_o) \xi_l^{-\,\prime}(nx_o) - \xi_l^{+\,\prime}(x_o) \xi_l^{-}(nx_o), \\
&l = 1, 2, \cdots
\end{aligned}
\right\} \tag{3.5-11}
$$

Comparing the form of the scattering coefficients $b_l^{(j)}$ with the form for the amplitudes $A^{(j)}$ from geometric optics given in Eq. (3.4-4) for the jth-degree scattered wave, we see the correspondence

$$
\left.
\begin{aligned}
&-\frac{\mathcal{W}_l^+}{\mathcal{W}_l} \frac{\xi_l^-(nx_o)}{\xi_l^+(nx_o)} \Leftrightarrow R[\varphi', n^{-1}] \\
&-\frac{2i}{\mathcal{W}_l} \frac{\xi_l^+(nx_o)}{\xi_l^+(x_o)} \Leftrightarrow n^{-1} T[\varphi', n^{-1}]
\end{aligned}
\right\} \tag{3.5-12a}
$$

Here we have the relation

$$R\left[\varphi', n^{-1}\right] + 1 = n^{-1} T\left[\varphi', n^{-1}\right] \tag{3.5-12b}$$

which, using Eqs. (3.5-11) and (3.5-6), is readily shown to hold.

3.5.3 Aggregate Scattering

The total field (scattered plus incident) outside of the sphere is given by

$$\begin{Bmatrix} e \\ m \end{Bmatrix} \Pi_{\text{Total}} = \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(i)} + \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(s)} = \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(i)} + \sum_{j=0}^{\infty} \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(s_j)}, \quad r \geq r_o \tag{3.5-13}$$

where the coefficients $b_l^{(j)}$, $j = 0,1,2,\cdots$, are obtained from Eq. (3.5-11), and $\Pi^{(i)}$ and $\Pi^{(s_j)}$ are given by Eqs. (3.2-6) and (3.3-2), respectively. Carrying out the summation in Eq. (3.5-13) and using the form for the scattering coefficients given in Eq. (3.5-11), one obtains for the aggregate scattering

$$\begin{aligned} \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(s)} = \sum_{j=0}^{\infty} \begin{Bmatrix} e \\ m \end{Bmatrix} \Pi^{(s_j)} &= -\frac{E_o}{rk^2} \begin{Bmatrix} \cos\phi \\ \sin\phi \end{Bmatrix} \\ &\times \sum_{l=1}^{\infty} \begin{Bmatrix} e \\ m \end{Bmatrix} S_l i^{l-1} \frac{2l+1}{l(l+1)} \xi_l^+(kr) P_l^1(\cos\theta), \quad r \geq r_o \end{aligned} \tag{3.5-14}$$

where the aggregate scattering coefficient S_l is

$$\begin{aligned} \begin{Bmatrix} e \\ m \end{Bmatrix} S_l &= \left(i^{l-1} \frac{2l+1}{l(l+1)} \right)^{-1} \sum_{j=0}^{\infty} \begin{Bmatrix} e \\ m \end{Bmatrix} b_l^{(j)} \\ &= -\left[\frac{2n}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l + \begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^+} + \frac{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l + \begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l^-}{2} \right] \left(\frac{1}{\begin{Bmatrix} e \\ m \end{Bmatrix} \mathcal{W}_l} \right) \end{aligned} \tag{3.5-15a}$$

This form includes the "–1/2" contribution from the $b_l^{(0)}$ term, which completely cancels the $\xi_l^+(kr)$ component of the scalar potential series [Eq. (3.2-6)]) for the incident plane wave, leaving only the $\xi_l^-(kr)$ component in that series. Equation (3.5-15a) can be simplified somewhat by using the Wronskian forms in Eqs. (3.5-6) and (3.5-11). It becomes

$$2^{\left\{\begin{smallmatrix}\mathrm{e}\\ \mathrm{m}\end{smallmatrix}\right\}}S_l + 1 = -\frac{{}^{\left\{\begin{smallmatrix}\mathrm{e}\\ \mathrm{m}\end{smallmatrix}\right\}}\tilde{\mathcal{W}}_l + {}^{\left\{\begin{smallmatrix}\mathrm{e}\\ \mathrm{m}\end{smallmatrix}\right\}}\mathcal{W}_l^-}{{}^{\left\{\begin{smallmatrix}\mathrm{e}\\ \mathrm{m}\end{smallmatrix}\right\}}\mathcal{W}_l + {}^{\left\{\begin{smallmatrix}\mathrm{e}\\ \mathrm{m}\end{smallmatrix}\right\}}\mathcal{W}_l^+} \tag{3.5-15b}$$

where $\tilde{\mathcal{W}}_l$ is given by

$$\left.\begin{aligned} {}^{\mathrm{e}}\tilde{\mathcal{W}}_l &= \xi_l^-(x_o)\xi_l^{+\prime}(nx_o) - n\xi_l^{-\prime}(x_o)\xi_l^+(nx_o) \\ {}^{\mathrm{m}}\tilde{\mathcal{W}}_l &= n\xi_l^-(x_o)\xi_l^{+\prime}(nx_o) - \xi_l^{-\prime}(x_o)\xi_l^+(nx_o) \end{aligned}\right\} \tag{3.5-16}$$

$\tilde{\mathcal{W}}_l$ is the complex conjugate of $\mathcal{W}_l$ when x and l are real.

As a credibility check, first let us determine the limiting form of $\Pi^{(s)}$ as $nx_o \to x_o$. In this case, $\mathcal{W}_l^{\pm} \to 0$ and $\mathcal{W}_l \to -2i$. Therefore, $S_l \to 0$ and $\Pi^{(s)} \to 0$. What about the individual scattering coefficients? The non-zero coefficients have the limiting forms $b_l^{(0)} \to -(1/2)i^{l-1}(2l+1)/l(l+1)$, $a_l^{(0)} \to i^{l-1}(2l+1)/l(l+1)$, and $b_l^{(1)} \to (1/2)i^{l-1}(2l+1)/l(l+1)$. Thus, $b_l^{(0)}$ and $b_l^{(1)}$ cancel each other and $a_l^{(0)}$ assumes the form of the coefficient for the incident plane wave, as expected. A second check is for $n \to \infty$. Here ${}^{\mathrm{e}}b_l^{(0)} \to -[\psi_l'(x_o)/\xi_l^{+\prime}(x_o)]i^{l-1}(2l+1)/l(l+1)$, ${}^{\mathrm{m}}b_l^{(0)} \to -[\psi_l(x_o)/\xi_l^{+}(x_o)]i^{l-1}(2l+1)/l(l+1)$, and all other coefficients approach zero, which is as expected for an in-plane polarized wave scattered from a perfectly reflecting sphere [8]. If one introduces absorption into the sphere but with a finite refractivity, then n would become complex and $\xi_l^{\pm}(nx_o)$ must be expressed in terms of modified Bessel and Neumann functions. It will be shown in this case that the enormity of x_o forces all coefficients to zero except $b_l^{(0)}$ when only a minuscule absorption coefficient is introduced.

Because the differences in the electric and magnetic scattering coefficients are small when $N << 1$, and to simplify the notation, we will suppress the superscripts "e" and "m" on the scattering coefficients; we retain only the electric coefficients in the following sections. The magnetic coefficients can easily be reconstituted. GPS signals are principally right-hand circular polarized; therefore, to study polarization effects from the refracting sphere, we also would need the scattering terms for cross-plane polarization. To obtain the scattered wave from a cross-plane polarized incident wave, we use the $\boldsymbol{H}$ field from the in-plane case and the symmetry in Maxwell's equations mentioned earlier, in which the definitions of the field and its medium parameters are exchanged: $(\boldsymbol{E}, \boldsymbol{H}, \varepsilon, \mu) \Leftrightarrow (-\boldsymbol{H}, \boldsymbol{E}, \mu, \varepsilon)$.

3.6 The Problem of Slow Convergence

The series in Eqs. (3.2-4), (3.3-2), and (3.5-14) converge rapidly only when kr is small or moderate; but in our problem, $kr_o \sim 10^8$, so in this case another technique must be used to evaluate these series rather than direct summation. When $l << x$, it is easily shown for large x that [10]

$$\psi_l(x) \to \sin(x - l\pi/2), \quad l << x \tag{3.6-1}$$

This is easily seen from the differential equation for $\psi_l(x) = (\pi x/2)^{1/2} J_{l+1/2}(x)$, which is given from Bessel's equation by

$$\psi_l'' + \left(1 - \frac{l(l+1)}{x^2}\right)\psi_l = 0 \tag{3.6-2}$$

Because $x >> 1$, we can linearize Eq. (3.6-2) by writing $x = kr_o + y$, where y is to be considered as a relatively small excursion of x from its nominal value of $x_o = kr_o$. Equation (3.6-2) becomes

$$\left.\begin{aligned} \psi_l'' + (K_l + M_l y)\psi_l &= 0 \\ K_l &= 1 - \frac{l(l+1)}{x_o^2} \\ M_l &= \frac{2l(l+1)}{x_o^3} \end{aligned}\right\} \tag{3.6-3}$$

For $l << x_o$, $\psi_l(y)$ behaves sinusoidally with y. For $l >> x_o$, $\psi_l(y)$ behaves exponentially. For $l = x_o$, $K_l = 0$, and the solutions to Eq. (3.6-2) are the Airy functions of the first and second kind. In the neighborhood where $K_l + M_l y \approx 0$, $\psi_l(y)$ undergoes a transition from a sinusoidal behavior to an exponential one. The Wentzel–Kramer–Brillouin (WKB) method [11] provides a useful technique for evaluating the asymptotic solutions to this equation.

It follows that for $l << x_o$ the coefficients in the series for the incident and scattered waves in Eqs. (3.2-6) and (3.3-2) chatter wildly throughout this regime. Using the asymptotic forms for the Hankel functions, $\xi_l^{\pm}(x) \sim (\mp i)^{l+1} \exp(\pm ix)$, which apply when $l << x_o = kr_o >> 1$, we find that the Wronskian-like terms $\mathcal{W}_l^{\pm}$ and $\mathcal{W}_l$ behave as $\mathcal{W}_l^{\pm} \to 0$ and $\mathcal{W}_l \to -2i\exp(-iNx_o)$. Therefore, the aggregate scattering coefficient S_l becomes

$$2S_l \to \exp(2iNx_o) - 1, \quad l << x_o = kr_o >> 1 \tag{3.6-4}$$

Let the partial sum for a scalar potential series, Π_{l^*}, be given by

$$\Pi_{l^*} = \sum_{l=1}^{l^*} f(x,\theta,\phi,l) \tag{3.6-5}$$

where f stands for the summand in any one of the series representations for the incident or scattered waves. It follows that the scalar potential for the spectral components of the scattered wave in the regime where $l << x_o$ asymptotically approaches

$$\Pi_{l^*}^{(s)} \sim \frac{1}{2}\Pi_{l^*}^{(i)}\left(\exp\left(2iNx_o\right) - 1\right),\ l^* << x_o \tag{3.6-6}$$

Without much restriction (because x_o is so large), we may choose N so that Nx_o equals an integer multiple of π, thereby nulling out completely the contribution from this asymptotic component.

When $l >> x$, it can be shown (see, for example, [10]) that

$$\xi_l^{\pm}(x) \rightarrow \sqrt{\frac{x}{l}}\left[\frac{1}{2}\left(\frac{ex}{2l}\right)^l \mp i\left(\frac{2l}{ex}\right)^l\right],\ l >> x \tag{3.6-7}$$

Thus, the contributions from the scattering coefficients to the series in Eq. (3.3-2) attenuate rapidly with increasing l when $l > x_o$. We will show that the principal contributions to these series come from that regime where $l \approx x_o$.

3.7 The Sommerfeld–Watson Transformation

A conventional method for solving such slowly converging series is to use a transformation due principally to Arnold Sommerfeld [3,12]. This transformation converts a sum, such as that in Eq. (3.5-14), into an equivalent but perhaps a more rapidly converging sum in the complex plane. Here the sum is first converted into a closed-contour integral representation in the complex l-space plane. An example of this is shown in Fig. 3-9(a); the closed contour is the path L_1, which lies near the positive real axis. When this contour integration is evaluated in terms of the sum of the residues around the poles of the integrand on the real axis, it yields the original sum to be evaluated (assuming that the *f* function in Fig. 3-9(a) has no poles of its own on the real axis).

This contour L_1 then is deformed into another closed contour L_2 that excludes the poles on the real axis as interior points but encloses the poles of the integrand located elsewhere in the complex plane; in Fig. 3-9(b), these poles are shown at $l_0, l_1, l_2, \cdots$. If the integrand is properly constituted, i.e., symmetric

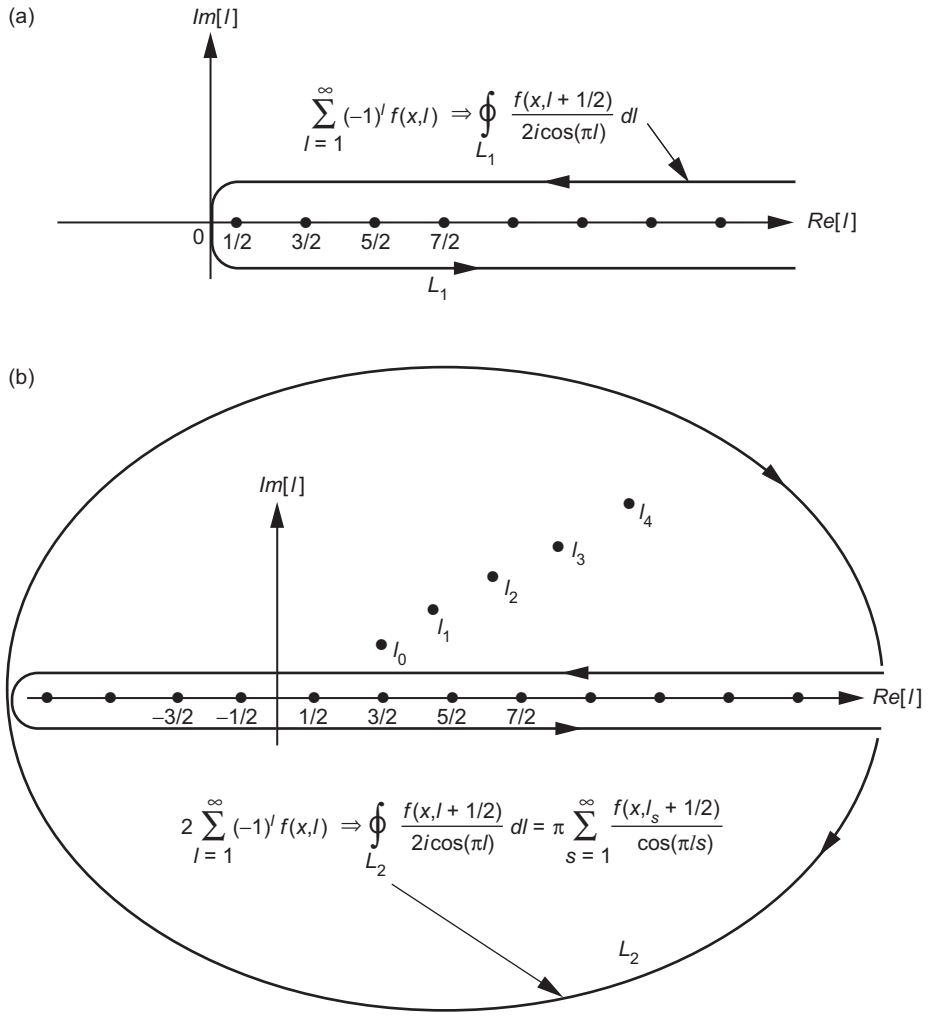

Fig. 3-9. Contour integration in the complex spectral number plane to evaluate the spectral series by summing the residues of the poles in the plane: (a) residues of the simple poles interior to the path L_1 at the 1/2-integer points along the real axis are summed, and (b) equivalent contour integration around L_2, which is evaluated by summing the residues at the interior poles at l_0, l_1, l_2, ··· in the upper complex plane.

about the origin on the real axis (for this example), then the residue series from its off-real-axis poles in the complex plane may converge rapidly under certain conditions. For example, if the residues at the poles in the complex plane diminish rapidly with altitude or the number of poles is limited, then convergence by summing in the complex spectral number plane may be more efficient than summing over integer spectral number on the positive real axis. Thus, one replaces a spectral distribution of partial wavelets in real-integer l

spectral number space by an equivalent sum of residue wavelets in complex l-space. This technique has been very useful in radio propagation problems [13] because the residues from the poles of the integrand located in the complex plane away from the real axis diminish rapidly for certain geometrical configurations of the emitter and receiver. The Hankel functions of the first kind, appearing in the series expressions for the scattered wave [see Eq. (3.3-2)], vanish exponentially with increasing positive altitude in complex l-space above the real axis. This approach has been used in the study of scattering and rainbow effects from spherical objects and also in resonance studies in high-energy physics (Regge poles). Appendix G provides a brief discussion of this technique.

3.8 Evaluating Scattering Coefficients with Asymptotic Expansions

In our problem, the situation is simpler because of the small discontinuity in refractivity that is assumed here. Moreover, the enormity of x_o allows us to use asymptotic expressions for the Bessel and Neumann functions (due to Hankel) and for the Legendre polynomials, which are relatively simple, becoming more accurate with increasing x_o. We will use a numerical integration technique aided by the stationary-phase technique to evaluate the scattering coefficients. In addition, the method of stationary phase will be used for interpretation. Use of stationary-phase concepts is instructive because the asymptotic forms for the Hankel functions and Legendre polynomials explicitly reveal the correspondence between the Mie solution to scattering from a sphere and the scalar diffraction integral formulation of the thin-screen model. They also reveal the correspondence between wave theory and geometric optics.

For large values of $\nu = l + 1/2$, the Hankel functions are given by asymptotic expansions in terms of the Airy functions and their derivatives [10]. Airy functions themselves are expressible in terms of Bessel functions of fractional order 1/3. Inasmuch as the significant scattering coefficients arise for spectral numbers only in the neighborhood near $\nu \approx x_o$ and because $x_o >> 1$, we can truncate the asymptotic expansions for the Hankel functions, retaining only the leading term in each expression. These are given by

$$H_\nu^{\pm}(x) = \left(\frac{4y}{\nu^2 - x^2}\right)^{1/4} \left(\mathrm{Ai}[y] \mp i\,\mathrm{Bi}[y]\right) + \mathrm{O}\left[\nu^{-5/3}\right] \tag{3.8-1}$$

The derivatives of the Hankel functions with respect to x are given by

$$H_\nu'^{\pm}(x) = -\frac{2}{x}\left(\frac{\nu^2 - x^2}{4y}\right)^{1/4} \left(\mathrm{Ai}'[y] \mp i\,\mathrm{B}i'[y]\right) + \mathrm{O}\left[\nu^{-4/3}\right] \tag{3.8-2}$$

Here $\mathrm{Ai}[y]$ and $\mathrm{Bi}[y]$ are the Airy functions of the first and second kind, and their argument y is given by

$$y = \nu^{2/3}\zeta\left[\frac{\nu}{x}\right] \tag{3.8-3}$$

The function $\zeta[\mu]$, which is defined to be real when $\mu = \nu / x$ is real and positive, is given by

$$\left.\begin{aligned} &\frac{2}{3}\left(\zeta[\mu]\right)^{3/2} = \int_1^{\mu} \frac{\sqrt{w^2-1}}{w^2}\,dw = \log\left(\mu + \sqrt{\mu^2-1}\right) - \frac{\sqrt{\mu^2-1}}{\mu}, \quad \mu \geq 1 \\ &\text{or} \\ &\frac{2}{3}\left(-\zeta[\mu]\right)^{3/2} = \int_{\mu}^{1} \frac{\sqrt{1-w^2}}{w^2}\,dw = \frac{\sqrt{1-\mu^2}}{\mu} - \cos^{-1}\mu, \quad \mu \leq 1 \end{aligned}\right\} \tag{3.8-4}$$

Useful series expansions for $\zeta[\mu]$ and y are given by

$$\left.\begin{aligned} \zeta &= \frac{\mu^2-1}{\mu^2}\left(\frac{3\mu^2}{2}\sum_{k=0}^{\infty}\frac{\left(1-\mu^{-2}\right)^k}{2k+3}\right)^{2/3} \quad |\mu|^2 \geq \frac{1}{2} \\ y &= \frac{1}{4}\left(\frac{2}{\nu}\right)^{4/3}\left(\nu^2 - x^2\right)\left(1 + \frac{2}{5}\cdot\frac{\nu^2-x^2}{\nu^2} + \cdots\right) \end{aligned}\right\} \tag{3.8-5}$$

When $x >> 1$, a near-linear relationship exists between y and $(\nu - x)$. By expanding Eq. (3.8-4), one can show that

$$\left.\begin{aligned} y &= \left(\frac{2}{x}\right)^{1/3}(\nu - x)\left(1 - \frac{\nu - x}{30x} + \cdots\right), \\ \nu - x &= \left(\frac{x}{2}\right)^{1/3} y\left(1 + \left(\frac{x}{2}\right)^{1/3}\frac{y}{30x} + \cdots\right) \end{aligned}\right\} \tag{3.8-6}$$

These truncated series for y and ν are very accurate when $x >> 1$ and ν is in the vicinity of x.

The Airy function $\mathrm{Ai}[y]$ is oscillatory for negative real y and rapidly damps to zero for positive real y [see Fig. 3-10(a)]. $\mathrm{Bi}[y]$ is also oscillatory for negative y, but it grows rapidly with positive y [see Fig. 3-10(b)]. The asymptotic forms for the Airy functions [10] applicable also for complex arguments z are given by

$$\left.\begin{aligned}
&\mathrm{Ai}[-z] \to \pi^{-1/2} z^{-1/4}\left[\sin\mathrm{X} - \frac{5}{48} z^{-3/2}\cos\mathrm{X} + \mathrm{O}\left[z^{-3}\right]\right], \ |\mathrm{Arg}[z]| < \frac{2\pi}{3} \\
&\mathrm{Ai}'[-z] \to \pi^{-1/2} z^{1/4}\left[-\cos\mathrm{X} + \frac{7}{48} z^{-3/2}\sin\mathrm{X} + \mathrm{O}\left[z^{-3}\right]\right], \ |\mathrm{Arg}[z]| < \frac{2\pi}{3} \\
&\mathrm{Bi}[-z] \to \pi^{-1/2} z^{-1/4}\left[\cos\mathrm{X} + \frac{5}{48} z^{-3/2}\sin\mathrm{X} + \mathrm{O}\left[z^{-3}\right]\right], \ |\mathrm{Arg}[z]| < \frac{2\pi}{3} \\
&\mathrm{Bi}'[-z] \to \pi^{-1/2} z^{1/4}\left[\sin\mathrm{X} + \frac{7}{48} z^{-3/2}\cos\mathrm{X} + \mathrm{O}\left[z^{-3}\right]\right], \ |\mathrm{Arg}[z]| < \frac{2\pi}{3}
\end{aligned}\right\} \tag{3.8-7}$$

where $\mathrm{X} = 2z^{3/2}/3 + \pi/4$. Also,

$$\left.\begin{aligned}
&\mathrm{Ai}[z] \to \frac{\pi^{-1/2}}{2} z^{-1/4}\left[\exp\left[-\frac{2}{3} z^{3/2}\right]\left(1 - \frac{5}{48} z^{-3/2} + \mathrm{O}\left[z^{-3}\right]\right)\right], \ |\mathrm{Arg}[z]| < \pi \\
&\mathrm{Ai}'[z] \to \frac{-\pi^{-1/2}}{2} z^{1/4}\left[\exp\left[-\frac{2}{3} z^{3/2}\right]\left(1 + \frac{7}{48} z^{-3/2} + \mathrm{O}\left[z^{-3}\right]\right)\right], \ |\mathrm{Arg}[z]| < \pi \\
&\mathrm{Bi}[z] \to \pi^{-1/2} z^{-1/4}\left[\exp\left[\frac{2}{3} z^{3/2}\right]\left(1 + \frac{5}{48} z^{-3/2} + \mathrm{O}\left[z^{-3}\right]\right)\right], \ |\mathrm{Arg}[z]| < \frac{\pi}{3} \\
&\mathrm{Bi}'[z] \to \pi^{-1/2} z^{1/4}\left[\exp\left[\frac{2}{3} z^{3/2}\right]\left(1 - \frac{7}{48} z^{-3/2} + \mathrm{O}\left[z^{-3}\right]\right)\right], \ |\mathrm{Arg}[z]| < \frac{\pi}{3}
\end{aligned}\right\} \tag{3.8-8}$$

Equation (3.8-7) applies in particular to negative real values of y, the argument of the Airy functions. Equation (3.8-8) also applies to positive real values of y. For $\nu \approx x_o$, Eq. (3.8-6) shows that $y \doteq (\nu - x_o)(2/x_o)^{1/3}$, which is very accurate for $x_o >> 1$. The Airy functions rapidly assume their respective asymptotic forms on the real axis when $|y| > 2$. At $|y| = 2$, the relative accuracy of these asymptotic forms is about 1 percent; at $|y| = 3$, it is 0.2 percent.

From Eqs. (3.2-8), (3.8-1), (3.8-2), and (3.8-5), it follows that when $l \approx x$ the spherical Hankel functions may be written in terms of the Airy functions in the form

$$\left.\begin{aligned}
\xi_l^{\pm}(x) &\sim \sqrt{\pi K_x}\left(1-\frac{y}{15K_x^2}+\mathrm{O}\left[x^{-5/3}\right]\right)\left(\mathrm{Ai}[y]\mp i\,\mathrm{Bi}[y]\right)\\
\xi_l^{\pm\prime}(x) &\sim -\sqrt{\frac{\pi}{K_x}}\left(1+\frac{y}{15K_x^2}+\mathrm{O}\left[x^{-5/3}\right]\right)\left(\mathrm{Ai}'[y]\mp i\,\mathrm{Bi}'[y]\right)\\
K_x &= \left(\frac{x}{2}\right)^{1/3},\ \nu = l+1/2
\end{aligned}\right\} \qquad (3.8\text{-}9)$$

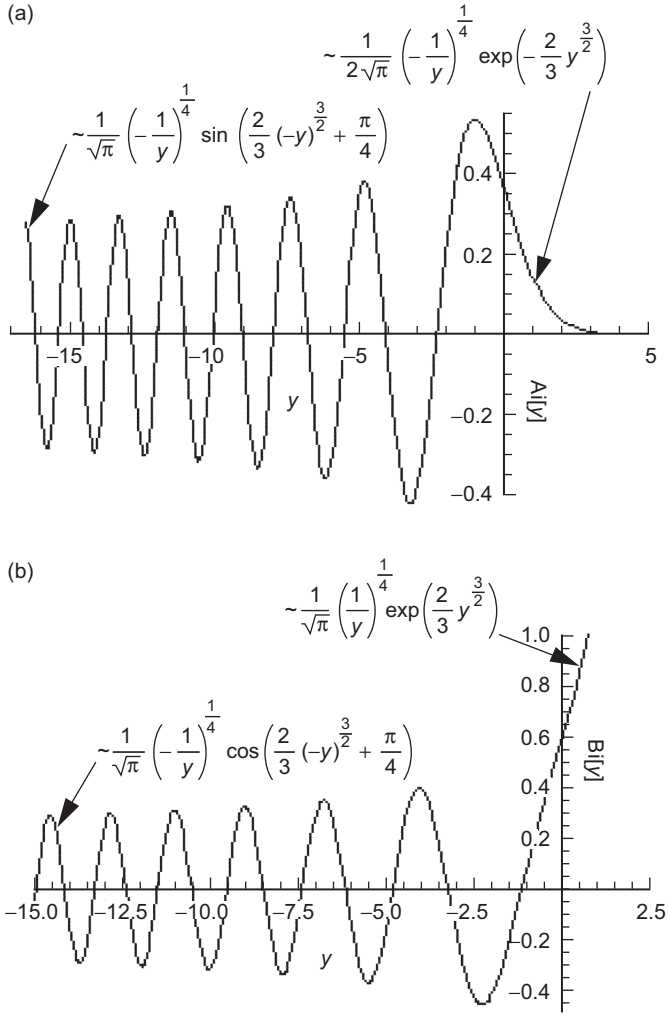

Fig. 3-10. Airy functions and asymptotic forms: (a) Airy function of the first kind and (b) Airy function of the second kind.

The quantity K_x appears often throughout the wave theory parts of this monograph. When x is expressed in phase units, then K_x provides the scale factor between y-space and spectral number or impact parameter space. It is a large quantity for our problem ($K_x \sim 500$). In most scattering problems, the principal contributions to the synthesis of the wave come from the spectral neighborhood where $l \approx x$. Therefore, K_x is essentially constant over that neighborhood. When $|l - x|/x << 1$, we may drop the y/K_x^2 terms in Eq. (3.8-9).

We will be interested in the asymptotic forms of the scattering coefficient quantities $\mathcal{W}_l^{\pm}$ and $\mathcal{W}_l$, particularly for negative y values. From Eq. (3.8-7), it follows that the negative-argument asymptotic forms for the spherical Hankel functions become

$$\left.\begin{aligned} &\left.\begin{aligned} &\xi_l^{\pm}(x) \to \mp i\left(\frac{x^2}{x^2-\nu^2}\right)^{1/4} \exp(\pm i\mathrm{X}), \\ &\xi_l^{\pm'}(x) \to \left(\frac{x^2-\nu^2}{x^2}\right)^{1/4} \exp(\pm i\mathrm{X}) \end{aligned}\right\}, \quad \nu << x \\ &\mathrm{X} = \sqrt{x^2-\nu^2} - \nu\cos^{-1}\left(\frac{\nu}{x}\right) + \frac{\pi}{4}, \quad \nu = l + 1/2 \end{aligned}\right\} \tag{3.8-10}$$

or in terms of the argument y (provided that $l \approx x$ so that $y = K_x^{-1}(\nu - x)$ from Eq. (3.6-6) is a valid approximation) they become

$$\left.\begin{aligned} &\left.\begin{aligned} &\xi_l^{\pm}(x) \to \mp i\left(\frac{K_x^2}{-y}\right)^{\frac{1}{4}} \exp(\pm i\mathrm{X}), \\ &\xi_l^{\pm'}(x) \to \left(\frac{-y}{K_x^2}\right)^{\frac{1}{4}} \exp(\pm i\mathrm{X}) \end{aligned}\right\}, \quad y < 0 \\ &\mathrm{X} = \frac{2}{3}(-y)^{3/2} + \frac{\pi}{4}, \quad y = \nu^{2/3}\zeta[\nu/x] \end{aligned}\right\} \tag{3.8-11}$$

For the scattering coefficients, it follows that

$$\left.\begin{aligned}
&\frac{\mathcal{W}_l^{\pm}}{\mathcal{W}_l} \to \left\{ \begin{aligned} &\pm \frac{n^2(x_o^2-\nu^2)^{1/2}-(n^2x_o^2-\nu^2)^{1/2}}{n^2(x_o^2-\nu^2)^{1/2}+(n^2x_o^2-\nu^2)^{1/2}} e^{2i\left\{{+\hat{\mathrm{X}} \atop -\mathrm{X}}\right\}}, \ \nu < x_o \\ &-1, \ \ \nu > x_o \end{aligned} \right\} \\
&\text{or} \\
&\frac{\mathcal{W}_l^{\pm}}{\mathcal{W}_l} \to \pm \frac{(-\hat{y})^{1/2}-(-y)^{1/2}}{(-\hat{y})^{1/2}+(-y)^{1/2}} e^{2i\left\{{+\hat{\mathrm{X}} \atop -\mathrm{X}}\right\}}, \ \ y < 0
\end{aligned}\right\} \tag{3.8-12}$$

Also,

$$\left.\begin{aligned}
&\mathcal{W}_l \to \left\{ \begin{aligned} &-i\left[\left(\frac{n^2\left(x_o^2-\nu^2\right)}{n^2x_o^2-\nu^2}\right)^{1/4}+\left(\frac{n^2\left(x_o^2-\nu^2\right)}{n^2x_o^2-\nu^2}\right)^{-(1/4)}\right] e^{i(\mathrm{X}-\hat{\mathrm{X}})}, \nu < x_o \\ &\infty, \ \nu > x_o, \end{aligned} \right\} \\
&\text{or} \\
&\mathcal{W}_l \to -i\left[\left(\frac{\hat{y}}{y}\right)^{\frac{1}{4}}+\left(\frac{\hat{y}}{y}\right)^{-(1/4)}\right] e^{i(\mathrm{X}-\hat{\mathrm{X}})}, \ \ y < 0
\end{aligned}\right\} \tag{3.8-13}$$

where

$$\left.\begin{aligned}
&y = \nu^{2/3}\zeta\left(\frac{\nu}{x}\right), \quad \hat{y} = \nu^{2/3}\zeta\left(\frac{\nu}{nx}\right) = y - 2NK_{x_o}^2(1+\mathrm{O}[x_o^{-1}]), \\
&\mathrm{X} = \frac{2}{3}(-y)^{3/2}+\frac{\pi}{4}, \ \ \hat{\mathrm{X}} = \frac{2}{3}(-\hat{y})^{3/2}+\frac{\pi}{4}, \quad y, \hat{y} < 0
\end{aligned}\right\} \tag{3.8-14}$$

To estimate the number of scattering terms required for $\nu > x_o$ in the spectral summations for the field, we note from Fig. 3-10(a) that $\mathrm{Ai}[y]$ damps to zero rapidly for $\nu > x_o$ and that $\mathrm{Bi}[y]$ grows large ($\sim \exp\left(2y^{3/2}/3\right)$). We note from Eqs. (3.5-11), (3.5-15), and (3.8-13) that the rapid growth in $\mathrm{Bi}[y]$ for $\nu > x_o$ drives $|S_l|$ to zero. In fact, for $\nu > x_o + 3(x_o/2)^{1/3}$, $|S_l|$ becomes less than 0.001. Therefore, the number of spectral terms $\Delta\nu$ greater than x_o that are required to reach $|S_l| = 0.001$ is about 1500 in our case ($x_o \approx 2\times 10^8$) in order to achieve better than 0.1 percent accuracy. That the scattering coefficients rapidly damp to zero for $l > x_o$ was first established 90 years ago [2]. A physical rationale of this damping in terms of spectral caustics in geometric optics is given in Section 3.11.

For $\nu < x_o$, the asymptotic forms for the Airy functions rapidly become sinusoidal [see Eq. (3.8-7)], which by choosing N so that $\sin(2Nx_o) = 0$, we know from Eq. (3.6-4) assures a zero contribution for $\nu << x_o$. However, the slight difference in arguments in the spherical Hankel functions in the $\mathcal{W}_l$ functions in Eq. (3.5-11), which arise because of the refractivity N, creates a variable phase difference in these scattering functions that only slowly damps out for $\nu < x_o$ (see Fig. 3-13). Much of the following discussion is concerned with the contribution of this phase difference to the computation of scattering effects. The stationary-phase technique will be used to isolate the appropriate neighborhoods in spectral number that significantly contribute to the spectral summations for the field.

3.9 Expressing Scattering Coefficients in Terms of Phasors

From Eqs. (3.4-3) through (3.4-7), (3.5-9), (3.5-11), (3.5-12), and (3.5-15a), or directly from Eq. (3.5-15b), it can be shown that

$$\left|2S_l + 1\right| = \left|\left(\frac{4n}{\mathcal{W}_l + \mathcal{W}_l^+} + \mathcal{W}_l^-\right)\frac{1}{\mathcal{W}_l}\right| = \left|\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+}\right| = 1 \tag{3.9-1}$$

for all real values of l and x. This is a manifestation of the conservation of energy principle; all incident radiation must be scattered. We know the asymptotic form [see Eq. (3.6-4)] that S_l assumes for $\nu << x_o$. Accordingly, for real values of l and x we define the scattering phase $\Omega^{(s)} = \Omega^{(s)}(N, x_o, \nu)$ in terms of the aggregate scattering coefficient S_l [see Eqs. (3.5-15a), (3.5-15b), and (3.5-16)] by

$$\exp\left(i2\Omega^{(s)}\right) = 2S_l + 1 = -\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+} \tag{3.9-2}$$

where, in this definition, $\Omega^{(s)}(N, x_o, \nu)$ is a real function for all real x and $\nu = l + 1/2$, but one has to account for its phase windup with diminishing ν. For $\nu > x_o$, one concludes, using Eq. (3.8-13) and noting that Bi[y] grows exponentially large with positive y, that $\Omega^{(s)}(N, x_o, \nu) \to 0$ as ν exceeds x_o; so, $S_l \to 0$ rapidly for $\nu > x_o$. We also know from Eq. (3.6-4) that $\Omega^{(s)}(N, x_o, \nu) \to Nx_o$ for $\nu << x_o$. Figure 3-11(a) shows $\Omega^{(s)}(N, x_o, \nu)$ over its entire range. For this example, Nx_o has been set to the value 83π exactly

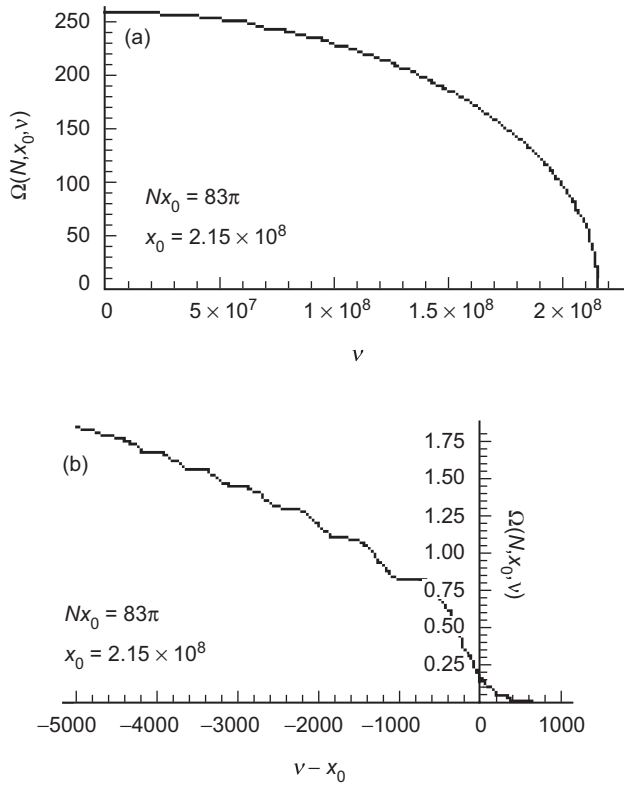

Fig. 3-11. Scattering phase $\Omega^{(S)}(N,x_0,\nu)$ versus spectral number: (a) variation over the entire range $0 \le \nu \le x_0$ and (b) variation near $\nu = x_0$.

($N \approx 1.2 \times 10^{-6}$), so that the scattering coefficients (S_l and also all of the individual coefficients $b_l^{(j)}$) decay to exactly zero with decreasing ν. Figure 3-11(b) displays $\Omega^{(s)}(N, x_o, \nu)$ in the vicinity of x_o and shows the diminishing undulations with decreasing ν below x_o. We will see that these undulations result in additional points of stationary phase in the series summation of the scattering coefficients for a certain range of angular positions of the LEO.

Similarly, we can define $\Omega^{(j)}(N, x_o, \nu)$ as the phase associated with the individual jth scattering coefficient $b_l^{(j)}$, $j = 0,1,2,\cdots$. From Eq. (3.5-11) we have

$$\left.\begin{aligned} &a^{(0)}\left(i^{l-1}n\frac{2l+1}{l(l+1)}\right)^{-1}\overset{\Delta}{=}\rho e^{i\Omega}=-\frac{2i}{\mathcal{W}_l^-},\\ &2S_l^{(0)}+1=2b_l^{(0)}\left(i^{l-1}\frac{2l+1}{l(l+1)}\right)^{-1}+1\overset{\Delta}{=}\rho^{(0)}e^{i2\Omega^{(0)}}=-\frac{\mathcal{W}_l^-}{\mathcal{W}_l^-},\\ &2S_l^{(j)}=2b_l^{(j)}\left(i^{l-1}\frac{2l+1}{l(l+1)}\right)^{-1}\overset{\Delta}{=}\rho^{(j)}e^{i2\Omega^{(j)}}=n\left(-\frac{\mathcal{W}_l^+}{\mathcal{W}_l^-}\right)^{j-1}\rho^2e^{i\Omega}\end{aligned}\right\} \quad (3.9\text{-}3)$$

Here the modulus $\rho=\rho(N,x_o,\nu)$ on the transmission coefficient is variable with spectral number. It follows from Eq. (3.8-13) that

$$\left.\rho=\left|\frac{-2i}{\mathcal{W}_l^-}\right|\to 2\left(\left(\frac{\hat{y}}{y}\right)^{1/4}+\left(\frac{y}{\hat{y}}\right)^{1/4}\right)^{-1}\to 1,\ \Omega\to\hat{\mathrm{X}}-\mathrm{X},\ y<0\right\} \quad (3.9\text{-}4)$$

Thus, $\rho\to 1$ for $\nu << x_o$, and $\rho\to 0$ for $\nu >> x_o$ (see Fig. 3-12). Similarly, for the external reflection it can be shown from Eq. (3.8-12) that

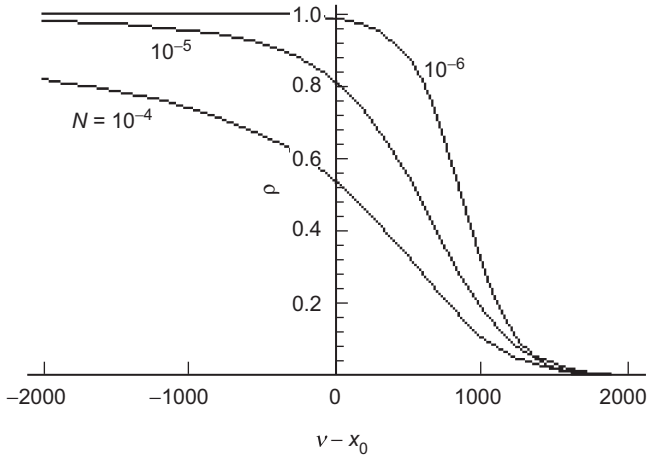

Fig. 3-12. Modulus $\rho\ (N,x_0\nu\ )$ **of** $(2i/W_l)$.

$$\left.\begin{aligned}
&\left\{\begin{aligned}
&2i\Omega^{(0)} = \log\left(\frac{\xi_l^-(x_o)}{\xi_l^+(x_o)}\right) \to 2i\mathrm{X} \\
&\rho^{(0)} = \left|\frac{\mathcal{W}_l^-}{\mathcal{W}_l}\right| \to \frac{n^2\left(x_o^2-\nu^2\right)^{1/2} - \left(n^2x_o^2-\nu^2\right)^{1/2}}{n^2\left(x_o^2-\nu^2\right)^{1/2} + \left(n^2x_o^2-\nu^2\right)^{1/2}}
\end{aligned}\right\}, \quad \nu < x_o \\
&\text{or} \\
&\rho^{(0)} \to \left\{\begin{aligned}
&\frac{\left|(-\hat{y})^{1/2} - (-y)^{1/2}\right|}{(-\hat{y})^{1/2} + (-y)^{1/2}}, \quad y, \hat{y} < 0 \\
&1, \quad y, \hat{y} > 0
\end{aligned}\right\}
\end{aligned}\right\} \tag{3.9-5}$$

The degree $j = 0$ modulus $\rho^{(0)}$ of the ratio $(\mathcal{W}_l^- / \mathcal{W}_l)$ is reminiscent of the Fresnel reflection formulas in Eq. (3.4-1a), except that it varies in ν-space rather than in impact parameter space. It rapidly approaches zero (for $N << 1$) with decreasing ν below x_o, and it rapidly approaches 1 for $\nu > x_o$ (see also Fig. 3-4). Also, $2\Omega^{(0)}$ is simply the phase difference between the incident and external reflected components of spectral number l. These scattering phasors will prove to be relevant in our stationary-phase treatment to follow.

For the aggregate scattering coefficient S_l, or its phase $\Omega^{(s)}$, we have the following asymptotic forms. For $\nu > x_o$, we have already noted that both S_l and $\Omega^{(s)}$ damp to zero rapidly. For $\nu < x_o$, it follows from Eq. (3.8-12) that $\mathcal{W}_l^{\pm} \to 0$ with decreasing ν. Therefore, it follows from Eq. (3.8-10) that $\mathrm{Exp}[i2\Omega^{(s)}] = 2S_l + 1 \to -\tilde{\mathcal{W}}_l / \mathcal{W}_l$ for decreasing $\nu < x_o$. From Eq. (3.8-13), it then follows that $\Omega^{(s)}$ becomes

$$e^{i2\Omega^{(S)}} \to -\frac{\tilde{\mathcal{W}}_l}{\mathcal{W}_l} \to e^{i2(\hat{\mathrm{X}}-\mathrm{X})} = e^{i2\Omega_A}, \; \nu < x_o \tag{3.9-6}$$

For $\nu < x_o$, $\Omega_A(N, x_o, \nu)$, which is the asymptotic form of $\Omega^{(s)}$, is given by

$$\Omega_A = \frac{2\nu}{3}\left(\left(-\xi\left[\frac{\nu}{nx_o}\right]\right)^{3/2} - \left(-\xi\left[\frac{\nu}{x_o}\right]\right)^{3/2}\right) + \mathrm{O}\left[y^{-5/2}\right], \; \nu < x_o \tag{3.9-7}$$

or upon using the form for $\xi[\nu / x_o]$ given in Eq. (3.8-4) and expanding in powers of N, one obtains

$$\Omega_A = \begin{cases} \dfrac{2\sqrt{2}}{3}\left[\left((1+N)x_o - \nu\right)^{3/2} - \left(x_o - \nu\right)^{3/2}\right], \ \nu \approx\leq x_o \\ N\sqrt{x_o^2 - \nu^2} + \nu \cdot \mathrm{O}\left[\left(N\nu / x_o\right)^2\right], \ \nu << x_o \end{cases} \qquad (3.9\text{-}8)$$

The value of $2\Omega_A$ at $\nu = x_o$ should be noted. It is $2x_o(2N)^{3/2}/3$, which is exactly equal to the zero point term in the expression for the phase delay $\varepsilon(h^*)$ in the thin-screen model for the uniform refracting sphere, as given in Chapter 2, Eq. (2.8-3). We will see that $\Omega^{(s)}$, which corresponds to a phase retardation of the νth spectral component of the scattered wave, plays an important role in the correspondence between Mie scattering theory and the thin-screen model. It can be shown using Eqs. (3.8-12), (3.8-13), and (3.9-2) that Ω, $\Omega^{(s)}$, and $\Omega^{(1)}$ converge to the same asymptotic form when $\nu < x_o$, that is, to Ω_A.

Figure 3-13 displays the difference between $\Omega^{(s)}$ and Ω_A. The source of these undulations is due to the ratio $(\mathcal{W}_l^{\pm} / \mathcal{W}_l)$, which Eq. (3.8-12) shows to vanish with decreasing spectral number and with $\nu < x_o$. Equation (3.9-8) shows that Ω_A and, therefore, $\Omega^{(s)}$ approach Nx_o quadratically with decreasing ν, as shown in Fig. 3-11(a).

Setting $Nx_o - \Omega^{(s)} = 0.001$ to achieve a 0.1 percent accuracy level, one obtains from Eq. (3.9-8) as a lower bound $\nu_L = \sqrt{0.002x_o / N} \approx 6 \times 10^6$ (with $N = 1.2 \times 10^{-6}$ and $x_o = 2.1 \times 10^8$). However, the range between the lower bound ν_L and the upper bound $x_o + \sim 1500$ is still formidable for a summation.

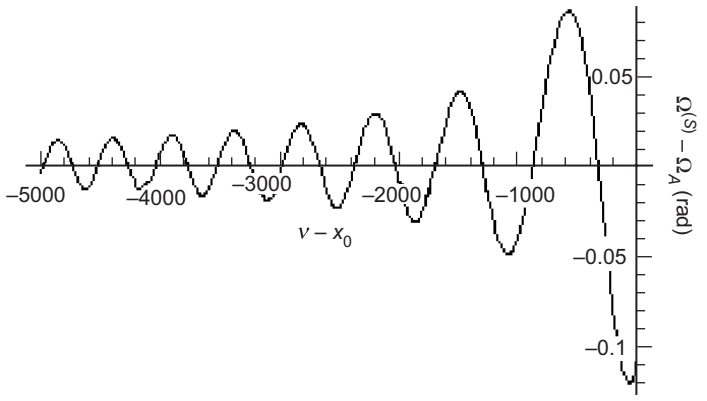

Fig. 3-13. $\Omega^{(S)}(N, x_0, \nu) - \Omega_A(N, x_0, \nu)$ **versus spectral number below** x_0. **The same conditions as used in Fig. 3-11.**

3.10 Asymptotic Forms for the Hankel and Legendre Functions Evaluated at the LEO

To evaluate the spectral summations, we also need the asymptotic forms for the Hankel functions and the Legendre polynomials that are valid for large spectral numbers but evaluated at the coordinates of the LEO. The asymptotic form for the Hankel function of the first kind, $H_\nu^+(x)$, is given in terms of the Airy functions through Eq. (3.8-1); in this case, the radial distance of the LEO is $x \approx 1.1x_o$. Therefore, the argument of the Airy functions for spectral numbers near x_o is given by $Y = \nu^{2/3}\zeta[\nu/1.1x_o] \approx -4\times 10^4$. Therefore, the higher-order terms in the asymptotic expansions for the Airy functions in Eq. (3.8-7) may be ignored. When $0 << \nu << x$, it follows from Eq. (3.8-7) that $\xi_l^\pm(x)$ is given by

$$\left.\begin{aligned} \xi_l^\pm(x) &= \sqrt{\frac{\pi x}{2}}H_\nu^+(x) \sim \mp i\left(\frac{x^2}{x^2-\nu^2}\right)^{1/4}\exp(\pm iY), \\ Y &= \frac{2}{3}\nu\left(-\zeta[\nu/x]\right)^{3/2} + \frac{\pi}{4},\ \ \nu = l+\frac{1}{2} \end{aligned}\right\} \tag{3.10-1}$$

When $x >> \nu$ (as would be the case if one were observing a rainbow), it follows from Eq. (3.8-4) that $Y \to x - \pi\nu/2$. Then, $\xi_l^\pm(x) \to (\mp i)^{l+1}\exp(\pm ix)$. In our problem, however, the geocentric radial distance of the LEO is only about 10 percent larger than the radius of the refracting sphere, that is, $x = kr \doteq 1.1kr_o$.

From Eq. (3.8-4) it follows that Y can be written as

$$Y = \left\{\begin{aligned} &\sqrt{x^2-\nu^2} - \nu\cos^{-1}\left(\frac{\nu}{x}\right) + \frac{\pi}{4},\ \ \nu \le x \\ &\nu\,\mathrm{sech}^{-1}\left(\frac{x}{\nu}\right) - \sqrt{\nu^2-x^2} + \frac{\pi}{4},\ \ \nu \ge x \end{aligned}\right\} \tag{3.10-2}$$

At the radial distance of the LEO, the scattering coefficients essentially will be zero for $\nu \ge x \approx 1.1x_o$, so we need only the expression for Y when $\nu \le x$. Referring to Fig. 3-14, we see from Eq. (3.10-2) that Y may be written as

$$Y = D_\nu - \nu\left(\frac{\pi}{2} - \theta_\nu\right) + \frac{\pi}{4},\quad \theta_\nu = \sin^{-1}\left(\frac{\nu}{x}\right),\ D_\nu = \sqrt{x^2-\nu^2} \tag{3.10-3}$$

where D_ν is the distance (in phase units) of the LEO from the limb of a sphere of radius ν, and θ_ν is the angle that the radius vector of length x from the

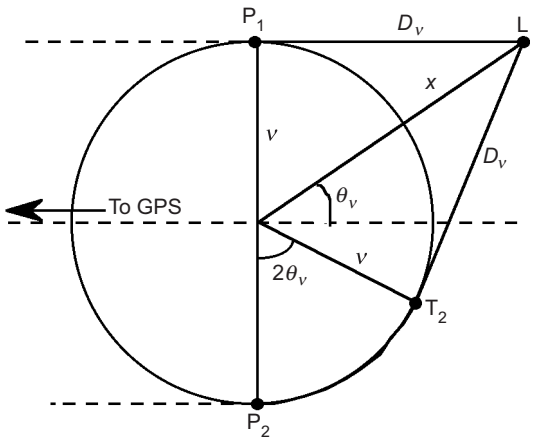

Fig. 3-14. Direct and retrograde wave geometry in spectral number space.

center of this sphere subtends. Since $x > \nu > 0$, it follows that $0 < \theta_\nu < \pi/2$. When $\nu = x_o$, we have from Eq. (3.10-3)

$$\left.\begin{aligned} D_{x_o} &= D = \sqrt{x^2 - x_o^2} = kr\cos\theta_{x_o}, \quad \theta_{x_o} = \theta_o = \sin^{-1}\left(\frac{x_o}{x}\right) \end{aligned}\right\} \tag{3.10-4}$$

The angle θ is the geocentric angle of the LEO position vector relative to the polar axis (the z-axis in Fig. 3-1). When $\theta = \theta_o$, the straight line path from the occulting GPS satellite to the LEO just grazes the surface of the sphere of radius r_o. In other words, the LEO is located on the shadow boundary of the sphere. The quantity D_{x_o} is just the distance in phase units of the LEO from the Earth's limb. Thus, D_{x_o} is the extra geometric phase term relative to the center line at $\theta = \pi/2$ accumulated by the scattered wave upon arrival at the LEO located at (r, θ_o).

The asymptotic form of the Legendre polynomial for large spectral numbers is given by [10]

$$\left.\begin{aligned} &P_l^m(\cos\theta) = \frac{\Gamma[l+1]}{\Gamma[l-m+1]}\sqrt{\frac{2}{\pi l \sin\theta}}\sin\left(\nu\theta + \frac{\pi}{4} - \frac{\pi m}{2}\right)\left(1 + \mathrm{O}\left[\frac{m^2}{l}\right]\right) \\ &m << l, \quad \nu = l + 1/2 \end{aligned}\right\} \tag{3.10-5}$$

Here $\Gamma[z]$ is the Gamma function. Equation (3.10-5) is valid except near $\theta = 0, \pi$. For our problem, the LEO is located at $r = 1.1 r_o$, $\theta \approx 65\,\mathrm{deg}$, and

$\phi = 0$ during the occultation. We assume that the orbit of the LEO is circular and remains essentially invariant during the occultation. Over the few seconds of time over a shadow boundary transition, θ will change by only several milliradians.

3.11 Geometric Optics Interpretation of Mie Scattering Theory

From Eqs. (3.5-14) and (3.9-2), we see that the phase of the lth spectral component of the scalar potential for the scattered wave at the LEO equals the sum of the phase $2\Omega^{(s)}$ of the scattering coefficient S_l plus the phase of the product $i^{l-1}\xi_l^+(x)P_l^1(\cos\theta)$. This product pertains to the geometry of the LEO relative to the scattering sphere. It follows from Section 3.10 that this product $i^{l-1}\xi_l^+(x)P_l^1(\cos\theta)$ contains the phasor $i^{1/2}\exp\left[i\Phi^{\pm}(x, x_o, \nu)\right]$ [see Eq. (3.12-1)], where the phase function $\Phi^{\pm}$ is given by

$$\Phi^{\pm} = D_\nu + \nu(\theta_\nu \pm \theta) + (1 \pm 1)\frac{\pi}{4} \tag{3.11-1}$$

Since ν will be near x_o and θ will be near θ_o, we let $\theta = \theta_\nu + \Delta\theta_\nu$, where $\Delta\theta_\nu$ marks the small departure of the angular position of the LEO from θ_ν during the occultation; thus, $\nu\Delta\theta_\nu$ is the extra phase due to the displacement of θ from θ_ν (see Fig. 3-15). Over the occultation region of interest, the maximum value of the magnitude of $\Delta\theta_\nu$ never exceeds a few milliradians for ν near x_o. Equation (3.11-1) may be written as

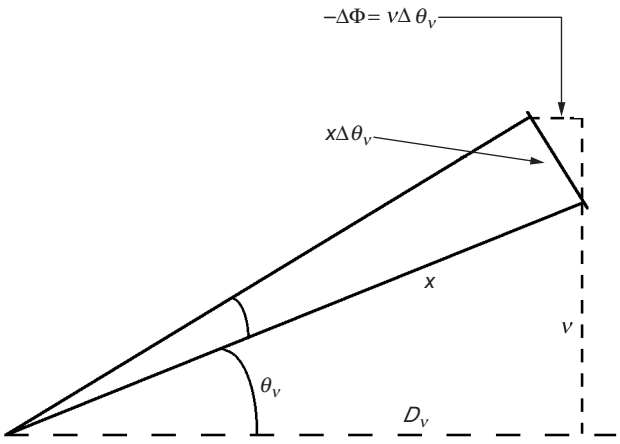

Fig. 3-15. Extra phase delay $\Delta\Phi$ due to $\Delta\theta_\nu$.

$$\left.\begin{aligned}\Phi^{+} &= D_{\nu} + 2\nu\theta_{\nu} + \nu\Delta\theta_{\nu} + \pi/2 \\ \Phi^{-} &= D_{\nu} - \nu\Delta\theta_{\nu},\ \ \Delta\theta_{\nu} = \theta - \theta_{\nu}\end{aligned}\right\} \tag{3.11-2}$$

Referring to Fig. 3-14, we see that Φ^{+} is the phase accumulated by a wave that in a geometric optics context has traveled to the LEO, having departed tangentially from the *far* side of the sphere of radius ν. On the other hand, Φ^{-} is the phase accumulation associated with the wave that arrives from the tangent point on the near side of the sphere. Thus, the sphere of radius ν represents in a geometric optics context a caustic surface for rays originating from the GPS satellite, dividing the region accessible to the rays ($\nu' \geq \nu$) from the region that is inaccessible ($\nu' < \nu$). If $\nu > x_o$, then no partial waves (in a geometric optics sense) with these spectral number values can interact with the surface of a sphere of radius x_o. Depicting the sphere of radius ν as a caustic surface for partial waves of spectral number ν provides a physical explanation for the rapid falloff in the magnitude of the scattering coefficients for $\nu > x_o$.

Equation (3.11-2) shows that Mie scattering theory encompasses the geometric optics concept of two partial waves traveling to the LEO: a direct wave and a retrograde wave. The line P_1L in Fig. 3-14 denotes the direct path. The line and arc labeled P_2T_2L denote the far-side ray path that arrives from the GPS satellite and tangentially impacts the sphere of radius ν at P_2, follows a meridian arc P_2T_2 of the sphere as an evanescent wave, and then departs tangentially from the sphere at T_2 along a straight-line path toward the LEO. The wave also can make multiple revolutions around the sphere and then depart tangentially at T_2 along the straight-line path toward the LEO. Moreover, the direct wave can make multiple revolutions as an evanescent wave and then depart at P_1 for the LEO. In geometric optics, these ray paths follow Fermat's principle.

There is a body of literature [14] that uses Fermat's principle and geometric optics to calculate diffraction effects from three-dimensional objects such as a sphere. A semi-empirical parameter in this approach is the rate of decay of the amplitude of the evanescent wave as it travels circumferentially along the surface of the scattering object, spinning off tangential rays as it progresses. When the decay parameter is properly chosen, close agreement between this geometric optics approach and scattering theory based on electrodynamics usually is attained for a variety of scattering objects in addition to spheres.

3.12 Evaluating Mie Scattering by Integration of the Scattering Phasor

The type of sinusoidal behavior in the scattering coefficients that $\Omega^{(s)}(N, x_o, \nu)$ produces, as well as their abrupt exponential decay for $\nu > x_o$, suggest that only those spectral components of the scattered wave with spectral numbers in the immediate vicinity of $\nu = x_o$ will contribute significantly to the overall summation. Because of the enormity of x_o, we may replace the scattering sum with an integral in ν-space and integrate over a limited range in the vicinity x_o. A numerical integration will be performed to obtain $\boldsymbol{E}^{(s)}$. To achieve this, we will use the asymptotic forms for the scattering coefficients (see the discussions in Sections 3.8 and 3.9). We will replace the required summations for the scattered wave by corresponding integrals that involve as integrands a scattering phasor times a relatively slowly varying function of spectral number. Stationary-phase concepts will be used to aid in the numerical evaluation of the integrals and for interpretation. Also, stationary-phase theory will be used to show the correspondence between Mie scattering results and the thin-screen Fresnel diffraction results.

In this and the following sections, we will focus on scattering from an incident wave with in-plane polarization observed in the same plane. The case of a cross-plane polarized incident wave is readily developed from the symmetry property in electrodynamics: $(\boldsymbol{E}, \boldsymbol{H}, \varepsilon, \mu) \Leftrightarrow (-\boldsymbol{H}, \boldsymbol{E}, \mu, \varepsilon)$.

Assembling all of the parts for the scalar potential of the scattered wave, we have from Eqs. (3.5-14) and (3.9-2)

$$\left\{\begin{matrix}\mathrm{e}\\ \mathrm{m}\end{matrix}\right\}\Pi^{(s)} = \frac{E_o}{xk}\left\{\begin{matrix}\cos\phi\\ \sin\phi\end{matrix}\right\}\sqrt{\frac{i}{2\pi\sin\theta}}\sum_{l=1}^{\infty}\left\{\begin{matrix}\mathrm{e}\\ \mathrm{m}\end{matrix}\right\}S_l\,\frac{2\nu^{3/2}}{\nu^2 - 1/4} \times\left(\frac{x^2}{x^2-\nu^2}\right)^{1/4}\left(\exp\left(i\Phi^{+}\right) - \exp\left(i\Phi^{-}\right)\right) + \mathrm{O}[N]\Bigg\} \qquad (3.12\text{-}1)$$

$$\nu = l + \frac{1}{2}$$

The aggregate scattering terms ${}^{\mathrm{e}}S_l$ and ${}^{\mathrm{m}}S_l$ are defined in Eqs. (3.5-15) and (3.9-2). For small N and grazing conditions, they have essentially the same value for the TM or TE mode. We will suppress the superscripts "e" and "m" and we use the TM form ${}^{\mathrm{e}}S_l$. Also, we can replace S_l in Eq. (3.12-1) with any one of the scattering terms $S_l^{(j)}$ associated with the jth-degree scattered wave that is given in Eq. (3.9-3) to obtain $\Pi^{(S_j)}$, the scalar potential for the jth-degree scattering.

Since the significant terms in the summation in Eq. (3.12-1) come from ν-values near x_o and because $x_o >> 1$, we can replace the sum with an integral. Equation (3.12-1) for the electric potential $\Pi^{(s)}(x,\theta,\phi)$ becomes

$$\Pi^{(s)} = \frac{2E_o}{xk}\sqrt{\frac{i}{2\pi\sin\theta}}\int_0^\infty S_l\left(\nu\cos\theta_\nu\right)^{-(1/2)}\left(e^{i\Phi^+} - e^{i\Phi^-}\right)d\nu\cos\phi \quad (3.12\text{-}2)$$

where $\cos\theta_\nu = (x^2 - \nu^2)^{1/2}/x$. We know that S_l becomes essentially zero for $\nu - nx_o > \sim 3(x_o/2)^{1/3} \approx 1500$ and for $\nu < 0.01x_o$; so, the limits on the integration are rather arbitrary so long as they exceed these bounds. For convenience we set them equal to 0 and ∞.

One can make a similar conversion of the series for the scattered electric field into an integral form. One applies the curl operations given in Eq. (3.2-2), $\nabla\times\nabla\times({}^{\mathrm{e}}\Pi\boldsymbol{r}) = (2 + \boldsymbol{r}\cdot\nabla)\nabla{}^{\mathrm{e}}\Pi - \boldsymbol{r}\nabla^2({}^{\mathrm{e}}\Pi)$ for a TM wave to the scalar potential series in Eq. (3.5-14) for the scattered wave evaluated at the LEO. One obtains for the radial component in the plane $\phi = 0$ and at the point (r,θ)

$$E_r^{(s)} = \frac{E_o}{x^2}\sum_{l=1}^{\infty} S_l i^{l-1}(2l+1)\xi_l^+(x)P_l^1(\cos\theta) + \mathrm{O}[N] \quad (3.12\text{-}3)$$

and for the in-plane transverse component

$$E_\theta^{(s)} = -\frac{E_o}{x}\sum_{l=1}^{\infty} i^{l-1}\frac{2l+1}{l(l+1)} \times\left({}^{\mathrm{e}}S_l\xi_l'^+(x)P_l^{1\,\prime}(\cos\theta)\sin\theta - {}^{m}S_l i\xi_l^+(x)\frac{P_l^1(\cos\theta)}{\sin\theta}\right) + \mathrm{O}[N] \quad (3.12\text{-}4)$$

For Legendre polynomials of high order, it is readily shown that $|P_\nu''| \sim \nu|P_\nu'| \sim \nu^2|P_\nu|$. Since $\nu \sim x_o >> 1$, Eq. (3.12-4) can be simplified by dropping the lower derivative term and, therefore, the contributions from the magnetic terms.

Assembling all of the asymptotic forms for these series and converting from a sum to an integral, we obtain for the radial component $E_r^{(s)}(r,\theta)$ of the scattered wave

$$E_r^{(s)} = E_o\sqrt{\frac{1}{2\pi ix\sin\theta}}\int_0^\infty S_l\left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu}\right)^{1/2} 2\left(e^{i\Phi^+} - e^{i\Phi^-}\right)d\nu \quad (3.12\text{-}5)$$

and for the $\hat{\theta}$ component $E_\theta^{(s)}(r,\theta)$

$$E_\theta^{(s)} = E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \int_0^\infty S_l \left(2\sin 2\theta_\nu\right)^{1/2} \left(e^{i\Phi^+} + e^{i\Phi^-}\right) d\nu \qquad (3.12\text{-}6)$$

Equations (3.12-5) and (3.12-6) provide the basic spectral representation of the scattered wave. We use these forms extensively in the subsequent discussions. Since the significant scattering contributions to the integrals in Eqs. (3.12-5) and (3.12-6) come mostly from where $\nu \approx x_o$ and $\theta_\nu \approx \theta_o$, it follows to a first approximation that $E_\theta^{(s)} \approx E_r^{(s)} \cot\theta_o$, which corresponds to a plane wave directed along the z-axis at the LEO position of (r,θ). It is the small deviations in which we are interested. We can replace S_l in these integrals with $S_l^{(j)}$ from Eq. (3.9-3) to obtain the expressions for the jth-degree scattered wave.

3.13 Interpreting Scattering Using the Stationary-Phase Technique

We shall evaluate the integrals in Eqs. (3.12-5) and (3.12-6) as a function of the position of the LEO by numerical integration aided by the stationary-phase technique, and we outline in Appendix G how these integrals can be calculated by using the Sommerfeld–Watson transformation. However, the stationary-phase concept provides valuable insights and enables a comparison with the thin-screen model for Fresnel diffraction.

3.13.1 Stationary Phase

Contributions to the scattering integrals in Eqs. (3.12-5) and (3.12-6) come from neighborhoods where their aggregate phasor is stationary. Let $\Psi^{(S)} = \Psi^{(S)}(x,\theta,N,x_o,\nu)$ be the aggregate phase, defined by

$$\Psi^{(S)} = 2\Omega^{(S)} + \Phi^- \qquad (3.13\text{-}1)$$

where the scattering phase $2\Omega^{(S)}$ is defined by Eq. (3.9-2) and the geometric phase delay Φ^- is given by Eq. (3.11-1). Similarly, we can define $\Psi^{(j)} = 2\Omega^{(j)} + \Phi^-$ as the phase associated with the jth-degree scattered wave, where $\Omega^{(j)}$ is defined in Eq. (3.9-3). We note again that Φ^- is just the extra phase accumulated by a wave arriving at the LEO from the near-side tangent point of the sphere of radius ν, whilst Φ^+ is the phase accumulated by the wave from the far-side tangent point $(+\pi/2)$. For limb sounders, we can dispense with the Φ^+ component because when x_o is very large it never yields stationary-phase points. In Φ^+, the angles θ and θ_ν add [see Eq. (3.11-1)], resulting in an enormously large rate of phase change with changing ν and, therefore, a negligible contribution to the scattering integrals. Whereas for Φ^-,

which involves the difference in these angles, stationary-phase points for $\Psi^{(S)}$ or for $\Psi^{(j)}$ are realizable.

The integrals in Eqs. (3.12-5) and (3.12-6) involve a convolution of the phasor $\exp(i\Phi^-)$ with a kernel of the form $\left(\exp(2i\Omega)-1\right)$ times a slowly varying factor, $\left(\sin\theta_\nu \cos\theta_\nu\right)^{1/2}$ for $E_\theta^{(s)}$, and $\left(\sin^3\theta_\nu / \cos\theta_\nu\right)^{1/2}$ for $E_r^{(s)}$. These phasor terms may be written as the difference: $2S_l \exp(i\Phi^-) = \exp[i(2\Omega^{(S)} + \Phi^-)] - \exp(i\Phi^-)$. Here the first integral uses the phasor $\exp(i\Psi^{(S)})$, and the second integral uses only the phasor $\exp(i\Phi^-)$. The latter integral involving only $\exp(i\Phi^-)$ (that is, the "-1" part of $2S_l$ [see Eq. (3.9-2)] and, therefore, the part that is independent of the scattering) is easily evaluated. We see that the phasor in the integrand will be stationary when $\partial\Phi^- / \partial\nu = 0$. Using Eq. (3.11-1), one can show that

$$\left.\begin{aligned} \frac{\partial\Phi^-}{\partial\nu} &= \theta_\nu - \theta \\ \frac{\partial^2\Phi^-}{\partial\nu^2} &= \frac{1}{D_\nu} \end{aligned}\right\} \tag{3.13-2}$$

It follows that $\partial\Phi^- / \partial\nu = 0$ when $\theta_\nu = \theta$, that is, when $\nu = \nu^* = x\sin\theta$; at this stationary-phase point, $D_\nu = D_{\nu^*} = x\cos\theta$ and $\Phi^{-*} = D_{\nu^*}$, which is the result to be expected from a geometric optics point of view. If Φ^- is expanded in a quadratic Taylor series about $\nu = \nu^*$, one obtains $\Phi^- \doteq D_{\nu^*} + (\nu - \nu^*)^2 / 2D_{\nu^*}$. In the stationary-phase technique, the quantities varying slowly with spectral number are evaluated at the stationary-phase point and moved outside of the integral, which leaves only the quadratic term. It follows that the contribution to the scattering integral for $E_\theta^{(s)}$ in Eq. (3.12-6) from only the phasor $\mathrm{Exp}[i\Phi^-]$ may be written in terms of a Fresnel integral in the form

$$\left.\begin{aligned} &E_o\sqrt{\frac{1}{2\pi i x\sin\theta}}\int_0^\infty \left(\sin\theta_\nu\cos\theta_\nu\right)^{1/2}\exp\left(i\Phi^-\right)d\nu \doteq \\ &E_o\sqrt{\frac{\sin\theta_{\nu^*}\cos\theta_{\nu^*}}{2\pi i x\sin\theta}}e^{iD_{\nu^*}}\int_0^\infty \exp\left(i\frac{(\nu-\nu^*)^2}{2D_{\nu^*}}\right)d\nu = \\ &\frac{E_o}{\sqrt{2i}}\cos\theta e^{ix\cos\theta}\times\int_{w^*\approx-\infty}^{\infty}\exp\left(i\pi\frac{w^2}{2}\right)dw = E_o\cos\theta e^{ix\cos\theta} \end{aligned}\right\} \tag{3.13-3}$$

The lower limit $w^* = -\nu^* / \left(\pi D_{\nu^*}\right)^{1/2}$ in the last integral is sufficiently large ($\approx -10^4$) that one effectively obtains a complete complex Fresnel integral, which equals $1 + i = (2i)^{1/2}$. The contribution of $\exp(i\Phi^-)$ in Eq. (3.12-5) to $E_r^{(s)}$ is

$$E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \int_0^\infty \left(\frac{\sin^3 \theta_\nu}{\cos\theta_\nu} \right)^{1/2} e^{i\Phi^-} d\nu \doteq E_o \sin\theta e^{ix\cos\theta} \tag{3.13-4}$$

The forms in Eqs. (3.13-3) and (3.13-4) are exactly the components for the phase and amplitude that one obtains in the absence of the scattering sphere for a plane harmonic wave traveling in a vacuum along the positive z-axis,[3] i.e., along $z = r\cos\theta$.

It follows that the total electric field (scattered plus incident, $\boldsymbol{E}^{(i)} + \boldsymbol{E}^{(S)}$) is given from Eqs. (3.12-5) and (3.12-6) with $2S_l$ replaced by the phasor $\exp\left(i2\Omega^{(S)}\right)$; that is, we drop the "–1" part of $2S_l$ in calculating the total field at (r, θ). We recall from Eqs. (3.9-1) and (3.9-2) that the aggregate scattering coefficient can be written in terms of its phasor as $2S_l = \exp\left(i2\Omega^{(S)}\right) - 1$. Dropping the –1 term (and dropping the Φ^+ term), it follows that the total field at the LEO can be written as

$$\left.\begin{aligned} E_r^{(i)} + E_r^{(S)} &= E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \int_0^\infty \left(\frac{\sin^3 \theta_\nu}{\cos\theta_\nu} \right)^{1/2} e^{i(2\Omega^{(S)} + \Phi^-)} d\nu \\ E_\theta^{(i)} + E_\theta^{(S)} &= E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \int_0^\infty \left(\frac{\sin 2\theta_\nu}{2} \right)^{1/2} e^{i(2\Omega^{(S)} + \Phi^-)} d\nu \end{aligned}\right\} \tag{3.13-5}$$

Here $\Omega^{(S)}$ is given in Eq. (3.9-2) or Eq. (3.9-3), and Φ^- is given in Eq. (3.11-1); θ_ν is defined in Eq. (3.10-3) and also shown in Fig. 3-14.

---

[3] This can also be seen by substituting the asymptotic forms for $\xi_l^\pm$ and $P_l(\cos\theta)$ into Bauer's identity in Eq. (3.2-3) for the incident plane wave. Converting this series into an integral and retaining only phasor terms for which stationary values are physically realizable (i.e., θ_ν must be in the range $0 \le \theta_\nu \le \pi/2$ because the spectral number must be positive, $0 < \nu \le\sim x_o \cong x/1.1$), one obtains

$$\exp(ikz) \doteq (2\pi i x \sin\theta)^{-1/2} \int_0^\infty \sqrt{\tan\theta_\nu}\, \exp\left[i\left(x\cos\theta_\nu - \nu\left(\theta_\nu - \theta\right)\right)\right] d\nu$$

Upon applying the stationary-phase technique to this integral, one obtains $\exp(ikz) \doteq \exp(ix\cos\theta)$. Only the ξ^+ term in Eq. (3.2-3) [see Eq. (3.3-1)] provides a stationary-phase contribution when $0 << \theta << \pi/2$ (i.e., θ is well into the first quadrant); only the ξ^- term contributes when $\pi/2 << \theta << \pi$.

3.13.2 Comparison with Diffraction from a Knife-Edge

It will be useful for comparison with thin-screen results to rewrite the scattering integrals in Eqs. (3.12-5) and (3.12-6) (dropping Φ^+) in the form

$$\begin{Bmatrix} E_r^{(s)} \\ E_\theta^{(s)} \end{Bmatrix} = \begin{Bmatrix} I_r + J_r \\ I_\theta + J_\theta \end{Bmatrix} \tag{3.13-6}$$

$$\begin{aligned} \begin{Bmatrix} I_r \\ I_\theta \end{Bmatrix} = E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \Bigg[& \int_0^{x_o} \begin{Bmatrix} \left(\sin^2\theta_\nu \tan\theta_\nu\right)^{1/2} \\ \left(\sin\theta_\nu \cos\theta_\nu\right)^{1/2} \end{Bmatrix} e^{i\Psi^{(S)}} d\nu \\ & + \int_{x_o}^{\infty} \begin{Bmatrix} \left(\sin^3\theta_\nu / \cos\theta_\nu\right)^{1/2} \\ \left(\sin\theta_\nu \cos\theta_\nu\right)^{1/2} \end{Bmatrix} d\nu \left(e^{i\Psi^{(S)}} - e^{i\Phi^-} \right) \Bigg] \end{aligned} \tag{3.13-7}$$

and

$$\begin{Bmatrix} J_r \\ J_\theta \end{Bmatrix} = -E_o \sqrt{\frac{1}{2\pi i x \sin\theta}} \left[\int_0^{x_o} \begin{Bmatrix} \left(\sin^2\theta_\nu \tan\theta_\nu\right)^{1/2} \\ \left(\sin\theta_\nu \cos\theta_\nu\right)^{1/2} \end{Bmatrix} e^{i\Phi^-} d\nu \right] \tag{3.13-8}$$

The J integrals in Eq. (3.13-8) essentially provide the scattering that one would obtain from a knife-edge, a result that Arnold Sommerfeld obtained 100 years ago [15]. This can be seen by evaluating Eq. (3.13-8) using the stationary-phase technique in the same manner as was used to obtain Eq. (3.13-3); thus, we obtain

$$\left.\begin{aligned} \begin{Bmatrix} J_r \\ J_\theta \end{Bmatrix} &\doteq -E_o \exp(ix\cos\theta) \begin{Bmatrix} \sin\theta \\ \cos\theta \end{Bmatrix} \frac{1}{1+i} \int_{\tilde{w}}^{\infty} \exp\left(i\frac{\pi}{2} w^2 \right) dw, \\ \tilde{w} &\doteq \left(\theta - \theta_o\right) \sqrt{x\cos\theta/\pi} \end{aligned}\right\} \tag{3.13-9}$$

On the other hand, consider the Fresnel diffraction from a semi-infinite plane mounted perpendicular to the z-axis and at a horizontal distance (in phase units) $D = x\cos\theta$ from the LEO (see Fig. 3-16). The upper boundary of the plane is located at a distance x_o above the z-axis. From the Rayleigh–Sommerfeld formulation of Fresnel diffraction for this knife-edge (see Chapter 2, Eq. (2.5-1), where now D, h, and h_{LG} are in phase units), we have

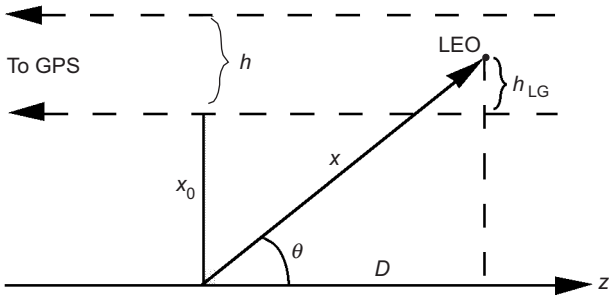

Fig. 3-16. Thin-screen geometry for a knife-edge mounted perpendicular to $\hat{z}$. A GPS satellite at infinite distance in the negative $\hat{z}$-direction.

$$\begin{Bmatrix} E_r^{(S)} \\ E_\theta^{(S)} \end{Bmatrix}_{\substack{\text{knife}- \\ \text{edge}}} = E_o \begin{Bmatrix} \sin\theta \\ \cos\theta \end{Bmatrix} \sqrt{\frac{1}{2\pi i D}} \int_0^\infty \exp\left(i\left(D + \frac{(h - h_{\text{LG}})^2}{2D} \right) \right) dh \qquad (3.13\text{-}10)$$

Noting from Fig. 3-16 that $h_{\text{LG}} = x\left(\sin\theta - \sin\theta_o\right) \doteq x(\theta - \theta_o)\cos\theta$ and that $D = x\cos\theta = D_{\nu^*}$, we convert the integral in Eq. (3.13-10) into a Fresnel integral to obtain the identity between Eqs. (3.13-9) and (3.13-10). In certain specific examples to follow, the difference between Mie scattering from a sphere and the diffraction from the thin-screen knife-edge are borne by the I integrals in Eq. (3.13-7).

3.13.3 Stationary-Phase Points for the Scattering Integrals

For the contribution to the scattering integrals in Eq. (3.13-5) from the phasor $\text{Exp}\left[i\Psi^{(S)}\right]$, the main contributions come from neighborhoods around points in spectral number space where $\partial\Psi^{(S)}/\partial\nu = 2\partial\Omega^{(S)}/\partial\nu + \partial\Phi^-/\partial\nu = 0$. To compute $\partial\Psi/\partial\nu$, we will need explicit forms for $\partial\xi_\nu^\pm/\partial\nu$. Recalling that $\nu = \mu x$ in the asymptotic forms for $\xi_\nu^\pm(x)$ [Eqs. (3.2-8), (3.8-1), and (3.8-2)], we have

$$\left.\begin{aligned} \xi_\nu^{\pm\prime} \equiv \left(\frac{\partial\xi_\nu^\pm}{\partial x}\right)_\nu &= \left(\frac{\partial\xi_\nu^\pm}{\partial\mu}\right)_\nu \left(\frac{\partial\mu}{\partial x}\right)_\nu \\ \left(\frac{\partial\xi_\nu^\pm}{\partial\nu}\right)_x &= \left(\frac{\partial\xi_\nu^\pm}{\partial\mu}\right)_\nu \left(\frac{\partial\mu}{\partial\nu}\right)_x + \left(\frac{\partial\xi_\nu^\pm}{\partial\nu}\right)_\mu \end{aligned}\right\} \qquad (3.13\text{-}11)$$

where $\mu = \nu / x$. From the asymptotic forms in Eqs. (3.8-1) and (3.8-2), it follows that

$$\left(\frac{\partial \xi_\nu^{\pm}}{\partial \nu}\right)_x = -F(\mu)\left(\frac{\partial \xi_\nu^{\pm}}{\partial x}\right)_\nu + \mathrm{O}[x^{-1}] \tag{3.13-12}$$

where $F(\mu)$ is defined by

$$F(\mu) = \frac{\tan^{-1}\left(\sqrt{\mu^{-2} - 1}\right)}{\sqrt{1 - \mu^2}} \tag{3.13-13}$$

Note that $F(\mu)$ also is real for $\mu > 1$. Similarly, one obtains

$$\left(\frac{\partial \xi_\nu^{\pm'}}{\partial \nu}\right)_x = \frac{\partial^2 \xi_\nu^{\pm}}{\partial x \partial \nu} = G(\mu)\xi_\nu^{\pm} + \mathrm{O}[x^{-1}] \tag{3.13-14}$$

where

$$G(\mu) = \sqrt{1 - \mu^2}\,\tan^{-1}\left(\sqrt{\mu^{-2} - 1}\right) \tag{3.13-15}$$

Here $G(\mu)$ also is real for $\mu > 1$.

From these expressions, it follows from Eq. (3.5-11) that the partial derivatives of the Wronskian forms are given by

$$\begin{aligned}\left(\frac{\partial \mathcal{W}_\nu}{\partial \nu}\right)_x = N\Big[&\big(F(\mu) + \mu F'(\mu)\big)\xi_\nu^{+\,\prime}(x_o)\xi_\nu^{-\,\prime}(nx_o) \\ &+\big(\mu G'(\mu) - G(\mu)\big)\xi_\nu^{+}(x_o)\xi_\nu^{-}(nx_o)\Big] + \mathrm{O}\Big[(N\mu)^2\Big]\end{aligned} \tag{3.13-16}$$

and similarly for $\left(\partial \mathcal{W}_\nu^{\pm} / \partial \nu\right)_x$. Here $\mu_o = 1$. With these partial derivatives for the Wronskian forms, it is possible to write an explicit expression for $\partial \Omega^{(S)} / \partial \nu$ and also for any one of the individual jth-degree scattering phases $\partial \Omega^{(S_j)} / \partial \nu$, $j = 0, 1, 2, \cdots$, as given by Eqs. (3.9-2) and (3.9-3). Thus, we have all the parts assembled to study the behavior of $\partial \Psi^{(S)} / \partial \nu$ and $\partial \Psi^{(S_j)} / \partial \nu$.

Figure 3-17 shows the variation of $\Psi^{(S)}(N, x, \theta, x_o, \nu)$ with spectral number for specific values of the LEO angular coordinate $\theta_o - \theta$. These correspond to orbital positions ranging from 1.5 to 7 km into the geometric shadow of the sphere of radius r_o. Two different refractivity values, $Nx_o = 83\pi$ and

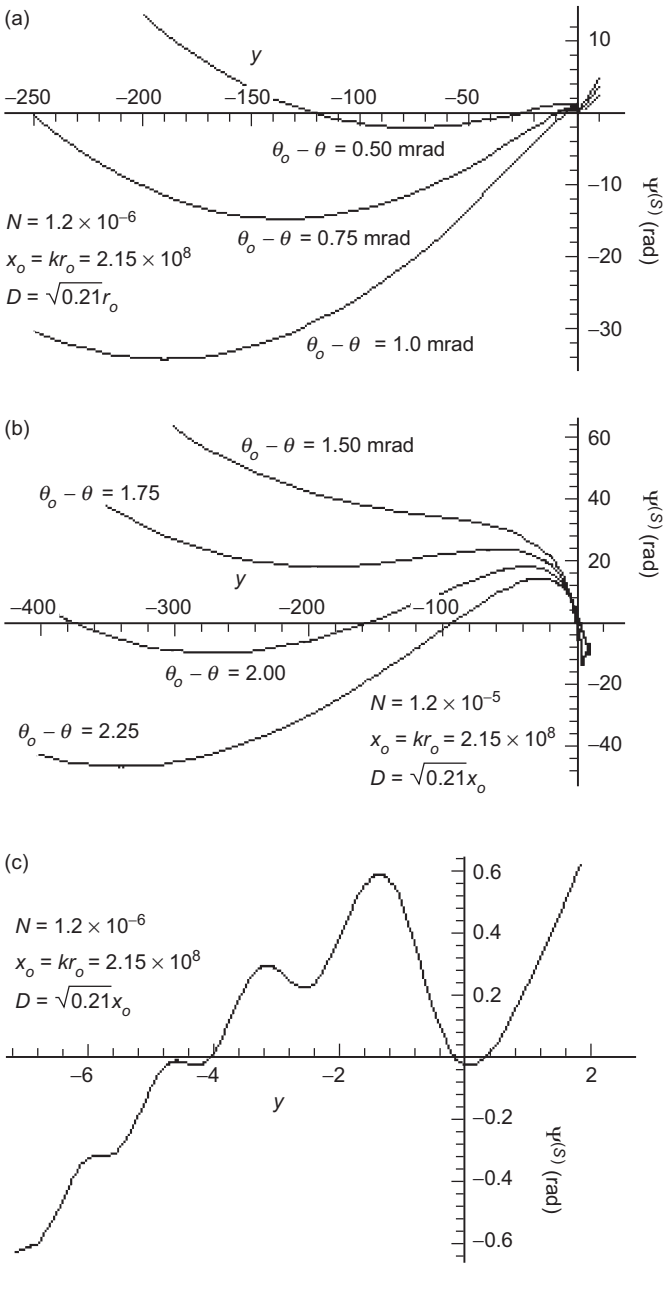

Fig. 3-17. Scattering phase function versus spectral number for different LEO orbital positions: (a) in the asymptotic regime with $Nx_o = 83\pi$, (b) in the asymptotic regime with $Nx_o = 830\pi$, and (c) near x_o with $Nx_o = 83\pi$, $\theta - \theta_o = -1.0$.

$Nx_o = 830\pi$, are used in Figs. 3-17(a) and 3-17(b). The abscissa is given in terms of the "dimensionless" spectral number difference y, which is the argument of the Airy functions and which in turn is given in terms of the difference between the spectral number ν and radial coordinate x_o. For small values of y, it follows from Eq. (3.8-6) that

$$\left.\begin{aligned} y \doteq K_{x_o}^{-1}(\nu - x_o), \;\; \nu \doteq x_o + K_{x_o} y, \;\; K_{x_o} = \left(\frac{x_o}{2}\right)^{1/3} \end{aligned}\right\} \qquad (3.13\text{-}17)$$

For $\theta - \theta_o < 0$, we will show later that the locations of the principal stationary-phase points lying in the regime where the negative-argument asymptotic forms for the Airy functions apply are given by the positive real roots to the cubic polynomial $\mathrm{P}[\xi]$

$$\left.\begin{aligned} &\xi^3 + K_{x_o}^{-1} D(\theta - \theta_o)\xi + 2DN + \mathrm{O}[N^2] = 0, \\ &y = -\xi^2, \quad D = kr\cos\theta_o \end{aligned}\right\} \qquad (3.13\text{-}18)$$

For example, Fig. 3-17(a) shows that the stationary values of $\Psi^{(S)}$ with $Nx_o = 83\pi$ and $\theta - \theta_o = -1$ mrad are located at $y \cong -190$ and $y \cong -1.5$. The more negative point is located near where the $-\partial\Phi^-/\partial\nu$ curve crosses zero. We also will show later that these points yield the phase delay and bending predicted by geometric optics for the degree $j = 1$ scattered wave (see Fig. 3-2), without accounting for rays arising from external or internal reflections. Figure 3-17(b) shows $\Psi^{(S)}$ for $Nx_o = 830\pi$; the stationary points occur deeper into the geometric shadow because the refractivity is larger and the refractive bending greater. The contribution of the more negative stationary-phase point dwarfs (for the magnitudes of Nx_o used in these figures) the combined contribution from the less negative point as well as the contributions from higher-degree scattered waves. Their effect is shown in Fig. 3-17(c), which displays $\Psi^{(S)}$ versus spectral number in the vicinity of x_o for the same conditions used in Fig. 3-17(a) with $\theta - \theta_o = -1$ mrad. We see that the predicted stationary-phase point here ($y \cong -1.5$) based on asymptotic theory [Eq. (3.13-18)] is located among a complex of such points. This complex arises because the Airy functions depart significantly from their negative argument asymptotic forms in this region.

Figure 3-18 shows the points for spectral numbers in the vicinity of x_o where $\partial\Psi^{(S)}/\partial\nu = 0$. These are the intersections of the $2\partial\Omega^{(S)}/\partial\nu$ curve with the $-\partial\Phi^-/\partial\nu$ curve [the very slightly negatively sloped line denoted as

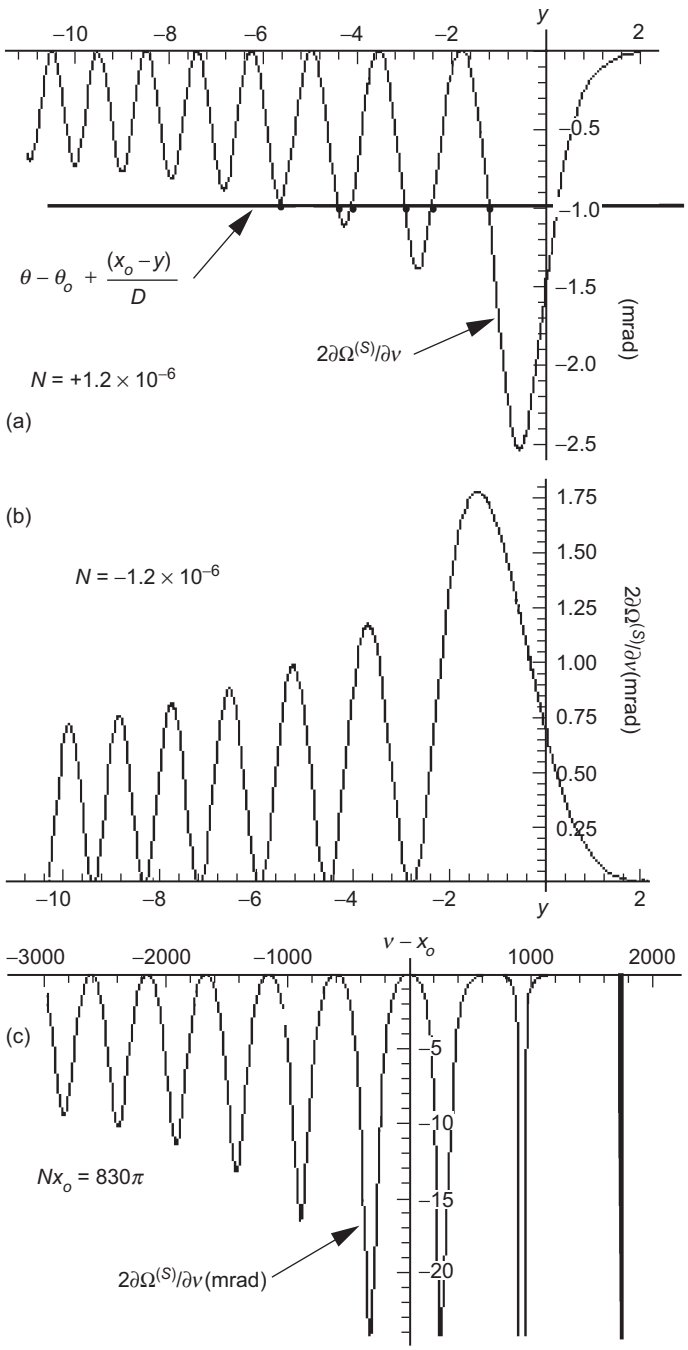

Fig. 3-18. Stationary phase points for the total scattering integral near $\nu = x_o$: (a) $Nx_o = +83\pi$, (b) near $Nx_o = -83\pi$, and (c) $Nx_o = 830\pi$. Points occur at intersections of the $2\partial\Omega^{(S)}/\partial\nu$ and $\partial\Phi^{-}/\partial\nu$ curves.

$\theta - \theta_o + (x_o - \nu)/D$ in Fig. 3-18(a)]. The contributions to the scattering integral from the stationary points in the immediate vicinity of x_o (and shown in Fig. 3-18) are the vestigial effects (because N is so small) from higher-degree scattering; their contribution is very small compared to that from the degree $j = 1$ wave. The dominant stationary-phase point providing most of the contribution to the scattering integral is off-scale in this figure ($y \approx -190$). As the occultation evolves in time, the near-horizontal line in Fig. 3-18(a) will migrate downward (for an immersion) at a rate of roughly 1 mrad/s. For $\theta - \theta_o > 0$, no stationary-phase points occur, but near-stationary points will become more pronounced as the LEO enters the geometric shadow zone and $\theta - \theta_o$ becomes more negative; the neighborhoods in ν-space about these near-stationary points will contribute to the scattering integrals. The asymptotic form for the $2\partial\Omega^{(S)}/\partial\nu$ curve is easily obtained from Eqs. (3.8-12) and (3.9-2).

Figure 3-18(b) is similar to Fig. 3-18(a) except that the refractivity is negative and the $-\partial\Phi^{-}/\partial\nu$ curve has been omitted. For $N > 0$, Fig. 3-18(a) shows that diffraction effects will begin rather abruptly as $\theta - \theta_o \to 0^{+}$ and that they taper off more gradually as $\theta - \theta_o$ becomes more negative. The reverse is true in Fig. 3-18(b), where $N < 0$. One sees these same characteristics in the scalar diffraction/thin-screen model. (For examples, see Chapter 2, Fig. 2-11.) The extent in θ-space over which the principal diffraction effects will prevail is about 3 mrad for $N = +1.2 \times 10^{-6}$ and about 2 mrad for $N = -1.2 \times 10^{-6}$, the former being close to the extent shown in Chapter 2, Figs. 2-17 and 2-18, for the Fresnel effects in phase and amplitude, as predicted by the thin-screen model for the same values of the parameters N, r_o, and λ. When $\theta - \theta_o = 0$ exactly, it marks the edge of the geometric shadow zone, and we can see from Fig. 3-18 that results from Mie scattering theory will depend on the curvature in the $\partial\Phi^{-}/\partial\nu$ curve to avoid numerical integration difficulty here due to slow convergence with decreasing spectral number.

3.13.4 Vestigial Rainbow Effects

What has become of the rainbow effects in Mie scattering theory? For degree $j = 2$, these are predicted by geometric optics to be most pronounced around $\theta - \theta_o = -3.5\sqrt{2N}$ [see Eq. (3.4-9)] or at about -5.4 mrad for $N = 83\pi / x_o = 1.2 \times 10^{-6}$. Figure 3-17 shows that they and the other higher-degree caustics have been almost completely washed out when all of the scattering degrees are summed up. This is due to the relatively small size of $N(r_o/\lambda)^{2/3}$ [~0.1 in Fig. 3-17(a) and ~1 in Fig. 3-17(b)] and the relatively large size of the first Fresnel zone. The caustic surfaces within the sphere for

$j \geq 2$ are all crowded into a region below the surface of the sphere that is only a few meters thick, much smaller than the first Fresnel zone and, therefore, washed out when observed at a distance by the LEO. By contrast, for a raindrop observed in the visual spectrum, $N(r_o/\lambda)^{2/3} \sim 100$. Moreover, rainbows are observed from afar spread over a cloud of coherently scattering raindrops, which has a spatial extent that is much larger than the first Fresnel zone.

Figures 3-19(a) and 3-19(b) show $\partial\Omega^{(S_j)}/\partial\nu$ versus $\nu - x_o$ for the individual scattered wave of degree $j = 2$ for $Nx_o = 83\pi$ and $Nx_o = 830\pi$. The stationary points in these figures define points where $\partial^2\Psi^{(S_2)}/\partial\nu^2 = 0$, the analog of a caustic surface but in ν-space rather than h-space. These figures show the potential for brightening when $\theta - \theta_o \approx -3.5\sqrt{2N}$, which these figures place at roughly the correct scattering angle [–6 to –8 mrad in Fig. 3-19(a)]. However, the situation is more complicated because two caustics rather than one occur in ν-space in Mie scattering theory, and the rays at these caustic contact points interfere with each other. Moreover, these contributions tend to be swamped by the effects of the dominant scattering coefficients $b^{(0)}$ and $b^{(1)}$ when N is small. As N grows larger, the caustic features at $\partial^2\Psi^{(S_2)}/\partial\nu^2 = 0$ become prominent and isolated.

As N is set to progressively larger values, the stationary-phase points near the shadow boundary become increasingly complex. Figure 3-18(c) shows $\partial\Omega^{(S)}/\partial\nu$ with $Nx_o = 830\pi$, ten times larger than the magnitude used in Figs. 3-18(a) and 3-18(b). When N is large and when $\theta - \theta_o = 0$, or when it is sufficiently close to zero, the stationary-phase intervals for $\Psi^{(S)}$ will be numerous and extensive, as Fig. 3-18(c) suggests. For larger values of *N*, these may cause convergence problems for both the stationary-phase and numerical integration techniques when the LEO is at or very near the shadow boundary of the sphere.

From Eq. (3.9-2), it follows that these spikes [and the resulting stair-step pattern they create in $\Omega^{(S)}$, see Fig. 3-11(b)] arise because of near-zero crossings in the imaginary part of $\mathcal{W}_l^+ + \mathcal{W}_l$ when the real part is already of diminished value [see Eqs. (3.8-7), (3.8-8), (3.8-12), and (3.8-13)]. When N is sufficiently large, there is a significant transition zone in spectral number, roughly defined by $x_o < \nu < nx_o$. Here the argument $y = v^{2/3}\zeta[\nu/x_o]$ in the Airy functions is positive and the Airy functions have become exponential-like as a function of real ν, but the Airy functions with their argument $\hat{y} = v^{2/3}\zeta[\nu/nx_o]$, which is negative in this zone, are still oscillatory. It is easily shown that $\hat{y} \doteq y - Nx_o K_{x_o}^{-1}$. For sufficiently large N, these near-singular points occur within the transition zone where ν satisfies the condition

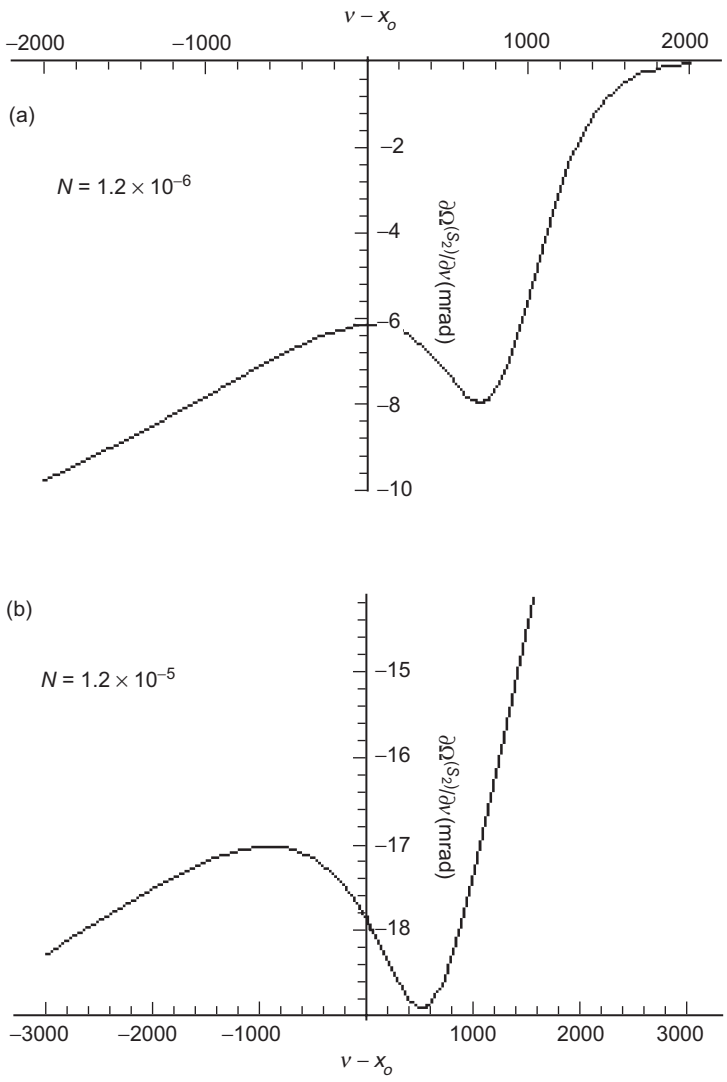

Fig. 3-19. $\partial\Omega^{(j)}/\partial\nu$ **for the degree** $j = 2$ **scattering coefficient: (a)** $N = 1.2 \times 10^{-6}$ **and (b)** $N = 12 \times 10^{-6}$. **Stationary points in spectral number are caustic points. Geometric optics predicts** $\theta - \theta_o = -3.5\sqrt{2N}$ **as the angular position of the caustic ray.**

$$\tan X \cong -n^{-2}\sqrt{\frac{\hat{y}}{y}}, \quad x_o < \nu < nx_o \tag{3.13-19}$$

and whose spacing on the ν axis becomes quite regular until $\nu >\sim nx_o$.

We can estimate the threshold of N above which the staircase features in $\Omega^{(S)}$ begin to appear. From Fig. 3-10, we note that when the argument y (y real) of the Airy functions is less than -2, the Airy functions are decidedly sinusoidal, and when $y > +2$, they are exponential. A value of $|y| = 2$ corresponds to a spectral number separation of $|\nu - x_o| \doteq 2K_{x_o}$. Equating twice this separation to Nx_o yields a threshold $N \geq 2K_{x_o}^{-2} \approx 1 \times 10^{-5}$. The larger value $Nx_o = 830\pi$ used in Figs. 3-17 and 3-18 is at that threshold.

3.13.5 Asymptotic Solutions for a Transparent Refracting Sphere

When the LEO is in the immediate vicinity of the shadow boundary, $\theta \cong \theta_o$, a numerical solution to the scattering problem for the refracting sphere is necessary because of the multiple stationary-phase points for ν in the vicinity of x_o (see Figs. 3-17 and 3-18) and because the Airy functions are not well-represented by their asymptotic forms in this neighborhood. As the LEO descends deeper into the shadow, two stationary-phase points quickly emerge whose location in ν space is well below x_o. Figure 3-17(b) shows the variation of $\Psi^{(S)}$ with spectral number for a range of LEO angular coordinates and with $Nx_o = 830\pi$. The asymptotic stationary points are evident in this figure. We will show here that these stationary-phase points yield the phase delays and bending angles predicted by geometric optics, plus a diffraction pattern that agrees closely with the Fresnel diffraction predicted by the thin-screen model.

Let ν^* denote a stationary-phase point. When $\nu^* - x_o <\sim -2K_{x_o}$, the negative-argument asymptotic forms for the Airy functions grow increasingly more accurate. From Eqs. (3.8-3), (3.8-4), and (3.9-7), the asymptotic form Ω_A for the aggregate scattering phase $\Omega^{(S)}$ is given by

$$\left.\begin{aligned} 2\Omega^{(S)} \to 2\Omega_A &\doteq \frac{4}{3}\nu\left(\left(-\zeta(\mu/n)\right)^{\frac{3}{2}} - \left(-\zeta(\mu)\right)^{\frac{3}{2}}\right) \\ &= 2\left(\sqrt{n^2x_o^2 - \nu^2} - \nu\cos^{-1}\left(\frac{\nu}{nx_o}\right) - \sqrt{x_o^2 - \nu^2} + \nu\cos^{-1}\left(\frac{\nu}{x_o}\right)\right) \end{aligned}\right\} \tag{3.13-20}$$

where $\mu = \nu / x_o < 1$. We have already noted that this asymptotic form for the scattering phase also applies to the degree $j = 1$ scattered wave, which corresponds in geometric optics to the principal ray passing through the sphere without internal reflection. The stationary-phase condition (dropping the Φ^+ term) is given by

$$\frac{\partial \Psi^{(S)}}{\partial \nu} \to 2\frac{\partial \Omega_A}{\partial \nu} + \frac{\partial \Phi^-}{\partial \nu} = 0 \tag{3.13-21}$$

Let ν^* denote the value(s) for which Eq. (3.13-21) holds. From Eqs. (3.11-1) and (3.13-20), it follows that

$$\left.\begin{aligned} &-2\left\{\cos^{-1}\left(\frac{\nu^*}{nx_o}\right) - \cos^{-1}\left(\frac{\nu^*}{x_o}\right)\right\} + \theta_{\nu^*} - \theta = 0, \\ &x_o \sin\theta_{\nu^*} = \nu^* \sin\theta_o \end{aligned}\right\} \tag{3.13-22}$$

To solve this expression for ν^* explicitly in terms of $\theta - \theta_o$, we note that ν^* is close in value to nx_o. Therefore, we write $\nu^* / (nx_o)$ as

$$\left.\frac{\nu^*}{nx_o} \doteq 1 + \frac{\hat{y}^*}{2K_{\rho_o}^2},\ K_{\rho_o} = \left(\frac{\rho_o}{2}\right)^{1/3},\ \rho_o = nx_o,\ -\frac{\hat{y}^*}{2K_{\rho_o}^2} << 1,\right\} \tag{3.13-23}$$

Expanding Eq. (3.13-22) in powers of $\hat{y}^* / K_{\rho_o}^2$ and also noting that $\theta_{\nu^*} \doteq \theta_o + \hat{y}^* \operatorname{Tan}\theta_o / \left(2K_{\rho_o}^2\right)$, one obtains

$$\left.2\left\{\sqrt{-\frac{\hat{y}^*}{K_{\rho_o}^2}} - \sqrt{2N - \frac{\hat{y}^*}{K_{\rho_o}^2}}\right\} + \theta - \theta_o - \frac{1}{2}\frac{y^*}{K_{\rho_o}^2}\tan\theta_o \doteq 0\right\} \tag{3.13-24}$$

This yields a quartic polynomial in $(-\hat{y}^* / K_{\rho_o}^2)^{1/2}$, the real roots of which provide the negative stationary-phase values of $\hat{y}^*$. Assuming that $-\hat{y}^* >> NK_{\rho_o}^2$, one can expand Eq. (3.13-24) to obtain

$$\left.\begin{aligned}
&P[\xi] = \xi^3 + K_{\rho_o}^{-1} D(\theta - \theta_o)\xi + 2DN + \mathrm{O}\left[N^2\right] = 0, \\
&\hat{y} = -\xi^2, \quad \hat{y}^* = \left(-\xi_1^2, -\xi_2^2\right), \quad \xi \geq 0, \\
&\nu^* = nx_o\left(1 + \frac{\hat{y}^*}{2K_{\rho_o}^2} + \mathrm{O}\left[\left(K_{\rho_o}^{-2}\hat{y}^*\right)^2\right]\right), \quad K_{\rho_o} = \left(\frac{\rho_o}{2}\right)^{\frac{1}{3}}, \quad \rho_o = nx_o \\
&\xi_1 = 2\sqrt{D\frac{\theta_o - \theta}{3K_{\rho_o}}}\sin\Theta, \quad \xi_2 = 2\sqrt{D\frac{\theta_o - \theta}{3K_{\rho_o}}}\cos(\pi/6 + \Theta), \\
&\Theta = \frac{1}{3}\sin^{-1}\left(3N\sqrt{\frac{3\tan\theta_o}{2(\theta_o - \theta)^3}}\right), \quad D = nkr\cos\theta_o
\end{aligned}\right\} \qquad (3.13\text{-}25)$$

This cubic polynomial $P[\xi]$ in Eq. (3.13-25) is the same as that given earlier in Eq. (3.13-18). The quantities ξ_1 and ξ_2 given by their expressions in Eq. (3.13-25) are the positive roots of $P[\xi]$.

Figure 3-17(b) shows that no real positive stationary-phase asymptotic solution exists in Mie scattering theory when $\theta_o - \theta$ drops below a certain critical value, which corresponds to $P[\xi]$ having no real positive roots. This critical value can be obtained by adjusting $\theta_o - \theta$ so that the positive roots in Eq. (3.13-25) merge, $P'[\xi] = P[\xi] = 0$. One obtains

$$\theta^\dagger - \theta_o = -3\left(\frac{nx_o}{2D}\right)^{1/3} N^{2/3} \qquad (3.13\text{-}26)$$

which corresponds to a LEO angular position that is slightly inside of the shadow boundary. For the conditions shown in Fig. 3-17(b), Eq. (3.13-26) predicts that $\theta^\dagger - \theta_o = -1.63$ mrad and the stationary-phase point in this case is given by $y^\dagger = -(DN)^{2/3} \cong -112$. Because $\partial^2\Psi^{(S)}/\partial\nu^2$ also equals zero at this stationary-phase point when $\theta^\dagger - \theta_o = -1.63$ mrad, this value of θ corresponds in geometric optics to the first contact with the caustic surface generated by the ray system associated with the system of stationary-phase points shown in Fig. 3-17(b) nearer to x_o (at $y^* = -(\xi_1)^2$). When the scattering integrals for $E_r^{(S)}$ and $E_\theta^{(S)}$ in Eqs. (3.12-5) and (3.12-6) are evaluated, we should expect significant flaring for LEO angular coordinates near $\theta^\dagger$ because of the broad band of spectral numbers about the stationary-phase point that constructively contributes to the scattering integral when $\partial^2\Psi^{(S)}/\partial\nu^2$ is near zero at the stationary-phase point. The range of angular coordinates $\theta_o \geq \theta \geq \theta^\dagger$ corresponds in geometric optics to the LEO lying in a shadow zone where no

stationary-phase points exist in impact parameter space. This is a super-refractivity effect. In Mie scattering theory, we see that this zone yields a situation where there are no stationary-phase points in spectral number space (except for the vestigial effects of higher-degree scattering); so, we should expect considerable darkening in this region. For $\theta \le \theta^{\dagger}$, the scattering integral supports two stationary-phase points based on asymptotic theory, plus the effects of higher-degree scattering nearer to $y = 0$ if N is sufficiently large. The contributions from this pair of asymptotic points ξ_1 and ξ_2 will interfere with each other just as in geometric optics, where the branching and anomalous rays interfere with each other [see Chapter 2, Fig. 2-2(b)].

Equation (3.13-26) can be compared with the point of maximum signal amplitude predicted by the thin-screen model. From Chapter 2, Eq. (2.8-2), we have from Snell's law the refractive bending angle $\alpha_S(h)$ from a sphere of radius r_o with a uniform refractivity $N = n - 1$:

$$\left.\begin{aligned} \alpha_S &= 2\left(\sin^{-1}\left(\frac{nx_o + h}{x_o}\right) - \sin^{-1}\left(\frac{x_o + h}{x_o}\right)\right) \\ &\doteq \sqrt{8}\left(\sqrt{\frac{-h}{x_o}} - \sqrt{\frac{-h}{x_o} - N}\right) \\ h \le h^c &= -(n-1)r_o = -Nx_o, \quad x = kr, \quad n > 1 \end{aligned}\right\} \tag{3.13-27}$$

where h is the altitude of the impact parameter of the ray above the sphere (see Fig. 3-16) and, therefore, negative for Eq. (3.13-27) to apply. Here h^c is the altitude of critical refraction. In geometric optics, no rays occur with turning points in the interval $x_o(1 - N) < x_* < x_o$. From the thin-screen model (see Fig. 2-3),

$$h \doteq h_{\mathrm{LG}} + D\alpha_S(h) \tag{3.13-28}$$

We know that a point where $dh_{\mathrm{LG}} / dh = 0$ marks the first contact with a caustic and also the point where both the first and second derivatives of the Fresnel phase in the scalar diffraction integral [Eq. (2.5-1)] are zero. Using Eq. (3.13-27) to find $h^{\dagger}$, which is the value of h at the point where $dh_{\mathrm{LG}} / dh = 0$, and assuming that $|h^{\dagger}| >> Nx_o$, one obtains

$$h^{\dagger} = -K_{\rho_o}(ND)^{2/3} \tag{3.13-29}$$

where $\rho_o = nx_o$. It follows, using Eq. (3.13-28), that

$$h_{\mathrm{LG}}(h^{\dagger}) \doteq D(\theta - \theta_o) = -3K_{\rho_o}\left(ND\right)^{2/3} \tag{3.13-30}$$

which is in perfect agreement with the asymptotic Mie scattering prediction in Eq. (3.13-26). One can further refine this by using the third-order Fresnel theory applied to the thin-screen model (see Appendix D and Section 3.16) to calculate the small offset from the inflection point ($\partial^2\Psi^{(S)}/\partial\nu^2 = 0$) to obtain maximum flaring.

Returning to Eqs. (3.13-21) through (3.13-23), we note that Eq. (3.13-22) may be rewritten as

$$\delta = \theta_{\nu^*} - \theta = 2\left\{\sin^{-1}\left(\frac{\nu^*}{x_o}\right) - \sin^{-1}\left(\frac{\nu^*}{nx_o}\right)\right\},\ \ \nu^* = x_o\,\frac{\sin(\theta+\delta)}{\sin\theta_o}\right\} \tag{3.13-31}$$

A comparison of Eq. (3.13-31) with the refractive bending angle given by Eq. (3.13-27) suggests that $\theta_{\nu^*} - \theta \leftrightarrow \alpha_S$, and at least to first order in N, $\nu^* \leftrightarrow h + nx_o$.

3.13.6 Correspondence Between the Stationary-Phase Value for Spectral Number and Snell's Law

An alternate view notes the correspondence between the stationary-phase point ν^* in Mie scattering theory and the law of refraction in geometric optics, in particular, the correspondence between ν^* and Snell's law for the refracting sphere. Figure 3-20 shows the corresponding refracting sphere expressed in spectral number space. Here the ray from the GPS satellite impacts the sphere at an angle of incidence φ and it is bent by an angle ε as it crosses the boundary surface. It continues through the uniform refracting sphere and by spherical symmetry it exits the sphere at an angle φ with respect to the local

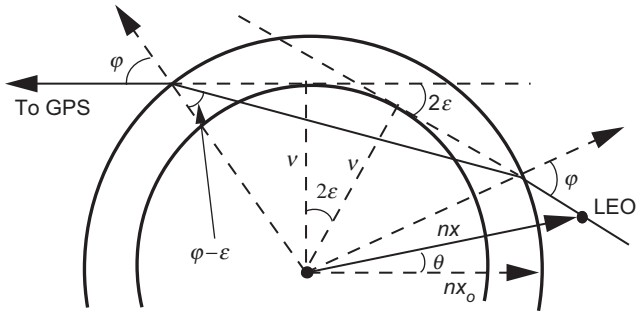

Fig. 3-20. Geometry for a refracting sphere expressed in spectral number space. The spectral number ν that provides a stationary phase point in wave theory and also yields Snell's law, $\sin\varphi = n\sin(\varphi-\varepsilon)$ **in geometric optics.**

normal vector as it continues on to the LEO. This particular path incurs a total bending angle of 2ε. In Fig. 3-20, the two radius vectors of length ν are drawn perpendicular to the straight-line extensions of the incident and departing rays, respectively. From the geometry in Fig. 3-20, it follows that

$$\left.\begin{aligned} \nu = nx_o \sin\varphi = nx\sin(\theta + 2\varepsilon) \quad \text{or} \quad \sin\varphi = \frac{\sin(\theta + 2\varepsilon)}{\sin\theta_o} \end{aligned}\right\} \tag{3.13-32}$$

Here $x_o = x\sin\theta_o$. From Eqs. (3.13-31) and (3.13-32), when we set $\nu = \nu^* = nx_o \sin(\theta + \delta)/\sin\theta_o$, it follows that ε must assume the value $\varepsilon = \delta/2 = \alpha/2$. Snell's law is also satisfied when $\varepsilon = \alpha/2$: $n\sin(\varphi - \alpha/2) = \sin\varphi$. Moreover, upon noting in Fig. 3-20 that $nx_o \sin(\varphi - \alpha/2) = nx_o + h$, it follows through first order in N that

$$\left.\begin{aligned} \nu^* = nx_o + h, \quad y^* \doteq K_{\rho_o}^{-1}(h + Nx_o), \quad \rho_o = nx_o \end{aligned}\right\} \tag{3.13-33}$$

Therefore, to at least first order in N, the value of ν^* is identified as equivalent to the value of the impact parameter ka in geometric optics, which is a constant along the ray path (Bouguer's law) when spherical symmetry applies.

In summary, when the LEO is well inside the shadow boundary (so that the asymptotic forms for the Airy functions apply), Mie scattering theory at the stationary-phase point(s) in spectral number space ν^* yields values for the bending angle of the total wave after scattering (incident plus scattered wave) that to at least first order in N are the same as those given by geometric optics. This should not be too surprising. The stationary-phase point ν^* itself corresponds to the impact parameter in geometric optics, at least to first order in N. We will now show that ν^* also yields the geometric-optics-predicted values for the LEO-observed phase and amplitude when ν^* is well isolated in spectral number space so that the stationary-phase technique is valid. When the LEO is on or near the shadow boundary, where the asymptotic forms for the Airy functions do not apply with sufficient accuracy, then a numerical solution of the spectral integrals in Eqs. (3.12-5) and (3.12-6), or their equivalent, is required to obtain the amplitude and phase of the field.

3.13.7 Stationary Value for the Scattering Phase

The stationary value of the asymptotic form for the scattering phase $\Psi^{(S)} \to \Psi_A^{(S)} = 2\Omega_A^{(S)} + \Phi^-$ can be obtained by expanding $2\Omega_A^{(S)}$ in fractional powers of $\left(\nu^2 - x_o^2\right)/x_o^2$. From Eqs. (3.8-5), (3.9-8), (3.11-1), (3.13-20), and (3.13-21), we have

$$\Psi_A^{(S)}(\nu) = \frac{2}{3\nu^2}\left(\left(n^2x_o^2 - \nu^2\right)^{3/2} - \left(x_o^2 - \nu^2\right)^{3/2}\right) + D_\nu + \nu(\theta_\nu - \theta) + \mathrm{O}\left[\left(\left(x_o^2 - \nu^2\right)x_o^{-2}\right)^{5/2}\right] \tag{3.13-34}$$

where $D_\nu = x\cos\theta_\nu = x_o\cos\theta_\nu / \sin\theta_o$, which is the geometric phase delay term from the tangent point on a sphere of radius ν to the LEO if it were located at θ_ν, and $\nu(\theta_\nu - \theta)$ is the extra geometric phase delay because of the small offset $\theta_\nu - \theta$ to account for the actual location of the LEO at θ. We note that at the stationary-phase value ν^* the geometric delay term becomes

$$\begin{aligned} D_{\nu^*} + \nu^*(\theta_{\nu^*} - \theta) &= x\left(\cos(\theta + \alpha_S) + \alpha_S\sin(\theta + \alpha_S)\right) \\ &= D_{\theta+\alpha_S} + a\alpha_S \end{aligned} \tag{3.13-35}$$

where $D_{\theta+\alpha_S} = kr\cos(\theta + \alpha_S)$ and $a = nx_o + h$ is the impact parameter. These terms account for the extra delay from the refractive bending; they give the geometric delay between the LEO and the intersection point of the line $\theta = \pi/2$ with the co-centered sphere of radius a.

Evaluating $\Psi_A^{(S)}$ at its stationary-phase point and expressing it in terms of the thin-screen variables using Eq. (3.13-33), we obtain

$$\Psi_A^{(S)}\Big|_{\nu^*} = \frac{2^{5/2}}{3}x_o\left(\left(\frac{-h}{nx_o}\right)^{3/2} - \left(\frac{-h}{nx_o} - N\right)^{3/2}\right) + D_{\theta+\alpha_S} + ka\alpha_S \tag{3.13-36}$$

For comparison with the Fresnel phase $\Phi\left(h_S(h_{\mathrm{LG}}), h_{\mathrm{LG}}\right)$ from scalar diffraction theory applied to the thin-screen model used in Chapter 2, Eq. (2.5-1), we write that expression here with the geometric phase delay D from the phase screen to the LEO added:

$$\Phi + D = \frac{2^{5/2}}{3}x_o\left(\left(\frac{-h}{nx_o}\right)^{3/2} - \left(\frac{-h}{nx_o} - N\right)^{3/2} - N^{3/2}\right) + D\left(1 + \frac{\alpha_S^2}{2}\right) \tag{3.13-37}$$

where $h(h_{\mathrm{LG}})$ is the stationary-phase altitude of the ray at its turning point, which is a function of the thin-screen coordinate of the LEO, $h_{\mathrm{LG}} = nkr(\sin\theta - \sin\theta_o)$. In the thin-screen model discussed in Chapter 2, D was only specified as the distance from the LEO to the thin screen, which was mounted perpendicular to the LEO–GPS line. Since $D_{\theta+\alpha_S} + a\alpha_S = D_\theta\left(1 + \alpha_S^2/2 + \mathrm{O}\left[\alpha_S^3\right]\right)$, it is clear that to achieve correspondence between Eqs. (3.13-36) and (3.13-37) we should set $D = D_\theta$. Except for this arbitrariness

in D and the very small zero-point constant $N^{3/2}$ introduced to null the thin-screen phase at the critical refraction altitude, the phase profiles predicted by Eqs. (3.13-36) and (3.13-37) agree closely, differing by terms of the order $x_o\alpha_S^3$.

The second derivative of $\Psi^{(S)}$ is given by

$$\frac{\partial^2\Psi^{(S)}}{\partial\nu^2} = 2\left(\frac{1}{\sqrt{n^2x_o^2-\nu^2}} - \frac{1}{\sqrt{x_o^2-\nu^2}}\right) + \frac{1}{D_\nu} \tag{3.13-38}$$

and, when it is expanded in powers of N and evaluated at $\nu^* = nx_o + h$, it becomes

$$\left.\frac{\partial^2\Psi^{(S)}}{\partial\nu^2}\right|_{\nu=\nu^*} = \frac{1}{D_{\nu^*}}\left(1 - D_{\nu^*}\left(\frac{2N}{x_o}\right)\left(\frac{nx_o}{-2h}\right)^{3/2}\right) \tag{3.13-39}$$

We see that $\partial^2\Psi^{(S)}/\partial\nu^2 \to 0$ when $h = h^\dagger \doteq -(N^2D_\theta^2x_o/2)^{1/3}$, which is identical with the condition given in Eq. (3.13-29) with $D = D_\theta$, which was obtained from the thin-screen model, and also when $\xi_1 = \xi_2$ in Eq. (3.13-25), which is based on Mie scattering.

In the stationary-phase technique, the term $|\partial^2\Psi^{(S)}/\partial\nu^2|^{-1/2}$ appears in the expression for the amplitude of the scattered field. Its minimum, which enables the broadest range of spectral numbers around the stationary-phase point to contribute to the scattering integrals, should closely correspond to the maximum amplitude for the scattered wave and, therefore, to maximum interference with the direct wave or to maximum darkening (and brightening) while the LEO is in the shadow. Appendix D discusses darkening in terms of Airy functions by expanding about the value of ν where $\partial^2\Psi^{(S)}/\partial\nu^2 = 0$.

From Eq. (3.13-39), we note from differentiating Eq. (3.13-27) and expanding that

$$\left.\frac{\partial^2\Psi^{(S)}}{\partial\nu^2}\right|_{\nu=\nu^*} = \frac{1}{D_{\nu^*}}\left(1 - D_{\nu^*}\frac{d\alpha_S}{dh}\right) = \frac{1}{D_{\nu^*}}\zeta_{\nu^*}^{-1} \tag{3.13-40}$$

where ζ is the defocusing factor. This is the same factor that appears in scalar diffraction theory applied to the thin-screen model, or in geometric optics applied to a medium with a gradient in its refraction profile. There the second derivative of the Fresnel phase with respect to impact parameter also carries the factor ζ^{-1} [see Section 2.6, Eq. (2.6-3)].

3.13.8 Evaluating the Scattering Integrals Using the Stationary-Phase Technique

When $\theta - \theta_o < 0$ and it is sufficiently negative ($\theta_o - \theta >\sim \alpha_S + 2K_{x_o} / D_{x_o}$) so that the asymptotic forms for the Airy functions are applicable, the stationary-phase technique may yield accurate values for the scattering integrals. The total field is given by

$$\left.\begin{aligned} E_r^{(t)} &= E_r^{(S)} + E_r^{(i)} \\ E_\theta^{(t)} &= E_\theta^{(S)} + E_\theta^{(i)} \end{aligned}\right\} \tag{3.13-41}$$

From Eq. (3.13-5), we have

$$E_r^{(t)} = E_o \sqrt{\frac{\sin\theta_o}{2\pi i x_o \sin\theta}} \int_0^\infty \left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu}\right)^{1/2} e^{i\Psi^{(S)}} d\nu \tag{3.13-42}$$

When θ lies well into the shadow regime [where a *single* stationary phase is dominant, i.e., the one at ξ_2 in Eq. (3.13-25)] and, therefore, where the stationary-phase point is well isolated, then we can use the stationary-phase technique to evaluate Eq. (3.13-42). Expanding $\Psi^{(S)}$ in powers of $\nu - \nu^*$ about its stationary point, Eq. (3.13-42) becomes

$$E_r^{(t)} \doteq E_o \sqrt{\frac{\sin\theta_o \sin^3\theta_{\nu^*}}{2\pi i x_o \sin\theta \cos\theta_{\nu^*}}} e^{i\Psi^{(S)}(\nu^*)} \int_0^\infty e^{i\left((\nu-\nu^*)^2 \frac{1}{2}\left(\frac{\partial^2\Psi^{(S)}}{\partial\nu^2}\right)_{\nu^*}\right)} d\nu \tag{3.13-43}$$

Similarly, the θ component of the total field is given by

$$E_\theta^{(t)} \doteq E_o \sqrt{\frac{\sin\theta_o \sin\theta_{\nu^*} \cos\theta_{\nu^*}}{2\pi i x_o \sin\theta}} e^{i\Psi^{(S)}(\nu^*)} \int_0^\infty e^{i\left((\nu-\nu^*)^2 \frac{1}{2}\left(\frac{\partial^2\Psi^{(S)}}{\partial\nu^2}\right)_{\nu^*}\right)} d\nu \tag{3.13-44}$$

Here $\nu^* \doteq \rho_o(1 - (\xi_2 / K_{\rho_o})^2 / 2)$, where ξ_2, ρ_o, and K_{ρ_o} are given in Eq. (3.13-25). When N is sufficiently small so that polarization effects from scattering are negligible, then it follows from Eqs. (3.13-43) and (3.13-44) that

$$\frac{E_r^{(t)}(r,\theta)}{E_\theta^{(t)}(r,\theta)} \doteq \tan\left[\theta_{\nu^*}(\theta)\right] = \tan\left(\theta + \alpha_S\right) \tag{3.13-45}$$

Thus, the total field observed by the LEO, when it is well inside the shadow boundary, is characterized to first order in N by plane waves propagating in a

direction that is inclined at an angle α_S below the positive z-axis and in the plane of polarization. In this case, α_S is to be evaluated on the main branch as given by Eq. (3.13-27).

Figure 3-21 shows the profiles of the main branch and the anomalous branch of α_S versus $\theta - \theta_o$ as predicted by geometric optics. The main branch is associated with the more negative value of y providing a stationary value for $\Psi_A^{(S)}$ ($y^* = -(\xi_2)^2$); the anomalous branch is associated with the stationary-phase value nearer to $y = 0$ ($y^* = -(\xi_1)^2$). The ray systems associated with the main branch and with the anomalous branch are discussed in Chapter 2, Fig. 2-2. Although the anomalous branch causes some interference, it is a weak contributor because of defocusing, or in Mie scattering parlance, because of the relatively large magnitude of $\partial^2\Psi^{(S)}/\partial\nu^2$ in the vicinity of $y^* = -(\xi_1)^2$ (see Fig. 3-17). This weakness continues until θ nears first contact with the caustic, $\theta - \theta_o = -3\left(r_o N^2 / 2D\right)^{1/3}$, which in Fig. 3-21 is near –0.35 mrad. There $\partial^2\Psi^{(S)}/\partial\nu^2 = 0$.

If one evaluates the Fresnel integral appearing in Eqs. (3.13-43) and (3.13-44) using Eq. (3.13-40) for $(\partial^2\Psi^{(S)}/\partial\nu^2)_{\nu^*}$, and assumes that the stationary-phase point is well-isolated, dominant, and located well into the Airy function asymptotic regime ($y^* << -2$), then one obtains

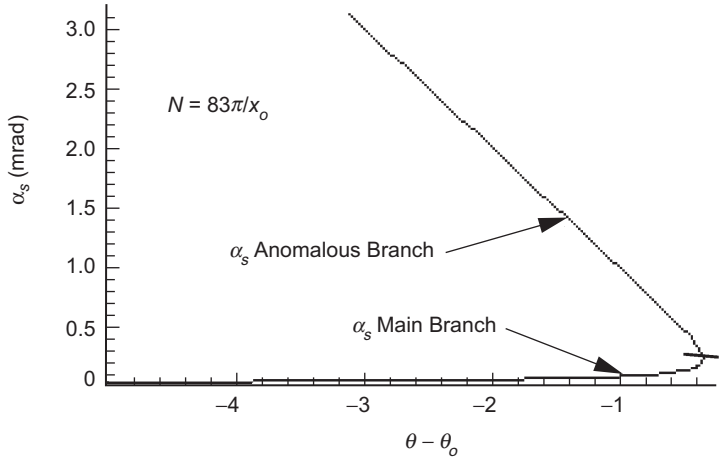

Fig. 3-21. Refractive bending angle based on geometric optics from the main and anomalous branches. The turning point is near the first contact with the caustic surface, $\partial\Psi^{(S)}/\partial\nu^2 = 0$.

$$\left.\begin{aligned} E_r^{(t)} &\doteq E_o \sin(\theta+\alpha_S)\sqrt{\frac{\sin(\theta+\alpha_S)\xi_{\nu^*}}{\sin\theta}}e^{i\Psi^{(S)}(\nu^*)} \\ E_\theta^{(t)} &\doteq E_o \cos(\theta+\alpha_S)\sqrt{\frac{\sin(\theta+\alpha_S)\xi_{\nu^*}}{\sin\theta}}e^{i\Psi^{(S)}(\nu^*)} \end{aligned}\right\} \tag{3.13-46}$$

One can use the stationary-phase technique to calculate the contribution to the scattering integrals from the minor stationary-phase point in spectral number space [see Fig. 3-17 and Eq. (3.13-25)] at $y^* = -(\xi_1)^2$. So long as this stationary-phase point exists, lies in the asymptotic regime, and is well separated from the dominant stationary-phase point so that sufficient phase accumulation occurs between stationary points and, in particular, so that the stationary-phase points are well away from inflection points [which holds, for example, when $\theta_o - \theta = 2.0, 2.25, \cdots$, in Fig. 3-19(b)], then the evaluation of the scattering integrals for this point using the stationary-phase technique may be sufficiently accurate. In this case, Eq. (3.13-46) would be modified to

$$\left.\begin{aligned} E_r^{(t)} &\doteq E_r^{(t)}\Big|_{\nu=\nu_1^*} + E_r^{(t)}\Big|_{\nu=\nu_2^*}, \quad \nu_k^* \doteq \rho_o\left(1-\left(\xi_k / K_{\rho_o}\right)^2/2\right) \\ E_\theta^{(t)} &\doteq E_\theta^{(t)}\Big|_{\nu=\nu_1^*} + E_\theta^{(t)}\Big|_{\nu=\nu_2^*} \qquad\qquad k = 1,2,\cdots \end{aligned}\right\} \tag{3.13-47}$$

On the other hand, if the phasor of the main contribution and/or the phasor of the secondary contribution have inaccuracies from application of the stationary-phase technique, then the error in the vector addition of these two contributions shown in Eq. (3.13-47) can be greatly exacerbated by these offsets. This especially holds when they interfere destructively. Moreover, we have noted that for $\theta > \theta^\dagger$ there are no stationary-phase points, so the application of the stationary-phase technique to this regime is somewhat moot. For a regime (for example, when the LEO is well into the geometric shadow) where the computations of the scattering from the farther and nearer stationary-phase points using the stationary-phase technique are sufficiently accurate, then they can be combined through vector addition to calculate the interference pattern in that regime from a transparent refracting sphere. However, we will present instead the numerical integration solutions (in spectral number space) for a range of angular positions of the LEO inside and outside of the shadow boundary. Chapter 5 discusses further the conditions for validity of geometric optics in calculating interference between rays.

3.14 Duality Between Stationary-Phase Concepts in Electrodynamics and in Geometric Optics

In summary, when asymptotic conditions apply, that is, when the LEO is sufficiently inside the shadow boundary so that asymptotic forms for the Airy functions apply, or, equivalently, when $\Omega^{(S)} \to \Omega_A$ asymptotically, then we have a kind of duality between the stationary-phase processes on one hand over spectral number space in electromagnetic theory and on the other hand in geometric optics over impact parameter space. For electromagnetic theory applied to a uniform refracting sphere in a vacuum on one hand, and for the scalar diffraction theory applied to the thin-screen proxy for the sphere on the other hand, the correspondence between these systems shown in Table 3-1 holds asymptotically at least to first order in N when stationary-phase values are assigned in each system.

3.15 Diffraction from a Large, Transparent, Refracting Sphere Using Mie Scattering Theory

3.15.1 Numerical Integrations

When "thin-atmosphere" conditions hold, a straight numerical integration of the scattering integrals in Eq. (3.13-5) performs well, although some caution is needed in dealing with the oscillatory nature of the kernels in these integrals to avoid aliasing effects and inaccuracies from under-sampling. For larger values of N, for example, such as found in a raindrop, the increasing number of episodes of high acceleration in phase with spectral number makes the numerical integration technique onerous. In this case, evaluating the scattering integrals in complex spectral number space using the Sommerfeld–Watson transformation is probably a better approach for angular positions in the geometric shadow of the refracting sphere.

Here our values of N are relatively small, and so the numerical integrations in Eq. (3.13-5) are used to obtain the scattered field $E^{(S)}(r,\theta)$ from a transparent refracting sphere; the integrations are aided by stationary-phase theory to define practical limits and limiting forms. In Eq. (3.13-5), we use the definition for $\Omega^{(S)}$ given in Eqs. (3.5-15) and (3.9-2), and we omit the Φ^+ term because it has a negligible effect on the integrals for large x_o and near-grazing conditions. Figures 3-22 and 3-23 display the results of this computation. Figure 3-22 shows the amplitude of the total field or vector sum of the incident wave and the scattered waves, $E^{(S)}(r,\theta) + E^{(i)}(r,\theta)$, at the LEO position, where $E^{(i)}(r,\theta) = \hat{\boldsymbol{x}} E_o \exp(ix\cos\theta)$. Here $\hat{\boldsymbol{x}}$ is a unit vector in the up direction;

Table 3-1. Duality between systems.

	Electrodynamics		Geometric Optics
Variational Parameter	Spectral Number		Impact Parameter
Stationary Point	ν^*	$\Leftrightarrow$	$nx_o + h^*$
Stationarity Condition	$\dfrac{\partial \Psi^{(S)}}{\partial \nu} = \dfrac{\partial\left(2\Omega_A + \Phi^-\right)}{\partial \nu} = 0$	$\Leftrightarrow$	$\dfrac{\partial}{\partial h}\left(\int_h^\infty \alpha dh + \dfrac{(h-h_{LG})^2}{2D}\right) = \dfrac{\partial \Phi}{\partial h} = 0$
Gradient	$-\left.\dfrac{\partial 2\Omega_A}{\partial \nu}\right\|_{\nu^*} = \theta_{\nu^*} - \theta$	$\Leftrightarrow$	α Bending Angle
Limb Distance	$D_{\nu^*} = \sqrt{x^2 - \nu^{*2}}$	$\Leftrightarrow$	$D = \sqrt{x^2 - n^2 x_o^2}$
Stationary Phase	$\Psi^{(S)}\Big\|_{\nu^*} = \left[2\left(\left(n^2x_o^2 - \nu^2\right)^{1/2} - \left(x_o^2 - \nu^2\right)^{1/2}\right) + D_\nu\right]_{\nu^*}$	$\Leftrightarrow$	$\Phi\left(h^*(h_{LG}), h_{LG}\right) + D$
Reflection Coefficient	$\left\|\rho^{(0)}\right\|_{\nu^*} = \left.\dfrac{n\left(x_o^2 - \nu^2\right)^{1/2} - \left(n^2x_o^2 - \nu^2\right)^{1/2}}{n\left(x_o^2 - \nu^2\right)^{1/2} + \left(n^2x_o^2 - \nu^2\right)^{1/2}}\right\|_{\nu^*}$	$\Leftrightarrow$	$\dfrac{n\cos\varphi - \cos\varphi'}{n\cos\varphi + \cos\varphi'} = R_\parallel$
Defocusing ζ:	$\left(D_\nu \dfrac{\partial^2 \Psi^{(S)}}{\partial \nu^2}\right)^{-1}_{\nu^*} = \left(1 + D_\nu \dfrac{\partial^2 2\Omega_A}{\partial \nu^2}\right)^{-1}_{\nu^*}$	$\Leftrightarrow$	$\left(1 - D\dfrac{d\alpha}{dh}\right)^{-1}_{h^*} = \left(D\dfrac{\partial^2 \Phi}{\partial h^2}\right)^{-1}_{h^*}$
Caustic Contact $\zeta \to \infty$	$\left(1 + D_\nu \dfrac{\partial^2 2\Omega_A}{\partial \nu^2}\right)_{\nu^*} \to 0$	$\Leftrightarrow$	$\left(1 - D\dfrac{d\alpha}{dh}\right)_{h^*} \to 0$

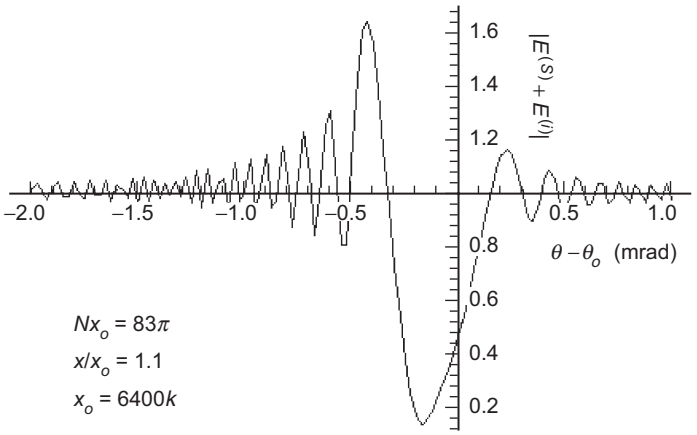

Fig. 3-22. Amplitude of the total field at the LEO (incident wave plus scattered wave from a refracting sphere); based on a numerical integration using Mie scattering theory.

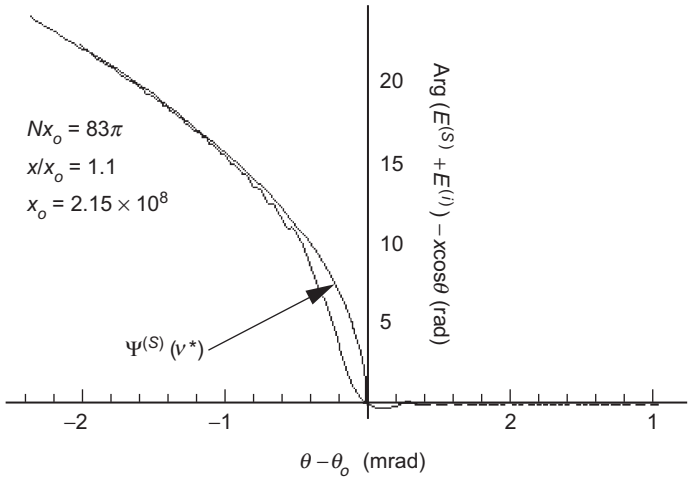

Fig. 3-23. Phase of the total field at the LEO (incident plus scattered wave from a refracting sphere) minus $x\cos\theta$; from Mie scattering theory.

it is the in-plane vector perpendicular to $\hat{z}$, which points along the direction of propagation (see Fig. 3-1). Similarly, $\boldsymbol{E}^{(S)} = \hat{\boldsymbol{r}}E_r^{(S)} + \hat{\boldsymbol{\theta}}E_\theta^{(S)}$. The refractivity used in this figure is $N = 83\pi / x_o$.

The phase of the total field is shown in Fig. 3-23 for $N = 83\pi / x_o$. The phase has been stopped, i.e., multiplied by the geometric phase delay phasor $\exp(-ix\cos\theta)$ to eliminate the phase rate from changing geometric delay with

changing θ; but for $\theta < \theta_o$, there is an additional phase rate because of the "lens" effect due to the refracting sphere. These Mie scattering solutions in Figs. 3-22 and 3-23 are in close agreement with the results from the thin-screen/scalar diffraction model for a refracting sphere, as presented in Chapter 2, Figs. 2-17 through 2-20 for $N = 83\pi / x_o$.

3.15.2 Comparison with Thin-Screen/Scalar Diffraction Results

Figures 3-24 and 3-25 compare Mie scattering results with the scalar diffraction thin-screen results for $N = 83\pi / x_o$ and $N = 830\pi / x_o$, respectively. In the thin-screen computation, D has been set equal to D_θ. The noise in these figures because of limited computational precision is about 0.001 in amplitude and about 0.01 rad in phase. The amplitude profiles [of $E^{(S)}(r,\theta) + E^{(i)}(r,\theta)$] are in virtual agreement with Mie scattering producing perhaps 1 to 2 percent deeper fringes. We see the very deep shadow zone but no rainbow effects; they are to be seen faintly at larger values of $\theta_o - \theta$, near –6 mrad for Fig. 3-24(a) and near –18 mrad for Fig. 3-25(a).

Figures 3-24(b) and 3-24(c) and Figs. 3-25(b) and 3-25(c) show the differences between these two systems in amplitude and in phase for $Nx_o = 83\pi$ and $Nx_o = 830\pi$, respectively. These figures show good agreement except in the midst of the shadow zone, which would be the most likely region for differences to arise because of the super-refracting conditions here. Geometric optics, which provides the basis for the thin-screen phase profile, denies the existence of rays in the shadow zone, but electromagnetic waves do pass through this region, albeit faintly, having already interacted with the bound charges of the dielectric medium in the sphere. Also, this is the region where no significant stationary-phase points exist in either spectral number space in electromagnetic theory or in impact parameter space in geometric optics, which might accentuate whatever differences exist between the two systems. Moreover, the main contributions to the Mie scattering integrals in this region partly come from spectral numbers where the asymptotic forms for the Airy functions do not apply. It would be surprising if these two systems did not disagree to some extent deep in the shadow zone.

The phase differences accumulate primarily in the shadow zone. The "bow-tie" pattern in the amplitude difference shown in Fig. 3-25(b) is a result of phase runoff between the two systems. A discontinuity in refractivity provides a relatively severe test of the fidelity of the thin-screen model. "Smoother" discontinuities, such as in the gradient of the refractivity, are less severe. This is discussed further in Chapter 5.

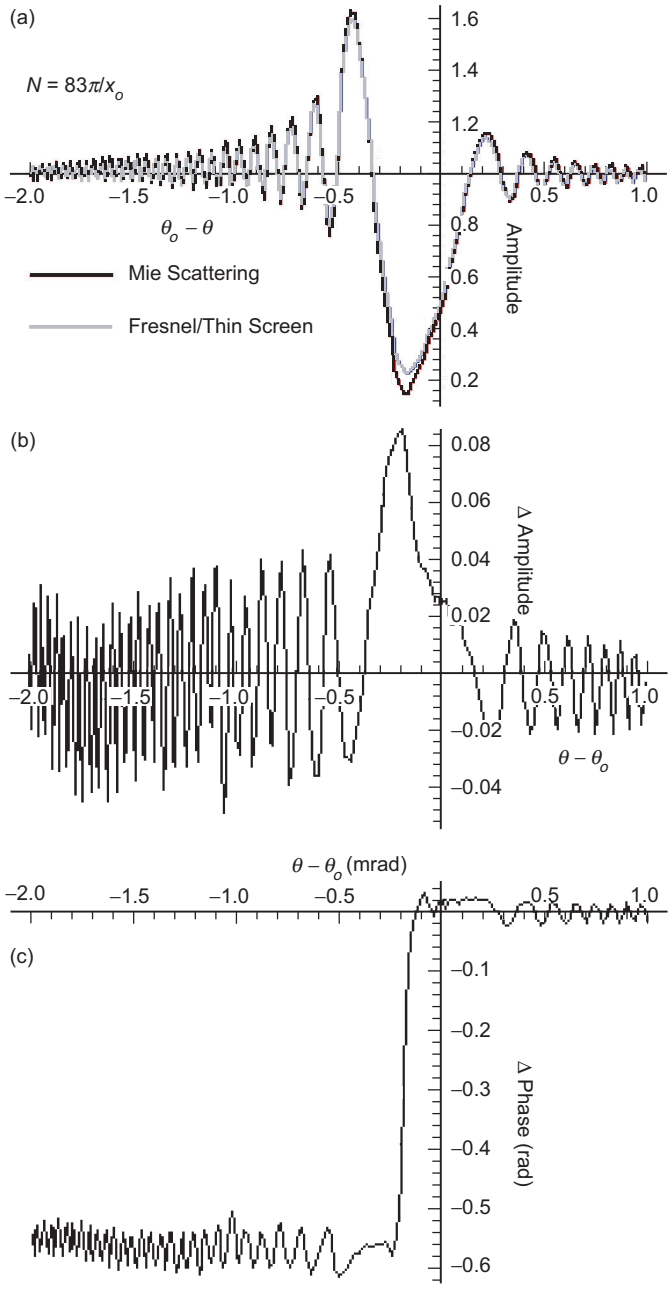

Fig. 3-24. Comparison of the total field at the LEO based on Mie scattering with a scalar diffraction/thin-screen model: (a) amplitude, (b) amplitude difference, and (c) phase difference. $N = 83\pi/x_o$, $x = 1.1x_o$, $x_o = kr_o \approx 2.15 \times 10^8$, **and** $r_o = 6478$ km.

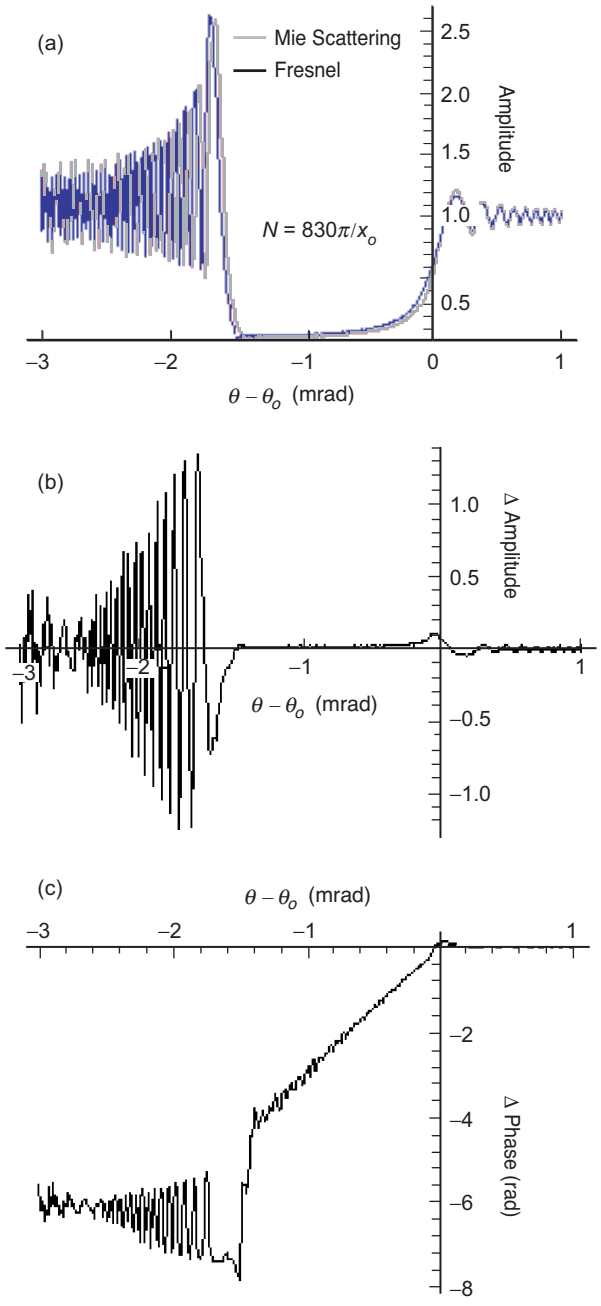

Fig. 3-25. Mie Scattering versus scalar diffraction/thin-screen model: (a) amplitudes, (b) amplitude difference, and (c) phase difference. The same conditions as used in Fig. 3-24, except $N = 830\,\pi\,/x_o$.

The thin-screen methodology is much less complex and computationally less intensive than Mie scattering theory; the ratio of the computational times to obtain the equivalent results shown in Figs. 3-24 and 3-25 is about 1:3.

3.16 Looking for Rainbows

3.16.1 Rainbow Effects on the Refracting Sphere

We can evaluate the scattered field on the surface of the sphere, $r = r_o$, to see if the $j = 2$ rainbow effects appear in the neighborhood of $\theta =\sim \theta_o - 2(6N)^{1/2}$, $\theta_o = \pi/2$. However, in this case where $r = r_o$, we cannot use the asymptotic forms for the Hankel functions given in Eq. (3.10-1) for the LEO radial position because the argument Y of the Airy functions will be small or even positive when $\nu \approx x_o$. Instead, we may use the Airy function forms for the spherical Hankel functions given in Eq. (3.8-9) with the understanding that when $\nu < x_o - 3K$ (for better than 0.1 percent accuracy), we can continue to use the asymptotic forms given by Eq. (3.10-1). If this form in Eq. (3.8-9) is substituted into the scattering series given in Eq. (3.12-3) for $E_r^{(S)}(r_o, \theta)$ and a conversion from the sums into integrals with y as the independent variable is made, one obtains the expressions for the scattering for an observer located on the surface of the refracting sphere. Retaining only the relevant terms, we obtain

$$\left.\begin{aligned} E_r^{(S)} &= \frac{E_o}{\sqrt{\sin\theta}} \exp\left(ix_o\tilde{\theta}\right) \int_{-\infty}^{\infty} S_l \left(\frac{\nu}{nx_o}\right)^{\frac{3}{2}} \mathrm{Ai}[\hat{y}] \exp\left(i\hat{y}K_{\rho_o}\tilde{\theta}\right) d\hat{y} \\ \frac{\nu}{\rho_o} &= 1 + \frac{\hat{y}}{2K_{\rho_o}^2} + \mathrm{O}\left[\frac{\hat{y}}{2K_{\rho_o}^2}\right]^2, \quad \tilde{\theta} = \frac{\pi}{2} - \theta, \quad \rho_o = nx_o = nkr_o \end{aligned}\right\} \tag{3.16-1}$$

$E_\theta^{(S)}$ will be very small because $\theta \approx \pi/2$. The scattering term S_l is given by Eqs. (3.5-11) and (3.5-15). It can be replaced by higher-degree scattering coefficients if one wishes to do so. Of particular interest is the degree $j = 2$ term, which provides the scattered field for the primary rainbow.

Figure 3-26 shows the amplitude of the total field (aggregate scattered plus incident) for the two cases: $N = 83\pi/x_o$ and $N = 830\pi/x_o$. These are obtained from a numerical integration of Eq. (3.16-1). Even for positions located on the surface of the scattering sphere a shadow zone develops, but it has a different shape and location than shown in Figs. 3-24 and 3-25. The appearance of ringing in Fig. 3-26(b) in the Fresnel fringe amplitude near $\theta - \theta_o \approx -20$ mrad presages the growth of rainbow resonances at certain

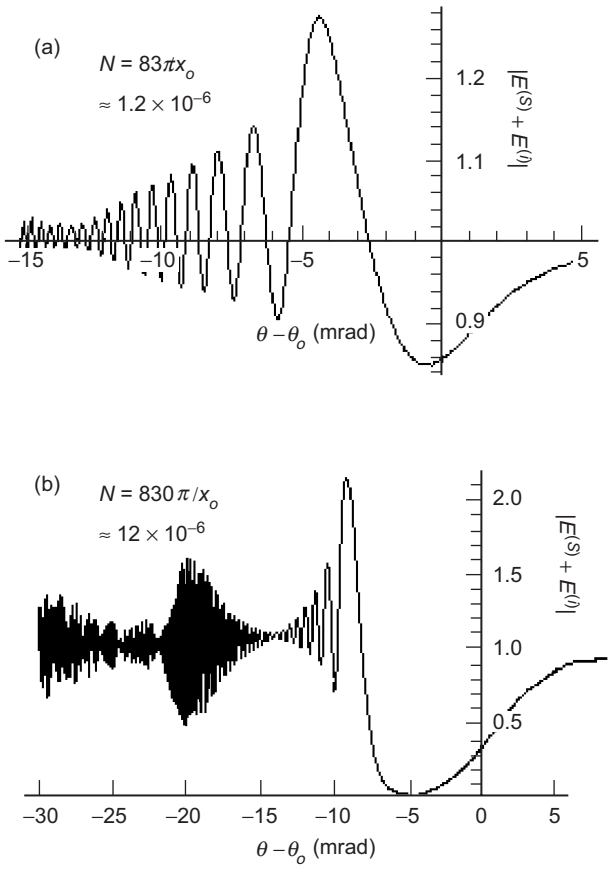

Fig. 3-26. Amplitude of the total field from a refracting sphere for an observer located on its surface: (a) $N = 83\pi / x_o$ **and** (b) $N = 830\,\pi / x_o$.

angular positions that would result from using larger values of the refractivity. Some under-sampling effects are evident in Fig. 3-26(b), particularly for $\theta - \theta_o < -15$ mrad.

One can apply the stationary-phase technique to Eq. (3.16-1). It can be shown that stationary-phase points exist near $\hat{y} = 0$, and when $\pi / 2 - \theta > \sim 8N$, a point also exists at

$$\hat{y}^* \doteq -\frac{K_{\rho_o}^2}{4}\left(\tilde{\theta} + \sqrt{\tilde{\theta}^2 - 8N}\right)^2 \tag{3.16-2}$$

Figure 3-27 shows the amplitude of the scattered field on the refracting sphere for only the degree $j = 2$ term, i.e., $\left|\boldsymbol{E}^{(S_2)}(r_o, \theta)\right|$. The scattering

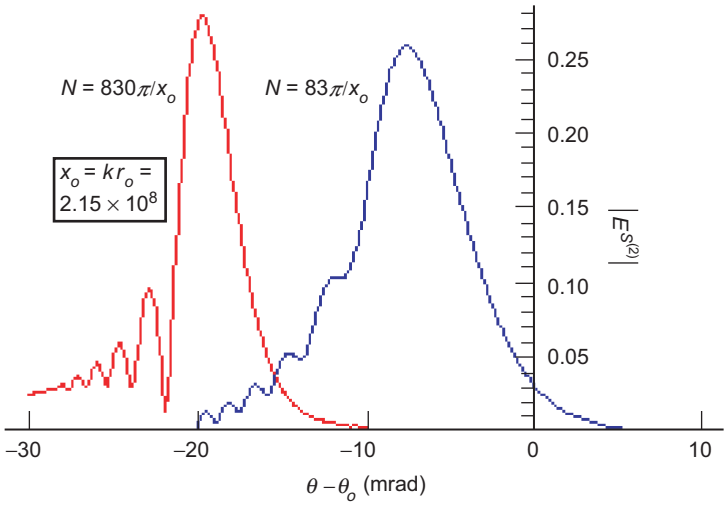

Fig. 3-27. Amplitude of the degree $j = 2$ scattered wave for an observer located on the surface of the refracting sphere.

coefficient $S_l^{(2)}$ replaces S_l in Eq. (3.16-1); it is obtained from Eqs. (3.5-11) and (3.9-3) and it is given by

$$S_l^{(2)} = 2n\frac{\mathcal{W}_l^+}{(\mathcal{W}_l)^3} \tag{3.16-3}$$

The locations of the peaks of the distributions in Fig. 3-27 agree moderately well with the degree $j = 2$ scattering angles predicted by geometric optics in Eq. (3.4-9). For the case $N = 830\pi / x_o$, the peak at –20 mrad falls in a relatively quiescent fringe region in Fig. 3-26(b); therefore, the interference of the $j = 2$ scattered wave with the aggregate scattered wave is quite noticeable in the form of expanded fringe amplitude in that neighborhood. On the other hand, for the case $N = 83\pi / x_o$, its peak lies near the location of the maximum power from the aggregate scattering in Fig. 3-26(a); its effect is scarcely noticeable. The angular location of the peaks in these amplitude distributions in Fig. 3-27 correspond to the scattering angle of the rainbow in geometric optics, which results from a ray undergoing a single internal reflection.

3.16.2 Rainbow Effects at LEO Distances

Figure 3-28 shows the amplitude of the scattered field for only the $j = 2$ term at the location of the LEO, namely, at $r = 1.1r_o$ for the case $N = 830\pi / x_o$. Here the peak amplitude is diminished by a factor of 6 relative to that shown in

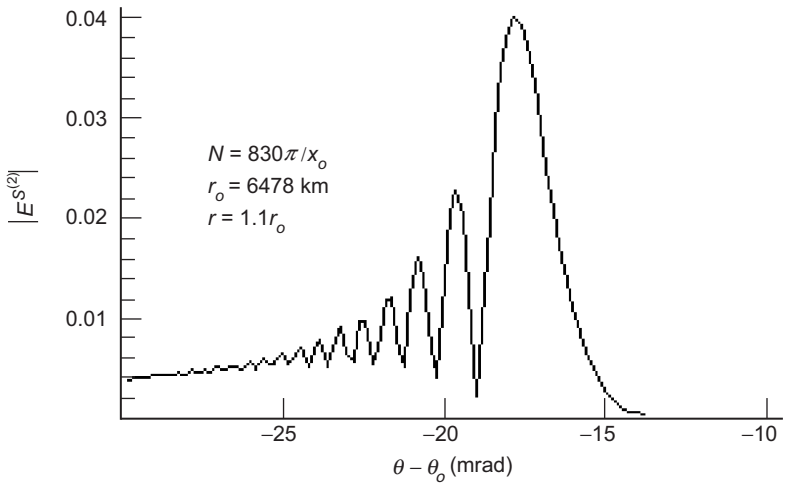

Fig. 3-28. Amplitude of the degree $j = 2$ scattered wave at the LEO.

Fig. 3-27 (for $N = 83\pi / x_o$, it is diminished by a factor of 25). The location of the peak amplitude in Fig. 3-28 agrees closely with the prediction from geometric optics given in Eq. (3.4-9), but the contribution of the degree $j = 2$ scattered wave to the aggregate scattering shown in Fig. 3-26(a) is very faint (a maximum of ~4 percent). For this case, $Nr_o << (\lambda D_S)^{1/2}$, the radius of the first Fresnel zone, which is about 750 m. In contrast, for the observer located on the sphere (to which Fig. 3-27 applies), the effective "Fresnel radius" is only about 20 m at the angular location where the caustic is dominant (~–20 mrad); which is comparable to the altitude of the turning point of the ray at the caustic point.

In terms of stationary-phase concepts, the difference in amplitude peaks shown in Figs. 3-27 and 3-28 can be explained in terms of the magnitude of the quadratic term $\partial^2 \Psi^{(S^{(2)})} / \partial y^2$ for the two scenarios. The location of the stationary-phase points in spectral number space for the $j = 2$ scattering coefficient corresponds for the $r = 1.1 r_o$ case to a situation where $\partial^2 \Psi^{(S^{(2)})} / \partial y^2$ is relatively large, which restricts the breadth of spectral number contributions to the scattering integrals (see Fig. 3-18). When the observer is located on the refracting sphere $r = r_o$, it can be shown that $\partial^2 \Psi^{(S^{(2)})} / \partial y^2$ is relatively smaller.

3.16.3 Assessing Rainbow Effects Using the Third-Order Stationary-Phase Technique

One can use third-order stationary-phase theory to solve for the angular location of the peak amplitude for the degree $j = 2$ scattered wave, such as that shown in Fig. 3-28, and also its magnitude. Using Eqs. (3.8-12), (3.8-14), (3.9-3), and (3.16-3), one obtains the asymptotic form of the scattering coefficient for negative values of $\hat{y}$ and y:

$$\left.\begin{aligned} S_l^{(2)} &\to -\left|S_A^{(2)}\right| \exp\left(i\left(4\hat{\mathrm{X}} - 2\mathrm{X}\right)\right), \quad y < 0 \\ \left|S_A^{(2)}\right| &= 2\sqrt{y\hat{y}}\, \frac{\left(-\hat{y}\right)^{1/2} - \left(-y\right)^{1/2}}{\left(\left(-\hat{y}\right)^{1/2} + \left(-y\right)^{1/2}\right)^3} \end{aligned}\right\} \tag{3.16-4}$$

where $\hat{\mathrm{X}}$, X, $\hat{y}$, and y are defined in Eq. (3.8-14). The degree $j = 2$ scattering phase $\Psi^{(2)}$ in the scattering integrals of Eqs. (3.12-5) and (3.12-6) has the asymptotic form for negative y that is given by

$$\Psi^{(2)} = 4\hat{\mathrm{X}} - 2\mathrm{X} + \Phi^{-} \tag{3.16-5}$$

To locate the peak amplitude, we set $\partial^2 \Psi^{(2)} / \partial y^2 = 0$, which yields

$$2\left(-\hat{y}\right)^{-1/2} - (-y)^{-1/2} + \frac{K_{x_o}^2}{D_v} = 0 \tag{3.16-6}$$

This condition defines the point $y^{\dagger}$, and thus the neighborhood in spectral number space providing the maximum contribution to the degree $j = 2$ scattered wave. The condition in Eq. (3.16-6) yields the single relevant root

$$y^{\dagger} \doteq -\frac{Nx_o}{3K_{x_o}}\left[1 - \frac{8}{3}\left(\frac{Nx_o}{3K_{x_o}}\right)^{1/2} \frac{K_{x_o}^2}{D_{x_o}}\right] \tag{3.16-7}$$

which for the case $N = 830\pi / x_o$ shown in Fig. 3-28 yields a value of $y^{\dagger} = -1.81$, just barely outside of the asymptotic range. From Eqs. (3.13-33) and (3.16-7), it follows that the altitude of the turning point of the corresponding ray is given by $h_S \doteq K_{x_o} y^{\dagger} \doteq -Nx_o / 3$, which agrees exactly with the prediction of geometric optics made in Eq. (3.4-10) and shown in Fig. 3-8.

From Eq. (3.16-5), it follows that

$$\frac{\partial \Psi^{(2)}}{\partial y} \doteq -4\left(-\hat{y}\right)^{1/2} + 2(-y)^{1/2} + K_{x_o}\left(\theta_\nu - \theta\right) \tag{3.16-8}$$

If we also adjust θ so that $\partial \Psi^{(2)} / \partial y = 0$ at the flexure point $y^\dagger$ given in Eq. (3.16-7), we obtain a stationary-phase refractive bending angle $\alpha^\dagger$ that corresponds closely to the direction of maximum scattering for the degree j =2 wave. From Eqs. (3.16-7) and (3.16-8), $\alpha^\dagger$ is given by

$$\alpha^\dagger = 2\sqrt{6N} + \mathrm{O}\left[N^{3/2}\right] \tag{3.16-9}$$

Using the thin-screen approximation (see Fig. 3-16) $\theta - \theta_o \doteq h_S / D_S - \alpha$, $h_S^\dagger / D_S \doteq K_{x_o} y^\dagger / D_{x_o}$, which is very small compared to $\alpha^\dagger$, it follows that

$$\theta^\dagger - \theta_o \doteq -\alpha^\dagger = -2\sqrt{6N} \tag{3.16-10}$$

which for the case $N = 830\pi / x_o$ yields a value of –17.1 mrad; this is near the location of the peak in Fig. 3-28 (–17.9 mrad), and it agrees exactly (to first order in N) with the stationary scattering angle predicted by geometric optics in Eq. (3.4-9) for the degree $j = 2$ scattered ray.

To estimate the magnitude of the peak in Fig. 3-28 and to refine the estimate of its angular direction given by Eq. (3.16-10), we can use the third-order stationary-phase technique on the scattering integral, which is discussed in Appendix D in connection with its application to the scalar diffraction integral for the thin screen. The treatments are formally identical for the two systems, i.e., for the Mie scattering integrals and for the scalar diffraction integral. So, without derivation, we use those results given in Appendix D on the Mie scattering integrals. From Eqs. (3.12-5) and (3.12-6) with S_l replaced by $S_l^{(2)}$, upon changing the integration variable from ν to y, and upon expanding $\Psi^{(2)}$ through third degree in y about the point of flexure $y^\dagger$ where $\partial^2 \Psi^{(2)} / \partial y^2 = 0$, the third-order stationary-phase technique yields for the amplitude of the degree $j = 2$ scattered wave

$$\left.\begin{aligned} \left|\boldsymbol{E}^{\boldsymbol{S}^{(2)}}\right| &\doteq 2E_o K_{x_o} \sqrt{\frac{2\pi}{D_{x_o}}} \left|S_A^{(2)}\right|_{y^\dagger} \left(2\left|\frac{\partial^3 \Psi^{(2)}}{\partial y^3}\right|_{y^\dagger}^{-1}\right)^{1/3} \mathrm{Ai}[z], \\ z &= \left(\frac{\partial \Psi^{(2)}}{\partial y}\right)_{y^\dagger} \left(2\left|\frac{\partial^3 \Psi^{(2)}}{\partial y^3}\right|_{y^\dagger}^{-1}\right)^{1/3} \end{aligned}\right\} \tag{3.16-11}$$

The Airy function reaches a local maximum $\mathrm{Ai}[z] = 0.5356$ at $z = -1.02$ (see Appendix D, Fig. D-1). From Eq. (3.16-8), if we adopt a value for θ so that $z = -1.02$ in Eq. (3.16-11), we should obtain the peak amplitude for the scattered wave. From Eqs. (3.16-4) and 3.16-6), it follows that $\left|S_A^{(2)}\right|_{y^\dagger} \doteq 4/27$, independent of the value of *N*. When Eqs. (3.16-4) through (3.16-8) are used to evaluate Eq. (3.16-11), one obtains a peak amplitude and location that are given by

$$\left.\begin{aligned} \left|\boldsymbol{E}_{\mathrm{MAX}}^{S^{(2)}}\right| &= 0.402 E_o \sqrt{N K_{x_o} \tan\theta_o}, \\ \theta_{\mathrm{MAX}} - \theta_o &= -2\sqrt{6N} - \frac{0.715}{N^{1/2} K_{x_o}^2} - \frac{1}{3} N \tan\theta_o \end{aligned}\right\} \qquad (3.16\text{-}12)$$

Equation (3.16-12) predicts a peak amplitude, $0.045E_o$, which agrees well with the peak (0.04) shown in Fig. 3-28. It predicts a location for the peak at $\theta_{\mathrm{MAX}} - \theta_o = -17.97$ mrad; the actual location of the peak in Fig. 3-28 is –17.9. The offset between the predictions of Mie scattering and geometric optics, which for the case shown in Fig. 3-28 is about 0.9 mrad, was first discovered by Airy in his investigation of rainbow caustics. He used essentially a third-order ray-optics approach, which results in a formalism that is similar to that described[4] in Appendix-D. Equation (3.16-12) predicts that the peak amplitude of the degree $j = 2$ scattered wave at the location of the LEO grows as $N^{1/2}$. For the case of $N = 83\pi / x_o$, this peak is located at about $\theta - \theta_o = -6$ mrad with an amplitude at the LEO that is about 1 percent of the total amplitude.

Incidentally, determination of the peak amplitude and location of the higher-degree scattered waves follows in an analogous way to that used for the degree $j = 2$ scattering. However, the stationary-phase points for the higher degrees are even closer to zero (see Fig. 3-18) and, therefore, recede further from the region of validity of the asymptotic forms for the Airy functions. With this caveat in mind, it is easily shown using the third-order stationary-phase

---

[4] In geometric optics, the scattering angles for the rainbow caustics are independent of the size of the scattering sphere [see Eq. (3.4-9)]. The third-order Mie scattering results in Eq. (3.16-11) show a dependence on the radius-to-wavelength ratio. This theory can be used to calculate the sensitivity of the scattering angle to a change in wavelength of the incident wave for a dispersive refractive medium. For water in the visual spectrum, Mie third-order theory (retaining higher powers of *N* to maintain sufficient accuracy) can be used to show that the scattering-angle sensitivity to wavelength change is greater for large raindrops than it is for small drops. This can be used to explain the brilliance and sharpness of the color separation in rainbows observed from large raindrops and the almost washed out colors from a fine mist. The size of a raindrop tends to increase during its descent.

technique and the negative argument asymptotic forms for the Airy functions that the peak amplitude and its location are given by

$$\left.\begin{aligned} \left|\boldsymbol{E}_{\mathrm{MAX}}^{S_j}\right| &= 0.536E_o\left[\frac{4j^2}{j^2-1}\right]^{1/3}\frac{4\left(j^2-1\right)j^{j-1}}{(j+1)^{2j}}\sqrt{\frac{2\pi NK_{x_o}\tan\theta_o}{3}}, \\ \theta_{\mathrm{MAX}}-\theta_o &= -N\frac{\tan\theta_o}{j^2-1}-2\sqrt{2\left(j^2-1\right)N} \\ &\quad -\left(\frac{1.47}{N^{1/2}K_{x_o}^2}\right)\left(\frac{j^2-1}{12j^2}\right)^{1/3}\left(\frac{j^2-1}{2}\right)^{1/2} \end{aligned}\right\} \quad (3.16\text{-}13)$$

This equation shows that the peak amplitude of the rainbow caustics attenuates rapidly with increasing degree.

It is remarkable how well Eqs. (3.16-6) through (3.16-12) predict the angular location and peak amplitude of the degree $j = 2$ scattered wave, considering the borderline validity of the asymptotic assumption used to derive them.

Figure 3-29 shows the Mie scattering and Fresnel diffraction extensions of Fig. 3-25(a) to the neighborhood of $\theta = \theta_{\mathrm{MAX}}$, or from –5 mrad out to –30 mrad. The expanded Mie scattering fringe amplitude near $\theta = \theta_{\mathrm{MAX}} = \theta_o - 18$ mrad and its banding between –19 and –25 mrad reflects significant interference of the degree $j = 2$ scattered wave near its peak amplitude with itself for $\theta < \theta_{\mathrm{MAX}}$ [see Fig. 3-5(b)]. This banding is the analog

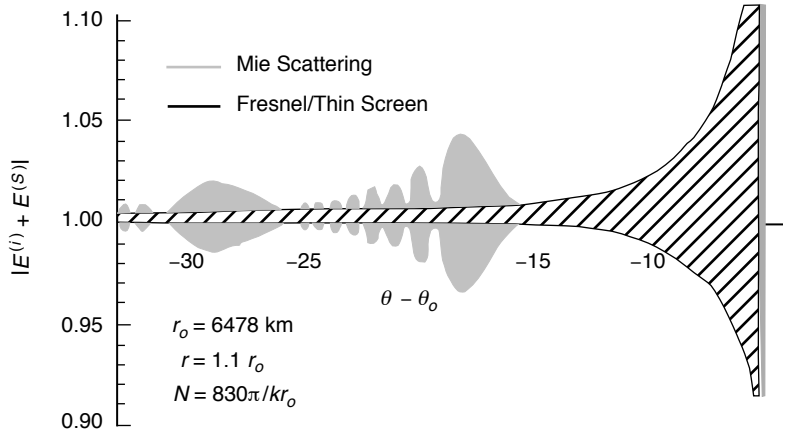

Fig. 3-29. Extension of the amplitude profile shown in Fig. 3-25(a) into the rainbow caustic region. Degrees j = 2, 3 **rainbow features occur at** $-25 \le \theta - \theta_o \le -15$ **and** $-35 \le \theta - \theta_o \le -25$**, respectively. See Fig. 3-28 for the** j **= 2 wave.**

of the supernumerary arcs in rainbows. The degree $j = 3$ scattered wave, which peaks at 0.007, becomes visible near $\theta - \theta_o = -29$ mrad. The noise in these figures due to computational precision is about 0.001.

In summary, except for very small discontinuities in refractivity, one should account for rainbow effects at least in the neighborhoods of their maximum amplitude. With very precise and clean data, it may be possible to see $j = 2$ rainbow effects below a super-refracting marine layer. The thin-screen model will not show these effects without further tinkering with the thin-screen phase profile per the suggestion outlined earlier in Footnote 2. On the other hand, in the vicinity of shadow zones and first contact with the caustic associated with the aggregate scattering, rainbow effects tend to be rather feeble, at least for the limited range of discontinuities in refractivity that has been considered here. For boundary layer studies, shadow zones, caustics, and the concomitant multipath effects in those regions will be the primary features of interest; rainbows should play a minor role, especially with softer discontinuities. However, for refractivity discontinuities 4 to 5 orders of magnitude larger than those considered here, for example, a raindrop, Eq. (3.16-12) shows that the peak amplitudes of these rainbow features will be 2 to 3 orders of magnitude larger. They will beam like searchlights at their peaks.

3.17 Limiting Cases

3.17.1 The Perfectly Reflecting Sphere

When $N \to \infty$, considerable simplification in the scattering coefficients results. In this case, the distinction between the "electric" and "magnetic" coefficients is maintained; but, we see from Eq. (3.5-11) that the only non-zero scattering coefficients are those for the zeroth-degree scattered waves, ${}^{\mathrm{e}}b_l^{(0)}$ and ${}^{\mathrm{m}}b_l^{(0)}$, which are the scattering coefficients for the external reflected wave. The electric coefficients reduce to

$$\left.\begin{aligned} {}^{\mathrm{e}}b_l^{(0)} &= -\frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)} \left(1 + \frac{\xi_l'^{-}}{\xi_l'^{+}}\right) \to -\frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)} \left(1 + \frac{\mathrm{Ai}'[y] + i\,\mathrm{Bi}'[y]}{\mathrm{Ai}'[y] - i\,\mathrm{Bi}'[y]}\right), \\ y &= \nu^{2/3} \zeta\left[\nu / x_o\right], \quad \nu = l + 1/2 \end{aligned}\right\} \tag{3.17-1}$$

where y and $\zeta\left[\nu / x_o\right]$ are defined in Eqs. (3.8-3) and (3.8-4). The magnetic coefficients reduce to

$$ {}^{m}b_l^{(0)} = -i^{l-1}\frac{2l+1}{l(l+1)}\left(\frac{\mathrm{Ai}[y]}{\mathrm{Ai}[y]-i\,\mathrm{Bi}[y]}\right) \tag{3.17-2} $$

However, it can easily be shown from Eq. (3.2-2) that the contributions from the magnetic coefficients to the scattered wave arise only in the $E_\theta^{(s)}$ component where they are multiplied by the factor $\mathrm{P}_l^1(\cos\theta)$; on the other hand, the electric coefficients are multiplied by $d\,\mathrm{P}_l^1/d\theta$. Because $\mathrm{P}_l^{1\prime}/\mathrm{P}_l^1 \sim x_o$ for ν in the vicinity of x_o, the relative contribution from the magnetic coefficients is diminished by $\mathrm{O}\left[x_o^{-1}\right]$ compared to the contribution from the electric coefficient. So, for in-plane observations, we ignore the magnetic case and drop the superscript "e" on the electric scattering coefficients.

The problem of computing the diffraction from a reflecting sphere was first worked in its entirety by G. N. Watson [3] using the Sommerfeld–Watson summation technique discussed earlier, which uses the poles of $\left(\mathrm{Ai}'[y]-i\,\mathrm{Bi}'[y]\right)$ in the upper complex plane to evaluate the scattering series.

Here we look at the problem from a stationary-phase point of view. Now $\left|\left(\mathrm{Ai}'[y]+i\,\mathrm{Bi}'[y]\right)/\left(\mathrm{Ai}'[y]-i\,\mathrm{Bi}'[y]\right)\right| = 1$ for all real values of y. It follows from Eq. (3.9-3) that the phase $\Omega^{(0)}$ for the zeroth-degree scattered wave is given by

$$ \left.\begin{aligned} b_l^{(0)} &= \frac{1}{2}i^{l-1}\frac{2l+1}{l(l+1)}\left(e^{i2\Omega^{(0)}}-1\right), \\ \Omega^{(0)}(y) &= \tan^{-1}\left(\frac{\mathrm{Bi}'[y]}{\mathrm{Ai}'[y]}\right)+\frac{\pi}{2} \end{aligned}\right\} \tag{3.17-3} $$

The phase in Eq. (3.13-5) for the total field, $\Psi^{(0)} = 2\Omega^{(0)} + \Phi^{-}$, can be written in the form

$$ \Psi^{(0)} = 2\Omega^{(0)} + D_\nu + \nu(\theta_\nu - \theta) \tag{3.17-4} $$

Setting $\partial\Psi^{(0)}/\partial\nu = 0$ to obtain the stationary-phase points in ν yields

$$ \left.\begin{aligned} 2\frac{\partial\Omega^{(0)}}{\partial\nu}+\frac{\partial\Phi^{-}}{\partial\nu} &\doteq \frac{-2y}{\pi\,\mathrm{M}^2[y]}\frac{\partial y}{\partial\nu}+\theta_\nu-\theta = 0 \\ \mathrm{M}^2[y] &= \mathrm{Ai}'^2[y]+\mathrm{Bi}'^2[y] \end{aligned}\right\} \tag{3.17-5} $$

When $\nu^* <\sim x_o - 2K_{x_o}$, we can use the asymptotic forms for the Airy functions. Here the stationary-phase condition in Eq. (3.17-5) can be rewritten in alternate forms. From Eqs. (3.8-3) through (3.8-5), we have for this condition

$$\left(2\frac{\partial\Omega^{(0)}}{\partial\nu}+\frac{\partial\Phi^{-}}{\partial\nu}\right)_{\nu^*} \doteq \left(-2\frac{\partial}{\partial\nu}\left(\frac{2}{3}(-y)^{3/2}\right)+\theta_\nu-\theta\right)_{\nu^*} \qquad (3.17\text{-}6a)$$
$$= -2\cos^{-1}\left(\nu^*/x_o\right)+\theta_{\nu^*}-\theta=0,\ \ \nu^* < x_o - 2K_{x_o}$$

or

$$y^* \doteq -\frac{1}{4}K_{x_o}^2\left(\theta-\theta_{\nu^*}\right)^2 \qquad (3.17\text{-}6b)$$

Noting that $|\nu^* - x_o|/x_o << 1$, we can expand these expressions about $\nu^* = x_o$ to obtain

$$\nu^* = x_o\left(1-\frac{1}{8}(\theta-\theta_o)^2\right)\left(1+\mathrm{O}\left[(\theta-\theta_o)^2\right]\right) \qquad (3.17\text{-}7a)$$

or

$$y^* \doteq -\frac{1}{4}\left(K_{x_o}^2(\theta-\theta_o)^2\right)\left(1+\mathrm{O}\left[\theta-\theta_o\right]\right) \qquad (3.17\text{-}7b)$$

It follows from Eq. (3.17-6b) for the perfectly reflecting sphere that diffraction effects as a function of angular position $\theta-\theta_o$ of the LEO relative to the geometric shadow boundary scale as $K_{x_o}^{-1}$. Providing that x_o is sufficiently large so that the Airy function asymptotic forms for the Hankel functions are valid, we need only solve the scattering integrals once and then scale the result by $K_{x_o}^{-1}$ in both amplitude and in the angular coordinate to obtain results for an arbitrary ratio r_o/λ.

Figure 3-30 shows $2\partial\Omega^{(0)}/\partial y$ in the vicinity of $y=0$. It has a global minimum of -0.964 at $y=+0.630$. Stationary-phase points occur at the intersections of the $2\partial\Omega^{(0)}/\partial y$ curve with the curve $\partial\Phi^{-}/\partial y \doteq K_{x_o}\left(\theta-\theta_o\right)$ $=$ constant. For -0.964 rad $\le K_{x_o}\left(\theta-\theta_o\right)<0$, two stationary-phase points exist, which will result in pronounced interference effects. For the scaling applicable here, $K_{x_o}\approx 475$; therefore, the range of $\left(\theta-\theta_o\right)$ values over which interference occurs is only a few milliradians; the elapsed time would be a few seconds. We see that these effects will attenuate rapidly for $K_{x_o}\left(\theta-\theta_o\right)<-0.964$ rad. For $\theta-\theta_o>0$, only one stationary-phase point exists but diffraction effects will persist for awhile as a result of the interference between the convolution integrals involving the phasors $\exp\left(i\Psi^{(0)}\right)$ and $\exp(i\Phi^{-})$. A caustic occurs at $K_{x_o}\left(\theta-\theta_o\right)=-0.964$ rad.

Figure 3-31 shows the variation of $\Psi^{(0)}$ with y parametrically in $(\theta - \theta_o)$. We can evaluate the scattering integrals in Eq. (3.13-5) for the extreme cases: $\theta - \theta_o >> 0$ and $\theta - \theta_o << 0$. In the latter case, the LEO is well into the shadow of the reflecting sphere and so the expressions for the scattered wave after the transition through the shadow boundary should approach the negative of the expression for incident plane wave. Figure 3-30 shows that no stationary-phase points exist for $\theta - \theta_o < -964K_{x_o}^{-1}$ mrad and that the phase rate in the phasor

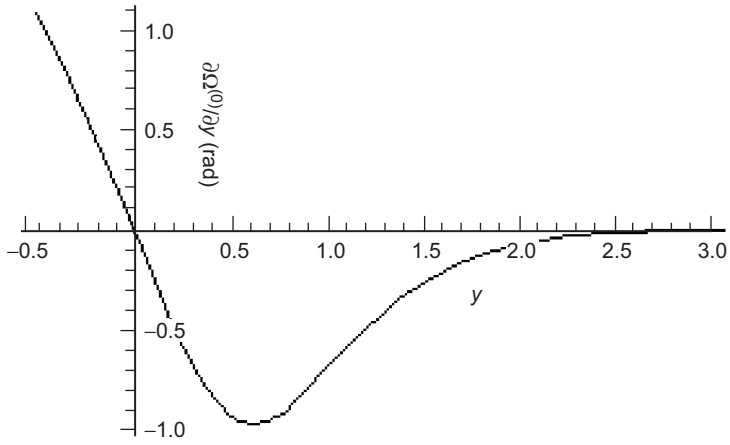

Fig. 3-30. Stationary points for the phase delay $\Psi^{(0)}$ for the reflected wave are defined by the intersections of the $2\partial\Omega^{(0)}/\partial y$ curve with a near-horizontal line (not shown); the latter is given by $\partial\Omega^{-}/\partial y \doteq K_{x_o}(\theta - \theta_o)$ = constant.

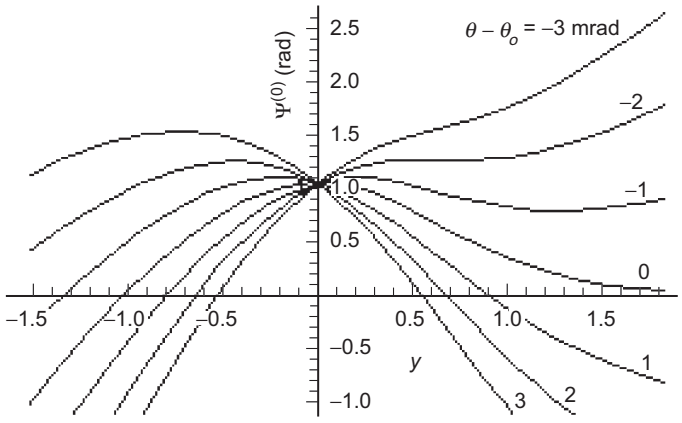

Fig. 3-31. Variation of $\Psi^{(0)}$ with spectral number.

$\exp\left(i\Psi^{(0)}\right)$ increases rapidly with decreasing $\theta-\theta_o$ below this threshold. Thus, we expect negligible contribution from the I integrals in Eq. (3.13-7) for $\theta-\theta_o$ in this region; the main contribution comes from J integrals in Eq. (3.13-8), which involve simple Fresnel integrals that are easily shown to approach the negative of the expression for the incident wave.

For $\left(\theta-\theta_o\right)>0$, the scattering results from near-grazing reflection of the incident wave from the sphere, and we will see that the scattering intensity increases as $\theta-\theta_o$ increases, that is, as the stationary point in y-space (see Fig. 3-31) becomes more negative. For this regime, the J integrals will contribute negligibly to the scattering (that is, scattering from a knife-edge is negligible for altitudes well above the altitude of the edge); we are concerned only with the I integrals for this limiting case. From Eq. (3.13-7), we have

$$\begin{Bmatrix} I_r \\ I_\theta \end{Bmatrix} \doteq E_o\sqrt{\frac{1}{2\pi i x\sin\theta}}\left[\int_{-\infty}^{x_o}\begin{Bmatrix}\left(\sin^2\theta_\nu\tan\theta_\nu\right)^{1/2} \\ \left(\sin\theta_\nu\cos\theta_\nu\right)^{1/2}\end{Bmatrix}e^{i\Psi^{(0)}}\,]d\nu\right.$$
$$\left.+\int_{x_o}^{\infty}\begin{Bmatrix}\left(\sin^2\theta_\nu\tan\theta_\nu\right)^{1/2} \\ \left(\sin\theta_\nu\cos\theta_\nu\right)^{1/2}\end{Bmatrix}\left(e^{i\Psi^{(0)}}-e^{i\Phi^-}\right)d\nu\right] \tag{3.17-8}$$

We note that the phasor in the second integral may be written as $\left(\exp\left(i2\Omega^{(0)}\right)-1\right)\exp(i\Phi^-)$, which decays to zero rapidly for $\nu>x_o$. We have already shown in Eqs. (3.17-6) and (3.17-7) for $\left(\theta-\theta_o\right)>0$ that the stationary-phase point in spectral number is located at $\nu^*<x_o$. Therefore, only the upper integral in Eq. (3.17-8) contributes significantly to the calculation of I for θ well above the shadow boundary.

To evaluate this upper integral using the stationary-phase technique, we expand $\Psi^{(0)}=2\Omega^{(0)}+\Phi^-$ about ν^* to obtain

$$\begin{aligned}\Psi^{(0)} &\doteq \Psi_{\nu^*}^{(0)}+\frac{1}{2}\left(\frac{\partial^2\Psi^{(0)}}{\partial\nu^2}\right)_{\nu^*}\left(\nu-\nu^*\right)^2 \\ &= \Psi_{\nu^*}^{(0)}+\frac{1}{2}\left(-\frac{2}{\left(x_o^2-\nu^{*2}\right)^{1/2}}-\frac{1}{\left(x^2-\nu^{*2}\right)^{1/2}}\right)\left(\nu-\nu^*\right)^2 \\ &\doteq \Psi_{\nu^*}^{(0)}-\frac{2}{x_o\left(\theta-\theta_o\right)}\left(\nu-\nu^*\right)^2\end{aligned} \tag{3.17-9}$$

Here the last line is obtained from Eq. (3.17-7a), noting that $\left(x_o^2 - \nu^{*2}\right)^{-1/2} >> \left(x^2 - \nu^{*2}\right)^{-1/2}$ for near-grazing reflections with $\left(\theta - \theta_o\right) > 0$. The upper integral in Eq. (3.17-8) becomes

$$\left.\begin{aligned} &\begin{Bmatrix} I_\theta \\ I_r \end{Bmatrix} \doteq \frac{E_o}{2} e^{i\Psi_{\nu^*}^{(0)}} \sqrt{\left(\theta - \theta_o\right)\tan\theta_o} \begin{Bmatrix} \cos\theta \\ \sin\theta \end{Bmatrix} \frac{1}{1+i} \int_{-\infty}^{w} e^{-i\frac{\pi}{2}w'^2}\, dw', \\ &w = \sqrt{K_{x_o}^3 \frac{(\theta - \theta_o)^3}{8\pi}}, \quad \Psi_{\nu^*}^{(0)} = \left(2\Omega^{(0)} + D_\nu + \nu(\theta_\nu - \theta)\right)_{\nu=\nu^*} \end{aligned}\right\} \tag{3.17-10}$$

Figure 3-32 shows the amplitude of the scattered wave versus $\left(\theta - \theta_o\right)$ from a perfectly reflecting sphere; a numerical integration of Eqs. (3.13-7) and (3.13-8) was used. Figure 3-33 is the amplitude of the scattering from a knife-edge pattern, that is, $\sqrt{J_r^2 + J_\theta^2}$ from Eq. (3.13-8). Figure 3-34 shows the amplitude of the difference between the scattering from the sphere and from a knife-edge, that is, the amplitude $\sqrt{I_r^2 + I_\theta^2}$ from Eq. (3.17-8). For $\theta - \theta_o > 0$, the Fresnel modulation is evident in Fig. 3-34 as a result of interference between wavelets with ν values above and below ν^*. The incomplete Fresnel integral in Eq. (3.17-10) provides this modulation. Also, the secular dependence

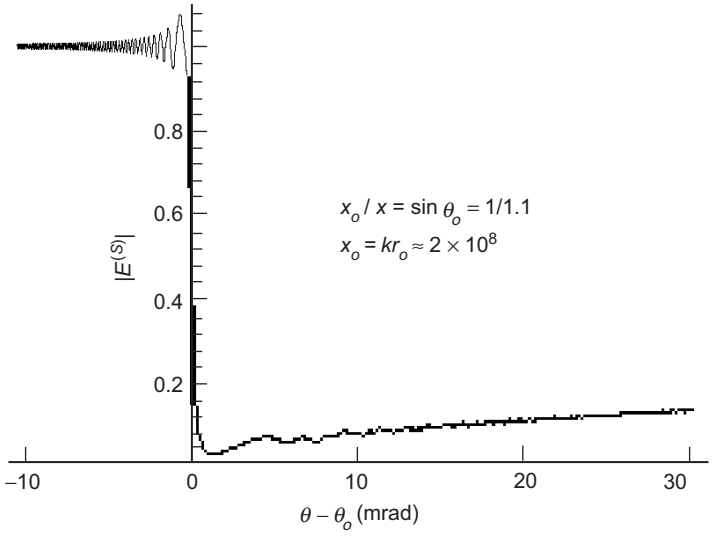

Fig. 3-32. Amplitude of the total scattering by a perfectly reflecting sphere.

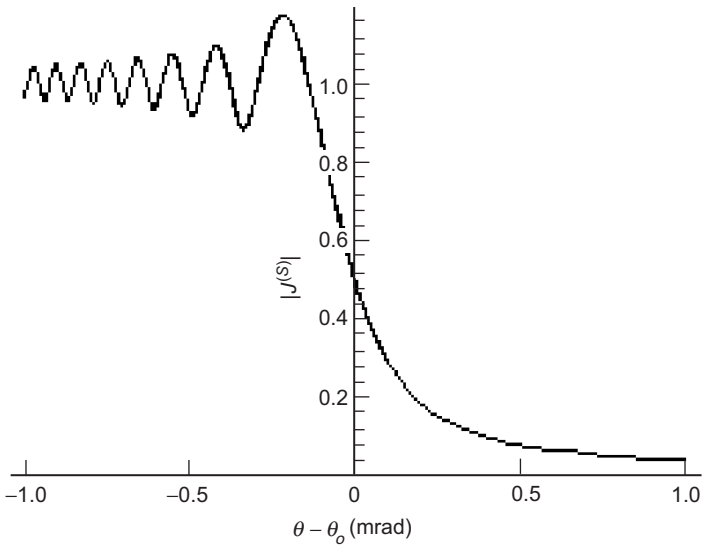

Fig. 3-33. Knife-edge component of scattering from a perfectly reflecting sphere. The same conditions as used in Fig. 3-32.

of the amplitude of the scattering on $\sqrt{(\theta-\theta_o)\tan\theta_o}\,/\,2$, which is predicted by Eq. (3.17-10), is clearly evident in this figure when $\theta-\theta_o>0$, indicative of a near-grazing reflection from the surface of the sphere. We see from these figures that the diffraction from a shiny sphere is accurately represented by the knife-edge diffraction pattern for $(\theta-\theta_o)<0$. Even for $\theta-\theta_o>0$, where near-grazing reflections occur, the knife-edge approximation is good to better than 10 percent for $K_{x_o}(\theta-\theta_o)<10$, or in our case for a LEO located at less than ~60 km, above the shadow boundary, or, equivalently, within about 20 to 30 s of elapsed time. We should expect better agreement for an opaque sphere, which should damp the scattering for $\theta-\theta_o>0$, except at near grazing conditions.

3.17.2 Geometric Optics Interpretation

The stationary-phase condition for y^* in Eq. (3.17-9) has a geometric optics interpretation in terms of the angle of incidence of the ray at the subreflection on the sphere. Referring to Fig. 3-14, let us define a new angular measure:

$$\theta_\nu^o = \sin^{-1}(\nu / x_o) \tag{3.17-11}$$

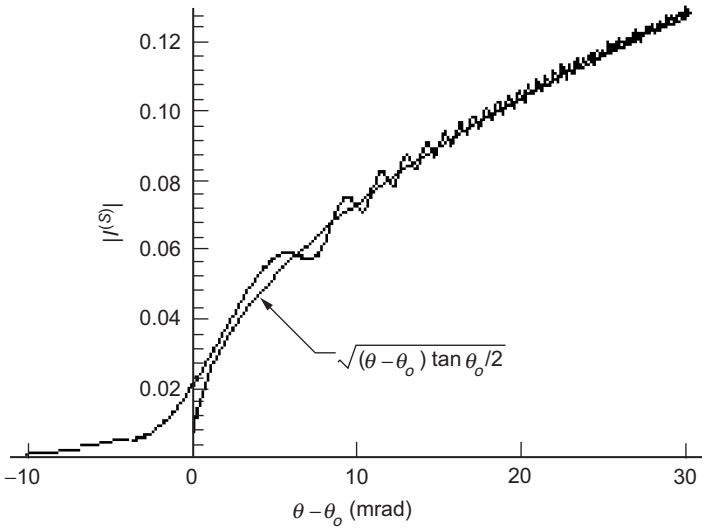

Fig. 3-34. Amplitude of the vector difference between scattering from a perfectly reflecting sphere and from a knife-edge. The same conditions as used in Fig. 3-32.

Hence $\sin\theta_\nu = \nu / x = \sin\theta_o \sin\theta_\nu^o$; θ_ν^o provides a measure of the spherical surface of radius ν relative to the reflecting sphere of radius x_o. When $\left(\theta-\theta_o\right) >\sim 2K_{x_o}^{-1} \approx 0.004$, so that $\Psi^{(0)} \to \pi/2 - 4(-y)^{3/2}/3 + \Phi^-$ with good accuracy, it follows from Eqs. (3.8-3) and (3.8-4) that the stationary-phase condition $\partial\Psi^{(0)}/\partial\nu = 0$ given in Eq. (3.17-6a) is achieved when ν assumes a value ν^*, which satisfies the expression

$$\left.\begin{aligned} &2\theta_{\nu^*}^o + \theta - \theta_{\nu^*} = \pi \\ &\text{or} \\ &\sin\left(2\theta_{\nu^*}^o + \theta\right) = -\sin\theta_o \sin\theta_{\nu^*}^o \end{aligned}\right\} \qquad (3.17\text{-}12)$$

However, referring to the geometry shown by Fig. 3-35, the ray from the GPS satellite that reflects from the surface of the sphere of radius x_o and continues to the LEO makes an angle of incidence φ_I with the normal to the sphere and reflects at an angle φ_R. When $\varphi_I = \varphi_R = \varphi$ and we replace θ_ν^o with φ in Eq. (3.17-12), we see from the geometry that this same condition holds for φ; hence, $\theta_{\nu^*}^o = \varphi = \pi/2 - (\theta-\theta_o)/2 + O[(\theta-\theta_o)^2]$. For the LEO located at an angular distance $(\theta-\theta_o)$ above the shadow boundary of the sphere, the ray from the tangent point on the caustic spherical surface of radius $\nu^* = x_o \sin\theta_{\nu^*}^o$

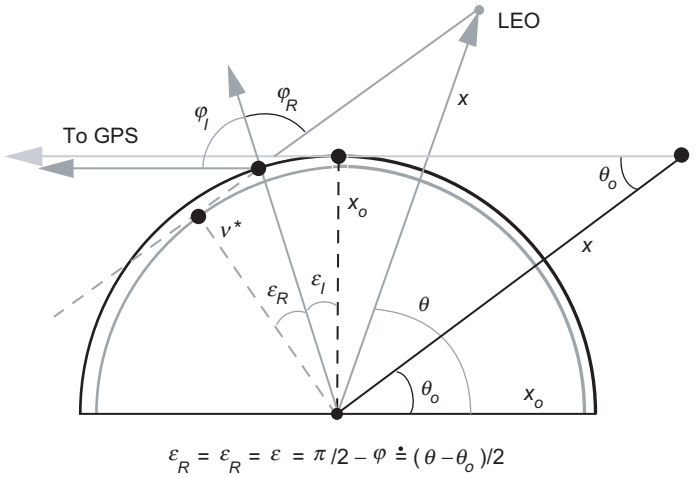

Fig. 3-35. Stationary phase geometry and the law of reflection. When the spectral number $\nu = \nu^*$ in wave theory, $\varphi_I = \varphi_R$ in geometric optics.

continues along a straight line to the LEO at a slope angle of $\pi - 2\varphi$ with respect to the LEO–GPS line. It intersects the reflecting sphere of radius x_o and makes an angle φ with respect to the radius vector; hence, this value ν^*, which provides a stationary value for $\Psi^{(0)}$, also yields the ray in geometric optics that satisfies the law of reflection $\varphi_I = \varphi_R$. The location $\tilde{\theta}$ of this new tangent point on the sphere of radius ν^* is given by $\tilde{\theta} \doteq \pi/2 + (\theta - \theta_o) + \tan\theta_o(\theta - \theta_o)^2/8$.

3.17.3 Limiting Cases: The Strongly Absorbing Sphere

In the case of a strongly absorbing sphere, the index of refraction $\hat{n}$ for the sphere contains an imaginary component giving rise to an extinction coefficient for waves penetrating its boundary [8]. We write this in the form $\hat{n} = n(1 + i\kappa)$, where n is the real component and $n\kappa$ is the imaginary component, and where κ is called the extinction coefficient and is real. Because $x_o >> 1$, the slightest hint of an extinction coefficient will result in the penetrating waves being completely damped before escaping from the sphere. It is easily shown from Maxwell's equations that the energy density of a penetrating wave falls to $1/e$ of its incident value after having traveled a distance $d = 1/2k\kappa n$ in a medium with a complex index of refraction $\hat{n}$. We want this distance to be much smaller than the radius of the sphere $d << r_o$. This inequality implies that $\kappa n >> 1/2x_o$, which is exceedingly small, and in this case the penetrating waves essentially are completely absorbed within the sphere for virtually all incident angles except very close to grazing. More precisely, to attain a "sharp" edge for the

limb of the sphere, we will want the distance d to be much less than the chord length of a sphere of radius r_o transected by the incident ray with an impact distance that is some specified small fraction of the first Fresnel zone; therefore, we want $\kappa >> x_o^{-3/4}/5 \sim 10^{-7}$. On the other hand, we want the sphere to reflect as little as possible if it is to emulate a knife-edge; therefore, $|\hat{n}|$ should be very close to unity in order to minimize reflection from the surface of the sphere at near-grazing angles of incidence. We may again take guidance from the Fresnel formulas in Eq. (3.4-1) and the reflection and transmission ratios shown in Fig. 3-4, where in this case the transmitted angle φ' will be complex, which gives rise to phase shifts in the reflected and transmitted waves. To avoid "glint" from near-grazing incident angles (that is, to optimize the agreement with the knife-edge results), we wish to keep κ as small as possible. We must do this consistent with maintaining essentially total absorption of the penetrating waves except within a very narrow range of impact parameter values that is a small fraction of the first Fresnel zone. We will see that there is a broad range for κ over which the diffraction results are nearly invariant and could be said to correspond to an "opaque" sphere.

The boundary conditions that have been applied to obtain the scattering coefficients in Eq. (3.5-11) are still valid for a complex index of refraction, but the arguments of the spherical Hankel functions $\xi_l^{\pm}(nx_o)$ are now complex and, therefore, involve in part the modified Bessel functions. The asymptotic forms for the Airy functions also hold in the complex plane, but certain restrictions apply to the argument of their complex variable $\hat{y}$ [see in Eq. (3.8-7)]. From Eq. (3.13-21), we have $\hat{y}(\nu) \doteq \left(2/\hat{n}x_o\right)^{2/3}\left(\nu - \hat{n}x_o\right)$; whenever $n\kappa$ is sufficiently large so that $\mathrm{Im}[\hat{y}(x_o)] \doteq -2\kappa\left(x_o/2\right)^{2/3} << -2$, the Airy functions $\mathrm{Ai}[\hat{y}]$ and $\mathrm{Bi}[\hat{y}]$ assume their respective asymptotic exponential forms to good accuracy even when $\nu = x_o$; it is easily shown in this case that when $\kappa >> (x_o/2)^{-2/3} \approx 4\times 10^{-6}$, $\xi_l^+\left(\hat{n}x_o\right) \to 0$, $\xi_l^-\left(\hat{n}x_o\right) \to \infty$. However, the ratio $\left(\xi_l'^-\left(\hat{n}x_o\right)/\xi_l^-\left(\hat{n}x_o\right)\right)$ is still finite and well-behaved, and is obtained from Eq. (3.8-10) when $\kappa n >> \left(2/x_o\right)^{2/3}$ and $\nu \sim x_o$.

From Eq. (3.5-11), we see that the only surviving scattering coefficients when $\kappa n >> \left(2/x_o\right)^{2/3}$ are $b_l^{(0)}$. From Eq. (3.9-3), these coefficients are given by

$$b_l^{(0)} = \frac{1}{2} i^{l-1} \frac{2l+1}{l(l+1)}\left(\rho^{(0)} \exp\left(i2\Omega^{(0)} - 1\right)\right) \tag{3.17-13}$$

where $\rho^{(0)}$ and $2\Omega^{(0)}$ are given by Eq. (3.9-5).

From Eq. (3.9-5), we have the asymptotic form for $2\Omega^{(0)}$, and it follows that $2\Omega^{(0)} \to \pi/2 - 4\left(-y(\nu)\right)^{3/2}/3$ for $\nu < x_o$, just as it does for the perfectly reflecting case. Here, y is real; the relationship between y and $\hat{y}$ when $n = 1$ and $\kappa << 1$ is given by $\hat{y} \doteq y - 2i\kappa\left(x_o/2\right)^{2/3}$. Therefore, the variation of $\Psi^{(0)}$ with y is very similar to that shown in Fig. 3-31 for the perfectly reflecting case, and in particular, when $\theta - \theta_o > 0$, $\Psi^{(0)}$ will have one stationary value, which is located in y-space at y^*, which in its asymptotic form is given by Eq. (3.17-8). However, unlike the case of the perfectly reflecting sphere where $|\rho^{(0)}(y)| \equiv 1$, in this case the modulus, which is the Mie analog of the Fresnel formula for the reflection coefficient, is a strong function of y and given by

$$\left|\rho^{(0)}(y)\right| \to \left\{ \begin{array}{l} \left\{ \left| \dfrac{\left(2i\kappa K_{x_o}^2 - y\right)^{1/2} - (-y)^{1/2}}{\left(2i\kappa K_{x_o}^2 - y\right)^{1/2} + (-y)^{1/2}} \right|, y < 0 \right\}, \\ 1,\ y > 0 \end{array} \right\} \qquad (3.17\text{-}14)$$

$$\hat{n} = 1 + i\kappa,\ \kappa << 1, \quad K_{x_o} = \left(\frac{x_o}{2}\right)^{1/3}$$

Figure 3-36 shows the behavior of $|\rho^{(0)}(y)|$ parametrically with different values of κ. For good absorption characteristics, we want $|\rho^{(0)}(y^*)| \approx 0$ whenever $\theta - \theta_o > 0$. This can be achieved by making κ sufficiently small so that $|\rho^{(0)}(y)|$ assumes a "Heaviside" functional form with a transition at $y = 0$, but not too small so that the sphere becomes partially transparent at near-grazing incidence angles. We see from Fig. 3-32 that this form is approached for $\kappa < \left(2/x_o\right)^{2/3} \sim 10^{-6}$.

Figure 3-37 shows the amplitude of the I integrals based on numerical integrations of Eq. (3.13-7); it shows the amplitude of I versus $\theta - \theta_o$ for different extinction coefficients. For the larger extinction coefficient, we would expect stronger scattering (and deeper interference) fringes for $\theta - \theta_o > 0$ and weaker scattering for $\theta - \theta_o < 0$.

For $\theta - \theta_o > 0$, the fringes in Fig. 3-37 can be largely recovered from the corresponding curve for the perfectly reflecting case (see Fig. 3-34) merely by multiplying the amplitudes for this case by $\rho^{(0)}(y^*[\theta - \theta_o])$. Figure 3-37 shows the amplitude of the vector difference between scattering from a strongly absorbing sphere and from a knife-edge; the difference is very slight for this

range of extinction values. For the scattering effects from an absorbing sphere, it is this close agreement between Mie scattering theory and scalar diffraction theory applied to a knife-edge that justifies the use of the latter to take advantage of its far simpler computational burden.

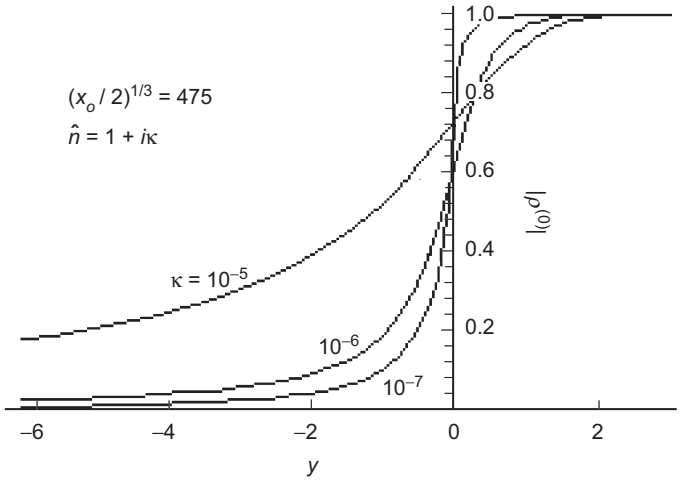

Fig. 3-36. Mie "reflection" coefficient for different extinction coefficients.

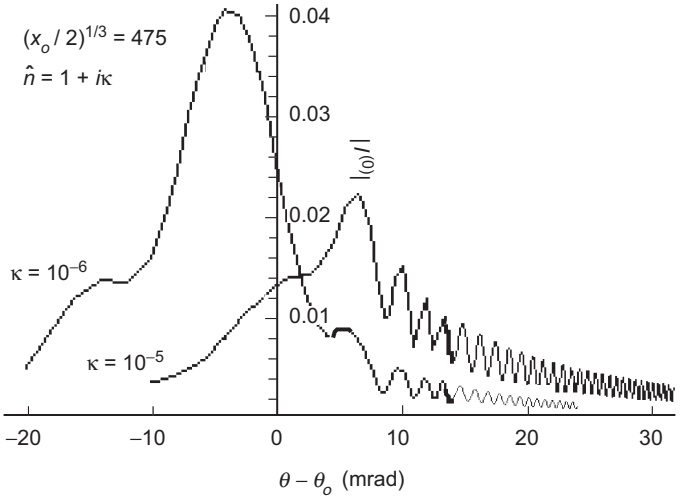

Fig. 3-37. Amplitude of the vector difference between the scattering from an opaque sphere with extinction coefficient κ and from a knife-edge.

3.17.4 Limiting Cases: Reflections from a Sphere of Infinite Radius

As the radius of the refracting sphere grows arbitrarily large relative to the wavelength of the incident wave, one can show that the stationary-phase conditions on the spectral number for the scattering series approach a simpler limiting form [16]. These become the Fresnel formulas for an infinite plane plus a phase offset to account for the location of the LEO.

References

[1] G. Mie, "Beiträge zur Optik Truber Medien Speziell Kolloidaler Metallösungen," *Annalen der Physik*, vol. 25, pp. 377–445, 1908.

[2] P. Debye, "Der Lichtdruck auf Kugeln von Beliebigem Material," *Annalen der Physik*, vol. 30, pp. 57–136, 1909.

[3] G. Watson, "The Diffraction of Electric Waves by the Earth," *Proceedings of the Royal Soc*iety, vol. A95, pp. 83–99, London, United Kingdom, 1918.

[4] B. Van der Pol and H. Bremmer, "The Diffraction of Electro-Magnetic Waves from an Electric Point Source Round a Finitely Conducting Sphere, with Applications to Radiotelegraphy and the Theory of the Rainbow, Part II," *Philosophical Magazine*, vol. 24, no. 164, *Supplement*, pp. 825–864, 1937.

[5] H. Bucerieus, "Theorie des Regenbogens und der Glorie," *Optik*, vol. 1, pp. 188–212, 1946.

[6] J. Stratton, *Electromagnetic Theory,* New York: McGraw-Hill, 1941.

[7] J. Jackson, *Classical Electrodynamics,* 2nd ed., New York: John Wiley & Sons, Inc., 1975.

[8] M. Born and E. Wolf, *Principles of Optics,* 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.

[9] G. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge, United Kingdom: Cambridge University Press, 1944.

[10] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards Series 55, Washington, DC, 1964.

[11] J. Mathews and R. Walker, *Mathematical Methods of Physics,* Menlo Park, California: Benjamin/Cummings Publishing Co., 1970.

[12] A. Sommerfeld, *Partial Differential Equations in Physics*, New York: Academic Press, 1967.

[13] H. Bremmer, *Terrestrial Radio Waves*, Amsterdam: Elsevier Publishing Co., 1949.

[14] J. Keller, "Geometric Theory of Diffraction," *The Journal of the Optical Society of America*, vol. 52, no. 2, pp. 116–130, 1962.

[15] A. Sommerfeld, "Matematische Theorie der Diffraction," *Math. Ann*. vol. 47**,** S319, pp. 317–374, 1896.

[16] J. Melbourne, *Sensing Atmospheric and Ionospheric Boundaries in GPS Radio Occultation Observations From a Low Earth Orbiter, Part 2*, JPL Publication 99-5, Jet Propulsion Laboratory, Pasadena, California, April 1999.

Chapter 4
Wave Propagation in a Stratified Medium: The Thin-Film Approach

4.1 Introduction

Chapter 4 changes gears, so to speak. It reviews a technique that uses the unitary state transition matrix for the system of first-order electromagnetic wave equations in a transmission medium that is thin, stratified, and linear [1–3]. This approach has proven useful for calculating the propagation of an electromagnetic wave through a thin film with Cartesian stratification. This chapter provides ready access to several propagation concepts that arise in the Mie scattering formulation: (1) the concepts of incoming and outgoing standing waves and their asymptotic forms, (2) turning points, (3) the osculating parameter technique in multiple Airy layers and the limiting forms of its solutions for a continuously varying refractivity, and (4) the accuracy of the osculating parameter technique within a Cartesian framework. The chapter also deals with the delicate problem of how to asymptotically match the incoming and outgoing solutions based on the osculating parameter technique. Sections 4.10 and 4.11 extend the thin-film approach and the unitary state transition matrix to cylindrical and spherical stratified media. Section 4.12 establishes a duality between spherical or cylindrical stratification and Cartesian stratification. This duality allows certain transformations to be applied to convert a problem with one type of stratification into another. The material in Chapters 5 and 6 does not depend on the material in this chapter. Therefore, this chapter may be skipped or skimmed. However, for the development of modified Mie scattering in Chapter 5, the background material herein may prove useful from time to time, particularly in the use of the Airy layer, a layer in which the gradient of the refractivity is a constant.

4.2 Thin-Film Concepts

Here we use the thin-film concepts [1–3] to develop the characteristic matrix, which describes the propagation of an electromagnetic wave through a stratified medium. When the "thin atmosphere" conditions hold [see Section 2.2, Eqs. (2.2-8) and (2.2-9)], this approach provides accurate results, and it also is instructive. In the usual thin-film approach, the stratified medium first is treated as a multi-layered medium. The index of refraction is held constant within each layer, but it is allowed to change across each boundary between the layers by an amount equal to some finite number (corresponding, for example, to the average gradient within the neighboring layers) times the thickness of the layer. Within each layer, the wave equations are readily solved in terms of sinusoid functions. The continuity conditions from Maxwell's equations allow one to tie the solutions together from neighboring layers across each layer boundary. Then the maximum thickness of each layer is driven to zero while the number of layers is allowed to grow indefinitely large so that the total thickness of the medium remains invariant. The resulting approximate solution from this ensemble of concatenated solutions can provide an accurate description of the electromagnetic field throughout the medium if the "thin atmosphere" conditions hold and if turning points are avoided.

4.2.1 Cartesian Stratification

We first develop the case for two-dimensional wave propagation in a medium with planar stratification. Many of the concepts developed here are applicable to the spherical stratified case, which we treat later. For the introduction of the Cartesian case, we follow closely the treatment presented in [3]. Here we assume that an electromagnetic wave, linearly polarized along the *y*-axis (see Fig. 4-1), i.e., a tranverse electric (TE) wave, is travelling through the medium. $\boldsymbol{H}$ lies in the *xz*-plane of incidence; $\boldsymbol{E}$ is parallel to the *y*-axis. $\boldsymbol{S}$ is the Poynting vector, also lying in the *xz*-plane. The angle φ is the angle of incidence of the wave. The wave is invariant in the y direction. The medium itself is stratified so that a planar surface of constant index of refraction is oriented perpendicular to the *x*-axis. The surfaces themselves are infinite in extent and parallel to the *yz*-plane. It will be sufficient to consider the TE case and its propagation in the *xz*-plane. Analyzing the TE case is preferable because in a medium where $\mu \equiv 1$ the equations are simpler [see Eq. (4.2-5)] than they are for the transverse magnetic (TM) case. To obtain results appropriate for the TM case, we can use the mathematical description for the $\boldsymbol{H}$ field that we will obtain for the TE case combined with use of the symmetry property in Maxwell's equations mentioned earlier. Maxwell's equations remain invariant when the definitions of the field vectors and their medium parameters are simultaneously exchanged according to the transformation

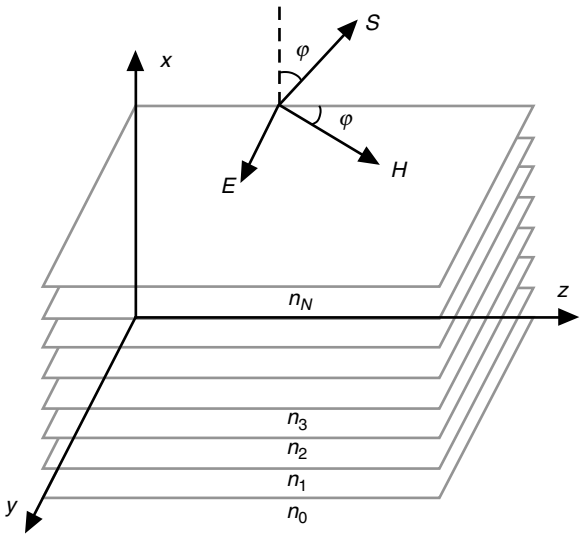

Fig. 4-1. Geometry for Cartesian stratification and a TE wave.

$(\boldsymbol{E},\boldsymbol{H},\varepsilon,\mu)\Leftrightarrow(-\boldsymbol{H},\boldsymbol{E},\mu,\varepsilon)$. This allows us to obtain $\boldsymbol{E}$ for the TM case from knowing $\boldsymbol{H}$ for the TE case.

In the plane of incidence of a planar wave, that is, the *xz*-plane, it follows for a TE wave that $E_x = E_z = 0$. The curl relations in Maxwell's equations for a TE harmonic wave give

$$\left.\begin{aligned} \frac{\partial E_y}{\partial x} &= ik\mu H_z \\ \frac{\partial E_y}{\partial z} &= -ik\mu H_x \\ H_y &= 0 \end{aligned}\right\} \tag{4.2-1a}$$

$$\left.\begin{aligned} \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} &= -ik\varepsilon E_y \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} &= 0 \\ \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} &= 0 \end{aligned}\right\} \tag{4.2-1b}$$

It follows from these equations that $H_y \equiv 0$ and that H_x, H_z, and E_y are functions of *x* and *z* only. From Eq. (3.2-1), it follows that the time-independent

component of the electric field for the TE case must satisfy the modified Helmholtz equation

$$\frac{\partial^2 E_y}{\partial x^2}+\frac{\partial^2 E_y}{\partial z^2}-\frac{d(\log\mu)}{dx}\frac{\partial E_y}{\partial x}+n^2k^2E_y=0 \tag{4.2-2}$$

Using the separation of variables technique, we set as a trial solution

$$E_y(z,x)=U(x)Z(z) \tag{4.2-3}$$

Then Eq. (4.2-2) becomes

$$\frac{1}{U}\frac{d^2U}{dx^2}+n^2k^2-\frac{d(\log\mu)}{dx}\frac{1}{U}\frac{dU}{dx}=\frac{1}{Z}\frac{d^2Z}{dz^2} \tag{4.2-4}$$

The left-hand side (LHS) of Eq. (4.2-4) is a function only of x while the right-hand side (RHS) is a function only of z. This can be true only if both sides are constant. Hence,

$$\frac{1}{Z}\frac{d^2Z}{dz^2}=-k^2n_o^2;\quad n_o=\text{constant} \tag{4.2-5}$$

Here n_o is a constant to be determined later. The solution for $Z(z)$ follows immediately:

$$Z=Z_0\exp(\pm ikn_oz) \tag{4.2-6}$$

Thus, kn_o is the rate of phase accumulation of the time-independent component of the wave along the z-direction, and it is an invariant for a particular wave. Without loss of generality, we can assume that the wave is trending from left to right in Fig. 4-1 (i.e., in the direction of positive z); hence, we adopt the positive sign in Eq. (4.2.6). Then the electric field for a harmonic wave is given by

$$E_y=U(x)\exp\big(i\big(kn_oz-\omega t\big)\big) \tag{4.2-7}$$

From Maxwell's equations [see Eq. (4.2-1a), it follows that the magnetic field components are expressed in terms of different functions of x but the same functions of z and t. These are given by

$$\left.\begin{aligned}H_z&=V(x)\exp\big(i\big(kn_oz-\omega t\big)\big)\\ H_x&=W(x)\exp\big(i\big(kn_oz-\omega t\big)\big)\end{aligned}\right\} \tag{4.2-8}$$

The wave equations also require that the U, V, and W functions satisfy certain conditions among themselves, which from Eq. (4.2-1) are given by

$$\left.\begin{aligned} U' &= ik\mu V \\ V' &= ik\left(n_o W + \varepsilon U\right) \\ \mu W + n_o U &= 0 \end{aligned}\right\} \tag{4.2-9}$$

Eliminating W, we obtain a coupled system of first-order differential equations for U and V:

$$\left.\begin{aligned} U' &= ik\mu V \\ V' &= ik\mu^{-1}\left(n^2 - n_o^2\right)U \end{aligned}\right\} \tag{4.2-10}$$

Alternatively, one can convert this coupled system into a pair of independent second-order differential equations, which from Eq. (3.2-1) are given by

$$\left.\begin{aligned} \frac{d^2U}{dx^2} - \frac{d(\log\mu)}{dx}\frac{dU}{dx} + k^2\left(n^2 - n_o^2\right)U &= 0 \\ \frac{d^2V}{dx^2} - \frac{d\left(\log\left[\left(n^2 - n_o^2\right)/\mu\right]\right)}{dx}\frac{dV}{dx} + k^2\left(n^2 - n_o^2\right)V &= 0 \end{aligned}\right\} \tag{4.2-11}$$

For the TM case ($H_x = H_z \equiv 0$), the transformation $(\boldsymbol{E}, \boldsymbol{H}, \varepsilon, \mu) \Leftrightarrow (-\boldsymbol{H}, \boldsymbol{E}, \mu, \varepsilon)$ yields

$$H_y = U(x)\exp\left(i\left(kn_o z - \omega t\right)\right) \tag{4.2-7′}$$

and

$$\left.\begin{aligned} E_z &= -V(x)\exp\left(i\left(kn_o z - \omega t\right)\right) \\ E_x &= -W(x)\exp\left(i\left(kn_o z - \omega t\right)\right) \end{aligned}\right\} \tag{4.2-8′}$$

The wave equations for the TM case become

$$\left.\begin{aligned} U' &= ik\varepsilon V \\ V' &= ik\varepsilon^{-1}\left(n^2 - n_o^2\right)U \\ \varepsilon W + n_o U &= 0 \end{aligned}\right\} \tag{4.2-9′}$$

or

$$\left.\begin{aligned} \frac{d^2U}{dx^2} - \frac{d\log\varepsilon}{dx}\frac{dU}{dx} + k^2\left(n^2 - n_o^2\right)U = 0 \\ \frac{d^2V}{dx^2} - \frac{d\left(\log\left[\left(n^2 - n_o^2\right)/\varepsilon\right]\right)}{dx}\frac{dV}{dx} + k^2\left(n^2 - n_o^2\right)V = 0 \end{aligned}\right\} \qquad (4.2\text{-}11')$$

Returning to the TE case, we note that in general U, V, and W are complex. From Eq. (4.2-7), a surface of constant phase for E_y (called the cophasal surface) is defined by

$$\psi(x) + kn_o z - \omega t = \text{constant} \qquad (4.2\text{-}12)$$

where $\psi(x)$ is the phase of $U(x)$. For an infinitesimal displacement $(\delta x, \delta z)$ at a fixed time and lying on the cophasal surface, we have from Eq. (4.2-12) the condition $\psi'\delta x + kn_o\delta z = 0$. Therefore, the angle of incidence φ that the cophasal surface makes with the *yz*-plane (Fig. 4-1) is given by

$$\tan\varphi = -\delta x / \delta z = kn_o / \psi' \qquad (4.2\text{-}13)$$

For the special case where the wave is planar, we have $\psi'(x) = kn\cos\varphi$, from which it follows that the constant, n_o, in Eq. (4.2-5) is given by

$$n_o = n\sin\varphi = \text{constant} \qquad (4.2\text{-}14)$$

which is Snell's law. It follows that the condition $n_o = \text{constant}$, obtained from the solution to the modified wave equation in Eq. (4.2-5), can be considered as a generalization of Snell's law. The value n_o, not to be confused with n_0 associated with the index of refraction of the 0th layer in Fig. 4-1, provides the index of refraction and therefore the layer(s) in which, according to geometric optics, $\varphi = \pi/2$, which marks a turning point for the wave.

4.3 The Characteristic Matrix

Returning to the coupled system in Eq. (4.2-10), we know that U and V both have two independent solutions to Eq. (4.2-11). Let these solutions be given by $F(x, x_0)$ and $f(x, x_0)$ for U, and by $G(x, x_0)$ and $g(x, x_0)$ for V. However, these solutions are constrained by the conditions in Eq. (4.2-10), which are given by

$$\left.\begin{aligned} \frac{dF}{dx} &= ik\mu G, \quad \frac{df}{dx} = ik\mu g \\ \frac{dG}{dx} &= ik\left(\varepsilon - n_o^2\mu^{-1}\right)F, \quad \frac{dg}{dx} = ik\left(\varepsilon - n_o^2\mu^{-1}\right)f \end{aligned}\right\} \tag{4.3-1}$$

Using a Green function-like approach, we construct these solutions so that the following specific boundary values are obtained:

$$\left.\begin{aligned} F(x_0, x_0) &= 1, \quad f(x_0, x_0) = 0 \\ G(x_0, x_0) &= 0, \quad g(x_0, x_0) = 1 \end{aligned}\right\} \tag{4.3-2}$$

Then it follows that in matrix form U and V can be written as

$$\begin{pmatrix} U(x) \\ V(x) \end{pmatrix} = \begin{bmatrix} F(x, x_0) & f(x, x_0) \\ G(x, x_0) & g(x, x_0) \end{bmatrix} \begin{pmatrix} U(x_0) \\ V(x_0) \end{pmatrix} \tag{4.3-3}$$

We define the characteristic matrix $\boldsymbol{M}[x, x_0]$ by

$$\boldsymbol{M}[x, x_0] = \begin{bmatrix} F(x, x_0) & f(x, x_0) \\ G(x, x_0) & g(x, x_0) \end{bmatrix} \tag{4.3-4}$$

Hence, Eq. (4.3-3) shows that the description of the electromagnetic wave through the stratified medium is borne solely by the initial conditions and by this state transition matrix $\boldsymbol{M}[x, x_0]$, a 2×2 unitary matrix. From the theory of ordinary differential equations, one can show that $\boldsymbol{M}[x, x_0]$ has a constant determinant, and in the special case where the boundary values given in Eq. (4.3-2) apply, $\mathrm{Det}\left[\boldsymbol{M}[x, x_0]\right] = 1$ for all values of x. This can be shown to result from the conservation of energy principle that applies to a non-absorbing medium where n is real. Henceforth, we will focus our attention on the properties of $\boldsymbol{M}[x, x_0]$ for different cases of stratification in the propagation medium.

4.4 The Stratified Medium as a Stack of Discrete Layers

An important transitive property of $\boldsymbol{M}[x, x_0]$ is obtained from the following observation. Consider two contiguous layers of different indices of refraction. The thickness of the first layer is $x_1 - x_0$, and its index of refraction is given by $n_1(x)$. The thickness of the second layer is $x_2 - x_1$, and its index of refraction is given by $n_2(x)$. Across a surface, Maxwell's equations require the tangential components of $\boldsymbol{E}$ to be continuous, and they also require the

tangential components of $\boldsymbol{H}$ to be continuous when surface currents are absent, which is assumed here. Since U and V describe the tangential components of the electromagnetic field vectors, U and V also must be continuous across the boundary. Hence, the relationships for the electromagnetic field are given by

$$\left.\begin{aligned} &\begin{pmatrix} U_2 \\ V_2 \end{pmatrix} = \boldsymbol{M}[x_2, x_1]\begin{pmatrix} U_1 \\ V_1 \end{pmatrix}, \ \begin{pmatrix} U_1 \\ V_1 \end{pmatrix} = \boldsymbol{M}[x_1, x_0]\begin{pmatrix} U_0 \\ V_0 \end{pmatrix} \\ &\begin{pmatrix} U_2 \\ V_2 \end{pmatrix} = \boldsymbol{M}[x_2, x_1]\boldsymbol{M}[x_1, x_0]\begin{pmatrix} U_0 \\ V_0 \end{pmatrix} = \boldsymbol{M}[x_2, x_0]\begin{pmatrix} U_0 \\ V_0 \end{pmatrix} \\ &\therefore \ \ \boldsymbol{M}[x_2, x_0] = \boldsymbol{M}[x_2, x_1]\boldsymbol{M}[x_1, x_0] \end{aligned}\right\} \tag{4.4-1}$$

This product rule can be generalized to N layers by

$$\boldsymbol{M}[x_N, x_0] = \boldsymbol{M}[x_N, x_{N-1}]\boldsymbol{M}[x_{N-1}, x_{N-2}]\cdots\boldsymbol{M}[x_1, x_0] = \prod_{k=1}^{N}\boldsymbol{M}[x_k, x_{k-1}] \tag{4.4-2}$$

If the form of the index of refraction $n(x)$ within each layer is such that the solutions for U and V can be expressed in terms of relatively simple functions, then in some cases it is possible to obtain a closed form for $\boldsymbol{M}[x, x_0]$ using the product rule. Approximate forms of sufficient accuracy also can be obtained in some cases.

4.4.1 The Characteristic Matrix when $n(x)$ = constant

When the index of refraction is constant within a layer, we obtain sinusoid solutions to Eqs. (4.2-10) and (4.2-11) for the TE case, which can be forced to satisfy the boundary conditions in Eq. (4.3-2). These solutions are the elements of the characteristic matrix that describes a plane wave traversing the layer. The functional elements of the characteristic matrix are given by

$$\left.\begin{aligned} &F(x, x_0) = \cos\left[k\varpi(x - x_0)\right], \quad f(x, x_0) = i\frac{1}{\varpi}\sin\left[k\varpi(x - x_0)\right] \\ &G(x, x_0) = i\varpi\sin\left[k\varpi(x - x_0)\right], \quad g(x, x_0) = \cos\left[k\varpi(x - x_0)\right] \\ &\varpi = \mu^{-1}\sqrt{n^2 - n_o^2} = \sqrt{\frac{\varepsilon}{\mu}}\cos\varphi \end{aligned}\right\} \tag{4.4-3}$$

Here φ is the angle of incidence of the plane wave in the layer (Fig. 4-1). It follows that $k\mu\varpi$ is the rate of phase accumulation of the time-independent component of the wave along the x-axis, perpendicular to the plane of stratification.

For the TM case, the solutions are of the same form except $f_{\mathrm{TM}} = (\varepsilon/\mu) f_{\mathrm{TE}}$ and $G_{\mathrm{TM}} = (\mu/\varepsilon) G_{\mathrm{TE}}$.

4.4.2 A Stack of Homogeneous Layers when $n(x)$ is Piecewise Constant

We apply these results to a stack of layers as shown in Fig. 4-1. The index of refraction varies from layer to layer, but within the jth layer the index of refraction is constant so that $\varpi_j = \mu_j^{-1}\left(n_j^2 - n_o^2\right)^{1/2}$ also is a constant within this layer. We also note that $\varpi_j = \left(\varepsilon_j/\mu_j\right)^{1/2}\cos\varphi_j$, where φ_j is the angle of incidence of the wave within the jth layer. Thus, within the jth layer the angle of incidence and the rate of phase accumulation remain constant.

Let us now define a reference characteristic matrix $\tilde{\boldsymbol{M}}$ by

$$\tilde{\boldsymbol{M}}\left[x_j, x_{j-1}\right] = \begin{bmatrix} \cos\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] & \dfrac{i}{\overline{\varpi}}\sin\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] \\ i\overline{\varpi}\sin\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] & \cos\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] \end{bmatrix} \tag{4.4-4}$$

where $\overline{\varpi}$ is a constant across all layers; its value will be set later. Here, x_j marks the upper boundary (Fig. 4-1) of the jth layer. We set

$$\boldsymbol{M}\left[x_j, x_{j-1}\right] = \tilde{\boldsymbol{M}}\left[x_j, x_{j-1}\right] + \delta\boldsymbol{M}\left[x_j, x_{j-1}\right] \tag{4.4-5}$$

where $\delta\boldsymbol{M}$ is defined as the difference between the actual characteristic matrix $\boldsymbol{M}\left[x_j, x_{j-1}\right]$ and the reference matrix $\tilde{\boldsymbol{M}}\left[x_j, x_{j-1}\right]$. This difference is due to $\delta\varpi_j = \varpi_j - \overline{\varpi}$. Then to first order in $\delta\varpi_j$, $\delta\boldsymbol{M}$ is given by

$$\delta\boldsymbol{M}\left[x_j, x_{j-1}\right] \doteq \begin{bmatrix} 0 & -\dfrac{i}{\overline{\varpi}}\sin\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] \\ i\overline{\varpi}\sin\left[k\varpi_j\left(x_j - x_{j-1}\right)\right] & 0 \end{bmatrix} \frac{\delta\varpi_j}{\overline{\varpi}} \tag{4.4-6}$$

Here $\delta\varpi_j$ is assumed to be small but not negligible.

From the product rule in Eq. (4.4-2) and truncating to first order, it follows that

$$\boldsymbol{M}[x_N, x_0] = \boldsymbol{M}_{N,0} = \prod_{j=1}^{N}\left(\tilde{\boldsymbol{M}}_{j,j-1} + \delta\boldsymbol{M}_{j,j-1}\right)$$

$$\doteq \prod_{j=1}^{N}\tilde{\boldsymbol{M}}_{j,j-1} + \sum_{j=2}^{N-1}\left(\left(\prod_{l=j}^{N-1}\tilde{\boldsymbol{M}}_{l+1,l}\right)\delta\boldsymbol{M}_{j,j-1}\left(\prod_{l=1}^{j-1}\tilde{\boldsymbol{M}}_{l,l-1}\right)\right) \tag{4.4-7}$$

$$+ \delta\boldsymbol{M}_{N,N-1}\left(\prod_{l=1}^{N-1}\tilde{\boldsymbol{M}}_{l,l-1}\right) + \left(\prod_{l=2}^{N}\tilde{\boldsymbol{M}}_{l,l-1}\right)\delta\boldsymbol{M}_{1,0}$$

Truncating the expansions in Eqs. (4.4-6) and (4.4-7) to first order in $\delta\varpi_j$ and $\delta\boldsymbol{M}$ should be accurate if $|\varpi'|$ is sufficiently small throughout the stack. The range of validity is discussed later. With regard to the reference characteristic matrix, it is easily shown that

$$\tilde{\boldsymbol{M}}_{m,l} = \prod_{j=l+1}^{m}\tilde{\boldsymbol{M}}_{j,j-1} = \begin{bmatrix} \cos\mathscr{A}_{m,l} & \dfrac{i}{\overline{\varpi}}\sin\mathscr{A}_{m,l} \\ i\overline{\varpi}\sin\mathscr{A}_{m,l} & \cos\mathscr{A}_{m,l} \end{bmatrix} \tag{4.4-8}$$

where

$$\left.\begin{aligned} \mathscr{A}_{m,l} &= k\sum_{j=l+1}^{m}\varpi_j\left(x_j - x_{j-1}\right) \\ \mathscr{A}_{m,m} &= 0 \end{aligned}\right\} \tag{4.4-9}$$

Also, defining $\tilde{\boldsymbol{M}}_{N,N} = \tilde{\boldsymbol{M}}_{0,0} = \boldsymbol{I}$, the identity matrix, the jth product in Eq. (4.4-7) for $j = 1, 2, \cdots, N$, becomes

$$\left(\prod_{l=j+1}^{N}\tilde{\boldsymbol{M}}_{l,l-1}\right)\delta\boldsymbol{M}_{j,j-1}\left(\prod_{l=1}^{j-1}\tilde{\boldsymbol{M}}_{l,l-1}\right) = \tilde{\boldsymbol{M}}_{N,j}\delta\boldsymbol{M}_{j,j-1}\tilde{\boldsymbol{M}}_{j-1,0}$$

$$= \frac{\delta\varpi_j}{\overline{\varpi}}\sin\left[k\varpi_j\left(x_j - x_{j-1}\right)\right]\begin{bmatrix} -\sin\mathscr{B}_j & \dfrac{i}{\overline{\varpi}}\cos\mathscr{B}_j \\ -i\overline{\varpi}\cos\mathscr{B}_j & \sin\mathscr{B}_j \end{bmatrix} \tag{4.4-10}$$

where

$$\mathscr{B}_j = \sum_{l=j+1}^{N}k\varpi_l\left(x_l - x_{l-1}\right) - \sum_{l=1}^{j-1}k\varpi_l\left(x_l - x_{l-1}\right) \tag{4.4-11}$$

Now we go to the limit, allowing $x_j - x_{j-1} \to 0$ and $N \to \infty$ so that

$$\sum_{j=1}^{N} k\varpi_j\left(x_j - x_{j-1}\right) \to k\int_{x_0}^{x_F} \varpi dx \tag{4.4-12}$$

It follows that

$$\left.\begin{aligned} &\mathscr{A}_{N,0} \to \mathscr{A}\left(x_F, x_0\right) = \int_{x_0}^{x_F} \varpi dx \\ &\mathscr{B}_j \to \mathscr{B}(x) = k\int_{x}^{x_F} \varpi dx - k\int_{x_0}^{x} \varpi dx = \mathscr{A}\left(x_F, x\right) - \mathscr{A}\left(x, x_0\right) \\ &\varpi(x) = \mu^{-1}\sqrt{n^2(x) - n_o^2}, \quad n_o = n(x)\big|_{x=x_o} \end{aligned}\right\} \tag{4.4-13}$$

where x_F is the final value of x, nominally where the electromagnetic field is to be evaluated. The quantity $\mathscr{A}\left(x_F, x_0\right)$ is the total phase accumulation of the time-independent component of the wave along the x-direction between x_0 and x_F. Note that $\mathscr{A}\left(x_F, x_0\right)$ is an implicit function of the refractivity profile of the medium, and it also is a function of n_o, (through ϖ) or the angle of incidence, for a specific wave.

Upon passing to the limit and integrating, Eq. (4.4-10) becomes

$$\begin{aligned} &\sum_{\boldsymbol{j=1}}^{N} \tilde{\boldsymbol{M}}_{N,j}\delta\boldsymbol{M}_{j,j-1}\tilde{\boldsymbol{M}}_{j-1,0} \to \int_{x_0}^{x_F} \tilde{\boldsymbol{M}}\left[x_F, x\right]\delta\boldsymbol{M}\left[x, x\right]\tilde{\boldsymbol{M}}\left[x, x_0\right]dx = \\ &\frac{k}{\overline{\varpi}}\int_{x_0}^{x_F} \left(\varpi(x) - \overline{\varpi}\right)\varpi(x)\begin{bmatrix} -\sin\mathscr{B}(x) & -\dfrac{i}{\overline{\varpi}}\cos\mathscr{B}(x) \\ i\overline{\varpi}\cos\mathscr{B}(x) & \sin\mathscr{B}(x) \end{bmatrix}dx \end{aligned} \tag{4.4-14}$$

Upon noting that $\mathscr{B}'(x) = -2k\varpi(x)$, we can integrate this integral by parts to obtain

$$\begin{aligned} &\int_{x_0}^{x_F} \tilde{\boldsymbol{M}}\left[x_F, x\right]\delta\boldsymbol{M}\left[x, x\right]\tilde{\boldsymbol{M}}\left[x, x_0\right]dx = \\ &\frac{1}{2\overline{\varpi}}\begin{bmatrix} \left(\varpi_0 - \varpi_F\right)\cos\mathscr{A} & \dfrac{i}{\overline{\varpi}}\left(2\overline{\varpi} - \varpi_0 - \varpi_F\right)\sin\mathscr{A} \\ i\overline{\varpi}\left(\varpi_0 + \varpi_F - 2\overline{\varpi}\right)\sin\mathscr{A} & \left(\varpi_F - \varpi_0\right)\cos\mathscr{A} \end{bmatrix} + \begin{bmatrix} I_1 & -\dfrac{i}{\overline{\varpi}}I_2 \\ i\overline{\varpi}I_2 & -I_1 \end{bmatrix} \end{aligned} \tag{4.4-14′}$$

where

$$\left.\begin{aligned} I_1 &= \frac{1}{2\overline{\varpi}}\int_{x_0}^{x_F}\frac{d\varpi}{dx}\cos\mathscr{B}(\zeta)dx \\ I_2 &= \frac{1}{2\overline{\varpi}}\int_{x_0}^{x_F}\frac{d\varpi}{dx}\sin\mathscr{B}(\zeta)dx \end{aligned}\right\} \tag{4.4-15}$$

We now set $\overline{\varpi}$ equal to its "average" value over the interval $[x_F, x_0]$. That is,

$$\overline{\varpi} = (\varpi_F + \varpi_0)/2 \tag{4.4-16}$$

Adding the resulting perturbation matrix in Eq. (4.4-14′) to the reference matrix $\tilde{\boldsymbol{M}}[x_F, x_0]$ given by Eq. (4.4-8), we obtain a first-order expression for the characteristic matrix $\boldsymbol{M}[x_F, x_0]$ applicable to the entire stratified medium for the TE case. This is given by

$$\boldsymbol{M}[x_F, x_0] \doteq \begin{bmatrix} \dfrac{\varpi_0}{\overline{\varpi}}\cos\mathscr{A} + I_1 & \dfrac{i}{\overline{\varpi}}(\sin\mathscr{A} - I_2) \\ i\overline{\varpi}(\sin\mathscr{A} + I_2) & \dfrac{\varpi_F}{\overline{\varpi}}\cos\mathscr{A} - I_1 \end{bmatrix} \tag{4.4-17}$$

4.4.3 Range of Validity

Let us estimate the range of validity of the linear perturbation approach used in Eq. (4.4-7) to obtain Eq. (4.4-17). We have noted that its accuracy will depend on $|\varpi'|$ being sufficiently small. From Eqs. (4.2-14) and (4.4-3), it follows that

$$\varpi' = n'\sec\varphi \tag{4.4-18}$$

The magnitude of the first term on the RHS of Eq. (4.4-14′) is of the order of $(x_F - x_0)\langle n'\sec\varphi\rangle$, where $\langle\ \rangle$ denotes an average over the interval $(x_F - x_0)$. Thus, if this term is small, the linear truncation should be valid. For the second term on the RHS of Eq. (4.4-14') involving the *I* integrals in Eq. (4.4-15), let us assume that ϖ' is a constant over the integration interval. In this case, $\mathscr{B}(x)$ is quadratic in x and Eq. (4.4-15) involves Fresnel integrals. It is easily shown that $|I_1| \approx |I_2| \approx (\varpi'\lambda)^{1/2}/\varpi$ for $(x_F - x_0)/\lambda >> 1$; these terms away from turning points are generally small for "thin atmospheres," and smaller than the first term. Although the *I* integrals will be small under these conditions, their integrands are highly oscillatory when $(x_F - x_0)/\lambda >> 1$. If retention of these terms is necessary, special integration algorithms using the rapid variation of $\exp[i\mathscr{B}(x)]$ and the slowly varying character of $d\varpi/dx$ are helpful.

We conclude for n' sufficiently small and for points located sufficiently far from turning points, where $\varphi = \pi/2$, that the linear truncation used in Eq. (4.4-7) will be sufficiently accurate. If these conditions hold, one can neglect the $N^2/2$ second-order terms $\delta\boldsymbol{M}_{j,j-1}\delta\boldsymbol{M}_{m,m-1}$ in the product rule expansion in Eq. (4.4-7), as well as the second-order terms in Eq. (4.4-6).

We note a special interpretation for the quantity $n'(x_F - x_0) = (n'x_0)\big((x_F - x_0)/x_0\big)$. In spherical coordinates, the first product is the ratio of the radius of curvature of the refracting surface to the radius of curvature of the ray path, which is one measure of atmospheric "thinness." For dry air at the Earth's surface, this quantity is about 1/4. The second product is merely the fraction of the total radius traversed by the ray, usually very small for a large sphere.

4.4.4 The TM Case

It is easy to show using the transformation $(\boldsymbol{E}, \boldsymbol{H}, \varepsilon, \mu) \Leftrightarrow (-\boldsymbol{H}, \boldsymbol{E}, \mu, \varepsilon)$ that the TM version of Eq. (4.4-17) is given by

$$\left.\begin{aligned} \boldsymbol{M}[x_F, x_0] &\doteq \begin{bmatrix} \dfrac{\varpi_{\mathrm{TM}}|_{x_F}}{\overline{\varpi}_{\mathrm{TM}}}\cos\mathscr{A}_{\mathrm{TM}} + I_1 & \dfrac{i}{\overline{\varpi}_{\mathrm{TM}}}\left(\sin\mathscr{A}_{\mathrm{TM}} + I_2\right) \\ i\overline{\varpi}_{\mathrm{TM}}\left(\sin\mathscr{A}_{\mathrm{TM}} - I_2\right) & \dfrac{\varpi_{\mathrm{TM}}|_{x_0}}{\overline{\varpi}_{\mathrm{TM}}}\cos\mathscr{A}_{\mathrm{TM}} - I_1 \end{bmatrix} \\ \varpi_{\mathrm{TM}} &= \frac{\mu}{\varepsilon}\varpi_{\mathrm{TE}} = \sqrt{\frac{\mu}{\varepsilon}}\cos\varphi \end{aligned}\right\} \quad (4.4\text{-}17')$$

The form for $\boldsymbol{M}[x_F, x_0]$ in Eq. (4.4-17), but without the I_1 and I_2 terms, first appears in [4]. It also can be obtained by applying the Wentzel–Kramer–Brillouin (WKB) method to Eq. (4.2-11). The WKB solution to Eq. (4.2-11) was almost certainly known during Lord Rayleigh's time because of his studies of acoustic waves in a refracting medium.

One could generalize this problem to include a stratified medium with an embedded discontinuity. Within the medium a surface parallel to the stratification is embedded. This surface acts as a boundary between two regions. Within each region, $n(x)$ is continuous, but across the boundary $n(x)$ or its gradient is discontinuous. Within each region, a characteristic matrix of the form in Eq. (4.4-17) applies, and the product of these two matrices provides the characteristic matrix that spans the entire medium, including the discontinuity. One also can calculate the reflection and transmission

coefficients across the discontinuous boundary in terms of the elements of the characteristic matrices at the boundary. Since a modified Mie scattering approach will be used in Chapter 5 to address the problem of a scattering spherical surface embedded in a refracting medium, it will not be pursued further here. See [3] and [5] for further discussion of this case.

4.5 The Characteristic Matrix for an Airy Layer

We can check the characteristic matrix given in Eq. (4.4-17), which results from a linear theory applied to an infinite stack of infinitesimal layers, with an essentially exact result, which can be obtained when the gradient of the index of refraction is constant within the medium. We designate a layer with a constant gradient an "Airy layer," because the Airy functions form the solution set for such a layer. We let

$$n^2 = n_0^2 + 2n_0 n'(x - x_0) \tag{4.5-1}$$

where n_0 and n' are constants throughout the layer and $n'(x_F - x_0)$ is sufficiently small so the term $\left(n'(x_F - x_0)\right)^2$ can be neglected. This quasi-linear form for the index of refraction has application in atmospheric propagation studies. There the continuous profile for $n(x)$ is approximated by a series of piecewise constant-gradient segments [6].

Returning to Eqs. (4.2-10) and (4.2-11), we have for the TE case in a single layer:

$$\left.\begin{aligned} &\frac{d^2U}{dx^2} + k^2\left(n_0^2 - n_o^2 + 2n_0 n'(x - x_0)\right)U = 0 \\ &\frac{dU}{dx} = ik\mu V \end{aligned}\right\} \tag{4.5-2}$$

Without loss of generality through reorientation of our coordinate frame, we can assume that $n' \geq 0$. Next, we make the transformation

$$\hat{y} = -\gamma^{-2}\left(n^2 - n_o^2\right) = -\varpi_0^2\gamma^{-2} - k\gamma(x - x_0) \tag{4.5-3}$$

where the constants ϖ_0 and γ are given by

$$\left.\begin{aligned} \varpi_0^2 &= n_0^2 - n_o^2 \\ \gamma &= \left(2k^{-1}n_0 n'\right)^{1/3} \end{aligned}\right\} \tag{4.5-4}$$

Equation (4.5-4) allows the possibility of ϖ_0^2 being negative; thus, ϖ_0 would be imaginary in this case. We show later that a negative value for ϖ_0^2 corresponds to a region where quantum tunneling applies; there $\hat{y}$ is positive and the amplitude of the electromagnetic field exponentially decays with increasing $\hat{y}$.

With this transformation in Eq. (4.5-3), the wave equations in Eq. (4.5-2) become

$$\left.\begin{aligned} \frac{d^2U}{d\hat{y}^2} - \hat{y}U = 0 \\ \frac{dU}{d\hat{y}} + i\gamma^{-1}V = 0 \end{aligned}\right\} \tag{4.5-2′}$$

Here we have set $\mu \equiv 1$. The solutions to these differential equations are the Airy functions and their derivatives; that is,

$$\left.\begin{aligned} U(\hat{y}) &= \left\{\mathrm{Ai}[\hat{y}], \mathrm{Bi}[\hat{y}]\right\} \\ V(\hat{y}) &= i\gamma\left\{\mathrm{Ai}'[\hat{y}], \mathrm{Bi}'[\hat{y}]\right\} \end{aligned}\right\} \tag{4.5-5}$$

In matrix form, the solutions are given by

$$\begin{pmatrix} U \\ V \end{pmatrix} = \boldsymbol{M}[\hat{y}, \hat{y}_0]\begin{pmatrix} U_0 \\ V_0 \end{pmatrix} \tag{4.5-6a}$$

The elements of the characteristic matrix $\boldsymbol{M}[\hat{y}, \hat{y}_0]$ are given by

$$\begin{aligned} \boldsymbol{M}[\hat{y}, \hat{y}_0] &= \begin{bmatrix} F(\hat{y}, \hat{y}_0) & f(\hat{y}, \hat{y}_0) \\ G(\hat{y}, \hat{y}_0) & g(\hat{y}, \hat{y}_0) \end{bmatrix} \\ &= \pi\begin{bmatrix} \begin{vmatrix} \mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}] \\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0] \end{vmatrix} & \frac{i}{\gamma}\begin{vmatrix} \mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}] \\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0] \end{vmatrix} \\ i\gamma\begin{vmatrix} \mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}] \\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0] \end{vmatrix} & -\begin{vmatrix} \mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}] \\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0] \end{vmatrix} \end{bmatrix} \end{aligned} \tag{4.5-6b}$$

where $\hat{y}$ is given in terms of x through Eq. (4.5-3) and $\hat{y}(x_0) = \hat{y}_0$. We note that $Fdg/d\hat{y} - fdG/d\hat{y} = 0$, $Gdf/d\hat{y} - gdF/d\hat{y} = 0$, or $d(Fg - Gf)/d\hat{y} = 0$, which is a consequence of the determinant of $\boldsymbol{M}[\hat{y}, \hat{y}_0]$ being a constant. We have

used the Wronskian of $\mathrm{Ai}[\hat{y}]$ and $\mathrm{Bi}[\hat{y}]$, $\mathrm{Ai}[\hat{y}]\mathrm{Bi}'[\hat{y}]-\mathrm{Ai}'[\hat{y}]\mathrm{Bi}[\hat{y}]=\pi^{-1}$, so that $\boldsymbol{M}\left[\hat{y}_0,\hat{y}_0\right]=\boldsymbol{I}$.

For this special case where the index of refraction is given by Eq. (4.5-1), we have

$$\varpi^2(x)=n^2(x)-n_o^2=\varpi_0^2+k\gamma^3\left(x-x_0\right)=-\gamma^2\hat{y} \tag{4.5-7}$$

and also

$$\frac{d\varpi}{dx}=k\gamma^2\left(-4\hat{y}\right)^{-1/2} \tag{4.5-8}$$

and

$$k\int_{x_0}^{x}\varpi(\zeta)d\zeta=\frac{2}{3}\left(\left(-\hat{y}\right)^{3/2}-\left(-\hat{y}_0\right)^{3/2}\right) \tag{4.5-9}$$

When $\hat{y}<<0$, $\hat{y}$ may be interpreted in terms of the angle of incidence φ by $\pm\gamma\left(-\hat{y}\right)^{1/2}=n\cos\varphi$. In geometric optics, $\hat{y}=\hat{y}_o\equiv 0$ corresponds to a turning point where $\varphi=\pi/2$. We will discuss turning points in more detail later for the TE wave.

When $\left(\varpi_0/\gamma\right)^2>>-k\gamma\left(x-x_0\right)$, then $y\le y_0<<0$ and we can use the negative argument asymptotic forms for the Airy functions given in Eq. (3.8-4) [7]. The elements of the characteristic matrix become

$$\left.\begin{aligned}
&F\left(\hat{y},\hat{y}_0\right)\to\left(\frac{\hat{y}_0}{\hat{y}}\right)^{1/4}\cos\left(X-X_0\right),\quad f\left(\hat{y},\hat{y}_0\right)\to\frac{i}{\gamma}\left(\frac{1}{\hat{y}\hat{y}_0}\right)^{1/4}\sin\left(X-X_0\right)\\
&G\left(\hat{y},\hat{y}_0\right)\to i\gamma\left(\hat{y}\hat{y}_0\right)^{1/4}\sin\left(X-X_0\right),\quad g\left(\hat{y},\hat{y}_0\right)\to\left(\frac{\hat{y}}{\hat{y}_0}\right)^{1/4}\cos\left(X-X_0\right)
\end{aligned}\right\} \tag{4.5-10}$$

where from Eq. (4.5-9) we have

$$\left.\begin{aligned}
&X=2/3\left(-\hat{y}\right)^{3/2}+\pi/4,\quad X-X_0=\mathscr{A}\left(x,x_0\right)=k\int_{x_0}^{x}\varpi(u)du,\\
&\hat{y}=-\gamma^{-2}\varpi^2(x),\quad \hat{y}\left(x_0\right)=\hat{y}_0=-\gamma^{-2}\left(n_0^2-n_o^2\right),\ \hat{y}\left(x_o\right)=\hat{y}_o=0,\\
&\varpi\left(x_0\right)=\varpi_0
\end{aligned}\right\} \tag{4.5-11}$$

Thus, the asymptotic form for the characteristic matrix for negative values of $\hat{y}$ becomes

$$\boldsymbol{M}\left[\hat{y},\hat{y}_0\right]\sim\begin{bmatrix}\left(\dfrac{\varpi_0}{\varpi_F}\right)^{1/2}\cos\mathscr{A} & \dfrac{i}{\left(\varpi_0\varpi_F\right)^{1/2}}\sin\mathscr{A}\\ i\left(\varpi_0\varpi_F\right)^{1/2}\sin\mathscr{A} & \left(\dfrac{\varpi_F}{\varpi_0}\right)^{1/2}\cos\mathscr{A}\end{bmatrix} \tag{4.5-12}$$

If we define the mean value of ϖ as $\overline{\varpi}=\left(\varpi_F\varpi_0\right)^{1/2}$, then Eq. (4.5-12) becomes

$$\boldsymbol{M}\left[\hat{y},\hat{y}_0\right]\sim\begin{bmatrix}\dfrac{\varpi_0}{\overline{\varpi}}\cos\mathscr{A} & \dfrac{i}{\overline{\varpi}}\sin\mathscr{A}\\ i\overline{\varpi}\sin\mathscr{A} & \dfrac{\varpi_F}{\overline{\varpi}}\cos\mathscr{A}\end{bmatrix} \tag{4.5-13}$$

It follows (away from turning points) that $\boldsymbol{M}\left[\hat{y}_F,\hat{y}_0\right]$ in Eq. (4.5-13) matches to first order in $\left(\varpi_F-\varpi_0\right)$ the characteristic matrix given in Eq. (4.4-17), which is derived from a more general but linear theory based on the thin-film approach.

Near a turning point, the limiting form in Eq. (4.5-13) breaks down when $\hat{y}>\sim -2$, or at $x=x_0$, when $\varpi_0\approx\varpi^*=\gamma\sqrt{2}$. For a typical value for γ in the Earth's lower troposphere, $\varpi^*\approx 0.002$. We note that in geometric optics $\hat{y}=-\left(n\cos\varphi/\gamma\right)^2$ and that the altitude $\hat{y}=\hat{y}_o=0$ corresponds to a turning point. For $\varpi^*\approx 0.002$, it follows for the spherical case that when θ lies within ~0.1 deg of the turning point of a ray, the full Airy solution in Eq. (4.5-6) or its equivalent provides better accuracy. However, for angles of incidence in the range $0\le\varphi<\sim\pi/2-\cos^{-1}\left[\gamma\sqrt{2}\right]$, the asymptotic form for the characteristic matrix is adequate.

The TM case for constant gradient in the index of refraction follows in a similar way, but the equations are somewhat more complicated because of the presence of the $\nabla\left[\boldsymbol{E}\cdot(\nabla(\log\ \varepsilon)\right]$ term in the modified wave equation, which is absent in the TE case. However, for small gradients, it can be shown to first order in n' that the characteristic matrix elements for the TM case are given by

$$\left.\begin{aligned} F(\hat{y},\hat{y}_0) &= \pi\frac{n}{n_0}\begin{vmatrix}\mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}]\\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0]\end{vmatrix}, \qquad f(\hat{y},\hat{y}_0) = i\pi\frac{nn_0}{\gamma}\begin{vmatrix}\mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}]\\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0]\end{vmatrix},\\ G(\hat{y},\hat{y}_0) &= i\frac{\pi\gamma}{nn_0}\begin{vmatrix}\mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}]\\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0]\end{vmatrix}, \qquad g(\hat{y},\hat{y}_0) = -\pi\frac{n_0}{n}\begin{vmatrix}\mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}]\\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0]\end{vmatrix},\\ \hat{y} &= -\varpi_{TM}^2\gamma^{-2} - \gamma(x-x_0), \qquad y(x_0) = \hat{y}_0 \end{aligned}\right\} \quad (4.5\text{-}14)$$

To convert these matrix elements into field components, Eqs. (4.2-7′) and (4.3-3) apply. For negative $\hat{y}$ values, it is readily shown that the asymptotic forms for the matrix elements in Eq. (4.5-14) approach to first order in $\varpi - \overline{\varpi}$ the values for the elements in Eq. (4.4-17′), which apply for the TM case.

4.6 Incoming and Outgoing Waves and Their Turning Points

Figure 4-2 shows ray paths for waves that are trending from left to right, and which are initially planar during their approach phase. In this example, we have two stratified layers infinite in extent parallel to the *yz*-plane with the index of refraction continuous across their boundary. The index of refraction is linearly varying in the lower medium, and it is constant in the upper medium. Across the boundary, n is continuous. The height $x = x_*$ marks the boundary between these two regimes; it is sufficiently far above the turning point height at $x = x_o$ (or, equivalently, with an angle of incidence sufficiently steep) so that asymptotic forms for the Airy functions can be applied to the characteristic

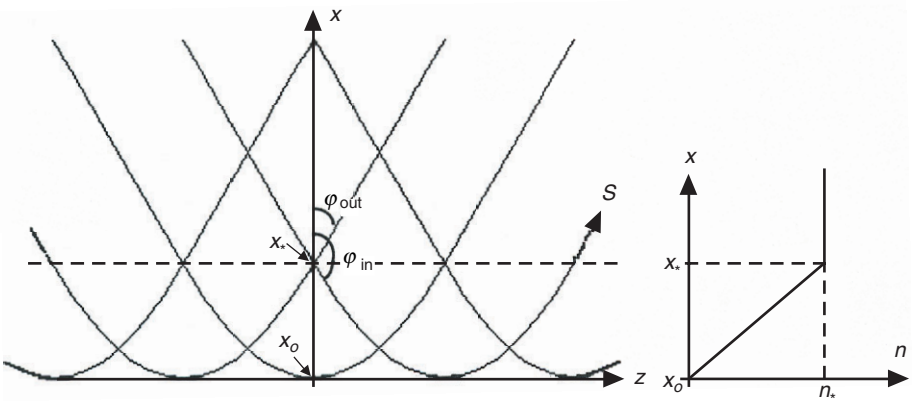

Fig. 4-2. Eikonal paths of incoming and outgoing waves in a two-layered Cartesian stratified medium. In the lower layer, the gradient of *n* is constant; in the upper layer, $n = n_* =$ constant.

matrix for heights at and above this boundary. Passing through any point above the turning point boundary at $x = x_o$ are both incoming and outgoing waves. The total field is given by the vector sum of these two waves.

For a planar TE wave impinging at an angle of incidence φ on a stratified surface in the homogeneous medium at or above the boundary $x = x_*$, the time-independent component of the electric field is given by

$$E = \exp\bigl(ikn_*\bigl(x\cos\varphi + z\sin\varphi + \text{constant}\bigr)\bigr) \tag{4.6-1}$$

Applying Maxwell's equations in Eq. (4.2-1a) to Eq. (4.6-1), one may express the angle of incidence in terms of the field components by

$$\tan\varphi = -\frac{H_x}{H_z} \tag{4.6-2}$$

For an incoming wave, H_x and H_z have the same polarity: they are either both positive or both negative (see Fig. 4-1). Therefore, it follows from Eq. (4.6-2) that $\pi/2 < \varphi \le \pi$. For an outgoing wave, H_x and H_z have opposite polarities, and it follows that $0 \le \varphi < \pi/2$.

At the boundary $x = x_*$, we have from Eqs. (4.2-8) and (4.2-9) for a planar wave

$$\tan\varphi_* = -\frac{W_*}{V_*} = n_o\frac{U_*}{V_*}, \quad \mu \equiv 1 \tag{4.6-3}$$

Using Snell's law in Eq. (4.2-14), it follows from Eq. (4.6-3) that

$$V_*^{\pm} = \pm n_* \,|\cos\varphi_*\,| U_*^{\pm} = \pm\gamma_*\sqrt{-\hat{y}_*}\,U_*^{\pm} \tag{4.6-4}$$

where the second relationship follows from Eq. (4.5-3). Here "+" denotes an outgoing wave where $0 \le \varphi < \pi/2$, and "–" denotes an incoming wave where $\pi/2 < \varphi \le \pi$.

When these constraining conditions between U_* and V_* are applied to the characteristic matrix solution for U and V given in Eq. (4.5-6), one obtains

$$\begin{pmatrix} U^{\pm} \\ V^{\pm} \end{pmatrix} = \pi\begin{pmatrix} \bigl(\mathrm{AiBi}_*' - \mathrm{Ai}_*'\mathrm{Bi}\bigr) \pm i\sqrt{-\hat{y}_*}\bigl(\mathrm{AiBi}_* - \mathrm{Ai}_*\mathrm{Bi}\bigr) \\ \gamma\Bigl(i\bigl(\mathrm{Ai}'\mathrm{Bi}_*' - \mathrm{Ai}_*'\mathrm{Bi}'\bigr) \mp \sqrt{-\hat{y}_*}\bigl(\mathrm{Ai}'\mathrm{Bi}_* - \mathrm{Ai}_*\mathrm{Bi}'\bigr)\Bigr) \end{pmatrix} U_*^{\pm} \tag{4.6-5}$$

Here $\mathrm{Ai} = \mathrm{Ai}[\hat{y}]$ and $\mathrm{Ai}_* = \mathrm{Ai}[\hat{y}_*]$; similarly for Bi. We now apply at the boundary $x = x_*$ the asymptotic forms in Eq. (3.8-4) for the Airy functions, which require that $\hat{y}_*$ be sufficiently negative. One obtains

$$\begin{pmatrix} U^{\pm} \\ V^{\pm} \end{pmatrix} \doteq \sqrt{\pi}\left(-\hat{y}_*\right)^{1/4} \begin{pmatrix} \mathrm{Ai}\left[\hat{y}\right] \mp i\,\mathrm{Bi}\left[\hat{y}\right] \\ \mathrm{i}\,\gamma\left(\mathrm{Ai}'\left[\hat{y}\right] \mp i\,\mathrm{Bi}'\left[\hat{y}\right]\right) \end{pmatrix} U_*^{\pm} \exp\left[\pm iX_*\right] \tag{4.6-6}$$

Here X_* is the phase accumulation along the x direction from the turning point at x_o up to x_* (plus $\pi/4$), which from Eq. (4.5-11) is given by

$$X_* = k\int_{x_o}^{x_*} \varpi dx + \frac{\pi}{4} \tag{4.6-7}$$

We note from Eq. (4.6-6) that the phases of $U^{\pm}$ and $V^{\pm}$ approach constant values for increasing $\hat{y} > 0$.

The total field at any point is given by adding the incoming and outgoing components. This sum must be devoid of the Airy function of the second kind, $\mathrm{Bi}\left[\hat{y}\right]$, in order to satisfy the physical "boundary condition" that there be a vanishing field below the turning point line, that is, in the region where $\hat{y}$ is positive; only $\mathrm{Ai}\left[\hat{y}\right]$ vanishes for positive $\hat{y}$. A turning point represents a kind of grazing reflection. From Eq. (4.6-6), we see that we can null the $\mathrm{Bi}\left[\hat{y}\right]$ component in the sum of the incoming and outgoing components if we set

$$U_*^{+} \exp\left[iX_*\right] = U_*^{-} \exp\left[-iX_*\right] \tag{4.6-8}$$

Thus, U_*^{+} equals U_*^{-} plus a phase delay Φ that is given by

$$\Phi = 2k\int_{x_o}^{x_*} \varpi dx + \frac{\pi}{2} \tag{4.6-9}$$

which is the round-trip phase delay along the x-direction (plus $\pi/2$).

Alternatively, from symmetry considerations it follows that U_*^{+} is the complex conjugate of U_*^{-} plus a correction arising from the variability of the index of refraction in the lower medium. It is easily shown from Eqs. (4.6-7) and (4.6-8) that if the incoming planar wave at the boundary has the form $U_*^{-} = \exp[ik\varpi_* x_* + \pi/4]$, then the outgoing planar wave at the boundary has the form

$$U_*^{+} = \exp\left[-i\left(k\varpi_* x_* + \frac{\pi}{4}\right)\right] \exp\left[i2k\int_{x_o}^{x_*} \varpi' x dx\right] \tag{4.6-10}$$

There is another way of interpreting incoming and outgoing waves using the characteristic matrix. From Eq. (4.5-6) and the Wronskian for the Airy functions, it follows that

$$\left.\begin{aligned} &\text{If } \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} = C \begin{pmatrix} \mathrm{Ai}[\hat{y}_0] \\ i\gamma\, \mathrm{Ai}'[\hat{y}_0] \end{pmatrix}, \\ &\text{then } \begin{pmatrix} U \\ V \end{pmatrix} = \boldsymbol{M}[\hat{y}, \hat{y}_0] \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} = C \begin{pmatrix} \mathrm{Ai}[\hat{y}] \\ i\gamma\, \mathrm{Ai}'[\hat{y}] \end{pmatrix} \quad \forall \quad \hat{y} \end{aligned}\right\} \tag{4.6-11}$$

Here C is a constant. If we let $\hat{y}$ be sufficiently negative so that the asymptotic forms for the Airy functions apply, and we set $C = 2i\pi^{1/2}$, then it follows that

$$\left.\begin{aligned} &\begin{pmatrix} U \\ V \end{pmatrix} \to (-\hat{y})^{\frac{1}{4}} \begin{pmatrix} \exp(i\,\mathrm{X}) - \exp(-i\,\mathrm{X}) \\ \big(\exp(i\,\mathrm{X}) + \exp(-i\,\mathrm{X})\big) n \cos\varphi \end{pmatrix} \\ &\qquad\qquad\quad \uparrow \qquad\quad \uparrow \\ &\qquad\qquad \text{Outgoing} \quad \text{Incoming} \\ &\mathrm{X} = \frac{2}{3}(-\hat{y})^{3/2} + \frac{\pi}{4} \end{aligned}\right\} \tag{4.6-12}$$

Thus, U and V at the position $\hat{y}$ can be interpreted as consisting of the *superposition* of an incoming wave with a phase $-X(\hat{y}) - \pi$ and an angle of incidence $\pi - \varphi$, and an outgoing wave with a phase $X(\hat{y})$ and an angle of incidence φ.

Using Eq. (4.6-8), the total field at any point is given in terms of the incoming planar wave at the boundary by the expressions

$$\begin{pmatrix} U \\ V \end{pmatrix} = \begin{pmatrix} U^+ \\ V^+ \end{pmatrix} + \begin{pmatrix} U^- \\ V^- \end{pmatrix} = 2\sqrt{\pi}(-\hat{y}_*)^{1/4} \begin{pmatrix} -i\,\mathrm{Ai}[\hat{y}] \\ \gamma\, \mathrm{Ai}'[\hat{y}] \end{pmatrix} U_*^- \exp(+iX_*) \tag{4.6-13}$$

We note that V describes the tangential component of the magnetic field H_z in the TE case, that is, the component parallel to the plane of stratification. It must change sign near a turning point; in fact, in geometric optics it changes sign precisely at a turning point. Here this occurs when $\mathrm{Ai}'[\hat{y}] = 0$ for the last time before $\hat{y}$ becomes positive, which occurs at $\hat{y} = -1.02$. For heights below the turning point, that is, where $\hat{y}$ becomes positive, Eq. (4.6-13) shows that both U and V decay exponentially with increasing $\hat{y}$. In fact, these asymptotic forms for positive $\hat{y}$ are given in Eq. (3.8-4). They are

$$\left.\begin{aligned} \mathrm{Ai}[\hat{y}] &\to \frac{1}{2}\pi^{-1/2}\hat{y}^{-1/4} \exp\left(-\frac{2}{3}\hat{y}^{3/2}\right) \\ \mathrm{Ai}'[\hat{y}] &\to -\frac{1}{2}\pi^{-1/2}\hat{y}^{1/4} \exp\left(-\frac{2}{3}\hat{y}^{3/2}\right) \end{aligned}\right\} \tag{4.6-14}$$

Another way of viewing turning points is to consider the angle φ which the surface of constant phase (the cophasal surface) of the electric field for an incoming wave makes with the *yz*-plane. This angle is given by Eq. (4.2-13), where ψ is the phase of $U^{-}(y)$. From Eq. (4.6-6), it follows that the phase of the electric field for an incoming TE wave is given by

$$\psi = \tan^{-1}\left(\frac{\mathrm{Bi}[\hat{y}]}{\mathrm{Ai}[\hat{y}]}\right) - \mathrm{X}_* - kn_* x_* \cos\varphi_* \tag{4.6-15}$$

Using Eq. (4.5-3) for the relationship between $\hat{y}$ and x, and also the Wronskian for the Airy functions, it follows that

$$\frac{d\psi}{dx} = -\frac{k\gamma}{\pi\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)} \tag{4.6-16}$$

Therefore, from Eq. (4.2-13), φ is given by

$$\tan\varphi = -\frac{\pi n_o}{\gamma}\left(\mathrm{Ai}^2[\hat{y}] + \mathrm{Bi}^2[\hat{y}]\right) = -\pi\tan\varphi_*\left(\mathrm{Ai}^2[\hat{y}] + \mathrm{Bi}^2[\hat{y}]\right)\left(-\hat{y}_*\right)^{1/2} \tag{4.6-17}$$

Thus, $\varphi \to \pi - \varphi_*$ as $\hat{y} \to \hat{y}_*$, as it must for an incoming plane wave with $\hat{y}$ sufficiently negative. Also, it follows that $\varphi \to \pi^+/2$ rapidly for increasing $\hat{y} > 0$.

From Eq. (4.6-6), we can evaluate the field components at any point, in particular at x_0 to obtain U_0 and V_0. Then we can write the field components evaluated at x in the form

$$\left.\begin{aligned} U^{\pm} &= \frac{\mathrm{Ai}[\hat{y}] \mp i\,\mathrm{Bi}[\hat{y}]}{\mathrm{Ai}[\hat{y}_0] \mp i\,\mathrm{Bi}[\hat{y}_0]} U_0^{\pm} \\ V^{\pm} &= \frac{\mathrm{Ai}'[\hat{y}] \mp i\,\mathrm{Bi}'[\hat{y}]}{\mathrm{Ai}'[\hat{y}_0] \mp i\,\mathrm{Bi}'[\hat{y}_0]} V_0^{\pm} \end{aligned}\right\} \tag{4.6-18}$$

We note the form $\mathrm{Ai}[\hat{y}] - i\,\mathrm{Bi}[\hat{y}]$ that is associated with an outgoing wave, and also the form $\mathrm{Ai}[\hat{y}] + i\,\mathrm{Bi}[\hat{y}]$ that is associated with an incoming wave. These are, of course, the same expressions contained in the asymptotic forms for the spherical Hankel functions [see Eq. (3.8-1)]. Thus, for spherical Hankel functions of the first kind, $\xi_l^+(\rho)/\rho$ is associated with outgoing waves because in the limit for large $\rho >> \nu$ it asymptotically approaches the form of an outgoing spherical wave. Similarly, for spherical Hankel functions of the

second kind, $\xi_l^-(\rho)/\rho$ approaches the form of an incoming wave, i.e., $\xi_l^{\pm}(\rho)/\rho \to i^{\pm(l+1)}\exp[\pm i\rho]/\rho,\ \rho >> l$. This formalism of identifying $\xi_l^+(\rho)$ with outgoing waves and $\xi_l^-(\rho)$ with incoming waves has already been used in Chapter 3 for scattering from a sphere. It was first pointed out by Hermann Hankel himself about 140 years ago.

4.6.1 Eikonal and Cophasal Normal Paths

In geometric optics, the optical path length $\mathscr{S}$ for a ray connecting two points is defined by

$$\mathscr{S} = \int n ds \tag{4.6-19}$$

where s is path length along a ray, and the integral is a path integral along the ray between the initial point and the end point. The path length vector infinitesimal $d\mathbf{s}$ at any point on the path is defined by

$$d\mathbf{s}/ds = \lim_{\lambda \to 0}[\mathbf{S}/S] \tag{4.6-20}$$

where $\boldsymbol{S} = c(\boldsymbol{E}\times\boldsymbol{H})/4\pi n$ is the Poynting vector and λ is the wavelength of the electromagnetic wave. The Poynting vector is perpendicular to the wave front at any point, and its limiting form follows the path defined by the eikonal equation in geometric optics. (See [3] for a discussion of the foundations of geometric optics.)

$\mathscr{S}$ may be considered as a field quantity $\mathscr{S} = \mathscr{S}(x,y,z)$, which is associated with a family of ray paths passing through space. By varying the initial values of the rays, one generates a family of rays—for example, the family shown in Fig. 4-2. $\mathscr{S}$ is akin to the action integral in Hamilton-Jacobi theory or to the Feynman path integral in the sum-over-histories approach to quantum electrodynamics. $\mathscr{S}$ is a function only of the end point of the trajectory along which the path integral is evaluated. In geometric optics, it is the phase accumulated by following a Fermat path, that is, a path of stationary phase. Thus, the phase that would be accumulated by following any alternative path neighboring the Fermat path, but having the same end coordinates, assumes a stationary value $\mathscr{S}$ when the Fermat path is in fact followed. The evolution of $\mathscr{S}(x,y,z)$ as a field variable is governed by the eikonal equation[1]

[1] The eikonal equation is related to the Hamilton-Jacobi partial differential equation, which arises in the Hamiltonian formulation of the Calculus of Variations problem for a Fermat path. In this formulation, a six-dimensional system of first-order ordinary differential equations in coordinate/conjugate momentum space determines a Fermat path in this six-dimensional space. The Hamilton-Jacobi equation describes the

$$|\nabla \mathscr{S}|^2 = n^2 \tag{4.6-21}$$

A surface $\mathscr{S}(x,y,z) = \text{constant}$ defines a wave front (in a geometric optics context), that is, a surface of constant phase across which a continuum of eikonal paths transect. The eikonal path through any point on the surface $\mathscr{S}(x,y,z) = \text{constant}$ is normal to it. The gradient $\nabla\mathscr{S}/|\nabla\mathscr{S}|$ is the unit tangent vector for an eikonal path. $\nabla\mathscr{S}/|\nabla\mathscr{S}|$ and the limiting form for the Poynting vector $\boldsymbol{S}$, as the wavelength of the wave approaches zero, are parallel; their magnitudes are related through a scale factor that equals the average electromagnetic power density of the wave.

For the case of Cartesian stratification with n' a constant, $\nabla\mathscr{S}$ is given by

$$\nabla\mathscr{S} = k^{-1}\left(\hat{\boldsymbol{x}}\ \frac{d\mathscr{A}}{dx} + \hat{\boldsymbol{z}}\frac{d\mathscr{C}}{dz}\right) = \hat{\boldsymbol{x}}\sqrt{n^2 - n_o^2} + \hat{\boldsymbol{z}}n_o \tag{4.6-22}$$

where $\mathscr{C}(z)$ is the phase accumulation of the time-independent component of the wave along the z-direction, that is, $\mathscr{C} = kn_o z$. $\mathscr{A}(x)$ is the phase accumulation along the x-direction, and it is given by Eq. (4.4-13). One can obtain $\mathscr{S}(x,y,z)$ from an integration of this gradient equation.

Note also that $\mathscr{S}(x,z)$ may not be unique, as is the case for the family of ray paths shown in Fig. 4-2. Since two ray paths pass through every point (x,z) above the altitude of the turning point, there are two functions, $\mathscr{S}^-(x,z)$ for the incoming path and $\mathscr{S}^+(x,z)$ for the outgoing path. Since the outgoing path has already touched the turning point line, which is a caustic surface in geometric optics and, therefore, an envelope to the system of ray paths, the outgoing path violates the Jacobi condition from the Calculus of Variations. This is a necessary condition that a stationary path must satisfy to provide a local minimum in the action integral, in this case the phase accumulation $\mathscr{S}$. In this example, $\mathscr{S}^+(x,z)$ provides a local maximum at the point (x,z).

For the incoming TE wave shown in Fig. 4-2, we can obtain the path generated by the normal to the cophasal surface of the electric field at any point, which essentially matches the eikonal path except near the turning point. In the plane of incidence, the coordinates for the normal path are given by its differential equation $dx/dz = \cot\varphi$, where φ is the angle of incidence and is given by Eq. (4.2-13). When the gradient of n is a constant, we obtain from Eqs. (4.5-4) and (4.6-17) for an Airy layer

---

behavior of the stationary phase at the end point of the Fermat path over a region in this space that is spanned by a family of rays. The eikonal equation provides similar information in three-dimensional coordinate space for this family of rays.

$$\frac{dx}{dz} = -\gamma\left(\left(\mathrm{Ai}^2[\hat{y}] + \mathrm{Bi}^2[\hat{y}]\right)\pi n_o\right)^{-1} = -\gamma^{-1}k^{-1}\frac{d\hat{y}}{dz} \tag{4.6-23}$$

from which it follows that

$$\left.\begin{aligned} z &= \frac{\pi n_o}{k\gamma^2}\int\left(\mathrm{Ai}^2[\hat{y}] + \mathrm{Bi}^2[\hat{y}]\right)d\hat{y} + \text{constant} \\ &= \frac{\pi n_o}{k\gamma^2}\left(\left(\mathrm{Ai}^2[\hat{y}] + \mathrm{Bi}^2[\hat{y}]\right)\hat{y} - \left(\mathrm{Ai}'^2[\hat{y}] + \mathrm{Bi}'^2[\hat{y}]\right)\right) \\ &\xrightarrow[\hat{y}<<0]{} -\frac{2n_o}{k\gamma^2}\sqrt{-\hat{y}} \end{aligned}\right\} \tag{4.6-24}$$

Here we have chosen a particular normal path defined by the condition that $z = 0$ when $\hat{y} = \hat{y}^\dagger$, where $\hat{y}^\dagger$ is given by

$$\hat{y}^\dagger = \frac{\mathrm{Ai}'^2\left[\hat{y}^\dagger\right] + \mathrm{Bi}'^2\left[\hat{y}^\dagger\right]}{\mathrm{Ai}^2\left[\hat{y}^\dagger\right] + \mathrm{Bi}^2\left[\hat{y}^\dagger\right]} = 0.44133\cdots \tag{4.6-25}$$

It follows from Eqs. (4.5-3) and (4.6-24) that in the region where $\hat{y} << 0$, but in the lower layer, the normal path is given by

$$x - x_o = \frac{n'}{2n_o}z^2, \quad \hat{y} << 0 \tag{4.6-26}$$

From geometric optics, we have Snell's law $n\sin\theta = n_o$; it follows for an incoming wave that the eikonal path is given by

$$\frac{dx}{dz} = -\frac{\sqrt{n^2 - n_o^2}}{n} = -\frac{\sqrt{2n_o n'(x - x_o)}}{n_o + n'(x - x_o)} \tag{4.6-27}$$

Integrating yields

$$z = -2\sqrt{\frac{n_o(x - x_o)}{2n'}} - \frac{2}{3}\sqrt{\frac{n'(x - x_o)^3}{2n_o}} \tag{4.6-28}$$

When $n' << 1$, we can drop the second term, and this leads to essentially the same result given in Eq. (4.6-26) for the normal path for large negative $\hat{y}$. Only in the vicinity of the turning point (and below) does a significant divergence occur between an eikonal or ray path and the path generated by a normal to the cophasal surface of the electromagnetic wave.

Geometric optics is a second-order ray theory. The accuracy of geometric optics as a ray theory (and, therefore, as an approximate description of the cophasal normal path in wave theory) depends on the second variation of $\mathscr{S}(x,z)$ with respect to deviations from the nominal ray path being sufficiently large. This is equivalent to requiring the Fresnel approximation in stationary phase theory to be sufficiently accurate when applied to the nominal ray path. It can be shown in the example given in Fig. 4-2 that at $\hat{y} = \hat{y}_o = 0$ a caustic surface is generated where the second variation of $\mathscr{S}(x,z)$ is zero. A ray path having an interior contact point with a caustic surface usually is troublesome for second-order ray theory. Note that the equation in Eq. (4.6-24) forces z for the cophasal normal path to become very positive (for an incoming wave) when $\hat{y} > \hat{y}^{\dagger}$.

4.6.2 Defocusing

We also can calculate the "defocusing" caused by this refracting medium. We have represented the electromagnetic field of the harmonic wave by complex forms such as $U(x)\exp\left[i\left(kn_o z - \omega t\right)\right]$. However, we want a given physical property of these forms, such as electromagnetic energy, to be real. It is convenient for harmonic waves to use the complex Poynting vector [8] defined by $\boldsymbol{S} = c(\boldsymbol{E} \times \boldsymbol{H}^*)/8\pi$, where $\boldsymbol{H}^*$ is the complex conjugate of $\boldsymbol{H}$. The definition of $\boldsymbol{S}$ here includes a 1/2 term to reflect the root-mean-square energy flow of a harmonic wave over a cycle. The real part of $\boldsymbol{S}$, which is defined by $\mathrm{Re}[\boldsymbol{S}] = (\boldsymbol{S} + \boldsymbol{S}^*)/2$, gives the time-averaged flow of electromagnetic energy of the wave across a mathematical surface normal to $\boldsymbol{S}$. In Gaussian units, the dimensions of $\boldsymbol{E}$ and $\boldsymbol{H}$ are $(\text{mass}/\text{length})^{1/2}/\text{time}$. Therefore, $\boldsymbol{S}$ has the dimensions of power per unit area. At the boundary $x = x_*$ between the layers, this average energy flow for the incoming and outgoing planar waves is given by

$$8\pi\left\langle S_*^{\pm}\right\rangle = cn_*\left(\hat{z}\sin\theta_* \mp \hat{x}\cos\theta_*\right) \tag{4.6-29}$$

Here $\left|U_*^-\right|^2$, which has the dimensions of $\text{mass}/\left(\text{length}*\text{time}^2\right)$, has been set to unity. The energy flow of the superposition of the two waves is given by

$$8\pi\left\langle \mathrm{Re}\left[S_*^{\pm}\right]\right\rangle = \hat{z}cn_*\sin\theta_* \tag{4.6-30}$$

Thus, the time-averaged energy flow at $x = x_*$ is equal to the root-mean-square energy density of the wave times the component of its velocity along the z-axis.

The ratio R of the average energy flows at two different altitudes is given by

$$R = \frac{\langle \mathrm{Re}[\boldsymbol{S}_o] \rangle}{\langle \mathrm{Re}[\boldsymbol{S}_*] \rangle} = \frac{\mathrm{Ai}[\hat{y}_o]^2}{\mathrm{Ai}[\hat{y}_*]^2} \tag{4.6-31a}$$

At the altitude $\hat{y} = -1.02, V = i\gamma \,\mathrm{Ai}'[y] = 0$, which corresponds to a "turning point." At this point, we have

$$\langle R \rangle = 0.287 \,\mathrm{Ai}[\hat{y}_*]^{-2} \underset{\hat{y}_* << 0}{\rightarrow} 1.8\sqrt{-\hat{y}_*} \rightarrow 1.8\gamma^{-1} \cos\theta_* \tag{4.6-31b}$$

This will be a large number for thin atmospheres except at near-grazing conditions. The square root of R yields the average voltage ratio of the electric fields at the two altitudes.

Figure 4-3 shows a family of eikonal paths generated from Eq. (4.6-24) by varying the angle of incidence φ_* at a given altitude x_*. The defocusing for a given path at a given point on the path is obtained by varying φ_*. To obtain the ratio of the displacement in x at the turning point and the perpendicular displacement of the path at another given point, one varies φ_* and calculates these two displacements. Their ratio gives the defocusing. The bold lines in Fig. 4-3 are the envelopes for this family of curves. These are caustic surfaces along which the defocusing is zero. A third-order or higher ray theory is required here to obtain the defocusing.

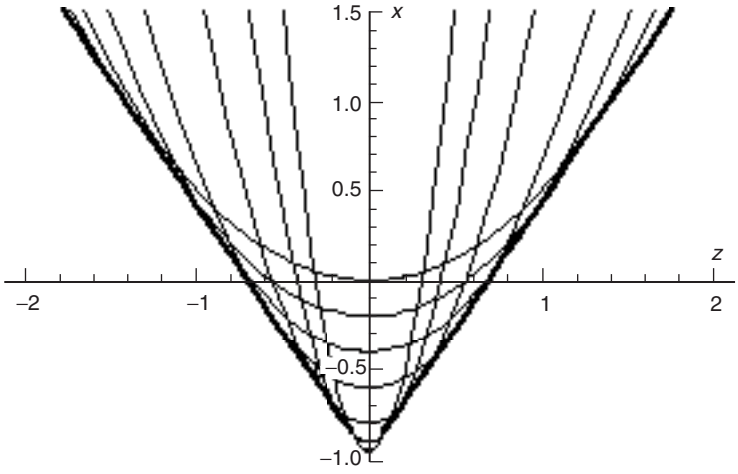

Fig. 4-3. Family of eikonal paths generated by varying the angle of incidence at a given altitude in an Airy layer.

4.7 Concatenated Airy Layers[2]

We now concatenate successive Cartesian layers; within a given layer the gradient of the index of refraction is held constant, but it is allowed to change from layer to layer. Because the elements of the characteristic matrix for each layer involve Airy functions, we call such a layer an "Airy layer." By varying n and n' across the boundaries between Airy layers and by allowing the thickness of the layers to shrink to zero while concomitantly allowing their number to grow infinitely large, we can create a discretely varying profile that matches a given continuously varying profile $n(x)$. In this approach, the gradient is stepwise constant, that is, it is constant within a layer but discontinuous across its boundary. In the former approach in Section 4.4, the gradient was zero but the index of refraction was stepwise constant. Using piecewise constant gradients in "onion skin" algorithms for recovery of atmospheric products from limb sounding data is considered to be more efficient than using zero gradient layers. This piecewise constant gradient approach may be useful near turning points.

Let us return to the wave equations for the Cartesian stratified case given in Eq. (4.2-10). We introduce the transformation

$$\left.\begin{aligned} \hat{y} &= -\gamma^{-2}\left(n^2 - n_o^2\right) \\ \gamma &= \left(2k^{-1}nn'\right)^{1/3} \end{aligned}\right\} \tag{4.7-1}$$

It follows that

$$\frac{d\hat{y}}{dx} = -k\gamma - \frac{2\hat{y}}{\gamma}\frac{d\gamma}{dx} \tag{4.7-2}$$

Within a given Airy layer, $\gamma_A(x)$ is a constant; hence, the wave equations in Eq. (4.5-4) and the Airy function formulation for the characteristic matrix given in Eq. (4.5-6) apply within that layer. The subscript "A" denotes the profile of the index of refraction and hence $\gamma_A(x)$ in the Airy layer. Within a layer, we match this piecewise constant function to the mean value of the actual profile $\gamma(x)$:

$$\gamma_A(x) = \frac{1}{2}\left(\gamma\left(x_{j-1}\right) + \gamma\left(x_j\right)\right), \quad x_{j-1} \le x < x_j \tag{4.7-3}$$

---

[2] This section offers an alternate basis to the Mie scattering approach for obtaining the forms of the osculating parameters in a stratified medium. Because the Mie scattering approach has been followed in Chapter 5, this section is not essential.

It follows from Eq. (4.7-2) that $\hat{y}_A$ varies linearly within an Airy layer because $\gamma_A(x)$ is a constant, but $\hat{y}_A$ is discontinuous across the boundary between layers. Examples of these profiles are shown in Fig. 4-4.

Within the *j*th Airy layer, $\hat{y}_A$ is given by

$$\hat{y}_A(x) = -\left(\gamma_A^{-2}\left(n^2 - n_o^2\right)\right)\Big|_{x_{j-1}} - k\gamma_A\big|_{x_{j-1}}\left(x - x_{j-1}\right), \quad x_{j-1} \le x < x_j \qquad \text{(4.7-4a)}$$

Across the upper boundary of the *j*th layer the discontinuity in $\hat{y}_A$ is given by

$$\Delta\hat{y}_A\big|_{x_j} = -\frac{2\hat{y}_A}{\gamma_A}\Delta\gamma_A\Big|_{x_j} \qquad \text{(4.7-4b)}$$

where $\Delta\gamma_A$ is the discontinuity required to maintain close tracking by $\gamma_A(x)$ of the actual profile $\gamma(x)$. $\Delta\gamma_A$ depends primarily on the curvature in $n(x)$. As

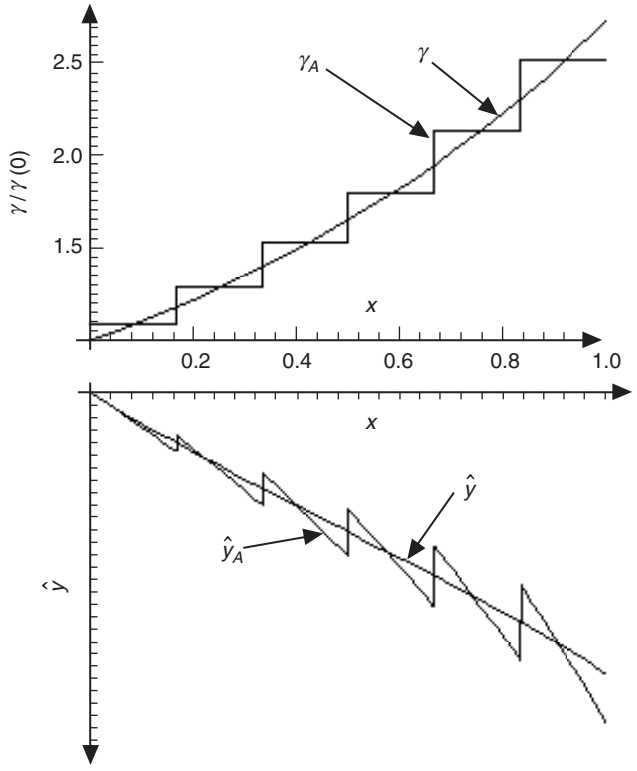

Fig. 4-4. Piecewise constant gradient profile for a series of Airy layers that approximate an exponential refractivity profile; x = 1 corresponds to 1 scale height above x = 0.

the maximum thickness of the Airy layers approaches zero and their number grows infinite, it follows that $\gamma_A(x) \to \gamma(x)$. From Eq. (4.7-4ab), it also follows that $\hat{y}_A$ will approach $\hat{y}(x)$ as defined by Eq. (4.7-1), but $\hat{y}'_A \neq \hat{y}'$. Accordingly, the characteristic matrix for the system of concatenated Airy layers spanning the space (x_0, x_N), and obtained from applying the product rule given in Eq. (4.4-2), will approach in the limit the characteristic matrix that applies to the actual medium with the profile $n(x)$.

In Cartesian stratification where n is held constant within a layer, Eq. (4.4-3) shows that sinusoidal functions result for the elements of the characteristic matrix. When the product rule in Eq. (4.4-2) is applied to obtain the reference characteristic matrix spanning multiple layers, Eq. (4.4-8) shows that sinusoids also result for the product. Here the argument of the sinusoids is the phase $\mathscr{A}$ accumulated along the x-axis. This "closed form" resulted from using the double angle formula, $\exp(iC) \times \exp(iD) = \exp[i(C + D)]$, to convert an infinite product of sinusoids into a single sinusoid with its argument containing an infinite sum, which can be represented by an integral. In this way, an infinite product of characteristic matrices is converted into a single characteristic matrix spanning all of the layers.

Unfortunately, no such "double angle" formula strictly applies to Airy functions. When $\hat{y}$ is sufficiently negative so that the sinusoidal asymptotic forms for the Airy functions can be used, then the double angle formula can be applied with sufficient accuracy. Near turning points where $\hat{y}$ is near zero, that is, where $x - x_o <\sim 2k^{-1}\gamma^{-1}$ (typically a few tens of meters for thin-atmosphere conditions with L-band signals), then the Airy functions themselves, or their equivalents, should be used. So, upon application of the product rule, we will need another strategy to convert an infinite product of characteristic matrices into a tractable form—in particular, into a single matrix spanning the entire sequence of layers.

The transitive property of characteristic matrices, $\boldsymbol{M}[\hat{y}_2, \hat{y}_1]\boldsymbol{M}[\hat{y}_1, \hat{y}_0] = \boldsymbol{M}[\hat{y}_2, \hat{y}_0]$, can be used even when the elements involve Airy functions. This is proved by using the Wronskian for the Airy functions. But this transitive property requires that the "$\hat{y}_1$" in $\boldsymbol{M}[\hat{y}_2, \hat{y}_1]$ have the same value as the "$\hat{y}_1$" in $\boldsymbol{M}[\hat{y}_1, \hat{y}_0]$. When one changes n' across the boundary $x = x_j$ between two Airy layers, Eq. (4.7-1) shows that $\hat{y}_A$ will change unless n_A and n'_A are concurrently changed to keep $\hat{y}_A$ invariant across the boundary. From Eq. (4.7-1), if we set $\delta\hat{y}_A = 0$, this requires the constraint

$$\left.\frac{\delta n'}{n'}\right|_{x_j} = \left(\frac{2n^2 + n_o^2}{n^2 - n_o^2}\right)\left.\frac{\delta n}{n}\right|_{x_j} \tag{4.7-5}$$

to hold between δn and $\delta n'$. It follows that

$$\left.\frac{\delta\gamma}{\gamma}\right|_{x_j} = \left.\left(\frac{n^2}{n^2-n_o^2}\frac{\delta n}{n}\right)\right|_{x_j} = \left.\left(\frac{n^2}{2n^2+n_o^2}\frac{\delta n'}{n'}\right)\right|_{x_j} \tag{4.7-6}$$

must hold across a boundary if $\hat{y}_A$ is to remain invariant.

Alternate strategies can be followed subject to the constraint in Eq. (4.7-5). Let us set a value for the change in the index of refraction across the boundary between Airy layers, Δn_A, so that at the layer boundary $x = x_j$ the value of the index of refraction for the piecewise constant gradient profile n_A matches the value $n(x_j)$ of the actual profile being modeled. One obtains a profile such as that shown in Fig. 4-5. In this scenario, it follows that at the jth boundary Δn_A is given by $\Delta n_A = \left(n'(x_{j-1}) - n'_{A_j}\right)(x_j - x_{j-1}) + \mathrm{O}\left[(x_j - x_{j-1})^2\right]$. Here n'_{A_j} is constant in the jth layer, and $\Delta n_A|_{x_j}$ is the required discontinuity at the jth boundary. Going to the limit, we obtain differential equations for $\hat{y}_A$ and γ_A that are given by

$$\left.\begin{aligned} &\frac{d\hat{y}_A}{dx} = -k\gamma_A, \quad \frac{1}{\gamma_A}\frac{d\gamma_A}{dx} = \frac{k}{2}\frac{\gamma^3-\gamma_A^3}{n^2-n_o^2}, \\ &\gamma_A^3 = 2k^{-1}n(x)n'_A(x), \quad \gamma^3 = 2k^{-1}n(x)n'(x) \end{aligned}\right\} \tag{4.7-7a}$$

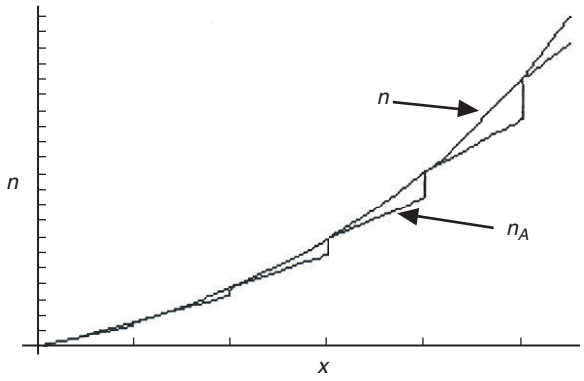

Fig. 4-5. Piecewise constant gradient profile for a series of Airy layers; $n_A(x)$ is matched to the actual profile $n(x)$, while keeping $\hat{y}$ continuous across the boundaries between layers.

It follows from Eqs. (4.7-1) and (4.7-7a) that

$$\hat{y}_A \gamma_A^2 = \hat{y}\gamma^2 \tag{4.7-7b}$$

The rate of phase accumulation along the *x*-axis is the same in the Airy layers as it is in the actual medium.

We are approximating the continuous index of refraction for the medium by an infinite stack of Airy layers. Suppose we set $\hat{y} = \hat{y}_0 = \hat{y}_o = 0$, and at this point we match the refractivity and its gradient to the actual profile that is being modeled, i.e., we equate $n_A(x_o) = n(x_o) = n_o$ and $n'_A(x_o) = n'(x_o)$; therefore, $\gamma_A(x_o) = \gamma(x_o)$. This provides boundary conditions for Eq. (4.7-7a) from which the profile for $\hat{y}_A$ and $d\hat{y}_A / dx$ can be obtained. In this manner, we can adjust n'_A in our piecewise constant gradient profile to best match the actual profile of the index of refraction. But it should be noted that $n'_A \neq n'(x)$, even in the limit; also, Eq. (4.7-7a) shows that $\gamma_A(x_j) \neq \gamma(x_j)$, except initially. Figure 4-6 shows an example of the differing profiles for γ_A and γ, in this case for an exponentially stratified medium. Their rate of divergence essentially depends only on the curvature in $n(x)$. For the exponential profile, it is independent of N_o; it depends only on H, the scale height.

For thin-atmosphere conditions, γ will be small and the difference $\gamma - \gamma_A$, which is zero at $\hat{y} = 0$, will be smaller. Accordingly, we let $\gamma_A = \gamma + w$, keeping only first-degree terms in *w*. A first-order differential equation for *w* truncated to first degree follows from Eq. (4.7-7a):

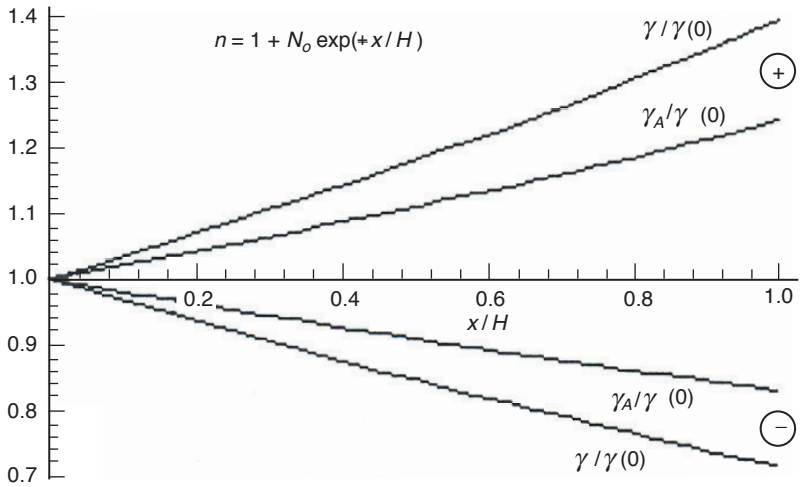

Fig. 4-6. Divergence between γ and γ_A from forcing $\hat{y}$ to be continuous across Airy layer boundaries.

$$\left.\begin{aligned} &\frac{dw}{dx} = -\frac{3w}{2}\frac{d}{dx}\log(|\, n^2 - n_o^2\,|) - \frac{d\gamma}{dx};\;\; w(x_o) = 0, \\ &\gamma_A = \gamma + w,\;\; \gamma_A^3 = 2k^{-1}n(x)n_A'(x),\;\; \gamma^3 = 2k^{-1}n(x)n'(x) \end{aligned}\right\} \tag{4.7-8}$$

This differential equation has as a solution

$$w = -\left(n^2 - n_o^2\right)^{-3/2}\int_{x_o}^{x}\left(n^2 - n_o^2\right)^{3/2}\frac{d\gamma}{dx}dx \tag{4.7-9}$$

Given the actual profile $n(x)$, a first-order correction term $w(x)$ can be obtained from Eq. (4.7-9), and, therefore, the profile $\gamma_A(x)$ is obtained for the medium being modeled by Airy layers. It follows from Eqs. (4.7-7) and (4.7-9) that $w(x)$ depends primarily on the curvature in $n(x)$. The profile $\hat{y}_A(x)$ follows from Eqs. (4.7-1) and (4.7-7b), that is,

$$\hat{y}_A(x) = -k\int_{x_o}^{x}\gamma_A(x')dx' = \hat{y}(x)\left(\gamma(x)/\gamma_A(x)\right)^2 \tag{4.7-10}$$

In an alternate strategy, one would set $n_A'(x_j) = n'(x_j)$ at each layer boundary with the initial conditions $n_A(x_o) = n(x_o)$. What strategy we should follow is dictated by the requirement that the results for the characteristic matrix from a given strategy match at least asymptotically the characteristic matrix from the first-order theory given in Section 3.5 for the case where n is piecewise constant profile. Without pinning down the strategy now, let us simply define $\gamma_A(x)$ as a piecewise constant function for a stack of Airy layers, and we will return to its form later.

With the profiles for $n_A(x_j)$ and $n_A'(x_j)$ that keep $\hat{y}$ continuous across the boundaries between Airy layers, one can develop a characteristic matrix for the stack. This is described in Appendix H.

4.8 Osculating Parameters

We return to a general Cartesian stratified medium (Fig. 4-1) for the TE case with $n = n(x)$. We have the coupled system

$$\left.\begin{aligned} &E_y = U(u),\;\; H_z = V(u),\;\; H_x = W(u), \\ &\frac{dU}{du} = i\mu V,\;\; \frac{dV}{du} = i\varpi^2 U,\;\;\; \mu W + n_o U = 0, \\ &\mu\varpi^2 = n^2 - n_o^2,\;\; n_o = n\sin\varphi = \text{constant} \end{aligned}\right\} \tag{4.8-1}$$

where $u = kx$. Here we set $\mu \equiv 1$ throughout the medium, but ϖ is variable.

Now we introduce an osculating parameter by solving for the reflection and transmission coefficients across a boundary that bears a discontinuity in n. To develop a functional form for these parameters, we first use the continuity conditions from Maxwell's equations that apply to the field components across a boundary. We obtain the changes in the parameters that result from a change in the index of refraction across a planar boundary, which is embedded in an otherwise homogeneous medium. After obtaining the transmission and reflection coefficients that apply across a boundary, we will use a limiting procedure to obtain a continuous version for these parameters. The continuity conditions are first-order condition, whereas the coupled equations in Eq. (4.8-1) are a second-order system. Therefore, these expressions provide approximate solutions. We will ascertain the limits of their validity.

For the case where the boundary carries neither charge density nor current density, and with $\mu \equiv 1$ throughout the medium, Maxwell's equations require for the TE case that across the boundary $E_y^{(i)} + E_y^{(r)} = E_y^{(t)}$, $H_z^{(i)} + H_z^{(r)} = H_z^{(t)}$, and also that $H_x^{(i)} + H_x^{(r)} = H_x^{(t)}$. Here, the superscript (i) denotes the incident wave, (r) the reflected wave, and (t) the transmitted wave. Reviewing Eq. (4.8-1), we see that these conditions are equivalent to requiring that $U(u)$ and $V(u)$ be continuous across the boundary at $u = u_*$. In Fig. 4-7, we assume that n_1 in the lower medium is constant, and similarly that n_2 is constant in the upper medium. But across the boundary, the index of refraction changes by $\Delta n = n_2 - n_1$. From Eq. (4.4-3), the solution for the field components in each homogeneous medium is given by

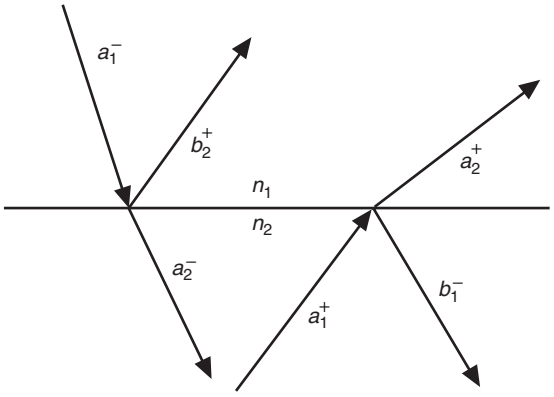

Fig. 4-7. Reflection and transmission coefficients across a boundary for upward and downward traveling waves.

$$\left.\begin{aligned} &\tilde{U}^{\pm} = e^{\pm i\varpi u},\ \tilde{V}^{\pm} = \pm\varpi e^{\pm i\varpi u},\ \tilde{W}^{\pm} = -n_o e^{\pm i\varpi u}, \\ &\varpi^2 = n^2 - n_o^2,\ n_o = n\sin\varphi = \text{constant}, \\ &n = n_1,\ u < u_*;\quad n = n_2,\ u > u_* \end{aligned}\right\} \tag{4.8-2}$$

In the solution $\tilde{U}^{\pm} = \exp(\pm i\varpi u)$, the plus sign is used for an upward-traveling wave, and the minus sign is used for a downward-traveling wave.

We write the solutions to Eq. (4.8-1) in terms of the basis functions $\tilde{U}^{\pm}$ and $\tilde{V}^{\pm}$ times an osculating parameter $a^{\pm}(u)$. Thus,

$$\left.\begin{aligned} &U^{\pm} = a^{\pm}\tilde{U}^{\pm} \\ &V^{\pm} = -i\frac{dU^{\pm}}{du} = a^{\pm}\left(\varpi + u\varpi' - i\frac{a'^{\pm}}{a^{\pm}}\right)\tilde{U}^{\pm} \\ &(*)' = \frac{d(*)}{du} \end{aligned}\right\} \tag{4.8-3}$$

We will show later that $(u\varpi' - a'/a)$ is small when evaluated away from a turning point and when thin-atmosphere conditions apply. For the case of a homogeneous medium, this term in Eq. (4.8-3) vanishes.

At a boundary, the incident radiation will split into a transmitted component and a reflected component. For an upward-traveling incident wave, the reflected component will be downward traveling and the transmitted component will be upward traveling. For each component, we have an osculating parameter, *a* for transmitted and *b* for reflected. The continuity conditions in this case of an upward traveling incident wave are

$$\left.\begin{aligned} &a_1^+\tilde{U}_1^+ + b_1^-\tilde{U}_1^- = a_2^+\tilde{U}_2^+ \\ &\varpi_1 a_1^+\tilde{U}_1^+ - \varpi_1 b_1^-\tilde{U}_1^- = \varpi_2 a_2^+\tilde{U}_2^+ \end{aligned}\right\} \tag{4.8-4}$$

Solving for a_2^+ and b_1^- in terms of a_1^+, one obtains

$$a_2^+ = \frac{2\varpi_1}{\varpi_1 + \varpi_2}\frac{1}{\tilde{U}_2^+\tilde{U}_1^-}a_1^+ \tag{4.8-5a}$$

$$b_1^- = \frac{\varpi_1 - \varpi_2}{\varpi_1 + \varpi_2}\frac{\tilde{U}_1^+}{\tilde{U}_1^-}a_1^+ \tag{4.8-5b}$$

which are essentially the Fresnel reflection and transmission formulas for the TE case.

For a series of layers, multiple internal reflections must be considered [9]. For example, a ray reflected downward from an upper -layer boundary will again be reflected upward at the boundary of interest. Figure 4-8 shows only rays that have been reflected twice to contribute to the upward-traveling main ray. Here N reflected rays, one from the boundary of each upward layer (the layers to the right of the jth layer in the figure), are then reflected again from the left-hand boundary of the jth layer. The reflection coefficients from these doubly reflected rays must be added to the transmission coefficient from main ray a_j^+ at the jth boundary to fully account for the total incident radiation at the left-hand boundary of the j + 1st layer. The second reflection also will occur from a lower boundary (to the left of the jth layer in Fig. 4-8), but reflections of this type will have already been folded into the value of a_j^+ at the left-hand boundary of the jth layer. However, Eq. (4.8-5b) shows that the reflection coefficients for doubly reflected rays will include a factor of the order of a_j^+ (here $\Delta\varpi$ is the average change in ϖ from layer to layer). Moreover, the phase of these secondary rays at the right-hand boundary of the jth layer will be randomly distributed when the span Δx of the ensemble of layers is such that $\Delta x >> \lambda$. It can be shown by vector summing up the contributions from all of these reflected rays with a second reflection from the left-hand boundary of the jth layer that the ratio of their combined contributions to the main ray contribution is given by $\lambda d\varpi / dx$, which is negligible for a thin atmosphere provided that turning points are avoided (where $\varpi = 0$ and $d\varpi / dx = nn' / \varpi \to \infty$). Therefore, in calculating the transmission coefficient for the wave, we can neglect secondary and higher-order reflections in our layer model when thin-atmosphere conditions apply and provided that we are sufficiently distant from a turning point.

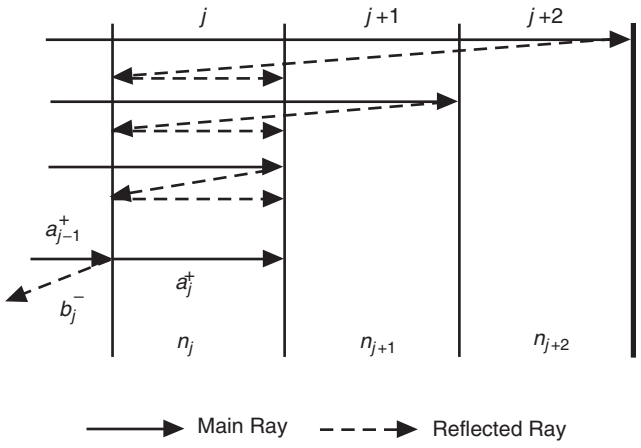

Fig. 4-8. Twice-reflected rays in the layered model.

The incident field at the $j+1$st boundary can be considered as the *product* of the transmission coefficients from the previous j layers. If we then expand that product and retain only the first-order terms, we will obtain a first-order equation for the transmission coefficient. The range of validity of this linear truncation is essentially the same as that found for the truncation of the characteristic matrix to linear terms given in Section 4.4; i.e., thin-atmosphere conditions and away from turning points.

Let us define a_j^+ to be the transmission coefficient of an upward-traveling wave for the jth layer. Then, from Eq. (4.8-5a) it follows that

$$a_{j+1}^+ = \frac{2\varpi_j}{\varpi_j + \varpi_{j+1}} \frac{1}{\tilde{U}_j^- \tilde{U}_{j+1}^+} a_j^+ \tag{4.8-6}$$

It follows for a series of layers that

$$\left.\begin{aligned} a_{j+1}^+ &\doteq a_1^+ \left(\prod_{k=1}^{k=j} \left[\frac{2\varpi_k}{2\varpi_k + \Delta\varpi_k} \right] \right) \exp\left[-i \sum_{k=1}^{k=j} \Delta\varpi_k u_k \right] \\ \Delta\varpi_k &= \varpi_{k+1} - \varpi_k \end{aligned}\right\} \tag{4.8-7}$$

Equation (4.8-2) for $\tilde{U}^{\pm}(\varpi u)$ has been used to obtain the exponential term in Eq. (4.8-7). The product in Eq. (4.8-7) can be evaluated as

$$\begin{aligned} \log\left[\prod_{k=1}^{k=j} \left(\frac{2\varpi_k}{2\varpi_k + \Delta\varpi_k} \right) \right] &= \sum_{k=1}^{k=j} \log\left(\frac{2\varpi_k}{2\varpi_k + \Delta\varpi_k} \right) \\ = -\sum_{k=1}^{k=j} \log\left(1 + \frac{\Delta\varpi_k}{2\varpi_k} \right) &\doteq -\frac{1}{2} \sum_{k=1}^{k=j} \frac{\Delta\varpi_k}{\varpi_k} \end{aligned} \tag{4.8-8}$$

which is valid provided that $\Delta\varpi_k / \varpi_k$ is small. Going to the limit, we obtain from Eq. (4.8-7)

$$a^+(u) \doteq a^+(u_0) \sqrt{\frac{\varpi_0}{\varpi}} \exp\left(-i \int_{\varpi_0}^{\varpi} u d\varpi \right) \tag{4.8-9}$$

From Eq. (4.8-4), it follows for an upward-traveling wave that

$$U^+ = a^+ \tilde{U}^+ \doteq a_0^+ \sqrt{\frac{\varpi_0}{\varpi}} \exp\left(-i \int_{\varpi_0}^{\varpi} u d\varpi - \varpi u \right) \tag{4.8-10}$$

Integrating by parts yields the WKB solution (for this special case of a Cartesian stratified medium):

$$U^{+}_{\mathrm{WKB}} = U^{+}_{0}\sqrt{\frac{\varpi_0}{\varpi}}\exp\left(i\int_{u_0}^{u}\varpi du\right) \tag{4.8-11}$$

Substituting the WKB solution in Eq. (4.8-11) into Eq. (4.8-1), one obtains

$$\left.\begin{aligned}\frac{d^2U^{+}_{\mathrm{WKB}}}{du^2} &= \left(\frac{3}{4}\left(\frac{\varpi'}{\varpi}\right)^2 - \frac{\varpi''}{2\varpi} - \varpi^2\right)U^{+}_{\mathrm{WKB}}\\ V^{+}_{\mathrm{WKB}} &= -i\frac{dU^{+}_{\mathrm{WKB}}}{du} = \left(\varpi + i\frac{\varpi'}{2\varpi}\right)U^{+}_{\mathrm{WKB}}\end{aligned}\right\} \tag{4.8-12}$$

Thus, these solutions essentially satisfy the wave equations when $|\varpi'/\varpi| << 1$ and $|\varpi''/\varpi| << 1$, which, except at turning points, holds for thin-atmosphere conditions. Note that $\varpi'/\varpi = (n'/n)\sec^2\varphi$, where φ is the angle of incidence of the wave with the x-axis. At a turning point, $\varphi = \pi/2$.

For a downward wave, we define a_j^- to be the transmission coefficient of a downward-traveling wave for the jth layer. Applying the continuity conditions from Maxwell's equations at the boundary, and following the same arguments that led to Eq. (4.8-5a), one obtains

$$a^{-}_{j-1} = \frac{2\varpi_j}{\varpi_j + \varpi_{j-1}}\frac{1}{\tilde{U}^{+}_{j}\tilde{U}^{-}_{j-1}}a^{-}_{j} \tag{4.8-13}$$

Following the same procedures in Eq. (4.8-6) to Eq. (4.8-9) one obtains

$$a^{-}(u) = a^{-}(u_0)\sqrt{\frac{\varpi_0}{\varpi}}\exp\left(i\int_{\varpi_0}^{\varpi} u d\varpi\right) \tag{4.8-14}$$

The downward wave is given by the WKB solution:

$$U^{-}_{\mathrm{WKB}} = a^{-}\tilde{U}^{-} = U^{-}_{0}\sqrt{\frac{\varpi_0}{\varpi}}\exp\left(-i\int_{u_0}^{u}\varpi du\right) \tag{4.8-15}$$

These are essentially the same forms given in the characteristic matrix in Section 3.5. From Eq. (4.8-9), we see that the phasor part of $a_j^{\pm}$ is just the phase accumulation resulting from a changing index of refraction between u_0 and u, i.e., $\varpi - \varpi_0$.

A further discussion of the higher-order correction terms from multiple reflections, including back-reflected waves, is given in [9].

4.8.1 At a Turning Point

At a turning point, these WKB solutions in Eqs. (4.8-11) and (4.8-15) fail. But we can approximate the turning point solution by use of the Airy solutions given in Section 4.5 when n' is a constant. Let us approximate the actual index of refraction in the vicinity of a turning point by $n = n_o + n_o'(u - u_o)$, where $n_o' = dn/du$ is evaluated at the turning point. It is assumed to remain constant in the vicinity of the turning points. The quadratic term in $n_o'(u-u_o)$ is negligible under our assumptions. Then the solution to the wave equations in the vicinity of the turning point is given by

$$\left.\begin{aligned} &U_A^{\pm} = K_1^{\pm}(\mathrm{Ai}[y] \mp i\,\mathrm{Bi}[y]), \quad V_A^{\pm} = i\gamma_o K_1^{\pm}(\mathrm{Ai}'[y] \mp i\,\mathrm{Bi}'[y]), \\ &\gamma_o = (2n_o n_o')^{1/3}, \quad y = -\left(\frac{\varpi}{\gamma_o}\right)^2, \quad \varpi^2 = n^2 - n_o^2 = \gamma_o^3(u-u_o) \end{aligned}\right\} \tag{4.8-16}$$

Here $u = kx$. At the turning point, $n = n_o$ and $y = 0$. Also, to ensure a vanishing field below the turning point, we must have $K_1^+ = K_1^- = K_1$, so that $U_A^+ + U_A^-$ involves only the Airy function of the first kind, $\mathrm{Ai}[y]$. At a point close to the turning point, but sufficiently above it so that $y <\sim -3$ (i.e., $u - u_o > 3/\gamma_o$, the equivalent of a few dekameters for the Earth's dry atmosphere at sea level), we may approximate the Airy functions with their negative argument asymptotic forms given in Eq. (3.8-7). The solutions in Eq. (4.8-16) become

$$U_A^{\pm} \doteq (-y)^{-1/4} K_1^{\pm} \exp\left[\pm i\left(\frac{2}{3}(-y)^{3/2} - \frac{\pi}{4}\right)\right] \tag{4.8-17}$$

But we recall from Section 4.5 that

$$\frac{2}{3}(-y)^{3/2} = \int_{u_o}^{u} \varpi\, du \tag{4.8-18}$$

Thus, Eq. (4.8-17) becomes

$$U_A^{\pm} \doteq K_1^{\pm} \sqrt{\frac{\mp i\gamma_o}{\varpi}} \exp\left(\pm i \int_{u_o}^{u} \varpi\, du\right) \tag{4.8-19}$$

The WKB solution at this point can be written in the form

$$U_{\mathrm{WKB}}^{\pm} = K_2^{\pm} \varpi^{-1/2} \exp\left(\pm i \int_{u_o}^{u} \varpi\, du\right) \tag{4.8-20}$$

Matching the coefficient in the Airy function asymptotic form with the coefficient in the WKB solution, we have

$$C_1^{\pm} \Leftrightarrow C_2^{\pm}\sqrt{\mp i\gamma_o} \tag{4.8-21}$$

The coefficients $C_2^{\pm}$ come from matching the WKB solutions with the boundary conditions for the problem, for example, those resulting from matching the solution to a plane incident or exiting wave at a particular altitude (see Fig. 4-2) above the turning point. With $C_1^{\pm}$ determined from Eq. (4.8-21), we then have an approximate solution in Eq. (4.8-16) valid at a turning point expressed in terms of Airy functions.

4.9 Airy Functions as Basis Functions

We can carry the use of osculating parameters a step further by defining the basis functions by

$$\left.\begin{aligned} &\tilde{U}_A^{\pm} = K_1^{\pm}\left(\mathrm{Ai}[y] \mp i\,\mathrm{Bi}[y]\right), \quad \tilde{V}_A^{\pm} = i\gamma_o K_1^{\pm}\left(\mathrm{Ai}'[y] \mp i\,\mathrm{Bi}'[y]\right), \\ &\gamma_o = \left(2n_o n_o'\right)^{1/3}, \quad y = -\left(\frac{\varpi}{\gamma_o}\right)^2, \quad \varpi^2 = n^2 - n_o^2 = \gamma_o^3\left(u - u_o\right) \end{aligned}\right\} \tag{4.9-1}$$

which are the exact solutions when $n' = dn/du$ is a constant in the medium and the quadratic term in n^2 can be ignored. When n' is variable, then γ is variable.

For the TE case, we recall that both $\tilde{U}$ and $\tilde{V}$ must be continuous across a boundary. At a boundary where n' changes discontinuously, the incident radiation will split into a transmitted component and a reflected component. For each component, we have an osculating parameter, a for transmitted and b for reflected. Therefore, the continuity conditions require

$$\left.\begin{aligned} a_1^{\pm}\tilde{U}_1^{\pm} + b_1^{\mp}\tilde{U}_1^{\mp} &= a_2^{\pm}\tilde{U}_2^{\pm} \\ i\gamma_1 a_1^{\pm}\tilde{U}_1'^{\pm} - i\gamma_1 b_1^{\mp}\tilde{U}_1'^{\mp} &= i\gamma_2 a_2^{\pm}\tilde{U}_2'^{\pm} \end{aligned}\right\} \tag{4.9-2}$$

Solving this system for $a_2^{\pm}$ and $b_1^{\mp}$ in terms of $a_1^{\pm}$, we obtain

$$\left.\begin{aligned} D^{\pm}a_2^{\pm} &= 2\gamma_1\pi^{-1}a_1^{\pm} \\ D^{\pm}b_1^{\mp} &= i\left(\gamma_1\tilde{U}_2^{\mp}\tilde{U}_1'^{\mp} - \gamma_2\tilde{U}_1^{\mp}\tilde{U}_2'^{\mp}\right)a_1^{\pm} \\ D^{\pm} &= \pm i\left(\gamma_1\tilde{U}_2^{\pm}\tilde{U}_1'^{\mp} + \gamma_2\tilde{U}_1^{\mp}\tilde{U}_2'^{\pm}\right) \end{aligned}\right\} \tag{4.9-3}$$

Now we let $\gamma_1 = \gamma - \Delta\gamma/2$, $\gamma_2 = \gamma + \Delta\gamma/2$, and $\Delta y = -2y\Delta\gamma/\gamma$. We expand $D^{\pm}$ in Eq. (4.9-3) to first order in $\Delta\gamma/\gamma$, using Eq. (4.9-1). We obtain

$$D^{\pm} = \pm\frac{2\gamma}{\pi}\left(1 + i\frac{\pi}{2}\frac{\Delta\gamma}{\gamma}\Gamma\right) \tag{4.9-4}$$

where $\Gamma[\hat{y}]$ is given by

$$\begin{aligned}\Gamma = &\left(\mathrm{Ai}[y]\mathrm{Ai}'[y] + \mathrm{Bi}[y]\mathrm{Bi}'[y]\right)\\ &+ 2y\left(\mathrm{Ai}'^{2}[y] + \mathrm{Bi}'^{2}[y] - y\left(\mathrm{Ai}^{2}[y] + \mathrm{Bi}^{2}[y]\right)\right)\end{aligned} \tag{4.9-5}$$

The incident field at the j + 1st boundary can be expressed as the product of the transmission coefficients from the previous j layers. We note that $\Delta\gamma/\gamma = (1/3)\Delta n'/n'$. Thus, the transmission coefficients can be written as

$$a_{j+1}^{\pm} = a_1^{\pm}\prod_{k=1}^{k=j}\left(2\gamma_k\left(\pi \mid D_k^{\pm} \mid\right)^{\mp 1}\right) \tag{4.9-6}$$

Going to the limit, we obtain

$$\left.\begin{aligned}a^{\pm} &= a_0^{\pm}\sqrt{\frac{\gamma_0}{\gamma}}\exp\left[\mp i\frac{\pi}{6}\left(\int_{u_0}^{u}\frac{d\log n'}{du}\Gamma[y]du\right)\right],\\ y &= -\left(\varpi/\gamma\right)^2,\ \gamma = (2nn')^{1/3},\ \varpi^2 = n^2 - n_o^2\end{aligned}\right\} \tag{4.9-7}$$

Therefore, this form for the osculating parameters should be good up to a turning point, provided $n' \neq 0$. It follows from Eq. (4.9-5) that $d\Gamma/dy = 0$ at $y = y^{\dagger} = 0.44331$. For $y > y^{\dagger}$, $\Gamma[y]$ grows large rapidly because it contains Airy functions of the second kind. This approach fails for increasing $y > y^{\dagger}$.

4.10 Wave Propagation in a Cylindrical Stratified Medium

We develop the characteristic matrix for cylindrical stratification. We again consider the TE case in two dimensions, that is, in the *xz*-plane in Fig. 4-9. The field is assumed to be invariant in the *y*-direction, along the axis of the cylinder. For the TE wave in this case, $E_r = E_\theta = 0$ and $H_y = 0$. When μ and ε are constant, Bessel functions of integer order l are solutions for the radial coordinate of the electromagnetic field. The mathematics is simpler for the cylindrical case than it is for the spherical case, but it is a simple extension to convert to spherical Bessel functions, once the cylindrical case is determined.

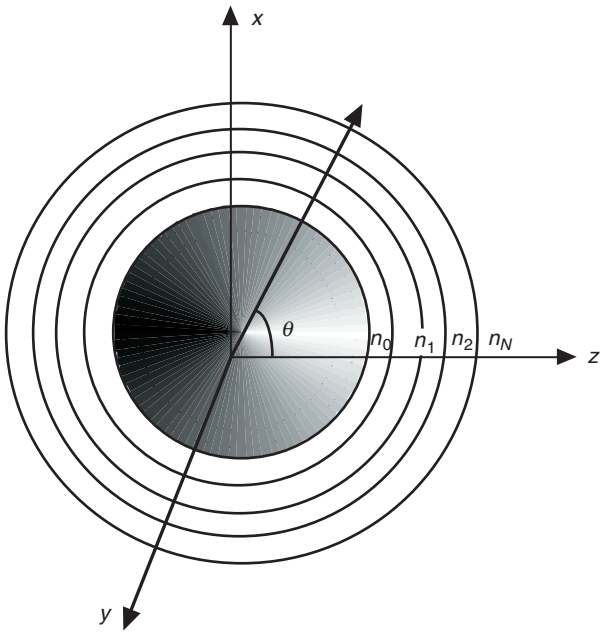

Fig. 4-9. Cylindrical stratified index of refraction.

Spherical stratification will be used to tie these results with those that follow from the spherical scattering coefficient approach, which is discussed later.

For a TE wave in the xz-plane ($\boldsymbol{E}$ along the y-axis), Maxwell's curl equations in Eq. (3.2-1a) expressed in cylindrical coordinates become

$$\left.\begin{aligned} \frac{1}{r}\frac{\partial E_y}{\partial \theta} &= ik\mu H_r \\ -\frac{\partial E_y}{\partial r} &= ik\mu H_\theta \\ 0 &= H_y \end{aligned}\right\} \tag{4.10-1a}$$

$$\left.\begin{aligned} 0 &= E_r \\ 0 &= E_\theta \\ \frac{1}{r}\left(\frac{\partial}{\partial r}(rH_\theta) - \frac{\partial H_r}{\partial \theta}\right) &= -ik\varepsilon E_y \end{aligned}\right\} \tag{4.10-1b}$$

We now use the separation-of-variables technique to solve these equations. We see from the second equation in Eq. (4.10-1a) that for the θ coordinate H_θ and

E_y must be represented by the same function times at most a constant. Accordingly, we set

$$\left.\begin{aligned} E_y &= U(r)\Phi(\theta) \\ H_\theta &= V(r)\Phi(\theta) \\ H_r &= W(r)\Theta(\theta) \end{aligned}\right\} \tag{4.10-2}$$

Inserting these forms into Eq. (4.10-1), it follows that conditions for these representations to be valid for all values of r and θ are that $\Phi' = a\Theta$ and $\Theta' = b\Phi$, where a and b are constants. Also, to ensure single-valuedness in these θ functions, the product ab must equal $-l^2$, where l must be integer-valued. Thus, both Φ and Θ are of the form $\exp(il\theta)$. It follows that $a = b = \pm il$. Separating the variables in Eq. (4.10-1) requires that the following system of equations must hold:

$$\mu \frac{d}{dr}\left(\frac{1}{\mu} r \frac{dU}{dr}\right) + \left(k^2 n^2 - \frac{l^2}{r^2}\right) rU = 0 \tag{4.10-3a}$$

$$\left.\begin{aligned} \frac{1}{knr}\frac{d(rV)}{dr} &= -i\sqrt{\frac{\varepsilon}{\mu}}\left(1 - \frac{l^2}{(knr)^2}\right)U \\ \frac{1}{kn}\frac{dU}{dr} &= -i\sqrt{\frac{\mu}{\varepsilon}}V \end{aligned}\right\} \tag{4.10-3b}$$

$$nkrW = l\sqrt{\frac{\varepsilon}{\mu}}U \tag{4.10-3c}$$

$$\frac{d^2\Theta}{d\theta^2} = \frac{d^2\Phi}{d\theta^2} = -il^2\Theta \tag{4.10-3d}$$

Thus, the eigenfunction solutions to Eq. (4.10-3d), irrespective of the functional form of $n(r)$, are given by

$$\Theta_l = \Phi_l = \exp(il\theta), \quad l = 0, \pm 1, \pm 2, \cdots \tag{4.10-3d′}$$

The complete solution will be a weighted summation of the individual spectral components, $U_l\Phi_l$, $V_l\Phi_l$, and $W_l\Theta_l$, over all integer values of l, which for each value of l are solutions to Eq. (4.10-3). For the special case of $l = 0$, we see from Eqs. (4.10-1) and (4.10-2) that $\Phi_0 \equiv 1$ and that $W_0 \equiv 0$. U_0 and V_0 are given by Eqs. (4.10-3a) and (4.10-3b) with $l = 0$.

Equation (4.10-3a) provides the modified wave equation for $U_l(r)$, which describes the lth spectral component of the electric field for a TE wave. It is the radial part of the modified vector wave equation given in Eq. (3.2-1b). In a medium with $\mu \equiv 1$, this wave equation in Eq. (4.10-3a) simplifies further to

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dU_l}{dr}\right)+\left(k^2n^2-\frac{l^2}{r^2}\right)U_l=0 \qquad (4.10\text{-}3a')$$

which is the modified Bessel equation discussed in Section 5.2.

When μ and ε are constant, one obtains from Eq. (4.10-3) as spectral solutions

$$\left.\begin{aligned} &U_l(\rho)=\{J_l(\rho),Y_l(\rho)\},\ V_l(\rho)=i\sqrt{\frac{\varepsilon}{\mu}}\{J_l'(\rho),Y_l'(\rho)\},\quad \rho=nkr \\ &\rho W_l(\rho)=l\sqrt{\varepsilon/\mu}\,U_l(\rho),\ \Theta_l(\theta)=\exp(il\theta),\quad l=0,\pm1,\pm2,\cdots \end{aligned}\right\} \qquad (4.10\text{-}4)$$

where $J_l(\rho)$ and $Y_l(\rho)$ are the Bessel functions of the first and second kind, discussed earlier.

If we write in characteristic matrix form the solutions to the coupled system in Eq. (4.10-3b), we obtain

$$\begin{pmatrix} U_l(\rho) \\ \rho V_l(\rho) \end{pmatrix}=\boldsymbol{M}_l[\rho,\rho_0]\begin{pmatrix} U_l(\rho_0) \\ \rho_0 V_l(\rho_0) \end{pmatrix}=\begin{bmatrix} F_l(\rho,\rho_0) & f_l(\rho,\rho_0) \\ G_l(\rho,\rho_0) & g_l(\rho,\rho_0) \end{bmatrix}\begin{pmatrix} U_l(\rho_0) \\ \rho_0 V_l(\rho_0) \end{pmatrix} \qquad (4.10\text{-}5)$$

One can show that the elements of the characteristic matrix for the TE case, which are solutions to the coupled system of equations in Eq. (4.10-3b), are given by

$$\left.\begin{aligned} &F_l(\rho,\rho_0)=\frac{\pi\rho_0}{2}\begin{vmatrix} J_l(\rho) & Y_l(\rho) \\ J_l'(\rho_0) & Y_l'(\rho_0) \end{vmatrix},\ f_l(\rho,\rho_0)=i\frac{\pi}{2}\sqrt{\frac{\mu}{\varepsilon}}\begin{vmatrix} J_l(\rho) & Y_l(\rho) \\ J_l(\rho_0) & Y_l(\rho_0) \end{vmatrix}, \\ &G_l(\rho,\rho_0)=i\frac{\pi\rho\rho_0}{2}\sqrt{\frac{\varepsilon}{\mu}}\begin{vmatrix} J_l'(\rho) & Y_l'(\rho) \\ J_l'(\rho_0) & Y_l'(\rho_0) \end{vmatrix},\ g_l(\rho,\rho_0)=-\frac{\pi\rho}{2}\begin{vmatrix} J_l'(\rho) & Y_l'(\rho) \\ J_l(\rho_0) & Y_l(\rho_0) \end{vmatrix} \end{aligned}\right\} \qquad (4.10\text{-}6)$$

Here $J_l(\rho)Y_l'(\rho)-J_l'(\rho)Y_l(\rho)=(\pi\rho/2)^{-1}$, which is the Wronskian of the Bessel functions, has been used to set the scale factors so that $F_l(\rho,\rho_0)=1$ and $g_l(\rho,\rho_0)=1$. It is easily shown that when the elements of $\boldsymbol{M}_l[\rho,\rho_0]$ are given by Eq. (4.10-6), then $\mathrm{Det}[\boldsymbol{M}_l[\rho,\rho_0]]=1$ for all values of ρ.

For the TM case, starting from Maxwell's curl equations in Eqs. 4.2-1(a) and 4.2-1(b with $H_r = H_\theta = 0$ and $E_y = 0$, we use the separation of variables technique and set

$$\left.\begin{aligned} H_y &= U(r)\Phi(\theta) \\ E_\theta &= V(r)\Phi(\theta) \\ E_r &= W(r)\Theta(\theta) \end{aligned}\right\} \tag{4.10-7}$$

One can readily show for a TM wave that the equation set analogous to Eq. (4.10-3) is given by

$$\frac{\varepsilon}{r}\frac{d}{dr}\left(\frac{r}{\varepsilon}\frac{dU_l}{dr}\right)+\left(k^2n^2-\frac{l^2}{r^2}\right)U_l = 0 \tag{4.10-8a}$$

$$\left.\begin{aligned} \frac{1}{knr}\frac{d(rV)}{dr} &= -i\sqrt{\frac{\varepsilon}{\mu}}\left(1-\frac{l^2}{(knr)^2}\right)U \\ \frac{1}{kn}\frac{dU}{dr} &= -i\sqrt{\frac{\mu}{\varepsilon}}V \end{aligned}\right\} \tag{4.10-8b}$$

$$W_l = \frac{-l}{nkr}\sqrt{\frac{\mu}{\varepsilon}}U_l \tag{4.10-8c}$$

$$\Theta_l = \Phi_l = \exp(il\theta),\ l = 0.\pm1,\pm2,\cdots \tag{4.10-8d}$$

Equation (4.10-8a) shows for the TM case that a medium with $\mu \equiv 1$ does not offer further simplification of the modified wave equation for $U_l(r)$; not only is n variable, but we have the extra term involving $\nabla\left[\boldsymbol{E}\cdot(\nabla(\log\ \varepsilon)\right]$. This term can be avoided through the renormalization of n; see Appendix I.

For the TM case with constant n, the elements of the characteristic matrix have the same forms given in Eq. (4.10-6) except that the reciprocals of $\sqrt{\mu/\varepsilon}$ and $\sqrt{\varepsilon/\mu}$ are used in f and G_l.

The complete solution for either a TE wave or a TM wave is obtained by summing these spectral components weighted by the appropriate spectral coefficients on the integer spectral number l. The spectral coefficients might be obtained from asymptotic boundary conditions on $U(\rho_0,\rho_0)$ and $V(\rho_0,\rho_0)$, such as those for an approaching plane wave located at a large distance from the cylinder. Bauer's identity for a plane wave in cylindrical coordinates is given by

$$\exp(i\rho\cos\theta)=\sum_{l=-\infty}^{l=\infty} i^l J_l(\rho)\exp(il\theta) \tag{4.10-9}$$

For a TE wave, Eq. (4.10-9) provides the asymptotic form for E_y/E_o. Differentiating with respect to ρ and θ yields

$$\left.\begin{aligned}\exp(i\rho\cos\theta)\cos\theta&=\sum_{l=-\infty}^{l=\infty} i^{l-1}J_l'(\rho)\exp(il\theta)\\ \exp(i\rho\cos\theta)\sin\theta&=-\sum_{l=-\infty}^{l=\infty} i^l\frac{l}{\rho}J_l(\rho)\exp(il\theta)\end{aligned}\right\} \tag{4.10-10}$$

For a TE plane wave traveling in the positive z-direction in Fig. 4-9, these are the forms for $-H_\theta/H_o$ and $-H_r/H_o$, respectively. These forms also are obtained from Maxwell's equation $ik\mu\boldsymbol{H}=\nabla\times\boldsymbol{E}$ when $\boldsymbol{E}=\hat{\boldsymbol{y}}E_o\exp(i\rho\cos\theta)$. According to Eqs. (4.10-2) and (4.10-4), we set

$$\left.\begin{aligned}E_y&=\sum_{l=-\infty}^{l=\infty}(a_lJ_l+b_lY_l)\Theta_l=E_o\exp(i\rho\cos\theta)\\ H_\theta&=i\sqrt{\frac{\varepsilon}{\mu}}\sum_{l=-\infty}^{l=\infty}(c_lJ_l'+d_lY_l')\Theta_l=-H_o\exp(i\rho\cos\theta)\cos\theta\\ H_r&=\sqrt{\frac{\varepsilon}{\mu}}\sum_{l=-\infty}^{l=\infty}\frac{l}{\rho}(a_lJ_l+b_lY_l)\Theta_l=-H_o\exp(i\rho\cos\theta)\sin\theta\end{aligned}\right\} \tag{4.10-11}$$

Equating these forms to their corresponding forms in Eqs. (4.10-9) and (4.10-10), we obtain for the coefficients

$$\left.\begin{aligned}&a_l=E_oi^l,\ \ b_l\equiv 0\\ &c_l=i^lH_o\sqrt{\mu/\varepsilon}=E_oi^l\\ &d_l\equiv 0\end{aligned}\right\} \tag{4.10-12}$$

Similarly, suppose that the boundary conditions correspond to a plane wave along the line $\theta=-\alpha$. Then the time-independent form for the wave is given by $E_y=E_o\exp[i\rho\cos(\theta+\alpha)]$. It follows that Eq. (4.10-9) is modified to

$$\exp(i\rho\cos(\theta+\alpha))=\sum_{l=-\infty}^{l=\infty} i^l J_l(\rho)\exp(il(\theta+\alpha)) \tag{4.10-13}$$

Then by matching the coefficients in Eq. (4.10-11) using Eqs. (4.10-9) and (4.10-10), it follows that the spectral coefficients for this case are given by

$$\left.\begin{aligned} a_l &= E_o i^l \exp(il\alpha) \\ b_l &\equiv 0 \\ c_l &= H_o i^l \sqrt{\mu/\varepsilon}\exp(il\alpha) = E_o i^l \exp(il\alpha) \\ d_l &\equiv 0 \end{aligned}\right\} \tag{4.10-14}$$

For a TM plane wave traveling in the positive z-direction in Fig. 4-9, H_y / H_o is given by Eq. (4.10-9), and E_θ / E_o and E_r / E_o are given by the forms in Eq. (4.10-10).

Using the property of the Bessel functions given by $J_{-l}(\rho) = (-1)^l J_l(\rho)$, Eq. (4.10-9) may be written as

$$\exp(i\rho\cos\theta) = J_0(\rho) + 2\sum_{l=1}^{l=\infty} i^l J_l(\rho)\cos(l\theta) \tag{4.10-9′}$$

Using the stationary phase technique, we will show in a later section that for locations well into the upper half-plane, $0 < \theta < \pi$, only spectral components involving the $\exp(-il\theta)$ part of $\cos(l\theta)$ contribute to the summations in Eq. (4.10-11) when $r/\lambda >> 1$. Moreover, these contributions come from neighborhoods around $l \approx \rho$.

For large l and ρ, we may replace the integer-order Bessel functions with their Airy function asymptotic forms (see Eqs. (3.8-1) and (3.8-2). For the TE case, the asymptotic forms for the elements of the characteristic matrix are

$$
\left.
\begin{aligned}
F_l(\hat{y},\hat{y}_0) &\doteq \pi \frac{K_{\rho_0}}{K_\rho} \begin{vmatrix} \mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}] \\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0] \end{vmatrix} \\
f_l(\hat{y},\hat{y}_0) &\doteq \frac{-i\pi}{2K_\rho K_{\rho_0}} \sqrt{\frac{\mu}{\varepsilon}} \begin{vmatrix} \mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}] \\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0] \end{vmatrix} \\
G_l(\hat{y},\hat{y}_0) &\doteq -i2\pi K_\rho K_{\rho_0} \sqrt{\frac{\varepsilon}{\mu}} \begin{vmatrix} \mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}] \\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0] \end{vmatrix} \\
g_l(\hat{y},\hat{y}_0) &\doteq \frac{\pi K_\rho}{K_{\rho_0}} \begin{vmatrix} \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0] \\ \mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}] \end{vmatrix} \\
\hat{y} &= K_\rho^{-1}(l-\rho)\left(1 - \frac{l-\rho}{60K_\rho^3} + \cdots\right), \quad K_\rho = \left(\frac{\rho}{2}\right)^{1/3}
\end{aligned}
\right\} \qquad (4.10\text{-}15)
$$

4.11 Wave Propagation in a Spherical Stratified Medium

Here we develop the characteristic matrix for spherical stratification, which is pursued in later sections using a modified Mie scattering theory. Spherical harmonic functions in the angular coordinates will result as solutions for the electromagnetic field when $n = n(r)$, and spherical Bessel functions also will result when the index of refraction is constant within a layer.

We consider an approaching wave that is planar with its electric vector parallel to the *x*-axis. Figure 4-10 shows the geometry. The central angle between the radius vector **r** and the z-axis is given by θ. The Poynting $\boldsymbol{S}$ for the approaching plane wave is directed along the *z*-axis. The azimuth angle ϕ lies in the *yx*-plane. The plane of incidence is given by $\phi = 0$.

Maxwell's curl equations for the TM wave become

$$
\left.
\begin{aligned}
\frac{1}{r\sin\theta}\left(\frac{\partial E_r}{\partial\phi} - \sin\theta\frac{\partial(rE_\phi)}{\partial r}\right) &= ik\mu H_\theta \\
H_r = \left(\frac{\partial(E_\phi\sin\theta)}{\partial\theta} - \frac{\partial E_\theta}{\partial\phi}\right) &\equiv 0 \\
\frac{1}{r}\left(\frac{\partial(rE_\theta)}{\partial r} - \frac{\partial E_r}{\partial\theta}\right) &= ik\mu H_\phi
\end{aligned}
\right\} \qquad (4.11\text{-}1a)
$$

and

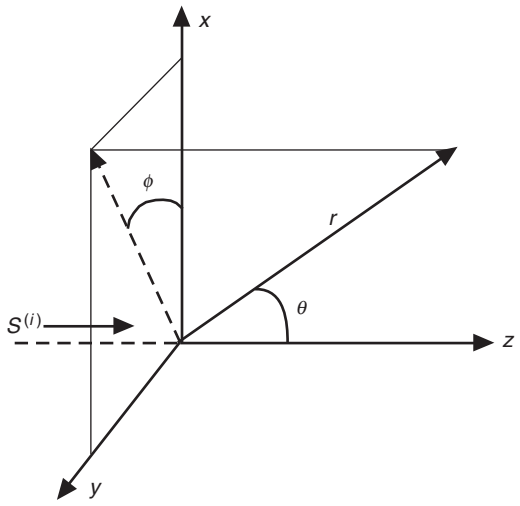

Fig. 4-10. Coordinate frame for a spherical stratified medium.

$$\left.\begin{aligned}
&\frac{1}{r\sin\theta}\left(\frac{\partial\left(H_\phi\sin\theta\right)}{\partial\theta}-\frac{\partial H_\theta}{\partial\phi}\right)=-ik\varepsilon E_r\\
&\frac{1}{r}\left(-\frac{\partial\left(rH_\phi\right)}{\partial r}\right)=-ik\varepsilon E_\theta\\
&\frac{1}{r}\frac{\partial\left(rH_\theta\right)}{\partial r}=-ik\varepsilon E_\phi
\end{aligned}\right\} \qquad (4.11\text{-}1b)$$

Since we are only concerned with in-plane propagation, it follows for the TM wave that $H_\theta = H_r = E_\phi = 0$ in the plane $\phi = 0$. In this case, Eq. (4.11-1) simplifies to

$$\frac{1}{r\sin\theta}\frac{\partial\left(H_\phi\sin\theta\right)}{\partial\theta}=-ik\varepsilon E_r \qquad (4.11\text{-}2a)$$

$$\frac{1}{r}\frac{\partial\left(rH_\phi\right)}{\partial r}=ik\varepsilon E_\theta \qquad (4.11\text{-}2b)$$

$$\frac{1}{r}\left(\frac{\partial\left(rE_\theta\right)}{\partial r}-\frac{\partial E_r}{\partial\theta}\right)=ik\mu H_\phi \qquad (4.11\text{-}2c)$$

Using the separation of variables technique again, we see from Eq. (4.11-2b) that in the θ coordinate H_ϕ and E_θ must be represented by the same function.

But Eq. (4.11-2c) shows that this function must be equal to the derivative of the θ function used for E_r times a constant. Accordingly, we set

$$\left.\begin{aligned} H_\phi &= U(r)\Theta'(\theta) \\ E_\theta &= V(r)\Theta'(\theta) \\ E_r &= W(r)\Theta(\theta) \end{aligned}\right\} \tag{4.11-3}$$

Equations (4.11-2) and (4.11-3) require that the following coupled system of first-order differential equations be satisfied:

$$\frac{U}{r}\frac{1}{\sin\theta}\left(\frac{d(\Theta'\sin\theta)}{d\theta}\right) = -ik\varepsilon W\Theta \tag{4.11-4a}$$

$$\frac{d(rU)}{dr}\Theta' = ik\varepsilon rV\Theta' \tag{4.11-4b}$$

$$\frac{1}{r}\left(\frac{d(rV)}{dr}\Theta' - W\Theta'\right) = ik\mu U\Theta' \tag{4.11-4c}$$

We see that Eqs. (4.11-4b) and (4.11-4c) have a common factor, $\Theta'(\theta)$, which can be factored, leaving only terms involving the radial coordinate. But, for Eq. (4.11-4a) to hold for all possible values of r and θ, it is required that

$$\left.\begin{aligned} \frac{1}{\Theta\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) &= \text{ constant} = -l(l+1) \\ ik\varepsilon rW(r) &= -l(l+1)U(r) \end{aligned}\right\} \tag{4.11-5}$$

The top line, of course, is the differential equation for a Legendre polynomial $P_l(\cos\theta)$. The constant in Eq. (4.11-5) has the value of $-l(l+1)$, with l restricted to integer values to ensure physical solutions to the differential equation and also single-valuedness for $\Theta(\theta)$. For negative integer values for l, we set $l = -\tilde{l} - 1$ with $\tilde{l} \geq 0$. It follows that $l(l+1) = (\tilde{l}+1)\tilde{l}$ and, therefore, $P_l = P_{-l-1}$. It follows that we need only the non-negative integers for l. The second solution to the Legendre differential equation $Q_l(\cos\theta)$ has a singularity at $\theta = 0$. Because it does not match the boundary conditions, we discard it.

From Eqs. (4.11-4) and (4.11-5), we obtain for the TM case

$$\varepsilon\frac{d}{dr}\left(\frac{1}{\varepsilon}\frac{d(rU_l)}{dr}\right) + \left(k^2n^2 - \frac{l(l+1)}{r^2}\right)(rU_l) = 0, \quad l = 0,1,2,\cdots \tag{4.11-6a}$$

$$\left.\begin{aligned} \frac{1}{kn}\frac{d(rV_l)}{dr} &= i\sqrt{\frac{\mu}{\varepsilon}}\left(1-\frac{l(l+1)}{(knr)^2}\right)(rU_l) \\ \frac{1}{kn}\frac{d(rU_l)}{dr} &= i\sqrt{\frac{\varepsilon}{\mu}}(rV_l) \end{aligned}\right\} \tag{4.11-6b}$$

$$nkrW_l = -il(l+1)\sqrt{\frac{\mu}{\varepsilon}}U_l \tag{4.11-6c}$$

$$\Theta_l = P_l(\cos\theta) \tag{4.11-6d}$$

The solution set for the TE wave follows the same course, starting with the TE version of Maxwell's curl equations that have been given in Eq. (4.11-1) for the TM case, and ending with the TE version of Eq. (4.11-6). For the TE wave, ε will be replaced by μ in the modified wave equation in Eq. (4.11-6a); the factor $-\sqrt{\varepsilon/\mu}$ replaces $\sqrt{\mu/\varepsilon}$ in the top line of Eq. (4.11-6b), and $-\sqrt{\mu/\varepsilon}$ replaces $\sqrt{\varepsilon/\mu}$ in the bottom line of Eq. (4.11-6b) and in Eq. (4.11-6c).

The superposition of the fields for the TM wave and the TE wave yields the complete field. Hence, a plane wave, invariant along the y-direction in Fig. 4-10, and with its electric field vector pointed along the x-axis and its magnetic vector along the y-axis, can be expressed in terms of the superposition of its TE and TM components. However, when $r/\lambda >> 1$ and $l >> 1$, the TE terms can be ignored for this case when the electric field vector is directed along the x-axis. This follows from inspecting the asymptotic forms for the Legendre polynomials for large l. For a TM wave, the series representation for the electric field involves $dP_l/d\theta$ terms. For large l, one can readily show from the asymptotic form for the Legendre polynomial that $dP_l/d\theta \sim lP_l$. On the other hand, for the TE wave, the series representation for the electric field involves only P_l terms, which are negligible compared to $dP_l/d\theta$. For the cylindrical representation, the TE terms are completely absent for a plane wave with its electric field vector along the *x*-axis.

When ε and μ are constant, the solutions to Eq. (4.11-6a) are spherical Bessel functions of the first and second kind and of integer-order l. Thus, the general solution to the system in Eq. (4.11-6) for the lth spectral component of the TM wave is given by

$$\left.\begin{aligned} &\rho U_l = \{\psi_l(\rho), \chi_l(\rho)\},\ \ \rho V_l = -i\sqrt{\frac{\mu}{\varepsilon}}\{\psi_l'(\rho), \chi_l'(\rho)\},\ \ \rho = nkr, \\ &\rho W_l = -il(l+1)\sqrt{\frac{\mu}{\varepsilon}}U_l,\ \ \Theta_l = P_l(\cos\theta),\ \ l = 0,1,2\cdots \end{aligned}\right\} \tag{4.11-7}$$

Here $\psi_l(\rho)$ is the spherical Bessel function of the first kind and of order l, and $\chi_l(x)$ is the spherical Bessel function of the second kind. Equation (3.2-5) defines $\psi_l(x)$ and $\chi_l(x)$.

The characteristic matrix for the lth spectral component is defined by

$$\rho\begin{pmatrix} U_l(\rho) \\ V_l(\rho) \end{pmatrix} = \boldsymbol{M}_l[\rho, \rho_0]\begin{pmatrix} U_l(\rho_0) \\ V_l(\rho_0) \end{pmatrix}\rho_0 \tag{4.11-8}$$

It is given for the in-plane TM case by

$$\begin{aligned} \boldsymbol{M}_l[\rho, \rho_0] &= \begin{bmatrix} F_l(\rho, \rho_0) & f_l(\rho, \rho_0) \\ G_l(\rho, \rho_0) & g_l(\rho, \rho_0) \end{bmatrix} \\ &= \begin{bmatrix} \begin{vmatrix} \psi_l(\rho) & \chi_l(\rho) \\ \psi_l'(\rho_0) & \chi_l'(\rho_0) \end{vmatrix} & -i\sqrt{\frac{\varepsilon}{\mu}}\begin{vmatrix} \psi_l(\rho) & \chi_l(\rho) \\ \psi_l(\rho_0) & \chi_l(\rho_0) \end{vmatrix} \\ -i\sqrt{\frac{\mu}{\varepsilon}}\begin{vmatrix} \psi_l'(\rho) & \chi_l'(\rho) \\ \psi_l'(\rho_0) & \chi_l'(\rho_0) \end{vmatrix} & -\begin{vmatrix} \psi_l'(\rho) & \chi_l'(\rho) \\ \psi_l(\rho_0) & \chi_l(\rho_0) \end{vmatrix} \end{bmatrix} \end{aligned} \tag{4.11-9}$$

Here the Wronskian $\psi_l\chi_l' - \psi_l'\chi_l = 1$ has been used to ensure that the determinant of $\boldsymbol{M}_l[\rho, \rho_0]$ is unity for all values of ρ. For the TE wave in the plane $\phi = 0$, i.e., the case where the electric field vector for the incident plane wave is perpendicular to the *xz*-plane in Fig. 4-7 ($H_y \equiv 0$), the characteristic matrix is the same as that given in Eq. (4.11-9) except that the reciprocals of $\sqrt{\varepsilon/\mu}$ and $\sqrt{\mu/\varepsilon}$ are used in the off-diagonal elements.

For large l and ρ, we may replace the spherical Bessel functions with their Airy function asymptotic forms [see Eqs. (3.8-1) and (3.8-2)]. The asymptotic forms for the elements of the characteristic matrix for the case of the electric field vector of the incident planar wave directed along the *x*-axis in Fig. 4-10 are

$$\left.\begin{aligned} F_l(\hat{y},\hat{y}_0) &\doteq \pi\sqrt{\frac{K_\rho}{K_{\rho_o}}}\begin{vmatrix}\mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}]\\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0]\end{vmatrix}\\ f_l(\hat{y},\hat{y}_0) &\doteq i\pi\sqrt{K_\rho K_{\rho_o}}\sqrt{\frac{\varepsilon}{\mu}}\begin{vmatrix}\mathrm{Ai}[\hat{y}] & \mathrm{Bi}[\hat{y}]\\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0]\end{vmatrix}\\ G_l(\hat{y},\hat{y}_0) &\doteq i\frac{\pi}{\sqrt{K_\rho K_{\rho_o}}}\sqrt{\frac{\mu}{\varepsilon}}\begin{vmatrix}\mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}]\\ \mathrm{Ai}'[\hat{y}_0] & \mathrm{Bi}'[\hat{y}_0]\end{vmatrix}\\ g_l(\hat{y},\hat{y}_0) &\doteq -\pi\sqrt{\frac{K_{\rho_o}}{K_\rho}}\begin{vmatrix}\mathrm{Ai}'[\hat{y}] & \mathrm{Bi}'[\hat{y}]\\ \mathrm{Ai}[\hat{y}_0] & \mathrm{Bi}[\hat{y}_0]\end{vmatrix}\end{aligned}\right\} \tag{4.11-10}$$

where

$$\left.\begin{aligned} \hat{y} &= \nu^{2/3}\zeta[\nu/\rho] \quad \text{(see Section 3.8)}\\ &= K_\rho^{-1}(\nu-\rho)(1-K_\rho^{-3}(\nu-\rho)/60+\cdots\\ K_\rho &= (\rho/2)^{1/3};\quad \nu = l+1/2\end{aligned}\right\} \tag{4.11-11}$$

Given specified beginning and end points (r, r_0), one would perform a weighted integration of each of the matrix elements in Eq. (4.11-10) with respect to $\hat{y}$ to obtain the complete asymptotic form for the characteristic matrix. The weights might be the spectral coefficients of the wave at r_0. For example, for an approaching plane wave at (r_0, θ_0) along the *z*-axis, one has

$$\left.\begin{aligned} \rho_0 U_l(\rho_0) &= H_0\, i^l\,\frac{2l+1}{l(l+1)}\psi_l(\rho_0)\\ \rho_0 V_l(\rho_0) &= E_0\, i^{l-1}\,\frac{2l+1}{l(l+1)}\psi_l'(\rho_0)\\ \rho_0^2 W_l(\rho_0) &= E_0\, i^{l-1}(2l+1)\psi_l(\rho_0)\end{aligned}\right\} \tag{4.11-12}$$

Substituting these forms into Eq. (4.11-8) for $U_l(\rho_0)$ and $V_l(\rho_0)$ leads, of course, to the identical forms for $U_l(\rho)$ and $V_l(\rho)$ with ρ_0 replaced by ρ. Since we started with a planar form for the approaching wave, it must remain planar when *n* is held constant within the medium.

The characteristic matrix in Eq. (4.11-10) for the spherical case is essentially the same as the TM version for the cylindrical case given in Eq. (4.11-8), except for the scale factors K_ρ.

4.12 Correspondence between Characteristic Matrices for Cartesian and Spherical Stratified Airy Layers

The elements of the characteristic matrix in Eq. (4.11-10) are the asymptotic forms for large values of l and ρ; they provide a two-dimensional description in polar coordinates of the propagation of an in-plane polarized wave ($E_y \equiv 0$) through a medium of constant index of refraction. Similarly, the elements given in Eq. (4.5-14) describe a TM wave ($E_y \equiv 0$) propagating through a Cartesian stratified Airy layer. We can establish a duality between these two systems. If we assume in the spherical framework that the difference between the beginning and end points of the ray is small compared to the radius of curvature r_0 of the stratified surface, that is, $(r - r_0)/r_0 << 1$, then the same mathematics applies to both systems; a correspondence between parameters is given by

$$\left.\begin{array}{lll} \gamma = (2k^{-1}nn')^{1/3} & \leftrightarrow & nK_\rho^{-1} = n(2/\rho)^{1/3} \\ \text{Cartesian} \uparrow \text{Stratified} & & \text{Spherical} \uparrow \text{Stratified} \\ \text{Airy Layer} & & n = \text{constant within layer; } \rho,\ l >> 1 \end{array}\right\} \quad (4.12\text{-}1)$$

The transformation between the arguments of the Airy functions in the two systems, the $\hat{y}$ values, is given by

$$\left.\begin{array}{lll} -k\int_{x_o}^{x} \gamma dx = \hat{y} & \leftrightarrow & \hat{y} \doteq \left\{\begin{array}{l} K_\rho^{-1}(\nu - \rho) \\ \text{or} \\ K_\nu^{-4}\left(\nu^2 - \rho^2\right)/4 \end{array}\right. \\ \text{Cartesian} \uparrow \text{Airy} & & \text{Large} \uparrow \text{Spherical} \\ \text{Layer (CAL)} & & \text{Layer (LSL)} \end{array}\right\} \quad (4.12\text{-}2a)$$

where γ is given by Eqs. (4.5-3) and (4.5-4). When $\nu \approx \rho >> 1$, $\hat{y}$ in Eq. (4.12-2a) may be expressed as

$$\left.\begin{array}{lll} \hat{y} = -\gamma^{-2}\left(n^2 - n_o^2\right) & \leftrightarrow & \hat{y} \doteq -K_\nu^{-4}\left(\rho^2 - \nu^2\right)/4 \\ \hat{y} = -\gamma^{-2}n^2\cos^2\varphi & \leftrightarrow & \hat{y} \doteq -K_\nu^2\cos^2\theta_\nu \\ \text{CAL} \uparrow & & \uparrow \text{LSL} \end{array}\right\} \quad (4.12\text{-}2b)$$

Here we have set the spectral number $\nu = \rho\sin\theta_\nu$. Thus, θ_ν, which is the analog of φ in the Cartesian framework, may be interpreted as the "angle of

incidence" of the νth spectral component in the spherical framework. Chapters 3 and 5 discuss θ_ν (see Fig. 3-14) in a wave theory context. These chapters also address the role of θ_ν in the connections between wave theory and geometric optics, and in stationary phase theory.

In Eq. (4.12-1), we can write γ for an Airy layer in the form

$$\gamma = \left|\frac{2nn'}{k}\right|^{1/3} = n\left(\frac{2}{\rho}\right)^{1/3}\left|\frac{n'r}{n}\right|^{1/3} = \frac{n\beta^{1/3}}{K_\rho} \tag{4.12-3}$$

In a geometric optics framework with spherical stratification, $|n'r/n|$ is the ratio of the radius of curvature of the stratified surface r to the radius of curvature of the refracted ray $|n/n'|$. In Chapter 2, it is defined as $(-\beta)$. For the Earth's atmosphere at sea level, this ratio has a value of about 1/4 for the dry air component. For water vapor layers, it can exceed unity.

This correspondence between the forms in Eqs. (4.12-1) and (4.12-2) allows us to perform a "flattening" transformation. We may convert a series of large concentric spherical layers (with the refractivity constant within a layer but discontinuous across boundaries) into an equivalent Cartesian stratified medium consisting of a stack of Airy layers.

A similar correspondence applies when one has an Airy layer with spherical stratification. Here we have a linearly varying index of refraction, $n = n_0 + n'(r - r_0)$, where n_0 and n' are constants within a layer. Referring to the modified Bessel equation in Eq. (4.10-3), if we expand ρ and l about $\rho_0 = kn_0r_0$ while retaining only first-order terms, we obtain

$$\left.\begin{aligned} \frac{d^2U_l}{d\rho^2} &\doteq \frac{2}{\rho_0}\Big((l-\rho_0)-(\rho-\rho_0)+\beta(\rho-\rho_0)\Big)U_l \\ \beta &= -n'r_0/n_0 \end{aligned}\right\} \tag{4.12-4}$$

Here U_l is a modified Bessel function of order l, and $\beta = -n'r_0/n_0$ is a constant, which, as mentioned earlier, will be substantially smaller than unity when the thin-atmosphere assumption applies. Also, we assume that $n'(r - r_0)$ is sufficiently small so that the quadratic term for $n(r)$ can be ignored in Eq. (4.12-4). If we now make the transformation

$$\tilde{y} = \frac{(2/\rho_0)^{1/3}}{|1-\beta|^{2/3}}\big(l - \rho + \beta(\rho - \rho_0)\big) \tag{4.12-5}$$

it follows that Eq. (4.12-4) becomes

$$\frac{d^2U}{d\tilde{y}^2} - \tilde{y}U = 0 \tag{4.12-6}$$

which is the differential equation for the Airy functions. It follows that the correspondence in Eq. (4.12-2) is modified for the spherical stratified Airy case. In particular, the parameters $K_v^{-1} \doteq \partial\tilde{y}/\partial\nu$ and $K_\rho^{-1} \doteq \partial\tilde{y}/\partial\rho$ for the case of constant refractivity are modified for an Airy layer to

$$\left.\begin{aligned} \frac{\partial\tilde{y}}{\partial\rho} &\doteq -K_{\rho_o}^{-1}(1-\beta)^{1/3} \\ \frac{\partial\tilde{y}}{\partial\nu} &\doteq K_{\rho_o}^{-1}(1-\beta)^{-2/3} \end{aligned}\right\} \tag{4.12-7}$$

Notice that this transformation in Eq. (4.12-5) for $\tilde{y}$ also works for $\beta > 1$, a region of super-refractivity. In this case, we have

$$\left.\begin{aligned} \tilde{y} &= \frac{1}{K_{\rho_o}(\beta-1)^{2/3}}\left(l - \rho + \beta(\rho - \rho_0)\right) \\ \frac{\partial\tilde{y}}{\partial\rho} &\doteq K_{\rho_o}^{-1}(\beta-1)^{1/3} > 0 \end{aligned}\right\} \tag{4.12-8}$$

In summary, for spherical or cylindrical stratification with $r_0/\lambda >> 1$, a stack of concentric layers with the refractivity held constant within a layer, but allowed to vary from layer to layer, leads to the same mathematical expressions for the *l*th spectral component of wave propagation as those expressions obtained from a stack of Airy layers with Cartesian stratification with an angle of incidence determined from Eq.(4.12-2b). Moreover, to the extent that a stack of Airy layers can be made through the first-order approximation of the characteristic matrix to closely replicate wave propagation through a Cartesian stratified medium with an arbitrarily varying refractivity profile, this duality should yield similar accuracy for the spherical or cylindrical stratification, also with an arbitrarily varying refractivity profile. Thus, the problem of wave propagation in large cylindrical or spherical stratified media can be transformed to an analogous problem of wave propagation in a Cartesian stratified medium. Also, summing over *l* in cylindrical or spherical stratification is equivalent to integrating over n_o in an Airy layer in the Cartesian stratified medium. Integrating over n_o is equivalent to integrating over angle of incidence or "impact parameter."

For a given spectral number in the curvilinear system, the relationship between angles of incidence in the two systems is given by

$$\nu = \rho \sin\theta_v, \quad \cos\theta_v = /\, n' r / n \,/^{-1/3} \cos\varphi = \beta^{-1/3} \cos\varphi \qquad (4.12\text{-}9)$$

If we force equality in angles of incidence in the two systems, then Eq. (4.12-9) yields a value for n' to be used in the surrogate Cartesian stratified model using Airy layers. This value is $n' = n / r$, the magnitude of the super-refractivity critical gradient. In this case, it follows that γ for the Cartesian Airy layer is given by $\gamma = n_o / K_{\rho_o}$.

References

[1] F. Abelès, "Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Applications to Thin Films," *Annales de Physique*, vol. 5, pp. 596–640, September–October 1950 (in French).

[2] F. Abelès, "Investigations on the Propagation of Sinusoidal Electromagnetic Waves in Stratified Media. Application to Thin Films. II. Thin Films," *Annales de Physique*, vol. 5, pp. 706–782, November–December 1950 (in French).

[3] M. Born and E. Wolf, *Principles of Optics,* 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.

[4] S. Schelkunoff, "Solution of Linear and Slightly Nonlinear Differential Equations," *Quarterly of Applied Mathematics*, vol. 3, no. 4, pp. 348–355, 1946.

[5] R. Jacobsson, "Light Reflection from Films of Continuously Varying Refractive Index," in *Progress in Optics*, E. Wolf, ed., vol. 5, New York: John Wiley, pp. 247-286, 1966.

[6] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, United Kingdom: The Institution for Electrical Engineers, 2000.

[7] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards Series 55, Washington, DC, 1964.

[8] J. Jackson, *Classical Electrodynamics,* 2nd ed., New York: John Wiley & Sons, Inc., 1975.

[9] H. Bremmer, "The W.K.B. Approximation as the First Term of a Geometric-Optical Series," *The Theory of Electromagnetic Waves, A Symposium*, pp. 105–115, Interscience, New York, 1951.

Chapter 5
Propagation and Scattering in a Spherical-Stratified Refracting Medium

5.1 Introduction

We extend the Mie scattering treatment of Chapter 3 to the case of a scattering spherical surface embedded in a refracting medium that is laminar but otherwise spherical symmetric. Figure 5-1 depicts this scenario. At the boundary, n or one of its derivatives may be discontinuous with a refraction profile $n^{+}(r)$ in the overlying medium and $n^{-}(r)$ inside the sphere. The angle α_{L} is the refractive bending angle at the low Earth orbiting (LEO) spacecraft. The angle $\tilde{\alpha}$ is the cumulative bending angle along the ray. A thin phase screen approach to this problem is discussed in Chapter 2. There the scalar diffraction integral, which is a convolution integral over the vertical continuum of emitters in a thin screen model, is used to calculate the phase and amplitude of the emitted wave at a point some distance from the thin screen. In that essentially wave/optics approach, the thin screen serves as a surrogate for the actual atmosphere. The radiation field from the emitters in the screen mimics the phase and amplitude effects on the electromagnetic wave resulting from propagating through the actual atmosphere, including the effects from the refractive gradient and from the embedded scattering surface. Fresnel diffraction, interference, shadowing, caustics, etc., all can be evaluated using the thin phase screen approach combined with the scalar diffraction integral.

A wave theory approach is based on solutions to Maxwell's equations applied to a spherical atmosphere. A convolution integral also appears in wave theory, but it is over spectral number instead of thin screen altitude. The wave theory approach can be considered more accurate, albeit computationally more expensive. In both approaches, thin-screen/scalar diffraction theory and wave

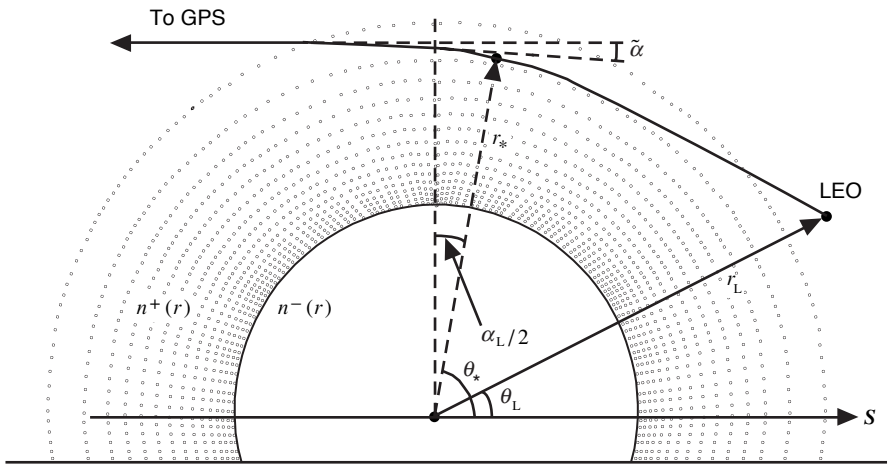

Fig. 5-1. Ray path geometry for a scattering sphere embedded in a stratified medium. The profiles for the index of refraction are given by $n^-(r)$ inside the sphere and by $n^+(r)$ outside the sphere.

theory, one ends up with a prediction of the observed phase and amplitude of the wave at some point as a result of its passing through an intervening atmosphere and perhaps encountering an embedded scattering surface. One question addressed here is the level of agreement between these two approaches, and how that level depends on the adversity of the wave propagation conditions in the atmosphere.

One concludes from a review of wave propagation literature that scattering theory in a sphere is potentially a very complicated problem. For example, see the survey article by Chapman and Orcutt [1] on wave propagation problems in seismology. There one finds refracted rays reflecting from multiple surfaces, Rayleigh and Love waves skittering along boundaries, super-refracted waves with multiple reflections ducted along between layers, and so on. Here we specifically rule out ducting, evanescence, or other confounding propagation effects, except for the effects resulting from the class of discontinuities under study here, which would include interference, shadow zones, caustics, diffraction, and super-refractivity. We assume that embedded in and co-centered with this refracting medium is a single large spherical scattering surface. Across this surface a discontinuity in the refractivity model is assumed to exist. This discontinuity can take different forms, ranging from a discontinuity in n itself, or in its gradient, or to a discontinuity merely in one of its higher derivatives. We must account for the effects of the refractive gradient in the overlying medium surrounding the sphere on the phase and amplitude of the electromagnetic wave. Therefore, scattering in this context includes external reflection from the scattering sphere, transmission through and refraction by the scattering sphere, including the possibility that the scattered wave has

undergone one or more reflections inside the scattering sphere, and finally the refractive bending of the scattered ray from the overlying medium.

For the purpose of comparing results from the full-wave theory approach with the scalar diffraction/thin phase screen approach, we assume that the local gradient of the refractivity is sufficiently small throughout the medium surrounding the scattering sphere so that the "thin atmosphere" conditions [see Section 2.2, Eqs. (2.2-8) and (2.2-9)] apply. Where rapid changes in refractivity are encountered—for example, at the boundary of a super-refracting water vapor layer—we assume that such changes are sufficiently localized that ray optics is still valid, i.e., rays do exist that pass through such a barrier, for at least a certain range of tangency points.

The wave theory approach followed here is derived from Mie scattering theory, but it is adapted to a medium with a continuously changing refractivity. The original formulation of Mie scattering theory [2] deals with a single spherical scattering surface in an otherwise homogeneous medium. Numerical wave theory approaches involve approximations to the solutions of Maxwell's equations in one form or another. In this chapter, we use an osculating parameter technique for dealing with the spectral integrals associated with wave theory. The accuracy of such a technique and its range of applicability are important questions that need to be addressed. The accuracy and range of applicability depend on the choice of basis functions used in the osculating parameter technique. For example, in a Cartesian-stratified medium, the use of sinusoids as the basis functions results in an osculating parameter solution that is identical to the Wentzel–Kramer–Brillouin (WKB) solution. The conditions for attaining a given accuracy and ascertaining its range of applicability are well established for WKB solutions. There also is a wealth of literature on the connection problem in WKB solutions across the transition zone between the oscillatory and exponential-decaying branches, important for quantum tunneling processes and other applications. Fortunately, we are concerned with the electromagnetic field away from turning points; therefore, asymptotic forms applicable to the oscillatory branch play an important role here. For a different choice of basis functions, the osculating parameter solutions do not reduce to the WKB forms and have a different range of applicability. Here the favored basis functions are the spherical Bessel functions or their Airy function surrogates, which are asymptotically equivalent when the radius of the scattering surface is very large compared to the wavelength of the electromagnetic wave. These particular basis functions offer a wide range of applicability for the osculating parameter solutions. Even at a turning point, a bête noir for wave theory, these basis functions provide a useful, if not completely successful, approach.

The question arises concerning the many sections to follow as to which parts are essential to this wave theory approach. Sections 5.2 and 5.3 provide a brief review of the basic general concepts in classical electrodynamics

involving harmonic waves: Maxwell's equations, scalar potentials for generating the electromagnetic field vectors, and series solutions using the separation of variables technique. These series involve spherical harmonic functions, which apply for spherical symmetry, but special functions are needed for the radial component. For the homogeneous case, these radial solutions become the spherical Bessel functions, but in general the radial functions depend on the refractivity profile of the medium. It is here that techniques like the WKB method or the osculating parameter technique arise.

Section 5.4 briefly summarizes the asymptotic approximations that are used in this chapter. A fuller account is found in Section 3.8. This section is referred to frequently in the later text.

Section 5.5 begins the adaptation of Mie scattering theory from a single large spherical surface to a concatenated series of concentric layers that in their limiting form approach a medium with a continuously varying refractivity. This section introduces a spectral density function for the phase delay induced by the refractive gradient in the medium. This quantity (defined as $G[\rho,\nu]$ in that section) essentially accounts for the extra phase delay at the radial position ρ experienced by a radial wave component of integer spectral number $l=\nu-1/2$, which results from the refractive gradient of the medium. In a homogeneous medium, $G[\rho,\nu]\equiv 0$. We consider there the propagation of an incident wave that asymptotically is planar at large approaching distances relative to the scattering sphere. The adjustments to account for a spherical approaching wave (when the emitting Global Positioning System (GPS) satellite is placed only a finite distance away) are noted.

Section 5.6 reviews several important concepts from geometric optics that are needed later when correspondences are established between these concepts and certain properties from wave theory when the spectral number assumes a stationary phase value. Geometric optics is discussed in Appendix A, but here, in addition to discussing the stationary phase property of a ray and its bending angle and phase delay, this section also introduces the concept of a cumulative bending angle along the ray, which mainly arises when evaluating the electromagnetic field within the refractive medium. This section also discusses Bouguer's law and the impact parameter of a ray, the geometric optics equivalent of the conservation of angular momentum in a conservative force field. This section also covers defocusing and the first Fresnel zone. Limitations in second-order geometric optics, which arise in association with caustics or when two or more rays have impact parameter separations that are less than the first Fresnel zone, are discussed in Section 5.12.

Section 5.7 develops more asymptotic forms needed later. The ratio of the radius of curvature of the stratified surface to the wavelength of the incident wave, r_o/λ, is sufficiently large that asymptotic forms for the Bessel functions apply. Also, because only spectral numbers that are of the same order of

magnitude as r_o / λ contribute significantly to the spectral integrals representing the field, we also can use the asymptotic forms for the harmonic functions that apply for large spectral number. Section 5.7 shows the close correspondence between certain geometric optics quantities—for example, the cumulative bending angle of a ray with an impact parameter value, ν, and evaluated at a radial position, ρ, and a certain spectral quantity from wave theory, $\partial G[\rho,\nu]/\partial\nu$. The issue of a breakdown in accuracy of the osculating parameter technique near a turning point $\nu = \rho$ also is addressed here. The behavior of the WKB solutions near a turning point is used to provide guidance in dealing with this breakdown. An asymptotic matching technique is developed to set the value of $G[\rho,\nu]$ for the regime $\nu > \rho$.

Section 5.8 begins the representation of the electromagnetic field in terms of the spectral integrals involving the spectral components of the radial osculating parameter functions and the harmonic functions for the angle coordinates.

This discussion is continued in Section 5.9, where a phasor representation for the integrands in these spectral integrals is introduced. The stationary phase technique also is introduced here. It is used to determine spectral number points that yield stationary values of the phasor, thereby aiding the numerical evaluation of the spectral integrals.

Section 5.10 compares results from wave theory with results from a thin phase screen model combined with the scalar diffraction integral. Correspondences between stationary phase values of certain wave theoretic quantities and their analogs in geometric optics are discussed.

Section 5.11 deals with the turning point problem using an Airy layer.

Section 5.12 discusses caustics and multipath from a wave theoretic point of view in a spectral number framework. It also discusses caustics and multipath in a second-order geometric optics framework, including its shortcomings near caustics or in dealing with ray pairs with nearly merged impact parameters. Third-order stationary phase theory is introduced to develop a ray theory that can accurately deal with these situations. Beginning in Section 5.12, numerical solutions for the spectral representation of the field at the LEO are presented. Here the numerical integrations have been aided by the stationary phase technique to identify contributing neighborhoods in spectral number, greatly improving the efficiency of the technique.

Section 5.13 deals with a spherical scattering surface embedded in an overlying refracting medium.

Finally, Section 5.14 discusses the perfectly reflecting sphere that is embedded in an overlying refracting medium. This section also discusses the correspondence between geometric optics quantities and wave theory quantities when stationary phase values are used in each system. For example, at a LEO position sufficiently away from the shadow boundary of the reflecting sphere,

the stationary phase values in spectral number in wave theory closely correspond to ray path impact parameter values that satisfy the law of reflection.

5.2 Maxwell's Equations in a Stratified Linear Medium

We follow closely the development given in Section 3.2 for the homogeneous case; the relevant symbols are defined in that section (see also the Glossary). Here Gaussian units are used. A harmonic electromagnetic wave may be written in the form

$$\left.\begin{aligned} \boldsymbol{E} &= \boldsymbol{E}(\boldsymbol{r})\exp(-i\omega t) \\ \boldsymbol{H} &= \boldsymbol{H}(\boldsymbol{r})\exp(-i\omega t) \end{aligned}\right\} \tag{5.2-1}$$

Maxwell's equations for the time-independent components in a linear medium, free of charge and current densities, are given by

$$\nabla \times \boldsymbol{E} = ik\mu\boldsymbol{H} \tag{5.2-2a}$$

$$\nabla \times \boldsymbol{H} = -ik\varepsilon\boldsymbol{E} \tag{5.2-2b}$$

$$\nabla \cdot (\varepsilon\boldsymbol{E}) = 0 \tag{5.2-2c}$$

$$\nabla \cdot (\mu\boldsymbol{H}) = 0 \tag{5.2-2d}$$

Here ε is the electrical permittivity of the propagation medium, μ is its magnetic permeability, and $k = 2\pi/\lambda$; k is the wave number of the harmonic wave in a vacuum, $\omega = kc$, where c is the velocity of light. These equations in Eq. (5.2-2) may be recast through successive vector calculus operations into separate vector wave equations that $\boldsymbol{E}$ and $\boldsymbol{H}$ must individually satisfy [3]. These are given by

$$\nabla^2\boldsymbol{E} + \mu\varepsilon k^2\boldsymbol{E} + \nabla(\log\mu)\times(\nabla\times\boldsymbol{E}) + \nabla\left[\boldsymbol{E}\cdot(\nabla(\log\varepsilon)\right] = 0 \tag{5.2-3a}$$

$$\nabla^2\boldsymbol{H} + \mu\varepsilon k^2\boldsymbol{H} + \nabla(\log\varepsilon)\times(\nabla\times\boldsymbol{H}) + \nabla\left[\boldsymbol{H}\cdot(\nabla(\log\mu)\right] = 0 \tag{5.2-3b}$$

Here the identity $\nabla\times\nabla\times\boldsymbol{A} = \nabla(\nabla\cdot\boldsymbol{A}) - \nabla^2\boldsymbol{A}$ is used. These are the modified wave equations that the time-invariant component of a harmonic wave must satisfy in a linear medium.

We assume now that the medium is spherical stratified. In this case, the index of refraction is a function of only the radial coordinate,

$$n(r) = \sqrt{\mu(r)\varepsilon(r)} \tag{5.2-4}$$

It follows for this case that the gradient vectors of $\varepsilon(r)$ and $\mu(r)$ are radial directed, which simplifies Eq. (5.2-3).

For the special case where $\mu \equiv 1$ throughout the medium, which is essentially the case for L-band radio signals in the neutral atmosphere, Eq. (5.2-3) is further simplified. In the special case where $\boldsymbol{E}$ is perpendicular or transverse to $\boldsymbol{r}$, which is the so-called transverse electric (TE) wave, then $\nabla\varepsilon \cdot \boldsymbol{E}_{\mathrm{TE}} \equiv 0$ and Eq. (5.2-3a) becomes

$$\nabla^2 \boldsymbol{E}_{\mathrm{TE}} + n^2 k^2 \boldsymbol{E}_{\mathrm{TE}} = 0 \tag{5.2-5}$$

Equation (5.2-5) is nearly the Helmholtz equation [see Section 3.2, Eq. (3.2-1c)] except for the radial dependency of $n(r)$. This variation of $n(r)$ will be very slight in our case of a thin atmosphere, except possibly at a boundary. But, because r_o / λ is so large, even a small variation, δn, results in a significant change, $k\delta n$, in the gradient of the phase accumulation of the wave.

5.2.1 Scalar Potential Functions

Following the approach in Section 3.2 for Mie scattering theory, we use the scalar potential functions for the electromagnetic field in a stratified medium expressed as a series summed over integer spectral number. It is convenient to express the electromagnetic field vectors in terms of vector calculus operations on a pair of scalar potentials, ${}^{\mathrm{e}}\Pi(r,\theta,\phi)$ and ${}^{\mathrm{m}}\Pi(r,\theta,\phi)$. In Section 3.2, it is shown [3] for the case where ε and μ are constant that these two scalar potentials are linearly independent solutions to the Helmholtz equation:

$$\nabla^2 \Pi + k^2 n^2 \Pi = 0 \tag{5.2-6}$$

where, in Section 3.2, n is a constant. Each solution for a homogeneous medium can be represented using the technique of separation of variables in spherical coordinates as a series expansion. The series is expressed in terms of spherical Bessel functions of integer order l, which are a function of the radial coordinate $\rho = nkr$, and the spherical harmonic functions of degree m and order l, which are functions of the angular coordinates θ and ϕ. Here θ is the angle between $\boldsymbol{r}$ and the z-axis. The latter is the axis of propagation [the asymptotic direction of the Poynting vector $\boldsymbol{S}$ (Fig. 5-1)] for the approaching wave. Also, ϕ is the azimuth angle about the z-axis. See Fig. 4-10 for the definition of the coordinate frame.

The electromagnetic field vectors for the homogeneous medium are obtained from a particular vector form for the scalar potentials (first introduced by Hertz). These are given by

$$\left.\begin{aligned} \boldsymbol{E} &= \nabla\times\nabla\times\left({}^{\mathrm{e}}\Pi\boldsymbol{r}\right)+ik\mu\nabla\times\left({}^{\mathrm{m}}\Pi\boldsymbol{r}\right) \\ \boldsymbol{H} &= \nabla\times\nabla\times\left({}^{\mathrm{m}}\Pi\boldsymbol{r}\right)-ik\varepsilon\nabla\times\left({}^{\mathrm{e}}\Pi\boldsymbol{r}\right) \end{aligned}\right\} \tag{5.2-7}$$

In the electrodynamics literature, the TE and transverse magnetic (TM) waves are generated from linearly independent solutions to the Helmholtz equation in Eq. (5.2-6). In Eq. (5.2-7), the term $ik\mu\nabla\times\left[{}^{\mathrm{m}}\Pi\boldsymbol{r}\right]$ generates the electric field $\boldsymbol{E}_{\mathrm{TE}}$, which is perpendicular to $\boldsymbol{r}$, that is, a transverse electric field; this wave is known in the literature as the TE wave. Similarly, the term $-ik\varepsilon\nabla\times\left[{}^{\mathrm{e}}\Pi\boldsymbol{r}\right]$ generates a transverse magnetic field $\boldsymbol{H}_{\mathrm{TM}}$, or the TM wave. One can readily show (see Appendix I) that these expressions in Eq. (5.2-7) yield field vectors that satisfy Maxwell's equations in Eq. (5.2-2) when ε and μ are constant.

For the stratified medium with $n=n(r)$, the scalar potentials are solutions to a modified Helmholtz equation. In classical electrodynamics, there is a certain degree of arbitrariness in the definition of the scalar electric potential Φ and the vector magnetic potential $\boldsymbol{A}$ from which $\boldsymbol{E}$ and $\boldsymbol{H}$ are derived. Specifically, the electromagnetic field remains invariant if Φ and $\boldsymbol{A}$ are transformed together to some other pair of functions through a so-called gauge transformation; that is, the transformation is effected while Φ and $\boldsymbol{A}$ are constrained to satisfy a gauge condition such as that provided by the Lorentz condition [4]. The electromagnetic field is called gauge invariant. It is rooted in the symmetries in the electrodynamics equations when they are expressed in the space-time framework of Special Relativity. There, the form of the electrodynamics equations for the 4-vector $(\Phi,\boldsymbol{A})$ remains invariant under a Lorentz transformation; the 4-vector $(\Phi,\boldsymbol{A})$ is called covariant in a relativistic framework.

Similarly, the scalar potentials for the stratified medium have some degree of freedom in their definition. For the case where $n=n(r)$, it is shown in Appendix I that the electromagnetic field can be expressed through vector calculus operations on the modified scalar potentials, $\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]$ and $\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]$. These expressions are given by

$$\left.\begin{aligned} \boldsymbol{E} &= \varepsilon^{-1}\nabla\times\nabla\times\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]+ik\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right] \\ \boldsymbol{H} &= \mu^{-1}\nabla\times\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]-ik\nabla\times\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right] \end{aligned}\right\} \tag{5.2-8}$$

The factors $\varepsilon^{1/2}$ and $\mu^{1/2}$ have been inserted into the potential terms in Eq. (5.2-8) to simplify the resulting modified Helmholtz equation that each of the scalar potentials must satisfy. These scalar potentials must satisfy modified Helmholtz equations, which are given by

$$\left.\begin{aligned}\nabla^{2\,\mathrm{e}}\Pi + k^2\tilde{n}^{2\ \mathrm{e}}_{\mathrm{TM}}\Pi = 0\\ \nabla^{2\,\mathrm{m}}\Pi + k^2\tilde{n}^{2\ \mathrm{m}}_{\mathrm{TE}}\Pi = 0\end{aligned}\right\} \tag{5.2-9}$$

Here the modified indices of refraction are

$$\left.\begin{aligned}\tilde{n}^2_{\mathrm{TM}} &= n^2\left[1-\frac{r^2\varepsilon^{1/2}}{k^2n^2}\frac{d}{dr}\left(\frac{1}{r^2}\frac{d}{dr}\left(\frac{1}{\varepsilon^{1/2}}\right)\right)\right]\\ \tilde{n}^2_{\mathrm{TE}} &= n^2\left[1-\frac{r^2\mu^{1/2}}{k^2n^2}\frac{d}{dr}\left(\frac{1}{r^2}\frac{d}{dr}\left(\frac{1}{\mu^{1/2}}\right)\right)\right]\end{aligned}\right\} \tag{5.2-10}$$

For the case where $\mu(r)\equiv 1$ throughout the medium, we note from Eq. (5.2-10) that the modified index of refraction for the TE wave reduces to the regular index of refraction. When the conditions $|\nabla n|<<1$ and $kr_o >> 1$ apply, which do apply for L-band propagation in the Earth's thin neutral atmosphere, it follows that $\tilde{n}(r)$ differs from $n(r)$ by a small amount of the order of n''/k^2; for L-band signals in dry air, this translates into a fractional difference in refractivity of roughly parts in 10^{11}. So, for computations in neutral atmosphere conditions, we may simply use $n(r)$ in the modified Helmholtz wave equation. Therefore, we herewith drop the distinction between $\tilde{n}_{\mathrm{TE}}$ or $\tilde{n}_{\mathrm{TM}}$ and $n(r)$, and simply use $n(r)$ in the modified Helmholtz equation in the following discussion. For the ionosphere, these differences may be more significant.

5.3 Modified Spherical Bessel Functions

We assume now that our stratified medium satisfies the asymptotic condition $n(r)\to 1$ as $r\to\infty$, so that the scalar potential series solutions for the homogeneous medium in Section 3.2 can be used as asymptotic boundary conditions for the stratified case. For the stratified medium, we again seek solutions to the modified Helmholtz equation using the technique of separation of variables of the form

$$\Pi = R(r)\Theta(\theta)\Phi(\phi) \tag{5.3-1}$$

where $\Pi(r,\theta,\phi)$ may be taken as any spectral component of a scalar potential. For large values of r, where $n(r)\to 1$, we know that these solutions must approach the forms given in Chapter 3 for the homogeneous case. We also conclude because of the spherical symmetry of the propagation medium that spherical harmonic functions will be applicable; that is, the $\Theta(\theta)$ functions will be the same associated Legendre polynomials $\mathrm{P}_l^m(\cos\theta)$ of order l and

degree m, and the $\Phi(\phi)$ functions will be sinusoids of the form $\exp(\pm im\phi)$. For an electromagnetic vector field, the m values are restricted to $m = \pm 1$. This follows from Bauer's identity, applicable to a plane wave in a homogeneous medium (see, for example, Chapter 3, Eq. (3.2-3)), and also [3]. Referring to Fig. 1-6, Bauer's identity is obtained from the multipole expansion [4] for a spherical wave centered at the point G and evaluated at the point L. The amplitude and phase of the time-independent part of the spherical wave is given by $\exp(inkr_{\mathrm{LG}})/(nkr_{\mathrm{LG}})$. Its expansion in terms of spherical Bessel and spherical harmonic functions of the transmitter and receiver coordinates is given by

$$\left.\begin{aligned} &\frac{\exp(i\rho_{\mathrm{LG}})}{\rho_{\mathrm{LG}}} = i\sum_{l=0}^{\infty}(2l+1)\frac{\psi_l(\rho_{\mathrm{L}})}{\rho_{\mathrm{L}}}\frac{\xi_l^{+}(\rho_{\mathrm{G}})}{\rho_{\mathrm{G}}}\mathrm{P}_l\left(\cos(\theta_{\mathrm{G}}-\theta_{\mathrm{L}})\right), \\ &\rho_{\mathrm{L}} < \rho_{\mathrm{G}},\ \ \rho = nkr \end{aligned}\right\} \qquad (5.3\text{-}2)$$

which is obtained by applying the addition theorem for spherical harmonic functions. If we now let $\rho_{\mathrm{G}} \to \infty$ and $\theta_{\mathrm{G}} = \pi$, then $\rho_{\mathrm{LG}} \to \rho_{\mathrm{G}} + \rho_{\mathrm{L}}\cos\theta_{\mathrm{L}}$ and we can replace $\xi_l^{+}(\rho_{\mathrm{G}})$ with its asymptotic form for large $\rho_{\mathrm{G}} >> l$, $\xi_l^{+}(\rho_{\mathrm{G}}) \to (-i)^{l+1}\exp(i\rho_{\mathrm{G}})$. We substitute these forms into the above expansion for the spherical wave, cancel terms, and note that $\mathrm{P}_l(-x) = (-1)^l\,\mathrm{P}_l(x)$. It follows that, for a plane harmonic wave traveling along the z-axis in a homogeneous medium, the time-independent component is given by $\exp(i\rho\cos\theta)$ and that Bauer's identity is given by

$$\exp(i\rho\cos\theta) = \sum_{l=0}^{\infty} i^l(2l+1)\frac{\psi_l(\rho)}{\rho}\mathrm{P}_l(\cos\theta),\ \ \rho = nkr \qquad (5.3\text{-}3)$$

The vector version (for a plane wave with its electric field vector directed along the x-axis in Fig.4-10) is given by multiplying Eq. (5.3-3) by $\left(\hat{\boldsymbol{r}}\sin\theta\cos\phi + \hat{\boldsymbol{\theta}}\cos\theta\cos\phi - \hat{\boldsymbol{\phi}}\sin\phi\right)$. When the coefficients of the basis functions $R(r,l)\,\mathrm{P}_l^m(\cos\theta)\exp(\pm im\phi)$ in the series solution to Eq. (5.2-6) for a given vector component of the field are matched on a term-by-term basis with the corresponding coefficients in the Bauer series for the same vector component (and using the property $\partial\,\mathrm{P}_l/\partial\theta = -\mathrm{P}_l^1$), one finds that m is indeed restricted to $m = \pm 1$. This restriction is perpetuated to the scattered field by the continuity conditions in electrodynamics that apply to the field components

across a scattering boundary.[1] The form of the series solutions in this case must approach, as $r \to \infty$, the same form as that given in Chapter 3, Eq. (3.2-4). Only the $R(r)$ functions will differ from the spherical Bessel functions that apply to the homogeneous case, and these modified functions will approach the Bessel function form as $r \to \infty$. Thus, we have

$$rR = \tilde{\xi}_l^{\pm}(kr) \tag{5.3-4}$$

where $\tilde{\xi}_l^{\pm}$ is related to the spherical Hankel functions of the first (+) and second (–) kind, but modified for the stratified medium. These functions must satisfy the modified differential equation for spherical Bessel functions, which is given by

$$\tilde{\xi}_l^{\pm\,\prime\prime} + \left(n(r)^2 - \frac{l(l+1)}{u^2} \right) \tilde{\xi}_l^{\pm} = 0 \tag{5.3-5}$$

Here, $u = kr$ and $(*)' = d(*)/du$. See Section 3.2, Eq. (3.2-8), for the definition of these spherical Hankel functions in the case of a homogenous medium in terms of the integer Bessel functions of the first and second kind. In particular, the relationships between the modified spherical Bessel functions of the first and second kind, $\tilde{\psi}_l(u)$ and $\tilde{\chi}_l(u)$, for the stratified medium and the modified spherical Hankel functions are given by

$$\left.\begin{aligned} \tilde{\xi}_l^{+}(u) &= \tilde{\psi}_l(u) + i\tilde{\chi}_l(u) \\ \tilde{\xi}_l^{-}(u) &= \tilde{\psi}_l(u) - i\tilde{\chi}_l(u) \\ \tilde{\psi}_l(u) &= \frac{1}{2}\tilde{\xi}_l^{+}(u) + \frac{1}{2}\tilde{\xi}_l^{-}(u) \\ \tilde{\chi}_l(u) &= \frac{1}{2i}\tilde{\xi}_l^{+}(u) - \frac{1}{2i}\tilde{\xi}_l^{-}(u) \end{aligned}\right\} \tag{5.3-6}$$

For $u \to 0$, $\tilde{\psi}_l(u) \to 0$ and $\tilde{\chi}_l(u) \to \infty$.

The form of the modified spherical Hankel functions $\tilde{\xi}_l^{\pm}$ will depend on the functional form of $n(r)$. For example, let the special function for the index of refraction be given by

$$n^2 = 1 + \eta + \beta\left(\frac{r_o}{r} \right)^2 \tag{5.3-7}$$

---

[1] The quantum mechanical analog of this restriction in *m* values for a photon is that its angular momentum vector is restricted to a unit value times Planck's constant parallel or anti-parallel to $\boldsymbol{S}$, the Poynting vector.

where η and β are constants. This was introduced in [5]. From Eq. (5.3-5), it can be shown that this form offsets the spherical Hankel function in argument and spectral number:

$$\left.\begin{aligned} \tilde{\xi}_l^{\pm}(u) &= \xi_{\tilde{l}}^{\pm}(\tilde{u}) \\ \tilde{u} &= u\sqrt{1+\eta} \\ \tilde{l}^2 + \tilde{l} &= l^2 + l + \beta u_o^2 \end{aligned}\right\} \tag{5.3-8}$$

For a thin atmosphere, $\eta + \beta \approx 0$; these parameters may be individually chosen to match the index of refraction and its gradient at $u = u_o$. For example, for dry air at the Earth's surface, $\eta + \beta \approx 1/4 \times 10^{-3}$ and $\beta \doteq -r_o n' \approx 0.2$. On the other hand, Eq. (5.3-7) does not satisfy our asymptotic boundary condition of $n(r) \to 1$ as $r \to \infty$. This form for $n(r)$ in Eq. (5.3-7) is useful for regional applications or over thin layers with boundaries on the top and bottom sides, and it has been used to study ducting, tunneling, super-refractivity, and other propagation effects in a strongly refracting medium.

Another technique, applicable when $n(r)$ assumes a general form, uses the WKB method to obtain an approximate expression for $\tilde{\xi}_l^{\pm}(u)$. We define $f_l(u)$ by

$$f_l = \frac{n^2 u^2 - l(l+1)}{u^2} \tag{5.3-9}$$

The WKB approximate solution, $W_l^{\pm}(u) \doteq \tilde{\xi}_l^{\pm}(u)$, to Eq. (5.3-5) is given by [6]

$$W_l^{\pm}(u) = \left(f_l\right)^{-(1/4)} \exp\left(\pm i \int_{u_o}^{u} \sqrt{f_l}\, du\right) \tag{5.3-10}$$

Depending on the sign of $f_l(u)$, $W_l^{\pm}(u)$ has either an exponential form or a sinusoidal form. The WKB method has very widespread applicability. For examples in seismology, see Chapman and Orcutt's review [1]. It also has been mentioned in Chapter 4 in regard to wave propagation through a Cartesian-stratified medium.

We will use an osculating parameter technique here. When $n(r)$ is variable, we may write

$$\left.\begin{aligned} \tilde{\xi}_l^{\pm}(u) &= a_l^{\pm}(\rho)\xi_l^{\pm}(\rho) \\ \rho &= krn(r) \\ u &= kr \end{aligned}\right\} \tag{5.3-11}$$

where $a_l^{\pm}(\rho)$ is a so-called osculating parameter. It carries the deviation in amplitude and phase of $\tilde{\xi}_l^{\pm}(u)$ from these quantities in $\xi_l^{\pm}(\rho)$ due to the variability of $n(r)$.

The general series solution for the scalar potentials in a spherical stratified medium using this osculating parameter approach is given by

$$\left.\begin{aligned}\Pi^{\pm}(r,\theta,\phi) &= \sum_{l=0}^{\infty}\left(a_l^{\pm}(\rho)\cos\phi + b_l^{\pm}(\rho)\sin\phi\right)\frac{\xi_l^{\pm}(\rho)}{\rho}\mathrm{P}_l^1(\cos\theta)\\ \rho &= krn(r)\end{aligned}\right\} \quad (5.3\text{-}12)$$

In a homogeneous medium, these spectral coefficients $a_l^{\pm}$ and $b_l^{\pm}$ are functions only of the spectral number, and their form depends on the asymptotic boundary conditions for the waveform; see Eqs. (3.2-4) through (3.2-6). In the inhomogeneous but spherical symmetric medium, these spectral coefficients $a_l^{\pm}(\rho)$ and $b_l^{\pm}(\rho)$ vary also with ρ. The technique for obtaining their variability with ρ is rather similar to one of the parabolic equation techniques [7], but here their variability with ρ is due only to the gradient of the refractivity; the geometric component of the delay is retained by the $\xi_l^{\pm}(\rho)$ functions. Our task is to determine the form of these osculating spectral coefficients in a refracting medium in which a discontinuity also may be embedded, and to evaluate the series solutions for the electromagnetic field.

5.4 Asymptotic Forms

Because the spherical Bessel functions will be used extensively in later sections, we will need their asymptotic forms in terms of the Airy functions that are applicable for very large values of $\rho = knr$ and l. These have already been presented in Chapter 3, Section 3.8. There we established that the principal contributions to the spectral coefficients come from spectral number values in the vicinity of $l = \rho$. Therefore, asymptotic forms that exploit the relatively small value of $|l-\rho|/\rho$ but the large value of ρ are appropriate. All of the asymptotic forms presented in Section 3.8 carry over to the stratified case here with the transformation $x = kr \to \rho = nkr$ in the argument of the Bessel functions, and with $y \to \hat{y}$ in the argument of the Airy functions. The argument $\hat{y}$ is a function of $\rho = knr(r)$ and ν. We have placed the caret over y to indicate its dependence on n through ρ. The key asymptotic forms used later are summarized here.

From Eqs. (3.2-8) and (3.8-9), we have for the spherical Hankel functions when $l \approx \rho$:[2]

$$\left.\begin{aligned} \xi_l^{\pm}(\rho) &\sim \sqrt{\pi K_\rho}\left(1 - K_\rho^{-2}\frac{\hat{y}}{15} + \mathrm{O}\left[\rho^{-5/3}\right]\right)\left(\mathrm{Ai}[\hat{y}] \mp i\,\mathrm{Bi}[\hat{y}]\right) \\ \xi_l^{\pm\prime}(\rho) &\sim -\sqrt{\frac{\pi}{K_\rho}}\left(1 + K_\rho^{-2}\frac{\hat{y}}{15} + \mathrm{O}\left[\rho^{-5/3}\right]\right)\left(\mathrm{Ai}'[\hat{y}] \mp i\,\mathrm{Bi}'[\hat{y}]\right) \end{aligned}\right\},\ l \approx \rho \qquad (5.4\text{-}1)$$

where $\rho = knr(r)$ and $\nu = l + 1/2$, and where $\mathrm{Ai}[\hat{y}]$ and $\mathrm{Bi}[\hat{y}]$ are the Airy functions of the first and second kind, respectively. See also [8]. For convenience we will use the spectral number l and $\nu = l + 1/2$ more or less interchangeably. The distinction between them is inconsequential because of the enormity of their values in the stationary phase neighborhoods. The argument $\hat{y}$ is given by

$$\hat{y} = \nu^{2/3}\zeta\left(\frac{\nu}{\rho}\right) \qquad (5.4\text{-}2)$$

Here the auxiliary function $\zeta(\mu)$, $\mu = \nu/\rho$, and it series expansions in powers of $\left[\left(\rho^2 - \nu^2\right)/\nu^2\right]$ and in powers of $[(\rho - \nu)/\nu]$ are defined in Eqs. (3.8-4) and (3.8-5) for both regimes $\mu \geq 1$ and $\mu \leq 1$. Using these expansions, we summarize the key relationships between $\hat{y}$ and ρ and ν:

$$\left.\begin{aligned} \hat{y} &= \frac{1}{4K_\nu^4}\left(\nu^2 - \rho^2\right)\left[1 + \frac{2}{5}\frac{\nu^2 - \rho^2}{\nu^2} + \cdots\right] \\ \hat{y} &= \frac{1}{K_\rho}(\nu - \rho)\left(1 - \frac{\nu - \rho}{60K_\rho^3} + \cdots\right) \\ \nu &= \rho + K_\rho\hat{y}\left(1 + \frac{\hat{y}}{60K_\rho^2} + \cdots\right),\quad K_\nu = \left(\frac{\nu}{2}\right)^{1/3},\quad K_\rho = \left(\frac{\rho}{2}\right)^{1/3} \end{aligned}\right\} \qquad (5.4\text{-}3)$$

These truncated series expansions for $\hat{y}$ and ν are very accurate for large values of ρ with $\nu \approx \rho$. For most stationary phase neighborhoods, the value of $\hat{y}$ will be small compared to K_ρ. Therefore, the term $K_\rho^{-2}\hat{y}/15$ in Eq. (5.4-1)

---

[2] When $\rho >> l$, Eq. (5.4-1) is not appropriate. Starting from Eq. (3.8-10), it follows that $\xi_l^{\pm}(\rho) \to \mp i\exp(\pm i\mathrm{X})$, and $\mathrm{X} = (\rho^2 - \nu^2)^{1/2} - \nu\cos^{-1}(\nu/\rho) + \pi/4 \to \rho - l\pi/2$. Hence, $\xi_l^{\pm}(\rho) \to (\mp i)^{l+1}\exp(\pm i\rho)$ for $\rho >> l$.

can be dropped in the applications here. For GPS wavelengths, $K_\rho^{-2}/15 \approx 3\times10^{-7}$.

The quantity $K_\rho = (\pi nr/\lambda)^{1/3}$, a quasi-constant, appears frequently throughout this monograph. For GPS wavelengths at sea level, $K_\rho \approx 500$ and $2K_\rho/k \approx 30$ m. The latter turns out to be the spatial distance over which the Airy functions asymptotically transform from exponential functions to sinusoidal functions.

We also will need the asymptotic forms for the Airy functions. See [8] for a comprehensive discussion. They also are given in Chapter 3, Eq. (3.8-7), for negative values of $\hat{y}$ and by Eq. (3.8-8) for positive values.

5.5 Modified Mie Scattering in a Spherical Stratified Medium

The central task in this section is to derive the spectral density function for the phase delay incurred by the lth spectral component of the wave as a result of the refractive gradient of the medium. This function $G[\rho,\nu]$, with $\nu = l+1/2$, accounts for the extra phase delay in the lth spectral coefficient induced by only the refractive gradient of the medium. The geometric component of the phase delay is carried by the spherical Hankel function.

To follow a Mie scattering approach, we use the scalar potentials for the approaching, transmitted, and scattered wave. Electric and magnetic scalar potentials, ${}^{\mathrm{e}}\Pi^{(i)}$ and ${}^{\mathrm{m}}\Pi^{(i)}$, were discussed in Chapter 3 and also in Section 5.2. An incoming planar harmonic wave with in-plane polarization and with zero phase at $\theta = \pi/2$ can be represented by series solutions in terms of spherical Bessel functions and spherical harmonic functions. For a non-conducting homogeneous medium, these representations are given by

$$\left.\begin{aligned} {}^{\mathrm{e}}\Pi_l &= \frac{E_0}{nk}\, i^{l-1}\,\frac{(2l+1)}{l(l+1)}\,\frac{\psi_l(nu)}{nu}\,\mathrm{P}_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\Pi_l &= \frac{H_0}{nk}\, i^{l-1}\,\frac{(2l+1)}{l(l+1)}\,\frac{\psi_l(nu)}{nu}\,\mathrm{P}_l^1(\cos\theta)\sin\phi \\ E_0\sqrt{\varepsilon} &= H_0\sqrt{\mu};\quad u = kr,\quad n = \text{constant} \end{aligned}\right\} \tag{5.5-1}$$

Here ${}^{\mathrm{e}}\Pi_l$ and ${}^{\mathrm{m}}\Pi_l$ are the lth spectral components of the electric and magnetic scalar potentials, respectively.

To obtain the electromagnetic field from these scalar potentials, one uses the vector curl operations on their vector form given in Eq. (5.2-7). This vector form, $(\Pi_l \boldsymbol{r})$, is known as the Hertz potential. Here E_o is the amplitude of the electric field vector that lies in the plane defined by $\phi = 0$, that is, along the $\hat{\boldsymbol{x}}$

direction in Fig. 4-10. Similarly, H_o is the amplitude of the magnetic field vector, which points in the $\hat{\boldsymbol{y}}$ direction. From Maxwell's equations, it follows that $E_o\sqrt{\varepsilon} = H_o\sqrt{\mu}$.

Following that treatment for the homogeneous case, we obtain the series expansion solutions for the scalar potentials of the incoming wave in the spherical symmetric stratified medium with an index of refraction $n(u)$. Here the scalar potentials are given by

$$\left.\begin{aligned} {}^{\mathrm{e}}\tilde{\Pi} &= \frac{E_o}{nk}\sum_{l=1}^{\infty} {}^{e}a_l(nu)\frac{\psi_l(nu)}{nu}\mathrm{P}_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\tilde{\Pi} &= \frac{H_o}{nk}\sum_{l=1}^{\infty} {}^{m}a_l(nu)\frac{\psi_l(nu)}{nu}\mathrm{P}_l^1(\cos\theta)\sin\phi \\ E_o\sqrt{\varepsilon} &= H_o\sqrt{\mu}, \qquad u = kr \end{aligned}\right\} \tag{5.5-2}$$

The main difference from the homogeneous case is that we have introduced the spectral coefficients $a_l(nu)$, which are now variable with u, to account for the effects of the variability in $n(r)$. Each spectral component of these series satisfies the modified Helmholtz equation in Eq. (5.2-8); thus, the product $a_l(nu)\psi_l(nu)$ constitutes a formal solution to the modified spherical Bessel equation in Eq. (5.3-3). For each integer value of l, $a_l(nu)$ is an osculating parameter. The osculating parameter technique already has been discussed in Section 4.8 for a Cartesian-stratified medium. This technique is useful for solving certain ordinary differential equations where the rapidly varying component is carried by the basis function, $\psi_l(nu)$ in this case, and the more slowly (sometimes) varying component is carried by $a_l(nu)$.

We will need the asymptotic form for $a_l(nu)$ corresponding to an incoming wave well outside the atmosphere and its refractivity or scattering effects. The asymptotic form depends on where we place the emitting GPS satellite, either a finite or an infinite distance away, but always in the direction $\theta = \pi$. For the infinite case, the incoming waves are planar, and it follows from Eq. (5.5-1) that $a_l(nu)$ has the limit

$$a_l(nu) \underset{u\to\infty,\ n\to 1}{\to} i^{l-1}\frac{2l+1}{l(l+1)} \tag{5.5-3a}$$

This form satisfies the asymptotic boundary condition that the approaching wave must be planar at large distances and traveling in the direction of the positive z-axis (see Eq. (4.11-1) and Fig. 4-10). The form of the approaching wave is $\exp\left(iu_{\mathrm{L}}\cos\theta_{\mathrm{L}}\right)$. This is referenced to the phase on the line $\theta = \pi/2$.

For the case of the GPS satellite at a finite distance, we have to account for the arrival of a spherical wave, with its center at the transmitting GPS satellite instead of a planar or collimated wave. Referring to Fig. A-3 in Appendix A, this spherical wave is given by $\exp(iu_{\mathrm{LG}})/u_{\mathrm{LG}}$. In this case, the asymptotic form for $a_l(nu)$ is more complicated than that given in Eq. (5.5-3a) because it must correspond to the spectral component of the spherical waveform, which explicitly includes the location of the transmitter. From Section 5.3, where Bauer's identity is derived from the multipole expansion for a spherical wave, one can work out the correct asymptotic form for the spherical case. It is given by[3]

$$a_l(nu) \to A i^{l-1} \frac{2l+1}{l(l+1)} \left(i^{l+1} \frac{\xi_l^+(u_{\mathrm{G}})}{u_{\mathrm{G}}} \right) \tag{5.5-3b}$$

---

[3] Equation (5.5-3b) follows from the homogeneous case, $n \equiv 1$, by first noting that

$$\frac{i}{u_{\mathrm{L}}} \frac{\partial}{\partial \theta_{\mathrm{L}}} \left(\frac{\exp[iu_{\mathrm{LG}}]}{u_{\mathrm{LG}}} \right) = \left(\frac{\exp[iu_{\mathrm{LG}}]}{u_{\mathrm{LG}}} \right) \left(1 + \frac{i}{u_{\mathrm{LG}}} \right) \sin(\theta_{\mathrm{L}} - \chi_{\mathrm{G}})$$

where χ_{G} is the deflection angle of the straight line between the transmitting GPS satellite and the LEO (see Fig. A-3). The GPS satellite is located at $(r_{\mathrm{G}}, \theta_{\mathrm{G}})$, but always in the direction $\theta_{\mathrm{G}} = \pi$. The radial component of the electric field at the LEO from the spherical wave centered at $(r_{\mathrm{G}}, \theta_{\mathrm{G}})$ is given by

$$E_r(r_{\mathrm{L}}, \theta_{\mathrm{L}}) = \left(\frac{\exp[iu_{\mathrm{LG}}]}{u_{\mathrm{LG}}} \right) \sin(\theta_{\mathrm{L}} - \chi_{\mathrm{G}})$$

Dropping the i/u_{LG} term, it follows that

$$E_r(r_{\mathrm{L}}, \theta_{\mathrm{L}}) \doteq \frac{i}{u_{\mathrm{L}}} \frac{\partial}{\partial \theta} \left(\frac{\exp[iu_{\mathrm{LG}}]}{u_{\mathrm{LG}}} \right) = \sum_{l=0}^{\infty} (-1)^l (2l+1) \frac{\psi_l(u_{\mathrm{L}})}{u_{\mathrm{L}}^2} \frac{\xi_l^+(u_{\mathrm{G}})}{u_{\mathrm{G}}} \mathrm{P}_l^1(\cos\theta_{\mathrm{L}})$$

The right-hand side (RHS) of this equation comes from the multipole expansion for $(\exp[iu_{\mathrm{LG}}]/u_{\mathrm{LG}})$ given in Eq. (5.3-2). Equating this series form for $E_r(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ to the form obtained from the corresponding vector calculus operations on the trial scalar potential series [see Eqs. (5.5-7) and (5.5-8)] yields the asymptotic form for the spectral coefficients given Eq. (5.5-3b). Getting the coefficients for one component of the field, $E_r(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ in this case, is sufficient.

Here the phase in this asymptotic form is now referenced to the position r_G of the transmitting GPS satellite.[4] The amplitude A is a constant. For example, if we renormalize the amplitude by setting $A = u_{LG}$, then in the limit as $r_G, r_{LG} \to \infty$, the asymptotic form for $a_l(nu)$ in Eq. (5.5-3b), but referenced to the $\theta = \pi/2$ line, approaches the form given in Eq. (5.5-3a) for the collimated wave. In any case, we will assume the collimated form in Eq. (5.5-3a) subsequently. The correction for the case of an incident spherical wave appears straightforward, and it is noted in Section 5.10 and Chapter 6.

To develop a functional form for $a_l(nu)$, we first will obtain the change in $a_l(nu)$ that results from a change in the index of refraction across a spherical boundary, which is embedded in an *otherwise homogeneous* medium and located at $r = r_o$. By applying the continuity conditions from Maxwell's equations, the spectral coefficients for the transmitted and reflected waves are expressed in terms of the spectral coefficients of the incident wave at the boundary and the change in refractivity. After obtaining the changes in the spectral coefficients that apply across a boundary, we will use a limiting procedure to obtain a continuous version for these spectral coefficients.

The change in $a_l(nu)$ obtained in this manner is characterized by a first-order differential equation. On the other hand, Maxwell's equations comprise a second-order system for this essentially two-dimensional problem. (See Section 4.11.) Therefore, this approach involves an approximation, the accuracy of which we will establish. We saw in the Cartesian case discussed in Sections 4.8 and 4.9 that this approximation works well for points sufficiently distant from a turning point and when thin atmosphere conditions apply. The same conclusions hold here, although the concept of a turning point in a wave theory approach has to be expressed in terms of both the radial coordinate ρ and the spectral number l.

Chapter 3, Sections 3.3 and 3.5, and also Chapter 4, Section 4.6 for the Cartesian case, discuss the formalism for treating standing electromagnetic waves in terms of a spectral composition of incoming and outgoing waves. In Chapter 3, the spherical Bessel function was bifurcated into the spherical

---

[4] We can use the asymptotic form

$$i^{l+1}\xi_l^+(u_G) \to (u_G^2/(u_G^2 - \nu^2))^{1/4} \exp[i\left(\sqrt{u_G^2 - \nu^2} + \nu \sin^{-1}(\nu/u_G)\right)], \quad \nu = l + 1/2$$

in Eq. (5.5-3b) because u_G will be very much larger than the range of spectral numbers yielding stationary values for the spectral series. If the phase terms here are added to the spectral density function for the phase delay through the atmosphere given from the collimated case, we have the correct form for the phase for the case where the incident wave is spherical. See Section 5.10, Eq. (5.10-12). The term $(u_G^2/(u_G^2 - \nu^2))^{1/4}$ is related to the reduced limb distance used to convert the geometry with u_G finite to an equivalent geometry with u_G infinite.

Hankel function of the first kind to represent outgoing waves, and into its equally weighted complex conjugate, the spherical Hankel function of the second kind, to represent incoming waves. Specifically, the spherical Hankel functions of the first (+) and second (−) kinds, $\xi_l^{\pm}$, are defined by $\xi_l^{\pm} = \psi_l \pm i\chi_l$, where $\psi_l(x) = (\pi x/2)^{1/2} J_{l+1/2}(x)$ and $\chi_l(x) = (\pi x/2)^{1/2} Y_{l+1/2}(x)$, where $J_l(x)$ and $Y_l(x)$ are the integer Bessel functions of the first and second kinds, respectively. Using the asymptotic forms for the Bessel functions applicable when $x >> l$, one can readily show that ξ_l^+ assumes the form that describes an outgoing spherical wave, and that ξ_l^- describes an incoming spherical wave. In a homogeneous medium, outgoing waves interior to the scattering boundary are generated from incoming waves that reflect around the origin, which the scattering coefficients b_l show as $r \to 0$. This formalism was necessary to treat internal reflections at the boundary of the scattering sphere and to isolate the scattering coefficients for an emerging wave that has undergone a specific number of internal reflections.

We adopt the same formalism here. Thus, the electric field at any point will be treated as a spectral composition of radial incoming and radial outgoing wavelets, which are combined in a weighted summation over all spectral numbers. They also are combined in such a way as to eliminate the singularity at the origin arising from the Bessel function of the second kind.

5.5.1 Incoming Waves

Let us first consider an incoming incident wave. Here the scalar potentials [see Eq. (5.5-1)] that generate $\boldsymbol{E}^{(i)}$ and $\boldsymbol{H}^{(i)}$ are given by

$$\left.\begin{aligned} {}^{\mathrm{e}}\Pi^{(i)} &= \frac{E_o}{2n_1 k} \sum_{l=1}^{\infty} {}^{\mathrm{e}}a_l^{(i)} \frac{\xi_l^-(n_1 u)}{n_1 u} \mathrm{P}_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\Pi^{(i)} &= \frac{H_o}{2n_1 k} \sum_{l=1}^{\infty} {}^{\mathrm{m}}a_l^{(i)} \frac{\xi_l^-(n_1 u)}{n_1 u} \mathrm{P}_l^1(\cos\theta)\sin\phi \end{aligned}\right\} \tag{5.5-4}$$

Here $u = kr$ is the radial coordinate expressed in phase units. The scattering boundary is located at u_o; ε_1 and μ_1 are constants that define the index of refraction [see Eq. (5.2-4)] in the homogeneous medium on the incident side of the boundary where $u \geq u_o$; $a_l^{(i)}$ is the spectral coefficient for the incoming incident wave. Because $\boldsymbol{E}^{(i)}$ and $\boldsymbol{H}^{(i)}$ are the fields for an incoming wave at the boundary, we must use the spherical Hankel functions of the second kind,

$\xi_l^-/2$, for the radial function instead of ψ_l for determining the spectral coefficients at the boundary.[5]

Similarly, the scalar potentials for the scattered or reflected wave are given in terms of the scattering coefficients b_l by

$$\left.\begin{aligned} {}^{\mathrm{e}}\Pi^{(S)} &= \frac{E_o}{n_1 k}\sum_{l=1}^{\infty} {}^{\mathrm{e}}b_l \frac{\xi_l^+(n_1 u)}{n_1 u}\mathrm{P}_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\Pi^{(S)} &= \frac{H_o}{n_1 k}\sum_{l=1}^{\infty} {}^{\mathrm{m}}b_l \frac{\xi_l^+(n_1 u)}{n_1 u}\mathrm{P}_l^1(\cos\theta)\sin\phi \end{aligned}\right\} \tag{5.5-5}$$

Because the scattered wave is outgoing, we must use the spherical Hankel functions of the first kind, ξ_l^+, in this representation in order to match the asymptotic boundary condition as $r \to \infty$, which requires a spherical wave front from a scattering surface (and which the ξ_l^+ function indeed provides in its asymptotic form for large r). Finally, the scalar potentials for the transmitted wave, which is incoming, are given by

$$\left.\begin{aligned} {}^{\mathrm{e}}\Pi^{(T)} &= \frac{E_o}{2n_2 k}\sum_{l=1}^{\infty} {}^{\mathrm{e}}a_l^T \frac{\xi_l^-(n_2 u)}{n_2 u}\mathrm{P}_l^1(\cos\theta)\cos\phi \\ {}^{\mathrm{m}}\Pi^{(T)} &= \frac{H_o}{2n_2 k}\sum_{l=1}^{\infty} {}^{\mathrm{m}}a_l^T \frac{\xi_l^-(n_2 u)}{n_2 u}\mathrm{P}_l^1(\cos\theta)\sin\phi \end{aligned}\right\} \tag{5.5-6}$$

Here $a_l^{(T)}$ is the spectral coefficient for the wave transmitted across the boundary located at u_o; ε_2 and μ_2 are constants that define the index of refraction on the transmitted side of the boundary where $u \le u_o$.

To obtain the continuity conditions, consider first the electromagnetic field generated by the scalar potential ${}^{\mathrm{e}}\Pi(r,\theta,\phi)$, which generates the TM wave. From Eq. (5.2-7) and using the identity $\nabla\times\nabla\times\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}$, one obtains

---

[5] Recall that $\psi_l = (\xi_l^+ + \xi_l^-)/2$. If we did use ψ_l in the scalar potential series for the incoming incident wave, we would find upon applying the continuity conditions at the boundary that the scattering coefficients $b_l(nu)$ would carry an extra "–1" term that would exactly cancel the $\xi_l^+/2$ part of ψ_l, effectively leaving only the $\xi_l^-/2$ part to represent the entire field, incident plus scattered. For this case where ψ_l is used in the incident series, as $(n_2 - n_1)u \to 0$, $b_l(nu) \to -a_l^{(i)}(nu)/2$.

$$\left.\begin{aligned} \boldsymbol{E}_{\mathrm{TM}} &= (2+\boldsymbol{r}\cdot\nabla)\nabla\left({}^{\mathrm{e}}\Pi\right)-\boldsymbol{r}\nabla^2\left({}^{\mathrm{e}}\Pi\right) \\ \boldsymbol{H}_{\mathrm{TM}} &= -ik\varepsilon\nabla\left({}^{\mathrm{e}}\Pi\right)\times\boldsymbol{r} \end{aligned}\right\} \tag{5.5-7}$$

For example, using Eq. (5.5-4) for the incident wave (and using the differential equation for the spherical Bessel function, $d^2\xi_l/d\rho^2+\left(1-l(l+1)/\rho^2\right)\xi_l=0$, for E_r), the field components in Eq. (5.5-7) become

$$\left.\begin{aligned} \left[E_r^{(i)}\right]_{\mathrm{TM}} &= \left(\frac{\partial^2}{\partial r^2}\left({}^{\mathrm{e}}\Pi^{(i)}r\right)+n_1^2k^2\,{}^{\mathrm{e}}\Pi^{(i)}r\right) \\ &= \frac{E_o^{(i)}}{n_1u}\sum_{l=1}^{\infty}l(l+1)\,{}^{\mathrm{e}}a_l^{(i)}\,\frac{\xi_l^{-}\left(n_1u\right)}{2n_1u}\,\mathrm{P}_l^1\cos\phi \\ \left[E_\theta^{(i)}\right]_{\mathrm{TM}} &= \frac{1}{r}\frac{\partial^2}{\partial r\partial\theta}\left({}^{\mathrm{e}}\Pi^{(i)}r\right)=E_o^{(i)}\sum_{l=1}^{\infty}{}^{\mathrm{e}}a_l^{(i)}\,\frac{\xi_l'^{-}\left(n_1u\right)}{2n_1u}\,\frac{d\,\mathrm{P}_l^1}{d\theta}\cos\phi \\ \left[E_\phi^{(i)}\right]_{\mathrm{TM}} &= \frac{1}{r}\frac{\partial^2}{\partial r\partial\phi}\left({}^{\mathrm{e}}\Pi^{(i)}r\right)=-E_o^{(i)}\sum_{l=1}^{\infty}{}^{\mathrm{e}}a_l^{(i)}\,\frac{\xi_l'^{-}\left(n_1u\right)}{2n_1u}\,\mathrm{P}_l^1\sin\phi \\ \left[H_r^{(i)}\right]_{\mathrm{TM}} &\equiv 0 \\ \left[H_\theta^{(i)}\right]_{\mathrm{TM}} &= \frac{k\varepsilon}{ir}\frac{\partial}{\partial\phi}\left({}^{\mathrm{e}}\Pi^{(i)}r\right)=\frac{iH_o^{(i)}}{\sin\theta}\sum_{l=1}^{\infty}{}^{\mathrm{e}}a_l^{(i)}\,\frac{\xi_l^{-}\left(n_1u\right)}{2n_1u}\,\mathrm{P}_l^1\sin\phi \\ \left[H_\phi^{(i)}\right]_{\mathrm{TM}} &= \frac{ik\varepsilon}{r}\frac{\partial}{\partial\theta}\left({}^{\mathrm{e}}\Pi^{(i)}r\right)=iH_o^{(i)}\sum_{l=1}^{\infty}{}^{\mathrm{e}}a_l^{(i)}\,\frac{\xi_l^{-}\left(n_1u\right)}{2n_1u}\,\frac{d\,\mathrm{P}_l^1}{d\theta}\cos\phi \\ E_o^{(i)}\sqrt{\varepsilon_1} &= H_o^{(i)}\sqrt{\mu_1},\quad \mathrm{P}_l^1=\mathrm{P}_l^1(\cos\theta) \end{aligned}\right\} \tag{5.5-8}$$

We can write a set of expressions of a similar form for the scattered fields, $\boldsymbol{E}^{(S)}$ and $\boldsymbol{H}^{(S)}$, and for the transmitted fields, $\boldsymbol{E}^{(T)}$ and $\boldsymbol{H}^{(T)}$. Using the symmetry properties of the electromagnetic field discussed in Section 3.2, one also can readily develop a set of expressions from ${}^{\mathrm{m}}\Pi$ for the TE wave. The complete field is given by the sum of these TM and TE expressions.

To obtain the required relationships between the spectral coefficients, we use the continuity conditions from Maxwell's equations that the various field components must satisfy. Across a boundary with neither surface charges nor surface currents, Maxwell's equations require the components of the electromagnetic field to satisfy the following continuity conditions:

$$\left.\begin{aligned} \boldsymbol{E}_{\text{tang}}^{(i)} + \boldsymbol{E}_{\text{tang}}^{(S)} &= \boldsymbol{E}_{\text{tang}}^{(T)}, \quad \varepsilon_1\left(\boldsymbol{E}_r^{(i)} + \boldsymbol{E}_r^{(S)}\right) = \varepsilon_2 \boldsymbol{E}_r^{(T)} \\ \boldsymbol{H}_{\text{tang}}^{(i)} + \boldsymbol{H}_{\text{tang}}^{(S)} &= \boldsymbol{H}_{\text{tang}}^{(T)}, \quad \mu_1\left(\boldsymbol{H}_r^{(i)} + \boldsymbol{H}_r^{(S)}\right) = \mu_2 \boldsymbol{H}_r^{(T)} \end{aligned}\right\} \tag{5.5-9a}$$

Here ε_1 and μ_1 apply to the incident side of the boundary; ε_2 and μ_2 apply to the transmitted side.

We apply these continuity conditions to the vector fields generated from the scalar potentials in Eqs. (5.5-4) through (5.5-6) for a boundary located at $r = r_o$. The electromagnetic field for the incident TM wave is shown in Eq. (5.5-8), but because they are all similar, we forego writing the other five sets for the scattered and transmitted TM waves and for all TE modes. Applying the continuity conditions in Eq. (5.5-9a) to these waves at all applicable points on the boundary of the sphere located at $r = r_o$, we obtain an equivalent set of continuity conditions that involve only the individual spectral coefficients and their Hankel functions. These conditions written in matrix form become

$$\left.\begin{aligned} \begin{pmatrix} \dfrac{\xi_l^-(n_2u_o)}{\mu_2} & \dfrac{-\xi_l^+(n_1u_o)}{\mu_1} \\ \dfrac{\xi_l^{-\prime}(n_2u_o)}{n_2} & \dfrac{-\xi_l^{+\prime}(n_1u_o)}{n_1} \end{pmatrix} \begin{pmatrix} {}^{\mathrm{e}}a_l^{(T)} \\ 2^{\mathrm{e}}b_l \end{pmatrix} &= \begin{pmatrix} \dfrac{\xi_l^-(n_1u_o)}{\mu_1} \\ \dfrac{\xi_l^{-\prime}(n_1u_o)}{n_1} \end{pmatrix} {}^{\mathrm{e}}a_l^{(i)} \\ \begin{pmatrix} \dfrac{\xi_l^-(n_2u_o)}{n_2} & \dfrac{-\xi_l^+(n_1u_o)}{n_1} \\ \dfrac{\xi_l^{-\prime}(n_2u_o)}{\mu_2} & \dfrac{-\xi_l^{+\prime}(n_1u_o)}{\mu_1} \end{pmatrix} \begin{pmatrix} {}^{\mathrm{m}}a_l^{(T)} \\ 2^{\mathrm{m}}b_l \end{pmatrix} &= \begin{pmatrix} \dfrac{\xi_l^-(n_1u_o)}{n_1} \\ \dfrac{\xi_l^{-\prime}(n_1u_o)}{\mu_1} \end{pmatrix} {}^{\mathrm{m}}a_l^{(i)} \end{aligned}\right\}, \quad l = 1, 2, \cdots \tag{5.5-9b}$$

Solving the linear system of equations in Eq. (5.5-9b) for the transmission and scattering coefficients in terms of the incident coefficients, we obtain

$$\left.\begin{aligned} {}^{\mathrm{e}}a_l^{(T)} &= \left[\frac{-2i}{n_1\mu_1\,{}^{\mathrm{e}}\mathcal{W}_l}\right] {}^{\mathrm{e}}a_l^{(i)}, \quad 2^{\mathrm{e}}b_l = -\left[\frac{{}^{\mathrm{e}}\mathcal{W}_l^-}{{}^{\mathrm{e}}\mathcal{W}_l}\right] {}^{\mathrm{e}}a_l^{(i)} \\ {}^{\mathrm{m}}a_l^{(T)} &= \left[\frac{-2i}{n_1\mu_1\,{}^{\mathrm{m}}\mathcal{W}_l}\right] {}^{\mathrm{m}}a_l^{(i)}, \; 2^{\mathrm{m}}b_l = -\left[\frac{{}^{\mathrm{m}}\mathcal{W}_l^-}{{}^{\mathrm{m}}\mathcal{W}_l}\right] {}^{\mathrm{m}}a_l^{(i)} \end{aligned}\right\}, \quad l = 1, 2, \cdots \tag{5.5-10}$$

where

$$\left.\begin{aligned}
{}^{\mathrm{e}}\mathcal{W}_l &= \frac{\xi_l^+(n_1u_o)\xi_l^{-\prime}(n_2u_o)}{n_2\mu_1} - \frac{\xi_l^{+\prime}(n_1u_o)\xi_l^-(n_2u_o)}{n_1\mu_2} \\
{}^{\mathrm{e}}\mathcal{W}_l^{\pm} &= \frac{\xi_l^{\pm}(n_1u_o)\xi_l^{\pm\prime}(n_2u_o)}{n_2\mu_1} - \frac{\xi_l^{\pm\prime}(n_1u_o)\xi_l^{\pm}(n_2u_o)}{n_1\mu_2} \\
{}^{\mathrm{m}}\mathcal{W}_l &= \frac{\xi_l^+(n_1u_o)\xi_l^{-\prime}(n_2u_o)}{n_1\mu_2} - \frac{\xi_l^{+\prime}(n_1u_o)\xi_l^-(n_2u_o)}{n_2\mu_1} \\
{}^{\mathrm{m}}\mathcal{W}_l^{\pm} &= \frac{\xi_l^{\pm}(n_1u_o)\xi_l^{\pm\prime}(n_2u_o)}{n_1\mu_2} - \frac{\xi_l^{\pm\prime}(n_1u_o)\xi_l^{\pm}(n_2u_o)}{n_2\mu_1}
\end{aligned}\right\}, \quad l = 1,2,\cdots \qquad (5.5\text{-}11)$$

The Wronskian of the spherical Hankel functions,

$$\mathscr{W}\left[\xi_l^+(z),\xi_l^-(z)\right] = \xi_l^+\xi_l^{-\prime} - \xi_l^{+\prime}\xi_l^- = -2i \qquad (5.5\text{-}12)$$

has been used in Eq. (5.5-9b) to obtain the transmission coefficients in Eq. (5.5-10).

The "electric" coefficients $({}^{\mathrm{e}}a_l, {}^{\mathrm{e}}b_l)$ and the "magnetic" coefficients $({}^{\mathrm{m}}a_l, {}^{\mathrm{m}}b_l)$ differ from their counterparts by a small quantity of the order $N = n-1$. Because we have assumed a thin atmosphere, $N(r) << 1$; we will ignore this difference herewith and, in the interest of simplifying the notation, we will suppress the superscripts "e" and "m" on the scattering coefficients and retain only the electric coefficients in the following. These small differences can readily be reconstituted to obtain the scattered wave from the vector calculus operations on both the electric and magnetic scalar potentials. Also, for the case where $\boldsymbol{E}$ lies in the plane $\phi = 0$, one can show that for large spectral numbers the magnetic coefficients provide a negligible contribution to the field. This follows from noting that the magnetic coefficients involve $\mathrm{P}_l^1(\cos\theta)$, whereas the electric coefficients involve $d\mathrm{P}_l^1/d\theta$. However, $\mathrm{P}_l^1/(d\mathrm{P}_l^1/d\theta) \sim l^{-1} << 1$.

We note that the GPS signals are principally right-hand circular polarized; therefore, to study polarization effects from the refracting sphere, we would need to retain the cross-plane polarization ($\phi = \pi/2$) scattering terms also, which are appropriately offset in phase to secure the proper elliptical or circular polarization. However, for $N << 1$, the scattering for the two linear polarization modes differs by an amount of the order N. Also, because of the previously mentioned relativistic covariance of the electrodynamics equations, we can exploit that symmetry to convert the solution for $\boldsymbol{H}^{(S)}$ for the in-plane

polarized case discussed here directly into a solution for $\boldsymbol{E}^{(S)}$ for the cross-plane polarization case.

For outgoing waves—for example, for waves that have passed through the scattering sphere or, in a geometric optics context, rays that have passed their point of tangency with an arbitrary spherical boundary at radius $r = r_*$—one would obtain a system of transfer equations analogous to those given in Eq. (5.5-10). The only difference is that the scalar potential series for the incident and transmitted waves would each carry the $\xi_l^+ / 2$ functions instead of the $\xi_l^- / 2$ functions because they are outgoing. Also, the scalar potential series for the waves reflected from the inner side of the boundary would carry the $\xi_l^- / 2$ functions because they are incoming after being reflected.

5.5.2 Evaluating the Spectral Coefficients in a Stratified Medium

We now set $\mu(r) \equiv 1$ in the following discussion, which further simplifies the notation, albeit at the price of losing the symmetries in Eqs. (5.5-10) and (5.5-11).

Next, we treat the continuously varying refractivity in the medium as a series of concentric shells. Within each spherical shell the refractivity is a constant, but it changes discontinuously across the boundary of each shell. So, the refractivity varies in the radial direction in a stepwise manner. This is the thin-film model, or one version of the so-called onionskin model. Across each boundary, the transition equations for the spectral coefficients in Eq. (5.5-10) apply. After obtaining these spectral coefficients across the boundary of each shell, we will let the number of shells grow infinite while requiring their individual widths to become infinitesimal in such a way that the ensemble spans the appropriate physical space or range.

At the boundary located at $u_o = kr_o$, we let $n_1 = n - \Delta n / 2$ and $n_2 = n + \Delta n / 2$, where Δn is sufficiently small that $u\Delta n$ can be considered as an infinitesimal. Expanding $n_1 n_2 \mathcal{W}_l$ and $n_1 n_2 \mathcal{W}_l^{\pm}$ in powers of $u\Delta n$, we obtain

$$\left.\begin{aligned} n\mathcal{W}_l &= -2i - \frac{\Delta n}{2n}\Big[nu\Big(2\xi_l^{+\prime}\xi_l^{-\prime} - \xi_l^-\xi_l^{+\prime\prime} - \xi_l^+\xi_l^{-\prime\prime}\Big) \\ &\qquad + \xi_l^{+\prime}\xi_l^- + \xi_l^+\xi_l^{-\prime}\Big]_{nu_o} + \mathrm{O}\big[(u\Delta n)^2\big] \\ n\mathcal{W}_l^{\pm} &= -\frac{\Delta n}{n}\Big[nu\Big(\xi_l^{\pm\prime}\xi_l^{\pm\prime} - \xi_l^{\pm}\xi_l^{\pm\prime\prime}\Big) + \xi_l^{\pm\prime}\xi_l^{\pm}\Big]_{nu_o} + \mathrm{O}\big[(u\Delta n)^2\big] \end{aligned}\right\} \qquad (5.5\text{-}13)$$

It follows that as $\Delta n \to 0$, $\mathcal{W}_l^{\pm} \to 0$, but $n\mathcal{W}_l \to -2i$. From Eq. (5.5-11), it follows that $b_l \to 0$ and that $a_l^{(T)} \to a_l^{(i)}$ when $\Delta n \to 0$.

For a series of concatenated shells, multiple internal reflections should be considered. For example, outward-reflected rays from inner shell boundaries again will be reflected inward at the boundary of interest. We already have discussed this in Section 4.8 for Cartesian layers, and Fig. 4-8 in that section applies here as well. Specifically, we can use the discussion in Section 4.12 to transform our spherical geometry here into an equivalent Cartesian-stratified geometry involving Airy layers. By this means, conclusions drawn from the Cartesian case can be applied here. In Section 4.8, we showed that the ensemble of doubly reflected rays that adds to the incident wave each involves a factor of the order of Δn^2 (here Δn is the average change in index of refraction from layer to layer). Moreover, the phase of these secondary rays (at the right-hand boundary of the *j*th layer in Fig. 4-8) will be distributed randomly when the span Δr of the layers is such that $\Delta r >> \lambda$. It can be shown by vector summing up the contributions from all of these reflected rays with a second reflection from the left-hand boundary of the *j*th layer that the ratio of their combined contributions to the main ray contribution is given by $n'\lambda$, which is negligible for a thin atmosphere. Therefore, in calculating the spectral coefficients for the transmitted wave through a transparent medium, we can neglect secondary and higher-order reflections in our shell model when thin-atmosphere conditions apply and provided that we avoid turning points.

The incident field at the j + 1st boundary can be considered as the product of the transmission coefficients from the previous j layers. If we then expand that product and retain only the first-order terms, we can obtain a first-order differential equation for the spectral coefficients. The range of validity of this linear truncation is essentially the same as that found for the truncation of the characteristic matrix to linear terms given in Section 4.4. There we found for a thin atmosphere that the accuracy of this truncation was satisfactory provided that we stay clear of turning points.

Let us define $(a_l^-)_j$ to be the lth spectral coefficient of an incoming transmitted wave for the *j*th layer. The superscript "–" on a_l^- denotes an incoming wave. We drop herewith the superscripts "*i*" and "*T*". Then, using Eqs. (5.5-10) and (5.5-13), it follows that

$$\left(a_l^-\right)_{j+1} \doteq \left(a_l^-\right)_j \left[\frac{1+\Delta n/2n}{1-ig_l\Delta n/n}\right]_j \tag{5.5-14}$$

where $g_l(\rho_j)$ is a function of the spherical Hankel functions obtained from Eq. (5.5-13), which is defined in Eq. (5.5-19) and will be discussed shortly. Here we define $\rho = un(r) = krn(r)$. For a series of layers, it follows from Eq. (5.5-14) that

$$\left(a_l^-\right)_{k+1} \doteq \left(a_l^-\right)_1 \prod_{j=1}^{k} \left(\frac{1+\frac{\Delta n_j}{2n_j}}{1-ig_l(\rho_j)\frac{\Delta n_j}{n_j}} \right) \tag{5.5-15}$$

To evaluate Eq. (5.5-15), we note that $\log\left[\prod(*)_j\right] = \sum\left[\log(*)\right]_j$. When $g_l \Delta n_j << 1$, we can expand $\log\left[1 - ig_l(\rho_j)\left(\Delta n_j / n_j\right)\right]$, retaining only first-order terms in Δn_j. Thus, Eq. (5.5-15) becomes

$$\log\left(\frac{\left(a_l^-\right)_{k+1}}{\left(a_l^-\right)_1} \right) \doteq \sum_{j=1}^{k} \left(\left(\frac{1}{2} + ig_l(\rho_j) \right) \frac{\Delta n_j}{n_j} \right) \tag{5.5-16}$$

We set[6] $\Delta n = (dn/d\rho)\Delta\rho$. Also, we define $\Delta n_j = n_{j+1} - n_j$ to be the change in the index of refraction across the *j*th boundary (Fig. 4-8), and we define $\Delta\rho_j = \rho_{j+1} - \rho_j$ to be the optical thickness of the *j*th layer. From Eq. (5.5-16), it follows that in the limit as $\Delta\rho \to 0$, we obtain

$$\frac{1}{a_l^-}\frac{da_l^-}{d\rho} = \left(\frac{1}{2} + ig_l(\rho) \right) \frac{d\log n}{d\rho} \tag{5.5-17}$$

Here $g_l(\rho)$ is defined by

$$g_l(\rho) = \left[\frac{\rho}{2}\left(\xi_l^{\pm\prime}\xi_l^{\mp\prime} - \xi_l^{\pm}\xi_l^{\mp\prime\prime} \right) + \frac{1}{4}\left(\xi_l^{+}\xi_l^{-\prime} + \xi_l^{+\prime}\xi_l^{-} \right) \right]_{\rho = un} \tag{5.5-18}$$

Bessel's equation in Eq. (5.3-3) has implicitly been used in Eq. (5.5-18). The enormity of $\rho \sim 10^8$ allows us to ignore the second term, $(\xi_l^+\xi_l^{-\prime} + \xi_l^{+\prime}\xi_l^-)/4$. Using Bessel's equation to replace $\xi_l^{\pm\prime\prime}$ and dropping the relatively small term in Eq. (5.5-18), one obtains

---

[6] Note that $d\log n/d\rho = (n + udn/du)^{-1} d\log n/du$. Also, $\rho(d\log n/d\rho) = u(d\log n/du)/(1 + u(d\log n/du))$. The quantity $u \mid d\log n/du \mid$ is the ratio of the radius of curvature (*r*) of the spherical boundary to the local radius of curvature of the ray ($n/\mid dn/du \mid$). It is the parameter β defined in Chapter 2, Eq. (2.2-9), which is small for a thin atmosphere (for dry air in the Earth's atmosphere at sea level this ratio is about 0.2). In a super-refracting medium, occasionally caused by a water vapor layer in the lower troposphere, $d\rho/du < 0$. Across a boundary, $d\rho/du = 0$, which requires reverting to the variable u in Eq. (5.5-17).

$$g_l(\rho) \doteq \frac{\rho}{2}\left(\xi_l^{+\prime}\xi_l^{-\prime} + \left(1 - \frac{l(l+1)}{\rho^2}\right)\xi_l^{+}\xi_l^{-} \right) \tag{5.5-19}$$

Figure 5-2 shows $g_l(\rho)$ versus $\hat{y}$, including its asymptotic forms. Here $\hat{y}$ is the argument of the Airy functions. The relationship between $\hat{y}$, l, and $\rho = un$ was discussed earlier in Section 5.4, Eqs. (5.4-2) and (5.4-3). It suffices here to note that $\nu = l + 1/2 \doteq \rho + \hat{y}(\rho/2)^{1/3}$ to very high accuracy when ρ is large and $\hat{y}$ is relatively small. For $\hat{y}$ values greater than about +2, $g_l(\rho)$ is dominated by the spherical Bessel function of the second kind, and it breaks sharply to very large negative values.

The derivation for $da_l^- / d\rho$ fails for this regime, $\hat{y} > 0$, because the basic assumption that $g_l \Delta n_j << 1$ in Eq. (5.5-16) is invalid when $g_l(\rho) \to \infty$ for increasing $\nu > \rho$. In fact, the correct form for $g_l(\rho)$ rapidly approaches zero for $\nu > \rho$, rather than blowing up, as the form for $g_l(\rho)$ given in Eq. (5.5-19) does. The modified Mie scattering derivation that we have used did not account for curvature terms, and it assumes that $g_l(\rho)\Delta n$ can be made a small quantity, which is not valid below a turning point. We return to this issue in Section 5.7, after a discussion of asymptotic forms. There we present one method for asymptotic matching of the $g_l(\rho)$ function given in Eq. (5.5-19) with a version that does hold for $\hat{y} > 0$.

In general the initial condition for a_l^- in Eq. (5.5-17) depends on the boundary conditions for the electromagnetic field. In a geometric optics context, the initial condition for a_l^- is a ray-specific quantity; that is, it depends

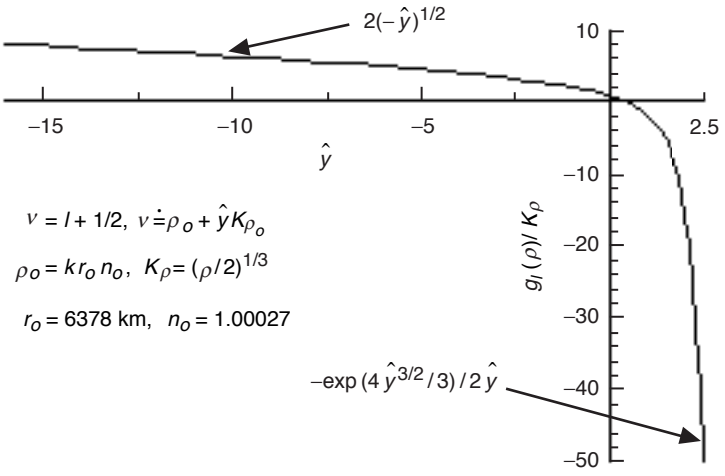

Fig. 5-2. Phase rate term $g_l(\rho)$ for spectral numbers near ρ.

at least in part on the impact parameter of the ray (or cophasal normal path) associated with the wave as it propagates through the medium. Therefore, the constant of integration obtained from integrating Eq. (5.5-17) will depend on ray-specific boundary conditions. However, in the special case where the approaching rays are collimated before encountering the medium, they all have the same asymptotic boundary condition as $u \to \infty$; in this case, the constant of integration will be invariant with impact parameter. For departing waves, this symmetry is spoiled[7] by the intervening refracting medium, and the asymptotic boundary conditions as $u \to \infty$ will vary with the impact parameter of the approaching ray.

We define the functional $G[\rho, \nu]$ by

$$G[\rho,\nu] = \int_{\rho}^{\infty} \left(\frac{d \log n}{d\rho'} \right) g_l(\rho') d\rho' \tag{5.5-20}$$

For convenience in this and in the following sections, we use the spectral number l and the parameter $\nu = l + 1/2$ interchangeably. The distinction between them is inconsequential because of the enormity of their values in their stationary phase neighborhoods. It is understood here that the form for $g_l(\rho)$ given in Eq. (5.5-19) must be modified so that $g_l(\rho) \to 0$ for increasing $\nu > \rho$.

Using the asymptotic boundary condition for $a_l^-(\rho)$ given in Eq. (5.5-3a) and noting that $\rho \to u = kr$ asymptotically with large r, the solution $a_l^-(\rho)$ can be obtained by integrating Eq. (5.5-17), and it can be written as

$$a_l^-(\rho) = n^{1/2} i^{l-1} \frac{2l+1}{l(l+1)} \exp\left(-iG[\rho,\nu]\right) \tag{5.5-21}$$

Thus, $-G[\rho, \nu]$ is the phase retardation induced by the refractive gradient in the lth spectral component of an incoming wave, which results from traveling through a transparent, spherical symmetric, refracting medium from infinity down to a radial distance r. Initially, as $r \to \infty$, the incoming wave is planar, and its spectral coefficient is given by Eq. (5.5-3a). For a homogeneous medium, $G[\rho, \nu] \equiv 0$.

---

[7] We could, however, form a symmetric problem merely by forcing the electromagnetic wave to be planar along the line $\theta = \pi/2$. The boundary conditions for this case are $a_l^{\pm}|_{\theta=\pi/2} = i^{l-1}(2l+1)/l(l+1)$, and $a_l^{\pm}$ at (ρ, θ) is

$$a_1^{\pm}(\rho,\theta,\rho_*) = i^{1-1}(2l+1)/l/(l+1)(\exp[\mp i(G[\rho_*,\nu] - G[\rho,\nu])]),$$

where $\rho_* = \rho_*(\rho,\theta)$ from Eq. (5.6-3), which is Bouguer's law.

For thin atmospheres, the term $n^{1/2}$ in Eq. (5.5-21) is essentially unity, and it will be ignored in subsequent discussions.

5.5.3 Outgoing Waves

We have a similar expression for a radial outgoing wave. In this case, we let $\Delta a_l^+ = a_l^{(T)} - a_l^{(i)}$, where $a_l^{(i)}$ is the spectral coefficient of spectral number l for the outward traveling wave incident on the inner side of the boundary, and $a_l^{(T)}$ is the coefficient for the outward-directed transmitted wave. The scalar potential series for both of these waves uses the ξ_l^+ functions because they are outgoing waves. Also, in Eq. (5.5-9b) we must change the ξ_l^- functions to ξ_l^+ functions because $a_l^{(i)}$ and $a_l^{(T)}$ are now the spectral coefficients for outgoing waves; similarly, we must change from ξ_l^+ to ξ_l^- for b_l because the reflected wave is incoming. Working through the same boundary conditions applicable to an outgoing wave and applying the same limit procedures that held for the inward case [see Eqs. (5.5-8) through-(5.5-17)], one obtains a differential equation for the spectral coefficients of the outward-directed wave:

$$\frac{1}{a_l^+}\frac{da_l^+}{d\rho} = -i\frac{d\log n}{d\rho}g_l(\rho) \tag{5.5-22}$$

Comparing Eq. (5.5-22) with Eq. (5.5-17) (and dropping the $n^{1/2}$ term), we see that the gradients of a_l^- and a_l^+ have opposite polarities. In other words, the spatial derivative of the spectral coefficient along the radial direction of propagation is invariant to whether the wave is incoming or outgoing. This must be true from a physical consideration: the rate of phase accumulation at a given site should be the same for the radial-traveling incoming and outgoing wavelets.

We see upon integrating Eq. (5.5-22) that a_l^+ will depend on the adopted value of a constant of integration. Let us fix that constant at $r = r_*$. We write a_l^+ in the form $a_l^+[\rho, \rho_*]$ to express this dual dependency; here $\rho_* = kr_* n(\rho_*)$. Integrating Eq. (5.5-22) and using Eq. (5.5-20), we obtain

$$a_l^+[\rho, \rho_*] = a_l^+[\rho_*, \rho_*]\exp\left[-i\left(G[\rho_*, \nu] - G[\rho, \nu]\right)\right] \tag{5.5-23}$$

If we let $r \to \infty$, which would be appropriate when observing the refracted wave from outside the refracting medium, such as the neutral atmosphere observed from a LEO, then $G[\rho, \nu] \to 0$ and one obtains

$$a_l^+[\infty, \rho_*] = a_l^+[\rho_*, \rho_*]\exp\left(-iG[\rho_*, \nu]\right) \tag{5.5-24}$$

The phase retardation incurred by the νth wavelet in traveling outward from r_* to infinity is $-G[\rho,\nu]$, which is the same retardation incurred by the inward traveling wavelet from infinity down to r_*.

The actual value(s) of $a_l^+[\rho_*,\rho_*]$ will depend in part on the physical properties assumed for the refracting and perhaps scattering atmosphere, and also in part on the impact parameter(s) associated with the ray(s). For example, if $dn/dr \equiv 0$ for $r < r_o$, then Eqs. (5.5-17) and (5.5-22) show that both a_l^- and a_l^+ will be constant in that region. They also must be equal there to avoid the Hankel function singularity at the origin. (Recall the definition of the spherical Bessel function of the first kind, $\psi_l = (\xi_l^+ + \xi_l^-)/2$, which is well-behaved at the origin.) It follows in this case that $a_l^+[\rho_o,\rho_o] = a_l^-(\rho_o)$, where $a_l^-(\rho_o)$ is given from Eq. (5.5-21) and it is the applicable spectral coefficient for an incoming wave that was initially planar. At the LEO, in this case we would have

$$a_l^+[\infty,\rho_o] = i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-i2G[\rho_o,\nu]\right) \tag{5.5-25}$$

Thus, $-2G[\rho_o,\nu]$ is the total phase delay incurred by the lth spectral coefficient of an initial plane wave with an impact parameter ρ_o as a result of the wave passing completely through an intervening medium. We will return to this topic in a later section, where specific refracting and scattering models are discussed. We also will discuss later the accuracy of this particular spectral representation in terms of osculating parameters.

5.5.4 Correspondence between Cartesian and Spherical-Stratified Phase Quantities

In Chapter 4, we applied the thin-film concepts to a Cartesian-stratified medium to solve the wave equations expressed in terms of the unitary state transition matrix $\boldsymbol{M}[x_2,x_1]$. Central quantities in that presentation, which are given in Eq. (4.4-13), are the phase accumulation $\mathscr{A}(x,x_o)$ and its rate $\varpi(x)$ (with $\mu \equiv 1$) that results from the profile $n(x)$ in that Cartesian-stratified medium. These are

$$\left.\begin{aligned} \mathscr{A}(x,x_o) &= k\int_{x_o}^{x}\varpi(x')dx', \\ \varpi(x) &= \left(n^2(x)-n_o^2\right)^{1/2},\ \ n_o = n(x_o) \end{aligned}\right\} \tag{5.5-26}$$

Note that $\mathscr{A}(x,x_o)$ provides the total phase accumulation of the wave along the x-axis, perpendicular to the plane of stratification, from the turning point at

x_o up to the altitude at x. It is an implicit function of the refractivity profile and the "angle of incidence" φ of the wave through the value of n_o, which is a constant for a particular wave (generalized Snell's law, $n_o = n\sin\varphi$), analogous to the impact parameter ρ_* for the spherical geometry. Thus, both $\mathscr{A}(x, x_o)$ and $\varpi(x)$ depend on the angle of incidence of the wave. Defining $\rho = kxn(x)$ for the Cartesian-stratified case, it follows from Eq. (5.5-26) that $\mathscr{A}$ may be rewritten in the form

$$\mathscr{A}(\rho, \rho_o) = \int_{\rho_o}^{\rho} \frac{\varpi}{n} d\rho' - \int_{\rho_o}^{\rho} \frac{d\log n}{d\rho'} \frac{\varpi\rho'}{n} d\rho' \tag{5.5-27}$$

The first integral provides the "geometric" phase delay ($\varpi / n = \cos\varphi$), and the second integral provides the additional phase delay resulting from the gradient of the refractivity over the interval ρ_o to ρ. The correspondence between the spectral quantities derived in this section for spherical stratification, $g_l(\rho)$ and $G[\rho, \nu]$, and their counterparts in Cartesian stratification should be clear. It is

$$\left.\begin{aligned} &G[\rho_o, \nu] - G[\rho, \nu] \Leftrightarrow \mathscr{A}(\rho, \rho_o) - \int_{\rho_o}^{\rho} \frac{\varpi}{n} d\rho' \\ &g_l(\rho) \Leftrightarrow \frac{\rho\varpi(\rho, \rho_o)}{n(\rho)} \Rightarrow \rho\cos\varphi, \quad \sin^{-1}\left(\frac{\nu}{\rho}\right) \Leftrightarrow \varphi \end{aligned}\right\} \tag{5.5-28}$$

Note that this correspondence applies only for $\rho > \nu$. Here the angle of incidence φ in the Cartesian frame is related to the spectral number l in the spherical frame through the relation given in Eq. (4.12-8).

5.5.5 Absorption

The modified Mie scattering approach used here lends itself easily to a medium with mild absorption. Here the index of refraction has the form $\hat{n} = n(1 + i\kappa)$, where $n(r)$ is the real component and $n\kappa$ is the imaginary component; κ is the extinction coefficient, and it is real. Because the refracting sphere is so large, κ must be a very small quantity or else the penetrating waves will be completely damped before escaping from the sphere. In any case, it follows from Eq. (5.5-17) that, when $\kappa \neq 0$, $a_l^-(\rho)$ will have an exponentially damping component in addition to a phase delay. In this case, the constant E_o, which is the amplitude of the incident wave, must be treated more carefully to account for the actual absorption through the medium. Also, in the case where the emitting GPS satellite is located at a finite distance away, ρ_{LG}, then E_o must account for the space loss in amplitude that the spherical wave emitted from the GPS satellite incurs in traveling to the LEO.

5.6 More Geometric Optics: Cumulative Bending Angle, Bouguer's Law, and Defocusing

We need a few more concepts from geometric optics for incoming and outgoing waves in order to interpret these wave theory results using the stationary phase technique. Appendix A briefly discusses deriving the ray path in geometric optics from Fermat's principle and the Calculus of Variations. We know that the path integral for the phase delay along the ray from the observed GPS satellite to the LEO, $\int n ds$, is stationary with respect to the path followed by the signal. That is, the actual path provides a stationary value for the phase delay compared with the phase delay that would be obtained by following any neighboring path with the same end points. Here s is path length. If one applies the Calculus of Variations to this phase delay integral, then one obtains Euler's equation, which is a second-order differential equation. This equation provides a necessary condition that the path must satisfy to yield a stationary value for the phase delay path integral. When the path integral is expressed in polar coordinates with r as the independent variable, then $ds = \left(1 + r^2\theta'^2\right)^{1/2} dr$, and Euler's equation becomes

$$\frac{d}{dr}\left(\frac{\partial}{\partial\theta'}\left(n\sqrt{1+r^2\theta'^2}\right)\right) - \frac{\partial}{\partial\theta}\left(n\sqrt{1+r^2\theta'^2}\right) = 0 \qquad (5.6\text{-}1a)$$

Provided that n is a function only of r, this equation may be integrated once to obtain a constant of integration:

$$\frac{nr^2\theta'}{\sqrt{1+r^2\theta'^2}} = n_* r_*, \quad \theta' = \frac{d\theta}{dr}, \quad n_* = n(r_*), \quad \frac{r\theta'}{\sqrt{1+r^2\theta'^2}} = \sin\gamma \qquad (5.6\text{-}1b)$$

Here θ' is related to the slope of the ray at the point (r,θ), and γ is the angle between the radius vector and the tangent vector of the ray. For planar approaching waves, i.e., for the case where the occulted GPS satellite is set infinitely far away in the $\theta = \pi$ direction, then $\gamma = \tilde{\alpha} + \theta$. Here $\tilde{\alpha}$ is the cumulative bending angle up to the point (r,θ) incurred by the ray relative to its original direction $(\theta = 0)$ as an approaching planar wave. In Appendix A, Fig. A-1, a positive value for $\tilde{\alpha}$ corresponds to a clockwise rotation of the tangent vector of the ray relative to the line $\theta = 0$. Along a ray path satisfying Euler's equation, the impact parameter $\rho_* = k n_* r_*$ must be constant when n is not a function of θ. From geometric optics, the differential bending angle $d\tilde{\alpha}$ over an infinitesimal length ds along the ray path expressed in polar coordinates is given by $d\tilde{\alpha} = (dn/dr)\sin\gamma ds$. Upon applying Eq. (5.6-1b) and integrating

$d\tilde{\alpha}$ along the ray path from the GPS satellite (assumed to be at infinity) to an approaching point (r,θ), one obtains

$$\left.\begin{aligned}&\tilde{\alpha}[r,\theta]=\tilde{\alpha}(\rho,\rho_*)=-\rho_*\int_\rho^\infty \frac{d\log n}{d\rho'}\frac{d\rho'}{\sqrt{\rho'^2-\rho_*^2}}\\&\rho=krn(r),\ \theta_*<\theta<\pi\end{aligned}\right\} \tag{5.6-2}$$

Here θ_* marks the angular coordinate of the tangency or turning point for a particular ray with an impact parameter value of ρ_*. See Fig. A-1. We note that $\theta_*=\theta_*(\rho_*)$, and also $\tilde{\alpha}$ at any approaching point (r,θ) may be considered to be a function of $\rho=rn(r)$ and the impact parameter ρ_* for the ray passing through that point. Thus, $\tilde{\alpha}=\tilde{\alpha}(\rho,\rho_*)$ and also $\tilde{\alpha}=\tilde{\alpha}[r,\theta]$. From Eqs. (5.6-1) and (5.6-2), it follows that the impact parameter $\rho_*=kr_*n(r_*)$ is given in terms of (r,θ) and $\tilde{\alpha}$ by

$$\left.\begin{aligned}&\text{Incoming: } \pi>\theta\ge\theta_*:\quad \rho_*=\rho\sin\left[\theta+\tilde{\alpha}(\rho,\rho_*)\right]\\&\text{Outgoing: } \theta_*\ge\theta>0:\quad \rho_*=\rho\sin\left[\theta+2\tilde{\alpha}(\rho_*,\rho_*)-\tilde{\alpha}(\rho,\rho_*)\right]\\&\theta_*=\frac{\pi}{2}-\tilde{\alpha}(\rho_*,\rho_*);\ \ \rho=krn(r)\end{aligned}\right\} \tag{5.6-3}$$

This is a version of Bouguer's law, which has been expressed for both an incoming ray ($\pi>\theta\ge\theta_*$) and for an outgoing ray ($\theta_*\ge\theta>0$). By symmetry, the bending angle for an outgoing ray is $2\tilde{\alpha}(\rho_*,\rho_*)-\tilde{\alpha}(\rho,\rho_*)$, where $\tilde{\alpha}(\rho_*,\rho_*)$ is the cumulative bending up to the turning point. The constant value of ρ_* along a given ray path is the geometric optics analog of the conservation of angular momentum in a classical mechanical system with spherical symmetry in its force field. One can solve Eqs. (5.6-2) and (5.6-3) simultaneously to obtain the values of both ρ_* and $\tilde{\alpha}(\rho,\rho_*)$ for a given position (r,θ) [which may or may not be uniquely determined, depending on the profile of $n(r)$].

We denote the coordinates (with the radial coordinate in phase units) of the LEO by $(\rho_{\mathrm{L}},\theta_{\mathrm{L}})$. When the LEO is outside of the atmosphere where $n\equiv 1$, $\tilde{\alpha}(\rho,\rho_*)\to 0$ as $\rho\to\infty$. Therefore, Bouguer's law becomes

$$\rho_{\mathrm{L}}\sin\left(\theta_{\mathrm{L}}+\alpha_{\mathrm{L}}\right)=\rho_*=\text{constant},\quad \alpha_{\mathrm{L}}=2\tilde{\alpha}(\rho_*,\rho_*) \tag{5.6-4}$$

Here $\alpha_{\mathrm{L}}(\rho_*)$ is the total refractive bending angle observed by the LEO. Thus, for a given LEO position $(\rho_{\mathrm{L}},\theta_{\mathrm{L}})$, there is a one-to-one correspondence (when

spherical symmetry applies) between α_{L} and the impact parameter $\rho_* = kr_* n_*$. These relations are given by

$$\left.\begin{aligned} \rho_* &= \rho_{\mathrm{L}} \sin\left(\theta_{\mathrm{L}} + \alpha_{\mathrm{L}}(\rho_*)\right) \\ \alpha_{\mathrm{L}}(\rho_*) &= 2\tilde{\alpha}[\rho_*, \theta_*] = 2\tilde{\alpha}(\rho_*, \rho_*) = -2\rho_* \int_{\rho_*}^{\infty} \frac{d\log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - \rho_*^2}} \end{aligned}\right\} \qquad (5.6\text{-}5)$$

where (r_*, θ_*) is the turning point for a ray with an impact parameter value of ρ_*. Given a position $(\rho_{\mathrm{L}}, \theta_{\mathrm{L}})$ for the LEO, Eqs. (5.6-4) and (5.6-5) yield the values ρ_* and $\alpha_{\mathrm{L}}(\rho_*)$ that must follow in order for the LEO to observe the ray (or rays if ρ_* and α_{L} are not uniquely determined at that position) from the GPS satellite that has been deflected as a result of refractive bending. The point of tangency of the ray on the boundary is located at an angular position $\theta = \pi/2 - \alpha_{\mathrm{L}}/2$ (for an infinitely distant GPS) and at a radial position of $r = r_*$.

Figure 5-3 shows an example of the solution to Eq. (5.6-5) for an exponential refractivity profile. In this case, the bending angle decreases monotonically with increasing impact parameter so that the solution is unique. If the refractivity profile causes the bending angle to exhibit reversals in slope, then there may be multiple solutions for a certain range of impact parameter values. We will discuss the question of uniqueness of the impact parameter and

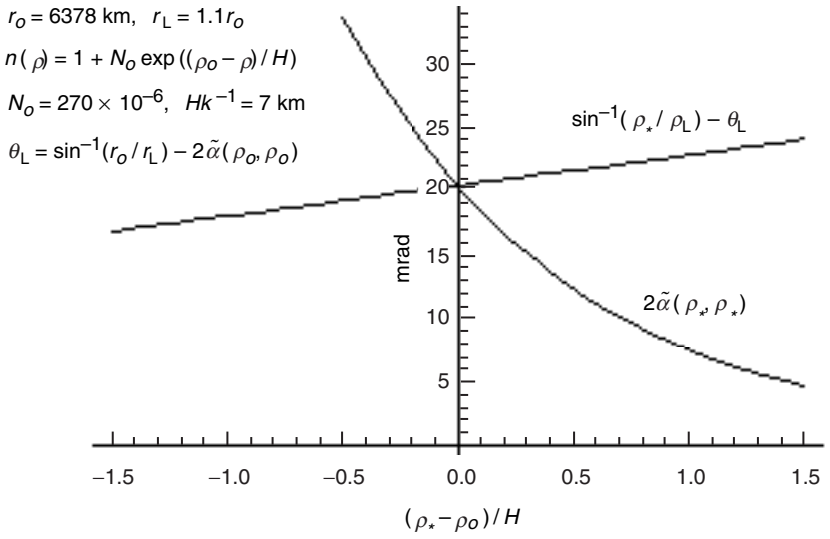

Fig. 5-3. Solution for refractive bending angle and impact parameter at the intersection of curves. Because $2\tilde{\alpha}$ is monotonic, the solution is unique.

total bending angle for a given LEO position (ρ_L, θ_L) in Section 5.12, which also addresses multipath and ray caustics.

5.6.1 Defocusing

Because defocusing will arise when we apply the stationary phase technique in wave theory, we review it here in somewhat more detail than given in Appendix A. Defocusing arises because of dispersive refraction. Incoming rays, which were collimated prior to encountering the atmosphere, are dispersed or spread out after entering the atmosphere because of the refractive gradient. To obtain a measure of the defocusing at a given point (r,θ), we compute the ratio of the signal flux density of an incoming wave prior to encountering the atmosphere with the signal flux density at the point (r,θ). This ratio is readily obtained from Bouguer's law and by invoking the principle of conservation of energy. Let $\Delta\sigma$ be the perpendicular displacement in phase units between two rays (Fig. A-1) in the neighborhood of the point (r,θ) that results from changing the impact parameter by an amount $\Delta\rho_*$. Conservation of energy requires (assuming complete transparency) that the power through a cross section of width $\Delta\rho_*$ in the collimated beam prior to atmospheric entry must equal the power through the cross section of width $\Delta\sigma$ at the point (r,θ) after atmospheric entry. Thus, the ratio $\zeta = \Delta\rho_* / \Delta\sigma$ gives us the defocusing, and its square root gives the ratio of the signal amplitudes: the amplitude of the wave at (r,θ) divided by the amplitude of the collimated wave prior to atmospheric entry.

Upon differentiating Bouguer's law given in Eq. (5.6-3) with respect to ρ_*, and defining $\gamma = \theta + \tilde{\alpha}$, one obtains

$$\left.\begin{aligned}
&\text{Incoming: } \pi > \gamma \geq \pi/2\text{: } \Delta\rho_* = \\
&\Delta r(n + \rho n'/n)\sin\gamma + \left(\Delta\theta + \Delta\tilde{\alpha}(\rho,\rho_*)\right)\cos\gamma, \\
&\text{Outgoing: } 0 < \gamma \leq \pi/2\text{: } \Delta\rho_* = \\
&\Delta r(n + \rho n'/n)\sin\gamma + \left(\Delta\theta + 2\Delta\tilde{\alpha}(\rho_*,\rho_*) - \Delta\tilde{\alpha}(\rho,\rho_*)\right)\rho\cos\gamma
\end{aligned}\right\} \tag{5.6-6}$$

Here

$$\left.\begin{aligned}
&\Delta\tilde{\alpha}(\rho,\rho_*) = \partial\tilde{\alpha}/\partial\rho_*\Delta\rho_* + \partial\tilde{\alpha}/\partial\rho\Delta\rho, \\
&\Delta\tilde{\alpha}(\rho_*,\rho_*) = d\tilde{\alpha}/d\rho_*\Delta\rho_*, \\
&\Delta\rho = k(n + rn')\Delta r,\ \ n' = \frac{dn}{dr},\ \ k\frac{dn}{d\rho} = \frac{n'}{n + rn'},\ \ \gamma = \theta + \tilde{\alpha}
\end{aligned}\right\} \tag{5.6-7}$$

From Eq. (5.6-2), it follows that

$$\left.\begin{aligned}\frac{\partial\tilde{\alpha}}{\partial\rho_*} &= -\int_\rho^\infty \frac{\rho'}{n}\frac{dn}{d\rho'}\frac{\rho' d\rho'}{\left(\rho'^2-\rho_*^2\right)^{3/2}} = \\ -\frac{\rho}{n}\frac{dn}{d\rho}&\frac{1}{\left(\rho^2-\rho_*^2\right)^{1/2}} - \int_\rho^\infty \frac{d}{d\rho'}\left(\frac{\rho'}{n}\frac{dn}{d\rho'}\right)\frac{d\rho'}{\left(\rho'^2-\rho_*^2\right)^{1/2}}\end{aligned}\right\} \tag{5.6-8}$$

and

$$\frac{\partial\tilde{\alpha}}{\partial\rho} = \frac{\rho_*}{n}\frac{dn}{d\rho}\frac{1}{\left(\rho^2-\rho_*^2\right)^{1/2}} \tag{5.6-9}$$

On the lower line of Eq. (5.6-8) the integral is well-behaved and bounded as $\rho_* \to \rho$. For an exponential refractivity profile, this integral is very closely equal to $-\tilde{\alpha}(\rho,\rho_*)/H$. But Eqs. (5.6-8) and (5.6-9) show that $\partial\tilde{\alpha}/\partial\rho_*$ and $\partial\tilde{\alpha}/\partial\rho$ have a $(\rho^2-\rho_*^2)^{-1/2}$ singularity at $\rho=\rho_*$. In the defocusing expression, these partial derivatives are multiplied by D, which is given by

$$D = \rho\cos\gamma = \pm\left(\rho^2-\rho_*^2\right)^{1/2} \tag{5.6-10}$$

D is essentially the optical distance between the point (r,θ) and the tangency point on the circle of radius ρ_*. Here the plus sign is used for an outgoing ray and the minus sign for an incoming ray. Multiplying the expressions in Eq. (5.6-8) or Eq. (5.6-9) by D removes the singularity.

In Eq. (5.6-6), Δr and $r\Delta\theta$ are constrained to follow a displacement direction at (r,θ) that is perpendicular to the ray. From Fig. A-1, it follows that

$$\left.\begin{aligned} k\Delta r &= \Delta\sigma\sin\gamma \\ kr\Delta\theta &= \Delta\sigma\cos\gamma\end{aligned}\right\} \tag{5.6-11}$$

If these quantities in Eqs. (5.6-8) through (5.6-11) are substituted into Eq. (5.6-6) and Bouguer's law is used, one can show that the defocusing ratios are given by

$$\left.\begin{aligned}
&\text{Incoming: } \pi > \gamma \geq \frac{\pi}{2}: \\
&\quad n\frac{d\sigma}{d\rho_*} = 1 - \frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\rho\cos\gamma \\
&\text{Outgoing: } 0 < \gamma \leq \frac{\pi}{2}: \\
&\quad n\frac{d\sigma}{d\rho_*} = 1 - \left(2\frac{d\tilde{\alpha}(\rho_*,\rho_*)}{d\rho_*} - \frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\right)\rho\cos\gamma
\end{aligned}\right\} \tag{5.6-12}$$

For a point located at the LEO outside of the atmosphere, $\tilde{\alpha}(\rho,\rho_*) \to 0$ as $\rho \to \infty$, and the outgoing form in Eq. (5.6-12) becomes

$$\left.\begin{aligned}
\frac{d\sigma}{d\rho_*} &= 1 - 2D\frac{d\tilde{\alpha}}{d\rho_*} \\
\tilde{\alpha} &= \tilde{\alpha}(\rho_*,\rho_*)
\end{aligned}\right\} \tag{5.6-13}$$

Here D becomes the distance (in phase units) of the LEO from the Earth's limb minus $\rho_*\alpha_{\mathrm{L}}(\rho_*)$. In practice a GPS satellites is not infinitely distant; its orbit radius is only about 4 Earth radii. To compensate for the wave-front curvature resulting from this finite distance, it is customary to use the "reduced distance" in Eq. (5.6-13), which is defined in the same way as the "reduced mass" in two-body dynamical systems. This is given by

$$D^{-1} = D_{\mathrm{L}}^{-1} + D_{\mathrm{G}}^{-1} \tag{5.6-14}$$

This definition for D follows directly from the Fresnel approximation in the thin phase screen theory (see Section 5.11). This form for D gives a slightly more accurate measure of defocusing. It also is very useful in thin phase screen analyses for diffraction and multipath.

For a circular LEO orbit, it follows from Fig. A-1 that $\Delta\sigma = D\Delta\theta_{\mathrm{L}}$, where $\Delta\theta_{\mathrm{L}}$ is the displacement in orbital position of the LEO required to intercept two nearby rays separated in impact parameter by $\Delta\rho_*$. In this case of a circular orbit for the LEO, the defocusing equation in Eq. (5.6-13) can be written as

$$\zeta^{-1} = \frac{d\sigma}{d\rho_*} = D\frac{d\theta_{\mathrm{L}}}{d\rho_*} = 1 - 2D\frac{d\tilde{\alpha}}{d\rho_*} \tag{5.6-13′}$$

From Eq. (5.6-2) it follows that

$$\frac{d\tilde{\alpha}(\rho_*,\rho_*)}{d\rho_*} = \frac{\tilde{\alpha}}{\rho_*} - \rho_*^2 \int_{\rho_*}^{\infty} \frac{d}{d\rho}\left(\frac{1}{\rho}\frac{d\log n}{d\rho}\right)\frac{d\rho}{\sqrt{\rho^2-\rho_*^2}}$$
$$\doteq -\rho_* \int_{\rho_*}^{\infty} \frac{1}{n}\frac{d^2n}{d\rho^2}\frac{d\rho}{\left(\rho^2-\rho_*^2\right)^{1/2}}, \quad \rho_* >> 0 \tag{5.6-15}$$

For an exponential refractivity profile, $d^2n/d\rho^2 = -H^{-1}dn/d\rho$, and it follows that $d\tilde{\alpha}/d\rho_* \doteq -\tilde{\alpha}/H$. In this case, the defocusing is related to the deflection $2\tilde{\alpha}D$ at the LEO expressed as the number of scale heights $2\tilde{\alpha}D/H$ that refractive bending induces.

At a turning point, $\gamma = \theta + \tilde{\alpha} \to \pi/2$. One could naïvely conclude that the defocusing factor would reduce to $1/(n_* + r_* n_*')$, which is the reciprocal of the derivative of the impact parameter ρ_* with respect to the radial coordinate kr at the turning point (r_*, θ_*). Equations (5.6-8) and (5.6-10) show that the singularity in $\partial\tilde{\alpha}(\rho,\rho_*)/\partial\rho_*$ as $\rho \to \rho_*$ yields a finite contribution to the defocusing ratio at a turning point. From Eqs. (5.6-8) and (5.6-10), Eq. (5.6-12) becomes at a turning point

$$n_* \frac{d\sigma}{d\rho_*} = 1 - \frac{\rho_*}{n_*}\left.\frac{dn}{d\rho}\right|_{\rho_*} = \frac{n_*}{n_* + r_* n_*'} \tag{5.6-16}$$

thus, confirming our intuition.

5.7 More Asymptotic Forms

We will need asymptotic forms for the functions $g_l(\rho)$ and $G[\rho,\nu]$, defined in Eqs. (5.5-19) and (5.5-20). We also must complete the unfinished business of fixing that form for $g_l(\rho)$. The derivation obtained in Section 5.5 based on modified Mie scattering fails when $\rho < \rho^{\dagger}(\nu)$. We know from Chapter 3 that the principal contributions to the scattering integrals come from spectral coefficients with wavenumber values in the near vicinity of $l = \rho_*$. Here asymptotic forms that exploit the relatively small value of $|l - \rho_*|/\rho_*$ but the large value of ρ_* are appropriate. Therefore, we use the asymptotic forms for the spherical Hankel functions in terms of the Airy functions of the first and second kind that have been given in Eq. (5.4-1).

It then follows upon replacing the spherical Hankel functions with their Airy function asymptotic forms that for $g_l(\rho)$ one obtains the following from Eq. (5.5-19):

$$g_l(\rho) \to g(\hat{y}) =$$
$$\pi K_\nu^2\left(\mathrm{Ai}'[\hat{y}]^2 + \mathrm{Bi}'[\hat{y}]^2\right) - \hat{y}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)\left(1 - \frac{\hat{y}}{5K_\nu^2} + \cdots\right) \tag{5.7-1}$$

Here $\nu = l + 1/2$. The quantities $\hat{y}$ and K_ν are defined in Eq. (5.4-3) in terms of ν and ρ. When $\nu \approx \rho$, we can drop the $\hat{y}/K_\nu^2$ term in Eq. (5.7-1) because of the enormity of ρ_*.

It follows from Eq. (5.5-20) that the applicable asymptotic form for $G[\rho,\nu]$ is given by

$$G[\rho,\nu] \doteq \int_\rho^\infty \left(\frac{d\log n}{d\rho'}\right) g(\hat{y}')d\rho', \quad \hat{y}' = \zeta\left(\frac{\nu}{\rho'}\right), \quad \rho > \rho^\dagger(\nu) \tag{5.7-2}$$

Here $\hat{y}'$ and ρ' are integration variables, and they are connected through the expression for $\hat{y}$ given in Eq. (5.4-3).

If one uses the asymptotic forms applicable to negative arguments for the Airy functions given by Eq. (3.8-7), it follows from Eq. (5.7-1) that

$$g(\hat{y}) = \left(\rho^2 - \nu^2\right)^{1/2}\left(1 + 1/\left(16\hat{y}^3\right) + \cdots\right), \quad \nu < \rho \tag{5.7-3}$$

Similarly, from Eq. (5.7-2) it follows for negative arguments that

$$G[\rho,\nu] \doteq \int_\rho^\infty \frac{d\log n}{d\rho'}\sqrt{\rho'^2 - \nu^2}\,d\rho', \quad \nu < \rho \tag{5.7-4}$$

Even at $\nu = \rho$ this asymptotic form for $G[\rho,\nu]$ is very accurate. The difference between the values of $G[\rho,\nu]$ from Eqs. (5.7-2) and (5.7-4) is roughly $0.2\beta/(1-\beta)$, yielding a relative accuracy in most conditions of a few parts in 10^5. A super-refractive region (where $\beta \geq 1$ requires special treatment; see Sections 5.8 and 6.4.

From the cumulative bending angle given by Eq. (5.6-2), we have upon integrating on ρ_*

$$\begin{aligned}\int_\nu^\rho \tilde{\alpha}(\rho,\nu')d\nu' &= \int_\rho^\infty \frac{d\log n}{d\rho'}\left[\sqrt{\rho'^2 - \rho^2} - \sqrt{\rho'^2 - \nu^2}\right]d\rho' \\ &= -\int_\rho^\infty \tilde{\alpha}(\rho',\rho')d\rho' - \int_\rho^\infty \frac{d\log n}{d\rho'}\sqrt{\rho'^2 - \nu^2}\,d\rho'\end{aligned} \tag{5.7-5}$$

From Eqs. (5.7-4) and (5.7-5), it follows that

$$G[\rho,\nu] \doteq -\int_{\nu}^{\rho} \tilde{\alpha}(\rho,\omega)d\omega - \int_{\rho}^{\infty} \tilde{\alpha}(\omega,\omega)d\omega, \quad \nu < \rho \tag{5.7-6}$$

We recall that $-G[\rho,\nu]$ is the phase delay induced by the spherical symmetric atmosphere upon a radial wavelet of wavenumber ν at the radial position $\rho = krn$. Equation (5.7-6) provides a physical interpretation for these integrals of the bending angle. This is similar in form to the stationary phase condition that we found for the thin-screen model in Chapter 2. We note from Eq. (5.7-6) that $-2G[\rho,\rho]$ corresponds to the thin-screen phase profile $\varphi(h)$ given in Eq. (2.5-1).

Differentiating $G[\rho,\nu]$ given in Eq. (5.7-2) with respect to ν and using the defining differential equation for the Airy functions, $\mathrm{Ai}'' = \hat{y}\,\mathrm{Ai}, \mathrm{Bi}'' = \hat{y}\,\mathrm{Bi}$, [and the near-linear relationship between $\hat{y}$ and ν given in Eq. (5.4-3)], one obtains

$$\frac{\partial G}{\partial \nu} = \int_{\rho}^{\infty} \left(\frac{d\log n}{d\rho'} \right) \frac{dg}{d\hat{y}} \frac{\partial \hat{y}}{\partial \nu} d\rho' \doteq -\pi K_{\nu} \int_{\rho}^{\infty} \left(\frac{d\log n}{d\rho'} \right) \left(\mathrm{Ai}[\hat{y}']^2 + \mathrm{Bi}[\hat{y}']^2 \right) d\rho' \tag{5.7-7}$$

For negative arguments of $\hat{y}$, we can use the negative asymptotic forms for the Airy functions; Eq. (5.7-7) becomes [alternatively, one can take the partial derivative of Eq. (5.7-4)]

$$\frac{\partial G}{\partial \nu} \doteq -\nu \int_{\rho}^{\infty} \frac{d\log n}{d\rho'} \frac{1}{\sqrt{\rho'^2 - \nu^2}} d\rho' = \tilde{\alpha}(\rho,\nu), \quad \nu < \rho \tag{5.7-8}$$

Comparing Eqs. (5.6-2) and (5.7-8), we conclude that, for negative values of $\hat{y}$, that is, for $\nu < \rho$, $\partial G[\rho,\nu]/\partial \nu$ *may be interpreted as the cumulative bending angle* $\tilde{\alpha}(\rho,\nu)$ *of an incoming ray at the radial position* r $(\rho = krn)$ *and with an impact parameter value* ν. The accuracy of this correspondence deteriorates only when ν lies in the immediate vicinity of ρ, that is, near a turning point (r_*, θ_*) where $\theta_* = \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$.

Figure 5-4 shows the level of this agreement in the vicinity of a turning point for an exponential refractivity profile that decreases with altitude. It roughly corresponds to the dry air refractivity profile for the Earth near sea level. Only for spectral numbers in the range $\nu > \rho - \sim 2K_{\nu}$ does the agreement deteriorate. Note that $\tilde{\alpha}(\rho,\nu)$ is not defined for $\nu > \rho$ $(\hat{y} > 0)$ and that $\partial\tilde{\alpha}(\rho,\nu)/\partial\nu \to \infty$ as $\nu \to \rho^-$, whereas $G[\rho,\nu]$ and its derivatives are well-behaved in this neighborhood. However, the form for $G[\rho,\nu]$ derived in Section 5.5 and its asymptotic form, given in Eqs. (5.7-1) and (5.7-2), change rapidly with increasing $\nu > \rho$ because of the behavior of the Airy function of

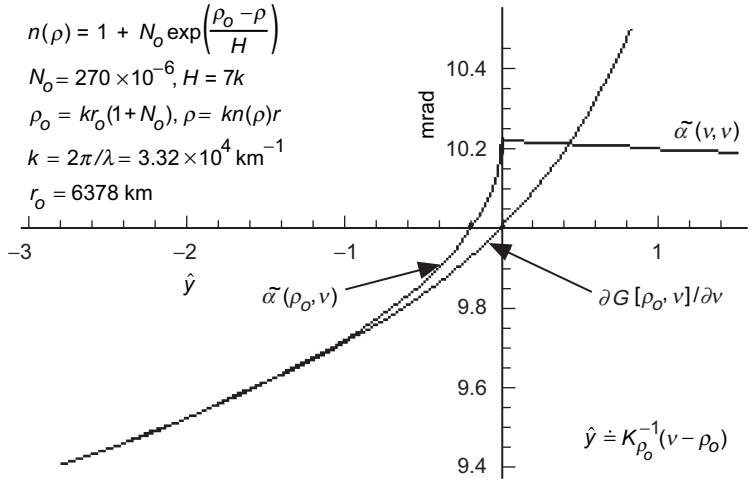

Fig. 5-4. Comparison of the gradient of the wave theory phase function $\partial G[\rho,v]/\partial v$ with the cumulative refractive bending angle $\tilde{\alpha}$ from geometric optics, in the vicinity of a turning point.

the second kind. This form fails for $v > \rho$. In fact, we show later that $g(\hat{y}) \to 0$ for increasing $\hat{y} > 0$, and that $G[\rho, v]$ approaches a constant value. The small discrepancies near the turning point result from the deviations of the Airy functions from their asymptotic forms for negative arguments, and also from the breakdown in accuracy of the spectral coefficients near a turning point when they are derived from Eq. (5.5-21). Figure 5-5 compares $\partial^2 G / \partial v^2$, a defocusing quantity from wave theory, with the analogous quantity $\partial \tilde{\alpha} / \partial v$ from geometric optics.

That there should be this very close, although not perfect, agreement between ray quantities and spectral coefficients from wave theory when the latter are evaluated at their stationary phase values should not be too surprising. In wave theory, the stationary phase process, which is discussed later, is effected over spectral number. The value $v = \rho_*$, which we will show to be very close to a stationary phase point in wave theory, also provides an equivalent ray in geometric optics between the GPS satellite and the point (r, θ), and with an impact parameter value of ρ_*. The ray path from geometric optics is in fact a path of stationary phase. Using geometric optics, we may vary the impact parameter over impact parameter space for a ray with constrained end points. The path has a stationary value of the phase delay when the impact parameter takes on the value $\rho_* = v$. Any other path with the same end points in the neighborhood of the actual ray path would present the observer at (r, θ) with a phase delay that differed from the observed phase delay by an amount that has only a second-order dependency on the coordinate and slope deviations

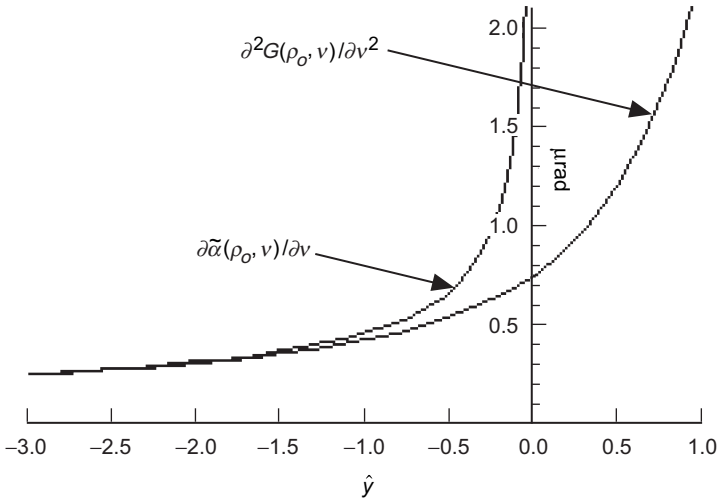

Fig. 5-5. Comparison of defocusing quantity $\partial^2 G/\partial\nu^2$ from wave theory with the analogous quantity $\partial\tilde{\alpha}/\partial\nu$ from geometric optics, in the vicinity of a turning point.

of the alternate path. This second-order variation would be due to a deviation in the value of the impact parameter. We will show that, when stationary values are assumed by the spectral number in wave theory and by the impact parameter in geometric optics, a close correspondence results.

Consider next the variability of $G[\rho,\nu_o]$ with r for a fixed value of the wavenumber ν_o. At a point (r,θ) on the approaching side at large distances where $n\to 1$, we set $kr\sin\theta=\nu_o$; that is, ν_o becomes the impact parameter for the ray passing through the point (r,θ). Figure 5-6 shows an example of $G[\rho,\nu_o]$ versus ρ in the vicinity of $\rho=\nu_o$. Here an exponential refractivity profile decreasing with altitude is used, but the altitude range shown in the figure is very narrow, $\sim 7K_{\nu_o}/k$, or about 100 meters for the parameter values shown. This is a small fraction of the refractivity scale height H (7 km) adopted in the figure. On the left side of the figure where $\hat{y}<0$, that is, where $\rho>\nu_o$, the $g(\hat{y})$ function is slowly decreasing with increasing $\hat{y}$ (decreasing ρ) (see Fig. 5-2), and $G[\rho,\nu_o]$ decreases with increasing $\hat{y}$ because the gradient of the refractivity is negative. On the right side where $\hat{y}>0$, $g(\hat{y})$ breaks sharply negative because of the rapid growth of the Airy function of the second kind in this region. Therefore, $G[\rho,\nu]$, according to the (erroneous) form given in Eq. (5.7-2), rapidly increases with increasing $\hat{y}$ because the refractivity gradient is negative in the example shown in Fig. 5-6. Therefore, $G[\rho,\nu]$ has a stationary value at the zero crossing of $g(\hat{y})$ near $\rho=\nu_o$. The correct form for $G[\rho,\nu]$ approaches a constant for increasing $\hat{y}>0$. This point

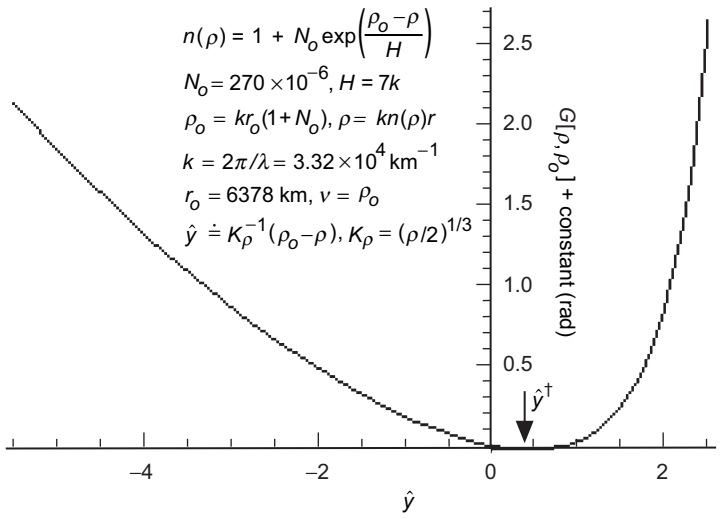

Fig. 5-6. Variation of $G[\rho,\nu_o]$ with ρ near $\rho = \nu_o = \rho_o$ obtained from Eqs. (5.7-1) and (5.7-2). This form is invalid for increasing $\hat{y} > \hat{y}^\dagger$.

where $g(\hat{y}) = 0$ marks a stationary phase point for $a_l^-(\rho)$ (with $l = \nu_o + 1/2$). It also is a stationary point for $a_l^+(\rho)$. Later, we will set $a_l^+ = a_l^-$ at this point to ensure no singularity at the origin from the spherical Hankel function.

For other refractivity profiles, there may be other stationary points for $G[\rho,\nu]$, but their occurrence and location are dependent on the functional form of the index of refraction. However, there is always one near $\rho = \nu_o$ (unless $dn/d\rho \equiv 0$), reflecting the "deepest penetration" by the corresponding ray into the sphere. This region ($\hat{y} > 0$) corresponds to refractivity features lying *below* the impact parameter, or the point of closest approach of the associated ray. Such features lying below the point of closest approach are not "felt" by the ray. Refractivity features lying near or above the point of closest approach ($\hat{y} \le 0$) are "felt" twice by the ray, incoming and outgoing; these can have a prominent effect on the shape of $G[\rho,\nu]$, depending on the actual refractivity profile.

From the definition of $G[\rho,\nu]$ in Eq. (5.7-2) and the asymptotic form for $g_l(\rho) \to g(\hat{y})$ in Eq. (5.7-1), it follows that

$$\begin{aligned}\frac{\partial G}{\partial \rho} &= -\frac{d\log n}{d\rho}\, g(\hat{y}) \\ &\doteq -\pi K_\nu^2 \frac{d\log n}{d\rho}\left(\text{Ai}'[\hat{y}]^2 + \text{Bi}'[\hat{y}]^2 - \hat{y}\left(\text{Ai}[\hat{y}]^2 + \text{Bi}[\hat{y}]^2\right)\right)\end{aligned} \tag{5.7-9}$$

Upon setting $\partial G/\partial\rho = 0$, it follows that either $g(\hat{y}) = 0$, or $dn/d\rho = 0$, or both are zero. Consider first $g(\hat{y}) = 0$ in Eq. (5.7-9). This yields a stationary point for $\hat{y}$ that is given by

$$\hat{y}^\dagger = \frac{\mathrm{Ai}'\left[\hat{y}^\dagger\right]^2 + \mathrm{Bi}'\left[\hat{y}^\dagger\right]^2}{\mathrm{Ai}\left[\hat{y}^\dagger\right]^2 + \mathrm{Bi}\left[\hat{y}^\dagger\right]^2} = 0.44133\cdots \tag{5.7-10}$$

From the properties of the Airy functions, it can be shown that this value of $\hat{y}^\dagger$ is unique. We have seen this quantity before in Section 4.6 in regard to turning points for waves traveling in a Cartesian-stratified Airy medium.

Upon using Eq. (5.4-3), Eq. (5.7-9) yields a stationary phase point $\rho = \rho^\dagger$ that is a function of ν_o and which is given by

$$\left.\begin{aligned} \rho^\dagger &\doteq \nu_o - K_{\rho^\dagger}\hat{y}^\dagger \\ K_{\rho^\dagger} &= \left(\frac{\rho^\dagger}{2}\right)^{1/3} \end{aligned}\right\} \tag{5.7-11}$$

Thus, $|\rho^\dagger - \nu_o|/\nu_o = y^\dagger/2K^2_{\rho^\dagger} \sim 10^{-6}$ for $\nu_o \sim 10^8$.

Continuing to the second derivative, $\partial^2 G/\partial\rho^2$, and evaluating it at $\rho = \rho^\dagger(\nu_o)$, it follows that

$$\left.\frac{\partial^2 G}{\partial\rho^2}\right|_{\rho^\dagger} \doteq -\pi K_\nu \left.\frac{d\log n}{d\rho}\right|_{\rho^\dagger} \left(\mathrm{Ai}\left[\hat{y}^\dagger\right]^2 + \mathrm{Bi}\left[\hat{y}^\dagger\right]^2\right) \tag{5.7-12}$$

Setting ν_o equal to $kr_o n(r_o)$, adopting the Earth's dry air conditions at sea level, and using 19 cm for the GPS wavelength, $\partial^2 G/\partial\rho^2|_{\rho^\dagger} \sim 10^{-6}$ rad. It follows that G will vary from its stationary phase value by the order of 1 radian when $G|\rho - \rho^\dagger|/\rho^\dagger \sim 10^{-5}$; in other words, the stationary phase neighborhood about $\rho = \rho^\dagger$ is very narrow (a few dekameters) when $\rho^\dagger$ is so large ($\sim 2\times 10^8$).

In stationary phase theory (to be discussed in a later section), we will use $\partial^2 G[\rho,\nu]/\partial\nu^2$, which is related to defocusing. A comparison of this second derivative with $\partial\tilde{\alpha}(\rho,\nu)/\partial\nu$, also used in geometric optics for defocusing, is shown in Fig. 5-5 for the same conditions given in Fig. 5-4. These also agree closely except near a turning point. Accordingly, we expect the accuracy of the amplitude predicted by the osculating parameter technique to degrade for $\hat{y} > -2$. Numerical results verify this threshold. At a turning point, the correct

value for the defocusing in a refracting medium without scattering is given by Eq. (5.6-16). At a turning point, the stationary phase value in spectral number using the osculating parameter approach is equivalent to $\hat{y} \approx 0$; it predicts unity there for the defocusing.

5.7.1 Equating $dG\left[\rho^{\dagger}(\nu),\nu\right]/d\nu$ to $\tilde{\alpha}(\nu,\nu)$ and $d^2G\left[\rho^{\dagger}(\nu),\nu\right]/d\nu^2$ to $d\tilde{\alpha}(\nu,\nu)/d\nu$

We note in Fig. 5-4 at the turning point that the value of $\partial G[\rho_o,\nu]/\partial\nu$ at $\nu = \nu^{\dagger} = \rho_o + K_{\rho_o}\hat{y}^{\dagger}$ is very close to the value of $\tilde{\alpha}(\rho_o,\rho_o)$. In fact, the actual numbers for Fig. 5-4 are 10.114 mrad for $\tilde{\alpha}(\rho_o,\rho_o)$ and 10.115 mrad for $\partial G[\rho_o,\nu]/\partial\nu$. Since $\rho = \rho^{\dagger}(\nu)$ marks a stationary point for $G[\rho,\nu]$ where $\partial G/\partial\rho = 0$, it follows that

$$\frac{dG^{\dagger}(\nu)}{d\nu} \equiv \frac{dG\left[\rho^{\dagger}(\nu),\nu\right]}{d\nu} = \left(\frac{\partial G[\rho,\nu]}{\partial\nu}\right)_{\rho=\rho^{\dagger}} \tag{5.7-13}$$

Hence, $dG^{\dagger}(\nu)/d\nu$ also will be very close in value to $\tilde{\alpha}(\nu,\nu)$. We show in Appendix J that

$$\left.\begin{aligned} \frac{dG^{\dagger}(\nu)}{d\nu} &= \tilde{\alpha}(\nu,\nu) + \mathrm{O}\left[\rho\frac{d^2n}{d\rho^2}\right] \\ G^{\dagger}(\nu) &= G\left[\rho^{\dagger}(\nu),\nu\right] \end{aligned}\right\} \tag{5.7-14}$$

Therefore, the difference between $dG^{\dagger}(\nu)/d\nu$ and $\tilde{\alpha}(\nu,\nu)$ depends on the curvature in the refractivity profile. For the thin-atmosphere conditions shown in Fig. 5-4, it can be shown that $\rho_o\left(d^2n/d\rho^2\right) \approx 10^{-3}$ mrad, which we can ignore. The difference becomes significant when near-super-refractivity conditions are encountered. See Appendix J.

Similarly, it can be shown that

$$\begin{aligned} \frac{d^2G^{\dagger}(\nu)}{d\nu^2} \equiv \frac{d^2G\left[\rho^{\dagger}(\nu),\nu\right]}{d\nu^2} &\doteq \left(-\frac{\partial^2G[\rho,\nu]}{\partial\rho^2} + \frac{\partial^2G[\rho,\nu]}{\partial\nu^2}\right)_{\rho=\rho^{\dagger}} \\ &= \frac{d\tilde{\alpha}(\nu,\nu)}{d\nu} + \mathrm{O}\left[\frac{d^2n}{d\rho^2}\right] \end{aligned} \tag{5.7-15}$$

For thin-atmosphere conditions, the curvature term here also can be dropped. Equations (5.7-14) and (5.7-15) will play key roles when we apply the

stationary phase technique in wave theory to outgoing waves. Equation (5.7-14) is related to the ray bending angle, and Eq. (5.7-15) is related to defocusing.

Regarding the near-equivalence of $dG[\rho^{\dagger}(\nu),\nu]/d\nu$ and $\tilde{\alpha}(\nu,\nu)$, we note a property of Bauer's identity. It is given in cylindrical coordinates in Eq. (4.10-9). We can transform $\theta \to \theta + \alpha$ to obtain

$$\exp\left[i\rho\cos(\theta+\alpha)\right] = \sum_{l=-\infty}^{l=\infty} i^l J_l(\rho)\exp\left[il(\theta+\alpha)\right] \tag{5.7-16}$$

Noting that $G[\rho_o,\nu]$ shows up in the exponential term in the spectral series representations for the electromagnetic field, its variability with spectral number will be related to the angle α above. Thus, we would expect $\partial G[\rho_o,\nu]/\partial\nu$ to be closely related to an angle, which turns out to be the bending angle.

5.7.2 Fixing the Form for $g(\hat{y})$ when $\hat{y} > 0$ by Asymptotic Matching

We have noted the failure of the modified Mie scattering approach to secure the correct form for $g(\hat{y})$ when $\hat{y} >\sim 0$. Equation (5.7-1) predicts that $g(\hat{y}) \to \infty$ for increasing $\hat{y} > 0$; in fact, $g(\hat{y})$ should approach zero. We can use the form for $g(\hat{y})$ obtained from an Airy layer as guidance. In an Airy layer with a boundary at $r = r_0$, the profile for the index of refraction is given by $n^2 = n_0^2 + 2n_0 n'(r - r_0)$, where n_0 and n' are constants. The quadratic term $(n'r_0)^2((r - r_0/r_0))^2$ is negligible.

In Chapter 4, Section 4.12, we showed that the solutions to the wave equations in a spherical Airy layer are given by the Airy functions. Let $U^{\pm}(\tilde{y})$ be a solution for the scalar field, with the top sign for an outgoing wave and the bottom sign for an incoming wave. When $\beta < 1$, from Eq. (4.10-3) and Eqs. (4.12-4) through-(4.12-6) we have

$$U^{\pm}(\tilde{y}) = c^{\pm}\left(\mathrm{Ai}[\tilde{y}] \mp i\,\mathrm{Bi}[\tilde{y}]\right) \tag{5.7-17}$$

where $c^{\pm}$ are complex constants obtained from matching this Airy function solution and its derivative at the boundary $r = r_0$ with the incoming and outgoing wave forms applicable on the other side of the boundary. The argument of the Airy functions is given by

$$\left.\begin{aligned} \tilde{y} &= \frac{1}{K_{\rho_0}|1-\beta|^{2/3}}\left((\beta-1)(\tilde{\rho}-\tilde{\rho}_0)+l-\tilde{\rho}_0\right) \\ \beta &= -\frac{n'r_0}{n_0} \\ \tilde{\rho} &= kn_0 r \end{aligned}\right\} \tag{5.7-18}$$

These forms also apply in a super-refracting medium where $\beta > 1$. Also, it is easily shown that in an Airy layer

$$\begin{aligned} \rho = krn(r) &= \rho_0 + k\left(n_0 + n'r_0\right)\left(r-r_0\right) + kn'\left(r-r_0\right)^2 \\ &\doteq \tilde{\rho}_0 + (1-\beta)\left(\tilde{\rho}-\tilde{\rho}_0\right) - \beta\left(\tilde{\rho}-\tilde{\rho}_0\right)^2 / \tilde{\rho}_0 \end{aligned} \tag{5.7-19}$$

The quadratic term is negligible, and it is dropped. Note that $\rho_0 = \tilde{\rho}_0$. It follows that the arguments of the Airy functions are related by

$$\hat{y} \doteq \frac{1}{K_{\rho_0}}\left(l-\rho\right) = \tilde{y}|1-\beta|^{2/3} \tag{5.7-20}$$

From Eq. (5.7-17), it follows that the phase $\psi_l^{A^-}$ of the incoming Airy function solution (for $\beta < 1$) is given by

$$\psi_l^{A^-} = \tan^{-1}\left(\frac{\mathrm{Bi}[\tilde{y}]}{\mathrm{Ai}[\tilde{y}]}\right) + \text{constant} \tag{5.7-21}$$

Similarly, by expressing an incoming spherical Hankel function $\xi_l^-(\rho)$ in terms of its Airy function asymptotic form, its phase is given by

$$\psi_l^{H^-} = \tan^{-1}\left(\frac{\mathrm{Bi}[\hat{y}]}{\mathrm{Ai}[\hat{y}]}\right) + \text{constant} \tag{5.7-22}$$

The difference $\psi_l^{A^-} - \psi_l^{H^-}$ is the phase accumulation in the lth spectral coefficient $a_l^-(\rho)$ for an Airy layer. We denote this phase of the spectral coefficient by the function $\vartheta_l^-(\rho) = \psi_l^{A^-}(\rho) - \psi_l^{H^-}(\rho)$. Taking the derivative $\partial\vartheta_l^- / \partial\rho$, we obtain

$$\frac{\partial\vartheta_l^-}{\partial\rho} = \frac{-1}{\pi K_{\rho_0}}\left(\frac{\tilde{y}}{\mathrm{Ai}[\tilde{y}]^2 + \mathrm{Bi}[\tilde{y}]^2} - \frac{\hat{y}}{\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2}\right)\frac{1}{\hat{y}} \tag{5.7-23}$$

We note that $\partial\vartheta_l^- / \partial\rho$ rapidly approaches zero for increasing $\hat{y} > 0$.

We compare this expression with $\partial G[\rho, \nu] / \partial\rho$ given in Eq. (5.7-9). For negative $\hat{y}$, we replace the Airy functions in Eq. (5.7-23) with their asymptotic forms. Then, using Eq. (5.7-20), we obtain

$$\frac{\partial\vartheta_l^-}{\partial\rho} \to \frac{-1}{K_{\rho_0}}\left((-\tilde{y})^{3/2} - (-\hat{y})^{3/2}\right)\frac{1}{\hat{y}} = \frac{\beta}{(1-\beta)K_{\rho_0}}\sqrt{-\hat{y}} \to \frac{-n'}{1-\beta}\sqrt{\rho^2 - \nu^2} \tag{5.7-24}$$

whereas, from Eq. (5.7-9), $\partial G[\rho, \nu] / \partial\rho$ becomes

$$\frac{\partial G}{\partial\rho} = -\frac{d\log n}{d\rho}\, g(\hat{y}) \to -2K_\nu^2 \frac{d\log n}{d\rho}\sqrt{-\hat{y}} = \frac{-n'}{1-\beta}\sqrt{\rho^2 - \nu^2} \tag{5.7-25}$$

These asymptotic forms are identical for decreasing $\hat{y} < 0$.

Equation (5.7-23) for $\partial\vartheta_l^- / \partial\rho$ applies only to an Airy layer, but it is essentially exact and it applies without restriction on the value of $\hat{y}$. On the other hand, $\partial G[\rho, \nu] / \partial\rho$ obtained from Eq. (5.7-9) is a general form applying to any physical profile for $n(r)$, but it fails for increasing $\hat{y} > 0$. Clearly, we have a potential match made in heaven. Over the troublesome interval, say beginning at $\hat{y} > -2$, we can use the form for $\partial\vartheta_l^- / \partial\rho$ given in Eq. (5.7-23). The range $-2 \le \hat{y} \le 2$ corresponds to a spatial range of $4k^{-1}K_{\rho_0} \approx 60$ m. Approximating the index of refraction profile by a constant gradient n' should be fairly accurate in most circumstances, especially since the phase variability is rapidly dying out with increasing $\hat{y} > 0$.

If a single Airy layer is not sufficient, then we can form a concatenated series of Airy layers near a turning point. Any physical refractivity profile can be approximated by a series of Airy layers. In this case, β and $\tilde{y}$ would be discontinuous, according to Eqs. (5.7-18) and (5.7-20), across each boundary between Airy layers (see Section 4.7). Applying the continuity conditions to the wave functions and their derivatives across each boundary ties the Airy function solutions together for the different layers, which enables one to derive the phase in any layer. For a given spectral number, whenever one reaches downward through successive Airy layers to a radial distance $\rho < \rho^\dagger(\nu)$, $\tilde{y}$ will become positive there, and $\partial\vartheta_l^- / \partial\rho$ for that Airy layer rapidly approaches zero for increasing $\hat{y} > 0$.

Figure 5-7 shows two comparisons of $\partial\vartheta_l^- / \partial\rho$ and $\partial G[\rho, \nu] / \partial\rho$ versus $\hat{y}$. In Fig. 5-7(a), the refractivity profile corresponds to dry air at sea level, but in

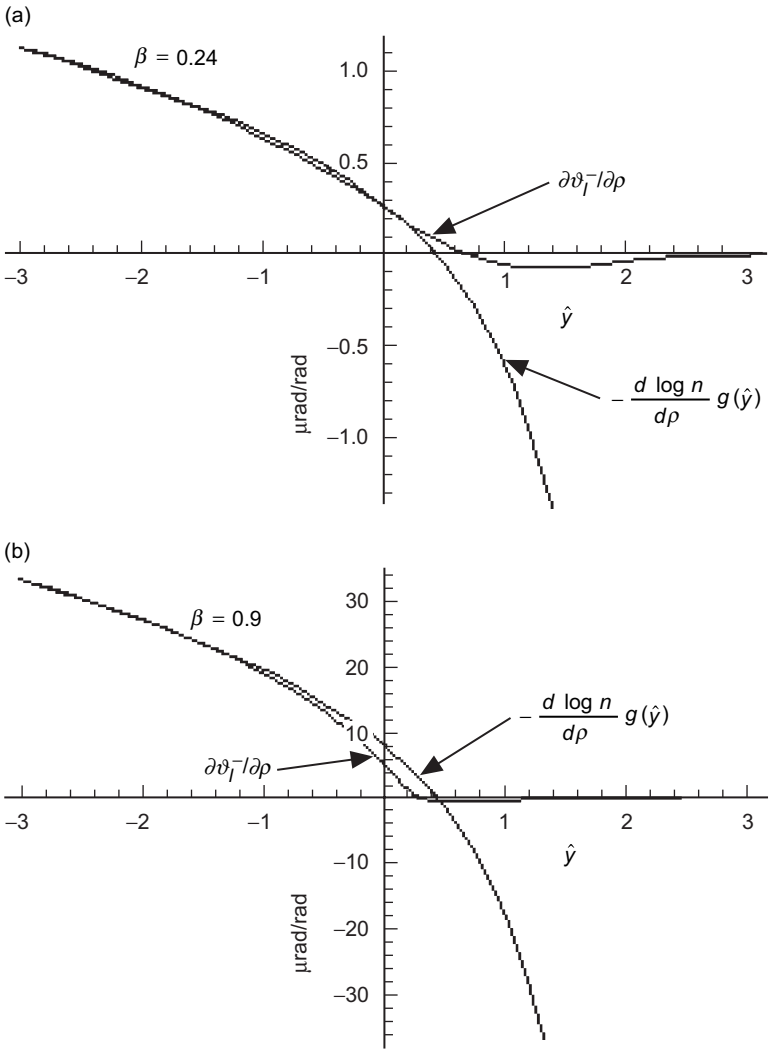

Fig. 5-7. Gradient of the phase of $a_l^-(\rho)$ in an Airy layer: (a) $\beta = 0.24$ and (b) $\beta = 0.9$. $\partial\vartheta_l^-/\partial\rho$ is the exact value; $-(d \log n/d\rho)g(\hat{y})$ is from modified Mie scattering theory.

Fig. 5-7(b), a rather hefty value of $\beta = 0.9$ has been used, hardly thin-atmosphere material.

5.7.3 Behavior of the WKB Solution at a Turning Point

The phase of a wave approaches a constant value as a function of ρ below a turning point. Its behavior can be quantified by examining the WKB solutions

to the radial wave equation in Eq. (5.3-5), which are given in Eq. (5.3-10) with $f_l(u) = (n^2u^2 - l(l+1))/u^2$. The incoming and outgoing WKB solutions, $W_l^-(u)$ and $W_l^+(u)$, can be linearly combined to yield in the $f_l < 0$ regime where $\rho < \nu$ an exponentially damping solution for decreasing $\rho < \nu$, and also an exponentially increasing solution. Using the connection formulas [6] between the WKB solutions for these two regimes, $\rho < \nu$ and $\rho > \nu$, we have

$$\left.\begin{array}{lll} \underline{f_l(u) < 0} & \underline{f_l(u_o) = 0} & \underline{f_l(u) > 0} \\ \dfrac{1}{2(-f_l)^{1/4}} \exp\left(-\int_u^{u_o} \sqrt{-f_l}\, du\right) & \rightarrow & \dfrac{1}{(f_l)^{1/4}} \cos\left(\int_{u_o}^{u} \sqrt{f_l}\, du - \dfrac{\pi}{4}\right) \\ \dfrac{\sin\left(\varphi + \dfrac{\pi}{4}\right)}{(-f_l)^{1/4}} \exp\left(\int_u^{u_o} \sqrt{-f_l}\, du\right) & \leftarrow & \dfrac{1}{(f_l)^{1/4}} \cos\left(\int_{u_o}^{u} \sqrt{f_l}\, du + \varphi\right) \end{array}\right\} \quad (5.7\text{-}26)$$

where φ is an arbitrary phase that is not too close in value to $-\pi/4$.[8] For $f_l = u$ and $\varphi = +\pi/4$, these are the leading terms of the asymptotic forms for the Airy functions of the first and second kind.

It follows that one solution to the wave equation is exponentially increasing for decreasing $\rho < \nu$, and the other is exponentially damping to zero. The asymptotic form of the incoming wave, U_l^-, in the $\rho < \nu$ regime is a linear combination of these two exponential solutions. Therefore, the phase of U_l^- must approach a constant value with decreasing $\rho < \nu$. The Airy layer analysis just discussed shows that the phase of U_l^- *rapidly* approaches a constant; therefore, $\partial G / \partial \rho$ must rapidly approach zero. Expanding f_l in powers of $u - u_o$ and setting $n_o u_o = l$ leads to $f_l = -\hat{y}/K_{\rho_o}^2 + \mathrm{O}[\hat{y}/K_{\rho_o}^2]^2$. The exponential terms in the WKB solutions become $\exp(\pm 2\hat{y}^{3/2}/3)$. For $\hat{y} = 2$, the ratio of their amplitudes is about 100:1; for $\hat{y} = 3$, it is 2000:1.

---

[8] The arrow in each of these two connection formulas indicates the applicable direction of information transfer. For example, continuing the exponentially damping solution in the $f < 0$ regime into the $f > 0$ regime leads to a stable sinusoidal solution in that regime with "twice" the amplitude and a phase offset of $-\pi/4$. But the reverse cannot be guaranteed. Inaccuracy in the numerical computation of the solution to the wave equation starting in the $f > 0$ regime and integrating downward into the $f < 0$ regime, even with the phase set equal to $-\pi/4$ exactly, leads inevitably to a numerical solution that blows up for decreasing $\rho < \nu$.

5.7.4 Setting $G[\rho,\nu]$ for $\rho \le \rho^{\dagger}(\nu)$

In the discussions to follow, we have taken a simpler approach in view of Eq. (5.7-6) and also the very close agreement between $dG[\rho^{\dagger}(\nu),\nu]/d\nu$ and $\tilde{\alpha}(\nu,\nu)$. We let $G[\rho,\nu]$ run its course based on Eq. (5.7-2) for $\rho^{\dagger}(\nu) \le \rho < \infty$. It has a stationary value at $\rho = \rho^{\dagger}(\nu)$. Then we set $G[\rho,\nu] = G[\rho^{\dagger}(\nu),\nu] = G^{\dagger}(\nu)$ for $\rho \le \rho^{\dagger}(\nu)$, a constant value for a given value of ν. In summary, we modify Eq. (5.7-2) as follows:

$$G[\rho,\nu] = \begin{cases} \displaystyle\int_{\rho}^{\infty} \left(\frac{d\log n}{d\rho'}\right) g(\hat{y}')d\rho', & \rho \ge \rho^{\dagger}(\nu) \\ \displaystyle\int_{\rho^{\dagger}}^{\infty} \left(\frac{d\log n}{d\rho'}\right) g(\hat{y}')d\rho', & \rho \le \rho^{\dagger}(\nu) \end{cases} \tag{5.7-27}$$

Also, we have

$$\frac{\partial G[\rho,\nu]}{\partial \rho} = \begin{cases} -\left(\dfrac{d\log n}{d\rho'}\right) g(\hat{y}'), & \rho \ge \rho^{\dagger}(\nu) \\ 0, & \rho \le \rho^{\dagger}(\nu) \end{cases} \tag{5.7-28}$$

This arrangement, which involves some approximation (see Fig. 5-7), also simplifies and clarifies the strategy for matching the incoming and outgoing spectral coefficients to eliminate the Hankel function singularity at the origin. We discuss this further in Section 5.8.

5.8 Spectral Representation of an Electromagnetic Wave in a Spherical Stratified Medium

We will need to distinguish between an incoming region in the medium and an outgoing region. Figure 5-8 provides an example of the simplest topology for these regions. The boundary between these regions is given by the locus of points (r_*,θ_*), which defines the turning point for each ray, where for a spherical symmetric medium $E_{\theta}(r_*,\theta_*) = 0$. In geometric optics $\theta_* = \pi/2 - \tilde{\alpha}(\rho_*,\rho_*)$, which is obtained from Eq. (5.6-5). By "incoming" we mean the field at any point where the Poynting vector $\boldsymbol{S} = c(\boldsymbol{E}\times\boldsymbol{H})/4\pi$ for the wave is such that $\boldsymbol{S}\cdot\boldsymbol{r} < 0$. For the simple topology shown by Fig. 5-8, for example, the incoming region includes any point with a radial position r and an angular position θ that satisfies the condition $\theta_* < \theta < \pi$.

When the profile of the refractivity gradient has reversals in polarity, these regions may not be so simply connected. When a scattering surface is present,

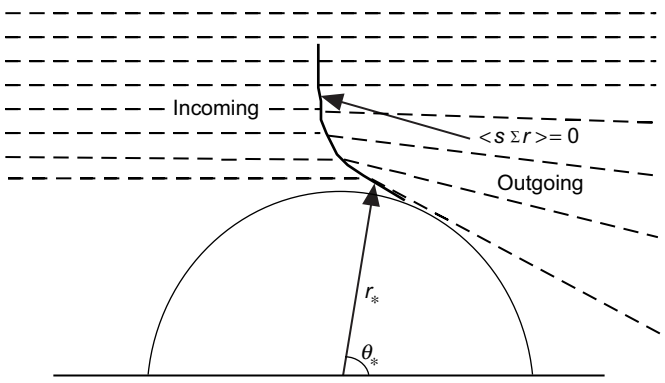

Fig. 5-8. Schematic of simple incoming and outgoing regions.

one could obtain both incoming and outgoing waves at the same point. Figure 5-9 sketches a more complicated refractivity profile that also produces coincident incoming and outgoing waves in certain regions.

We first develop the spectral coefficients for the pure refraction case, and we assume that no scattering surfaces are present in the medium. Scattering or diffraction occurs where sharp changes in gradient or discontinuities in the refractivity occur. We will deal with those cases later. In particular, we assume that $n(r)$ and its derivatives are continuous throughout the medium and that the simple topology of the kind shown in Fig. 5-8 applies.

Let us evaluate the electric field vector $\boldsymbol{E}(r,\theta)$ at a point within the refracting medium in terms of its incoming and outgoing spectral coefficients. We assume that the planar asymptotic boundary condition applies to the approaching wave and, therefore, a_l^- is obtained from Eq. (5.5-21). The proper functional form for a_l^+ is more problematical, and we defer that to later in this section. From Eqs. (5.3-6), (5.5-2), and (5.5-8), we have the following spectral representation for the in-plane ($\phi = 0$) radial and transverse components of the electric field for the TM wave:

$$\left.\begin{aligned} E_r &= \frac{E_o}{2\rho^2}\sum_{l=1}^{\infty} l(l+1)\left(a_l^-\xi_l^- + a_l^+\xi_l^+\right)\mathrm{P}_l^l(\cos\theta) \\ E_\theta &= \frac{E_o}{2\rho}\sum_{l=1}^{\infty}\left(a_l^-\xi_l'^- + a_l^+\xi_l'^+\right)\frac{\partial}{\partial\theta}\mathrm{P}_l^l(\cos\theta) \end{aligned}\right\} \qquad (5.8\text{-}1a)$$

Because of the enormity of ρ and of l from where essentially all contributions to these summations originate, we can replace these summations by integrals, which are given by

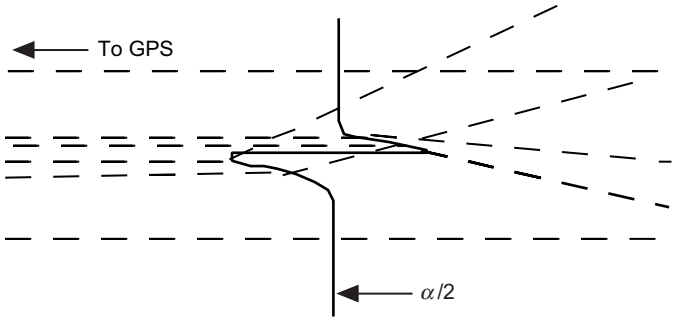

Fig. 5-9. Example of a topology with overlapping incoming and outgoing regions, from a sharp change in refractivity gradient.

$$\left.\begin{aligned} E_r &= \frac{E_o}{2\rho^2}\int_0^\infty \nu^2\left(a_l^-\xi_l^- + a_l^+\xi_l^+\right)\mathrm{P}_l^l(\cos\theta)d\nu \\ E_\theta &= \frac{E_o}{2\rho}\int_0^\infty \left(a_l^-\xi_l'^- + a_l^+\xi_l'^+\right)\mathrm{P}_l^{l\,\prime}(\cos\theta)d\nu \end{aligned}\right\} \qquad (5.8\text{-}1b)$$

A similar set holds for the magnetic field. In the limit as n assumes a constant value throughout the medium, $a_l^+ \to a_l^- \to a_l$, and Eq. (5.8-1) reduces to the collimated form in Eq. (5.5-1) with $a_l = i^{l-1}(2l+1)/l(l+1)$, as it must. Also, using stationary phase arguments, we will show later that in the incoming region well away from a turning point only the $a_l^-(\rho)\xi_l^-(\rho)$ term contributes significantly to the scattering integral; in the outgoing region, only the $a_l^+\xi_l^+$ term contributes. Only for points very near a turning point must we include contributions from both spherical Hankel functions. This means that, except near a turning point, we may replace the terms $a_l^\pm\xi_l^\pm/2$ with $a_l^\pm\psi_l$ for computational purposes, which decays to zero rapidly with increasing $l > \rho$.

5.8.1 Behavior of $\partial G/\partial\nu$

The spectral representation for the electric field in Eq. (5.8-1b) involves integrals over spectral number. To evaluate these integrals using the stationary phase technique as an aid, we need to find those spectral neighborhoods where the phase accumulation of the integrands is stationary. In this regard, we study the variability of $G[\rho,\nu]$ with spectral number, $\partial G/\partial\nu$, in terms of the refractivity profile that determines it. We first look at $\partial G/\partial\nu$ for two different refractivity profiles. Both examples adhere to the thin-atmosphere assumptions. For Case A, n is exponentially distributed so that $d\alpha/dr$ is monotonic negative with altitude; therefore, no multipath or shadow zone situations arise. For this case, $\partial G/\partial\nu$ already has been shown in Fig. 5-4. In Case B, n has a

Gaussian distribution; multipath, shadow zones, and caustics are prominent features for this distribution. These two profiles for the index of refraction are given by

$$n = 1 + N_o \exp\left[-(\rho - \rho_o)/H\right] \tag{5.8-2a}$$

$$n = 1 + N_w \exp\left[-(\rho - \rho_w)^2/2H_w^2\right] \tag{5.8-2b}$$

Case B is useful for study of spherical shell structures embedded in an ambient profile such as that given by Case A. Case B could be used to describe the refractivity profile of a sporadic E-layer in the ionosphere or a marine layer in the lower troposphere. Case A can be used to describe the Earth's refractivity profile for dry air, and the values used for N_o and H in these examples correspond roughly to dry air refractivity ($N_o = 270 \times 10^{-6}$) and scale height ($k^{-1}H = 7\,\text{km}$) at sea level. For computational convenience, we now have written these refractivity profiles in terms of $\rho = krn(\rho)$, i.e., r is an implicit function of ρ.[9]

For the exponential refraction profile given by Case A, it can be shown using the thin-atmosphere assumption in the defining path integral for $\tilde{\alpha}(\rho, \nu)$, given in Eq. (5.6-2) [see also Appendix A, Eq. (A-30)], that $\tilde{\alpha}(\rho, \nu)$ is given by

---

[9] The extraction of n given a value of $u = kr$ through iteration of Eq. (5.8-2) is cumbersome because $\rho = un(\rho)$. However, this form for an exponential distribution in terms of ρ has some advantages. It results in the very simple form for $\tilde{\alpha}(\rho, \nu)$ given in Eq. (5.8-3), which closely agrees with the numerical integration version for $\tilde{\alpha}(\rho, \nu)$ given in Eq. (5.6-2). Provided $|un'/n| < 1$, there is a unique relationship between u and ρ. Equation (5.8-3) includes the effects of second-order ray path bending in the path integral in Eq. (5.6-2). When a positional exponential form is used instead, $n(u) = 1 + N_o \exp[-(u - u_o)/H_u]$, the form for $\tilde{\alpha}(u, \nu)$ requires a series in powers of (un'/n) to account for higher-order ray path bending. This has been discussed in Appendix A. For a given value of scale height $H_\rho = H$, the version in Eq. (5.8-2a) gives a smaller radial gradient than the positional exponential version (about 20 percent less for dry air at sea level). This is seen by noting that, for the two versions of the exponential distributions, one obtains the following for dry air at sea level:

$$H_u/H_\rho \doteq (d\log n/d\log\rho)\big/(d\log n/d\log u) = d\log u/d\log\rho = 1 - \rho n'/n \approx 1.2$$

One can adjust the values of the parameters N_o and H in Eq. (5.8-2a) to attain a close, but not exact, match with the profile from the positional exponential form; also, the bending-angle profiles from the two versions can be matched rather closely.

For near-super-refractivity situations where $un'/n \to -1$, then the functional form $n(\rho)$ becomes inconvenient because $dn/d\rho = n'/(1 + un') \to \infty$. But the defining integral for $G[\rho, \nu]$ readily allows a change of variable to u.

$$\tilde{\alpha}(\rho,\nu) \doteq \big(n(\rho)-1\big)\sqrt{\frac{\pi\rho}{2H}}\left(1-\operatorname{erf}\left[\sqrt{\frac{\rho^2-\nu^2}{2\rho H}}\right]\right),\ \nu<\rho \qquad (5.8\text{-}3)$$

where $n(\rho)$ is given by Eq. (5.8-2a). This expression is very accurate when compared to the value from the path integral in Eq. (5.6-2). For the decreasing exponential distribution given in Eq. (5.8-2a), the difference is less than 1 percent for dry air. Equation (5.8-3) accounts for both the first- and second-order ray path bending effects when the refractivity profile in Eq. (5.8-2a) is used. For the dry air component of the Earth's atmosphere, the first-order bending term, $\big(\sqrt{2}-1\big)un'/n$, contributes less than 10 percent of the total. The second-order ray path bending effect for dry air at sea level amounts to roughly 1 percent of the total.

Figure 5-10 shows the variation of $\partial G/\partial\nu$ with ν while holding ρ fixed for Case A. It was obtained from the integral in Eq. (5.7-7). For $\nu<\rho$, we have already noted from Eqs. (5.6-2) and (5.7-8), and from Fig. 5-4, that $\partial G/\partial\nu \doteq \tilde{\alpha}(\rho,\nu)$, which holds very accurately for all values of ν nearly up to $\nu=\rho$. Numerical integrations of $\partial G/\partial\nu$ from Eq. (5.7-7) and $\tilde{\alpha}(\rho,\nu)$ from Eq. (5.6-2) show microradian-level agreement for both Cases A and B when the inequality $\nu-\rho <\sim -2K_\rho$ and the thin-atmosphere assumptions are satisfied.

The three curves for $\partial G/\partial\nu$ in Fig. 5-10 correspond to three different radial positions, which are defined, respectively, by $\rho-\rho_o=0, H/2, H$, and $\rho=un(\rho)$. The dashed curves in Figs. 5-10 and 5-11 mark the value of $\partial G/\partial\nu$ for $\nu=\rho_*+\hat{y}^\dagger K_{\rho_*}$ as a function of ρ_*, which is a stationary point for $G[\rho_*,\nu]$.

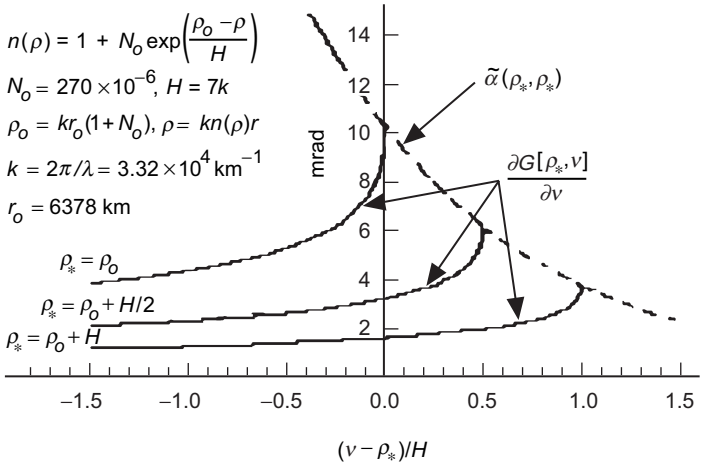

Fig. 5-10. $\partial G/\partial\nu$ **curves for an exponential refractivity profile. The dashed curve shows** $\partial G/\partial\nu$ **at** $\nu=\rho_*+\hat{y}^\dagger K_{\rho_*}$ **as** ρ_* **varies or, equivalently,** $\tilde{\alpha}(\rho_*,\rho_*)$.

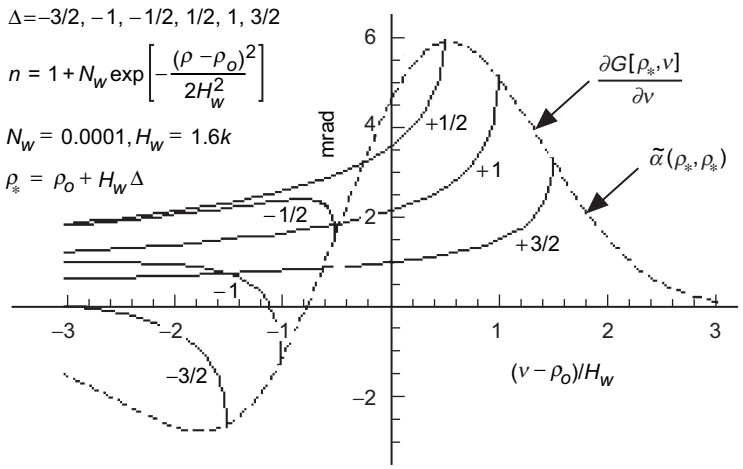

Fig. 5-11. $\partial G/\partial \nu$ **curves for a Gaussian refractivity profile for different locations of its center. The dashed curve is the value of** $\partial G/\partial \nu$ **at** $\nu = \rho_* + \hat{y}^{\dagger} K_{\rho_*}$ **or, equivalently,** $\tilde{\alpha}(\rho_*, \rho_*)$.

We have shown that at this stationary point $dG/d\nu = \partial G/\partial \nu \doteq \tilde{\alpha}(\rho, \rho)$ with high accuracy. Here, $\tilde{\alpha}(\rho, \rho)$ is given from either Eq. (5.6-2), or from Eq. (5.8-3) when Case A applies.

Figure 5-11 shows the resulting curves for $\partial G/\partial \nu$, obtained from the integral in Eq. (5.7-7) when the Gaussian refractivity profile in Eq. (5.8-2b) is used. If this profile also is applied in Eq. (5.6-2) to obtain $\tilde{\alpha}(\rho, \rho)$, a bipolar refractive bending-angle profile results, which is shown in Appendix E, Fig. E-1, and also as the dashed curve in Fig. 5-11. The bending-angle profile mimics the shape of $dn/d\rho$, but it is modified to reflect the geometry of a ray transecting a spherical shell. The intersection of the $\partial G/\partial \nu$ curve with the $\tilde{\alpha}(\rho, \rho)$ curve in Fig. 5-11 also occurs very near the point $\nu = \rho + \hat{y}^{\dagger} K_{\rho}$. The point where the polarity change occurs for $\partial G/\partial \nu$ at its initial break point depends on the location of the center of the Gaussian distribution relative to ρ_o. For $\rho_w < \rho_o$ the initial break is negatively directed, but for $\rho < \rho_w$ the slope of $n(\rho)$ becomes negative and $\partial G/\partial \nu$ will then break positive. Physically, this regime where the variability of $G[\rho, \nu]$ approaches zero, see Eqs. (5.7-26) through-(5.7-28), corresponds to the Gaussian layer being located below the level, $\rho_o = \nu$; the layer cannot be "sensed" by a ray with an impact parameter $\nu > \rho_o$. We will show later that there are no stationary phase points in spectral number for this regime.

5.8.2 Accuracy of the Osculating Parameter Technique

To check the accuracy of the spectral representation used in this section, as given in Eq. (5.8-1), we again use the Airy layer model for a refracting medium with spherical stratification. We embed this layer in an otherwise homogeneous medium. These analytic solutions can be compared with the osculating parameter and numerical solutions. We let

$$\left.\begin{aligned} & n^2 = n_A^2 + 2n_A \frac{dn}{d\tilde{\rho}}\left(\tilde{\rho} - \rho_A\right), \quad \frac{dn}{d\tilde{\rho}} = \begin{cases} \text{constant}, & \tilde{\rho} \le \rho_A \\ 0, & \tilde{\rho} > \rho_A \end{cases} \\ & \tilde{\rho} = kn_A r, \quad \tilde{\rho}_A = \rho_A = kn_A r_A, \quad n_A = \text{constant} \end{aligned}\right\} \tag{5.8-4}$$

Here $\tilde{\rho} = \rho_A$ marks the upper boundary of the Airy layer, and n_A is the value of the index of refraction there. From Sections 4.10 and 4.11, we know that the wave equations for this case are given by

$$\frac{d^2\left(\tilde{\rho}U_l\right)}{d\tilde{\rho}^2} + \left(\left(\frac{n}{n_A}\right)^2 - \frac{l(l+1)}{\tilde{\rho}^2}\right)\left(\tilde{\rho}U_l\right) = 0, \quad \tilde{\rho}V_l = -i\frac{d\left(\tilde{\rho}U_l\right)}{d\tilde{\rho}} \tag{5.8-5}$$

$U(\tilde{\rho})$ describes the field along the normal to the plane of propagation (the y-direction in Fig. 4-10). For a given value of l, $U(\tilde{\rho})$ also provides the radial spectral component of the electric field, and $V(\tilde{\rho})$ describes the tangential component parallel to the plane of stratification and in the plane of propagation (the θ-direction in Fig. 4-10).

This model has been discussed in Section 4.12, where a correspondence between spherical and Cartesian stratification was established, and also in Section 5.7 to obtain an asymptotic form for $G[\rho, v]$ valid for $v > \rho$. There the solutions to the wave equations in Eq. (5.8-5) are given to a good approximation by the Airy functions with their argument $\tilde{y}$ given by Eq. (4.12-5). Thus, in the medium described by Eq. (5.8-4) for $r \le r_A$, we have

$$\left.\begin{aligned} & \tilde{\rho}U_l \doteq \left\{\text{Ai}[\tilde{y}], \text{Bi}[\tilde{y}]\right\}, \quad \tilde{\rho}V_l \doteq i\left|1-\beta_A\right|^{1/3} K_{\tilde{\rho}_A}^{-1}\left\{\text{Ai}'[\tilde{y}], \text{Bi}'[\tilde{y}]\right\} n_A, \\ & \tilde{\rho} = kn_A r, \quad \beta_A = -\left(\frac{1}{n}\frac{dn}{d\tilde{\rho}}\tilde{\rho}\right)_A, \\ & \tilde{y} = -\frac{1}{\left|1-\beta_A\right|^{2/3} K_{\tilde{\rho}_A}}\left(\left(1-\beta_A\right)\left(\tilde{\rho}-\tilde{\rho}_A\right) - \left(l-\tilde{\rho}_A\right)\right) \end{aligned}\right\} \tag{5.8-6}$$

For the region $r > r_A$, the solutions are the spherical Bessel functions of the first and second kinds, $\psi_l(\tilde{\rho}), \chi_l(\tilde{\rho})$. We note here that n_A is a constant in the Airy layer. Hence, $-\beta = (dn/d\tilde{\rho})\tilde{\rho}/n = (dn/dr)r/n$.

At the boundary $r = r_A$, we must match these two solution sets for each integer value of l to ensure the continuity of the tangential components of the electromagnetic field across the boundary. We first match the solutions when the $\mathrm{Ai}[\tilde{y}]$ solution applies in the region $r \le r_A$, that is, the solution that vanishes for decreasing $\tilde{\rho}$ with $\tilde{\rho} < l$. Upon noting from Eq. (5.8-6) that $\partial\tilde{y}/\partial\tilde{\rho} = -|1-\beta_A|^{1/3}/K_{\tilde{\rho}_A}$ (with $\beta_A < 1$), we set

$$\left.\begin{aligned} &\mathrm{Ai}[\tilde{y}_A] = c_l\psi_l(\tilde{\rho}_A) + d_l\chi_l(\tilde{\rho}_A) \\ &-K_{\tilde{\rho}_A}^{-1}(1-\beta_A)^{1/3}\,\mathrm{Ai}'[\tilde{y}_A] = c_l\psi_l'(\tilde{\rho}_A) + d_l\chi_l'(\tilde{\rho}_A) \end{aligned}\right\} \tag{5.8-7}$$

Using the asymptotic forms for both the spherical Bessel functions and the Airy function for $-\tilde{y} >> 1$ and for $-\hat{y} >> 1$, we obtain values for the matching coefficients that are given by

$$\left.\begin{aligned} &c_l = \frac{|1-\beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}}\cos(\tilde{X}_l - \hat{X}_l),\quad d_l = -\frac{|1-\beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}}\sin(\tilde{X}_l - \hat{X}_l), \\ &\tilde{X}_l = \frac{2}{3}(-\tilde{y}_A)^{3/2} + \frac{\pi}{4},\quad \hat{X}_l = \frac{2}{3}(-\hat{y}_A)^{3/2} + \frac{\pi}{4} \end{aligned}\right\} \tag{5.8-8}$$

Here $\tilde{y}$ is given in Eq. (5.8-6), and $\hat{y}$ is given by Eq. (5.4-3). Their ratio $\hat{y}/\tilde{y}$ is equal to $|1-\beta_A|^{2/3}$.

Similarly, for a $\mathrm{Bi}[\tilde{y}]$ solution in the region $r \le r_A$, that is, the solution that blows up for decreasing $\tilde{\rho}$ with $\tilde{\rho} < l$, we set

$$\left.\begin{aligned} &\mathrm{Bi}[\tilde{y}_A] = e_l\psi_l(\tilde{\rho}_A) + f_l\chi_l(\tilde{\rho}_A) \\ &-\frac{|1-\beta_A|^{1/3}}{K_{\tilde{\rho}_A}}\mathrm{Bi}'[\tilde{y}_A] = e_l\psi_l'(\tilde{\rho}_A) + f_l\chi_l'(\tilde{\rho}_A) \end{aligned}\right\} \tag{5.8-9}$$

When the asymptotic forms apply, that is, at a location well away from a turning point, we obtain

$$\left.\begin{aligned} e_l &= -\frac{|1-\beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}}\sin\left(\tilde{X}_l - \hat{X}_l\right) = d_l \\ f_l &= -\frac{|1-\beta_A|^{1/6}}{\sqrt{\pi K_{\tilde{\rho}_A}}}\cos\left(\tilde{X}_l - \hat{X}_l\right) = -c_l \end{aligned}\right\} \tag{5.8-10}$$

Referring to our discussion in Section 4.6 on incoming and outgoing waves in a Cartesian Airy layer, it follows in this case that outgoing and incoming waves at the boundary can be expressed in the form

$$\mathrm{Ai}\left[\tilde{y}_A\right] \mp i\,\mathrm{Bi}\left[\tilde{y}_A\right] = \left(c_l \mp i d_l\right)\xi_l^{\pm}\left(\rho_A\right) \tag{5.8-11}$$

The top sign applies to an outgoing wave, and the bottom sign to an incoming wave. Also, the complex coefficients $\left(c_l \mp i d_l\right)$ provide the phase delay (modulo 2π) between the incoming and outgoing waves at the boundary due to the refracting medium below. This phase delay offset remains invariant for $r > r_A$ because the medium is taken to be homogeneous for $r \geq r_A$. From Eqs. (5.8-8) and (5.8-10), it follows that the ratio

$$\frac{c_l - i d_l}{c_l + i d_l} = \exp\left(2i\left(\tilde{X}_l - \hat{X}_l\right)\right) \tag{5.8-12}$$

provides the round-trip phase delay between the incoming and outgoing wavelets of spectral number l at the boundary $r = r_A$. From Eqs. (5.8-6) and (5.8-8) and noting again that in the Airy layer $\hat{y} / \tilde{y} = |1-\beta_A|^{2/3}$, it follows that

$$\begin{aligned} \tilde{X}_l - \hat{X}_l &= \frac{2}{3}\left(\left(-\tilde{y}_A\right)^{3/2} - \left(-\hat{y}_A\right)^{3/2}\right) \\ &= \frac{2}{3}\cdot\frac{\beta_A}{1-\beta_A}\left(-\hat{y}_A\right)^{3/2} \doteq \frac{1}{3}\cdot\frac{\rho_A\beta_A}{1-\beta_A}\left(\frac{2(\rho_A - l)}{\rho_A}\right)^{3/2} \end{aligned} \tag{5.8-13}$$

We return to this expression momentarily after discussing the form of $G[\rho, \nu]$ for this model of the refracting medium.

Now we use the osculating parameter technique in the spectral representation given in Eq. (5.8-1) to describe the wave in the refracting Airy medium where $r \leq r_A$. The basis functions are the spherical Hankel functions $\{\xi_l^+(\rho), \xi_l^-(\rho)\}$, where $\rho = knr$ and n now is variable and given by Eq. (5.8-4). For a given spectral number, the form for the radial term from Eq. (5.8-1) is given by

$$a_l^- \xi_l^- + a_l^+ \xi_l^+ = C_l \left(\exp(-iG[\rho, \nu]) \xi_l^-(\rho) + \exp(iG[\rho, \nu]) \xi_l^+(\rho) \right) \qquad (5.8\text{-}14)$$

Here C_l is a spectral number-dependent complex factor that depends on boundary conditions, which we will discuss later in regard to turning points; it is not of interest here.

From Eq. (5.7-2), it follows that $G[\rho, \nu]$ for an Airy layer is given by

$$\left.\begin{aligned} G[\rho,\nu] &\doteq \frac{\pi}{6}\left(\frac{1}{n}\frac{dn}{d\rho}\rho\right)_A \left(\Gamma(\hat{y}) - \Gamma(\hat{y}_A)\right), \quad \rho^\dagger(\nu) < \rho \le \rho_A, \\ G[\rho,\nu] &\to \frac{2}{3}\left(\frac{1}{n}\frac{dn}{d\rho}\rho\right)_A \left((-\hat{y}_A)^{3/2} - (-\hat{y})^{3/2}\right), \quad \tilde{y}_A, \hat{y} << -1, \\ G[\rho,\nu] &\equiv 0, \quad \rho \ge \rho_A \end{aligned}\right\} \qquad (5.8\text{-}15)$$

Here $\Gamma(\hat{y})/3$ is the integral of $g(\hat{y})$ given in Eq. (5.7-1). $\Gamma(\hat{y})$ has been discussed in Section 4.9 with regard to osculating parameters in a Cartesian-stratified medium, and it is given by Eq. (4.9-5). It is shown in Fig. 5-12. $\Gamma(\hat{y})$ has only the two roots shown in this figure, and it is monotonic elsewhere.

For a given value of l in both Eqs. (5.8-13) and (5.8-15), what is the value of $\hat{y}$ in Eq. (5.8-15), that is, what value for ρ must we use to match the round-trip phase delay $-2G[\rho, \nu]$ with the asymptotically exact value of this phase delay $2(\tilde{X}_l - \hat{X}_l)$, given from Eq. (5.8-13)? Noting that n is variable in Eq. (5.8-15), it follows from Eq. (5.8-4) that

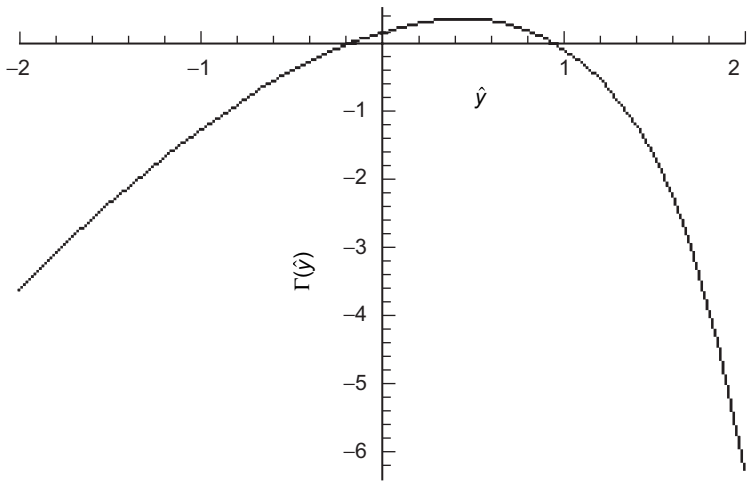

Fig. 5-12. Variation of $\Gamma(\hat{y})$ in the vicinity of a turning point.

$$\left(\frac{1}{n}\frac{dn}{d\rho}\rho\right)_A = -\frac{\beta_A}{1-\beta_A} \tag{5.8-16}$$

We see upon comparing Eq. (5.8-15) with Eq. (5.8-13) that $\hat{y}$ must be zero. However, the negative argument asymptotic forms for the Airy functions do not apply at $\hat{y}=0$; instead, we must set $\Gamma(\hat{y})=0$ in Eq. (5.8-15). This is discussed in Appendix I, Eq. (I-11), with regard to relating $dG[\rho^\dagger(\nu),\nu]/d\nu$ to $\tilde{\alpha}(\nu,\nu)$. The roots of $\Gamma(\hat{y})$ occur at $\hat{y}\doteq -0.2$ and $+0.9$, and $\Gamma(\hat{y})$ reaches a stationary value of almost zero at $\hat{y}=\hat{y}^\dagger$. It is nearly zero there compared to the value of $(-\hat{y}_A)^{3/2}$ when $-\hat{y}_A >> 1$ (see Fig. 5-12). Thus, we set $\rho=\rho^\dagger=\nu-K_\rho y^\dagger$ to force a near alignment of Eq. (5.8-15) with Eq. (5.8-13). The relative error of this choice is $\Gamma(\hat{y}^\dagger)/\Gamma(\hat{y}_A)\approx 0.001$ for this example.

In Section 5.5, Eqs. (5.5-22) through (5.5-25), we discussed the spectral coefficients a_l^+ for an outgoing wave in terms of $G[\rho,\nu]$. The equation for a_l^+ involved a constant of integration $a_l^+[\rho_*,\rho_*]$. The question of the appropriate value of ρ_* to use in this constant of integration was left rather moot in that discussion. Here we set $\rho_*=\rho^\dagger$ and

$$a_l^+\left[\rho^\dagger,\rho^\dagger\right] = i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG[\rho^\dagger,\nu]\right) \tag{5.8-17}$$

This provides a rationale for setting $\rho=\rho^\dagger(\nu)$ in $2G[\rho,\nu]$ for the outgoing wave, at least when well away from a turning point so that $\hat{y}_A<0$. We will return later to this question of linking the incoming and outgoing spectral coefficients.

5.8.3 Numerical Comparisons

Although the close comparison between $2G[\rho^\dagger,\nu]$ and $(c_l-id_l)/(c_l+id_l)$ at the boundary of the Airy layer is encouraging, we should compare the osculating parameter representation for the wave given in Eq. (5.8-14) with the exact solution obtained from integrating the wave equations in Eq. (5.8-5). Their level of agreement as a turning point is approached is of particular interest. Figure 5-13(a) shows the solution for $U_l(\rho)$, which is the solution to the wave equations in Eq. (5.8-5), and it also shows the osculating parameter solution from Eq. (5.8-14). They are virtually doppelgangers over the range shown for $\hat{y}$. In this example, $\nu=l+1/2=\rho_o=6400k$, $\rho_A=6432k$, $k=100\,\mathrm{km}^{-1}$, $n'=1.56\times10^{-4}$, and $n_o=1$. The solution is closely approximated by $\mathrm{Ai}[\tilde{y}]$. It yields a single-sided bending angle at

$\rho_* = \rho_o = 6400k$ of 10 mrad, and this medium has a moderate ray path curvature index β of about 0.1. The general procedure in Figs. 5-13 through 5-16 for comparing these solutions is as follows. For initial conditions, we set $U_l(\rho^\dagger) = \mathrm{Ai}[\tilde{y}^\dagger]$, and we also equate their derivatives at this point. This generates a numerical solution for $U_l(\rho)$ that exponentially damps to zero for decreasing ρ in the vicinity of the turning point at $\rho_o = \nu$ (but it blows up for $\rho << \rho_o$ because of limited numerical precision). Then at $\rho = \rho_A$, which in the example shown in Fig. 5-13 corresponds to a $\hat{y}$ value of about –50, we set $C_l(a_l^-(\rho)\xi_l^-(\rho) + a_l^+(\rho)\xi_l^+(\rho))$ equal to the numerical solution for $U_l(\rho)$ there, and we also equate their derivatives there. This sets the complex value of C_l in Eq. (5.8-14) and gives the osculating parameter solution for decreasing ρ.

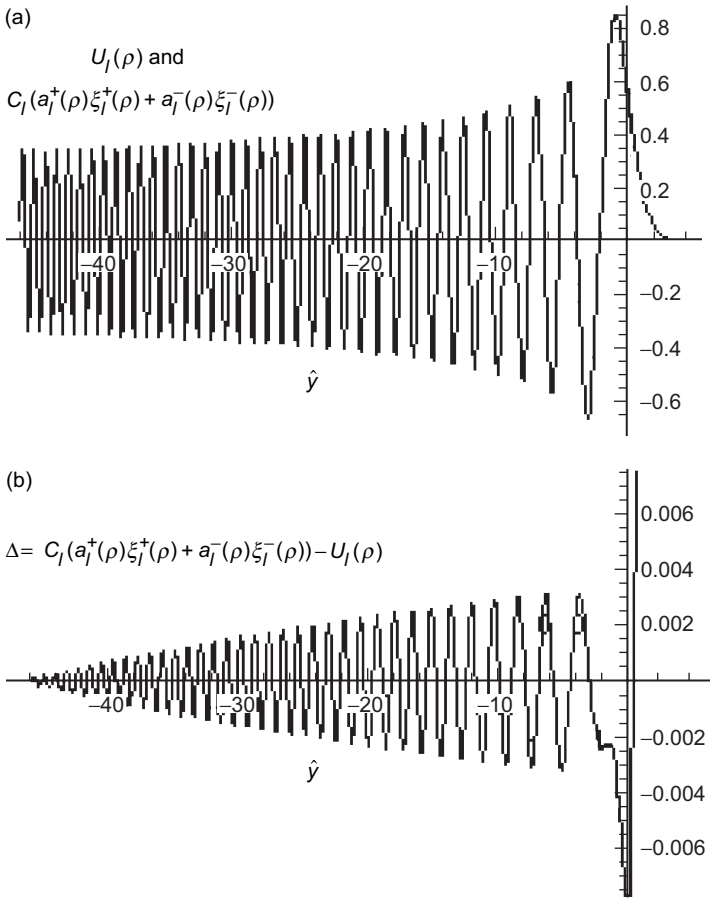

Fig. 5-13. Comparison of $U_l(\tilde{\rho})$ from a numerical integration of Eq. (5.8-5) in an Airy layer with the osculating parameter solution: (a) wave equation solutions and (b) the difference between the solutions.

Figure 5-13(b) shows the difference between $U_l(\tilde{\rho})$ and the osculating parameter form with the same initial conditions at the boundary $r = r_A$. Here $G[\rho, \nu]$ is obtained from Eq. (5.5-20), integrated on ρ over the Airy layer from ρ_A to ρ. In the osculating parameter solution for these two panels, $G[\rho, \rho_*]$ accumulates about 3 1/2 cycles between r_A and $r^\dagger$, where $r^\dagger = k^{-1}\rho^\dagger / n(\rho^\dagger)$. This accumulation of 3 1/2 cycles is required to keep the solution $C_l(a_l^- \xi_l^- + a_l^+ \xi_l^+)$ in Eq. (5.8-14) aligned in phase with the exact solution $U_l(\rho)$ over the entire Airy layer. Better than 1 percent numerical agreement holds between solutions except very near a turning point. As expected, for $\hat{y} >\sim -2$ the osculating parameter solution begins to deteriorate, but even at $\hat{y} = 0$ it still is moderately accurate for this example; the difference is 0.007. The differential equations in Eq. (5.8-5) become numerically unstable for $\hat{y} > 0$. Any small numerical errors in matching the boundary conditions or in the numerical integration will magnify greatly in the region $\hat{y} > 0$; that is, Bi[$\tilde{y}$] begins leaking into the numerical solution. See the discussion of the connection formula for the WKB solutions in Eq. (5.7-26). Matching boundary conditions closer to $\hat{y} = 0$, for example at $\hat{y} = -5$, improves the overall agreement, but the osculating parameter solution still deteriorates rapidly for $\hat{y} > 0$. But, the numerical solution for U_l also deteriorates for $\hat{y} > 0$ because of limited precision.

The wavenumber $k = 100$ km^{-1} used in Fig. 3-14 is more than 300 times smaller than the wavenumbers of the GPS navigation signals. The smaller value is used here to save computational time and to preserve computational accuracy—for given values of r_A and r_*, the number of cycles in the solution depends linearly on k. It is difficult to maintain coherency between solutions by matching both the solution and its derivative at the same point for highly oscillatory systems. The runoff error is roughly proportional to k. Small errors in the computation of U_l and $G[\rho, \nu]$ inevitably lead to runoff. An alternate matching strategy to reduce runoff is to pick two separate points and match solutions there, but not their derivatives.

Figure 5-14 shows another comparison between the numerical and osculating parameter solutions. Here a moderately strong refractivity gradient has been used, leading to single-sided bending-angle variations of more than 30 mrad over a tangency point altitude range of 8 km. Figure 5-14(a) shows the profile of the refractivity change, and Fig. 5-14(b) shows the resulting bending-angle profile. In the lowest layer, $\rho dn / d\rho = -0.8$, which gives a ray path curvature index of $\beta = 4/9$. But in the middle layer, $\rho dn / d\rho = +0.8$, which yields $\beta = -4$, a rather extreme negative ray path curvature. Figure 5-14(c) compares the two solutions, $U_l(\rho)$ and $a_l^-(\rho)\xi_l^-(\rho) + a_l^+(\rho)\xi_l^+(\rho)$. Here the

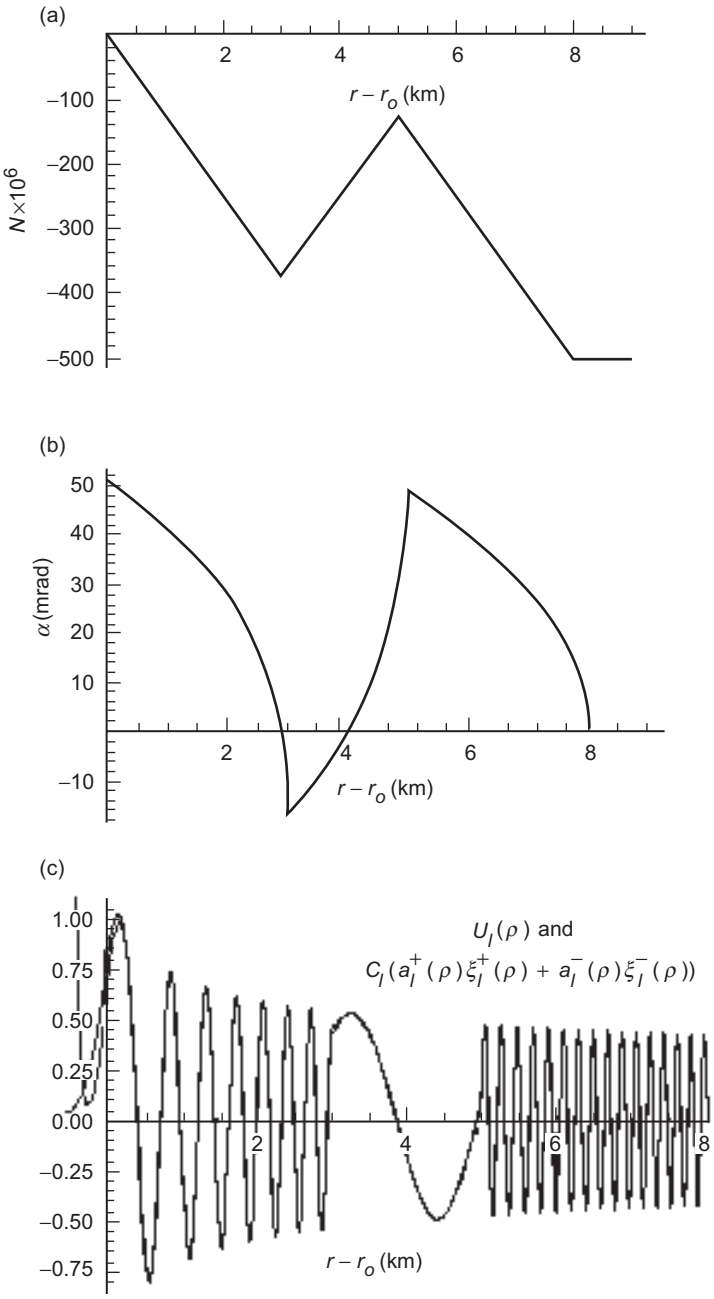

Fig. 5-14. Comparison of wave equation solutions in a strong refracting medium: (a) refractivity profile, (b) bending angle profile, and (c) wave equation solutions, $U_l(\rho)$ **and** $C_l(a_l^+(\rho)\xi_l^+(\rho) + a_l^-(\rho)\xi_l^-(\rho))$.

radial distance r, with a range of 8 km and with $k = 400\,\mathrm{km}^{-1}$, is used for the abscissa instead of $\hat{y}$. Per the discussion in Section 5.7 concerning Eqs. (5.7-25) and (5.7-26), by forcing the osculating parameter solution to follow the Airy layer solution for $\hat{y}$ values greater than ~ -2, roughly 1/2 km above the turning point for $k = 400\,\mathrm{km}^{-1}$ (~30 m for GPS wavelengths), one can greatly improve the solution below the turning point, until the numerical solution itself begins to fail.

Figure 5-15 compares the wave equation solutions for a severely refracting medium that includes a super-refracting layer, 1-km thick in the range $r_d < r < r_u$. The refractivity profile is shown in Fig. 5-15(a). Within the super-refracting layer, the ray path curvature index has a value of $\beta = 2$. Above the layer, $\beta = 0.4$, and below it, $\beta = 0.3$. These refractive gradients lead to enormous swings in the ray path bending angle. Figure 5-15(b) shows the resulting single-sided bending-angle profile versus tangency point radius, including the super-refracting zone $r^c < r_* < r_u$. For tangency points within this range, no rays can occur when spherical symmetry applies. Figure 5-15(c) shows the variation in impact parameter in the vicinity of this super-refracting layer. The impact parameter has a negative slope within the layer. Figure 5-15(d) compares the wave equation solutions. In this example, the turning point is well below the critical radius, i.e., $r_o < r^c$. Good agreement holds except near the turning point. A wavenumber value of $k = 1000\,\mathrm{km}^{-1}$ is used in Figs. 5-15 and 5-16. The wave equation solutions in Fig. 5-15(d) show a slight amplitude bulge and a slight reduction in frequency in the vicinity of the super-refracting layer, which is expected. Referring to the WKB solution connection formulas in Eq. (5.7-26), the amplitude and frequency factor expressed in length units is $f_l(\rho) \doteq -k^2 \hat{y} / K_v^2 \approx -46\hat{y}$ km^{-2}. When $\beta < 1$, $\partial \hat{y} / \partial r < 0$; but in a super-refracting medium where $\beta > 1$, then $\partial \hat{y} / \partial r > 0$. It follows that $\partial f_l / \partial r > 0$ for $r < r_d$; but within the super-refracting layer, $r_d < r < r_u$, $\partial f_l / \partial r < 0$. For a turning point below the critical tangency point, i.e., for $\nu < kn(r^c) r^c$, $f_l(\rho)$ remains positive for all $r > k^{-1}\nu / n(r)$, but it becomes less positive with increasing r within the super-refracting layer, which is reflected in the figure by the increasing amplitude ($\sim 1/\sqrt[4]{f_l}$) and decreasing frequency ($\sim \sqrt{f_l}$) within the layer. Above the layer, $\partial f_l / \partial r$ returns to positive territory. When $\nu < kn(r^c) r^c$, the lth spectral component of the wave effectively powers through the super-refracting layer, as shown in Fig. 5-15(d). The corresponding ray, with an impact parameter value $\rho_* = \nu \le kn(r^c) r^c = kn(r_u) r_u$, also powers through, although it is severely refracted as $\rho_* \to kn(r^c) r^c$ from below.

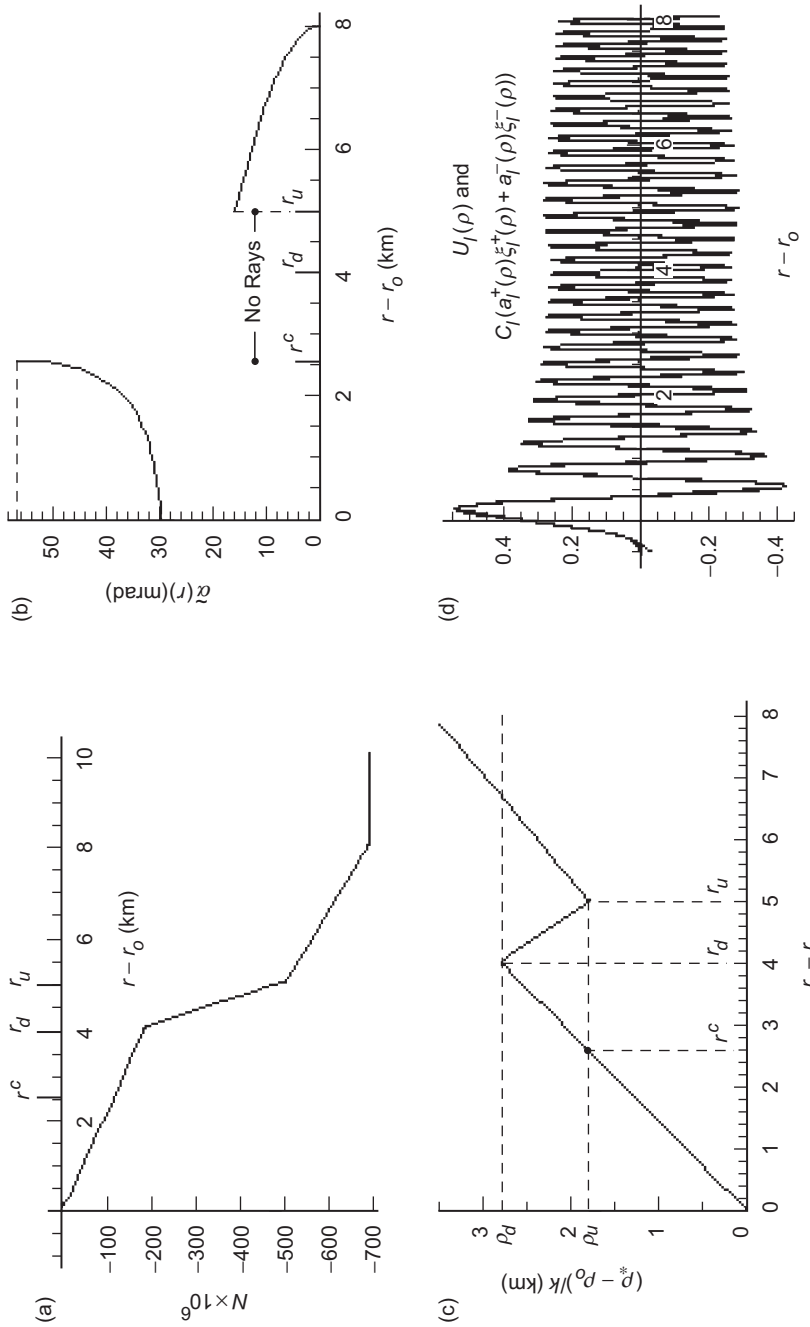

Fig. 5-15. Comparison of wave equation solutions in a severely refracting medium with a super-refracting layer: (a) refractivity profile, (b) bending angle profile, (c) impact parameter profile, and (d) wave equation solutions; spectral number $\nu = kn(r_o)r_o < kn(r^c)r^c$.

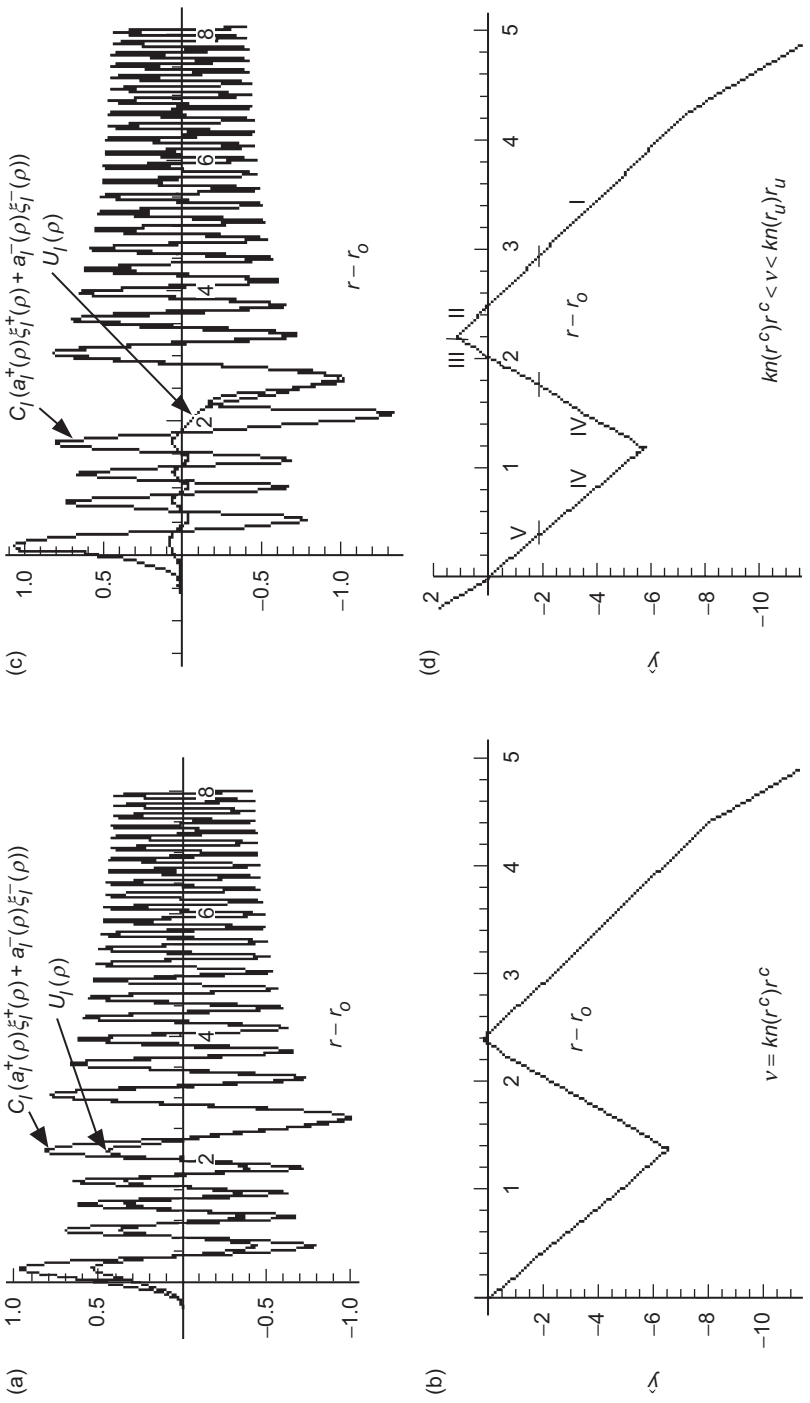

Fig. 5-16. Comparison of wave equation solutions with the spectral number located in the super-refracting zone: (a) wave equation solutions with $\nu = kn(r^c)r^c$, (b) profile for $\hat{y}$ for (a), (c) wave equation solutions with $kn(r^c)r^c < \nu < kn(r_u)r_u$, and (d) profile for $\hat{y}$ for (c).

As the ray path tangency point approaches the critical radius $r_* = r^c$, the bending angle approaches a limit, shown in Fig. 5-15(b), and the wave solution shows a marked transition from an oscillatory form to an exponential form in the vicinity of the layer. A delicate situation occurs when the turning point is such that $kn(r^c)r^c \le \nu \le kn(r_d)r_d$. Here $f_l(\rho)$ reverses sign at $kn(r)r = \nu$, with the lower root to the equality lying in the range $r^c < r < r_u$, and the upper root at $r > r_u$. In this case, we have two zones where $U_l(\rho)$ must be exponential-like: in that region in and above the super-refracting layer where $f_l < 0$, and below the turning point where $f_l < 0$ again. Inasmuch as there are no rays for tangency points in the range $r^c < r_* < r_u$, we expect the wave equation equations to give a damped amplitude for the field.

Lastly, Fig. 5-16 shows the case where the spectral number has been increased (actually, the refractivity profile was lowered in altitude relative to a fixed turning point) so that the spectral number lies in the range $kn(r^c)r^c \le \nu = kn(r_o)r_o \le kn(r_d)r_d$. This figure shows the wave equation solutions and the profile for $\hat{y}$. Figures 5-16(a) and 5-16(b) are for the case where $\nu = kn(r^c)r^c = kn(r_u)r_u$, the critical impact parameter value marking the boundaries of the zone corresponding to the altitude range $r^c < r_* < r_u$, within which no ray path tangency points may lie. Figures 5-16(c) and 5-16(d) are for the case where the spectral number lies within the critical range $kn(r_u)r_u < \nu < kn(r_d)r_d$, the equivalent of about 150 m in radial distance above r^c for $k = 1000\,\mathrm{km}^{-1}$. Figures 5-16(a) and 5-16(c) were normalized so that the amplitudes of all of the wave equation solutions in Figs. 5-16(a) and 5-16(c) have the same asymptotic value with increasing r well above the atmosphere. Here the agreement between $U_l(\rho)$ and the osculating parameter solutions degrades significantly in the vicinity and below the super-refracting layer because $\hat{y}$ becomes positive over a limited interval about the upper boundary at $r = r_u$. It also becomes positive again below the turning point. Although the amplitudes of the two solutions diverge below the layer, their phases remain aligned right down to the turning point.

A super-refracting layer acts like a second turning point when the spectral number lies in the critical zone $kn(r^c)r^c \le \nu \le kn(r_d)r_d$. We note from Eq. (5.7-26) that the slightest hint of a positive value for $\hat{y}$ in the super-refractive zone ($\sqrt{-f_l} \doteq k\sqrt{\hat{y}} / K_\nu \approx 7\sqrt{\hat{y}}$) causes $U_l(\rho)$ to damp to near-zero below the lower critical point where $\hat{y} = 0$. The easiest way to see this is to consider the reverse sense of propagation. Consider a solution for $U_l(\rho)$ that is forced to damp to zero below the turning point at $\rho = \rho_o$. With these initial conditions on $U_l(\rho)$ and its derivative, the solution then is propagated outward.

When the region above the turning point where $f_l(\rho)$ first becomes negative is encountered, the solution becomes exponential-like. This causes the amplitude of $U_l(\rho)$, which still has an Airy function-like character (of the first kind) below this point, to be catapulted to an astronomical magnitude if $f_l(\rho)$ remains negative for a sufficient interval.[10] But we have normalized the asymptotic values of the outbound amplitudes to correspond to the amplitude of an incoming wave, which is essentially invariant over the narrow range of spectral numbers considered here. This forces the amplitude of $U_l(\rho)$ in the region below the lower point at which $f_l(\rho) = 0$ to be greatly diminished when $\hat{y}_{\text{MAX}} = \hat{y}_u$ is positive. In Fig. 5-16(a), $\hat{y}_{\text{MAX}} \doteq 0$, and the amplitude of $U_l(\rho)$ below the layer is beginning to attenuate. But in Fig. 5-16(c), $\hat{y}_{\text{MAX}}$ is barely positive, 1.1, and $f_l(\rho)$ is negative only over a 40-m interval about $r = r_u$. Already for these slight intrusions the amplitude of $U_l(\rho)$ is an order of magnitude diminished below the layer. For GPS wavelengths, $\Delta\hat{y} = 1$ corresponds to $\Delta r = 14\,\text{m}$ and $\sqrt{-f_l} \approx 70\sqrt{\hat{y}}$. Therefore, the cutoff in amplitude of $U_l(\rho)$ below a super-refracting layer for spectral numbers greater than the limiting value yielding $\hat{y}_{\text{MAX}} = 0$ is extremely sharp.

The lower point for $r < r_u$ in the super-refracting layer where $\hat{y} = 0$ is also the first point where the bending-angle integral for a hypothetical ray with an impact parameter value of $\rho_* = \nu$ would become complex, if such a ray were to exist. Here $kn(r)r$ becomes smaller than ρ_* for increasing r above this point until the upper point where $\hat{y} = 0$ is reached at $r < r_u$.

A better approximation approach (but not used here) for the osculating parameter solution to obtain closer alignment with $U_l(\rho)$ in and below the super-refraction zone is to break the spherical medium into regimes. In Fig. 5-16(d), these regimes are designated by Roman numerals, and they have boundaries in the radial coordinate r at the points where $\hat{y} \approx -2$. In I and IV, the osculating parameter solution applies because $\hat{y} < -2$ throughout these regimes. In each of the tunneling regimes, II, III, and V, the refractivity profile

---

[10] An easy refractivity model with which to see this amplification process is given by $n' = \text{constant} < -157 \times 10^{-6}\ \text{km}^{-1}$ in the super-refracting layer $r_d \le r \le r_u$, and $n' \equiv 0$ elsewhere. The wave equation solutions are: below the layer, $\text{Ai}[\hat{y}]$; within the layer, $a\text{Ai}[\tilde{y}] + b\text{Bi}[\tilde{y}]$; and above the layer, $c\text{Ai}[\hat{y}] + d\text{Bi}[\hat{y}]$. Equating these solutions and their derivatives across their respective boundaries, at $r = r_d$ and at $r = r_u$ where $\hat{y} = \hat{y}_{\text{max}}$, leads to an explicit evaluation of the coefficients c and d. It can be shown that the amplitude $|\, c^2 + d^2\,|^{1/2} \to 2(\beta - 1)^{1/6} \exp(2(\beta/(\beta - 1))\hat{y}_{\text{max}}^{3/2}/3)$, when $\hat{y}_{\text{max}} > 0$. The extreme case $n' \to -\infty$ with Δn finite is addressed in Chapter 3 (Mie scattering theory) and in Section 5.13.

is approximated by an Airy layer, which has a wave equation solution given by Eq. (5.8-6). Since $\hat{y}$ need not be too positive before the solution below the layer is essentially damped to zero, this Airy layer approximation should be valid. The coefficients of these different solutions are tied together by using the continuity conditions on the solutions and their derivatives across each boundary; these boundaries are marked in Fig. 5-16(d). The amplitude of the osculating parameter solution in IV then will damp to zero rapidly below the layer, which $U_l(\rho)$ also does, except for spectral numbers in the transition zone where $\hat{y}_{\text{MAX}} \approx 0$. On the other hand, in region IV below the super-refraction zone, $G[\rho, \nu]$ contains valid phase information because $\hat{y} <\sim -2$.

In summary, Figs. 5-13 through 5-16 show good agreement between the exact and osculating parameter solutions over almost all regions except those in the immediate neighborhood of a turning point or in that delicate transition across the critical spectral number range $\nu \approx kn(r^c)r^c$, below a super-refracting layer. Here $\hat{y}_{\text{MAX}} \approx 0$, but it is not yet positive enough to rapidly damp $U_l(\rho)$ to zero below the layer. In calculating the electric field from the integrals over spectral number that are given in Eq. (5.8-1b), Section 3.16 shows that the principal contribution to these integrals comes from a neighborhood in spectral number where the phasor in the integrand is varying the least. This is a stationary phase neighborhood, which may or may not be unique, depending on the refractivity profile. If the field is being evaluated at an incoming point well away from a turning point, then a stationary phase neighborhood will not be located near $\hat{y} = 0$, but rather in negative territory. Therefore, except for incoming positions near a turning point, we expect the accuracy of the osculating parameter technique applied to the spectral integrals in Eq. (5.8-1b) to be adequate. Moreover, near a turning point, one can use the Airy layer approximation to greatly improve the accuracy there, which is discussed in Section 5.11. The issue of outgoing points remains, and it is discussed later.

5.8.4 Comparison of Phase Delays in an Airy Layer from Wave Theory and Geometric Optics

According to geometric optics, the single-sided phase delay for an initially collimated incident ray with an impact parameter $\rho_* \le \rho_A$ is given by

$$\begin{aligned} ck\tau = k\int_{r_*}^{r_A} \frac{n^2 r dr}{\sqrt{n^2 r^2 - n_*^2 r_*^2}} &= -\int_{\rho_*}^{\rho_A} \frac{d\log n}{d\rho}\sqrt{\rho^2 - \rho_*^2}\, d\rho \\ &+ \sqrt{\rho_A^2 - \rho_*^2} + \rho_*\left(\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_A, \rho_*)\right) \end{aligned} \tag{5.8-18}$$

The second equality follows from integrating by parts and using Eq. (5.6-2) for the bending angle. Using Bouguer's law, we see that the bottom line simply

equals $\rho_A \cos\theta_A \sec(\Delta\tilde{\alpha})$ through second order in $\Delta\tilde{\alpha} = \tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_A, \rho_*)$. Therefore, the extra single-sided phase delay caused by the refractive gradient for a ray with an impact parameter $\rho_* \le \rho_A$ is given by

$$ck\Delta\tau = -\int_{\rho_*}^{\rho_A} \frac{d\log n}{d\rho}\sqrt{\rho^2 - \rho_*^2}\, d\rho \doteq \frac{\beta_A \rho_*}{3(1-\beta_A)}\left(\frac{2(r_A - r_*)}{r_*}\right)^{3/2} \quad (5.8\text{-}19)$$

The approximate expression on the RHS applies to the Airy layer model, and it is accurate when thin-atmosphere conditions apply and $r_*/\lambda >> 1$ (see Appendix A). This expression also gives about 3 1/2 cycles for the conditions shown in Fig. 5-13, and it matches the phase delay expression given in Eq. (5.8-13).

Equation (5.8-13), which is an asymptotically exact result for an Airy layer, Eq. (5.8-15), which gives $G[\rho, \nu]$ from use of the osculating parameter technique, and Eq. (5.8-19), which is from geometric optics, all essentially agree on the phase induced by the refractive gradient on a wave passing through the Airy layer. We return to this model in Section 5.11, where calculating the field at a turning point is discussed.

Equation (5.8-19) provides us with further insight into the character of $G[\rho, \nu]$. When $\nu^* = \rho_*$, $G[\rho, \rho_*]$ provides the extra path delay for a ray with an impact parameter value of ρ_*, both from the curvature component of the ray that is induced by the refractive gradient and from the change in wave velocity along the ray in the medium.

5.8.5 Asymptotic Matching the Spectral Coefficients for Incoming and Outgoing Waves

Let (r_*, θ_*) mark a turning point and $\rho_* = kr_* n_*$ be the impact parameter associated with the ray passing the turning point. We know from Bouguer's law for a spherical symmetric medium that ρ_* is a constant when evaluated along that ray path.

When the planar asymptotic boundary conditions apply to the approaching wave, then the incoming spectral coefficient $a_l^-(\rho)$ is uniquely determined from Eq. (5.5-21). It is independent of the impact parameter ρ_* associated with the point (r, θ) at which $a_l^-(\rho)$ is evaluated because the asymptotic boundary conditions in Eq. (5.5-3) for a planar approaching wave are independent of impact parameter. Because the wave front surface is not symmetric about the turning-point boundary (the approaching waves are collimated prior to impacting the medium; the departing waves are dispersed), the outgoing coefficient depends on the value of ρ_*. For the outgoing wave, it follows from Eq. (5.5-23) that we can obtain $a_l^+[\rho, \rho_*]$ if we know the value of $a_l^+[\rho, \rho_*]$.

Consider first the symmetric problem mentioned earlier, where the electromagnetic wave is planar along the line $\theta = \pi/2$, $\phi = 0$, i.e., along the x-axis in Fig. 4-10. From Bauer's identity in spherical coordinates, we have along the x-axis

$$1 = \sum_{l=1}^{\infty} i^l (2l+1)\left(\frac{\psi_l(\rho)}{\rho}\right) \mathrm{P}_l(0) \tag{5.8-20}$$

which holds for all values of ρ. As the wave propagates away from the line $\theta = \pi/2$, which is along the x-axis, the cophasal normal path in the plane $\phi = 0$ will depend on its initial position ρ along the x-axis from which it started. Thus, for every value $\rho = \rho_*$, the spectral coefficients must have the form

$$a_l^{\pm}[\rho, \rho_*] = i^{l-1} \frac{2l+1}{l(l+1)} \exp\left(\pm i\left(G[\rho, \nu] - G[\rho_*, \nu]\right)\right) \tag{5.8-21}$$

Using these forms for the spectral coefficients in the scalar potentials given in Eq. (5.8-1), it is readily shown with the Helmholtz equation in Eq. (5.2-6) that the radial component of the electric field is invariant along the x-axis and is given by the RHS of Eq. (5.8-20).

To convert this symmetric form for the electromagnetic wave into a non-symmetric form, that is, the version where the wave is asymptotically collimated on the incoming side ($\pi > \theta > \theta_*$), we merely multiply both a_l^- and a_l^+ in Eq. (5.8-21) by $\exp\left(-iG[\rho_*, \nu]\right)$ (thus preserving their equality at $\rho = \rho_*$). We will show later that this is equivalent to rotating clockwise each cophasal normal path by the angle $\tilde{\alpha}(\rho_*, \rho_*)$, where ρ_* is the impact parameter of the path. For this case, as $\rho \to \infty$

$$\left.\begin{aligned} a_l^- &\to i^{l-1} \frac{2l+1}{l(l+1)} \\ a_l^+ &\to i^{l-1} \frac{2l+1}{l(l+1)} \exp\left(-2iG[\rho_*, \nu]\right) \end{aligned}\right\}, \quad \rho \to \infty \tag{5.8-22}$$

The question arises as to what value to use for ρ_* in these expressions in either Eq. (5.8-21) or Eq. (5.8-22). Consider the variability of $a_l^-(\rho)$ with ρ for the collimated case when the spectral number is set to a fixed value ν_o. At a point (r, θ) on the approaching side at large distances where $n \to 1$, we set $\nu_o = kr\sin\theta$; that is, ν_o becomes the impact parameter for the cophasal normal passing through the point (r, θ). From Eq. (5.5-21), the evolution of $a_l^-(\rho)$

with ν while traveling along this particular ray path can be obtained by studying the behavior of $G[\rho,\nu]$ with ν fixed. Figure 5-6 shows an example of $G[\rho,\nu_o]$ versus ρ in the close vicinity of $\rho=\nu_o$. $G[\rho,\nu]$ has a stationary value at $\rho=\nu_o-\hat{y}^\dagger K_\rho$, where $\hat{y}^\dagger$ is given by Eq. (5.7-10). This point marks a stationary phase point for $a_l^-(\rho)$ [and for $a_l^+(\rho)$] with respect to ρ. We know that the forms for $G[\rho,\nu]$ given in Eq. (5.5-20) or Eq. (5.7-2) begin to fail for increasing $\hat{y}>0$, and that the correct form for $G[\rho,\nu]$ rapidly assumes a constant value. Figure 5-7 shows a comparison between the exact phase rate $\partial\vartheta_l^-/\partial\rho$ and $\partial G[\rho,\nu]/\partial\rho$ for an Airy layer.

Consistent with the approximation chosen in Eq. (5.7-27) for $G[\rho,\nu]$, we now set $a_l^+=a_l^-$ at the point where $\hat{y}=\hat{y}^\dagger$ to ensure no singularity at the origin from the Hankel functions and to attain a close match with the refractive bending angle. Note that this stationary phase point in ρ space varies with the value of ν_o. At this stationary point, $\rho^\dagger=\rho^\dagger(\nu)=\nu-K_{\rho^\dagger}y^\dagger$, we set

$$a_l^+\left[\rho^\dagger,\rho^\dagger\right]=a_l^-\left(\rho^\dagger\right)=i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG\left[\rho^\dagger,\nu\right]\right) \tag{5.8-23}$$

This is exactly the same form for $a_l^+[\rho_*,\rho_*]$ that we obtained in Eq. (5.8-17) at the boundary of an Airy layer to force alignment between the round-trip delay obtained from the osculating parameter technique with the round-trip delay from the Airy layer solution. Here $G[\rho,\nu]$ would include the delay from the Airy layer and from the overlying medium. Of course, our Airy layer can be made as thin as we please relative to the extent of the overlying medium. Thus, we asymptotically match the incoming and outgoing spectral coefficients.

From Eq. (5.5-23) it follows that $a_l^+(\rho)$ is given by

$$\begin{aligned}a_l^+(\rho)&=a_l^+\left[\rho^\dagger,\rho^\dagger\right]\exp\left(-i\left(G\left[\rho^\dagger,\nu\right]-G[\rho,\nu]\right)\right)\\&=i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-i\left(2G\left[\rho^\dagger(\nu),\nu\right]-G[\rho,\nu]\right)\right)\end{aligned} \tag{5.8-24}$$

At the LEO, which is located at $(r_{\mathrm{L}},\theta_{\mathrm{L}})$ and assumed to be outside of the refracting medium where $n\equiv 1$, the second term in Eq. (5.8-24) is zero; that is, $G[\rho_{\mathrm{L}},\nu]\equiv 0$. It follows that

$$\left.\begin{aligned}a_l^+(\rho_{\mathrm{L}})&=i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-i2G\left[\rho^\dagger(\nu),\nu\right]\right)\\\rho^\dagger&=\nu-y^\dagger K_{\rho^\dagger}\end{aligned}\right\} \tag{5.8-25}$$

Thus, $-2G^{\dagger}(\nu) \equiv -2G[\rho^{\dagger}(\nu),\nu]$ is the two-way phase accumulation of the l th spectral coefficient from the stationary point $\rho^{\dagger} = kr^{\dagger}n(\rho^{\dagger})$ to outside of the refracting medium. *It corresponds in geometric optics to the extra phase accumulation induced by the refractivity gradient while traveling along a ray that has completely transected the atmosphere and that has an impact parameter value of* ν.

We will show later that the value of ν that yields a stationary phase in the wave theory spectral representation in Eq. (5.8-1) is essentially equal to the value of the impact parameter for the ray passing through the LEO provided that the inequality $\rho_* \mid d^2 n / d\rho_*^2 \mid << \tilde{\alpha}(\rho_*, \rho_*)$ is satisfied. The latter value is given from Bouguer's law by

$$\nu = \rho_* = kr_* n_* = kr_{\mathrm{L}} \sin\left(\theta_{\mathrm{L}} + \alpha_{\mathrm{L}}(\rho_*)\right) \tag{5.8-26}$$

where $\alpha_{\mathrm{L}}(\rho_*) = 2\tilde{\alpha}(\rho_*, \rho_*)$ is obtained from Eq. (5.6-5). When multipath situations occur, this value may not be unique; we will discuss that later. It follows that the stationary point for the spectral number in wave theory is given by

$$\nu = \nu^{\dagger} = \rho_* + \hat{y}^{\dagger} K_{\rho_*} \tag{5.8-27}$$

A continuing issue for outgoing points concerns the accuracy of the adopted form in Eq. (5.7-27) for $G[\rho,\nu]$ and, therefore, the accuracy of $-2G[\rho^{\dagger}(\nu),\nu]$ in representing the round-trip phase delay. Because of the failure of the osculating parameter technique for decreasing $\rho < \rho^{\dagger}$, this adopted form has some error. But the adopted form for $G[\rho,\nu]$ does have a stationary value at $\rho = \rho^{\dagger}$, and the correct form must rapidly approach a fixed value for $\rho < \rho^{\dagger}$ (see Fig. 5-7). In addition, we have already noted in Section 5.7 the close correspondence between $G[\rho^{\dagger}(\nu),\nu]$ and the component of the phase delay in the eikonal equation induced by the refractivity gradient for a ray with an impact parameter value of ν. Also noted there and in Appendix J is the high accuracy of the relationship, $dG[\rho_*^{\dagger}(\rho_*),\rho_*] / d\rho_* \doteq \tilde{\alpha}(\rho_*, \rho_*)$, provided that the curvature in the refractivity profile is moderate. In this section, the close comparison between $-2G[\rho^{\dagger}(\nu),\nu]$ and the exact solution for the round-trip phase delay in an Airy layer has been noted. Moreover, a byproduct of the numerical solutions presented in Figs. 5-13 through 5-16 (along with the second numerical solution for $U_l(\rho)$ with the boundary conditions generated from $\mathrm{Bi}[\tilde{y}^{\dagger}]$) is the exact value of the extra phase delay induced by the refractivity gradient of the medium. $G[\rho^{\dagger}(\nu),\nu]$, based on Eq. (5.7-27) and calculated from Eq. (5.7-2), may be directly compared to this numerical result for the different

refractivity models assumed in these examples. The agreement is generally in the range of 0.5 to 0.1 percent when the curvature index β is not too close to unity. The computational imprecision in these results also is of the order of 0.1 percent. Section 6.5 compares the adopted form for $dG[\rho^{\dagger}(\nu),\nu]/d\nu$, a key spectral quantity in the recovery of the refractivity profile from the LEO amplitude and phase observations, with the exact form for an Airy layer. It gives the error as a function of the ray path curvature index.

An alternative rationale for picking the form given in Eq. (5.8-24) for a_l^+ is based on probabilistic arguments derived from summing over all possible ray paths. The rationale is similar to the Feynman sum-over-histories technique in quantum electrodynamics to calculate the probability of a quantum event. On the incoming side, for an initial planar wave, the probability density distribution of impact parameter values for the rays is a flat curve; each value is equally likely to occur. Bouguer's law requires that the flat distribution be preserved after atmospheric encounter. Therefore, the values of ρ_* at a turning point are uniformly distributed. If we set the outgoing spectral coefficient to be the spatial average over all possible impact parameter values to be used in Eq. (5.8-22), we end up with an averaging integral to evaluate. The stationary phase value of this integral yields $\rho_* = \rho^{\dagger}(\nu)$. Therefore, the stationary value of $G[\rho,\nu]$ at $\rho = \rho^{\dagger}(\nu)$ should be adopted in Eq. (5.8-22).

5.9 Interpreting Wave Theory in a Refracting Medium Using the Stationary Phase Technique

We now return to the wave theory spectral representation in Eqs. (5.8-1a) and (5.8-1b) for the electric field at a point (r,θ) in the refracting medium. Before presenting numerical examples using this spectral representation, we apply stationary phase concepts to aid in the interpretation of those numerical results to compare with geometric optics. We follow Chapter 3 closely here; in particular, we refer to the material in Sections 3.10 through 3.13. The integrals in Eq. (5.8-1b) are characterized by a slowly varying factor multiplied by the sum of four phasors that are rapidly oscillating over most of the spectral number space. The main contributions to these integrals come from neighborhoods where any one of these phasors is varying the least. Our task now is to use the stationary phase technique on these integrals to identify the possible stationary phase neighborhoods for each phasor and to calculate the values for E_r and E_θ at a given point (r,θ).

5.9.1 Geometric Interpretation of the Phasors

We rewrite the asymptotic forms for the Hankel and Legendre functions in Eq. (5.8-1) in a phasor form that provides a useful geometric interpretation.

Except for a point (r,θ) located very near a turning point, we will show that the stationary phase neighborhoods in ν space are sufficiently below the value $\rho = krn(\rho)$ that the negative argument asymptotic forms for the Airy functions given in Eq. (3.8-7) can be used. In this case, from Eqs. (3.10-1) through (3.10-4), the spherical Hankel functions can be rewritten as

$$\left.\begin{aligned} \xi_l^{\pm}(\rho) &\sim \left(\frac{\rho}{D_\nu}\right)^{1/2} \exp\left(\pm i\left(D_\nu + \nu(\theta_\nu - \pi/2) - \pi/4\right)\right) \\ \xi_l^{\pm\prime}(\rho) &\sim \left(\frac{D_\nu}{\rho}\right)^{1/2} \exp\left(\pm i\left(D_\nu + \nu(\theta_\nu - \pi/2) - \pi/4\right)\right) \end{aligned}\right\} \qquad (5.9\text{-}1)$$

where

$$D_\nu = \sqrt{\rho^2 - \nu^2}, \quad \theta_\nu = \sin^{-1}\left(\frac{\nu}{\rho}\right), \quad \rho = krn(\rho), \quad \nu < \rho \qquad (5.9\text{-}2)$$

Similarly, from Eq. (3.10-5), the asymptotic form for the Legendre polynomial $\mathrm{P}_l^1(\cos\theta)$ is given by

$$\mathrm{P}_l^1(\cos\theta) \sim -\sqrt{\frac{l}{2\pi\sin\theta}}\left(\exp\left[i\left(\nu\theta + \frac{\pi}{4}\right)\right] + \exp\left[-i\left(\nu\theta + \frac{\pi}{4}\right)\right]\right) \qquad (5.9\text{-}3)$$

The geometric interpretation of θ_ν as the central angle to the point (ρ,θ_ν) and with D_ν as its tangential distance in phase units from a spherical caustic surface of radius ν has been given in Chapter 3, Eqs. (3.11-1) and (3.11-2), and in Figs. 3-14 and 3-15. Those figures are applicable to an outgoing wave, but the concepts are the same, whether incoming or outgoing. Figure 3-14 shows the geometric relationships given by Eq. (5.9-2). Two rays, originally collimated from the GPS direction, reach the point L located at (ρ,θ). The direct path is a straight path to the tangential point P_1 and then straight on to the point L. Along the retrograde path via P_2, the ray arrives tangentially at P_2, travels along the arc $\mathrm{P}_2\,\mathrm{P}_2'$, and then departs tangentially from P_2' going straight on to the point L. We note that, when it is assumed that the stationary phase value ν^* lies in the range $0 < \nu^* < \rho$, it follows that $0 < \theta_{\nu^*} < \pi/2$. As ν increases through its range of values in the spectral integrals in Eq. (5.8-1b), the radius of this caustic sphere expands, and its center descends.

The four phasors appearing in the integrals in Eq. (5.8-1b) result from the product of the spherical Hankel and Legendre functions in Eqs. (5.9-1) and (5.9-3) times the incoming and outgoing spectral coefficients given in

Eqs. (5.5-21) and (5.8-24), respectively. To further interpret these phasors geometrically, we let

$$\left.\begin{array}{l}\text{outgoing: } \theta = \theta_\nu + \Delta\theta_\nu \\ \text{incoming: } \theta = \pi - (\theta_\nu + \Delta\theta_\nu)\end{array}\right\} \tag{5.9-4}$$

Because the stationary neighborhoods for ν will be close to $\nu^* = \rho\sin(\theta+\alpha)$, we would expect $\Delta\theta_\nu$ to be small in these neighborhoods. Figure 3-15 provides a geometric interpretation of $\Delta\theta_\nu$ in terms of the extra phase $\nu\Delta\theta_\nu$ along the $\theta = 0$ direction that results from the offset $\theta - \theta_\nu$ for an outgoing wave (and $\theta - (\pi - \theta_\nu)$ for an incoming wave).

5.9.2 Stationary Phase Conditions

We now insert the asymptotic forms in Eqs. (5.9-1) through-(5.9-3) into Eq. (5.8-1b), and we substitute the forms for the spectral coefficients a_l^- and a_l^+ given by Eqs. (5.5-21) and (5.8-24), respectively. After some manipulation of Eq. (5.8-1b), the spectral representation for $E_r(r,\theta)$ and $E_\theta(r,\theta)$ becomes

$$\left.\begin{array}{l}E_r = \\ \dfrac{E_o}{\sqrt{2\pi\rho\sin\theta}}\displaystyle\int_0^\infty \left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu}\right)^{\frac{1}{2}}\left(e^{i\Psi(+,+)} + e^{i\Psi(+,-)} + e^{i\Psi(-,+)} + e^{i\Psi(-,-)}\right)d\nu \\ \\ E_\theta = \\ \dfrac{E_o}{\sqrt{2\pi\rho\sin\theta}}\displaystyle\int_0^\infty \left(\frac{\sin 2\theta_\nu}{2}\right)^{\frac{1}{2}}\left(e^{i\Psi(+,+)} + e^{i\Psi(+,-)} + e^{i\Psi(-,+)} + e^{i\Psi(-,-)}\right)d\nu\end{array}\right\} \tag{5.9-5}$$

where the four phases are given by

$$\left.\begin{array}{l}\Psi(+,+) = D_\nu + \nu(\theta_\nu + \theta) + \dfrac{\pi}{4} + G[\rho,\nu] - 2G\left[\rho^\dagger,\nu\right] \\ \Psi(+,-) = D_\nu + \nu(\theta_\nu - \theta) - \dfrac{\pi}{4} + G[\rho,\nu] - 2G\left[\rho^\dagger,\nu\right] \\ \Psi(-,+) = -D_\nu - \nu(\theta_\nu - \theta - \pi) + \dfrac{3\pi}{4} - G[\rho,\nu] \\ \Psi(-,-) = -D_\nu - \nu(\theta_\nu + \theta - \pi) + \dfrac{\pi}{4} - G[\rho,\nu]\end{array}\right\} \tag{5.9-6}$$

Here (±,±) designates use of certain combinations of spherical Hankel functions and spherical harmonic asymptotic functions. For example in (+, –) the plus sign in the first argument designates use of the spherical Hankel function of the first kind ξ_l^+, and the minus sign in the second argument designates use of the negative exponential term in the asymptotic form for the Legendre polynomial. In the integrals in Eq. (5.9-5), θ_ν varies slowly with ν, but the phasor terms are rapidly varying except at stationary phase points. The principal contribution to these integrals comes from neighborhoods about stationary phase points. In the stationary phase technique, the phase of each phasor is expanded in a Taylor series through second degree about possible stationary phase values of ν. Thus, the first-degree term is zero, and the Taylor series contains only a zeroth-degree term and a quadratic term. Upon evaluating the slowly varying terms at a stationary phase point, the integral reduces to a Fresnel integral.

To see if a stationary phase point exists for the four phasors given in Eq. (5.9-6), we substitute Eq. (5.9-4) into Eq. (5.9-6) to obtain

$$\left.\begin{aligned}
\Psi(+,+) &= D_\nu + \nu\left(2\theta_\nu + \Delta\theta_\nu\right) + \frac{\pi}{4} + G[\rho,\nu] - 2G\left[\rho^\dagger,\nu\right] \\
\Psi(+,-) &= D_\nu - \nu\Delta\theta_\nu - \frac{\pi}{4} + G[\rho,\nu] - 2G\left[\rho^\dagger,\nu\right] \\
\Psi(-,+) &= -D_\nu - \nu\left(2\theta_\nu - 2\pi - \Delta\theta_\nu\right) + \frac{3\pi}{4} - G[\rho,\nu] \\
\Psi(-,-) &= -D_\nu + \nu\Delta\theta_\nu + \frac{\pi}{4} - G[\rho,\nu]
\end{aligned}\right\} \qquad (5.9\text{-}6')$$

Comparing the terms in Eq. (5.9-6′) with Fig. 3-14, we conclude that (+,+) is associated with an outgoing wave from the far side of the scattering sphere—that is, it is associated with the retrograde path; (−,−) is associated with the incoming wave on the near side, and so on. We can eliminate by inspection the phase terms $\Psi(+,+)$ and $\Psi(-,+)$, because these are associated with waves that travel around the far side of the sphere in Fig. 3-14 via point P_2. They will provide negligible contributions to the spectral integrals for the electric field when r_o/λ is large and the point L is located well into the first or second quadrants in θ. The term $2\nu\theta_\nu$ in their arguments results in very high rates of phase accumulation for essentially all values of ν.

5.9.3 Stationarity of $\Psi(\pm,\pm)$

Returning to Eq. (5.9-6), we take the partial derivative of the phase $\Psi(\pm,\pm)$ with respect to ν in each of these four phasor combinations that appear in Eq. (5.9-6) and attempt to set the resultant equal to zero. We conclude

(remembering that $0 < \theta_v \le \pi/2$ and that $0 < \theta < \pi$) the following with regard to the possibility of achieving a stationary value for each of these phases:

$$\partial\Psi(+,+)/\partial\nu = 0:$$

no, if $\theta > \partial\left(2G\left[\rho_*^{\dagger},\nu\right] - G[\rho,\nu]\right)/\partial\nu$,

$$\partial\Psi(+,-)/\partial\nu = 0:$$

yes, if $\theta < \pi/2 - \partial\left(2G\left[\rho_*^{\dagger},\nu\right] - G[\rho,\nu]\right)/\partial\nu$,

$$\partial\Psi(-,+)/\partial\nu = 0:$$

no, if $\theta \ge 0$

$$\partial\Psi(-,-)/\partial\nu = 0:$$

yes, if $\theta > \pi/2 - \partial G\left[\rho_*^{\dagger},\nu\right]/\partial\nu$.

As already mentioned, for an occultation from a LEO, we can effectively rule out the $(+,+)$ and $(-,+)$ combinations. These are contributions to the integrals in Eq. (5.9-5) that originate from the far side of the sphere (Fig. 3-14). They are negligible when θ lies well into the upper quadrants and when the ratio r_o/λ is very large, both of which are assumed here. The combination $(+,-)$ corresponds to an outgoing wave on the near side of the sphere where the spherical Hankel function $\xi_l^+(\rho)$ is used; the combination $(-,-)$ corresponds to an incoming wave on the near side where $\xi_l^-(\rho)$ applies.

5.9.4 Plane Waves

As an illustrative case, consider the stationary phase possibilities for a planar wave in a homogeneous medium. Here $n' \equiv 0$ and, therefore, $G[\rho,\nu] \equiv 0$. An appropriate spectral representation for this case is given from Bauer's identity (Section 5.3) in spherical coordinates:

$$\exp(i\rho\cos\theta) = \sum_{j=0}^{\infty} i^l(2l+1)\frac{\psi_l(\rho)}{\rho}\mathrm{P}_l(\cos\theta) \qquad (5.9\text{-}7)$$

Here, expressing $i^l\psi_l\,\mathrm{P}_l$ in terms of phasors (with $\psi_l = (\xi_l^+ + \xi_l^-)/2$) leads to the same combinations $\Psi(\pm,\pm)$ given in Eq. (5.9-6) but with $G[\rho,\nu] \equiv 0$. Attempting to set $\partial\Psi(\pm,\pm)/\partial\nu = 0$ for the four possible sign combinations in the region $0 < \theta < \pi$ produces the following conditions on θ_v:

$$(+,+):\ \theta_v = -\theta;\ \text{impossible}$$
$$(+,-):\ \theta_v = \theta;\ \text{possible if}\ 0 < \theta < \pi/2$$
$$(-,+):\ \theta_v = \theta + \pi;\ \text{impossible}$$
$$(-,-):\ \theta_v = \pi - \theta;\ \text{possible if}\ \pi > \theta > \pi/2$$

It follows that the $(+,-)$ case corresponds to a departing planar wave and that the $(-,-)$ case corresponds to an approaching wave. In either case, a stationary value for $\Psi(\pm,-)$ is achieved when the spectral number $\nu = \nu^* = \rho_* = \rho\sin\theta$ (see Eq. 5.9-6 with $G[\rho,\nu] \equiv 0$). For this value, $D_{\nu^*} = |D_{\rho_*}| = |\rho\cos\theta|$. The second derivative of $\Psi(\pm,-)$ is $\partial^2\Psi/\partial\nu^2 = \pm D_{\nu^*}^{-1}$.

We can evaluate the summation in Bauer's identity using the stationary phase technique. Expanding $\Psi(\pm,-)$ in a Taylor series about the stationary phase point $\nu^* = \rho_*$ through the quadratic term yields

$$\left.\begin{aligned} &\Psi(\pm,-) \doteq D_{\nu^*} \pm \frac{1}{2D_{\nu^*}}(\nu-\nu^*)^2 \mp \frac{\pi}{4}, \\ &\nu^* = \rho_* = \rho\sin\theta,\ \ D_{\nu^*} = |\rho\cos\theta| \end{aligned}\right\} \tag{5.9-8}$$

Replacing the summation in Bauer's identity by an integral and using Eq. (5.9-8), we obtain

$$\begin{aligned} \sum_{j=0}^{\infty} i^l(2l+1)\frac{\psi_l}{\rho}\mathrm{P}_l(\cos\theta) &\doteq \sqrt{\frac{1}{2\pi}}\int_0^{\infty} D_\nu^{-1/2}\exp(\pm i\Psi(\pm,-))d\nu \\ &\doteq \sqrt{\frac{\pm 1}{2\pi i D_{\nu^*}}}\exp(\pm iD_{\nu^*})\int_0^{\infty}\exp\left(\pm i\frac{(\nu-\rho_*)^2}{2D_{\nu^*}}\right)d\nu = \exp(\pm iD_{\nu^*}) \end{aligned} \tag{5.9-9}$$

By a change of variable $\pi s^2 = (\nu-\rho_*)^2/D_{\nu^*}$, the integral in Eq. (5.9-9) is transformed into a complex Fresnel integral. Since $\rho_* >> 1$, this is essentially a complete Fresnel integral equal to $(\pm 2\pi i D_{\nu^*})^{1/2}$. Thus, Eq. (5.9-9) reduces to $\exp(\pm iD_{\nu^*}) = \exp(iD_{\rho_*})$, which equals the left side of Bauer's identity.

Note again that $2\psi_l = \xi_l^+ + \xi_l^-$. However, ξ_l^- appears in $\Psi(-,-)$, and ξ_l^+ appears in $\Psi(+,-)$. Therefore, only ξ_l^- contributes to the spectral sum in Eq. (5.9-9) for a position well within the incoming region, i.e., for an incoming wave. Only ξ_l^+ contributes for an outgoing wave. This is consistent with the asymptotic forms that ξ_l^+ and ξ_l^- assume when $\rho >> \nu$. Here

$\xi_l^{\pm} / \rho \to (\mp i)^{l+1} \exp(\pm i\rho) / \rho$. Thus, ξ_l^{+} / ρ corresponds to an outgoing spherical wave, and ξ_l^{-} / ρ corresponds to an incoming spherical wave.

5.9.5 The Electric Field for an Incoming Wave

We now evaluate the integrals in Eq. (5.9-5) for the electric field vector using the stationary phase technique. We first assume that the point (r,θ) is in the incoming region well away from a turning point, so that the negative argument asymptotic forms for the Airy functions apply. In an incoming region, $\pi > \theta > \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$. If we set as a criterion for "well away" that the point (r,θ) must be such that $\hat{y} < -2$ at its stationary phase point, then by applying Bouguer's law in Eq. (5.6-3), one can show that a suitable criterion is $\theta > \pi/2 - \tilde{\alpha}(\rho_*, \rho_*) + \sqrt{2} / K_{\rho_*}$. For GPS signals in the Earth's atmosphere, $\sqrt{2} / K_{\rho_*} \approx 3\,\text{mrad}$. In an incoming region well away from a turning point, only the phase ${}^{(i)}\Psi = \Psi(-,-)$ contributes significantly to the spectral integrals.

We denote the field in an incoming region by ${}^{(i)}\boldsymbol{E}(\rho,\theta)$, and we let ${}^{(i)}\Psi = \Psi(-,-)$. Thus, ${}^{(i)}\Psi$ is the spectral density function for the phase delay at the point (ρ,θ) for an incoming wave. It follows from Eqs. (5.9-5) and (5.9-6) for $\rho > \rho_*$ that

$$\left.\begin{aligned} {}^{(i)}E_r(\rho,\theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^{\infty} \left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu} \right)^{1/2} \exp\left(i\,{}^{(i)}\Psi(\rho,\theta,\nu) \right) d\nu, \\ {}^{(i)}E_\theta(\rho,\theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^{\infty} \left(\sin\theta_\nu \cos\theta_\nu \right)^{1/2} \exp\left(i\,{}^{(i)}\Psi(\rho,\theta,\nu) \right) d\nu \end{aligned}\right\} \tag{5.9-10}$$

To evaluate these integrals using the stationary phase technique, we expand ${}^{(i)}\Psi$ in a Taylor's series through second order in spectral number about its stationary phase value. Using Eqs. (5.9-2) and (5.9-6), $\partial\,{}^{(i)}\Psi / \partial\nu$ becomes

$$\frac{\partial}{\partial\nu}\left({}^{(i)}\Psi \right) = -\left(\theta_\nu + \theta - \pi + \frac{\partial G}{\partial\nu} \right) = 0 \tag{5.9-11}$$

and the second derivative becomes

$$\frac{\partial^2}{\partial\nu^2}\left({}^{(i)}\Psi \right) = -\frac{1}{D_\nu} - \frac{\partial^2 G}{\partial\nu^2} \tag{5.9-12}$$

Let ν^* be a spectral number value that provides a stationary phase; that is, it satisfies the stationarity condition in Eq. (5.9-11). Then from Eqs. (5.9-2) and (5.9-11), one obtains

$$\left.\begin{aligned} \theta_{\nu^*} &= \pi - \theta - \left(\frac{\partial G[\rho,\nu]}{\partial \nu}\right)_{\nu=\nu^*} \\ \nu^* &= \rho \sin\theta_{\nu^*} \end{aligned}\right\} \tag{5.9-13}$$

We define $\delta(\rho,\theta,\nu)$ by

$$\left.\begin{aligned} \theta_\nu &= \pi - \theta - \delta \\ \delta &= \pi - \theta - \sin^{-1}\left(\frac{\nu}{\rho}\right) \\ \rho &= krn(\rho) \end{aligned}\right\} \tag{5.9-14}$$

where for a thin atmosphere $\delta^* = \delta(\rho,\theta,\nu^*)$ generally will be a small quantity. When ν assumes a value so that $\delta^* = \partial G / \partial \nu$, Eq. (5.9-14) shows that we have a stationary phase point. In our discussion in Section 3.12 on stationary phase processes in Mie scattering theory, we noted that δ^* should be very close in value to the refractive bending angle. Here δ^* should be close in value to the cumulative bending angle $\tilde{\alpha}(\rho,\rho_*)$ for a ray passing through the point (ρ,θ) with an impact parameter value of ρ_*. Here $\rho_* = \rho\sin[\theta + \tilde{\alpha}(\rho,\rho_*)]$. From Eq. (5.7-8), it follows that $\partial G[\rho,\nu]/\partial\nu = \tilde{\alpha}(\rho,\nu)$ to high accuracy provided that $\nu < \rho - \sim 2K_\nu$ (see Fig. 5-4). The differences between δ^* and $\tilde{\alpha}(\rho,\rho_*)$ and between ν^* and ρ_* can be obtained by expanding Eq. (5.9-14) in a power series for $G[\rho,\nu]$ about the point $\tilde{\nu}$ where $\partial G/\partial\nu = \tilde{\alpha}$, exactly. One obtains

$$\left.\begin{aligned} \nu^* - \rho_* &\doteq (\delta^* - \tilde{\alpha})\rho\cos(\theta + \tilde{\alpha}) \\ \delta^* - \tilde{\alpha} &\doteq (\nu^* - \tilde{\nu})\left(\frac{\partial^2 G[\rho,\nu]}{\partial \nu^2}\right)_{\nu=\nu^*} \end{aligned}\right\} \tag{5.9-15}$$

We expect that $\tilde{\nu}$ will be close to ρ_*. Upon setting $\partial^2 G/\partial\nu^2 = \partial\tilde{\alpha}/\partial\rho_*$, it follows that

$$\left.\begin{aligned} v^* - \rho_* &= D_{\rho_*}\frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\left(1 - D_{\rho_*}\frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\right)^{-1} \\ v^* - \tilde{v} &= (\rho_* - \tilde{v})\left(1 - D_{\rho_*}\frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\right)^{-1} \end{aligned}\right\} \tag{5.9-16}$$

It follows that

$$\delta^* - \tilde{\alpha} \doteq (\rho_* - \tilde{v})\frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\left(1 - D_{\rho_*}\frac{\partial\tilde{\alpha}(\rho,\rho_*)}{\partial\rho_*}\right)^{-1} \tag{5.9-17}$$

We note the defocusing factor showing up in these expressions for the offsets. It follows from Eqs. (5.9-14) and (5.9-15) that, if v^* is near the value ρ_*, which is the value(s) of the impact parameter for the ray(s) passing through the point (r,θ), then δ^* will equal the corresponding value of $\tilde{\alpha}$ with close accuracy. If the ray system from the actual refractivity profile generates a caustic surface, when the position (r,θ) is such that v^* lies near a contact point with that caustic, $1-(\partial\tilde{\alpha}/\partial\rho_*)D_{\rho_*} \approx 0$. Equations (5.9-16) and (5.9-17) are not valid in that neighborhood.

Figure 5-17 shows the stationary phase solutions for the exponential refractivity profile in Eq. (5.8-2a) at the intersections of the $\partial G/\partial v$ and $\delta(\rho,\theta,v)$ curves [see Eqs. (5.7-7) and (5.9-14)]. The figure shows the

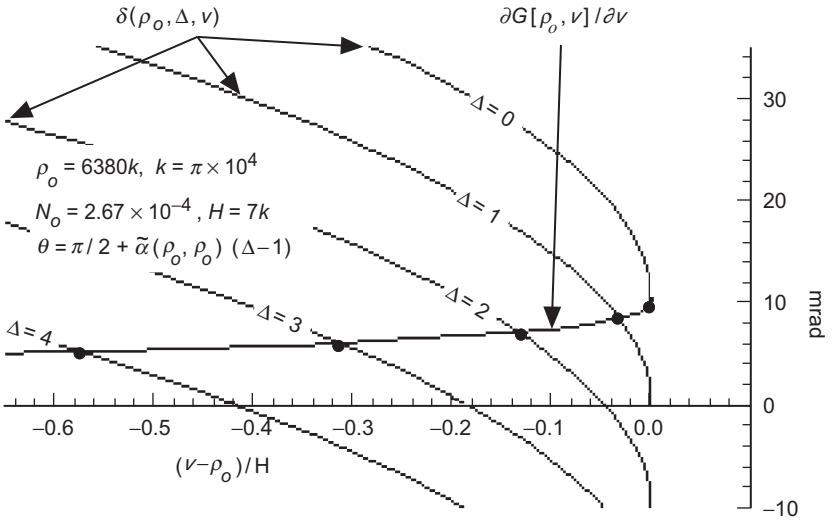

Fig. 5-17. Stationary phase solution v^* to Eq. (5.9-13) for the Case A refractivity profile [Eq. (5.8-2a)].

stationary phase solution for four position points (r,θ), all at the same radial distance r_o and with θ increasing from the turning point at $\theta = \theta_* = \pi/2 - \tilde{\alpha}(\rho_o, \rho_o)$ in four equally spaced increments up to $\theta = \pi/2 + 3\tilde{\alpha}(\rho_o, \rho_o)$. Even at $\theta = \pi/2$, which is very close to the turning point and where one might worry that either (1) the asymptotic form at negative arguments for the Airy functions might break down or (2) the approximations for the spectral coefficients given in Eq. (5.5-21) for a_l^- and Eq. (5.8-24) for a_l^+ might fail, the agreement is close. At $\theta = \pi/2$, the intersection point yields a stationary phase value that corresponds to an Airy function argument of $\hat{y}^* \approx -20$. This value for $\hat{y}^*$ also is corroborated by evaluating $K_\rho^{-1}(\rho_* - \rho)$ from Bouguer's law in Eq. (5.6-3). This justifies our use of the negative argument asymptotic forms for the Airy functions in the expressions for $\Psi(-,-)$ in Eq. (5.9-1) when θ is at least $\pi/2$ or greater. The intersection point for $\theta = \pi/2$ in Fig. 5-17 yields a value for $\partial G/\partial\nu$ of 8.056 mrad, which differs by only $\sim 5\,\mu$rad from the value for $\tilde{\alpha}(\rho_o, \rho_*)$, $\rho_* = \rho_o \cos\Delta$, predicted by ray theory in Eq. (5.6-2). A lower bound on θ already has been established from Bouguer's law, where we set $\hat{y}^* = -2$; this gave $\theta \geq \pi/2 - \tilde{\alpha}(\rho_o, \rho_o) + \varepsilon$, $\varepsilon = \sqrt{2}K_{u_o}^{-1} \approx \tilde{\alpha}/3$. Thus, except for a very narrow range of θ values near the turning point, the stationary phase point ν^* should be very close to ρ_* when the curvature in the refractivity is relatively slight, i.e., when $K_{\rho_o}/(kH)$ is small. $K_{\rho_o}/(kH)$ essentially is the ratio of the Airy function transition scale (a change in ρ and/or ν corresponding to $\Delta\hat{y} = 1$) to the refractivity scale height (in phase units), which for the example shown in Fig. 5-17 is about 0.002.

Even at the turning point at $\theta = \theta_* = \pi/2 - \tilde{\alpha}(\rho_o, \rho_o)$, the intersection point in Fig. 5-17 (where $\hat{y} \cong 0$ in this case) yields a value $\delta(\rho, \theta, \nu^*) = \partial G/\partial\nu|_{\nu^*} = 9.894$ mrad; the ray theory value from Eq. (5.6-2) for $\tilde{\alpha}(\rho_o, \rho_o)$ is 10.110 mrad. We already have noted the levels of agreement between these quantities, which are shown in detail in Figs. 5-5 and 5-6. These figures show the close agreement between the wave-theoretic and geometric optics representations of bending angle under thin-atmosphere conditions except in the immediate vicinity of a turning point.

5.9.6 Evaluating the Electric Field Vector Using the Stationary Phase Technique

Using the stationary phase technique, we now evaluate the integrals in Eq. (5.9-10) for the electric field vector in the incoming region. We assume that the point (r,θ) is well into the incoming region so that the negative argument asymptotic forms for the Airy functions apply. We have just seen that the

criterion $\theta > \pi/2 - 2\tilde{\alpha}(\rho_*, \rho_*)/3$ should suffice. Inserting into Eq. (5.9-10) the Taylor's series expansion through second-order terms for ${}^{(i)}\Psi(\rho,\theta,v) = \Psi(-,-)$ evaluated at the stationary phase point, and using Eqs. (5.9-11) through-(5.9-15), one obtains for $E_r(r,\theta)$ and $E_\theta(r,\theta)$

$$\left.\begin{aligned}
&{}^{(i)}E_{\left\{\begin{matrix}r\\ \theta\end{matrix}\right\}} \doteq E_o e^{i^{(i)}\Phi^*} \begin{Bmatrix}\sin\left({}^{(i)}\gamma^*\right)\\ \cos\left({}^{(i)}\gamma^*\right)\end{Bmatrix} \sqrt{\frac{-\sin\left({}^{(i)}\gamma^*\right)}{2\pi i D_{v^*}\sin\theta}} \int_0^\infty e^{-i\frac{(v-v^*)^2}{2\zeta D_{v^*}}}\, dv,\\
&{}^{(i)}\Phi^* = {}^{(i)}\Psi\left(\rho,\theta,v^*\right) - \frac{\pi}{4} = \left(-D_v + v\frac{\partial G}{\partial v} - G\right)_{v=v^*},\\
&\left(\frac{\partial {}^{(i)}\Psi}{\partial v}\right)_{v=v^*} = \sin^{-1}\left(\frac{v^*}{\rho}\right) - \left(\pi - \theta - \frac{\partial G}{\partial v}\right)_{v=v^*} = 0,\\
&{}^{(i)}\zeta = \left(1 + {}^{(i)}D_v \frac{\partial^2 G}{\partial v^2}\right)^{-1}_{v=v^*},\quad G = G[\rho, v],\\
&{}^{(i)}\gamma^* = \theta + \left(\frac{\partial G}{\partial v}\right)_{v=v^*},\quad \pi > {}^{(i)}\gamma^* > \frac{\pi}{2},\\
&{}^{(i)}D_{v^*} = -\rho\cos\left({}^{(i)}\gamma^*\right) \geq 0,\quad v^* = \rho\sin\left({}^{(i)}\gamma^*\right)
\end{aligned}\right\} \quad (5.9\text{-}18)$$

The integrals in Eq. (5.9-18) for ${}^{(i)}E_r$ and ${}^{(i)}E_\theta$ (upon noting that $-v^* \cong -\infty$) are readily transformed through a change of variable into the form $\int_{-\infty}^{\infty} \exp\left(-i\pi t^2/2\right) dt$, which is the complex conjugate of the complete Fresnel integral, with a value of $1 - i = (-2i)^{1/2}$. Hence, the electric field for the incoming wave in Eq. (5.9-18) may be written as

$$\left.\begin{aligned}
{}^{(i)}E_r &\doteq E_o \sqrt{\frac{{}^{(i)}\zeta \sin\left({}^{(i)}\gamma^*\right)}{\sin\theta}} \sin\left({}^{(i)}\gamma^*\right) \exp\left(i^{(i)}\Phi^*\right)\\
{}^{(i)}E_\theta &\doteq E_o \sqrt{\frac{{}^{(i)}\zeta \sin\left({}^{(i)}\gamma^*\right)}{\sin\theta}} \cos\left({}^{(i)}\gamma^*\right) \exp\left(i^{(i)}\Phi^*\right)
\end{aligned}\right\} \quad (5.9\text{-}19)$$

We recognize in Eqs. (5.9-18) and (5.9-19) the phase delay term ${}^{(i)}\Phi^*$, which here is largely a phase advance for the incoming wave because it is referenced to the $\theta = \pi/2$ line. Recalling the asymptotic relationships between $G[\rho, v]$ and $\partial G/\partial v$ and bending-angle quantities given by Eqs. (5.7-6) and

(5.7-8), respectively, we see that the equivalent form for ${}^{(i)}\Phi^*$ in geometric optics is

$$\left.\begin{aligned} {}^{(i)}\Phi^* &\doteq \rho\cos\big(\theta+\tilde{\alpha}(\rho,\rho_*)\big)+\rho_*\tilde{\alpha}(\rho,\rho_*) \\ &\quad +\int_{\rho_*}^{\rho}\tilde{\alpha}(\rho,\nu)d\nu+\int_{\rho}^{\infty}\tilde{\alpha}(\omega,\omega)d\omega \\ {}^{(i)}\gamma^* &\doteq \theta+\tilde{\alpha}(\rho,\rho_*) \end{aligned}\right\} \tag{5.9-20}$$

The leading term $\rho\cos\big(\theta+\tilde{\alpha}(\rho,\rho_*)\big)$ is the optical path delay (negative for this incoming case) from the tangent point on the spherical surface radius ρ_* to the point (ρ,θ). The second term is an arc-length term along a sphere of radius ρ_* due to refractive bending. The third and fourth terms—equal to $-G[\rho,\nu^*]$ in wave theory [see Eq. (5.7-6)]—account for the extra phase delay due to the refractive gradient-induced bending through the atmosphere that a ray of impact parameter value ρ_* from the GPS satellite (assumed to be at infinity) undergoes in reaching the incoming point (r,θ).

Equation (5.9-20) may be compared to the relationship with the eikonal equation given in Eq. (5.10-11). Here the spectral density function for the complete phase delay at (r,θ), ${}^{(i)}\Psi(\rho,\theta,\nu)$, has a close correspondence to the eikonal $\mathscr{S}(\rho)$ associated with the path delay along a ray. When ${}^{(i)}\Psi(\rho,\theta,\nu)$ is evaluated at a stationary phase point $\nu^*\doteq\rho_*$, then ${}^{(i)}\Psi(\rho,\theta,\rho_*)\Leftrightarrow\mathscr{S}(\rho)-\mathscr{S}(\rho_*)$.

We also recognize the defocusing factor ${}^{(i)}\zeta$ in Eq. (5.9-18), which has its analog in geometric optics to account for the dispersive effect of the refraction gradient on signal power (see Section 5.6). From Eq. (5.7-7), $\partial^2 G/\partial\nu^2$ is given by

$$\frac{\partial^2 G}{\partial\nu^2}\doteq 2\pi\int_{\rho}^{\infty}\frac{d\log n}{d\rho'}\big(\mathrm{Ai}[\hat{y}]\,\mathrm{Ai}'[\hat{y}]+\mathrm{Bi}[\hat{y}]\,\mathrm{Bi}'[\hat{y}]\big)d\rho' \tag{5.9-21}$$

Using the asymptotic forms for the Airy functions given in Eq. (3.8-10) when $\rho>\nu$, one obtains

$$\left.\frac{\partial^2 G}{\partial\nu^2}\doteq -\nu^2\int_{\rho}^{\infty}\frac{d\log n}{d\rho'}\frac{d\rho'}{\big(\rho'^2-\nu^2\big)^{3/2}}=\frac{\partial\tilde{\alpha}}{\partial\nu},\quad \tilde{\alpha}=\tilde{\alpha}(\rho,\nu)\right\} \tag{5.9-22}$$

The integral here can be integrated by parts to eliminate its near singularity. Figure 5-5 compares $\partial^2 G/\partial\nu^2$ and $\partial\tilde{\alpha}/\partial\nu$ for the exponential refractivity

profile given in Eq. (5.8-2a). It follows that in terms of geometric optics quantities the defocusing factor can be written as

$$ {}^{(i)}\zeta^{-1} \doteq \left(1 - \rho \cos\left(\theta + \tilde{\alpha}(\rho,\nu)\right) \frac{\partial \tilde{\alpha}(\rho,\nu)}{\partial \nu} \right)_{\nu=\rho_*} \tag{5.9-23} $$

When $\partial\tilde{\alpha}/\partial\nu > 0$, then ${}^{(i)}\zeta < 1$ because $\pi > \theta + \tilde{\alpha}(\rho,\rho_*) > \pi/2$.

When the refractivity profile results in a unique stationary phase path from the GPS satellite to the point (r,θ) [e.g., Case A], then Eq. (5.9-19) shows that the refracted wave, which was planar before entering the atmosphere, is still nearly planar (for a thin atmosphere) at the point (r,θ). However, its Poynting vector $\boldsymbol{S}$ is pointed away from the original axis of propagation by an angular displacement $\tilde{\alpha}$. Its phase is delayed by an additional amount $({}^{(i)}\Phi^* - \rho\cos\theta)$. Its amplitude is modified by the factor $({}^{(i)}\zeta \sin(\theta+\tilde{\alpha})/\sin\theta)^{1/2}$, which may be significantly greater or less than unity depending on the magnitude of $\partial^2 G/\partial\nu^2$ at the stationary phase point.

If we have a multipath situation where, for example, three different total bending-angle values satisfy the boundary conditions, then Eq. (5.9-19) will appear in triplet form (or even in higher odd multiples for more complicated refractivity profiles). The Gaussian refractivity profile given in Eq. (5.8-2b) will produce triplets for a certain range of altitudes. The total field in this case would be obtained by vector addition of the field components from each contribution. We note again the failure of the stationary phase technique when the stationary phase points are too near each other or, alternatively, when they are too near the first contact points with the caustic surfaces for the complete ray system generated by varying the position of the point (r,θ). In wave theory, the accuracy of the stationary phase technique for evaluating the integrals in Eq. (5.9-10) depends on the magnitude of $\partial^{2\,(i)}\Psi/\partial\nu^2$ being sufficiently large. Caustics occur when stationary phase points in spectral number space also coincide, or nearly so, with $\partial^{2\,(i)}\Psi/\partial\nu^2 = 0$. The conditions for the validity of the stationary phase approach and the third-order approach leading to the Airy function of the first kind are discussed in Appendix D. Multipath scenarios and caustics are discussed more fully in Section 5.12.

Finally, we note the breakdown in accuracy of this osculating parameter approach if one attempts to use it exactly at a turning point. As we approach a turning point, $D_{\nu^*} \to 0$. Because $\partial^2 G[\rho,\nu]/\partial^2\nu$ is finite at $\nu = \rho$, we see from Eq. (5.9-18) that the defocusing factor ζ predicted by this technique approaches unity at a turning point. But from geometric optics [Eq. (5.6-16)], we know that $\zeta^{-1} \to 1 + u_* n_*'/n_*$, which results from the singularity in $\partial\tilde{\alpha}(\rho,\nu)/\partial\nu$ as the turning point is approached, i.e.,

$D_v \partial\tilde{\alpha}/\partial v \to -(dn/d\rho)(\rho/n)$ as $D_v \to 0$. Turning points using an Airy layer approach are discussed in Section 5.11.

5.9.7 An Outgoing Electric Field

For a point in the outgoing region well away from a turning point, only the phase combination ${}^{(o)}\Psi = \Psi(+,-)$ in Eq. (5.9-6) contributes significantly to the scattering integrals for the electric field representation given in Eq. (5.9-5). For an outgoing region, $0 < \theta \le \pi/2 - \tilde{\alpha}(\rho_*, \rho_*)$, and we have

$$\left.\begin{aligned} {}^{(o)}E_r(\rho,\theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty \left(\frac{\sin^3\theta_v}{\cos\theta_v}\right)^{1/2} \exp\left(i^{(o)}\Psi(\rho,\theta,v)\right) dv \\ {}^{(o)}E_\theta(\rho,\theta) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta}} \int_0^\infty \left(\sin\theta_v \cos\theta_v\right)^{1/2} \exp\left(i^{(o)}\Psi(\rho,\theta,v)\right) dv \end{aligned}\right\} \quad (5.9\text{-}24)$$

Referring to Eq. (5.9-6), we see that ${}^{(o)}\Psi = \Psi(+,-)$ is obtained from ${}^{(i)}\Psi = \Psi(-,-)$ through the transformation

$${}^{(o)}\Psi = -{}^{(i)}\Psi, \;\; G[\rho,v] \Rightarrow 2G\left[\rho^\dagger(v),v\right] - G[\rho,v]$$

${}^{(o)}\Psi$ is the spectral density function for the phase delay at the point (ρ,θ) for an outgoing wave. We need only apply this transformation to the incoming forms in Eq. (5.9-18) to obtain the stationary phase evaluation of the spectral integrals for an outgoing wave. Carrying through the stationary phase computations in Eq. (5.9-24) yields the electric field for the outgoing wave. It may be written as

$$\left.\begin{aligned} {}^{(o)}E_r &\doteq E_o \sin\left({}^{(o)}\gamma^*\right) \sqrt{\frac{{}^{(o)}\zeta \sin\left({}^{(o)}\gamma^*\right)}{\sin\theta}} \exp\left(i^{(o)}\Phi^*\right) \\ {}^{(o)}E_\theta &\doteq E_o \cos\left({}^{(o)}\gamma^*\right) \sqrt{\frac{{}^{(o)}\zeta \sin\left({}^{(o)}\gamma^*\right)}{\sin\theta}} \exp\left(i^{(o)}\Phi^*\right) \\ {}^{(o)}\gamma^* &= \theta + \left(2\frac{dG^\dagger}{dv} - \frac{\partial G}{\partial v}\right)_{v=v^*} \end{aligned}\right\} \quad (5.9\text{-}25)$$

where the outgoing phase ${}^{(o)}\Phi^*$ is given by

$$\left.\begin{aligned} {}^{(o)}\Phi^{*} &= \left(D_{\nu} + \nu\left(2\frac{dG^{\dagger}}{d\nu} - \frac{\partial G}{\partial \nu} \right) - (2G^{\dagger} - G) \right)_{\nu=\nu^{*}} \\ &\doteq D_{\rho_{*}} + \rho_{*}\left(2\tilde{\alpha}(\rho_{*},\rho_{*}) - \tilde{\alpha}(\rho,\rho_{*}) \right) \\ &\quad + 2\int_{\rho_{*}}^{\infty} \tilde{\alpha}(\omega,\omega)d\omega - \int_{\rho_{*}}^{\rho} \tilde{\alpha}(\rho,\omega)d\omega - \int_{\rho}^{\infty} \tilde{\alpha}(\omega,\omega)d\omega, \\ G^{\dagger} &= G\left[\rho^{\dagger}(\nu),\nu\right],\ G = G[\rho,\nu],\ D_{\rho_{*}} = \sqrt{\rho^{2} - \rho_{*}^{2}} \end{aligned}\right\} \quad (5.9\text{-}26)$$

The leading term D_{ρ_*} is the optical path delay from the tangent point on the spherical surface of radius ρ_* to the point (ρ,θ). The second term is an arc-length term along a sphere of radius ρ_* due to cumulative refractive bending $2\tilde{\alpha}(\rho_*,\rho_*) - \tilde{\alpha}(\rho,\rho_*)$ to the outgoing point (ρ,θ). The first and second terms combined reference the phase to the line $\theta = \pi/2$ for an incident-collimated wave originating from the direction $\theta = \pi$. The third, fourth, and fifth terms in wave theory account for the extra phase delay due to the refractive gradient-induced bending through the atmosphere that a ray of impact parameter value ρ_* from the GPS satellite (assumed to be at infinity) undergoes in reaching the outgoing point (ρ,θ).

The defocusing factor ${}^{(o)}\zeta$ in Eq. (5.9-25) also has its analog in geometric optics to account for the dispersive effect of the refraction gradient on signal power (see Section 5.6). It is given by

$$\left.\begin{aligned} {}^{(o)}\zeta &= \left(1 - D_{\nu}\left(2\frac{d^{2}G[\rho^{\dagger}(\nu),\nu]}{d\nu^{2}} - \frac{\partial^{2}G[\rho,\nu]}{\partial\nu^{2}} \right) \right)^{-1} \Bigg|_{\nu=\nu^{*}} \\ &\doteq \left(1 - D_{\rho_{*}}\left(2\frac{d}{d\rho_{*}}\left(\tilde{\alpha}(\rho_{*},\rho_{*})\right) - \frac{\partial}{\partial\rho_{*}}\left(\tilde{\alpha}(\rho,\rho_{*})\right) \right) \right)^{-1} \end{aligned}\right\} \quad (5.9\text{-}27)$$

Forms for $2d^{2}G[\rho^{\dagger},\rho^{\dagger}]/d\nu^{2} - \partial^{2}G[\rho,\nu]/\partial\nu^{2}$ are given in Appendix J and Eq. (5.9-22).

As has already been shown for the incoming case, Eq. (5.9-25) shows the (small) deflections in angular displacement $2\tilde{\alpha}_* - \tilde{\alpha}$. Its phase is delayed by an additional amount, ${}^{(o)}\Phi^{*} - \rho\cos\theta$. Its amplitude is modified by the factor $(\zeta\sin(\theta + 2\tilde{\alpha}_* - \tilde{\alpha})/\sin\theta)^{1/2}$, and so on.

We have already commented in the incoming case about multipath and the possible non-uniqueness of these solutions, depending on the profile of the refractivity.

5.9.8 The Electric Field at the LEO

We assume that the LEO is well out of the refracting medium and receiving a signal from an outgoing wave. The equations given for the outgoing case also describe the field at the LEO except that they are somewhat simplified because $\tilde{\alpha}(\rho,\rho_*)\to 0$ as $\rho\to\infty$. They become

$$\left.\begin{aligned} E_r(\rho_{\mathrm{L}},\theta_{\mathrm{L}}) &\doteq E_o \sin\gamma^* \sqrt{\frac{\zeta\sin\gamma^*}{\sin\theta_{\mathrm{L}}}}\exp(i\Phi^*) \\ E_\theta(\rho_{\mathrm{L}},\theta_{\mathrm{L}}) &\doteq E_o \cos\gamma^* \sqrt{\frac{\zeta\sin\gamma^*}{\sin\theta_{\mathrm{L}}}}\exp(i\Phi^*) \\ \Phi^* &= D_{\nu^*} + 2\left(\nu\frac{dG^\dagger}{d\nu} - G^\dagger\right)_{\nu^*} \doteq D_{\rho_*} + 2\rho_*\tilde{\alpha}_* + 2\int_{\rho_*}^{\infty}\tilde{\alpha}(\nu,\nu)d\nu \end{aligned}\right\} \quad (5.9\text{-}28)$$

where

$$\left.\begin{aligned} &D_{\nu^*} = \rho_L\cos\gamma^* \doteq D_{\rho_*} = \rho_L\cos(\theta+2\tilde{\alpha}_*), \\ &\zeta = \zeta(\nu^*,\rho_L) = \left(1-2D_{\nu^*}\frac{d^2G^\dagger}{d\nu^{*2}}\right)^{-1} \doteq \left(1-2D_{\rho_*}\frac{d\tilde{\alpha}_*}{d\rho_*}\right)^{-1}, \\ &\gamma^* = \theta+\left(2\frac{dG^\dagger}{d\nu^*}\right)\doteq\theta+2\tilde{\alpha}_*,\ 0>\gamma^*<\frac{\pi}{2},\ \nu^* = \rho\sin^{(o)}\gamma^*, \\ &G^\dagger = G\left[\rho^\dagger(\nu^*),\nu^*\right],\ \tilde{\alpha}_* = \tilde{\alpha}(\rho_*,\rho_*),\ \rho_* = \rho_L\sin(\theta+2\tilde{\alpha}_*) \end{aligned}\right\} \quad (5.9\text{-}29)$$

Here it is understood that both $\tilde{\alpha}_*$ and the impact parameter ρ_* are implicit functions of the LEO position coordinates $(r_{\mathrm{L}},\theta_{\mathrm{L}})$ through the application of Bouguer's law and the total refractive bending angle given in Eq. (5.6-5).

5.10 Comparison of Geometric Optics and Wave Theory

Here we briefly review the scalar diffraction results applied to a thin phase screen model. These results then are compared with those from wave theory. Thin-screen models were discussed in Chapter 2, notably in Section 2.3. A thin phase screen model provides a proxy for the actual atmospheric medium through which the electromagnetic wave travels. A thin phase screen model mimics the transmission effects of the real atmosphere on a traversing electromagnetic wave. The closeness of this match is limited, of course, by limitations in the thin-screen model itself. However, it is surprising how

accurate a description of certain electromagnetic processes can be achieved from the model in most situations when thin-atmosphere conditions apply. Thin-atmosphere conditions are defined in Chapter 2, Eqs. (2.2-8) and (2.2-9). Basically, thin-atmosphere conditions imply that the curvature of the ray path is much smaller than the curvature of the ambient equipotential surface, and that the length along the ray where refracting or bending occurs is small compared to the radius of curvature of the equipotential surface. Both of these conditions are related to the refractive gradient in the atmosphere. A somewhat strengthened thin-screen requirement for a perpendicularly mounted thin screen is that no caustics may occur in that screen. The condition for this is given in Eq. (2.3-13). Thin phase screen models offer considerable simplification in calculations of electromagnetic processes, including refractive bending and path delay, but also of other important properties, such as diffraction, multipath, caustics, and shadow zones. Chapter 2 discusses two alternate thin-screen models, a planar screen oriented perpendicularly to the LEO/GPS line near the point of tangency of the actual ray in the Earth's atmosphere and an impact parameter "screen," actually the post-encounter impact parameter space curve generated by varying the impact parameter $\rho_* = ka$.

In Chapter 3, we discussed rainbow caustic effects that would be difficult, but not impossible, for a thin-screen model to predict. Processes at turning points also would be difficult for a thin-screen model, as well as super-refractivity situations. Also, the thin planar screen model will encounter difficulty when caustics occur in the screen, i.e., when $1 + a\alpha d\alpha / da = 0$. For an exponentially distributed refractivity, this is equivalent to requiring $\beta < (2\pi)^{-1/2} \approx 0.4$, where β is the ray path curvature parameter given in Eq. (2.2-9) as part of the thin-atmosphere definition. This threshold $\beta = (2\pi)^{-1/2}$ is exceeded across some marine layer boundaries in the lower troposphere, but rarely is exceeded at higher altitudes.

5.10.1 Comparison of Wave Theory with Geometric Optics

We now compare stationary phase terms in spectral number from wave theory with phase terms from geometric optics. In ray theory, the phase at the point (r_L, θ_L) relative to the phase at the line $\theta = \pi/2$ for a collimated incident wave traveling along the direction $\theta = 0$ is given by [see Eqs. (A-55) through (A-57)]

$$\varphi = \rho_L \cos(\theta_L + \alpha_L) + \rho_* \alpha_L + \int_{\rho_*}^{\infty} \alpha_L(\omega) d\omega \qquad (5.10\text{-}1)$$

In wave theory, a stationary value of the spectral density function for the phase delay at the LEO, Φ_L^*, referenced to the line $\theta = \pi/2$, is given from Eq. (5.9-28) by

$$\left.\begin{aligned} \Phi_{\mathrm{L}}^{*} &= \left(\sqrt{\rho_{\mathrm{L}}^2 - \nu^2} + 2\nu \frac{dG^{\dagger}}{d\nu} - 2G^{\dagger} \right)\Bigg|_{\nu=\nu^*} \\ &\left(\theta_{\nu} - \theta_{\mathrm{L}} - 2\frac{dG^{\dagger}}{d\nu} \right)\Bigg|_{\nu=\nu^*} = 0, \\ \theta_{\nu} &= \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}} \right), \quad G^{\dagger}(\nu) = G\left[\rho^{\dagger}(\nu), \nu \right] \end{aligned}\right\} \tag{5.10-2}$$

The second line is the stationary phase condition that the spectral number must satisfy. We have seen in Section 5.7, Eqs. (5.7-4) and (5.7-14) (see also Appendix J), that

$$\left.\begin{aligned} 2\frac{dG^{\dagger}(\nu)}{d\nu} &= \alpha_{\mathrm{L}}(\nu) + \mathrm{O}\left[\rho \frac{d^2 n}{d\rho^2}\Big|_{\rho=\nu} \right] \\ 2G^{\dagger}(\nu) &\doteq -\int_{\nu}^{\infty} \alpha_{\mathrm{L}}(\omega) d\omega \end{aligned}\right\} \tag{5.10-3}$$

This means that the term $\rho(d^2n/d\rho^2)$ must be much smaller than α_{L} to maintain close agreement between spectral number in wave theory and impact parameter in ray theory. This translates into a scale height in distance (i.e., not a scale height in impact parameter) for an exponential refractivity profile that must be greater than about 1 km. This scale height limit already is super-refractive in the lower troposphere, i.e., $\beta > 1$; rays won't exist for a certain critical altitude range of tangency points lying in and below the super-refractive layer. (See Section 6.4.) If this stationary phase condition in Eq. (5.10-2) can be satisfied, and if the curvature condition is met, i.e., $\rho_* \mid d^2n/d\rho^2 \mid_* << \alpha_{\mathrm{L}}(\rho_*)$, then $\nu^* \doteq \rho_{\mathrm{L}} \sin(\theta_{\mathrm{L}} + \alpha_{\mathrm{L}}(\nu^*))$, and it follows that the stationary value for the spectral density function is given by

$$\Phi_{\mathrm{L}}^{*} \doteq \rho_{\mathrm{L}} \cos\left(\theta_{\mathrm{L}} + \alpha_{\mathrm{L}}\left(\nu^*\right)\right) + \nu^* \alpha_{\mathrm{L}}\left(\nu^*\right) + \int_{\nu^*}^{\infty} \alpha_{\mathrm{L}}(\omega) d\omega \tag{5.10-4}$$

The spectral number ν^* will be close to ρ_* if the inequality $\rho_* \mid d^2n/d\rho^2 \mid_* << \alpha_{\mathrm{L}}(\rho_*)$ holds; then the stationary value of phase delay spectral density function in Eq. (5.10-4) is essentially identical to the phase delay or eikonal function from geometric optics in Eq. (5.10-1). Therefore, when a stationary phase value for the spectral number exists, then the two systems give stationary values for the phase delay that are essentially the same

when the impact parameter in ray theory is set equal to that spectral number in wave theory.

Concerning the correspondence with the thin-screen models, we have seen in Chapter 2, Eq. (2.3-12), that the phase delay evaluated at the point $(\rho_*, \theta_{\rho_*})$ on the impact parameter space curve (with $\theta_{\rho_*} = \pi/2 - \alpha_{\rm L}(\rho_*)$) is given by

$$\varphi_{\rho_*} = \rho_* \alpha_{\rm L} + \int_{\rho_*}^{\infty} \alpha_{\rm L}(\omega) d\omega \tag{5.10-5}$$

The difference between this impact parameter phase delay and the stationary value of the spectral density function $\Phi_{\rm L}^*$ is simply due to the geometric delay (see Fig. 5-18) between the LEO and the point $(r_{\rho_*}, \theta_{\rho_*})$ on the impact parameter space curve, that is, $\rho_{\rm L} \cos(\theta_{\rm L} + \alpha_{\rm L})$.

Regarding the planar thin-screen model mounted perpendicularly to the LEO/GPS line, the Fresnel phase function in Chapter 2, Eq. (2.5-1) is given by

$$\Phi(h, h_{\rm LG}) = \frac{\pi}{\lambda D}(h - h_{\rm LG})^2 + \int_h^{\infty} \alpha_{\rm L}(\omega)\, d\omega \tag{5.10-6}$$

where h and D are in phase units. Referring to Eq. (5.10-4), the first term, $\rho_{\rm L} \cos(\theta + \alpha_{\rm L})$, is the optical distance from the LEO to the tangent point on a sphere of radius ρ_*, which is the impact parameter of the ray passing through the point $(\rho_{\rm L}, \theta_{\rm L})$. The second term is the optical distance $\rho_* \alpha_{\rm L}$ along the circumference of a sphere of radius ρ_* from $\theta = \pi/2 - \alpha_{\rm L}$ to $\theta = \pi/2$. The sum of these two terms is given by

$$\rho_{\rm L} \cos(\theta_{\rm L} + \alpha_{\rm L}) + \rho_* \alpha_{\rm L} = \rho_{\rm L} \cos\theta_{\rm L} \sec\alpha_{\rm L} - \alpha_{\rm L}^3 \rho_{\rm L} \sin\frac{\theta_{\rm L}}{3} + \rho_{\rm L}\, {\rm O}\left[\alpha_{\rm L}^4\right] \tag{5.10-7}$$

Thus, through second order in $\alpha_{\rm L}$, the first two terms in the stationary phase value for the spectral density function in Eq. (5.10-4) sum to the phase accumulation along the ray asymptote from the LEO to the line $\theta = \pi/2$.

The planar thin-screen distance D was rather loosely defined as the perpendicular distance from the LEO to the thin screen; see Fig. 5-18. In this figure, $h = \rho_{\rm L} \sin(\theta_{\rm L} + \alpha_{\rm L}) \sec\alpha_{\rm L} - \rho_o$ and $h_{\rm LG} = \rho_{\rm L} \sin\theta_{\rm L} - \rho_o$, where ρ_o provides the reference altitude. The impact parameter space curve is defined by the point $(\rho_*, \theta_{\rho_*})$ as ρ_* varies. Here $\theta_{\rho_*} = \pi/2 - \alpha_{\rm L}(\rho_*)$. If in Eq. (5.10-6) we now set $D = D_S = \rho_{\rm L} \cos\theta_{\rm L}$, it follows upon comparison with Eq. (5.10-4) that through second order in $\alpha_{\rm L}$ we have matched the geometric components of the phase delay in the two systems. This is about as close an agreement as we should expect because the version of the scalar diffraction integral that we have

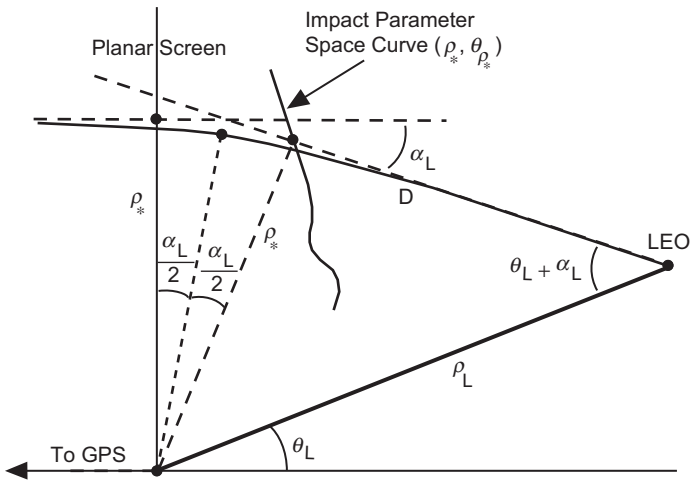

Fig. 5-18. Geometric relationships between ray path and phase screens for collimated incident rays.

used in Chapter 2 (the Rayleigh–Sommerfeld is typical) is accurate only through second order in α_L.

Next to be reconciled are the phase delays in the two systems induced by the refractivity gradient, which are the integral terms in Eqs. (5.10-4) and (5.10-6). We also have not been too specific about the orientation of the thin screen. For convenience here, we place it along the line $\theta = \pi/2$. Comparing these terms requires

$$\int_{h_S^*}^{\infty} \alpha_L(h_S')dh_S' = \int_{\rho_*}^{\infty} \alpha_L(\rho_*')d\rho_* \tag{5.10-8}$$

This must hold for all values of h_{LG}, the altitude of the LEO/GPS line in the planar thin-screen model above the reference plane, or, equivalently, for all locations (r_L, θ_L) of the LEO during the occultation. This condition may or may not be feasible, depending on whether or not h is a single-valued function of ρ_*. Differentiating with respect to ρ_* yields

$$\alpha_L(h)\frac{dh}{d\rho_*} = \alpha_L(\rho_*) \tag{5.10-9}$$

If no caustics occur in the screen, then at every altitude we can equate the bending angle in the thin screen at a thin-screen altitude h to the atmospheric bending angle $\alpha_L(\rho_*)$. If we place the thin screen along the line $\theta = \pi/2$, then it follows that

$$h = \rho_* \sec\alpha_{\mathrm{L}} - \rho_{\mathrm{L}} \sin\theta_o = \rho_* \sec\alpha_{\mathrm{L}} - \rho_o \qquad (5.10\text{-}10)$$

In a medium with large refractive gradients, thin-screen caustics where $dh/d\rho_* = 0$ can arise. The perpendicularly mounted thin screen is not suitable for these situations.

5.10.2 Duality between Systems

It follows from the previous discussions that when stationary values of the spectral density function exist with respect to spectral number, and when super-refractivity situations are avoided, then we have a kind of duality between the stationary phase processes over spectral number in wave theory and over impact parameter in geometric optics. For thin-atmosphere conditions away from turning points, rainbow caustics, etc., we may establish a correspondence between wave theory and geometric optics when stationary phase values in each system are assumed. For wave theory applied to a spherical refracting medium on one hand, and for the scalar diffraction integral applied to a thin-screen proxy for this refracting medium on the other hand, Table 5-1 shows the correspondence between these systems when stationary phase values are assigned in each system. Table 5-1 shows this correspondence at a LEO position. Here $G^\dagger = G^\dagger(\nu) \equiv G[\rho^\dagger(\nu), \nu]$, and $\rho^\dagger(\nu) = \nu - \hat{y}^\dagger K_{\rho^\dagger}$.

In summary, a thin phase screen model combined with scalar diffraction theory gives results that closely match those from wave theory applied to a spherical geometry when (1) caustics do not occur in the thin screen, $dh/d\rho_* \neq 0$, which is related to the thin-atmosphere conditions cited earlier, (2) the observer is relatively far from the refraction zone, and (3) certain LEO rainbow caustic and reflection locations are avoided.

On the other hand, there are situations where wave theory and wave/optics approaches give disparate results. One example is when no stationary spectral numbers occur; this corresponds to super-refractivity in geometric optics or to transition across a shadow boundary. Another example is a caustic contact point where second-order geometric optics predicts infinity for the amplitude of the ray and it errs in predicting the exact location of maximum flaring. Wave theory accurately addresses these cases. Although addressed later, one also can make a close correspondence between wave theory and geometric optics results when reflections occur.

Table 5-1. Duality between systems at the LEO.

	Electrodynamics in a Spherical Medium	⇔	Geometric Optics/Scalar Diffraction
Variational Parameter:	Spectral Number, ν		Impact Parameter, $kn_*r_* = \rho_*$
Atmospheric Phase Delay	$-\left(2G^{\dagger}\right)_{\nu=\nu^*} = -2G[\rho^{\dagger}(\nu^*), \nu^*]$	⇔	$\varphi(\rho_*)$, Thin Screen Phase
Stationary Phase Condition	$\left(\sin^{-1}\left(\frac{\nu}{\rho}\right) - \theta - 2\frac{dG^{\dagger}}{d\nu}\right)_{\nu=\nu^*}$	$= 0 =$	$\left(\alpha_L(\rho_*) + \frac{d\varphi}{d\rho_*}\right)$
Stationary Point	$\nu^* = \rho_L \sin\left(\theta_L + 2\frac{dG^{\dagger}}{d\nu}\right)_{\nu=\nu^*}$	⇔	$\rho_* = \rho_L \sin\gamma_L$
Gradient	$2\left(\frac{dG^{\dagger}}{d\nu}\right)_{\nu=\nu^*}$	⇔	$\alpha_L = 2\tilde{\alpha}(\rho_*, \rho_*)$, Bending Angle
Stationary Phase at LEO	$\left(D_\nu + 2\nu\frac{dG^{\dagger}}{d\nu} - 2G^{\dagger}\right)_{\nu=\nu^*}$	⇔	$\rho_L\cos\gamma_L + \rho_*\alpha_L + \int_{\rho_*}^{\infty}\alpha_L(v)dv$
Defocusing ζ	$\left(1 - 2D_\nu\frac{d^2G^{\dagger}}{d\nu^2}\right)^{-1}_{\nu=\nu^*}$	⇔	$\left(1 - \rho_L\cos\gamma_L\frac{d\alpha_L}{d\rho_*}\right)^{-1}$
Caustic $\zeta \to \infty$	$\left(1 - 2D_\nu\frac{d^2G^{\dagger}}{d\nu^2}\right) = 0$	⇔	$\left(1 - \rho_L\cos\gamma_L\frac{d\alpha_L}{d\rho_*}\right) = 0$

5.10.3 Amendments to Account for Wave-Front Curvature from the Finite Distance of the Emitting GPS Satellite

Almost all of the discussion in this monograph has assumed an infinite distance for the emitting GPS satellite. The wave-front curvature effects from an emitting GPS satellite at a finite distance can be accommodated as follows. From geometric optics, see Appendix A, Eq. (A-55), the phase delay along a ray, now referenced to the emitting GPS satellite, is given by

$$\varphi = \sqrt{\rho_{\mathrm{L}}^2 - \rho_*^2} + \sqrt{\rho_{\mathrm{G}}^2 - \rho_*^2} + \rho_* \alpha_{\mathrm{L}}(\rho_*) + \int_{\rho_*}^{\infty} \alpha_{\mathrm{L}}(\omega) d\omega \qquad (5.10\text{-}11)$$

Here it is assumed the emitter and the receiver are well out of any refracting medium. Otherwise, we would have to break the terms involving α_{L} into separate parts, one for δ_{L}, the ray path deflection angle at the LEO, and one for δ_{G}, the ray path deflection angle at the GPS, with $\alpha_{\mathrm{L}} = \delta_{\mathrm{L}} + \delta_{\mathrm{G}}$; see Fig. A-3.

This form for the phase delay in Eq. (5.10-11) is the eikonal function $\mathscr{S}(\boldsymbol{r})$ in geometric optics. The eikonal equation is $|\nabla \mathscr{S}| = n$, and a constant value for $\mathscr{S}$ defines the geometric optics equivalent of a surface of constant phase delay of the electromagnetic wave (essentially a wave-front or cophasal surface).

The main change in wave theory for a finite GPS distance is in the asymptotic form that the incoming spectral coefficients assume for large radial distances out of the atmosphere. This difference between collimated and spherical incident waves was briefly discussed in Section 5.5, Eqs. (5.5-3a) and (5.5-3b). When one accounts for wave-front curvature, the spectral density function $\Psi(+,-)$ for an outgoing wave at the LEO [see Eq. (5.9-6)] is amended to the form (always with $\theta_{\mathrm{G}} \equiv \pi$)

$$\left.\begin{aligned} \Psi &= \sqrt{\rho_{\mathrm{G}}^2 - \nu^2} + \sqrt{\rho_{\mathrm{L}}^2 - \nu^2} + \nu\left(\theta_\nu^{\mathrm{G}} + \theta_\nu^{\mathrm{L}} - \theta_{\mathrm{L}}\right) - 2G^{\dagger}(\nu) - \frac{\pi}{4} \\ \theta_\nu^{\mathrm{G}} &= \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{G}}}\right) \\ \theta_\nu^{\mathrm{L}} &= \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}}\right) \end{aligned}\right\} \qquad (5.10\text{-}12)$$

Then the stationary value of $\Psi(+,-)$, if one exists, is obtained by setting $\partial\Psi(+,-)/\partial\nu = 0$. This yields

$$\frac{\partial \Psi}{\partial \nu} = \left(\theta_\nu^{\mathrm{G}} + \theta_\nu^{\mathrm{L}} - \theta_{\mathrm{L}}\right) - 2\frac{dG^{\dagger}}{d\nu} = 0 \qquad (5.10\text{-}13)$$

Letting $\Phi_{\mathrm{L}}^{*} = \Psi(+,-)\big|_{\nu=\nu^{*}} + \pi/4$ and using Eq. (5.10-3), we obtain

$$\begin{aligned}\Phi_{\mathrm{L}}^{*} &= \sqrt{\rho_{\mathrm{G}}^{2} - \nu^{*2}} + \sqrt{\rho_{\mathrm{L}}^{2} - \nu^{*2}} + 2\left(\nu \frac{dG^{\dagger}}{d\nu}\right)_{\nu^{*}} - 2G^{\dagger}(\nu^{*}) \\ &\to \sqrt{\rho_{\mathrm{G}}^{2} - \rho_{*}^{2}} + \sqrt{\rho_{\mathrm{L}}^{2} - \rho_{*}^{2}} + \rho_{*}\alpha_{\mathrm{L}}(\rho_{*}) + \int_{\rho_{*}}^{\infty} \alpha_{\mathrm{L}}(\omega)\, d\omega\end{aligned} \tag{5.10-14}$$

The lower line is obtained by letting $\nu^{*} \to \rho_{*}$, and it is in agreement with the eikonal form in Eq. (5.10-11) given from geometric optics.

5.11 The Electric Field at a Turning Point

At a turning point, $r = r_{*}$ and $\theta = \theta_{*}$. We have discussed in Section 5.8 the breakdown in accuracy of the representation of the field by osculating parameters when the field is evaluated in the immediate neighborhood of a turning point. When certain scale factors permit, we can accurately approximate the medium by a spherical Airy layer in the neighborhood of a turning point. We have seen in Sections 4.7 and 4.8 for a Cartesian stratified medium that the osculating parameter approach works well at a turning point when the Airy functions are used as basis functions. For the boundary $r = r_A$ between the Airy layer below and the medium above, Section 5.8 uses the continuity conditions from electrodynamics to match the osculating parameter solution for $r \geq r_A$ to the Airy solution for $r \leq r_A$.

The Airy layer approximation works well when the effect of curvature in the refractivity profile is sufficiently small over the width of the layer. Two key scale factors can permit this approximation. The first is related to n''/n'. For an exponential medium, this factor becomes H^{-1}, where H is the scale height (here in length units). In this case, we want the width of the Airy layer Δr to be small enough that, in the power series expansion of the exponential representation of the medium, the quadratic and higher-order terms in $\Delta r / H$ are negligible, i.e., $\Delta r / H << 1$. On the other hand, we don't want the boundary $r = r_A$ of the Airy layer to be chosen so close to the turning point radius that the accuracy of the osculating parameter approach has begun to deteriorate, which occurs when $\hat{y} >\sim -2$. Using Eq. (5.4-3) to express $\Delta\hat{y}$ in terms of Δr over the width of the Airy layer, we have

$$\Delta\hat{y} = -\frac{kH}{K_{\rho_A}} \frac{\Delta r}{H} \tag{5.11-1}$$

Thus, the two dimensionless factors permitting an Airy layer approximation are kH and $K_{\rho_A} = (kn_A r_A / 2)^{1/3}$; the latter is related to r_A / λ. When H

corresponds to the Earth's dry air conditions ($H \approx 7\,\text{km}$) and for GPS wavelengths, $kH / K_{\rho_A} \approx 500$. In this case, we could set $\Delta r / H \sim 0.01$, achieving better than 0.01 percent accuracy in the representation of the refractivity by an Airy layer, with a boundary above which the osculating parameter approach also is sufficiently accurate. We also want to keep $\Delta\hat{y}$ small enough that phase runoff in the Airy function approximate solution, given in Eq. (5.8-6), as compared to the exact solution is negligible. This also is related to the size of n''/n'.

In matching the solution in the Airy layer involving the Airy function of the first kind with the osculating parameters/spherical Hankel functions that hold above the layer, we apply the continuity conditions from electrodynamics. This results in the matching coefficients c_l and d_l given in Eq. (5.8-8). We may multiply these coefficients by any complex factor that we wish [for example, $i^l(2l+1)/l(l+1)$], provided that we apply that factor to both coefficients or, equivalently, to both the incoming and outgoing components of the field. Also, at the boundary the outgoing phase term $2G\left[\rho^{\dagger}, \nu\right] - G[\rho_A, \nu]$, which is given in Eq. (5.8-24), must equal the exact value of the round-trip phase delay through the Airy layer. This fixes $\rho^{\dagger}$ at the Airy layer boundary per Eq. (5.8-15). For a given value for l, the asymptotically exact value for the round-trip phase delay through an Airy layer is given by Eqs. (5.8-12) and (5.8-13). Fortunately, we need not explicitly apply this condition; we need only the representation by the Airy function of the first kind at the turning point. But in practice the numerical integration of these wave equations has a stability problem below the turning point. The slightest phase error in the initial conditions at the boundary $r = r_A$ results in a blowup below the turning point or, equivalently, to a leakage of the Airy function of the second kind into the numerical solution. See the discussion of the WKB solutions and Eq. (5.7-26).

Near a turning point, we start with the integral version of the spectral representation for the field given in Eq. (5.8-1b). Because we are evaluating the field near a turning point, we do not have unrestricted use of the negative argument asymptotic form for the Airy function. But we still can use the negative exponential form for the Legendre function, which generally is applicable for large spectral numbers and for $0 < \theta < \pi$.

We replace the form $a_l^- \xi_l^- + a_l^+ \xi_l^+$ in Eq. (5.8-1) with $\text{Ai}[\tilde{y}]$ times a spectral number-dependent coefficient. When the negative argument asymptotic forms hold at the boundary of the Airy layer, we have from Eq. (5.8-8) the connecting relationship between the solution below the boundary and the solutions above the boundary. This is given by

$$\left.\begin{aligned} 2\left(\mathrm{Ai}[\tilde{y}]\right)_A &= \frac{(1+\varepsilon)^{1/6}}{\sqrt{\pi K_A}}\left(\xi_l^+ e^{+i\left(\tilde{\mathrm{X}}_l-\hat{\mathrm{X}}_l\right)}+\xi_l^- e^{-i\left(\tilde{\mathrm{X}}_l-\hat{\mathrm{X}}_l\right)}\right)_A \\ \tilde{X}_l &= \frac{2}{3}\left(-\tilde{y}_A\right)^{3/2}+\frac{\pi}{4} \\ \hat{X}_l &= \frac{2}{3}\left(-\hat{y}_A\right)^{3/2}+\frac{\pi}{4} \end{aligned}\right\} \qquad (5.11\text{-}2)$$

Here $\tilde{y}$ is given by Eq. (5.8-6), and $\tilde{y}=\hat{y}\left|1-\beta\right|^{-2/3}$. We also have seen in Section 5.8 that

$$\left(\tilde{X}_l-\hat{X}_l\right)_A = G[\rho_A,\nu]-G\left[\rho^\dagger,\nu\right] \qquad (5.11\text{-}3)$$

where $\rho^\dagger$ in this case is adjusted to force this equality to hold, which is a value very close to ν. In general, it follows from Eq. (5.8-8) that we may set

$$\left(a_l^-\xi_l^- + a_l^+\xi_l^+\right)_A = C_l\frac{2\sqrt{\pi K_{\rho_A}}}{\left|1-\beta\right|^{1/6}}\mathrm{Ai}\left[\tilde{y}_A\right] \qquad (5.11\text{-}4)$$

where C_l is chosen to account for the asymptotic boundary conditions on the wave. For an initially collimated wave, we use the asymptotic boundary conditions as $r\to\infty$ for a planar incoming wave. For this case, $a_l^-\left(\rho_A\right)$ is given by Eq. (5.5-21), and $a_l^+\left(\rho_A\right)$ is given by Eq. (5.8-24). It then follows from Eqs. (5.11-2) and (5.11-3) that C_l is given by

$$\left.\begin{aligned} C_l &= i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG^\dagger(\nu)\right) \\ G^\dagger(\nu) &= G\left[\rho^\dagger(\nu),\nu\right] \end{aligned}\right\} \qquad (5.11\text{-}5)$$

We now have the replacement form at the boundary of the Airy layer for an originally collimated incoming wave. It is given by

$$\left(a_l^-\xi_l^- + a_l^+\xi_l^+\right)_A = \frac{2\sqrt{\pi K_{\rho_A}}}{\left|1-\beta\right|^{1/6}}i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG^\dagger(\nu)\right)\left(\mathrm{Ai}[\tilde{y}]\right)_A \qquad (5.11\text{-}6)$$

We can use this form on the RHS to evaluate the field at any point within the Airy layer. For a given spectral number and radial position, $\tilde{y}$ is obtained from Eq. (5.8-6). Upon using the spectral form in Eq. (5.8-1b), we obtain

$$\left.\begin{aligned}
E_r(\tilde{\rho},\theta) &= \frac{|1-\beta|^{-1/6}}{K_{\rho_*}}\frac{E_o}{\sqrt{\sin\theta}}\int_0^\infty\left(\frac{\nu}{\rho}\right)^{3/2}\mathrm{Ai}[\tilde{y}]e^{i\left(\nu\left(\frac{\pi}{2}-\theta-G^\dagger(\nu)\right)\right)}d\nu,\\
E_\theta(\tilde{\rho},\theta) &= \frac{|1-\beta|^{-1/6}}{K_{\rho_*}^2}\frac{iE_o}{\sqrt{\sin\theta}}\int_0^\infty\left(\frac{\nu}{\rho}\right)^{1/2}\mathrm{Ai}'[\tilde{y}]e^{i\left(\nu\left(\frac{\pi}{2}-\theta-G^\dagger(\nu)\right)\right)}d\nu,\\
\tilde{y} &= -\frac{1}{K_{\rho_*}|1-\beta|^{2/3}}\left((1-\beta)(\tilde{\rho}-\rho_*)-(l-\rho_*)\right),\\
\tilde{\rho} &= kn_*r,\ \ \beta=-\left(\frac{1}{n}\frac{dn}{d\tilde{\rho}}\tilde{\rho}\right)_{\rho_*},\quad K_{\rho_*}=\left(\frac{\rho_*}{2}\right)^{1/3},\ \ \tilde{\rho}\le\tilde{\rho}_A
\end{aligned}\right\}\quad (5.11\text{-}7)$$

The constant gradient of the refractivity in the Airy layer, $n' = dn/d\tilde{\rho}$, is set by the actual values of $n(\rho)$ at the turning point r_* and at the boundary at $r = r_A$. Thus, the gradient of the refractivity will be discontinuous at the boundary, but refractivity itself will be continuous.

5.11.1 Fourier Transform of the Airy Function

We have seen in previous sections on stationary phase processes that, near a turning point and at the stationary phase value in spectral number, $dG\left[\rho^\dagger(\rho_*),\rho_*\right]/d\rho_* \doteq \tilde{\alpha}(\rho_*,\rho_*)$ with very high accuracy for thin-atmosphere conditions. But for now let us consider the integrals in Eq. (5.11-7) as Fourier transforms. When throughout the medium $n(\rho) = \text{constant}$, then $G[\rho,\nu] \equiv 0$ or at most a constant; it can be removed from the integrands in Eq. (5.11-7). From [8], we have the integral representation for the Airy function:

$$\mathrm{Ai}[\hat{y}] = \frac{(3a)^{1/3}}{2\pi}\int_{-\infty}^{\infty}\exp\left(i\left(at^3+(3a)^{1/3}\hat{y}t\right)\right)dt \qquad (5.11\text{-}8)$$

where a is any positive constant. It is easy to show that this integral form satisfies the differential equation for the Airy function $\mathrm{Ai}'' = \hat{y}\,\mathrm{Ai}$. Using the Dirac delta function,

$$\delta(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(i\omega t)dt \qquad (5.11\text{-}9)$$

it follows from Eqs. (5.11-8) and (5.11-9) that the Fourier transform of $\mathrm{Ai}[\hat{y}]$ is given by

$$\int_{-\infty}^{\infty}\mathrm{Ai}[\hat{y}]\exp(i\omega\hat{y})d\hat{y} = \exp\left(\frac{-i\omega^3}{3}\right) \qquad (5.11\text{-}10)$$

If we set $\omega = K_\rho(\pi/2-\theta)$ and we use the approximation $\nu \doteq \rho + K_\rho \hat{y} + \hat{y}^2/60K_\rho$ [see Eq. (5.4-3)], then we obtain

$$\int_{-\infty}^{\infty} \mathrm{Ai}[\hat{y}] e^{i\nu(\pi/2-\theta)} d\hat{y} = e^{i\rho\left(\pi/2-\theta-(\pi/2-\theta)^3/3!+(\pi/2-\theta)^5/5!\right)} \doteq e^{i\rho\cos\theta} \tag{5.11-11}$$

Thus, the Fourier transform of the Airy function of the first kind, at least near a turning point, is the phasor associated with the optical path length along the ray measured from the turning point, hardly a surprising result.

We can carry the Fourier transform approach a bit further and apply it to the radial component of the electric field in Eq. (5.11-1) for the case of the homogeneous medium where $G[\rho,\nu] \equiv 0$. Again using the Dirac delta function and its derivatives, it can be shown upon expanding the term $(l/\rho)^{3/2}$ in powers of $\hat{y}$ that

$$\frac{1}{\sqrt{\sin\theta}} \int_{-\infty}^{\infty} \left(\frac{l}{\rho}\right)^{3/2} \mathrm{Ai}(\hat{y}) e^{i\nu(\pi/2-\theta)} d\hat{y} \doteq \sin\theta e^{i\rho\cos\theta} \tag{5.11-12}$$

which we already know from physical considerations to be true.

Returning to the stratified case in Eq. (5.11-7), we now evaluate the radial component of the field at the turning point $E_r(\rho_*,\theta_*)$, where $\theta_* = \pi/2 - \tilde{\alpha}(\rho_*,\rho_*)$. We expand $G[\rho^\dagger,\nu]$ in powers of spectral number about the given value ρ_* where $(dG[\rho^\dagger(\nu),\nu]/d\nu)_{\rho_*} \doteq \tilde{\alpha}(\rho_*,\rho_*)$. We also make a change of variable in the integration from ν to $\tilde{y}$ using Eq. (5.11-7). We obtain

$$\left.\begin{aligned}
&E_r(\rho_*,\theta_*) \doteq E_o \sqrt{\frac{1-\beta}{\cos\tilde{\alpha}_*}} e^{i\Phi_{\rho_*}} \int_{-\infty}^{\infty} \mathrm{Ai}[\tilde{y}] \exp\left(-\frac{i}{2}\frac{d\tilde{\alpha}_*}{d\rho_*}(\nu-\rho_*)^2\right) d\tilde{y}, \\
&\Phi_{\rho_*} = \rho_*\tilde{\alpha}_* - G\left[\rho^\dagger(\rho_*),\rho_*\right], \quad \rho_* = kr_*n(\rho_*), \\
&\theta_* = \frac{\pi}{2} - \tilde{\alpha}_*, \quad \frac{d\tilde{\alpha}}{d\rho} = \frac{d\tilde{\alpha}(\rho,\rho)}{d\rho}, \quad \tilde{\alpha}_* = \tilde{\alpha}(\rho_*,\rho_*), \\
&\nu \doteq \rho_* + \tilde{y}(1-\beta)^{2/3} K_{\rho_*}, \quad \beta = -\left(\frac{1}{n}\frac{dn}{d\tilde{\rho}}\tilde{\rho}\right)_{\rho_*}, \quad \tilde{\rho} = kn_*r
\end{aligned}\right\} \tag{5.11-13}$$

We note the phase term Φ_{ρ_*}. It may be written as

$$\Phi_{\rho_*} = \rho_*\tilde{\alpha}(\rho_*,\rho_*) - G\left[\rho^\dagger(\rho_*),\rho_*\right] = \rho_*\tilde{\alpha}(\rho_*,\rho_*) + \int_{\rho_*}^{\infty} \tilde{\alpha}(\omega,\omega) d\omega \tag{5.11-14}$$

The first term is simply an optical path length along the circular arc of radius ρ_* measured from the reference point ($\theta = \pi/2$) to the angular position $\theta = \pi/2 - \tilde{\alpha}_*$. The second term, from Eq. (5.7-6), is the phase retardation induced by the gradient in the refractivity that the ray incurs in traveling through the atmosphere to the tangency point at r_*.

The quadratic term in the integral in Eq. (5.11-13) involving $d^2G/dv^2 = d\tilde{\alpha}/dv$ is related to the defocusing of the incident collimated wave resulting from the gradient of the refractivity.

5.11.2 Fresnel Transform of the Airy Function

To obtain the Fresnel transform of the Airy function, we again use the integral form for the Airy function given by Eq. (5.11-8). We obtain

$$\int_{-\infty}^{\infty} \mathrm{Ai}[\tilde{y}] \exp\left(i\left(\sigma\tilde{y} + \tau\tilde{y}^2\right)\right) d\tilde{y} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{iz^3/3}\left(\int_{-\infty}^{\infty} e^{i\left((z+\sigma)\tilde{y}+\tau\tilde{y}^2\right)} d\tilde{y}\right) dz \quad (5.11\text{-}15)$$

Here σ and τ are constants. From Eq. (5.11-13), it follows for our case that

$$\tau = -\frac{1}{2} K_{\rho_*}^2 (1-\beta)^{4/3} \frac{d\tilde{\alpha}}{d\rho_*} \quad (5.11\text{-}16)$$

Equation (5.11-15) is a more general version of the Fresnel transform; the latter is obtained by setting $\sigma = 0$. Upon completing the square in the inner integral of Eq. (5.11-15), one can write it in terms of the complete Fresnel integral. Then with a change of variable and the use again of the integral form for the Airy function given in Eq. (5.11-8), one finally obtains

$$\int_{-\infty}^{\infty} \mathrm{Ai}[\tilde{y}] \exp\left(i\left(\sigma\tilde{y} + \tau\tilde{y}^2\right)\right) d\tilde{y} = \sqrt{\frac{\pi}{|\tau|}}\, \mathrm{Ai}[b] \exp(ic) \quad (5.11\text{-}17)$$

where

$$\left.\begin{aligned} b &= -\frac{1+8\sigma\tau}{16\tau^2} \\ c &= \pm\frac{\pi}{4} + \frac{1}{192\tau^3} - \frac{(1+4\sigma\tau)^2}{64\tau^3} \end{aligned}\right\} \quad (5.11\text{-}18)$$

Here $+\pi/4$ is used in the expression for c when $\tau > 0$, and $-\pi/4$ is used when $\tau < 0$. Thus, the Fresnel transform of the Airy function yields an Airy function again but with offsets in phase and argument and with a modified amplitude. We note the defocusing term $\sqrt{|\tau|}$ appearing in the denominator of Eq. (5.11-17).

Using the Fresnel transform of the Airy function, the value of the integral in Eq. (5.11-13) becomes

$$E_r(\rho_*,\theta_*) = \frac{E_o|1-\beta|^{1/2}}{\sqrt{\cos\tilde{\alpha}_*}}\sqrt{\frac{\pi}{|\tau|}}\,\mathrm{Ai}[b]\exp\left(i\left(\Phi_{\rho_*}+c\right)\right) \qquad (5.11\text{-}19)$$

We note that, for thin-atmosphere conditions, $\tau << 1$. For an exponential atmosphere,

$$\tau = -\frac{1}{4}\left(\frac{\pi K_{\rho_*}}{kH}\right)^{1/2}\beta(1-\beta)^{4/3} \qquad (5.11\text{-}20)$$

For GPS signals with dry air at sea level, $\tau \sim 10^{-3}$. In this case, Eq. (5.11-17) essentially collapses to unity as $\tau \to 0$ and for $\sigma = 0$. In this case, Eq. (5.11-19) reduces to the form

$$E_r(\rho_*,\theta_*) = \frac{E_o|1-\beta|^{1/2}}{\sqrt{\cos\tilde{\alpha}_*}}\exp\left(i\Phi_{\rho_*}\right) \qquad (5.11\text{-}21)$$

We see the defocusing factor $|1-\beta|^{1/2}$ appearing here. Squaring this factor yields the defocusing in signal power, which essentially agrees with Eq. (5.6-16), the prediction from geometric optics for the defocusing at a turning point.

For E_θ, one can use the Fresnel transform of $\mathrm{Ai}'[\hat{y}]$ to derive similar results. E_θ should be essentially zero at a turning point. Also, the integral of $\mathrm{Ai}'[\hat{y}]$ over the real axis is zero.

5.12 Caustics and Multipath

Whether or not there is a unique ray, or possibly no ray, arriving at the point (r,θ) depends on the nature of the profile of $n(\rho)$, which may or may not result in caustics for the ray system and consequent multiple ray paths within the multipath zone. We have discussed caustics and their effects in Chapters 2 and 3, and they arise again here. These chapters also show examples of the converse of a multipath zone: the shadow zone, which in second-order geometric optics is a zone devoid of rays in the limit and extreme defocusing in actuality. A shadow zone based on geometric optics has an analog in wave theory. In shadow zones, wave theory using the spectral integrals in Eqs. (5.8-1) or (5.9-24) predicts much-diminished amplitudes for the field, but not zero. Also, wave theory accounts for diffraction. Using the stationary phase technique, it is readily shown that, in a strict shadow zone in geometric optics, the wave theory spectral integrals have no stationary phase points in spectral

number; thus, a small but non-zero fluctuating amplitude results. In the thin-screen model, geometric optics predicts no rays within a shadow zone, but the scalar diffraction integral from a thin screen yields a small but non-zero amplitude in this zone. The scalar diffraction integral applied to a thin screen integrates the path delay phasor over impact parameter space in the thin screen. Section 5.10 discussed the close correspondence between stationary phase values of spectral number in wave theory and stationary phase values of impact parameter in the scalar diffraction/thin screen. Under most conditions, the scalar diffraction/thin-screen model generates essentially the same values for the electromagnetic field at a point well away from the screen, including diffraction effects, that are obtained from wave theory using the spectral technique described here.

We have already shown a specific example without caustics in Fig. 5-10 for a monotonic profile for the refractivity, the exponential profile given in Eq. (5.8-2a). In that example, for a given position (r_o, θ), the stationary phase condition on wavenumber in Eq. (5.9-11) is satisfied uniquely at the intersection of the δ and $\partial G / \partial \nu$ curves versus ν. These curves are generated from Eqs. (5.9-14) and (5.7-7), respectively, for several values of θ, and they apply to an incoming region. The curve for $\partial G / \partial \nu$ (and implicitly the δ curve) in Fig. 5-10 uses a refraction profile that corresponds closely to the physical conditions for the dry air component of the Earth's atmosphere with r_o corresponding to sea level. The value of ν at the intersection point of the $\partial G / \partial \nu$ and δ curves is essentially the impact parameter at the Earth's surface. Fixing (r_o, θ) to a different pair of values would cause displacements of the δ curve (see Eq. (5.9-14)) and the $\partial G / \partial \nu$ curve (see Eq. (5.7-7)) and, therefore, to an intersection at a different bending angle $\tilde{\alpha}(\rho_o, \nu)$, which yields a different ν value.

5.12.1 A Numerical Example of Multipath and Caustics

As a simple first example, we consider the refractivity profile from Eq. (5.8-2b), where $n(\rho)$ has a non-super-refracting Gaussian distribution. A more realistic case, to be considered later, embeds this distribution in a background distribution corresponding to the near-exponential form for dry air. This dry air distribution causes significant defocusing, which has the effect of greatly compressing the bending angle spread from the multiple rays.

In the case of a Gaussian refractivity profile, $d\tilde{\alpha} / d\rho$ is not monotonic, and, therefore, caustics and multipath arise for certain positions (r, θ). We use the refractivity profile given by Eq. (5.8-2b) to obtain the electric field at the LEO. When this Gaussian profile is used in Eq. (5.6-5), a bipolar bending-angle profile results, $\tilde{\alpha}(\rho_*, \rho_*)$ (see Appendix E). The total refractive bending angle at the LEO is $\alpha_{\mathrm{L}}(\rho_*) = 2\tilde{\alpha}(\rho_*, \rho_*)$.

Now, we use this Gaussian profile for $n(\rho)$ in the integral expression for $2dG^{\dagger}(\nu)/d\nu = 2dG[\rho^{\dagger}(\nu),\nu]/d\nu$ given by Eqs. (5.7-7) and (5.7-13), and we generate the curve in Fig. 5-19 by varying ν. In this example, a particular set of values, $N_{\mathrm{W}} = 0.0001$, $H_{\mathrm{W}} = 1.6k$, and $\rho_w = \rho_o$, has been used. H_{W} is the $1-\sigma$ width of the Gaussian refractivity distribution. Here, $\rho_o \approx 2\times 10^8$, which corresponds to near sea level for a GPS wavelength. The peak refractivity is N_{W}, which is about 40 percent of the refractivity from dry air there.

We know from our earlier discussion that this $2dG^{\dagger}(\nu)/d\nu$ curve coincides with $2\tilde{\alpha}(\rho_*,\rho_*)$ to very high accuracy. From the stationarity condition on $\Psi(+,-)$ [see Eq. (5.9-6)] for an outgoing wave, we have, upon setting $\partial\Psi/\partial\nu = 0$, the relationships

$$\left.\begin{aligned} \delta &= \theta_{\nu} - \theta_{\mathrm{L}} = \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}}\right) - \theta_{\mathrm{L}} \\ \delta^* &= \theta_{\nu^*} - \theta_{\mathrm{L}} = 2\left(\frac{dG^{\dagger}(\nu)}{d\nu}\right)_{\nu^*} \end{aligned}\right\} \tag{5.12-1}$$

where ν^* denotes a stationary value of the spectral number. By varying ν while fixing $\rho_{\mathrm{L}} = kr_{\mathrm{L}} = 1.1kr_o$ and θ_{L} to specific orbital values, we generate the δ curves in Fig. 5-19 for three specific orbital values of θ_{L}. The intersections of the δ and $2dG[\rho^{\dagger}(\nu),\nu]/d\nu$ curves correspond to the δ^* and ν^* values given in Eq. (5.12-1), which are stationary phase values for $\Psi(+,-)$, i.e., $\partial\Psi/\partial\nu = 0$ at these points. The two particular δ curves tangent to the $2dG^{\dagger}/d\nu$ curve bound the range of θ_{L} values where the effects of multipath are evident. In this example, that range is about 17 mrad between the upper and lower caustic contacts for a typical LEO orbit. An intermediate value for θ_{L} yields a δ curve with three intersections with the $2dG^{\dagger}/d\nu$ curve and, therefore, three stationary phase points for $\Psi(+,-)$ in ν space. Within these two bounding δ curves, three rays with different bending angles from the GPS satellite will concurrently converge to the LEO. Outside of this zone, there is only a single ray and its corresponding bending angle.

The upper and lower δ curves, corresponding to tangency points on the $2dG^{\dagger}/d\nu$ curve with bending-angle values of 12.42 and -5.68 mrad, respectively, define the first contact points with the upper and lower caustic surfaces bounding this region. The θ_{L} values corresponding to these upper and lower tangency points are positions for the LEO where episodes of intense flaring can occur. Each tangency point is near the point of maximum observed amplitude. At these points, the tangency condition $2d^2G/d\nu^2 = \partial\delta/\partial\nu$ must

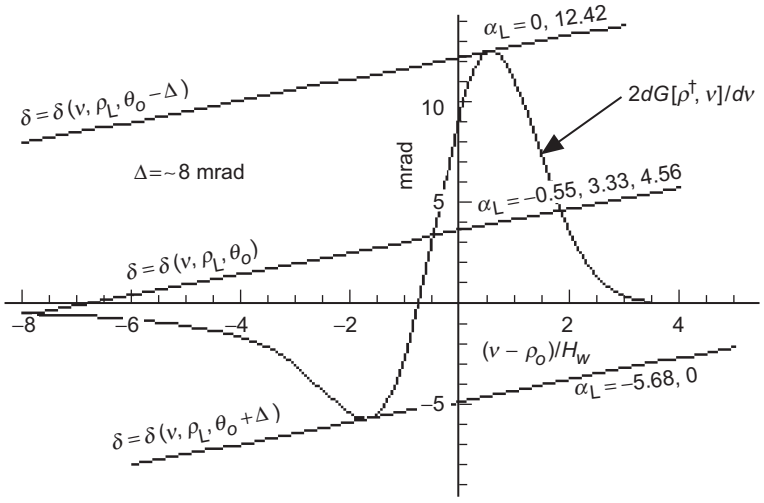

Fig. 5-19. Wave theory multipath zone in spectral number for a Gaussian refractivity profile. The phasor $\exp(i\Psi(+,-))$ **in the spectral integral is stationary at an intersection of the** δ **and** $2dG^\dagger/d\nu$ **curves, which gives the bending angle** $2\tilde{\alpha}$ **of the corresponding ray. Tangent points are caustic contacts where** $\partial\Psi/\partial\nu = \partial^2\Psi/\partial\,^2\nu = 0$.

hold. From Eq. (5.12-1), we have $\partial\delta / \partial\nu = 1/D_\nu$, where $D_\nu = \sqrt{\rho_L^2 - \nu^2}$. The defocusing factor at the LEO for a tangency point is given by $\zeta^{-1} = 1 - 2D_\nu d^2G/d\nu^2 = 0$ (see Eq. (5.9-29)), which defines a caustic contact point. Setting $\nu = \nu^* = \rho_*$, we conclude that a condition for no caustics is given by $2d^2G^\dagger / d\nu^2 \doteq 2d\tilde{\alpha}/d\rho_* < 1/D_{\rho_*}$ for all ρ_* values. Therefore, the gradient of the bending angle can be positive, but not too positive, less than about 0.3 mrad/ km for a GPS wavelength and with $r_L \approx 1.1 r_o$. We also note that, in cases where the refractivity gradient is continuous, the number of multipath rays is odd, except at a caustic contact point.

If these stationary phase points are adequately separated in wavenumber (see Appendix D for an "adequacy" index), then geometric optics as an approximate technique should suffice, which treats them as corresponding to separate rays. Geometric optics can be used with Eq. (5.6-5) to generate the ray for the impact parameter value $\rho_* = \rho_*(\nu^\dagger)$ corresponding to each of these intersection points. The phasors for three rays can be combined as vectors, i.e., taking into account their relative phases and amplitudes to obtain the aggregate electric field at the point (r_L, θ_L). However, near the caustic limits, two of these points will merge, and geometric optics will become inadequate for stationary phase points in spectral number that are in the neighborhood of the caustic contact point. In wave theory, the second-order stationary phase technique for

evaluating the integrals in Eq. (5.8-1b) over spectral number also becomes inadequate for the same reason. These spectral integrals should be directly evaluated, or a third-order stationary phase technique can be used, such as that described in Appendix D and discussed later in this chapter.

A caustic contact historically is associated with laser-like rays, for example, a rainbow spectral component from a water drop. Whether or not that kind of flaring in signal amplitude actually occurs at a caustic depends on the curvature of the $2dG^{\dagger}/d\nu$ curve; that is, it depends on the magnitude of $\partial^3\Psi(+,-)/\partial\nu^3$ at the caustic contact point $\nu = \nu^*$, where $\partial\Psi(+,-)/\partial\nu = \partial^2\Psi(+,-)/\partial\nu^2 = 0$. If the magnitude of $\partial^3\Psi(+,-)/\partial\nu^3$ is relatively small at the caustic point, then a broader neighborhood of spectral numbers about the stationary phase point contributes constructively to the spectral integral evaluation. This results in a larger amplitude. In other words, the spectral width about a stationary phase point across which phase coherence in the integrand is preserved depends inversely on $\partial^3\Psi(+,-)/\partial\nu^3$ when both $\partial\Psi(+,-)/\partial\nu$ and $\partial^2\Psi(+,-)/\partial\nu^2$ are at or near zero. We will show later that the amplitude of the flaring at a caustic contact point is proportional to $(\partial^3\Psi(+,-)/\partial\nu^3)^{-1/3}$. If the amplitudes of the new rays created at a caustic contact point are relatively strong, then signal flaring is to be expected even if interference is present from a pre-existing ray. If the amplitudes of the new rays are relatively weak, then the pre-existing ray will be dominant and the signal flaring, such as it is, will show up on the envelope of the resulting interference fringes. The Gaussian refractivity profile used in this section yields the latter scenario. Section 5.13 and Chapter 6 provide examples of the former scenario with very strong nascent rays.

Figure 5-20 shows the relationship between LEO orbital angle θ_{L} and impact parameter ρ_* from geometric optics for the same Gaussian refractivity profile and orbit model used in Fig. 5-19. This impact parameter diagram is obtained directly from Bouguer's law, $\theta_{\mathrm{L}} = \sin^{-1}(\rho_*/\rho_{\mathrm{L}}) - 2\tilde{\alpha}(\rho_*, \rho_*)$. Once the refractivity profile is specified, $\tilde{\alpha}(\rho_*, \rho_*)$ is obtained from Eq. (5.6-5). For $-12.2 \le \theta_{\mathrm{L}} - \theta_o \le 4.8$, a triplet of impact parameter values and bending angles simultaneously satisfies the boundary conditions provided by the position $(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ of the LEO. Significant multipath occurs within this zone. Caustic contacts lie at the points where $d\theta_{\mathrm{L}}/d\rho_* = 0$. These points correspond to episodes of signal flaring, and they mark the boundaries of the multipath zone, in this example at $\theta_{\mathrm{L}} - \theta_o = 4.8$ and at $\theta_{\mathrm{L}} - \theta_o = -12.2$. Above $\theta_{\mathrm{L}} - \theta_o = 4.8$, the rays labeled (a) and (b) are non-existent in a second-order geometric optics context. Only the ray labeled (m) exists. Similarly, rays (m) and (a) do not exist below $\theta_{\mathrm{L}} - \theta_o = -12.2$; only (b) rays exist.

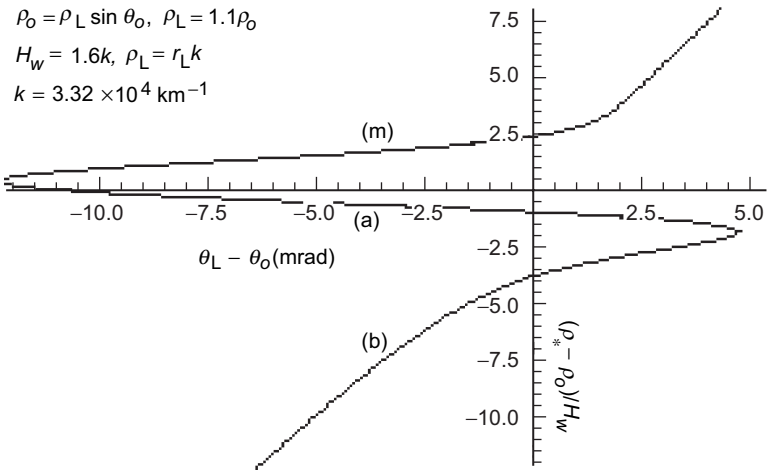

Fig. 5-20. Impact parameter diagram from geometric optics. This figure shows ρ_* versus LEO orbit angle θ_L in the neighborhood of the same refracting layer described in Fig. 5-19. Labels (m), (a), and (b) identify the corresponding rays: main, anomalous, and branching.

For the occulted GPS satellite lying in the orbit plane of the LEO and setting , θ_L decreases with time at a rate of roughly 1 mrad/s. Therefore, events unfold in Fig. 5-20 from right to left for a setting occultation. Above $\theta_L - \theta_o = 4.8$, only one stationary phase value in spectral number is realized. There is a unique value for any given orbit angle above this limit. This unique sequence of stationary phase values in spectral number versus θ_L above this limit corresponds in geometric optics to the impact parameter branch (m) in the figure, which defines the tangency points or impact parameter values generated by the (m) family of stationary phase paths, the (m) rays. When $\theta_L - \theta_o$ drops below 4.8 mrad, the first contact with a caustic is made. This results in flaring of the signal and the onset of interference from the triplet of competing stationary phase points in spectral number. In geometric optics, this corresponds to the existence of three competing ray systems, (m), (a), and (b); a member from each family passes through the position of the LEO. See Fig. 2-3 for a ray diagram. Ray (a) is the anomalous ray, and although it is a path of stationary phase, it can be shown that this path provides a local maximum in phase delay. Paths (m) and (b) provide local minima. This multipath episode continues until the lower caustic contact point is reached at $\theta_L - \theta_o = -12.2$. Below this point, a single stationary phase value in spectral number resumes, corresponding to the lower main ray (b) and the resumption of a smooth signal.

Figure 5-21(a) shows the signal amplitude $|\boldsymbol{E}(r_L, \theta_L)|$ at the LEO versus orbit angle $\theta_L - \theta_o$ for the same models used in Figs. 5-19 and 5-20. This figure

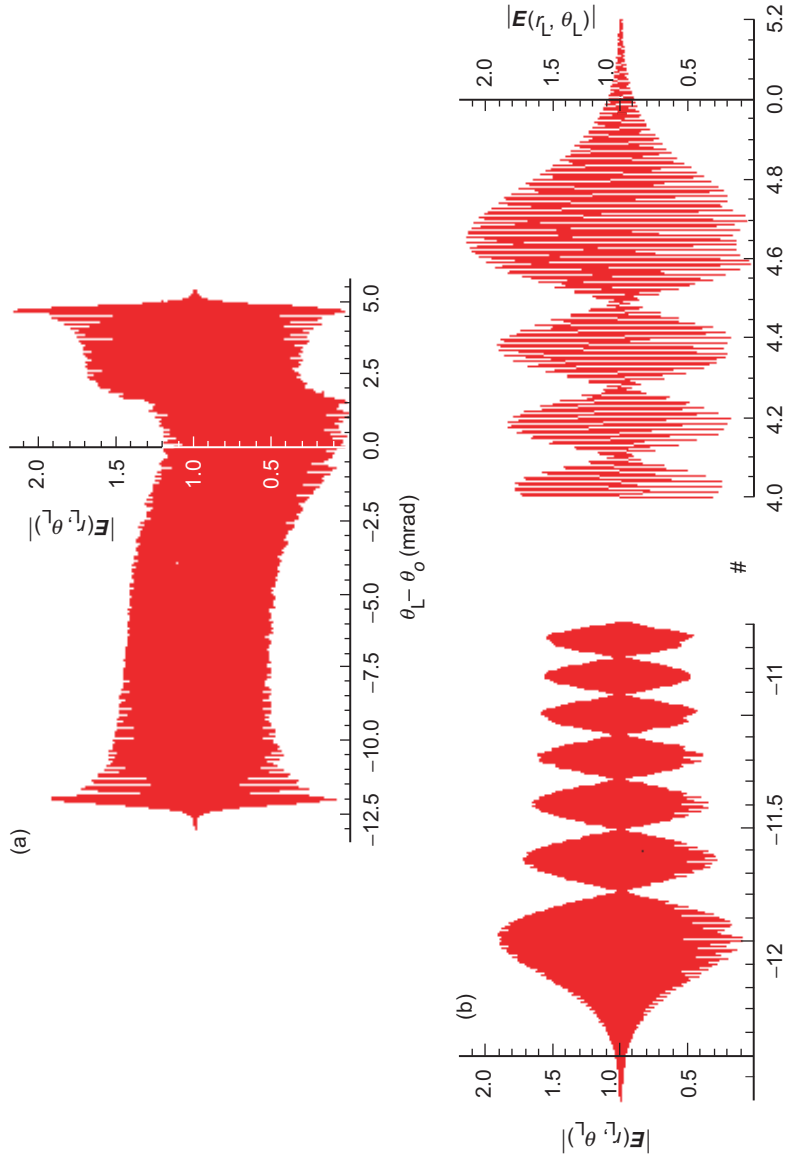

Fig. 5-21. Amplitude of the field at the LEO versus orbit angle for the same orbit model and Gaussian refractivity profile used in Figs. 5-19 and 5-20: (a) amplitude over the entire multipath zone and (b) amplitude in the vicinity of the upper and lower caustic contact points.

was obtained from a numerical integration of the spectral integral for E_r given in Eq. (5.8-1b) or, equivalently, Eq. (5.9-10). The numerical integration was aided by the stationary phase technique to isolate stationary phase neighborhoods, thereby reducing the computational burden[11] (see Section 2.6). The spectral coefficient $a_l^+(\rho)$ used in the integration is given in Eq. (5.8-25), and $G^\dagger(\nu) = G\left[\rho^\dagger(\nu), \nu\right]$ is given from Eq. (5.7-2), using the refractivity profile given in Eq. (5.8-2b) and the parameter values used for Figs. 5-19 and 5-20. Figure 5-21(a) shows the smooth and nearly constant signal amplitude outside of the multipath zone, and it clearly shows the abrupt onset of a multipath episode at the upper caustic contact point at about $\theta_L - \theta_o \approx 4.8$, continuing until the lower contact point at $\theta_L - \theta_o \approx -12.0$. This figure demonstrates the violent interference from a triplet set of stationary phase points that can result from a non-monotonic refractivity profile if the refractive gradient is even moderate in magnitude. Caustic flaring at the contact points is almost immediately overwhelmed by phase interference among the three full-fledged stationary phase paths. In this example, inspection of the gradients in bending angle in the middle of the multipath zone in Fig. 5-20 shows that none of the paths is overly defocused relative to its competitors. Defocusing is evident in the neighborhood of $\theta_L - \theta_o = 0$ because of the steep gradients in bending angle on all three rays there. Above $\theta_L - \theta_o = 5$, the principal contribution to $E(r_L, \theta_L)$ comes from the main ray (m) without interference from the (a) and (b) rays (see Fig. 5-20). Below this point, three mutually interfering ray paths contribute to the spectral integrals. Contact with the lower caustic occurs at $\theta_L - \theta_o \approx -12.2$. Abrupt termination of the scintillation episode follows. Below this only the (b) ray remains to contribute to $E(r_L, \theta_L)$.

Figure 5-21(b) is a blowup of the neighborhoods around the upper and lower caustic contact points to show the fast and slow fringes in amplitude of the field at the LEO. The high-frequency fringes from interference between rays are resolved at the upper contact point; at the lower point, their spacing is less than the resolution of the figure. The lower contact point shows somewhat less flaring than the upper point. Figure 5-20 shows that $d^2\theta_L / d\rho_*^2$ is somewhat larger at the lower point than at the upper point. It will be shown that $d^2\theta_L / d\rho_*^2 = \partial^3 \Psi(+,-) / \partial \nu^3$ when $\nu = \rho_*$. It follows that when $d^2\theta_L / d\rho_*^2$ is larger, the near-tangency condition between the $2dG^\dagger(\nu)/d\nu$ curve and the δ

[11] Here one uses the stationary phase technique to find all spectral numbers that yield a stationary value in $\Psi(+,-)$, in this example a maximum of three. Then in Eq. (5.9-24) it is necessary only to carry out a numerical integration over a suitable neighborhood around each of these spectral numbers, taking care to properly phase-connect these separate converged integration intervals.

curve near a caustic contact point runs over a shorter interval in spectral number. Therefore, a smaller contribution to the spectral integral for $E(r_L, \theta_L)$ comes from the spectral neighborhood around the lower caustic point in Fig. 5-21(b) than from the neighborhood around the upper caustic point. The peak amplitudes near caustic contacts are slightly offset, 0.1 to 0.2 mrad, from the geometric optics predictions for caustic contact (i.e., where $\zeta \to \infty$).

Figure 5-22 shows the difference in phase near the upper caustic point between the complete field at the LEO and the field from only the (m) ray. The spikes of large phase acceleration in this figure correspond to orbit angles in Fig. 5-21(b) where the complete field is almost totally annihilated through mutual interference between the three rays. The frequency of the phase fringes in Fig. 22 scales with impact parameter separation of the rays, which in turn scales roughly linearly with the $1-\sigma$ width of the Gaussian refractivity profile. Fast fringes are from interference between the main ray (m) and the nascent rays (a) and (b). Slow modulation or banding results from interference between the (a) and (b) rays as their impact parameters begin to separate near the caustic point.

5.12.2 A Geometric Optics Interpretation of Multipath

Geometric optics can accurately predict many of the fringe features for multipath situations, such as those shown in Figs. 5-21 and 5-23. One must take into account the amplitudes of the competing rays and their phase interference as the orbit angle or, equivalently, time evolves. A notable exception is the failure of geometric optics to accurately predict the field in the immediate

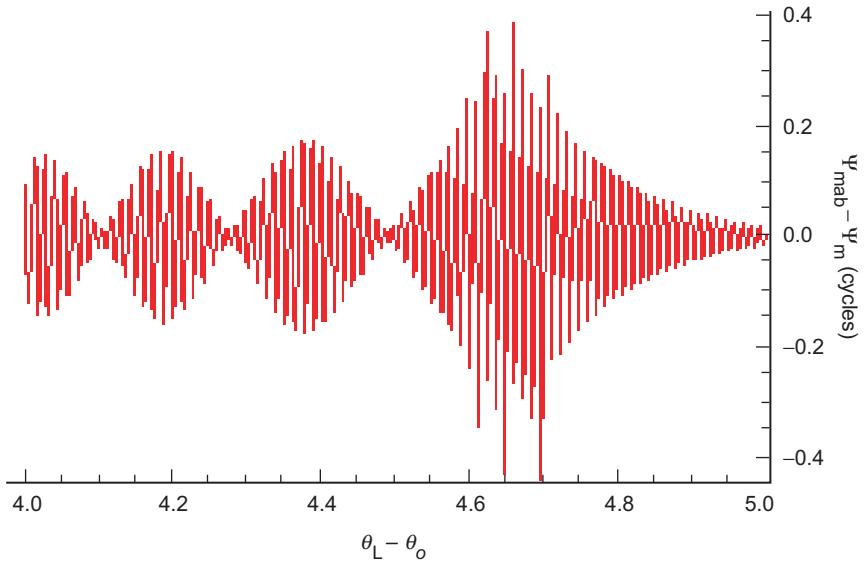

Fig. 5-22. Phase difference near the upper caustic, expressed in cycles of the complete field at the LEO minus the field of the (m) ray.

vicinity of a caustic contact. Second-order geometric optics predicts an infinite amplitude at such places, whereas wave theory gives the correct values.

We now apply geometric optics to see how well it does at predicting the fringe amplitudes and frequencies shown in Fig. 5-21. Starting with Bouguer's law, and given our specific refractivity and orbit models, the impact parameter curve for each competing ray system can be calculated. From the impact parameter diagram in Fig. 5-20, one can determine the defocusing of any ray versus orbit angle and also the interference frequencies between any two rays. To interpret the fringes, we need the relative phase changes and amplitudes of these competing rays evaluated at the LEO.

For the phase change or excess Doppler, we start with the stationary value of the spectral density of the phase at the LEO. From Eqs. (5.9-6) and (5.12-1), it is given by

$$\left.\begin{aligned} \Psi^*(+,-) &= D_{\nu^*} + \nu^*\left(\theta_{\nu^*} - \theta_{\mathrm{L}}\right) - 2G^{\dagger}\left(\nu^*\right) - \frac{\pi}{4}, \\ D_{\nu^*} &= \sqrt{\rho_{\mathrm{L}}^2 - \nu^{*2}}, \quad \nu^* = \rho_{\mathrm{L}} \sin\theta_{\nu^*} \end{aligned}\right\} \tag{5.12-2}$$

where $G^{\dagger}(\nu) = G[\rho^{\dagger}(\nu), \nu]$ is given in Eq. (5.7-2) with $\rho^{\dagger} = \nu - \hat{y}^{\dagger} K_{\rho^{\dagger}}$. Here we evaluate $\Psi(+,-)$ at a stationary value $\Psi^*(+,-)$ with respect to spectral number. That spectral number ν^* is allowed to assume the value of the impact parameter ρ_* at a given epoch for a specific ray, either (m), (a), or (b) as shown in Fig. 5-20. We know from spectral theory that ν^* will be very close to ρ_* for that ray, away from super-refracting conditions. Now we differentiate $\Psi^*(+,-)$ with respect to time. Both the orbital position of the LEO and the impact parameter change as time evolves.[12] Since we are interested only in the difference in phase between rays, we can neglect the Doppler term from the observed GPS satellite. Thus, we have

[12] We also should include the obliquity effect resulting from the GPS satellites generally not lying in the LEO orbit plane. This correction factor is discussed in Chapter 6. It can reduce Doppler values by up to about 35 percent for inclination angles up to 30 deg. Here we assume that the occulted GPS satellite lies in the orbit plane of the LEO. This assumption essentially impacts only the scale factor between LEO orbit angle change and elapsed time.

$$\begin{aligned}\frac{d\Psi^*}{dt} &= \frac{\partial\Psi^*}{\partial\rho_{\mathrm{L}}}\dot{\rho}_{\mathrm{L}} + \frac{\partial\Psi^*}{\partial\theta_{\mathrm{L}}}\dot{\theta}_{\mathrm{L}} + \frac{\partial\Psi^*}{\partial\nu}\dot{\rho}_* \\ &= \frac{D_{\rho_*}}{\rho_{\mathrm{L}}}\dot{\rho}_{\mathrm{L}} - \rho_*\dot{\theta}_{\mathrm{L}} + \dot{\rho}_*\left(\theta_\nu - \theta_{\mathrm{L}} - 2\frac{dG^\dagger}{d\nu}\right)_{\nu^*=\rho_*}\end{aligned} \tag{5.12-3}$$

But, at a stationary phase point in spectral number, $\theta_\nu - \theta_{\mathrm{L}} - 2dG^\dagger / d\nu = 0$, and for a circular orbit $\dot{\rho}_L \equiv 0$. It follows in this case that

$$\frac{d\Psi^*}{dt} = -\rho_*\dot{\theta}_{\mathrm{L}} \quad \text{or} \quad \frac{d\Psi^*}{d\theta_{\mathrm{L}}} = -\rho_* \tag{5.12-4}$$

Thus, $d(\Delta\Psi_*)/d\theta_L = -\Delta\rho_*$, where $\Delta\rho_*$ is the difference in impact parameter values between any two of the three rays, (m), (a), or (b), at a given epoch or orbit angle, and $\Delta\Psi_*$ is the difference in phase at the LEO between these two rays. These impact parameter differences can be read directly from Fig. 5-20. The high-frequency fringes in Fig. 5-21(b) near $\theta_{\mathrm{L}} - \theta_o = 5$ result from interference between the combined but relatively weak[13] (a) and (b) rays with the strong (m) ray. At $\theta_{\mathrm{L}} - \theta_o = 4.8$, the difference in altitudes of the impact parameters between the (m) ray and the nascent (a) and (b) rays is 16.3 km. Thus, the frequency of the fringes resulting from phase interference between these two ray systems is given by $(2\pi)^{-1} d\left(\Psi_{*\mathrm{m}} - \Psi_{*\mathrm{ab}}\right)/d\theta_{\mathrm{L}} = 16.3/\lambda$, or about 82 cycles/mrad. Although difficult to measure precisely in Fig. 5-21(b), the high-frequency fringe rate is indeed about 80 cycles/mrad. At the lower caustic contact point near $\theta_{\mathrm{L}} - \theta_o = -12$, the high-frequency fringes are caused by interference between the (b) ray and the dying (m) and (a) rays (or nascent rays for a rising occultation). Figure 5-20 shows a much wider separation in impact parameter values there between the (b) ray and the (m) or (a) ray. The high-frequency fringe rate is around 180 cycles/mrad. We deal with fringe frequency compression from defocusing later.

The low-frequency modulation in Fig. 5-21(b) comes from interference between the (a) and (b) rays near the upper caustic contact point at $\theta_{\mathrm{L}} - \theta_o = 4.8$, and between the (m) and (a) rays near the lower contact at $\theta_{\mathrm{L}} - \theta_o = -12$. These ray pairs have narrowly different impact parameter values

[13] Caustic rays are renowned for beaming like searchlights at their nascence, but in this example the (a) and (b) rays are relatively dim (a combined amplitude of 1.14 compared to the (m) ray amplitude of 1.0). This is because of the relatively large magnitude of $d^2\theta_{\mathrm{L}}/d\rho_*^2$ at the caustic contact point. The upper Gaussian wing of the refractivity profile leads to a rapid falloff in the gradient of the bending angle with increasing altitude and, therefore, to a main ray (m) not significantly defocused here.

at these locations in this figure, but they have about the same amplitude initially. For the right-hand side of Fig. 5-21(b) near $\theta_L - \theta_o = 4$, $(2\pi)^{-1} d(\Psi_a^* - \Psi_b^*)/d\theta_L$ is about 5 cycles/mrad, but this modulation frequency is accelerating rapidly with decreasing θ_L. We can infer this acceleration from Fig. 5-20, which shows the impact parameter values of these two rays separating rapidly with decreasing θ_L as they move away from the caustic contact point where these rays were created. At the lower caustic point, the low frequency and weakening modulation results from interference between the (m) and (a) rays, which are dying out completely below $\theta_L - \theta_o = -12$.

5.12.3 Amplitude Variability from Geometric Optics

Except in the immediate neighborhood of the caustic points, the amplitude variability shown in Fig. 5-21 can be predicted using geometric optics. In this approach, one adds up the complex amplitudes of the three rays taking into account their respective defocusing. The amplitude of the slow modulation in this figure can be obtained from the defocusing factor $\zeta^{-1} = 1 - 2D_{\rho_*} d\tilde{\alpha}_* / d\rho_*$ for each ray. The slow modulation peaks and valleys in amplitude shown in Fig. 5-21(b) are accurately predicted from the four combinations $|\zeta_m^{1/2} \pm \zeta_b^{1/2} \pm |\zeta_a|^{1/2}|$. ($\zeta_a$ is negative, a hallmark of an anomalous ray; the path delay along this ray is a local maximum.) For example, in Fig. 5-21(b) at $\theta_L - \theta_o = 4.2$, the (m) ray is still not significantly defocused; its amplitude is unity. For the (a) and (b) rays, we have $|\zeta_a|^{1/2} = 0.382$ and $\zeta_b^{1/2} = 0.461$. The four combinations yield $1 \pm 0.461 \pm 0.382 =$ 1.843, 0.157, 1.079, and 0.921. The corresponding peak values of the envelope in Fig. 5-21(b) are about 1.84 and 0.16 for the points of maximum amplitude difference, and 1.08 and 0.93 for the nodes. This is very close agreement, considering that the envelope isn't well defined in that figure.

At the upper caustic contact point in Fig. 5-21(b), $\zeta_{ab} \to \infty$; therefore, the agreement cited above must break down as we near such a point. From geometric optics, this occurs at $\theta_L - \theta_o = 4.77$, but the peak amplitude from wave theory in Fig. 5-21(b) occurs at $\theta_L - \theta_o = 4.65$. At the lower caustic point, $\zeta_{am} \to \infty$ at $\theta_L - \theta_o = -12.18$, but the peak amplitude from wave theory occurs at $\theta_L - \theta_o = -11.98$.

5.12.4 Third-Order Theory

The small offsets in location of the peak and the major discrepancy in amplitude prediction from geometric optics can be reconciled through use of a third-order stationary phase technique. Here, to evaluate the spectral integrals in

Eq. (5.9-10), we expand $\Psi(+,-)$ in spectral number about the caustic contact point where $\partial^2\Psi(+,-)/\partial\nu^2 = 0$, retaining third-order terms. We obtain

$$\left.\begin{aligned} &\Psi(+,-) \doteq \Psi|_{\nu_{\mathscr{C}}} + (\nu-\nu_{\mathscr{C}})\left.\frac{\partial\Psi}{\partial\nu}\right|_{\nu_{\mathscr{C}}} + \frac{(\nu-\nu_{\mathscr{C}})^3}{3!}\left.\frac{\partial^3\Psi}{\partial\nu^3}\right|_{\nu_{\mathscr{C}}} \\ &\left.\frac{\partial^2\Psi}{\partial\nu^2}\right|_{\nu_{\mathscr{C}}} = 0 \end{aligned}\right\} \tag{5.12-5}$$

Here $\nu_{\mathscr{C}}$ is the value of the spectral number at the caustic contact point. It satisfies the tangency condition between the $2dG^{\dagger}/d\nu$ and δ curves shown in Fig. 5-19, and we know that its value is close in value to $\rho_{\mathscr{C}}$, which is the impact parameter value at the caustic contact point. Therefore, for the purpose of evaluating the amplitude of the field at the LEO, we set $\nu_{\mathscr{C}} = \rho_{\mathscr{C}}$. The partial derivatives $\partial\Psi(+,-)/\partial\nu$ and $\partial^3\Psi(+,-)/\partial\nu^3$ are evaluated at that point. We obtain

$$\left.\begin{aligned} &\left.\frac{\partial\Psi}{\partial\nu}\right|_{\nu=\rho_{\mathscr{C}}} \hat{=} \Psi'_{\mathscr{C}} = \theta_{\rho_{\mathscr{C}}} - \theta_{\mathrm{L}} - 2\tilde{\alpha}_{\mathscr{C}}, \\ &\left.\frac{\partial^2\Psi}{\partial\nu^2}\right|_{\nu=\rho_{\mathscr{C}}} \hat{=} \Psi''_{\mathscr{C}} = \left.\frac{d\theta_{\mathrm{L}}}{d\rho_*}\right|_{\rho_*=\rho_{\mathscr{C}}} = \left(\frac{1}{D_{\rho_*}} - 2\frac{d\tilde{\alpha}}{d\rho_*}\right)_{\rho_{\mathscr{C}}} = 0, \\ &\left.\frac{\partial^3\Psi}{\partial\nu^3}\right|_{\nu=\rho_{\mathscr{C}}} \hat{=} \Psi'''_{\mathscr{C}} = \left.\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right|_{\rho_*=\rho_{\mathscr{C}}} = \left(\frac{\rho_*}{D_{\rho_*}^3} - 2\frac{d^2\tilde{\alpha}}{d\rho_*^2}\right)_{\rho_{\mathscr{C}}}, \\ &\tilde{\alpha}_{\mathscr{C}} = \tilde{\alpha}(\rho_{\mathscr{C}},\rho_{\mathscr{C}}),\ 2\tilde{\alpha} = \alpha_{\mathrm{L}} \end{aligned}\right\} \tag{5.12-6}$$

Here the relationships, $\Psi'' = d\theta_L/d\rho_*$ and $\Psi''' = d^2\theta_L/d\rho_*^2$, follow from Bouguer's law in Eq. (5.12-5). Differentiating that expression successively with respect to $\rho_{\mathscr{C}}$ yields the relationships given in Eq. (5.12-6). Since we seek only the modulation amplitude near the caustic contact point (but not the phase), we can simplify the spectral integrals in Eq. (5.9-10) by placing the slowly varying quantities, i.e., $(\sin^3\theta_\nu/\cos\theta_\nu)^{1/2}$ and $(\sin\theta_\nu\cos\theta_\nu)^{1/2}$, outside of the integrals and setting them equal to their values at the caustic contact point. When the radial and angular components of the field are combined, Eq. (5.9-10) yields the field at the LEO from the nascent rays near the caustic contact in the form

$$E_{\mathscr{C}}(r_{\mathrm{L}},\theta_{\mathrm{L}}) \doteq E_o\left(\frac{\tan\theta_{\mathscr{C}}}{2\pi\rho_{\mathrm{L}}\sin\theta_{\mathrm{L}}}\right)^{\frac{1}{2}} e^{i\Psi_{\nu_{\mathscr{C}}}} \int_0^{\infty} e^{i\left((\nu-\nu_{\mathscr{C}})\Psi'_{\nu_{\mathscr{C}}}+\frac{1}{3!}(\nu-\nu_{\mathscr{C}})^3\Psi'''_{\nu_{\mathscr{C}}}\right)} d\nu \qquad (5.12\text{-}7)$$

Here $\sin\theta_{\mathscr{C}} = \rho_{\mathscr{C}} / \rho_{\mathrm{L}}$, where $\rho_{\mathscr{C}}$ is the impact parameter at the caustic contact point. The subscript "$\mathscr{C}$" on $E_{\mathscr{C}}(r_{\mathrm{L}},\theta_{\mathrm{L}})$ is to remind us that it represents the nascent field only. The Taylor series expansion in Eq. (5.12-7) is not intended to span the spectral number range needed to include the stationary point corresponding to the main ray. Making a change of integration variable to $(\nu-\nu_{\mathscr{C}}) = \mathrm{sign}[\Psi'''_{\nu_{\mathscr{C}}}](2/|\Psi'''_{\nu_{\mathscr{C}}}|)^{1/3} t$, and noting that $\nu_{\mathscr{C}} \doteq \rho_{\mathscr{C}}$ is a very large number, we obtain for the amplitude of the nascent field at the LEO

$$\begin{aligned}\left|\boldsymbol{E}_{\mathscr{C}}(r_{\mathrm{L}},\theta_{\mathrm{L}})\right| &\doteq E_o\left(\frac{2\pi\tan\theta_{\mathscr{C}}}{\rho_{\mathrm{L}}\sin\theta_{\mathrm{L}}}\right)^{\frac{1}{2}}\left(\frac{2}{|\Psi'''_{\mathscr{C}}|}\right)^{\frac{1}{3}}\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left(i(zt+t^3/3)\right)dt = \\ E_o\left(\frac{2\pi\tan\theta_{\mathscr{C}}}{\rho_{\mathrm{L}}\sin\theta_{\mathrm{L}}}\right)^{\frac{1}{2}}&\left(\frac{2}{|\Psi'''_{\mathscr{C}}|}\right)^{\frac{1}{3}}\mathrm{Ai}[z] = E_o\left(\frac{2\pi\tan\theta_{\mathscr{C}}}{\rho_{\mathrm{L}}\sin\theta_{\mathrm{L}}}\right)^{\frac{1}{2}}\left(\frac{1}{2}\left|\frac{d^2\theta_{\mathrm{L}}}{d\rho_{\mathscr{C}}^2}\right|\right)^{-\frac{1}{3}}\mathrm{Ai}[z]\end{aligned} \qquad (5.12\text{-}8)$$

where z is given by

$$\begin{aligned} z &= \Psi'_{\mathscr{C}}\left(\frac{2}{|\Psi'''_{\mathscr{C}}|}\right)^{\frac{1}{3}}\mathrm{sign}\left[\Psi'''_{\mathscr{C}}\right] \\ &= \left(\theta_{\rho_{\mathscr{C}}} - \theta_{\mathrm{L}} - 2\tilde{\alpha}_{\mathscr{C}}\right)\left(\frac{1}{2}\left|\frac{d^2\theta_{\mathrm{L}}}{d\rho_{\mathscr{C}}^2}\right|\right)^{-\frac{1}{3}}\mathrm{sign}\left[\frac{d^2\theta_{\mathrm{L}}}{d\rho_{\mathscr{C}}^2}\right]\end{aligned} \qquad (5.12\text{-}9)$$

In Eq. (5.12-8), the integral form for the Airy function given in Eq. (5.11-8) has been used.

Equation (5.12-8) tells us a few things. First, near a caustic contact, the profile of the signal flaring envelope with orbit angle closely follows the shape of the Airy function of the first kind (first established by George Airy). Figure 5-23 provides a comparison of this third-order prediction of amplitude at the LEO from the nascent (a) and (b) rays with their actual combined amplitude from wave theory for the same case shown in the right side of Fig. 5-21(b). Here the contribution from the main ray (m) is removed. The shape of the third-order curve follows the absolute value of the Airy function. The agreement is very close in the neighborhood of the caustic contact point. Third-order theory predicts zero amplitude at the nodes, which actually occurs only when the (a) and (b) rays have identical amplitudes and opposite phase. In Fig. 5-23, we see

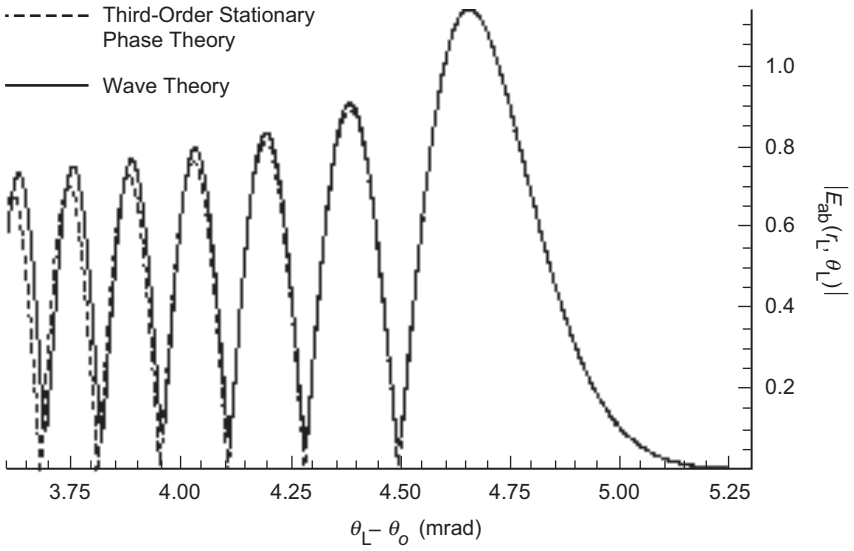

Fig. 5-23. Comparison of the amplitude of the field at the LEO from only the nascent rays (a) and (b) based on wave theory versus third-order stationary phase theory. Refractivity and orbit models are the same as those used in Figs. 5-19 through 5-22.

slight differences from the wave theory and also a gradual phase misalignment building up as θ_L recedes from the caustic contact. The accuracy of third-order theory with decreasing θ_L depends on the degree of symmetry in the θ_L versus ρ_* curve for these two nascent rays about the caustic point. If $|\, d^3\theta_L / d\rho_*^3 \,|_{\mathscr{C}}$ and the higher derivatives are relatively small, then the defocusing for the two nascent rays will remain roughly equal in magnitude as the impact parameters of the nascent rays move away from the caustic contact point. When close symmetry holds, these rays will continue with decreasing θ_L to mostly null each other at the nodes of the secondary lobes, and they will double each other's amplitude at the peaks. The Airy function approximation will be valid in this case over a wider range of θ_L values. Second-order geometric optics predicts an infinite amplitude at the caustic contact point at $\theta_L - \theta_o = 4.77$, which is 0.13 mrad greater than the actual location of the peak. But the accuracy of geometric optics improves rapidly away from the caustic contact point. Even at the first node, it accurately predicts its amplitude and location.

The second conclusion from Eq. (5.12-8) is that the amplitude of the nascent field at the LEO is proportional $(|\, d^2\theta_L / d\rho_*^2 \,|_{\mathscr{C}})^{-1/3}$, a quantity that is readily obtained from Bouguer's law and the bending-angle equation if the refractivity profile is given. We see that caustic flaring comes in two flavors. When $|\, d^2\theta_L / d\rho_*^2 \,|_{\mathscr{C}}$ is small enough, we will have strong flaring near a caustic

contact, on which high-frequency fringes from interference with the main ray will cause a minor modulation. Here there is a wider range of spectral numbers associated with the nascent rays that coherently contribute to the spectral integral. An example of this is shown in Section 5.13, Figs. 5-31 and 5-32. Rainbow caustics have a similar form. However, when $|d^2\theta_L / d\rho_*^2|_{\mathscr{C}}$ is larger, as is the case in Fig. 5-20, then there is a narrower range of spectral numbers contributing coherently to the spectral integral. In this case, the amplitude of the nascent field will be smaller; interference fringes from the main ray will become the dominant feature, with the nascent field providing the envelope of the amplitude fringes. Two examples of relatively weak caustics are shown in Fig. 5-21.

The third conclusion concerns the location of maximum amplitude. Neither $\Psi''_{\mathscr{C}}$ nor $\Psi'''_{\mathscr{C}}$ is variable; their values are set at the caustic contact point. But, z is a variable, and it depends on the value of θ_L. We now adjust θ_L to achieve the maximum value for $\left|\boldsymbol{E}_{\mathscr{C}}(r_L, \theta_L)\right|$ given in Eq. (5.12-8). The value $z = 0$ corresponds to the stationary phase point ($\theta_{\rho_{\mathscr{C}}} - \theta_{L\mathscr{C}} - 2\tilde{\alpha}_{\mathscr{C}} = 0$ in the second-order stationary phase formulation) also occurring at the caustic point where $1 - 2D_{\mathscr{C}}\tilde{\alpha}_{\mathscr{C}} = 0$, which is the prediction from geometric optics concerning maximum amplitude (i.e., infinity). This condition yields a determination of $\theta_{L\mathscr{C}}$, which is at $\theta_o + 4.77$ for the upper caustic contact point, and at $\theta_o - 12.18$ for the lower point. But, we see that the Airy function achieves greater values away from the stationary phase point at the caustic, which has a value of $\mathrm{Ai}[0] = 0.355$. The maximum value of $\mathrm{Ai}[z]$ nearest to $z = 0$ is 0.536, and it occurs at $z = -1.019$. If we let $\hat{\theta}_L$ be the value of θ_L where $z = -1.019$, then the offset between the geometric optics prediction of the location $\theta_{L\mathscr{C}}$ of maximum amplitude and the third-order stationary phase prediction, $\hat{\theta}_L$, is given by

$$\hat{\theta}_L - \theta_{L\mathscr{C}} = 1.019\left(\frac{1}{2}\left|\frac{d^2\theta_L}{d\rho_*^2}\right|\right)_{\mathscr{C}}^{-(1/3)} \operatorname{sign}\left[\frac{d^2\theta_L}{d\rho_*^2}\right]_{\mathscr{C}} \tag{5.12-10}$$

This expression predicts offsets in Fig. 5-21 for the upper caustic point at $\theta_L - \theta_o = 4.77$ of $\hat{\theta}_L - \theta_{L\mathscr{C}} = -0.13$ mrad, or $\hat{\theta}_L - \theta_o = -4.64$. At the lower caustic point at $\theta_L - \theta_o = -12.18$, the offset is +0.16 mrad, or $\hat{\theta}_L - \theta_o = -12.02$. Figure 5-21(b) shows that these predictions are very close to the actual offsets obtained from wave theory.

The maximum amplitudes predicted by Eq. (5.12-8) for the nascent field at the LEO are 1.09 near $\theta_L - \theta_o = 4.8$ and 0.90 near $\theta_L - \theta_o = -12$. The actual values are 1.14 and 0.91. Because the amplitudes of the nascent fields are

comparable to the main ray, the (m) ray at the upper caustic and the (b) ray at the lower caustic, there will be high-frequency fringes of comparable amplitude superimposed on this nascent radiation. Thus, it is the envelope of the field at the caustic points that we should examine. Setting the field amplitude of the main ray to unity, Eq. (5.12-10) predicts the peaks in the envelope to be 2.09 maximum and –0.09 minimum near $\theta_{\rm L} - \theta_o = 4.7$, and 1.90 and +0.10 near $\theta_{\rm L} - \theta_o = -12$. Since we are displaying amplitudes here, the negative minimum peak near $\theta_{\rm L} - \theta_o = 4.8$ is "reflected" about the $\theta_{\rm L}$ axis and becomes +0.09. These compare very closely, if not perfectly, with the actual peaks in the lobe nearest the caustic contact points.

An index to characterize the two flavors of caustic flaring, strong and weak, can be formulated. We take the ratio $E_{\mathscr{C}}(r_{\rm L}, \theta_{\rm L}) / E_o \zeta_{\rm m}^{1/2}$, which is the ratio of the amplitude of the field at the LEO from the nascent (a) and (b) rays at the caustic point $\theta_{\rm L} = \theta_{{\rm L}\mathscr{C}}$ divided by the field from the main ray (m) at the same orbit angle. It follows from Eqs. (5.12-6) and (5.12-8) that this ratio is given by

$$\left\{ \mathscr{R} = \frac{E_{\mathscr{C}}}{E_o \zeta_{\rm m}^{1/2}} = \sqrt{2\pi}\ \mathrm{Ai}[-1.02] \left| \frac{d\theta_{\rm L}}{d\rho_*} \right|_{\rm m}^{1/2} \left| \frac{1}{2} \frac{d^2\theta_{\rm L}}{d\rho_*^2} \right|_{\mathscr{C}}^{-(1/3)} \right. \tag{5.12-11}$$

If $\left| d^2\theta_{\rm L} / d\rho_*^2 \right|_{\mathscr{C}} >> 5 \left| d\theta_{\rm L} / d\rho_* \right|_{\rm m}^{3/2}$, then $\mathscr{R} << 1$, and vice versa.

Even the fine structure in the minimum envelope in Fig. 5-21(b) at $\theta_{\rm L} - \theta_o = 4.7$ is predicted well by Eq. (5.12-8). As $\theta_{\rm L}$ is moved away from its maximum amplitude point at $\hat{\theta}_{\rm L}$, z varies and $\mathrm{Ai}[z]$ diminishes. If we adjust z so that the amplitude of the nascent field at the LEO is unity, this will mark a pair of points along the $\theta_{\rm L}$ axis where the nascent field and the main field from the (m) ray can completely cancel each other. Setting $E_{\mathscr{C}}(r_{\rm L}, \theta_{\rm L}) = 1$ yields z values of –1.40 and –0.58. This yields $\theta_{\rm L} - \hat{\theta}_{\rm L} = +0.05, -0.06$, or $\theta_{\rm L} - \theta_o = 4.69,\ 4.58$, which corresponds almost exactly with the $\theta_{\rm L}$ values where cancellation is almost complete.

The first node in the envelope nearest the peak corresponds in Eq. (5.12-8) to the first zero of $\mathrm{Ai}[z]$, which occurs at $z = -2.34$. This yields $\theta_{\rm L} - \hat{\theta}_{\rm L} = -0.16$, or $\theta_{\rm L} = \theta_o + 4.48$, which is very close to the location of the actual node in Fig. 5-21(b), $\theta_{\rm L} = \theta_o + 4.49$. Here third-order theory predicts zero for the nodal amplitude, but we see a slight departure from this value. The accuracy of the third-order theory deteriorates the further away from the caustic point one applies it. Actually, third-order theory does well for several lobes at predicting amplitude peaks and nodes, but eventually it falls out of phase with wave theory results.

The defocusing factors of the (a) and (b) rays at the first node at $\theta_{\mathrm{L}} = \theta_o + 4.49$ are given by $|\zeta_{\mathrm{a}}|^{1/2} = 0.462$ and $\zeta_{\mathrm{b}}^{1/2} = 0.526$. This node marks the point where the (a) and (b) rays first become 180 deg out of phase. The difference of these defocusing factors is 0.06, which is very close to the actual nodal amplitude at $\theta_{\mathrm{L}} = \theta_o + 4.49$.

One also can predict the location of the first node from geometric optics. We know from Eq. (5.12-2) that at $\theta_{\mathrm{L}} = \theta_{\mathrm{L}\mathscr{C}}$ the phase function difference $\Psi_{\mathrm{a}}^* - \Psi_{\mathrm{b}}^* = 0$ because, according to geometric optics, the (a) and (b) rays are merged at that caustic contact point. We now adjust θ_{L} away from the caustic point until the phase difference is $\Psi_{\mathrm{a}}^* - \Psi_{\mathrm{b}}^* = \pi$, exactly. The expressions for Ψ^* in Eq. (5.12-2) can be used for this calculation, although care in taking the difference must be exercised because of the enormity of Ψ^* itself. When this program is undertaken for the example shown in Fig. 5-21(b), one obtains $\theta_{\mathrm{L}} = \theta_o + 4.56$ for the location of the node where the (a) and (b) rays have exactly opposite phase, a somewhat less accurate result than the third-order stationary phase prediction, but still a rather good prediction.

An alternate approach is to use the third-order expansion for Ψ^* around the caustic contact point $(\rho_{\mathscr{C}}, \theta_{\mathrm{L}\mathscr{C}})$ in the impact parameter diagram (see Fig. 5-24). We use this expansion to calculate the separation distance $\rho_{\mathrm{a}} - \rho_{\mathrm{b}}$ and the

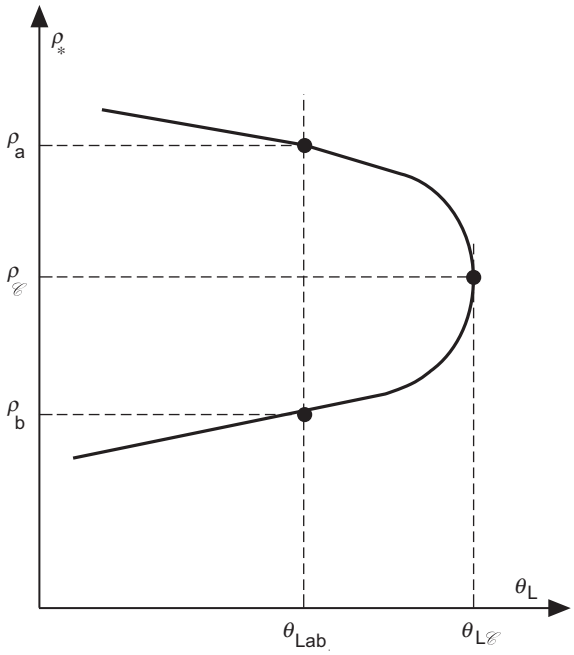

Fig. 5-24. Quadratic behavior of ρ_* for nascent rays near caustic.

change in orbit angle $\theta_{\mathrm{La}} - \theta_{\mathrm{Lb}}$ required to achieve exactly π rad phase difference between the (a) and (b) rays at the LEO since their nascence. At the caustic contact point $\theta_{\mathrm{L}} = \theta_{\mathrm{L}\mathscr{C}}$, and from Eqs. (5.12-5) and (5.12-6), it follows that $\Psi'_{\mathscr{C}} = 0$ and $\Psi''_{\mathscr{C}} = \left(d\theta_{\mathrm{L}} / d\rho_*\right)_{\mathscr{C}} = 0$. Therefore, we may write the difference in phase between rays (a) and (b) in terms of the third-order expansion as

$$\begin{aligned} \Psi_{\mathrm{a}} - \Psi_{\mathrm{b}} &\doteq \frac{1}{3!}\left(\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right)_{\mathscr{C}} \left((\rho_{\mathrm{a}} - \rho_{\mathscr{C}})^3 - (\rho_{\mathrm{b}} - \rho_{\mathscr{C}})^3\right) \\ &= \frac{1}{24}\left(\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right)_{\mathscr{C}} (\rho_{\mathrm{a}} - \rho_{\mathrm{b}})^3 = \pi \end{aligned} \tag{5.12-12}$$

It follows that the separation distance in impact parameters of the (a) and (b) rays at their first null since nascence, $\Delta\rho_{*\mathrm{ab}}$, is given by

$$\Delta\rho_{*\mathrm{ab}} = (\rho_{\mathrm{a}} - \rho_{\mathrm{b}})\big|_{\Delta\Psi_{\mathrm{ab}} = \pi} = (24\pi)^{1/3} \left|\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right|_{\mathscr{C}}^{-(1/3)} \mathrm{sign}\left[\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right]_{\mathscr{C}} \tag{5.12-13}$$

For $\Delta\theta_{\mathrm{L}} = (\theta_{\mathrm{L}} - \theta_{\mathrm{L}\mathscr{C}})\big|_{\Delta\Psi_{\mathrm{ab}} = \pi}$, we obtain

$$\begin{aligned} \Delta\theta_{\mathrm{L}} &\doteq \frac{1}{2}\left(\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right)_{\mathscr{C}} (\rho_{\mathrm{a}} - \rho_{\mathscr{C}})^2 \doteq \frac{1}{2}\left(\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right)_{\mathscr{C}} (\rho_{\mathrm{b}} - \rho_{\mathscr{C}})^2 \\ &\doteq \frac{1}{8}\left(\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right)_{\mathscr{C}} (\rho_{\mathrm{a}} - \rho_{\mathrm{b}})^2 = \frac{1}{2}(3\pi)^{2/3} \left|\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right|_{\mathscr{C}}^{1/3} \mathrm{sign}\left[\frac{d^2\theta_{\mathrm{L}}}{d\rho_*^2}\right]_{\mathscr{C}} \end{aligned} \tag{5.12-14}$$

We can compare the coefficients in Eqs. (5.12-9) and (5.12-14) for predicting the location of the first node. The coefficient in Eq. (5.12-9) for $\Delta\theta_{\mathrm{L}} = \theta_{\mathrm{L}} - \theta_{\mathrm{L}\mathscr{C}}$ is 2.34, which corresponds to the first zero of the Airy function Ai[z]. Equation (5.12-9) predicts the location of the node at $\theta_{\mathrm{L}} = \theta_o + 4.48$, which is very close to the actual first node shown in Fig. 5-23. Equation (5.12-14) gives a coefficient of $(3\pi)^{2/3} / 2 = 2.23$, and it predicts the location at $\theta_{\mathrm{L}} = \theta_o + 4.50$.

We recall that the semi-minor axis of the first Fresnel zone in second-order geometric optics is the separation distance of the impact parameters for two rays that arrive at the LEO exactly π rad out of phase. Thus, at a point where $\Psi'' \neq 0$, one obtains

$$\Psi_{\mathrm{a}} - \Psi_{\mathrm{b}} \doteq \frac{1}{2}\Psi''(\rho_{\mathrm{a}} - \rho_{\mathrm{b}})^2 = \pi \tag{5.12-15}$$

or

$$k\mathscr{F} = \left(\rho_{\mathrm{a}} - \rho_{\mathrm{b}}\right)_{\Delta\Psi=\pi} = \left(\frac{2\pi}{|\Psi''|}\right)^{1/2} = \left|\frac{1}{2\pi}\frac{d\theta_{\mathrm{L}}}{d\rho_*}\right|^{-(1/2)}$$
$$= \left(\frac{2\pi D_{\rho_*}}{\left|1 - 2D_{\rho_*}\tilde{\alpha}'_*\right|}\right)^{1/2} = \left(2\pi D_{\rho_*}|\zeta_*|\right)^{1/2} \tag{5.12-16}$$

At any rate, $\mathscr{F}$ is infinite at the caustic point. For comparison, for the example given in Fig. 5-21 at the upper caustic point at $\theta_{\mathrm{L}} = \theta_o + 4.47$, $k^{-1}\Delta\rho_{*\mathrm{ab}}$ in Eq. (5.12-13) has a value of 880 m. In a vacuum, $\mathscr{F}$ is about 750 m at the LEO radial distance of $r_{\mathrm{L}} = 1.1r_o$, whereas at sea level for a defocusing dry air medium, $\mathscr{F}$ is about 250 m. We show in the next section that the defocusing medium significantly compresses the separation distance between the impact parameters of caustic rays (see Fig. 5-26). At sea level for dry air $k^{-1}\Delta\rho_{\mathrm{ab}}$ is reduced to about 290 m. The offset $\Delta\theta_{\mathrm{L}}$ between the primary peak amplitude at the upper caustic point and the first node is about 0.29 mrad in Fig. 5-21 and about 0.80 mrad in a defocusing dry air medium (see Fig. 5-26). (This corresponds to roughly 300 ms and 800 ms of elapsed time, respectively, plenty of time for a phase-locked loop GPS receiver to decide that it has got the right stuff!).

From these and previous discussions, we conclude that even at the first node after the caustic point, second-order ray optics does a good job at predicting the field amplitude, but it gives a somewhat less accurate prediction of the phase. This agreement with wave theory results improves as the impact parameters for the nascent rays recede further from the caustic point.

In summary, geometric optics works well for predicting the amplitude of the field in multipath conditions provided that (1) the separation in altitudes between impact parameters of the competing rays exceeds a certain minimum distance and (2) caustic contact neighborhoods are avoided. Condition (1) is equivalent to the requirement in spectral theory that the stationary points in spectral number be sufficiently separated. When this is satisfied, the spectral integrals, one for each stationary phase point, do not significantly co-mingle; that is, the spectral neighborhood providing significant contributions to the spectral integral from one ray does not overlap the spectral neighborhood for the other ray. This is equivalent to saying that the two rays corresponding to these two spectral integrals are separated adequately in impact parameter altitude. An accuracy index based on this separation concept can be derived from second- and third-order stationary phase theory. This has been given in Appendix D. One can derive from that discussion an accuracy-dependent index for minimum separation altitudes of the impact parameters (or spectral numbers in wave theory) for which second-order ray theory will be adequate. For a given

accuracy, it can be shown that this index is proportional to $\left(d\theta_{\mathrm{L}} / d\rho_{*}\right)^{-1/2}$, or to the first Fresnel zone. From the agreements in amplitude fringes and modulations discussed above, we also can infer that the relative phases between rays are handled accurately by geometric optics away from caustic neighborhoods. Getting absolute phase from ray theory to agree closely with the phase from wave theory is somewhat more challenging because of the extreme sensitivity of absolute phase to boundary conditions; this results from the fact that $r_o / \lambda >> 1$. But, even the phases between the two systems can be aligned by renormalization.

5.12.5 Reduction of Multipath Spectral Width by Defocusing

The frequencies of the high-frequency interference fringes shown in Figs. 5-21 and 5-22 have a temporal equivalent in the range from 80 to 180 Hz. These high frequencies result from using a solitary Gaussian refraction model in these figures without the presence of a defocusing background medium. These frequencies exceed by an order of magnitude the actual maximum bandwidth (~10 to 15 Hz) of transient signatures in the excess Doppler observed by a LEO sounding the Earth's atmosphere. This disparity in the fringe frequency bandwidth appearing in Figs. 5-21 and 5-22 versus realistic bandwidths is largely due to the defocusing coming from the gradient of the dry air refractivity, which was omitted in these early figures. In a background medium that is defocusing, the impact parameter spread between multipath rays is greatly compressed. We present now a simple expression that accounts for the defocusing, and from which the qualitative aspects of the interference fringes from a Gaussian model with dry air added can be inferred.

From Bouguer's law given in Eq. (5.6-5), we have

$$\left.\begin{aligned} \theta_{\mathrm{L}} - \theta_o &= \sin^{-1}\left(\frac{\rho_*}{\rho_{\mathrm{L}}}\right) - \sin^{-1}\left(\frac{\rho_o}{\rho_{\mathrm{L}}}\right) - 2\left(\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_o, \rho_o)\right) \\ &\doteq D_{\rho_o}^{-1}(\rho_* - \rho_o) - 2\left(\tilde{\alpha}(\rho_*, \rho_*) - \tilde{\alpha}(\rho_o, \rho_o)\right) \\ D_{\rho_o} &= \rho_{\mathrm{L}} \cos\left(\theta_o + 2\tilde{\alpha}(\rho_o, \rho_o)\right) \end{aligned}\right\} \qquad (5.12\text{-}17)$$

It follows that the shape of the θ_{L} versus ρ_* curve, such as that shown in Fig. 5-20, is essentially determined by the shape of $\tilde{\alpha}(\rho_*, \rho_*)$.

Suppose that $\tilde{\alpha}(\rho_*, \rho_*)$ is composed of two parts. The first part is from a refractivity gradient for a local medium arising from, for example, a localized water vapor layer. The second part is from a smooth background refractivity

gradient arising from, for example, dry air. We write $\tilde{\alpha} = \tilde{\alpha}_{\mathrm{W}} + \tilde{\alpha}_{\mathrm{A}}$.[14] We expect that the rapid variations and the caustic points in $\theta_{\mathrm{L}}(\rho_*)$ come from rapid variations in $\tilde{\alpha}_{\mathrm{W}}$. On the other hand, $\tilde{\alpha}_{\mathrm{A}}$ is relatively slowly varying, with a monotonic negative gradient. Therefore, over a sufficiently narrow range of impact parameter values, the linear expression $\tilde{\alpha}_{\mathrm{A}} \doteq \tilde{\alpha}_{\mathrm{A}o} + \tilde{\alpha}'_{\mathrm{A}}(\rho_* - \rho_o)$ will suffice. Because of the non-linearity in $\tilde{\alpha}_{\mathrm{W}}$ versus ρ_*, we have to keep the full expression. In Eq. (5.12-17), we substitute $\tilde{\alpha}_{\mathrm{W}} + \tilde{\alpha}_{\mathrm{A}}$ for $\tilde{\alpha}$, and we hold θ_{L} fixed while varying ρ_*. This gives us an expression for the breadth $\Delta\rho_*$, or the altitude difference at any given epoch between impact parameter values associated with any two competing rays arriving at the LEO. Upon expanding Eq. (5.12-17) through first degree in $\Delta\rho_*$, we have

$$\Delta\rho_*\big|_{\Delta\theta_{\mathrm{L}}=0} = \rho_{*2} - \rho_{*1} = \frac{2D_{\rho_o}\left(\tilde{\alpha}_{\mathrm{W}2} - \tilde{\alpha}_{\mathrm{W}1}\right)}{1 - 2D_{\rho_o}\tilde{\alpha}'_{\mathrm{A}}} \tag{5.12-18}$$

Here $\tilde{\alpha}_{\mathrm{W}2} - \tilde{\alpha}_{\mathrm{W}1}$ is the difference in one-way bending angles resulting from the local intrusive medium at a given orbit angle θ_{L} between multipath rays no. 1 and no. 2. At sea level, the defocusing factor from dry air is $\zeta_{\mathrm{A}}^{-1} = 1 - 2D_{\rho_o}\tilde{\alpha}'_{\mathrm{A}} \approx 10$. It follows that, given the change in bending angle $(\tilde{\alpha}_{\mathrm{W}2} - \tilde{\alpha}_{\mathrm{W}1})$, $\Delta\rho_*$ and, therefore, from Eq. (5.12-4) the multipath Doppler spread $\Delta(d\Psi_*/dt)/2\pi$ will be reduced by about a factor of ten relative to the spread that would be realized without the background refractivity.

5.12.6 Combined Water Vapor and Dry Air Refractivity Model

Figure 5-25 shows $dG^{\dagger}(\nu)/d\nu$ (or in a geometric optics context, $\tilde{\alpha}(\nu)$) where Cases A and B for the refractivity profile have been combined into the form

$$n - 1 = N_o \exp\left(\frac{\rho_o - \rho}{H}\right) + N_w \exp\left(-\frac{(\rho - \rho_w)^2}{2H_w^2}\right) \tag{5.12-19}$$

The curves for $dG^{\dagger}/d\nu$ are for different altitudes of the center of the Gaussian distribution near ρ_o; the center is given by ρ_w. Here $N_w = 0.2N_o$ and

---

[14] Strictly speaking, we should write $\tilde{\alpha}_{\mathrm{W}} = \tilde{\alpha}_{\mathrm{W+A}} - \tilde{\alpha}_{\mathrm{A}}$. In other words, expressing $\tilde{\alpha}_{\mathrm{W}}$ and $\tilde{\alpha}_{\mathrm{A}}$ as separate integrals in Eq. (5.6-5) is not strictly legitimate because of their non-linear dependence on refractivity. Two rays with the same impact parameter value, one in a medium of W+A and one in a medium of W only, follow different paths. However, for a thin atmosphere, Eq. (5.12-18) is fairly accurate.

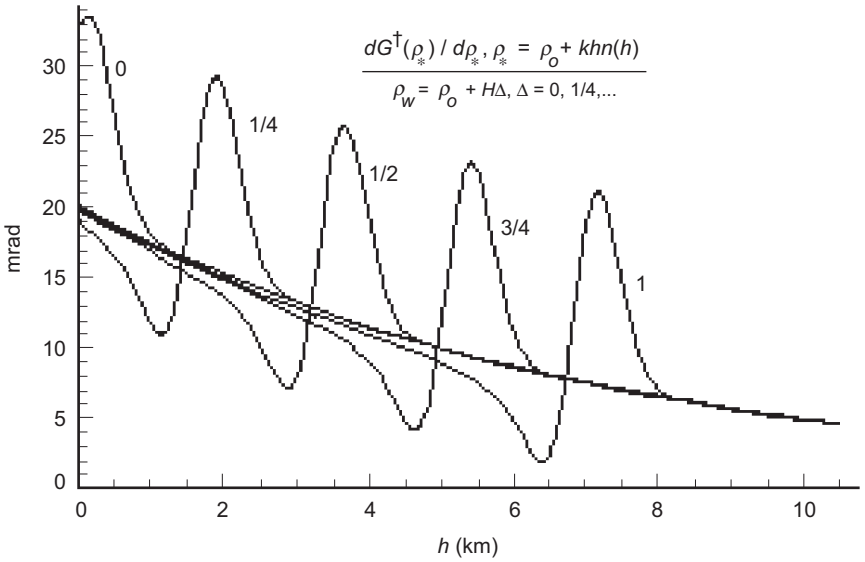

Fig. 5-25. Bending angle by a refractivity profile from dry air plus a narrow Gaussian marine layer that is located at different altitudes above sea level.

$H_w = 0.05H$. The values chosen for the dry air component are $H = 7k$, $N_o = 270 \times 10^{-6}$, and $\rho_o = 6378(1 + N_o)k$. These yield a refractivity profile at sea level that closely matches the dry air profile for a standard atmosphere. This dry air profile yields a two-sided bending angle at sea level of 20.4 mrad. A marine layer with substantial water vapor that is several hundred meters thick could have a refractivity profile similar to the Gaussian one used in this example. The value used for N_w in this example could correspond to a peak water vapor number density of about 1 percent of the local density of dry air. This composite moist refractivity profile yields a narrow (~1-km) transient with a peak $|n'\rho|$ value of 0.8. This bending-angle profile also results in two caustics.

Figure 5-26 is an impact parameter diagram showing the relationship between impact parameter ρ_* and the LEO orbit angle θ_{L}. This curve results from the combined exponential distribution for dry air plus the Gaussian distribution located in the lower troposphere, given in Eq. (5.12-19). Here $k^{-1}\rho_w = 1.75\,\mathrm{km}$ and ρ_o is the impact parameter of a ray with a tangency point at sea level. Also, $\theta_{\mathrm{L}o} = \sin^{-1}(\rho_o / \rho_{\mathrm{L}}) - 2\tilde{\alpha}(\rho_o, \rho_o)$, which is the orbit angle of the LEO at the refracted shadow boundary. For a setting occultation, θ_{L} decreases with time at a rate of roughly 1 mrad/s. This figure shows the extensive range ($\Delta\theta_{\mathrm{L}} \sim 20\,\mathrm{mrad}$, or about 1400 km of LEO orbital motion) or duration (~20 s) of multipath at the LEO compared to the half-width H_w

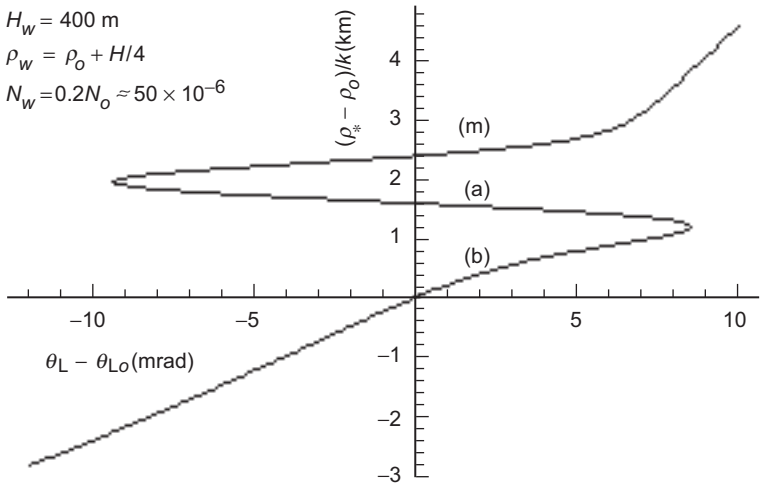

Fig. 5-26. Impact parameter diagram with defocusing. The LEO orbit radius is $\rho_L = 1.1\,\rho_o$. This figure shows ρ_* versus θ_L in the multipath zone from a narrow Gaussian refractivity profile from a marine layer, plus a more slowly varying dry air exponential distribution with $N_0 = 260 \times 10^{-6}$, $hk^{-1} = 7$ km.

(350 m) of the layer. Scintillation in amplitude and phase at the LEO first occurs at ray path tangency altitudes that are about 3 km above the altitude of the layer itself or about 10 s earlier.

Whereas Figs. 5-19 through 5-23 use a solitary Gaussian refractivity distribution and include no defocusing effects from dry air, Fig. 5-26 includes defocusing effects arising from the background refractivity due to dry air, which manifests itself in the much narrower impact parameter separations for the multiple rays. This defocused scenario produces maximum high-frequency fringes of around 15 Hz, much closer to a realistic case. Over most of the multipath zone, the bandwidth of the interference spectrum is less than 10 Hz for that example. Defocusing from the dry air refractivity profile compresses the bandwidth of the interference spectrum, but defocusing does not compress the duration of the multipath episode. In Appendix E, it is shown that the scale for $\left(\tilde{\alpha}_{W2} - \tilde{\alpha}_{W1}\right)$ for any two competing rays within the multipath zone depends nearly linearly on H_W, the $1-\sigma$ width of the Gaussian distribution, and it depends only weakly on the peak refractivity value N_W. The width of the multipath zone in orbit angle or in elapsed time—for example, the difference in θ_L between the upper and lower caustic points—depends on $H_W^{-3/2}$; it depends linearly on N_W. As we already have noted in Eq. (5.12-18), the fringe frequency scales, fast and slow, are compressed by defocusing from the slowly varying negative refractivity gradient from the background medium, dry air in this case. This compression ratio from dry air near sea level is about ten to one.

The excess Doppler difference between rays is $\dot{\theta}_{\mathrm{L}}\Delta\rho_{*}/2\pi \approx 5\Delta r_{*}$ Hz, where the difference in tangency points between rays, Δr_{*}, is in kilometers. Therefore, the fast fringe frequency between the main ray (m) and the nascent rays (a) and (b) shown in Fig. 5-26 is about 15 Hz at the midpoint in the multipath zone. This may be compared to the fast fringe frequency range of 80 to 180 Hz in Fig. 5-20, which does not include any defocusing from an ambient medium. A simple rule of thumb to convert an impact parameter diagram from ρ_{*} versus θ_{L} into a Doppler spread versus elapsed time is as follows: Multiply the ordinate in kilometers by the factor $5\dot{\theta}_{\mathrm{L}}$ to obtain hertz, and divide the abscissa in milliradians by $\dot{\theta}_{\mathrm{L}}$ to obtain seconds. Here $\dot{\theta}_{\mathrm{L}}$ is the rate of change of the subtending angle between the emitting GPS satellite and the LEO in the plane of propagation, typically 0.6 to 1.0 mrad/s, mostly depending on an obliquity factor between the LEO orbit and propagation planes.

Figure 5-27 shows the signal amplitude $\left|\boldsymbol{E}(r_{\mathrm{L}},\theta_{\mathrm{L}})\right|$ at the LEO versus orbit angle $\theta_{\mathrm{L}}-\theta_{\mathrm{L}o}$ during entry into the multipath zone for the same refractivity and orbit model used in Fig. 5-26. Knife-edge diffraction effects from the Earth's limb are ignored in these figures. This figure was obtained from a numerical integration of the spectral integral in Eq. (5.9-10). Above $\theta_{\mathrm{L}}-\theta_{\mathrm{L}o} \approx 8.5$, the main ray (m) provides the principal contribution to $\boldsymbol{E}(r_{\mathrm{L}},\theta_{\mathrm{L}})$, without interference from the (a) and (b) rays. For the $\theta_{\mathrm{L}}-\theta_{\mathrm{L}o}$ range shown here, the main ray impact parameter is located 2 to 4 km above sea level. Because of defocusing from the dry air refractivity gradient there, the amplitude of the (m) ray

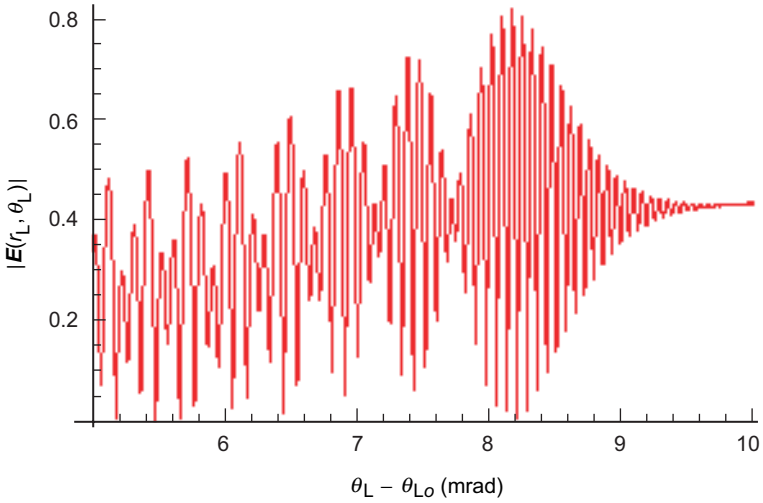

Fig. 5-27. Amplitude of the field at the LEO versus orbit angle near entry into the multipath zone for the defocused case using the same refractivity and orbit models used in Fig. 5-26. The equivalent elapsed time shown in this figure is about 5 s.

has been reduced to about 0.4 from its vacuum value of unity. The contact with the caustic surface where $d\theta_L / d\rho_* = 0$ occurs at $\theta_L - \theta_{Lo} \approx 8.54$, but the actual maximum flaring from wave theory occurs at $\theta_L - \theta_{Lo} \approx 8.17$. Geometric optics predicts infinite amplitude at the caustic contact point, but the actual value for the field contribution from the nascent rays (a) and (b) is 0.407. We have already noted that these differences between geometric optics and wave theory predictions can be reconciled using third-order stationary phase theory. The fast fringe frequency near maximum flaring is about 13 Hz, which is due to the 2.6-km separation between the impact parameter of the (a) and (b) rays at their nascence and the impact parameter of the (m) ray. This rate reduces to a minimum of about 10 Hz near $\theta_L - \theta_{Lo} = 4$. The slow modulation of the envelope, initially at roughly 1 Hz, is due to interference between the (a) and (b) rays. The node at $\theta_L - \theta_{Lo} = 7.8$ marks the first occasion where the (a) and (b) rays are π radians out of phase since their nascence. As the impact parameter separation distance between the (b) ray and the other two increases, the fast frequency gradually increases to a maximum of about 20 Hz near the lower caustic contact at $\theta_L - \theta_{Lo} = -9.4$.

An additional consequence follows from Figs. 5-26 and 5-27. Multipath from a relatively narrow refracting feature low down in the atmosphere, given a sufficient gradient, can be felt by the LEO for nominal ray path tangency points far above (this height difference scales roughly as $H_w^{-3/2}$). For a setting occultation, flaring of the observed signal at some otherwise benign point is a harbinger of things to come. We will return to this case later in connection with the effect of an embedded reflecting surface on the electric field at the LEO.

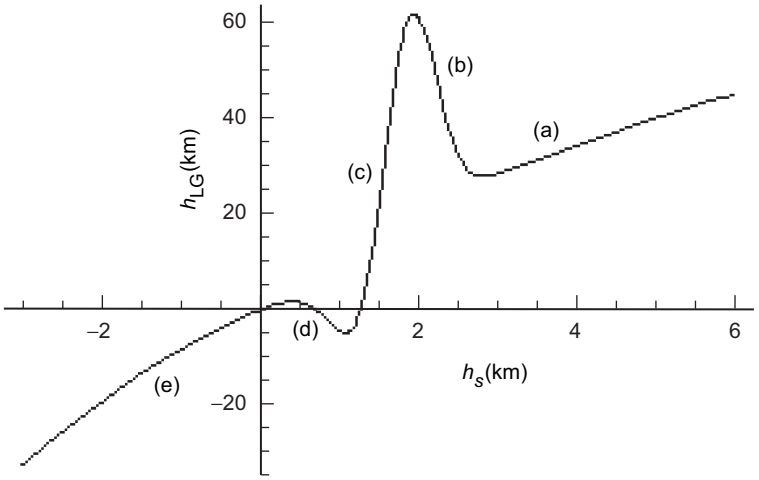

Fig. 5-28. Impact parameter diagram in the thin phase screen for the same refractivity profile used in Fig. 5-26, except that $N_w = -0.2N_o$.

Finally, Fig. 5-28 shows in the thin phase screen the impact parameter altitude h_S versus h_{LG} using the same Gaussian-plus-dry-air refractivity profile used in Fig. 5-26, except that $N_w = -0.2N_o$, an unlikely scenario in the lower troposphere. We have discussed the thin screen phase model in Section 5.10. This refractivity profile results in four caustics, five separate ray systems, and also a quasi-shadow zone around $h_S = 1.5$ km as a result of local defocusing there. Figure 5-29 shows the amplitude of the field at the LEO that results from the refractivity profile used to produce Fig. 5-28. Multiplying the abscissa by 3 gives the altitude h_{LG} in Fig. 5-28. The steep gradient in h_S versus h_{LG} in the thin screen, or equivalently in the θ_L versus ρ_* curve, results in the shadow zone. These rays in this example don't concurrently interfere with each other as quintuplets, but only as triplets. But, it would be easy enough to adjust the local refractivity gradient so that the lower caustic point at the nascence of the (d) and (e) rays in Fig. 5-28, for example, is raised to an altitude in h_{LG} that is higher than the caustic point at the nascence of the (b) and (c) rays. This would create a quintuplet multipath episode.

5.13 Spectral Coefficients in a Spherical Refracting Medium with an Embedded Discontinuity

The case where the refractivity is discontinuous at $r = r_o$ in a large, homogeneous, spherical symmetric medium was discussed in Chapter 3. Mie scattering [1,2] forms the basis of that discussion. It employed the stationary

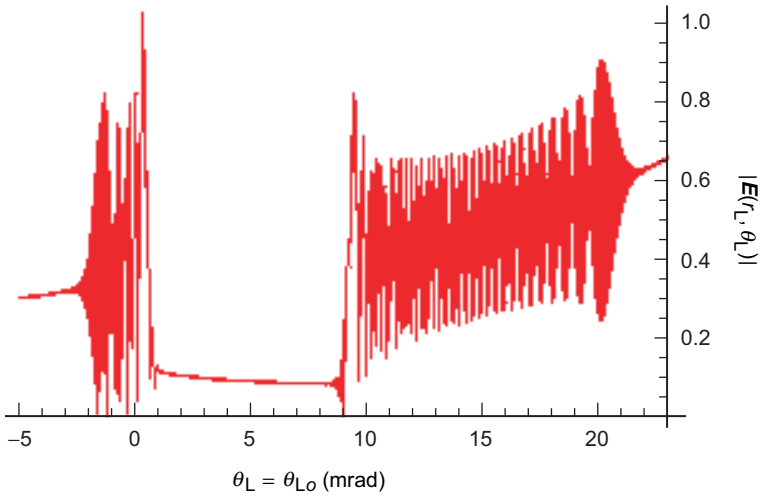

Fig. 5-29. Amplitude of the field at the LEO for the refractivity and orbit model used in Fig.5-28. Double reversal in the polarity of the bending-angle gradient causes four caustics and two multipath zones.

phase technique for interpretation and to aid the numerical evaluation of the scattering integrals. All of the scattering equations carry over to the case of a scattering sphere in a spherical symmetric stratified medium if we replace the spectral coefficients for an incident planar wave in those equations in Chapter 3, $i^{l-1}(2l+1)/(l(l+1))$, by the spectral coefficients $a_l^-(\rho)$ for an incoming wave. From Eq. (5.5-21), these spectral coefficients are given by

$$a_l^-(\rho) = i^{l-1}\frac{2l+1}{l(l+1)}\exp(-iG[\rho,\nu]) \tag{5.13-1}$$

These coefficients are applicable to an initially planar wave that has been subsequently refracted and retarded by the overlying spherical stratified medium before striking the scattering surface.

Let us now consider the spectral coefficients for the outgoing wave, $a_l^+(\rho)$. From Eq. (5.8-24), we have

$$a_l^+(\rho) = i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-i\left(2G\left[\rho^\dagger(\nu),\nu\right]-G[\rho,\nu]\right)\right) \tag{5.13-2}$$

Here the actual functional form of $a_l^+(\rho)$ will depend in part on the physical properties assumed for the refracting and perhaps scattering atmosphere, and also in part on the impact parameter(s) associated with the ray(s). Suppose that the scattering surface is located at $r = r_o$. For $\rho^\dagger > \rho_o$, these coefficients in Eq. (5.13-2) are still applicable. What happens when $\rho^\dagger \le \rho_o$? That depends on the medium lying below.

5.13.1 A Medium with a Discontinuity in its Refractive Gradient

We consider a spherical shell with $n(\rho)$ variable for $\rho \ge \rho_o$ and with $dn/d\rho \equiv 0$ for $\rho < \rho_o$. For this case Eqs. (5.5-18) and (5.5-22) show that both a_l^- and a_l^+ are constant with ρ when $\rho < \rho_o$. They also must be equal there to avoid the Hankel function singularity at the origin. Recall that the definition of the spherical Bessel function of the first kind is $\psi_l = \left(\xi_l^+ + \xi_l^-\right)/2$, which is well-behaved at the origin. It follows in this case that $a_l^+(\rho_o) = a_l^-(\rho_o)$. This also follows from the defining integral for $G[\rho,\nu]$ given in Eq. (5.5-20), or by its Airy function form given in Eq. (5.7-2), which we use here. Thus, $G\left[\rho^\dagger,\nu\right] = G[\rho_o,\nu]$, $\rho^\dagger \le \rho_o$. It follows that the spectral coefficients for the incoming wave are given by

$$a_l^-(\rho)=\begin{cases} i^{l-1}\dfrac{2l+1}{l(l+1)}\exp(-iG[\rho,\nu]), & \rho\ge\rho_o \\ i^{l-1}\dfrac{2l+1}{l(l+1)}\exp(-iG[\rho_o,\nu]), & \rho\le\rho_o \end{cases} \tag{5.13-3a}$$

At the LEO, $G[\rho_L,\nu]\equiv 0$, and the spectral coefficients for the outgoing wave are given by

$$a_l^+(\rho_{\mathrm{L}})=\begin{cases} i^{l-1}\dfrac{2l+1}{l(l+1)}\exp\left(-i2G\left[\rho^\dagger,\nu\right]\right), & \rho^\dagger\ge\rho_o \\ i^{l-1}\dfrac{2l+1}{l(l+1)}\exp(-i2G[\rho_o,\nu]), & \rho^\dagger\le\rho_o \end{cases} \tag{5.13-3b}$$

For a given value of $\nu=l+1/2$ such that $\rho^\dagger(\nu)\le\rho_o$, $-2G[\rho_o,\nu]$ is the total phase delay incurred by the lth spectral coefficient of an initially plane wave as a result of that wave passing completely through an intervening refracting shell with an inner boundary at $r=r_o$. When the stationary phase technique is applied to the spectral integrals, it yields a stationary phase value for ν that corresponds in geometric optics to a ray with an impact parameter $\rho_*\doteq\nu^*\le\rho_o$.

The total field at the LEO consists of both the incoming and outgoing radial components, as given by Eq. (5.13-3a) for $\rho=\rho_{\mathrm{L}}$ and Eq. (5.13-3b), respectively. But, we have already seen from our discussion in Section 5.9 of the stationary phase technique that, at the LEO, for an occultation, only the outgoing components contribute to the scattering integrals given in Eq. (5.9-5). Because the LEO is so far away from a turning point, only the outgoing components yield stationary phase neighborhoods in spectral number. So, we can ignore the incoming components at the LEO for an occultation geometry.

We also note the relative ease with which this spectral technique using osculating parameters can deal with a discontinuity in the gradient of the refractivity at $r=r_o$. Such a discontinuity might correspond to a discontinuity in scale height of the atmosphere, for example, at the boundary of a marine layer, or in the lapse rate of the troposphere, for example, at the tropopause. From Eq. (5.5-20), it follows that $G[\rho,\nu]$ is continuous with ρ, even though n' is discontinuous at $r=r_o$. Also, $dG\left[\rho^\dagger(\nu),\nu\right]/d\nu$ is continuous. It follows that, using the stationary phase technique to evaluate the field from the integrals in Eq. (5.9-5), the same stationary phase point in spectral number is obtained as that obtained from the ambient medium without the discontinuity in n'. We know that, for a given the position $(r_{\mathrm{L}},\theta_{\mathrm{L}})$ of the LEO, the stationary phase point ν^* is near the impact parameter ρ_*. But the second derivative

$d^2G[\rho^\dagger,\nu]/d\nu^2$ is discontinuous at $\nu^* = \rho_o$. This means that, when using the stationary phase technique, it is necessary to break the integral for the field over spectral number into two parts: one part for $\nu > \rho_o - \hat{y}^\dagger K_{\rho_o}$, and a second part for $\nu < \rho_o - \hat{y}^\dagger K_{\rho_o}$. When the impact parameter is close to ρ_o, these two parts when summed interfere in phase, resulting in fringes in the amplitude and phase of the field.

We choose the simple model in Eq. (5.13-4) with which to compare wave theory scattering results with Fresnel diffraction from a thin-screen model. Here we assume that $n' \equiv 0$ for $r < r_o$. For $r \geq r_o$, we assume that $n = n(\rho)$, which is arbitrary other than satisfying the thin-atmosphere conditions. As a specific example, we assume that n follows an exponential law for $r > r_o$ and is a constant for $r < r_o$; thus,

$$\left.\begin{aligned} n^+ &= 1 + N_o \exp\left(-\frac{\rho - \rho_o}{H}\right), \ \rho \geq \rho_o \\ n^- &= 1 + N_o, \ \rho \leq \rho_o \end{aligned}\right\} \tag{5.13-4}$$

Figure 5-30 shows the $2dG^\pm/d\nu$ curves from Eq. (5.7-7) and certain δ curves from Eq. (5.12-1) for this particular refraction profile. From Eq. (5.7-7), it follows that $dG^\pm/d\nu$ is defined by

$$\left.\begin{aligned} \frac{dG^+}{d\nu} &= -\pi K_{\rho^\dagger} \int_{\rho^\dagger}^{\infty} \frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho, \ \rho^\dagger \geq \rho_o \\ \frac{dG^-}{d\nu} &= -\pi K_{\rho_o} \int_{\rho_o}^{\infty} \frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho, \ \rho^\dagger \leq \rho_o \\ \frac{dG^+}{d\nu} &= \frac{dG[\rho^\dagger(\nu),\nu]}{d\nu}, \quad \frac{dG^-}{d\nu} = \frac{\partial G[\rho_o,\nu]}{\partial \nu}, \quad \hat{y} \doteq K_\rho^{-1}(\nu - \rho) \end{aligned}\right\} \tag{5.13-5}$$

When $\nu^* - \rho_o < - \sim 2K_{\rho_o}$, we may replace the Airy functions in these integral expressions for $2dG^-/d\nu$ with their respective asymptotic forms. From Eq. (5.7-8), we obtain

$$\frac{dG^-}{d\nu} \doteq -\nu \int_{\rho_o}^{\infty} \frac{d\log n}{d\rho} \frac{d\rho}{\sqrt{\rho^2 - \nu^2}} = \tilde{\alpha}(\rho_o, \nu), \ \rho^\dagger \leq \rho_o - K_{\rho^\dagger}\hat{y}^\dagger \tag{5.13-6}$$

Using the refractivity profile given in Eq. (5.13-4), an explicit form $\tilde{\alpha}(\rho_o,\nu)$ is given in Eq. (5.8-3). Here $\tilde{\alpha}(\rho_o,\nu)$ is the cumulative bending on a ray at

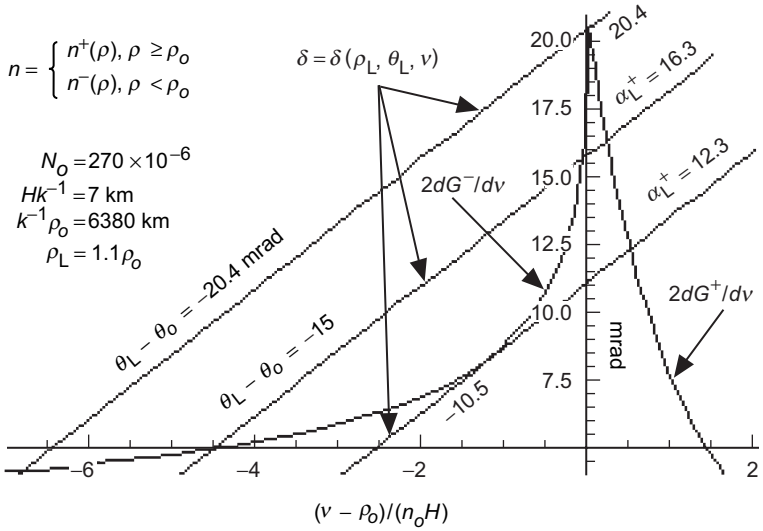

Fig. 5-30. Stationary phase curves for a spherical shell with a discontinuity in $dn/d\rho$ at r_o per Eq. (5.13-4). Stationary phase points occur at the intersections of the δ and $2dG^{\pm}/d\nu$ curves.

$\rho = \rho_o$ with an impact parameter value of $\nu \leq \rho_o$. It corresponds to $\alpha^-(\nu)/2$ given in Section 2.3, Eq. (2.3-10). As was discussed in Section 5.7, the small difference between $dG^-/d\nu$ and $\tilde{\alpha}(\rho_o, \nu)$ shows up only in the immediate vicinity of a turning point.

We may write the integral form as $dG^{\pm}/d\nu$

$$2\frac{dG^+}{d\nu} \doteq -2\nu\int_\nu^\infty \frac{d\log n}{d\rho}\frac{d\rho}{\sqrt{\rho^2-\nu^2}} = 2\tilde{\alpha}(\nu,\nu), \quad \rho^\dagger > \rho_o \qquad (5.13\text{-}7)$$

Here $2\tilde{\alpha}(\nu,\nu) = \alpha_L^+(\nu)$ corresponds to $\alpha^+(\nu)$ in Eq. (2.3-11), the bending angle for a transecting ray through the upper medium with a value ν for its impact parameter such that $\nu \geq \rho_o$. This ray completely transects the atmosphere unperturbed by the discontinuity lying below.

The δ curves in Fig. 5-30 are obtained from Eq. (5.12-1), and they depend uniquely on the LEO angular position θ_L as a parameter when the orbital radius r_L is fixed. From Bouguer's law in Eq. (5.6-5), this dependence may be recast in terms of bending angle, except that in this example bending angle is not unique over a certain range of angular positions. For that reason, we have chosen $\alpha_L^+(\rho_*)$, the bending angle in the upper regime, as the parameter. It is unique. By expanding Eq. (5.12-1) in powers of $(\nu - \rho_*)$, and upon noting that

$\partial\theta_\nu / \partial\nu = D_\nu^{-1}$, it follows that over a sufficiently narrow range of spectral numbers $\delta \doteq \alpha_{\mathrm{L}}^{+}(\rho_*^{+}) + (\nu - \rho_*^{+}) / D_{\rho_*}$. The point of first contact with the caustic surface is located at the point of tangency of the $2dG^{-} / d\nu$ curve with the δ curve for $\alpha_{\mathrm{L}}^{+} = 12.3\,\mathrm{mrad}$, or for $\theta_{\mathrm{L}} - \theta_o = -10.5$ mrad for this example. For LEO angular positions from this value down to $\theta_{\mathrm{L}} = \theta_o - 20.4$, there are three stationary phase points in spectral number and, therefore, three bending angles. For orbital angles outside of $-20.4 \le \theta_{\mathrm{L}} - \theta_o \le -10.5\,\mathrm{mrad}$, i.e., for δ curves above the uppermost curve or below the lowest one in Fig. 5-30, there is just one intersection point, or one stationary phase point and only one value for the bending angle. The LEO-observed phase and amplitude in this region between the caustic contact point and the cusp in bending angles are marked by strong multipath interference effects and by diffraction in the transition regions. In a wave theory context, fringes in the observed amplitude and phase result from interference among the spectral components for spectral numbers in the immediate neighborhood of ρ_o when the impact parameter is nearby. In geometric optics, these fringes come from interference among multiple rays.

At the point of tangency of the δ curve with the $2dG^{-} / d\nu$ curve, the condition $2d^2G^{-} / d\nu^2 = \partial\delta / \partial\nu$ must hold. The geometric optics equivalent is $d\alpha_{\mathrm{L}}^{-} / d\nu = \partial\delta / \partial\nu = D_\nu^{-1}$, which is equivalent to the condition that the defocusing factor ζ must be infinite. This wave theory tangency condition in Fig. 5-30 is equivalent to the geometric optics condition on the impact parameter (for a circular LEO orbit) $d\rho_* / d\theta_{\mathrm{L}} \to \infty$, which is shown in Fig. 5-31. At this tangency point, flaring in the observed signal will be evident,

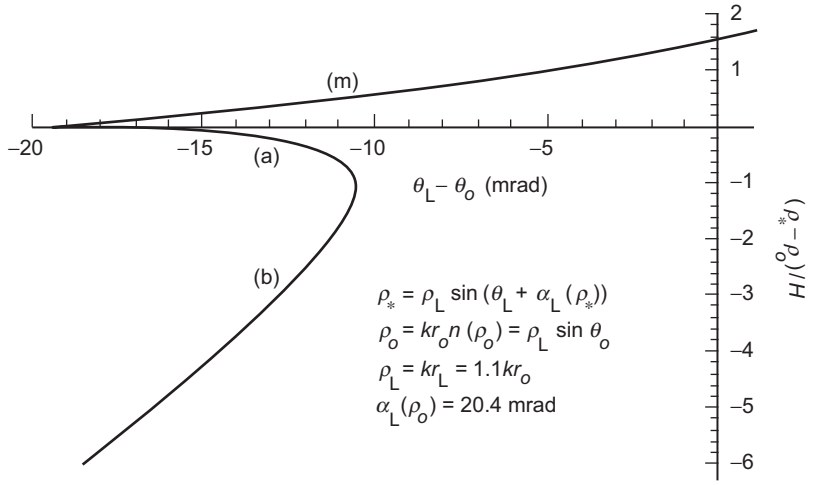

Fig. 5-31. Impact parameter diagram for the refracting shell described in Fig. 5-30. Boundary encountered at $\theta_{\mathrm{L}} - \theta_o = -20.4$.

although the actual maximum in flaring occurs at a point slightly offset from the tangency point. This offset has been discussed in Section 5.12 using third-order stationary phase theory.

Figure 5-32 shows the amplitude of the field at the LEO in the vicinity of the first contact with the caustic ($\theta_L - \theta_o = -10.5$ mrad) using the same conditions used in Fig. 5-30. The flaring is quite prominent in this example. Figures 5-32 through 5-34 were obtained from a numerical integration of the spectral integrals in Eq. (5.9-24) aided by the stationary phase technique to isolate the contributing stationary phase neighborhoods in spectral number. At $\theta_L - \theta_o = -10.5$ mrad, a pair of powerful rays, (a) and (b) in Fig. 5-31, are created. In this example, their individual amplitudes are roughly six times the amplitude of the original (m) ray at this orbital position. Therefore, their contributions to the field at the LEO largely overwhelm the contribution from the (m) ray by an order of magnitude. Third-order stationary phase theory is needed at the caustic point in Fig. 5-32 to accurately predict the amplitude and location of the peak. Second-order geometric optics predicts an infinite amplitude at the caustic point, whereas Fig. 5-32 shows that wave theory yields a peak amplitude of about 5.3. In Section 5.12, it is shown that the peak amplitude associated with a caustic is proportional to $(d^2\theta_L / d\rho_*^2)^{-1/3}$, evaluated near the caustic contact point where $d\theta_L / d\rho_* = 0$. This curvature term is readily obtained from Bouguer's law in Eq. (5.12-5) and the bending profile, which is given in Eq. (5.6-5) in terms of the refractivity profile. Equation (5.12-13) provides a ratio for the amplitudes of the nascent rays compared with the amplitude of the main ray, expressed in terms of the first

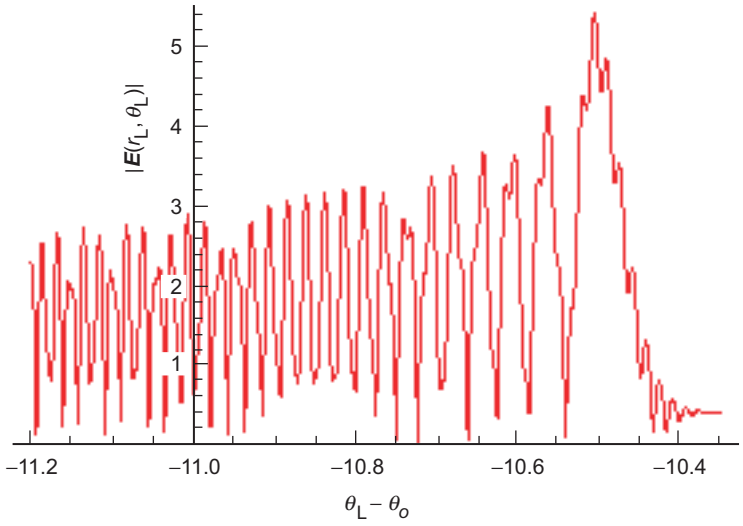

Fig. 5-32. Amplitude of the field at the LEO near the caustic contact using the same refractivity and orbit models used in Fig. 5-30.

and second derivatives of the curve, θ_{L} versus ρ_* as shown in Fig. 5-31. For the specific exponential shell model in Eq. (5.13-4), which we are using in this section, Eq. (5.12-13) yields a flaring from the nascent rays alone that is 12.3 times the amplitude of the (m) ray. This yields a maximum amplitude of 5.3 from all three rays combined, very close to the actual peak amplitude shown in Fig. 5-32 near the caustic contact. The amplitude of the main ray (m) is normalized to unity outside the medium, but defocusing at $\theta_{\mathrm{L}} - \theta_o = -10.5$ mrad has reduced it to 0.4.

These new rays (a) and (b) begin to mutually interfere with each other, as is evidenced by the onset of fringes for $\theta_{\mathrm{L}} - \theta_o \le -10.5$ mrad. As discussed in Section 5.12, Eq. (5.12-4), the frequency of these fringes can be obtained by multiplying the difference in impact parameter values in kilometers by $5\dot{\theta}_{\mathrm{L}}$. At $\theta_{\mathrm{L}} = \theta_o - 10.6$ mrad for rays (a) and (b), that frequency already has grown to about 40 Hz (with $\dot{\theta}_{\mathrm{L}} = -1$ mrad/s). Figure 5-31 shows that the impact parameter values for the (a) and (b) rays promptly separate with decreasing θ_{L}, and their defocusing factors begin to steeply decrease from initially infinite values. This rapid separation in impact parameter values leads to a rapidly increasing fringe frequency with decreasing θ_{L}. This interference initially results principally from the mutual interference between the (a) and (b) rays with a smaller modulation from the weaker (m) ray. Later, as shown in Fig. 5-33, as the (a) ray becomes defocused, the high-frequency interference is between the (b) and (m) rays with a weaker modulation of much lower frequency from interference between the (m) and (a) rays.

Figure 5-33 shows the amplitude of the field near the end of the interference zone. Here the (m) and (b) rays are well separated and the anomalous ray (a) has become a minor contributor because of its defocusing (see Fig. 5-31). The impact parameter differences read from Fig. 5-31 at the end of the interference zone predict very high fringe frequencies here. At $\theta_{\mathrm{L}} - \theta_o = -19$, the frequency of the high-frequency fringes between the (m) and (b) rays is about 250 cycles/mrad. The low-frequency modulation in Fig. 5-33 comes from interference between the (m) and (a) rays, which have narrowly different impact parameter values in this figure. Here $d(\Psi_{*\mathrm{a}} - \Psi_{*\mathrm{m}})/d\theta_{\mathrm{L}}$ is about 5 cycles/mrad at $\theta_{\mathrm{L}} - \theta_o = -17$, about 2 cycles/mrad at $\theta_{\mathrm{L}} - \theta_o = -19$, and zero at $\theta_{\mathrm{L}} - \theta_o = -20.4$, the end of the multipath zone.

As also was discussed in Section 5.12, the amplitude of the slow modulation in Fig. 5-33 can be obtained from the defocusing factor $\zeta^{-1} = 1 - D_{\rho_*} d\alpha_{\mathrm{L}}/d\rho_*$ for each ray. The modulation peaks and valleys are accurately predicted from the four combinations $||\zeta_{\mathrm{a}}|^{1/2} \pm |\zeta_{\mathrm{b}}|^{1/2} \pm |\zeta_{\mathrm{m}}|^{1/2}|$. The slow modulation ($\Delta\theta_{\mathrm{L}} \sim 0.2$ mrad per cycle at $\theta_{\mathrm{L}} - \theta_o = -17$) results from interference between the narrowly separated (m) and (a) rays. The (a) ray

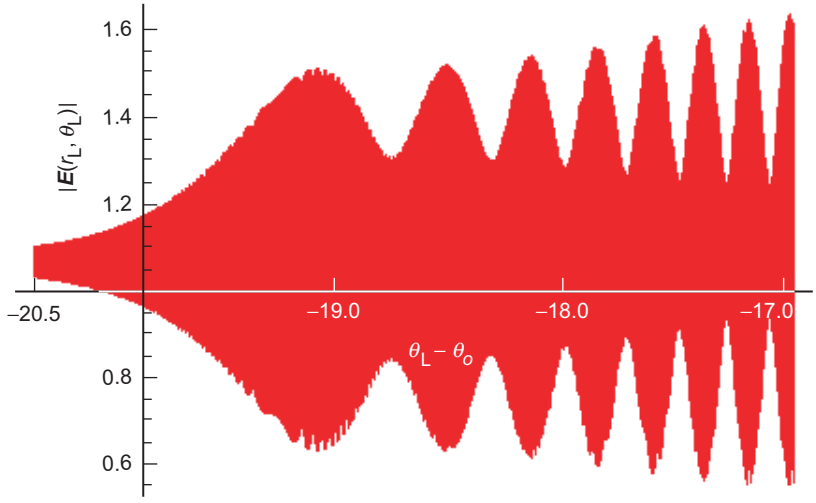

Fig. 5-33. Continuation of Fig. 5-32 to the lower boundary of the interference zone at $\theta_L - \theta_o = -20.4$.

becomes very defocused with decreasing θ_L. At $\theta_L - \theta_o = -17$, the amplitude contributions from the three rays based on their defocusing factors have the ratios b:m:a ~1:1/3:1/5. At $\theta_L - \theta_o = -19$, these ratios are ~1:1/3:1/11. For $\theta_L - \theta_o < -20.4$, they are 1:0:0.

The mean amplitude of the field for the range of $\theta_L - \theta_o$ values shown in Fig. 5-33 is about 1.1. This is about a factor of three greater than the amplitude of the (m) ray here because of the lens-like property of the refracting shell given by Eq. (5.13-4).

Figure 5-34 shows a section of the de-trended phase of the complete field $\Delta\Phi^*_{mba}$ at the LEO, expressed in cycles for the range $-11.35 \le \theta_L - \theta_o \le -10.85$. It may be compared to the amplitude fringes shown in Fig. 5-32, which cover most of this region. This figure shows the bursts of rapid acceleration in phase that correspond to local neighborhoods where substantial destructive interference occurs, principally between the newly created (a) and (b) rays. These points correspond to the troughs in amplitude in Fig. 5-32. Phase changes of 1/2 cycle occur over a change in θ_L of less than $2\ \mu$rad, or less than 2 ms of elapsed time. For a more realistic refractivity model with multipath impact parameter separations that are 1/4 the size of those in Fig. 5-31, the elapsed time for a 1/2-cycle change from these brief bursts would be closer to 5 ms, which is still a significant operational problem for a closed-loop receiver.

Finally, Fig. 5-35 provides a much more benign scenario. It shows the amplitude of the field at the LEO from the spectral integral in Eq. (5.8-1b) for the case where the gradient of the refractivity has a discontinuity of the opposite

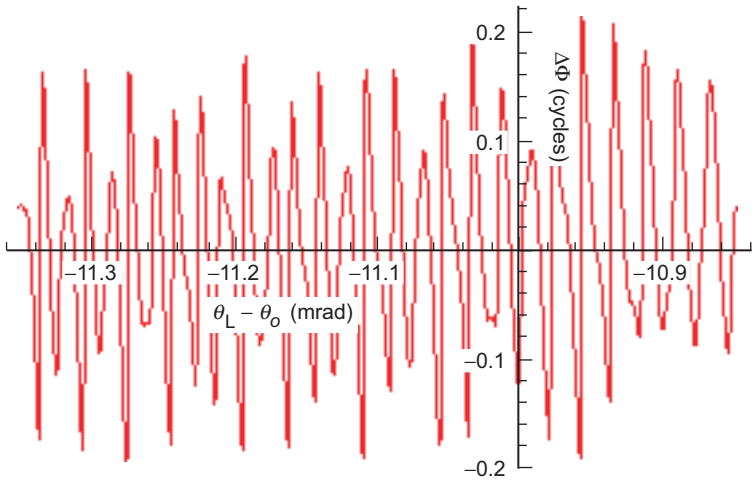

Fig. 5-34. De-trended phase at the LEO near onset of multipath. Figures 5-30 through 5-34 use the same refractivity and orbit models.

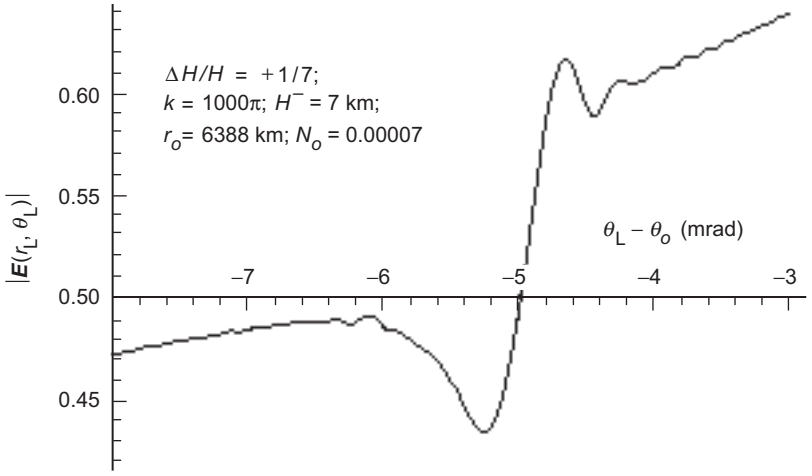

Fig. 5-35. Amplitude at the LEO from a discontinuity in scale height.

polarity to that used in Fig. 5-32 (and also significantly smaller). The index of refraction for this case is given by

$$\left.\begin{aligned} n^{+} &= 1 + N_o \exp[-(\rho - \rho_o)/H^{+}], \quad \rho > \rho_o \\ n^{-} &= 1 + N_o \exp[-(\rho - \rho_o)/H^{-}], \quad \rho \le \rho_o \end{aligned}\right\} \tag{5.13-8}$$

Here $\Delta H = H^+ - H^- > 0$. This profile results in a mild shadow zone in the neighborhood $-6 \le \theta_L - \theta_o \le -5$ mrad with minor diffraction effects. The bending-angle profile for this example is given in Section 2.3 by the Case B scenario, in particular Eq. (2.3-21). See also Section 2.4. In this case, ΔH is positive. When the tangency point of the ray descends through the surface of the discontinuity, the ray experiences an increased bending rate because of the increased gradient in refractivity in the lower layer. The amplitude profile closely follows the square root of the defocusing function except in the immediate vicinity of the shadow zone. One can compare Fig. 5-35 with Fig. 2-10 from the thin-screen/scalar diffraction approach. Figure 2-10 applies to a discontinuity in the lapse rate at the tropopause of $\Delta(dT/dr) = -7$ K/km, whereas Fig. 5-35 applies to a discontinuity in scale height of $\Delta H = +1$ km. There is a rough correspondence between these two quantities in terms of their perturbations on bending angle, which is given by $\Delta H / H \leftrightarrow -H\Delta(dT/dr)$. When this becomes an equality, the perturbations on the bending angle are about the same, and the resulting diffraction pattern is about the same. Here $\Delta H = +1$ km is equivalent to a lapse rate discontinuity of 4 to 5 K/km.

5.13.2 A Transparent Sphere Embedded in a Refracting Medium

For a transparent sphere with a discontinuity in $n(\rho)$ at $\rho = \rho_o$ and $n' = 0$ inside, the spectral coefficients evaluated at the scattering sphere for the total scattering, including the scattering effects of multiple internal reflections, are obtained by modifying the discussion in Section 3.5, Eqs. (3.5-11) and (3.5-15b), to account for the phase delay $-G\left[\rho_o^+, \nu\right]$ in the incident wave induced by the refractivity gradient in the overlying medium. The term $b_l^{(0)}$ is called the zeroth-degree reflection coefficient because it applies only to an external reflection from the sphere; $b_l^{(j)}$ is the jth-degree reflection coefficient; it applies to a wave that has undergone $j-1$ internal reflections within the sphere. Thus, $j = 1$ corresponds to a wave entering and exiting the sphere without undergoing any internal reflections, $j = 2$ corresponds to a wave with one internal reflection, and so on. Summing these coefficients over degree yields the total or aggregate scattering coefficient S_l, which is given by

$$S_l = \sum_{j=0}^{\infty} b_l^{(j)} = -a_l^-\left(\rho_o^+\right)\left(\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+}\right) \tag{5.13-9}$$

where $\tilde{\mathcal{W}}_l$ is the complex conjugate of $\mathcal{W}_l$. Here

$$a_l^-\left(\rho_o^+\right) = i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG\left[\rho_o^+, \nu\right]\right), \quad \rho_o^+ = n_o^+ u_o \tag{5.13-10}$$

and

$$\left.\begin{aligned} \mathcal{W}_l &= \xi_l^+\left(n_o^+ u_o\right)\xi_l^{-\prime}\left(n_o^- u_o\right) - \left(n_o^- - n_o^+\right)\xi_l^{+\prime}\left(n_o^+ u_o\right)\xi_l^-\left(n_o^- u_o\right) \\ \mathcal{W}_l^{\pm} &= \xi_l^{\pm}\left(n_o^+ u_o\right)\xi_l^{\pm\prime}\left(n_o^- u_o\right) - \left(n_o^- - n_o^+\right)\xi_l^{\pm\prime}\left(n_o^+ u_o\right)\xi_l^{\pm}\left(n_o^- u_o\right) \end{aligned}\right\} \tag{5.13-11}$$

Here n_o^+ is the index of refraction on the outer side of the boundary at $r = r_o$; n_o^- is the value on the inner side. For a transparent sphere, $n(\rho)$ is constant and real for $\rho < \rho_o$. Upon propagating the aggregate scattering coefficients upward through the refracting medium to the LEO, one obtains

$$S_l(\rho_{\mathrm{L}}) = -i^{l-1}\frac{2l+1}{l(l+1)}\left(\frac{\tilde{\mathcal{W}}_l + \mathcal{W}_l^-}{\mathcal{W}_l + \mathcal{W}_l^+}\right)\exp\left(-i2G\left[\rho_o^+, \nu\right]\right) \tag{5.13-12}$$

Once again, we see that the scattering coefficients are phase delayed relative to their vacuum counterparts by an amount $-2G\left[\rho_o^+, \nu\right]$. The total outgoing field is obtained by adding these scattering coefficients, given by Eqs. (5.13-10) and (5.13-12), to the spectral coefficients in Eq. (5.13-3) for the direct field. This combination then is used in the spectral integrals in Eq. (5.8-1b).

Rainbow caustics through a refracting atmosphere with a sharp transparent boundary underneath can be obtained from this approach.

5.14 The Scattered Field from a Perfectly Reflecting Sphere Embedded in a Refracting Medium

Continuing the subject of embedded discontinuities begun in Section 5.13, we now consider a perfectly reflecting sphere embedded in a refracting medium. This example has some application to grazing, near-specular reflections from the ocean [9,10], especially at very shallow angles of incidence where the reflected signal is essentially linearly polarized. The transmitted GPS signals are right-hand circularly polarized (RHCP). Therefore, for grazing angles, roughly half the original power is potentially available from reflected signals, especially if the sea surface is smooth.

The spectral coefficients $b_l^{(0)}$ for a very large, perfectly reflecting sphere in a homogeneous medium are given in Chapter 3, Eq. (3.17-1). Those coefficients were obtained from the more general case of a transparent sphere with a finite discontinuity in the refractivity at its boundary $\rho = \rho_o$. By letting $n \to \infty$ at $\rho = \rho_o^-$, one obtains in the limit the reflection coefficients for a

perfectly reflecting sphere. Therefore, in the case of a reflecting sphere in a non-homogenous medium, $G[\rho,\nu]$ is defined only for $\rho \geq \rho_o$. There is essentially infinite phase windup in the spectral coefficients traveling inward across that boundary. If we modify the scattering equation in Eq. (3.17-1) to account for the sphere being embedded in a stratified refracting medium, we have the following for the reflection coefficients at the boundary:

$$b_l^{(0)}(\rho_o) = -\frac{1}{2}\left(a_l^+(\rho_o) + a_l^-(\rho_o)\frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)} \right), \quad \rho_o = kn_o r_o \tag{5.14-1}$$

At the boundary, the spectral coefficients are

$$a_l^+(\rho_o) = a_l^-(\rho_o) = i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-iG[\rho_o,\nu]\right) \tag{5.14-2}$$

These apply to an initially planar incoming wave. The modification to this equation to account for an initially spherical incoming wave is found in Section 5.3, Eqs. (5.5-3a) and (5.5-3b). We recall from Chapter 3 that the superscript "(0)" on $b_l^{(0)}$ denotes the zeroth-degree reflection coefficient—that is, the coefficient for the ray with only a surface reflection, i.e., no internal reflections within the sphere. This is the only non-zero term for a perfectly reflecting sphere. To obtain $b_l^{(0)}(\rho)$ for $\rho > \rho_o$, we use arguments that are similar to those used in Section 5.8 to propagate the outgoing spectral coefficients through the medium. We define $b_l^{(0)}(\rho_{\mathrm{L}})$ as the reflection coefficient at the LEO. Assuming that the LEO is out of the medium so that $G[\rho_{\mathrm{L}},\nu] \equiv 0$, we obtain

$$b_l^{(0)}(\rho_{\mathrm{L}}) = b_l^{(0)}(\rho_o)\exp\left(-iG[\rho_o,\nu]\right) \tag{5.14-3}$$

It follows from Eqs. (5.8-24) and (5.13-1) that the reflection coefficients at the LEO for a perfect spherical reflector in a stratified medium are given by

$$b_l^{(0)}(\rho_{\mathrm{L}}) = -\frac{1}{2}i^{l-1}\frac{2l+1}{l(l+1)}\left(1+\frac{\xi_l'^-(\rho_o)}{\xi_l'^+(\rho_o)}\right)\exp\left(-i2G[\rho_o,\nu]\right) \tag{5.14-4}$$

In other words, for the stratified medium, the reflection coefficient of spectral number l at the LEO is delayed in phase by an amount $-2G[\rho_o,\nu]$ as compared to the pure vacuum case discussed in Chapter 3.

The total field at the LEO is obtained from the spectral integrals in Eq. (5.8-1b), which are composed of the reflection coefficients in Eq. (5.14-4)

plus the spectral coefficients for the incident field. For a collimated approaching wave, the spectral coefficients for the incident field are given by

$$\left.\begin{aligned} a_l^-(\rho_{\mathrm{L}}) &= i^{l-1}\frac{2l+1}{l(l+1)}\exp(-iG[\rho_{\mathrm{L}},\nu]), \\ a_l^+(\rho_{\mathrm{L}}) &= i^{l-1}\frac{2l+1}{l(l+1)}\exp\left(-i2G\left[\rho^{\dagger},\nu\right]\right),\ \ \rho^{\dagger}\geq\rho_o \\ a_l^+(\rho_{\mathrm{L}}) &= i^{l-1}\frac{2l+1}{l(l+1)}\exp(-i2G[\rho_o,\nu]),\ \ \rho^{\dagger}\leq\rho_o \end{aligned}\right\} \qquad (5.14\text{-}5)$$

Referring to Fig. 5-36, we see that for LEO orbital positions above the shadow boundary there are two paths, a direct path with an impact parameter $\rho_d \geq \rho_o = kr_o n(r_o)$ and a reflected path with an impact parameter $\rho_{\mathrm{r}} \leq \rho_o$. For regimes where the stationary phase condition in spectral number, $\nu^* \doteq \rho_d \leq \rho_o$, would have held in the absence of the reflecting sphere, there are no stationary phase points for ν. Here one obtains a diffracted knife-edge decay in amplitude as the GPS satellite becomes occulted by the reflecting sphere.

From Section 5.9, we know that at the LEO the incoming coefficients $a_l^-(\rho_{\mathrm{L}})$ for the direct ray (d) may be ignored because they provide no stationary phase points in spectral number. The sum of the reflection and outgoing direct coefficients gives the total field at the LEO. They may be recombined into the form

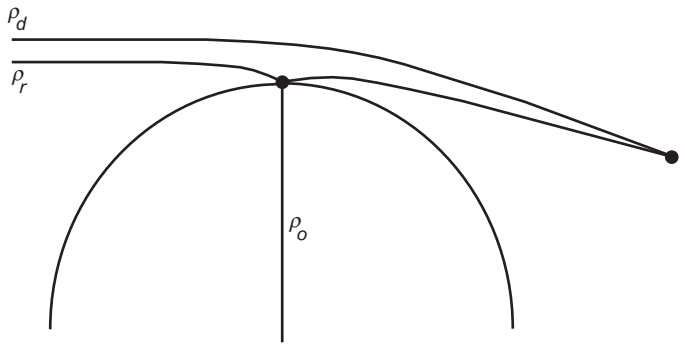

Fig. 5-36. Direct and reflected paths for a sphere of radius r_o embedded in a medium with a monotonic refractive gradient.

$$b_l^{(0)}(\rho_{\mathrm{L}})+\frac{1}{2}a_l^{+}(\rho_{\mathrm{L}})= \\ \left\{\begin{array}{l} i^{l+1}\dfrac{1}{2}\dfrac{2l+1}{l(l+1)}\left(\left(1+\dfrac{\xi_l^{-'}(\rho_o)}{\xi_l^{+'}(\rho_o)}\right)e^{-i2G[\rho_o,\nu]}-e^{-i2G[\rho^{\dagger},\nu]}\right),\ \rho^{\dagger}>\rho_o, \\ \dfrac{1}{2}i^{l+1}\dfrac{2l+1}{l(l+1)}\dfrac{\xi_l^{-'}(\rho_o)}{\xi_l^{+'}(\rho_o)}e^{-i2G[\rho_o,\nu]},\ \rho^{\dagger}<\rho_o \end{array}\right\} \quad (5.14\text{-}6)$$

Consider the regime $\rho^{\dagger}(\nu)>\rho_o$, the middle line in Eq. (5.14-6). For $\nu>\rho_o$, $\xi_l^{-'}(\rho_o)/\xi_l^{+'}(\rho_o)\to -1$ rapidly because of the runoff of Bi[$\hat{y}$] for $\hat{y}\doteq K_{\rho_o}^{-1}(\nu-\rho_o)>0$. Consequently, $b_l^{(0)}(\rho_o)\to 0$ for increasing $\nu>\rho_o$. Already at $\hat{y}=+3$, $|\xi_l^{-}(\rho_o)/\xi_l^{+}(\rho_o)+1|=0.001$. This leaves only the term $i^{l-1}(2l+1)/l(l+1)\exp(-i2G[\rho^{\dagger},\nu])/2$ in this regime where $\rho^{\dagger}(\nu)>\rho_o$. This is just the spectral coefficient for a direct ray, but applied to a spectral integral that has a lower cutoff at $\rho^{\dagger}(\nu)>\rho_o$. For $\rho_d>>\rho_o$, there will be a stationary phase point ν^* well above the cutoff. Assuming that the overlying medium has a monotonic refractivity gradient, i.e., no multipath, there will be only the one stationary phase point near $\nu=\rho_d$. It follows that when $\rho_d>>\rho_o$ this integral yields the field from the direct path, unencumbered by the reflection barrier; the lower cutoff at $\rho^{\dagger}(\nu)>\rho_o$ has no effect on the value of the integral. This is the field for a wave that has traveled through a refracting medium without influence from the reflecting surface below.

On the other hand, for a LEO orbital position that is lower than that shown in Fig. 5-36, for which the direct ray has an impact parameter value $\rho_d\approx\rho_o$, the stationary phase point ν^* will be near ρ_o. Here the lower cutoff in the spectral integral near $\nu=\rho_o$ does affect the calculation of the field at the LEO. It yields the knife-edge diffraction pattern that the wave from the direct ray will exhibit as the GPS satellite appears to approach the limb of the reflecting sphere and becomes eclipsed. If the LEO is located so that $\rho_d\leq\sim\rho_o$, then no stationary phase contributions occur for any value of ν; the LEO is in the refracted shadow. We will show these properties later using stationary phase theory.

The spectral coefficients on the lower line in Eq. (5.14-6) where $\rho^{\dagger}(\nu)>\rho_o$ provide the principal contribution to the reflected field at the LEO. Referring to Fig. 5-36, if the orbital position of the LEO is such that $\rho_d>\rho_o$, then there is also a reflected ray with an impact parameter value $\rho_r<\rho_o$. There will be a stationary phase point in spectral number, $\nu=\nu^*$, near $\nu=\rho_r$, and the spectral integral over a neighborhood about this point provides the principal

contribution to the reflected field. For spectral numbers $\nu < \rho_o$, $\xi_l^{-'} / \xi_l^{+'} \to -i\exp\left(-4i(-\hat{y})^{3/2}/3\right)$, which also winds up rapidly with decreasing $\nu < \rho_o$. Therefore, for near-grazing reflections, only spectral numbers very near but below ρ_o contribute significantly to the spectral integrals for the reflecting part, but they do so only when the reflected ray exists with $\rho_r \le \rho_o$. Figure 5-38, which we discuss later, shows the impact parameter diagram for both the direct and reflected rays for a reflecting surface embedded in a refracting medium. It shows the narrow separation between ρ_r and ρ_o for near-grazing conditions.

Applying the spectral coefficients for the outgoing part of the wave given in Eq. (5.14-6) to the spectral integrals in Eq. (5.8-1b), one obtains

$$\left.\begin{aligned} E_r^+(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\rho_{\mathrm{L}}^2}\int_0^\infty \nu^2\left(\frac{a_l^+(\rho_{\mathrm{L}})}{2} + b_l^{(0)}(\rho_{\mathrm{L}})\right)\xi_l^+(\rho_{\mathrm{L}})\,\mathrm{P}_l^1(\cos\theta_{\mathrm{L}})\,d\nu \\ E_\theta^+(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\rho_{\mathrm{L}}}\int_0^\infty \left(\frac{a_l^+(\rho_{\mathrm{L}})}{2} + b_l^{(0)}(\rho_{\mathrm{L}})\right)\xi_l^{+'}(\rho_{\mathrm{L}})\frac{\partial}{\partial\theta}\mathrm{P}_l^1(\cos\theta_{\mathrm{L}})\,d\nu \end{aligned}\right\} \quad (5.14\text{-}7)$$

Here the forms for the spectral coefficients for the direct and reflected waves are given in Eqs. (5.14-5) and (5.14-6). The spectral coefficient for the incoming part for the direct wave at the LEO, $a_l^-(\rho_{\mathrm{L}})$, is excluded, because the LEO is located well into the outgoing region.

Asymptotic forms can be used everywhere in these integrals except in the reflection coefficients for spectral numbers near $\nu = \rho_o$. However, the Airy functions may be used in place of the spherical Hankel functions, $\xi_l^{-'} / \xi_l^{+'} \doteq (\mathrm{Ai}'[\hat{y}] + i\,\mathrm{Bi}'[\hat{y}]) / (\mathrm{Ai}'[\hat{y}] - i\,\mathrm{Bi}'[\hat{y}])$.

We now consider the stationary phase points in spectral number in the phasor form of the integrands in Eq. (5.14-7). For an outgoing wave, these phasor forms are given by Eqs. (5.9-6) and (5.9-24) but are augmented here by the inclusion of the reflection coefficients. We note that $\left|\xi_l^- / \xi_l^+\right| = 1$ for real values of l and ρ. Therefore, we may write this ratio in phasor form:

$$\left.\begin{aligned} \frac{\xi_l^-}{\xi_l^+} &= \exp(i2\Omega), \quad \Omega = \tan^{-1}\left(\frac{\mathrm{Bi}'[\hat{y}]}{\mathrm{Ai}'[\hat{y}]}\right) \end{aligned}\right\} \quad (5.14\text{-}8)$$

Let $I(\rho_{\mathrm{L}},\theta_{\mathrm{L}})$ denote the part of the field at the LEO for a perfectly reflecting sphere that is due to the ratio $\xi_l^-(\rho_o)/\xi_l^+(\rho_o)$. It follows from Eqs. (5.14-6) through (5.14-8) that Eq. (5.9-24) for this part becomes

$$
\left.
\begin{aligned}
I_r(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta_{\mathrm{L}}}}\int_0^\infty \left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu}\right)^{1/2} e^{i\left(2\Omega+\Psi_o(+,-)\right)}d\nu,\\
I_\theta(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta_{\mathrm{L}}}}\int_0^\infty \left(\sin\theta_\nu\cos\theta_\nu\right)^{1/2} e^{i\left(2\Omega+\Psi_o(+,-)\right)}d\nu,\\
\Psi_o(+,-) &= D_\nu + \nu(\theta_\nu-\theta_{\mathrm{L}}) - 2G[\rho_o,\nu] - \frac{\pi}{4},\\
D_\nu &= \sqrt{\rho_{\mathrm{L}}^2-\nu^2},\quad \theta_\nu = \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}}\right),\\
\Omega &= \tan^{-1}\left(\frac{\mathrm{Bi}'[\hat{y}]}{\mathrm{Ai}'[\hat{y}]}\right),\ \ \hat{y} \doteq K_{\rho_o}^{-1}(\nu-\rho_o)
\end{aligned}
\right\} \qquad (5.14\text{-}9)
$$

The total direct and reflected field at the LEO is given by

$$
\left.
\begin{aligned}
E_r(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= I_r(r_{\mathrm{L}},\theta_{\mathrm{L}}) + J_r(r_{\mathrm{L}},\theta_{\mathrm{L}})\\
E_\theta(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= I_\theta(r_{\mathrm{L}},\theta_{\mathrm{L}}) + J_\theta(r_{\mathrm{L}},\theta_{\mathrm{L}})
\end{aligned}
\right\} \qquad (5.14\text{-}10)
$$

where from Eq. (5.14-6) we have

$$
\left.
\begin{aligned}
J_r(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta_{\mathrm{L}}}}\int_{\nu_o^\dagger}^\infty \left(\frac{\sin^3\theta_\nu}{\cos\theta_\nu}\right)^{1/2}\left(e^{i\Psi^\dagger(+,-)} - e^{i\Psi_o(+,-)}\right)d\nu,\\
J_\theta(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\sqrt{2\pi\rho\sin\theta_{\mathrm{L}}}}\int_{\nu_o^\dagger}^\infty \left(\frac{\sin 2\theta_\nu}{2}\right)^{1/2}\left(e^{i\Psi^\dagger(+,-)} - e^{i\Psi_o(+,-)}\right)d\nu,\\
\Psi^\dagger(+,-) &= D_\nu + \nu\left(\theta_\nu-\theta_{\mathrm{L}}\right) - 2G\left[\rho^\dagger(\nu),\nu\right] - \frac{\pi}{4},\\
\Psi_o(+,-) &= D_\nu + \nu\left(\theta_\nu-\theta_{\mathrm{L}}\right) - 2G[\rho_o,\nu] - \frac{\pi}{4},\\
\nu_o^\dagger &\doteq \rho_o + \hat{y}^\dagger K_{\rho_o},\ \ \rho^\dagger(\nu) \doteq \nu - \hat{y}^\dagger K_{\rho_o}
\end{aligned}
\right\} \qquad (5.14\text{-}11)
$$

The contribution from the $\Psi_o(+,-)$ term to these J integrals mostly will be negligible when $\nu > \rho_o$. Moreover, because $2\Omega \to \pi$ rapidly for $\nu > \rho_o$, this region in the I integral given in Eq. (5.14-9) will essentially completely cancel this $\Psi_o(+,-)$ term in the J integral. But the $\Psi^\dagger(+,-)$ part of the J integrals in Eq. (5.14-11) accounts for the direct field at the LEO. When $\rho_d >> \rho_o$, this field is undisturbed by the reflecting surface and the stationary phase point in this integral $\nu^* >> \rho_o$; but when $\rho_d \approx \rho_o$, $\nu^* \approx \rho_o$ and J yields a knife-edge diffraction pattern.

5.14.1 Stationary Phase Analysis

Let us now examine the possible stationary phase neighborhoods for the reflection integrals I_r and I_θ given in Eq. (5.14-9). We seek the zero points for $\partial\left(2\Omega + \Psi_o\right)/\partial\nu$. From Eqs. (5.14-8) and (5.14-9), it follows that the stationary phase condition is given by

$$\left.\begin{aligned} &\frac{\partial}{\partial\nu}\left(2\Omega + \Psi_o\right) \doteq -\frac{2\hat{y}}{\pi\,\mathrm{M}[\hat{y}]K_{\rho_o}} + \theta_\nu - \theta_{\mathrm{L}} - 2\frac{\partial G[\rho_o,\nu]}{\partial\nu}, \\ &\mathrm{M}[\hat{y}] = \mathrm{Ai}'^2[\hat{y}] + \mathrm{Bi}'^2[\hat{y}], \quad \hat{y} \doteq K_{\rho_o}^{-1}(\nu - \rho_o) \end{aligned}\right\} \tag{5.14-12}$$

We assume here that the overlying medium has a monotonic refractivity gradient. Therefore, the direct ray system is unique (no multipath). One can show in this case that, for LEO orbital angular positions such that $\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o,\rho_o) \geq \theta_{\rho_o} \doteq \theta_o + N_o\tan\theta_o$, where $\theta_o = \sin^{-1}(r_o/r_{\mathrm{L}})$, there is a single negative stationary phase point, i.e., $\hat{y}^* < 0$. Otherwise, that point lies in positive territory, and the stationary phase contribution from there will largely cancel the $\Psi_o(+,-)$ contribution in the J integral. The latter situation corresponds to the LEO lying in the refracted shadow of the reflecting sphere. The LEO orbital position $\theta_{\mathrm{L}o} = \theta_{\rho_o} - 2\tilde{\alpha}(\rho_o,\rho_o)$ marks the shadow boundary, the beginning (for the setting case) of the eclipse of the GPS satellite by the reflecting sphere. Thus, except in the immediate vicinity of this shadow boundary, the stationary phase value for the spectral number will be less than ρ_o, that is, $\nu^* \doteq \rho_o + K_{\rho_o}\hat{y}^*$ with $\hat{y}^* < 0$, when $\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o,\rho_o) \geq \theta_{\rho_o}$.

For the J integral, essentially only the $\Psi^\dagger(+,-)$ term contributes to the field when combined with the I integral. Except in the interval $\nu^\dagger \leq \nu \leq\sim \nu^\dagger + 3K_{\rho_o}\sim$, the $\Psi_o(+,-)$ term in the J integral rapidly winds up, contributing negligibly to the integral. It follows that, to calculate the total field at the LEO using the stationary phase technique, one needs only the contribution to the I integrals in Eq. (5.14-9) from the stationary phase neighborhood provided by the condition in Eq. (5.14-12), plus the $\exp(i\Psi^\dagger(+,-))$ contribution from the J integral in Eq. (5.14-11). For $\rho_d >> \rho_o$, the latter integral gives essentially the direct field at the LEO unperturbed by the reflecting sphere but refracted by the overlying medium. The spectral treatment for this direct ray has been previously discussed in Section 5.9, and its stationary phase solution is presented there. Note the lower limit ρ_o in spectral number for the integration in Eq. (5.14-11). If the LEO orbital position is such that the impact parameter for a ray unperturbed by the reflecting surface is less than ρ_o, i.e., $\rho_d < \rho_o$, then this J integral will not contribute

significantly to the field. Moreover, if $\theta_{\mathrm{L}} < \theta_{\rho_o} - 2\tilde{\alpha}(\rho_o, \rho_o)$, the I integrals also have no stationary phase points for negative values of $\hat{y}$, and the contribution from the positive stationary phase points in $\hat{y}$ will be essentially completely canceled by the $\exp(i\Psi_o(+,-))$ term in the J integral. For orbital positions well into this region, it will be dark.

If $\hat{y}$ is sufficiently negative, i.e., $\hat{y} <\sim -2$, we may use the negative argument asymptotic forms in M[$\hat{y}$]. In this case, $\pi\,\mathrm{M}[\hat{y}] \to (-\hat{y})^{1/2}(1 - 7/(32\hat{y}^3) + \cdots)$. It follows that $2\Omega + \Psi_o(+,-)$ in the I integrals is given by

$$\left.\begin{aligned}
&2\Omega + \Psi_o(+,-) = -\frac{4}{3}(-\hat{y})^{3/2} - \frac{\pi}{2} + \Psi_o(+,-) \\
&\text{or} \\
&2\Omega + \Psi_o(+,-) = 2\left(\nu\left(\frac{\pi}{2} - \theta_\nu^o\right) - \sqrt{\rho_o^2 - \nu^2}\right) \\
&+\sqrt{\rho_{\mathrm{L}}^2 - \nu^2} + \nu(\theta_\nu - \theta_{\mathrm{L}}) - 2G[\rho_o, \nu] - \frac{3\pi}{4}, \\
&\theta_\nu = \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}}\right),\ \ \theta_\nu^o = \sin^{-1}\left(\frac{\nu}{\rho_o}\right), \\
&\frac{2}{3}(-\hat{y})^{3/2} = \sqrt{\rho_o^2 - \nu^2} - \nu\cos^{-1}\left(\frac{\nu}{\rho_o}\right), \quad \nu \le \rho_o
\end{aligned}\right\} \qquad (5.14\text{-}13)$$

Although the power series expression for $\hat{y}$ in terms of spectral number ν and ρ_o for the reflecting sphere, which is given in Eq. (5.14-12), is adequate, its exact form is given in Eq. (5.14-13) for $\hat{y} < 0$. See Eq. (5.4-3) for further discussion. The exact form is useful in a later discussion showing the correspondence between stationary phase in spectral number and the law of reflection.

It follows upon setting $\partial G[\rho_o, \nu]/\partial\nu = \tilde{\alpha}(\rho_o, \nu)$ and $\partial\left(2\Omega + \Psi_o\right)/\partial\nu = 0$ that the stationary phase point for $\hat{y}^* < 0$ is given by the condition

$$2\theta_{\nu^*}^o - \theta_{\nu^*} + \theta_{\mathrm{L}} + 2\tilde{\alpha}\left(\rho_o, \nu^*\right) = \pi \qquad (5.14\text{-}14a)$$

or

$$\hat{y}^* \doteq -\frac{1}{4}K_{\rho_o}^2\left(\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o, \nu^*) - \theta_{\nu^*}\right)^2 \qquad (5.14\text{-}14b)$$

To obtain an explicit value for ν^* in terms of θ_{L}, we note that for near-grazing reflections $\theta_{\mathrm{L}} + 2\tilde{\alpha}\left(\rho_o, \nu^*\right)$ will be close in value to θ_o, as shown in Fig. 5-37. We expand θ_{ν^*} in a Taylor series expansion about θ_{ρ_o}. Here θ_{ρ_o} and θ_{L}^A are defined by

$$\left.\begin{aligned} \theta_{\rho_o} &= \sin^{-1}\left(\frac{\rho_o}{\rho_{\mathrm{L}}}\right) \doteq \theta_o + N_o \tan\theta_o \\ \theta_o &= \sin^{-1}\left(\frac{r_o}{r_{\mathrm{L}}}\right) \\ \theta_{\mathrm{L}}^A &= \theta_{\mathrm{L}} + 2\tilde{\alpha}\left(\rho_o, \nu^*\right) \end{aligned}\right\} \tag{5.14-15}$$

The angle θ_{L}^A is the apparent orbit angle for the LEO. Because of the incoming and outgoing bending, $\tilde{\alpha}\left(\rho_o, \nu^*\right)$ on each leg, our coordinate frame will have been effectively rotated clockwise by $2\tilde{\alpha}\left(\rho_o, \nu^*\right)$ (see Fig. 5-36). Thus, the angles of incidence and reflection will be less by an amount $\tilde{\alpha}\left(\rho_o, \nu^*\right)$ than they would be for the case of a reflecting sphere in a homogeneous medium.

We rewrite Eq. (5.14-14a) as $2\theta_{\nu^*}^o - \left(\theta_{\nu^*} - \theta_{\rho_o}\right) + \left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right) = \pi$. Expanding θ_{ν^*} about θ_{ρ_o} in powers of $\nu^* - \rho_o$, we obtain

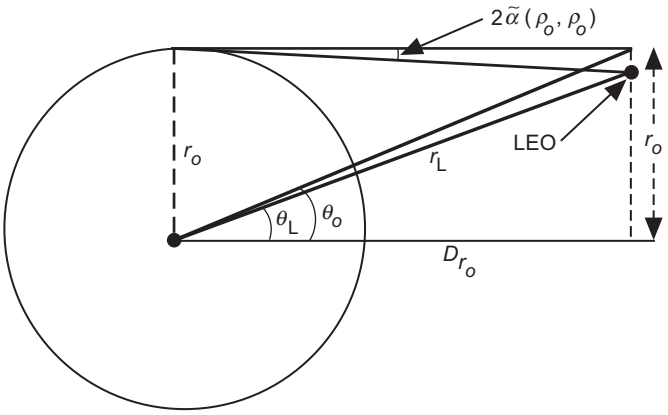

Fig. 5-37. Geometry for the LEO near the shadow boundary of a reflecting sphere embedded in a refracting medium.

$$\theta_{\nu^*} = \theta_{\rho_o} + \frac{\nu^* - \rho_o}{\sqrt{\rho_{\mathrm{L}}^2 - \rho_o^2}} + \cdots = \theta_{\rho_o} + \left(\sin\theta^o_{\nu^*} - 1\right)\tan\theta_{\rho_o} + \cdots \tag{5.14-16}$$

Noting that for near-grazing reflections $\theta^o_{\nu^*} \approx \pi/2$, it follows from Eq. (5.14-14) that $2\Omega + \Psi_o(+,-)$ has a stationary value in spectral number when $\nu = \nu^*$ or $\hat{y} = \hat{y}^*$, which are given by the conditions

$$\left.\begin{aligned} \theta^o_{\nu^*} &= \frac{\pi}{2} - \frac{\theta^A_{\mathrm{L}} - \theta_{\rho_o}}{2} - \frac{\left(\theta^A_{\mathrm{L}} - \theta_{\rho_o}\right)^2}{16}\tan\theta_{\rho_o} + \cdots \\ \nu^* &= \rho_o \sin\theta^o_{\nu^*} = \rho_o\left(1 - \frac{\left(\theta^A_{\mathrm{L}} - \theta_{\rho_o}\right)^2}{8} + \cdots\right) \\ \hat{y}^* &= -\frac{K^2_{\rho_o}}{4}\left(\theta^A_{\mathrm{L}} - \theta_{\rho_o}\right)^2 + \cdots \end{aligned}\right\} \tag{5.14-17}$$

Because $\theta^A_{\mathrm{L}} = \theta_{\mathrm{L}} + 2\tilde{\alpha}\left(\rho_o, \nu^*\right)$, an iteration is required in Eq. (5.14-17) to determine ν^* once a specific form for $2\tilde{\alpha}(\rho_o, \nu)$ is given. For the specific exponential refractivity profile used in the numerical examples here for dry air, $2\tilde{\alpha}(\rho_o, \nu)$ for the reflected ray is given accurately by Eq. (5.8-3). It can be shown that in this case

$$\left.\begin{aligned} 2\tilde{\alpha}\left(\rho_o, \nu^*\right) &= 2\tilde{\alpha}_o - \frac{\beta}{1+\beta}\left(\theta_{\mathrm{L}} + 2\tilde{\alpha}_o - \theta_{\rho_o}\right) \\ &\times\left(1 + \frac{\rho_o}{24H(1+\beta)^3}\left(\theta_{\mathrm{L}} + 2\tilde{\alpha}_o - \theta_{\rho_o}\right)^2 + \cdots\right), \\ \tilde{\alpha}_o &= \tilde{\alpha}(\rho_o, \rho_o), \quad \beta = \frac{N_o \rho_o}{H} \end{aligned}\right\} \tag{5.14-18}$$

This expression is accurate to 1 percent for near-grazing reflections, i.e., $\left|\theta_{\mathrm{L}} + 2\tilde{\alpha}_o - \theta_{\rho_o}\right| < 0.05$. It follows for this case that

$$\nu^* \doteq \rho_o\left(1 - \frac{1}{8}\left(\frac{\theta_{\mathrm{L}} + 2\tilde{\alpha}_o - \theta_{\rho_o}}{1+\beta}\right)^2\right) \tag{5.14-19}$$

Figure 5-38 shows the impact parameter diagram for this reflection case. This figure uses the same refractivity model for the overlying medium as that used in Figs. 5-30 through 5-34. A circular LEO orbit was used with $r_{\mathrm{L}} = 1.1 r_o$.

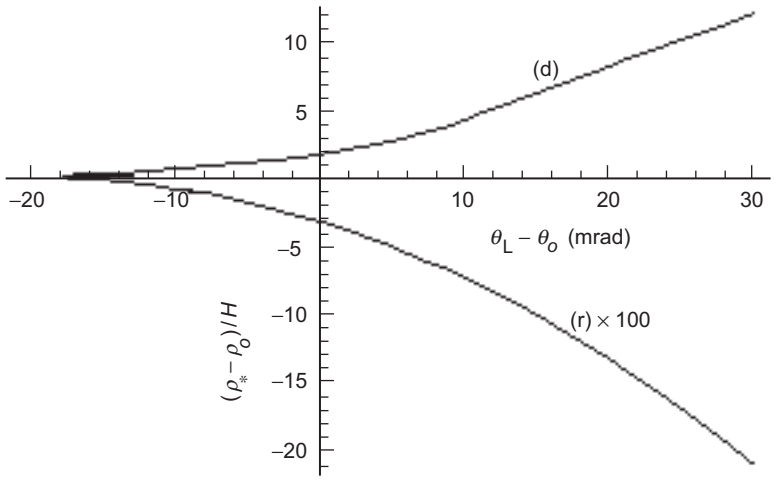

Fig. 5-38. Impact parameter diagram for a reflecting sphere embedded in a refracting medium versus orbit angle near grazing. The direct curve (d) is identical to the (m) curve in Fig. 5-31.

Therefore, the (d) curve in Fig. 5-38 for the direct ray is identical to the (m) curve in Fig. 5-31. The impact parameters for the (d) and (r) rays merge at $\theta_L - \theta_{\rho_o} = -20.4$, the grazing point. Divide the impact parameter value given in Fig. 5-38 for the reflected ray (r) by 100 to obtain the correct value. The reflected ray is very defocused for near-grazing conditions. At $\theta_L - \theta_{\rho_o} = 0$, it is about an order of magnitude more defocused than the (d) ray. This can be calculated by recalling from geometric optics that for a circular LEO orbit the defocusing factor is given by $\zeta = \left| D_{\rho_*} \left(d\theta_L / d\rho_* \right) \right|^{-1}$. If we set $\nu^* = \rho_*$ in the stationary phase condition given in Eq. (5.14-14a), then we can form the derivative $d\theta_L / d\rho_*$ while still satisfying the stationary phase condition as θ_L and ρ_* vary. For the defocusing factor ζ_r for the (r) ray, we obtain

$$\zeta_r^{-1} = D_{\rho_*} \left| \frac{d\theta_L}{d\rho_*} \right|_r^{-1} = \frac{2\cos\theta_{\rho_*}}{\sin\left(\left(\theta_L^A - \theta_{\rho_*} \right) / 2 \right) \sin\theta_{\rho_o}} - 1 + 2D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*} \quad \text{(5.14-20a)}$$

For near-grazing reflections, the term $\left(\theta_L^A - \theta_{\rho_*} \right)$ is small. Also, $\theta_{\rho_*} - \theta_{\rho_o} \doteq -\left(\theta_L^A - \theta_{\rho_o} \right)^2 / 8D\rho_o$, which is a very small term. Expanding about $\left(\theta_L^A - \theta_{\rho_o} \right)$ yields

$$\zeta_{\mathrm{r}}^{-1} \doteq \frac{4\cot\theta_{\rho_o}}{\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o, \rho_*) - \theta_{\rho_o}} - \left(1 - 2D_{\rho_o}\frac{\partial\tilde{\alpha}(\rho_o, \rho_*)}{\partial\rho_*}\right) \qquad (5.14\text{-}20\text{b})$$

The first term on the RHS is dominant for near-grazing reflections. At $\theta_{\mathrm{L}} - \theta_o = 0$ for the case shown in Fig. 5-38, the values of the first and second terms are about 120 and 20, respectively. Thus, Eq. (5.14-20) predicts about 0.08 for the amplitude $E_o\zeta_{\mathrm{r}}^{1/2}$ of the reflected ray at $\theta_{\mathrm{L}} - \theta_o = 0$. The amplitude of the (d) ray there is about 0.6, which decreases to less than 0.4 at the grazing point at $\theta_{\mathrm{L}} - \theta_o = -20.4$ mrad.

Figure 5-39 shows the impact parameter diagram for the reflected and direct rays for $\theta_o - 2\tilde{\alpha}(\rho_o, \rho_o) \le \theta_{\mathrm{L}} \le \pi$, which covers the entire range of LEO orbit angles for which a reflected ray exists, from the shadow boundary to a vertical reflection. Here Eq. (5.14-14a) has been used to solve for ρ_{r} using the same refractivity for the overlying medium and orbit models that are used in Fig. 5-30.

We note that $\left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right)K_{\rho_o}$ has a scale invariance provided that third- and higher-order terms in $\left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right)$ given for $\hat{y}^*$ in Eq. (5.14-17) can be ignored. The stationary phase points for a reflecting sphere of another radius $\tilde{\rho}_o$ is obtained from our problem merely by applying the scale factor $(\tilde{\rho}_o / \rho_o)^{1/3}$ to the results obtained here for $\left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right)K_{\rho_o}$. One of the practical aspects of this is that the a priori value for $\tilde{\rho}_o$—for example, in the case of the topography of the ocean surface—is uncertain to some extent in an actual observation

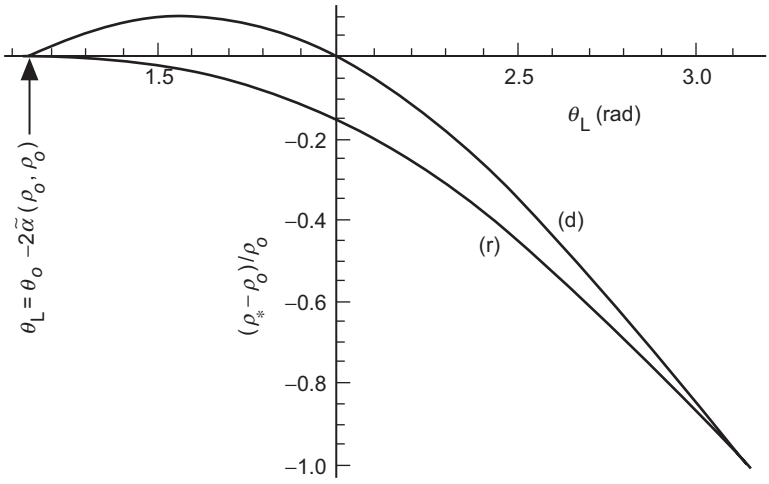

Fig. 5-39. Large scale version of the impact parameter diagram over the entire LEO orbit above the grazing point.

sequence of reflected and direct signals. This is particular true for near-grazing conditions. This scale-invariance property may be useful.

For $\hat{y} < 0$, we can determine from the I integrals in Eq. (5.14-9), using the stationary phase technique, the contribution from the stationary phase point to the field at the LEO from the reflecting sphere. We need the second derivative $\partial^2\left(2\Omega + \Psi_o\right)/\partial\nu^2$ evaluated at the stationary phase point, given by either Eq. (5.14-14) or Eq. (5.14-17). This is given by

$$\left(D_\nu \frac{\partial^2(2\Omega+\Psi)}{\partial\nu^2}\right)_{\nu^*} = \left(\frac{-2D_\nu}{\rho_o \cos\theta_\nu^o} + 1 - 2D_\nu \frac{\partial^2 G}{\partial\nu^2}\right)_{\nu^*} \tag{5.14-21a}$$

Replacing $\left(\partial^2 G/\partial\nu^2\right)_{\nu^*}$ with $\partial\tilde{\alpha}/\partial\rho_*$ and also $\theta_{\nu^*}^o$ with its expansion in terms of $\left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right)$, which is given in Eq. (5.14-27), we obtain

$$\left(\frac{\partial^2(2\Omega+\Psi)}{D_\nu \partial\nu^2}\right)_{\nu^*} \doteq \frac{-4\cot\theta_{\rho_*}}{\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o,\rho_*) - \theta_{\rho_*}} + 1 - 2D_{\rho_*}\frac{\partial\tilde{\alpha}(\rho_o,\rho_*)}{\partial\rho_*} \tag{5.14-21b}$$

which is the same as the defocusing factor from geometric optics given in Eq. (5.14-20b). For orbital positions near but above the shadow boundary ($\hat{y}^* < 0$), it follows from the stationary phase technique that the field at the LEO from the reflecting sphere is given by

$$\begin{aligned} I &\doteq \frac{E_o \exp\left(i\Phi^*\right)}{\sqrt{-2\pi i\rho_{\mathrm{L}}\cos\theta_{\rho_o}}}\int_0^\infty \exp\left\{ i\frac{1}{2}\left.\frac{\partial^2\left(2\Omega+\Psi_o\right)}{\partial\nu^2}\right|_{\nu^*}\left(\nu-\nu^*\right)^2\right\} d\nu \\ &\doteq \frac{E_o \exp\left(i\Phi^*\right)}{\sqrt{-2\pi i\rho_{\mathrm{L}}\cos\theta_{\rho_o}}}\left\{\frac{1}{\pi}\left|\frac{\partial^2\left(2\Omega+\Psi_o\right)}{\partial\nu^2}\right|_{\nu^*}\right\}^{-1/2}\int_{-\infty}^{\infty}\exp\left(-i\frac{\pi}{2}x^2\right)dx \end{aligned} \tag{5.14-22a}$$

or

$$I\left(\rho_{\mathrm{L}},\theta_{\mathrm{L}}\right) \doteq E_o \exp\left(i\Phi^*\right)\sqrt{\xi_{\mathrm{r}}^*} \tag{5.14-22b}$$

where

$$
\left.
\begin{aligned}
&\left(\xi_r^*\right)^{-(1/2)} = \left(D_\nu \left| \frac{\partial^2 \left(2\Omega + \Psi_o\right)}{\partial \nu^2} \right| \right)_{\nu^*}^{1/2} \\
&\doteq \sqrt{\left| \frac{4\cot\theta_{\rho_*}}{\theta_{\mathrm{L}} + 2\tilde{\alpha}(\rho_o, \rho_*) - \theta_{\rho_*}} - \left(1 - 2D_{\rho_*} \frac{\partial \tilde{\alpha}(\rho_o, \rho_*)}{\partial \rho_*}\right) \right|}, \\
&\Phi^* = \left(2\Omega + \Psi_o - \pi/4\right)^* \\
&= -\pi + \sqrt{\rho_{\mathrm{L}}^2 - \nu^{*2}} - 2\sqrt{\rho_o^2 - \nu^{*2}} + 2\left(\nu^* \tilde{\alpha}\left(\rho_o, \nu^*\right) - G\left[\rho_o, \nu^*\right]\right)
\end{aligned}
\right\} \qquad (5.14\text{-}23)
$$

and also

$$
\left.
\begin{aligned}
&\theta_{\rho_o} = \sin^{-1}\left(\frac{\rho_o}{\rho_{\mathrm{L}}}\right) \doteq \theta_o + N_o \tan\theta_o, \quad \theta_o = \sin^{-1}\left(\frac{r_o}{r_{\mathrm{L}}}\right), \\
&\theta_{\mathrm{L}}^A = \theta_{\mathrm{L}} + 2\tilde{\alpha}\left(\rho_o, \nu^*\right), \quad \nu^* = \rho_o \left(1 - \frac{\left(\theta_{\mathrm{L}}^A - \theta_{\rho_o}\right)^2}{8} + \cdots\right), \\
&\rho_o = k n_o r_o, \quad \rho_{\mathrm{L}} = k r_{\mathrm{L}}, \quad D_{\rho_o} = \sqrt{\rho_{\mathrm{L}}^2 - \rho_o^2}
\end{aligned}
\right\} \qquad (5.14\text{-}24)
$$

The stationary phase term Φ^* in Eq. (5.14-23) gives the stationary value of the phase at the LEO from the reflecting sphere relative to the phase of the field (neither refracted nor reflected) at the centerline at $\theta = \pi/2$ in Fig. 5-43(a). The π term gives the phase reversal from the reflection. The term $\sqrt{\rho_{\mathrm{L}}^2 - \nu^{*2}} - 2\sqrt{\rho_o^2 - \nu^{*2}}$ accounts for the difference in optical path length relative to the centerline in Fig. 5-43(a). The first term, $\sqrt{\rho_{\mathrm{L}}^2 - \nu^{*2}}$, gives the length along a straight line from the tangent point on the circle of radius ν^* centered at the origin to the LEO. The second term, $-2\sqrt{\rho_o^2 - \nu^{*2}}$, subtracts the length ($\sqrt{\rho_o^2 - \nu^{*2}}$) between that tangent point and the reflection point along this straight line, and it also subtracts the distance ($\sqrt{\rho_o^2 - \nu^{*2}}$) that the reflection point has moved counterclockwise from the centerline. The third term, $2\nu^* \tilde{\alpha}\left(\rho_o, \nu^*\right)$, accounts for the extra path length along an arc of radius ν^* resulting from the refractive bending. The fourth term, $-2G\left[\rho_o, \nu^*\right]$, accounts for the extra round-trip delay from infinity down to a radius ν^* resulting from the refractive gradient in the overlying medium. These third and fourth

terms may be rewritten in the form $2\left(\nu^*\tilde{\alpha}\left(\rho_o,\nu^*\right)-G\left[\rho_o,\nu^*\right]\right)$ $\doteq 2\left(\rho_o\tilde{\alpha}\left(\rho_o,\nu^*\right)-G[\rho_o,\rho_o]\right)$.

The form of the amplitude signature $\sqrt{\zeta_{\mathrm{r}}}$ given in Eq. (5.14-23) is dominated by the $\left(\theta_{\mathrm{L}}^A-\theta_{\rho_o}\right)^{-1}$ term for near-grazing reflections. This dominance continues with increasing θ_{L} until the term $\left(1-2D_{\rho_{\mathrm{r}}}\partial\tilde{\alpha}/\partial\nu\right)$ in $\partial^2\left(2\left|+\Psi_o\right.\right)/\partial\nu^2$ given in Eq. (5.14-21) becomes dominant. Equating these two terms yields a threshold of $\theta_{\mathrm{L}}^A-\theta_{\rho_o}\approx 4\zeta^*\cot\theta_{\rho_o}$, or when θ_{L} becomes about a quarter of a radian greater than $\theta_o-2\tilde{\alpha}(\rho_o,\rho_o)$. Here the impact parameter of the direct ray is well above the atmosphere. For impact parameters above this threshold, $\zeta_{\mathrm{r}}^{-1}\to 1-2D_\nu\partial\tilde{\alpha}(\rho_o,\nu)/\partial\nu$ is the defocusing factor at the reflecting surface for a ray with an impact parameter value of ν. In this case, the amplitude of the reflected signal at the LEO reduces to $|I|\to E_o\zeta_{\mathrm{r}}^{1/2}$, the same form that the amplitude of the direct ray has except that the defocusing factor is based on the bending angle $\tilde{\alpha}\left(\rho_o,\nu^*\right)$ for the reflected ray rather than $\tilde{\alpha}(\rho_*,\rho_*)$ for the direct ray. However, polarization effects from reflections at these steeper angles must be considered. The GPS receiver onboard the LEO is configured for RHCP.

5.14.2 Results from Wave Theory

Figures 5-40 through 5-42 show the amplitude of the field at the LEO from a GPS satellite that is being occulted by a perfectly reflecting sphere embedded in a refracting atmosphere. Figure 5-40 shows the amplitude $E^{(\mathrm{d})}$ from the direct ray only. This corresponds to the (d) ray in the impact parameter diagram shown in Fig. 5-38. Figure 5-41 shows the amplitude $E^{(\mathrm{r})}$ from the reflected ray (r) only, and Fig. 5-42 shows the amplitude of the complete field. These figures show the transition over about 25 mrad in orbit angle, or roughly 1/2 minute in time, as the LEO moves into the shadow. Here an exponential profile [Eq. (5.8-2a)] has been used for the overlying medium with $N_o=270\times10^{-6}$, $k^{-1}H=7$ km, $r_o=6378$ km, and $k=3.31\times10^4$ km^{-1}. Also, $r_{\mathrm{L}}=1.1r_o$ and $\theta_o=\sin^{-1}\left(r_o/r_{\mathrm{L}}\right)$. From Eq. (5.6-5) or Eq. (5.8-3), one can show that $2\tilde{\alpha}(\rho_o,\rho_o)=20.4$ mrad for these parameter values. Figures 5-40 through 5-42 were obtained from a numerical integration of Eq. (5.8-1b) aided by the stationary phase technique. The impact parameter diagram in Fig. 5-38 effectively provides those stationary phase points in spectral number for a given value of θ_{L}.

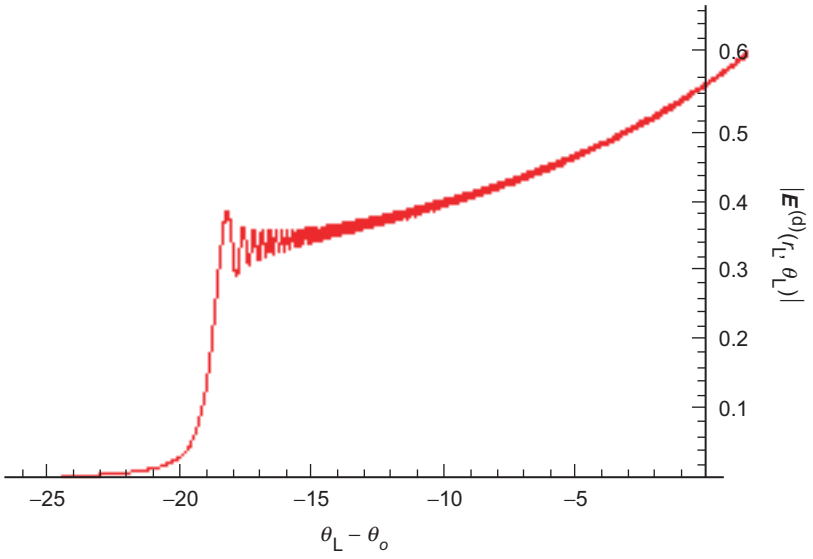

Fig. 5-40. A perfectly reflecting sphere embedded in a refracting medium. Wave theory prediction of the amplitude $|E^{(d)}(r_L, \theta_L)|$ of the field at the LEO from the direct ray only.

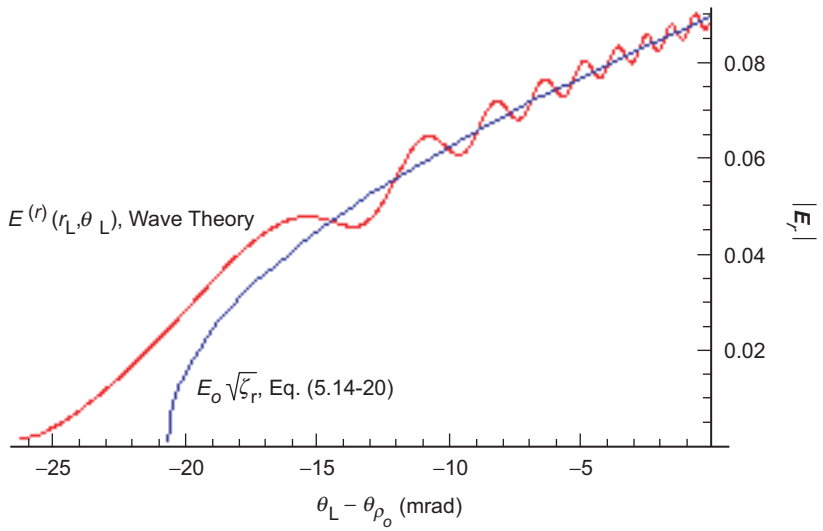

Fig. 5-41. Amplitude of the field at the LEO from a perfectly reflecting sphere embedded in a refracting medium, reflected ray (r) only.

The mean amplitude in Fig. 5-40 shows a rapid decay for θ_L values such that $\rho_d < \rho_o$, or $\theta_L - \theta_o < -2\tilde{\alpha}(\rho_o, \rho_o)$. This is the penumbra region. The GPS satellite being eclipsed by the reflecting sphere causes the knife-edge diffraction pattern in Fig. 5-40. For $\rho_d > \rho_o$, Fig. 5-41 clearly shows the $\sqrt{\zeta_r}$ signature in the mean amplitude for the reflected wave that is predicted from the stationary phase technique in Eq. (5.14-23). This agrees with the geometric optics prediction in Eq. (5.14-20) for LEO orbit angles above the shadow boundary, $\theta_L > \theta_o - 2\tilde{\alpha}(\rho_o, \rho_o)$.

Figure 5-42, which gives the wave theory prediction of the total field at the LEO over the same near-grazing orbit angles shown in Figs. 5-40 and 5-41, shows the fringes from interference between the direct and reflected rays. Very high fringe frequencies develop as the separation in altitude between the impact parameters of the reflected and direct rays increases with increasing θ_L. The single-sided amplitude of the interference fringes here nearly equals the amplitude of the reflected ray (r) in Fig. 5-41. The fringe amplitude would be significantly reduced for an imperfectly reflecting sphere. The fringe frequency here ranges from about 60 Hz at $\theta_L - \theta_{\rho_o} = 0$, about 20 s before eclipse for the setting case, to zero at $\theta_L - \theta_o = -2\tilde{\alpha}(\rho_o, \rho_o) = -20.4$ mrad, at the refracted shadow boundary.

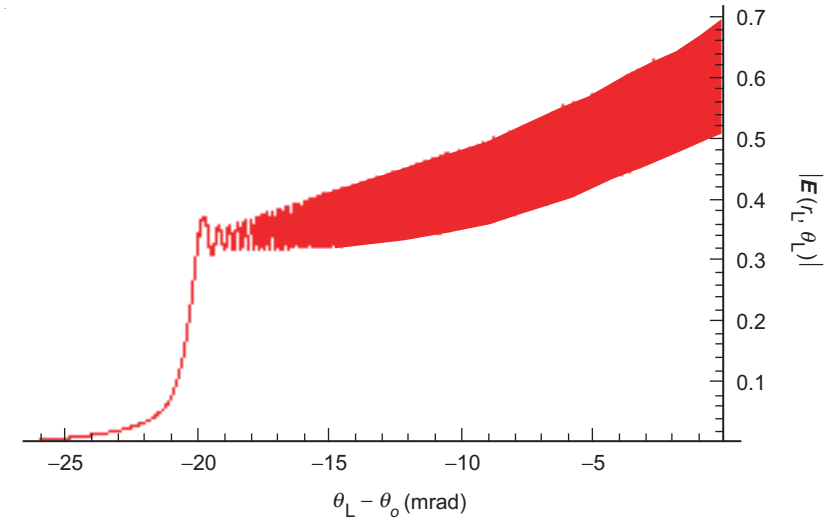

Fig. 5-42. Amplitude of the complete field $(d)+(r)$ at the LEO from a perfectly reflecting sphere embedded in a refracting medium, the direct ray from Fig. 5-40 plus the reflected ray from Fig. 5-41.

5.14.3 Law of Reflection

Chapter 3 provides a geometric optics interpretation of the stationary phase point for a reflecting sphere in a homogeneous medium. There it was shown for a point $(r_{\mathrm{L}},\theta_{\mathrm{L}})$ sufficiently away from the shadow boundary (so that the negative argument asymptotic terms for the Airy functions are valid) that, when the spectral number assumes its stationary phase point, the condition for stationarity in spectral number becomes identical with the law of reflection in geometric optics. We also can similarly interpret the stationary phase results for a reflecting sphere embedded in a refracting medium. Referring to Figs. 5-43(a) and 5-43(b), which are representations in spectral number space, we have two concentric circles in each figure. The outer circle corresponds to the reflecting sphere, and it has a fixed radius ρ_o in spectral number space. The inner one of variable radius is a circle of radius ν, the spectral number. In Fig. 5-43(a), $\nu=\nu^*$, the stationary phase value. In Fig. 5-43(b), the provisional reflection point has been moved clockwise; here $\nu<\nu^*$, and the angle of incidence ι is greater than the angle of reflection $\rho=\theta_\nu^o$. Noting the triangle ABC in Fig. 5-43(b), we see that the sum of the interior angles of this triangle satisfies the relationship

$$\iota+\rho-\theta_\nu+\theta_{\mathrm{L}}^A=\pi \tag{5.14-25}$$

for any provisional point of reflection on the circle of radius ρ_o in spectral number. But, at the actual reflection point where $\iota=\rho$, Eq. (5.14-25) becomes the same relationship given in Eq. (5.14-14a) for the stationary phase condition on the spectral number. We conclude that $\theta_{\nu^*}^o$, which always equals the angle of reflection ρ by construction, also equals the angle of incidence ι when the spectral number assumes its stationary phase value, thus establishing the law of reflection. It follows that

$$\iota=\rho=\theta_{\nu^*}^o \doteq \frac{\pi}{2}-\frac{\theta_o^A-\theta_{\rho_*}}{2} \tag{5.14-26}$$

Lastly, we see in Fig. 5-43(a) that the position of the reflection point on the reflecting sphere is located at a point counterclockwise relative to the midpoint. This rotation, ε, is given by

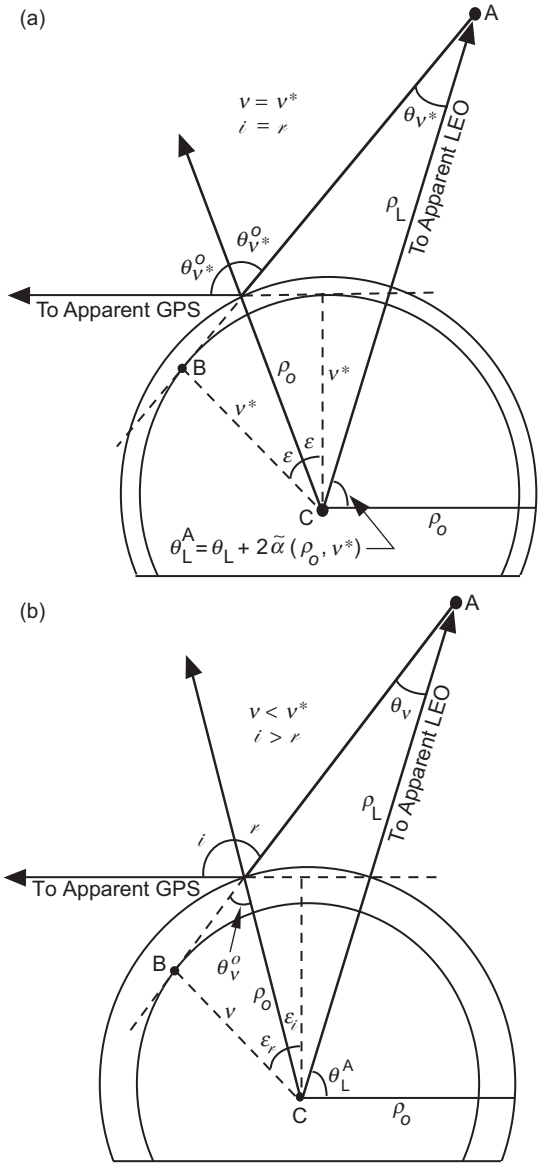

Fig. 5-43. The law of reflection; the geometry in spectral number space for a reflecting sphere embedded in a refracting medium. The outer circle describes a reflecting sphere of radius r_o, which maps in spectral number space into a fixed circle of radius ρ_o. The inner circle has a variable radius value, ν: (a) the radius of the inner circle is the stationary phase value, ν^*, which gives $i = r$, the law of reflection, and (b) $\nu < \nu^*$, a non-stationary phase value, which gives $i > r$.

$$\left.\begin{aligned}\varepsilon &= \frac{\pi}{2} - \theta_{\nu^*}^{o} = \frac{\theta_{\mathrm{L}}^{A} - \theta_{\rho_o}}{2} + \frac{\left(\theta_{\mathrm{L}}^{A} - \theta_{\rho_o}\right)^2}{16}\tan\theta_{\rho_o} + \cdots \\ &\doteq \frac{\theta_{\mathrm{L}} - \theta_o}{2} + \tilde{\alpha}\left(\rho_o, \nu^*\right) - \frac{1}{2}N_o \tan\theta_o\end{aligned}\right\} \qquad (5.14\text{-}27)$$

For the exponential refractivity model given in Eq. (5.8-2) and from the resulting bending-angle expression given in Eq. (5.8-3), we have

$$\varepsilon \doteq \frac{\theta_{\mathrm{L}} - \theta_o}{2} + \tilde{\alpha}(\rho_o, \rho_o) - \frac{N_o \rho_o}{2H}\left(\theta_{\mathrm{L}} - \theta_o\right) - \frac{1}{2}N_o \tan\theta_o \qquad (5.14\text{-}28)$$

Thus, the refractive bending (for a negative refractivity gradient, i.e., $\tilde{\alpha}\left(\rho_o, \nu^*\right) > 0$) acts to increase the counterclockwise rotation of the reflection point. This is qualitatively indicated in Fig. 5-36. For sea-level conditions and for $\theta_{\mathrm{L}} - \theta_o = 10$ mrad, the third and fourth terms on the RHS of Eq. (5.14-28) provide a correction of about 10 percent.

For a backward reflection case, that is, when $\theta_{\mathrm{L}} > \pi/2$, the incoming spectral coefficients $a_l^-(\rho_{\mathrm{L}})$ also would have to be taken into account. When $\theta_{\mathrm{L}} >> \pi/2$, this geometry would correspond more closely to a near-vertical reflection geometry than to the near-grazing one discussed here.

References

[1] C. Chapman and J. Orcutt, "The Computation of Body Wave Synthetic Seismograms in Laterally Homogeneous Media," *Reviews of Geophysics*, vol. 23, no. 2, pp. 105–163, 1985.

[2] G. Mie, "Beiträge zur Optik Truber Medien Speziell Kolliodaler Metallösungen," *Annales de Physique* (Leipzig), vol. 25, pp. 377–452, 1908.

[3] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference, and Diffraction of Light,* 6th ed., New York, New York: Pergamon Press, 1980.

[4] J. Jackson, *Classical Electrodynamics,* 2nd ed., Oxford, United Kingdom: J. Wiley & Sons, Inc., 1975.

[5] T. Eckersley and G. Millington, "Application of the Phase Integral Method to the Analysis of the Diffraction and Reflection of Wireless Waves Round the Earth." *Philos. Trans*., vol. 237, pp. 273–309, 1938.

[6] J. Mathews and R. Walker, *Mathematical Methods of Physics*, Menlo Park, California: The Benjamin/Cummings Publishing Co., 1970.

[7] M. Levy, *Parabolic Equation Methods for Electromagnetic Wave Propagation*, London, United Kingdom: The Institution for Electrical Engineers, 2000.

[8] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* U.S. Department of Commerce, National Bureau of Standards, U.S. Government Printing Office, Washington, D.C., 1964.

[9] A. Pavelyev, A. Zakharov, A. Kucherjavenkov, E. Molotov, I. Siderenko, I. Kucherjavenkova, and D. Pavelyev, "Propagation of Radio Waves Reflected from Earth's Surface at Grazing Angles between a Low-Orbit Space Station and a Geostationary Satellite," *Journal of Communications Technology and Electronics*, vol. 42, no. 1, pp. 51–57, 1997.

[10] G. Beyerle, K. Hocke, J. Wickert, T. Schmidt, C. Marquardt, and Ch. Reigber, "GPS Radio Occultations with CHAMP: A Radio Holographic Analysis of GPS Signal Propagation in the Troposphere and Surface Reflections," *Journal of Geophysical Research*, vol. 107, no. D24, 10.1029/ 2001JD001402, 2002.

Chapter 6
The Inverse Problem: Using Spectral Theory to Recover the Atmospheric Refractivity Profile

6.1 Introduction

Throughout Chapter 5, we mostly were concerned with the use of a full spectral representation of the electromagnetic field. We are not unmindful, however, of the potential suitability of this particular spectral technique to the inverse problem: determining the bending-angle and refractivity profiles from a time sequence of measurements of amplitude and phase of the received signal. This is especially interesting when adverse signal conditions prevail, that is, when the received signal exhibits significant amplitude and phase interference from multiple rays. In these situations, classical recovery algorithms, such as those using the excess Doppler with the Abel transform, can run into difficulty because of the non-uniqueness or even non-existence of the ray path. Caustic points also lead to a breakdown in the validity of geometric optics, i.e., second-order ray theory, on which the Abel transform algorithm is based. Spectral techniques, whether they are rooted in geometric optics or in a full-spectrum wave theory, can deal with these types of propagation problems. Although multiple rays with different bending angles can arrive at the low Earth orbiting (LEO) spacecraft at the same time, they cannot arrive simultaneously with the same excess Doppler values when spherical symmetry applies; those must be distinct. The transformation of the time series of observations into a spectral series provides a means to uniquely recover bending-angle and refractivity profiles because of this one-to-one relationship between bending angle and excess Doppler.

In this chapter, we briefly outline the use of the particular full-spectrum wave theory technique developed in Chapter 5 for the inverse problem. The spectral density function $G[\rho, \nu]$, which is the extra phase delay at the radial position r induced by the refractivity gradient on the νth spectral component of the wave, plays a central role. We will not discuss in detail the computational aspects using this particular spectral technique, nor the concomitant use of estimation theory techniques on noisy data.

We assume that the LEO has received an ordered time series of amplitude and phase measurements, appropriately stripped of the geometric LEO/Global Positioning System (GPS) Doppler signature and any other contributory error sources. For example, we assume perfect orbit knowledge from precision orbit determination (POD). We assume that the effects of oscillator variations in the various clocks affecting the phase measurements have been eliminated through redundant differential tracking, and we assume that ionospheric effects can be eliminated through dual-band L1 and L2 tracking, possibly supplemented with modeling to eliminate third-order effects. We also assume that the signal-to-noise ratio (SNR) of the signal is sufficient for a Nyquist sample rate to be practicable with respect to the bandwidth of the particular atmospheric signature under study.

Although this chapter does discuss a stand-alone Fourier approach for the inverse problem with radio occultation data, we note again that a principal utility of radio occultation data is for meteorology and numerical weather prediction. There the occultation data are merged in a timely way into a much richer and broader multi-sensor database, which is constrained by a comprehensive model that characterizes the atmosphere and controls atmospheric processes. This already is mentioned in Chapter 1. In this context, the difference between an actual radio occultation observation (phase, amplitude, etc.) and a predicted observation obtained by forward propagation of the GPS signal through the atmospheric model becomes a constraint among the free parameters of the model. The four-dimensional variational analysis (4DVAR) methodology for minimizing a quadratic cost function involving disparate competing databases constrained by the model is central to accurate weather prediction [1]. In this application, one usually would not recover the refractivity profile from a radio occultation profile alone, but only from within the 4DVAR context. Nevertheless, the stand-alone approach has some important uses, and it is discussed in the next sections.

6.2 GPS Receiver Operations

It is helpful to understand some basic operational aspects of the GPS receiver in measuring phase and amplitude. The BlackJack GPS receiver series is a modern, high accuracy, dual-band, digital receiver developed by the Jet Propulsion Laboratory (JPL) for scientific applications in space. As its

development has evolved over the past several years, it has successfully flown on over one-half dozen Earth satellites for navigation and timekeeping, precision orbit determination, geopotential mapping, ocean reflections, and limb sounding [2].

The following is a rather high-level account of this particular receiver. It omits important details, which are crucial to the actual recovery of atmospheric parameters, but which are less relevant to the discussion here, e.g., dual-frequency tracking to remove ionosphere effects, dealing with clock errors, digital signal processing, and correlation operations on noisy data [3,4].

In normal operations, the BlackJack receiver uses a closed-loop phase model before extracting the phase delay information to reduce the frequency of the received radio frequency (RF) signal from the GPS satellite to base band, a few tens of hertz. First, the received RF signal (carriers in the 1.2 to 1.6 GHz range) is 1-bit digital sampled in-phase and separately in-quadrature, that is, with the received RF phase shifted by 90 deg. To facilitate the signal processing by the receiver, the carrier frequency in these two parallel bit streams is reduced to an intermediate frequency (IF) of around 200 kHz, which is then further reduced to the base band frequency using an in-receiver phase model. This model is generated from a cubic polynomial fit to previous phase measurements made by the receiver over the past few tens of milliseconds, and in the closed-loop mode it is updated after each observational epoch with the latest phase measurement [4].

The L1 (1575 MHz) and L2 (1226 MHz) carriers of the navigation signals from each GPS satellite are derived from the same onboard master oscillator, and, therefore, they are initially coherent. The L1 carrier is phase modulated coherently with the clear access (C/A) and precision (P) ranging codes. The L2 carrier is phase-modulated with only the P code (on the present generation of GPS satellites), but coherently with the L1. These ranging codes are pseudorandom, phase-modulating square waves that fully suppress the carrier tone. Each transition of a code, occurring at a frequency or chip rate of 1.023 MHz for the C/A code and at 10.23 MHz for the P code, involves a change in phase of the carrier of either 0 deg or 180 deg in accordance with the pseudorandom algorithm specific to that particular code. Both carriers also are phase modulated with a header code operated at a 50-Hz chip rate. This very low rate code carries the almanac and timing information for the tracked satellite and other satellite health and housekeeping data. Each satellite broadcasts distinctive codes that are unique to that satellite. These codes are mutually orthogonal, and they also are orthogonal between GPS satellites. In the limit, cross-correlating two different codes yields a null result. Moreover, because the codes are pseudorandom, their auto-correlation function is triangular across an alignment offset of up to ± 1 chip period, and its value is zero outside of this range. This means that multiplying the same code with itself but time shifted also yields a null result unless the two components are aligned

within ±1 chip period, the spatial equivalent of about ±30 m for the P code. Therefore, cross-correlating the received signal from a specific GPS satellite with the appropriately time- and Doppler-shifted C/A and P code replicas effectively filters out the signals received from all other satellites and also signals with phase delay offsets greater than the chip period. Also, by aligning the codes, it enables a determination of the propagation delay between that satellite and the receiver plus any timekeeping difference between their clocks.

The base band signal embedded in the in-phase and the quadrature bit streams being processed by the receiver is a composite of all broadcasting GPS satellites in view of the receiver's antennas. Also, the sign of each bit, +1 or –1, is dominated by the inevitable thermal or SNR noise on the original RF sampling process. Each of these bi-level bit streams is multiplied on a bit-by-bit basis by the time-delayed and Doppler-shifted base band replica of the actual signal broadcast from a specific GPS satellite. These products then are averaged over a given time interval, i.e., they are cross-correlated. The nominal averaging time for this cross-correlation is 20 ms. The receiver can perform this operation in parallel on the dual-band L1 and L2 signals from up to a dozen individual GPS satellites. The cross-correlation operation accomplishes three objectives. First, the signals from all other satellites are essentially filtered out. Second, when the time-delayed and Doppler-shifted base band replica is aligned with the arriving signal from that specific satellite, the effective SNR noise on the correlation products for that satellite is averaged down. Third, the alignment returning the maximum value of the correlation product provides the difference of the reception epoch of the receiver clock minus the transmission epoch of the clock onboard the GPS satellite. Synchronizing the transmitter and receiver clocks then allows the propagation delay, or group delay, from that satellite to the receiver to be obtained. Clock synchronization is achieved through redundant concurrent tracking of selected GPS satellites, including the occulted satellite, from the LEO receiver and from the global network of ground station receivers. Concurrent tracking of multiple GPS satellites from multiple receivers allows determination of the clock epoch differences. This cross-correlation between the received signal and its replica from a given satellite is performed on the in-phase bit stream and also separately on the quadrature bit stream. Thus, during 20 ms on each bit stream, the signal is averaged over 20,000 transition points or chip periods of the C/A code and 200,000 chip periods of the P code. The noise error in the average is inversely proportional to the square root of the averaging interval.

The propagation delay between the broadcasting GPS satellite and the receiver may be determined in two ways. The alignment of the bit streams, so that the autocorrelation function described above from the received and replica codes is maximized, yields the group delay plus SNR error. After the autocorrelation function is maximized, a far more precise determination of propagation delay is obtained from the measurement of the phase of the carrier

itself, which has a wavelength of only about 20 cm; in comparison, the "wavelength" of the P code is about 30 m. This propagation delay of the carrier phase is obtained from the two time-averaged correlation coefficients, the in-phase coefficient (I) and the quadrature coefficient (Q). The arctangent of I divided by Q gives a measurement of the difference in true phase minus the phase predicted from the in-receiver model, modulo 2π.[1] If no cycles are lost between the epochs of the current measurement and the immediate previous measurement 20 ms earlier, adding the predicted phase from the model to the measured phase difference gives the true phase plus the averaged-down SNR error at that observation epoch. An important property of this scheme (adding the predicted phase to the measured difference) is that the SNR measurement error of the true phase at each observation epoch is statistically independent of the SNR error at any other epoch. The root-sum-square of the in-phase and quadrature correlation coefficients gives the amplitude of the signal. The SNR error on the phase measurement is statistically uncorrelated with the SNR error on the amplitude measurement. These measurements of true phase and amplitude then are reported by the receiver at a nominal sample rate of every 20 ms; this is a convenient reporting rate because of the 50-Hz header code. Other multiples of that canonical sample interval can be used.

This closed-loop correlation and reporting scheme for the phase works well when the root-mean-square (rms) difference between the predicted and measured phase is small, substantially less than 1/4 cycle. In this case, there is a very high probability that no complete cycles, of either a positive or negative integer number, have been unaccounted for between the two successive measurement epochs. Using Gaussian statistics, it is easy to show why this is so. Suppose that the measured phase at a particular epoch differed from the predicted value by exactly 1/2 cycle. Then we would have no way of determining whether that measurement was a cycle above or a cycle below the predicted value; all cycles look alike. Regarding this measured 1/2 cycle difference, do we add it to or subtract it from the predicted phase given by the model? Suppose that the statistical difference between measured and predicted

---

[1] In early versions of the BlackJack receiver, a simpler 2-quadrant arctangent routine without complete 4-quadrant resolution was used on the I and Q correlation coefficients to extract phase. This design choice eliminated the requirement to determine during the signal processing the sign of the 50-Hz header code bit ± 1. An error in its sign affects the sign of the I and Q correlation products the same way, and cancels in their ratio. But, this results in a 1/2-cycle ambiguity. This exacerbates the cycle slipping problem in noisy and/or loop-stressed conditions. Unfortunately, this 2-quadrant arctangent operation has resided in the digital signal processing firmware—the application-specific integrated circuits (ASICs)—of the receiver, which is not easily modified without incurring the expense of redesigning and fabricating the ASIC at a silicon foundry. BlackJack designers at JPL plan to incorporate a 4-quadrant discriminator in a future version. The discussion here assumes the 4-quadrant version.

phase, arising from either SNR errors or from unknown phase acceleration or from both, turns out to be 1/4 cycle, $1-\sigma$. Then the probability of getting less than 1/2 cycle difference at each measurement epoch, if the errors are Gaussian distributed, is $\text{erf}[2/2^{1/2}]$, or 0.9545. It follows that the probability of having at least one difference that is greater than 1/2 cycle after n successive statistically independent samples is $1-0.9545^n$. After 1 s of elapsed time at a rate of 1 sample per 20 ms, $n=50$; therefore, for $\sigma=1/4$ cycle, the probability of getting at least one phase difference greater than 1/2 cycle in 50 trials is virtually certain. If at a given measurement epoch the measured phase difference is 1/2 cycle, then there is at least a 50 percent chance of choosing the wrong integer for the cumulative cycle count. Therefore, for $\sigma=1/4$ cycle, a non-zero integer number of cycle slips becomes virtually certain within 1 s with a 20 ms sample interval. On the other hand, suppose the $1-\sigma$ value of the statistical difference between the measured and predicted phase is 1/8 cycle. Then the probability of getting one or more phase differences that are greater than 1/2 cycle in n successive samples is $1-\text{erf}[4/2^{1/2}]^n=1-0.9968^n$, still nearly zero for $n=50$. In good signal conditions with small differences between measured and predicted phase, this scheme (of adding the predicted phase to the measured difference modulo 2π) usually works well.

6.2.1 Adverse Signal Conditions

A problem arises in adverse signal conditions, for example, when more than one ray arrives concurrently at the LEO, leading to interference and to sharp accelerations in phase across the troughs in the amplitude scintillation. Another adverse condition is where the LEO enters a quasi-shadow zone where no rays or at most highly defocused rays are present. Interference scenarios have been described earlier in Figs. 5-19 through 5-34. Figure 5-34 shows a specific example of fringe frequencies of roughly 50 Hz, 1 cycle change over 20 ms. Reducing the refractivity gradient in this model by a factor of four to obtain more realistic conditions, that is, ones that are more closely aligned with the Earth's atmosphere, quadruples the length of time. Thus, 1/4 cycle change occurs for this relaxed case in 20 ms, which still is a potential problem for closed-loop operations. This level of phase acceleration is serious enough to cause, with some probability on each correlation interval, at least one cycle not to be properly added or subtracted from the integer count book kept by the in-receiver phase model. Over the many successive 20-ms sample intervals spanning a few seconds, this could amount to a significant number of systematically lost cycles.

Consider the Gaussian refractivity model used in Fig. 5-20 and also the composite Gaussian/exponential model in Fig. 5-26. For a Gaussian distribution, the impact parameter separations between rays in the multipath

zone scale roughly linearly with the $1-\sigma$ width H_{W} of the distribution, but the width of the multipath zone (in θ_{L} or in elapsed time) scales roughly as $H_{\mathrm{W}}^{-3/2}$. The separations between rays scale only weakly with N_{W}, but the width of the zone scales nearly linearly with N_{W}. For the impact parameter diagram shown in Fig. 5-20, the values used are $N_{\mathrm{W}} = 0.0001$ and $H_{\mathrm{W}} = 1.6\,\mathrm{km}$, which results in impact parameter separations of roughly 10 km, rather large compared to realistic scenarios. But, in Fig. 5-26, the values used in the Gaussian component are $N_{\mathrm{W}} = 0.00005$ and $H_{\mathrm{W}} = 350\,\mathrm{m}$, corresponding to a peak water vapor number density of about 1 percent of the local dry air density. More importantly, for Fig. 5-26, this component is superimposed on a background refracting medium that is defocusing, an exponential profile corresponding to dry air. This composite model leads to impact parameter separations in the multipath zone of up to 3 km or to excess Doppler differences of up to about 15 Hz, or 1/3 cycle in 20 ms.

Another adverse signal situation occurs when the LEO encounters a shadow zone with weak signal conditions followed by flaring and strong interference. Consider a local refracting medium (e.g., a water vapor layer) embedded in an ambient medium (e.g., dry air) that gives rise to the transient in Fig. 6-1, which shows bending angle versus impact parameter. An abrupt increase in refractivity below a spherical boundary could yield this form for the bending-angle profile. Here $\rho_o = k(1+N_o)r_o$ corresponds to sea level. N_o is the refractivity for dry air at sea level. The exponential model in Eq. (5.8-2a) is used here for the dry air component of the refractivity with $N_o = 0.00027$ and the scale height $Hk^{-1} = 7\,\mathrm{km}$. Below the boundary at $r_o + Hk^{-1}/4$, about 2 km above sea level, the total refractivity abruptly increases, but then with decreasing altitude its gradient gradually approaches the dry air gradient. (Figure 6-7 later in this chapter shows the recovered refractivity profile for this water vapor layer.)

The impact parameter diagram shown in Fig. 6-2 results from the same bending-angle profile shown in Fig. 6-1. The values of the refractivity and orbit parameters are the same for Figs. 6-1 through 6-7. The LEO orbit radius is $r_{\mathrm{L}} = 1.1r_o$. Here θ_{ρ_o} marks the angular position of the LEO when located on the refracted shadow boundary, and it is defined by Bouguer's law, $\rho_o = \rho_{\mathrm{L}} \sin\left(\theta_{\rho_o} + \alpha_{\mathrm{L}}(\rho_o)\right)$. Here $\rho_{\mathrm{L}} = kr_{\mathrm{L}}$, and $\alpha_{\mathrm{L}}(\rho_o)$ is the refractive bending angle observed by the LEO at $\theta_{\mathrm{L}} = \theta_{\rho_o}$. This figure, expressed in terms of impact parameter altitude in kilometers versus orbit angle, provides an example of a shadow zone (where $d\rho_*/d\theta_{\mathrm{L}} \approx 0$) followed by caustic flaring and multipath. At $r \approx r_o + H/4k$, the main ray (m) encounters a sharp increase in refractivity as its tangency point descends below a boundary there. The abrupt increase in refractive bending at this boundary causes a shadow zone. As θ_{L}

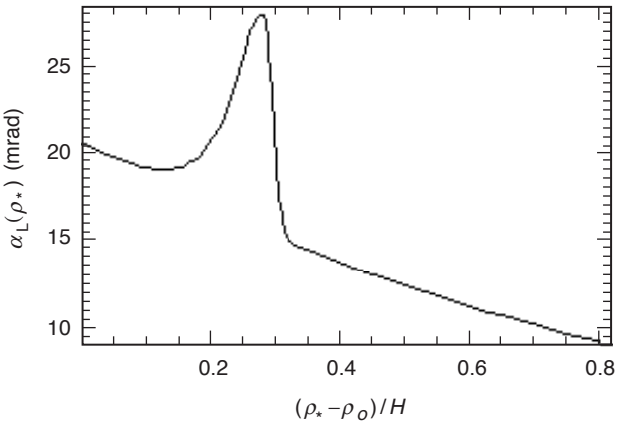

Fig. 6-1. Hypothetical bending angle at the LEO for a layer of higher refractivity embedded in dry air near sea level. Figure 6-7 shows the recovered refractivity profile for this water vapor layer.

further decreases, the first caustic is encountered at $\theta_L - \theta_{\rho_o} \approx 2$ mrad, leading to the creation of two more rays (a) and (b) in addition to the main ray (m). For a coplanar geometry, the Doppler difference between these new rays and the (m) ray is 6 to 7 Hz at the first contact point; it gradually increases as the rays separate. This difference is given by $\Delta f_{b-m} = (\rho_{*m} - \rho_{*b})\dot{\tilde{\theta}}_L / 2\pi$, where $\dot{\tilde{\theta}}_L$ is the LEO angular velocity in the plane of incidence. Interference continues until the (a) and (m) rays disappear below the lower caustic point at $\theta_L - \theta_{\rho_o} \approx -7$. This scenario has been discussed in more detail in Chapter 2 using a thin phase screen and scalar diffraction theory—in particular, in Figs. 2-2(c) and 2-11. In the case shown here in Fig. 6-2, the (a) and (b) rays created at the right-hand caustic point have phases at the LEO that differ by many cycles from the phase of the main ray (m). Moreover, they also have substantially lower excess Doppler frequencies. In this figure, the difference is 6 to 7 Hz initially for a coplanar geometry, and it grows to 15 Hz as the altitude of the (b) ray impact parameter separates from the other two. When the occulted GPS satellite does not lie in the LEO orbit plane, obliquity effects reduce this Doppler level by up to about 30 percent; see Eq. (6.3-4) and Fig. 6-8.

Figure 6-3 shows the amplitude of the field at the LEO that results from this transient in the bending angle shown in Fig. 6-1. The amplitude of the main field, which has a vacuum value of unity, has been defocused down to about 0.4 by the dry air refractivity profile. In this example, the upper caustic (right) yields strong signal flaring, but flaring from the lower caustic (left) is more muted because of the larger curvature in the impact parameter curve there (see Fig. 6-2). The voltage SNR well into the shadow zone is about 11 dB below the

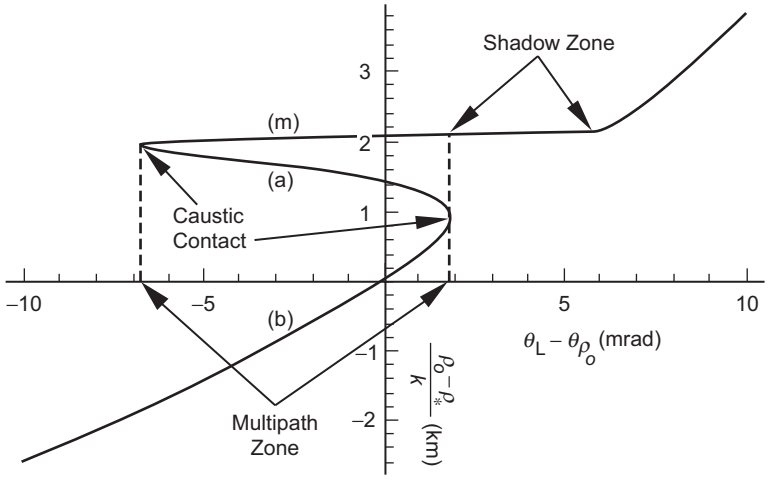

Fig. 6-2. Impact parameter diagram for the refractive bending-angle profile given in Fig. 6-1.

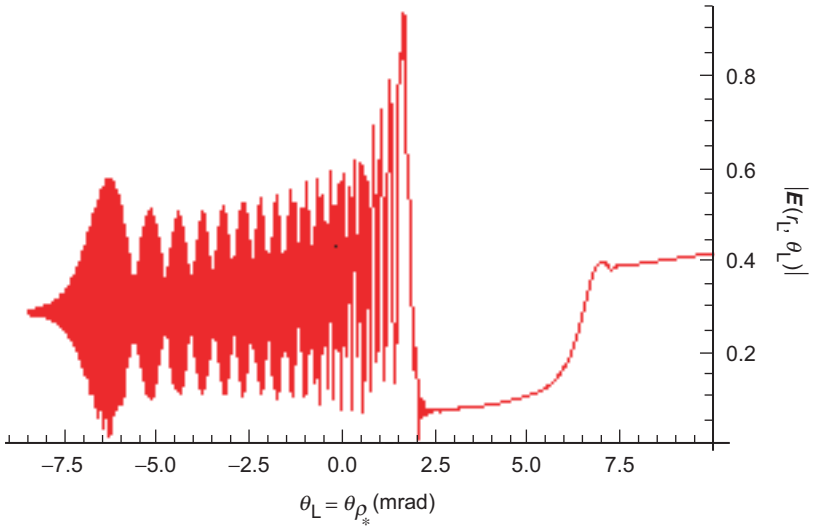

Fig. 6-3. Amplitude of the field at the LEO from the bending-angle profile given in Fig. 6-1.

voltage SNR for GPS signals in a vacuum. Diffraction creates edge fringes, but it also softens rough edges, resulting in the more gradual decay of SNR at the beginning of the shadow boundary.

Figure 6-4 shows the phase difference between the complete field and the field from the main ray (m) only. This figure begins near the end of the shadow zone and includes the contact with the first caustic. For a coplanar geometry, this figure covers about 3 to 4 s of elapsed time. The excess Doppler for the (m) ray becomes nearly constant at entry into the shadow zone. In fact, a ray with nearly constant excess Doppler is a very defocused ray [see Eq. (6.3-11)]. In the lower troposphere, defocusing from the dry air refractivity gradient compresses the wider altitude differences of these impact parameters. Nevertheless, this figure shows an abrupt change in Doppler, within a 20- to 30-msec interval, from 0 Hz to about 7 Hz for a coplanar geometry.

Figure 6-5 shows a blowup of the amplitude and phase of the field at the LEO in Figs. 6-3 and 6-4 in the vicinity of $\theta_L - \theta_{\rho_o} = 2.1$ around a very deep amplitude trough in the interference fringes. Here the emerging (a) and (b) rays, still essentially coherent and at a point slightly earlier than the geometric optics prediction of the caustic contact point, have strengthened so that their combined amplitude at the point where their phase is opposite the phase of the (m) ray nearly matches its amplitude. This causes a near-complete cancellation of the total field and a short burst of rapid phase acceleration. The black dots in the figure denote measurement epochs on 20-ms centers (for $d\tilde{\theta}_L / dt = -1$ mrad/s). The gray-scale dots are predicted phase values at future epochs. The SNR error bars are relative; their actual values also depend on signal strength and processing technique. In these situations,

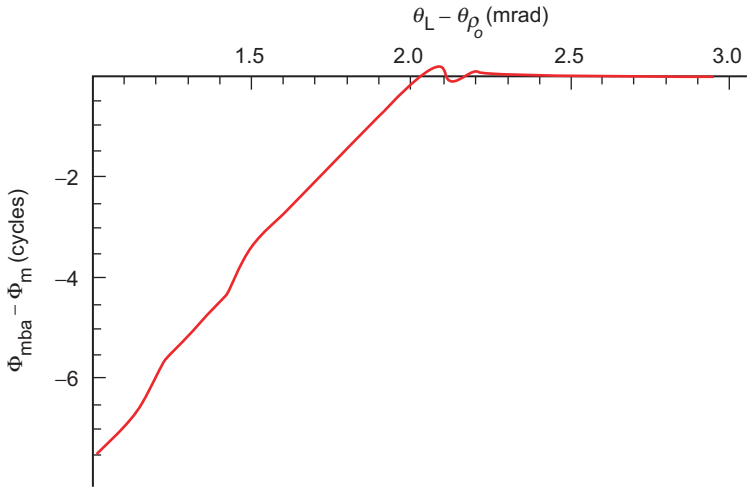

Fig. 6-4. Phase of the complete field at the LEO minus the phase of the main ray (m) resulting from the Fig. 6-1 bending-angle profile.

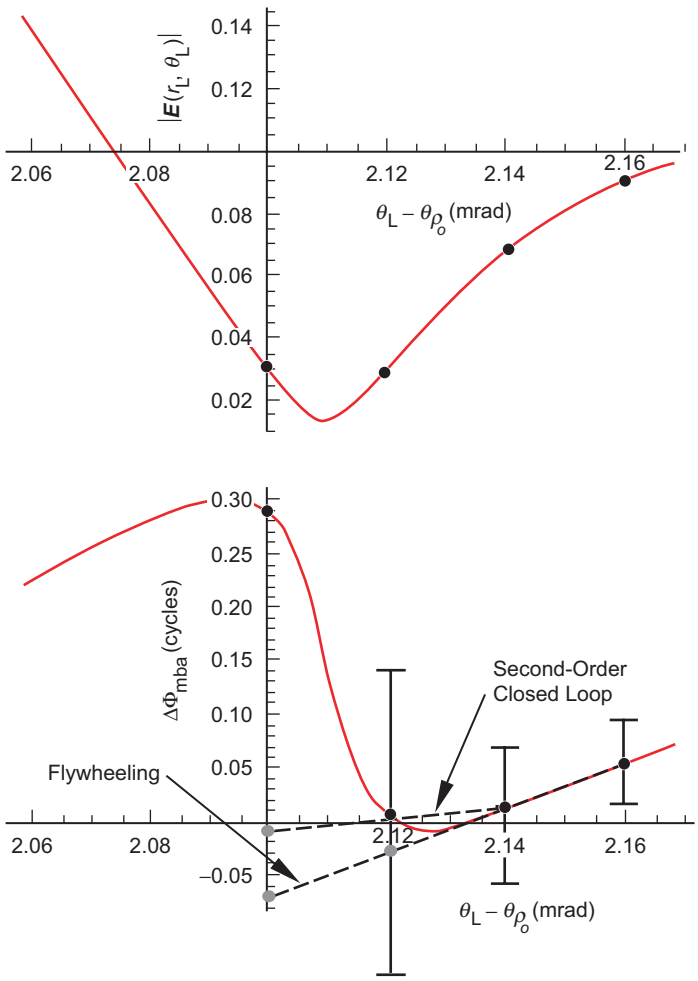

Fig. 6-5. Blowup of Figs. 6-3 and 6-4 at $\theta_L - \theta_{\rho_o} \approx 2.1$ showing the burst of phase acceleration at $\theta_L - \theta_{\rho_o} = 2.12$ resulting from near-complete cancellation of the field caused by interference between the main ray and the opposing caustic rays.

the rms disparity between predicted and measured phase can exceed 1/4 cycle primarily because of the inability of the in-receiver phase model to anticipate adequately the phase acceleration from interference between these multiple rays, and because of poor SNR noise in the phase measurement. Reducing the sample time to mitigate phase acceleration effects and to catch stray cycles adversely impacts the effective SNR of the sampled measurement, which further exacerbates the potential disparity.

Multipath situations like these have led the BlackJack designers to experiment with alternate sampling and reporting schemes, such as flywheeling the receiver. An additional technique is to use a realistic excess Doppler algorithm combined with the receiver fractional phase measurement to connect phase. Other open-loop schemes are feasible; the receiver uses a realistic in-receiver model and simply reports multiple time-lagged correlation coefficients in an open-loop mode. This approach leaves it to the data analysts to recover the various phases from the multiple rays relative to the realistic model and also to extract their respective amplitudes.

6.2.2 Flywheeling

Flywheeling uses an extrapolation from the in-receiver phase model, set at an earlier time where the phase was considered to be unambiguous, to predict the phase at a future epoch, without updating the model with new information from the most recent epoch because of its uncertainty. Both the closed-loop and the flywheeling modes are depicted in Fig. 6-5. To simplify the figure, only a second-order loop is shown, which is indicated by the sloped straight lines. But the receiver actually uses a third-order technique to predict forward; the extrapolation lines in this figure should be curved. In this simpler form of the in-receiver model, the phase measurements from the previous two measurement epochs are used to set the Doppler and the phase. In Fig. 6-5, these two points are located at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.14$ and $2.16\,\mathrm{mrad}$. In this figure, time evolves to the left for a setting occultation at a rate of roughly –1 s/mrad for a coplanar occultation. Thus, the abscissa can be read directly in seconds of time. For a non-coplanar case, the time scale would be compressed by the obliquity; for a 30-deg LEO orbit-plane inclination relative to the plane of incidence, the elapsed time interval in this figure would be 30 percent greater (see Fig. 6-8). In the closed-loop mode, the receiver uses the measured phase at these two earlier epochs at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.14$ and 2.16 to predict the phase at the next epoch at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.12$. Upon obtaining the new phase measurement at this later epoch, it then updates the Doppler and phase of the model to predict the phase at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.10$. In the flywheeling mode, the receiver does not update the model at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.12$, and it extrapolates the model fixed at the earlier epochs to predict the phase at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.10$.

Although this example of near-complete extinction of the field may seem pathological, it happens. The sharp phase acceleration combined with increased SNR error in the measurement is problematical for the receiver. To show this, redraw the phase prediction lines in Fig. 6-5 to run from the opposite ends of the $1-\sigma$ phase measurement error bars at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.14$ and 2.12 for the closed-loop mode and at 2.16 and 2.14 for the flywheeling mode. The worst-

case combination for this example runs from the upper end of the error bar for the earlier point to the lower end of the error bar for the later point. For these combinations, the new phase predictions at $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.10$ for both the closed-loop and the flywheeling modes are more than 1/2 cycle different from the actual phase at this point. Both the closed-loop and flywheeling modes have difficulty with this example.

The receiver automatically changes to the flywheeling mode according to certain pre-set signal conditions related to SNR and measured phase residuals, typically around $\mathrm{SNR_V} \approx 30$, and it can revert back to closed-loop tracking according to another set of conditions. But, as many of the figures in this monograph suggest, tracking in either the closed-loop or the flywheeling mode can result in cycle losses. If, for example, the earlier Doppler from the still strong and unique (m) ray in Fig. 6-2 is used in a flywheeling mode to carry the struggling receiver through the later poorer SNR conditions across a shadow zone, what happens at the contact point with the upper caustic? It depends on the strength of that caustic. A weak caustic [see Fig. 5-21(b)] reveals itself at the LEO as the envelope defining the amplitude of relatively high frequency interference fringes. There the main ray (m) is still dominant, and the frequency of the interference fringes depends on the difference in altitudes of the impact parameter of the (m) ray and the impact parameter of the caustic rays. The receiver may have difficulty tracking either one of these nascent (a) and (b) rays because of their continuing interference with the (m) ray. Figure 5-22 shows the phase acceleration spikes that result when the rays become comparable in amplitude.

On the other hand, in the examples given in Fig. 5-32 and Fig. 6-3, the extraordinary signal strength at the caustic contact is likely to induce the GPS receiver to lock onto the phase of the field there. For strong caustics, the nascent (a) and (b) rays, which are temporarily coherent in their early stage, become the principal contributors to the field at the LEO. This is the case in Fig. 6-3 at the caustic contact near $\theta_{\mathrm{L}} - \theta_{\rho_o} = 2.0$. Until the impact parameters of these two nascent rays have had time to separate after the caustic contact point, there is very little interference between them. The amplitude of the field can be very strong at the LEO, depending on the curvature of the impact parameter curve θ_{L} vs ρ_* at the caustic contact point [see Eq. (5.12-11)]. The width of a caustic peak for strong nascent rays can be hundreds of milliseconds, roughly given by $\Delta t \sim 3 \left| d^2\theta_{\mathrm{L}} / d\rho_*^2 \right|_{\mathscr{C}}^{1/3} / \dot{\theta}_{\mathrm{L}}$. This width can be several to many 20-ms correlation intervals. For examples, see Fig. 5-32 ($\Delta t \approx 150$ ms) and Fig. 6-3 ($\Delta t \approx 500$ ms). These strong and locally stable caustic fields can easily lure the GPS receiver to lock on to them. In this case, the receiver reverts from

the flywheeling mode[2] to closed-loop mode and takes on the Doppler of the new field formed by the composite of all the rays. In the case of Fig. 6-4, the new Doppler abruptly (<50 ms) becomes 7 Hz less than the original Doppler for the (m) ray. The (m) ray has not yet disappeared, but it is no longer "recognized." One has a 7-Hz negative bias relative to the (m) ray, resulting from the receiver jumping from one ray to the other, from (m) to (b). However, it should be noted in this example, especially if the receiver were in fact to remain in the closed-loop mode through this delicate transition shown in Fig. 6-5, that no cycles were lost. The reported phase by the receiver in this case would be the measured phase of the complete field at the LEO from all rays. The problem is with the interpretation of the measurements resulting from the implicit adoption of a single-ray paradigm.

Probably a prudent rule of thumb declares that, whenever a caustic-like feature is encountered in the amplitude data, this probably means that it is a strong caustic in order to stand out in noisy data. This feature then is a harbinger for subsequent multipath and for the possibility that certain rays may not be properly accounted for with the usual ray theory approach, i.e., the Abel transform. In fact, the mere event of the receiver converting to the flywheeling mode should serve as an alarm announcing multipath and that spectral techniques may be required.

6.2.3 Refractivity Error from a Single-Ray Paradigm

Figure 6-6 shows the receiver in the flywheeling mode using the excess Doppler from the (m) ray just prior to entering the shadow zone ($\theta_{\mathrm{L}} - \theta_{\rho_o} \approx 7$) to power through the shadow zone with poor SNR. Closed-loop operations resume with the return of strong signal near the first caustic contact point ($\theta_{\mathrm{L}} - \theta_{\rho_o} \approx 2$) a few seconds later. The continued existence of the (m) and (a) rays after the (b) ray is "tracked," and their contributions to refractivity recovery, are essentially ignored in the multipath zone with the single-ray paradigm. It is straightforward to calculate the error in recovered refractivity caused by

[2] For the specific refractivity profile used in Figs. 6-1 through 6-5, it is doubtful that a high-performance receiver like the BlackJack actually would drop out of the closed-loop mode upon entering the shadow zone. Diffraction softens the rate of falloff in amplitude at the shadow boundary. Also, the average minimum amplitude in this example is still 20 percent of the original amplitude. At $\theta_{\mathrm{L}} - \theta_{\rho_o} \approx 2.1$, there is a single episode of almost complete cancellation between the (m) ray and the combined field from the emerging (a) and (b) rays. This event could force the receiver into the flywheeling mode, but it is somewhat irrelevant. In both modes, flywheeling or closed loop, the resulting error comes more from the analysis of the tracking data than from the receiver. Following a single-ray paradigm, the error in the refractivity recovery would be essentially the same regardless of mode.

ignoring this multipath. For example, one can adopt as the true refractivity model the same model used to generate Figs. 6-1 through 6-5. This includes an exponential model for dry air plus a localized Heaviside-like component for the water vapor layer. This causes the multipath zone shown in Fig. 6-2 with the three rays, (m), (a), and (b). The adopted model for recovery would be the exponential-only model, which has no additional Heaviside-like refractivity component and which allows no multipath. If we then extend the impact parameter curve from the (b) ray near the caustic up to the (m) ray, as shown in Fig. 6-6, we have essentially ignored the entire anomalous (a) ray and a segment of the (m) ray. These sections are shown as the dashed loop in this figure. A best fit of N_o and H from the exponential-only model to this region can be done. Comparison of the recovered refractivity profile from this fitted model with the true refractivity profile gives the error resulting from overlooking the dashed segments of the (a) and (m) rays. This is shown in Fig. 6-7. Here the true refractivity profile $N(\rho_*)$ generates the bending-angle profile shown in Fig. 6-1, but the modeled refractivity profile $\hat{N}(\rho_*)$ is only an exponential fit to the observations. This figure shows the difference in refractivity profiles between the fitted model and the true model. In this case, the error in the refractivity is negatively biased because the excess Doppler from the (b) ray, including its extrapolation backward, is on average less than the combination of the (a) and (m) rays. This systematic loss of counted cycles relative to the (m) and (a) rays is equivalent to a negative error in the Doppler, and this effectively leads in this example to a negative error in the recovered refractivity profile. See Appendix F for a further discussion of the error in recovered refractivity expressed in terms of an error in excess Doppler or bending angle.

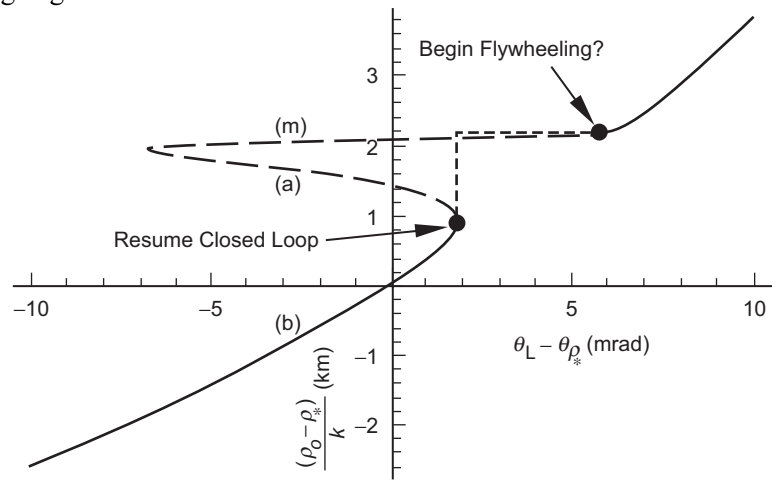

Fig. 6-6. Multipath with a single ray paradigm.

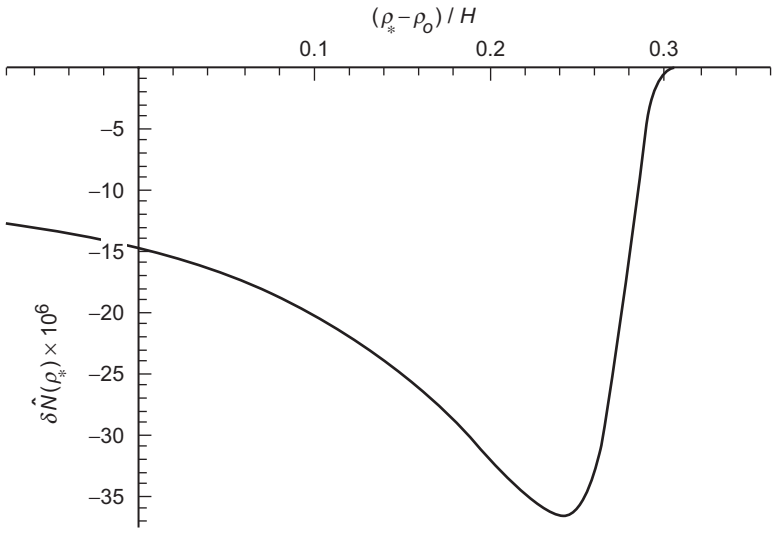

Fig. 6-7. Error in recovered refractivity $\hat{N}(\rho_*)$ as a result of ignoring the dashed sections of the (m) and (a) rays in Fig. 6-6.

An examination of actual SNR profiles from LEO occultation observations reveals a seemingly endless number of scintillation episodes, particularly as the signal passes through the water-vapor-laden lower troposphere or through various layers in the ionosphere. Flywheeling does not appear to provide a comprehensive remedy when deep scintillation occurs. The Blackjack tracking techniques probably will be augmented by JPL soon in favor of some open-loop scheme, such as reporting correlation coefficients using a realistic Doppler model, and with multiple time lags between the predicted and received time series.

6.3 Spectral Representation of the Field at the LEO

With these caveats, we now assume that an appropriate signal processing scheme has been implemented so that the amplitude and connected phase have been recovered from the receiver, plus the unavoidable SNR noise. We start from the spectral representation for an outgoing wave evaluated at the LEO located at $(r_{\mathrm{L}}, \theta_{\mathrm{L}})$. For the emitting GPS satellite located at $(r_{\mathrm{G}}, \theta_{\mathrm{G}})$ with $\theta_{\mathrm{G}} = \pi$, a fixed value, the spectral integral representation for the field is given from Eqs. (5.9-5) and (5.9-6), modified to account for the finite value of ρ_{G}. We have

$$\left.\begin{aligned} E(r_{\mathrm{L}},\theta_{\mathrm{L}}) &= \frac{E_o}{\sqrt{2\pi\rho_{\mathrm{L}}\sin\theta_{\mathrm{L}}}}\int_0^{\infty}\left(\frac{\sin\theta_\nu^{\mathrm{L}}}{\cos\theta_\nu^{\mathrm{G}}\cos\theta_\nu^{\mathrm{L}}}\right)^{1/2}\exp\big(i\Psi(+,-)\big)d\nu, \\ \Psi(+,-) &= D_\nu^{\mathrm{G}} + D_\nu^{\mathrm{L}} + \nu\big(\theta_\nu^{\mathrm{G}} + \theta_\nu^{\mathrm{L}} - \theta_{\mathrm{L}}\big) - 2G^{\dagger}(\nu) - \frac{\pi}{4}, \\ \theta_\nu^{\mathrm{G}} &= \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{G}}}\right),\ \ \theta_\nu^{\mathrm{L}} = \sin^{-1}\left(\frac{\nu}{\rho_{\mathrm{L}}}\right),\ \ \nu < \rho_{\mathrm{L}} < \rho_{\mathrm{G}}, \\ D_\nu^{\mathrm{G}} &= \sqrt{\rho_{\mathrm{G}}^2 - \nu^2},\ \ D_\nu^{\mathrm{L}} = \sqrt{\rho_{\mathrm{L}}^2 - \nu^2},\quad \rho = krn \end{aligned}\right\} \qquad (6.3\text{-}1)$$

where[3]

---

[3] The spectral representation for the field at the LEO when the GPS satellite is located at a finite distance, about 4.5 Earth radii, must account for the wave front curvature of the incident wave. This is a spherical wave of the form $e^{i\rho_{\mathrm{LG}}}/\rho_{\mathrm{LG}}$ centered at the emitting GPS satellite (see Fig. A-3). Although we did not derive this form given in Eq. (6.3-1), its extra terms compared with the form given in Eqs. (5.9-5) and (5.9-6) for the collimated case can easily be inferred from the difference in the asymptotic forms at large distances out of the atmosphere for the incoming spectral coefficients $a_l^-(\rho)$ for these two cases. These are given in Eq. (5.5-3a) for the case of a collimated or planar incident wave and in Eq. (5.5-3b) for the case of a spherical incident wave. For the latter, $a_l^-(\rho)$ carries the extra factor $i^{l+1}\xi_l^+(\rho_{\mathrm{G}})/\rho_{\mathrm{G}}$, which is derived from the multipole spectral expansion for a spherical wave [5] combined with the addition theorem for spherical harmonic functions. But, for $\rho_{\mathrm{G}} >> \nu$, we may use the asymptotic form for the spherical Hankel function $\xi_l^+(\rho_{\mathrm{G}})$. The factor $i^{l+1}\xi_l^+(\rho_{\mathrm{G}})$ has the asymptotic form

$$i^{l+1}\xi_l^+(\rho_{\mathrm{G}}) \to \left(\rho_{\mathrm{G}}^2/\left(\rho_{\mathrm{G}}^2 - \nu^2\right)\right)^{1/4}\exp\left[\,i\left(\left(\rho_{\mathrm{G}}^2 - \nu^2\right)^{1/2} + \nu\theta_\nu^{\mathrm{G}}\right)\right]$$

which coincides with the extra terms in Eq. (6.3-1). Here the phase delay spectral density function $\Psi(+,-)$ is referenced to the emitting GPS satellite, whereas $\Psi(+,-)$ in Eqs. (5.9-5) and (5.9-6) for the collimated case is referenced to the line $\theta = \pi/2$. Also, the constant E_o has a different meaning from E_o in Eq. (5.9-5). Essentially it must account for the $1/\rho_{\mathrm{LG}}$ space loss that the amplitude of the spherical wave emitted from the GPS satellite incurs in traveling to the LEO. This is inconsequential in recovering the refractivity profile because it is the variability of the amplitude and phase over an occultation episode that contains the atmospheric information. The product $(\cos\theta_\nu^{\mathrm{L}}\cos\theta_\nu^{\mathrm{G}})$ in Eq. (6.3.1) is related to the reduced limb distance, $D = D_{\mathrm{G}}D_{\mathrm{L}}/(D_{\mathrm{G}} + D_{\mathrm{L}})$, with $D_{\mathrm{G}} = \rho_{\mathrm{G}}\cos\theta_{\rho^*}^{\mathrm{G}}$ and $D_{\mathrm{L}} = \rho_{\mathrm{L}}\cos\theta_{\rho^*}^{\mathrm{L}}$.

The ray theory interpretation of θ_ν^{L} and θ_ν^{G} is as follows. Let $\nu = \nu^*$ be a spectral number at which Ψ assumes a stationary value. When super-refracting situations are avoided, we know that to high accuracy $\nu^* \doteq \rho_*$, the impact parameter of the

$$
\left.
\begin{aligned}
&2G^{\dagger}(\nu) \doteq 2\int_{\rho^{\dagger}}^{\infty} \frac{d\log n}{d\rho}\, g(\hat{y})d\rho \doteq -\int_{\nu}^{\infty} \alpha_{\mathrm{L}}(\omega)d\omega, \\
&g(\hat{y}) = \pi K_{\nu}^{2}\left(\mathrm{Ai}'[\hat{y}]^2 + \mathrm{Bi}'[\hat{y}]^2 - \hat{y}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)\right), \\
&\hat{y} \doteq \frac{1}{4K_{\nu}^{4}}\left(\nu^2 - \rho^2\right), \quad K_x = \left(\frac{x}{2}\right)^{1/3}, \\
&\nu = l + \frac{1}{2}, \quad \rho^{\dagger} = \nu - \hat{y}^{\dagger}K_{\rho^{\dagger}}, \qquad \hat{y}^{\dagger} = 0.44133\cdots
\end{aligned}
\right\} \tag{6.3-2}
$$

The spectral density function $G^{\dagger}(\nu) = G[\rho^{\dagger}(\nu), \nu]$ for the phase delay from the refractive gradient and $\rho^{\dagger}(\nu)$ are discussed more fully in Section 5.7, and their forms are given in Eqs. (5.7-2) and (5.7-11), respectively.[4] They are the spectral density forms applicable to an outgoing wave at the position of the LEO. We have taken minor license with Eq. (5.9-5) by consolidating the radial and transverse components into a single scalar form for the electric field $E(r_{\mathrm{L}}, \theta_{\mathrm{L}})$. The resulting error is negligible for occultations.

We note that the spectral density for the phase, $\Psi(+,-)$, is a function of $(\rho_{\mathrm{G}}, \rho_{\mathrm{L}}, \theta_{\mathrm{L}}, \nu)$. $\Psi(+,-)$ is the appropriate spectral density function for a position located well into the upper half-plane, $\pi >> \theta_{\mathrm{L}} >> 0$, and in the outgoing quadrant, $\pi/2 - \alpha_{\mathrm{L}} >> \theta_{\mathrm{L}} >> 0$. $\Psi(+,-)$ gives the spectral density of the complete phase delay at the LEO position $(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ relative to the emitting GPS satellite located at $(r_{\mathrm{G}}, \theta_{\mathrm{G}})$ with $\theta_{\mathrm{G}} = \pi$. It includes the geometric delay terms and the term $-2G^{\dagger}(\nu)$ for the delay from the refractive gradient (see Fig. 1-14). The geometric delay term, D_{ν}^{L}, gives the delay in phase along a straight line between the LEO and the tangency point of the line on a sphere of radius $\nu < \rho_{\mathrm{L}}$ centered at the origin. The term $\nu(\theta_{\mathrm{L}} - \theta_{\nu}^{\mathrm{L}})$ is an arc length along this sphere of radius ν, and it is subtracted from D_{ν}^{L} to correct it to the intersection of the sphere with the line $\theta = \pi/2$, which is the fixed reference line for computing phase delays at the LEO. Similarly, the terms $D_{\nu}^{\mathrm{G}} + \nu\theta_{\nu}^{\mathrm{G}}$ give the geometric phase delay from the emitting GPS satellite along the straight line

---

corresponding ray. Then from Eq. (6.3-1) and Bouguer's law, it follows that $\theta_{\rho_*}^{\mathrm{L}}$ becomes the angle between the ray path tangent vector and the radius vector of the LEO, $\chi_{\mathrm{L}} + \delta_{\mathrm{L}}$, in Fig. A-3. Similarly, $\theta_{\rho_*}^{\mathrm{G}} \to \chi_{\mathrm{G}} + \delta_{\mathrm{G}}$. Note that, from Fig. A-3 with $\theta_{\mathrm{G}} = \pi$, it follows that $\theta_{\rho_*}^{\mathrm{G}} + \theta_{\rho_*}^{\mathrm{L}} - \theta_{\mathrm{L}} \to \delta_{\mathrm{G}} + \delta_{\mathrm{L}} = \alpha_{\mathrm{L}}$.

[4] Incidentally, the difference between $\rho^{\dagger}(\nu)$ and ν is very small for a large sphere, $r_o/\lambda >> 1$. Here that spatial difference is about 7 m.

to the tangency point on the sphere of radius ν and thence along the sphere to the line $\theta = \pi/2$. For the LEO located at a given point (r_L, θ_L), the spectral neighborhoods around the stationary phase points ν^*, that is, where $\partial\Psi/\partial\nu|_{\nu^*} = 0$, provide the principal contributions to this spectral integral in Eq. (6.3-1).

$E(r_L, \theta_L)$ is the scalar field at the LEO; the phase of $E(r_L, \theta_L)$ includes an integer multiple of 2π driven by the absolute phase represented by the spectral density function $\Psi(+,-)$. But, as we just discussed, it is problematic whether or not the actual phase measurements can yield unambiguous connected phase at all times. Nevertheless, it is important to have the correct phase change between observation epochs, uncorrupted by systematic cycle slipping by the receiver or by the post-measurement data editing. Occasional cycle breaks, although undesirable, are probably inevitable. We assume here that cycle breaks have been fixed.

6.3.1 Stopped Field at the LEO

Let $\hat{E}(r_L, \theta_L)$ be defined as the "stopped" or "counter-rotated" signal received by the LEO. In this case, "stopped" means that the orbital Doppler tone between the LEO and the observed GPS satellite has been removed. Also, the excess Doppler based on a first-order model for the atmospheric refractivity signature has been removed. Thus,

$$\hat{E}(r_L, \theta_L) = E(r_L, \theta_L)\exp(-i\varphi_m(t)) \tag{6.3-3}$$

where the model phase $\varphi_m(t)$ is a known function that describes the time history of the stopping phase. Dealing with $\hat{E}(r_L, \theta_L)$ rather than $E(r_L, \theta_L)$ alleviates aliasing problems in finite sampling techniques and sharpens the resolution. To get an idea of the magnitude of the variability of $\varphi_m(t)$, we will form the time derivative of $\Psi(+,-)$, the spectral density for the phase delay at the LEO, given in Eq. (6.3-1). We evaluate Ψ at a stationary point in spectral number $\nu^* \doteq \rho_*$, where $\partial\Psi/\partial\nu = \theta_\nu - \theta_L - 2dG^\dagger/d\nu = 0$, and then we differentiate it with respect to time to obtain $\dot{\Psi}_*$. This already has been discussed in Section 5.12 for a circular LEO orbit with a coplanar geometry, that is, with the GPS satellite located in the orbit plane of the LEO.

6.3.2 The Obliquity Factor

However, we also should allow for the obliquity effect because in general coplanarity does not apply. From Fig. 6-8, we have two angles, θ_L and $\tilde{\theta}_L$ defining the angular position of the LEO. Here θ_L gives the angular position in

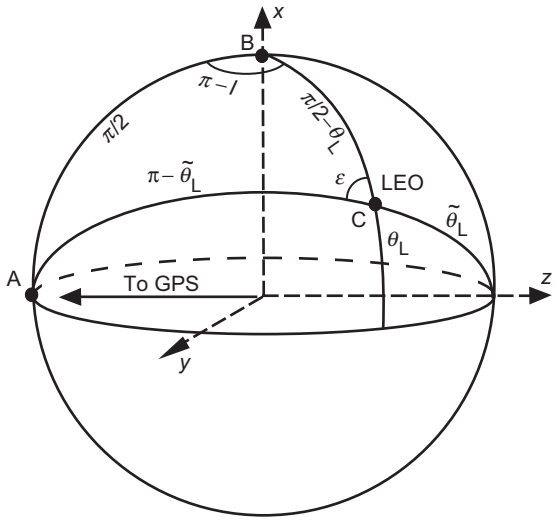

Fig. 6-8. A spherical geometry for non-coplanar LEO and GPS orbits. GPS satellite is located infinitely afar along the negative z-axis.

the LEO orbit plane, but $\tilde{\theta}_L$ gives it in the propagation plane, which is defined by the ray from the GPS satellite located in the negative z-direction[5] to the LEO. This propagation plane includes the "geocenter" (the center for the local geoid) and defines the great circle arc AC on the unit sphere in Fig. 6-8. The LEO orbit plane defines the great circle arc BC. The departure from coplanarity is given by the inclination angle I of the LEO orbit plane about the x-axis in the figure relative to the direction to the GPS satellite. The inclination angle I is satellite-position dependent, but it is readily expressed in terms of the orbit elements for the two satellites. The boresight-offset angle is ε. This is the azimuthal angle about the radial axis relative to the in-orbit plane direction at which the GPS satellite would be seen from the LEO. The spherical triangle ABC in Fig. 6-8 gives the relationship between θ_L and $\tilde{\theta}_L$. We have

[5] We assume here for calculation of the obliquity factor that the emitting GPS satellite is located infinitely afar in the negative z-direction, $\theta_G = \pi$.

$$\left.\begin{aligned} &\sin\varepsilon = \frac{\sin I}{\sin\tilde{\theta}_{\mathrm{L}}},\ \ \tan\tilde{\theta}_{\mathrm{L}} = \tan\theta_{\mathrm{L}}\cos\varepsilon, \\ &\cos\tilde{\theta}_{\mathrm{L}} = \cos\theta_{\mathrm{L}}\cos I, \\ &\frac{d\tilde{\theta}_{\mathrm{L}}}{d\theta_{\mathrm{L}}} = \cos^2 I\cos\varepsilon = \frac{\cos^2 I}{\sin\tilde{\theta}_{\mathrm{L}}}\sqrt{\sin^2\tilde{\theta}_{\mathrm{L}} - \sin^2 I} \end{aligned}\right\} \tag{6.3-4}$$

This provides the relationships between θ_{L} and $\tilde{\theta}_{\mathrm{L}}$, as well as the obliquity factor $d\tilde{\theta}_{\mathrm{L}} / d\theta_{\mathrm{L}}$ to reduce the Doppler for the effect of non-coplanarity. It follows that

$$\dot{\tilde{\theta}}_{\mathrm{L}} = \left(\frac{d\tilde{\theta}_{\mathrm{L}}}{d\theta_{\mathrm{L}}}\right)\dot{\theta}_{\mathrm{L}} \tag{6.3-5}$$

This obliquity factor is essentially constant over an occultation episode; thus, the relationship between $\tilde{\theta}_{\mathrm{L}}$ and θ_{L} is essentially linear. The obliquity factor is shown in Fig. 6-9 as a function of *I*. Here $\tilde{\theta}_{\mathrm{L}o} = \sin^{-1}\left(r_o / r_{\mathrm{L}}\right)$ with $r_{\mathrm{L}} / r_o = 1.1$, which gives about the correct value of $\tilde{\theta}_{\mathrm{L}}$ during an occultation. Thus, an inclination of 30 deg reduces the excess Doppler to about two-thirds of the coplanar value. It is $\tilde{\theta}_{\mathrm{L}}$ that should be used in Bouguer's law and in the spectral density functions involving phase. Accordingly, we adopt the following convention in the subsequent discussion. In any expression that involves a sensitive variable, such as phase or angular velocity, we shall replace θ_{L} with $\tilde{\theta}_{\mathrm{L}}$; otherwise, we leave the notation as is.

6.3.3 Doppler Variability

We now estimate the variability of $\varphi_m(t)$. For the purpose of calculating the Doppler variability, we assume circular orbits for the satellites. From Eq. (5.12-4), it follows after this replacement of θ_{L} with $\tilde{\theta}_{\mathrm{L}}$ in Bouguer's law that

$$\dot{\Psi}_* \doteq -\rho_*\dot{\tilde{\theta}}_{\mathrm{L}} \tag{6.3-6}$$

where the phase $\Psi_* = \Psi\left(\rho_{\mathrm{G}}, \rho_{\mathrm{L}}, \tilde{\theta}_{\mathrm{L}}, \rho_*\right)$, and ρ_* is the impact parameter corresponding to a specific ray, not necessarily unique. To simplify this calculation, we assume that the emitting GPS satellite is at an infinite distance. Its actual finite distance and orbital motion have a minor effect on our estimate

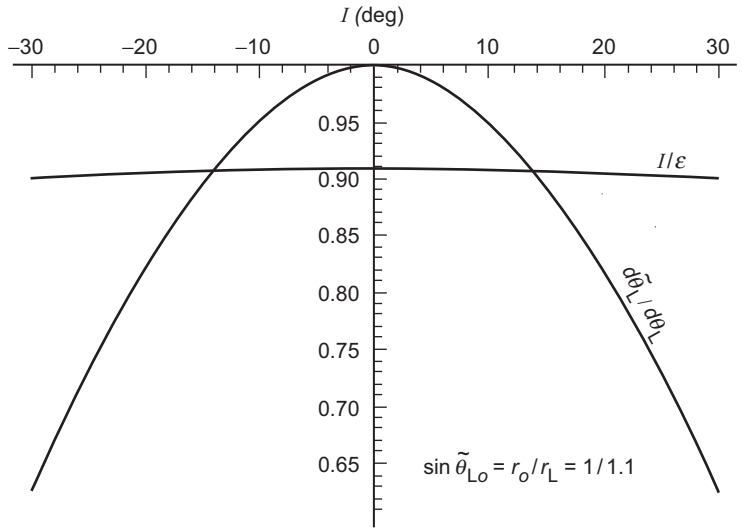

Fig. 6-9. Obliquity factor $d\tilde{\theta}_{\mathrm{L}}/d\theta_{\mathrm{L}}$ and I/ε versus LEO orbit plane inclination angle I.

here. From Bouguer's law in Eq. (5.6-5) for the GPS satellite at an infinite distance, we have

$$\rho_* = \rho_{\mathrm{L}} \sin\left(\tilde{\theta}_{\mathrm{L}} + \alpha_{\mathrm{L}}\right) = \rho_o + D_{\mathrm{L}}\alpha_{\mathrm{L}} + \mathrm{O}\left[\alpha_{\mathrm{L}}^2\right] \tag{6.3-7}$$

where $\rho_o = \rho_{\mathrm{L}} \sin\tilde{\theta}_{\mathrm{L}}$ and $D_{\mathrm{L}} = \rho_{\mathrm{L}} \cos\left(\tilde{\theta}_{\mathrm{L}} + \alpha_{\mathrm{L}}\right)$. The quantity D_{L} is close to the distance (or reduced distance) in phase units (i.e., distance $\times 2\pi/\lambda$) from the LEO to the Earth's limb, even with a modest cosine effect from orbit inclination angle I factored in. For the 20-cm wavelength of the GPS signal, $\rho_o \approx 2\times 10^8$ at sea level, and for a LEO orbit radius $r_{\mathrm{L}} = 1.1 r_o$, $D_{\mathrm{L}} \approx 1\times 10^8$. It follows that

$$\dot{\Psi}_* \doteq -\rho_o \dot{\tilde{\theta}}_{\mathrm{L}} - \left(D_{\mathrm{L}} \dot{\tilde{\theta}}_{\mathrm{L}}\right)\alpha_{\mathrm{L}} \tag{6.3-8}$$

The first term on the right-hand side (RHS) of Eq. (6.3-8) produces the orbital Doppler term due to the LEO. For a typical LEO orbit, $\dot{\theta}_{\mathrm{L}} \approx 1$ mrad/s, and it is essentially constant with time; thus, for the LEO part, $\rho_o \dot{\tilde{\theta}}_{\mathrm{L}} / 2\pi \approx 30$ kHz times the obliquity factor $d\tilde{\theta}_{\mathrm{L}} / d\theta_{\mathrm{L}}$, which is shown in Fig. 6-9 as a function of the inclination angle I of the LEO orbit plane.

The orbital velocity of a GPS satellite is about half the LEO velocity because its orbit radius is about four times larger. During an occultation, the position of a GPS satellite is located about 4.5 Earth radii away from the LEO

on the far side of the Earth; so, only about a quarter of its velocity vector projects plus or minus in the direction of the LEO. Also, the GPS satellite orbit planes are inclined differently to the LEO orbit plane, and the limb of the Earth as seen from the LEO is offset downward from the LEO orbit velocity direction by roughly 25 deg. The upshot is that all of these factors combine to yield a maximum Doppler during an occultation from both LEO and GPS orbital kinematics of around 35 kHz.

The second term in Eq. (6.3-8) is the excess Doppler from the bending angle α_{L} caused by the atmospheric refractivity; here $D_{\mathrm{L}}\dot{\tilde{\theta}}_{\mathrm{L}} / 2\pi \approx 15\left(d\tilde{\theta}_{\mathrm{L}} / d\theta_{\mathrm{L}}\right)$ Hz/mrad. Dry air yields a bending angle for a ray path tangency point at sea level of about 20 mrad; so, this component reaches about 300 Hz. But, bending angles through water vapor in the lower troposphere can exceed twice this level. The water vapor contribution is largely unknown a priori, but it can be characterized statistically by latitude and season. Excess Doppler signatures typically range over several hundred hertz.

What about the variability of these Doppler frequencies? Let us assume that the LEO is in a circular orbit. Then from Eq. (6.3-6) it follows that

$$\ddot{\Psi}_{*} \doteq -\dot{\rho}_{*}\dot{\tilde{\theta}}_{\mathrm{L}} \tag{6.3-9}$$

The term $\rho_{*}\ddot{\tilde{\theta}}_{\mathrm{L}} << \dot{\rho}_{*}\dot{\tilde{\theta}}_{\mathrm{L}}$, and it is ignored here. The acceleration contribution from the GPS satellite will be small over the relatively short time intervals of interest here. Upon differentiating ρ_{*} from Bouguer's law in Eq. (6.3-7), it follows that

$$\dot{\rho}_{*} = D_{\mathrm{L}}\dot{\tilde{\theta}}_{\mathrm{L}}\zeta_{\mathrm{L}} \tag{6.3-10}$$

where $\zeta_{\mathrm{L}} = \left(1 - D_{\mathrm{L}} d\alpha_{\mathrm{L}} / d\rho_{*}\right)^{-1}$ is the defocusing factor. It follows that the acceleration in the phase term is given by

$$\ddot{\Psi}_{*} \doteq -D_{\mathrm{L}}\dot{\tilde{\theta}}_{\mathrm{L}}^{2}\zeta_{\mathrm{L}} \tag{6.3-11}$$

Similarly, it follows that the acceleration from the orbital motion of the LEO is given by

$$\frac{d^{2}D_{\mathrm{L}}}{dt^{2}} \doteq D_{\mathrm{L}}\dot{\tilde{\theta}}_{\mathrm{L}}^{2} \tag{6.3-12}$$

Subtracting Eq. (6.3-12) from Eq. (6.3-11) gives the acceleration in excess phase:

$$2\pi \dot{f}_D = \ddot{\Psi}_* - \frac{d^2 D_L}{dt^2} \doteq -D_L^2 \dot{\tilde{\theta}}_L^2 \frac{d\alpha_L}{d\rho_*} \zeta_L \underset{\alpha_L' \to -\infty}{\to} D_L \dot{\tilde{\theta}}_L^2 \qquad (6.3\text{-}13)$$

It follows for strong defocusing that the excess Doppler rate approaches a constant value with time, 10 to 15 Hz/s, depending on the obliquity of the orbit and propagation planes. (Therefore, the bending angle rate of a given ray in strong defocusing conditions will approach a constant, $d\alpha_L / dt \to\sim 1$ mrad/s, or, equivalently, $d\alpha_L / d\tilde{\theta}_L \to\sim 1$.)

Over the course of a few tens of seconds that we will be interested in recovering the refractivity profile under adverse signal conditions, how well can we fit the stopping phase with a simple linear polynomial in time, $\varphi_m(t) = \omega_m t$? Equation (6.3-13) suggests that in strong defocusing this linear form should be good. We set $\omega_m = \dot{\Psi}_*$. Then $\dot{\Psi} - \dot{\varphi}_m(t) \approx \ddot{\Psi}_* t$. From Eq. (6.3-11), it follows that $\ddot{\Psi}_* / 2\pi <\sim 2$ Hz/s in the lower troposphere where the defocusing factor from dry air has a value around 1/10. The defocusing in the lower troposphere causes the impact parameter separations between multipath rays to be proportionately compressed, thereby reducing their Doppler differences by a factor of about 10. Thus, for a 10-s single-sided sample interval, we can use a linear term in time to stop the Doppler in the signal with a frequency runoff of a few tens of hertz. Sampling the signal at a 50-Hz rate usually should satisfy the Nyquist criterion for this sample interval. At the altitude of the sporadic E-layer, the ambient value of the defocusing factor is essentially unity; we would need a narrower sample interval there.

The last question concerns the spread in Doppler tones from different multipath rays. We already have seen that the spread in altitudes of the tangency points typically is less than 5 km. Thus, the maximum spread in Doppler is less than $\left(D_L \dot{\theta}_L / 2\pi\right)\left(5k / D_L\right) \approx 25$ Hz; the bandwidth of most interference spectra is less than 10 Hz. A sample rate of 50 Hz should suffice.

Incidentally, Eq. (6.3-10), which gives the velocity of the impact parameter of a ray, shows the retardation caused by the defocusing factor ζ_L. In the lower troposphere, the dry air component of the defocusing systematically compresses the altitude separation between multipath rays and narrows the maximum bandwidth of the interference spectrum; see Eq. (5.12-18).

6.4 Refractivity Recovery

Let us apply a discrete Fourier transform to the stopped LEO observations given in Eq. (6.3-3) and to the spectral representation for the stopped wave given in Eq. (6.3-1). For the latter, we have

$$\left.\begin{aligned}\hat{E}[\omega] &= \sum_{k=-M/2}^{M/2} \hat{E}\left(r_{\mathrm{L}}, \theta_{\mathrm{L}}(t_k)\right) \exp\left(i\omega \frac{kT}{M}\right), \\ \tilde{\theta}_{\mathrm{L}}(t_k) &= \tilde{\theta}_{\mathrm{L}o} + \dot{\tilde{\theta}}_{\mathrm{L}}(t_k - t_o) = \tilde{\theta}_{\mathrm{L}o} + \dot{\tilde{\theta}}_{\mathrm{L}} \frac{kT}{M}, \ \theta_{\mathrm{G}} \equiv \pi\end{aligned}\right\} \tag{6.4-1}$$

Here $\tilde{\theta}_{\mathrm{L}o} = \tilde{\theta}_{\mathrm{L}}(t_o)$ is the LEO orbit angle at the center of the data interval measured in the plane of incidence; see Fig. 6-8. Equation (6.3-4) provides the relationship between $\tilde{\theta}_{\mathrm{L}}$ and θ_{L}. T is the total time span covered by the observations, and $M+1$ is the total number (odd) of samples; for example, $M = 50T$ for a 50-Hz sample rate. Using the spectral representation in Eq. (6.3-1) for $\tilde{E}(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ and $\varphi_m(t) = \omega_m(t - t_o)$ for the stopping phase, from Eq. (6.4-1) we have

$$\left.\begin{aligned}\hat{E}[\omega] &= \sum_{k=-M/2}^{M/2} \hat{E}\left(r_{\mathrm{L}}, \theta_{\mathrm{L}}(t_k)\right) e^{i\omega \frac{kT}{M}} \\ &= \sum_{k=-M/2}^{M/2} e^{i(\omega-\omega_m)\frac{Tk}{M}} \left(\frac{E_o}{\sqrt{2\pi\rho_{\mathrm{L}} \sin\tilde{\theta}_{\mathrm{L}k}}} \int_0^{\infty} \left(\frac{\sin\tilde{\theta}_{\nu}^{\mathrm{L}}}{\cos\tilde{\theta}_{\nu}^{\mathrm{G}} \cos\tilde{\theta}_{\nu}^{\mathrm{L}}} \right)^{1/2} e^{i\Psi_{\mathrm{L}}} d\nu \right) \\ &\doteq \frac{E_o e^{i\Psi_{\mathrm{L}o}}}{\sqrt{2\pi\rho_{\mathrm{L}} \sin\tilde{\theta}_{\mathrm{L}o}}} \int_0^{\infty} \left(\frac{\sin\tilde{\theta}_{\nu}^{\mathrm{L}}}{\cos\tilde{\theta}_{\nu}^{\mathrm{G}} \cos\tilde{\theta}_{\nu}^{\mathrm{L}}} \right)^{1/2} \left(\sum_{k=-M/2}^{M/2} e^{ik\left(\omega - \omega_m - \nu\dot{\tilde{\theta}}_{\mathrm{L}}\right)\frac{T}{M}} \right) d\nu, \\ \Psi_{\mathrm{L}o} &= D_{\nu}^{\mathrm{G}} + D_{\nu}^{\mathrm{L}} + \nu(\tilde{\theta}_{\nu}^{\mathrm{G}} + \tilde{\theta}_{\nu}^{\mathrm{L}} - \tilde{\theta}_{\mathrm{L}o}) - 2G^{\dagger}(\nu) - \frac{\pi}{4}\end{aligned}\right\} \tag{6.4-2}$$

Now we replace the discrete Fourier transform with the integral transform using the fact that

$$\sum_{k=-M/2}^{M/2} \exp(iuk) = \frac{\sin((M+1)u/2)}{\sin(u/2)} \underset{M\to\infty}{\to} 2\pi\delta(u), \ -\pi \le u \le \pi \tag{6.4-3}$$

Here $\delta(u)$ is the Dirac delta function. Setting $u = \pi$ defines the Nyquist limit, $|\omega| \le \pi M / T$. Upon replacing the discrete "delta function" in Eq. (6.4-2) with the Dirac delta function, it follows that the Fourier transform of $\tilde{E}(r_{\mathrm{L}}, \theta_{\mathrm{L}})$ is given by

$$\left.\begin{aligned}
&\hat{E}[\omega_\nu] = 2\pi E_o \sqrt{\frac{1}{2\pi\rho_{\mathrm{L}} \sin\tilde{\theta}_{\mathrm{L}o}}} \left(\frac{\sin\tilde{\theta}_\nu^{\mathrm{L}}}{\cos\tilde{\theta}_\nu^{\mathrm{G}} \cos\tilde{\theta}_\nu^{\mathrm{L}}} \right)^{1/2} \exp\left(i\Psi_{\mathrm{L}o}\right), \\
&\Psi_{\mathrm{L}o} = D_\nu^{\mathrm{G}} + D_\nu^{\mathrm{L}} + \nu\left(\tilde{\theta}_\nu^{\mathrm{G}} + \tilde{\theta}_\nu^{\mathrm{L}} - \tilde{\theta}_{\mathrm{L}o}\right) - 2G^\dagger(\nu) - \frac{\pi}{4}, \\
&\omega_\nu = \omega_m + \dot{\tilde{\theta}}_{\mathrm{L}} \nu, \;\; D_\nu^{\mathrm{G}} = \sqrt{\rho_{\mathrm{G}}^2 - \nu^2}, \;\; D_\nu^{\mathrm{L}} = \sqrt{\rho_{\mathrm{L}}^2 - \nu^2}
\end{aligned}\right\} \tag{6.4-4}$$

For a finite sample interval, the discrete Fourier transform yields the discrete delta function, which has the $\sin Nx / \sin x$ character shown in Eq. (6.4-3), and the actual time series would have additive noise. The granularity threshold from the discrete transform can be obtained from the first zero of the discrete delta function in Eq. (6.4-3), which occurs at $u = 2\pi/(M+1)$. This is equivalent to $\Delta\omega = 2\pi/T$, which is essentially the uncertainty inequality for a discrete process. In impact parameter space, this relationship maps into a granularity in altitude of $\lambda\Delta\rho_* / 2\pi = \lambda\dot{\tilde{\theta}}_{\mathrm{L}}^{-1}\Delta\omega / 2\pi = \lambda/(\dot{\tilde{\theta}}_{\mathrm{L}} T) \approx 200/T$ m, or about 20 m for $T = 10$ s. Instead of a Fourier transform, one could use any one of several other transform schemes, such as the Morlet wavelet transform, which treats the time series as a spectral composition of wave packets. The subsequent equations will differ, and their efficiencies in recovering the refractivity profile might vary, but recovery should still be feasible. For the purpose of outlining this particular spectral approach for recovering the refractivity profile, we avoid further discussion of these important computational and related stochastic issues.

Note in Eq. (6.4-4) that $\omega_m \sim 2\times10^5$ rad/s, or 30 to 35 kHz. On the other hand, $\omega_\nu / 2\pi$ varies over only a few tens of hertz within the time interval for which the Fourier transform is applied. We have used the slowly varying character of $\tilde{\theta}_\nu^{\mathrm{L}}$ and $\tilde{\theta}_\nu^{\mathrm{G}}$ to simplify the spectral expressions given in Eq. (6.4-4). Over the bandwidth spanned by ω_ν, $\tilde{\theta}_\nu^{\mathrm{L}} + \tilde{\theta}_\nu^{\mathrm{G}}$ changes by less than 0.1 percent. Thus, we may set $\tilde{\theta}_\nu^{\mathrm{L}} + \tilde{\theta}_\nu^{\mathrm{G}} = \tilde{\theta}_{\mathrm{L}o} + \alpha_{\mathrm{L}o} \doteq \tilde{\theta}_{\mathrm{L}o}$ in the slowly varying terms (but not in $\Psi_{\mathrm{L}o}$). The error here is roughly 1 percent or smaller, the ratio $\alpha_{\mathrm{L}}(t_k) / \tilde{\theta}_{\mathrm{L}}(t_k)$. The Fourier transform in Eq. (6.4-4) further simplifies to

$$\left.\begin{aligned}
&\hat{E}[\omega_\nu] = 2\pi E_o \sqrt{\frac{1}{2\pi D}} \exp\left(i\Psi_{\mathrm{L}o}\right), \\
&\Psi_{\mathrm{L}o} = D_\nu^{\mathrm{G}} + D_\nu^{\mathrm{L}} + \nu\left(\tilde{\theta}_\nu^{\mathrm{G}} + \tilde{\theta}_\nu^{\mathrm{L}} - \tilde{\theta}_{\mathrm{L}o}\right) - 2G^\dagger(\nu) - \frac{\pi}{4}, \\
&\omega_\nu = \omega_m + \dot{\tilde{\theta}}_{\mathrm{L}} \nu, \;\; D^{-1} = D_{\mathrm{L}}^{-1} + D_{\mathrm{G}}^{-1}
\end{aligned}\right\} \tag{6.4-5}$$

We now take the Fourier transform over the occultation sequences of stopped phase and amplitude measurements made by the LEO, given in Eq. (6.3-3). The temporal breadth of this sequence T would depend on one's goals for refractivity recovery. We equate these two Fourier transforms. The left-hand side (LHS) on the upper line of Eq. (6.4-5) becomes the Fourier transform of the sampled amplitude and stopped phase of the field measured by the LEO; the RHS is from wave theory.

We note the one-to-one correspondence between ω_ν or excess Doppler and the spectral number ν in wave theory, or the impact parameter in ray optics, which holds when spherical symmetry applies. Also, we note that the bending angle may not be a unique function of time (or $\tilde{\theta}_{\mathrm{L}}$), but it is a unique function of excess Doppler, and when spherical symmetry holds it is a unique function of impact parameter through Bouguer's law. The impact parameter diagrams in Figs. 5-20, 5-26, 5-28, 5-31, and 6-2 all show implicitly this uniqueness property of the bending angle versus impact parameter when spherical symmetry holds. See also Eqs. (6.3-6) and (6.3-8), which apply to a circular LEO orbit. We have converted through a Fourier transform the time series of phase and amplitude measurements of the field, in which the bending angle may not be unique, into a spectral series in which the bending angle is unique (when spherical symmetry applies). It follows that we should be able to unambiguously determine the bending-angle profile versus excess Doppler from the Fourier transform $\hat{E}[\omega]$, given, of course, the limitations imposed by measurement errors.

Returning to Eq. (6.4-5), the LHS is the Fourier transform $\hat{E}[\omega]$ from the observations. It is a known quantity. The RHS is from wave theory; it also contains quantities that are known a priori or from POD information, except for $G^\dagger(\nu) = G\left[\rho^\dagger(\nu), \nu\right]$. Forcing equality between these two Fourier transforms forces the phase of $\hat{E}[\omega]$ to equal $\Psi_{\mathrm{L}o}$. Therefore, we can determine values for $G^\dagger(\nu)$ from Eq. (6.4-5) over the Fourier bandwidth spanned by ω.

However, a more suitable platform for extracting values for $\log n$ is $dG^\dagger(\nu)/d\nu$. Recalling that $g\left(\hat{y}^\dagger\right) = 0$, it follows from Eq. (6.3-2) that

$$\frac{dG^\dagger}{d\nu} \doteq -\pi K_{\nu_\omega} \int_{\rho^\dagger}^{\infty} \frac{d\log n}{d\rho} \left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right) d\rho \tag{6.4-6}$$

Let us now differentiate the Fourier transforms in Eq. (6.4-5) with respect to ω. Noting that $d\omega_\nu / d\nu = \dot{\tilde{\theta}}_{\mathrm{L}}$, we obtain

$$\frac{d\log \hat{E}[\omega_\nu]}{d\omega_\nu} = i\left(\tilde{\theta}_\nu - \tilde{\theta}_{\mathrm{L}o} - 2\frac{dG^\dagger}{d\nu}\right)\dot{\tilde{\theta}}_{\mathrm{L}}^{-1} \tag{6.4-7}$$

However, from Eq. (6.4-1), it follows that

$$\frac{d}{d\omega}\left(\hat{E}[\omega]\right) = i\int_{-\infty}^{\infty} t\hat{E}\left(\rho_L, \theta_L(t)\right)\exp(i\omega t)dt \tag{6.4-8}$$

No explicit differentiation of the observed phase and amplitude with respect to time is required to obtain the derivative of $\hat{E}[\omega]$ with respect to ω. In Eq. (6.4-6), we note that $\partial(\partial G[\rho,\nu]/\partial\nu)/\partial\rho \neq 0$ at $\rho = \rho^{\dagger}(\nu)$ (whereas $\partial G[\rho,\nu]/\partial\rho = 0$). It follows that Eq. (6.4-6) provides a stable means for determining the profile for $d\log n/d\rho$ from the determination of $dG\left[\rho^{\dagger}(\nu),\nu\right]/d\nu$.

Let the Fourier observation function $F[\omega]$ be defined by

$$\dot{\tilde{\theta}}_{\mathrm{L}} F\left[\omega_\nu\right] = i\dot{\tilde{\theta}}_{\mathrm{L}}\frac{d\log\hat{E}\left[\omega_\nu\right]}{d\omega_\nu} + \left(\tilde{\theta}_{\nu_\omega} - \tilde{\theta}_{\mathrm{L}\,o}\right) \tag{6.4-9}$$

$F\left[\omega_\nu\right]$ is a determined spectral quantity from the measurements and the POD information. From Eqs. (6.4-6) through (6.4-11), it follows that

$$\left.\begin{aligned} &-2\pi K_\nu\int_{\rho^{\dagger}}^{\infty}\frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho = 2\frac{dG^{\dagger}(\nu)}{d\nu} = \dot{\tilde{\theta}}_{\mathrm{L}} F\left[\omega_\nu\right] \\ &\rho^{\dagger} = \nu - \hat{y}^{\dagger}K_\nu = \frac{\omega_\nu - \omega_m}{\dot{\tilde{\theta}}_{\mathrm{L}}} - \hat{y}^{\dagger}K_\nu \end{aligned}\right\} \tag{6.4-10}$$

Equation (6.4-10) is in effect a linear system, an integral equation from which the profile $d\log n/d\rho$ may be recovered from a spectral sequence of known values for $F[\omega]$ based on the spectral derivative of the Fourier transform of the observations. Recalling Eqs. (5.4-3) and (3.8-7), the negative argument asymptotic forms for the Airy functions are $\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \to \pi^{-1}(-\hat{y})^{-1/2}$ with $\hat{y} \doteq K_\nu^{-4}\left(\nu^2 - \rho^2\right)/4$. Equation (6.4-10) becomes

$$\begin{aligned} &-2\pi K_\nu\int_{\rho^{\dagger}}^{\infty}\frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho \to \\ &\qquad -2\nu\int_{\nu}^{\infty}\frac{d\log n}{d\rho}\frac{d\rho}{\sqrt{\rho^2 - \nu^2}} = \alpha_{\mathrm{L}}(\nu) \end{aligned} \tag{6.4-11}$$

It follows from Eqs. (6.4-10) and (6.4-11) that

$$\left.\begin{aligned} &\alpha_{\mathrm{L}}(\nu) \leftrightarrow F[\omega_\nu]\dot{\tilde{\theta}}_{\mathrm{L}} \\ &\omega_\nu = \omega_m + \dot{\tilde{\theta}}_{\mathrm{L}}\nu \end{aligned}\right\} \tag{6.4-12}$$

To the extent to which the asymptotic forms for the Airy functions are applicable, $F[\omega_\nu]$ is proportional to the bending angle for an impact parameter value of ν. It follows for a setting occultation ($\dot{\tilde{\theta}}_{\mathrm{L}} < 0$) that, if $\alpha_{\mathrm{L}}(\nu) \to 0,\ \nu \to \infty$, then $F[\omega_\nu] \to 0,\ \omega_\nu \to -\infty$.

More generally, setting $\nu = \rho_*$, corresponding to a value of $\omega_\nu = \omega_m - \rho_* \mid \dot{\tilde{\theta}}_{\mathrm{L}} \mid$ within the Fourier bandwidth, one can form

$$\left.\begin{aligned} &\int_{\nu^\dagger}^{\infty} \frac{\dot{\tilde{\theta}}_{\mathrm{L}} F[\omega_\nu]}{\sqrt{\nu^2 - \rho_*^2}} d\nu = \\ &-2\pi \int_{\nu^\dagger}^{\infty} \frac{K_\nu d\nu}{\sqrt{\nu^2 - \rho_*^2}} \left(\int_{\rho^\dagger}^{\infty} \frac{d\log n}{d\rho} \left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \right) d\rho \right) \\ &= -2\pi \int_{\rho_*}^{\infty} \frac{d\log n}{d\rho} \left(\int_{\rho_*}^{\rho} \frac{\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \right)}{\sqrt{\nu^2 - \rho_*^2}} K_\nu d\nu \right) d\rho \\ &= \pi \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho, \quad \nu^\dagger = \rho_* + \hat{y}^\dagger K_{\rho_*}, \quad \rho^\dagger = \nu - \hat{y}^\dagger K_\nu \end{aligned}\right\} \tag{6.4-13}$$

The second integral on the third line involving the Airy functions is completely deterministic. The fourth line is obtained from integrating by parts. The weighting function $W(\rho, \rho_*)$ is given by

$$\left.\begin{aligned} &W(\rho, \rho_*) = \\ &2K_\rho \frac{\mathrm{Ai}[0]^2 + \mathrm{Bi}[0]^2}{\sqrt{\rho^2 - \rho_*^2}} - 4\int_{\rho_*}^{\rho} \frac{\mathrm{Ai}[\hat{y}]Ai'[\hat{y}] + \mathrm{Bi}[\hat{y}]\mathrm{Bi}'[\hat{y}]}{\sqrt{\nu^2 - \rho_*^2}} K_\nu^2 d\nu, \\ &\int_{\rho_*}^{\rho} W(\rho, \rho_*) d\rho = 2\int_{\rho_*}^{\rho} \frac{\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2}{\sqrt{\nu^2 - \rho_*^2}} K_\nu d\nu \xrightarrow[\rho \to \infty]{} 1, \\ &\hat{y} \doteq \frac{1}{4K_\nu^4} \left(\nu^2 - \rho^2 \right) \end{aligned}\right\} \tag{6.4-14}$$

In the limit as $\lambda \to 0$, $W(\rho, \rho_*)$ resembles a Dirac delta function, having similar properties. $W(\rho, \rho_*)$ and its integral, which rapidly approaches unity

with increasing $\rho > \rho_*$, are shown in Fig. 6-10. The decay profile of $W(\rho, \rho_*)$ is determined by the span in $\hat{y}$ space over which the Airy functions make their transition to negative argument asymptotic forms, $\sim 2k^{-1}K_{\rho_*} = 2(\lambda^2 n_* r_* / 4\pi^2)^{1/3} \approx 30$ m. Therefore, $W(\rho, \rho_*)$ strongly weights the contribution from $\log n(\rho)$ in the convolution integral in Eq. (6.4-13) at $\rho = \rho_*$, and it attenuates rapidly for $\rho > \rho_*$ to nearly zero within 30 m. The half-area point is at 7 m. This spatial interval, $\sim 2(\lambda^2 n_* r_* / 4\pi^2)^{1/3}$, is where the differences between wave theory and ray theory mainly arise. If we approximate $W(\rho, \rho_*)$ by a Dirac delta function, we have

$$\frac{1}{\pi}\int_{\nu^\dagger(\rho_*)}^{\infty} \frac{\dot{\tilde{\theta}}_L F[\omega_\nu]}{\sqrt{\nu^2 - \rho_*^2}}\, d\nu = \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho \to \log n(\rho_*) \tag{6.4-15}$$

Equation (6.4-15) is the wave theory equivalent of the Abel integral transform. The Abel transform owes its existence to a remarkable property of the integral,

$$\int_a^b \frac{2x dx}{\sqrt{(x^2 - a^2)(b^2 - x^2)}} = \pi \tag{6.4-16}$$

for all real positive values of a and $b > a$. Unfortunately, there does not seem to be the crisp equivalent of the Abel transform in wave theory. From

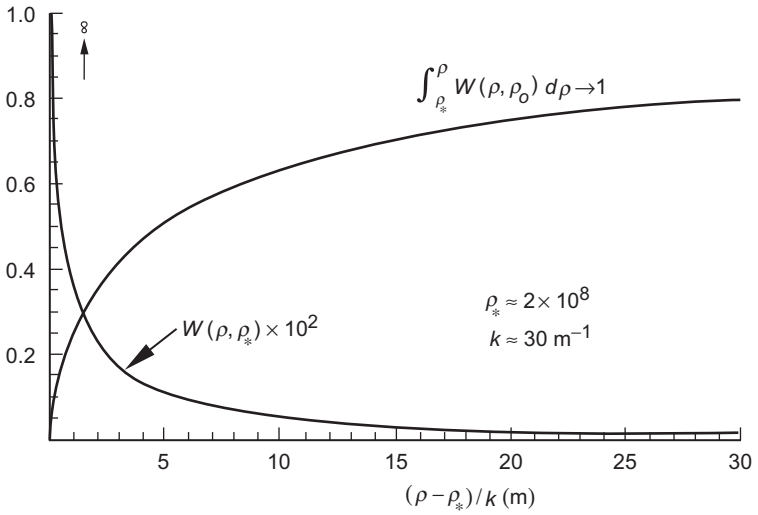

Fig. 6-10. Weighting function $W(\rho, \rho_*)$ in the wave theory analog of the Abel transform.

Eq. (6.4-13) and upon setting the Airy functions to their negative argument asymptotic forms, $\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \to \pi^{-1}(-\hat{y})^{-1/2}$, we have

$$2\pi\int_{\rho_*}^{\rho} \frac{\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)}{\sqrt{\nu^2 - \rho_*^2}} K_\nu d\nu \to \int_{\rho_*}^{\rho} \frac{2\nu d\nu}{\sqrt{\left(\nu^2 - \rho_*^2\right)\left(\rho^2 - \nu^2\right)}} = \pi \qquad (6.4\text{-}17)$$

By solving the convolution integral in Eq. (6.4-15), the refractivity profile is recovered from an integral operation on the spectral quantity $F[\omega_\nu]$, which is related directly to the spectral derivative of the Fourier transform of the stopped observations $\hat{E}(r_\mathrm{L}, \theta_\mathrm{L}(t))$ through Eqs. (6.3-3) and (6.4-9), and which is essentially proportional to the bending angle associated with an impact parameter value ν.

The time span T used in the Fourier treatment just described has not been specified. When the SNR permits, one can partition the entire data set into a time-ordered series of contiguous subsets or data packets of temporal width ΔT_k. Over each packet, a Fourier transform can be applied, and the spectrum for each of these strips can be assembled contiguously and displayed as a function of time, or as a function of nominal ray path tangency altitude, and so on. Figure 1-11, which is from [6], shows an example of this approach. This is one example of the so-called sliding spectrum technique [7].

6.4.1 Super-Refractivity

The integral for $2dG^\dagger(\nu)/d\nu$ in terms of the refractivity gradient in Eq. (6.4-6) requires special treatment to handle super-refractivity conditions. We have used $\rho = knr$ as the integration variable for convenience, but implicit in its use is the assumption that $d\rho/dr$ is positive throughout. Within a super-refracting medium, the ray curvature index, $\beta = -n'r/n$, is greater than unity. For a super-refracting spherical layer, the radius of curvature of the ray is less than the radius of curvature of the refracting surface. Super-refractivity occurs when $d\rho/dr = nk(1-\beta) \le 0$. The region where $d\rho/dr < 0$ should be evident from the data, and the critical value of dn/dr at the boundary is known, $dn/dr = -n/r \approx -157\times10^{-6}$ km^{-1}. Figure 6-11 shows a profile for a hypothetical refractivity gradient in the vicinity of a super-refracting layer. It could correspond to a narrow marine layer in the lower troposphere. Over the range $r_d \le r \le r_u$, the profile is super-refracting.

Figure 6-12 is a schematic showing the ray geometry for a super-refracting spherical layer with an upper boundary at $r = r_u$ and a lower boundary at $r = r_d$. Such layers are called ducts in ground-based radio transmission, and the word "ducting" is often used instead of "super-refracting." Two critical rays are shown in this figure. The upper critical ray just grazes the top of the super-

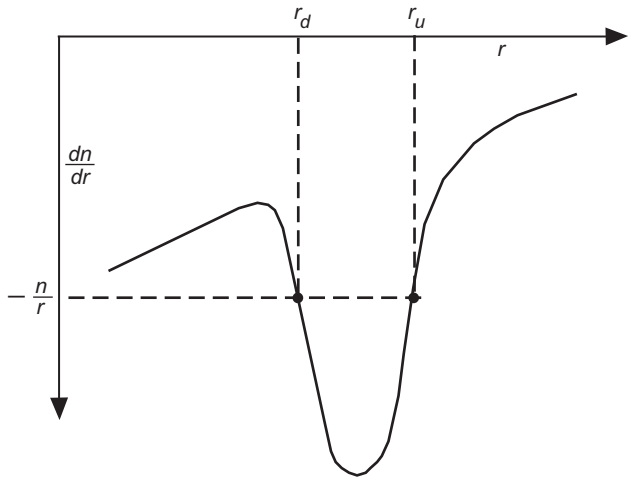

Fig. 6-11. Gradient of $n(r)$ near a super-refracting medium.

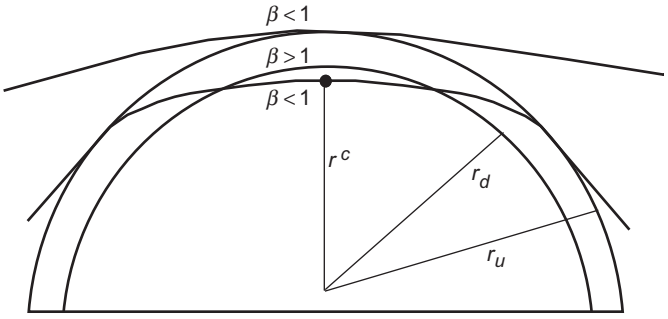

Fig. 6-12. Ray Geometry for a super-refracting spherical layer.

refracting layer; it has an impact parameter value of $n(r_u)r_u$. The lower critical ray has a tangency point at $r_* = r^c$ and an impact parameter value of $n(r^c)r^c$; it manages to escape from the top of the layer just before its ray path would have been turned inward by the strong refractive gradient in the layer. This lower critical ray escapes tangentially to the surface at $r = r_u$. It follows from Bouguer's law that the impact parameter for this critical ray is $n(r^c)r^c = n(r_u)r_u$. Thus, in geometric optics we have a discontinuity in the bending-angle profile versus impact parameter at this critical impact parameter value, $\rho_* = kn(r_u)r_u$. Two rays, one just grazing the top of the layer with $r_* = r_u$

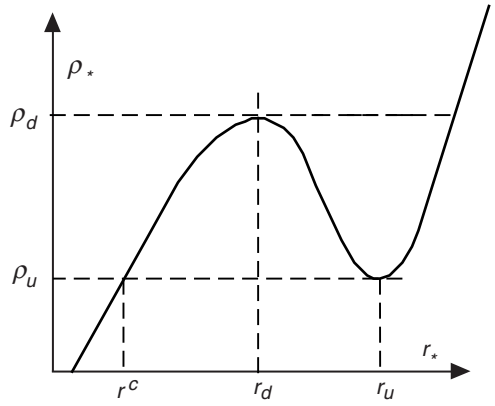

Fig. 6-13. Impact parameter curve in a super-refracting zone.

and the other at the lower critical tangency point $r_* = r^c$, both have the same impact parameter value but different bending angles. A ray with its turning point radius in the range $r^c < r_* < r_u$ cannot escape; for a ray to exist, the turning point radius r_* must either equal or exceed $r = r_u$, or it must be equal to or less than r^c.

Figure 6-13 shows an impact parameter curve $\rho_* = kn(r_*)r_*$ versus turning point radius r_* in the vicinity of a super-refracting layer. Note that $\rho_u < \rho_d$. A hypothetical ray with its tangency point in the range $r^c < r_* < r_u$ would have an impact parameter value in the range $\rho_u < \rho_* < \rho_d$. But from Fig. 6-13 we see that, in traveling along such a ray in the range $r_* \le r \le r_u$, one would eventually come to a point on the ray in the super-refracting layer past which $\rho(r) < \rho_*$, which is not allowed in geometric optics for a spherical geometry. See Appendix A, Eq. (A-4b). A necessary condition for the existence of a real ray between specified endpoints is that $\rho \ge \rho_*$ at *all* points along the ray. It is no good for $\rho \ge \rho_*$ to hold part of the way; it must hold all the way between end points, or else the term $(\rho^2 - \rho_*^2)^{1/2}$ in the bending angle and phase delay equations becomes imaginary at some point.

Figure 6-13 shows that the value of r^c relative to r_d depends on the difference $\rho_d - \rho_u$ and on the average slope of $d\rho/dr$ below the super-refracting layer. A first approximation is given by

$$r^c \approx r_d - (r_u - r_d)\frac{\overline{\beta}_{SR} - 1}{1 - \overline{\beta}^c} \tag{6.4-18}$$

where $\overline{\beta}^c < 1$ is the average value of β below the lower boundary of the super-refracting layer in the vicinity of $r = r^c$, and $\overline{\beta}_{SR} > 1$ is the average value of β in the super-refracting layer. If $\overline{\beta}^c$ is near unity, that is, nearly super-refracting, then the difference $r_d - r^c$ can be much larger than the thickness of the super-refracting layer itself.

For geometric optics, then, the refractivity profile in the range $r^c < r < r_u$ is terra incognito. Moreover, geometric optics already begins to fail[6] before reaching these critical tangency point limits at $r_* = r^c$ and $r_* = r_u$.

No such restriction applies in wave theory, but a super-refracting layer does complicate matters. The point where $d\rho / dr = 0$ marks the boundary of a super-refractive layer, which necessitates breaking the integral in Eq. (6.4-6) for $2dG^{\dagger}(\nu)/d\nu$ into three sections, one section above the upper boundary at $\rho_u = kn(r_u)r_u$, one below the lower boundary at $\rho_d = kn(r_d)r_d$, and one through the layer where $d\rho / dr < 0$.[7] Alternatively, one can write the integral in a less convenient form in terms of the radial coordinate itself. The end point of the integral becomes a function of $\rho^{\dagger}(\nu)/kn(r^{\dagger})$. This has three roots for $r^{\dagger}$ when ν lies within the range $\rho_u < \nu < \rho_d$: one above, $r > r_u$; one below, $r < r_d$; and one within the super-refractive layer, $r_d < r < r_u$.

Let us define the index of refraction profile according to regime:

$$n(r) = \begin{cases} n_1(r), & r \le r_d \\ n_2(r), & r_d \le r \le r_u \\ n_3(r), & r_u \le r \end{cases} \tag{6.4-19}$$

---

[6] For a geometric optics version of a super-refractive boundary caused by a 5 percent discontinuity in refractivity, see Figs. 2-2(b) and 2-7(a). The predicted amplitude is exactly zero in the shadow zone. Figure 2-11 shows the scalar diffraction version. Figures 3-24 and 3-25 show the Mie scattering version.

[7] Recall in Section 5.7, Eq. (5.7-27), that for a fixed spectral number ν we have set $\partial G[\rho,\nu]/\partial\rho = 0$ for $\rho \le \rho^{\dagger} = \nu - \hat{y}^{\dagger}K_{\rho}$. This is an approximation that exploits the near-equality of $2dG[\rho^{\dagger}(\nu),\nu]/d\nu$ and $\alpha_{\mathrm{L}}(\nu)$ in a medium with a moderate refractive gradient (see Appendix J). The error is small, but it depends on the curvature of the actual refractivity profile in the immediate vicinity of this turning point where the phase of the incoming wave is rapidly becoming stationary for decreasing $\rho < \nu$. See Fig. 5-7 for the exact phase profile $\partial\vartheta_l^{-}/\partial\rho$ in an Airy layer compared to $\partial G[\rho,\nu]/\partial\rho$. In a super-refractivity zone where $\rho_u < \nu < \rho_d$, we need the integral for $2dG^{\dagger}(\nu)/d\nu = 2dG[\rho^{\dagger}(\nu),\nu]/d\nu$ only for those sections where $\rho \ge \rho^{\dagger}(\nu)$. Hence the three sections.

Figure 6-14 is a sketch for $n(\rho)$ in the super-refracting zone. We have the continuity constraints $n_1(r_d) = n_2(r_d)$, $n_2(r_u) = n_3(r_u)$. If we now apply the wave theory version of the Abel transform in Eq. (6.4-15) to the Fourier observation function $F[\omega_\nu]$, we obtain

$$\frac{1}{\pi}\int_{\nu^\dagger(\rho_*)}^{\infty} \frac{\dot{\tilde{\theta}}_L F[\omega_\nu]}{\sqrt{\nu^2 - \rho_*^2}} d\nu = \int_{\rho_*}^{\infty} \log n(\rho) W(\rho, \rho_*) d\rho =$$

$$\int_{\rho_*}^{\infty} \log n_3 W d\rho, \quad \rho_* \geq \rho_d,$$

$$\int_{\rho_*}^{\rho_d} (\log n_1 - \log n_2 + \log n_3) W d\rho + \int_{\rho_d}^{\infty} \log n_3 W d\rho, \quad \rho_u \leq \rho_* \leq \rho_d,$$

$$\int_{\rho_*}^{\rho_u} \log n_1 W d\rho + \int_{\rho_u}^{\rho_d} (\log n_1 - \log n_2 + \log n_3) W d\rho + \int_{\rho_d}^{\infty} \log n_3 W d\rho, \rho_* \leq \rho_u$$

(6.4-20)

If we replace $W(\rho, \rho_*)$ by the Dirac delta function, Eq. (6.4-20) becomes

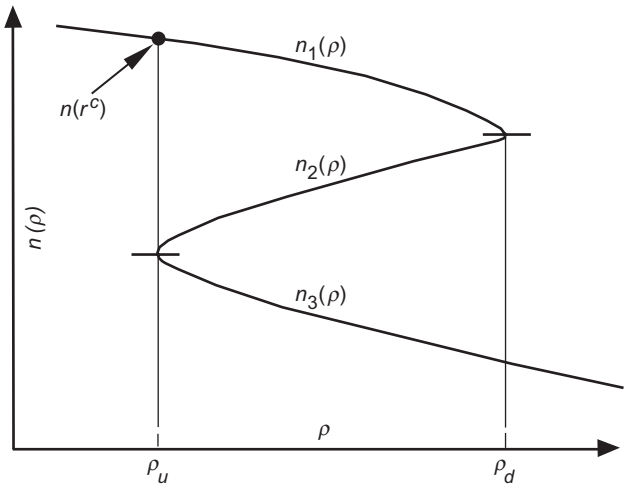

Fig. 6-14. Profile for n versus ρ in a super-refracting zone.

$$\frac{1}{\pi}\int_{\nu^{\dagger}(\rho_*)}^{\infty}\frac{\dot{\tilde{\theta}}_{\mathrm{L}}F[\omega_\nu]}{\sqrt{\nu^2-\rho_*^2}}d\nu \doteq$$

$$\left\{\begin{array}{l}\log n_3(\rho_*),\ \ \rho_* \geq \rho_d,\\ \log n_1(\rho_*)-\log n_2(\rho_*)+\log n_3(\rho_*),\ \ \rho_u \leq \rho_* \leq \rho_d,\\ \log n_1(\rho_*),\ \ \rho_* \leq \rho_u\end{array}\right\} \tag{6.4-21}$$

For spectral numbers lying in the region $\rho_u \leq \nu \leq \rho_d$, only the combination $N_1(\rho_*)-N_2(\rho_*)+N_3(\rho_*)$ is recoverable with this approach.

6.4.2 Improving the Accuracy of $G^{\dagger}(\nu)$

One can refine the approximation in Eq. (5.7-27) for $G^{\dagger}(\nu)$ by forcing $G[\rho,\nu]$ to align with the exact form for the phase delay in an Airy layer. Per the discussion concerning Eqs. (5.7-24) and (5.7-25), $\vartheta_l^-(\rho)$ gives the phase delay of the lth spectral coefficient $a_l^-(\rho)$ for an incoming wave in an Airy layer. One aligns $G[\rho,\nu]$ with $\vartheta_l^-(\rho)$ at a radial distance ρ where $G[\rho,\nu]$ is still accurate. The form for $\vartheta_l^-(\rho)$ is given by

$$\vartheta_l^-(\rho)=\tan^{-1}\left(\mp\frac{\mathrm{Bi}[\tilde{y}]}{\mathrm{Ai}[\tilde{y}]}\right)-\tan^{-1}\left(\frac{\mathrm{Bi}[\hat{y}]}{\mathrm{Ai}[\hat{y}]}\right)+\text{constant} \tag{6.4-22}$$

where $\tilde{y}$ is defined in Eq. (5.7-18) for an Airy layer, and $\tilde{y}/\hat{y}=|1-\beta|^{-2/3}$. The top sign in Eq. (6.4-22) applies to a super-refracting medium where $\beta>1$; the bottom sign applies when $\beta<1$. One can readily show that the asymptotic forms for $\vartheta_l^-(\rho)$ and $G[\rho,\nu]$ for negative values of $\hat{y}$ are identical in an Airy layer. See Fig. 5-7. For decreasing $\rho<\nu$, note that $\vartheta_l^-(\rho)-\vartheta_l^-(0)$ rapidly approaches zero.

From Eq. (6.4-22), we have

$$\frac{\partial\vartheta_l^-(\rho)}{\partial\nu}\doteq\frac{1}{\pi K_\nu}\left(\frac{\mp\tilde{y}}{\mathrm{Ai}[\tilde{y}]^2+\mathrm{Bi}[\tilde{y}]^2}-\frac{\hat{y}}{\mathrm{Ai}[\hat{y}]^2+\mathrm{Bi}[\hat{y}]^2}\right)\frac{1}{\hat{y}}+\ \text{constant} \tag{6.4-23}$$

To align $\partial G[\rho,\nu]/\partial\nu$ with $\partial\vartheta_l^-/\partial\nu$ in the vicinity of $\rho=\nu$, one sets $\partial G[\rho,\nu]/\partial\nu$ according to the following schedule:

$$\frac{\partial G[\rho,\nu]}{\partial \nu} \Rightarrow \left\{ \begin{array}{l} \dfrac{\partial G[\rho,\nu]}{\partial \nu}, \;\; \rho \geq \rho^{\Diamond} = \nu + \hat{y}^{\Diamond} K_{\nu} \\ \dfrac{\partial G\left[\rho^{\Diamond},\nu\right]}{\partial \nu} + \left(\dfrac{\partial \vartheta_l^{-}(\rho)}{\partial \nu} - \dfrac{\partial \vartheta_l^{-}\left(\rho^{\Diamond}\right)}{\partial \nu} \right), \;\; \rho \leq \rho^{\Diamond} \end{array} \right\} \tag{6.4-24}$$

The chosen value of $\hat{y}^{\Diamond}$ is a compromise, taking into account on one hand the impending failure of $g(\hat{y})$ to provide the correct phase delay for $a_l^{-}(\rho)$ for increasing $\hat{y}$ near zero, and on the other hand the decreasing accuracy of the Airy layer approximation to the actual refractivity profile if applied over too wide an altitude range. Equation (6.4-23) shows that $\partial\left(\vartheta_l^{-}(\rho) - \vartheta_l^{-}(0)\right)/\partial \nu$ also rapidly approaches zero for decreasing $\rho < \nu$. Thus, we are concerned about the Airy layer approximation over a relatively narrow altitude range. The Airy layer approximation should be valid over the altitude range $\sim -2K_{\nu} \leq \rho - \nu \leq -\hat{y}^{\Diamond} K_{\nu}$, about 50 m. Here we have set $\hat{y}^{\Diamond} = -2$. With this modified form, the derivative of the spectral density function becomes

$$\begin{aligned} &\frac{dG^{\dagger}}{d\nu} \Rightarrow \frac{\partial G[0,\nu]}{\partial \nu} \Rightarrow \frac{\partial G\left[\rho^{\Diamond},\nu\right]}{\partial \nu} + \left(\frac{\partial \vartheta_l^{-}(0)}{\partial \nu} - \frac{\partial \vartheta_l^{-}\left(\rho^{\Diamond}\right)}{\partial \nu} \right) = \\ &\frac{\partial G\left[\rho^{\Diamond},\nu\right]}{\partial \nu} - \frac{1}{\pi K_{\nu}} \left(\frac{\mp \tilde{y}^{\Diamond}}{\mathrm{Ai}\left[\tilde{y}^{\Diamond}\right]^2 + \mathrm{Bi}\left[\tilde{y}^{\Diamond}\right]^2} - \frac{\hat{y}^{\Diamond}}{\mathrm{Ai}\left[\hat{y}^{\Diamond}\right]^2 + \mathrm{Bi}\left[\hat{y}^{\Diamond}\right]^2} \right) \frac{1}{\hat{y}^{\Diamond}} \end{aligned} \tag{6.4-25}$$

The weighting function $W(\rho, \rho_*)$ also is slightly modified. In this regard, note that the corrective term from the Airy layer in Eq. (6.4-25) approaches zero as β approaches zero with increasing impact parameter.

The difference in the two forms for $dG^{\dagger}/d\nu$ in Eqs. (6.4-2) and (6.4-25) is given by

$$\begin{aligned} &\Delta\left(\frac{dG^{\dagger}}{d\nu} \right) = -\pi K_{\nu} \int_{\rho^{\dagger}}^{\rho^{\Diamond}} \frac{d \log n}{d\rho} \left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2 \right) d\rho + \frac{\partial \vartheta_l^{-}\left(\rho^{\Diamond}\right)}{\partial \nu} - \frac{\partial \vartheta_l^{-}(0)}{\partial \nu} \\ &\doteq \frac{d \log n}{d\rho} g\left(\hat{y}^{\Diamond}\right) + \frac{1}{\pi \hat{y}^{\Diamond} K_{\nu}} \left(\frac{\mp \tilde{y}^{\Diamond}}{\mathrm{Ai}\left[\tilde{y}^{\Diamond}\right]^2 + \mathrm{Bi}\left[\tilde{y}^{\Diamond}\right]^2} - \frac{\hat{y}^{\Diamond}}{\mathrm{Ai}\left[\hat{y}^{\Diamond}\right]^2 + \mathrm{Bi}\left[\hat{y}^{\Diamond}\right]^2} \right) \\ &\rightarrow \frac{1}{K_{\nu}} \left(\frac{\beta}{\beta - 1} \right) \left(\frac{6 - 5\beta}{32} \right) \left(-\hat{y}^{\Diamond} \right)^{-5/2} + \mathrm{O}\left[\left(-\hat{y}^{\Diamond} \right)^{-4} \right] \end{aligned} \tag{6.4-26}$$

To the extent that the Airy layer approximation is valid, this form gives the error in Eq. (6.4-2) for $dG^{\dagger}/d\nu$. For $\beta = 0.5$, this difference is about 0.1 mrad

at $\hat{y}^{\diamond} = -2$; for $\beta = 2$, it is about 0.2 mrad. The error increases as $\beta \to 1$. The Airy layer analysis in Section 5.7 fails at $\beta = 1$, exactly.

With this modification, one still ends up with an integration interval over ρ that is effectively truncated from below at $\rho = \nu - \hat{y}^{\diamond} K_{\nu}$ instead of at $\rho = \nu - \hat{y}^{\dagger} K_{\nu}$. A super-refractivity zone still yields three separate integration sections for $dG^{\dagger}(\nu)/d\nu$ when the spectral number lies in the super-refractivity zone $kn(r_u)r_u < \nu < kn(r_d)r_d$: one section above $r = r_u$, one below $r = r_u$ down to $r = r_d$, and one below $r = r_d$ down to the turning point. See also the discussion in Section 5.8 on comparison of the wave equation solutions in a super-refracting medium.

6.4.3 Resolution Issues

One could infer from Eq. (6.4-15) that away from a super-refracting region the resolution with which $n(\rho)$ may be recovered is limited only by the SNR measurement errors present in the Fourier transform quantity $F[\omega]$. But there are a number of implicit assumptions embedded in Eq. (6.4-15), not the least of which are the assumptions of spherical symmetry and error-free ionosphere effects. These and other resolution-limiting factors have been discussed in [8,9]. Departures from spherical symmetry arise in two forms: a first part, largely deterministic, is due to departures of the geopotential of the Earth from spherical symmetry, the oblateness being the principal term. The latitude-dependent Earth-flattening term is factored into the spectral formulation presented here by a small adjustment to the value of r_o and to the satellite coordinates. The second part is due to imperfect a priori knowledge of the topography of any given surface of constant refractivity—along-track water vapor variability, for example. Another contributor is the geostrophic effect from winds aloft on a surface of constant pressure. Ad hoc calibrations could be used to correct for these usually small effects for each occultation, using some local model from European Centre for Medium-Range Weather Forecasts (ECMWF), for example—not an impossible task, but surely a tedious one. In a thin phase screen model, these adjustments are equivalent to adjusting D, the limb distance, by an amount δD. It can be shown [8,9] that an uncertainty or error δD in the adopted value of D degrades the resolution Δh in the screen and that $\delta(\Delta h)/\mathscr{F}_o$ is at least as large as $0.45(\delta D/D)^{1/2}$, where $\mathscr{F}_o$ is the vertical diameter of the free space first Fresnel zone, about 1.5 km. Thus, a 1 percent error in D maps into a limiting vertical resolution that is about 5 percent of $\mathscr{F}_o$; a 4 percent error maps into about 10 percent, and so on. With respect to the local Fresnel limit, these percentages would be greater by the factor $\zeta^{-1/2}$ because of defocusing. In other words, the limiting resolution is quite sensitive to this type of error.

Similar conclusions follow for other resolution-limiting factors. The horizontal resolution is limited by the vertical resolution [see Eq. (2.2-10)]. But the vertical resolution is limited in effect by the horizontal resolution, as just discussed. Equating these two expressions containing δD yields

$$\delta D \approx 90\left(\frac{\zeta}{(1-\beta)^2}\right)^{1/3} \text{ km} \tag{6.4-27}$$

This gives a minimum horizontal resolution of about 50 km where the defocusing is 1/10 and $\beta = 0.2$, and of about 100 km where the defocusing is 1. This limiting horizontal resolution yields a limiting vertical resolution that is 15 to 20 percent of the local Fresnel diameter. That the Fresnel limit on vertical resolution can be substantially surpassed using scalar diffraction theory techniques (30 to 100 m) has already been demonstrated [8,10–12], but a realistic limit is yet to be established.

Another consideration is the cross-track and along-track drifts of the tangency point during its descent or ascent. The velocity vector of the tangency point of an occultation is rarely exactly vertical in the upper atmosphere; it is canted off-vertical depending on the orbit geometry of the satellites. The cross-track velocity of the tangency point is proportional to the tangent of this off-vertical angle. Moreover, the vertical component of the tangency point velocity vector slows because of defocusing. But, the cross-track component of the drift is not appreciably slowed, and the along-track component accelerates with depth, approaching the rate $-\rho_* d\tilde{\theta}_L / dt$ in strong defocusing. These factors result in a progressively more shallow descent ratio with depth for the tangency point. In other words, the horizontal displacement of the tangency point during an occultation is a significant factor in resolution questions. Of course, in a 4DVAR context, these resolution issues should be addressed in terms of the information content already extant in the 4DVAR system.

6.5 Summary

This section outlines one spectral technique based on wave theory for recovering the refractivity profile $N(\rho)$ from the spectral derivative of the Fourier transform of the received amplitude and stopped phase measurements made by the LEO over time. Coincidentally, the bending-angle profile $\alpha_{\mathrm{L}}(\nu)$ is recovered. Spectral techniques in general facilitate recovery of bending angles and refractivity fundamentally because of their ability to uniquely sort received rays according to their excess Doppler or to their impact parameter values in a spherical symmetric medium. Spectral techniques seem essential when adverse signal conditions prevail because of the concurrent reception of multiple rays. Spectral techniques also are efficacious, but perhaps not essential if a third-

order ray theory is used in near-caustic situations where the validity of second-order geometric optics breaks down.

Because of the close correspondence noted in Table 5-1 between the phase delay spectral density function $2G^{\dagger}(\nu)$ evaluated at its stationary value in spectral number and the scalar diffraction/thin-screen phase function $\varphi(\rho_*)$ discussed in Chapter 2, one need not start from wave theory as the framework for obtaining these spectral results. We noted in Section 5.10 that the phase delay spectral density function $G^{\dagger}(\nu)$ from wave theory and the thin-screen phase function $\varphi(h)$ are related by

$$-2G^{\dagger}(\nu) \doteq 2\int_{\nu}^{\infty} \tilde{\alpha}(\nu,\nu)d\nu = \int_{\nu}^{\infty} \alpha_{\mathrm{L}}(\nu)d\nu \Leftrightarrow \varphi(\nu) \tag{6.5-1}$$

The analog of $\Psi_{\mathrm{L}o}$ in the thin screen is $\Phi\left(h,h_{\mathrm{LG}}\right)$, the Fresnel phase function given in Eq. (2.5-1). It follows that essentially the same form given in Eq. (6.4-5) would come from the scalar diffraction integral in a thin phase screen model, provided the screen is so constructed that there is a one-to-one relationship between h and the impact parameter. Therefore, with this caveat in mind, the wave-theoretic Fourier approach followed here in Chapter 6 to recover the refractivity profile could be converted into a thin-screen Fourier approach by replacing $-2G^{\dagger}(\nu)$ with $\varphi(\rho_*)$ and its derivative $2dG^{\dagger}(\nu)/d\nu$ with $-d\varphi(\rho_*)/d\rho_*$, that is, with $\alpha_{\mathrm{L}}(\rho_*)$. One also can start from a thin phase screen model, whose surface is defined by the impact parameter, and use scalar diffraction theory to obtain essentially the same results. This offers an easier approach when the thin-screen model is caustic-free.

Finally, an intriguing prospect for future in-receiver signal processing operations in support of limb sounding is the incorporation of advanced processing techniques, such as essentially real-time Fourier transform algorithms. With current POD information, a smart receiver not only could power through adverse signal episodes and perform backward (in time) reconstruction, it also could report bending angles for multiple rays and refractivity profiles directly, along with essential statistics, basic phase and amplitude data, and other housekeeping information.

One should not underestimate the potential cost savings of an in-receiver automated system using the GPS, especially in adverse signal conditions. For an analogy, one need look no further than the navigation of a typical LEO with an onboard GPS receiver. Here one can find highly automated processes resulting in centimeter-level orbit accuracy [13]. Once operational, the workforce required to routinely maintain and use this capability is an order of magnitude smaller than that required for most ground-based tracking systems [14]. Investment now in the necessary research and development to enable future automated in-receiver operations for GPS-based limb sounding would

seem to be a wise programmatic option. In the future, other global navigation satellite system programs besides the GPS will become operational. Use of these systems also should be incorporated in future flight receiver designs.

References

[1] R. Anthes, C. Rocken, and Y. H. Kuo, "Applications of COSMIC to Meteorology and Climate," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 115–156, 2000.

[2] T.Yunck, C. H. Liu, and R. Ware: "A History of GPS Sounding," *Terrestrial, Atmospheric and Ocean Sciences*, vol. 11, pp. 1–20, 2000.

[3] J. Thomas, *Signal-Processing Theory for the TurboRogue Receiver*, JPL Publication 95-6, Jet Propulsion Laboratory, Pasadena, California, 1995.

[4] C. Ao, T. Meehan, G. Hajj, A. Mannucci, and G. Beyerle, "Lower Troposphere Refractivity Bias in GPS Occultation Retrievals," submitted to *Journal of Geophysical Research*, 2003.

[5] J. Jackson, *Classical Electrodynamics,* 2nd ed., New York, New York: J. Wiley & Sons, Inc., 1975.

[6] G. Lindal, J. Lyons, D. Sweetnam, V. Eshelman, D. Hinson, and G. Tyler, "The Atmosphere of Uranus: Results of Radio Occultation Measurements with Voyager 2," *Journal of Geophysical Research,* vol. 92, pp. 14,987–15,001, 1987.

[7] S. Sokolovskiy, "Modeling and Inverting Radio Occultation Signals in the Moist Troposphere," *Radio Science*, vol. 36, no. 3, pp. 441–458, 2001.

[8] E. Marouf, G. Tyler, and P. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120–166, 1986.

[9] W. Melbourne, E. Davis, C. Duncan, G. Hajj, K. Hardy, E. Kursinski, T. Meehan, L. Young, and T. Yunck, *The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring*, JPL Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, 1994.

[10] E. Karayel and D. Hinson, "Sub-Fresnel Scale Vertical Resolution in Atmospheric Profiles from Radio Occultation," *Radio Science*, vol. 32, no. 2, pp. 411–423, 1997.

[11] M. Gorbunov, A. Gurvich, and L. Bengtsson, *Advance Algorithms of Inversion of GPS/MET Satellite Data and Their Application to Reconstruction of Temperature and Humidity*, Max Planck Institute for Meteorology, Report 211, Hamburg, Germany, 1996.

[12] K. Igarashi, K. Pavelyev, J. Wickert, K. Hocke, and D. Pavelyev, "Application of Radio Holographic Method for Observation of Altitude Variations of the Electron Density in the Mesosphere/Lower Thermosphere Using GPS/MET Radio Occultation Data," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 64, pp. 959–969, 2002.

[13] R. Muellerschoen, A. Reichert, D. Kuang, M. Heflin, W. Bertinger, and Y. Bar-Sever, "Orbit Determination with NASA's High Accuracy Real-Time Global Differential GPS System," ION GPS2001, September 11–14, 2001, Salt Lake City, Utah.

[14] W. Melbourne, T. Yunck, W. Bertiger, B. Haines, and E. Davis, "Scientific Applications of GPS on Low Earth Orbiters," *Zeitschrift für Satellitengestützte Positionierung, Navigation und Kommunikation*, vol. 4, pp. 131–145, 1993.

Appendix A:
Miscellaneous Derivations

A.1 Bending Angle and Bouguer's Law, Eq. (2.2-2)

In geometric optics, a ray, if it exists, must be a path of stationary phase. This is Fermat's principle. In other words, the cumulative phase along the path integral between the positions A and B,

$$\varphi = k\int_A^B n ds = k\int_A^B n\sqrt{1+(r\theta')^2}\,dr \tag{A-1}$$

must be stationary with respect to all neighboring paths passing through the end points A and B.[1] Here φ is the phase accumulation along the path; $n = n(r,\theta)$ is the index of refraction; s is path length; and $\theta' = d\theta/dr$, which gives the slope of the ray path. The term "all neighboring paths" means all continuous paths within some suitably defined local region about the actual ray path. Fermat's principle is a local condition, not a global one. From the Calculus of Variations a necessary condition for stationarity is that the ray at all points must

[1] Here we have assumed a Euclidian metric for space-time curvature. In practice, however, with the precise calculation of the predicted phase to be compared with actual data, general relativistic space-time curvature from the mass of the Earth and special relativistic effects from the motions of the satellites have to be accounted for to achieve satisfactory accuracy. The path integral for the ray in this case involves a general relativistic curvature tensor, and the integral is better written in a four-dimensional parametric form. A ray is a geodesic in relativistic four-dimensional space-time. In practice, relativity effects usually are included as small differential corrections to the Euclidian form for the phase delay. They sometimes are referred to as "post-Newtonian" corrections [4].

satisfy Euler's differential equation or its integral equivalent at points of discontinuity in n [1–3]. Euler's equation in polar coordinates is given by

$$\frac{d}{dr}\frac{\partial}{\partial\theta'}\left(n\sqrt{1+(r\theta')^2}\right)-\frac{\partial}{\partial\theta}\left(n\sqrt{1+(r\theta')^2}\right)=0 \tag{A-2}$$

When n is a function of r only, one can integrate Eq. (A-2) to obtain

$$\frac{\partial}{\partial\theta'}\left(n\sqrt{1+(r\theta')^2}\right)=\text{constant}=a \tag{A-3}$$

Carrying out the differentiation, one obtains

$$\frac{nr^2\theta'}{\sqrt{1+(r\theta')^2}}=nr\sin\gamma=a=r_*n(r_*) \tag{A-4a}$$

$$\frac{d\theta}{dr}=\pm\frac{a}{r}\frac{1}{\sqrt{n^2r^2-a^2}} \tag{A-4b}$$

where γ is the angle between the radius vector and the direction of propagation of the ray at the point (r,θ) (see Fig. A-1). When we set the occulted GPS satellite infinitely afar in the direction $\theta=\pi$, then $\gamma=\theta+\tilde{\alpha}$, where $\tilde{\alpha}$ is the cumulative bending angle of the ray at the point (r,θ) relative to its asymptotic direction prior to encountering the atmosphere. Equation (A-4a) is Bouguer's law, and a is the impact parameter, which is a ray-specific constant when spherical symmetry applies. In Eq. (A-4b), the minus sign is used in the first quadrant and the plus sign in the second quadrant. At the turning point(s) where $d\theta/dr\to\infty$, it follows that $a=r_*n(r_*)$, where r_* denotes a turning point.

To obtain a differential equation for the bending angle, we note that a small incremental change in position Δs along the ray results in a small incremental

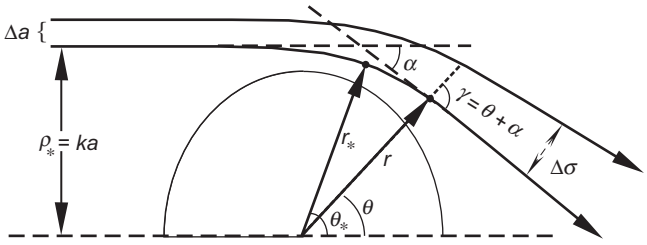

Fig. A-1. Defocusing geometry. Small cross-track displacement Δa of the ray in the plane of propagation before encountering the atmosphere results in a post-encounter cross-track displacement $\Delta\sigma$.

change $\Delta\gamma$ in the angle between the radius vector and the ray path tangent vector. But $\Delta\gamma = \Delta\theta + \Delta\tilde{\alpha}$, even when the occulted Global Positioning System (GPS) satellite is only a finite distance away. By differentiating Bouguer's law, $nr\sin\gamma = a$, with respect to r, replacing $d\gamma / dr$ with $d\theta / dr + d\tilde{\alpha} / dr$, and using Eq. (A-4b) for $d\theta / dr$, one obtains

$$\frac{d\tilde{\alpha}}{dr} = -a\frac{n'}{n}\frac{1}{\sqrt{n^2r^2 - a^2}} \tag{A-5}$$

Integrating Eq. (A-5) to infinity along both sides of the turning point yields Eq. (2.2-2).

We note from Eq. (A-4b) that a necessary condition for a real solution to Euler's equation and, therefore, for the existence of a ray between specified endpoints, is that $nr \geq a$ at all points along the ray.

A.2 Defocusing

The defocusing factor may be obtained by differentiating the equation for Bouguer's law with respect to a. We assume that the low Earth orbiting (LEO) satellite is located at the point (r,θ), which is outside of the refracting medium so that $n \equiv 1$ in this neighborhood. Let σ be a cross-track distance perpendicular to the ray at the point (r,θ). To compute the defocusing, we need the ratio of the signal flux densities for incoming and outgoing rays. Therefore, we need the ratio $\Delta a / \Delta\sigma$, where $\Delta\sigma$ is the displacement of the ray in the cross-direction at the point (r,θ) due to a displacement of the incoming ray at a point before encountering the atmosphere, which is a change in impact parameter Δa (see Fig. A-1). Upon differentiating $nr\sin\gamma = a$ with respect to a while constraining the LEO to lie on the cross-track perpendicular to the ray ($\Delta r = \Delta\sigma\sin\gamma$, $r\Delta\theta = \Delta\sigma\cos\gamma$), we obtain

$$\frac{d\sigma}{da} = 1 - r\cos\gamma\frac{d\alpha}{da} \tag{A-6}$$

But $r\cos\gamma = \sqrt{r^2 - a^2} = D$, which is essentially the distance of the LEO from the Earth's limb minus a relatively small arc length $a\alpha$. Hence,

$$\zeta = \frac{da}{d\sigma} = \left(1 - D\frac{d\alpha}{da}\right)^{-1} \tag{A-7}$$

A.3 Excess Doppler

When one removes the Doppler shift on the measured carrier frequency due to the relative motion between the LEO and the occulted GPS satellite, the remainder due to refraction is called the "excess Doppler." This is equivalent to removing the relative velocity component along the LEO–GPS line, leaving only the cross-velocity component. The component of the cross-velocity lying in the plane of propagation, that is, the plane containing the GPS–LEO line and the geocenter (assuming spherical symmetry and, therefore, coplanar propagation), is designated as $V_\perp$. Thus, $V_\perp$ is the geocentric closing velocity of the turning point or tangency point of the GPS–LEO straight line. For a LEO orbit radius of $1.1R_E$, $V_\perp \approx 2.5$ km/s during an occultation through the atmosphere. The exact value differs depending primarily on the angle between $\boldsymbol{V}_\perp$ and the orbital cross-velocity vector. When this angle is less than 30 deg, $V_\perp$ ranges roughly between 2 and 3 km/s. $V_\perp$ is shown pictorially in Fig. A-2, which also shows pre- and post-encounter planar wave fronts of the ray. (Here, the GPS satellite is assumed to be infinitely afar.)

Here is a heuristic derivation of the excess Doppler expressed in terms of bending angle. To eliminate the Doppler shift from the relative velocity between the LEO and the occulted GPS satellite, we freeze the motion of the wave fronts of the signal and consider only the motion of the LEO perpendicular to the approaching wave fronts. Referring to Fig. A-2, we see that a small displacement $V_\perp \Delta t$ of the LEO perpendicular to the GPS–LEO line in a short time interval, Δt, results in a change in measured phase at the LEO of

$$\Delta\varphi = kV_\perp \Delta t \sin\alpha \tag{A-8}$$

By dividing $\Delta\varphi$ by $k\Delta t$ and letting $\Delta t \to 0$, one obtains the excess observed Doppler f_D in hertz. It is related to the observed bending angle through the near-linear expression

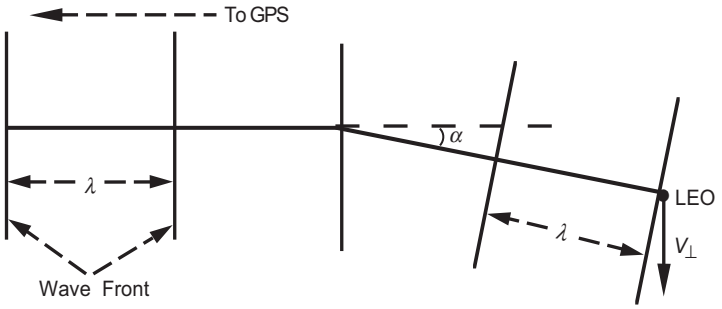

Fig. A-2. Geometry for the relationship between excess Doppler f_D and bending angle α. To first order, $\lambda f_D \doteq \alpha V_\perp$.

$$\lambda f_D = \alpha V_\perp + \mathrm{O}[\alpha^2] \tag{A-9}$$

For dry air at sea level, α is about 20 mrad. Thus, the excess Doppler from dry air ranges up to a few hundred hertz. The Doppler from the relative motion between the two satellites ranges up to 35 kHz.

A more comprehensive version of this expression in Eq. (A-9) can be derived using the transversality condition from the Calculus of Variations [3,5]. The transversality condition results from considering the given ray passing through the positions of the GPS and LEO satellites as being embedded in a family of neighboring rays generated by varying the values of one or more boundary conditions, such as the position coordinates of the LEO and GPS satellites. The phase delay, which is stationary on each ray, then becomes a field variable, in this case a function of the six satellite position coordinates. Displacement of the GPS and LEO satellites over a short time interval Δt to new nearby positions causes a new ray to intersect the new points. The phase delay on the new ray between the new positions of the two satellites minus the phase delay on the old ray between the old positions, $\delta\varphi$, is given from the transversality condition. Let

$$F(r,\theta,\theta') = n(r)\left(1 + r^2\theta'^2\right)^{1/2} \tag{A-10}$$

which is the integrand in Eq. (A-1) for the path integral giving the phase delay φ. The transversality condition gives the change in φ resulting from small displacements in the end point positions A and B of the integral in Eq. (A-1) *when the integral is evaluated along a path that gives it a stationary value*. Thus, the path between A and B must satisfy the Euler equation

$$\frac{d}{dr}\left(\frac{\partial F}{\partial \theta'}\right) - \frac{\partial F}{\partial \theta} = 0 \tag{A-11}$$

at every point, or its integral equivalent. By varying the end points, it can be shown from the calculus of variations that the change in φ is given from the transversality condition by[2]

---

[2] The transversality condition is related to the Hamilton–Jacobi partial differential equation, which describes the behavior of φ as a field variable, that is, as a function of (r,θ) at the end points. To keep the matter simpler, we consider only the variability of φ at the single end point at B. We define the conjugate momentum as $p = \partial F / \partial \theta'$, and also the Hamiltonian as $H(r,\theta,\theta') = F - \theta' \partial F / \partial \theta'$. Then in terms of r, p, and θ, it follows that $r\theta' = p/(n^2 r^2 - p^2)^{1/2}$ and the Hamiltonian becomes $r\mathscr{H}(r,\theta,p) = (n^2 r^2 - p^2)^{1/2}$. A specific ray path in the field is described by a pair of first-order differential equations in terms of the independent variable r:

$$\delta\varphi = \left[\left(F - \theta' \frac{\partial F}{\partial \theta'}\right)\delta r + \frac{\partial F}{\partial \theta'}\delta\theta\right]_A^B \tag{A-12}$$

where δr and $\delta\theta$ are the variations in end point values. Upon setting $\delta r = \dot{r}\Delta t$ and $\delta\theta = \dot{\theta}\Delta t$, we obtain

$$\delta\varphi = k\left(n_{\mathrm{L}}\boldsymbol{T}_{\mathrm{L}} \cdot \boldsymbol{V}_{\mathrm{L}} - n_{\mathrm{G}}\boldsymbol{T}_{\mathrm{G}} \cdot \boldsymbol{V}_{\mathrm{G}}\right)\Delta t \tag{A-13}$$

where n is the index of refraction, $\boldsymbol{V}$ is the velocity vector of the satellite, and $\boldsymbol{T}$ is the unit tangent vector of the ray. The subscript "L" on these quantities denotes their evaluation at the position of the LEO at the time of reception of the signal, and the subscript "G" denotes evaluation at the occulted GPS

$d\theta / dr = -\partial\mathscr{H} / \partial p$ and $dp / dr = \partial\mathscr{H} / \partial\theta \equiv 0$. The conjugate momentum $p = nr\sin\gamma$ is the impact parameter, and it is a constant along a ray when spherical symmetry applies. But it, along with r and θ, are field variables in Hamilton–Jacobi theory. Also, the Hamiltonian is given by $r\mathscr{H} = nr\cos\gamma$, which is essentially the limb distance minus the small bending-angle offset $a\alpha$. Hamilton–Jacobi theory is the basis for the canonical transform technique in ray theory [6] for converting spatial coordinates for rays (r,θ) into phase space, or canonical coordinates (r,θ,p). In spatial coordinates, multiple rays may pass through a point (r,θ) in multipath situations, that is, different values of θ' can hold. But, in canonical coordinates, only a single ray can pass through a given point (r,θ,p), provided that spherical symmetry holds (and super-refracting situations are avoided). One uses a Fourier integral transform similar to the diffraction integral given in Eq. (A-22) to transform the phase and amplitude measurements made by the LEO along a path in spatial coordinates into an equivalent single-valued set in canonical coordinates to obtain the phase and amplitude as a function of p. One then can apply the Abel transform to the derived bending-angle profile in the canonical surface to recover the refractivity profile.

From the transversality condition, it also follows that the behavior of φ as a field variable evaluated at the end point (r,θ) at B is given by $\partial\varphi / \partial r - \mathscr{H} = 0$ and $\partial\varphi / \partial\theta - p = 0$. Upon replacing p by $\partial\varphi / \partial\theta$ in the Hamilton–Jacobi equation, $\partial\varphi / \partial r - \mathscr{H}(r,\theta,p) = 0$, we have $(\partial\varphi / \partial r)^2 + r^{-2}(\partial\varphi / \partial\theta)^2 = |\nabla\varphi|^2 = n^2$, which is the eikonal equation.

satellite at the time of transmission.[3] Because φ is stationary on each ray, $\delta\varphi$ depends to first order only on the end point values, which is somewhat remarkable considering that the value of φ comes from a path integral that explicitly accounts for the refractivity along the entire ray path. From $\delta\varphi$, we subtract the term $k\dot{r}_{\mathrm{LG}}\Delta t = k\Delta t(\boldsymbol{V}_{\mathrm{L}} - \boldsymbol{V}_{\mathrm{G}})\cdot(\boldsymbol{r}_{\mathrm{L}} - \boldsymbol{r}_{\mathrm{G}})/|\boldsymbol{r}_{\mathrm{L}} - \boldsymbol{r}_{\mathrm{G}}|$, which is the phase change from the radial velocity between the two satellites over the time interval Δt. Then dividing by $k\Delta t$ and letting $\Delta t \to 0$ yields $f_D = (d\varphi/dt)/2\pi$, the excess Doppler from the intervening medium, which is given by

$$\lambda f_D = n_{\mathrm{L}}\boldsymbol{T}_{\mathrm{L}}\cdot\boldsymbol{V}_{\mathrm{L}} - n_{\mathrm{G}}\boldsymbol{T}_{\mathrm{G}}\cdot\boldsymbol{V}_{\mathrm{G}} - \frac{(\boldsymbol{V}_{\mathrm{L}} - \boldsymbol{V}_{\mathrm{G}})\cdot(\boldsymbol{r}_{\mathrm{L}} - \boldsymbol{r}_{\mathrm{G}})}{|\boldsymbol{r}_{\mathrm{L}} - \boldsymbol{r}_{\mathrm{G}}|} \tag{A-14}$$

This form in Eq. (A-14) accounts for the velocities of both the transmitting and receiving satellites. It is quite general, applying to three-dimensional situations, to finite satellite distances, and to the case where the satellites are located in a refracting medium with or without spherical symmetry.

From Fig. A-3, we obtain a geometric representation of Eq. (A-14). Figure A-3 applies to a coplanar, spherical symmetric geometry. Here triangle OLG defines the instantaneous plane of propagation of the ray from the occulted GPS satellite to the LEO. This plane osculates with time depending on the three-dimensional orientation of the satellite orbit planes and on the satellite orbital positions. The interior angles of this triangle OLG, χ_{L}, χ_{G}, and $\theta_{\mathrm{G}} - \theta_{\mathrm{L}}$, and its sides, r_{L}, r_{G}, and r_{LG}, are completely determined from the precision orbit determination (POD) information about the orbits of the LEO and the occulted GPS satellite. The refraction-related quantities, which are the impact parameter a and the bending angle $\alpha = \delta_{\mathrm{L}} + \delta_{\mathrm{G}}$, are to be determined from the excess Doppler measurement. From Fig. A-3, the ray path tangent vectors are given by $\boldsymbol{T}_{\mathrm{L}} = \hat{\boldsymbol{r}}_{\mathrm{L}}\cos\gamma_{\mathrm{L}} - \hat{\boldsymbol{\theta}}_{\mathrm{L}}\sin\gamma_{\mathrm{L}}$ and $\boldsymbol{T}_{\mathrm{G}} = -\hat{\boldsymbol{r}}_{\mathrm{G}}\cos\gamma_{\mathrm{G}} - \hat{\boldsymbol{\theta}}_{\mathrm{G}}\sin\gamma_{\mathrm{G}}$, where $\gamma = \chi + \delta$. Also, $\boldsymbol{V}_{\mathrm{L}} = \hat{\boldsymbol{r}}_{\mathrm{L}}\dot{r}_{\mathrm{L}} + \hat{\boldsymbol{\theta}}_{\mathrm{L}}r_{\mathrm{L}}\dot{\theta}_{\mathrm{L}}$ and $\boldsymbol{V}_{\mathrm{G}} = \hat{\boldsymbol{r}}_{\mathrm{G}}\dot{r}_{\mathrm{G}} + \hat{\boldsymbol{\theta}}_{\mathrm{G}}r_{\mathrm{G}}\dot{\theta}_{\mathrm{G}}$. Here $\dot{r}$ and

---

[3] Besides accounting for light-time effects from the finite speed of light, one also would have to distinguish in a special relativistic framework [4] between the proper time intervals, $\Delta\tau_{\mathrm{L}}$ and $\Delta\tau_{\mathrm{G}}$, kept by the clocks onboard the moving satellites, and the coordinate time interval, Δt, kept in the stationary coordinate frame center assumed here (at position O in Fig. A-3). The transmitted and received carrier frequencies depend on proper time. The differences between these time intervals, and also the difference between the relativistic expression and the classical form for the Doppler shift used to obtain Eq. (A-14), can be expressed as a power series of quadratic and higher powers of (V/c). Here V is the velocity of the satellite and c is the speed of light. The magnitude of (V/c) is 10 to 25 parts per million. It turns out that because of term cancellation the form given in Eq. (A-14) is relativistically accurate, at least through quadratic terms in (V/c).

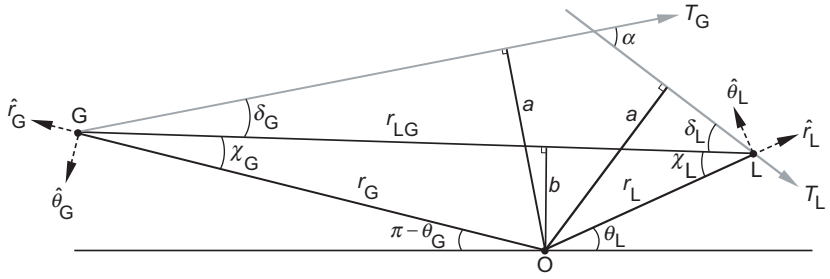

Fig. A-3. Ray path geometry from point G to point L in the plane of propagation for a spherical symmetric medium.

$r\dot{\theta}$ are the radial and transverse components of the orbital velocity vector projected into the plane of propagation defined by the triangle OLG. Also, $\hat{\boldsymbol{r}}$ and $\hat{\theta}$ are orthogonal unit vectors defining the local rotating coordinate frame. The radial velocity between the satellites $\dot{r}_{LG}$ is given by $\dot{r}_{LG} = \dot{r}_L \cos\chi_L + \dot{r}_G \cos\chi_G + b\left(\dot{\theta}_G - \dot{\theta}_L\right)$, where b is the impact parameter of the straight line between the LEO and the occulted GPS satellite. In practice, $\dot{\theta}_L$ is roughly an order of magnitude larger than $\dot{\theta}_G$. Also, for near-circular orbits, $r\dot{\theta} >> \dot{r}$.

By using these geometric expressions and applying Bouguer's law, $a = nr\sin\gamma = \text{constant}$, one obtains from Eq. (A-14) (with $n \equiv 1$)

$$\lambda f_D = \left(\sqrt{r_L^2 - a^2} - \sqrt{r_L^2 - b^2}\right)\frac{\dot{r}_L}{r_L} - \left(\sqrt{r_G^2 - a^2} - \sqrt{r_G^2 - b^2}\right)\frac{\dot{r}_G}{r_G} + (a-b)\left(\dot{\theta}_G - \dot{\theta}_L\right) \tag{A-15}$$

The only quantity on the right-hand side (RHS) of Eq. (A-15) not known from POD information is the impact parameter a. Therefore, the measurement of the excess Doppler f_D enables a determination of a. Also, from Bouguer's law we have

$$\left.\begin{aligned} a &= r_L \sin\left(\delta_L + \chi_L\right) \\ a &= r_G \sin\left(\delta_G + \chi_G\right) \end{aligned}\right\} \rightarrow \delta_L, \delta_G \Bigg\} \qquad \alpha = \delta_L + \delta_G \rightarrow \alpha \tag{A-16}$$

from which a determination of the bending angle α follows.

If spherical symmetry does not apply, then these expressions in Eqs. (A-15) and (A-16) become somewhat more complicated, and Bouguer's law does not

apply. Here some recursive ray-tracing technique is required to obtain the turning point position of the ray (r_*, θ_*) and its bending angle. Equation (A-14) shows that the one-to-one relationship between bending angle and excess Doppler still holds, but the relationship between bending angle and turning point must be qualified when spherical symmetry does not apply. In other words, one can conjure up a strong along-track refractive gradient scenario that leads to another turning point value yielding the same bending angle, whereas a gradual departure from spherical symmetry, for example, an oblate spheroid, may not. Figure A-4 shows this former point in a thin phase screen context. Here the limb distance D, or distance of the LEO from the screen, varies with altitude. The same bending angle occurs at two different altitudes in the screen, or at two different impact parameter values.

In Eqs. (A-8) and (A-9), we have simplified the matter further by placing the occulted GPS satellite infinitely afar in the $\theta = \pi$ direction, and we have used the reduced limb distance D to compensate for this simplification. In this case, we set $\boldsymbol{T}_G = (\boldsymbol{r}_L - \boldsymbol{r}_G)/|\boldsymbol{r}_L - \boldsymbol{r}_G| \rightarrow -\hat{\boldsymbol{r}}_G$, $\boldsymbol{V}_G = 0$, $\chi_G = 0$, and $\delta_G = 0$. Also, $\boldsymbol{T}_L = \hat{\boldsymbol{r}}_L \cos\gamma_L - \hat{\theta}_L \sin\gamma_L$, where in this case $\gamma_L = \theta_L + \alpha$. Equation (A-14) becomes

$$\lambda f_D = (\boldsymbol{T}_L - \boldsymbol{T}_G) \cdot \boldsymbol{V}_L = V_\perp \alpha + \mathrm{O}[\alpha^2] \tag{A-17}$$

where $V_\perp = \boldsymbol{V}_L \cdot \boldsymbol{T}_\perp = -\dot{r}_L \sin\theta_L - r_L \dot{\theta}_L \cos\theta_L$. Equation (A-17) yields the same relationship between excess Doppler and bending angle that is given in Eq. (A-9). For a rising occultation, $\dot{\theta}_L > 0$, and for a setting occultation, $\dot{\theta}_L < 0$.

We note that in the case where r_G is finite, as shown in Fig. A-3, the generalized version of Eq. (A-17) is given by

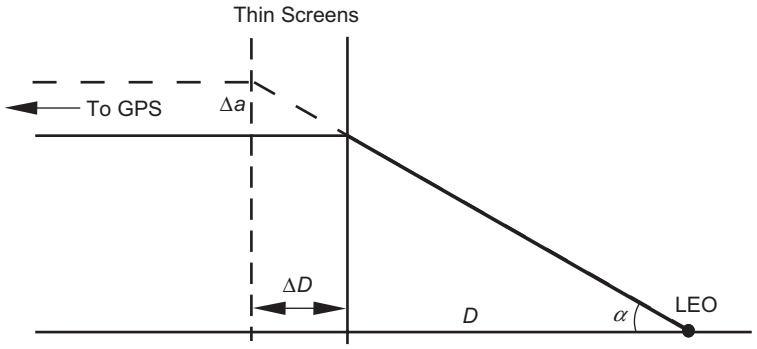

Fig. A-4. Altitude-dependent thin-screen geometry. Because of the assumed horizontal variability in refractivity, the limb distance *D* is a function of the altitude of the ray path turning point.

$$\lambda f_D = \boldsymbol{T}_\perp \cdot \left(\boldsymbol{V}_\mathrm{L}\delta_\mathrm{L} - \boldsymbol{V}_\mathrm{G}\delta_\mathrm{G}\right) + \mathrm{O}\left[\delta_\mathrm{L}^2, \delta_\mathrm{G}^2\right] \tag{A-18}$$

where $\boldsymbol{T}_\perp = -\hat{\boldsymbol{r}}_\mathrm{L}\sin\chi_\mathrm{L} - \hat{\theta}_\mathrm{L}\cos\chi_\mathrm{L} = -\hat{\boldsymbol{r}}_\mathrm{G}\sin\chi_\mathrm{G} + \hat{\theta}_\mathrm{G}\cos\chi_\mathrm{G}$, which is the unit vector in the plane of propagation perpendicular to the straight line between the LEO and the occulted GPS satellite.

Finally, we define $\boldsymbol{a}$ to be the vector version of the impact parameter directed perpendicular to the plane of propagation. Thus, $\boldsymbol{a}$ is analogous to the angular momentum vector in classical mechanics. From Eq. (A-4b), it follows that $\boldsymbol{a} = n(\boldsymbol{T}\times\boldsymbol{r})$, which is a constant vector along a specific ray. The satellite positions evolve with time, and different rays intersect those evolving positions. These rays have different values for the impact parameter and bending angle. At the LEO, we set $n_\mathrm{L} \equiv 1$. Then differentiating the expression $\boldsymbol{a} = \boldsymbol{T}_\mathrm{L}\times\boldsymbol{r}_\mathrm{L}$ with respect to time, one obtains $\dot{\boldsymbol{a}} = \zeta\left(\boldsymbol{T}_\mathrm{L}\times\boldsymbol{V}_\mathrm{L}\right)$ for the rate of change of the impact parameter. Also, $\dot{\alpha} = \pm\zeta(d\alpha/da)\left|\boldsymbol{T}_\mathrm{L}\times\boldsymbol{V}_\mathrm{L}\right|$, with the plus sign for a rising occultation and the negative sign for a setting occultation. Note that the defocusing factor appears in these expressions for $\dot{\boldsymbol{a}}$ and $\dot{\alpha}$, and through Eq. (A-7) the defocusing factor also appears in $\dot{f}_D$. For dry air, $\zeta \approx 0.1$ at sea level. At the tropopause, $\zeta = 0.3 - 0.4$. As a consequence, the variability in f_D resulting from the local variability in the refractivity is increasingly more compressed as the turning point of the ray sinks deeper into the troposphere.

A.4 Scalar Diffraction Theory

Fresnel diffraction effects are computed using scalar diffraction theory. It is based on the Helmholtz–Kirchoff integral theorem from classical electrodynamics, which relates the amplitude and phase distributions of an electromagnetic wave over a surface to the amplitude and phase at a point interior to the surface [7,8]. One form of this integral theorem is given by

$$E(\boldsymbol{r}_1, t) = \frac{1}{2\pi c}\iint_S \frac{1}{r_{12}}\frac{\partial}{\partial t}E\left(\boldsymbol{r}_2, t_\mathrm{R}\right)\left(\hat{\boldsymbol{r}}_{12}\cdot\hat{\boldsymbol{n}}(\boldsymbol{r}_2)\right)dS \tag{A-19}$$

Here $\boldsymbol{r}_1$ is the position vector from the origin to point 1, where the field is to be evaluated; $\boldsymbol{r}_2$ is the position vector to point 2 on the surface element dS; and $\hat{\boldsymbol{r}}_{12}$ is a unit vector directed along $\boldsymbol{r}_{12} = \boldsymbol{r}_2 - \boldsymbol{r}_1$. The unit vector $\hat{\boldsymbol{n}}(\boldsymbol{r}_2)$ is directed outward along the normal to dS (away from the position $\boldsymbol{r}_1$). The quantity t_R is the retarded time, $t_\mathrm{R} = t - r_{12}/c$, which accounts for the extra time r_{12}/c required for a wavelet from dS to reach the point at $\boldsymbol{r}_1$. This surface integral is valid when the wavelength of the wave is small compared to the scale of the radiating surface.

Let us assume that we have a harmonic wave with an angular frequency ω. Then

$$E(\boldsymbol{r}_2, t_R) = E(\boldsymbol{r}_2)\exp(\pm i\omega t_R) = E(\boldsymbol{r}_2)\exp(\pm i\omega t)\exp(\mp ikr_{12}) \qquad \text{(A-20)}$$

where $k = 2\pi/\lambda$, the wave number of the wave in a vacuum. Here, the top or bottom sign to be adopted depends on the propagation sense. Because we assume homogeneity in the transverse horizontal direction, we have coplanar propagation; $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are coplanar and $E(\boldsymbol{r}_2)$ is invariant in the transverse horizontal direction. Then, the two-dimensional surface integral in Eq. (A-19) may be reduced to a one-dimensional path integral by integrating along the transverse direction. Let C be the path defined by the intersection of the plane of propagation and the radiating surface S. Let y be the displacement perpendicular to the plane of propagation. Let the three-dimensional distance $(r_{12})_{3D}$ be written in terms of the two-dimensional distance $(r_{12})_{2D}$ plus the offset y in the transverse direction. Then using the Fresnel approximation, we have $(r_{12})_{3D} \doteq (r_{12})_{2D} + y^2/2(r_{12})_{2D}$, which is valid when $|y|/r_{12} << 1$, or, equivalently, when the scale along y from which significant contributions to the integral occur is small compared to r_{12}. For the time-independent part of the wave, Eq. (A-19) becomes

$$E(\boldsymbol{r}_1) \doteq \frac{\pm i}{\lambda}\int_C \left(\frac{E(\boldsymbol{r}_2)}{r_{12}}\exp(\mp ikr_{12})ds\right)\int_{-\infty}^{\infty}\left(\hat{\boldsymbol{r}}_{12}\cdot\hat{\boldsymbol{n}}(\boldsymbol{r}_2)\right)\exp\left(\mp ik\frac{y^2}{2r_{12}}\right)dy \qquad \text{(A-21)}$$

Here $ds = |d\boldsymbol{r}_2|$ is a line element on the coplanar path C. Also, we have dropped the subscript "$2D$" denoting a two-dimensional vector. Because the wavelength is very small compared to the scale of the surface, the obliquity term $\hat{\boldsymbol{r}}_{12}\cdot\hat{\boldsymbol{n}}(\boldsymbol{r}_2)$ can be moved out of the transverse integral because the phasor $\exp(iky^2/2r_{12})$ winds up so rapidly. The value of this integral in the transverse direction is given by a complete complex Fresnel integral, and it is equal to $(1\pm i)\sqrt{\lambda r_{12}}$. For the two-dimensional case, Eq. (A-21) becomes, for the time-independent part of the wave,

$$E(\boldsymbol{r}_1) = \sqrt{\frac{\pm i}{\lambda}}\int_C\left(\frac{E(\boldsymbol{r}_2)}{r_{12}^{1/2}}\exp(\mp ikr_{12})\left(\hat{\boldsymbol{r}}_{12}\cdot\hat{\boldsymbol{n}}(\boldsymbol{r}_2)\right)\right)ds \qquad \text{(A-22)}$$

By adopting the top sign in Eq. (A-22), we have the basis for the back plane propagation method used to map the actual LEO observations of phase and amplitude backwards (toward the emitter) to an equivalent set in a virtual plane

close to the limb of the Earth. Multipath in the back plane can be significantly reduced with this technique [9,10].

Adopting the bottom sign in Eq. (A-22) leads directly to the Rayleigh–Sommerfeld scalar diffraction integral introduced in Section 2.5. In a thin-screen model (see Fig. 2-3) where the screen is planar and mounted perpendicular to the GPS–LEO line, we use the Fresnel approximation again, $r_{12} \doteq D + (h - h_{\mathrm{LG}})^2 / 2D$, where D is the distance of the LEO from the screen and $h - h_{\mathrm{LG}}$ is the vertical offset. Because $D >> h$, the obliquity term $\hat{\boldsymbol{r}}_{12} \cdot \hat{\boldsymbol{n}}(\boldsymbol{r}_2)$ can be set to unity in our qualitative treatment. Placing these substitutions into Eq. (A-22) and replacing $E(\boldsymbol{r}_2)$ with the thin-screen phase function in Eq. (2.3-16) yields the form for the Rayleigh–Sommerfeld integral given in Eq. (2.5-1).

A.5 The First Fresnel Zone

The definition of the first Fresnel zone is a straightforward byproduct from the Fresnel phase term in the Rayleigh–Sommerfeld integral [Eq. (2.5-1)]. The first Fresnel zone can be defined as a solid angle observed by the LEO or, equivalently, as an area in the thin screen containing all points that produce an observed phase at the LEO that differ from the stationary phase value by π rad or less. The second Fresnel zone contains those points that produce an observed phase that differs from the stationary phase value by π to 2π rad, and so on for higher-order zones. The second- and higher-order zones typically are simply connected and roughly concentric around the first zone. The idea is that the second- and higher-order zones produce radiation at the LEO that largely cancels out through mutual destructive interference, effectively leaving only radiation emanating from the first zone, which more or less arrives in-phase at the LEO. In any case, the first zone contains the stationary phase point as its centroid.

The first Fresnel zone provides a resolution limit for a single observation made by the LEO in a manner similar to the way the Airy disk provides a resolution limit for an optical instrument. Also, the first Fresnel zone provides a threshold index for the presence of diffraction effects. When the transition length of a propagation feature, such as a refraction perturbation, is short compared to the radius of the first Fresnel zone, the feature is called “sharp” and geometric optics tends to become less accurate; diffraction effects usually should be included. On the other hand, if the transition length is long, geometric optics with interference effects from multiple rays may suffice (see Section 2.6 for further discussion).

From the definition given above, we require that the boundary of the first Fresnel zone be defined by the locus of points in h-space that yields the condition on the Fresnel phase function [see Eq. (2.5-1)]

$$\Phi\left(\tilde{h}, h_{\mathrm{LG}}\right) - \Phi\left(h\left(h_{\mathrm{LG}}\right), h_{\mathrm{LG}}\right) = \pi \tag{A-23}$$

where $h\left(h_{\mathrm{LG}}\right)$ is the stationary phase altitude in the thin screen for a given value of h_{LG}. If we expand Eq. (A-23) in a quadratic Taylor series about the stationary phase altitude, we obtain

$$\frac{1}{2}\frac{\partial^2\Phi}{\partial h^2}\bigg|_{\tilde{h}=h}\left(\tilde{h} - h\left(h_{\mathrm{LG}}\right)\right)^2 = \frac{\pi}{\lambda D}\left(1 - D\frac{d\alpha}{dr}\right)_{\tilde{h}=h}\left(\tilde{h} - h\left(h_{\mathrm{LG}}\right)\right)^2 \doteq \pi \tag{A-24}$$

from which it follows that $\tilde{h} = \mathscr{F}\left(h\left(h_{\mathrm{LG}}\right)\right)$. Thus, the "radius" of the first Fresnel zone is given by

$$\mathscr{F} = \sqrt{\zeta\lambda D} \tag{A-25}$$

where ζ is the defocusing factor given by Eq. (A-7) and evaluated at the stationary phase point determined from $h = h\left(h_{\mathrm{LG}}\right)$. It follows from this discussion that the first Fresnel zone is an ellipse with a semi-major axis length of $\sqrt{\lambda D}$ and a semi-minor (when $|\zeta| < 1$) axis length of $(\zeta\lambda D)^{1/2}$.

A.6 The Abel Transform

The Abel integral transform has been a data analysis tool in seismology for nearly a century. Under the assumption of spherical symmetry, it is used to recover the radial profile of seismic velocity from propagation delay data taken from seismometer arrays. The first suggestion of its use in a radio occultation is found in [11].

From the Doppler observations of the occulted GPS satellite and the POD information, one obtains the excess Doppler and the relative coordinates and velocities of the LEO and GPS satellites. One converts this information into a bending-angle profile $\alpha(a)$, that is, at each observational epoch one obtains a determined bending-angle value $\hat{\alpha}$ and an impact parameter value $\hat{a}$, per our discussion earlier on excess Doppler, specifically Eqs. (A-8) through (A-16). One uses this observed profile to form the integral

$$I(a) = \int_a^\infty \frac{\alpha(\xi)}{\sqrt{\xi^2 - a^2}}\, d\xi \tag{A-26}$$

where $\alpha(\xi)$ is the profile of the bending angle given from the observation sequence $\left(\hat{\alpha}_\kappa, \hat{a}_\kappa\right)$, $\kappa = 1, 2, \cdots$. Then, using Eq. (2.2-2) for $\alpha(\xi)$ and assuming that $N(\infty) = 0$, it is easily shown that

$$I(a) = \int_a^\infty \left(-2\xi \int_\xi^\infty \frac{1}{n}\frac{dn}{d\rho} \frac{d\rho}{\sqrt{(\xi^2 - a^2)(\rho^2 - \xi^2)}} \right) d\xi$$
$$= -2\int_a^\infty \frac{1}{n}\frac{dn}{d\rho} d\rho \left(\int_a^\rho \frac{\xi d\xi}{\sqrt{(\xi^2 - a^2)(\rho^2 - \xi^2)}} \right) = \pi \log[n(a)] \doteq \pi N(a) \quad \text{(A-27)}$$

Thus, $N(a)$ from Eq. (A-27) and $\alpha(a)$ from Eq. (2.2-2′) essentially form an Abel Transform pair.

In practice with actual data, α_κ approaches zero as a_κ increases. Therefore, the measured value of α_κ becomes increasingly dominated by measurement noise. Usually the measured profile $\hat{\alpha}_\kappa$ is assisted by use of an a priori atmospheric model at thinning altitudes, for example, above 30 to 40 km.

A.7 Sensitivity of the Recovered Refractivity to an Error Source

Since the Abel transform is a linear operator, it is straightforward to use it to evaluate the sensitivity of the recovered refractivity to an error in the observations from essentially any source, provided that it can be expressed in terms of an equivalent error in the bending-angle profile, and provided that it is indeed sufficiently small that the reference ray path is minimally perturbed. It follows from Eq. (A-27) that the error in the recovered refractivity profile $\delta\hat{N}(a)$ is given in terms of the error profile in bending angle $\delta\alpha(a)$ by

$$\delta\hat{N}(a) = \frac{1}{\pi}\int_a^\infty \frac{\delta\alpha(\xi)}{\sqrt{\xi^2 - a^2}} d\xi \quad \text{(A-28)}$$

This expression is valid for a thin medium. However, in general the perturbation in $\alpha(a)$ results in a different ray path being followed, limiting the accuracy of Eq. (A-28). This expression is used in Appendix F on the effect of cycle slips on the recovered refractivity profile.

A.8 Perturbation of the Bending-Angle Profile Near a Boundary

To evaluate Eq. (2.2-2), we will need expressions for $\alpha(r_*)$ and for its sensitivity to changes in the parameters that describe the refractivity profile. Here the sensitivities of $\alpha(r_*)$ to a discontinuous change at $r = r_o$ in scale height and in lapse rate are developed. We will obtain expressions for the change in bending angle while holding the point of tangency of the ray r_* fixed

at $r_* = r_o$. For the isothermal case, when H_{p_o} remains constant throughout the atmosphere, Eq. (2.2-2) becomes

$$\alpha(r_o) = \frac{2aN_o}{H_{p_o}} \int_{r_o}^{\infty} \exp\left(\frac{r_o - r}{H_{p_o}}\right) \frac{dr}{n\sqrt{n^2 r^2 - a^2}}, \quad a = r_o n(r_o) \tag{A-29}$$

If one makes certain approximations in Eq. (A-29) that are applicable to a thin exponential atmosphere [12], it can be shown that $\alpha(r_o)$ is given by

$$\alpha(r_o) = \sqrt{\frac{2\pi r_o}{H_{p_o}}} N(r_o)\left((\sqrt{2} - 1)\beta_o + 0.28\beta_o^2 + \cdots\right) \tag{A-30}$$

where $\beta_o = |n' r / n|_o = N_o r_o / H_{p_o}$ is the ratio of the impact parameter of the ray to its radius of curvature. Equation (A-30) is fairly accurate for a thin exponentially distributed atmosphere. The error is less than 1 percent for dry air at sea level. It takes into account bending terms along the ray path to second order in β_o. Equation (A-30) begins to break down when β_o exceeds roughly 0.5. By differentiating Eq. (A-30), one obtains

$$-\frac{H_{p_o}}{\alpha} \frac{\partial \alpha}{\partial H_{p_o}} \hat{=} -K_H = \frac{1}{2} + 0.41\beta_o + 0.39\beta_o^2 + \cdots \tag{A-31}$$

Equation (A-31) also can be obtained by explicitly differentiating Eq. (A-29) with respect to H_{p_o} and carrying out the required integrations. For example, if one ignores the first- and higher-order terms in β_o, differentiating Eq. (A-29) yields

$$\frac{\partial \alpha}{\partial H_{p_o}} = -\frac{\alpha}{H_{p_o}} + \frac{2r_o N_o}{H_{p_o}^3} \int_{r_o}^{\infty} \exp\left(\frac{r_o - r}{H_{p_o}}\right) \frac{(r - r_o)dr}{n\sqrt{n^2 r^2 - a^2}} + \mathrm{O}[\beta_o] \frac{\alpha}{H_{p_o}} \tag{A-32}$$

or

$$\frac{\partial \alpha}{\partial H_{p_o}} \doteq -\frac{\alpha}{H_{p_o}} + \frac{1}{H_{p_o}} \sqrt{\frac{2r_o}{H_{p_o}}} N(r_o) \int_{r_o}^{\infty} \sqrt{\frac{(r - r_o)}{H_{p_o}}} \exp\left(\frac{r_o - r}{H_{p_o}}\right) \frac{dr}{H_{p_o}} = -\frac{\alpha}{2H_{p_o}} \tag{A-33}$$

At $r_o - R_{\mathrm{E}} = +10\ \mathrm{km}$, with $H_{p_o} = 7\ \mathrm{km}$, and $N_o = N(r_o) = 70 \times 10^{-6}$, Eq. (A-31) gives $K_H \cong -0.53$.

Next, an expression for α is obtained that is an explicit function of the lapse rate, which is assumed to be a constant above $r = r_o$. Let that function be given by $\alpha[\gamma, r]$. We will need to evaluate

$$\frac{\partial \alpha}{\partial \gamma} = -2a \int_{r_o}^{\infty} \frac{\partial}{\partial \gamma} \left[\frac{n'}{n\sqrt{n^2 r^2 - a^2}} \right] dr \tag{A-34}$$

when the refractivity profile has an explicit dependency on the lapse rate. (Here the refractivity at r_o is assumed to be invariant to a change in γ.) From the assumption of hydrostatic equilibrium and use of the perfect gas law, one has

$$\left. \begin{aligned} -\frac{1}{\rho}\frac{d\rho}{dr} &= H_\rho^{-1} = H_p^{-1} + \gamma \\ H_p^{-1} &= \frac{\mu g}{kT}, \quad \gamma = \frac{1}{T}\frac{dT}{dr} \end{aligned} \right\} \tag{A-35}$$

Here ρ is the mass density of the atmosphere. Let the temperature vary linearly with altitude and be given by

$$T = T_o\left[1 + \gamma\left(r - r_o\right)\right] \tag{A-36}$$

and γ is assumed to be constant above $r = r_o$. Inserting Eq. (A-36) into Eq.(A-35) and integrating upward from $r = r_o$, one obtains for the density

$$\left. \begin{aligned} \log\left(\frac{\rho}{\rho_o}\right) &= -\frac{1}{\gamma H_{\rho_o}} \log\left[1 + \gamma\left(r - r_o\right)\right] \\ H_{\rho_o}^{-1} &= H_{p_o}^{-1} + \gamma \end{aligned} \right\} \tag{A-37}$$

where ρ_o is the density at $r = r_o$. Now, the refractivity at tropopause altitudes and higher is given by

$$N(r) = N_o \frac{\rho(r)}{\rho_o} \tag{A-38}$$

This form for $N(r)$ is used in Eq. (2.2-2), which is expanded to second order in γ. Then, certain approximations are made that are consistent with capturing at least all of the second-order terms in β_o. Therefore, this approach must take into account bending along the ray path of integration, which was also included in the derivation of Eq. (A-30). One replaces the integration variable r with arc

length s in Eq. (2.2-2) and uses the ray path curvature relationship,[4] expanded to second order in β_o. It can be shown that the integration yields

$$\left.\begin{aligned}\alpha\left(r_o,\gamma,H_{p_o}\right) = \sqrt{\frac{2\pi r_o}{H_{p_o}}}N_o\Bigg[1+0.41\beta_o+\frac{3}{8}\left(\gamma H_{p_o}\right)\\ +0.28\beta_o^2+0.37\beta_o\left(\gamma H_{p_o}\right)-\frac{7}{128}\left(\gamma H_{p_o}\right)^2+\cdots\Bigg]\end{aligned}\right\} \quad \text{(A-39)}$$

which, with Eqs. (A-37) and (A-38), provides a vertical profile of the bending angle for a constant lapse rate in a dry and spherical symmetric atmosphere that is in hydrostatic equilibrium.

By differentiating Eq. (A-39) with respect to γ, one obtains

$$\begin{aligned}\frac{1}{\alpha H_p}\frac{\partial\alpha}{\partial\gamma}\bigg|_{r=r_o} &\hat{=} K_\gamma\\ &= \frac{3}{8}+\left(\frac{31\sqrt{2}-32}{32}\right)\beta_o-\frac{1}{4}\left(\gamma H_{p_o}\right)+\ \mathrm{O}\left[\beta_o,\left(\gamma H_{p_o}\right)\right]^2\end{aligned} \quad \text{(A-40)}$$

By differentiating Eq. (A-39) with respect to H_{p_o}, one obtains an improved value for K_H that accounts for a non-zero lapse rate,

$$\frac{H_p}{\alpha}\frac{\partial\alpha}{\partial H_p}\bigg|_{r_o} = K_H = -\frac{1}{2}-0.41\beta_o+\frac{3}{8}\gamma H_{p_o}+\ \mathrm{O}\left[\beta_o,\gamma H_{p_o}\right]^2 \quad \text{(A-41)}$$

At $r_o - R = 10$ km, one obtains

$$\frac{1}{\alpha H_{p_o}}\frac{\partial\alpha}{\partial\gamma}\bigg|_{\gamma=0} = 0.40 \quad \text{(A-42)}$$

Although a change in lapse rate at $r = r_o$ does produce an "immediate" change in H_{ρ_o} and, therefore, in the gradient of the refractivity, the consequent change

---

[4] Let $z(x,r_*)$ be the vertical deflection of the ray path from its horizontal tangent line at the turning point. For an exponential refractivity profile, it is given by

$$z(x,r_*) = (\pi)^{1/2}\beta(r_*)H_{p_o}\left(|x|\,\mathrm{erf}(|x|)-\left(1-\exp(-x^2)\right)/(\pi)^{1/2}\right)+\ \mathrm{O}\left[\beta^2 H_{p_o}\right]$$

where x is the normalized arc length from the tangent point, $x = s/(2r_o H_{p_o})^{1/2}$.

in the temperature profile above $r = r_o$ induces a change in the density scale height profile in this regime that has a polarity opposite to that of the incremental change in H_{p_o}. This partially offsets the effect of the initial change in the refractivity gradient at $r = r_o$. Hence, one would expect the smaller value (0.4) for K_γ (for $\gamma = 0$) relative to the value (0.53) for K_H.

A.9 Bending-Angle Perturbation

Approximate expressions for the integrals in Eq. (2.3-20) valid through the zeroth order in β (no ray path bending corrections applied) can be readily obtained through simplifications of the integrands. More accurate expressions valid to higher powers in β and γ also can be obtained, but these calculations, like those carried out in the case of the derivations of Eqs. (A-30) and (A-39), are more tedious. The zeroth-order expressions have been compared with a numerical integration of Eq. (2.3-20) for $\gamma = 0$; good agreement has been obtained for "thin atmosphere" conditions. Here one represents the ray path along which the integrations in Eq. (2.3-20) follow by the straight line approximation:

$$r = r_o + \frac{s^2}{2r_o} \hat{=} r_o + x^2 H_{p_o} \tag{A-43}$$

where s is arc length measured from the point of tangency. The radial integration variable *r* in Eq. (2.3-20) is replaced by *s*, and all terms of degree 1 and higher in β are dropped. This is carried out for Cases B and C.

With the simplifications in ray path representation in place, the "+" and "–" integrals in Eq. (2.3-20) each have the form

$$\begin{aligned} I &= -2a^- \int_{r_o}^{\infty} \frac{1}{n}\frac{dn}{dr'} \frac{dr'}{\sqrt{(nr')^2 - (a^-)^2}} \\ &= \sqrt{\frac{8r_o}{H_{p_o}}} N(r_o)\left[1 + (\gamma H_{p_o})\right] \int_0^{\infty} \exp(-x^2)\left[1 - (\gamma H_{p_o})\left(2x^2 - \frac{1}{2}x^4\right) + \mathrm{O}\left[(\gamma H_{p_o})^2\right]\right] \frac{x\,dx}{\sqrt{\sigma^2 + x^2}} \\ &= \sqrt{\frac{2\pi r_o}{H_{p_o}}} N(r_o)\left[F(\sigma) + (\gamma H_{p_o})\left\{\left(K_\gamma + \frac{3\sigma^2}{2} + \frac{\sigma^4}{2}\right)F(\sigma) - \frac{1}{\sqrt{\pi}}\left(\frac{5\sigma}{4} + \frac{\sigma^3}{2}\right)\right\} + \mathrm{O}\left[(\gamma H_{p_o})^2\right]\right] \end{aligned} \tag{A-44}$$

where $F(\sigma)$ is defined by Eq. (2.3-28) and σ is given by Eq. (2.3-26).

For Case B with $\gamma = 0$, Eq. (2.3-20) becomes

$$\alpha(r) \doteq \begin{cases} \alpha^-(r) + \alpha^+(r_o)F(\sigma^+) - \alpha^-(r_o)F(\sigma^-), & r \le r_o \\ \alpha^+(r), & r > r_o \end{cases} \tag{A-45}$$

where $\sigma^\pm = (h_o - h)/H_{p_o}^\pm$ and $\alpha^\pm(r)$ is given by Eq. (A-25). By expanding Eq. (A-45) in a Taylor series in powers of $[H_{p_o}^+ - H_{p_o}^-]$ and using Eq. (A-39) for $\alpha^+(r_o) - \alpha^-(r_o)$, one obtains Eq. (2.3-30).

For Case C, the quantity H_{p_o} is held fixed and γ is varied. Equation (2.3-31) follows directly from inspecting the coefficient of the (γH_{p_o}) term in Eq. (A-44). Differential expressions for these cases also can be developed that account for the effects of ray path bending. A first-order treatment is given below for Case A.

A.10 Bending-Angle Perturbation by a Discontinuous Refractivity, Case A

Obtaining the change in bending angle when n is discontinuous across a boundary in an otherwise smoothly refracting medium follows a course similar to that followed for Cases B and C. However, in Case A, one also has to account for the discontinuity in the impact parameter a when the ray contacts the boundary at $r = r_o$. This results in the σ_+ term in Eq. (2.3-29). Here one has to account for effects of the different impact parameters that arise in evaluating Eq. (2.3-20): $a^\pm = rn^\pm(r)$.

Both the continuous and discrete form of Snell's law must be applied in the neighborhood across a boundary marking a discontinuity in the index of refraction. Also, accounting for ray path bending will improve the accuracy of the analytic expressions for the overall refractive bending. Ray path bending will alter the intersection point of the ray with the boundary relative to the intersection point that holds for the straight-line assumption. Therefore, both the direction of the local normal to the boundary surface and the direction of the impinging ray are altered by ray path bending. Most of the first-order ray path bending can be accounted for by expanding $z[x, r_o]$, given in Footnote 4, in powers of s. This yields $z = \beta s^2 / 2r_o + \mathrm{O}[s^3]$, where z is the departure of the ray from a straight-line path. Accordingly, the amended version of Eq. (A-43) that accounts for first-order bending (most of it) is given by

$$r = r_o + x^2(1 - \beta_o)H_{p_o} \tag{A-46}$$

From this it is easily shown that the change $\Delta\theta$ in the direction of the local normal at the amended intersection point is given to first order by

$$\Delta\theta \doteq \sqrt{\frac{r_o - r}{2r_o}}\,\frac{\beta_o}{(1-\beta_o)^{3/2}} \tag{A-47}$$

Also the change in direction of the ray at the intersection point due to ray path bending can be shown to be just $2\Delta\theta$. Thus, the change in angle of incidence on the interior side of the boundary is $\Delta\theta$. The change in bending angle across the boundary due to the discontinuity in refractivity, including first-order ray path bending within the interior of the sphere, is given from Bouguer's law by

$$\Delta\alpha_{\mathrm{S}}(r) = \sin^{-1}\left(\frac{rn^-}{r_o n^+}\right) - \sin^{-1}\left(\frac{r}{r_o}\right) - \frac{\Delta N}{n^-}\Delta\theta, \quad r \le \min\left[r_o, r_o\frac{n^+}{n^-}\right] \tag{A-48}$$

However, $\Delta\theta$ is of the order of a milliradian for the ray tangency point located within a few kilometers of the boundary, and also ΔN is small. Consequently, the last term in Eq. (A-48), which is due to ray path bending, can be ignored for the discrete part of the bending angle.

For the continuous or integral portion of the bending angle, we follow a line similar to that followed for Cases B and C except that we take into account ray path bending. Equation (2.2-2) is a *path* integral, that is, it is assumed implicitly that the integrand is evaluated along the ray path [see Eq. (A-5)]. This means that, when one makes a change of variable from the radial coordinate to path length, one must explicitly include the curvature of the ray path in that relationship. Using Eq.(A-46) to account for the effect of ray path bending, one obtains for the overall change in bending, including the discrete component,

$$\alpha(r) \doteq \begin{cases} \alpha^+(r)e^{-\kappa}\,\mathrm{erfc}\left(\sqrt{\sigma^2-\kappa}\right) - \alpha^-(r)\,\mathrm{erf}(\sigma) + \Delta\alpha_{\mathrm{S}}, & r \le \tilde{r} \le r_o \\ \alpha^+(r), & r > r_o \end{cases} \tag{A-49}$$

Here $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, $\kappa = -r_o\Delta N/(H_{p_o}(1-\beta_o))$, and $\tilde{r}$ is the critical reflection radius. Expanding Eq.(A-49) to first order in ΔN and setting $\beta_o = 0$ yields Eq. (2.3-27).

Finally, when the ray lies only in the "+" regime,

$$\alpha(r) = \alpha^+(r), \quad r > r_2 \tag{A-50}$$

A.11 The Fresnel Transform Pair

The Fresnel transform pair is a special case of the two-dimensional scalar diffraction formula given in Eq. (A-22) that is obtained when the Fresnel approximation is used and the integration is along a straight line. It should be

clear that reciprocity applies to the two forms of the scalar diffraction integral given in Eq. (A-22); we can use the top signs to integrate $E(\boldsymbol{r}_2)$ along the curve C to obtain $E(\boldsymbol{r}_1)$, or we can use the bottom signs to integrate $E(\boldsymbol{r}_1)$ along another curve to get $E(\boldsymbol{r}_2)$. Therefore, these two forms are the inverses of each other.

Similarly, from Eq. (2.10-3), we have for the Rayleigh–Sommerfeld integral, expressed in Fresnel variables, $u = h(2/\lambda D)^{1/2}$ and $u_{\mathrm{LG}} = h_{\mathrm{LG}}(2/\lambda D)^{1/2}$,

$$E(u_{\mathrm{LG}})\exp\big(i\psi(u_{\mathrm{LG}})\big) = \frac{1}{1+i}\int_{-\infty}^{\infty} A(u)\exp\left(i\left(\varphi(u) + \frac{\pi}{2}(u-u_{\mathrm{LG}})^2\right)\right)du \qquad \text{(A-51)}$$

We multiply the left-hand side (LHS) of Eq. (A-51) by $\exp(-i\pi(\tilde{u}-u_{\mathrm{LG}})^2/2)$ and integrate on u_{LG} to obtain

$$\begin{aligned}
&\int_{-\infty}^{\infty} E(u_{\mathrm{LG}})e^{i\left(\psi(u_{\mathrm{LG}}) - \frac{\pi}{2}(\tilde{u}-u_{\mathrm{LG}})^2\right)}du_{\mathrm{LG}} \\
&= \frac{1}{1+i}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(u)e^{i\left(\varphi(u) + \frac{\pi}{2}(u-u_{\mathrm{LG}})^2 - \frac{\pi}{2}(\tilde{u}-u_{\mathrm{LG}})^2\right)}du\,du_{\mathrm{LG}} \\
&= \frac{1}{1+i}\int_{-\infty}^{\infty}\left(A(u)e^{i\left(\varphi(u) + \frac{\pi}{2}(u^2-\tilde{u}^2)\right)}\int_{-\infty}^{\infty} e^{i\pi(\tilde{u}-u)u_{\mathrm{LG}}}du_{\mathrm{LG}}\right)du \\
&= \frac{2}{1+i}\int_{-\infty}^{\infty} A(u)e^{i\left(\varphi(u) + \frac{\pi}{2}(u^2-\tilde{u}^2)\right)}\delta(\tilde{u}-u)du = \frac{2}{1+i}A(\tilde{u})e^{\varphi(\tilde{u})}
\end{aligned} \qquad \text{(A-52)}$$

It follows that

$$A(u)\exp\big(i\varphi(u)\big) = \frac{1}{1-i}\int_{-\infty}^{\infty} E(u_{\mathrm{LG}})\exp\left(i\left(\psi(u_{\mathrm{LG}}) - \frac{\pi}{2}(u-u_{\mathrm{LG}})^2\right)\right)du_{\mathrm{LG}} \qquad \text{(A-53)}$$

which is Eq.(2.10-4). The Dirac delta function,

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(ixy)dy \qquad \text{(A-54)}$$

has been used in Eq. (A-52).

A.12 Ray Path Phase Delay

To compare wave theory and wave/optics results, it is helpful to have an expression for the phase delay along a ray. For a spherical symmetric

atmosphere, it follows from Eq. (A-1), Bouguer's law in Eq. (A-4a), and the bending-angle gradient in Eq. (A-5) that the phase delay between the emitting GPS satellite and the LEO is given by

$$\begin{aligned}\varphi &= k\int nds = k\int_{r_*}^{r_L} n\sqrt{1+(r\theta')^2}\,dr + k\int_{r_*}^{r_G} n\sqrt{1+(r\theta')^2}\,dr \\ &= k\int_{r_*}^{r_L}\frac{n^2 r dr}{\sqrt{n^2r^2-a^2}} + k\int_{r_*}^{r_G}\frac{n^2 r dr}{\sqrt{n^2r^2-a^2}} \\ &= \int_{\rho_*}^{\rho_L}\frac{\rho d\rho}{\sqrt{\rho^2-\rho_*^2}} + \int_{\rho_*}^{\rho_G}\frac{\rho d\rho}{\sqrt{\rho^2-\rho_*^2}} - 2\int_{\rho_*}^{\infty}\frac{d\log n}{d\rho}\frac{\rho^2 d\rho}{\sqrt{\rho^2-\rho_*^2}} \\ &= \sqrt{\rho_L^2-\rho_*^2} + \sqrt{\rho_G^2-\rho_*^2} + \rho_*\alpha_L(\rho_*) + \int_{\rho_*}^{\infty}\alpha_L(\omega)d\omega\end{aligned} \tag{A-55}$$

where $\rho = nkr$ and $\rho_* = kn(r_*)r_* = ka$. We have assumed here that $d\log n/d\rho \to 0$ as $\rho \to \infty$.

Referring to Fig. A-3, the sides and angles of triangle OLG are completely determined from the precision orbit determination information, independent of the ray bending. This figure shows the ray asymptotes passing through the satellite positions G and L. These asymptotes are defined by the unit tangent vectors, $\boldsymbol{T}_G$ and $\boldsymbol{T}_L$; their deflection angles, δ_G and δ_L, with $\alpha = \delta_G + \delta_L$; and their impact parameter, a, or ρ_* in phase units. The value of the impact parameter is the same for both asymptotes when spherical symmetry applies. Therefore, the phase delay in Eq. (A-55) can be written as

$$\left.\begin{aligned}\varphi &= \rho_G\cos(\chi_G+\delta_G) + \rho_L\cos(\chi_L+\delta_L) + \rho_*\alpha_L + \int_{\rho_*}^{\infty}\alpha_L(\omega)d\omega \\ \alpha_L &= \delta_L + \delta_G\end{aligned}\right\} \tag{A-56}$$

Here χ_G and χ_L are strictly orbit-related angles. When $d\log n/d\rho \equiv 0$, $\delta_G = \delta_L = \alpha_L \equiv 0$ and $\rho_* \to \rho_L\sin\chi_L = \rho_G\sin\chi_G$. The phase delay is just $\rho_G\cos\chi_G + \rho_L\cos\chi_L = \rho_{LG}$. Thus, $\varphi - \rho_{LG}$ is the excess phase delay from the refracting atmosphere.

For the special case of when the emitting GPS satellite is placed at infinity with $\theta_G = \pi$, then from Fig. A-3 it follows that $\delta_G + \chi_G \to 0$, $\chi_L + \delta_L \to \theta_L + \alpha_L$, and $\rho_{LG} \to \rho_G + \rho_L\cos\theta_L$. The phase relative to the line $\theta = \pi/2$ becomes

$$\varphi = \rho_L\cos(\theta_L+\alpha_L) + \rho_*\alpha_L + \int_{\rho_*}^{\infty}\alpha_L(\omega)d\omega \tag{A-57}$$

References

[1] J. Mathews and R. Walker, *Mathematical Methods of Physics*, Menlo Park, California: The Benjamin/Cummings Publishing Co., 1970.

[2] C. Cartheodory, *Calculus of Variations and Partial Differential Equations of the First Order*, Part 2: Calculus of Variation, San Francisco: Holden-Day, Inc., 1967.

[3] L. Pars, *An Introduction to the Calculus of Variations*, New York: John Wiley, 1962.

[4] T. D. Moyer, *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation*, New York: John Wiley, 2003.

[5] I. Gelfand and S. Fomin, *Calculus of Variations*, Englewood Cliffs, New Jersey: Prentice–Hall, 1963.

[6] M. E. Gorbunov, "Canonical Transform Method for Processing Radio Occultation Data in the Lower Troposphere," *Radio Science,* vol. 37, no. 5, 1076, doi:10.1029/2000RS002592, 2002.

[7] J. Goodman, *Introduction to Fourier Optics*, New York: McGraw-Hill, 1968.

[8] M. Born and E. Wolf, *Principles of Optics,* 6th ed., Oxford, United Kingdom: Pergamon Press, 1980.

[9] E. Karayel and D. Hinson, "Sub-Fresnel Scale Vertical Resolution in Atmospheric Profiles from Radio Occultation," *Radio Science*, vol. 32, no. 2, pp. 411–423, 1997.

[10] M. E. Gorbunov, A. S. Gurvich, and L. Bengtsson, *Advanced Algorithms of Inversion of GPS/MET Satellite Data and Their Application to Reconstruction of Temperature and Humidity*, Report 211, Max Planck Institute for Meteorology, Hamburg, Germany, 1996.

[11] R. Phinney and D. Anderson, "On the Radio Occultation Method for Studying Planetary Atmospheres," *Journal of Geophysical Research*, vol. 73, no. 5, pp. 1819–1827, 1968.

[12] W. G. Melbourne, E. S. Davis, C. B. Duncan, G. A. Hajj, K. R. Hardy, E. R. Kursinski, T. K. Meehan, L. E. Young, and T. P. Yunck, *The Application of Spaceborne GPS to Atmospheric Limb Sounding and Global Change Monitoring*, JPL Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, April 1994.

Appendix B
Caustic Surfaces

A caustic surface or "burning curve" in geometric optics is a boundary separating accessible and inaccessible regions for a given family of light rays. The rays within that defined family can "pile up" against the boundary but they never cross it. The boundary, therefore, is an envelope with respect to that family. Caustics are very common in everyday life—for example, the double crescent-shaped cusp of reflected light from a point source at the bottom of a coffee cup or inside a wedding band.

Another caustic phenomenon is the rainbow [1], where for a given color the scattering angle from the raindrop assumes a stationary value with respect to the impact parameter or distance that the incident ray makes relative to the center of the raindrop (at a local minimum of about 138 deg for the primary and a local maximum of about 130 deg for the secondary). Since light rays are equally likely to impinge on a raindrop at any impact distance, stationarity of the scattering angle (which occurs for the primary at an impact distance of about 80 percent of the radius of the raindrop) means that the light rays "pile up" at that exit angle. And because their scattering angle and, therefore, their phase at exit from the raindrop are stationary with respect to the impact parameter at this angle, they all are essentially in phase and reinforce each other in the vicinity of this stationary point upon arrival at the observer. This results in the "caustic" phenomenon. Light rays from a raindrop at scattering angles slightly less than the local maximum of the primary rainbow can have impact distances that are slightly higher and lower than the impact distance that provides the stationary scattering angle. Because these higher and lower rays travel through the raindrop along slightly different paths, their travel times to the observer differ, which can result in their arriving at the observer both in and out of phase; both constructive and destructive interference can result. These are the supernumerary bands observed with some rainbows. Supernumerary

bands in rainbows are the analog to the amplitude and phase variability that we observe in a radio occultation signal near a caustic.

The stationarity property of the contact point of a ray with its caustic surface defines the mathematical form of the caustic surface. A caustic surface is an envelope generated by a family of curves, which in the case of geometric optics are themselves stationary phase paths. For convenience, we use here a thin-screen framework to develop a mathematical description of a caustic surface. Let $y = f(x,h)$ describe a one-parameter family of ray paths, where y is the altitude of a point on the ray path, x is its perpendicular distance from the thin screen, and h is the parameter whose value identifies the family member. The boundary conditions are $y = h = f(0,h)$ at the thin screen and $y = h_{\mathrm{LG}} = f(D,h)$ at the low Earth orbiting (LEO) satellite. It is straightforward to work out the explicit functional form of $f(x,h)$ from the thin-screen relationship in Eq. (2.2-5). Thus,

$$y = f(x,h) = h - x\alpha\big(a(h)\big) \tag{B-1}$$

where in this model $a(h)$ provides the relationship between impact parameter and thin-screen altitude used to generate the bending angle α.

As we vary h, we generate the family of rays that satisfies the boundary conditions at $x = 0$ and at $x = D$. An envelope, if one exists, is defined by the condition that, at its contact point with a ray, it must be tangent to that member of the family. Also, there must be a continuum of contact points over at least a subset of the family of rays. Let the functional form defining the envelope be given by $y = g(x)$, and let the contact point be designated by $\left(x^{\dagger}, y^{\dagger}\right)$, which is an implicit function of the parameter h. At a contact point, we require that

$$y^{\dagger} = g\left(x^{\dagger}\right) = f(x,h)\big|_{x=x^{\dagger}} \tag{B-2}$$

As the contact point varies due to varying h, the first-order variations of g and f with respect to h are given by

$$\frac{dy^{\dagger}}{dh} = \frac{dg}{dx^{\dagger}}\frac{dx^{\dagger}}{dh} = \left(\frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial h}\right)_{x=x^{\dagger}} \tag{B-3}$$

The tangency condition on the ray and the envelope at the contact point requires that $dg/dx^{\dagger} = (\partial f/\partial x)_{x=x^{\dagger}}$. It follows that the position y of the ray member must be stationary with respect to the parameter h at the contact point. That is,

$$\left.\frac{\partial f}{\partial h}\right|_{x=x^{\dagger}} = 0 \tag{B-4}$$

These contact and stationarity conditions in Eqs. (B-2) and (B-4) enable one to solve for the functional form of the envelope, $y = g(x)$. The stationarity condition in Eq. (B-4) implies that $\partial y/\partial h = 0$ in Eq. (B-1), which is equivalent to letting the defocusing factor $\zeta \to \infty$, which in turn is equivalent to setting both the first and second partial derivatives of the Fresnel phase function [Eq. (2.5-1)] with respect to h to zero. At the contact point with the envelope, one obtains from Eq. (B-1)

$$x^{\dagger} = \left(\frac{d\alpha}{dh}\right)^{-1} \tag{B-5}$$

From this relation, one obtains $h^{\dagger} = h\left(x^{\dagger}\right)$ through the definition of $\alpha(a(h))$, for example, by Eqs. (A-39) and (A-40) for Case C. It follows that

$$y^{\dagger} = f\left(x^{\dagger}, h^{\dagger}\right) = g\left(x^{\dagger}\right) = h\left(x^{\dagger}\right) - x^{\dagger}\alpha\left(a\left(h\left(x^{\dagger}\right)\right)\right) \tag{B-6}$$

which is the functional form for the envelope.

The caustic for Case C is shown (not to scale) as the limiting concave arc in Fig. B-1, which shows the local families of the ordinary (b) rays and the anomalous (a) rays. These are the "–" rays that are applicable below h_o. In this figure, h_o is the altitude where the discontinuity in the lapse rate occurs [see also Fig. 2-7(c)]. The LEO plane is on the left side of Fig. B-1; the thin screen is on the right side. The (a) family, whose rays begin from the thin screen in the altitude range $h_o \geq h_a \geq h(2)$, generates the envelope. These are the so-called anomalous rays. The ordinary family of "–" rays, the (b) rays, which originates

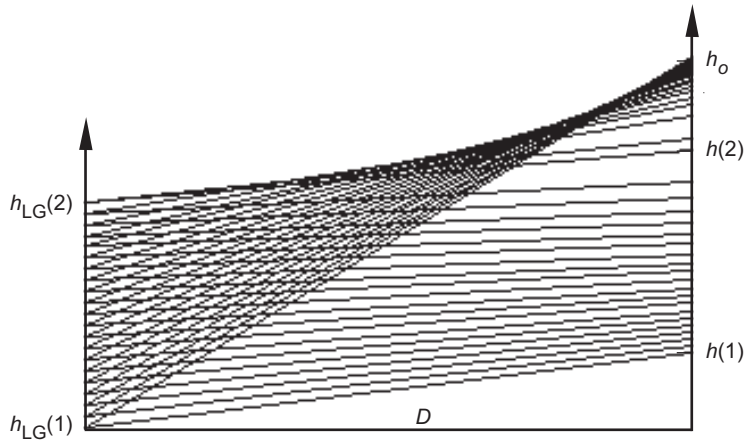

Fig. B-1. Limiting arc is the caustic surface generated by a family of anomalous rays.

from the thin screen in the altitude range $h(2) \geq h_b \geq h(1)$, does not generate an envelope. Figure B-1 shows that one ray from each of these families intercepts the h_{LG} plane at the same point when $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$. To improve the clarity in the figure, the local family of "+" rays originating from thin-screen altitudes above h_o is not shown. These are the (m) rays, the original ray system before impacting the discontinuity. Adding these (m) rays to the figure would produce a triplet system of rays arriving at the LEO when the altitude of the Global Positioning System (GPS)–LEO line is in the range $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$.

The first caustic contact is at $h^{\dagger} = h(2)$ with the GPS–LEO line at $h_{\mathrm{LG}} = h_{\mathrm{LG}}(2)$. This marks the birth of the ray systems (a) and (b). As the altitude of the GPS–LEO line migrates downward in the range $h_{\mathrm{LG}}(1) \leq h_{\mathrm{LG}} \leq h_{\mathrm{LG}}(2)$, the altitude of the (a) ray in the thin screen migrates upward in the range $h_o \geq h_a \geq h(2)$. The defocusing factor for the (a) ray is negative; hence, the appellation "anomalous." Concurrently, the altitude of the (b) ray migrates downward in the range $h(2) \geq h_b \geq h(1)$; its defocusing factor is positive. When the altitude of the GPS–LEO line drops below $h_{\mathrm{LG}}(1)$, only the (b) ray survives.

If a ray contacts the caustic surface along a tangential between the thin screen and the LEO, which every member of the anomalous branch does, then the Calculus of Variations tells us that this interior point of contact is a conjugate point. A ray with a conjugate point, although a path of stationary phase, is not necessarily a path along which the path delay is a local minimum; it can be, and in this case it is a local maximum, as Fig. 2-9(c) shows.[1] The anomalous branch violates the Jacobi condition from the Calculus of Variations, which requires that the ray have no interior contact point with its envelope. The Jacobi condition is an additional necessary condition that the ray path must satisfy to yield a local minimum in elapsed signal time, or phase delay. Caustics are well known in seismology [2].

---

[1] A simple example of a conjugate point is taken from the system of geodesics on a spherical surface. Here the stationary phase path is a great circle. Consider the family of great circle paths, all originating from the north pole and generated by varying the longitude parameter, for example, the longitude of the path when it makes its southward crossing of the equator. The distance from the north pole to any geographical location on the sphere is a global minimum along a great circle route provided that the path does not first pass through the south pole. If it does, there is an alternate family member, the great circle path with its longitude parameter 180 deg different, that provides a shorter distance to the same geographical location. The south pole is a conjugate surface for this family of geodesics originating from the north pole. Unfortunately, the conjugate surface in this example is somewhat pathological, shrinking to a single point at the South Pole.

References

[1] H. M. Nussenzveig, "The Theory of the Rainbow," *Scientific American*, vol. 236, no. 4, pp. 116–127, 1977.

[2] K. Bullen and B. Bolt, *An Introduction to the Theory of Seismology*, Cambridge, United Kingdom: Cambridge University Press, 1993.

Appendix C
Multiple Ray Path Separation Altitudes

We develop the expressions for the differences in altitudes of the ray path tangency points between the multiple rays generated by the Case A, Case B, and Case C scenarios. We begin with Case C.

C.1 Case C: A Discontinuous Lapse Rate

Figure 2-7(d) shows the multipath relationship between h and h_{LG}, which results from a discontinuous lapse rate with the appropriate polarity. Differential expressions for the altitude differences between these multiple rays at their tangency points with the Earth's limb can be estimated for Case C from the power series expansion for $\delta\alpha(r_*) = \alpha(r_*) - \alpha^+(r_*) = (\alpha(r_*) - \alpha^-(r_*)) + (\alpha^-(r_*) - \alpha^+(r_*))$. From Eqs. (2.3-21), (2.3-23), and (2.3.31) with $K_\gamma = 3/8$, we have for $r_* \le r_o$

$$\delta\alpha(r) = -\left(\Delta\gamma H_{p_o}\right)\alpha^-(r_*)\left(\frac{2}{\sqrt{\pi}}\sigma - \frac{3}{2}\sigma^2 + \frac{4}{\sqrt{\pi}}\sigma^3 - 2\sigma^4 + \cdots\right) \qquad \text{(C-1)}$$

where σ is given by $\sigma^2 = \left(r_o - r_*\right)/H_{p_o}$. From the thin-screen model given by Eq. (2.2-5), we have

$$h_{\mathrm{LG}} - h_{\mathrm{LG}}(1) \doteq h - h_o - D\left[\alpha(r_*) - \alpha\left(r_o\right) + \delta\alpha(r_*)\right] \qquad \text{(C-2)}$$

We expand $\alpha\left(r_o\right)$ about r_* and define $\xi(r_*)$ by the relation

$$\left.\begin{aligned} \xi(r_*) &= \zeta(r_*)\alpha(r_*)\frac{D}{H_{p_o}}\left(\Delta\gamma H_{p_o}\right) \\ \zeta(r_*) &= \left(1+\frac{D}{H_{p_o}}\alpha(r_*)\right)^{-1} \end{aligned}\right\} \tag{C-3}$$

The quantity ξ has a value of approximately $5\Delta\gamma$ (i.e., small) for nominal conditions at the tropopause. Truncating terms in Eq. (C-1) above a quadratic in σ, one obtains

$$\sigma = \frac{\xi}{\sqrt{\pi}\left(1+3\xi/2\right)} \pm \sqrt{\frac{\xi^2}{\pi\left(1+3\xi/2\right)^2} - \frac{\left[h_{\mathrm{LG}} - h_{\mathrm{LG}}(1)\right]\zeta}{\left(1+3\xi/2\right)H_{p_o}}} + \cdots \tag{C-4}$$

Since $\sigma \geq 0$, it follows that Eq. (C-4) admits solutions only when $\xi > 0$; that is, caustics occur only when $\Delta\gamma > 0$. When $h_{\mathrm{LG}} = h_{\mathrm{LG}}(1)$, the altitude in h_{LG}-space of maximum ray path separation, we obtain from Eq. (C-4) the two roots

$$\left.\begin{aligned} \sigma &= 0 \\ \sigma &= \frac{2\xi}{\sqrt{\pi}\left(1+3\xi/2\right)} \end{aligned}\right\} \tag{C-5}$$

Since ξ is small, truncation in Eq. (C-1) to σ^2 terms is valid, and Eq. (C-4) should be fairly accurate in the neighborhood of the caustic zone. Equation (C-5) yields the maximum altitude separation, which occurs when $h_{\mathrm{LG}} = h_{\mathrm{LG}}(1)$, of

$$\frac{h_o - h^-(1)}{H_{p_o}} = \frac{4\xi^2}{\pi\left(1+3\xi/2\right)^2} + \mathrm{O}\left[\xi^3\right] \tag{C-6}$$

To obtain the spread in Doppler between the corresponding two tones, one would multiply Eq. (C-6) by the factor $\dot{h}_{\mathrm{LG}} H_{p_o} / (\lambda D)$.

When $h_{\mathrm{LG}} = h_{\mathrm{LG}}(2)$, which is where $dh / dh_{\mathrm{LG}} \to \infty$, Eq. (C-4) has but one root, and

$$\frac{h_o - h^-(2)}{H_{p_o}} = \frac{\xi^2}{\pi\left(1+3\xi/2\right)^2} + \mathrm{O}\left[\xi^3\right] \tag{C-7}$$

This marks the first contact point with the caustic surface and the epoch of onset of the multiple rays. The ratio dh / dh_{LG} is the defocusing factor. Thus,

$$\frac{dh^+}{dh_{\mathrm{LG}}} = \zeta \tag{C-8}$$

It follows that the nominal upper altitude at the onset of ray path separation is given by

$$\frac{h^+(2) - h_o}{H_{p_o}} \cong \zeta \frac{h_{\mathrm{LG}}(2) - h_{\mathrm{LG}}(1)}{H_{p_o}} = \frac{\xi^2}{\pi\left(1 + 3\xi/2\right)} \; + \; \mathrm{O}\left[\xi^3\right] \tag{C-9}$$

Adding Eqs. (C-7) and (C-9) yields the separation in altitudes between the nominal upper ray and the lower ray at the onset of ray path separation, and multiplying this sum by $\dot{h}_{\mathrm{LG}} H_{p_o} / (\lambda D)$ yields the Doppler spread. If this Doppler spread is well outside the first Fresnel zone, then distinct tones should appear in a power spectrum (see Fig. 1-11); if not, a diffraction treatment is required. In any case, a diffraction treatment (or at least a third-order ray optics approach) is required in the neighborhood of the point where $dh / dh_{\mathrm{LG}} \to \infty$ (see Appendix D). Equation (2.4-1) also can be applied to obtain the phase offsets between the multiple rays. The ray path separation scale varies essentially quadratically with $\Delta\gamma$.

C.2 Case B: A Discontinuous Scale Height

From Eqs. (2.3-21) and (2.3-30), we obtain $\delta\alpha(r_*) = \alpha(r_*) - \alpha^+(r_*)$ for the case of a discontinuous scale height at $r = r_o$:

$$\delta\alpha(r_*) = \frac{\Delta H_{p_o}}{H_{p_o}} \alpha^-(r_o) \left[\frac{2\sigma}{\sqrt{\pi}} - \sigma^2 + \; \cdots \right], \; r_* \le r_o \tag{C-10}$$

Defining $\eta(r_*)$ by

$$\left.\begin{aligned} \eta(r_*) &= -(1-\zeta)\frac{\Delta H_{p_o}}{H_{p_o}} \\ \zeta(r_*) &= \left(1 + \frac{D}{H_{p_o}} \alpha(r_*)\right)^{-1} \end{aligned}\right\} \tag{C-11}$$

and solving for σ, we obtain

$$\sigma = \frac{\eta}{\sqrt{\pi}(1+\eta)} \pm \sqrt{\frac{\eta^2}{\pi(1+\eta)^2} - \frac{\left(h_{\mathrm{LG}} - h_{\mathrm{LG}}(1)\right)\zeta}{(1+\eta)H_{p_o}}} + \cdots \tag{C-12}$$

from which the relative altitudes of the multiple ray paths may be obtained as in Case C. Since $\sigma > 0$, it follows that ΔH_{p_o} must be negative to obtain a caustic surface. The multipath separation scales quadratically with $\Delta H_{p_o} / H_{p_o}$.

C.3 Case A: A Discontinuous Refractivity

An abrupt discontinuity in an exponentially varying background refractivity, ΔN, can be treated by following an approach similar to that taken for the other cases. It can be shown from Eqs. (2.2-5) and (A-49) that we need the roots of the equation

$$0 = \sigma^2 \zeta^{-1} + \left(\frac{D}{H_{p_o}}\right)\left(\alpha_o^+ + \left(\frac{2}{\sqrt{\pi}}\right)\left(\alpha_o^- \sigma - \alpha_o^+ \sqrt{\sigma^2 - \kappa}\right) + \Delta\alpha_{\mathrm{S}}\right) + \frac{h_{\mathrm{LG}} - h_o}{H_{p_o}} + \mathrm{O}[\sigma^3] \tag{C-13a}$$

If one expands the Snell and radical parts and drops small terms, one obtains an alternate form:

$$\left.\begin{aligned} &\sigma^2 \zeta^{-1} + 2\frac{g}{\sqrt{1-\beta}}\sigma + \left(\frac{D}{H_{p_o}}\right)\alpha_o^+ + \frac{g}{\sigma} + \frac{h_{\mathrm{LG}} - h_o}{H_{p_o}} \doteq 0 \\ &g = -\Delta N_o \left(\frac{D}{H_{p_o}}\right)\left(\frac{2r_o}{H_{p_o}}\right)^{1/2} \end{aligned}\right\} \tag{C-13b}$$

When $dh_{\mathrm{LG}} / dh = 0$, it follows that the condition

$$\sigma^3 \zeta^{-1} + \frac{g\sigma}{(1-\beta)^{1/2}} - \frac{g}{2} \doteq 0 \tag{C-14}$$

defines the value of $\sigma(2)$ and also $h_{\mathrm{LG}}(2)$ and $h(2)$, the first contact point with the caustic. Equation (C-14) has but one root for $g > 0$. That root is given to good approximation by the form

$$\sigma(2) \doteq \left(\frac{g\xi}{2}\right)^{1/3} \tag{C-15}$$

Therefore, $h_o - h(2)$ grows approximately as $\left(-\Delta N_o\right)^{2/3}$ when $\Delta N_o < 0$.

To obtain the maximum separation, we set $h_{\mathrm{LG}} = h_{\mathrm{LG}}(1) = h_o - D\left[\alpha^+(r_o) + \Delta\alpha(r_o)_{\mathrm{S}}\right]$ and solve for the roots in Eq. (C-13a). It can be shown from Eq. (C-13b) that one obtains the approximate expression

$$\sigma^3\zeta^{-1} + 2g\frac{2g}{(1-\beta)^{1/2}}\sigma^2 - 2\sqrt{-\Delta N_o}\,\frac{D}{H_{p_o}}\sigma + g \doteq 0 \qquad \text{(C-16)}$$

which yields the upper root for $\sigma(1)$ and, therefore, an estimate for the maximum separation in altitude, $h_o - h(1)$. Except for small values of $-\Delta N_o / N_o$, a higher-degree expansion is required for accuracy. It is better to use Eqs. (2.2-5) and (A-49) without expanding in a power series.

The separation altitude at onset of multipath (first contact point), the maximum altitude separation (where the main branch terminates), and the width of the shadow zone are shown in Fig. C-1, which is based on a numerical integration of Eq. (2.2-2). The low Earth orbiting (LEO) satellite's orbit radius is $r_{\mathrm{L}} = 1.1 r_o$. Because of the higher ambient refractivity and lower altitude, this

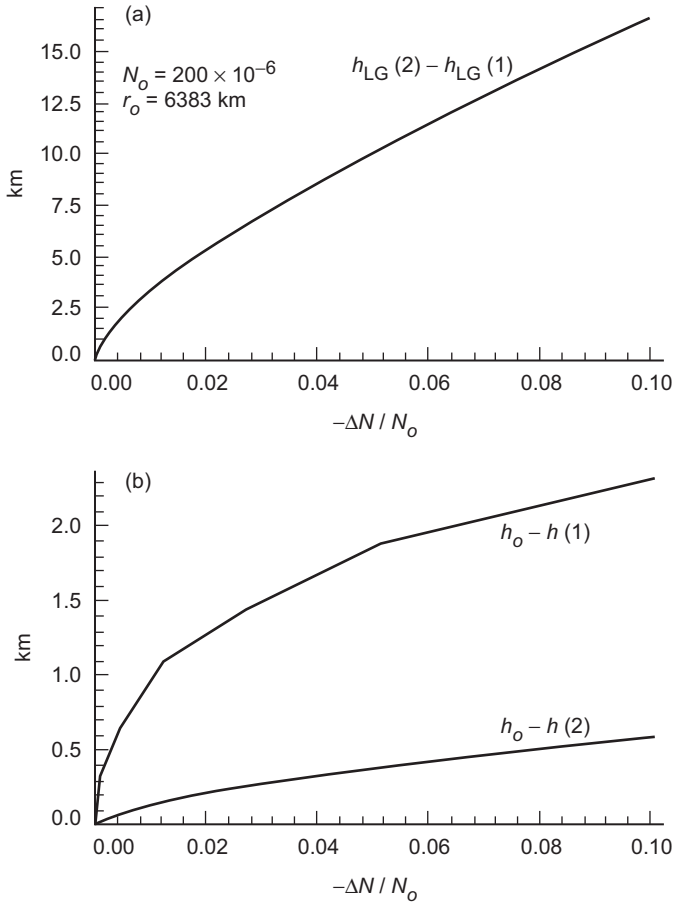

Fig. C-1. Altitude separations and shadow zone width for Case A: (a) width of shadow zone and (b) altitude separation of ray path tangency points.

example would be more applicable to a water vapor layer in the lower troposphere. Here the abscissa is the fractional change $-\Delta N_o / N_o$ across the boundary for the case where $\Delta N_o < 0$, and, therefore, the case where both a shadow zone and a caustic are encountered below the boundary surface. The maximum multipath ($h_o - h(1)$) separation altitude at the tangency points of the rays scales approximately as $(-\Delta N_o)^{1/3}$; the length of the shadow zone scales roughly as $(-\Delta N_o)^{2/3}$ at the boundary.

For a positive ΔN_o, one obtains a doublet ray without a caustic and without a shadow zone. This can be visualized by flipping about the horizontal axis (at h_{LG}) the ray path altitude branch in Fig. 2-7(a). Both the cases $\Delta N_o < 0$ and $\Delta N_o > 0$ are shown in Fig. C-2, which provides the (h_{LG}, h) relationship for these cases. A Fresnel treatment of these cases is shown in Fig. 2-11. When $\Delta N_o > 0$, interference effects are seen several seconds before the main ray hits the boundary. An estimate of the offset of this earlier epoch where multipath first commences—corresponding to t_1 in Fig. 2-2(a)—is given by

$$\Delta t = \frac{D}{V_\perp} \Delta \alpha_{\mathrm{S}} \tag{C-17}$$

For the example shown in Fig. C-2 ($\Delta N_o / N_o = +0.05$), this would result in an initiation of multipath interference about 10 s earlier, that is, when the main ray is still $\zeta^+(r_o) V_\perp \Delta t$, or about 5 km *above* the boundary, which is well outside

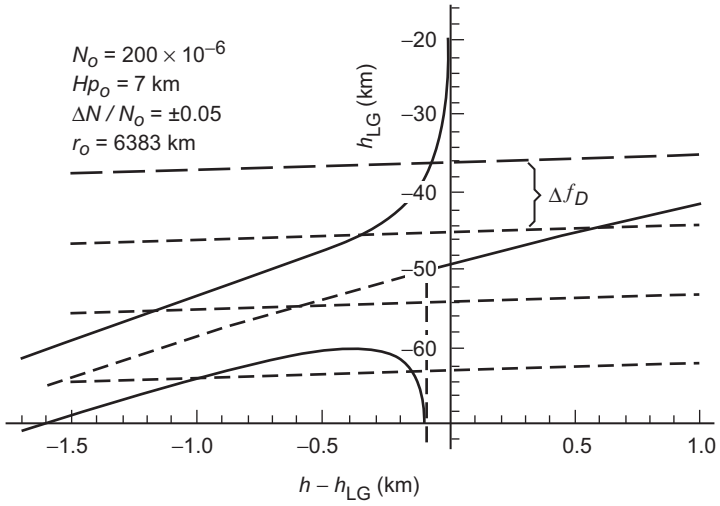

Fig. C-2. Case A example of multipath altitudes and Doppler tones for both a positive and negative discontinuity in refractivity at r_o.

the first Fresnel zone. The interference would be slight at first because of the severe defocusing just below the boundary.

The nearly horizontal dashed lines in Fig. C-2 are constant Doppler contours. The separation between contours ranges between ~30 Hz and ~45 Hz depending on the obliquity between the propagation plane and the LEO orbit plane. For the $\Delta N / N = -0.05$ curve, no stationary phase paths exist with turning point values in the range $r_o - 64\ \mathrm{m} < r_* \le r_o$, resulting also in a shadow zone. Fresnel diffraction for this case is shown in Fig. 2-11. The Mie scattering version for this case is given in Chapter 3, Fig. 3-25. There it is shown that the shadow zone corresponds to a region where no stationary phase points exist in the scattering series representation over spectral number space.

Appendix D
Third-Order Stationary Phase Theory

D.1 Maximum Flaring

At the first contact of a ray with a caustic surface, the radius of the first Fresnel zone becomes infinite. Both the first and second derivatives of the Fresnel phase function vanish at this point. In a geometric optics framework, flaring in observed signal amplitude will be a maximum (in fact, infinite) at this point [see Eqs. (2.2-5) and (2.2-6)]. In a Fresnel framework, the signal amplitude will be large in this neighborhood, but not infinite, and the point where it reaches a maximum may be slightly offset from the first contact point where geometric optics based on a second-order theory predicts infinite signal power.

The maximum and the offset can be estimated using a third-order stationary phase treatment and the thin phase screen model discussed in Chapter 2. Let us expand the Fresnel phase function $\Phi\left(h, h_{\mathrm{LG}}\right)$ defined in Eq. (2.5-1) in a Taylor series about $h^{\dagger}$, the point of zero convexity, where $\partial^2\Phi/\partial h^2 = 0$. We assume that the stationary phase points where $\partial\Phi/\partial h = 0$ are nearby; thus, the principal contributions to the Rayleigh–Sommerfeld integral for the observed signal will come from the immediate neighborhood about $h^{\dagger}$. When $h^{\dagger}$ is well away from an integration end point, the integral $I\left(h_{\mathrm{LG}}\right) = E\left(h_{\mathrm{LG}}\right)\exp\left[i\psi\left(h_{\mathrm{LG}}\right)\right]$ in Eq. (2.5-1) becomes, to a good approximation,

$$I\left(h_{\mathrm{LG}}\right) \doteq \sqrt{\frac{2}{\lambda D}}\,\frac{1}{1+i}\exp\left(i\Phi^{\dagger}\right)\int_{-\infty}^{\infty}\exp\left(i\left(\Phi'^{\dagger}\left(h-h^{\dagger}\right)+\frac{1}{6}\Phi'''^{\dagger}\left(h-h^{\dagger}\right)^3\right)\right)dh \quad \text{(D-1)}$$

From the stationary phase and thin-screen methodology in Chapter 2, we have

$$\left.\begin{aligned}
\Phi^{\dagger} &= \frac{\pi}{\lambda D}\left(h^{\dagger} - h_{\mathrm{LG}}\right)^2 + \frac{2\pi}{\lambda}\int_{h^{\dagger}}^{\infty}\alpha(h)dh \\
\Phi'^{\dagger} &= \left.\frac{\partial\Phi}{\partial h}\right|_{h^{\dagger}} = \frac{2\pi}{\lambda D}\left[h^{\dagger} - h_{\mathrm{LG}} - D\alpha\left(h^{\dagger}\right)\right] \\
\Phi''^{\dagger} &= \left.\frac{\partial^2\Phi}{\partial h^2}\right|_{h^{\dagger}} = \frac{2\pi}{\lambda D}\left(1 - D\alpha'\left(h^{\dagger}\right)\right) = 0 \\
\Phi'''^{\dagger} &= \left.\frac{\partial^3\Phi}{\partial h^3}\right|_{h^{\dagger}} = -\frac{2\pi}{\lambda}\alpha''\left(h^{\dagger}\right)
\end{aligned}\right\} \qquad \text{(D-2)}$$

The Fresnel phase and its partial derivatives are evaluated at the zero convexity point in h-space. We note that $\Phi^{\dagger}$ and $\Phi'^{\dagger}$ vary with h_{LG} and, therefore, with time. However, $\Phi''^{\dagger}$ (and, therefore, $h^{\dagger}$) and also $\Phi'''^{\dagger}$ are independent of h_{LG}. We also note in passing that a zero convexity point requires that the radial gradient of the bending angle be positive, which is of course the same necessary (but not sufficient) condition for the existence of a caustic. The integral in Eq. (D-1) can be evaluated in terms of the Airy function of the first kind [1–3]. We make a change of integration variable in Eq. (D-1), $\Phi'''^{\dagger}\left(h - h^{\dagger}\right)^3 / 6 = z^3$, to obtain for the signal amplitude

$$\left|I\left(h_{\mathrm{LG}}\right)\right| \doteq \sqrt{\frac{1}{\lambda D}}\left(\frac{6}{\left|\Phi'''^{\dagger}\right|}\right)^{1/3} J\left[a\left(h_{\mathrm{LG}}\right)\right] \qquad \text{(D-3)}$$

where the function $J\left[a\left(h_{\mathrm{LG}}\right)\right]$ is given by

$$J(a) = \int_{-\infty}^{\infty}\exp\left(i\left(az + z^3\right)\right)dz = 2\pi 3^{-1/3}\,\mathrm{Ai}\left[a3^{-1/3}\right] \qquad \text{(D-4)}$$

where $\mathrm{Ai}[y]$ is the Airy function of the first kind. The quantity a is given by

$$a\left(h_{\mathrm{LG}}\right) = \Phi'^{\dagger}\left(\frac{6}{\left|\Phi'''^{\dagger}\right|}\right)^{1/3} = \frac{2\pi}{\lambda D}\left[h^{\dagger} - h_{\mathrm{LG}} - D\alpha\left(h^{\dagger}\right)\right]\left(\frac{3\lambda}{\pi\left|\alpha''\left(h^{\dagger}\right)\right|}\right)^{1/3} \qquad \text{(D-5)}$$

The function $J(x)$ in Eq. (D-4) is generated by the differential equation $d^2J/dx^2 = xJ/3$. The solutions are oscillatory for negative x, and, for the case where the boundary values are set by the exact numerical values of $J(0)$ and $J'(0)$ from Eq. (D-4), $J(x)$ damps to zero exponentially with increasing positive x. In addition to diffraction problems, Airy functions arise in classical electrodynamics in connection with the spectral properties of synchrotron

radiation, and also in quantum mechanical potential well problems where the positive z-regime corresponds to quantum tunneling processes. Airy functions can be solved in terms of certain Bessel functions of fractional order $\pm 1/3$, which are tabulated. Airy functions of the first and second kind (the latter grows exponentially large for increasing positive x) provide the asymptotic forms for the Bessel functions for large spectral number and argument.

Figure D-1 shows the behavior of $J(a)$ as a function of the parameter a around zero, including the applicable asymptotic form for the positive regime. The oscillatory behavior of the Airy function $\mathrm{Ai}[y]$ for negative y is shown in [1]. Positive real values of a correspond to that range of h_{LG} values where no real stationary phase points exist near $h^{\dagger}$. For negative values of a, two stationary phase points for $\Phi\left(h, h_{\mathrm{LG}}\right)$ exist at $h - h^{\dagger} = \pm\sqrt{2\Phi'^{\dagger} / |\Phi'''^{\dagger}|}$. The oscillatory behavior of $J(a)$ for the negative regime arises from the phase interference between the contributions to Eq. (D-1) from the neighborhoods around those two points.

A caustic contact point occurs when both $\Phi''^{\dagger}$ and $\Phi'^{\dagger}$ (and a) are zero. At this point, $J(0) = 2\pi / (3\Gamma[2/3]) = 1.547$. As time varies during an occultation, h_{LG} varies nearly linearly and, consequently, so also does a. From Eq. (D-4), setting $dJ/da = 0$ yields the point of maximum flaring, which occurs at $a = -1.469$; the maximum value of $J\left[a\left(h_{\mathrm{LG}}\right)\right]$ is 2.334. Hence,

$$\left|I\left(h_{\mathrm{LG}}\right)\right|_{\mathrm{MAX}} = \left|I\left(\hat{h}_{\mathrm{LG}}\right)\right| \doteq 2.33\sqrt{\frac{1}{\lambda D}}\left(\frac{6}{|\Phi'''^{\dagger}|}\right)^{1/3} \tag{D-6}$$

and

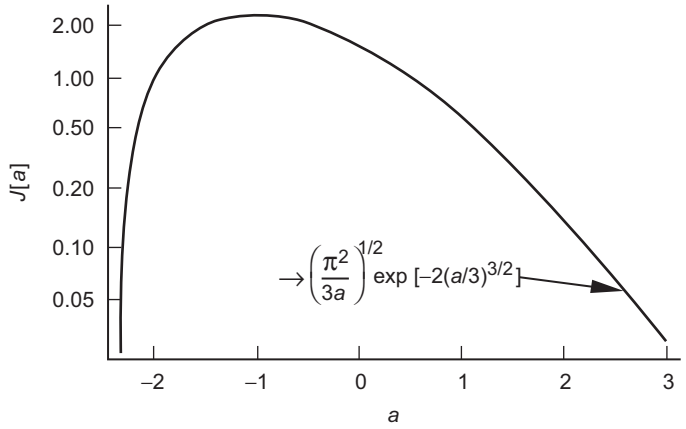

Fig. D-1. First lobe of the Airy function of the first kind.

$$\left.\begin{aligned} \hat{h}_{\mathrm{LG}} - h_{\mathrm{LG}}^{\dagger} &= \hat{h}_{\mathrm{LG}} - \left(h^{\dagger} - D\alpha\left(h^{\dagger}\right)\right) \\ &= 1.47\frac{\lambda D}{2\pi}\left(\frac{\left|\Phi'''^{\dagger}\right|}{6}\right)^{1/3} = 1.47D\left(\frac{\lambda^2\left|\alpha''\left(h^{\dagger}\right)\right|}{24\pi^2}\right)^{1/3} \\ \alpha'\left(h^{\dagger}\right) &= \frac{1}{D} \end{aligned}\right\} \qquad \text{(D-7)}$$

Thus, maximum flaring occurs at $\hat{h}_{\mathrm{LG}}$, which is offset from the first contact with the caustic at $h_{\mathrm{LG}}^{\dagger}$, where $h_{\mathrm{LG}}^{\dagger} = h^{\dagger} - D\alpha(h^{\dagger}) = h^{\dagger} - \alpha(h^{\dagger})/\alpha'(h^{\dagger})$. For example, the first contact point for the upper caustic in Fig. 2-25 (i.e., the one where $h_{\mathrm{LG}} - h_o > 0$) is predicted using the model in Eq. (2.8-33) to be located at $h_{\mathrm{LG}}^{\dagger} - h_o = 600\,\mathrm{m}$, but Eq. (D-7) predicts an additional third-order offset of $\hat{h}_{\mathrm{LG}} - (h^{\dagger} - D\alpha(h^{\dagger})) = 278\,\mathrm{m}$ for maximum flaring, or a total of 878 m above the boundary. The exact value of the location of the maximum flaring point above the boundary is +846 m, which is the location shown in Fig. 2-25. This is good agreement, considering that Eq. (D-1) is a truncated version of the complete convolution integral given by Eq. (2.5-1). The value of $|I|_{\mathrm{Max}}$ predicted by Eq. (D-6) for this example is 1.52, which agrees with the exact result shown in Fig. 2-25 to better than 1 percent.

This offset of the local maximum in intensity from the position of the caustic has its analog in the theory of the rainbow. In 1838, George Airy first demonstrated using ray optics that the scattering angle of any given color in the rainbow, which is a caustic phenomenon, is similarly offset. The Airy function of the first kind originates from his study of this problem [3].

D.2 Minimum Signal Amplitude in a Shadow Zone

A fade-out in signal amplitude occurs if the stationary phase points of the Fresnel phase function $\Phi\left(h, h_{\mathrm{LG}}\right)$ are located in neighborhoods of very large convexity: $|\partial^2\Phi/\partial h^2| = |\left(2\pi/\lambda D\right)\left(1 - D\,d\alpha/dh\right)| >> 1$, so that only a small contribution to the convolution integral in Eq. (2.5-1) is obtained from such points. Thus, fade-outs are associated with very large gradients in the bending angle of a certain polarity. In the limiting case, $\Phi\left(h, h_{\mathrm{LG}}\right)$ has no stationary values anywhere in h-space within the integration limits of Eq. (2.5-1). Since $\Phi\left(h, h_{\mathrm{LG}}\right)$ must grow infinite with increasing $|h|$, it follows for this extreme case that $\partial\Phi/\partial h$ must be discontinuous at some point, and so also must be the bending angle. Two examples of such behavior in the Fresnel phase function are shown in Fig. D-2. For this situation, the principal contributions to the integral in Eq. (2.5-1) will come from those neighborhoods in h-space where

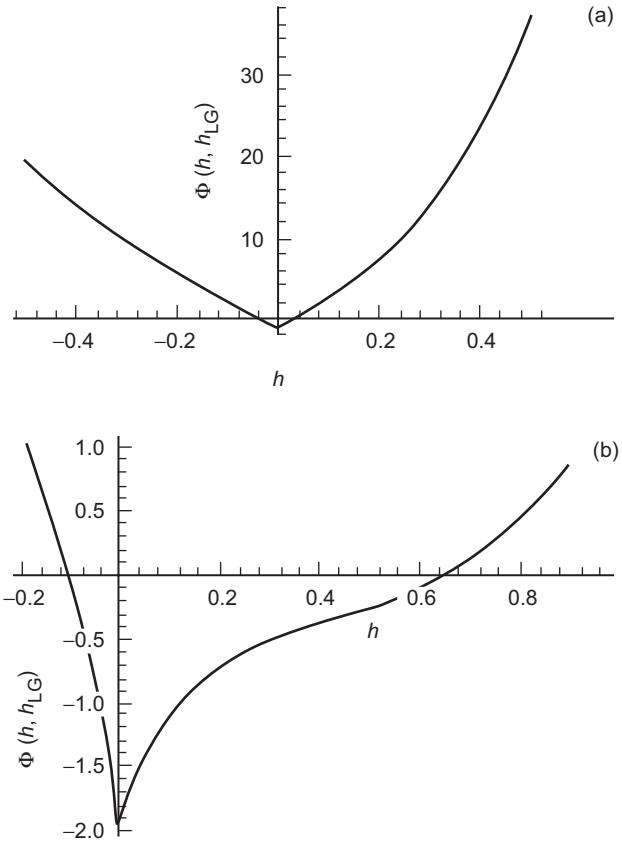

Fig. D-2. Examples of Fresnel phase function convexity for a discontinuous refractivity: (a) positive convexity throughout and (b) positive and negative convexity.

$|\partial\Phi/\partial h|$ is a minimum, which will occur either at the cusp when the convexity is positive throughout, as shown in Fig. D-2(a), or at one or more interior points where $\partial^2\Phi/\partial h^2 = 0$, an example of which is shown in Fig. D-2(b). It is possible for $\Phi(h, h_{LG})$ to have a proper stationary phase point, but with a convexity so large there that the principal contribution to the convolution integral comes from some other neighborhood where $\partial\Phi/\partial h$ is small but not zero. An example is obtained from Fig. D-2(b) by simply "rounding" the cusp point. Geometric optics fails when this type of condition applies.

Let us consider the simplest case of a single cusp and positive convexity throughout, as shown by Fig. D-2(a). In this case, most of the contribution to the Fresnel integral comes from the neighborhood around the cusp. Expanding the Fresnel phase function in a Taylor series through second order about the

cusp position h_o, and using the stationary phase approximation technique, Eq. (2.5-1) becomes

$$I(h_{\mathrm{LG}}) \doteq \sqrt{\frac{2}{\lambda D}}\,\frac{e^{i\Phi_o}}{1+i}\left(\int_{-\infty}^{0} e^{\,i\left(\frac{\partial\Phi_o^-}{\partial x}x+\frac{1}{2}\frac{\partial^2\Phi_o^-}{\partial x^2}x^2\right)}dx+\int_{0}^{\infty} e^{\,i\left(\frac{\partial\Phi_o^+}{\partial x}x+\frac{1}{2}\frac{\partial^2\Phi_o^+}{\partial x^2}x^2\right)}dx\right) \quad \text{(D-8)}$$

where $x = h - h_o$. The superscripts "+" and "–" on the derivatives of Φ denote, respectively, the Fresnel phase functions applicable to each regime, which are evaluated just above (for +) and just below (for –) the cusp position at $x = h - h_o = 0$. Because we are dealing with a trough in signal-to-noise ratio (SNR), the magnitude of the ratio $\partial\Phi_o / \partial x \big/ \partial^2\Phi_o / \partial x^2$ is necessarily large, and the ratio itself will be negative in the "–" regime and positive in the "+" regime. Therefore, we can use an asymptotic expansion for each Fresnel integral. After completing the square and retaining the leading term in the asymptotic expansion, Eq. (D-8) becomes

$$I(h_{\mathrm{LG}}) \doteq \frac{1}{1+i}\sqrt{\frac{2}{\lambda D}}\,\frac{e^{i\Phi_o}}{i\pi}\left[\left(\frac{\partial\Phi^-}{\partial x}\right)^{-1}-\left(\frac{\partial\Phi^+}{\partial x}\right)^{-1}\right]\Bigg|_{x=0} \quad \text{(D-9)}$$

If we use the thin-screen relation $\partial\Phi_o^{\pm} / \partial h = (2\pi / \lambda D)\left[h_o - h_{\mathrm{LG}} - D\alpha^{\pm}(h_o)\right]$, take the absolute value of Eq. (D-9), and minimize $\left|I(h_{\mathrm{LG}})\right|$ with respect to h_{LG}, we obtain a minimum at

$$h_o - h_{\mathrm{LG}}^{\dagger} = \frac{D}{2}\left(\alpha^-(h_o)+\alpha^+(h_o)\right) \quad \text{(D-10)}$$

The minimum value is given by

$$|I|_{\mathrm{Min}} \doteq \frac{2}{\pi}\sqrt{\frac{\lambda}{D}}\left[\frac{1}{\alpha^+(h_o)-\alpha^-(h_o)}\right] \quad \text{(D-11)}$$

For a class of discontinuity as shown in Fig. D-2(a), the minimum of the trough in SNR will be inversely proportional to the discontinuity in bending angle at the boundary at $h = h_o$. This form for the darkening in the discontinuous case should be compared with the case where $\Phi(h, h_{\mathrm{LG}})$ has a single stationary phase point in h-space (the position of which is a function of h_{LG}) but where $\Phi(h, h_{\mathrm{LG}})$ has a very large convexity and the convexity is positive throughout the range of integration in h-space. For this case, we obtain from the stationary phase technique the form

$$\left.\begin{aligned} |I|_{\text{Min}} &= \left|I\left(h_{\text{LG}}^{\dagger}\right)\right| \doteq \min_{h_{\text{LG}}}\left[\sqrt{\xi\left[h\left(h_{\text{LG}}\right)\right]}\right] = \min_{h_{\text{LG}}}\left[\left|1 - D\frac{d\alpha}{dh}\right|^{-(1/2)}_{h\left(h_{\text{LG}}\right)}\right] \\ h\left(h_{\text{LG}}\right) &- h_{\text{LG}} - D\alpha\left[h\left(h_{\text{LG}}\right)\right] = 0 \end{aligned}\right\} \tag{D-12}$$

For a discontinuity as shown in Fig. D-2(b), where the principal part of the contribution to the Fresnel integration comes from interior point(s) away from the cusp, we will need to expand the Fresnel phase function about the points where $\partial^2\Phi / \partial h^2 = 0$, as given in Eq. (D-1). As an example, let us assume that there is only one zero convexity point in each regime located at h^+ in the "+" regime and at h^- in the "–" regime, which is taken to be below the "+" regime. From the thin-screen model, h^+ and h^- are defined respectively by the conditions

$$\left.\begin{aligned} 1 - D\frac{d\alpha^+}{dh}\bigg|_{h=h^+} &= 0 \\ 1 - D\frac{d\alpha^-}{dh}\bigg|_{h=h^-} &= 0 \end{aligned}\right\} \tag{D-13}$$

For this case,

$$\begin{aligned} |I\left(h_{\text{LG}}\right)|^2 = \frac{1}{\lambda D}\Bigg[&\left(\frac{6}{\Phi^{+'''}}\right)^{\frac{2}{3}} J^2\left[a^+\left(h_{\text{LG}}\right)\right] + \left(\frac{6}{\Phi^{-'''}}\right)^{\frac{2}{3}} J^2\left[a^-\left(h_{\text{LG}}\right)\right] \\ &+ 2\left(\frac{6}{\Phi^{+'''}}\right)^{\frac{1}{3}}\left(\frac{6}{\Phi^{-'''}}\right)^{\frac{1}{3}} J\left[a^+\left(h_{\text{LG}}\right)\right] J\left[a^-\left(h_{\text{LG}}\right)\right]\cos(\Phi^+ - \Phi^-)\Bigg] \end{aligned} \tag{D-14}$$

where $a^{\pm}\left(h_{\text{LG}}\right)$ is defined in Eq. (D-5) for each regime. From the thin-screen model, Φ^+ and Φ^- are given by

$$k^{-1}\Phi^{\pm} = \begin{cases} \dfrac{\left(h^+ - h_{\text{LG}}\right)^2}{2D} + \displaystyle\int_{h_o}^{\infty} \alpha^+(h')dh', \quad \text{"+" regime} \\ \dfrac{\left(h^- - h_{\text{LG}}\right)^2}{2D} + \displaystyle\int_{h_o}^{\infty} \alpha^+(h')dh' + h\int_{r}^{h_o} \alpha^-(h')dh', \quad \text{"–" regime} \end{cases} \tag{D-15}$$

One needs to find the minimum of $\left|I\left(h_{\text{LG}}\right)\right|$ in Eq. (D-14) with respect to h_{LG}, but the formal solution is tedious. It is easier to work with specific models in

hand. Note also the interference arising from the $\cos(\Phi^+ - \Phi^-)$ term, which varies with h_{LG}. Equation (D-14) may be considered as providing the magnitude of the vector addition of the Fresnel integration vector component from each regime. $\left|I(h_{\mathrm{LG}})\right|$ will be minimized at that value of h_{LG}, where these two vectors maximally cancel each other upon addition.

For the ionospheric model used in Eq. (2.8-2), for which the Fresnel effects are shown in Figs. 2-17, 2-18, and 2-19, the Fresnel phase function in the trough of Fig. 2-19 (at $h_o - h_{\mathrm{LG}} = \sim 600\,\mathrm{m}$) is shown schematically in Fig. D-2(b). Here only one zero convexity point exists (for h_{LG} in the vicinity of the trough). Thus, in this case the minimum of the SNR trough involves a trade-off between the contribution to the integration from the zero convexity neighborhood in the "–" regime and the contribution from the "+" regime at the boundary. For this specific model, Eq. (D-14) becomes

$$\left|I(h_{\mathrm{LG}})\right|^2 =$$

$$\frac{1}{\lambda D}\left[\left(\frac{1}{\Phi'_o}\right)^2 + \left(\frac{6}{\Phi^{-'''}}\right)^{\frac{2}{3}} J^2\left[a^-(h_{\mathrm{LG}})\right] + 2\left(\frac{1}{\Phi'_o}\right)\left(\frac{6}{\Phi^{-'''}}\right)^{\frac{1}{3}} J\left[a^-(h_{\mathrm{LG}})\right]\cos\left[\Phi_o - \Phi^-\right]\right] \tag{D-16}$$

where, for $h_o - h_- >> -r_o\Delta N$, we have from Eqs. (2.5.1) and (2.8-3)

$$\left.\begin{aligned} \Phi'_o &\doteq \frac{2\pi}{\lambda D}\left(h_o - h_{\mathrm{LG}}\right) \\ \Phi^{-'} &\doteq \frac{2\pi}{\lambda D}\left(h_o - h_{\mathrm{LG}} + r_o\varpi^{1/3}\right) \\ \Phi^{-'''} &\doteq \frac{3\pi\Delta N}{\sqrt{2\lambda r_o^2}}\varpi^{5/6} \end{aligned}\right\} \tag{D-17}$$

where

$$\left.\begin{aligned} \Phi_o &\doteq \frac{\pi}{\lambda D}\left(h_o - h_{\mathrm{LG}}\right)^2, \\ \Phi^- &\doteq \frac{\pi}{\lambda D}\left(h_o - h_{\mathrm{LG}} + r_o\varpi^{1/3}\right)^2 + \frac{4\pi r_o}{3\lambda}\left(-(-2\Delta N)^{3/2} - 3\Delta N\varpi^{1/6}\right) \end{aligned}\right\} \tag{D-18}$$

and where

$$\varpi = \frac{D^2(\Delta N)^2}{2r_o^2} \tag{D-19}$$

Finding the minimum value of $|I(h_{LG})|$ with respect to h_{LG} from Eq. (D-16) using these constraining relations in Eqs. (D-4), (D-17), and (D-18) is straightforward, but in view of the complexity of the stationary phase technique in this case, a straight integration of Eq. (2.5-1) seems simpler.

D.3 Accuracy of the Stationary Phase Technique

The stationary phase technique for integrating Eq. (2.5-1) traditionally has been applied only at stationary points of the Fresnel phase function $\Phi(h, h_{LG})$ or at integration end points. Referring to Fig. 2-9, we see a case in Fig. 2-9(a) where the technique works well, in spite of the reversals in polarity of the convexity of $\Phi(h, h_{LG})$ in h-space. In Fig. 2-9(b), the technique is compromised by the additional contribution from the neighborhood around the zero convexity point (near an altitude of 10 km) where $|\partial\Phi/\partial h|$ is a minimum, but not zero. Lastly, in Fig. 2-9(c), we see a virtually hopeless case for the stationary phase technique.

When a worrisome zero convexity point is well isolated from end points and stationary phase points, we can use Eq. (D-3) to estimate its contribution to the overall diffraction integral in Eq. (2.5-1). Let us call this contribution $|I^{\dagger}|$, which is given by Eq. (D-3) with the partial derivatives of $\Phi(h, h_{LG})$ evaluated at the zero convexity point $h^{\dagger}$ where $|\partial\Phi/\partial h|$ is a local minimum.

Let $\Phi(h, h_{LG})$ for a given value of h_{LG} have a stationary value at $h = h^*(h_{LG})$, where $h^*(h_{LG})$ is defined by $h_{LG} - h^* + D\alpha(h^*) = 0$. Then, applying the stationary phase technique at h^*, one obtains for the amplitude of the signal

$$\left.\begin{aligned} |I^*| &= \left(\frac{2\pi}{\lambda D|\Phi^{*\prime\prime}|}\right)^{1/2} \\ \Phi^{*\prime\prime} &= \left.\frac{\partial^2\Phi}{\partial h^2}\right|_{h=h^*(h_{LG})} \end{aligned}\right\} \tag{D-20}$$

Using Eqs. (D-3) and (D-20), the ratio $|I^{\dagger}|/|I^*|$ is given by

$$\frac{|I^{\dagger}|}{|I^*|} = \sqrt{\frac{2|\Phi^{\prime\prime *}|}{\pi}}\left(\frac{6}{|\Phi^{\prime\prime\prime\dagger}|}\right)^{1/3} J(a) \tag{D-21}$$

We require that $|I^{\dagger}|/|I^*| < \varepsilon$, where ε, for example, might be 1 percent. Let $a^{\dagger}$ be that value of $a > 0$ such that

$$J(a^{\dagger}) = \varepsilon \sqrt{\frac{\pi}{2\Phi''^{*}}} \left(\frac{|\Phi'''^{\dagger}|}{6} \right)^{1/3} \Bigg|_{a=a^{\dagger}} \doteq \frac{1.27}{(a^{\dagger})^{1/4}} \exp\left(-2\left(\frac{a^{\dagger}}{3} \right)^{3/2} \right) \quad \text{(D-22)}$$

Then, from Eq. (D-5), the condition on $\Phi'^{\dagger}$ to achieve a relative accuracy of ε with the stationary phase technique applied only to the stationary phase point is given by

$$|\Phi'^{\dagger}| < a^{\dagger} \left(\frac{|\Phi'''^{\dagger}|}{6} \right)^{1/3} \quad \text{(D-23)}$$

For 1 percent accuracy, $a^{\dagger}$ typically would be in the range from 5 to 10.

References

[1] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards Series 55, Washington, DC, 1964.

[2] F. Oliver, *Introduction to Asymptotics and Special Functions*, New York: Academic Press, 1974.

[3] G. Airy, "On the Intensity of Light in the Neighborhood of a Caustic," *Trans. Cambridge Phil. Soc.*, vol. 6, pp. 379–404, 1838.

Appendix E
Bending by a Gaussian Electron Density Distribution

Suppose that the electron number density distribution $n_e(r)$ is locally spherical symmetric and Gaussian-distributed around some altitude $r_o - R_E$. Then $n_e(r)$ is given by

$$n_e(r) = \frac{\mathrm{TEC}}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(r-r_o)^2}{2\sigma^2}\right) \tag{E-1}$$

where TEC is the total columnar electron content and σ provides the width of the distribution. Equation (E-1) can be considered as the leading term in a Chapman distribution of the form

$$n_e(r) = A\exp\left(1 - \frac{r-r_o}{\sigma} - \exp\left(\frac{r-r_o}{\sigma}\right)\right) \tag{E-2}$$

Using Eqs. (2.2-2) and (2.10-1), it follows that the bending-angle profile from this distribution is given by

$$\alpha(r) \doteq \frac{\kappa\,\mathrm{TEC}}{\mathrm{f}^2}\sqrt{\frac{2r}{\sigma^3}}I(\gamma) \tag{E-3}$$

where $I(\gamma)$ is given by

$$I(\gamma) = -\int_{\gamma}^{\infty}\frac{y}{\sqrt{y-\gamma}}\exp\left(-\frac{y^2}{2}\right)dy, \quad \gamma = \frac{r-r_o}{\sigma} \tag{E-4}$$

The integral $-I(\gamma)$ is shown in Fig. E-1. For $\gamma > 0$, the ray traverses only through the upper wing of the electron distribution, and α damps to zero rapidly with increasing altitude. For $\gamma < 0$, the ray traverses through the peak of the distribution twice. The zero crossing occurs at $\gamma \cong -3/4$; this is the altitude of the tangency point where the positive bending contribution from the lower wing of the electron density distribution just cancels the entire contribution from the upper wing. Below this altitude, the bending angle remains positive and damps to zero more gradually with decreasing altitude.

This electron distribution and the bending-angle profile are used in Chapter 5 in connection with a discussion there on multipath and caustics in a wave theoretic framework.

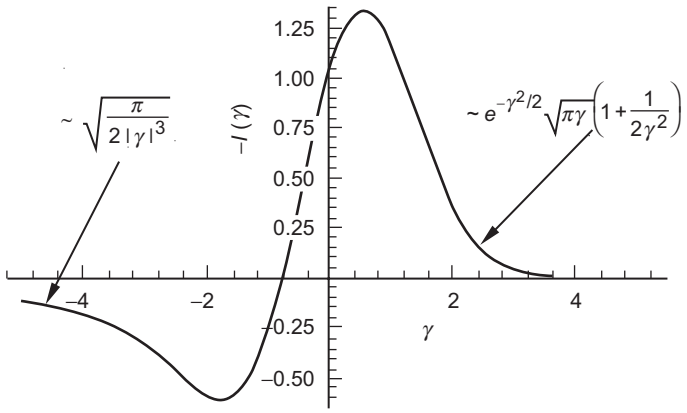

Fig. E-1. Normalized bending-angle profile for a Gaussian refractivity distribution.

Appendix F
The Effect of Cycle Slips on Recovered Refractivity

Equation (A-28) gives the error in the recovered refractivity expressed in terms of the error profile in the bending angle. We use that expression to estimate the error in recovered refractivity that results from a "flywheeling" receiver failing to account for a transient in excess Doppler and, hence, bending angle. To simplify the analysis, we consider the case of a ray encountering a hard discontinuity ΔN at a radius r_o; this is the Case A scenario discussed in the main text. Also, Fig. C-2 shows examples of the transient Doppler profile for this case. To further simplify the analysis, we set the reference refractivity profile to zero. Therefore, we need only the Snell term $\Delta\alpha_S$ for the transient incurred when the main ray hits the boundary bearing the discontinuity. From Eq. (A-48), $\Delta\alpha_S$ may be written in the form

$$\Delta\alpha_S \doteq 2\sqrt{2}\left[\sqrt{\nu} - \sqrt{\nu + \Delta N}\right], \quad \nu = \frac{r_o - r}{r_o} \tag{F-1}$$

Let us consider the case where $\Delta N = N^+ - N^- < 0$, which leads to a shadow zone where receiver operations might be difficult. In this case, we assume that the receiver fails to "track" the excess Doppler associated with this discontinuity; thus, the phase error made by the receiver results in an excess Doppler profile that is equivalent to a bending angle $-\Delta\alpha_S(a)$. If we assume that the receiver fails to track any portion of the transient from its onset at the boundary down to an impact parameter value *a*, then it follows from Eq. (A-28) that

$$\delta\hat{N}(a) = -\frac{1}{\pi}\int_a^{\tilde{a}} \frac{\Delta\alpha_S(\xi)}{\sqrt{\xi^2 - a^2}}\, d\xi \tag{F-2}$$

where $\tilde{a}$ is the impact parameter at the critical refraction altitude $\tilde{a} = r_o(1 + \Delta N)$. If we insert $\Delta\alpha_S$ from Eq. (F-1) into Eq. (F-2) and make some simple but accurate approximations that exploit the large magnitude of a, it can be shown that

$$\delta\hat{N}(\nu) \doteq \Delta N - \frac{2}{\pi}\left(\nu \sin^{-1}\left[\sqrt{\frac{-\Delta N}{\nu}}\right] - \sqrt{-\Delta N(\nu + \Delta N)}\right) \tag{F-3}$$

If we let $\nu \to \infty$, we obtain

$$\delta\hat{N} \to \Delta N \tag{F-4}$$

The recovered refractivity asymptotically assumes a bias equal to the discontinuity, a not too surprising result, and a negative bias in this case.

The reader may object on the grounds that the receiver almost surely would be able to track part of the transient, particularly after the signal returns upon exiting the shadow zone. Suppose that the receiver resumes normal operations at the flaring point associated with the first contact with the caustic (the significant interference at this point notwithstanding; see, for example, Fig. 2-11) and continues to "track" error free at all altitudes below this point. At the contact point, we have from the thin-screen model and Eq. (2.5-1) that $dh_{\mathrm{LG}} / dh = 0$. From Eq. (F-1), it follows that the value of ν at this point is given by the condition

$$\frac{2D}{r_o}\left(\nu^{-1/2} - (\nu + \Delta N)^{-1/2}\right) = 1 \tag{F-5a}$$

where D is the limb distance. When $\nu >> |\Delta N|$, it follows upon expanding Eq. (F-5a) in powers of ν that at the contact point

$$\nu^{\dagger} \doteq \frac{1}{2}\left(-2\Delta N \frac{D}{r_o}\right)^{2/3} \tag{F-5b}$$

which is indeed very much larger than $|\Delta N|$ for small $|\Delta N|$. Hence, the asymptotic form for $\delta\hat{N}$ in Eq. (F-4) still applies even though the receiver is able to partially recover.

In summary, most of the error from cycle slips is incurred in the shadow zone for this example, and the recovered refractivity for subsequent altitudes below this point suffers a negative bias of ΔN. One might expect to find cycle

slips of opposite polarity associated with scenarios where $\Delta N > 0$. Although a ΔN of positive polarity can produce significant interference, it does not cause shadow zones. Also, we note from Figs. 2-11, 3-23, and 3-24 that the "tracking-loop stress" at and just below the boundary (i.e., phase acceleration, which would result from a mis-modeled phase profile in the receiver software, or equivalently, the difference between the extrapolated Fresnel phase from the main ray and the actual Fresnel phase in the shadow zone) differs sharply, depending on the polarity and magnitude of ΔN, with the negative polarity showing a larger phase windup because of the critical refraction or super-refracting situation that can occur for this polarity. In a true shadow zone, there are no rays arriving at the low Earth orbiting (LEO) satellite and, therefore, no stationary phase path to be tracked, which provides a limiting scenario that will be approached for large gradients in refractivity near the boundary. There will be a hiatus in tracking for such zones. On the other hand, when $\Delta N > 0$, there is no hiatus; there is at all points at least one stationary phase path to be "tracked" well before the tangency point of the main ray reaches the boundary. The challenge for the analyst in this case is to assign the correct phase and amplitude to each tone in this triplet zone (see Fig. 2-2a) when there is abundant signal power but significant interference.

We can obtain a rough estimate of the potential cycle slipping in a shadow zone by evaluating the difference between the stationary value of the Fresnel phase at a given epoch and the extrapolated phase that would be observed at the same epoch from the main ray. The latter would be the observed phase produced by the reference phase profile in the thin screen. Let us choose as the epoch the first contact with the caustic where the signal power will be near its maximum value. The extrapolated phase from the main ray would be zero in this example because we have set the thin-screen reference phase profile to zero to simplify the analysis. From Eqs.(2.5-1), (2.8-3), (F-1), and (F-5b) it follows that the stationary value of the Fresnel phase at first contact with the caustic is given by

$$\Phi\left(h^{\dagger}, h_{\mathrm{LG}}^{\dagger}\right) = \frac{2^{5/2}}{3} k r_o (-\Delta N)^{4/3} \left\{ 3\left(\frac{D}{16 r_o}\right)^{1/3} - (-\Delta N)^{1/6} \right\} \tag{F-6}$$

The point of actual maximum in signal-to-noise ratio (SNR) is slightly offset from $h_{\mathrm{LG}}^{\dagger}$, which is located at $\hat{h}_{\mathrm{LG}}$. The correction to the geometric optics prediction for the point of maximum flaring is given in Appendix D. From Eq. (D-7), it follows that

$$\hat{h}_{\mathrm{LG}} - h_{\mathrm{LG}}^{\dagger} = 1.47 \left[\frac{\lambda^2}{4\pi^2} \left(\frac{D^4}{(-\Delta N)^2 r_o} \right)^{1/3} \right]^{1/3} \tag{F-7}$$

and one could as easily evaluate the Fresnel phase at $\hat{h}_{\mathrm{LG}}$ instead of at $h_{\mathrm{LG}}^{\dagger}$. At any rate, Eq. (F-6) gives for the Fresnel phase a value of about 1/2 cycle for $\Delta N = -1 \times 10^{-6}$, and 10 cycles for $\Delta N = -1 \times 10^{-5}$, which are the same levels of cycle slipping seen in Figs. 2-11, 3-23, and 3-24.

A "smart" receiver could be characterized by its ability to account for adverse signal conditions, i.e., multiple tones, interference, and low SNR, and by its capability for near-real-time backward signal recovery and phase connection, etc. Whether these capabilities are found in near-real-time data products from an advanced receiver design or obtained in non-real time from advanced (e.g., Fourier) data analysis techniques, is immaterial. In either case, over a number of trials, one would expect even "smart" designs to have more difficulty with shadow zones than with strong interference zones. This would result statistically in a tendency to incur a net loss of cycles over a series of layers presenting adverse signal conditions involving both polarities for ΔN. If this surmise is accurate, there would be a tendency for recovered refractivity profiles to be negatively biased in the lower troposphere for setting occultations when one cannot reliably deal with adverse signal conditions either by smart receiver design or by advanced techniques for signal recovery.

Appendix G
Using the Sommerfeld–Watson Transformation

G.1 Introduction

The complex spectral number summation technique already has been described briefly in Section 3.7. Here, for illustrative purposes, we outline the technique for computing the scattering from a perfectly reflecting sphere and from a transparent refracting sphere. The technique offers efficient convergence provided that (1) a complete set of poles can be isolated, (2) their residues can be calculated accurately, and (3) the set of poles offers a converging solution. By use of contour integration, we replace the sum for the scattering along the real axis in real-integer spectral number space with another sum in complex spectral number space (see Fig. 3-9). From Eq. (3.12-3), the radial component of the scattered wave in series form is

$$E_r^{(j)} = -\frac{E_o}{x^2}\sum_{l=1}^{\infty} S_l^{(j)}(2l+1)i^{l-1}\xi_l^+(x)P_l^1(\cos\theta) \qquad \text{(G-1)}$$

Here the superscript "(j)" denotes the jth degree scattered wave and $S_l^{(j)}$ is its scattering coefficient. Eq. (3.9-3) gives $S_l^{(0)} = -\left(1+\mathcal{W}_l^-/\mathcal{W}_l\right)/2$, $S_l^{(1)} = n\left(-2i/\mathcal{W}_l\right)^2/2$, $S_l^{(2)} = 2n\mathcal{W}_l^+/\left(\mathcal{W}_l\right)^3$, etc. For the transverse component, appropriately simplified per discussion after Eq. (3.12-4), a similar expression is obtained:

$$E_\theta^{(j)} = \frac{E_o}{x}\sum_{l=1}^{\infty} S_l^{(j)}\frac{2l+1}{l(l+1)}i^{l-1}\xi_l^{+\prime}(x)P_l^{1\prime}(\cos\theta)\sin\theta \qquad \text{(G-2)}$$

Let us consider just the radial component in Eq. (G-1).

We define the function

$$F[l] = S_l^{(j)} i^l (2l+1)\xi_l^+(x) P_l^1\left(\cos(\theta+\pi)\right) \tag{G-3}$$

Consider the contour integral I^+ in the complex plane along a closed path enclosing the positive real axis and lying at an infinitesimal distance ε above and below the real axis [see Fig. 3-9(a)]:

$$I^+ = \frac{1}{2}\left[\int_{0-i\varepsilon}^{\infty-i\varepsilon} \frac{F[l'-1/2]}{\cos(\pi l')} dl' + \int_{\infty+i\varepsilon}^{0+i\varepsilon} \frac{F[l'-1/2]}{\cos(\pi l')} dl'\right] \tag{G-4}$$

We can use the theory of residues in contour integration of analytic functions to evaluate this integral by summing the residues at the simple poles of the integrand, which on the real axis are located only at the half-integer points in l'. Upon summing the residues from the complex contour integrations around each of the poles in Eq. (G-4), one obtains

$$\begin{aligned} I^+ &= -i\sum_{\kappa=0}^{\infty}(-1)^\kappa F[k] = \sum_{l=1}^{\infty}(-1)^{l-1} iF[l] \\ &= \sum_{l=1}^{\infty} S_l^{(j)} i^{l+1}(2l+1)\xi_l^+(x)(-1)^{l-1} P_l^1\left(\cos(\theta+\pi)\right) \\ &= -\sum_{l=1}^{\infty} S_l^{(j)} i^{l-1}(2l+1)\xi_l^+(x) P_l^1(\cos\theta) = E_r^{(s)}\frac{x^2}{E_o} \end{aligned} \tag{G-5}$$

To obtain the result in Eq. (G-5), note that $P_l^1(\cos\ \theta) = (-1)^{l-1} P_l^1(\cos\ (\theta+\pi))$, and also that $F[0] = 0$ because $P_0^1(\cos\theta) \equiv 0$. A similar relation holds for $E_\theta^{(s)}$. In this manner, we convert the sum in Eq. (G-1) into the contour integral given in Eq. (G-4).

Next, we need to extend the contour integration in Eq. (G-4) to include the entire real axis so that we can use the vanishing property of the integrand when it is evaluated on a semi-circular arc bounding the upper complex plane whose radius approaches infinity. Therefore, we will need to invoke the symmetry properties of the integrand in this contour integral. From the defining differential equation for the spherical Hankel functions of order $l-1/2$ [see Eq. (3.6-2)], it follows that these functions must be either symmetric or anti-symmetric in l. It is easily shown from the defining equation for the Bessel function that

$$\xi_{l-1/2}^{\pm}(x) = (-1)^l \xi_{-l-1/2}^{\pm}(x) \tag{G-6}$$

It follows from Eq. (3.5-11) that all of the Wronskian scattering terms, $\mathcal{W}^{\pm}_{l-1/2}$ and $\mathcal{W}_{l-1/2}$, each of which involves products of a pair of spherical Hankel functions, are symmetric in l. Therefore, it follows from Eq. (3.5-11) that all scattering coefficients of spectral number $l-1/2$, $S^{(0)}_{l-1/2}$, $S^{(1)}_{l-1/2}$, $S^{(2)}_{l-1/2}$,$\cdots$, are symmetric in l. Next, from the defining differential equation for the Legendre functions, we have

$$P^1_{l-1/2}(\cos\theta) = P^1_{-l-1/2}(\cos\theta) \tag{G-7}$$

Also, $P^1_0(\cos\theta) = P^1_{-1}(\cos\theta) \equiv 0$. Noting that $2(l-1/2)+1 = 2l$, it follows that

$$(-2l)i^{-l-1/2} = (-1)^{l+1}(2l)i^{l-1/2} \tag{G-8}$$

Assembling all of the parts, we see that

$$\frac{F[-l-1/2]}{\cos(-\pi l)} = -\frac{F[l-1/2]}{\cos(\pi l)}, \quad F[0] = F[-1] = 0 \tag{G-9}$$

Now, the contour integral I^- enclosing the negative real axis is to be taken in the same counter-clockwise sense that also applied to I^+ in Eq. (G-4). Therefore, we obtain for I^-

$$I^- = \frac{1}{2}\left[\int_{0+i\varepsilon}^{-\infty+i\varepsilon} \frac{F[l'-1/2]}{\cos(\pi l')}dl' + \int_{-\infty-i\varepsilon}^{0-i\varepsilon} \frac{F[l'-1/2]}{\cos(\pi l')}dl'\right] \tag{G-10}$$

Upon making a change of integration variable to $v = -l'$ and applying the anti-symmetry property in Eq. (G-9), the resulting contour integrals in Eq. (G-10) become identical with those in Eq. (G-4). Hence, $I^- = I^+$. It follows that we may extend our contour integration in Eq. (G-4) to enclose the entire real axis. We obtain

$$E^{(s)}_r = \frac{E_o}{4x^2}\left[\int_{-\infty-i\varepsilon}^{\infty-i\varepsilon} \frac{F[l-1/2]}{\cos(\pi l)}dl + \int_{\infty+i\varepsilon}^{-\infty+i\varepsilon} \frac{F[l-1/2]}{\cos(\pi l)}dl\right] \tag{G-11}$$

The same symmetry holds for $E^{(s)}_\theta$.

It is convenient to change from the spectral number variable l in Eq. (G-11) to the argument of the Airy functions y because we will be using the asymptotic forms for the Hankel functions in terms of Airy functions and we will be concerned with the zeros of certain combinations of these functions. The defining relationships between y and l are given by Eqs. (3.8-2) and (3.8-3) (here $l = \nu - 1/2$), which can be greatly simplified in our case where

$x_o >> 1$ to the near-linear relationships given in Eq. (3.13-15). In this case, Eq. (G-11) becomes

$$\left.\begin{aligned} E_r^{(s)} &= \frac{E_o K_{x_o}}{4x^2}\left[\int_{-\infty-i\varepsilon}^{\infty-i\varepsilon}\frac{F[l(y)-1/2]}{\cos(\pi l(y))}dy+\int_{\infty+i\varepsilon}^{-\infty+i\varepsilon}\frac{F[l(y)-1/2]}{\cos(\pi l(y))}dy+\mathrm{O}\left[x_o^{-1}\right]\right] \\ l &= x_o + yK_{x_o} \end{aligned}\right\} \tag{G-12}$$

and similarly for $E_\theta^{(s)}$.

Next, we deform this closed contour encompassing the entire real axis for y into one that excludes the real axis as interior points but encompasses the upper complex plane [see Fig. 3-9(b)]. For Hankel functions of the first kind, it can be shown that the integral along the outer boundary of this deformed contour, whose radius approaches infinity, is zero. Thus, the integral around the closed path of the deformed contour may be evaluated in terms of the residues at the poles of the integrand located anywhere in the complex plane except on the real axis. The potential advantage of this approach may be a much more rapidly converging series in the complex plane than the original series along the positive real axis.

To obtain expressions for $E_r^{(s)}$ and $E_\theta^{(s)}$, we must find the poles of $F[l(y)-1/2]/\cos(\pi l(y))$ other than those on the real axis and sum up the residues at these poles. Let y_κ, $\kappa = 1,2,\cdots$ define the location of those poles of $F[l(y)-1/2]/\cos(\pi l(y))$ in the complex plane away from the real axis. From Eqs. (3.5-11) and (G-3), we see that these poles occur at the zeros of $\mathcal{W}_{l-1/2}$. Moreover, Eq. (3.5-11) also shows that when $F[l-1/2]$ represents the scattering coefficients for the jth-degree scattered wave, it has poles of order $j+1$ at these zeros. Thus, $F[l-1/2]$ for an externally reflected wave has only simple poles; $F[l-1/2]$ for the primary wave, which is refracted twice as it passes through the sphere without any internal reflections, has second-order poles; and so on.

Let us write

$$\frac{F[l(y)-1/2]}{\cos(\pi l(y))} = \frac{U[y]}{V[y]} \tag{G-13}$$

where $U[y]$ has no poles for $\mathrm{Im}[y] > 0$ and is not zero at the zeros of $V[y]$, which occur at the points $y = y_\kappa$, $\kappa = 1,2,\cdots$. If $F[l-1/2]$ has poles of order m at these points, then $V[y_\kappa] = V'[y_\kappa] = \cdots = V^{(m-1)}[y_\kappa] = 0$, but $V^{(m)}[y_\kappa] \neq 0$. From the theory of analytic functions, it follows that the residue for $U[y]/V[y]$ at a pole of order m is given by

$$\oint_{y=y_\kappa} \frac{U[y]}{V[y]} dy = 2\pi i \frac{mU^{(m-1)}}{V^{(m)}}\bigg|_{y=y_\kappa} \tag{G-14}$$

G.2 Application to a Perfectly Reflecting Sphere

Let us first consider the case of the externally reflected wave from a large, perfectly reflecting sphere, which has only simple poles. From residue theory, it follows from Eqs. (G-12) and (G-14) that the radial component of the scattered field is given by

$$E_r^{(s)} = \frac{i\pi E_o}{2x^2} K_{x_o} \sum_{\kappa=1}^{\infty} \frac{U[y_\kappa]}{V'[y_\kappa]} \tag{G-15}$$

Here the sequence $y_\kappa, \kappa = 1,2,\cdots$ $y_\kappa, \kappa = 1,2,\cdots$ defines the locations in the complex plane where $V[y]$ is zero.

Recalling the asymptotic forms for the scattering coefficients, we have from Eqs. (3.9-3) and (3.17-1)

$$-S_{l-1/2}^{(0)} = \frac{1}{2}\left(1 + \frac{\mathcal{W}_{l-1/2}^{-}}{\mathcal{W}_{l-1/2}}\right) \underset{n\to\infty}{\to} \frac{\psi'_{l-1/2}(x_o)}{\xi'^{+}_{l-1/2}(x_o)} \to \frac{\mathrm{Ai}'[y]}{\mathrm{Ai}'[y] - i\,\mathrm{Bi}'[y]} \tag{G-16}$$

where y and l are given by Eq. (3.13-15) with ν replaced by $l+1/2$, i.e., $y = K_{x_o}^{-1}(l+1/2-x_o)+\cdots$ and $l+1/2 = x_o + K_{x_o} y + \cdots$. It follows that U and V are given by

$$\left.\begin{aligned} U(y) &= -2i^{\mu}\, l\xi_{\mu}^{+}(x)\,\mathrm{P}_{\mu}^{1}\big(\cos(\theta+\pi)\big)\,\mathrm{Ai}'[y]/\cos(\pi l) \\ V(y) &= \mathrm{Ai}'[y] - i\,\mathrm{Bi}'[y] \\ \mu &= l - 1/2 \end{aligned}\right\} \tag{G-17}$$

For $x_o >> 1$, the asymptotic forms for the Hankel functions in terms of Airy functions [see Eq. (3.8-1)] apply. Also, for an observer some distance from the reflecting sphere, $x >> x_o >> 1$, the asymptotic forms for the Airy functions in terms of complex exponential functions apply [see Eq. (3.8-10)]. In this case,

$$i^{\mu}\xi_{\mu}^{+}(x) \underset{x>x_o>>1}{\to} \left(\frac{x}{D_l}\right)^{1/2} \exp\left(i\left(D_l + l\theta_l - \frac{\pi}{2}\right)\right),\ \mu = l - \frac{1}{2} \tag{G-18}$$

where D_l and θ_l, now complex, are still defined in Eq. (3.10-3) (see Fig. 3-14). Similarly, the product $\mathrm{P}_{\mu}^{1}\big[\cos(\theta+\pi)\big]/\cos(l\pi)$ has the asymptotic form

$$\frac{\mathrm{P}_{\mu}^{1}\left[\cos(\theta+\pi)\right]}{\cos(l\pi)} \underset{\mu>>1}{\rightarrow} \left(\frac{2\mu}{\pi\sin\theta}\right)^{1/2} \frac{\cos\left[l(\theta+\pi)+\pi/4\right]}{\cos(l\pi)},\ \mu = l-1/2 \qquad \text{(G-19)}$$

The zeros of $\left[\mathrm{Ai}'[\hat{y}]-i\,\mathrm{Bi}'[\hat{y}]\right]$ only lie in the first quadrant of the complex plane where $\mathrm{Re}[y]>0$ and $\mathrm{Im}[y]>0$. It follows that $\mathrm{Re}[l]=x_o+\mathrm{Re}[y]K_{x_o}>>1$ and $\mathrm{Im}[l]=\mathrm{Im}[y]K_{x_o}>>1$. Therefore, $\exp(i\pi l)\rightarrow 0$, and we can simplify Eq. (G-19) to

$$\frac{\mathrm{P}_{\mu}^{1}\left[\cos(\theta+\pi)\right]}{\cos(l\pi)} \rightarrow \left(\frac{2\mu}{\pi\sin\theta}\right)^{1/2} \exp\left(-i\left(l\theta+\frac{\pi}{4}\right)\right),\ \mu = l-\frac{1}{2} \qquad \text{(G-20)}$$

It follows that

$$U(y) \underset{\substack{|y|>>\sim 2/K \\ x>>x_o>>1}}{\rightarrow} 4\left(\frac{x^3\sin^3\theta_l}{2\pi i\cos\theta_l\sin\theta}\right)^{1/2} \exp\left[i\left(D_l+l(\theta_l-\theta)\right)\right]\mathrm{Ai}'[y] \qquad \text{(G-21)}$$

When this asymptotic form for $U(y)$ plus $V(y)$ is used in Eq. (G-11), it is easily shown that the result is virtually identical to the scattering integral for $E_r^{(S_o)}$ using the stationary phase approach as given in Eq. (3.12-5) (with the phasor $\exp(i\Phi^{+})$ deleted). Thus, we could have deformed the scattering integral in Eq. (3.12-5) directly into one that spanned the upper complex plane to arrive at Eq. (G-15) without enduring the foregoing discussion leading up to Eq. (G-15).

To evaluate Eq. (G-15), we need the zero crossings of $\left(\mathrm{Ai}'[\hat{y}]-i\,\mathrm{Bi}'[\hat{y}]\right)$, which are exhibited in Fig. G-1. They lie only along the straight line defined by $y=\beta\exp(i\pi/3)$ in the positive half of the complex plane, and their values, given by β_{κ}, $\kappa=1,2\cdots$, are shown in this figure. Because of the high sensitivity of l and μ for large x_o to changes in y, $|\partial l/\partial y|=|\partial\mu/\partial y|\doteq K_{x_o}\approx 475$, high precision is required in determining the values of y_{κ} in order to obtain accurate values for the phasor part of $U[y_{\kappa}]$. Asymptotic formulas for obtaining the roots of $\mathrm{Ai}'[y]-i\,\mathrm{Bi}'[y]=0$ are found in [1]. The values $-\beta_{\kappa}$, $\kappa=1,2\cdots$ are also the zeros of $\mathrm{Ai}'[y]$ along the negative real axis [see Eq. (G-25)]. From the defining differential equation for the Airy function, $z''=xz$, it follows that to calculate the residue at the pole we have $V'(y)=y\left(\mathrm{Ai}[y]-i\,\mathrm{Bi}[y]\right)$. We have, therefore, all of the parts required to calculate the electric field of the reflected wave at the low Earth orbiting (LEO) satellite as a function of θ using the theory of residues and Eq. (G-15).

In order for the summed series in Eq. (G-15) to converge practicably, $U(y_\kappa)$ must approach zero with increasing κ in an efficient way. Inspection of Eq. (G-21) shows that this convergence question hinges on the behavior of the imaginary part of the complex phase term $\Phi_l = D_l + l(\theta_l - \theta)$. By expanding Φ_l in powers of $(l - x_o)$, we obtain

$$\Phi_l = \Phi_{x_o} + (\theta_o - \theta)K_{x_o} y + \frac{(K_{x_o} y)^2}{2D_o} \tag{G-22}$$

A sign change in $d(\mathrm{Im}[\Phi_l])/d(\mathrm{Im}[y])$ occurs at $\mathrm{Im}[y] = D_{x_o}(\theta - \theta_o)/K_{x_o} \approx 200(\theta - \theta_o)$, where $(\theta - \theta_o)$ is expressed in milliradians. Therefore, when the angular position of the LEO is above the geometric shadow boundary, i.e., $(\theta - \theta_o) > 0$, ever so slightly, one obtains a very slowly converging series, which is impracticable. On the other hand, for $(\theta - \theta_o) \le 0$, the series converges rapidly. The latter is, of course, the shadow region where no stationary phase points exist. Thus, two methods for summing the spectral series, the stationary phase technique and the contour integration technique in the complex spectral number plane, complement each other to some extent.

G.3 Application to a Refracting Sphere

The extension to a refracting sphere with a finite index of refraction is straightforward. Here one replaces the derivatives of the Airy functions shown in Eqs. (G-17) and (G-21) with the corresponding Wronskian forms from Eq. (3.5-11). For example, for the external reflected component, $j = 0$, and one obtains

$$\left.\begin{aligned} U(y) &= \left(\frac{x^3 \sin^3\theta_l}{2\pi i \cos\theta_l \sin\theta}\right)^{1/2} \exp\left[i\left(D_l + l(\theta_l - \theta)\right)\right]\left(\mathcal{W}_l^- + \mathcal{W}_l\right) \\ V(y) &= \mathcal{W}_l', \quad l + \frac{1}{2} = x_o + K_{x_o} y + \cdots \end{aligned}\right\} \tag{G-23a}$$

For the refracted wave passing through the sphere without internal reflections, $j = 1$, and $U(y)$ and $V(y)$ become

$$\left.\begin{aligned} U(y) &= 4\left(\frac{x^3 \sin^3\theta_l}{2\pi i \cos\theta_l \sin\theta}\right)^{1/2} \exp\left[i\left(D_l + l\left(\theta_l - \theta\right)\right)\right] \\ V(y) &= \left(\mathcal{W}_l\right)^2, \;\; l + \frac{1}{2} = x_o + K_{x_o} y + \cdots \end{aligned}\right\} \tag{G-23b}$$

which has poles of order two at the zeros of $\mathcal{W}_l$.

In the case of either $j = 0$ or $j = 1$, we need to isolate the zeros of $\mathcal{W}_l$, as defined in Eq. (3.5-11). Using the Airy function asymptotic forms for the Hankel functions, it follows from Eqs. (3.8-1)–(3.8-4), that the Wronskian scattering forms in Eq. (3.5-11) are given by

$$\left.\begin{aligned} \mathcal{W}_l^{\pm} &= \pi n^{1/2}\left(n(\mathrm{Ai}'[y] \mp i\,\mathrm{Bi}'[y])(\mathrm{Ai}[\hat{y}] \mp i\,\mathrm{Bi}[\hat{y}])\right. \\ &\quad \left. - n^{-1}(\mathrm{Ai}[y] \mp i\,\mathrm{Bi}[y])(\mathrm{Ai}'[\hat{y}] \mp i\,\mathrm{Bi}'[\hat{y}])\right) + \mathrm{O}\left[N, \nu^{-5/3}\right] \\ \mathcal{W}_l &= \pi n^{1/2}\left(n(\mathrm{Ai}'[y] - i\,\mathrm{Bi}'[y])(\mathrm{Ai}[\hat{y}] + i\,\mathrm{Bi}[\hat{y}])\right. \\ &\quad \left. - n^{-1}(\mathrm{Ai}[y] - i\,\mathrm{Bi}[y])(\mathrm{Ai}'[\hat{y}] + i\,\mathrm{Bi}'[\hat{y}])\right) + \mathrm{O}\left[N, \nu^{-5/3}\right] \\ y &= \nu^{2/3}\zeta\left[\frac{\nu}{x_o}\right] \doteq K_{x_o}^{-1}\left(\nu - x_o\right) \\ \hat{y} &= \nu^{2/3}\zeta\left[\frac{\nu}{nx_o}\right] \doteq n^{-1/3} y - N x_o K_{nx_o}^{-1}\left(1 + \mathrm{O}\left[\left(nx_o\right)^{-1}\right]\right) \end{aligned}\right\} \tag{G-24}$$

The zero points of $\mathcal{W}_l$ also are shown in Figure G-1 for refractivity values ranging continuously from $Nx_o = 83\pi$ down to $Nx_o = \pi/10$. As N increases, we note from Eq. (G-24) that $\left(\mathcal{W}_l + \mathcal{W}_l^-\right)/\mathcal{W}_l \to 2\,\mathrm{Ai}'[y]/\left(\mathrm{Ai}'[y] - i\,\mathrm{Bi}'[y]\right)$; therefore, the zero points of $\mathcal{W}_l$ should approach the zeros on the line $\mathrm{Im}[y] = \tan[\pi/3]\mathrm{Re}[y]$, the zeros for the perfectly reflecting case.

However, as N grows large, the form for $\mathcal{W}_l$ given in Eq. (G-24) loses numerical precision. Although the magnitude of y may be small, $|\hat{y}| \doteq \left|n^{-1/3} y - N x_o K_{nx_o}^{-1}\right|$ can be large if $Nx_o >> K_{x_o}$. Because y is complex, $|\mathrm{Ai}[\hat{y}]|$ and $|\mathrm{Bi}[\hat{y}]|$ can grow very large. Equation (G-24), which involves differences between these terms, can lose numerical precision. This can be alleviated to a certain extent by noting several identities involving the Airy functions in the complex plane [1]. These are given by

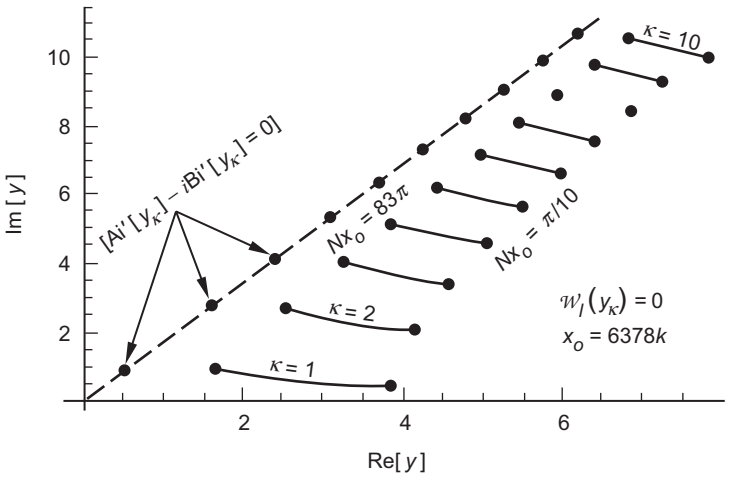

Fig. G-1. Location of zeros for $\mathrm{Ai}'[y] - i\mathrm{Bi}'[y]$ **and also for** $\mathcal{W}_l$.

$$\left.\begin{aligned} \mathrm{Ai}[y] \pm i\,\mathrm{Bi}[y] &= 2(p^{\pm})^{1/2}\,\mathrm{Ai}\left[yp^{\mp}\right] \\ \mathrm{Ai}'[y] \pm i\,\mathrm{Bi}'[y] &= 2(p^{\mp})^{1/2}\,\mathrm{Ai}'\left[yp^{\mp}\right] \\ \mathrm{Ai}[y] + p^{+}\,\mathrm{Ai}\left[yp^{+}\right] + p^{-}\,\mathrm{Ai}[yp^{-}] &= 0 \\ \mathrm{Bi}[y] + p^{+}\,\mathrm{Bi}\left[yp^{+}\right] + p^{-}\,\mathrm{Bi}[yp^{-}] &= 0 \\ p^{\pm} &= \exp\left[\pm 2\pi i/3\right] \end{aligned}\right\} \qquad \text{(G-25)}$$

and

$$\left.\begin{aligned} &\mathscr{W}\left[\mathrm{Ai}\left[yp^{+}\right], \mathrm{Ai}[yp^{-}]\right] = i(2\pi)^{-1},\ \mathscr{W}\left[\mathrm{Ai}[y], \mathrm{Ai}\left[yp^{\pm}\right]\right] = \pm\left(p^{\mp}\right)^{1/4}(2\pi)^{-1} \\ &\mathscr{W}\left[\mathrm{Ai}[y], \mathrm{Bi}[y]\right] = \pi^{-1} \end{aligned}\right\} \qquad \text{(G-26)}$$

where $\mathscr{W}[X,Y]$ is the Wronskian of X and Y. Thus, Eq. (G-24) can be rewritten in the form

$$\left.\begin{aligned} \mathcal{W}_l^{\pm} &= 4\pi n^{1/2}\left(n\,\mathrm{Ai}[yp^{\pm}]\,\mathrm{Ai}'[\hat{y}p^{\pm}] - n^{-1}\,\mathrm{Ai}'[yp^{\pm}]\,\mathrm{Ai}[\hat{y}p^{\pm}]\right) \\ \mathcal{W}_l &= 4\pi n^{1/2}\left(np^{-}\,\mathrm{Ai}[yp^{+}]\,\mathrm{Ai}'[\hat{y}p^{-}] - n^{-1}p^{+}\,\mathrm{Ai}'[yp^{+}]\,\mathrm{Ai}[\hat{y}p^{-}]\right) \end{aligned}\right\} \qquad \text{(G-27)}$$

When N is small, even though $Nx_o >> K_{x_o}$, we can factor out the n^{-1} term in Eq. (G-27) and expand the remaining n^2 term in powers of N. The problem term is the first involving the zeroth power of N. We can write this term in the form

$$p^- \operatorname{Ai}[yp^+]\operatorname{Ai}'[\hat{y}p^-] - p^+ \operatorname{Ai}'[yp^+]\operatorname{Ai}[\hat{y}p^-] = \operatorname{Ai}[yp^+]^2 \frac{d}{dy}\left(\frac{\operatorname{Ai}[\hat{y}p^-]}{\operatorname{Ai}[yp^+]}\right) \quad \text{(G-28)}$$

Also, in this case when $\operatorname{Im}[\hat{y}p^-] \neq 0$, one can use the exponential asymptotic forms for $\operatorname{Ai}[\hat{y}p^-]$ and $\operatorname{Bi}[\hat{y}p^-]$ given in Eqs. (3.8-7) and (3.8-8).

For decreasing N, the zeros of $\mathcal{W}_l$ for a fixed value of κ drift logarithmically downward and to the right in Fig. G-1. For very small N, $Nx_o >> K_{x_o}$, the locations of the zeros are given by the asymptotic condition

$$\pi N K_{x_o}^2 \left(y\left(\operatorname{Ai}[y]^2 + \operatorname{Bi}[y]^2\right) - \operatorname{Ai}'[y]^2 - \operatorname{Bi}'[y]^2\right) = i \quad \text{(G-29)}$$

Figure G-1 shows that, for a fixed value of N, the zeros of $\mathcal{W}_l$ migrate upward from the real axis for increasing integer values of the index κ. As was the case for the perfectly reflecting sphere, here it also can be readily shown that $U(y) \to 0$ exponentially for increasing $\operatorname{Im}[y_\kappa]$ when $(\theta - \theta_o) \leq 0$. In this case the individual contribution of the zeros to the scattering coefficients diminishes exponentially as their distance above the real axis increases.

G.4 Aggregate Scattering

For the aggregate scattering coefficient, which is given by Eq. (3.5-15b), we see that its poles are located where $\mathcal{W}_l + \mathcal{W}_l^+ = 0$. From Eqs. (G-24) and (G-26), it follows that

$$\mathcal{W}_l + \mathcal{W}_l^+ \doteq \begin{cases} 2\pi n^{1/2}\left(n \operatorname{Ai}[\hat{y}](\operatorname{Ai}'[y] - i \operatorname{Bi}'[y]) - n^{-1} \operatorname{Ai}'[\hat{y}](\operatorname{Ai}[y] - i \operatorname{Bi}[y])\right) \\ \text{or} \\ 4\pi n^{1/2}\left((p^+)^{1/2} n \operatorname{Ai}[\hat{y}]\operatorname{Ai}'[yp^+] - (p^-)^{1/2} n^{-1} \operatorname{Ai}'[\hat{y}]\operatorname{Ai}[yp^+]\right) = \\ 4\pi (np^+)^{-1/2} \operatorname{Ai}[\hat{y}]\left(\operatorname{Ai}[\hat{y}] \frac{d}{dy}\left(\frac{\operatorname{Ai}[yp^+]}{\operatorname{Ai}[\hat{y}]}\right) + \left(2N + N^2\right) p^+ \operatorname{Ai}'[yp^+]\right) \end{cases} \quad \text{(G-30)}$$

Unfortunately, the zeros of $\mathcal{W}_l + \mathcal{W}_l^+$ are located in an infinite string along and slightly above the negative real y-axis. Figure G-2 shows the first several zero points for two fixed values of N. Here, contrary to the case shown in Fig. G-1, $\operatorname{Im}[y_\kappa]$ either decreases or does not increase sharply with increasing κ. Also, the numerator term $U(y)$ for S_l does not converge to zero for large values of κ. Therefore, numerical convergence problems result in attempting to evaluate the aggregate scattering field using the poles of S_l in complex spectral number

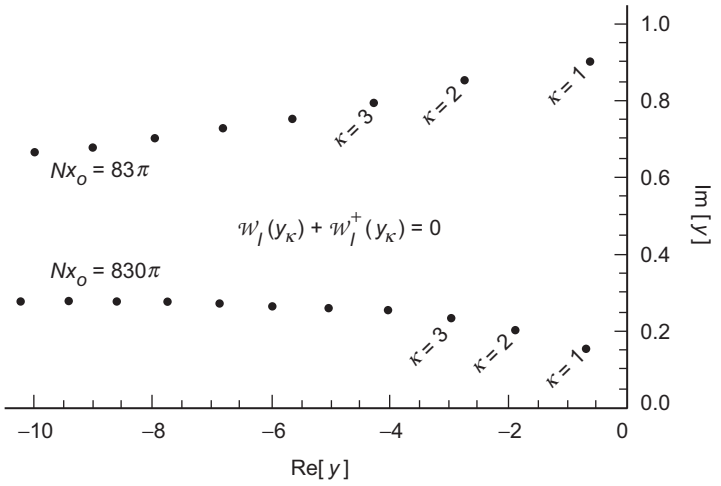

Fig. G-2. Zeros of $\mathcal{W}_l + \mathcal{W}_l^+$ in the complex plane, from Eq. (G-30)

space. One can sum up the contributions from the individual jth-degree scattered waves up to some specified truncation degree using the zeros of $\mathcal{W}_l$ and the residues from the jth-order pole approach outlined earlier. This would require successive numerical differentiation of the appropriate expressions for $U(y)$ and $V(y)$ per the requirements given in Eq. (G-14) for a *j*th-order pole.

A numerical integration approach aided by the stationary phase technique, which has been used extensively in this monograph, seems simpler for smaller magnitudes of *N*, roughly $N <\sim 2K_{x_o}^{-2}$. But it requires high precision and also dense sampling to avoid aliasing, and when multiple rays are involved, it also requires care in the vector addition of the field contributions from these separate stationary phase neighborhoods in spectral number. For larger magnitudes of *N,* the complex spectral number summation technique is preferable in the region where $(\theta - \theta_o) \leq 0$, even with the aforementioned numerical precision problems and the convergence problem for the aggregate scattering coefficient. Also, near shadow boundaries, numerical integration of the scattering integrals when *N* is large becomes a struggle because of the high accelerations in phase in the scattering coefficients and the large number of stationary points in spectral number.

The bottom line is that there seems to be no free lunch in wave theory for calculating accurately in all regimes the complete field vector of an electromagnetic wave that has passed through a transparent refracting sphere. Parabolic equation methods such as the multiple phase screen approach to propagate the wave through a refracting sphere is perhaps the most promising, but multiple internal reflections from the sphere ($j = 2,3\cdots$) would require

careful treatment. For limb sounding from a LEO, a thin-screen/scalar diffraction model offers an attractive alternative, provided in-screen caustics are avoided.

Reference

[1] M. Abramowitz and I. Stegun, eds., *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables,* National Bureau of Standards Series 55, Washington, DC, 1964.

Appendix H
Characteristic Matrix in a Stack of Airy Layers

We model an actual profile for the index of refraction $n(x)$ by a stack of Airy layers (see Sections 4.5 and 4.7). We consider a stack of Cartesian stratified layers where n' is a constant within any given layer but is discontinuous across the boundary between layers. Within a given layer, n'_A is held constant, but across a boundary between two layers, both n_A and n'_A are adjusted to ensure that $\hat{y}$ remains invariant and to best match the profile of the actual index of refraction.

The elements of the characteristic matrix $\boldsymbol{M}[\hat{y}_{Aj}, \hat{y}_{Aj-1}]$ for a transverse electric (TE) wave within the jth Airy layer are given in Eq. (4.5-6). We introduce a shorthand notation here to reduce the size of the equations:

$$\left.\begin{aligned} F_{j,j-1} &= F\left(\hat{y}_{Aj}, \hat{y}_{Aj-1}\right) = \pi \begin{vmatrix} \mathrm{Ai}\left[\hat{y}_{Aj}\right] & \mathrm{Bi}\left[\hat{y}_{Aj}\right] \\ \mathrm{Ai}'\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj-1}\right] \end{vmatrix} \\ f_{j,j-1} &= f\left(\hat{y}_{Aj}, \hat{y}_{Aj-1}\right) = \pi \frac{i}{\gamma_{Aj}} \begin{vmatrix} \mathrm{Ai}\left[\hat{y}_{Aj}\right] & \mathrm{Bi}\left[\hat{y}_{Aj}\right] \\ \mathrm{Ai}\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}\left[\hat{y}_{Aj-1}\right] \end{vmatrix} \\ G_{j,j-1} &= G\left(\hat{y}_{Aj}, \hat{y}_{Aj-1}\right) = \pi i \gamma_{Aj} \begin{vmatrix} \mathrm{Ai}'\left[\hat{y}_{Aj}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj}\right] \\ \mathrm{Ai}'\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj-1}\right] \end{vmatrix} \\ g_{j,j-1} &= g\left(\hat{y}_{Aj}, \hat{y}_{Aj-1}\right) = -\pi \begin{vmatrix} \mathrm{Ai}'\left[\hat{y}_{Aj}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj}\right] \\ \mathrm{Ai}\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}\left[\hat{y}_{Aj-1}\right] \end{vmatrix} \end{aligned}\right\} \tag{H-1}$$

Here γ_{Aj} is given by

$$\gamma_{Aj} = \left(2k^{-1} n_j n'_{Aj}\right)^{1/3} \tag{H-2}$$

It is a constant within the jth layer. Here n_j is the index of refraction in the actual medium at the bottom of the jth Airy layer, i.e., $n_j = n\left(x_{j-1}\right)$, and at this point, $n_A = n_j$. As we progress through the layer, $n_A(x)$, which is a linear function of x, will deviate from $n(x)$ because n'_A differs from $n'\left(x_{j-1}\right)$ and because the curvature and higher-order terms in $n(x)$ produce an additional divergence. Across the upper boundary between the jth and $j+1$st layers, both n and n' are jointly adjusted to keep $\hat{y}_A$ invariant but to readjust $n_A\big|_{x_j}$ so that $n_A\big|_{x_j} = n\left(x_j\right) = n_{Aj+1}$.

Following Eq. (4.4-4), we define a reference characteristic matrix by

$$\overline{\boldsymbol{M}}_{j,j-1} = \begin{bmatrix} F_{j,j-1} & \bar{f}_{j,j-1} \\ \overline{G}_{j,j-1} & g_{j,j-1} \end{bmatrix} \tag{H-3}$$

where

$$\left.\begin{aligned}\tilde{f}_{j,j-1} &= f\left(\hat{y}_{Aj},\hat{y}_{Aj-1}\right) = \pi\frac{i}{\bar{\gamma}_A}\begin{vmatrix}\mathrm{Ai}\left[\hat{y}_{Aj}\right] & \mathrm{Bi}\left[\hat{y}_{Aj}\right]\\ \mathrm{Ai}\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}\left[\hat{y}_{Aj-1}\right]\end{vmatrix}\\ \tilde{G}_{j,j-1} &= G\left(\hat{y}_{Aj},\hat{y}_{Aj-1}\right) = \pi i\bar{\gamma}_A\begin{vmatrix}\mathrm{Ai}'\left[\hat{y}_{Aj}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj}\right]\\ \mathrm{Ai}'\left[\hat{y}_{Aj-1}\right] & \mathrm{Bi}'\left[\hat{y}_{Aj-1}\right]\end{vmatrix}\end{aligned}\right\} \tag{H-4}$$

Here $\bar{\gamma}_A$ is a constant throughout all Airy layers. Its value is to be set later; however, following the discussion in Section 3.5, we would expect it to be some average value among the layers spanned by the characteristic matrix.

It follows that the difference between the actual characteristic matrix for the jth Airy layer and the reference matrix is given by

$$\delta\boldsymbol{M}_{j,j-1} = \boldsymbol{M}_{j,j-1} - \overline{\boldsymbol{M}}_{j,j-1} \doteq \begin{bmatrix}0 & -\bar{f}_{j,j-1}\\ \overline{G}_{j,j-1} & 0\end{bmatrix}\frac{\delta\gamma_{Aj}}{\bar{\gamma}_A} \tag{H-5}$$

where $\delta\gamma_{Aj} = \gamma_{Aj} - \bar{\gamma}_A$, a small but non-zero quantity in general. For thin layers, one can show by using the Wronskian for the Airy functions that $\delta\boldsymbol{M}_{j,j-1}$ can be written as

$$\delta\boldsymbol{M}_{j,j-1} \to \delta\boldsymbol{M}_j \doteq \begin{bmatrix}0 & i\bar{\gamma}_A^{-1}\\ i\bar{\gamma}_A\hat{y}_{Aj} & 0\end{bmatrix}\frac{\delta\gamma_{Aj}}{\bar{\gamma}_A}d\hat{y}_{Aj} \tag{H-6}$$

Here $d\hat{y}_{Aj} = \hat{y}_{Aj} - \hat{y}_{Aj-1}$, which will approach zero upon reducing the maximum thickness of the layers to zero while allowing $N \to \infty$.

We now use the product rule following the discussion in Eqs. (4.4-7) through (4.4-15) to obtain a first-order expression for the characteristic matrix spanning a stack of layers, which will be given in terms of the reference matrix plus a first-order correction matrix. It can be shown using the Wronskian property for Airy functions, $\mathrm{AiBi}' - \mathrm{Ai}'\mathrm{Bi} = 1/\pi$, that for the reference matrix we have

$$\overline{\boldsymbol{M}}_{m,l} = \prod_{\boldsymbol{j=l+1}}^{\boldsymbol{m}}\overline{\boldsymbol{M}}_{j,j-1} = \begin{bmatrix}F_{m,l} & \bar{f}_{m,l}\\ \overline{G}_{m,l} & g_{m,l}\end{bmatrix} \tag{H-7}$$

From Eq. (4.4-7), it follows that a first-order expression for $\boldsymbol{M}_{N,0}$, which spans the entire stack of Airy layers, is given by

$$\begin{aligned} \boldsymbol{M}_{N,0} &= \prod_{j=1}^{N}\left(\overline{\boldsymbol{M}}_{j,j-1} + \delta\boldsymbol{M}_j\right) \\ &\doteq \overline{\boldsymbol{M}}_{N,0} + \sum_{j=1}^{N} \overline{\boldsymbol{M}}_{N,j}\delta\boldsymbol{M}_j\overline{\boldsymbol{M}}_{j,0} \end{aligned} \tag{H-8}$$

where N is the number of layers in the stack. Upon carrying out the matrix multiplication in Eq. (H-8) and using the defining differential equation for Airy functions ($d^2w/dx^2 = xw$), it can be shown that the jth product in its limiting form is given by

$$\overline{\boldsymbol{M}}_{N,j}\delta\boldsymbol{M}_j\overline{\boldsymbol{M}}_{j,0} \rightarrow$$

$$\frac{-\pi^2}{2}\frac{d^2}{d\hat{y}_{Aj}^2}\begin{bmatrix} \begin{vmatrix}\mathrm{Ai}_N & \mathrm{Bi}_N\\ \mathrm{Ai}_j & \mathrm{Bi}_j\end{vmatrix}\cdot\begin{vmatrix}\mathrm{Ai}_j & \mathrm{Bi}_j\\ \mathrm{Ai}_0' & \mathrm{Bi}_0'\end{vmatrix} & \dfrac{i}{\bar{\gamma}_A}\begin{vmatrix}\mathrm{Ai}_N & \mathrm{Bi}_N\\ \mathrm{Ai}_j & \mathrm{Bi}_j\end{vmatrix}\cdot\begin{vmatrix}\mathrm{Ai}_j & \mathrm{Bi}_j\\ \mathrm{Ai}_0' & \mathrm{Bi}_0'\end{vmatrix} \\ i\bar{\gamma}_A\begin{vmatrix}\mathrm{Ai}_N' & \mathrm{Bi}_N'\\ \mathrm{Ai}_j & \mathrm{Bi}_j\end{vmatrix}\cdot\begin{vmatrix}\mathrm{Ai}_j & \mathrm{Bi}_j\\ \mathrm{Ai}_0' & \mathrm{Bi}_0'\end{vmatrix} & -\begin{vmatrix}\mathrm{Ai}_N' & \mathrm{Bi}_N'\\ \mathrm{Ai}_j & \mathrm{Bi}_j\end{vmatrix}\cdot\begin{vmatrix}\mathrm{Ai}_j & \mathrm{Bi}_j\\ \mathrm{Ai}_0 & \mathrm{Bi}_0\end{vmatrix} \end{bmatrix}\frac{\delta\gamma_{Aj}}{\bar{\gamma}_A}d\hat{y}_{Aj} \tag{H-9}$$

In shorthand notation, Eq. (H-9) can be written as

$$\overline{\boldsymbol{M}}_{N,j}\delta\boldsymbol{M}_j\overline{\boldsymbol{M}}_{j,0} \rightarrow i\frac{d^2}{d\hat{y}_{Aj}^2}\begin{bmatrix}\bar{f}_{Nj}F_{j0} & \bar{f}_{Nj}\bar{f}_{j0}\\ g_{Nj}F_{j0} & g_{Nj}\bar{f}_{j0}\end{bmatrix}\delta\gamma_{Aj}d\hat{y}_{Aj} \tag{H-10}$$

It follows that the limiting form for a first-order expression for the characteristic matrix is given by

$$\begin{aligned} &\boldsymbol{M}\left[\hat{y}_{AN},\hat{y}_{A0}\right] \rightarrow \overline{\boldsymbol{M}}\left[\hat{y}_{AN},\hat{y}_{A0}\right] + \\ &i\int_{\hat{y}_{A0}}^{\hat{y}_{AN}}\left(\gamma_A(z)-\bar{\gamma}_A\right)\frac{d^2}{dz^2}\begin{bmatrix}\bar{f}\left(\hat{y}_{AN},z\right)F\left(z,\hat{y}_{A0}\right) & \bar{f}\left(\hat{y}_{AN},z\right)\bar{f}\left(z,\hat{y}_{A0}\right)\\ g\left(\hat{y}_{AN},z\right)F\left(z,\hat{y}_{A0}\right) & g\left(\hat{y}_{AN},z\right)\bar{f}\left(z,\hat{y}_{A0}\right)\end{bmatrix}dz \end{aligned} \tag{H-11}$$

Here $\gamma_A = \left(2n(x)n_A'(x)/k\right)^{1/3} \neq \gamma$, so we have to be clear about the relationship between $\hat{y}_A$ and x in the integration in Eq. (H-11).

For negative values of $\hat{y} <\sim -2$, we can use the asymptotic forms for the Airy functions in the characteristic matrix. Thus, $\overline{\boldsymbol{M}}\left[\hat{y}_{AN},\hat{y}_{A0}\right]$ becomes

$$\overline{\boldsymbol{M}}\left[\hat{y}_{AN},\hat{y}_{A0}\right]\rightarrow\begin{bmatrix}\left(\dfrac{\hat{y}_{A0}}{\hat{y}_{AN}}\right)^{1/4}\cos\left(\mathscr{A}_A\right)_{N,0} & \dfrac{i\bar{\gamma}_A^{-1}}{\left(\hat{y}_{AN}\hat{y}_{A0}\right)^{1/4}}\sin\left(\mathscr{A}_A\right)_{N,0}\\ i\bar{\gamma}_A\left(\hat{y}_{AN}\hat{y}_{A0}\right)^{1/4}\sin\left(\mathscr{A}_A\right)_{N,0} & \left(\dfrac{\hat{y}_{AN}}{\hat{y}_{A0}}\right)^{1/4}\cos\left(\mathscr{A}_A\right)_{N,0}\end{bmatrix} \tag{H-12}$$

where $\mathscr{A}_A=\mathscr{A}_A\left(\hat{y}_A,\hat{y}_{A0}\right)$ is given by

$$\left.\begin{aligned}\mathscr{A}_A\left(\hat{y}_A,\hat{y}_{A0}\right)&=\mathrm{X}\left(\hat{y}_A\right)-\mathrm{X}\left(\hat{y}_{A0}\right)\\ \mathrm{X}\left(\hat{y}_A\right)&=\frac{2}{3}\left(-\hat{y}_A\right)^{3/2}+\frac{\pi}{4}\\ \mathscr{A}_A&=k\int_{x_0}^{x}\gamma_A\sqrt{-\hat{y}_A}\,dx'=k\int_{x_0}^{x}\gamma\sqrt{-\hat{y}}\,dx'\end{aligned}\right\} \tag{H-13}$$

The last relationship follows from Eq. (4.7-7b). From Eq. (4.5-7), it follows that $\mathscr{A}_A$ is still the phase accumulation by the wave along the x-axis, either for a wave traveling through the Airy layers or through the actual medium being modeled with Airy layers. The correction matrix in Eq. (H-11) becomes

$$\int_{\hat{y}_0}^{\hat{y}_N}\overline{\boldsymbol{M}}\left[\hat{y}_{AN},\xi\right]\left(\frac{d\boldsymbol{M}[\xi,\xi]}{d\xi}\right)\overline{\boldsymbol{M}}\left[\xi,\hat{y}_{A0}\right]d\xi\rightarrow$$

$$\int_{\hat{y}_{A0}}^{\hat{y}_{AN}}\frac{\gamma_A-\bar{\gamma}_A}{\bar{\gamma}_A}\sqrt{-\xi}\begin{bmatrix}\left(\dfrac{\hat{y}_{A0}}{\hat{y}_{AN}}\right)^{1/4}\sin\mathscr{B}_A & \dfrac{i\bar{\gamma}_A^{-1}}{\left(\hat{y}_{AN}\hat{y}_{A0}\right)^{1/4}}\cos\mathscr{B}_A\\ -i\bar{\gamma}_A\left(\hat{y}_{AN}\hat{y}_{A0}\right)^{1/4}\cos\mathscr{B}_A & -\left(\dfrac{\hat{y}_{AN}}{\hat{y}_{A0}}\right)^{1/4}\sin\mathscr{B}_A\end{bmatrix}d\xi \tag{H-14}$$

where $\mathscr{B}_A\left(\hat{y}_{AN},\hat{y}_{A0},\xi\right)$ is given by

$$\mathscr{B}_A=\mathscr{A}_A\left(\hat{y}_{AN},\hat{y}_{A0}\right)-2\mathscr{A}_A\left(\xi,\hat{y}_{A0}\right) \tag{H-15}$$

We now set

$$2\bar{\gamma}_A=\gamma_A\big|_{\hat{y}_{A0}}+\gamma_A\big|_{\hat{y}_{AN}} \tag{H-16}$$

and we define

$$\left.\begin{aligned} \varpi_A &= \gamma_A\sqrt{-\hat{y}_A} \\ \overline{\varpi}_A &= \overline{\gamma}_A\left(\hat{y}_{A0}\hat{y}_{AN}\right)^{1/4} \end{aligned}\right\} \tag{H-17}$$

Thus, $k\varpi_A$ is the rate of phase accumulation along the x-axis by a wave traveling through a series of Airy layers; $k\overline{\varpi}_A$ is the geometric average of the rate of phase accumulation.

From Eqs. (H-13) and (H-15), it follows that $2(-\xi)^{1/2}$ is the derivative of $\mathscr{B}_A$ with respect to the dummy variable ξ; thus, one can integrate Eq. (H-14) by parts:

$$\boldsymbol{M}\left[\hat{y}_{AN},\hat{y}_{A0}\right] \doteq \begin{bmatrix} \dfrac{\varpi_A\big|_{\hat{y}_{A0}}}{\overline{\varpi}_A}\cos\left(\mathscr{A}_A\right)_{N,0} + I_1 & \dfrac{i}{\overline{\varpi}_A}\left(\sin\left(\mathscr{A}_A\right)_{N,0} - I_2\right) \\ -i\overline{\varpi}_A\left(\sin\left(\mathscr{A}_A\right)_{N,0} + I_2\right) & \dfrac{\varpi_A\big|_{\hat{y}_{AN}}}{\overline{\varpi}_A}\cos\left(\mathscr{A}_A\right)_{N,0} - I_1 \end{bmatrix} \tag{H-18}$$

Here the integrals I_1 and I_2 are given by

$$\left.\begin{aligned} I_1 &= \frac{1}{2\overline{\gamma}_A}\int_{x_0}^{x_N}\frac{d\gamma_A}{dx}\cos\mathscr{B}_A\,dx \\ I_2 &= \frac{1}{2\overline{\gamma}_A}\int_{x_0}^{x_N}\frac{d\gamma_A}{dx}\sin\mathscr{B}_A\,dx \end{aligned}\right\} \tag{H-19}$$

These forms in Eqs. (H-18) and (H-19) are similar to those that we obtained for Cartesian stratification in Eqs. (4.4-16) and (4.4-17) for a constant intra-layer index of refraction. There, $k\varpi$ is the rate of phase accumulation of the wave along the x-axis; here $\varpi_A = k\gamma_A(x)\left(-\hat{y}_A(x)\right)^{1/2}$ also is the rate of phase accumulation along the x-axis, but for a wave traveling through a series of Airy layers. From Eq. (4.7-7b), it follows that $\varpi_A = k\gamma(x)\left(-\hat{y}(x)\right)^{1/2} = \varpi$, which may be considered as a necessary condition that the refractivity profiles in the Airy layers must satisfy to obtain a limiting form for the characteristic matrix spanning the layers that matches the exact form.

Appendix I
Field Equations in a Stratified Medium

The electromagnetic field vectors for the homogeneous medium are obtained from the vector calculus operations on the scalar potentials given in Eq. (5.2-7). To show that these expressions actually yield a vector field that satisfies Maxwell's equations, we take the curl and divergence of these forms. For example, for the transverse electric (TE) case, we have $\boldsymbol{E}_{\mathrm{TE}} = ik\mu\nabla\times[{}^{\mathrm{m}}\Pi\boldsymbol{r}]$ and $\boldsymbol{H}_{\mathrm{TE}} = \nabla\times\nabla\times[{}^{\mathrm{m}}\Pi\boldsymbol{r}]$. Taking the curl of $\boldsymbol{E}_{\mathrm{TE}}$ evidently satisfies Eq. (5.2-2a). Also, $\nabla\cdot\boldsymbol{H}_{\mathrm{TE}} \equiv 0$, which satisfies Eq. (5.2-2d). Also, $\nabla\cdot\boldsymbol{E}_{\mathrm{TE}} \equiv 0$; therefore, Eq. (5.2-2c) is satisfied. Finally, taking the curl of $\boldsymbol{H}_{\mathrm{TE}}$, $\nabla\times\boldsymbol{H}_{\mathrm{TE}} = \nabla\times(\nabla\times(\nabla\times[{}^{\mathrm{m}}\Pi\boldsymbol{r}]))$, one may use the identities $\nabla\times\nabla\times\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}$ and $\nabla^2(\nabla\times(\psi\boldsymbol{r})) = \nabla\times(\nabla^2(\psi\boldsymbol{r})) = -\boldsymbol{r}\times\nabla(\nabla^2\psi)$ and apply the Helmholtz equation in Eq. (5.2-6) to obtain Eq. (5.2-2b). By symmetry arguments, Maxwell's equations also hold for the transverse magnetic (TM) case where $\boldsymbol{E}_{\mathrm{TM}} = \nabla\times\nabla\times[{}^{\mathrm{e}}\Pi\boldsymbol{r}]$ and $\boldsymbol{H}_{\mathrm{TM}} = -ik\varepsilon\nabla\times[{}^{\mathrm{e}}\Pi\boldsymbol{r}]$.

For the stratified medium with $n = n(r)$, the electromagnetic vector field can be derived from scalar potentials that are solutions to a modified Helmholtz equation. As discussed in Section 5.2, the scalar potentials for the stratified medium have some degree of freedom in their definition. The electromagnetic field vectors for the case where $n = n(r)$ can be expressed through vector calculus operations on a vector version for the modified scalar potentials, $\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]$ and $\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]$. The field vectors are given by

$$\left.\begin{aligned}\boldsymbol{E} &= \varepsilon^{-1}\nabla\times\nabla\times\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right] + ik\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]\\ \boldsymbol{H} &= \mu^{-1}\nabla\times\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right] - ik\nabla\times\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]\end{aligned}\right\} \qquad \text{(I-1)}$$

The factors $\varepsilon^{1/2}$ and $\mu^{1/2}$ have been inserted into the potential terms in Eq. (I-1) to simplify the resulting modified Helmholtz equation that each of the scalar potentials must satisfy. These are given by

$$\left.\begin{aligned}\nabla^2({}^{\mathrm{e}}\Pi) + k^2\tilde{n}^2_{\mathrm{TM}}\,{}^{\mathrm{e}}\Pi = 0\\ \nabla^2({}^{\mathrm{m}}\Pi) + k^2\tilde{n}^2_{\mathrm{TE}}\,{}^{\mathrm{m}}\Pi = 0\end{aligned}\right\} \tag{I-2}$$

where the modified indices of refraction are given by

$$\left.\begin{aligned}\tilde{n}^2_{\mathrm{TM}} &= n^2\left[1 - \frac{r^2\varepsilon^{1/2}}{k^2n^2}\frac{d}{dr}\left(\frac{1}{r^2}\frac{d}{dr}\left(\frac{1}{\varepsilon^{1/2}}\right)\right)\right]\\ \tilde{n}^2_{\mathrm{TE}} &= n^2\left[1 - \frac{r^2\mu^{1/2}}{k^2n^2}\frac{d}{dr}\left(\frac{1}{r^2}\frac{d}{dr}\left(\frac{1}{\mu^{1/2}}\right)\right)\right]\end{aligned}\right\} \tag{I-3}$$

In the electrodynamics literature, the TE and TM waves are generated from linearly independent solutions to the Helmholtz equation. In Eq. (I-1), the term $ik\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]$ generates the electric field $\boldsymbol{E}_{\mathrm{TE}}$, which is perpendicular to $\boldsymbol{r}$, that is, a transverse electric field; this wave is known in the literature as the TE wave. Similarly, the term $-ik\nabla\times\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]$ generates a transverse magnetic field $\boldsymbol{H}_{\mathrm{TM}}$, or the TM wave. The converse polarization holds for the field vectors generated from the curl operations on the $\left[{}^{\mathrm{e}}\Pi\varepsilon^{1/2}\boldsymbol{r}\right]$ term.

If one forms the curl and divergence of $\boldsymbol{E}$ and $\boldsymbol{H}$ as defined by Eq. (I-1), one can show that these scalar potentials do indeed generate fields that satisfy Maxwell's equations in Eq. (5.2-2) when $\varepsilon = \varepsilon(r)$ and $\mu = \mu(r)$. For the TE case, we have $\boldsymbol{E}_{\mathrm{TE}} = ik\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]$ and $\boldsymbol{H}_{\mathrm{TE}} = \mu^{-1}\nabla\times\nabla\times\left[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}\right]$. Taking the curl of $\boldsymbol{E}_{\mathrm{TE}}$ evidently satisfies Eq. (5.2-2a). Also, $\nabla\cdot\left(\mu\boldsymbol{H}_{\mathrm{TE}}\right)\equiv 0$, which satisfies Eq. (5.2-2d). Noting that $\nabla\cdot\boldsymbol{E}_{\mathrm{TE}}\equiv 0$, we have $\nabla\cdot\left(\varepsilon\boldsymbol{E}_{\mathrm{TE}}\right) = \nabla\varepsilon\cdot\boldsymbol{E}_{\mathrm{TE}}$, which is zero because $\nabla\varepsilon$ is radial directed; therefore, Eq. (5.2-2c) is satisfied. Finally, for the case where $\nabla\times\boldsymbol{H}_{\mathrm{TE}} = \nabla\times(\mu^{-1}\nabla\times\nabla\times[{}^{\mathrm{m}}\Pi\mu^{1/2}\boldsymbol{r}])$, one may use the identities $\nabla\times\nabla\times\mathbf{A} = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A}$ and $\nabla^2(\nabla\times(\psi\boldsymbol{r})) = \nabla\times(\nabla^2(\psi\boldsymbol{r})) = -\boldsymbol{r}\times\nabla(\nabla^2\psi)$ and apply Eq. (I-2) to obtain Eq. (5.2-2b). By symmetry, the same proof applies to the TM case.

Appendix J
Conditions for Near-Equivalence between $dG^{\dagger}(\nu)/d\nu$ and $\tilde{\alpha}(\nu,\nu)$, and between $d^2G^{\dagger}(\nu)/d\nu^2$ and $d\tilde{\alpha}(\nu,\nu)/d\nu$

Since $\rho = \rho^{\dagger}(\nu)$ marks a stationary point for $G[\rho,\nu]$ where $\partial G/\partial\rho = 0$, it follows from Eq. (13.2) that

$$\left.\begin{aligned} \frac{dG^{\dagger}(\nu)}{d\nu} &\equiv \frac{dG\left[\rho^{\dagger}(\nu),\nu\right]}{d\nu} = \left(\frac{\partial G[\rho,\nu]}{\partial\nu}\right)_{\rho=\rho^{\dagger}} \\ \frac{d^2G^{\dagger}(\nu)}{d\nu^2} &\equiv \frac{d^2G\left[\rho^{\dagger}(\nu),\nu\right]}{d\nu^2} \doteq \left(-\frac{\partial^2 G[\rho,\nu]}{\partial\rho^2} + \frac{\partial^2 G[\rho,\nu]}{\partial\nu^2}\right)_{\rho=\rho^{\dagger}} \end{aligned}\right\} \qquad \text{(J-1)}$$

Carrying out this program in Eq. (J-1) and using the asymptotic form for $G[\rho,\nu]$ given in Eq. (5.7-2), we have

$$\left.\begin{aligned} &\frac{dG^{\dagger}(\nu)}{d\nu} \doteq \\ &-\nu\int_x^{\infty} \frac{d\log n}{d\omega}\frac{d\omega}{\sqrt{\omega^2-\nu^2}} - \pi K_{\nu}\int_{\rho^{\dagger}}^{x} \frac{d\log n}{d\omega}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\omega, \\ &\frac{d^2G^{\dagger}(\nu)}{d\nu^2} \doteq \\ &-\nu\int_x^{\infty} \frac{d^2\log n}{d\omega^2}\frac{d\omega}{\sqrt{\omega^2-\nu^2}} - \pi K_{\nu}\int_{\rho^{\dagger}}^{x} \frac{d^2\log n}{d\omega^2}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\omega \end{aligned}\right\} \qquad \text{(J-2)}$$

where $\hat{y} \doteq K_\omega^{-1}(\nu - \omega)$. Here x is a chosen point where the accuracy of the asymptotic forms for the Airy functions is deemed adequate when $\rho \geq x$. From geometric optics, we have from Eqs. (5.6-2) and (5.6-15)

$$\left.\begin{aligned} &\tilde{\alpha}(\rho,\nu) = -\nu\int_\rho^\infty \frac{d\log n}{d\omega}\frac{d\omega}{\sqrt{\omega^2-\nu^2}}, \quad \rho \geq \nu, \\ &\frac{d\tilde{\alpha}(\nu,\nu)}{d\nu} = \frac{\tilde{\alpha}}{\nu} - \nu^2\int_\nu^\infty \frac{d}{d\omega}\left(\frac{d\log n}{\omega d\omega}\right)\frac{d\omega}{\sqrt{\omega^2-\nu^2}} \doteq -\nu\int_\nu^\infty \frac{d^2\log n}{d\omega^2}\frac{d\omega}{\sqrt{\omega^2-\nu^2}} \end{aligned}\right\} \quad \text{(J-3)}$$

Comparison of Eqs. (J-2) and (J-3) yields for $dG\left[\rho^\dagger(\nu),\nu\right]/d\nu$

$$\begin{aligned} &\frac{dG^\dagger(\nu)}{d\nu} \doteq \\ &\tilde{\alpha}(\nu,\nu) - \left(\tilde{\alpha}(\nu,\nu) - \tilde{\alpha}(x,\nu)\right) - \pi K_\nu \int_{\rho^\dagger}^x \frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho \end{aligned} \quad \text{(J-4)}$$

From Eq. (5.6-2), we have

$$\tilde{\alpha}(\nu,\nu) - \tilde{\alpha}(x,\nu) = -\nu\int_\nu^x \frac{d\log n}{d\rho}\frac{d\rho}{\sqrt{\rho^2-\nu^2}} \quad \text{(J-5)}$$

Integrating by parts and using Eq. (5.4-3) to express the end value in terms of $\hat{y}$, one obtains

$$\begin{aligned} &\tilde{\alpha}(\nu,\nu) - \tilde{\alpha}(x,\nu) = \\ &2K_\rho^2\sqrt{-\hat{y}(x,\nu)}\left(\frac{d\log n}{d\rho}\right)_{\rho=x} - \nu\int_\nu^x \frac{d^2\log n}{d\rho^2}\log\left[\frac{\rho+\sqrt{\rho^2-\nu^2}}{\nu}\right]d\rho \end{aligned} \quad \text{(J-6)}$$

We can continue integrating Eq. (J-6) by parts. It is clear that by successive integrations we can build up a series of terms, all evaluated at $\rho = x$. Similarly, in Eq. (J-4) for $dG\left[\rho^\dagger(\nu),\nu\right]/d\nu$, we have

$$\begin{aligned}
&\pi K_\nu \int_{\rho^\dagger}^{x} \frac{d\log n}{d\rho'}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho' \doteq \\
&2K_\nu^2\left(\frac{d\log n}{d\rho}\frac{\pi}{2}\int_{\hat{y}(\rho,\nu)}^{\hat{y}^\dagger}\left(\mathrm{Ai}[\hat{y}']^2 + \mathrm{Bi}[\hat{y}']^2\right)d\hat{y}'\right)_{\rho=x} \\
&-\pi K_\nu^2 \int_{\rho^\dagger}^{x} \frac{d^2\log n}{d\rho'^2}\left(\int_{\ddot{y}^\dagger}^{\hat{y}(\rho',\nu)}\left(\mathrm{Ai}[\hat{y}'']^2 + \mathrm{Bi}[\hat{y}'']^2\right)d\hat{y}''\right)d\rho'
\end{aligned} \tag{J-7}$$

It is readily shown that

$$\begin{aligned}
&\frac{\pi}{2}\int_{\hat{y}}^{\hat{y}^\dagger}\left(\mathrm{Ai}[\hat{y}']^2 + \mathrm{Bi}[\hat{y}']^2\right)d\hat{y}' = \\
&\frac{\pi}{2}\left(\left(\mathrm{Ai}'[\hat{y}]^2 + \mathrm{Bi}'[\hat{y}]^2\right) - \hat{y}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)\right) \underset{\hat{y}\to-\infty}{\to} (-\hat{y})^{1/2}
\end{aligned} \tag{J-8}$$

In particular, when $x = \nu + 2K_\nu$,

$$\frac{\pi}{2}\int_{-2}^{y^\dagger}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\hat{y} = 1.419 \doteq \sqrt{2} \tag{J-9}$$

Thus, even for x as low as $x = \nu + 2K_\nu$, the end terms in Eqs. (J-6) and (J-7) are equal to three significant figures. When $n(\rho)$ is slowly varying, it follows that

$$\tilde{\alpha}(\nu,\nu) - \tilde{\alpha}(x,\nu) + \pi K_\nu \int_{\rho^\dagger}^{x} \frac{d\log n}{d\rho}\left(\mathrm{Ai}[\hat{y}]^2 + \mathrm{Bi}[\hat{y}]^2\right)d\rho \doteq 0 \tag{J-10}$$

The accuracy with which Eq. (J-10) holds depends on the curvature in $n(\rho)$, provided that we choose $x > \nu$ so that the asymptotic forms for the Airy functions are not significantly in error. For the examples shown in Figs. 5-4 and 5-5, $K_\nu / H \sim 10^{-3}$, that is, $dn/d\rho$ is slowly varying relative to the range of $\hat{y}$ values ($\sim -2 \le \hat{y} \le 2$) across which the Airy functions make their transition to asymptotic forms. This ratio is generally small for thin atmosphere conditions.

The accuracy of Eq. (J-10) can be checked by comparison of end terms at $\rho = x$ after successive integration by parts in Eqs. (J-6) and (J-7). For example, for the next integration by parts, it can be shown that

$$\frac{\pi}{2}\int_{\hat{y}}^{\hat{y}^\dagger}\left(\int_{\hat{y}}^{\hat{y}^\dagger}\left(\mathrm{Ai}[\hat{y}'']^2+\mathrm{Bi}[\hat{y}'']^2\right)d\hat{y}''\right)d\hat{y}'$$
$$=\frac{\pi}{6}\left(\Gamma(\hat{y}^\dagger)-\Gamma(\hat{y})\right)\underset{\hat{y}\to-\infty}{\to}\frac{2}{3}(-\hat{y})^{3/2}+0.195\cdots \tag{J-11}$$

where $\Gamma(\hat{y})$ has been given in Eq. (4.9-5) and shown in Fig. 5-12. Thus, the difference between Eqs. (J-6) and (J-7) in the end terms after a second integration by parts is about $0.2\nu n''/n$. If $\rho d^2n/d\rho^2 << \alpha$, then a close correspondence between spectral number in wave theory and impact parameter in ray theory should hold. For an exponential refractivity profile in terms of an impact parameter scale height H_ρ, the inequality $\rho d^2n/d\rho^2 << \alpha$ translates into the scale height inequality, $H_\rho >> k^{-1}\left(r_o/\lambda\right)^{1/3}\approx 0.01\,\mathrm{km}$. However, H_ρ is an impact parameter scale height. It relates to a distance scale height H_r by $H_\rho=(d\rho/dr)H_r \doteq H_r+Nr$. Therefore, a value $H_\rho=0$ corresponds to a boundary of a locally super-refracting medium; the critical gradient is $dn/dr=-n/r$, or $H_r\approx 1.5\,\mathrm{km}$. Bending angles are no longer defined for $dn/dr<-n/r$ when the tangency point of the corresponding ray lies within such a layer, or even below it if it is too near the lower boundary.

It follows that when $dn/d\rho$ is slowly varying relative to $\hat{y}$ (i.e., the change in refractivity gradient over the Airy function transition width, from an exponential form to a sinusoidal form, $4K_\rho$, is very small), and specifically when a super-refracting medium is avoided, this near-equivalence between $dG/d\nu$ and $\tilde{\alpha}(\nu,\nu)$ holds. We have from Eqs. (J-4) and (J-11)

$$\frac{dG^\dagger(\nu)}{d\nu}=\tilde{\alpha}(\nu,\nu)+\mathrm{O}\left[\rho\frac{d^2n}{d\rho^2}\right],\quad \rho^\dagger=\nu-\hat{y}^\dagger K_\nu \tag{J-12}$$

Similarly, it can be shown from Eqs. (J-1) through (J-11) that

$$\frac{d^2G^\dagger(\nu)}{d\nu^2}=\frac{d\tilde{\alpha}(\nu,\nu)}{d\nu}+\mathrm{O}\left[\frac{d^2n}{d\rho^2}\right] \tag{J-13}$$

Glossary of Terms

a	Chapter 2, impact parameter, $a = r\sin\gamma$; Appendix D, parameter in third-order stationary phase theory.
(a)	Anomalous ray in multipath scenario.
$a_l^{\pm}$	Transmission coefficient of spectral number l; a_l^{+} applies to an outgoing wave; a_l^{-} applies to an incoming wave.
AS	Anti-spoofing.
$\mathrm{Ai}[\hat{y}]$	Airy function of the first kind.
$\mathscr{A}(x, x_o)$	Phase accumulation from turning point at x_o to x in direction perpendicular to Cartesian-stratified plane in thin-film medium.
(b)	Branching ray in multipath scenario.
$b_l^{(j)}$	Reflection coefficient of spectral number l and reflection degree j. For $j \geq 1$, $j-1$ denotes the number of internal reflections of wave before exiting sphere; $j = 0$ denotes external reflection.
B	Chapter 2, half-width of shadow zone; Fresnel aliasing parameter.
$\mathrm{Bi}[\hat{y}]$	Airy function of the second kind.
$\mathscr{B}$	$\mathscr{B} = \mathscr{A}(x_2, x) - \mathscr{A}(x, x_o)$.
c	Speed of light.
$\mathscr{C}(z)$	In-plane phase accumulation along z-direction parallel to the plane of stratification in a Cartesian-stratified thin-film medium.

D	$D = r_{\mathrm{L}} \cos\gamma_{\mathrm{L}}$; essentially the distance of LEO from Earth's limb, generally the reduced distance; length or phase units apply, depending on context.
D	Reduced distance from LEO to thin phase screen, $D^{-1} = D_{\mathrm{G}}^{-1} + D_{\mathrm{L}}^{-1}$; $D_{\mathrm{G}}, D_{\mathrm{L}}$ = GPS, LEO actual limb distances.
D_ν	$D_\nu = \left(\rho^2 - \nu^2\right)^{1/2}$, the tangential distance in spectral number space from sphere of radius ν to point ρ; see Fig. 3-14.
$\boldsymbol{E}$	Electric field vector of electromagnetic wave; $E_{\parallel}$ is the in-plane component; $E_{\perp}$ is cross-plane.
$\hat{E}$	Stopped electric field with nominal RF removed.
$\hat{E}[\omega]$	Fourier transform of $\hat{E}$.
f	Frequency of harmonic wave.
f_d	Excess Doppler due to refraction.
$F[x]$	Chapter 2 and Appendix D, $F[x] = \left(1 - \mathrm{erf}\left[x^2\right]\right)\exp\left(x^2\right)$; Chapter 3, $F[x] = \left(1 - x^2\right)^{-1/2} \tan^{-1}\left(x^{-2} - 1\right)^{1/2}$; Chapter 6, spectral function, proportional to the bending angle.
$f(x_1, x_2)$	Off-diagonal element in the 2×2 characteristic matrix $\boldsymbol{M}[x_1, x_2]$ for the parallel component of a wave in a stratified thin-film medium.
$F(x_1, x_2)$	On-diagonal element in the 2×2 characteristic matrix $\boldsymbol{M}[x_1, x_2]$ for the parallel component of a wave in a stratified thin-film medium.
$\mathscr{F}$	Vertical radius of the first Fresnel zone.
$g_l(\rho)$	Phase-rate function for the lth spectral coefficient at position ρ.
$g(\hat{y})$	Asymptotic version of $g_l(\rho)$.
$g(x_1, x_2)$	On-diagonal element in the 2×2 characteristic matrix $\boldsymbol{M}[x_1, x_2]$ for the $\perp$ component of a wave in a stratified thin-film medium.
$G(x_1, x_2)$	Off-diagonal element in the 2×2 characteristic matrix $\boldsymbol{M}[x_1, x_2]$ for the $\perp$ component of a wave in a stratified thin-film medium.
$G[\rho, \nu]$	Refractive gradient-induced phase accumulation in the lth spectral coefficient, $a_l^{\pm}$, at the position $\rho = nkr$, $\nu = l + 1/2$.
$G[\mu]$	$G[\mu] = \left(1 - \mu^2\right)^{1/2} \tan^{-1}\left(\mu^{-2} - 1\right)^{1/2}$.

h	Chapter 2, thin-screen altitude in length units; Chapter 3, penetration depth in Fresnel formulas.
h_S	Thin-screen altitude of the ray impact parameter, in phase units.
h_{LG}	Thin-screen altitude of straight line between an occulted GPS satellite and a LEO.
$h(1)$	Lower boundary in thin phase screen for multipath.
$h(2)$	Upper boundary in thin phase screen for multipath.
$\boldsymbol{H}$	Magnetic field vector of electromagnetic wave.
H	Scale height in exponential refractivity distribution; H_p is pressure scale height; H_p is density scale height.
H_S	Refractivity scale height in the transition zone.
H_w	$1-\sigma$ half-width of a Gaussian refractivity distribution.
$H_l^{\pm}(x)$	Hankel function of first (+) and second (–) kinds.
$\mathscr{H}[\mathrm{x}]$	Heaviside function.
ι	Angle of incidence at a reflecting boundary.
I	Chapter 3, amplitude of incident ray in Fresnel formulas; Chapter 3, scattering from a sphere at LEO position minus diffraction from a knife-edge; Chapter 6, angle between the LEO orbit plane and the plane of incidence.
$I(\rho_{\mathrm{L}},\theta_{\mathrm{L}})$	Electric field at the LEO from a reflecting sphere.
$J(\rho_{\mathrm{L}},\theta_{\mathrm{L}})$	Electric field at the LEO from the direct wave with a reflecting sphere; total field $E = I + J$.
J	Knife-edge diffraction integral.
$J_l(x)$	Bessel function of the first kind of order l.
k	Wave number of harmonic wave $k = 2\pi / \lambda$.
K_H	Sensitivity of refractive bending angle to change in scale height.
K_γ	Sensitivity of refractive bending angle to change in lapse rate.
K_ρ	$K_\rho = (\rho/2)^{1/3};\ K_\nu = (\nu/2)^{1/3};\ \rho = nkr,\ \nu = l + 1/2.$
l	Spectral number, an integer.
L1	Chapter 1, GPS carrier at 1.575 GHz; Chapter 2, phase in range units.
L2	Chapter 1, GPS carrier at 1.228 GHz; Chapter 2, phase in range units.

(m)	Main ray in multipath scenario.
M	Sample size number.
$M^{@}$	Fresnel aliasing threshhold for discrete sampling.
$M[y]$	Modulus of the derivatives of the Airy functions.
$\boldsymbol{M}[x_1, x_2]$	Unitary state transition matrix for the electromagnetic field in a stratified refracting medium.
n	Index of refraction function in a refracting medium.
n_e	Electron number density.
N	Refractivity $\times 10^{-6}$; $N = n - 1$.
p	Pressure.
p_k	Generalized parameter in bending-angle function.
$P_l^m(x)$	Legendre polynomial of order l and degree m.
r	Radial coordinate; r_{L} applies to the LEO; r_{G} to the GPS satellite.
r_o	Geocentric distance of spherical boundary surface.
$\measuredangle$	Angle of reflection at a reflecting boundary.
R	Amplitude of the reflected wave from Fresnel formulas.
R_E	Radius of the Earth.
$\mathscr{R}$	Ratio of amplitudes at the LEO of the main ray and caustic ray.
s	Arc length along ray.
S_l	Aggregate scattering coefficient for spectral number l.
$S_l^{(j)}$	Scattering coefficient for a wave with $j-1$ internal reflections for $j \geq 1$; $S_l^{(0)}$ is the coefficient for an external reflected wave.
$\boldsymbol{S}$	Poynting vector; $\boldsymbol{S}^{(i)}$ applies to the incident wave; $\boldsymbol{S}^{(s)}$ applies to the scattered wave.
$\mathscr{S}$	Eikonal function; eikonal equation is $\lvert \nabla \mathscr{S} \rvert = n(x, y, z)$.
t	Time.
T	Time interval spanned by data set; Chapter 3, amplitude of the transmitted wave in Fresnel formulas; Chapter 2 and Appendix A, temperature.
u	Chapter 2, dimensionless thin screen altitude, $u = h(2/\lambda D)^{1/2}$; Chapter 5, dimensionless radial coordinate, $u = kr$.
$U(x)$	Solution to the electromagnetic wave equations for the transverse component of the field in a stratified medium.

$V(x)$	Solution to wave equations for the in-plane tangential component of the field in a stratified medium.
$W(x)$	Solution to wave equations for the in-plane $\perp$ component of the field in a stratified medium.
$W(\rho,\rho_*)$	Weighting function in the wave theory analog of the Abel transform.
$V_\perp$	Component of the geocentric radial velocity of the tangency point of the GPS–LEO straight line in the propagation plane.
$\mathscr{W}_l, \mathscr{W}_l^{\pm}$	Mie scattering functions in transfer equations across a boundary.
x	Chapter 3, in-plane Cartesian coordinate $\perp$ to $\boldsymbol{S}^{(i)}$; also $x = kr$; Chapter 4, in-plane Cartesian coordinate $\perp$ to the plane of stratification.
y	Argument of the Airy functions in a homogeneous medium; Chapter 3 and Chapter 5, cross-plane Cartesian coordinate $\perp$ to $\boldsymbol{S}^{(i)}$; Chapter 4, cross-plane coordinate parallel to the plane of stratification.
$\hat{y}$	Argument of the Airy functions in a refracting medium; $\hat{y} = \nu^{2/3}\zeta[\nu/\rho],\ \rho = nkr,\ \nu = l + 1/2.$
$\hat{y}^\dagger$	$\hat{y}^\dagger = 0.441\cdots$, the zero point of $g_l(\rho)$.
$Y_l(x)$	Bessel function of the second kind of order l.
z	Chapter 2, deviation of the ray path from a straight line; Chapter 3 and Chapter 5, in-plane Cartesian coordinate parallel to $\boldsymbol{S}^{(i)}$; Chapter 4, in-plane coordinate parallel to the plane of stratification; Chapter 5, argument of the Airy function in third-order theory.
α	Chapter 2, refractive bending angle at the LEO.
$\tilde{\alpha}$	$\tilde{\alpha}(\rho,\rho_*)$ is the cumulative bending angle on an incoming ray at position (r,θ) with $\rho = nkr$ and with an impact parameter value of ρ_*. For a spherical symmetric medium, the cumulative bending angle on an outgoing ray is $2\tilde{\alpha}(\rho_*,\rho_*) - \tilde{\alpha}(\rho,\rho_*)$.
α_{L}	Refractive bending angle at the LEO; $\alpha_{\text{L}}(\rho_*) = 2\tilde{\alpha}(\rho_*,\rho_*)$; ρ_* obtained from Bouguer's law.
β	$\beta = -r(dn/dr)/n$; β is the ratio of the radius of curvature of the iso-refractivity surface to the radius of curvature of the ray.

γ	Chapter 2 and Appendix A, normalized lapse rate, $\gamma = (dT/dr)/T$; Chapter 4, $\gamma = \left(2n_0 n'/k\right)^{1/3}$ in a Cartesian stratified Airy layer; Appendix A and Chapter 5, the angle between the radius vector and the tangent vector of the ray; $\gamma = \theta + \alpha$ when r_{G} is infinite.
$\Gamma[\mathrm{x}]$	Gamma function.
$\Gamma(\hat{y})$	$d\Gamma / d\hat{y} = 3g(\hat{y})$.
δ	$\delta = \theta_{\nu *} - \theta$ equals the refractive bending angle to first order in N.
$\delta(x)$	Dirac delta function.
ε	Chapter 2, in the thin-screen model, $\varepsilon(h)$ is the phase perturbation due to discontinuity in a refraction parameter; Chapters 3, 4, and 5, dielectric coefficient in a refracting medium, assumed linear; Chapter 6, azimuthal angle of the apparent direction to the occulted GPS satellite as seen from the LEO.
ζ	Defocusing factor $\zeta^{-1} = 1 - Dd\alpha / d\rho_*$.
$\zeta[\nu/\rho]$	Relates the argument $\hat{y}$ of the Airy functions to ν and ρ in the Bessel functions, $\hat{y} = \nu^{2/3}\zeta[\nu/\rho]$.
η	Auxiliary parameter in multipath altitude separation.
θ	Geocentric central angle in the orbit plane relative to the z-axis; θ_{L} applies to the LEO; θ_{G} applies to the GPS satellite.
$\tilde{\theta}_{\mathrm{L}}$	Geocentric central angle of the LEO relative to the z-axis in the plane of propagation; see Fig. 6-8(a).
θ_ν	$\theta_\nu = \sin^{-1}(\nu/\rho)$; see Fig. 3-14; θ_ν is the "angle of incidence" of the νth spectral component in a spherical framework.
κ	Chapter 2, the coefficient relating refractivity to n_e, $\kappa = 40.3\,\mathrm{m}^3/\mathrm{s}^2/e$; Chapter 3, the extinction coefficient in an absorbing medium.
λ	Wavelength of the harmonic wave, $\lambda = 2\pi / k$.
μ	Chapters 3, 4, and 5, magnetic permeability of the refracting medium; Chapter 3, $\mu = \nu / \rho$.
ν	Fractional spectral number, $\nu = l + 1/2$; Chapter 2, $\nu = (r_o - r)/r_o$.
ν^*	Stationary phase value of fractional spectral number ν.
ξ	Parameter in multipath altitude separation, Appendix C.

$\xi_l^{\pm}(x)$ Spherical Hankel function of the first (+) and second (–) kinds, $\xi_l^{\pm} = \sqrt{\pi x/2}H_\nu^{\pm}(x) = \sqrt{\pi x/2}(J_\nu(x) \pm iY_\nu(x))$.

Π Scalar potential for electromagnetic wave; ${}^{e}\Pi$ is electric potential; ${}^{m}\Pi$ is magnetic potential.

ρ Radial coordinate $\rho = krn(r)$; Appendix A, mass density.

$\tilde{\rho}$ $\tilde{\rho} = n_{\mathrm{A}} kr$; n_{A} is a constant.

ρ_{L} Radial position of the LEO in spectral number space.

ρ_* Ray path impact parameter.

$\rho_{\mathscr{C}}$ Impact parameter at a caustic contact point.

$\rho^{\dagger}$ $\rho^{\dagger} = \nu - \hat{y}^{\dagger} K_{\rho^{\dagger}}$, a stationary point for $G[\rho, \nu]$ with respect to ρ.

$\rho^{(0)}$ Wave theory analog of Fresnel reflection coefficient.

σ Chapter 2, $\sigma = \sqrt{(r_o - r)/H_{p_o}}$; Appendix E, parameter in the electron density model; Chapter 5 and Appendix A, $\Delta\sigma$ is the $\perp$ displacement between two rays with an impact parameter difference $\Delta\rho_*$; Chapter 5, parameter in the generalized Fresnel transform of $\mathrm{Ai}[y]$.

σ_+ Chapter 2, accounts for ray path bending in the case of a discontinuous refractivity, $\sigma_+^2 = \sigma^2 + \beta\Delta N / N$.

τ Time interval; Chapter 5, parameter in the generalized Fresnel transform of $\mathrm{Ai}[y]$.

ϕ Azimuthal angular coordinate around the z-axis.

$\varphi(h_S)$ Stationary phase profile embedded in the thin phase screen.

φ Angle of incidence in Fresnel formulas.

$\Phi(h_S, h_{\mathrm{LG}})$ Thin-screen Fresnel phase $= \varphi(h_S) + (h_S - h_{\mathrm{LG}})^2 / (2D)$.

$\Phi^{\pm}$ Chapters 3 and 5, geometric phase delay to the LEO from the tangency point on the near side (–) or far side (+) of the sphere of radius ν.

Φ^* Chapter 5, $\Phi^* = \Psi(\pm, -)\big|_{\nu=\nu^*} \pm \pi/4$, the stationary phase at point (r, θ) from wave theory.

$\chi_l(x)$ Spherical Bessel function of the second kind of order l, $\chi_l(x) = \sqrt{\pi x/2}Y_{l+1/2}(x)$.

$\hat{\mathrm{X}}$	$\hat{\mathrm{X}} = 2(-\hat{y})^{3/2}/3 + \pi/4$, variable in the negative argument asymptotic forms for the Airy functions.
$\psi(h_{\mathrm{LG}})$	Phase of wave at the LEO in the thin screen model.
$\psi_l(x)$	Spherical Bessel function of the first kind of order l, $\psi_l(x) = \sqrt{\pi x/2}J_{l+1/2}(x)$.
$\Psi(\pm,\pm)$	Chapter 5, spectral density function for the phase of the vth spectral component of the wave at the point (r,θ).
$(\pm,\pm)$	The four combinations $(\pm,\pm)$ yield the phase spectral densities $\Psi(\pm,\pm)$ for waves passing through either the near or the far side of a sphere, and for either incoming or outgoing waves. When $r_o/\lambda >> 1$, only the near-side combinations $(\pm,-)$ contribute to the field at (r,θ) when $0 << \theta << \pi$, and in this case only the outgoing combination $(+,-)$ applies at the LEO.
$\Psi^{(j)}$	Chapter 3, reflection angle from geometric optics of an exit ray after $j-1$ internal reflections; Chapter 3, phase of degree j scattering phasor plus near-side geometric delay $\Psi^{(j)} = 2\Omega^{(j)} + \Phi^{-}$.
$\Psi^{(S)}$	Chapter 3, phase of aggregate scattering plus near-side geometric phase delay at (r,θ), $\Psi^{(S)} = \Psi^{(S)}(x,\theta,N,x_o,v) = 2\Omega^{(S)} + \Phi^{-}$.
ω	Angular frequency of the harmonic wave, $\omega = ck$.
ϖ	Chapter 4, $\varpi(x) = \left(n(x)^2 - n_o^2\right)^{1/2}$ in a Cartesian stratified Airy layer; $n_o = n(x_o) = \text{constant}$; x_o is a turning point.
$\Omega^{(j)}$	Chapter 3, phase of jth degree scattering coefficient $S_l^{(j)}$; $2\Omega^{(S)}$ is the phase of S_l; Ω_A is the asymptotic form of $\Omega^{(S)}$ for $v << \rho$.

Acronyms

4DVAR	four-dimensional variational analysis
AS	anti-spoofing
ASIC	application-specific integrated circuit
C/A	clear access
C/A	coarse acquisition
CAL	Cartesian Airy layer
CHAMP	Challenging Minisatellite Payload
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
DoD	Department of Defense
ECMWF	European Centre for Medium-Range Weather Forecasts
FFT	fast Fourier transform
GLONASS	Global Navigation Satellite System
GNSS	global navigation satellite system
GPS	Global Positioning System
GPS/MET	Global Positioning System/Meteorology
GRACE	Gravity Recovery and Climate Experiment
I	in-phase
IF	intermediate frequency
IGS	International GPS Service

IUGG	International Union of Geodesy and Geophysics
JPL	Jet Propulsion Laboratory
L1	1575 MHz
L2	1226 MHz
LEO	low Earth orbiting (spacecraft; orbit altitude less than $\sim 0.2R_E$)
LHS	left-hand side
LSL	large spherical layer
NWP	numerical weather prediction
P	precision
POD	precision orbit determination
Q	quadrature
RF	radio frequency
RHCP	right-hand circularly polarized
RHS	right-hand side
rms	root-mean-square
SA	Selective Availability
SAC-C	Satellite de Aplicaciones Cientificas-C
SNR	signal-to-noise ratio (SNR_V is the voltage SNR, the ratio of signal amplitude to noise amplitude)
TE	transverse electric (electromagnetic wave with electric field vector directed $\perp$ to $\boldsymbol{r}$).
TEC	total along-track electron content per unit columnar area (1 TEC = 10^{16} e/m^2)
TM	transverse magnetic (electromagnetic wave with magnetic vector directed $\perp$ to $\boldsymbol{r}$).
USAF	U.S. Air Force
UTC	coordinated Universal time
WKB	Wentzel–Kramers–Brillouin