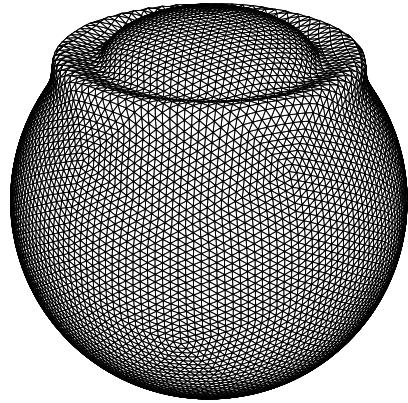


Waves on a Spherical Membrane

by

Carl Harlan Tape

Thesis submitted to the University of Oxford
for the Degree of Master of Science by Research



Department of Earth Sciences
and
Worcester College
University of Oxford
July 15, 2003

Advisor: Prof. John H. Woodhouse

Abstract

The purpose of this study is to develop a numerical model for wave propagation on a spherical membrane and to compare the results with predictions made from ray theory. Such membrane waves are an analogue for seismic surface waves. The two-dimensional wave equation on a sphere is solved using a finite-difference method on a spherical grid of several thousand hexagonal faces. The spherical grid is first tested using a homogeneous phase velocity field with an initial Gaussian source. The resulting numerical solutions, $u(\theta, \phi, t)$, agree with analytical solutions for the homogeneous case. Next a heterogeneous phase velocity field, $c(\theta, \phi)$, is used; the solutions reveal the variation in amplitude and phase due to the lateral heterogeneity.

Surface wave ray-tracing equations are then used to calculate the amplitude and phase anomalies for a set of source-receiver pairs. Exact ray theory calculates these values along the exact ray path, which may deviate considerably from the great-circle path between the source and receiver, which is the basis for the linearized ray theory calculations. We find that multipathing — multiple ray paths between a source and receiver — is quite common for the current resolution of long-period surface wave phase velocity maps, and it gives an indication of the divergence of linearized ray theory from exact ray theory.

Using the results from the numerical model, we are able to examine the regimes under which ray theory predictions are invalid, and we discuss the results in terms of the ray theory validity condition, $\lambda \ll \Lambda$, where λ is the wavelength of the waves in the numerical model and Λ is the minimum scalelength of heterogeneity in the phase velocity map. We show that exact ray theory is better than linearized ray theory at predicting both phase and amplitude anomalies. We find that ray theory predictions of phase are quite stable. Predictions of amplitude, however, are valid only when the validity condition is well-satisfied.

Waves on a Spherical Membrane

Extended abstract

The purpose of this study is to develop a numerical model for wave propagation on a spherical membrane and to compare the results with predictions made from ray theory. Such membrane waves are an analogue for seismic surface waves. Ray theory has been used in seismology to explain observations of path differences, amplitude differences, and phase differences of surface waves with respect to a spherical reference model (e.g., *Woodhouse and Wong*, 1986; *Laske et al.*, 1994). Furthermore, it is a standard tool used in creating 3D global models of the earth. However, ray theory is an approximate theory and cannot always be used to explain observations. Thus, numerical models are helpful, especially as earth models become increasingly more detailed. This study presents a two-dimensional numerical model aimed at investigating the regimes under which ray theory predictions are invalid.

Analytical solutions

Our study is concerned with analytical and numerical solutions to the wave equation on the sphere:

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = f,$$

where $u = u(\theta, \phi, t)$ is the wave field, $c = c(\theta, \phi)$ is the frequency-independent phase velocity, $f = f(\theta, \phi, t)$ is the forcing term, and $\nabla^2 u$ is the surface Laplacian of u on a sphere with radius a , where θ, ϕ are spherical coordinates.

For the analytical solutions, we specify a constant phase velocity function, as well as the forcing term $f(\theta, \phi, t)$, which describes the source. The source we use is simply a Gaussian in time (source-time function) multiplied by a Gaussian in space (“initial-shape function”).

Numerical solutions

We employed a finite-difference method (FDM) on a special spherical grid in solving the wave equation numerically, using a method developed in the field of climate modelling (*Williamson*, 1968). The grid was constructed by subdividing triangular patches on the sphere into four

new triangles at each level of subdivision. This process, which was carried out on grids with up to 122,882 points, achieves a relatively uniform distribution of gridpoints. The numerical approximation for the surface Laplacian (∇^2) in the wave equation is based on a local FDM which averages the function gradients at a point with its six adjacent neighbors.

We are able to test the numerical solutions by comparing them with the analytical solutions for the homogeneous case, that is, $c(\theta, \phi) = \text{const.}$ We find good agreement between the numerical and analytical solutions, and this gives us confidence in the numerical solutions for the heterogeneous case. For the heterogeneous case, we use surface wave phase velocity maps for the function $c(\theta, \phi)$ (*Trampert and Woodhouse, 1995*).

The solution to the wave equation is the displacement field, $u(\theta, \phi, t)$, i.e., a seismogram for a fixed surface point, or a snapshot of the distorted spherical membrane for a fixed time. Comparing the seismograms from the homogeneous earth case with those from the heterogeneous case, we calculate the phase anomaly P and amplitude anomaly A , where P and A are the values that minimize the residuals between the two seismograms for the first four arrivals, orbits R1–R4. This gives us our basis for comparison with the phase and amplitude anomalies calculated in ray theory.

Ray theory predictions

With surface wave ray theory, the rays, confined to the spherical surface, are the orthogonal trajectories of the wavefront of the surface wave. Ray theory can be used to calculate the amplitude anomalies and phase anomalies for a given heterogeneous earth model (with respect to a homogeneous earth model). We examine the results from ray theory using exact ray-tracing as well as a linearized ray theory, in which the calculations are carried out along the great-circle path between the source and receiver.

We can identify regimes under which ray theory predictions are inaccurate. The regime can be changed by any combination of the following parameters: l_{\max} , the truncation degree of the spherical harmonic expansion of the phase velocity map $c(\theta, \phi)$; ϵ , the power of the map; T , the period of the map; and Δ , the path length of the ray, which has no upper limit for the purposes of ray tracing.

We find that multipathing — multiple ray paths between a source and receiver — is quite common for the current resolution of long-period surface wave phase velocity maps (e.g.,

Trampert and Woodhouse, 1995). In particular, for paths beyond R1 for a map expanded in spherical harmonics up to $lmax=20$, we would expect at least three paths between the source and receiver; for R4 ray paths in an $lmax=40$ map, we find cases with up to 15 ray paths. Multipathing does not occur in linearized ray theory, which assumes that the ray deviates only slightly from great-circle path between the source and the receiver. Thus, multipathing gives an indicating of the divergence of linearized ray theory from exact ray theory.

A comparison of three models

The numerical model provides us with our “ground-truth” estimate of the phase and amplitude anomalies, as it is the only model that calculates wave propagation. Furthermore, it incorporates properties of a specified, albeit simplified, source. Ray theory assumes an infinitesimally-short-wavelength wave, and it is this approximation which leads to deviations from the actual picture of wave propagation. The accuracy of the numerical model is limited by the number of gridpoints per wavelength (gpw): too few gpw results in unwanted numerical dispersion.

Here we will state three conclusions:

1. The results from the numerical model — phase anomalies and amplitude anomalies — correlate well with the results from ray theory, and this correlation is best when $\lambda \ll \Lambda$, where λ is the wavelength of the waves in the numerical model and Λ is the minimum scalelength of the heterogeneity in the phase velocity map. We see that exact ray-tracing is a better predictor of phase and amplitude than linearized ray theory, based on the comparison with $lmax \leq 8$.
2. The phase anomalies predicted from ray theory are much better than the predictions of amplitudes, which are more sensitive to shorter-scalelength structure. Even in regimes where $\lambda > \Lambda$, the phase predictions are well-correlated between the numerical model and ray theory. This is one reason why phase remains the primary seismic data in constructing velocity models of the earth.
3. In regimes of extreme multipathing (i.e., R4 orbits, $T=40s$, $lmax=20$), exact ray theory predicts a large spread phase and amplitude anomalies, as each ray path carries information on the distorted wave traveling between source and receiver. It is possible

that linearized ray theory is better in these extreme cases, or perhaps there is a way of extracting one ray or combining the information from the rays get a single value. This warrants further investigation.

We show that, with increased computing power, we would be able to input shorter and shorter waves into the numerical model, thereby allowing us to see the wave's response to *all* of the structure in the more-detailed phase velocity maps. Thus, we could test ray theory in the regime where multipathing is most prevalent, and where we expect ray theory to fail.

Acknowledgements

I would like to thank my advisor, John Woodhouse, who provided guidance throughout my studies at Oxford. The basic idea behind the project was suggested by him, and he was instrumental in helping to develop the analytical solutions (Chapter 3) and the more generalized ray-tracing equations (Appendix E). It has been a pleasure working with him over the past two years, and I could not have asked for a better introduction into the field of seismology.

I would like to thank the Rhodes Trust at Oxford for making the experience possible from a financial standpoint. I am also grateful for their financial support in attending meetings at AGU 2002 in San Francisco and GSA 2001 in Boston. Worcester College provided support for both of these meetings as well.

I have benefitted from discussions with Ana Ferreira, and I look forward to reading her thesis in two years' time ("Surface Waves in a Laterally Varying Earth"). She also generated the data for the surface wave dispersion figures in Appendix F. Gareth Funning helped in downloading the IRIS data for Figure 1.3. I would also like to thank Paul Davis for helping me get started with wave propagation in Matlab, and my dad Walt for helping to clarify some of the details in Appendix A. Also, I have benefited from discussions on finite-difference methods with Nick Trefethen at Oxford, as well as from his two-semester course on scientific computing. Finally, thanks to my MSc examiners, Jeannot Trampert at Utrecht and Barry Parsons at Oxford, whose critical reading helped to improve the thesis.

Contents

Abstract	i
Extended abstract	ii
Acknowledgements	vi
Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
2 The two-dimensional wave equation	8
2.1 Introduction	8
2.2 Waves on a rectangular membrane	9
2.3 Waves on a circular membrane	10
2.4 Numerical solutions of waves on a planar membrane	12
2.5 Waves on a spherical membrane	14
3 Analytical spherical membrane waves	18
3.1 Introduction	18
3.2 The solution for given forcing $f = f_1(\theta, \phi, t)$	21
3.3 The solution for given forcing $f = \frac{\partial}{\partial t} f_1(\theta, \phi, t)$	24
3.4 The solution for given initial shape $g(\theta, \phi)$	25
3.5 Discussion	27
4 Spherical hexagonal grids	34
4.1 Introduction	34
4.2 Constructing the hexagonal grids	37
4.3 Properties of the hexagonal grids	39
4.4 Calculating the Laplacian on the hexagonal grids	42
4.5 Error in the numerical approximation of the Laplacian	44
4.6 Order of accuracy for the numerical approximation	46
4.7 Variations and optimizations on the hexagonal grids	47
4.8 Summary	50
5 Numerical spherical membrane waves	60
5.1 Introduction	60
5.2 The phase velocity function $c(\theta, \phi)$	61
5.3 Numerical iteration for wave propagation	64
5.4 Numerical solutions for homogeneous and heterogeneous phase velocity maps	64

5.5 Stability and accuracy in the numerical model	66
6 Ray calculations on the sphere	79
6.1 Introduction	79
6.2 Surface wave ray-tracing equations	81
6.3 Path anomaly	82
6.4 Phase anomaly	84
6.5 Amplitude anomaly	87
6.6 Linearized ray theory	88
6.7 Exact ray-tracing: Ray diagram cross-section curves	91
6.8 Exact ray-tracing: Variation with Δ , l_{max} , ϵ , and T	93
7 Numerical model vs ray theory	113
7.1 Introduction	113
7.2 Exact ray theory vs linearized ray theory	115
7.3 Obtaining phase and amplitude anomalies from the numerical model	117
7.4 Ray theory validity condition	117
7.5 Numerical model vs ray theory	119
7.6 Summary and conclusions	121
A Functions on the sphere	138
B Spherical harmonic functions	144
B.1 Spherical harmonic conventions	144
B.2 Spherical harmonic conventions in Matlab	146
B.3 Spherical harmonic expansions	148
C Euler angles and rotations	151
D Derivations for selected equations	155
D.1 Equation (3.12)	155
D.2 Equations (3.27) and (3.33)	156
D.3 Equations (3.20) and (3.62)	158
D.4 Some definitions and properties	159
E Ray-tracing equations	161
E.1 Receiver at a fixed longitude	161
E.2 Equation (6.19) — $ds/d\phi$	167
E.3 Equation (6.11)	168
F Wave parameters and dispersion	170
G Flow chart of computer programs	172
Bibliography	174

List of Figures

1.1	Rayleigh waves and Love waves	5
1.2	Global Rayleigh waves	6
1.3	Alaska earthquake of November 2002	7
2.1	Rectangular and circular membranes	9
2.2	Waves on a rectangular membrane	15
2.3	Waves on a square membrane	16
2.4	Waves on a circular membrane	17
3.1	Initial-shape function $g(\theta, \phi)$ and source-time function $h(t)$	31
3.2	The width of the Gaussian initial-shape function	32
3.3	Surface plot of $I_l(\mu)$	32
3.4	Analytical solutions for wave propagation on the sphere	33
4.1	Discretizations of the sphere	36
4.2	Sketch of integration parameters	43
4.3	Parameter for optimizing the gridpoint spacing	48
4.4	Relationship between successive orders of triangular grids	51
4.5	Construction of the triangular grids	51
4.6	The order $q=0$ triangular grid	51
4.7	Construction of a hexagonal grid from a triangular grid	52
4.8	Spherical hexagonal grids	53
4.9	Geometric properties of the hexagonal grids, I	54
4.10	Geometric properties of the hexagonal grids, II	55
4.11	Numerical method for calculating the Laplacian at a gridpoint	55
4.12	Error in the numerical approximation of the Laplacian, I	56
4.13	Error in the numerical approximation of the Laplacian, II	57
4.14	Error in the numerical approximation of the Laplacian, III	58
4.15	Order of accuracy of the numerical approximation of the Laplacian	59
5.1	Phase velocity map: Homogeneous earth	62
5.2	Surface wave phase velocity map (vary T)	69
5.3	Phase velocity maps: Vary l_{max}	70
5.4	Phase velocity maps: Vary ϵ	71
5.5	Amplitude spectra for phase velocity maps	72
5.6	Phase velocity maps: TW95 vs TW96	73
5.7	Seismograms for a homogeneous velocity map	74

5.8	Seismograms for a heterogeneous velocity map	75
5.9	Waveforms for a heterogeneous phase velocity map	76
5.10	Wavefront propagating on a heterogeneous earth	77
5.11	Variation in seismograms with varying l_{max}	78
6.1	Source and receivers	80
6.2	Great-circle ray paths	82
6.3	Introduction to ray-tracing	98
6.4	Phase velocity map and transverse derivatives	99
6.5	Ray diagram plus cross-sections	101
6.6	Ray diagram cross-sections	102
6.7	Path deviation from great circles.	103
6.8	Ray diagram cross-sections: Variation with l_{max} , ϵ , and T	104
6.9	Zoom-in on Figure 6.8	105
6.10	Variation in amplitude anomaly with l_{max} , ϵ , and T	106
6.11	Variation in phase anomaly with l_{max} , ϵ , and T	107
6.12	Ray diagram cross-sections: TW95 vs TW96	108
7.1	Flowchart of seismogram generation	114
7.2	Exact ray theory vs linearized ray theory: Vary l_{max}	124
7.3	Exact ray theory vs linearized ray theory: Vary ϵ	125
7.4	Exact ray theory vs linearized ray theory: Vary T	126
7.5	Exact ray theory vs linearized ray theory: Vary orbit	127
7.6	Obtaining the amplitude and phase anomalies from the numerical model	128
7.7	Wavelength as a function of source width	129
7.8	Numerical results vs ray theory: $l_{max}=6$	130
7.9	Numerical results vs ray theory: $l_{max}=12$	131
7.10	Numerical results vs exact ray theory: Vary l_{max}	132
7.11	Numerical results vs exact ray theory: Vary λ	133
7.12	Numerical results vs exact ray theory: Vary T	134
7.13	Numerical results vs exact ray theory: Vary ϵ	135
7.14	Numerical results vs exact ray theory: Vary orbit	136
7.15	Shorter wavelengths versus shorter scalelengths	137
A.1	Function mappings for spherical coordinates	138
A.2	Extending $\tilde{g}(\theta, \phi)$ to the ϕ - θ plane, I	142
A.3	Extending $\tilde{g}(\theta, \phi)$ to the ϕ - θ plane, II	143
A.4	Extending $\tilde{g}(\theta, \phi)$ to the ϕ - θ plane, III	143
B.1	Sample functions of $Y_{lm}(\theta, \phi)$	146
C.1	Euler angle rotation	153
C.2	Euler angle rotation (in reverse)	154
E.1	Sketches for ray-tracing equations	169
F.1	Dispersion of surface waves	171
G.1	Flowchart of computer programs	173

List of Tables

4.1	Spherical hexagonal grids used in other studies	36
4.2	The five platonic solids	37
4.3	Geometric properties of the spherical grids	40
4.4	Properties of the hexagonal grids used in this study	40
4.5	Properties of the hexagonal grids of <i>Heikes and Randall</i> (1995a)	41
4.6	Subdivided cubic grids used in other studies	47
4.7	Optimizing the hexagonal grids	49
5.1	Properties of phase velocity maps of <i>Trampert and Woodhouse</i> (1995)	68
5.2	Minimum scalelength in a phase velocity map	68
6.1	Phase anomaly convention	87
6.2	Great-circle path lengths for global surface waves	97
6.3	Dependence of multipathing on orbit and l_{max}	109
6.4	Dependence of multipathing on ϵ and T	111
6.5	Dependence of multipathing on TW95 vs TW96	112
F.1	Wave parameters and some related equations	170

Chapter 1

Introduction

This is a study in global seismology, and because it leans on the theoretical side, it is important to establish its context in the earth sciences before going into any details. The basic goal of global seismology is to improve our physical picture of the earth’s interior, from the center to the surface. Our pictures of the earth’s interior — tomographic images — are expressed in terms of seismic velocities, which serve as proxies for physical properties of the material through which the seismic waves travel¹. The tomographic images (e.g., *Ritsema et al.*, 1999) give us a snapshot of what lies deep beneath our feet, and these snapshots provide the best starting point for discussion of topics such as mantle convection and plate-driving mechanisms (e.g., *Albarede and van der Hilst*, 2002). Thus an important goal of global seismology is to map the earth’s interior with as much detail as possible, thereby establishing the images essential for discussing and modelling the dynamic processes of the earth’s interior.

Surface waves provide the best constraints on the structure² of one of the most interesting regions of the earth, the upper mantle, which extends to a depth of approximately 700 km. Long-period surface waves — periods approximately in the range $T = 20$ to 300s — sample the upper mantle, hence their label as “mantle waves”. These waves provide insight into

¹In particular, the expressions for the compressional velocity (v_p) and shear velocity (v_s) in an isotropic solid are given by:

$$v_p = \sqrt{\frac{\kappa + \frac{4}{3}\mu}{\rho}} \quad v_s = \sqrt{\frac{\mu}{\rho}},$$

where μ is the shear modulus, κ is the bulk modulus, and ρ is the density.

²By detail of structure, we mean the detail of the seismic velocity coverage. The detail in velocity structure could be a manifestation of the thermal structure or chemical structure (or a combination). The term “structure” in this thesis will generally refer to the “structure” of the phase velocity variation.

the lateral heterogeneity and anisotropy within the upper mantle. Lateral heterogeneity is helpful in characterizing subduction zones, plumes, and the fate of subducted material, while anisotropy in the upper mantle is one of the indicators of mantle flow direction, which is considered to be a principal driver of the plates.

There are two types of surface waves: Rayleigh waves, which propagate with vertical and radial displacements, and Love waves, which propagate with transverse displacements only (Figure 1.1). After large earthquakes, such as the one in central Alaska in November 2002 (M7.9), the surface waves may be detected traveling around the earth multiple times. Figure 1.2 shows a theoretical example of these global surface waves, and Figure 1.3 shows that these surface waves are readily identifiable in seismograms. It is data such as these to which theories in seismology are applied in order to construct the images of the earth's interior.

Our three-dimensional picture of the earth has improved drastically over the past 25 years. The development of the Preliminary Reference Earth Model (PREM), a 1D earth model that remains the standard, marked a starting point for tomography (*Dziewonski and Anderson*, 1981). *Woodhouse and Dziewonski* (1984) produced the first 3D global images of the Earth's interior, and today there exist several 3D models with greater detail of structure (e.g., *Ritsema et al.*, 1999). Many models of the upper mantle rely on surface wave data (an exception is *Grand* (2002)), and some are constructed from surface wave data alone³ (*Boschi and Ekström*, 2002). The Caltech–Oxford model (*Ritsema et al.*, 1999), one of the most accepted global models today, incorporates surface waves, including higher-mode surface wave arrivals. Improvements in the global models have come with improvements in theory and with improvements in data coverage and data quality. In this study, we look at one of the most commonly used tools in seismology, ray theory.

The purpose of this study is to develop a numerical model for wave propagation on a spherical membrane and to use the results to check predictions made from ray theory. We use this model as an analogue for surface wave propagation, in order to investigate the properties of ray theory for wave propagation on a sphere. Ray theory has been used in seismology to explain observations of path differences, amplitude differences, and phase differences of surface waves with respect to a spherical reference model (e.g., *Woodhouse and Wong*, 1986;

³A global model of the crustal thickness down to the Moho is crucial in these models (*Bassin et al.*, 2000; *Mooney et al.*, 1998).

Laske et al., 1994). Furthermore, it is a standard tool used in creating the global models of the earth. However, ray theory is an approximate theory and cannot always be used to explain observations. Thus, numerical models are helpful, especially as earth models become increasingly more detailed. This study presents a two-dimensional numerical model aimed at investigating the regimes under which ray theory predictions fail.

Here we will emphasize some key points, and then outline the remainder of this report.

1. This study is divided into two approaches to examining wave propagation on a spherical membrane: (1) a numerical model, in which the wave propagates on the spherical membrane according to the wave equation, and (2) surface wave ray theory, whereby the rays, confined to the spherical surface, are the orthogonal trajectories of the wavefront of the surface wave. The numerical solution provides the amplitude information in the form of a global “snapshot” at each timestep, while ray theory provides amplitude and phase information for a particular source-and-receiver pair. Our goal is to use the numerical model to examine the limitations of ray theory.
2. Our problem is mathematically a two-dimensional problem, as the term “membrane” suggests. The propagation of a Rayleigh wave induces displacements down to depths of hundreds of km, but we are only concerned with the displacement of the surface, which can be adequately modelled by the wave equation. Thus our membrane waves are an analogue for the surface waves. The surface wave phase velocity maps $c(\theta, \phi)$ that we input into the numerical model and into the ray tracing include information about the 3D structure of the upper mantle.
3. The mathematics of our problem is stripped of the complexity of wave propagation that we know to exist in response to the 3D, laterally heterogeneous, anisotropic, and anelastic structure of the upper mantle. In nature these effects give rise to dispersion, scattering, reflection, anisotropy, and attenuation of the surface waves⁴. Our wave equation simply includes a forcing term $f(\theta, \phi, t)$ for the source and a phase velocity map of the earth $c(\theta, \phi)$.

⁴We do, however, encounter several of these effects in our numerical model as a result of discretizing a continuous problem; numerical dispersion is the most predictable artifact of the numerical solution.

In this thesis we first look at analytical and numerical solutions to waves on a planar membrane (Chapter 2), which provides context for examining analytical solutions to waves on a spherical membrane (Chapter 3). In order to calculate numerical solutions to waves on a spherical membrane, we need a spherical grid that will allow us to numerically approximate the Laplacian (∇^2) at each time-step. After constructing and testing the grid (Chapter 4), we then compare the numerical solutions of waves on the spherical membrane with analytical solutions for the homogeneous case (i.e., phase velocity constant everywhere) (Chapter 5). We then find numerical solutions for the heterogeneous case, using phase velocity maps of *Trampert and Woodhouse (1995)* (Chapter 5). Lastly, we use ray theory to predict the phase and amplitude variation of the waves on the spherical membrane (Chapter 6), and we investigate how these predictions compare with those from the numerical model (Chapter 7).

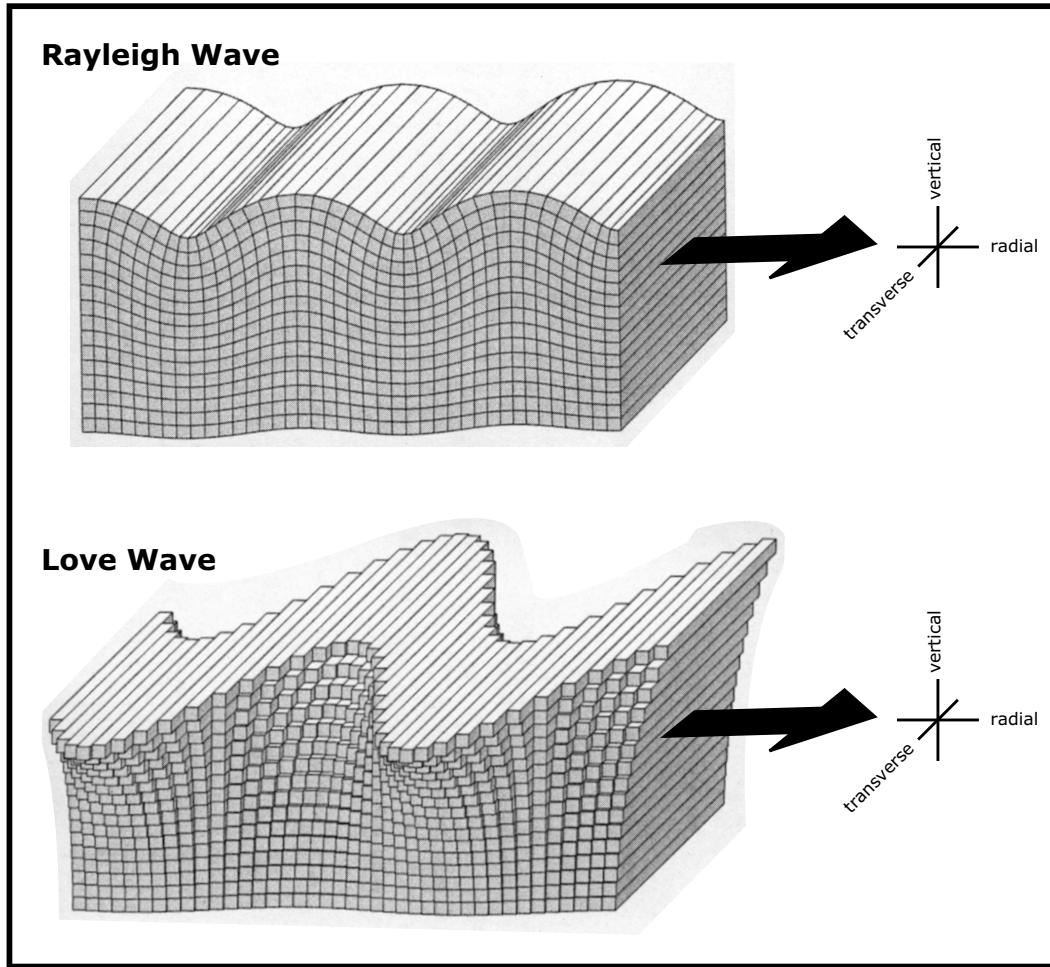


Figure 1.1: Ground displacements for Rayleigh waves and Love waves (after Shearer, 1999). The direction of propagation in each case is in the radial direction. Displacement of Rayleigh waves is restricted to the vertical and radial directions; displacement of Love waves is restricted to the transverse direction only. These surface waves travel 4–5 km/s — see Appendix F for more details.

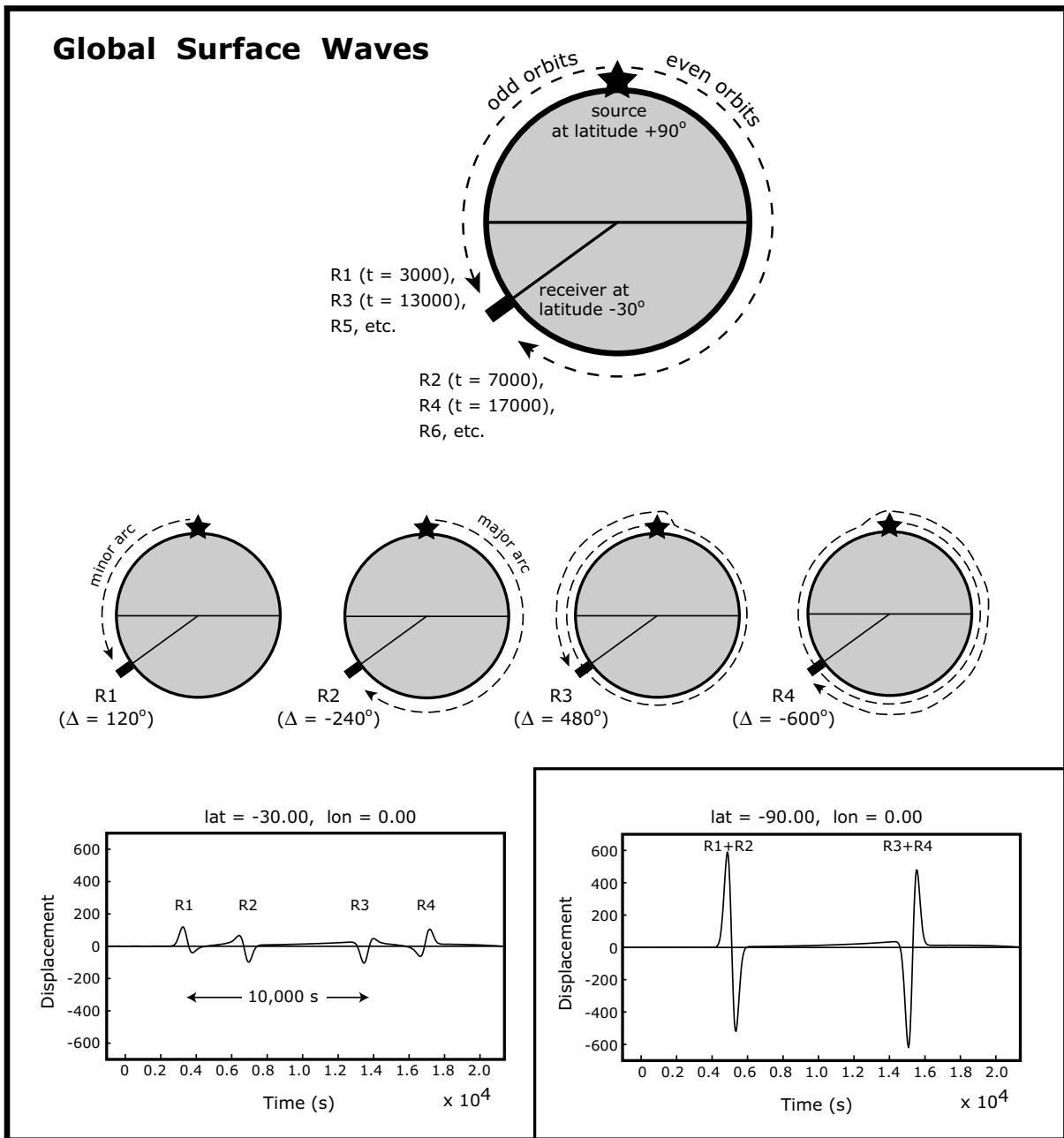


Figure 1.2: Global surface waves. Schematic diagram illustrating the nomenclature of global surface wave orbits, depicted for a source at the North Pole and a receiver at $\text{lat} = -30^\circ$. R denotes Rayleigh waves ($G = \text{Love waves}$), and each number corresponds to the relative arrival of the orbit at a particular receiver (see theoretical seismogram). The time between even or odd orbits is the time it takes for the wave to propagate around the earth ($\approx 10,000$ seconds, or ≈ 2.8 hours). Inset shows a seismogram for the special case of the receiver located at the antipode, the point on the earth directly opposite the source. In that case, the arrivals constructively interfere, and the observer experiences two pulses, rather than four.

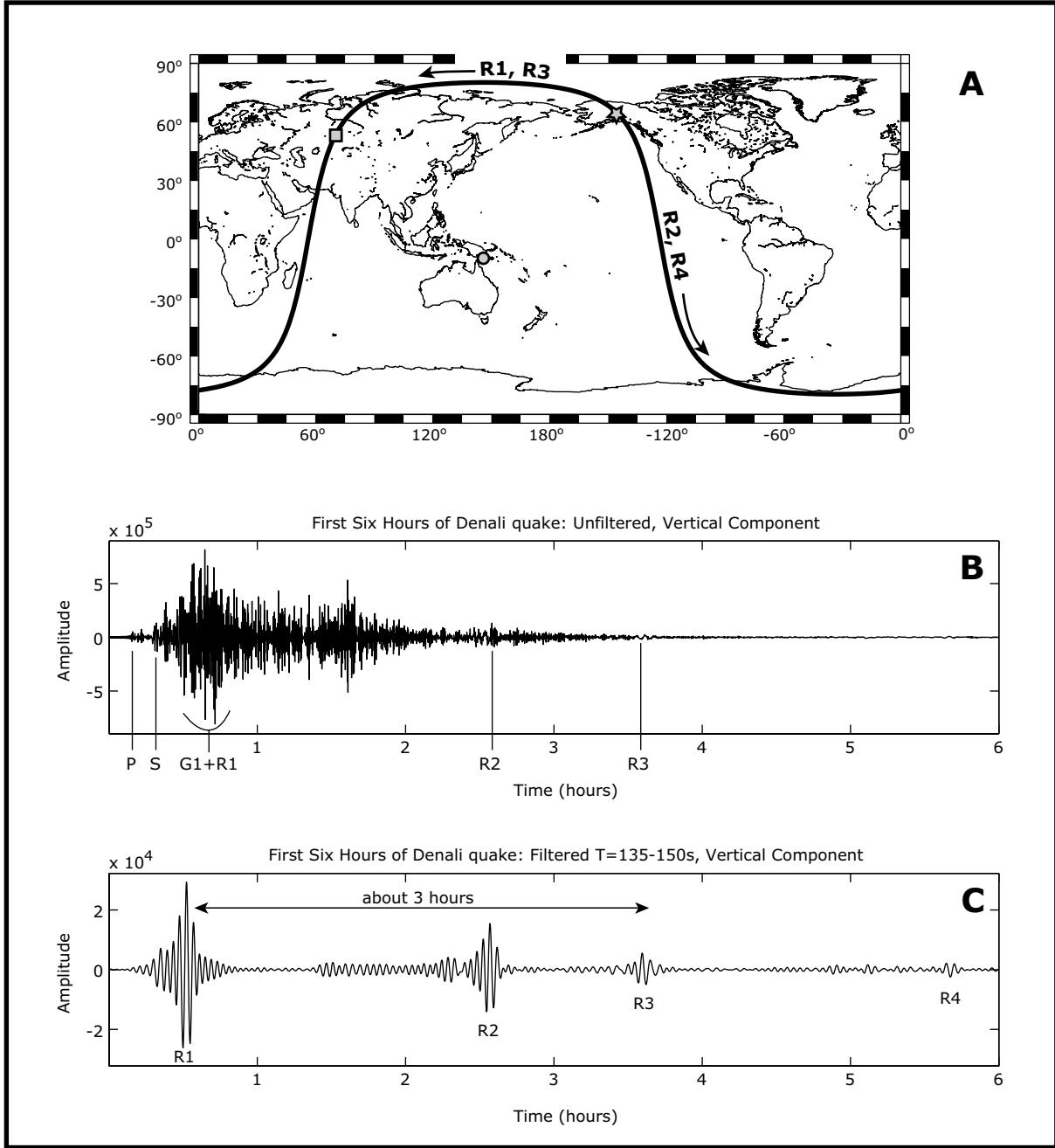


Figure 1.3: The M7.9 Alaska earthquake of November 3, 2002 (lat= -63.42° , lon= -144.96°). Here we show the event recorded in Borovoye, Kazakhstan (BRVK) (lat= 53.06° , lon= 70.28°). **A.** Great circle path containing the source (\star) and receiver (\square); also shown is the pole to the great circle (\circ). **B.** Unfiltered record of the first six hours. Note the relative arrival times and relative amplitudes of the P -wave, the S -wave, and the surface waves (R1, G1). **C.** Filtered record of the first six hours. The long-period Rayleigh waves are recorded orbiting the earth in both directions more than once (see Figure 1.2). The time for these elastic waves to circle the earth is about three hours. The great-circle path lengths are as follows: $\Delta_{R1} = 60.32^\circ$, $\Delta_{R2} = -299.68^\circ$, $\Delta_{R3} = 420.32^\circ$, $\Delta_{R4} = -659.68^\circ$.

Chapter 2

The two-dimensional wave equation

Contents

2.1	Introduction	8
2.2	Waves on a rectangular membrane	9
2.3	Waves on a circular membrane	10
2.4	Numerical solutions of waves on a planar membrane	12
2.5	Waves on a spherical membrane	14

2.1 Introduction

The wave equation is given by

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \nabla^2 u , \quad (2.1)$$

where c is the phase velocity of the wave field u . In the homogeneous case considered in this chapter, c is a constant.

In the following sections we will examine the solutions to Equation (2.1) for waves (i) on a rectangular membrane and (ii) on a circular membrane (Figure 2.1). The boundary condition in these cases is that the edge of the membrane is fixed for all $t \geq 0$. These cases will provide context for studying waves on a spherical membrane. Details of the following analytical solutions for waves on rectangular and circular membranes can be found in *Kreyszig* (1988) and references therein.

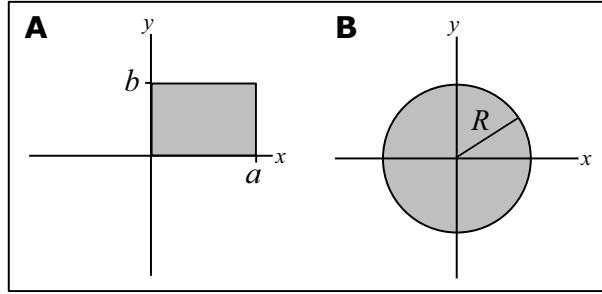


Figure 2.1: Rectangular and circular membranes for the two-dimensional wave equation.

2.2 Waves on a rectangular membrane

The two-dimensional wave equation for waves in the xy -plane is

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \quad (2.2)$$

where $u(x, y, t)$ is the displacement in the vertical direction. Our boundary condition is $u = 0$ for $t \geq 0$ on the boundary of a length a -by- b rectangular membrane (Figure 2.1A). Our initial conditions consist of a given initial displacement and initial velocity:

$$u(x, y, 0) = g(x, y) \quad (2.3)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = v(x, y). \quad (2.4)$$

The general solution is the two-dimensional Fourier sine series

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} u_{mn}(x, y, t) \quad (2.5)$$

$$= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} (B_{mn} \cos \lambda_{mn} t + B_{mn}^* \sin \lambda_{mn} t) \sin \left(\frac{m\pi x}{a} \right) \sin \left(\frac{n\pi y}{b} \right), \quad (2.6)$$

where

$$\lambda_{mn} = c\pi \sqrt{\frac{m^2}{a^2} + \frac{n^2}{b^2}}. \quad (2.7)$$

Using the initial conditions—Equations (2.3) and (2.4)—in Equation (2.6), we obtain

$$u(x, y, 0) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{mn} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) = g(x, y) \quad (2.8)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} B_{mn}^* \lambda_{mn} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) = v(x, y). \quad (2.9)$$

Supposing that $g(x, y)$ and $v(x, y)$ can each be developed in such a series, Equations (2.8) and (2.9) can be used to determine the coefficients B_{mn} and B_{mn}^* in Equation (2.6) (e.g., Kreyszig, 1988):

$$B_{mn} = \frac{4}{ab} \int_0^b \int_0^a g(x, y) \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) dx dy \quad (2.10)$$

$$B_{mn}^* = \frac{4}{ab \lambda_{mn}} \int_0^b \int_0^a v(x, y) \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) dx dy. \quad (2.11)$$

We can plot Equation (2.6) once we specify the initial conditions $g(x, y)$ and $v(x, y)$. Let us choose an initial shape as a Gaussian at the center of the (a -by- b) rectangular membrane, and choose an initial velocity of zero:

$$u(x, y, 0) = g(x, y) = \frac{1}{\mu^2} e^{-((x-a/2)^2 + (y-b/2)^2) / \mu^2} \quad (2.12)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = v(x, y) = 0, \quad (2.13)$$

where μ is included as a normalization constant. Plots of Equation (2.6) for three different times are shown in Figure 2.2. If we use a square membrane instead of a rectangle, then we see a more symmetric reflection of the wave for the initial Gaussian (Figure 2.3).

2.3 Waves on a circular membrane

We now look at the case with a circular membrane of radius R (Figure 2.1B). Using the polar-coordinate form of the Laplacian, the wave equation is

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}. \quad (2.14)$$

Considering solutions $u(r, t)$ that are radially symmetric (i.e., no θ -dependence), Equation (2.14) reduces to

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r}. \quad (2.15)$$

With the membrane fixed at $r = R$, we have the boundary condition $u(R, t) = 0$ for $t \geq 0$. The initial position and velocity do not depend on θ and are given by

$$u(r, 0) = g(r) \quad (2.16)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = v(r). \quad (2.17)$$

The general solution is the series

$$u(r, t) = \sum_{m=1}^{\infty} u_m(r, t) \quad (2.18)$$

$$= \sum_{m=1}^{\infty} (a_m \cos \lambda_m t + b_m \sin \lambda_m t) J_0 \left(\frac{\alpha_m r}{R} \right), \quad (2.19)$$

where $\lambda_m = c \alpha_m / R$, and α_m are the positive (real) roots of $J_0(x)$, which is the Bessel function of the first kind of order zero:

$$J_0(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(j!)^2} \left(\frac{x}{2} \right)^{2j} = 1 - \frac{x^2}{2^2} + \frac{x^4}{2^2 4^2 6^2} - \frac{x^6}{2^2 4^2 6^2 8^2} + \dots. \quad (2.20)$$

Using the initial conditions—Equations (2.16) and (2.17)—in Equation (2.19), we obtain

$$u(r, 0) = \sum_{m=1}^{\infty} a_m J_0 \left(\frac{\alpha_m r}{R} \right) = g(r) \quad (2.21)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = \sum_{m=1}^{\infty} b_m \lambda_m J_0 \left(\frac{\alpha_m r}{R} \right) = v(r). \quad (2.22)$$

Supposing that $g(r)$ and $v(r)$ can each be developed in such a series, Equations (2.21) and (2.22) can be used to determine the coefficients a_m and b_m of the Fourier-Bessel series in

Equation (2.19) (*Kreyszig*, 1988):

$$a_m = \frac{2}{R^2 (J_1(\alpha_m))^2} \int_0^R r g(r) J_0\left(\frac{\alpha_m r}{R}\right) dr \quad (2.23)$$

$$b_m = \frac{2}{c \alpha_m R (J_1(\alpha_m))^2} \int_0^R r v(r) J_0\left(\frac{\alpha_m r}{R}\right) dr . \quad (2.24)$$

We can plot Equation (2.19) once we specify the initial conditions $g(r)$ and $v(r)$. Let us choose an initial shape as a Gaussian at the center of the circular membrane, and choose an initial velocity of zero:

$$u(r, 0) = g(r) = \frac{1}{\mu^2} e^{-r^2/\mu^2} \quad (2.25)$$

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = v(r) = 0 , \quad (2.26)$$

where μ is included as a normalization constant. Plots of Equation (2.19) for three different times are shown in Figure 2.4.

2.4 Numerical solutions of waves on a planar membrane

We now examine numerical solutions of the wave equation on a planar membrane. This will serve as a basis for examining numerical solutions for waves on a spherical membrane (Section 5.3). Here we examine the case of a rectangular membrane, and include a forcing term in the two-dimensional wave equation:

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = f , \quad (2.27)$$

where $f = f(x, y, t)$ is the forcing term, $u = u(x, y, t)$ is the wave field, represented as a vertical displacement, and $\nabla^2 u = \partial^2 u / \partial x^2 + \partial^2 u / \partial y^2$. The solutions are subject to the boundary condition $u = 0$ on the boundary of the membrane for all $t \geq 0$:

$$u(x, 0, t) = u(x, b, t) = 0 , \quad 0 \leq x \leq a \quad (2.28)$$

$$u(0, y, t) = u(a, y, t) = 0 , \quad 0 \leq y \leq b ,$$

where a and b describe the boundary of the rectangular membrane (Figure 2.1A).

The wave equation represents a continuous problem in space and time. In order to obtain numerical solutions, we must discretize the problem in both space and time. This involves discretizing the differential equations (and the functions within), as well as the boundary conditions. Here we use a finite-difference method in space and time to solve the 2D wave equation. Using the second-order “leap-frog” finite-difference approximation for \ddot{u} , u_{xx} , and u_{yy} , Equation (2.27) becomes

$$\ddot{u} = c^2 (\nabla^2 u + f) \quad (2.29)$$

$$\ddot{u} = c(x, y)^2 (u_{xx} + u_{yy} + f(x, y, t)) \quad (2.30)$$

$$\frac{u_{t+1}^{i,j} - 2u_t^{i,j} + u_{t-1}^{i,j}}{dt^2} = (c^{i,j})^2 \left(\frac{u_t^{i+1,j} - 2u_t^{i,j} + u_t^{i-1,j}}{dx^2} + \frac{u_t^{i,j+1} - 2u_t^{i,j} + u_t^{i,j-1}}{dx^2} + f_t^{i,j} \right), \quad (2.31)$$

where $u_t^{i,j}$ is the displacement at each gridpoint, t is the time increment index¹, i is the spatial increment index in the x -direction, j is the spatial increment index in the y -direction, dt is the time increment, and dx is the space increment² (e.g., Trefethen, 1996). Thus, Equation (2.30) is the continuous form of the wave equation, and Equation (2.31) is our discrete form. For simplicity, we shall substitute ‘`de12`’ for the discrete finite difference approximation of the Laplacian operator in the xy -plane:

$$\frac{u_{t+1}^{i,j} - 2u_t^{i,j} + u_{t-1}^{i,j}}{dt^2} = (c^{i,j})^2 (\text{de12}(u_t^{i,j}) + f_t^{i,j}), \quad (2.32)$$

which, solving for u_{t+1} , gives

$$u_{t+1} = 2u_t - u_{t-1} + c^2 dt^2 (\text{de12}(u_t) + f_t), \quad (2.33)$$

where we have eliminated the i, j superscripts for clarity.

¹Note that the subscripts in the index notation do *not* refer to partial differentiation.

²The spatial increment dx is the same in both directions, even though we have a rectangular domain. We could put a dy in the second denominator of Figure 2.31 to indicate the *direction* of the spatial step.

We used Equation (2.33) in a program to produce numerical solutions to the wave equation on a rectangular membrane. First we specify the forcing term $f = f(x, y, t)$ and the initial conditions $u_{t-1} = u_t = g(x, y)$. For each time-step, the iteration is:

$$u_{t-1} = u_t \quad (2.34)$$

$$u_t = u_{t+1} \quad (2.35)$$

$$u_{t+1} = 2u_t - u_{t-1} + c^2 dt^2 (\text{del2}(u_t) + f_t) \quad (2.36)$$

$$u_t|_{\text{boundary}} = 0. \quad (2.37)$$

With initial conditions given by Equation (2.12) and no forcing term ($f(x, y, t) = 0$) and $c(x, y) = \text{const}$, the numerical solutions for the rectangular membrane waves agreed with those for the analytical case (i.e., Figures 2.2 and 2.3).

It should be noted that there are various ways of obtaining numerical solutions to the wave equation. An alternative approach to using finite difference methods is to use spectral methods (or pseudo-spectral methods). An excellent treatment of these two methods as applied to the wave equation can be found in works by Trefethen (*Trefethen*, 1982, 1996, 2000). A third approach it to use finite element methods.

2.5 Waves on a spherical membrane

The work of *Tanimoto* (1990) was one of the motivations for this study. *Tanimoto* (1990) modelled waves on a planar membrane using a second-order finite-difference method on a region bounded by constant latitudes and longitudes. He first looked at surface wave propagation through California, comparing his numerical results with ray theory predictions and with seismograms for the Whittier Narrows earthquake of 01 October 1987.

With the California example, the area was small enough so that the curvature of the Earth was neglected. In order to consider wave propagation on a global scale, *Tanimoto* (1990) solved the wave equation on the sphere and incorporated absorbing boundary conditions (*Clayton and Engquist*, 1977), which “work fairly well up to about $\pm 75^\circ$ [latitude]” (p. 100). Our goal was to model waves on a spherical membrane without using a Cartesian grid, which always requires special boundary conditions for high latitudes.

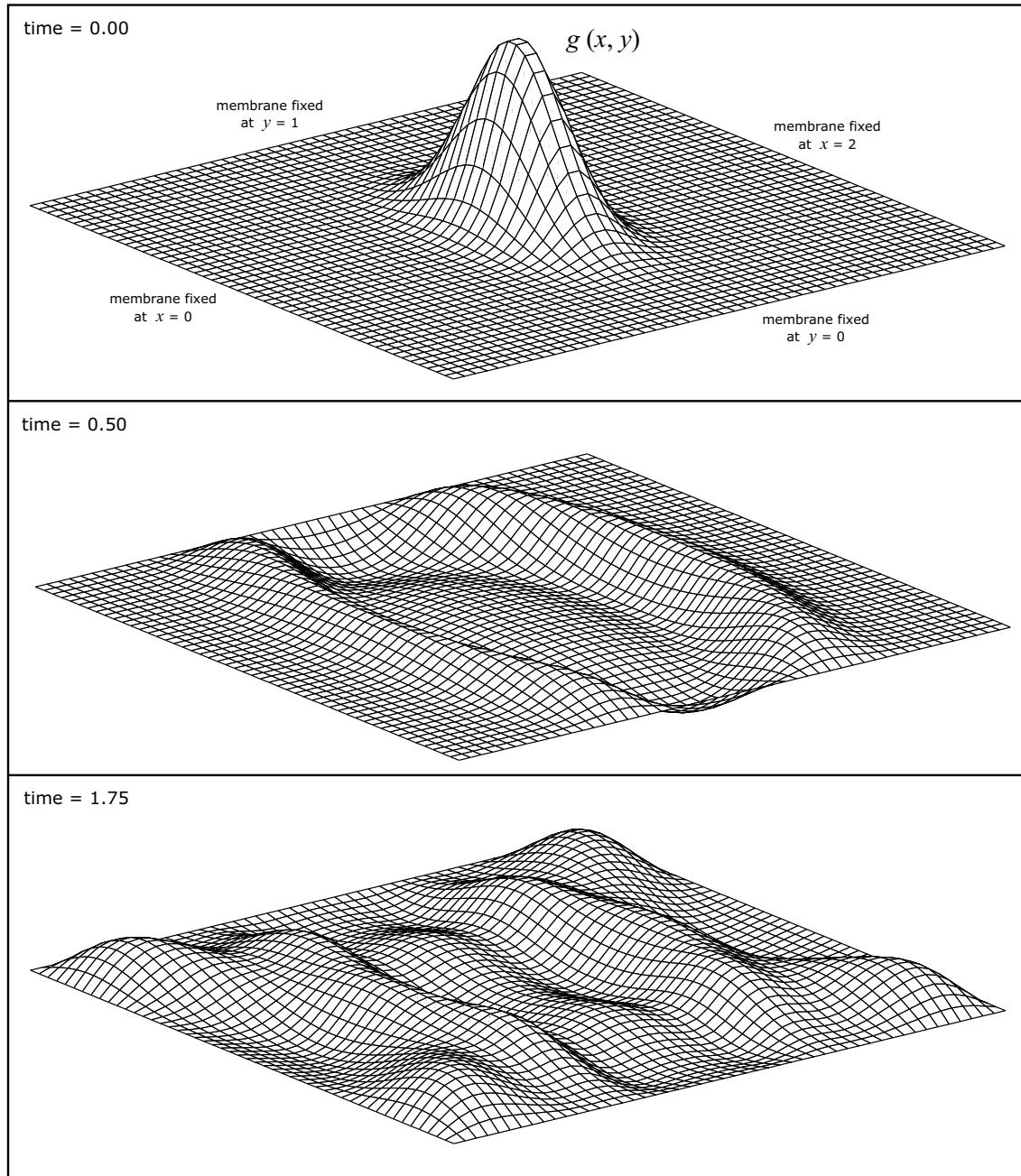


Figure 2.2: Analytical solutions of waves on a rectangular membrane. Three time frames for the initial conditions of (i) an initial shape of a Gaussian (Eq. 2.12 with $\mu = 0.1$; top figure) and (ii) an initial velocity of zero (Eq. 2.13). These are plots of Eq. 2.6 with $c = 1$, summing the modes up to $m = 20$ and $n = 20$. The boundary is fixed.

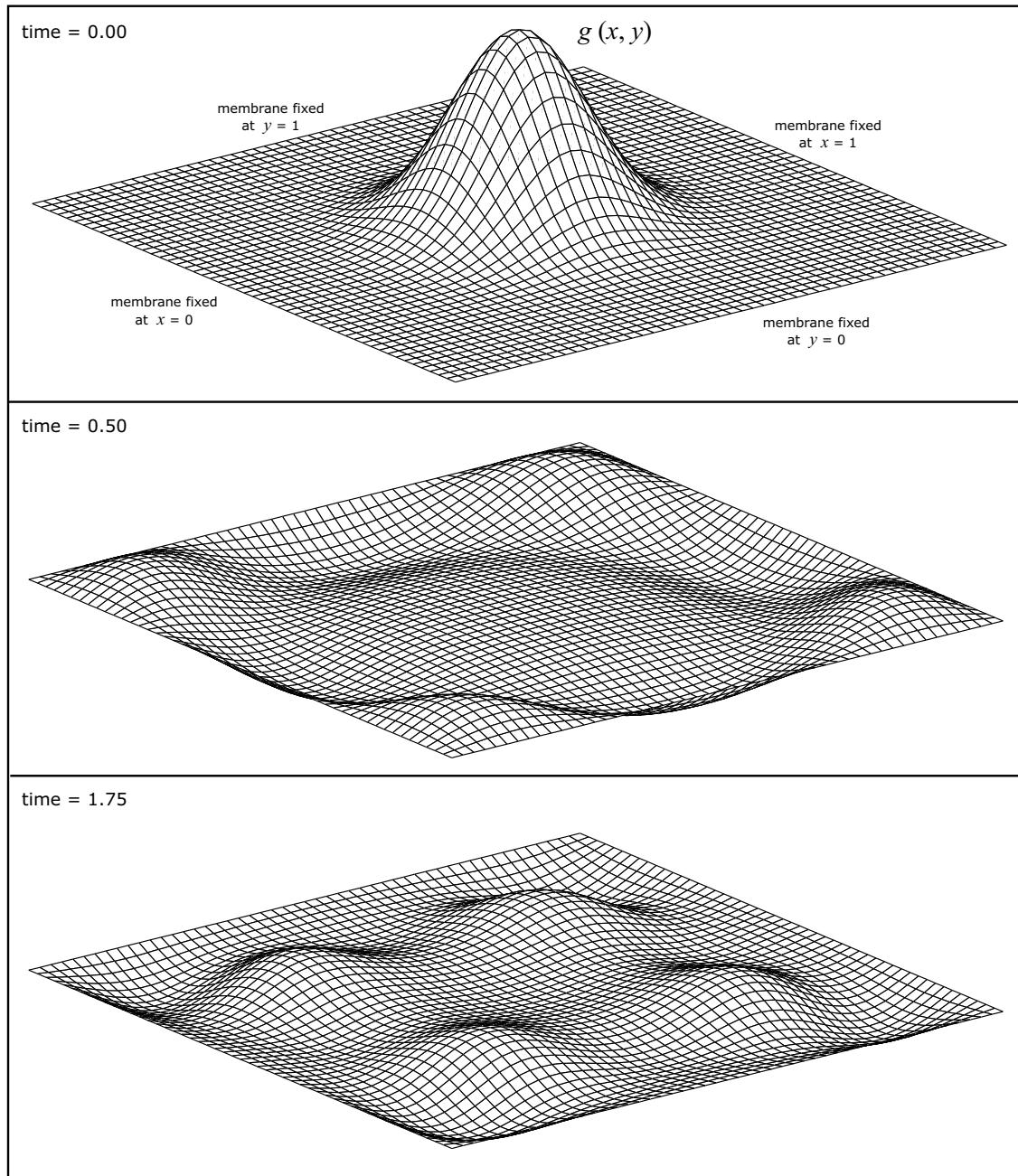


Figure 2.3: Analytical solutions of waves on a square membrane. The only difference from Figure 2.2 is that we have changed the width of the membrane to from $a = 2$ to $a = 1$. The initial Gaussian wave has a more symmetric reflection off the square boundary here than off the rectangular boundary in Figure 2.2, as expected.

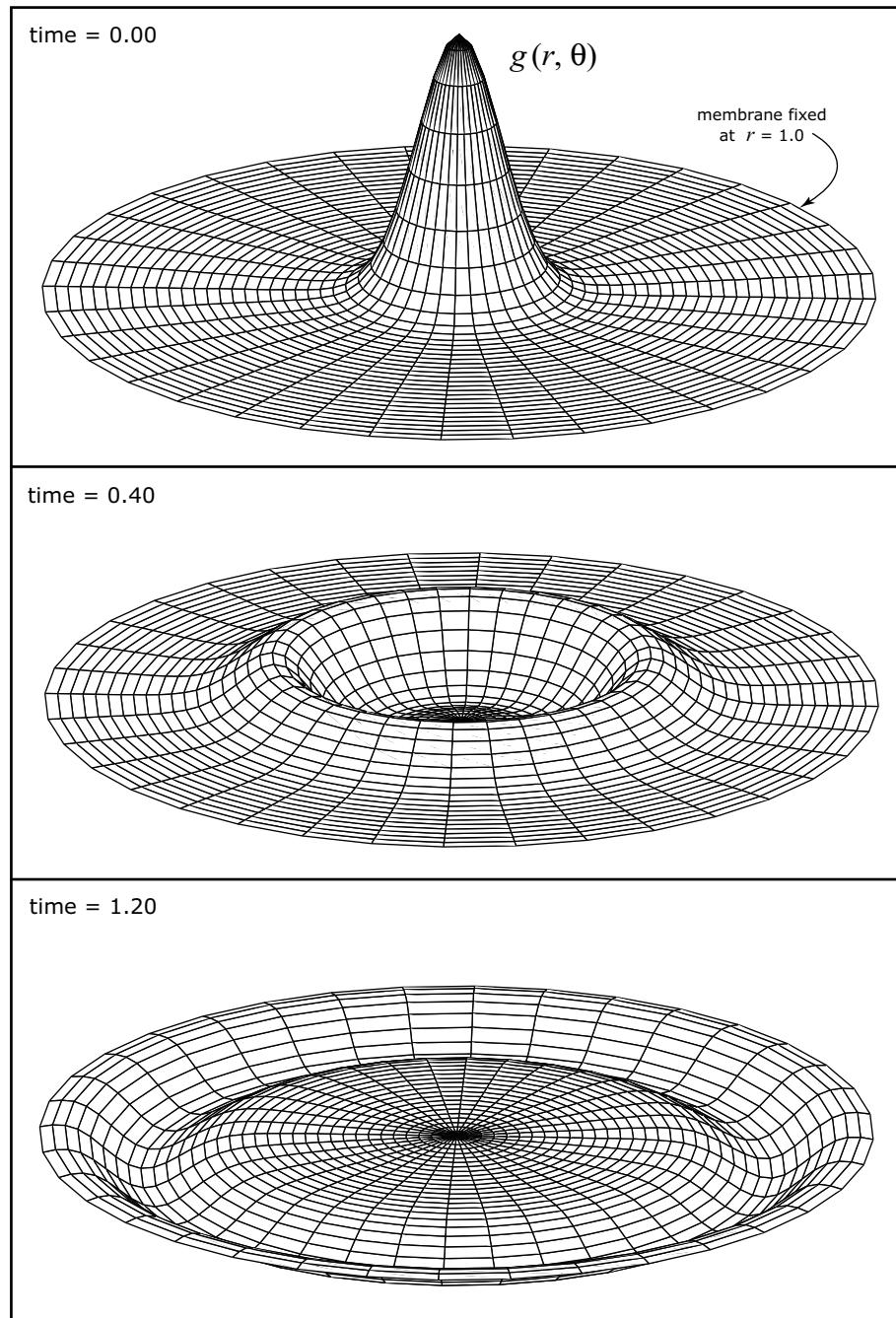


Figure 2.4: Analytical solutions of waves on a circular membrane. Three time frames for the initial conditions of (i) an initial shape of a Gaussian (Eq. 2.25 with $\mu = 0.1$; top figure) and (ii) an initial velocity of zero (Eq. 2.26). These are plots of Eq. 2.19 with $c = 1$, summing the modes up to $m = 20$. At $t = 1.20$, the wave has just reflected at the boundary and is returning on the underside. The boundary is fixed.

Chapter 3

Analytical solutions of waves on a spherical membrane

Contents

3.1	Introduction	18
3.2	The solution for given forcing $f = f_1(\theta, \phi, t)$	21
3.3	The solution for given forcing $f = \frac{\partial}{\partial t} f_1(\theta, \phi, t)$	24
3.4	The solution for given initial shape $g(\theta, \phi)$	25
3.5	Discussion	27

3.1 Introduction

This chapter presents analytical solutions to the wave equation on a spherical membrane. Some of the more important derivations are shown, while others are included in Appendix D. Appendix A examines the issue of functions being well-defined on the sphere, and Appendix B discusses spherical harmonic functions.

We wish to solve the following equation for $u = u(\theta, \phi, t)$:

$$\frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = f , \quad (3.1)$$

where u is the wave field, $c = c(\theta, \phi)$ is the frequency-independent, constant phase velocity,

f is the forcing term, and

$$\nabla^2 u = \frac{1}{a^2} \left(\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2} \right) \quad (3.2)$$

is the surface Laplacian of u on a sphere with radius a , where θ, ϕ are spherical coordinates (θ is colatitude). In our results, we plot the wave field u as a radial displacement on the sphere. It is assumed that the forcing term $f = f(\theta, \phi, t)$ is a given function, vanishing at (sufficiently large) negative times t . In this form, Equation (3.1) is the two-dimensional wave equation on the sphere. Note that attenuation and dispersion are not considered.

To solve Equation (3.1) we first expand f and u in spherical harmonics :

$$f = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_l^m(\theta, \phi) \quad (3.3)$$

$$u = \sum_{l=0}^{\infty} \sum_{m=-l}^l u_{lm} Y_l^m(\theta, \phi) \quad (3.4)$$

where $f_{lm} = f_{lm}(t)$ and $u_{lm} = u_{lm}(t)$ are given by

$$f_{lm} = \frac{1}{N_l} \int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) f(\theta, \phi, t) \sin \theta d\theta d\phi \quad (3.5)$$

$$u_{lm} = \frac{1}{N_l} \int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) u(\theta, \phi, t) \sin \theta d\theta d\phi , \quad (3.6)$$

where N_l is the normalization integral (assumed independent of m), such that

$$\int_0^{2\pi} \int_0^\pi Y_{l'}^{m'*}(\theta, \phi) Y_l^m(\theta, \phi) \sin \theta d\theta d\phi = N_l \delta_{l'l} \delta_{m'm} , \quad (3.7)$$

where

$$\delta_{l'l} \delta_{m'm} = \begin{cases} 1 & \text{if } l = l' \text{ and } m = m' \\ 0 & \text{otherwise .} \end{cases} \quad (3.8)$$

Equation (3.7) is a statement of the orthogonality of spherical harmonic functions, which are discussed in Appendix B. Substituting Equations (3.3) and (3.4) into Equation (3.1), we

obtain

$$\frac{1}{c^2} \frac{\partial^2}{\partial t^2} \left(\sum_{l=0}^{\infty} \sum_{m=-l}^l u_{lm} Y_l^m(\theta, \phi) \right) - \nabla^2 \left(\sum_{l=0}^{\infty} \sum_{m=-l}^l u_{lm} Y_l^m(\theta, \phi) \right) = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm} Y_l^m(\theta, \phi), \quad (3.9)$$

or, for each spherical harmonic term individually,

$$\begin{aligned} \frac{1}{c^2} \frac{d^2 u_{lm}}{dt^2} Y_l^m(\theta, \phi) - u_{lm} \nabla^2 Y_l^m(\theta, \phi) &= f_{lm} Y_l^m(\theta, \phi) \\ \frac{1}{c^2} \frac{d^2 u_{lm}}{dt^2} Y_l^m(\theta, \phi) + u_{lm} \frac{l(l+1)}{a^2} Y_l^m(\theta, \phi) &= f_{lm} Y_l^m(\theta, \phi) \\ \frac{1}{c^2} \frac{d^2 u_{lm}}{dt^2} + \frac{l(l+1)}{a^2} u_{lm} &= f_{lm} \\ \frac{d^2 u_{lm}}{dt^2} + \omega_l^2 u_{lm} &= c^2 f_{lm}, \end{aligned} \quad (3.10)$$

where

$$\omega_l \equiv \frac{c \sqrt{l(l+1)}}{a}. \quad (3.11)$$

The solution of Equation (3.10) can be obtained by the method of Green's functions, or by Laplace or Fourier transforms. The solution is, assuming that u_{lm} and du_{lm}/dt , like f_{lm} , vanish at (sufficiently large) negative times (see Appendix D):

$$u_{lm}(t) = c^2 \int_{-\infty}^t f_{lm}(\tau) \frac{\sin \omega_l(t-\tau)}{\omega_l} d\tau. \quad (3.12)$$

Substituting Equation (3.12) into Equation (3.4) gives

$$u = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(c^2 \int_{-\infty}^t f_{lm}(\tau) \frac{\sin \omega_l(t-\tau)}{\omega_l} d\tau \right) Y_l^m(\theta, \phi) \quad (3.13)$$

From Equation (3.5) we see that

$$f_{lm}(\tau) = \frac{1}{N_l} \int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) f(\theta, \phi, \tau) \sin \theta d\theta d\phi, \quad (3.14)$$

and we combine Equations (3.13), (3.14), and (3.7) to obtain

$$u = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(c^2 \int_{-\infty}^t \frac{\int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) f(\theta, \phi, \tau) \sin \theta d\theta d\phi}{\int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) Y_l^m(\theta, \phi) \sin \theta d\theta d\phi} \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) Y_l^m(\theta, \phi). \quad (3.15)$$

Equation (3.15) is our general expression for the wave field $u(\theta, \phi, t)$ on the surface of the sphere, given a forcing term f . In the next two sections we will examine the solutions for two different forcings, f_1 and f_2 .

3.2 The solution for given forcing $f = f_1(\theta, \phi, t)$

We assume that the forcing term $f = f(\theta, \phi, t)$ in Equation (3.1) has the form:

$$f_1(\theta, \phi, t) = h_1(t) g(\theta, \phi), \quad (3.16)$$

where

$$h_1(t) = \frac{e^{-t^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \quad (3.17)$$

is the *source-time function* and

$$g(\theta, \phi) = \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \quad (3.18)$$

is the *initial-shape function* (Figure 3.1). The denominators in Equations (3.17) and (3.18) are included for convenience only, to facilitate the consideration of the properties of the solutions as $\mu \rightarrow 0$ ($g \rightarrow \delta(\theta)$) or $\sigma \rightarrow 0$ ($h \rightarrow \delta(t)$) — the “point source” solutions (see Section 3.5). They are chosen by virtue of:

$$\int_{-\infty}^{\infty} \left(\frac{e^{-t^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) dt = 1 \quad (3.19)$$

$$\int_0^{\infty} \left(\frac{e^{-\theta^2/2\mu^2}}{\mu^2} \right) \theta d\theta = 1. \quad (3.20)$$

Intuitively, we would expect the wavelength of propagating the wave on the sphere to decrease if either σ or μ decrease. (This is shown later in Figure 7.7).

It is somewhat easier to interpret changes in μ on $g(\theta, \phi)$ if we write μ as a function of θ_{wid} , the angular width of the Gaussian at a specified fraction r of its maximum amplitude. If we set $g(\theta_r, \phi) = r/\mu^2$ and solve for μ in Equation (3.18), we get

$$\begin{aligned}\mu &= \frac{\theta_r}{\sqrt{-2 \ln r}} \\ \mu(\theta_{wid}, r) &= \frac{\theta_{wid}}{2\sqrt{-2 \ln r}},\end{aligned}\quad (3.21)$$

where we have set $\theta_{wid} \equiv 2\theta_r$ as shown in Figure 3.2. Thus, μ and θ_{wid} can be thought of interchangeably.

Because $f_1(\theta, \phi, t)$ is independent of ϕ , the solution will be cylindrically symmetric; thus $f_{lm} = 0$ for $m \neq 0$, and similarly $u_{lm} = 0$ for $m \neq 0$. With $m = 0$ and no ϕ -dependence, Equation (3.15) becomes

$$u = \sum_{l=0}^{\infty} \left(c^2 \int_{-\infty}^t \frac{2\pi \int_0^\pi Y_l^{0*}(\theta) f_1(\theta, \tau) \sin \theta d\theta}{2\pi \int_0^\pi Y_l^{0*}(\theta) Y_l^0(\theta) \sin \theta d\theta} \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) Y_l^0(\theta) \quad (3.22)$$

Writing $Y_l^0(\theta) = Y_l^{0*}(\theta) = P_l(\cos \theta)$ and substituting the forcing term (Eq. 3.16), we get

$$\begin{aligned}u_{f_1} &= \sum_{l=0}^{\infty} \left(c^2 \int_{-\infty}^t \frac{\int_0^\pi P_l(\cos \theta) h_1(\tau) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta d\theta}{\int_0^\pi P_l(\cos \theta) P_l(\cos \theta) \sin \theta d\theta} \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) P_l(\cos \theta) \\ &= c^2 \sum_{l=0}^{\infty} \left(\int_{-\infty}^t \frac{h_1(\tau) \int_0^\pi P_l(\cos \theta) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta d\theta}{\int_0^\pi (P_l(\cos \theta))^2 \sin \theta d\theta} \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) P_l(\cos \theta) \\ &= c^2 \sum_{l=0}^{\infty} \left(\frac{\int_0^\pi P_l(\cos \theta) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta d\theta}{\int_0^\pi (P_l(\cos \theta))^2 \sin \theta d\theta} \right) \left(\int_{-\infty}^t h_1(\tau) \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) P_l(\cos \theta).\end{aligned}\quad (3.23)$$

We use the notation u_{f_1} for this solution to emphasise that it is the solution corresponding to the forcing term $f_1(\theta, \phi, t)$. Clearly the result Equation (3.23) is independent of the normalization of the P_l functions. However, using the standard normalization, we have

$$\int_0^\pi (P_l(\cos \theta))^2 \sin \theta d\theta = \frac{1}{l + \frac{1}{2}}, \quad (3.24)$$

and Equation (3.23) becomes

$$u_{f_1} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \left(\int_{-\infty}^t h_1(\tau) \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) P_l(\cos \theta), \quad (3.25)$$

where

$$I_l(\mu) = \int_0^\pi P_l(\cos \theta) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta d\theta. \quad (3.26)$$

These integrals will need to be calculated numerically. Plots of these integrals are shown in Figure 3.3A, where we see that $I_l(\mu) \rightarrow 1$ as $l \rightarrow 0$ or $\mu \rightarrow 0$. Figure 3.3B shows how we select the upper limit l_{max} for the infinite summation in the displacement function u_f .

The time integral in Equation (3.25) will give a complicated result involving error functions of complex argument. However, the integral simplifies greatly if the upper limit is replaced by ∞ . This will be a good approximation for all times $t \gg \sigma$ — i.e., at times after the source has ceased to act. Using the source-time function $h_1(t)$ (Eq. 3.17), we find (see Appendix D)

$$\int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \right) \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau = \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} \quad (t \gg \sigma), \quad (3.27)$$

and Equation (3.25) becomes

$$u_{f_1}(\theta, \phi, t) = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma), \quad (3.28)$$

where the $I_l(\mu)$ are given by Equation (3.26). Note that the width of the source-time function, σ , and the width of the initial-shape function, μ , both influence the solution. The physical meaning of Equation (3.28) is that it is the seismogram for a fixed surface point, or a snapshot of the distorted spherical membrane for a fixed time.

The *spherical average* of the solution is calculated by integrating Equation (3.28) over the entire sphere. However, only the $l=0$ term gives a non-zero contribution:

$$\begin{aligned}
\bar{u}_{f_1}(t) &= \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi u_{f_1}(\theta, \phi, t) \sin \theta d\theta d\phi \\
&= \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi u_{f_1}(\theta, \phi, t)|_{l=0} \sin \theta d\theta d\phi \\
&= \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi c^2 \frac{1}{2} I_0(\mu) t \sin \theta d\theta d\phi \\
&= \frac{1}{2} c^2 \frac{1}{4\pi} I_0(\mu) t \cdot 2 \cdot 2\pi \\
&= \frac{1}{2} c^2 I_0(\mu) t,
\end{aligned} \tag{3.29}$$

where in the third step we have used

$$\begin{aligned}
u_{f_1}(\theta, \phi, t)|_{l=0} &= \lim_{l \rightarrow 0} \left[c^2 \left(l + \frac{1}{2} \right) I_l(\mu) \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \right] \\
&= c^2 \frac{1}{2} I_0(\mu) \lim_{l \rightarrow 0} \frac{\sin \omega_l t}{\omega_l} \\
&= c^2 \frac{1}{2} I_0(\mu) t.
\end{aligned}$$

Note that the spherical average increases linearly with time.

3.3 The solution for given forcing $f = \frac{\partial}{\partial t} f_1(\theta, \phi, t)$

Our forcing term thus far has been Equation (3.16). Let us now use a forcing term

$$f_2(\theta, t) = \frac{\partial}{\partial t} f_1(\theta, t) = h_2(t) g(\theta), \tag{3.30}$$

where

$$h_2(t) = \frac{d}{dt} h_1(t) = \frac{-t}{\sigma^2} \frac{e^{-t^2/2\sigma^2}}{\sqrt{2\pi}\sigma}. \tag{3.31}$$

The forcing term is again cylindrically symmetric, as we have not changed the initial-shape function $g(\theta)$; the source-time function $h_2(t)$ is shown in Figure 3.1. In this case the expression

for displacement, Equation (3.25), is given by

$$u_{f_2} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \left(\int_{-\infty}^t \frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \frac{\sin \omega_l(t-\tau)}{\omega_l} d\tau \right) P_l(\cos \theta), \quad (3.32)$$

where we have simply substituted the forcing term $h_2(\tau)$. In place of Equation (3.27) we have (see Appendix D)

$$\int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi}\sigma} \right) \frac{\sin \omega_l(t-\tau)}{\omega_l} d\tau = \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2} \quad (t \gg \sigma), \quad (3.33)$$

and in place of Equation (3.28) we have

$$u_{f_2}(\theta, \phi, t) = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma). \quad (3.34)$$

In this case, the spherical average (compare Eq. 3.29) is constant:

$$\bar{u}_{f_2}(t) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi u_{f_2}(\theta, \phi, t) \sin \theta d\theta d\phi = \frac{1}{2} c^2 I_0(\mu). \quad (3.35)$$

Figure 3.4 shows plots of Equation (3.34) at two particular times. As the wave pulse spreads out over the sphere, the amplitude of the pulse varies approximately as $K/\sqrt{\sin \theta}$, where K is the maximum displacement of the pulse at the equator ($\theta = \pi/2$).

3.4 The solution for given initial shape $g(\theta, \phi)$

We now look at the case with zero forcing, but with a prescribed initial shape. The solution in the case that $u(\theta, \phi, t)$ and $\partial u(\theta, \phi, t)/\partial t$ are given at time $t = 0$, will be similar, except that in place of Equation (3.10) we have

$$\frac{1}{c^2} \frac{d^2 u_{lm}}{dt^2} + l(l+1) u_{lm} = 0, \quad (3.36)$$

subject to initial conditions that u_{lm} and du_{lm}/dt have prescribed values at time $t = 0$. Let us assume that $u(\theta, \phi, 0) = g(\theta, \phi, t)$ is given, and that the time derivative is zero, at $t = 0$. (These are analogous to the initial conditions we set for waves on the circular membrane, e.g.,

Eqs. 2.16 and 2.17.) Then we require the solution, $u_{lm}(t)$ of Equation (3.36) satisfying

$$u_{lm}|_{t=0} = g_{lm} \quad (3.37)$$

$$\frac{du_{lm}}{dt}\Big|_{t=0} = 0, \quad (3.38)$$

where g_{lm} are the spherical harmonic coefficients of the given initial-shape function:

$$g_{lm} = \frac{1}{N_l} \int_0^{2\pi} \int_0^\pi Y_l^{m*}(\theta, \phi) g(\theta, \phi) \sin \theta \, d\theta \, d\phi. \quad (3.39)$$

In this case, the solution (to Eq. 3.36) is simply

$$u_{lm}(t) = g_{lm} \cos \omega_l t, \quad (3.40)$$

which can be compared to Equation (3.12). Let us assume we have the initial-shape function $g(\theta, \phi, t)$ given by Equation (3.18). Again, the terms for $m \neq 0$ will all vanish, and we find:

$$u_g(\theta, \phi, t) = \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t P_l(\cos \theta). \quad (3.41)$$

We use the notation u_g to emphasise that this is the solution corresponding to the given initial shape $g(\theta, \phi, t)$. Note that for $t=0$ the solution is simply $g(\theta)$.

The spherical average is again constant (compare Eq. 3.35):

$$\bar{u}_g(t) = \frac{1}{2} I_0(\mu). \quad (3.42)$$

Comparing Equations (3.29), (3.35), and (3.42), we find:

$$\bar{u}_{f_2} = c^2 \bar{u}_g \quad (3.43)$$

$$\bar{u}_{f_1} = t \bar{u}_{f_2} = t c^2 \bar{u}_g, \quad (3.44)$$

where the f_i -subscript denotes the forcing term used for the displacement function u_{f_i} .

3.5 Discussion

Equations (3.28), (3.34), and (3.41) give the solutions to three problems: (i) with forcing f_1 , (ii) with forcing $f_2 = \partial f_1 / \partial t$, and (iii) with given initial shape g (and no forcing):

$$u_{f_1}(\theta, \phi, t) = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma) \quad (3.45)$$

$$u_{f_2}(\theta, \phi, t) = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma) \quad (3.46)$$

$$u_g(\theta, \phi, t) = \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t P_l(\cos \theta), \quad (3.47)$$

where $I_l(\mu)$ is given by Equation (3.26) and ω_l is given by Equation (3.11).

In the case of the forcing term f_1 , the time-derivative of the solution is given by

$$\begin{aligned} \dot{u}_{f_1} &= \frac{d}{dt} \left(c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \right) \\ &= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{d}{dt} \left(\frac{\sin \omega_l t}{\omega_l} \right) e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \\ &= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma). \end{aligned} \quad (3.48)$$

Thus, examining Equations (3.46) and (3.48), we have the relationship

$$\dot{u}_{f_1} = u_{f_2} \quad (t \gg \sigma) \quad (3.49)$$

$$\dot{u}_f|_{h=h_1(t)} = u_f|_{h=\dot{h}_1(t)}, \quad (t \gg \sigma), \quad (3.50)$$

i.e., the time-derivative of the displacement due to a Gaussian source-time function $h_1(t)$ is equal to the displacement due to a time-derivative Gaussian source-time function $\dot{h}_1(t)$. Equation (3.50) comes naturally by differentiating Equation (3.1) with respect to time:

$$\frac{1}{c^2} \frac{\partial^2 \dot{u}}{\partial t^2} - \nabla^2 \dot{u} = \dot{f} = g(\theta, \phi) \dot{h}_1(t). \quad (3.51)$$

Solutions for an impulsive source ($\sigma = 0$)

We now look at the case where $\sigma = 0$ (impulsive force). In this limit, the Gaussian source-time function $h_1(t)$ becomes a delta function. Equation (3.45) and Equation (3.48) become

$$u_{f10} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{\sin \omega_l t}{\omega_l} P_l(\cos \theta) \quad (t \gg 0) \quad (3.52)$$

$$\dot{u}_{f10} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t P_l(\cos \theta) \quad (t \gg 0), \quad (3.53)$$

where $u_{f10} = u_{f10}(\theta, \phi, t)$ is the displacement due to an impulsive force with source-time function $h_1(t)$. The same process follows for the solutions due to the source-time function $h_2(t)$ (Eq. 3.31):

$$\dot{u}_{f2} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) (-\omega_l \sin \omega_l t) e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma) \quad (3.54)$$

$$u_{f20} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l t P_l(\cos \theta) \quad (t \gg 0) \quad (3.55)$$

$$\dot{u}_{f20} = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) (-\omega_l \sin \omega_l t) P_l(\cos \theta) \quad (t \gg 0). \quad (3.56)$$

Examining Equations (3.55), (3.53), and (3.47), we have the following relationship:

$$u_{f20} = \dot{u}_{f10} = c^2 u_g \quad (t \gg 0). \quad (3.57)$$

The effect of nonvanishing σ is to introduce a frequency decay, arising from the convolution of the instantaneous solution u_{f10} with the source-time function $h_1(t)$ (Eq. 3.17). Thus, for

general σ , we have

$$u_{f_1} = u_{f_10} * h_1(t) \quad (D.33)$$

$$\dot{u}_{f_1} = \frac{d}{dt} (u_{f_10} * h_1(t))$$

$$= \dot{u}_{f_10} * h_1(t) \quad (D.31)$$

$$= c^2 h_1(t) * u_g \quad (3.57), (D.27)$$

$$= c^2 \int_{-\infty}^t h_1(\tau) u_g(\theta, \phi, t - \tau) d\tau \quad (D.26)$$

$$= c^2 \int_{-\infty}^t h_1(\tau) \left(\sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \cos \omega_l(t - \tau) P_l(\cos \theta) \right) d\tau \quad (3.47)$$

$$= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \left(\int_{-\infty}^t h_1(\tau) \cos \omega_l(t - \tau) d\tau \right) P_l(\cos \theta), \quad (3.59)$$

where '*' denotes convolution (Eq. D.26), and it is understood that u_g is taken to vanish for negative times. Equation (3.59) is in agreement with the result obtainable by taking the time derivative in Equation (3.23):

$$\begin{aligned} \dot{u}_{f_1} &= \frac{d}{dt} \left(c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \int_{-\infty}^t h_1(\tau) \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau P_l(\cos \theta) \right) \\ &= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \frac{d}{dt} \left(\int_{-\infty}^t h_1(\tau) \frac{\sin \omega_l(t - \tau)}{\omega_l} d\tau \right) P_l(\cos \theta) \\ &= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \left(\int_{-\infty}^t h_1(\tau) \frac{d}{dt} \left(\frac{\sin \omega_l(t - \tau)}{\omega_l} \right) d\tau \right) P_l(\cos \theta) \\ &= c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) I_l(\mu) \left(\int_{-\infty}^t h_1(\tau) \cos \omega_l(t - \tau) d\tau \right) P_l(\cos \theta). \end{aligned} \quad (3.60)$$

The Fourier transform of $h_1(t)$ is (see Eq. D.25)

$$\tilde{h}(\omega_l) = e^{-\omega_l^2 \sigma^2 / 2}. \quad (3.61)$$

Note that, after making the approximation $t \gg \sigma$ (replace the upper limit of t with ∞ in Eq. 3.59 or Eq. 3.60), the effect of the convolution is to introduce a factor $\tilde{h}(\omega_l)$ in each term of u_{f_10} (Eq. 3.52) to get u_{f_1} (Eq. 3.45). The same is true for the time-derivative, introducing the factor $\tilde{h}(\omega_l)$ in each term of \dot{u}_{f_10} (Eq. 3.53) to get \dot{u}_{f_1} (Eq. 3.48).

Solutions for a point source ($\mu = 0$)

We now look at the case where $\mu = 0$ (point source). In this limit, the Gaussian initial-shape function $g(\theta)$ becomes a delta function. The integrals $I_l(\mu)$ (Eq. 3.26) simplify (see Appendix D):

$$I_l(0) = \lim_{\mu \rightarrow 0} \int_0^\pi P_l(\cos \theta) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta d\theta = P_l(1) = 1. \quad (3.62)$$

The spatial point-source solutions ($\mu = 0$) are thus (see 3.45–3.47):

$$u_{f_1}^* = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma) \quad (3.63)$$

$$u_{f_2}^* = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2} P_l(\cos \theta) \quad (t \gg \sigma) \quad (3.64)$$

$$u_g^* = \sum_{l=0}^{\infty} (l + \frac{1}{2}) \cos \omega_l t P_l(\cos \theta), \quad (3.65)$$

where the ‘*’ is used to distinguish the spatial point-source form of the solutions.

Lastly, we can present solutions for a point source in both time ($\mu = 0$) and space ($\sigma = 0$):

$$u_{f_{10}}^* = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) \frac{\sin \omega_l t}{\omega_l} P_l(\cos \theta) \quad (t \gg 0) \quad (3.66)$$

$$u_{f_{20}}^* = c^2 \sum_{l=0}^{\infty} (l + \frac{1}{2}) \cos \omega_l t P_l(\cos \theta) \quad (t \gg 0), \quad (3.67)$$

The solutions in Equations (3.65), (3.66), and (3.67), however, contain arbitrarily high frequencies and wavenumbers and will probably not be useful unless a frequency filter is applied.

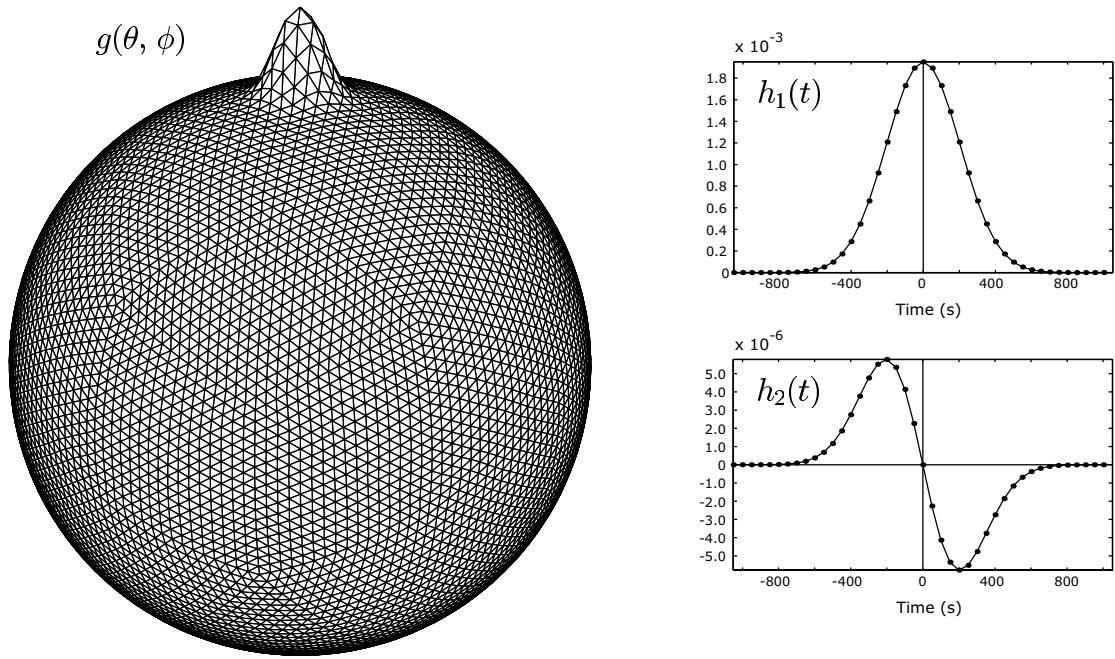


Figure 3.1: Initial-shape function $g(\theta, \phi)$, and two source-time functions, $h_1(t)$ and $h_2(t) = \partial h_1 / \partial t$ (Eqs. 3.18, 3.17, and 3.31). The two forcing terms discussed in this paper are $f_1(\theta, \phi, t) = h_1(t) g(\theta, \phi)$ and $f_2(\theta, \phi, t) = h_2(t) g(\theta, \phi)$. The initial-shape function is plotted on the order $q=4$ triangular grid. For the examples here, the parameters governing the width of the Gaussians are $\mu = 0.0713$ and $\sigma = 204.5$. For the analytical solutions, of course, the gridpoint separation $d\theta$ and the timepoint separation dt are not relevant.

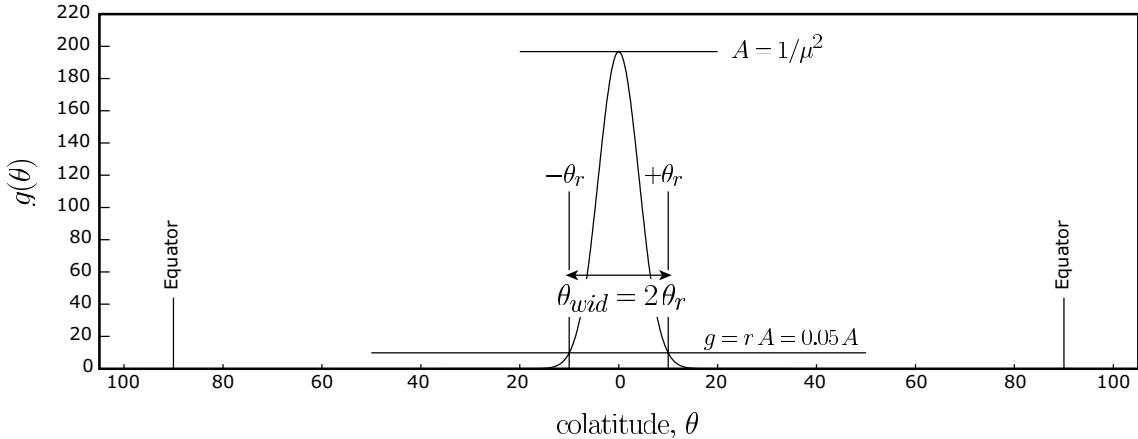


Figure 3.2: The relationship between μ and θ_{wid} for the Gaussian initial-shape function $g(\theta)$. In this example, $g(\theta)$ is plotted with $\theta_{wid} = 20^\circ$ at a height of $0.05A$ ($r = 0.05$), which gives $\mu(\theta_{wid}, r) = 0.0713$ (Eq. 3.21).

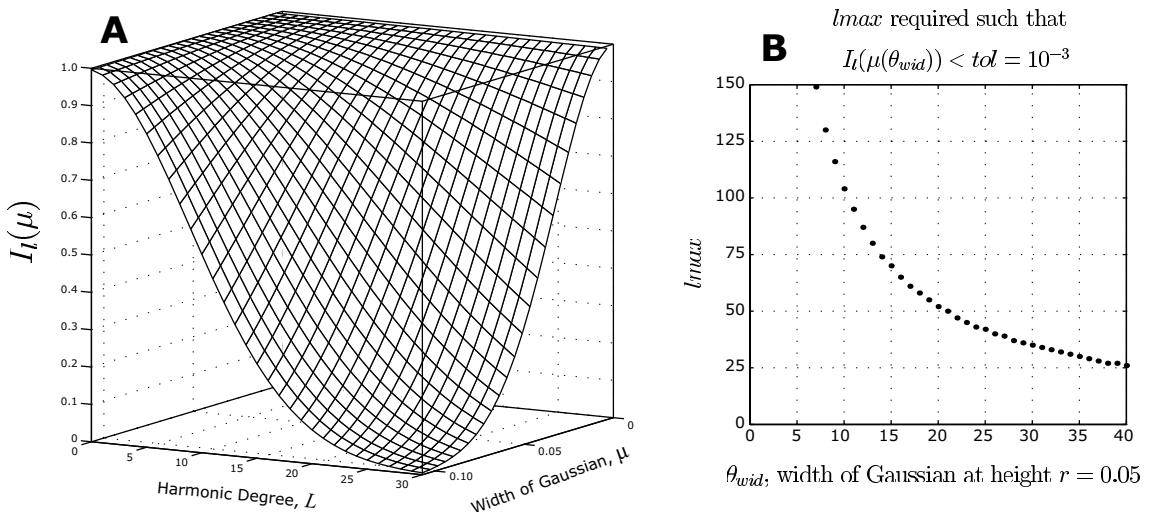


Figure 3.3: Surface plot of $I_l(\mu)$ (Eq. 3.26). **A.** We see graphically that $I_l(\mu) \rightarrow 1$ as $l \rightarrow \infty$ or $\mu \rightarrow 0$. The latter case represents the point source approximation (see Eq. 3.18). **B.** The narrower the Gaussian is in the initial-shape function, the greater the number of l terms required such that $I_{l_{max}}(\mu) < tol$, where $tol = 10^{-3}$ in this plot. μ and θ_{wid} can be thought of interchangeably (Figure 3.2).

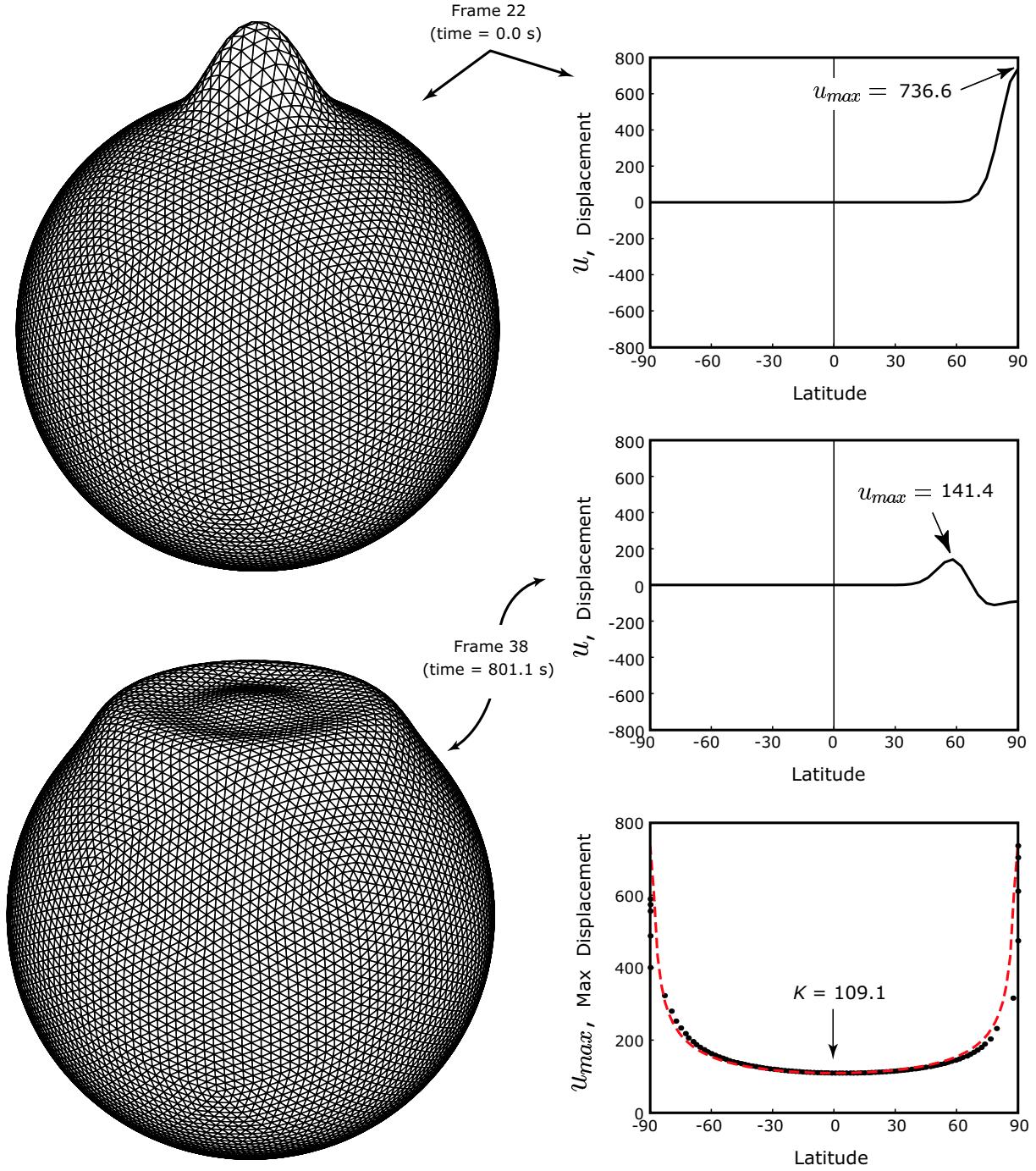


Figure 3.4: Analytical solutions for wave propagation on the sphere, using the forcing term $f_2(\theta, \phi, t)$ of Figure 3.1. The wave is shown both as a 3D figure and as a 2D plot for two different times t . The waveform in the 2D plots is Equation (3.34) for fixed t and varying θ . The bottom plot shows how the maximum values of the wave pulse varies as a function of latitude. The dashed curve is $K/\sqrt{\sin \theta}$, where K is the maximum displacement of the pulse at the equator and θ is the colatitude. As expected, the displacement is maximum at the poles and minimum on the Equator. See also Figure 5.7.

Chapter 4

Spherical hexagonal grids

Contents

4.1	Introduction	34
4.2	Constructing the hexagonal grids	37
4.3	Properties of the hexagonal grids	39
4.4	Calculating the Laplacian on the hexagonal grids	42
4.5	Error in the numerical approximation of the Laplacian	44
4.6	Order of accuracy for the numerical approximation	46
4.7	Variations and optimizations on the hexagonal grids	47
4.8	Summary	50

4.1 Introduction

Every numerical computation is carried out on a discrete number of gridpoints. The choice of the distribution of gridpoints will depend on the numerical method that is being used. The two basic types of numerical methods are global methods and local methods. Global methods refer to spectral methods, which use all the gridpoint values at each timestep iteration; local methods, such as finite difference methods or finite element methods, use only nearby gridpoint values at each iteration. Historically, these methods were developed roughly in the following periods (*Trefethen*, 2000): 1950s, finite difference methods; 1960s, finite element methods; and 1970s, spectral methods.

The use of spherical geodesic grids¹ in approaching wave propagation on a sphere was developed in the field of climate modelling in the late 1960s. The aim of these grids was to cover the sphere with grid cells as uniformly as possible, and in a manner that allowed for the implementation of a finite difference method to solve equations of fluid flow. Developing grids inspired by Buckminster Fuller's geodesic dome, *Williamson* (1968) and *Sadourny et al.* (1968) discretized the sphere into spherical triangles that were nearly uniform in area. Part of their motivation was to develop an alternative grid to the standard latitude-longitude grids, which have an unnecessarily high resolution of grid cells near the poles (Figure 4.1).

Although spectral methods are predominant in meteorology², they are far from ideal. *Stuhne and Peltier* (1996) point out three fundamental problems of the spectral approach to fluid dynamics: (1) it over-resolves polar regions, (2) its results may be less physical in that “it has the potential to manifest anomalous long range interactions”, and (3) it has “poor theoretical efficiency at high spatial resolution” (p. 58). These shortcomings have motivated a resurgence of the local methods over the past ten years, a debate which is reviewed in most of the introductions of the post-1990 studies in Table 4.1 (e.g., *Swarztrauber et al.*, 1997; *Ringler et al.*, 2000; *Randall et al.*, 2002).

Developments in local approaches have included variations on the grids of Williamson and Sadourny, notably those by *Masuda and Ohnishi* (1987) and *Heikes and Randall* (1995a), among others (Table 4.1). In these studies, once the sphere is discretized, the grid is tested using some version of equations describing fluid flow, such as the shallow-water equations. Many authors emphasize the need to develop and test numerical models (*Williamson et al.*, 1992).

Instead of modelling fluid waves in the atmosphere, we shall be modelling Rayleigh wave propagation on the surface of the earth. Our spherical geodesic grid and finite difference calculation is based on those used in the climate studies above (*Heikes and Randall*, 1995a). Various names for the grids have been used (Table 4.1); our nomenclature will be “triangular grid” for a spherical grid of triangles, and “hexagonal grid” for a spherical grid of predominantly hexagons.

¹This nomenclature refers to the fact that the grids are characterized by arcs between gridpoints on the unit sphere, i.e., geodesics (*Williamson*, 1968).

²The method which is cited as most successful at modelling fluid flow is in fact the *pseudo*-spectral method (*Stuhne and Peltier*, 1996).

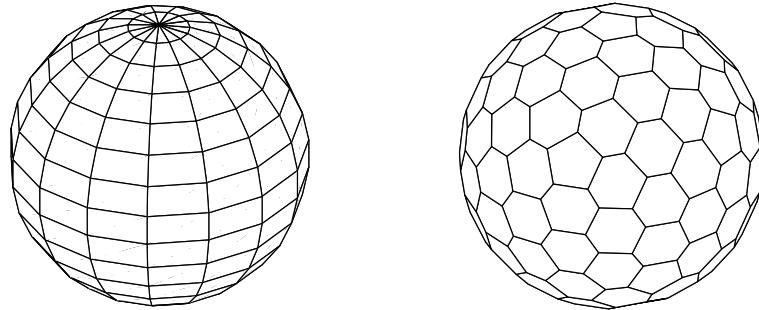


Figure 4.1: Two discretizations of the sphere. One problem of the latitude–longitude discretization (*left*) is that the areas of the grid cells go to zero as the cells approach the poles. Alternative discretizations (*right*) try to keep the area of the cells as uniform as possible.

Table 4.1: Spherical grids used in other studies, each which used an icosahedron as the initial triangular grid to subdivide. The third column indicates the type of numerical method used: FDM = finite difference method, FEM = finite element method. The fourth column indicates whether the study employed a discrete Laplacian operator on the grid, which is of interest to our study.

Study	Name of Grid	Method	∇^2
Williamson (1968, 1970)	spherical geodesic grid	FDM	yes
Williamson (1969)	triangular grid	FDM	yes
Sadourny <i>et al.</i> (1968)	icosahedral–hexagonal grid	FDM	yes
Sadourny and Morel (1969)	hexagonal grid	FDM	yes
Cullen (1974)	icosahedral grid	FEM	yes
Baumgardner and Frederickson (1985)	icosahedral grid	FEM	yes
Sword <i>et al.</i> (1986a,b)	icosahedral grid	FEM	yes
Masuda and Ohnishi (1987)	icosahedral–hexagonal grid	FDM	yes
Wang and Dahlen (1995a)	spherical tessellation grid	N/A	no
Wang <i>et al.</i> (1998)			
Heikes and Randall (1995a,b)	twisted icosahedral grid	FDM	yes
Stuhne and Peltier (1996, 1999)	icosahedral grid	FEM	yes
Swarztrauber <i>et al.</i> (1997)	icosahedral geodesic grid	FDM	yes
Thuburn (1997)	hexagonal–icosahedral grid	FDM	yes
Ringler <i>et al.</i> (2000)	spherical geodesic grid	FDM	yes
Randall <i>et al.</i> (2002)		FDM	yes
Ringler and Randall (2002a,b)			

Table 4.2: The five platonic solids, plus the “buckeyball” for comparison. The internal angle for each solid is given by α , with the number of faces meeting at each vertex in parenthesis. The sum of the interior angles at each vertex is given by $\sum \alpha$.

platonic solid	polygonal face	faces, f	vertices, v	edges, e	α	$\sum \alpha$
tetrahedron	triangle	4	4	6	60° (3)	180°
cube (hexahedron)	square	6	8	12	90° (3)	270°
octahedron	triangle	8	6	12	60° (4)	240°
dodecahedron	pentagon	12	20	30	108° (3)	324°
icosahedron	triangle	20	12	30	60° (5)	300°
buckeyball ($q=0$ hexagonal grid)	pentagon (12) hexagon (20)	32	60	90	108° (1) 120° (2)	348°

4.2 Constructing the hexagonal grids

Each point on a grid (triangular or hexagonal) is always found on the unit sphere, which represents the surface of the earth. The hexagonal grids are generated from the triangular grids. Below we explain how to generate triangular grids and hexagonal grids.

Triangular grids

The algorithm for generating successively finer triangular grids is simple: take each existing triangular patch, find the midpoints of the three sides, and then construct four new triangular patches. This process, shown in planar view in Figure 4.4, is carried out for triangular gridpoints on the unit sphere.

What type of spherical polyhedron do we choose for our lowest order triangular grid? It seems reasonable to construct our triangular grid from a polyhedron whose faces are uniform in area and whose vertices lie on the unit sphere: this limits us to five polyhedra, the platonic solids (Table 4.2). Properties of the platonic solids are: (i) the same number of faces meet at each vertex, (ii) each face is a regular congruent polygon, (iii) the sum of the interior angles at each vertex is $< 360^\circ$, and (iv) each solid satisfies Euler’s Polyhedron Theorem,

$$f + v - e = 2 , \quad (4.1)$$

where f is the number of faces, v is the number of vertices, and e is the number of edges

(see Table 4.2). To construct our order $q=0$ triangular grid, we begin with a dodecahedron, subdividing each face into five $72^\circ/54^\circ/54^\circ$ triangles and then subdividing each triangle into four new triangles (Figure 4.5). Our order $q=0$ grid contains the vertices of both the dodecahedron and icosahedron³ (Figure 4.6) and is not a platonic solid. It is clear from the criteria for the platonic solids that the maximum number of equal-area triangles that can form a spherical triangular grid is 20 (Table 4.2); our goal in constructing a finer mesh of points on the sphere is to keep the area of the triangles as uniform and small as possible.

The mechanism for constructing higher order triangular grids (Figure 4.4) shows that, beginning with our subdivided dodecahedron, the number of faces, f_t , of the q th-order triangular grid is

$$f_t(q) = 60 \cdot 4^q , \quad (4.2)$$

which can be seen in Table 4.3.

Hexagonal grids

The hexagonal grids are constructed from the triangular grids, a process shown schematically in Figure 4.7. The corner of each hexagonal face is found by calculating the central point of the three vertices forming the face of the corresponding triangular grid. The central point of the vectors \mathbf{T}_1 , \mathbf{T}_2 , and \mathbf{T}_3 is given by

$$\mathbf{V} = \frac{(\mathbf{T}_3 - \mathbf{T}_1) \times (\mathbf{T}_2 - \mathbf{T}_1)}{|(\mathbf{T}_3 - \mathbf{T}_1) \times (\mathbf{T}_2 - \mathbf{T}_1)|} = \frac{\mathbf{T}_1 \times \mathbf{T}_3 + \mathbf{T}_3 \times \mathbf{T}_2 + \mathbf{T}_2 \times \mathbf{T}_1}{|\mathbf{T}_1 \times \mathbf{T}_3 + \mathbf{T}_3 \times \mathbf{T}_2 + \mathbf{T}_2 \times \mathbf{T}_1|} , \quad (4.3)$$

where \mathbf{T}_1 , \mathbf{T}_2 , \mathbf{T}_3 , and \mathbf{V} are all on the unit sphere (*Augenbaum and Peskin, 1985*). Note that \mathbf{T}_1 , \mathbf{T}_2 , and \mathbf{T}_3 are ordered in a clockwise sense, looking from the exterior, so that \mathbf{V}_2 lies in the correct hemisphere (see Figure 4.7B).

The order $q=0$ hexagonal grid is in fact the well-known buckyball (C_{60} molecular structure), with 12 regular pentagonal faces and 20 regular hexagonal faces (having two types of faces, it is *not* a platonic solid — Table 4.1). The buckyball has the property that the dis-

³It is quite simple to construct and visualize the icosahedron. It can be constructed from three mutually perpendicular rectangles whose sides have the golden ratio $t = (1 + \sqrt{5})/2 \approx 1.62$. The vertices are the corners of the rectangles and are thus given by the points $(0, \pm t, \pm 1)$, $(\pm t, 0, \pm 1)$, $(\pm t, \pm 1, 0)$. (Note that these points are on the unit sphere.)

tance between one point and its neighbors is the same for all points. Higher-order hexagonal grids are shown in Figure 4.8A; the relationship between successive orders of hexagonal grids is shown in Figure 4.8B. The number of faces f_h of the q th-order hexagonal grid is

$$f_h(q) = 30 \cdot 4^q + 2 , \quad (4.4)$$

which can be seen in Table 4.3.

Here we note two important aspects of the grids. First, note that *the vertices of the triangular grids are the face centers on the hexagonal grids* (Figure 4.7). Similarly, the vertices of the hexagonal grids are the face centers of the triangular grids. This gives us the equalities

$$f_h = v_t \quad (4.5)$$

$$v_h = f_t . \quad (4.6)$$

Function values are calculated at the center of each hexagonal face, or, in other words, at each triangular vertex. Second, note that the triangular vertices accumulate as the grids become finer; in other words, the vertices of the coarser triangular grids represent a subset of the vertices of the finer grids (Figure 4.5). This is not the case with the hexagonal grids, whereby the gridpoints of the coarser grids do not coincide exactly with gridpoints of the finer grids (Figure 4.8).

So far we have qualitatively presented how to construct the spherical grids, but we have not presented an explicit algorithm for indexing the vertices, which is crucial. Our code was originally written by Woodhouse in 1995 and modified over the course of this study to accommodate the numerical method (Section 4.4).

4.3 Properties of the hexagonal grids

Table 4.4 shows some properties of the hexagonal grids. The values are for a sphere with earth's radius:

$$\text{earth radius } a = 6371 \text{ km}$$

$$\text{circumference } 2\pi a = 40,030 \text{ km}$$

$$\text{surface area } 4\pi a^2 = 5.1 \times 10^8 \text{ km}^2 .$$

Table 4.3: Geometric properties of the triangular grids (Figure 4.5) and the hexagonal grids (Figure 4.8A). The order $q=0$ triangular grid is a subdivided dodecahedron. Each hexagonal grid and triangular grid satisfies Equation (4.1).

order, q	triangular grid			hexagonal grid		
	faces, f_t	vertices, v_t	edges, e_t	faces, f_h	vertices, v_h	edges, e_h
0	60	32	90	32	60	90
1	240	122	360	122	240	360
2	960	482	1 440	482	960	1 440
3	3 840	1 922	5 760	1 922	3 840	5 760
4	15 360	7 682	23 040	7 682	15 360	23 040
5	61 440	30 722	92 160	30 722	61 440	92 160
6	245 760	122 882	368 640	122 882	245 760	368 640
q	$f_t = 60 \cdot 4^q$	$v_t = 30 \cdot 4^q + 2$	$e_t = 90 \cdot 4^q$	$f_h = v_t$	$v_h = f_t$	$e_h = e_t$

Table 4.4: Properties of the hexagonal grids used in this study. f_{eq} is the approximate number of faces on the equator. $d\theta$ is the average angular distance between face centers. Areas and lengths assume a spherical earth with radius $a = 6371$ km. Compare with Table 4.5.

order, q	faces, f_h	f_{eq}	area of faces (km^2)		distance between face centers (km)	
			a_{ave}	a_{min}/a_{max}	d_{ave} ($d\theta$)	d_{min}/d_{max}
0	32	10	15 939 515	0.9413	4320 (38.86°)	0.8940
1	122	19	4 180 856	0.9142	2208 (19.86°)	0.8608
2	482	39	1 058 225	0.9070	1111 (9.99°)	0.8520
3	1 922	78	265 382	0.8776	557 (5.00°)	0.8498
4	7 682	155	66 397	0.8700	278 (2.50°)	0.8492
5	61 440	310	16 603	0.8681	139 (1.25°)	0.8491
6	245 760	620	4151	0.8676	69 (0.63°)	0.8490

Table 4.5: Properties of the hexagonal grids (“twisted icosahedron polyhedron”) of *Heikes and Randall* (1995a). f_{eq} is the exact number of faces on the equator. $d\theta$ is the average angular distance between face centers. Areas and lengths assume a spherical earth with radius $a = 6371$ km. Compare with Table 4.4.

order, q	faces, f_h	f_{eq}	area of faces (km^2)		distance between face centers (km)	
			a_{ave}	a_{min}/a_{max}	d_{ave} ($d\theta$)	d_{min}/d_{max}
0	42	10	12 140 000	0.885	3755.5 (33.8°)	0.881
1	162	20	3 149 000	0.774	1916.4 (17.2°)	0.848
2	642	40	794 600	0.763	962.4 (8.7°)	0.840
3	2 562	80	199 100	0.742	481.7 (4.3°)	0.838
4	10 242	160	49 800	0.736	240.9 (2.2°)	0.837
5	40 962	320	12 450	0.733	120.5 (1.1°)	0.837
q	$40 \cdot 4^q + 2$	$5 \cdot 2^{q+1}$				

With the hexagonal grids, each corner of a hexagon (or pentagon) is a junction of three edges. And since each edge has exactly two (hexagonal) vertices, we have $2e = 3v$ (*Augenbaum and Peskin*, 1985). Using Equation (4.1), which is true for an arbitrary arrangement of points on the surface of a sphere, we see that $v = 2f - 4$ and $e = 3f - 6$ (*Augenbaum and Peskin*, 1985). These relationships are evident in Table 4.3 for the hexagonal grids.

Area calculations for each hexagonal face are done by subdividing each face into six (or five) triangles, and then summing the areas of the spherical triangles. The area of a spherical triangle with sides having arc lengths a , b , and c is

$$a_i = 4 \tan^{-1} \left(\sqrt{\tan\left(\frac{s}{2}\right) \tan\left(\frac{s-a}{2}\right) \tan\left(\frac{s-b}{2}\right) \tan\left(\frac{s-c}{2}\right)} \right), \quad (4.7)$$

where $s = (a + b + c)/2$. Thus, subdividing each face into spherical triangles, the area of hexagonal face with center \mathbf{T}_i is

$$A_i = \sum_{i=1}^N a_i, \quad (4.8)$$

where $N = 6$ or $N = 5$.

Figures 4.9 and 4.10 show the distribution of three geometric properties of the hexagonal grids: area of faces, distance to each neighbor, and length of hexagonal face edge. The

variation in these values is larger than one might have expected, based on our subdivision algorithm. We see that the distribution of areas is *not* particularly uniform — the deviation from the mean is about 10%.

Comparison of hexagonal grids with those used in other studies

Several climate modelling studies use a subdivided *icosahedron* as the order $q=0$ triangular grid (Table 4.1). A comparison of our order $q=0$ triangular grid — the subdivided dodecahedron — with the subdivided icosahedron shows the following differences: the subdivided dodecahedron has 60 triangular faces, each triangle having approximately $67.6^\circ/56.2^\circ/56.2^\circ$ internal angles (Figure 4.5); the subdivided icosahedron has 80 triangular faces, 60 having approximately $68.8^\circ/55.6^\circ/55.6^\circ$ internal angles and 20 being equilateral triangles.

For the subdivided icosahedron, the number of faces on the triangular grid is $f_t(q) = 80 \cdot 4^q$ (compare Eq. 4.2), and the number of faces on the hexagonal grid is $f_h(q) = 40 \cdot 4^q + 2$ (compare Eq. 4.4). Table 4.5 shows some properties of the initial-icosahedron hexagonal grids, which can be compared with the properties of the initial-dodecahedron hexagonal grids (ours) in Table 4.4. Based on the values in the two tables, our hexagonal grids appear to be better at having the greatest uniformity of face area ($a_{min}/a_{max} \approx 1$) and distance between faces ($d_{min}/d_{max} \approx 1$). Optimizations of the initial-icosahedron grids are discussed in Section 4.7.

4.4 Calculating the Laplacian on the hexagonal grids

The hexagonal grids are discretized in a way that makes it possible to calculate the Laplacian $\nabla^2 u$ of a given function $u(\theta, \phi)$ on the surface of the sphere by using a numerical approximation. We employed the approximation used in the climate studies to calculate the Laplacian on the sphere⁴ (*Masuda and Ohnishi*, 1987; *Heikes and Randall*, 1995a).

⁴Generally the Laplacian $\nabla^2 \alpha$ is one of several operators approximated on the grids, such as the Jacobian $J(\alpha, \beta)$ and the flux-divergence operator $\nabla \cdot (\alpha \nabla \beta)$, where α and β are arbitrary scalar fields (e.g., *Masuda and Ohnishi*, 1987).

From Gauss's theorem, we have

$$\int_A \nabla^2 u \, dA = \oint_C \frac{\partial u}{\partial n} \, ds , \quad (4.9)$$

where u is a scalar function and $\partial u / \partial n$ is the gradient of u , normal to a small segment ds of a curve C (Figure 4.2). Now, consider a particular face center $\mathbf{T}_0(\theta, \phi)$ on the hexagonal grid. The function $u(\mathbf{T}_0) = u_0$ is defined at the center of the face, which has a spherical area A_0 defined by Equation (4.8).

If the faces on the grid are small, then $\nabla^2 u$ can be considered constant across each face, and we can approximate the left-hand side of Equation (4.9) by

$$\int_A \nabla^2 u \, dA \approx \nabla^2 u|_{\mathbf{T}_0} \int_A dA = \nabla^2 u|_{\mathbf{T}_0} A_0 . \quad (4.10)$$

The line integral on the right-hand side of Equation (4.9) can be approximated by summing the u -gradients between \mathbf{T}_0 and its $N = 6$ (or $N = 5$) neighbors:

$$\oint_C \frac{\partial u}{\partial n} \, ds \approx \sum_{i=1}^N \frac{l_i}{L_i} (u_i - u_0) , \quad (4.11)$$

where L_i is the arc distance to the adjacent face-center, and l_i is the arc length of the corresponding hexagonal face edge (Figure 4.11). Using these results, we find

$$\mathbf{D2}(u)|_{\mathbf{T}_0} \equiv \nabla^2 u|_{\mathbf{T}_0} \approx \frac{1}{A_0} \sum_{i=1}^N \frac{l_i}{L_i} (u_i - u_0) \quad (4.12)$$

$$= \sum_{i=1}^N \left(\frac{l_i}{L_i A_0} \right) u_i - \left(\sum_{i=1}^N \frac{l_i}{L_i A_0} \right) u_0 , \quad (4.13)$$

where $\mathbf{D2}$ is our notation for our discrete Laplacian. Equation (4.13) emphasizes the dependence of the Laplacian on the function values u_i at neighboring points. Figure 4.11 shows the variables used in calculating the Laplacian at each hexagonal face.

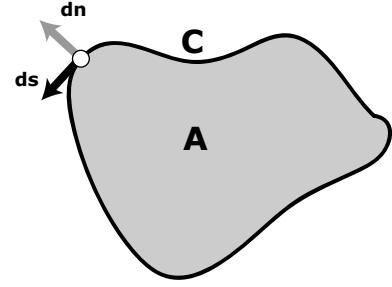


Figure 4.2: Sketch of integration parameters for Eq. 4.9. The scalar function u is defined everywhere in space.

4.5 Error in the numerical approximation of the Laplacian

The error in the numerical Laplacian approximation is determined by comparing the numerical value of $\nabla^2 u(\theta, \phi)$ with the “actual value”, which can be the (exact) analytical value, if we choose an appropriate test function $u(\theta, \phi)$. Spherical harmonic functions $Y_{lm}(\theta, \phi)$ are a convenient choice for $u(\theta, \phi)$, since they are well-defined on the sphere (Appendix A) and are eigenfunctions of the Laplacian operator:

$$\nabla^2 Y_{lm} = -l(l+1)Y_{lm}. \quad (4.14)$$

Thus, with $u \equiv Y_{lm}$, from Equation (4.12) we obtain $[\nabla^2 u]_{\text{nu}}$ and from Equation (4.14) we obtain $[\nabla^2 u]_{\text{an}}$. Both of these functions are defined at each gridpoint of the order q grid.

We define the error in the numerical calculation of the Laplacian at $\mathbf{T}_i(\theta, \phi)$ as

$$\text{D2error}(\theta, \phi) = \left| \frac{[\nabla^2 u]_{\text{nu}}|_{\mathbf{T}_i} - [\nabla^2 u]_{\text{an}}|_{\mathbf{T}_i}}{\max\{[\nabla^2 u]_{\text{an}}\}} \right|, \quad (4.15)$$

where `max` finds the maximum single value of $\nabla^2 u$ on the sphere. Thus, for each point on the sphere, the absolute difference between the model approximation and the actual value is normalized to the maximum actual value on the sphere⁵.

Figure 4.12A shows the distribution of the error over the $q=6$ grid, and we see that the error is not at all uniformly distributed. Although the average error at each point is small ($\text{D2error}(\theta, \phi) \approx 10^{-4}$), there is a small fraction of points having errors 10–20 times greater than the average.

Figure 4.12B-C takes a closer look to show how isolated the “bad error” gridpoints are. The error at gridpoint 112 ($\text{D2error}(112) = 6.807 \times 10^{-3}$) is on average 24 times greater than the error at one of its neighboring points. Furthermore, the error at 112 does not decrease with finer grids, which means that the numerical approximation for the Laplacian operator is zero-order accurate at this point. What makes these particular gridpoints so bad? The answer seems to lie in the geometric design of the grids. We see in Figure 4.9 that gridpoint 112 is a particularly “anisotropic” point, where the areas are relatively large in one direction and relatively small in the perpendicular direction. Our discrete approximation of the Laplacian

⁵This normalization gives the smallest value for the error. If we normalized to the analytical value at each point, however, we would have points with near-infinite error, where $Y_{lm}(\theta, \phi) = 0$.

(Eq. 4.12) is not effective at these gridpoints.

Spherically-averaged error

We are also interested in calculating the spherical average of the error. One way of calculating the spherically-averaged error is take the mean of all the error values, since the gridpoints are nearly uniformly distributed:

$$\overline{\text{D2error}}(l, m, q) = \text{mean} \{ \text{D2error}(\theta, \phi) \} . \quad (4.16)$$

We also adopt the norm conventions of *Heikes and Randall* (1995b):

$$\|f_i\| = \frac{1}{A_e} \sum_{\text{cells}} A_i |f_i| \quad (4.17)$$

$$\|f_i\|_2 = \left[\frac{1}{A_e} \sum_{\text{cells}} A_i (f_i)^2 \right]^{1/2} \quad (4.18)$$

$$\|f_i\|_\infty = \max \{|f_i|\} , \quad (4.19)$$

where f_i is a function defined at all the gridpoints (i.e., cells) on the sphere and $A_e = 4\pi a^2$ is the area of the sphere⁶. We use these norms to establish three more measurements of the spherically-averaged error⁷:

$$\overline{\text{D2one}}(l, m, q) = \frac{\| [\nabla^2 u]_{\text{nu}} - [\nabla^2 u]_{\text{an}} \|}{\| [\nabla^2 u]_{\text{an}} \|} \quad (4.20)$$

$$\overline{\text{D2two}}(l, m, q) = \frac{\| [\nabla^2 u]_{\text{nu}} - [\nabla^2 u]_{\text{an}} \|_2}{\| [\nabla^2 u]_{\text{an}} \|_2} \quad (4.21)$$

$$\overline{\text{D2inf}}(l, m, q) = \frac{\| [\nabla^2 u]_{\text{nu}} - [\nabla^2 u]_{\text{an}} \|_\infty}{\| [\nabla^2 u]_{\text{an}} \|_\infty} . \quad (4.22)$$

Thus, given a test function $u \equiv Y_{lm}$ on a given grid q , we have four equations for the spherically-averaged error in terms of l , m , and q . Figure 4.13 shows how three of the error values vary with grid order q for a particular Y_{lm} function. We immediately see that the error converges for the one-norm and two-norm, but not for the infinity norm, indicating that there must be at least one “bad” gridpoint (note Eq. 4.19).

⁶To be precise, A_e is the sum of the areas of all the faces, which is equal to $4\pi a^2$ to within 10^{-5} .

⁷These are essentially equivalent to the three standard normalized global errors put forth by *Williamson et al.* (1992).

In general, we see that the error in the numerical Laplacian increases as degree l increases, which is what one would expect, since the higher the degree, the larger the surface gradients. Also, we observe relative lows in error values for test functions with order $m = 5N$ ($N = 0, 1, \dots, 8$); this can be attributed to the symmetry of the test function matching the symmetry of the grids (Figure 4.14). (The vertical color streaks for these m -values in Figure 4.15 indicate this feature as well.) Our findings suggest that the orientation of the grids has an effect on the errors, something that was first noted by *Williamson* (1968).

The error in the numerical Laplacian was also compared between (1) the grid calculations using arc distances and spherical areas and (2) the grid calculations using chord distances and planar areas (see Figure 4.11). There is a negligible difference in the results for the finer grids ($q \geq 4$). We used arc distances and spherical areas, since there is no increase in the calculation time, as the geometric properties of the grids are calculated once prior to iterating the wave equation.

4.6 Order of accuracy for the numerical approximation

As *Heikes and Randall* (1995b) point out, it is difficult to analytically investigate the order of accuracy of a discrete operator on an irregular grid. Suppose our finite difference form of the Laplacian operator is m th order for the order- q spherical grid having average grid spacing h :

$$[\nabla^2 u]_{\text{an}} - [\nabla^2 u]_{\text{nu}} = O[h^m] \quad (4.23)$$

Suppose we have grids having two different spacings, h_1 and h_2 , where $h_1 = 2h_2$ (i.e., grid 2 is finer than grid 1). Then we have:

$$\frac{\|[\nabla^2 u]_{\text{nu}}(h_1) - [\nabla^2 u]_{\text{an}}\|}{\|[\nabla^2 u]_{\text{nu}}(h_2) - [\nabla^2 u]_{\text{an}}\|} = \frac{O[(h_1)^m]}{O[(h_2)^m]} = \frac{O[2^m(h_2)^m]}{O[(h_2)^m]} = 2^m, \quad (4.24)$$

where $\| \cdot \|$ is some norm (*Heikes and Randall*, 1995b), and m is the order of accuracy of the Laplacian operator. $[\nabla^2 u]_{\text{nu}}(h_1)$ is a vector of values of the numerical Laplacian evaluated at each gridpoint on the grid having average gridpoint spacing h_1 .

Table 4.6: Studies of wave propagation on the global scale that use subdivided cubic grids. FDM = finite difference method; SEM = (pseudo-)spectral element method.

Study	Name of Grid	Numerical Method
<i>Sadourny</i> (1972)	quasi-uniform spherical grid	FDM
<i>Rancic et al.</i> (1996)	expanded spherical cube	FDM
<i>Ronchi et al.</i> (1996)	cubed sphere	FDM
<i>Komatitsch and Tromp</i> (1999, 2002a,b)	cubed sphere	SEM
<i>Capdeville et al.</i> (2003)	cubed sphere	SEM

Geometrically, the order of accuracy is the slope of the error-vs-gridpoint-spacing curve (Figure 4.13B). We can plot the order of accuracy between successive grids q and $q+1$, and plot these values for each of the three norms (Figure 4.15). We see that our discrete approximation of the Laplacian operator for the finest grids ($q \geq 4$) is nearly second-order accurate for the one-norm and two-norms, but zero- or first-order accurate for the infinity norm (Figure 4.15). This can be seen in other figures as well. The leveling of the infinity-norm curve in Figure 4.13B indicates zero-order accuracy for at least one gridpoint, since the infinity norm measures the greatest absolute deviation of the numerical value from the analytical value (Eq. 4.22). In Figure 4.12 we can actually identify where these particularly bad points are — they happen to be at the gridpoints of the lower-order q grids.

In short, our numerical approximation for the Laplacian is nearly second-order, except at a few bad points on the grids. This is consistent with the findings of *Heikes and Randall* (1995b) (their Figure 6), who used the same numerical approximation but used slightly different grids and different test functions.

4.7 Variations and optimizations on the hexagonal grids

It is apparent that discretizing the sphere as uniformly as possible is a challenging problem (e.g., *Saff and Kuijlaars*, 1997; *Cui and Freeden*, 1997; *Augenbaum and Peskin*, 1985). Furthermore, the task of “communicating” among the points, as required by numerical methods, is an additional barrier. A wealth of studies in global climate modelling have used grids and numerical methods similar to ours (Table 4.1). However, these grids have their weaknesses, as we have shown that there are specific points in these grids that do not converge, supporting

the findings of *Heikes and Randall* (1995b) and *Sword et al.* (1986a).

Heikes and Randall (1995b) developed an improved grid by optimizing the distribution of gridpoints such that the value $(r_i)^4$ was minimized for a particular order q grid (Figure 4.3). This is a rather brute-force approach that requires a global minimization to a problem whose numerical method is local. If the finite difference approximation is bad at certain points, why not improve the method? Nevertheless, *Heikes and Randall* (1995b) show that with the optimized grid, the errors converge to zero at *all* gridpoints on the sphere.

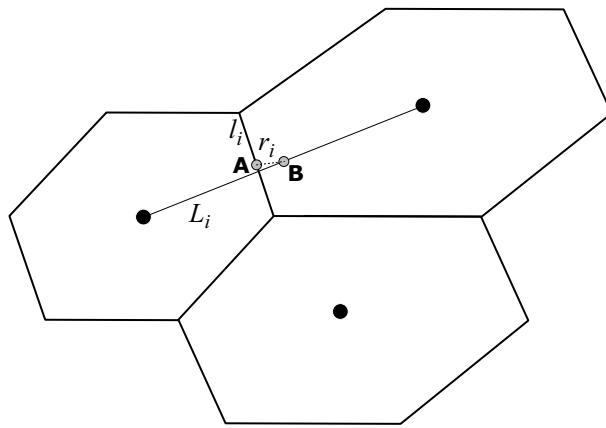


Figure 4.3: Schematic diagram illustrating the parameter r_i , defined as the distance between the midpoint of a face edge (A) and the midpoint of the corresponding distance-between-face-centers (B). This is the optimization parameter used by *Heikes and Randall* (1995b) for optimizing the gridpoint spacing.

There are negative consequences associated with optimizing a particular parameter of the grid. *Heikes and Randall* (1995b) perturbed the gridpoints to obtain a more uniform area coverage — or, equivalently, a more uniform density of gridpoints — but this was at the expense of increasing the range of distance-between-cells on the grid (Table 4.7). *Sword et al.* (1986a) found similar results when optimizing the area coverage⁸, reporting that “the regularization increased the measure of local anisotropy in some regions, while smoothing out sharp contrasts” (p. 72). They defined local anisotropy as the difference between the maximum and minimum distance-to-neighboring-cell for each triangular vertex.

⁸*Sword et al.* (1986a) do not explain specifically how they optimized the areas: “A least-squares problem was set up, with the goal of making the areas of the individual triangles as equal to each other as possible” (p. 67).

Table 4.7: Optimizing the hexagonal grids: results from Heikes & Randall. Ratios from the unoptimized grids are from *Heikes and Randall* (1995a) and shown earlier in Table 4.5. Ratios from the optimized grids are from *Heikes and Randall* (1995b) and *Randall et al.* (2002). After optimization in 1995, the area ratios were better (i.e., closer to 1) but the distance ratios were worse.

grid order q	hex faces f_h	Ratio of smallest cell to largest cell			Ratio of smallest distance to largest distance between cell centers		
		unopt- imized	optimized 1995	optimized 2002	unopt- imized	optimized 1995	optimized 2002
0	42	0.885	0.885	0.885	0.881	0.881	0.881
1	162	0.774	0.868	0.916	0.848	0.811	0.820
2	642	0.763	0.880	0.942	0.840	0.787	0.881
3	2562	0.742	0.877	0.948	0.838	0.778	0.799
4	10492	0.736	0.870	0.951	0.837	0.775	0.790
5	40962	0.733	0.867	0.952	0.837	0.776	0.788

Other studies have tried different discretizations of the sphere, such as a subdivided cube (Table 4.6), which has seen developments in the field of seismology (*Komatitsch and Tromp*, 2002b). *Cui and Freeden* (1997) is a comprehensive study of the equidistribution of points on the sphere, and they discuss the convergence of the subdivided tetrahedron, octahedron, and icosahedron, in comparison with other distributions of gridpoints. They state: “To our surprise the results are not very convincing for larger pointsets” (p. 606). Nevertheless, it seems that the distribution of gridpoints in these subdivided platonic solids makes it easier to implement the numerical approximations, which is, after all, is the purpose for using these grids.

4.8 Summary

The important points from this chapter are included here.

1. Numerical methods on the sphere are required to analyze global wave propagation, particularly in the fields of climate modelling, oceanography, and seismology. Present approaches are predominantly spectral (“global methods”), but finite difference and finite element methods (“local methods”) pose as appealing alternatives and have seen significant developments (Table 4.1).
2. Novel distributions of gridpoint systems on the sphere — such as the subdivided icosahedron (Table 4.1) or dodecahedron (ours) — have been developed in order to implement local methods of numerical approximation.
3. Triangular grids are constructed by subdividing each triangular face into four faces and then projecting the points to the unit sphere (Figure 4.5). Hexagonal grids are constructed from triangular grids (Figure 4.7).
4. The hexagonal and triangular grids should be considered interchangeable. The vertices of the triangular grids are the centers of the faces on the hexagonal grids. The function values are calculated at each hexagonal face center, or, equivalently, at each triangular vertex (Figure 4.7).
5. Our numerical approximation for the Laplacian operator (∇^2) — the discrete Laplacian — is a finite difference method and depends on the function values at the six (or five) neighboring points, as well as on the geometric properties, such as the area of the cell and distance-to-neighbors (Figure 4.11).
6. The discrete Laplacian is nearly second-order accurate at almost all of the gridpoints. However, at a few gridpoints, the discrete approximation is zero-order (Figure 4.12). These bad points may act as scatterers when we propagate waves on these grids.
7. It is possible to improve the discrete Laplacian by optimizing the gridpoint distribution, but we did not attempt this.

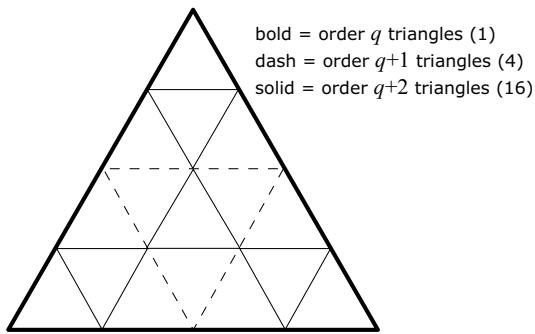


Figure 4.4: Relationship between successive orders of triangular grids. Each triangle is divided into four smaller triangles. This process is applied to spherical triangular patches on the unit sphere (see Figure 4.5).

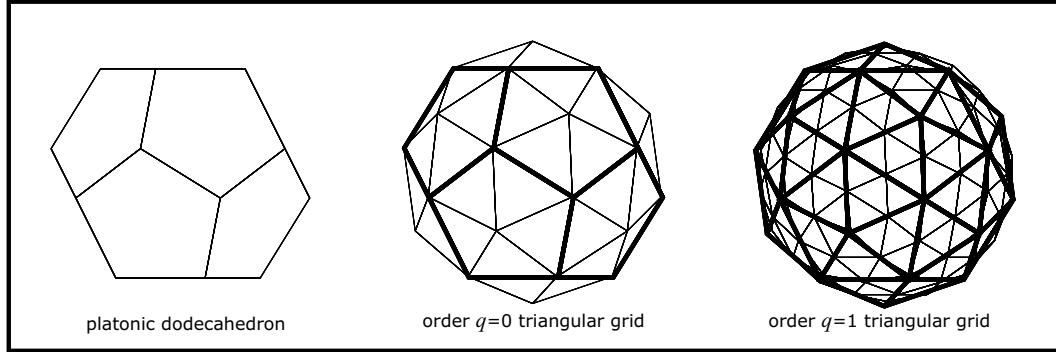


Figure 4.5: Construction of the $q=0$ triangular grid and higher order triangular grids. Beginning with the order $q=0$ grid, the midpoints between vertices are calculated and then projected to the surface of the unit sphere (see Figure 4.4).

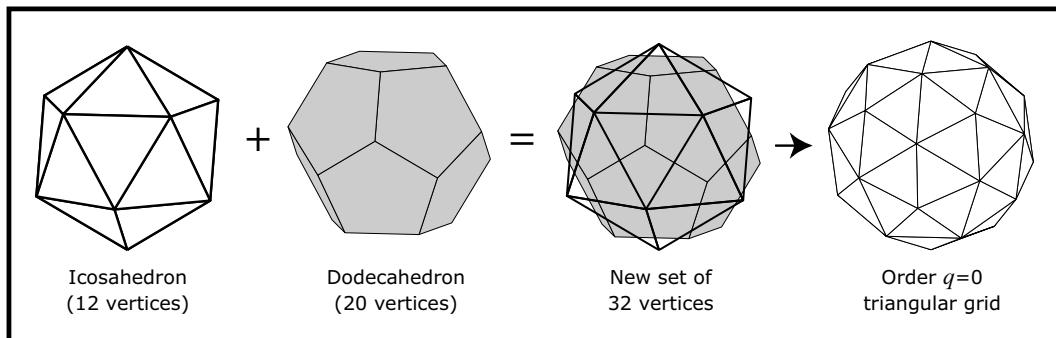


Figure 4.6: Geometric interpretation of the order $q=0$ triangular grid. The order $q=0$ triangular grid is the superposition of two platonic solids (see Table 4.2).

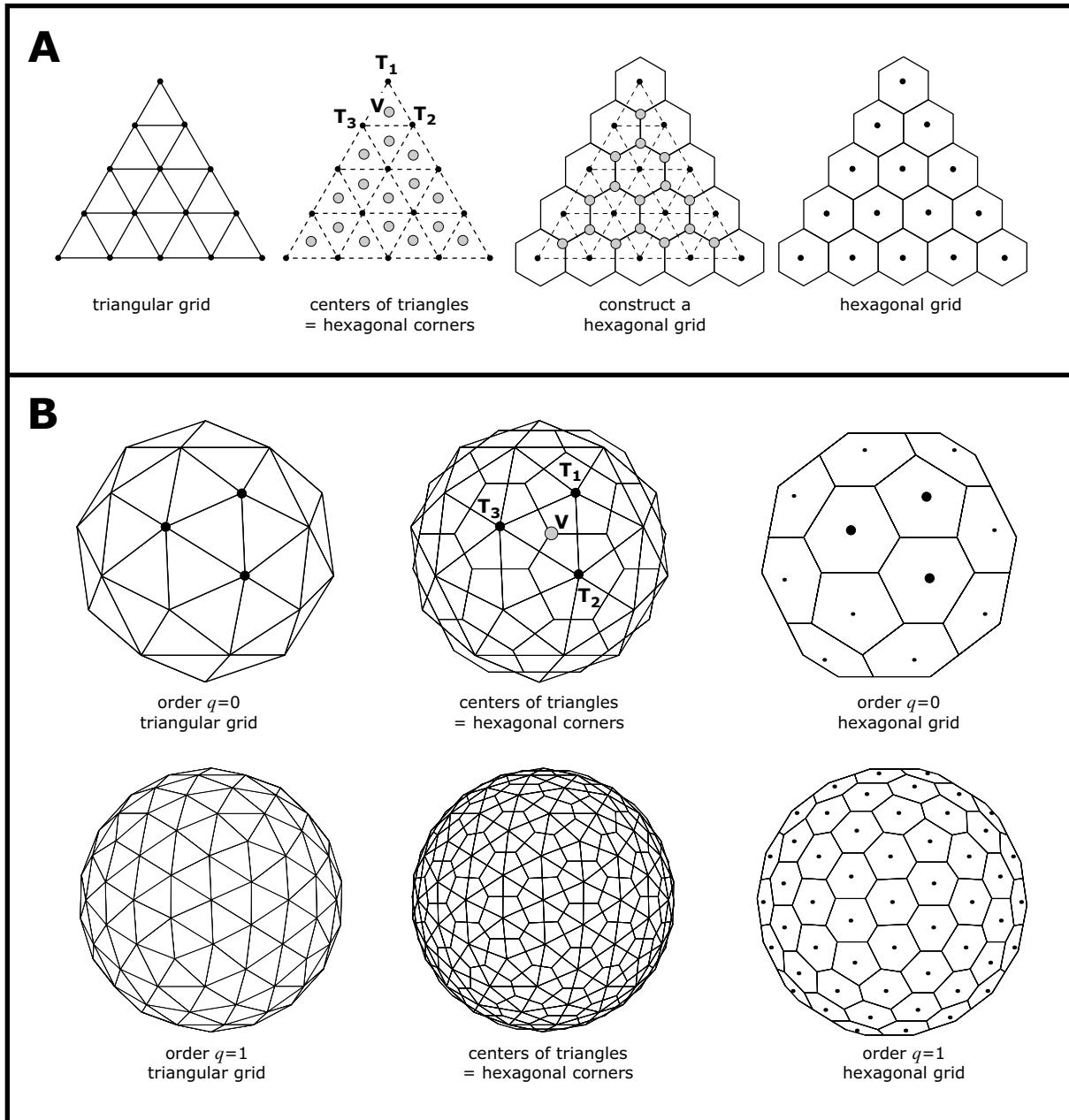


Figure 4.7: Construction of a hexagonal grid from a triangular grid. **A.** Hexagonal grid corners (gray dots) are calculated from the triangular gridpoints (black dots) via Eq. 4.3. **B.** The method in A applied to the order $q=0$ and order $q=1$ triangular grids. The centers of the hexagonal grid faces are the vertices from the corresponding triangular grid.

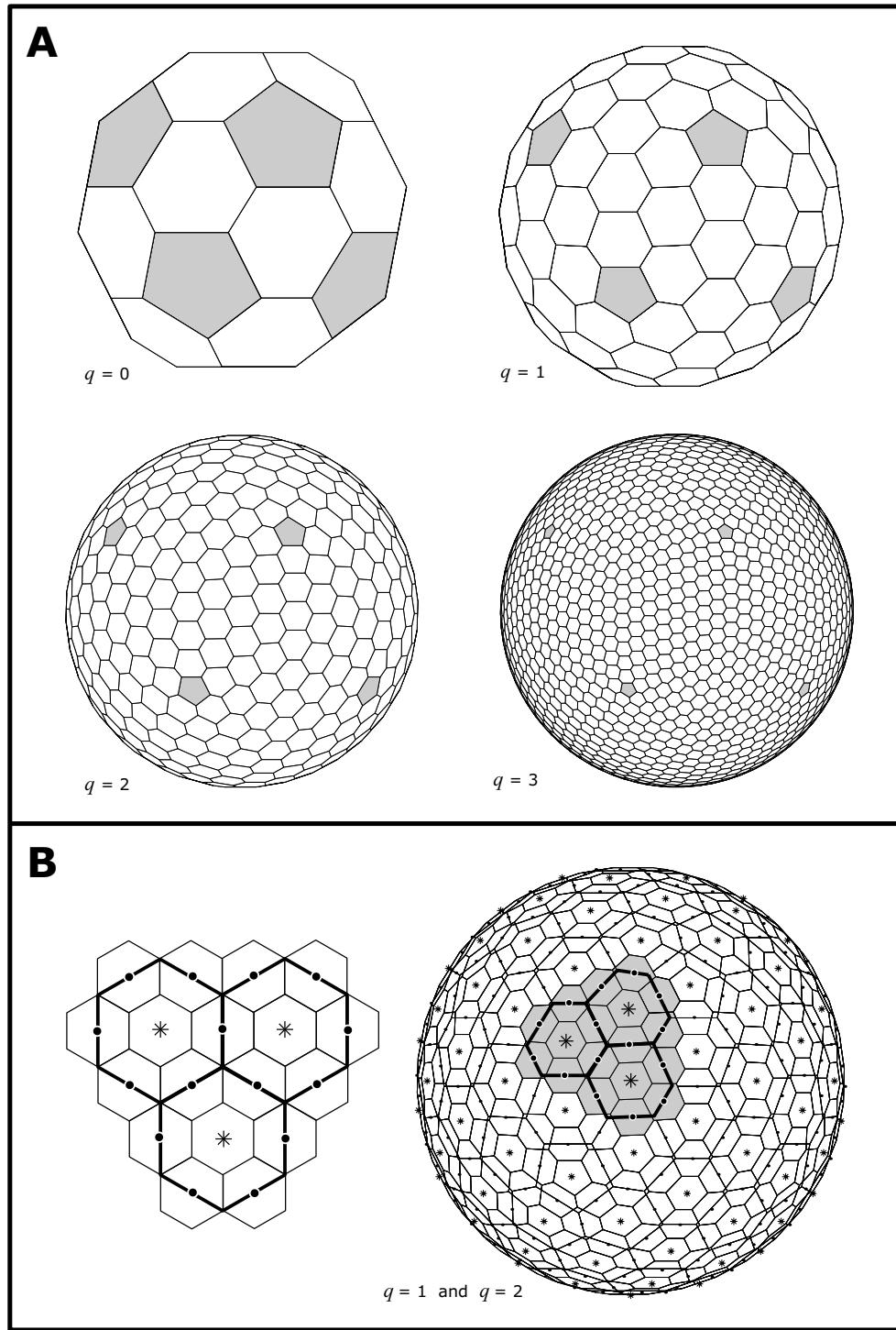


Figure 4.8: **A.** Hexagonal grids, orders $q=0$ to $q=3$. Each order q hexagonal grid has 12 pentagonal faces (gray) and $30 \cdot 4^q - 10$ hexagonal faces (see Table 4.3). The $q=0$ hexagonal grid is a buckyball. **B.** Relationship between successive orders of hexagonal grids. The lower-order grid ($q=1$) contains larger faces with '*' at each center; the higher-order grid ($q=2$) contains smaller faces with '●' at each center.

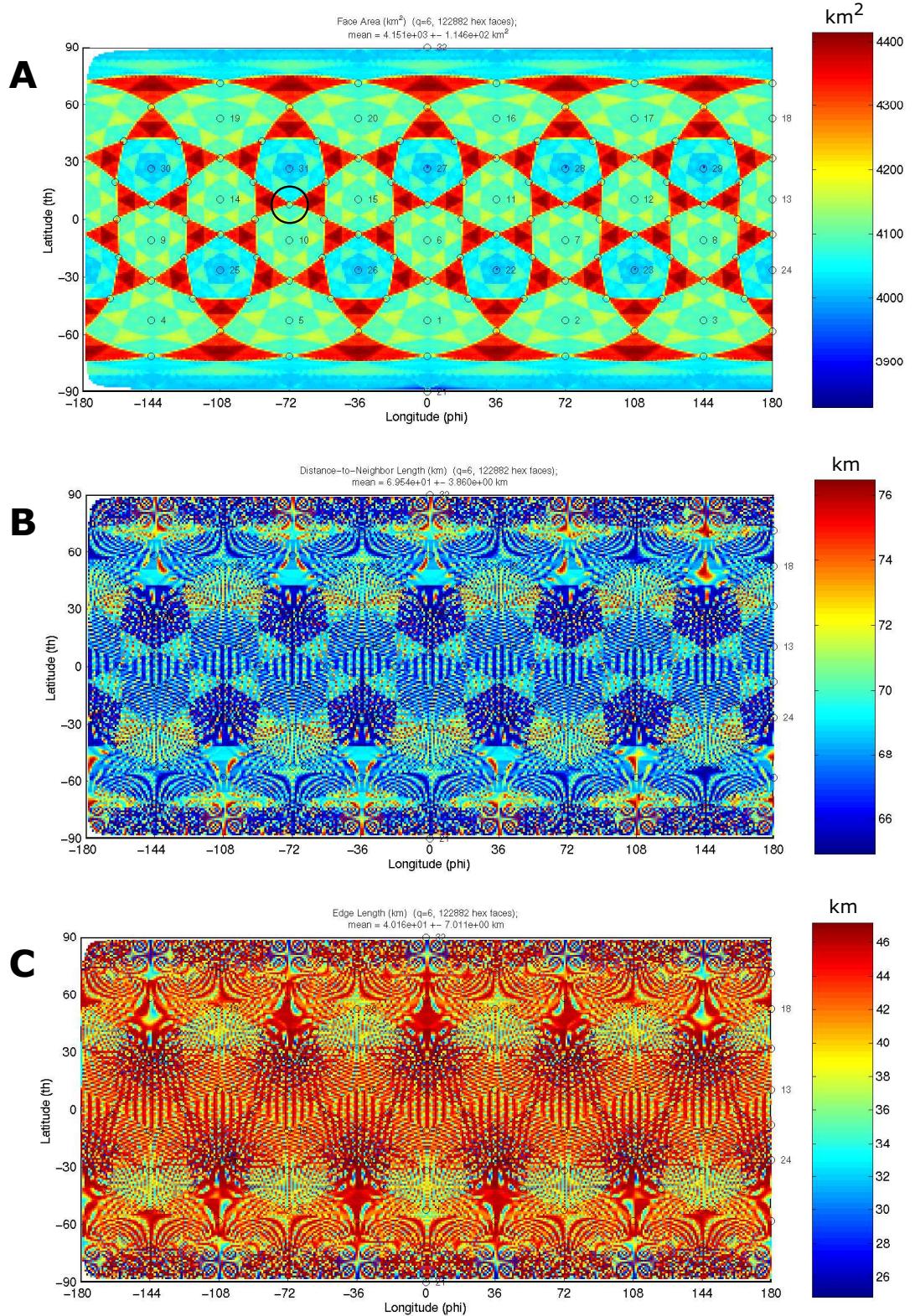


Figure 4.9: Geometric properties of the order $q=6$ hexagonal grid, showing the latitude-longitude variations of the hexagonal face areas (A), the distance between hexagonal faces (B), and the lengths of the hexagonal face edges (C). Superimposed circles denote the 122 order $q=1$ triangular vertices (i.e., hexagonal face centers); circles with labels denote the 32 order $q=0$ triangular vertices (a subset of the $q=1$ vertices). Figure 4.10 shows another view of these data. Circle in A indicates gridpoint 112 (see Figure 4.12).

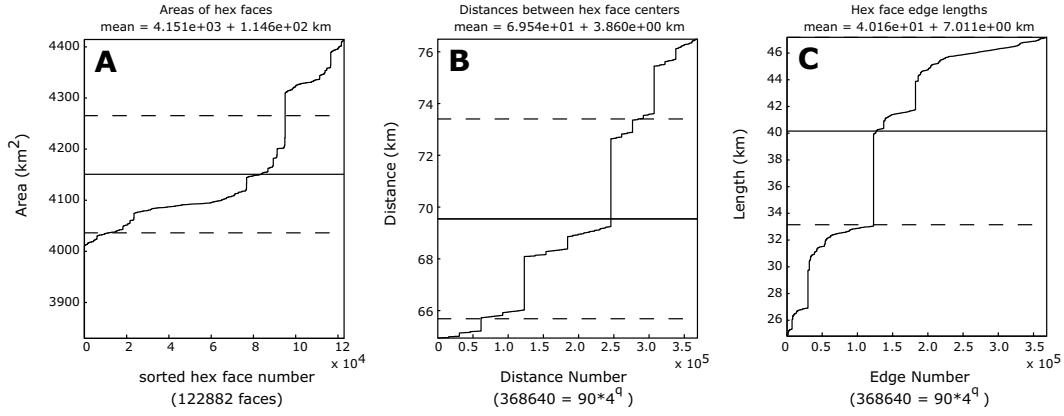


Figure 4.10: Geometric properties of the order $q=6$ hexagonal grid, showing the distributions of the hexagonal face areas (**A**), the distance between hexagonal faces (**B**), and the lengths of the hexagonal face edges (**C**). The data are the same as those shown in Figure 4.9, only here they are sorted in increasing order. Solid line is the mean; dashed lines show the standard deviation.

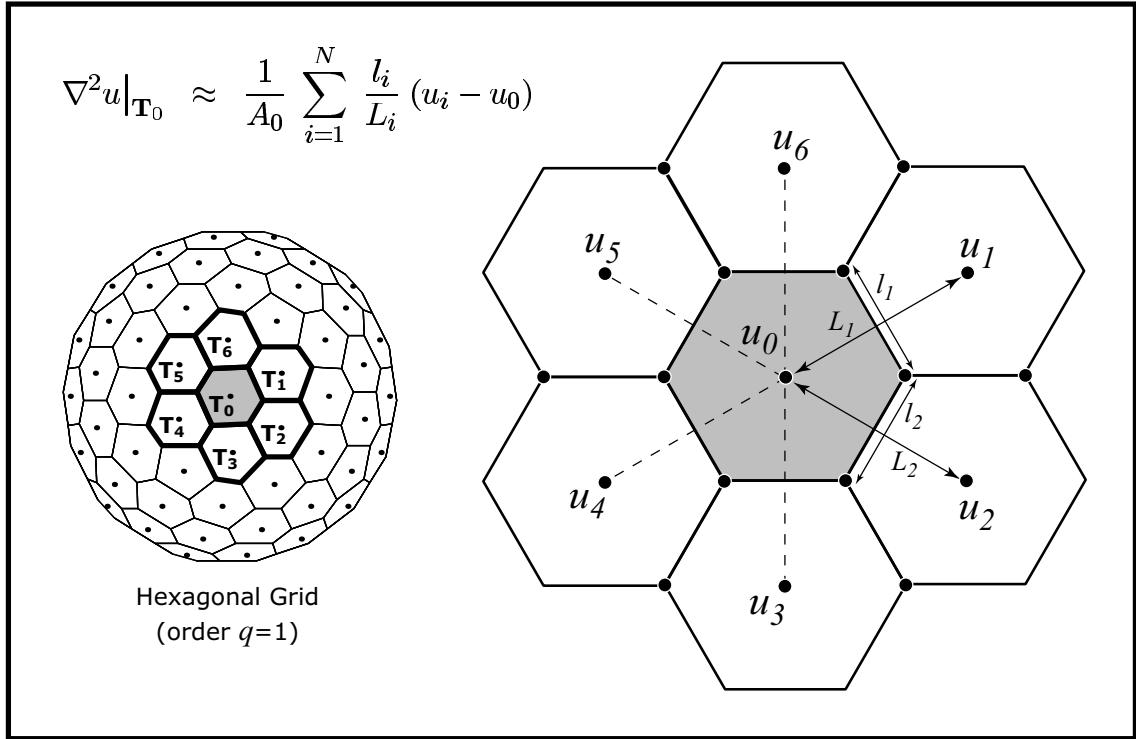


Figure 4.11: Finite-difference approximation for the Laplacian at a hexagonal face center \mathbf{T}_0 . The function $u(\theta, \phi)$ is defined at all face centers as $u(\mathbf{T}_i) = u_i$, and l_i and L_i are arc distances on the unit sphere. The Laplacian at \mathbf{T}_0 is calculated via Eq. 4.12 (shown here), where A_0 is the spherical area of the target face with center \mathbf{T}_0 on the unit sphere. As shown in Figure 4.7, the hexagonal face centers are the vertices of the triangular grids.

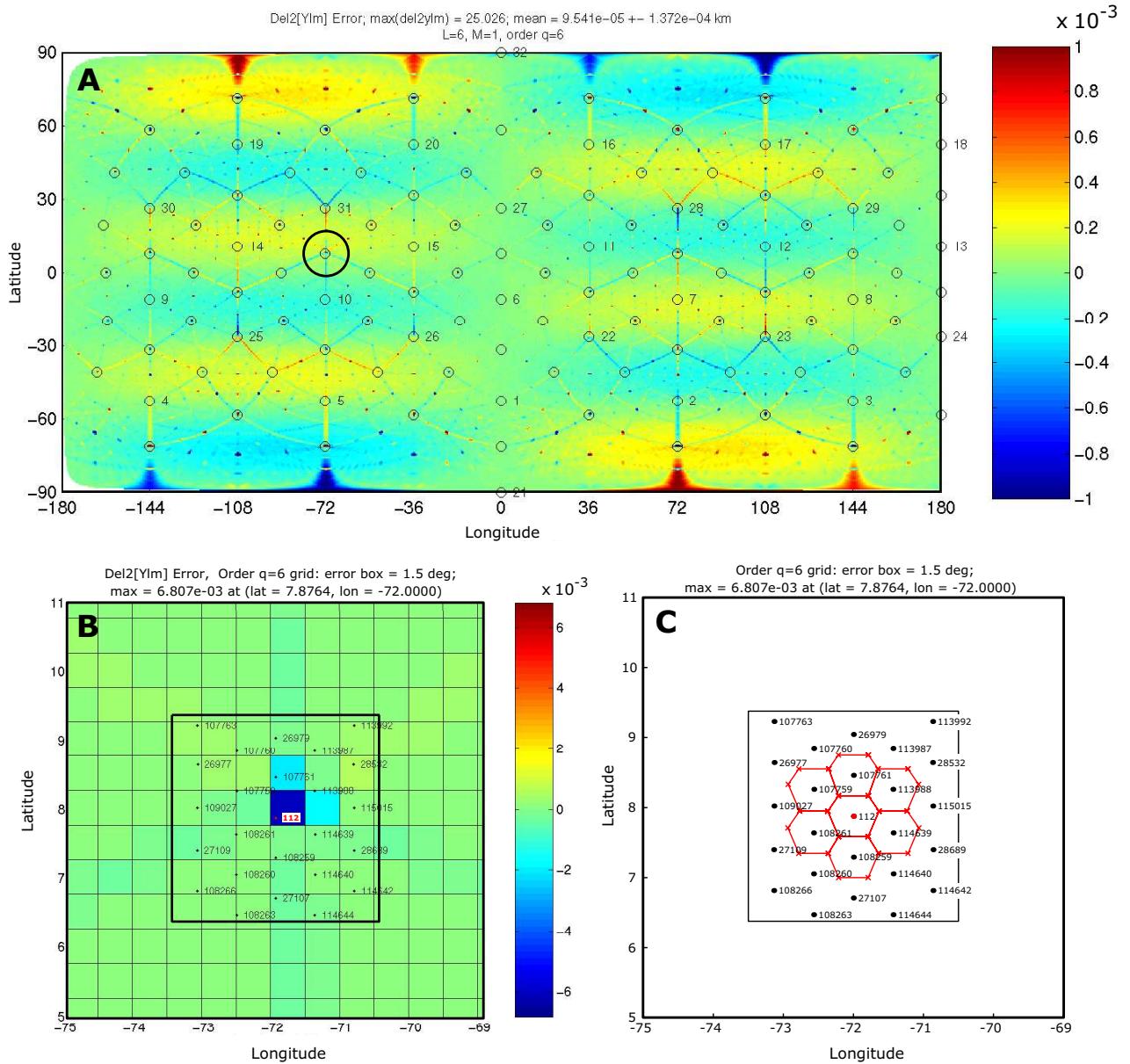


Figure 4.12: Error in the numerical Laplacian approximation for the order $q=6$ hexagonal grid. **A.** The error is calculated for each of the 122,882 gridpoints via Eq. 4.15, and then the values are interpolated onto a latitude-longitude grid with cells 1° -by- 1° . Lack of coverage near the poles in the interpolation produces the white areas. Superimposed circles denote the 122 order $q=1$ triangular vertices; circles with labels denote the 32 order $q=0$ triangular vertices (a subset of the $q=1$ vertices). Large circle indicates gridpoint 112 (see Figure 4.9). The color scale is saturated in order to accentuate the pattern of the error on the grid. **B.** Zoom-in on the maximum value of the error, which occurs at index 112, one of the $q=1$ vertices. Here the color scale is the default, with the range spanning twice the maximum value of the error. **C.** Configuration of hexagonal gridpoints used to calculate the error in the numerical Laplacian at 112. See Section 4.5.

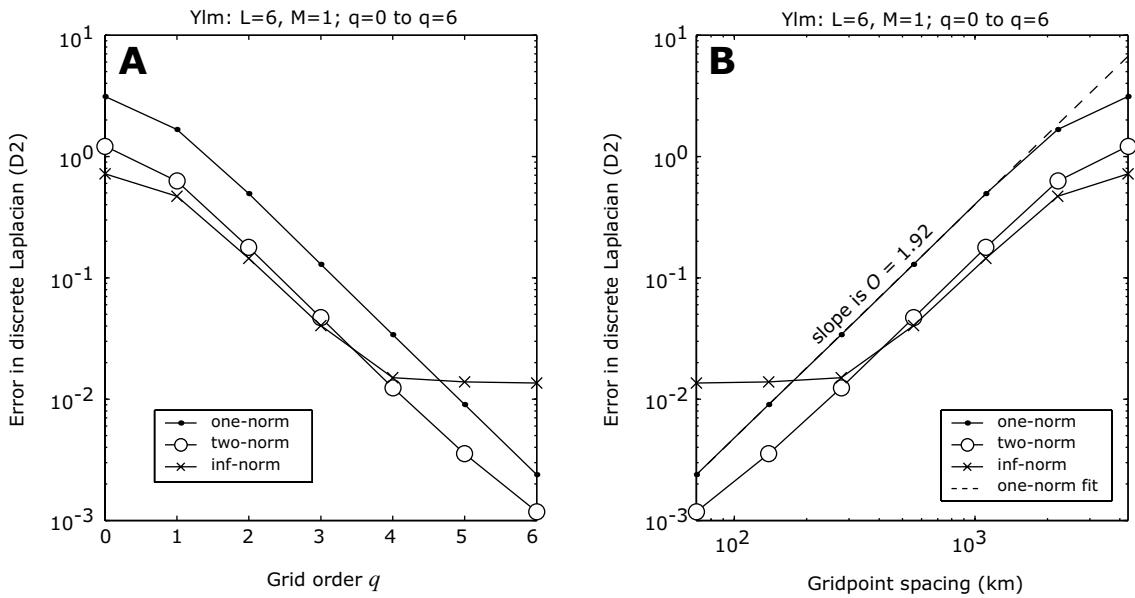


Figure 4.13: Error in the numerical approximation of the Laplacian. The error norms are calculated via Eqs. 4.20–4.22 for a $Y_{lm}(\theta, \phi)$ test function with $l = 6$ and $m = 1$. In **A** the x -axis data are plotted in terms of the grid order q ; in **B** they are plotted on a log-scale in terms of the average gridpoint spacing for a particular order q grid (these values are shown in the sixth column of Table 4.4). Notice that the error in the infinity norm does not appear to converge to zero for this test function. The slope of the curves in **B** gives the order of accuracy for a particular norm. Here the line is fit using the one-norm data for the gridpoint spacing for $q=3$ to $q=6$. See Section 4.5.

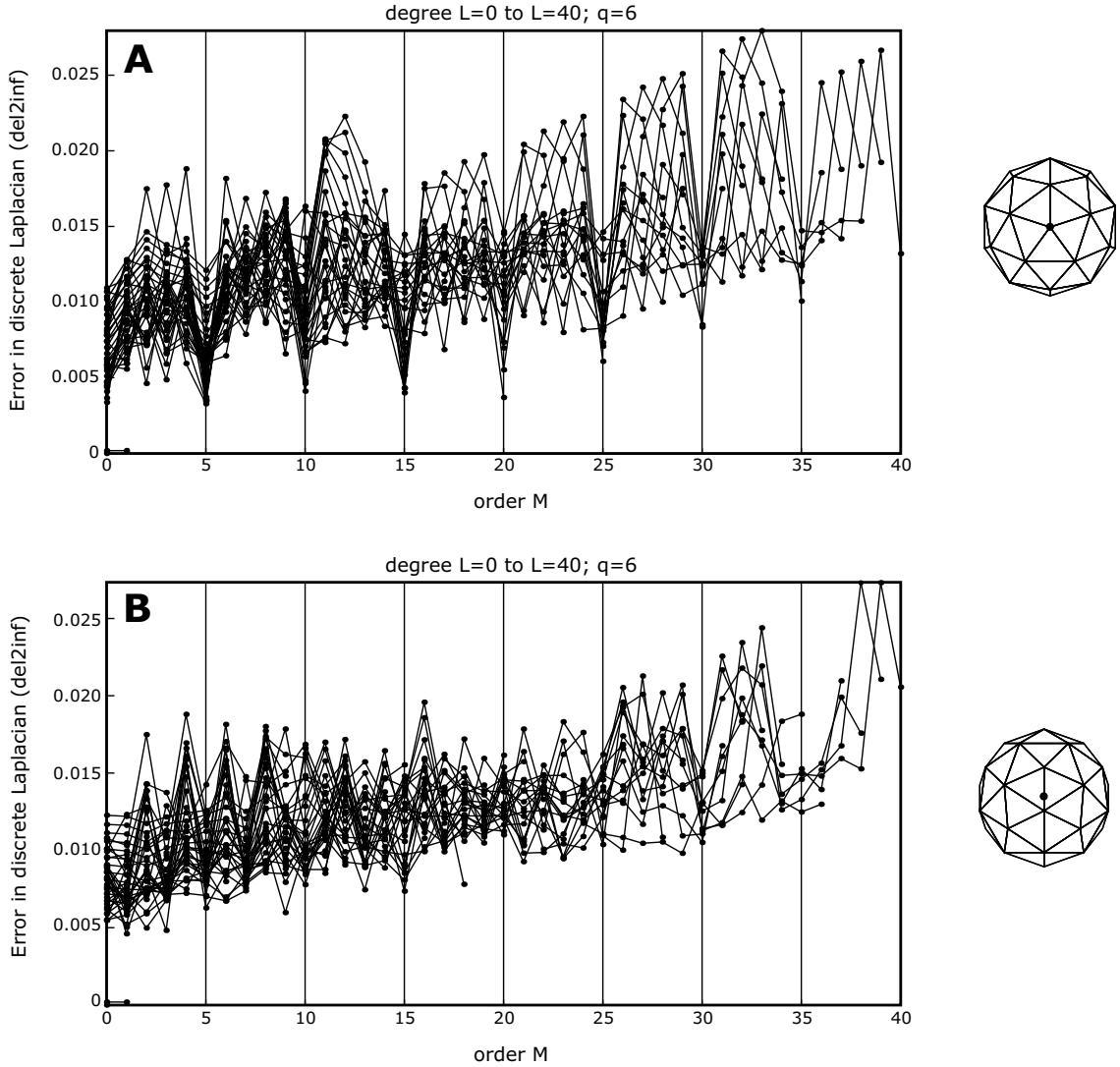


Figure 4.14: Error in the numerical approximation of the Laplacian, plotted using the infinity norm (Eq. 4.22) for the order $q=6$ grid for all Y_{lm} test functions with harmonic degree $0 \leq l \leq 40$ and harmonic order $0 \leq m \leq l$. A and B compare the infinity norms for two orientations of the order $q=6$ grid. **A.** Standard orientation of the grid (right), revealing the five-fold symmetry (with ϕ) about the z -axis — view is looking down the z -axis (order $q=0$ grid). The dips in error for the test functions having order $m = 5N$ ($N = 0, 1, \dots, 8$) are probably due to the test functions' symmetry coinciding with the symmetry of the grids. **B.** Rotated orientation of the grids. Here the grids have been rotated by $\beta = -90^\circ$ about the y -axis. The relative decreases in error are no longer present. See Section 4.5.

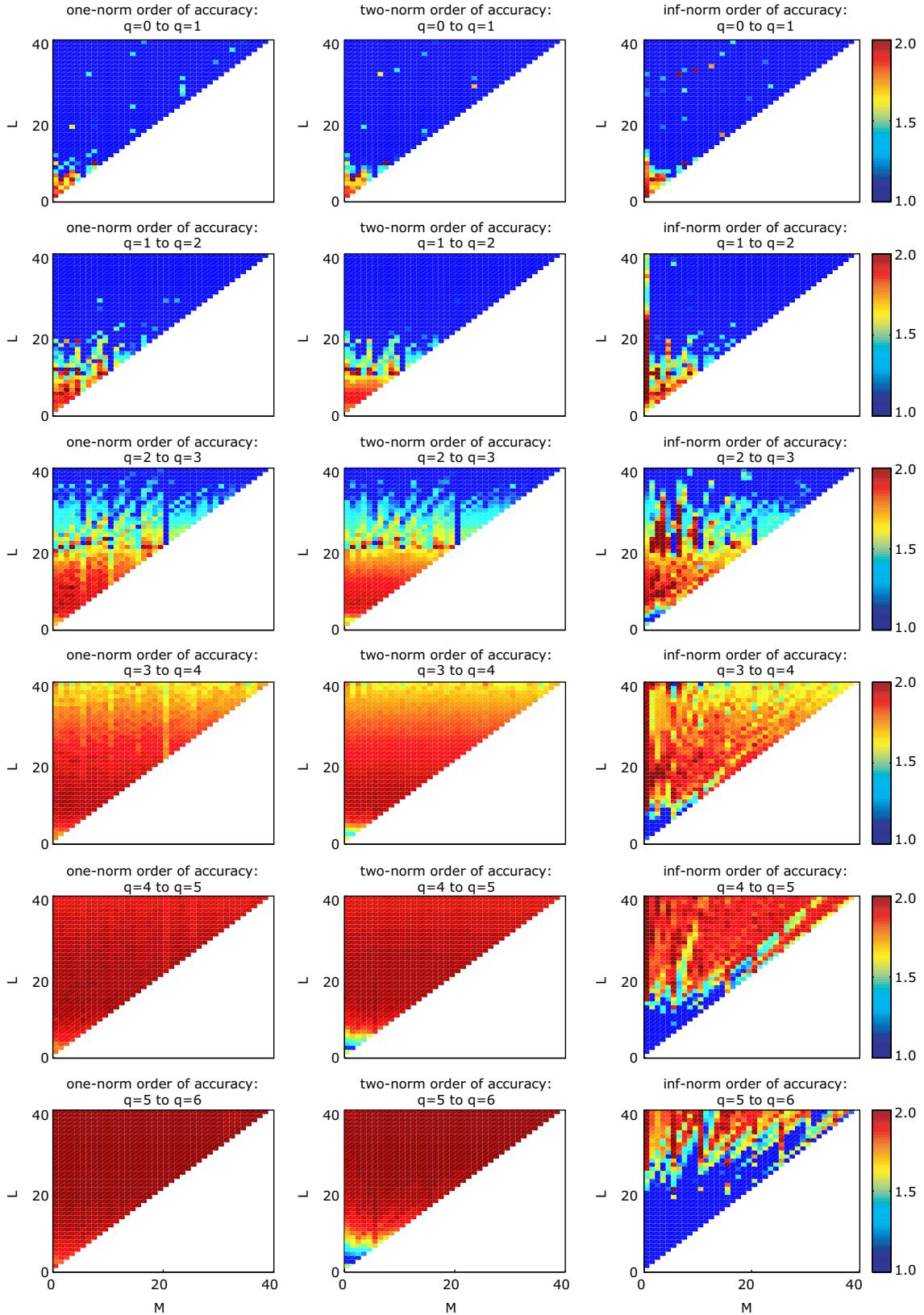


Figure 4.15: Order of accuracy of the numerical approximation of the Laplacian. The order of accuracy is a function of the grid order q , the Y_{lm} test function, and the norm that is the gauge for the error. Thus, each colored point is a comparison between the error in the discrete Laplacian of two spherical-hexagonal grids for a particular Y_{lm} test function. Note that the one-norms and two-norms are nearly second-order accurate (red) for the finest grids, while the infinity-norms are zero- or first-order accurate (blue). See Section 4.6.

Chapter 5

Numerical solutions of waves on a spherical membrane

Contents

5.1	Introduction	60
5.2	The phase velocity function $c(\theta, \phi)$	61
5.3	Numerical iteration for wave propagation	64
5.4	Numerical solutions for homogeneous and heterogeneous phase velocity maps	64
5.5	Stability and accuracy in the numerical model	66

5.1 Introduction

In the previous chapter we developed our spherical grids and presented our numerical method for calculating the discrete approximation of the Laplacian operator on the grids. The spatial part of the wave equation has thus been discretized, and now it is a matter of choosing a phase velocity function $c(\theta, \phi)$ (Section 5.2), and then iterating the wave equation (Section 5.3).

However, there are several complexities that arise in the numerical problem, most of which result from discretizing a continuous equation into discrete spatial units and temporal units (Section 5.5). As mentioned in the introduction, our wave equation (Eq. 3.1) is quite simple, as it is based on a non-dispersive, isotropic spherical membrane. Nevertheless, the *discrete*

form of the wave equation, including the irregular discretization of the spatial gridpoints, induces several of these effects, notably numerical dispersion and anisotropy (e.g., *Trefethen*, 1982). Thus we must take care in establishing stability and accuracy conditions, in order to establish numerical solutions that match the analytical solutions for the homogeneous case ($c(\theta, \phi) = \text{const}$).

5.2 The phase velocity function $c(\theta, \phi)$

Amplitude spectrum of $c(\theta, \phi)$

For the purposes of comparisons with ray theory, we can select *any* reasonable function to insert into the wave equation for $c(\theta, \phi)$. It is convenient to write $c(\theta, \phi)$ as an expansion in spherical harmonics (see Appendix B). The *amplitude spectrum* of $c(\theta, \phi)$, $R(l)$, is calculated for each degree l via

$$R(l) = (A_{l0})^2 + \sum_{m=1}^l \left[(A_{lm})^2 + (B_{lm})^2 \right], \quad (5.1)$$

where A_{lm} and B_{lm} are the spherical harmonic coefficients describing $c(\theta, \phi)$ (e.g., Eq. B.15). We normalize the amplitude spectra to the maximum coefficient value for all the spectra being compared in a given plot.

Homogeneous earth

The phase velocity for the heterogeneous earth, $c(\theta, \phi)$, is expressed as a perturbation from a homogeneous earth reference model, $c_{\text{hom}}(\theta, \phi) = \text{const}$:

$$\frac{\delta c(\theta, \phi)}{c_{\text{hom}}} = \frac{c(\theta, \phi) - c_{\text{hom}}}{c_{\text{hom}}}. \quad (5.2)$$

For our reference model, we use the values calculated from PREM (*Dziewonski and Anderson*, 1981), i.e., $c_{\text{hom}}(\theta, \phi) = c_{\text{PREM}}$. The amplitude spectrum of our reference model is, by definition¹, $R(l) = 0$; any different homogeneous earth model will have a nonzero $l=0$ component (Figure 5.1).

¹In other words, the deviation of the reference model from itself is zero.

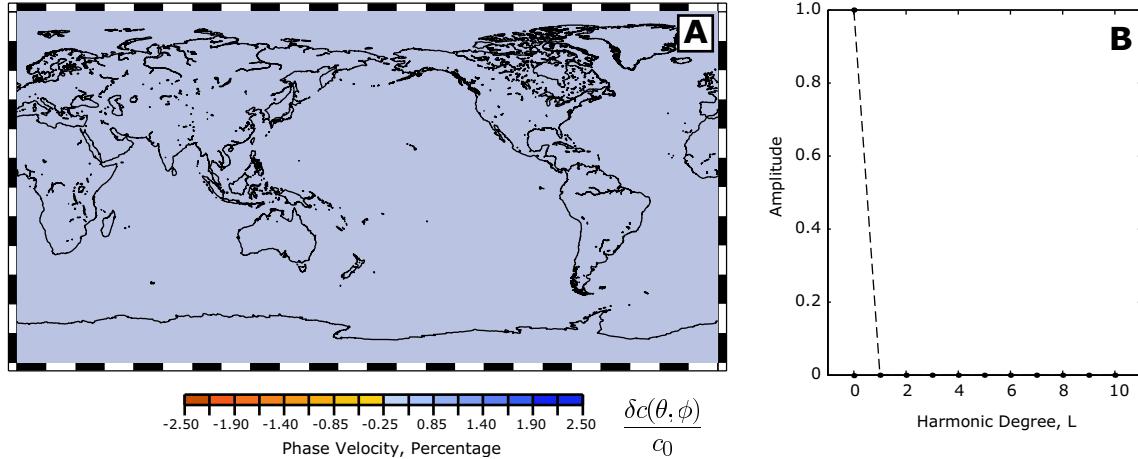


Figure 5.1: Homogeneous earth reference model. **A.** The phase velocity is constant on a homogeneous earth. Phase velocity maps are plotted as a phase velocity perturbation with respect to a homogeneous reference model. Thus, the light blue corresponds to a 0% perturbation from the reference model. **B.** The amplitude spectrum of the reference model is $R(l) = 0$. Any other homogeneous earth model would have a non-zero $l=0$ component (dashed line).

Heterogeneous earth

By selecting an actual phase velocity map for $c(\theta, \phi)$, our problem becomes particularly relevant to seismology, since ray theory has been used with these maps to make predictions of phase and amplitude for surface wave arrivals in seismograms (e.g., *Woodhouse and Wong*, 1986). We will be examining three variations on the phase velocity maps of *Trampert and Woodhouse* (1995): (1) the period T (Figure 5.2), (2) the truncation degree l_{max} of the spherical harmonic expansion (Figure 5.3), and (3) a perturbation factor ϵ that controls the power of the maps (Figure 5.4). Thus, with three parameters to test, we may think of the phase velocity as

$$c(\theta, \phi) = c(\theta, \phi, T, l_{max}, \epsilon). \quad (5.3)$$

Note that the color-scale in the phase velocity maps is always the same.

Vary the period T . *Trampert and Woodhouse* (1995) provide Rayleigh and Love wave phase velocity maps for periods of 40, 60, 80, 100, 150, and 200 seconds. Figure 5.2 shows three of these maps, and the corresponding amplitude spectra are shown in Figure 5.5A-B. The PREM phase velocity for each map is shown in Table 5.1. For each map, the spherically-averaged phase velocity value differs from the PREM value, indicating the presence of a

nonzero $l=0$ component.

Upon examining Figure 5.2, we see that the shorter-period map ($T=40\text{s}$) contains a greater amount of shorter-scalelength structure ($l > 20$). Table 5.1 shows the periods and corresponding wavelengths for each of the maps of *Trampert and Woodhouse* (1995). In general, the surface wave maps “average” the structure down to the depth of the corresponding wavelength of their period². For example, the $T=40\text{s}$ map is sensitive to lateral structure with scalelength as short as $\approx 160 \text{ km}$, and it shows the “averaged” vertical structure down to a depth of approximately 160 km.

Vary the truncation degree l_{max} . We can easily create smoother functions by truncating the spherical harmonic expansion at some $l_{max} < 40$. Figure 5.3 shows three maps, with differing l_{max} values, and the corresponding amplitude spectra are shown in Figure 5.5C. Table 5.2 shows how the minimum scalelength Λ changes as a function of l_{max} .

Vary the power ϵ . We now include a perturbation factor ϵ in the phase velocity maps:

$$\epsilon \left(\frac{\delta c(\theta, \phi)}{c_{\text{hom}}} \right), \quad 0 \leq \epsilon \leq 1. \quad (5.4)$$

Thus, the maps with $\epsilon=0$ represent the homogeneous earth reference model, and the maps with $\epsilon=1$ are exactly those of *Trampert and Woodhouse* (1995). Figure 5.3 shows three maps, with differing ϵ values, and the corresponding amplitude spectra are shown in Figure 5.5D.

Other sets of phase velocity maps. A review of studies producing surface wave phase velocity maps can be found in *Dahlen and Tromp* (1998, Section 16.9) (e.g., *Ekström et al.*, 1997; *Laske and Masters*, 1996), and further discussion is presented in *Trampert and Woodhouse* (2001). We also tested the maps of *Trampert and Woodhouse* (1996), which incorporate more major arc data (i.e., R2 orbits) than those of TW95. Figure 5.6 shows a map of TW96 in comparison with a map of TW95. We immediately notice that the map of TW96 contain a greater amount of shorter-scalelength structure (Figure 5.6C). *Trampert and Woodhouse* (1996) state that the models in TW95 are “long wavelength approximations” (p. 22) of the models in TW96.

²It is more accurate to talk of the *sensitivity kernels* of the Rayleigh waves (fundamental spheroidal mode). These functions give more insight into how the propagating waves sample the structure with depth (see *Woodhouse*, 1996).

5.3 Numerical iteration for wave propagation

The numerical iteration for wave propagation on the spherical membrane is essentially the same as that of the planar membrane (Eqs. 2.34–2.36). The only difference is that here, for the discrete Laplacian approximation $\nabla^2 u_t$, we use `D2` of Equation (4.12) instead of `del2`. Furthermore, we do not require a boundary condition, only the initial conditions, which are $u_{t-1} = u_t = u_{t+1} = 0$. Thus the iteration at each timestep is:

$$u_{t-1} = u_t \quad (5.5)$$

$$u_t = u_{t+1} \quad (5.6)$$

$$u_{t+1} = 2u_t - u_{t-1} + c^2 dt^2 (\text{D2}(u_t) + f_t) . \quad (5.7)$$

where $f_t = f(\theta, \phi, t)$ is the (discrete) forcing term, $u_t = u(\theta, \phi, t)$ is the displacement at time t at all θ - ϕ gridpoints on the sphere, and `D2` is the discrete Laplacian for the hexagonal grids. The phase velocity function is entered as $c(\theta, \phi) = c_{\text{PREM}}$ for the homogeneous case, or as a phase velocity map for the heterogeneous case.

5.4 Numerical solutions for homogeneous and heterogeneous phase velocity maps

This section contains numerical solutions to the wave equation on a spherical membrane for both a homogeneous earth and a heterogeneous earth. The solution, $u(\theta, \phi, t)$, can be thought of as a global snapshot of the wave field (fixed time) or as a seismogram at a receiver (fixed space). Figures 5.7–5.10 shows plots of the solutions³. For the homogeneous earth reference model (Figure 5.1), we have chosen the PREM model for $T=40$ s surface waves, which corresponds to $c(\theta, \phi) = 3.9280$ km/s (Table 5.1). For the heterogeneous earth model, we have chosen the $T=40$ s, $l_{\text{max}}=12$ ($\epsilon=1.0$) Rayleigh wave phase velocity map of *Trampert and Woodhouse* (1995) (middle map in Figure 5.3). The minimum scalelength of this map is $\Lambda = 3202$ km (Table 5.2). The forcing term for the solutions in this section is $f = f_2(\theta, \phi, t)$, with $\mu = 0.0713$ and $\sigma = 204.5$ (Figure 3.1).

In Figure 3.4 we introduced the analytical solutions to waves on the sphere for the forcing

³Our set of data for these figures consists of 450 timesteps of solutions on the order $q=4$ grid, which has 7682 gridpoints. In other words, we have 7682 seismograms, one for each gridpoint.

term $f_2(\theta, \phi, t)$. The solution, $u_{f_2}(\theta, \phi, t)$ (Eq. 3.32), provides a basis for comparison with the numerical solutions for the homogeneous case. Figure 5.7 shows the agreement between numerical and analytical seismograms for the homogeneous case. This agreement for the homogeneous case gives us confidence that the numerical solutions for the heterogeneous case are reasonably accurate.

In Figure 5.8 we use a heterogeneous phase velocity map, and the numerical solutions are significantly different from the homogeneous numerical solutions, especially at later times, when more interference of the waves has occurred. Figure 5.9 shows some of the waveforms for two selected times.

Figure 5.10 shows a snapshot of the wavefront for a source at the North Pole. In the case of the waves on the sphere, it is not obvious what the wavefront actually is. We define the wavefront as the maximum value of the displacement for a given longitude ϕ (see Figure 3.4). Thus we can plot the position of the wavefront and the amplitude of the wavefront, each as a function of longitude. Because of the sinusoidal nature of the pulse, we keep track of both the top and the bottom of the pulse (Figure 5.10).

For a homogeneous velocity field, and for a source at the North Pole, the wavefront will always have constant latitude; on a $\theta\text{-}\phi$ latitude-longitude grid, a wavefront propagating between the North Pole and South Pole will be a horizontal line. The horizontal lines in Figure 5.10A show the position of the homogeneous wavefront; the corresponding horizontal lines in Figure 5.10B show that there is no variation in the amplitude of the top and bottom of the pulse, as we would expect.

Some aspects of the position of the heterogeneous wavefront in Figure 5.10A are what we might expect. First, we see that the heterogeneous pulse is leading the homogeneous pulse, and this gap increases with time. This is due to a positive $l=0$ component in the phase velocity map, meaning that the spherical average of the map is greater than the PREM value, causing the wavefront to propagate faster than in PREM (e.g., *Trampert and Woodhouse*, 2001). (The amplitude is not similarly affected by the $l=0$ component.) Second, we see that the leading portion of the wavefront is along the longitude of $\approx 180^\circ$, corresponding to the high-velocity anomaly of the Pacific basin. In general, it is difficult to predict the shape (or position) of the wavefront, especially for a complicated phase velocity map and after many timesteps.

Scalelength and wavelength

In Figure 5.11 we show how the scalelength of the heterogeneity affects the numerical solutions of wave propagation. We pick one receiver at $\Delta \approx 90^\circ$ from the source, and we record the seismogram for four phase velocity maps: $lmax=4$, $lmax=8$, $lmax=12$, and $lmax=20$; the corresponding minimum scalelengths are $\Lambda = 8896$, 4709 , 3202 , and 1953 km, respectively (Table 5.2). We notice that the predominant shape of the arrivals associated with each orbit changes very little for $lmax > 8$. This is because in that regime, the minimum scalelength of the heterogeneity is less than the wavelength of the wave, which is $\lambda \approx 8300$ km in this example. In reality — and in the numerical model — we do not expect the wave to “see” heterogeneity with scalelengths longer than the wavelength of the wave. Nevertheless, we see that these relatively short-scalelength structures do affect the wave propagation (Figure 5.11).

Thus far we have looked at numerical results calculated using the order $q=4$ spherical grid (7682 gridpoints). In Chapter 7 we return to the numerical solutions and present results for the order $q=6$ grid (122,882 gridpoints), which allows for shorter wavelengths. As we discuss in Sections 7.4 and 7.5, shorter wavelengths allow us to explore the validity of ray theory predictions in shorter-scalelength earth models.

5.5 Stability and accuracy in the numerical model

An analysis of stability and accuracy is helpful for any numerical model (*Trefethen*, 1996). An unstable solution is one in which, for example, the solution “blows up” exponentially. An inaccurate solution is one in which, for example, dispersion arises due to numerical effects. Both of these are unwanted, and both can generally be predicted prior to running the numerical program.

There is a general stability condition that states the relationship between the timestep and spacestep for a numerical model (see *Trefethen*, 1996):

$$\frac{dx}{dt} \frac{1}{c} \approx R , \quad (5.8)$$

where R is some critical value. Using $dx = a d\theta$ (a is earth radius), Equation (5.8) becomes

$$dt \approx \frac{a d\theta}{c} R . \quad (5.9)$$

Table 4.4 shows $d\theta$ for the hexagonal grids, orders $q=0$ to $q=6$. Note that if we increase c by a factor of 4, then dt decreases by a factor of 4.

Given a uniform grid and a discretized wave equation, it is possible to calculate a value for R , or a more complicated stability condition. By trial-and-error, it appears that $R = 1/\sqrt{2}$ is a good value for choosing the time-step dt in our problem. Larger R values will cause the solutions to increase indefinitely; smaller R values are allowed, although they increase computation time without decreasing error. In addition, error introduced at each time-step may lead to a more inaccurate solution.

The inaccuracy of our numerical solutions is manifested primarily by numerical dispersion. The effects of numerical dispersion are entirely predictable *on regular grids* (e.g., Trefethen, 1982) (in fact, this “nuisance” can be rather interesting). A numerical dispersion relationship (for a particular FDM) will show the extent to which one would expect dispersion for a wave with a particular wavelength. On irregular grids, however, it is more a matter of trial-and-error in trying to determine the minimum wavelength that achieves a satisfyingly small amount of numerical dispersion⁴. We know that the dispersion is related to the wavelengths of the waves propagating on the grid, and thus to our source parameters: the source-time function $h(t)$ and the initial-shape function $g(\theta, \phi)$. The width of the source-time function must be scaled according to the time-step; thus we set the width of the function to be some multiple of dt . Similarly the initial-shape function $g(\theta, \phi)$ must be greater than or equal to some multiple of $d\theta$. These scaling values are chosen by trial-and-error in order to minimize the numerical dispersion in our problem. These aspects of numerical grids are generally discussed in terms of the number of gridpoints per wavelength (gpw). Our values of ~ 20 gpw is considerably greater than those of Komatitsch and Tromp (2002a) (~ 6 gpw) and Baig *et al.* (2003) (~ 7 gpw), two studies that model wave propagation using a pseudo-spectral method.

⁴The study of Sword *et al.* (1986a) suggests that it might be possible to approximate the accuracy conditions by assuming a regular planar grid of hexagons, and this would eliminate some of the trial and error.

Table 5.1: Wave parameters for the Rayleigh wave phase velocity maps of *Trampert and Woodhouse* (1995) (also shown for $T=200\text{s}$.). The wavelengths are calculated via $\lambda = 2\pi a/(l + \frac{1}{2})$; the phase velocity values are calculated using normal-mode summation with PREM (see Appendix F). Each map is expanded up to $l_{\max}=40$. Note that the values in this table are properties of the “probing” surface waves used to create the maps.

Period, T (s)		Harmonic Degree, l	Wavelength, λ (km)	PREM Phase Velocity, c_0 (km/s)
approx.	exact			
40	40.043	254	157.29	3.9280
60	60.147	167	238.99	3.9734
80	79.910	124	321.53	4.0236
100	100.393	97	410.57	4.0896
150	149.124	62	640.48	4.2950
200	200.837	43	920.23	4.5820

Table 5.2: Minimum scalelength in a phase velocity map expanded in spherical harmonics up to degree l_{\max} . The scalelengths are calculated via $\Lambda = 2\pi a/(l + \frac{1}{2})$ (see Appendix F).

l_{\max}	Scalelength, Λ (km)
4	8895.59
6	6158.49
8	4709.43
10	3812.40
12	3202.41
14	2760.70
16	2426.07
18	2163.79
20	1952.69
40	988.40

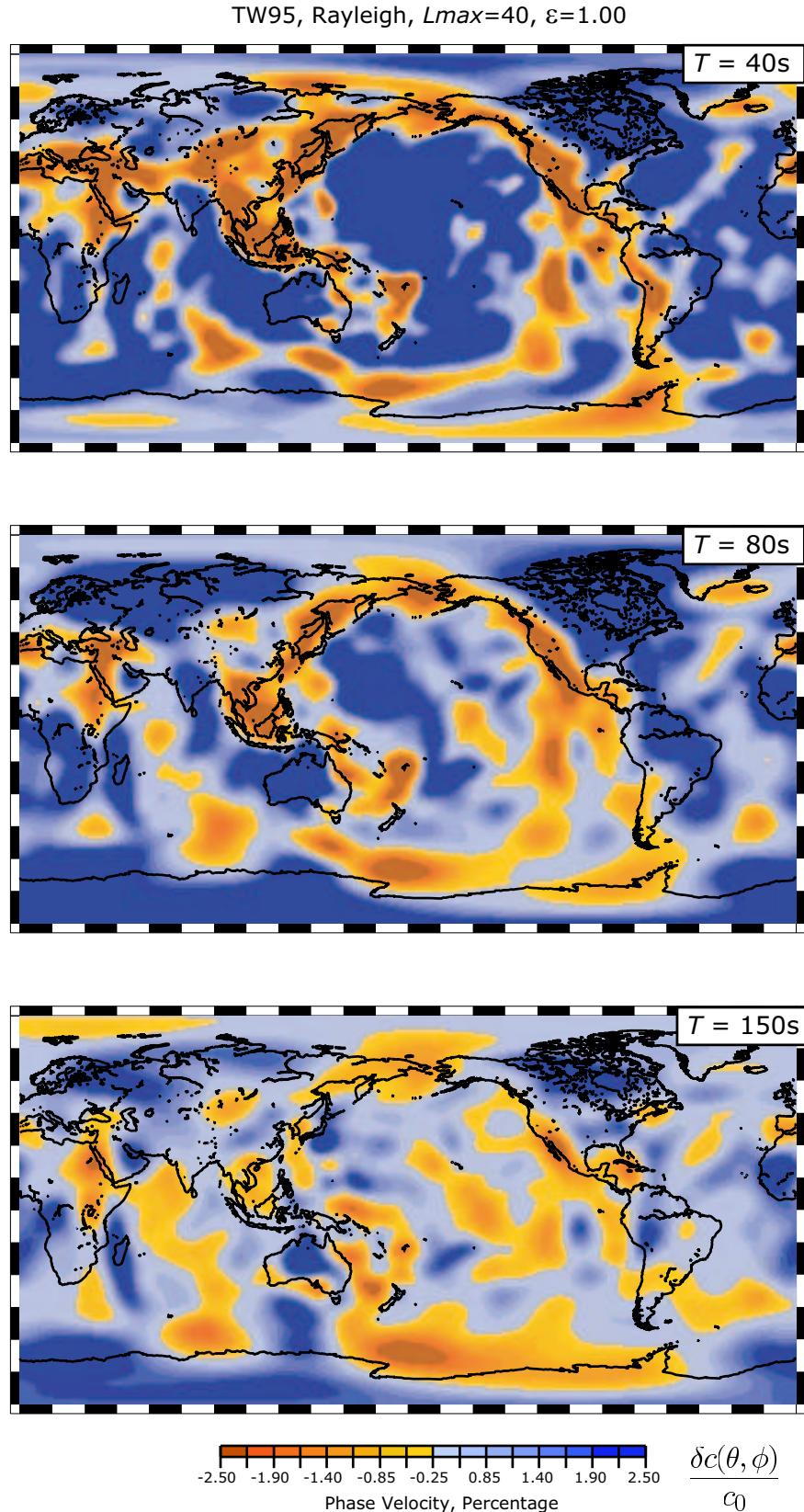


Figure 5.2: Phase velocity maps of *Trampert and Woodhouse* (1995) for Rayleigh waves, periods $T=40\text{s}$, $T=80\text{s}$, and $T=150\text{s}$. Each map is expanded in spherical harmonics up to degree $l_{max}=40$. Values are expressed as $100(c - c_0)/c_0$, where c_0 is from PREM (Table 5.1). See Figure 5.5A-B.

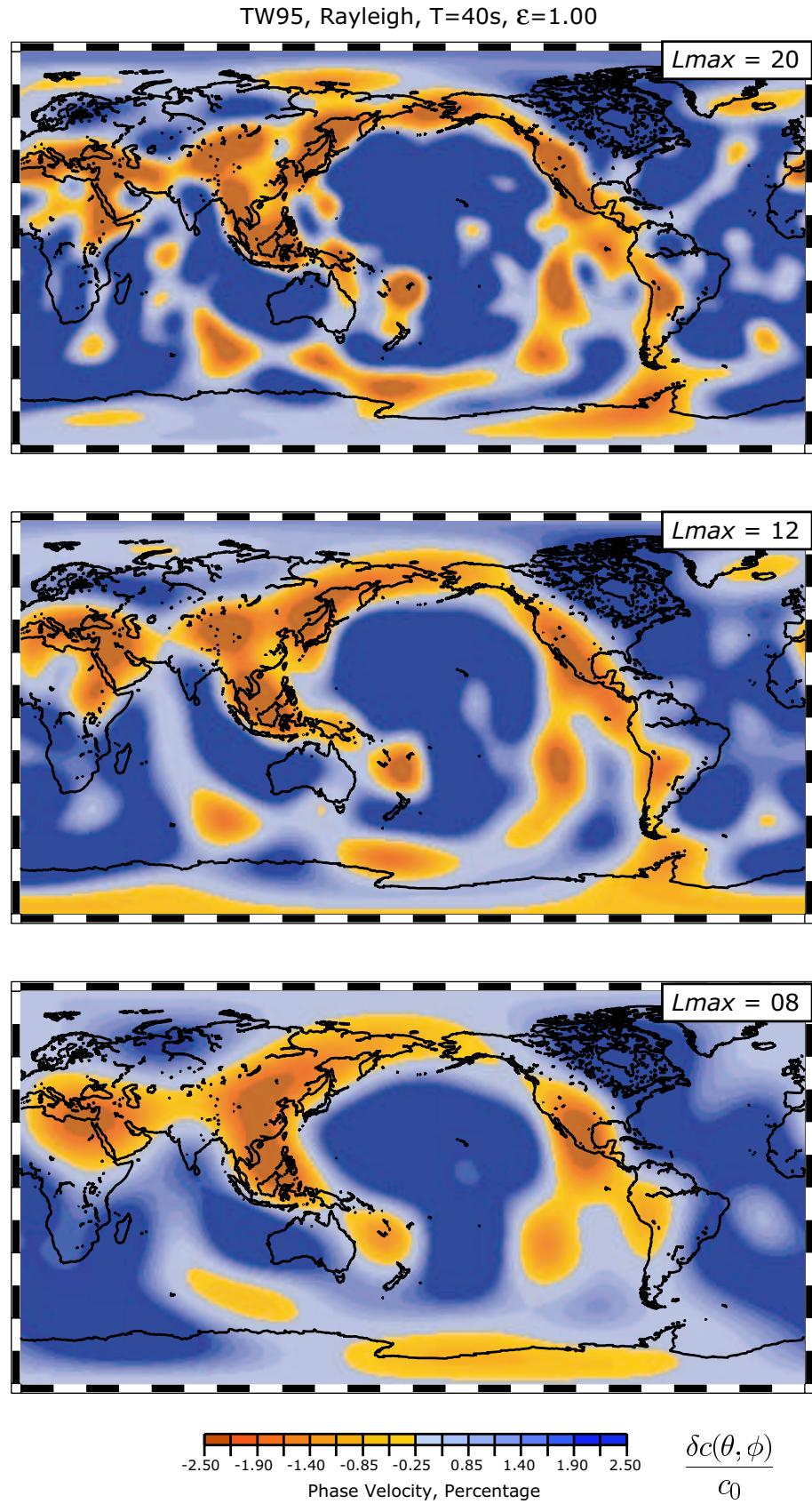


Figure 5.3: Varying l_{max} on the phase velocity maps. Three different l_{max} truncations of the $T=40s$, $l_{max}=40$ Rayleigh wave phase velocity map (Figure 5.2A). See Figure 5.5C.

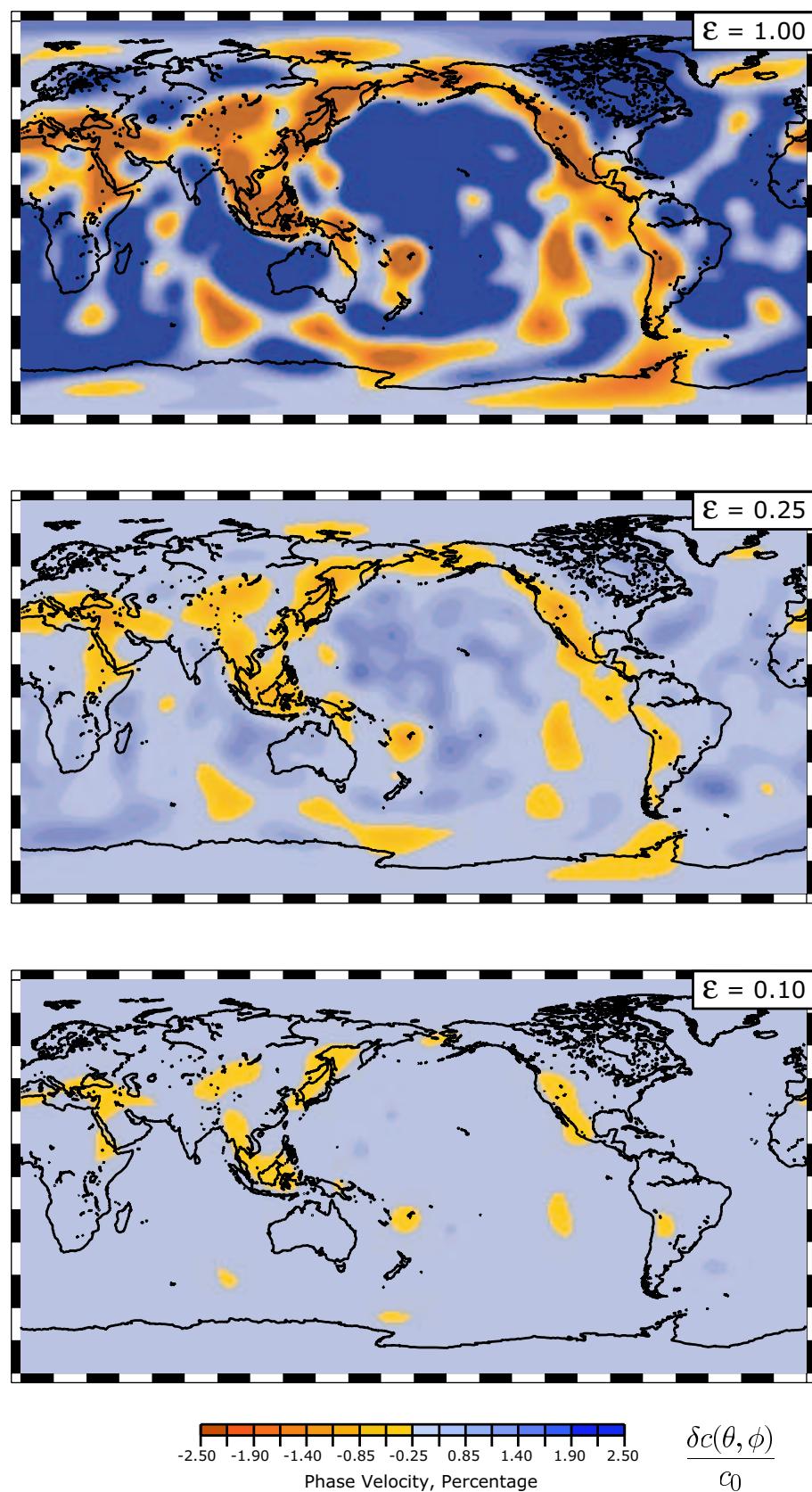


Figure 5.4: Varying ϵ on the phase velocity maps (see Eq. 5.4). Three different ϵ values for the $T=40s$ Rayleigh wave phase velocity map, truncated at $lmax=20$ (Figure 5.2A). See Figure 5.5D.

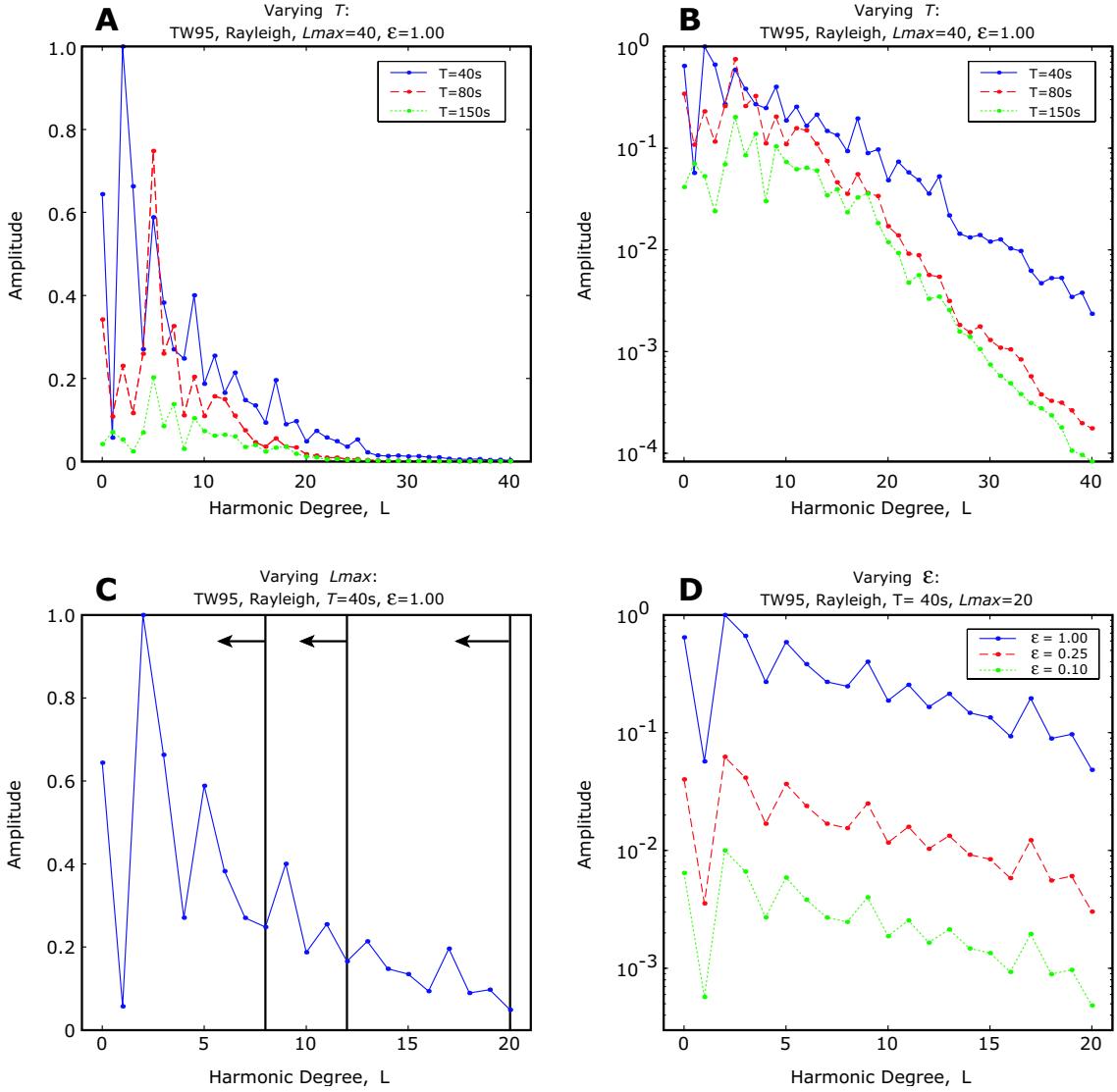


Figure 5.5: Amplitude spectra for phase velocity maps of *Trampert and Woodhouse* (1995) shown in Figures 5.2–5.4, which show variations in T , l_{max} , and ϵ . **A.** Variation in maps with period T (Figure 5.2). Spectra are normalized to the maximum coefficient value for all the maps ($l=2$ for $T=40s$). The decay for $l > 20$ is a primarily product of the inversion procedure and is not indicative of the earth structure (*Trampert and Woodhouse*, 1996; *Trampert and Snieder*, 1996). **B.** Same as A, with amplitude plotted on a logarithmic scale. **C.** Variation in maps with truncation degree l_{max} (Figure 5.3). **D.** Variation in maps with ϵ (Figure 5.4). Note that the amplitude is plotted on a logarithmic scale.

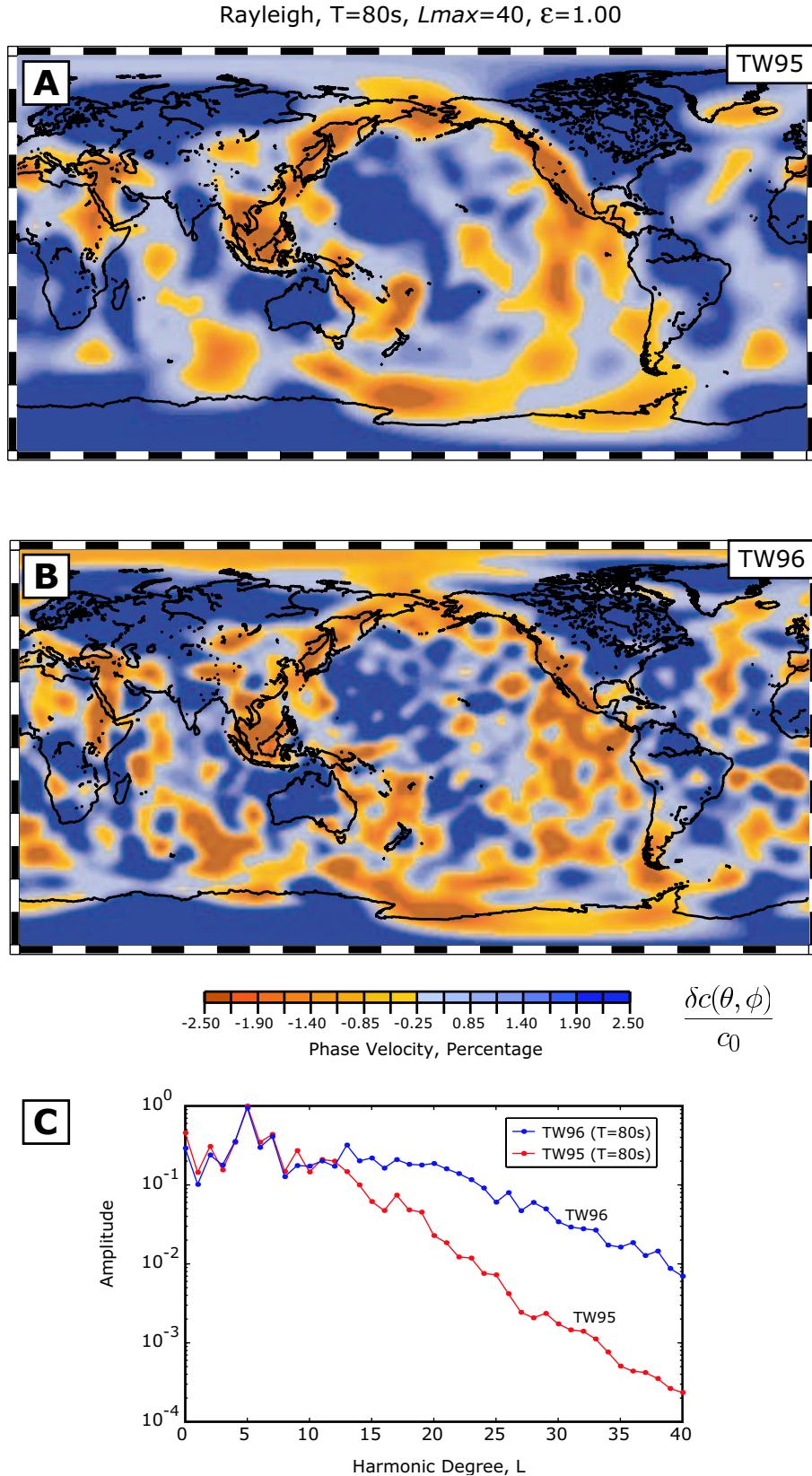


Figure 5.6: Surface wave phase velocity maps for $T=80s$, $l_{max}=40$, and $\epsilon=1.0$. **A.** Map from *Trampert and Woodhouse (1995)*. **B.** Map from *Trampert and Woodhouse (1996)*. **C.** Amplitude spectra of the two maps. Note the greater amount of shorter-scale length structure in TW96, especially for $l > 12$. See Section 5.2.

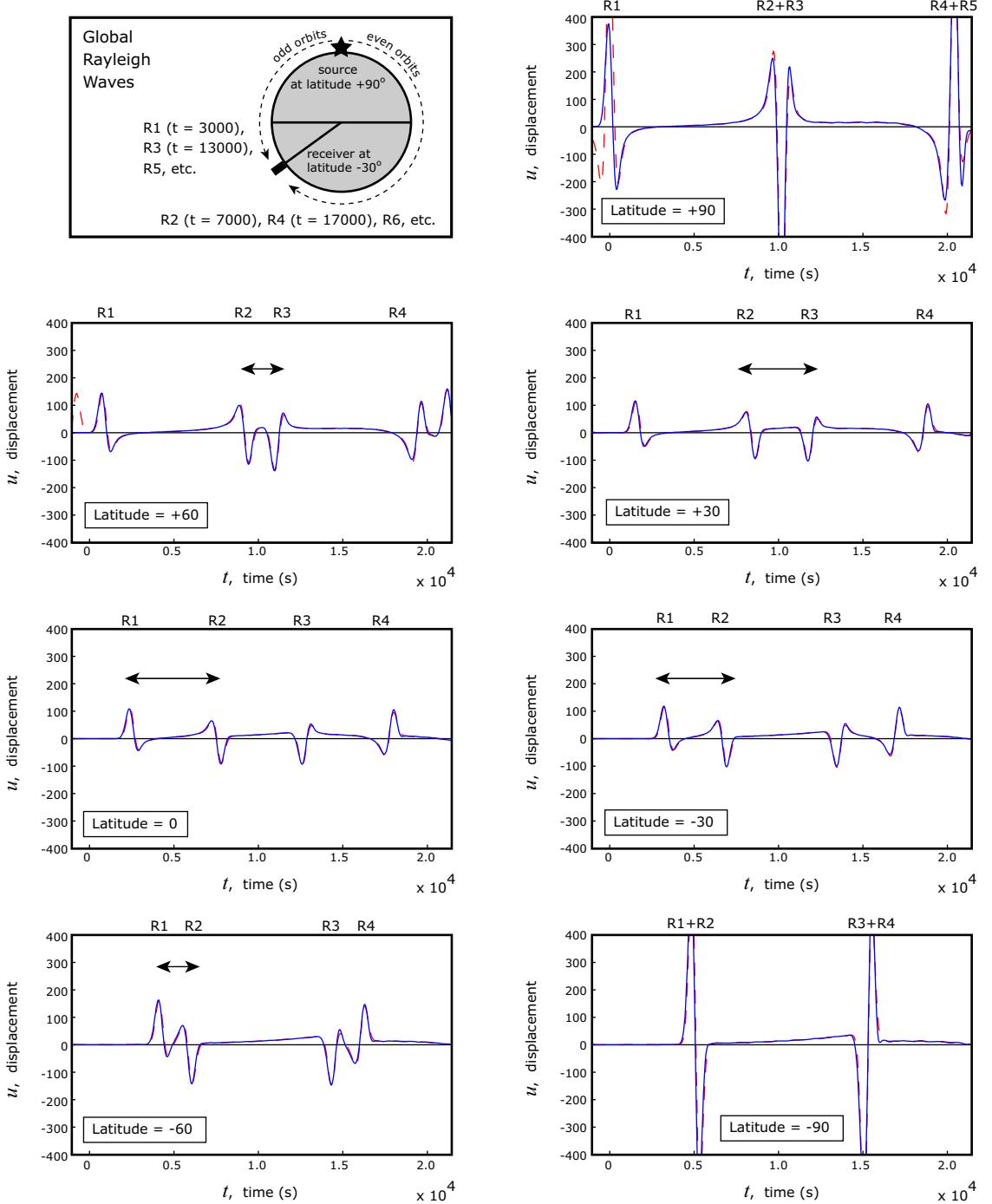


Figure 5.7: Analytical and numerical solutions for seismograms at selected latitudes for the wave propagation on the sphere using a homogeneous velocity map ($c(\theta, \phi) = 3.9280$ km/s). The dashed line is the analytical solution given by Eq. 3.34 for fixed (θ, ϕ) and varying t ; the solid line is the numerical solution for the model. The descriptive sketch showing the nomenclature is drawn for a seismogram at $\Delta = -30^\circ$ (see Figure 1.2). The double-arrow indicates the minimum time between surface wave arrivals in each seismogram. This value is a minimum for a receiver at the source ($\Delta = 0^\circ$; $\text{lat} = 90^\circ$) and antipode ($\Delta = 180^\circ$; $\text{lat} = -90^\circ$), and is a maximum for a receiver at $\Delta = 90^\circ$ ($\text{lat} = 0^\circ$). See Section 5.4.

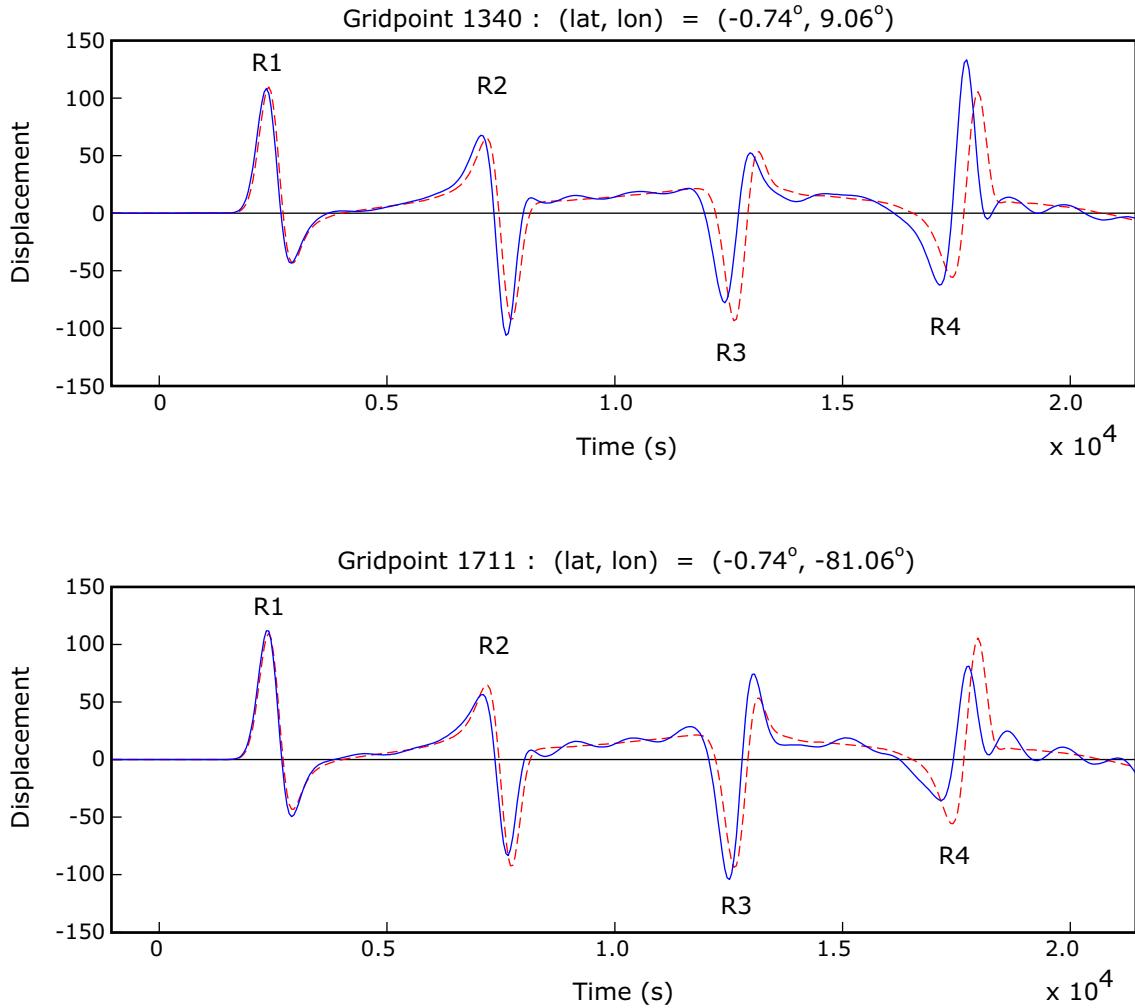


Figure 5.8: Seismograms for the wave propagation on the sphere using a heterogeneous phase velocity map ($T=40\text{s}$, $l_{\max}=12$, $\epsilon=1.0$, Trampert and Woodhouse (1995)). The dashed line is the numerical solution for the homogeneous case; the solid line is the numerical solution for the heterogeneous case. The homogeneous solution is the same in both plots, since the latitudes are the same. The agreement between the homogeneous analytical and numerical solutions can be seen in Figure 5.7 for $\text{lat}=0^\circ$. The variation here is due to the heterogeneity of the phase velocity map. The gridpoint number is the index into the order $q=4$ triangular grid (7862 gridpoints). See Section 5.4.

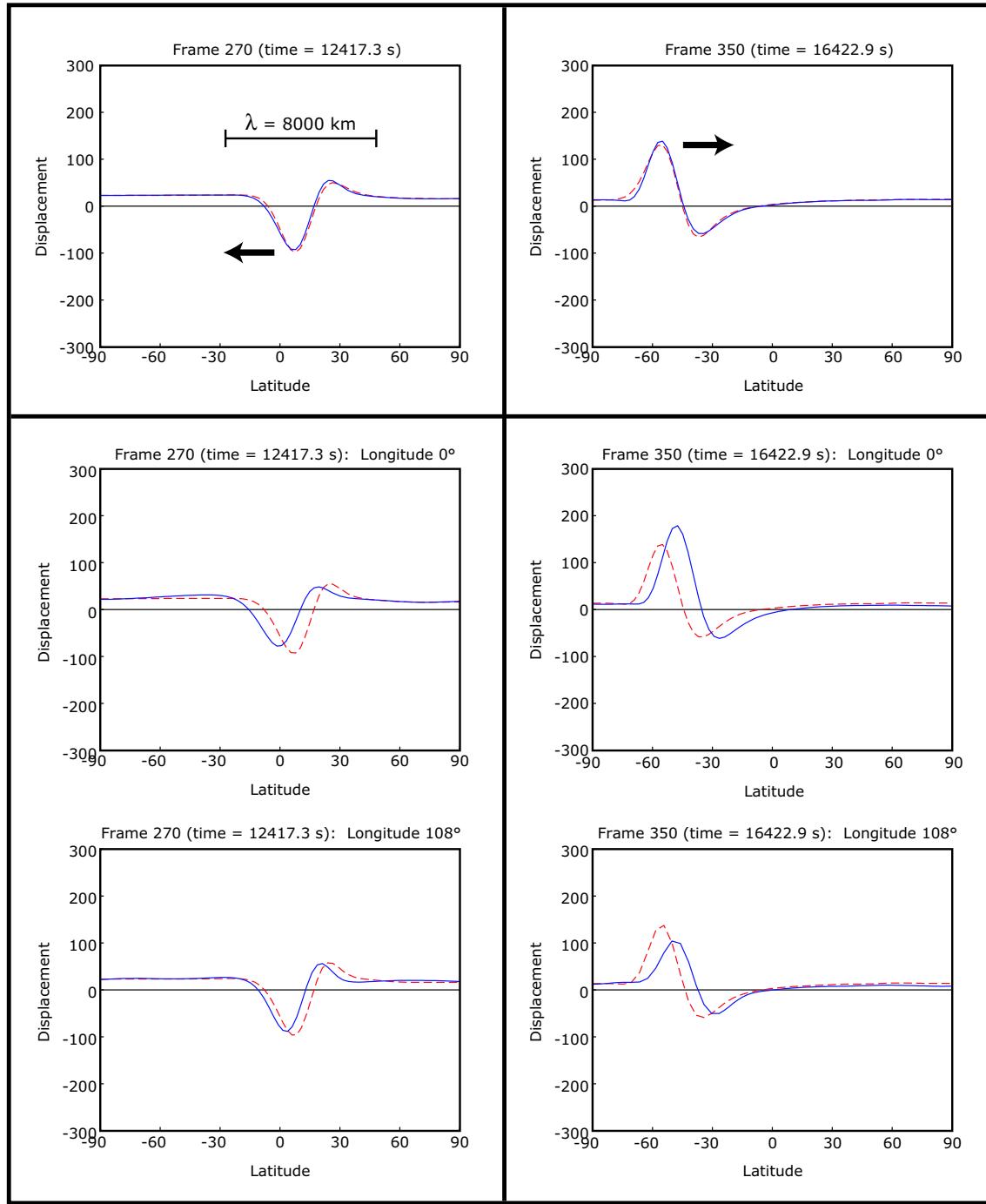


Figure 5.9: Waveforms for a heterogeneous phase velocity map ($T=40$ s, $l_{max}=12$, $\epsilon=1.0$, Trampert and Woodhouse (1995)). The top two plots show the agreement between numerical (solid) and analytical (dashed) solutions for the homogeneous case. Each of the bottom four plots compares the heterogeneous numerical solution (solid) with the homogeneous numerical solution (dashed). See Section 5.4.

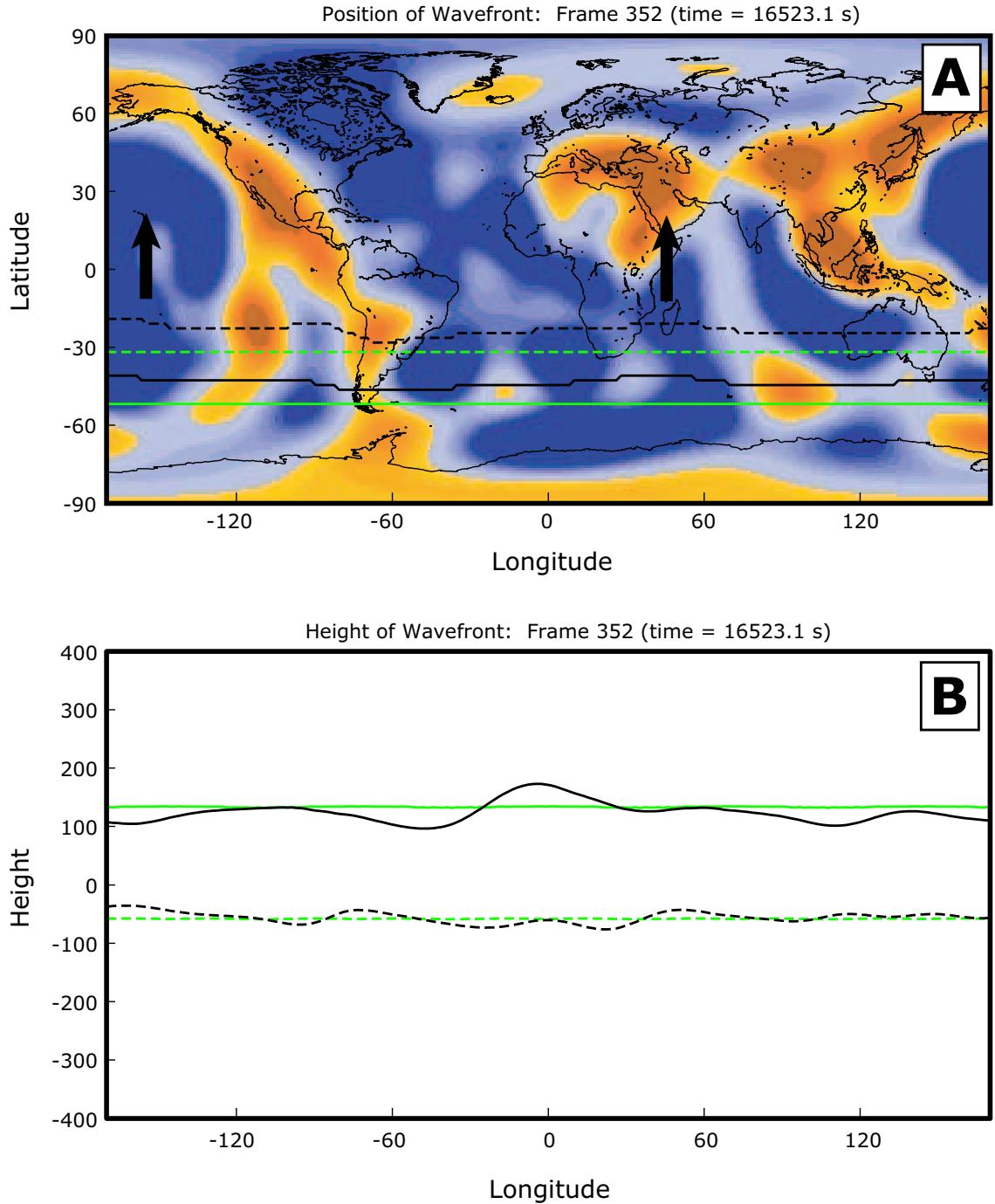


Figure 5.10: Position and amplitude of a wavefront due to a source at the North Pole (NP). The horizontal lines are wavefronts that are propagating through a homogeneous phase velocity map ($c_{PREM} = 3.9280$ km/s); the other curves are wavefronts propagating through a heterogeneous phase velocity map ($T=40$ s, $l_{max}=12$, $\epsilon=1.0$, Trampert and Woodhouse (1995)). **A.** Position of the wavefront, superimposed on the phase velocity map: the solid line is the top of the pulse, the dashed line is the bottom of the pulse. At this time, the wavefront is traveling northward, having travelled NP–SP–NP–SP→. **B.** Amplitude of the top (solid) and bottom (dashed) of the wavefront. This view is looking at a cross-section of the wavefront. See Section 5.4.

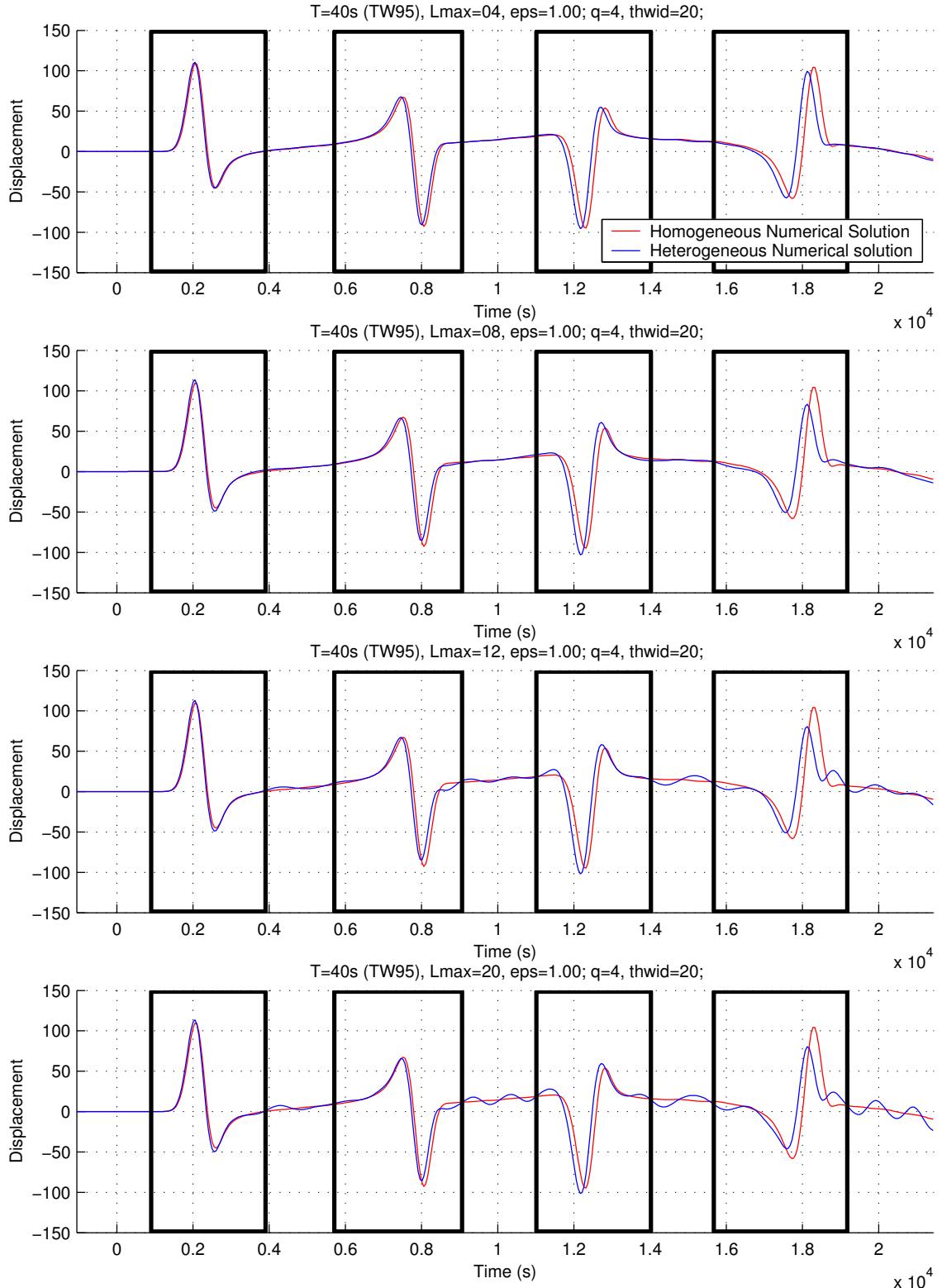


Figure 5.11: Variation in heterogeneous earth seismograms with varying l_{max} . The l_{max} -value of each phase velocity map increases from $l_{\text{max}}=4$ at the top to $l_{\text{max}}=20$ at the bottom ($T=40\text{s}$, $\epsilon=1.0$). The wave has a wavelength of $\lambda \approx 8300 \text{ km}$. Each box designates the arrival associated with each orbit, R1–R4. See Section 5.4.

Chapter 6

Ray calculations on the sphere

Contents

6.1	Introduction	79
6.2	Surface wave ray-tracing equations	81
6.3	Path anomaly	82
6.4	Phase anomaly	84
6.5	Amplitude anomaly	87
6.6	Linearized ray theory	88
6.7	Exact ray-tracing: Ray diagram cross-section curves	91
6.8	Exact ray-tracing: Variation with Δ , l_{max} , ϵ , and T	93

6.1 Introduction

In the previous section, we showed wavefronts that were produced from numerical solutions to wave propagation on the sphere using a heterogeneous phase velocity map. The results show us the position and the amplitude of the wavefront as a function of time (one time frame shown in Figure 5.10). Ray theory — whereby the rays are the orthogonal trajectories to the wavefronts — can be used to approximate the differences between heterogeneous and homogeneous solutions for (1) paths, (2) amplitude, and (3) phase. In this chapter we present the ray-tracing equations for waves on a sphere, as well as a linearized version of these equations (Section 6.6). In Section 6.8 we show how the amplitude and phase anomaly

predictions vary as we alter the path length Δ , as well as one of three parameters of the phase velocity maps $c(\theta, \phi)$: l_{max} , ϵ , and T (see Section 5.2).

Source and receivers

Orbit	Arc distance
R1	$0^\circ < \Delta \leq 180^\circ$
R2	$180^\circ \leq \Delta < 360^\circ$
R3	$360^\circ < \Delta \leq 540^\circ$
R4	$540^\circ \leq \Delta < 720^\circ$

For our purposes, we can essentially choose arbitrary sources and receivers. For simplicity, we will look at one (fictitious) source, situated at (lat, lon) = $(0^\circ, 0^\circ)$, and 25 receivers, situated approximately $\Delta = 90^\circ$ from the source (Figure 6.1). This is done in order to maximize the separation between arrivals of Rayleigh wave orbits (e.g., Figure 5.7), which is particularly important for the numerical solutions, but not for the ray-tracing. Based on the nomenclature for global surface waves (Figure 1.2), we can identify the arc-distance range for each orbit (see table above). In Table 6.2 we have listed the arc distances for R1–R4 for the 25 receivers in Figure 6.1.

Since we will be comparing the results from ray tracing with the results from the numerical model, we must select receivers that are located exactly at the gridpoints in the numerical model. The process of selecting the stations is discussed in Appendix C.

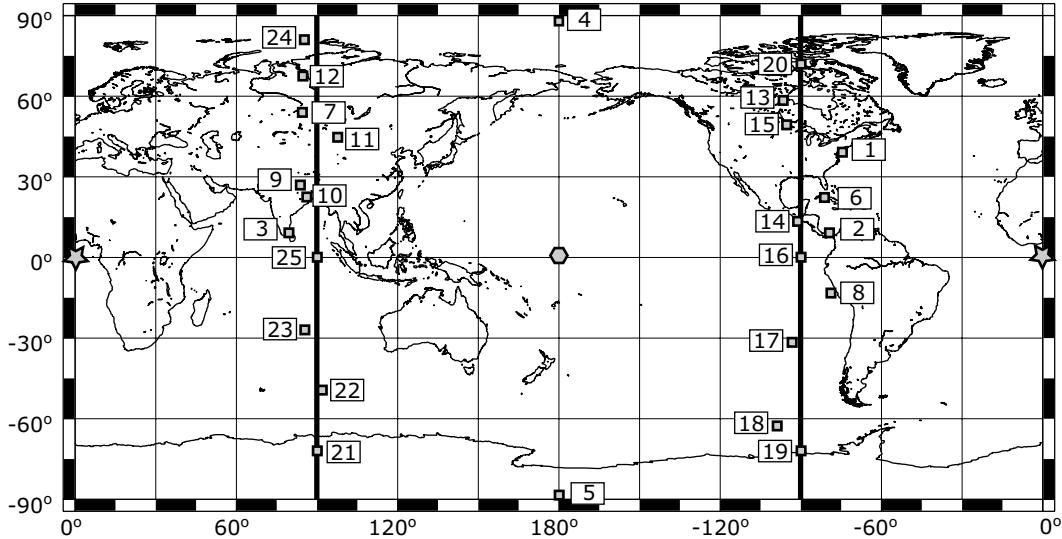


Figure 6.1: Source (star) and the 25 receivers used for the ray-tracing calculations. Receivers are clustered around the great circle that is $\Delta = 90^\circ$ from the source (bold lines). The hexagon indicates the antipode of the source. Path lengths for global Rayleigh wave orbits R1–R4 are listed in Table 6.2.

6.2 Surface wave ray-tracing equations

Ray theory was applied to surface waves by *Woodhouse* (1974), extended in *Woodhouse and Wong* (1986), and analyzed in numerous studies (e.g., *Tromp and Dahlen*, 1992a,b; *Wang et al.*, 1993; *Wang and Dahlen*, 1994, 1995b; *Laske et al.*, 1994). Below we present some results of *Woodhouse and Wong* (1986) that are pertinent to our study.

For waves on a curved surface, the surface displacement field may be approximated by:

$$u(\mathbf{x}, t) = A(\mathbf{x}, t) e^{i\psi(\mathbf{x}, t)}, \quad (6.1)$$

where $\mathbf{x} = (x^1, x^2)$ are the surface coordinates, t is time, A is the amplitude, and ψ is the phase (*Woodhouse*, 1974; *Woodhouse and Wong*, 1986). *Woodhouse and Wong* (1986) derive equations for ray tracing and amplitude calculations on a spherical earth. Their surface coordinates are given by $x^1 = \gamma \equiv \cot \theta$ and $x^2 = \phi$, where θ is colatitude and ϕ is longitude. The ray-tracing equations are written as a second-order differential equation for $\gamma = \gamma(\phi)$ (*Woodhouse and Wong*, 1986, Eq. 33):

$$\frac{d^2\gamma}{d\phi^2} + \gamma = \left(\frac{\nu^2}{1 + \gamma^2} \right) (\partial_\theta - \nu \partial_\phi) \ln c(\theta, \phi), \quad (6.2)$$

where $\nu \equiv -d\gamma/d\phi$ and $c = \omega/k$ is the phase velocity defined everywhere on the surface of the sphere. The initial conditions $\gamma(0) = \gamma_0$ and $\nu(0) = \nu_0$ must be specified; they give the starting colatitude, θ_0 (e.g., on the equator, $\gamma_0 = \cot(\pi/2) = 0$), and the take-off angle of the ray, α , via $\nu_0 = -\tan \alpha$. Note that with $\gamma = \cot \theta$, we have $1/(1 + \gamma^2) = \sin^2 \theta$.

HOMOGENEOUS CASE: On a homogeneous earth, we have $c(\theta, \phi) = c_{\text{hom}}$, and the solutions are great-circle ray paths. In this case, Equation (6.2) becomes

$$\frac{d^2\gamma}{d\phi^2} + \gamma = 0, \quad (6.3)$$

and the solutions are given by

$$\gamma(\phi) = -\nu_0 \sin \phi, \quad (6.4)$$

which plot great-circle paths on a Mercator projection. Other forms of this solution are shown

in Appendix E, Equations (E.15)–(E.19); Figure 6.2 shows some examples of great-circle ray paths.

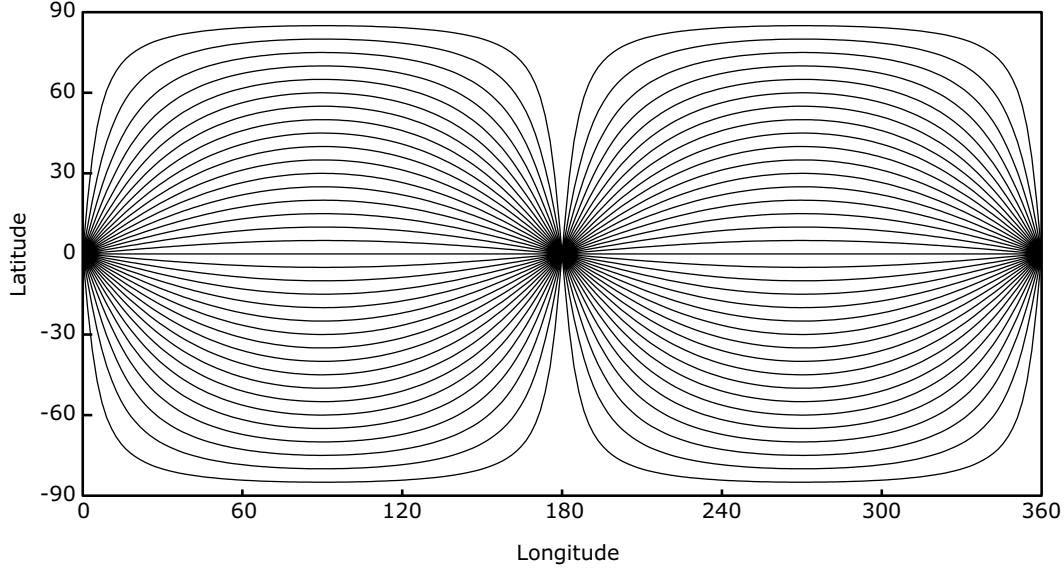


Figure 6.2: Great-circle ray paths: ray paths for the homogeneous earth. Shown here are plots of Eq. E.19 (Appendix E) for 35 paths, with take-off angles $\alpha = -85^\circ$ to $+85^\circ$.

6.3 Path anomaly

On a homogeneous earth, the ray paths are great-circle paths. Once the heterogeneity is added, the rays will deviate from great-circle paths. Figure 6.3 shows three rays for a heterogeneous phase velocity map, with a starting latitude of 0° ($\gamma_0 = 0$), with starting longitude $\phi_0 = 0^\circ$, and with finishing longitude $\phi_f = 180.0^\circ$. The homogeneous solutions are great circles, and the heterogeneous solutions are curves that deviate from the great circles. The rays are solutions to Equation (6.2), which is solved in Fortran using a fourth-order Runge–Kutta algorithm¹. Figure 6.3C shows that the deviation from the homogeneous ray paths (i.e., great-circle paths) starts at 0° , increases to a maximum, and then decreases to 0° , where the heterogeneous ray crosses the homogeneous ray, which is in the vicinity of the antipode.

It is helpful to modify the ray-tracing equations to find only those rays between a particular source and receiver. The calculations are made easier by shifting the phase velocity maps so

¹Originally I wrote the ray-tracing equations of *Woodhouse and Wong* (1986) in Matlab and used the differential equation solver `ode45`, which compares the results from a fourth-order algorithm and a fifth-order algorithm at each step. However my Matlab program proved to be relatively slow, so in the end all the ray-tracing was performed in Fortran.

that the source is on the “equator” (Appendix C). *Woodhouse and Wong* (1986) use a coordinate system in which both the source and the receiver are situated on the equator, i.e., $\gamma = 0$ at longitudes $\phi_0 = 0$ and $\phi_f = \Delta$. This imposes the boundary conditions

$$\gamma(0) = \gamma(\Delta) = 0 \quad (6.5)$$

on Equation (6.2). In Appendix E we show a generalized version of the equations, where only the source is on the equator and the receiver is at a particular longitude ϕ_f but not necessarily on the equator.

The goal is to determine how the finishing latitude at the receiver varies as a function of take-off angle, or, in this coordinate system, how $\gamma(\Delta, \nu_0)$ varies as ν_0 changes. With that information, we can use a convergence method to find the rays that hit the receiver. This gives rise to the notation (*Woodhouse and Wong*, 1986):

$$\gamma'(\phi) = \frac{\partial}{\partial \nu_0} \gamma(\phi, \nu_0) \quad (6.6)$$

$$\nu'(\phi) = \frac{\partial}{\partial \nu_0} \nu(\phi, \nu_0) = -\frac{\partial}{\partial \phi} \gamma'(\phi, \nu_0) . \quad (6.7)$$

Differentiating Equations (6.2) and (6.5) gives (*Woodhouse and Wong*, 1986, Eqs. 38–39):

$$\begin{aligned} \frac{d^2 \gamma'}{d\phi^2} + \gamma' &= \frac{2\nu}{1+\gamma^2} \left(\nu' - \frac{\nu\gamma\gamma'}{1+\gamma^2} \right) (\partial_\theta - \nu\partial_\phi) \ln c(\theta, \phi) \\ &\quad - \left(\frac{\nu^2}{1+\gamma^2} + 1 \right) \frac{\gamma'}{1+\gamma^2} (\partial_\theta^2 - \nu\partial_\theta\partial_\phi) \ln c(\theta, \phi) \\ &\quad - \left(\frac{\nu^2}{1+\gamma^2} + 1 \right) \nu' \partial_\phi \ln c(\theta, \phi) \end{aligned} \quad (6.8)$$

$$\gamma'(0) = 0 \quad (6.9)$$

$$\nu'(0) = 1 . \quad (6.10)$$

Equations (6.2) and (6.8) can be solved numerically to obtain both $\gamma(\phi)$ — the ray path — and $\gamma'(\phi)$.

Using Newton’s method (heterogeneous earth)

We use Newton’s method to determine the right take-off angles for the ray-tracing, i.e., the take-off angles of rays that converge on the receiver. If we shoot a number of rays from

a particular source (situated on the equator) and stop the rays at a particular longitude, then we can plot γ as a function of ν_0 , which can be thought of as the finishing latitude, $\bar{\theta}_f = 90^\circ - \theta_f$, as a function of take-off angle α , or simply $\bar{\theta}_f(\alpha)$. To obtain Newton's method of approximating the root ν_0 , we expand $\nu_0(\gamma_n)$ as a first-order Taylor series (see Appendix E)

$$\nu_0(\gamma_n) \approx \nu_0(\gamma_f) - \frac{\gamma_f}{\gamma'_f}, \quad (6.11)$$

where $\gamma_n = \gamma(\phi_f)$ is the finishing γ value for the *next* iteration and γ_f is the finishing γ value for the *current* iteration. Thus, using Equation (6.11) we can find the ν_0 of the ray that converges on the receiver. It is important to note that Newton's method only works in two cases: (1) for complicated curves when your initial guess of the root (ν_0) is sufficiently close to the actual root, and (2) for simple, quadratic curves, whereby your guess need not be sufficiently close to the root.

Therefore, what is needed is the take-off angle of the ray that will finish sufficiently close to the receiver, so that, using Newton's method, the next take-off angle brings the ray closer to finishing at the receiver. *Woodhouse and Wong* (1986) used a linearized ray theory prediction of initial take-off angle (Eq. 6.29), which worked well for the smooth ($lmax=8$) maps used in their study. However, we use phase velocity maps that are much more detailed ($lmax \leq 40$), and in many cases the linearized ray theory prediction of take-off angle is far from the actual take-off angle (of the exact ray path). In our study we shoot out a fan of rays in the direction of the receiver, effectively creating a segmented $\bar{\theta}_f(\alpha)$ curve, where $\bar{\theta}_f$ is the finishing latitude and α is the take-off angle. We then apply Newton's method over the $d\alpha$ -intervals where there is a change of sign. This way we do not use a linearized prediction of take-off angle, and we account for cases involving more than one path between the source and receiver (i.e., multipathing).

6.4 Phase anomaly

The phase is given by

$$\psi = \int_{\text{ray}} \frac{\omega}{c(\theta, \phi)} ds, \quad (6.12)$$

where ds is an infinitesimal arc length of the ray, $c(\theta, \phi)$ is the phase velocity along the ray, and ω is the (local instantaneous) angular frequency. Note that the phase is proportional to the frequency and inversely proportional to the phase velocity. Following *Woodhouse and Wong* (1986), we put the source and receiver on the equator ($\theta = \pi/2$), with the source at $\phi = 0$ and the receiver at $\phi = \Delta$. The phase and its derivative are defined at each $\phi = \phi_k$ along the ray as

$$\psi(\phi_k) = \int_0^{\phi_k} \left(\frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi} \right) d\phi \quad (6.13)$$

$$\frac{d\psi}{d\phi} = \frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi}. \quad (6.14)$$

HOMOGENEOUS CASE: We have $c(\theta, \phi) = c_{\text{hom}}$, which gives

$$\frac{ds}{d\phi} = a \quad (6.15)$$

$$\frac{d\psi}{d\phi} = \frac{\omega a}{c_{\text{hom}}} \quad (6.16)$$

$$\psi(\phi_k) = \int_0^{\phi_k} \left(\frac{\omega a}{c_{\text{hom}}} \right) d\phi = \left(\frac{\omega a}{c_{\text{hom}}} \right) \phi_k, \quad (6.17)$$

where a is the radius of the earth.

HETEROGENEOUS CASE: We specify a phase velocity map $\overline{c}(\theta, \phi)$ or $c(\theta, \phi)$:

$$\overline{c} = \frac{c(\theta, \phi) - c_{\text{hom}}}{c_{\text{hom}}} \iff c(\theta, \phi) = (1 + \overline{c}) c_{\text{hom}}. \quad (6.18)$$

In this case, we have

$$\frac{ds}{d\phi} = a \left[\frac{\nu(\phi)^2}{[1 + \gamma^2(\phi)]^2} + \frac{1}{1 + \gamma^2(\phi)} \right]^{1/2} \quad (6.19)$$

$$\frac{d\psi}{d\phi} = \frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi} \quad (6.20)$$

$$\psi(\phi_k) = \int_0^{\phi_k} \left(\frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi} \right) d\phi. \quad (6.21)$$

Equation (6.19) is derived in Appendix E, Equation (E.34).

The phase anomaly P along the ray is defined with respect to a homogeneous earth model:

$$P(\phi_k) \equiv \delta\psi(\phi_k) \equiv \psi_{\text{het}}(\phi_k) - \psi_{\text{hom}}(\phi_k). \quad (6.22)$$

Substituting Equations (6.21) and (6.17), we obtain

$$\begin{aligned}
 P(\phi_k) &\equiv \psi_{\text{het}}(\phi_k) - \psi_{\text{hom}}(\phi_k) \\
 &= \int_0^{\phi_k} \left(\frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi} \right) d\phi - \int_0^{\phi_k} \left(\frac{\omega a}{c_{\text{hom}}} \right) d\phi \\
 &= \int_0^{\phi_k} \left[\left(\frac{\omega}{c(\theta, \phi)} \frac{ds}{d\phi} \right) - \frac{\omega a}{c_{\text{hom}}} \right] d\phi \\
 &= \frac{\omega}{c_{\text{hom}}} \int_0^{\phi_k} \left[\left(\frac{c_{\text{hom}}}{c(\theta, \phi)} \frac{ds}{d\phi} \right) - a \right] d\phi \\
 &= \frac{\omega a}{c_{\text{hom}}} \int_0^{\phi_k} \left[\left(\frac{c_{\text{hom}}}{c(\theta, \phi)} \left[\frac{\nu(\phi)^2}{[1 + \gamma^2(\phi)]^2} + \frac{1}{1 + \gamma^2(\phi)} \right]^{1/2} \right) - 1 \right] d\phi, \quad (6.23)
 \end{aligned}$$

where $a\phi_k$ is the arc distance of the HOM-ray, and $ds/d\phi$ is given by Equation (6.19). Equation (6.23) is presented as Equation 42 in *Woodhouse and Wong* (1986).

In essence it is immaterial for our purpose what we choose for c_{hom} and ω , as long as they are the same for both the ray-tracing and the wave propagation on the hexagonal grids. In our study, the homogeneous surface wave phase velocity is calculated using the PREM model ($c_{\text{hom}} = c_{\text{PREM}}(T)$; see Appendix F)². We let $\omega = 1$, which means that the phase anomaly is simply a travel-time difference and has units in seconds. Values for $c_{\text{hom}} = c_0$ are listed in Table 5.1.

²A different choice for the reference model could be the mean value of the particular phase velocity map ($c_{\text{hom}} = c_{\text{mean}}(\omega)$).

Note that with this convention, a positive phase anomaly at the receiver arises when the ray has traveled through a slower region of phase velocity (late arrival); a negative phase anomaly arises when the ray has traveled through a faster region (early arrival) (Table 6.1).

It is important to keep in mind that the phase is essentially averaging the phase velocity map along the ray path. The phase at a particular point depends on the path the ray has taken from the source. Phase velocity variation (in other words, “structure”) in the vicinity of the receiver does not have as great an effect on the phase anomaly as it does on amplitude anomaly (*Um and Dahlen, 1992*).

Table 6.1: Phase anomaly convention. $P \equiv \psi_{\text{het}} - \psi_{\text{hom}}$, and $\psi \propto 1/c$.

Ray travels through a...	phase velocity	phase	phase anomaly	relative arrival
faster (blue) region	$c_{\text{het}} > c_{\text{hom}}$	$\psi_{\text{het}} < \psi_{\text{hom}}$	$P < 0$	early
homogeneous earth	$c_{\text{het}} = c_{\text{hom}}$	$\psi_{\text{het}} = \psi_{\text{hom}}$	$P = 0$	on time
slower (red) region	$c_{\text{het}} < c_{\text{hom}}$	$\psi_{\text{het}} > \psi_{\text{hom}}$	$P > 0$	late

6.5 Amplitude anomaly

The amplitude anomaly A is defined as the square root of the ratio of geometrical spreading in the homogeneous case to the geometrical spreading in the heterogeneous case:

$$A = \left| \frac{\frac{d\sigma}{d\alpha}|_{\text{hom}}}{\frac{d\sigma}{d\alpha}|_{\text{het}}} \right|^{1/2}, \quad (6.24)$$

where $d\alpha$ is an infinitesimal take-off angle width at the source, and $d\sigma$ is the ray-tube width at some distance Δ from the source (Appendix E, Figure E.1B). In other words, when the rays passing through a heterogeneous earth are converging relative to the rays in the homogeneous earth, the amplitude anomaly increases; when the rays are diverging, the amplitude anomaly decreases. In Appendix E (Eq. E.28) we derive the following expression for the amplitude anomaly:

$$A = \left| \frac{\frac{d\sigma}{d\alpha}|_{\text{hom}}}{\frac{d\sigma}{d\alpha}|_{\text{het}}} \right|^{1/2} = \frac{(1 - \sin^2 \theta_f \cos^2 \phi_f)^{1/4} (1 + \nu_f^2 \sin^2 \theta_f)^{1/4}}{|\sin \theta_f| (\gamma'_f)^{1/2} (1 + \nu_0^2)^{1/2}}, \quad (6.25)$$

which is a more general case than Equation 41 of *Woodhouse and Wong* (1986) (Eq. E.29). Equation (6.25) is a ray-theory prediction for the amplitude anomaly due the heterogeneity of the phase velocity maps.

6.6 Linearized ray theory

Linearized ray theory is an approximation for the (exact) ray-tracing equations in the case of a slight heterogeneity $\delta c = c(\theta, \phi) - c_0$, where $c_0 = c_{\text{hom}}$. The calculations for the anomalies are made *not* along the exact ray path but along the great circle connecting the source and receiver (i.e., $\theta = \pi/2$). Hence its alternative name, Path Integral Approximation.

The linearized ray theory offers a physical interpretation of the phase velocity maps, in terms of phase, path, and amplitude. In particular, the linearized ray theory states that the phase anomaly will be proportional to the integral of slowness, $1/c(\theta, \phi)$, the path will be proportional to $\partial_\theta c(\theta, \phi)$, and the amplitude anomaly will be proportional to $\partial_{\theta\theta} c(\theta, \phi)$ (see Eqs. 6.28, 6.29, 6.30, and 6.31 below). Thus, the maps of the derivatives of phase velocity in Figure 6.4 are relevant to the calculation of arrival-angle and amplitude anomalies for the case of a source and a receiver situated on the equator. The basic idea is that the ray path is affected only by gradients perpendicular to its propagation direction: the $\partial_\theta c$ and $\partial_{\theta\theta} c$ derivatives “push and pull” the ray off the great-circle path.

Phase anomaly

In the case of slight heterogeneity we can use a first-order approximation:

$$\frac{1}{c(\theta, \phi)} = \frac{1}{c_0 + \delta c} = \frac{1}{c_0} \left(\frac{1}{1 + \frac{\delta c}{c_0}} \right) \approx \frac{1}{c_0} \left(1 - \frac{\delta c}{c_0} \right). \quad (6.26)$$

In terms of δc the phase can thus be written as

$$\begin{aligned} \psi &= \int_{\text{path}} \left(\frac{\omega}{c(\theta, \phi)} \right) ds = \int_{\text{path}} \left(\frac{\omega}{c_0 + \delta c} \right) ds \\ &\approx \frac{\omega}{c_0} \int_{\text{path}} \left(1 - \frac{\delta c}{c_0} \right) ds \\ &= \frac{\omega a}{c_0} \int_0^\Delta \left(1 - \frac{\delta c}{c_0} \right) d\phi, \end{aligned} \quad (6.27)$$

where we have substituted $ds = a d\phi$.

Fermat's principle states that the travel time (or phase) between the source and receiver is stationary, and thus can be approximated by replacing the integral along the exact path by the integral along the unperturbed or great-circle path. Thus, we substitute Equations (6.27) and (6.17) into Equation (6.22) in order to get the approximation for the phase anomaly:

$$\begin{aligned} P_{\text{linear}} &\equiv \psi_{\text{het}} - \psi_{\text{hom}} \\ &= \frac{\omega a}{c_0} \int_0^\Delta \left(1 - \frac{\delta c}{c_0} \right) d\phi - \int_0^\Delta \left(\frac{\omega a}{c_0} \right) d\phi \\ &= -\frac{\omega a}{c_0} \int_0^\Delta \left(\frac{\delta c(\theta, \phi)}{c_0} \right) d\phi. \end{aligned} \quad (6.28)$$

This is Equation 43 of *Woodhouse and Wong* (1986). Keep in mind that in this equation the integration path is the great circle between source and receiver, whereas in Equation (6.23) it is over the actual ray path.

Path anomaly

The path anomaly refers to the deviation of rays from great-circle paths. The quantity ν gives the slope of the ray along the ray path and may be calculated using linearized ray theory, in order to estimate the ray path between the source and receiver. At the source and receiver, the linearized ray theory provides estimates for the initial take-off angle and the off-azimuth

arrival angle. These equations are given by

$$\nu_0 = -\frac{d\gamma}{d\phi}\Big|_{\phi=0} \cong \int_0^\Delta \csc \Delta \sin(\Delta - \phi) c_0^{-1} \frac{\partial}{\partial \theta} [\delta c\left(\frac{\pi}{2}, \phi\right)] d\phi \quad (6.29)$$

$$\nu_\Delta = -\frac{d\gamma}{d\phi}\Big|_{\phi=\Delta} \cong -\int_0^\Delta \csc \Delta \sin \phi c_0^{-1} \frac{\partial}{\partial \theta} [\delta c\left(\frac{\pi}{2}, \phi\right)] d\phi. \quad (6.30)$$

We should make an important distinction about the linearized ray theory prediction for the path. The calculation of the path is performed along the great-circle (“equatorial”) path, as indicated by the term $\delta c(\pi/2, \phi)$ in the two equations above. So even though the linearized ray theory prediction of the ray path is *not* a great-circle path, the *calculation* of it is based on information (i.e., the transverse derivative of the phase velocity) that lies exactly on the great-circle path. It is not accurate to say that linearized ray theory “predicts a great-circle path”, but it is true that the great-circle path is the only region of the phase velocity map that is incorporated into the linearized calculations³.

Laske has done extensive work in measuring and predicting off-azimuth arrival angles, and she has used these data in refining maps of surface wave phase velocity (*Laske et al.*, 1994; *Laske*, 1995; *Laske and Masters*, 1996). In Figure 7 of *Laske* (1995), she compares the off-azimuth arrival angle between exact ray tracing and the linearized ray tracing, noting that “the linear approximation is accurate enough to model the observed angles reliably” (p. 253). However, it is clear from the figure that the correlation decreases significantly for higher orbits (up to R4). Furthermore, the phase velocity map used (*Wong*, 1989) is expanded up to $l_{max}=12$; thus she was probably not concerned with multipathing (see Table 6.3). An additional finding of these studies is that Laske has identified several stations whose azimuths have been misaligned by as much as 6° , and this misalignment has been confirmed and corrected in a couple cases.

We do not use the ray theory predictions (linear or exact) of take-off angle or arrival angle in our comparison with the numerical results, mainly because this would be difficult to measure with the case of the numerical model. We instead focus on ray theory predictions of phase anomalies and amplitude anomalies.

³Of course, for a low- l_{max} map with broad structures, the transverse derivatives $\partial_\theta c$ and $\partial_{\theta\theta} c$ evaluated on the great-circle path will provide accurate information over a broader region than they would for a high- l_{max} map with short-scalelength structures.

Amplitude anomaly

The linearized amplitude anomaly is given by (*Woodhouse and Wong*, 1986, Eq. 52):

$$\ln A_{\text{linear}} = \frac{1}{2} \csc \Delta \int_0^\Delta \sin(\Delta - \phi) c_0^{-1} [\sin \phi \partial_\theta^2 - \cos \phi \partial_\phi] \delta c \left(\frac{\pi}{2}, \phi \right) d\phi. \quad (6.31)$$

The derivations and figures in *Woodhouse and Wong* (1986) show the behavior of the linearized amplitude anomaly as a function of orbit. The anomalies of the even orbits form a line whose slope is *opposite* to that of the line formed by the anomalies of the odd orbits.

6.7 Exact ray-tracing: Ray diagram cross-section curves

This section examines the variation in path, phase, and amplitude anomalies based on calculations from exact ray-tracing. Figure 6.5A-B shows a surface wave ray diagram superimposed on a $T=40\text{s}$, $l_{\text{max}}=20$ ($\epsilon=1.0$) phase velocity map. The figure shows both the homogeneous rays (i.e., rays traced through a homogeneous earth) and the heterogeneous rays, which deviate from the great-circle paths. In the remainder of this section we will be looking at *cross-section curves* of ray diagrams, as shown in Figure 6.5C-D. By cross-section, we mean this: take a source and receiver that are rotated to the equator with the source at $\phi_0 = 0^\circ$. Shoot a batch of rays over a range of take-off angles in the direction of the receiver, which is situated at ϕ_f , where ϕ_f will depend on the orbit (Table 6.2). Then stop the rays at any $|\phi| \leq |\phi_f|$ (typically we stop at ϕ_f) and plot the finishing latitude as a function of take-off angle at the longitudinal meridian ϕ . We denote the cross-section curves as $\bar{\theta}_f(\alpha)$, where α is the take-off angle and $\bar{\theta}_f$ is the finishing latitude. Similarly, we have cross-section curves for phase anomaly and amplitude anomaly given by $P_f(\alpha)$ and $A_f(\alpha)$, respectively.

The first thing that we notice in Figure 6.5 is that there is substantial deviation of the heterogeneous rays from the great-circle paths — even more than 25° with some of the finishing rays in Figure 6.5D. The reference cross-section curve is the (red) sloping straight line formed by the finishing homogeneous rays. Secondly, we see that *multipathing* occurs, that is, multiple ray paths existing between source and receiver (again, assuming the receiver is on the equator). The number of ray paths finishing at a particular latitude can be found by drawing a horizontal line ($\bar{\theta}_f = \text{const}$) through the cross-section curve and then counting the number of intersections; we denote these points with the green vertical lines for the

“equatorial” finishing points with $\bar{\theta}_f = 0^\circ$.

Each ray path carries information about the amplitude anomaly and phase anomaly. Thus we can plot amplitude and phase anomaly values along the cross-section curves, as in Figure 6.5E-H. Here we note a few observations based on this example. First, notice that the amplitude anomalies (A) approach infinity where the slope of the cross-section curve is zero, and that $A \rightarrow 0$ in areas where the slope is steep. This is consistent with our definition of amplitude anomaly (Eq. 6.24). Steep slopes in the cross-section curves correspond to regions of large geometrical spreading, and thus low amplitude anomalies; shallow slopes correspond to regions of small geometrical spreading, and thus high amplitude anomalies⁴. The regions of infinite-amplitude anomaly are *caustics*.

Secondly, we note that at none of the “equatorial” multipathing points do we observe near-infinite amplitudes. In other words, the multipathing rays in Figure 6.5F (vertical lines intersecting amplitude anomalies) do not fall in the regions of near-infinite A . This observation highlights the subtlety of the definition of caustics. As stated in the first sentence of *Wang et al.* (1993), “The geometrical theory of surface wave propagation on a laterally heterogeneous earth model diverges at caustics, where neighboring rays cross.” In general, multipathing presents crossing ray paths, but not crossing *neighboring* paths. In other words, a caustic is a case of multipathing, but the reverse is not necessarily true.

Thirdly, it appears that the phase anomalies (P) exhibit less variation than do the amplitude anomalies. Furthermore, there appears to be a correlation between the phase and the finishing latitude, which is particularly noticeable when the rays have undergone a relatively large amount of focusing and defocusing (high- l_{max} maps combined with long path lengths) (Figures 6.5D,F). The ray diagram suggests why this might be occurring. Notice that at $\phi_f = -640.5^\circ$ many of the rays are tending to the high-velocity anomaly (blue) beneath the eastern Indian Ocean. To be precise, we should say that the phase anomaly correlates to the *phase velocity* $c(\theta, \phi)$, which we showed in Equation (6.28). It appears that after significant focusing, the rays have been channeled into a high velocity zone, and thus the finishing latitudes are correlated to the finishing phase velocities and thus to the finishing phase anomalies P . This relationship between finishing latitude and phase was only recently

⁴Based on the definition of A (Eq. 6.24), we should get exact agreement between a scatter plot of $A(\bar{\theta}_f(\alpha))$ versus $1/\frac{d}{d\alpha}[\bar{\theta}_f(\alpha)]$. This shows that the amplitude anomaly can be calculated directly from the cross-section curve.

observed and warrants further study⁵.

To emphasize the importance and variation of the ray diagram cross-sections, we present a figure of cross-section curves for the first 15 receivers (Figure 6.6). The variability in these diagrams shows why it is necessary to calculate a range of take-off angles if one hopes to find the ray paths between the source and receiver. Also, note that there are, perhaps unexpectedly, very sharp changes in the cross-section curves, which, upon closer examination, appear as smooth curves (Figure 6.9). In Figure 6.7 we plot the difference between the finishing latitudes of the heterogeneous and homogenous rays; this difference is a measure of the deviation of the exact ray path from the great-circle path, which is the integration path for linearized ray theory. We see that at the finishing (R4) longitude ϕ_f , there are heterogeneous rays that deviate from the great-circle paths by more than 30°.

Figures 6.5–6.7 show the exact ray-tracing results for one particular phase velocity map. Now we present results for 15 different phase velocity maps, each one an alteration of five maps of *Trampert and Woodhouse* (1995). The results, which we will discuss later, are in Figures 6.8–6.11.

6.8 Exact ray-tracing: Variation with Δ , $lmax$, ϵ , and T

Now we will examine how the exact ray-tracing results vary in terms of four parameters: the path-distance⁶ Δ , the truncation level of the phase velocity map $lmax$, the power of the phase velocity map ϵ , and the period of the phase velocity map T . These are the parameters discussed and presented in figures in Section 5.2. It is particularly helpful to refer to the amplitude spectra of the phase velocity maps (Figures 5.5 and 5.6C) when analyzing these results.

⁵Intuitively, one might expect the paths which are deviating the most from the great circle to have travelled a longer path (in distance) and therefore arrive at a later time. However, the plots of deviation versus phase anomaly show no correlation, while plots of finishing latitude versus phase anomaly show a moderate correlation, suggesting that the influencing factor may be a quality of the map itself.

⁶ Δ is the ray-path distance between the source and the receiver, whether it be a great-circle path, a linearized ray theory path, or the actual ray path. We discuss the variation in path length in terms of ϕ_f , which is the finishing longitude of the rays. For the “equatorial” great-circle path, $\phi_f = \Delta$.

Variation with Δ

Figure 6.5 shows the effect on the path, amplitude, and phase anomalies of increasing the path-length Δ . Deviation from the great-circle paths (i.e., Figure 6.7) increases with path-length Δ . Intuitively, longer paths sample a greater amount of heterogeneity, and therefore the rays have more “room” for focusing and defocusing. For R4 orbits in our $l_{max}=8$ model, we find maximum great-circle path deviations of approximately 15° . For path lengths approximately twice as long, *Woodhouse and Wong* (1986) report deviations of nearly 25° in their $l_{max}=8$ model.

One way of discussing this is in terms of the variation with orbit, which was the motivation for *Woodhouse and Wong* (1986). Keep in mind, however, that odd and even orbits represent fundamentally different sets of rays, since they take off in opposite directions. *Woodhouse and Wong* (1986) provide a good summary of some of these results:

The character of amplitude and phase variations in mantle waves has been shown to be diagnostic of substantial lateral refraction. Sometimes this is so large that the great circle path can appear to be slow for orbits of one sense and fast for those of the opposite sense. There is a tendency for even and odd orbit paths to polarize and thus to sample somewhat different areas of the globe. Amplitude anomalies are predicted to have systematic and opposite trends for even and odd orbits and this is very commonly observed in the data. The slope of these trends is sensitive to the great circle integral of the second transverse derivative of phase velocity, with a range dependent kernel which also depends upon epicentral distance. (p. 772)

Although these observations were based on a comparison of actual data with ray theory predictions (exact and linear) using relatively broad phase velocity maps (M84C of *Woodhouse and Dziewonski* (1984), expanded up to $l_{max}=8$), it is possible that the general ideas will hold for maps with a greater amount of shorter-scalelength structure. We did not make a systematic study of the regions sampled by even and odd orbits, but we did analyze the effect of increasing the orbit on the ray theory predictions (see Chapter 7).

Variation with l_{max}

The left column of Figures 6.8–6.11 shows the variation in the path, amplitude, and phase with l_{max} . The value l_{max} controls the minimum-scalelength structure in the phase velocity maps. As an infinite-frequency approximation, ray theory will “see” all the structure in the

maps. Thus if we decrease l_{max} for a given map, the deviations of the path/amplitude/phase anomalies should decrease. This is exactly what we observe in the figures, for $l_{max} = 40, 20, 12, 8, 4$.

Variation with ϵ

The motivation for including ϵ as a testing parameter is that it allows us to investigate the convergence of the results of exact ray theory to those of the linearized ray theory, as $\epsilon \rightarrow 0$. The middle column of Figures 6.8–6.11 shows the variation in the path, amplitude, and phase with ϵ . The value ϵ controls the power of the phase velocity map. Thus if we decrease ϵ for a given map, the deviations of the path, amplitude, and phase anomalies should decrease. This is exactly what we observe in the figures, for $\epsilon = 1.0, 0.8, 0.6, 0.4, 0.2$.

Variation with T

Here we show some results for phase velocity maps corresponding to different periods. The right column of Figures 6.8–6.11 shows the variation in the path, amplitude, and phase with T . As discussed in Section 5.2, each phase velocity map has different power in each part of the amplitude spectrum (Figure 5.5A-B). The longer-period maps contain less shorter-scalelength structure, and thus the deviation from linearized theory is less. We observe that the variation in path/amplitude/phase anomalies decreases with increasing $T = 40, 60, 80, 150, 200$ s.

TW95 vs TW96

We present a comparison between the maps of *Trampert and Woodhouse* (1995) and *Trampert and Woodhouse* (1996) in Figure 5.6. Both of the maps are expanded up to $l_{max}=40$; however, the maps of TW96 contain more shorter-scalelength structure (i.e., greater power) for $l > 12$, as shown in Figure 5.6C. As expected, this gives rise to exact ray paths deviating considerably more from the great-circle paths for the maps of TW96, as shown in Figure 6.12. The correlation between the finishing latitudes of TW95 and TW96 diminishes for $l_{max} > 12$. Also note that the cross-section curves in Figure 6.12 are for the R2 orbits, not the R4 orbits, as shown earlier (Figures 6.6 and 6.8).

Multipathing

We are interested in documenting the influence of the aforementioned parameters on multipathing. Tables 6.3, 6.4, and 6.5 show some regimes under which multipathing occurs. The trends in Table 6.3 are consistent with what we might expect: multipathing is more common for longer path-lengths (i.e., $R4$) and for maps with more shorter-scalelength structure, that is, maps with greater $lmax$. It is clear from this table that the current resolution of our phase velocity maps, $lmax \geq 20$ poses a significant barrier to using ray theory for orbits beyond $R1$. Table 6.4 also shows what we would expect as well: as $\epsilon \rightarrow 0$, multipathing ceases to occur.

The occurrence of multipathing raises an important question. If there is more than one ray between the source and receiver, and if each ray carries separate information about the phase and amplitude anomaly, what can we do? This question will be discussed in the next chapter, when we introduce the numerical results for comparison with the results from ray theory.

Table 6.2: Receivers used in the ray-tracing calculations and the numerical model (Section 6.1, Figure 6.1). The arc distance and the orbit distances depend on the source, which we have set at $(0, 0)$. The receiver–11 boxed value refers to the scenario for the rays plotted in Figures 6.8–6.11; the receiver–5 boxed value refers to the scenario for rays plotted in Figure 6.12.

Receiver			Arc distance Δ	Path length ϕ_f			
#	lat	lon		R1	R2	R3	R4
1	39.13	-74.60	78.11	78.11	-281.89	438.11	-641.89
2	9.11	-79.45	79.58	79.58	-280.42	439.58	-640.42
3	9.11	79.45	79.58	79.58	-280.42	439.58	-640.42
4	88.53	0.00	88.53	88.53	-271.47	448.53	-631.47
5	-88.53	180.00	91.47	91.47	(-268.53)	451.47	-628.53
6	22.33	-81.16	81.83	81.83	-278.17	441.83	-638.17
7	53.88	84.56	86.80	86.80	-273.20	446.80	-633.20
8	-13.20	-78.77	79.07	79.07	-280.93	439.07	-640.93
9	26.95	83.86	84.53	84.53	-275.47	444.53	-635.47
10	22.49	86.23	86.52	86.52	-273.48	446.52	-633.48
11	44.66	97.71	95.47	95.47	-264.53	455.47	(-624.53)
12	67.44	84.78	88.00	88.00	-272.00	448.00	-632.00
13	58.36	-96.65	93.48	93.48	-266.52	453.48	-626.52
14	13.48	-91.30	90.00	90.00	-270.00	450.00	-630.00
15	49.35	-95.35	93.48	93.48	-266.52	453.48	-626.52
16	0.00	-90.00	90.00	90.00	-270.00	450.00	-630.00
17	-31.51	-93.22	92.74	92.74	-267.26	454.74	-627.26
18	-62.77	-98.74	93.99	93.99	-266.01	450.00	-630.00
19	-72.00	-90.00	90.00	90.00	-270.00	450.00	-630.00
20	72.00	-90.00	90.00	90.00	-270.00	450.00	-630.00
21	-72.00	90.00	90.00	90.00	-270.00	450.00	-630.00
22	-49.47	91.94	91.26	91.26	-268.74	451.26	-628.74
23	-26.97	85.53	86.01	86.01	-273.99	446.01	-633.99
24	80.91	85.29	89.26	89.26	-270.74	449.26	-630.74
25	0.00	90.00	90.00	90.00	-270.00	450.00	-630.00

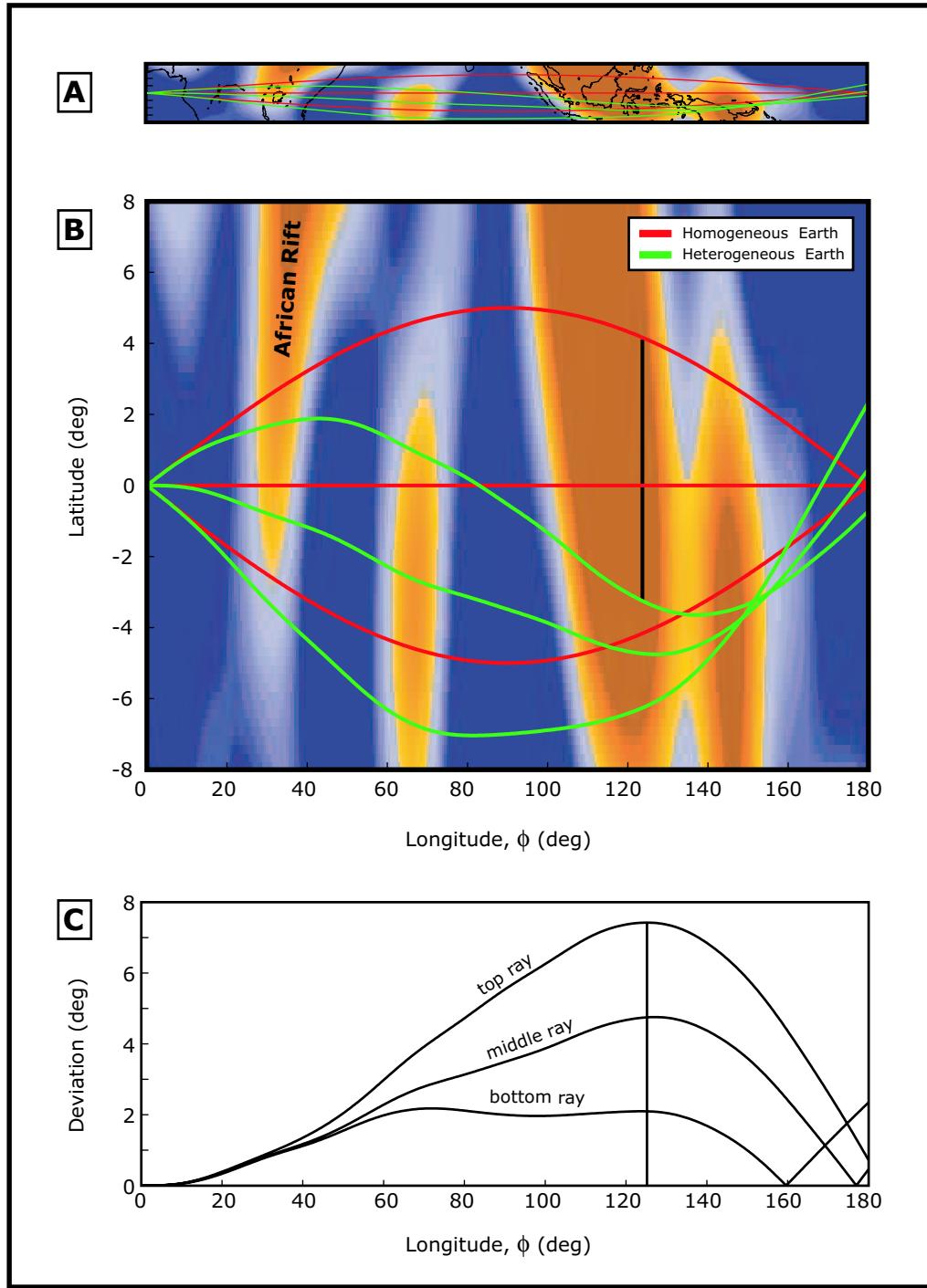


Figure 6.3: Introductory figure for ray-tracing. Three surface wave rays superimposed on the $T=40\text{s}$, $l_{\max}=20$ ($\epsilon=1.0$) phase velocity map (see Figure 5.3A). Here the source is at $(\text{lat}, \text{lon})=(0^\circ, 0^\circ)$, and we have shot three rays eastward with take-off angles of $\alpha = -5^\circ, 0^\circ, +5^\circ$. The homogeneous rays (red) are great-circle paths, and they all converge at the antipode $(\text{lat}, \text{lon})=(0^\circ, 180^\circ)$. The heterogeneous rays (green) deviate from the great-circle paths. **A.** Ray paths plotted to scale, superimposed on the outline of the land masses (Africa, Indonesia). **B.** Ray paths plotted on a vertically exaggerated scale. The black line indicates the maximum latitudinal deviation of the heterogeneous ray path from the homogeneous ray path. This occurs for the top ray at $\text{lon}=125.20^\circ$ and has a value of 7.42° . **C.** Magnitude of the deviation of the heterogeneous rays from the homogeneous rays. See Section 6.3.

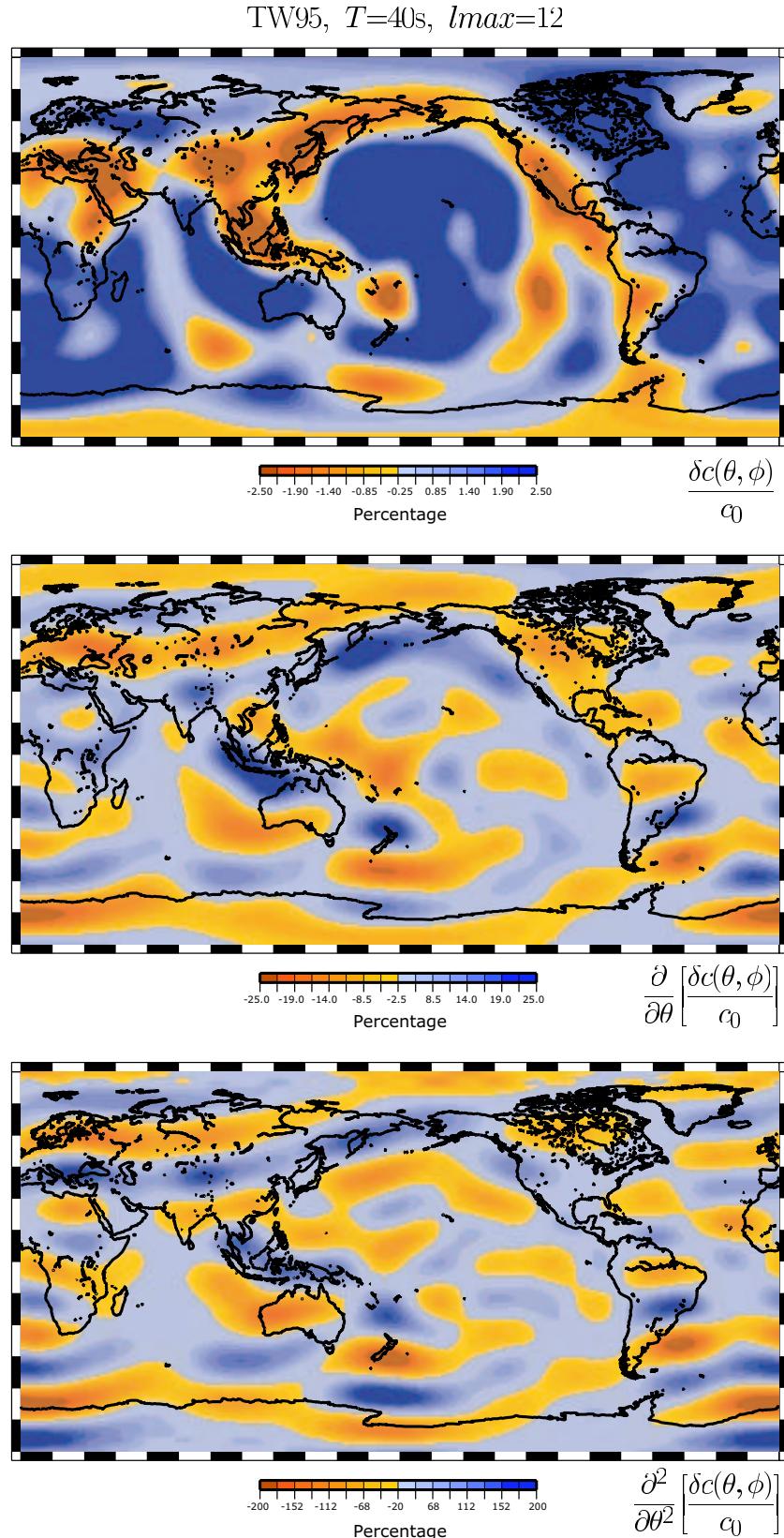
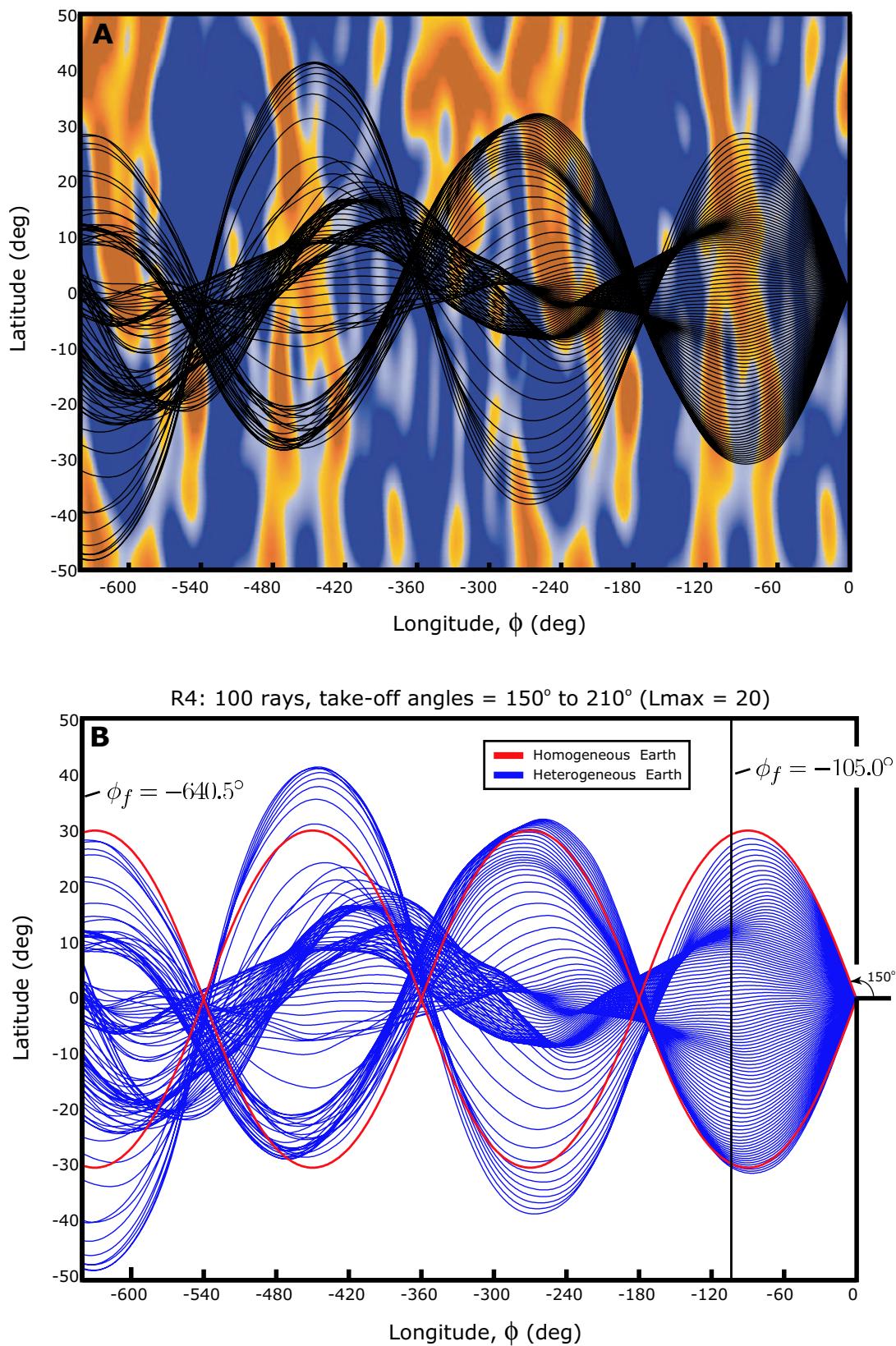


Figure 6.4: A phase velocity map δc , and its first and second derivative with respect to the colatitude θ . Note the different color scales. See Section 6.6.



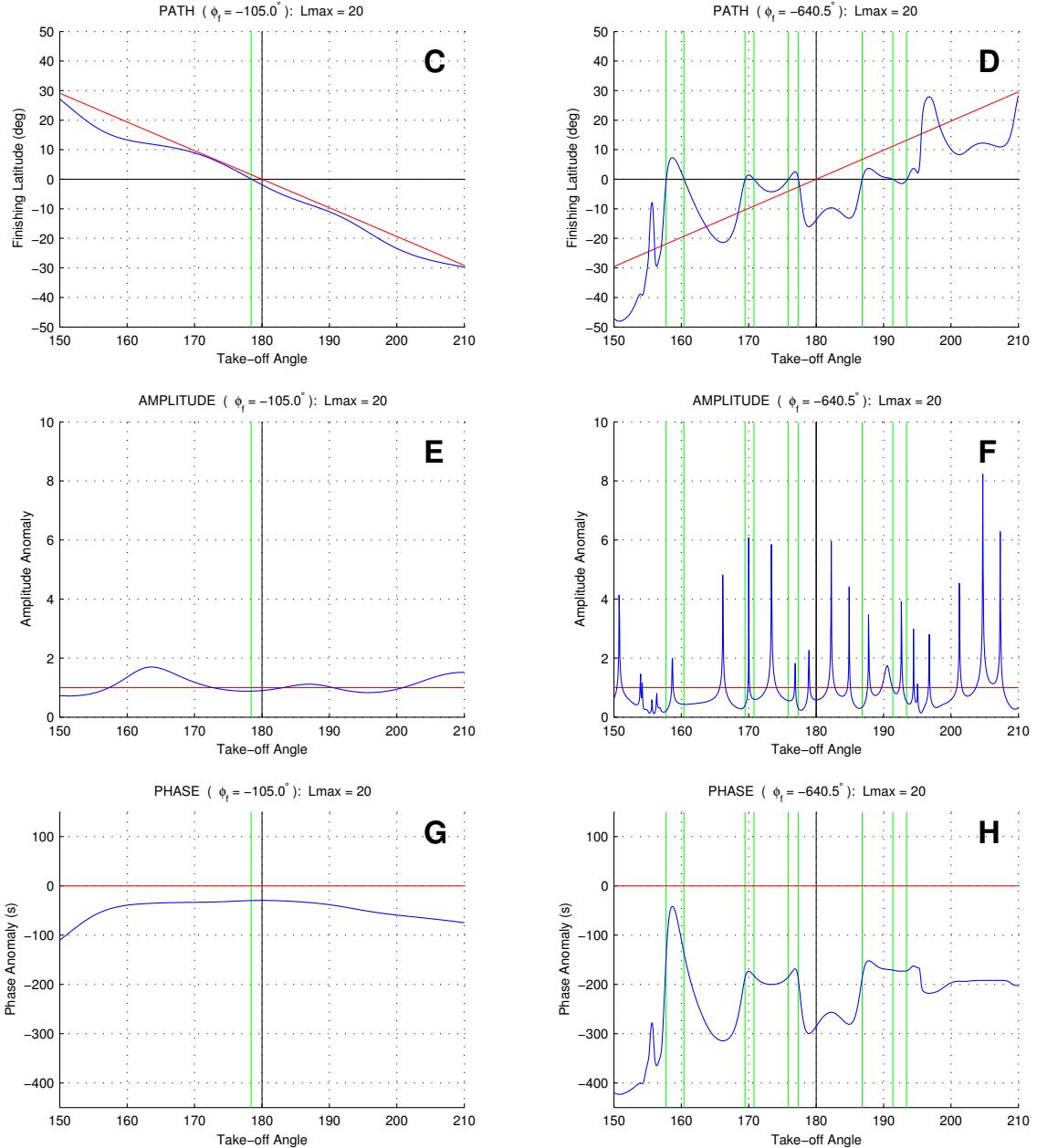


Figure 6.5: Ray theory predictions of path, amplitude, and phase. **A–B** show a ray diagram for Rayleigh waves (100 rays, $d\phi = 0.4^\circ$) superimposed on the phase velocity map of Figure 5.3A, which is expanded up to $l_{\text{max}}=20$. The source is at $(0, 0)$, the intersection of the Greenwich meridian and the equator; the finishing longitude-value for each ray is $\phi_f = -640.5^\circ$, equivalent to the R4 rays for a receiver on the equator at $\phi = +79.5^\circ$. Take-off angles range from 150° to 210° , as measured counter-clockwise from the equator. **B.** Same rays as in A. Red rays show the great circles that we would expect for the first and last ray of the homogeneous case. The two ϕ_f -values indicate the positions of the ray diagram cross-sections in C and D. **C.** Cross-section of the ray diagram at $\phi_f = -105.0^\circ$. The diagram indicates that one distinct ray path crosses the equator at the longitude of the receiver (vertical green line). **D.** Cross-section of the ray diagram at $\phi_f = -640.5^\circ$. The diagram indicates that nine distinct ray paths cross the equator (vertical green lines). In other words, ray theory predicts nine R4 rays between the source and receiver. **E–F.** Amplitude anomaly at $\phi_f = -105.0^\circ$ and $\phi_f = -640.5^\circ$. **G–H.** Phase anomaly at $\phi_f = -105.0^\circ$ and $\phi_f = -640.5^\circ$.

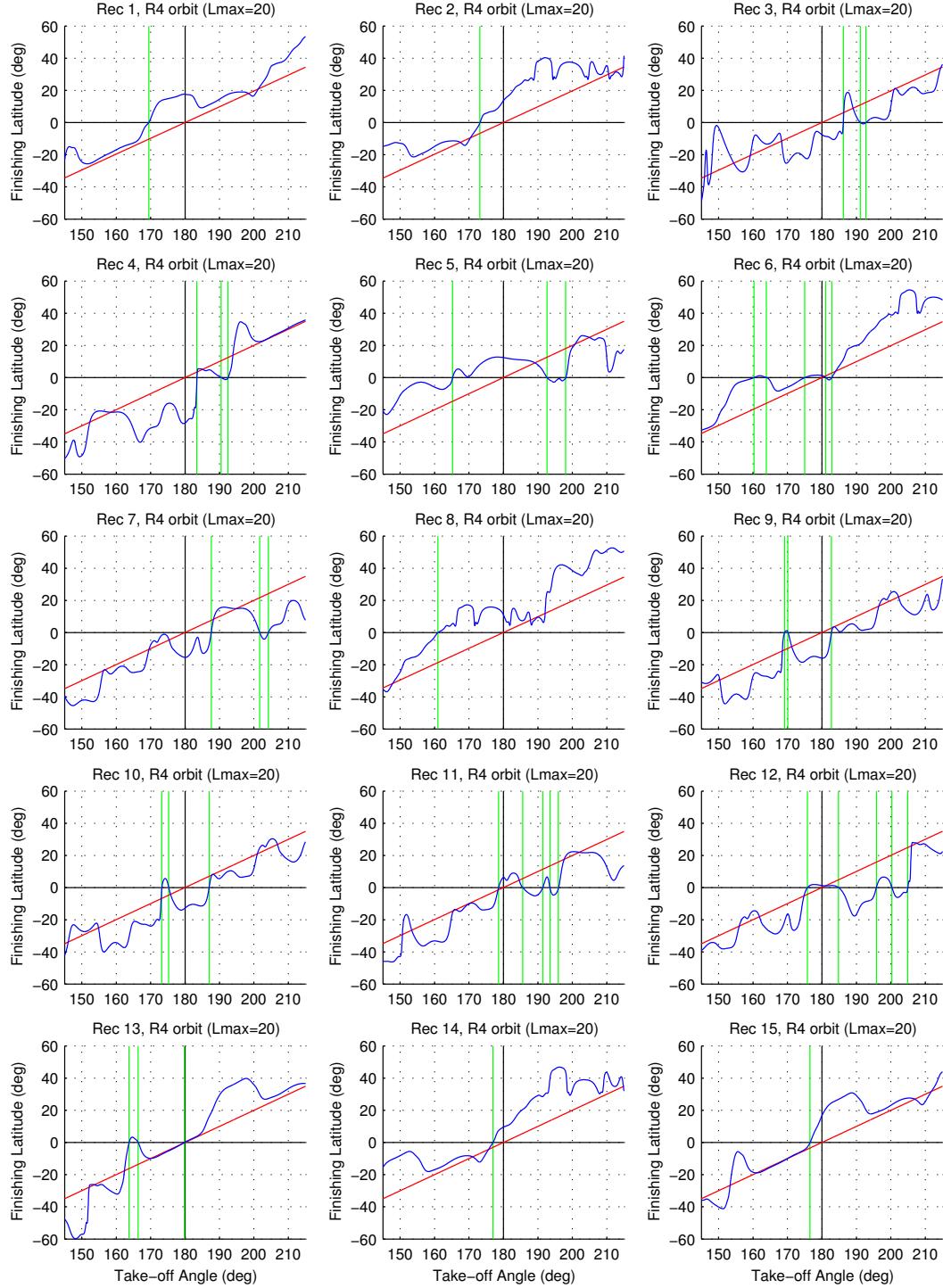


Figure 6.6: Ray diagram cross-sections. A cross-section is a plot of the finishing latitude at a specified longitudinal meridian ϕ_f as a function of take-off angle ($145^\circ \leq \alpha \leq 215^\circ$): red is for a homogeneous earth; blue is for a heterogeneous earth, which is indicated by the phase velocity map ($T=40s$, $l_{max}=20$, $\epsilon=1.0$). Shown here are the R4 ray diagram cross-sections for the first 15 receivers (see Figure 6.1). Vertical green lines show the rays that converge on the receiver, which had been rotated to the equator (finishing latitude is 0°). See Figure 6.7 and Section 6.7.

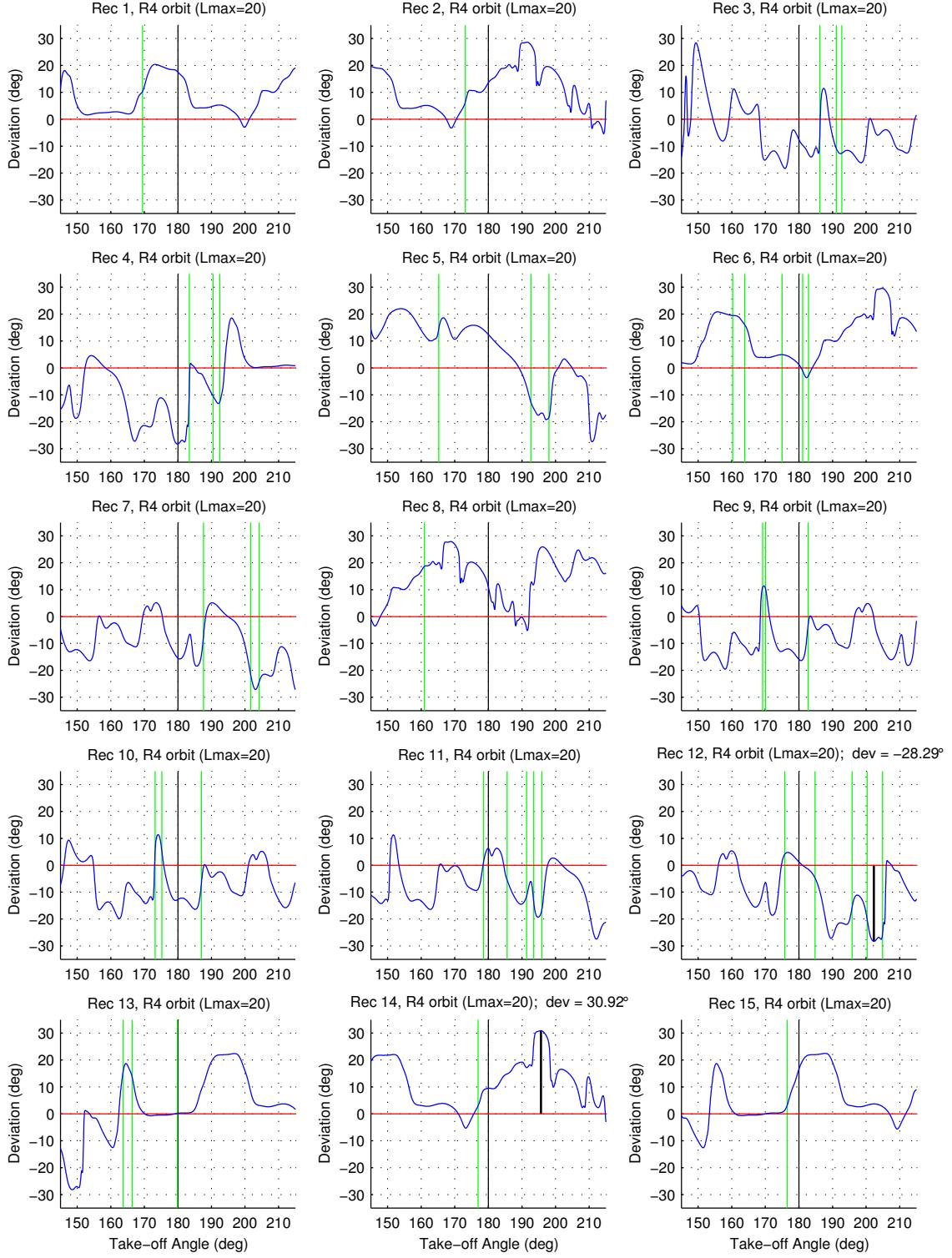


Figure 6.7: Path deviation from great circles. Each data point shows the finishing latitude for a ray in the heterogeneous model minus the finishing latitude for a ray in the homogeneous model having the *same* take-off angle at the source. These are obtained by subtracting the homogeneous cross-section curves from the heterogeneous cross-section curves in Figure 6.6. In each plot we are seeing the deviation of the R4 rays for the $T=40$ s, $l_{\text{max}}=20$, $\epsilon=1.0$ phase velocity map. The maximum positive and negative deviation from great circle paths for the 15,000 ray paths in this figure can be found in the plots for receivers 14 and 12, respectively.

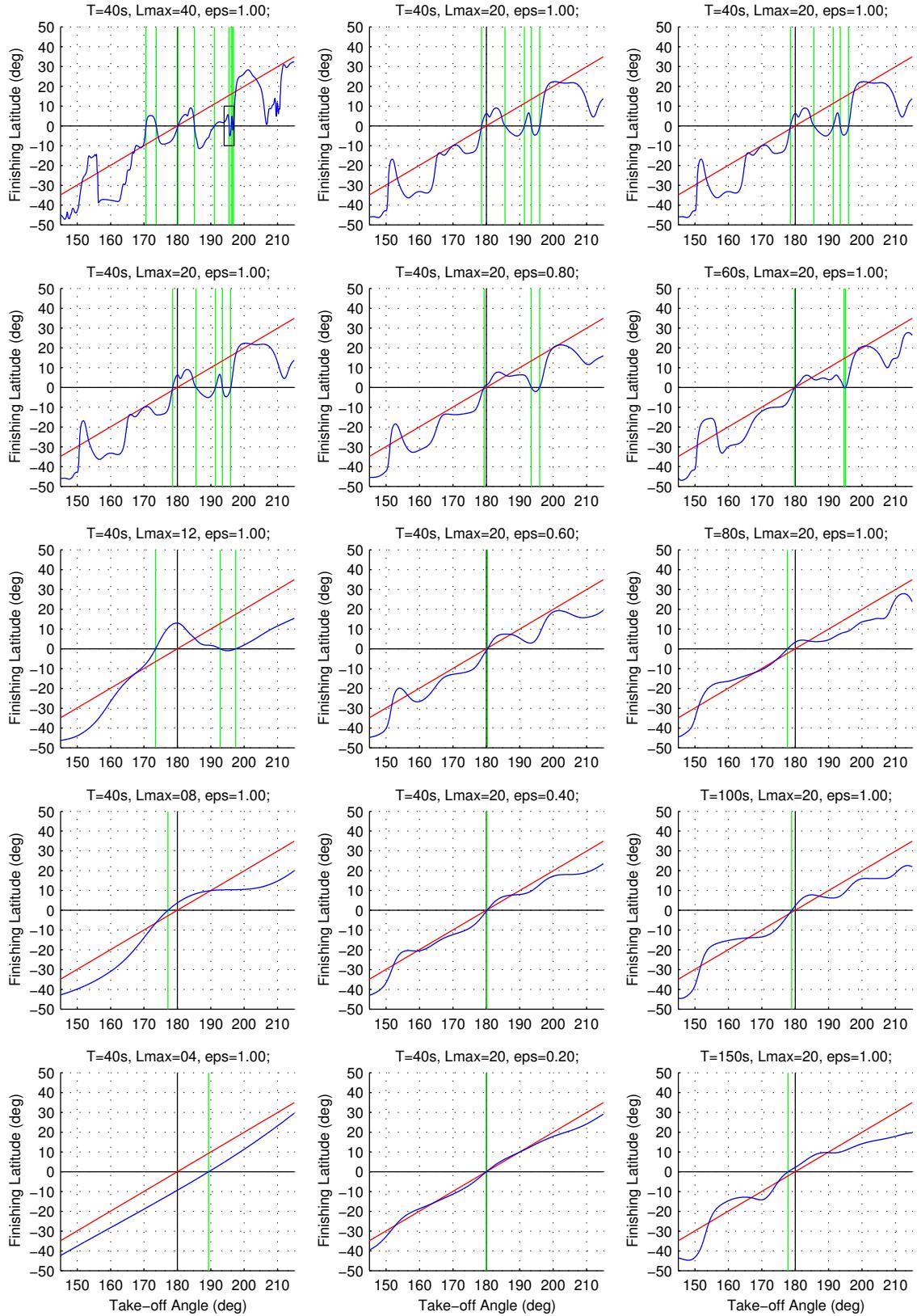


Figure 6.8: Ray diagram cross-sections: Variation with l_{max} , and ϵ , and T . Each plot is for the R4 rays between the source and receiver 11; the great-circle path length is $\phi_f = -624.53^\circ$. Left column shows variation with l_{max} ; middle column shows variation with ϵ ; right column shows variation with T . Thus each figure represents exact ray-tracing results for a different phase velocity map. See Figures 6.6 and 6.9 and Section 6.8.

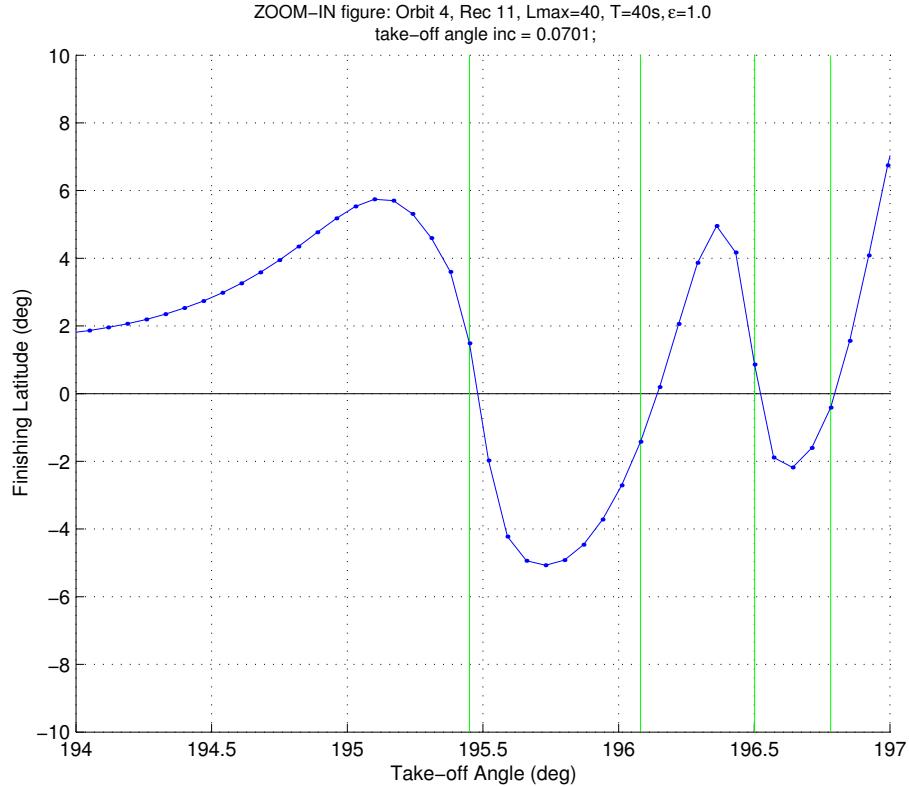


Figure 6.9: Zoom-in on the box shown in the upper-left plot of Figure 6.8, which is the phase velocity map with the greatest amount of short-scalelength structure ($T=40\text{s}$, $l_{\text{max}}=40$, $\epsilon=1.0$). The zoom-in box is shown to show that the “kinks” in the upper plots of Figure 6.8 are in fact smooth curves, and could be made smoother with more rays. There are 1000 points in each curve (i.e., 1000 rays, take-off angles, or finishing latitudes), so the take-off angle increment is $d\alpha = 70^\circ/1000 = 0.070^\circ$. The vertical green lines indicate the point closest to the zero and are only for graphing purposes; the actual converging rays are calculated by using a convergence method and not by using these data.

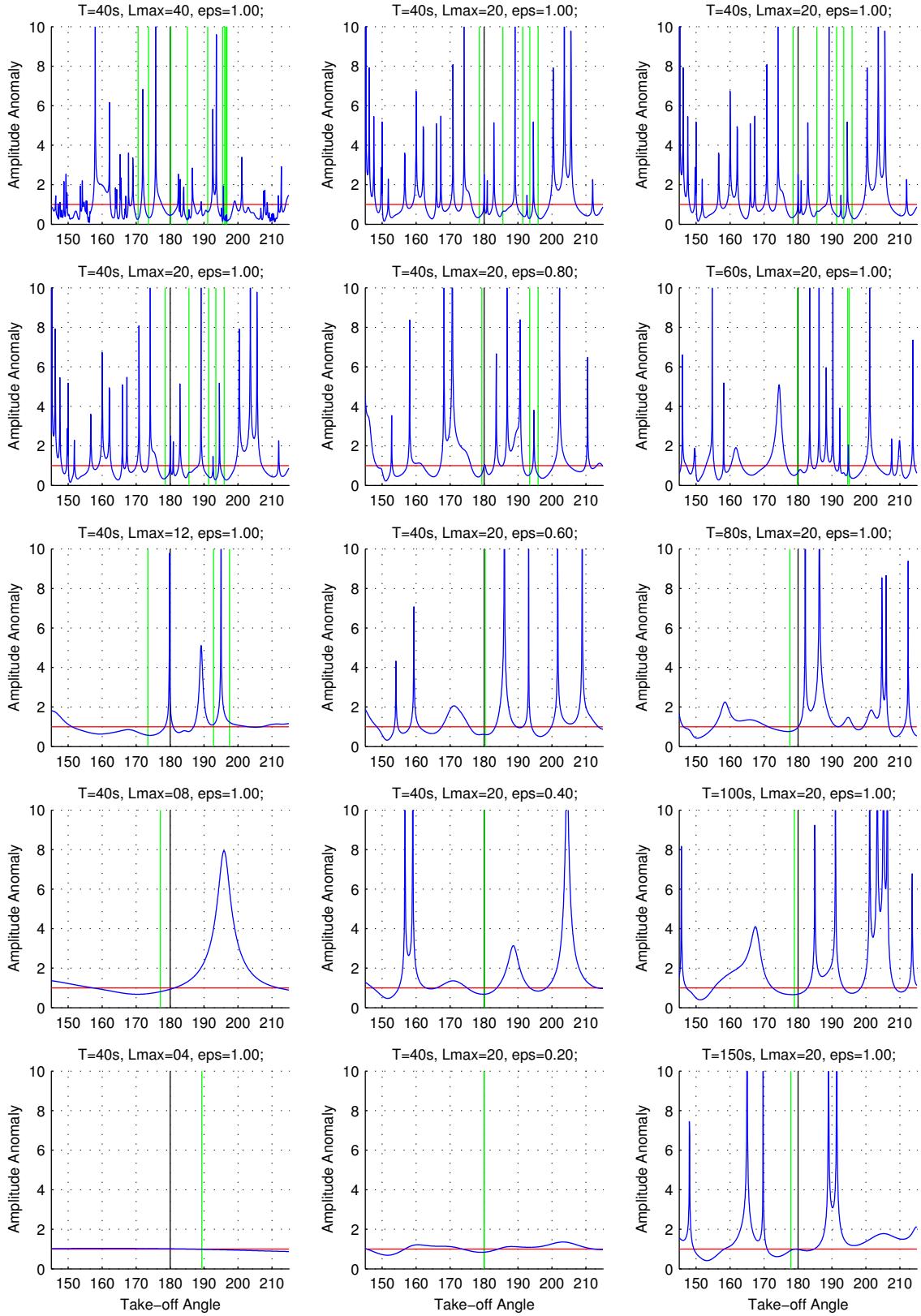


Figure 6.10: Same as Figure 6.8, only for amplitude anomaly. Left column shows variation with l_{max} ; middle column shows variation with ϵ ; right column shows variation with T . Vertical green lines indicate the amplitude values for the converging rays identified in Figure 6.8.

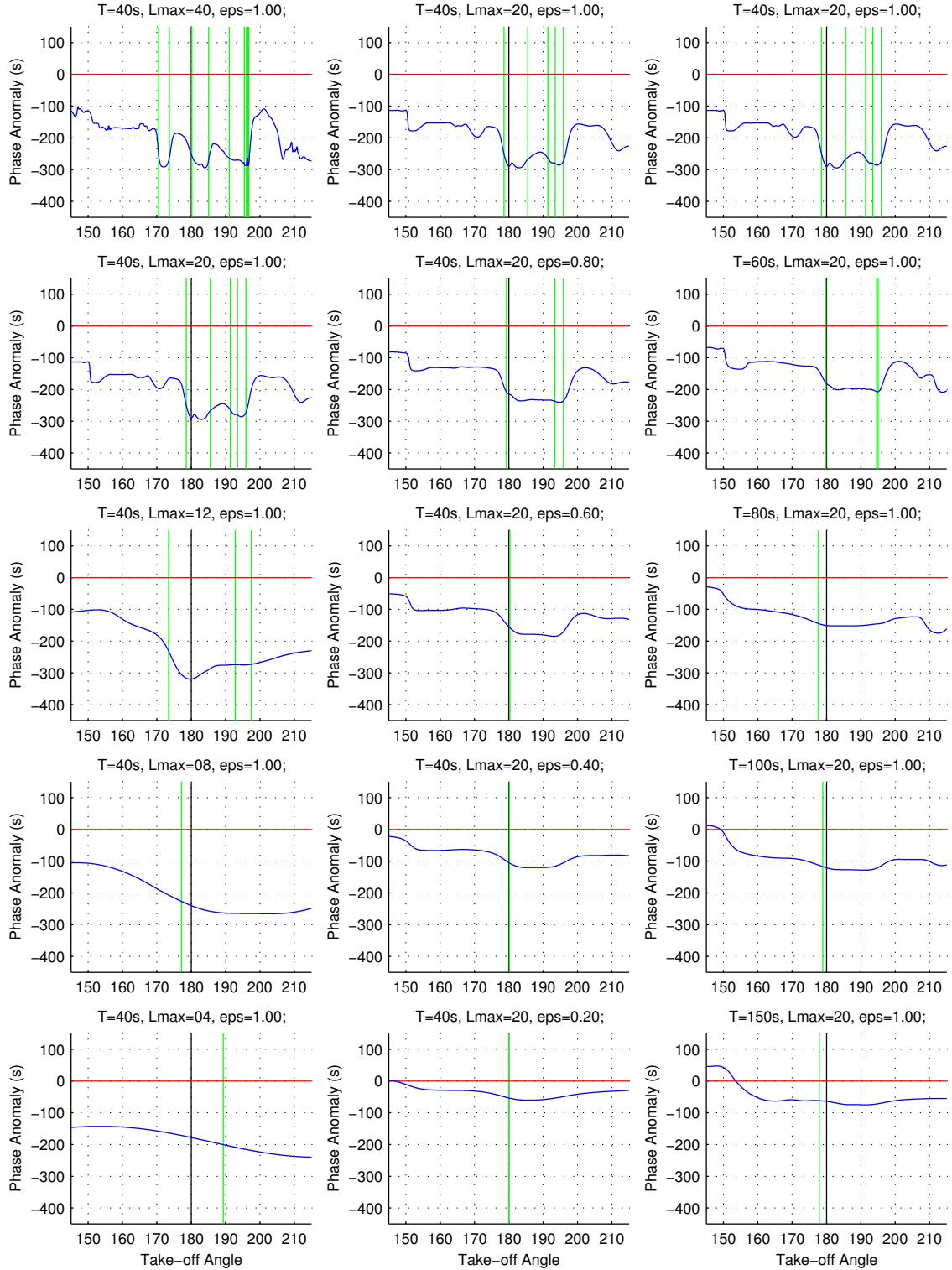


Figure 6.11: Same as Figure 6.8, only for phase anomaly. Left column shows variation with l_{max} ; middle column shows variation with ϵ ; right column shows variation with T . Note the correlation with the path anomalies in Figure 6.8, particularly for the higher l_{max} cases. Vertical green lines indicate the phase values for the converging rays identified in Figure 6.8.

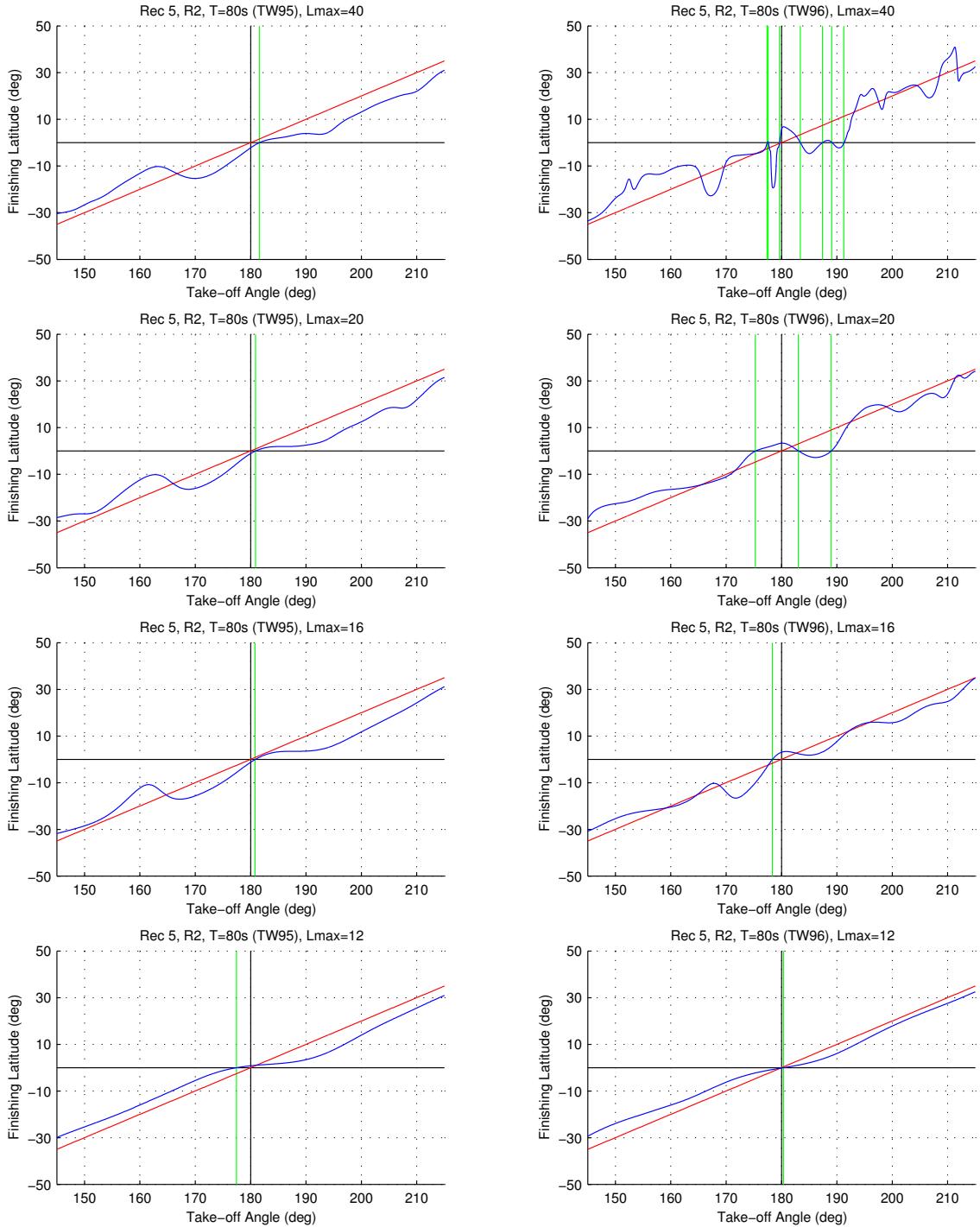


Figure 6.12: Ray diagram cross-sections (see Figure 6.6): TW95 vs TW96. Each plot is for the $T=80s$, $\epsilon=1.0$ map (Figure 5.6) for variable l_{max} (40, 20, 16, 12); left column is TW95, right column is TW96. Vertical green lines indicate distinct ray paths between the source and the receiver, which has been rotated to the equator. See Table 6.5 and Section 6.8.

Table 6.3: Dependence of multipathing on orbit and l_{max} ($T=40s$, $\epsilon=1.0$). Each number indicates the number of distinct ray paths for a particular source-receiver pair; a ‘-’ indicates a single ray (i.e., no multipathing). Source location and receiver locations are shown in Figure 6.1; path lengths are listed in Table 6.2. This page shows the results for receivers 1–15; the following page shows the results for receivers 16–25. The horizontal circled numbers refer to the plots in Figure 6.6; vertical circled numbers refer to the plots in the left column of Figure 6.8. See Section 6.8.

Orbit	l_{max}	Receiver Number														
		01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
R1	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
R2	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	14	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	16	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	18	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	20	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	40	-	-	-	-	-	-	-	-	-	-	3	3	-	3	-
R3	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	14	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	16	3	-	-	3	-	-	-	-	-	-	-	-	-	-	-
	18	3	-	-	3	-	-	-	-	-	-	-	-	-	-	3
	20	3	3	-	-	3	-	-	-	-	5	-	-	-	-	-
	40	3	-	-	3	-	5	-	3	3	3	-	3	-	5	3
R4	4	-	-	-	-	-	-	-	-	-	-	□	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-	□	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-	3	-	-	-	-
	12	-	-	-	-	3	-	3	-	-	-	3	-	3	-	-
	14	-	-	-	-	3	-	3	-	-	-	3	-	3	-	-
	16	-	-	-	-	3	3	-	-	-	-	□	-	3	-	-
	18	-	-	-	-	5	5	5	-	-	-	5	3	3	-	-
	20	□	□	3	3	3	5	3	□	3	3	5	5	3	3	□
	40	-	3	9	5	-5	7	7	-	5	3	9	11	3	5	-

Orbit	l_{max}	Receiver Number									
		16	17	18	19	20	21	22	23	24	25
R1	4	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-
	14	-	-	-	-	-	-	-	-	-	-
	16	-	-	-	-	-	-	-	-	-	-
	18	-	-	-	-	-	-	-	-	-	-
	20	-	-	-	-	-	-	-	-	-	-
	40	-	-	-	-	-	-	-	-	-	-
R2	4	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-
	14	-	-	-	-	-	-	-	-	-	-
	16	-	-	-	-	-	-	-	-	-	-
	18	-	-	-	-	-	-	-	-	-	-
	20	-	-	3	-	-	3	-	-	-	-
	40	-	3	5	-	-	-	-	3	-	3
R3	4	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	3	-	-
	14	-	-	-	-	-	-	-	-	-	-
	16	-	-	-	-	3	-	-	-	-	-
	18	-	-	3	-	-	-	-	-	-	-
	20	-	-	3	-	5	-	-	-	-	-
	40	3	3	5	5	-	-	-	9	5	-
R4	4	-	-	-	-	-	-	-	-	-	-
	6	-	-	-	-	-	-	-	-	-	-
	8	-	-	-	-	-	-	-	-	-	-
	10	-	-	-	-	-	-	-	-	-	-
	12	-	-	-	-	-	-	-	-	-	-
	14	-	-	-	-	-	-	-	-	-	-
	16	-	-	-	-	-	3	3	-	-	-
	18	-	-	3	-	-	3	-	-	-	3
	20	-	3	3	3	-	3	-	3	-	5
	40	3	5	13	7	3	9	5	11	-	15

Table 6.4: Dependence of multipathing on ϵ , a parameter that adjusts the power of the phase velocity map, and T , the period of the map. Each number indicates the number of distinct ray paths for a particular source-receiver pair; a ‘-’ indicates a single ray (i.e., no multipathing). Source location and receiver locations are shown in Figure 6.1. Shown here is the number of paths between the source and each receiver for the R4 orbit. The circled numbers refer to the plots in the middle column (ϵ) and right column (T) of Figure 6.8. See Section 6.8.

Receiver (R4)	Vary ϵ ($lmax=20$, $T=40s$)					Vary T ($lmax=20$, $\epsilon=1.0$)				
	1.0	0.8	0.6	0.4	0.2	40	60	80	100	150
	-	-	-	-	-	-	-	-	-	-
1	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-
3	3	-	-	-	-	3	-	-	-	-
4	3	-	-	-	-	3	3	3	-	-
5	3	3	-	-	-	3	3	3	3	-
6	5	-	-	-	-	5	-	-	-	-
7	3	-	-	-	-	3	-	-	-	-
8	-	-	-	-	-	-	-	-	-	-
9	3	-	-	-	-	3	3	-	-	-
10	3	3	3	-	-	3	3	-	-	-
11	5	3	-	-	-	5	3	-	-	-
12	5	5	-	-	-	5	5	3	-	-
13	3	3	-	-	-	3	3	3	3	-
14	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-	-	-
17	3	-	-	-	-	3	3	-	-	-
18	3	3	-	-	-	3	3	-	-	-
19	3	-	-	-	-	3	-	-	-	-
20	-	-	-	-	-	-	3	-	-	-
21	3	3	3	-	-	3	5	3	-	-
22	-	-	-	-	-	-	-	-	-	-
23	3	-	-	-	-	3	3	-	-	-
24	-	-	-	-	-	-	-	-	-	-
25	5	3	-	-	-	5	5	-	-	3

Table 6.5: Dependence of multipathing on tw95 vs tw96, i.e., the phase velocity of maps from *Trampert and Woodhouse* (1995) and *Trampert and Woodhouse* (1996) for the same T , l_{max} , and ϵ . Each number indicates the number of distinct ray paths for a particular source-receiver pair; a ‘-’ indicates a single ray (i.e., no multipathing). Source location and receiver locations are shown in Figure 6.1. Shown here is the number of paths between the source and each receiver for the R2 orbit. For the R1 orbit there are no cases of multipathing for the 25 source-receiver pairs. The circled numbers refer to the plots in Figure 6.12. See Section 6.8.

Chapter 7

Comparison of results from numerical model with results from ray theory

Contents

7.1	Introduction	113
7.2	Exact ray theory vs linearized ray theory	115
7.3	Obtaining phase and amplitude anomalies from the numerical model	117
7.4	Ray theory validity condition	117
7.5	Numerical model vs ray theory	119
7.6	Summary and conclusions	121

7.1 Introduction

In the previous chapter we illustrated how according to ray theory the amount of deviation from great-circle paths depends on the distance of propagation Δ and the properties of the phase velocity map (l_{max} , T , and ϵ). In this chapter we introduce an additional parameter that is extremely important when considering the results from the numerical model: the wavelength λ of the wave (Section 7.4). We compare some of the results from the ray theory

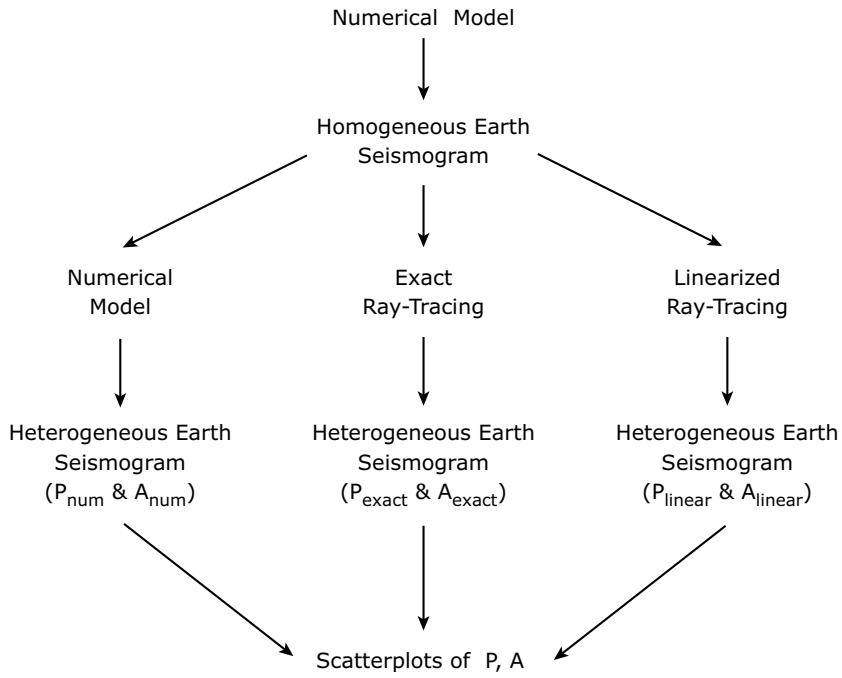


Figure 7.1: Flow chart for creating scatterplots comparing seismograms from three models: numerical model, exact ray theory, and linearized ray theory. Each heterogeneous earth seismogram is specified in terms of a phase anomaly and amplitude anomaly. The calculation of P_{num} and A_{num} is shown in Figure 7.6. (See Appendix G for more details.)

calculations — both exact ray theory and linearized ray theory — with results from the numerical model, and we discuss them in term of the ray theory validity condition (Section 7.5).

The results are plotted as scatterplots correlating the phase or amplitude anomalies predicted by one theory with those predicted by another theory. Thus, each point in the scatterplot represents a correlation between a seismogram predicted by one method versus a seismogram predicted by another method (Figure 7.1). In each scatterplot we show a minimum of 100 data points, representing R1–R4 for 25 receivers; the number of points will exceed 100 in cases of multipathing occurrences in exact ray theory predictions. The units of the phase anomaly are seconds¹. The limits on phase anomaly in the scatterplots are always the same; the limits on the amplitude anomaly vary from figure to figure.

In Section 7.2 we compare the results between exact ray theory and linearized ray theory.

In Section 7.5 we compare the results between the numerical model and exact ray theory.

¹Normally, phase is expressed in units of radians. With our single-frequency numerical calculations, the traveltime difference could be converted to a phase shift by linear scaling, knowing the wavelength in the numerical model. Any scaling, of course, would not affect the appearance of the scatterplot correlations.

7.2 Exact ray theory vs linearized ray theory

Here we will examine the agreement (and disagreement) between linearized ray theory predictions and exact ray theory predictions. In general these results should be correlated, and in some limiting cases, they should match exactly. However, as we will see, they are poorly correlated under many circumstances. As we did in Section 6.8, we will discuss how the correlation varies in terms of $lmax$, ϵ , T , and Δ . These are the parameters discussed and presented in figures in Section 5.2.

Linearized ray theory is based on slight perturbations of the ray path from the great-circle path. Thus, any effect that causes an increase in path deviation will make the results of linearized ray theory less valid, and hence the correlation with the exact ray-tracing results worse. In essence, the figures in this section give a different perspective on the material presented in Section 6.8. One of the key points here is that exact ray theory may (and often does) predict multiple ray paths, while linearized ray theory always predicts one ray path.

Variation with $lmax$

Figure 7.2 shows plots of exact ray theory predictions of amplitude anomalies and phase anomalies versus predictions from linearized ray theory. With increasing $lmax$, the deviations from the great-circle paths increase (Figure 6.8), and the correlation between exact and linear ray theory worsens (Figure 7.2).

We should note three points from Figure 7.2: (1) the correlation fails at very different $lmax$ levels for amplitude ($lmax \approx 8$) and phase ($lmax \approx 20$); (2) linearized ray theory predictions of amplitude anomalies are greater than those of exact ray theory, especially for $lmax \geq 12$; (3) exact ray theory predicts a very wide spread in phase anomalies in cases of extreme multipathing (blue dots). It is quite remarkable how “steady” the correlation is for phase; the correlation for $lmax=20$ is not much worse than for $lmax=4$. Also note that the range of phase anomalies remains the same for each value of $lmax$. This suggests that the phase anomaly is primarily sensitive to low-degree structure.

Variation with ϵ

The motivation for including ϵ as a testing parameter is that it should provide the precise scenario in which the results of exact ray theory converge to those of the linearized ray theory, i.e., linearized ray theory is guaranteed to be a good approximation for small enough values of ϵ . As $\epsilon \rightarrow 0$, the ray paths of the exact ray-tracing approach the great-circle paths used with linearized ray theory, as we observed in Figure 6.8 (middle column).

Figure 7.3 confirms our hypothesis nicely: as $\epsilon \rightarrow 0$, both the amplitude and phase anomalies approach agreement for the two ray theories. For the phase velocity map in Figure 7.3, the exact ray theory predictions begin to deviate from the linearized predictions when $\epsilon \approx 0.4$.

Variation with T

Finally, we show some results for different period phase velocity maps to illustrate the effects of the varying degree of structure of these maps. As discussed in Section 5.2, each phase velocity map has different power in each part of the amplitude spectrum (Figure 5.5A-B). Figure 7.4 shows how the correlation between linearized ray theory and exact ray theory increases with increasing period. This is because the longer-period maps contain less shorter-scalelength structure, and thus the deviation from linearized ray theory is less.

Variation with Δ

The deviation from great-circle paths increases with path length Δ , as we showed in Figure 6.5. Thus, we would expect the discrepancy between the linearized calculations and the exact calculations to increase with path length Δ as well. Figure 7.5A-P shows the influence of path length on the exact-vs-linear ray theory correlation. It is clear that for longer path lengths (i.e., higher orbits), the correlation is poorer (and is worse for higher l_{max} values). In particular, for the $T=40$ s map of *Trampert and Woodhouse* (1995), linearized ray theory predictions of phase agree with exact ray theory up to orbits of R3, even for $l_{max}=40$ (Figure 7.5O). Linearized ray theory predictions of amplitude agree with exact ray theory only for R1 for $l_{max}=40$ (Figure 7.5I) and up to R2 for $l_{max}=8$ (Figure 7.5B).

The average phase anomaly becomes increasingly negative as the orbit increases (Figure 7.5E-H). This is a result of the non-zero (and positive) $l=0$ component of the phase velocity map, meaning that the heterogeneous model is on average faster than the PREM model.

As a result, the rays arrive earlier for longer paths, indicated by the increasingly negative phase anomalies for increasing orbit (see Table 6.1). The only positive phase anomalies (i.e., late arrivals) occur for R1 orbits (Figure 7.5E,M). It would be helpful as an additional test to remove the $l=0$ component, which would remove the bias of the phase anomaly correlation (e.g., *Trampert and Woodhouse*, 2001).

7.3 Obtaining phase and amplitude anomalies from the numerical model

The numerical solutions $u(\theta, \phi, t)$ from the wave propagation on the hexagonal grids give both the homogeneous case (HOM) and the heterogeneous case (HET). The numerical solutions at the 25 receivers provide 25 pairs of HOM-HET seismograms, each containing the R1–R4 arrivals. In simplest terms, we obtain the phase and amplitude anomalies by shifting and amplifying HOM to match HET; this is described further in Figure 7.6.

In general, there is some discrepancy between the HET seismogram and the ray-theory-corrected HOM seismogram (Figure 7.6B). These discrepancies, generally very small, are effectively not considered in the comparison of numerical model results with ray theory results, since we specify the heterogeneous earth seismogram in terms of only two parameters, P and A . The discrepancies are minimal when the wavelength of the input wave is much less than the minimum scalelength of the heterogeneity (Figure 5.11), in other words, when the ray theory validity condition is satisfied.

7.4 Ray theory validity condition

In the next section we compare phase and amplitude predictions from the numerical model and ray theory; here we note where we expect ray theory predictions to fail. Ray theory predictions are valid when the wavelength of the wave, λ , is much less than the scalelength of the heterogeneity, Λ , in the earth model, i.e.,

$$\lambda \ll \Lambda , \tag{7.1}$$

which we will refer to as the *validity condition*.

From a practical standpoint, we would like to know *how much less* the wavelength needs to be in order to obtain a good correlation between our ground-truth model (numerical model) and exact ray theory. Furthermore, we want to examine the sensitivity of amplitude anomalies and phase anomalies to the validity condition. These questions can be examined by approaching the validity condition from either side, by varying the wavelength or the scalelength.

Changing the scalelength Λ

The easiest way of testing the validity condition is by varying the scalelength of the phase velocity maps. Table 5.2 shows the relationship between the l_{max} of a phase velocity map and the minimum possible scalelength, Λ_{\min} , which we will simply call the scalelength Λ .

Changing the wavelength λ

We can also test the validity condition by varying the input wavelength in the numerical model. As we mentioned in Section 5.5, the limiting factor on using shorter wavelengths in our numerical model is the gridpoint spacing $d\theta$. And based on the construction of our grids, this value will halve for every increase in grid-order q , which will result in a factor of four times as many gridpoints. Furthermore, we require twice as many time-steps if we halve the gridpoint spacing, so in total we have a substantial increase in computation time when increasing the grid order.

Figure 7.7 shows a plot of how the wavelength of the wave in the numerical model varies as a function of the source parameters $\mu(\theta_{wid})$ and σ . The range of values for $\lambda(\mu, \sigma)$ shows the approximate range of wavelengths that our numerical grids, orders $q = 4, 5, 6$ can sustain while managing accurate solutions. For example, the order $q=6$ grid, containing 122,882 points, can support the shortest wavelengths: $\lambda \approx 2050$ km. This is the value used in all of the numerical results, with the exception of the top two plots in Figure 7.11.

The validity condition of *Wang and Dahlen* (1995b)

Wang and Dahlen (1995b) extend the notion of the validity condition in Equation (7.1) by incorporating Fresnel areas around the ray path. They express the original validity condition, Equation (7.1), as

$$\frac{\bar{s}}{l} \ll 1 , \quad (7.2)$$

where \bar{s} is a calculation of the average scalelength of an earth model (*Wang and Dahlen*, 1995b) and l is the angular degree corresponding to the wavelength of surface waves. They derive a new validity condition,

$$\frac{\bar{s}}{\sqrt{4\pi l}} \ll 1 , \quad (7.3)$$

which is more stringent than the condition in Equation (7.2). *Wang and Dahlen* (1995b) emphasize that this condition does not depend on path length.

7.5 Numerical model vs ray theory

Finally we have reached the stage where we can compare results of our numerical model with results from ray theory (exact or linear). The numerical model can be thought of as our “ground-truth” model in the sense that it provides the most complete picture of wave propagation, and we have demonstrated good agreement between numerical and analytical solutions (e.g., Figure 5.7). It is the numerical model which we set as the standard to which ray theory predictions are compared.

Figure 7.8 shows results from all three models for a very broad, $lmax=6$ phase velocity map. The results from the numerical model correlate with the results from ray theory, both exact and linearized. The phase anomalies are extremely well correlated, and the agreement is significantly better than the amplitude anomalies. This is in support of our findings in Chapter 6 that phase anomalies are better behaved than amplitude anomalies. Thirdly, and very importantly, we see that the amplitude results from exact ray theory are better correlated with the numerical model than are the results from linearized ray theory.

Figure 7.9 shows the same results as Figure 7.8, only for a map with more short-scalelength structure ($lmax=12$). The presence of colored points indicates a regime of multipathing. In the plots comparing linearized ray theory and the numerical model, the colored points indicate the corresponding multipathing rays predicted by exact ray theory (of course, both linearized ray theory and the numerical model calculate only a single arrival associated with each orbit). Although in general the exact ray theory predictions are better than the linearized ray theory predictions, we see that in some cases — for example, the multipathing rays — the exact ray theory predictions are worse (Figure 7.9). The remaining scatterplots will compare predictions from the numerical model and exact ray theory.

Variation with $lmax$ (Λ)

Figure 7.10 shows the systematic worsening of the correlation of amplitude anomalies and phase anomalies between the numerical model and ray theory for increasing $lmax$. Beginning with $lmax=4$, we have good correlation for both amplitude and phase. As we increase $lmax$, thereby increasing the amount of short-scalelength structure, the correlation decreases. We see that the ray theory predictions for amplitude break down seriously, while the results for phase remain quite good. Thus, ray theory predictions of phase based on an $lmax=20$ map are probably accurate enough for most purposes (especially for R1 and R2).

In Figure 7.10 the correlation breaks down at $lmax \approx 10$ for amplitude and $lmax \approx 20$ for phase. Thus, in terms of the validity condition (Eq. 7.1), the ray theory predictions of amplitude are accurate down to scalelengths of $\Lambda \approx 3812$ km with $\lambda \approx 2050$ km, and the predictions of phase are accurate even when $\Lambda \approx \lambda$ (Table 5.2).

Variation with λ

By using the order $q=6$ spherical grids, we can introduce shorter wavelengths. Based the validity condition ($\lambda \ll \Lambda$), we would expect shorter wavelengths to increase the correlation between the numerical results and the ray theory results (both exact and linear). Indeed, this is what we see in Figure 7.11. The only difference between the two scenarios is the wavelength of the input wave, and the correlation is clearly better for the shorter wavelength.

Variation with T

Exact ray theory predictions of phase and amplitude are better for longer-period phase velocity maps, which contain less shorter-scalelength structure. This is illustrated in Figure 7.12, which shows the variation with period for $lmax=12$ phase velocity maps.

Variation with ϵ

Exact ray theory predictions are better for phase velocity maps with low power (Figure 7.13). The reduced-power maps diminish the effects of focusing and defocusing, thereby keeping the exact ray paths closer to the great-circle paths, along which the linearized ray theory predictions are calculated.

Variation with Δ

Figure 7.14A-X shows the dependence of the exact ray theory predictions on orbit for R1–R4, using the $T=40\text{s}$ ($\epsilon=1.0$) phase velocity map. Ray theory predictions of amplitude and phase are worse for higher orbits, and this is most noticeable for higher- $lmax$ maps, where the validity condition is less well-satisfied (Figure 7.14Q-X). In particular, the $lmax=12$ map is a regime in which it is not well-satisfied, considering the wavelengths in our numerical model ($\lambda \approx 2050 \text{ km}$, $\Lambda \approx 3202 \text{ km}$). *Wang and Dahlen* (1995b) emphasize that their more stringent validity condition (Eq. 7.3) does not depend on orbit number. However, even some of their figures show a worsening of exact ray theory predictions with increasing orbit, especially for higher degree models² ($lmax > 20$).

7.6 Summary and conclusions

Our results are based on global surface wave Rayleigh wave orbits up to R4, i.e., path lengths up to $\Delta \approx 640^\circ$ (Table 6.2). Here results from three models are summarized: numerical model, exact ray-tracing, linearized ray-tracing. We also raise some questions and point out directions of future work.

1. NUMERICAL MODEL

²See Figures 9 and 11 in *Wang and Dahlen* (1995b). The error of the ray theory predictions increases with orbit, and is especially noticeable for Love waves, G1–G4. Their ground-truth calculations are based on a coupled-mode method described by *Um and Dahlen* (1992).

- (a) With a non-standard distribution of gridpoints on a sphere (Section 4.2), and with a relatively simple finite-difference method (Figure 4.11), we are able to produce numerical solutions to the wave equation on a spherical membrane. The numerical model offers an appealing approach to examining surface wave propagation, and for comparing its results with predictions from ray theory.
- (b) Our numerical solutions to wave propagation on a spherical membrane agree with analytical solutions for the homogeneous case (Figure 5.7). This gives us confidence in the numerical solutions for the heterogeneous case.
- (c) The minimum number of gridpoints per wavelength on our spherical grids could be improved by optimizing the distribution of gridpoints, or perhaps by using a better finite difference approximation to the Laplacian. This would allow us to use shorter wavelengths in our grids without going to finer grids.

2. RAY THEORY

- (a) Linearized ray theory results — i.e., phase and amplitude anomalies — are calculated along the great-circle path between the source and receiver. Exact ray theory results are calculated along the actual ray path, and thus when the actual path deviates from the great-circle path, the “exact” results deviate from the linearized results. Deviation from linearized ray theory increases with: (1) increasing path-length Δ , (2) increasing l_{max} , (2) increasing ϵ , and (4) decreasing period T .
- (b) Multipathing — multiple ray paths between the source and receiver — tends to be the norm for path lengths longer than R_1 and for our current resolution of the phase velocity maps ($l_{max}=20$) (Tables 6.3–6.5). Multipathing is an indication of the divergence between linearized ray theory and exact ray theory; in a regime with no multipathing, linearized ray theory predictions agree with exact ray theory predictions.
- (c) Ray diagram *cross-section curves*, $\bar{\theta}_f(\alpha)$, show the variation of the finishing latitude as a function of take-off angle. The variability of these curves is surprisingly sharp and almost “kinked” (Figure 6.9). It is not obvious what controls the “sharpness” of the cross-section curves. Certainly the path length and the scalelength of the structure (l_{max}) play a role, but this topic warrants further analysis.

3. COMBINED RESULTS

- (a) Exact ray theory is better than linearized ray theory at predicting amplitude and phase anomalies for $l_{max} < 10$ phase velocity maps (Figure 7.8). In this regime multipathing is rare.
- (b) Linearized ray theory may be better than exact ray theory at predicting amplitude and phase anomalies in regimes of extreme multipathing (e.g., $l_{max} \geq 20$ phase velocity maps), where exact ray theory predicts seemingly implausible results (e.g., Figure 7.2). Based on the phase velocity maps used here, it appears that in the regime $\epsilon < 0.4$, linearized ray theory predictions of amplitude and phase agree with those from exact ray theory (Figure 7.3).
- (c) We check the validity of the predictions of ray theory by altering the two parameters of the *validity condition*, $\lambda \ll \Lambda$, where λ is the wavelength in the numerical model and Λ is the scalelength of the phase velocity map (Figures 7.10 and 7.11). Our

results show that the validity condition is more stringent for amplitudes than for phase (Figure 7.10), something we might have expected based on the results from linearized ray theory (Figure 7.2).

Physically, in the regimes approaching the validity condition ($\lambda \approx \Lambda$) or outside the validity condition ($\lambda > \Lambda$), the very-long-period waves in the numerical model effectively “smooth over” some of the heterogeneity in the earth model, but with ray theory no such smoothing occurs (Figure 7.15).

- (d) In other words, ray theory is particularly sensitive to the smallest scalelengths in the model, whereas physical waves are not — they tend to “see” only elements of the model comparable with or larger than their wavelength. It may be that a smoothed or filtered version of the model (i.e., not simply a truncated l_{max} version) could yield more stable and accurate results in ray theory, a conjecture that remains for future work.

4. EXTENDING THIS STUDY

- (a) An increase in computing power would allow us to compare results from ray-tracing with those from the numerical model for higher l_{max} phase velocity maps. With more computing power, we can use finer grids and thus propagate shorter-wavelength waves in the numerical model. These waves need to be sufficiently shorter than the scalelength structure, in order to approach the infinitely-short wavelength approximation employed by ray theory.
- (b) With finer grids, we would enter the regimes of multipathing ($l_{max} > 12$) and perhaps find a way to utilize the information from the multiple ray paths. It could be that, with a fine enough numerical grid, we could identify the surface wave arrivals corresponding to each ray path.
- (c) Real surface waves, such as those in Figure 1.3, include a range of frequencies. The wave propagation in our numerical model is non-dispersive (i.e., phase velocity, c , is independent of frequency). It would be desirable to introduce dispersion into the numerical model, in order to model surface wave properties more realistically.
- (d) It is also possible to use ray theory to create more realistic waveforms. For example, we could combine the phase and amplitude anomaly results from all periods to make a single waveform. Realistic waveforms from the numerical model or waveforms from ray theory could be compared with observed seismograms (e.g., *Woodhouse and Wong, 1986; Laske, 1995*).

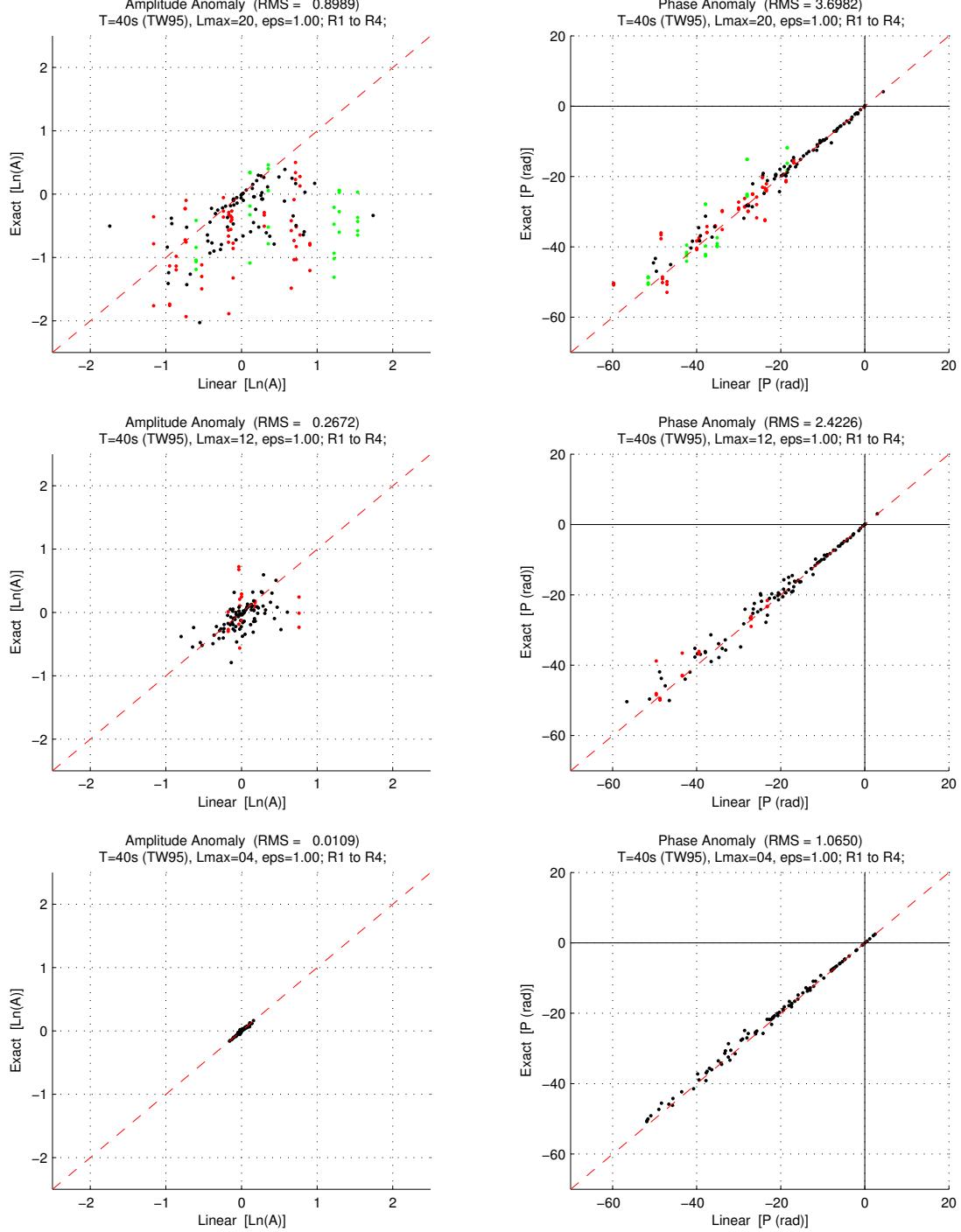


Figure 7.2: Vary l_{max} . Correlation between linear and exact ray theory for varying the l_{max} value of the phase velocity maps ($T=40\text{s}$, $\epsilon=1.0$). Each plot is of the exact ray theory predictions vs the linearized ray theory predictions; left column is for amplitude anomalies, right is for phase anomalies. The l_{max} value decreases from $l_{\text{max}}=40$ at the top to $l_{\text{max}}=4$ at the bottom. The correlation increases as l_{max} decreases. Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays, green = 5 rays, blue = ≥ 7 rays. See Section 7.2.

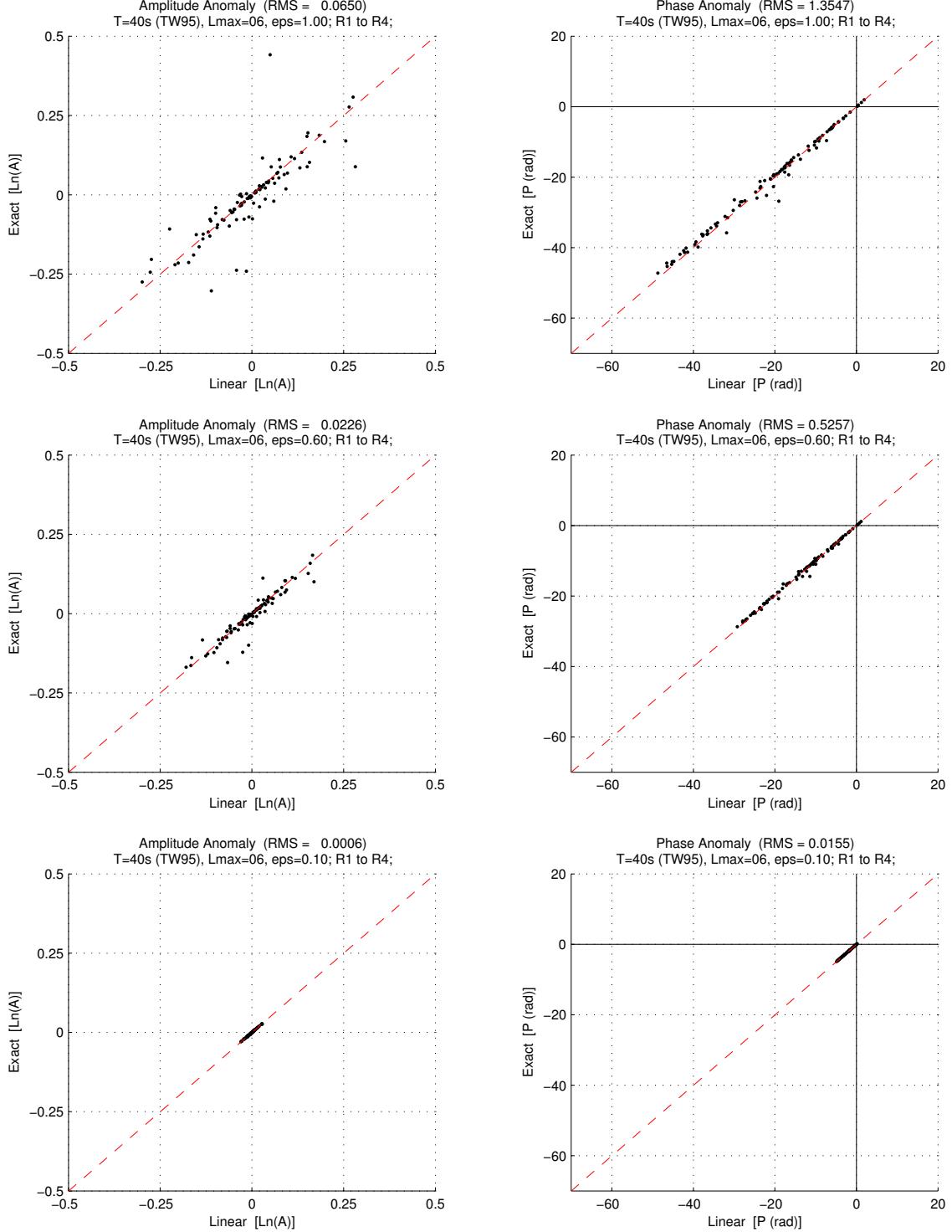


Figure 7.3: Vary ϵ . Correlation between linear and exact ray theory for varying the ϵ value of the phase velocity maps ($T=40s$, $lmax=6$). Each plot is of the exact ray theory predictions vs the linearized ray theory predictions; left column is for amplitude anomalies, right is for phase anomalies. The ϵ value increases from $\epsilon=1.0$ at the top to $\epsilon=0.1$ at the bottom. As expected, the correlation increases as $\epsilon \rightarrow 0$. There is no multipathing here, due to the relatively low $lmax=6$ of the map. See Section 7.2.

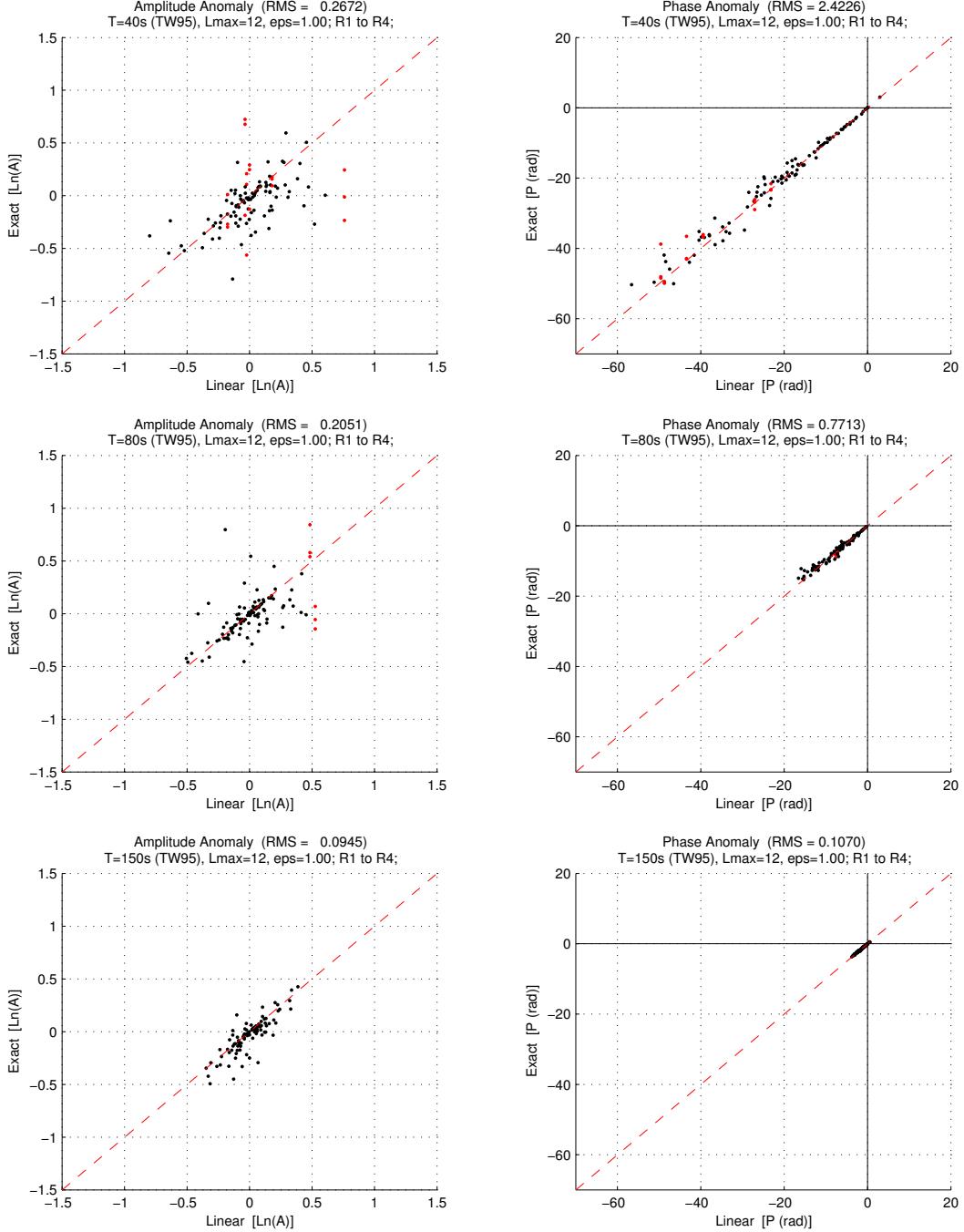


Figure 7.4: Vary T . Correlation between linear and exact ray theory for varying the period T of the phase velocity maps ($l_{\text{max}}=12$, $\epsilon=1.0$). Each plot is of the exact ray theory predictions vs the linearized ray theory predictions; left column is for amplitude anomalies, right is for phase anomalies. The period T increases from $T=40\text{s}$ at the top to $T=150\text{s}$ at the bottom. The correlation increases as T increases. Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays. See Section 7.2.

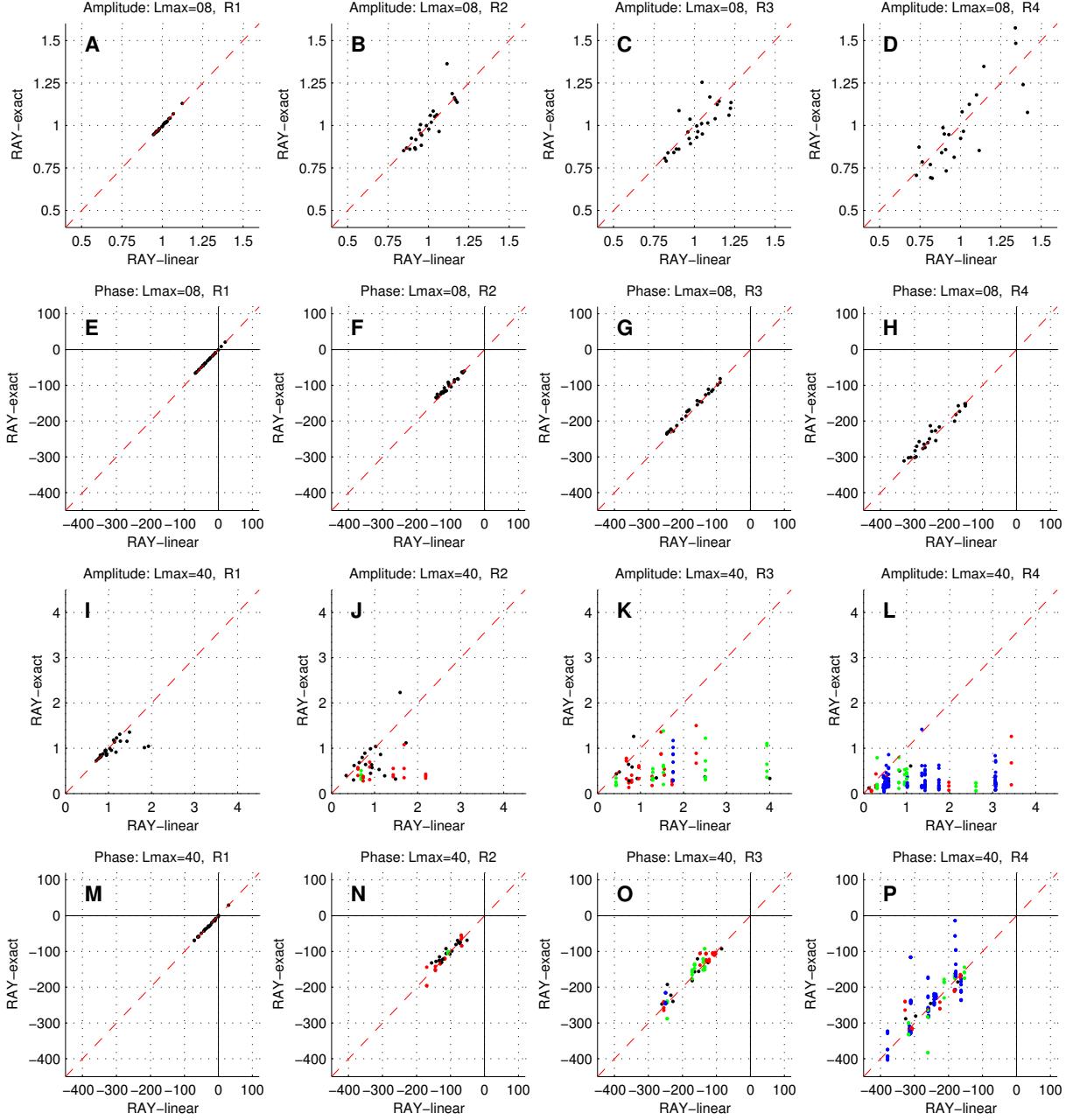


Figure 7.5: Vary orbit (and l_{max}). Comparison of amplitude anomalies (rows 1 and 3) and phase anomalies (rows 2 and 4), as predicted from exact ray theory (RAY-exact) and from linearized ray theory (RAY-linear). Each column shows the results for a particular orbit, R1 (left column) to R4. Each phase velocity map is $T=40\text{s}$ ($\epsilon=1.0$) for two l_{max} values. Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays, green = 5 rays, blue = ≥ 7 rays. Each scatterplot contains a minimum of 25 points. **A-H.** Variation with orbit for $l_{\text{max}}=8$ for amplitude anomalies (A-D) and for phase anomalies (E-H). **I-P.** Same as A-H, only for $l_{\text{max}}=40$. See Section 7.2.

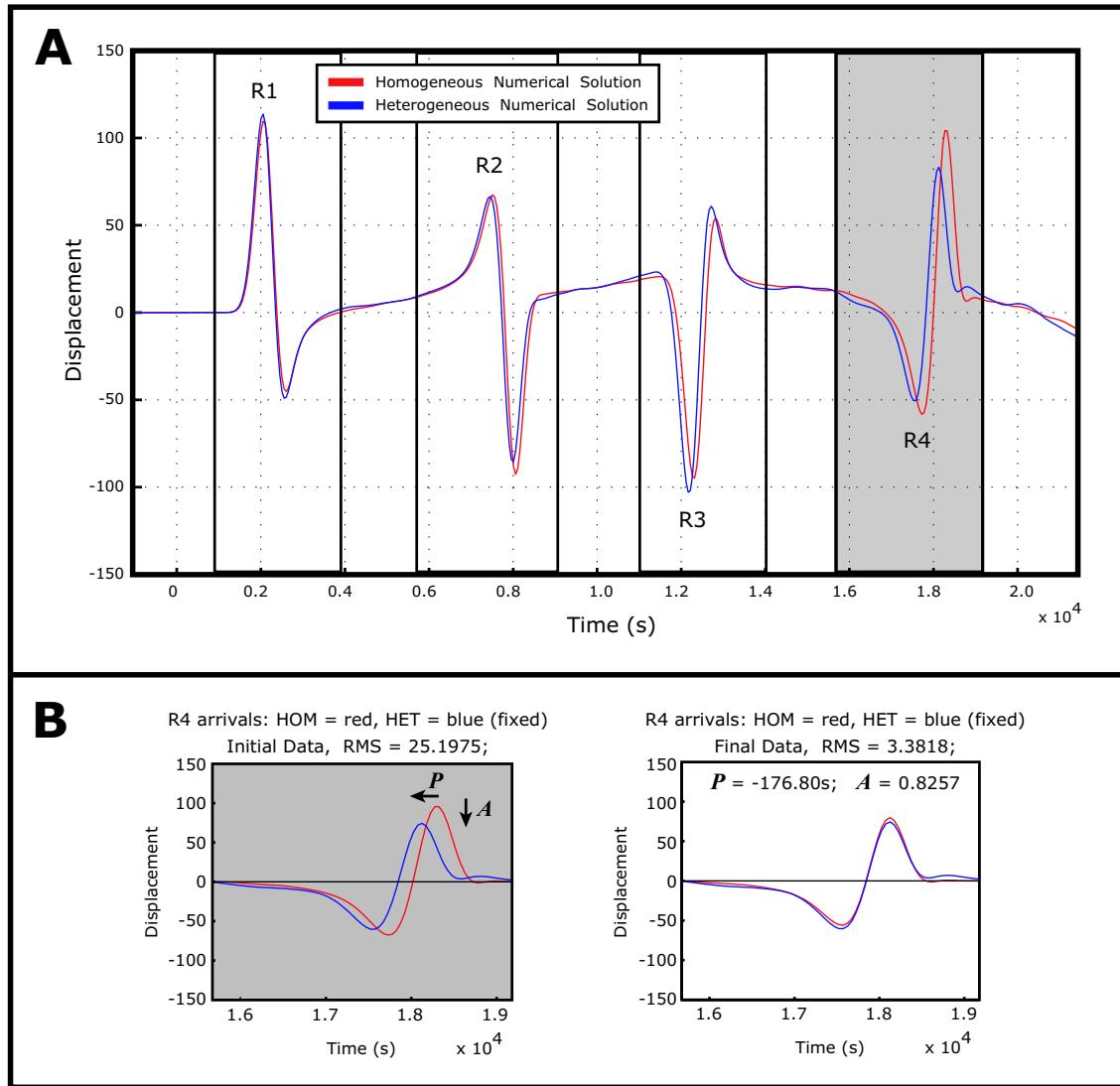


Figure 7.6: Obtaining the amplitude and phase anomalies from the numerical model. **A.** Seismogram at a receiver approximately 90° from the source. Boxes show the time intervals picked to designate the impulse associated with each orbit. In this example, the extremely-long period wave propagated through the $T=40\text{s}$, $l_{max}=8$ ($\epsilon=1.0$) phase velocity map. **B.** Method for determining the phase anomaly and amplitude anomaly in the numerical model, shown here for the R4 arrival in A. A non-linear least-squares program calculates the phase shift P (seconds) and amplitude change A that best fits the homogeneous earth seismogram (red) to the heterogeneous earth seismogram (blue). Minor discrepancy remains between the fitted seismogram and the heterogeneous numerical solution, as the two parameters P and A do not completely describe the heterogeneous numerical solution (see Section 7.3).

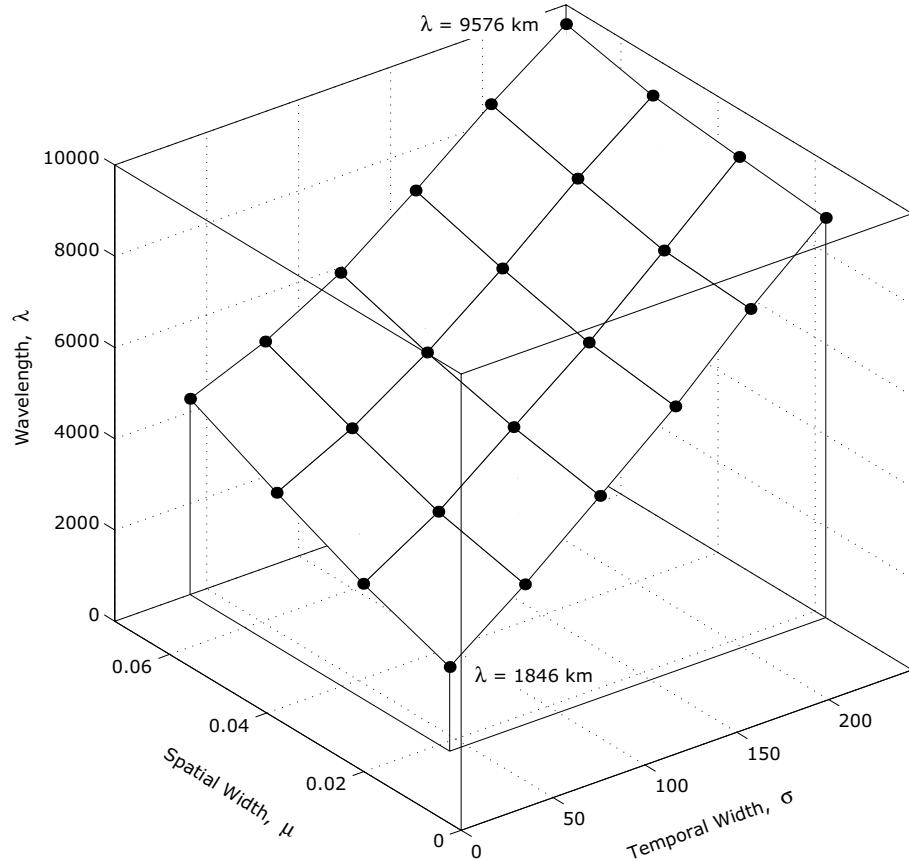


Figure 7.7: Wavelength $\lambda(\mu, \sigma)$ as a function of spatial source width μ and temporal source width σ . Each point represents the wavelength of the wave pulse when it reaches the equator, having originated at the North Pole. It is calculated by taking the Fourier transform of the pulse for the analytical solution $u_{f_2}(\theta, \phi, t)$ (Eq. 3.46); the λ corresponds to the frequency at the maximum of the frequency spectrum. In the point source limit, when $\mu \rightarrow 0$ and $\sigma \rightarrow 0$, we see that $\lambda \rightarrow 0$, as we would expect. The λ values shown here cover the range of wavelengths used in the numerical calculations. See Section 7.4.

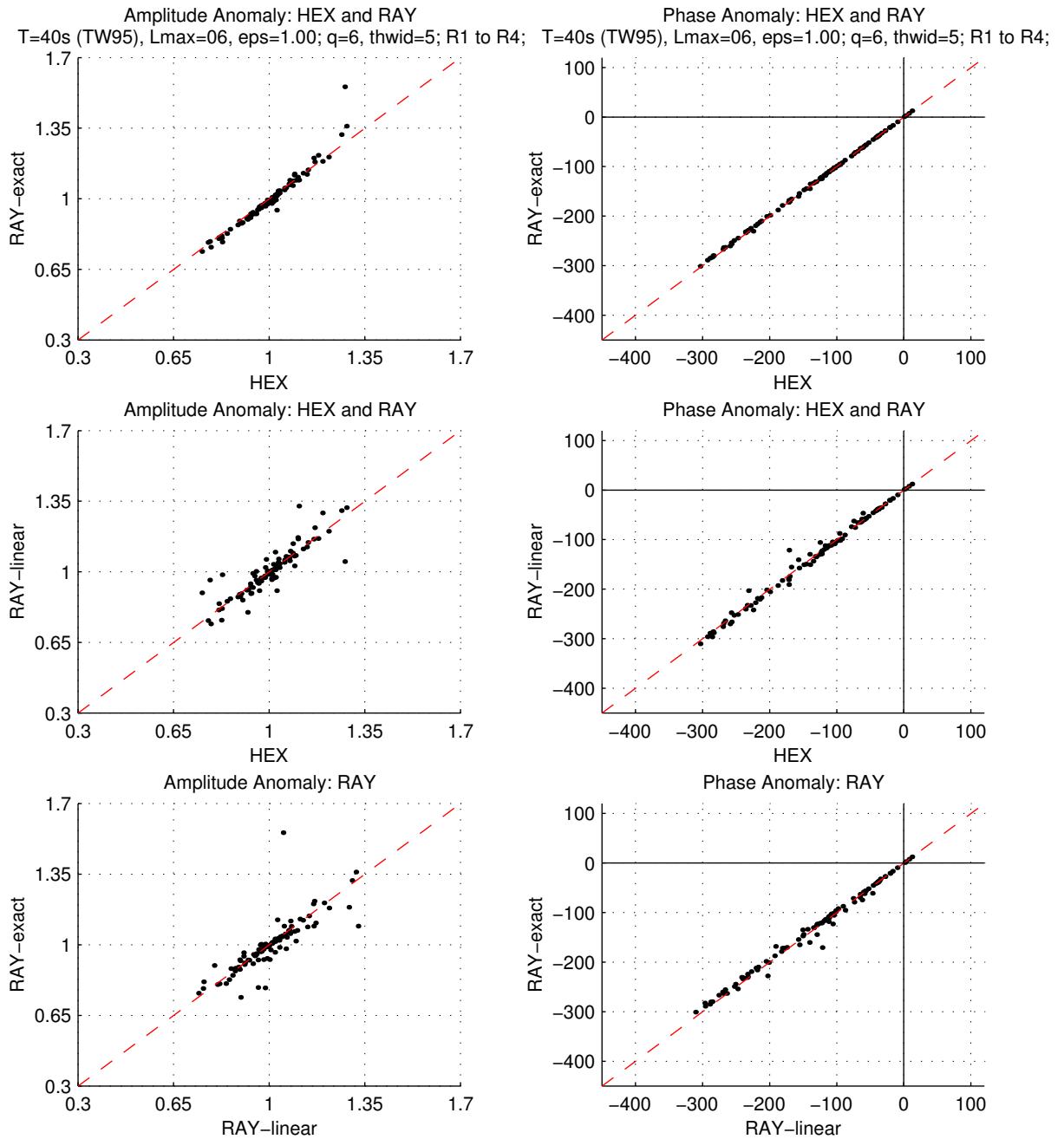


Figure 7.8: Comparison of phase and amplitude anomalies, as predicted from the numerical model (HEX) and from ray tracing, including both exact ray theory (RAY-exact) and linearized ray theory (RAY-linear). Each plot includes calculations carried out using an $l_{\text{max}}=6$ phase velocity map ($T=40\text{s}$, $\epsilon=1.0$). Note that the exact ray theory results are a better predictor of amplitude anomalies and phase anomalies than linearized ray theory. See Section 7.5.

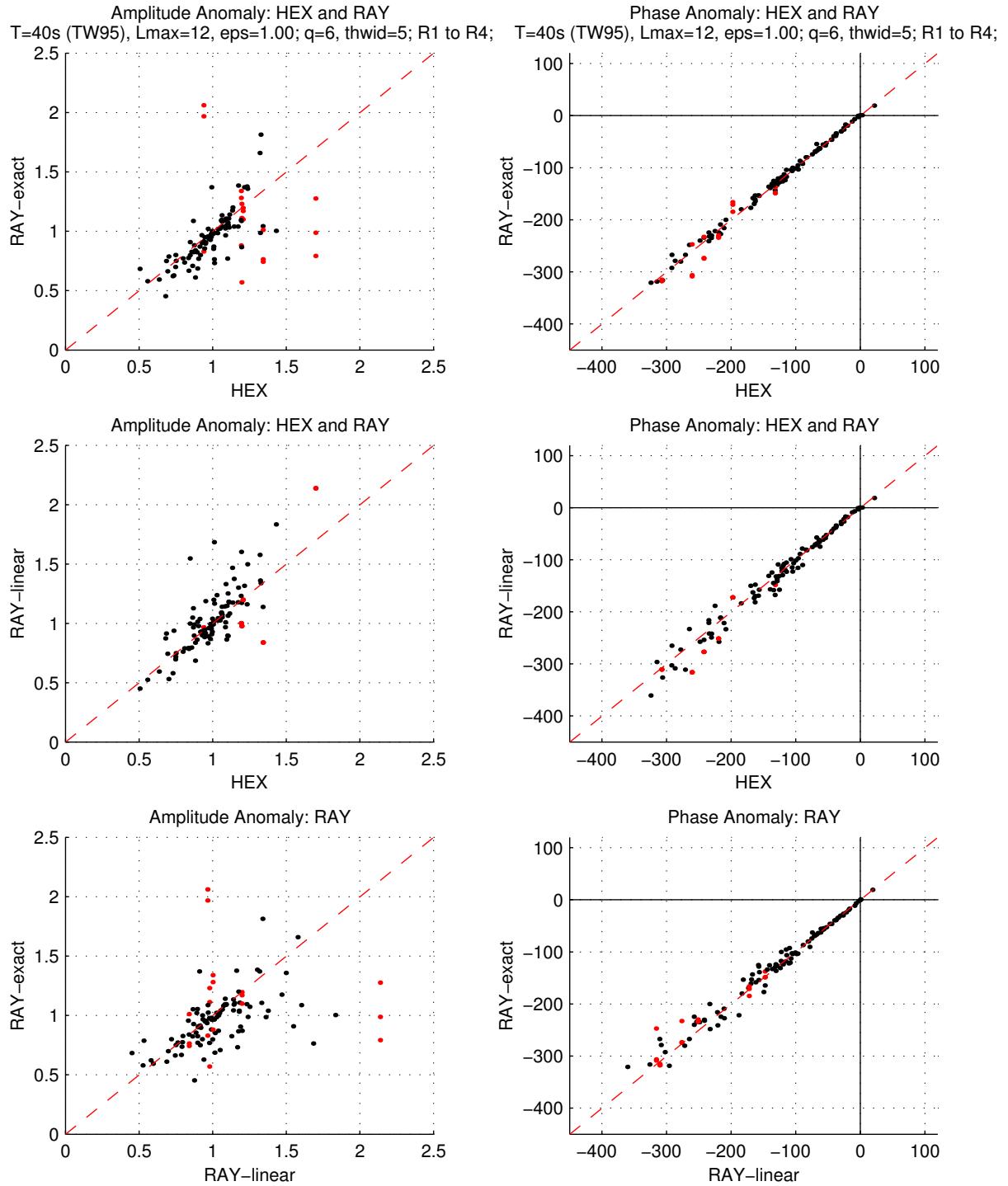


Figure 7.9: Same as Figure 7.8, only for an $l_{\text{max}}=12$ phase velocity map ($T=40\text{s}$, $\epsilon=1.0$). Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays. See Section 7.5.

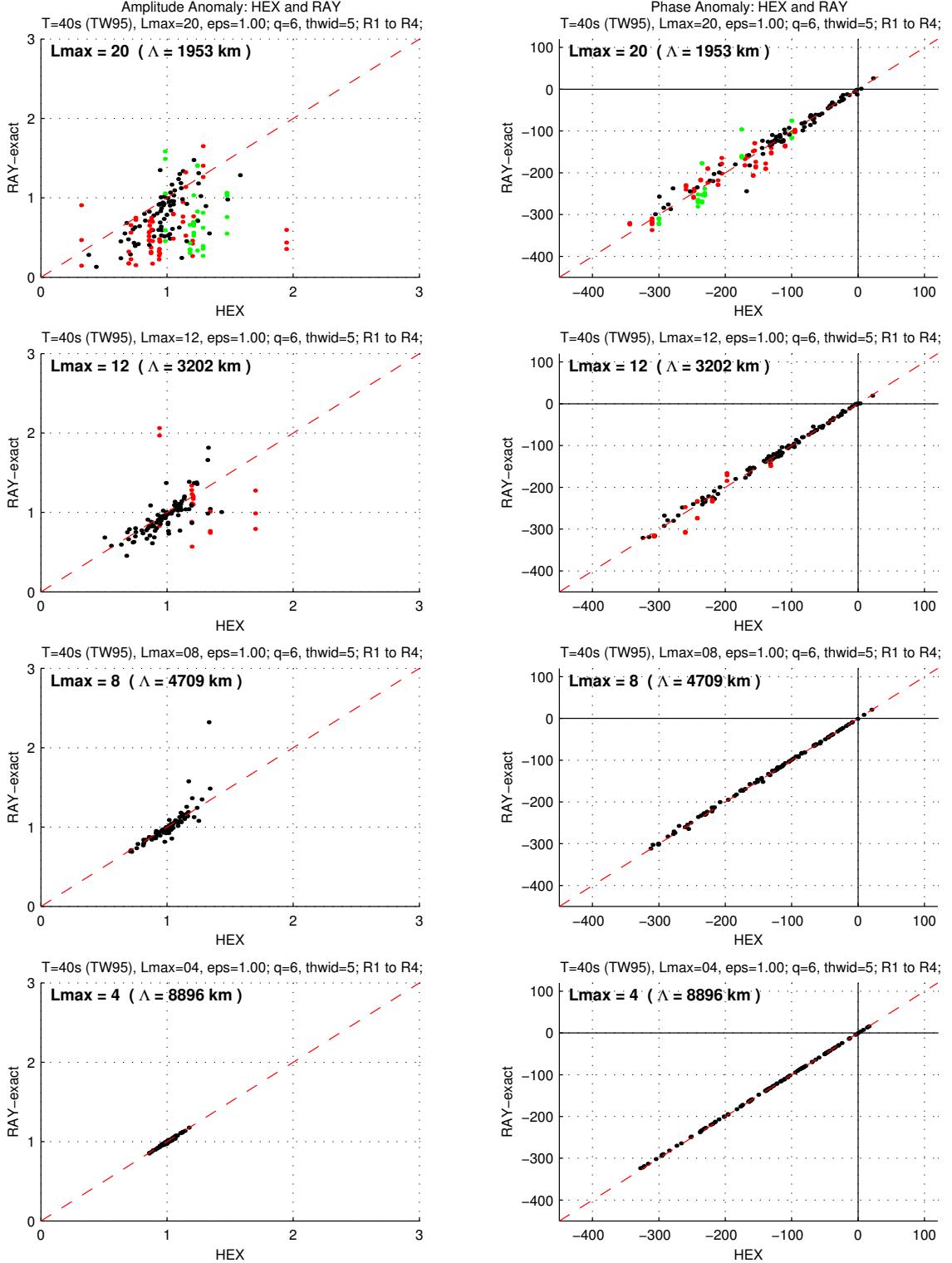


Figure 7.10: Vary l_{max} . Comparison of amplitude anomalies (left column) and phase anomalies (right column), as predicted from the numerical model (HEX) and from exact ray tracing (RAY-exact). The l_{max} value of each phase velocity map varies from $l_{\text{max}}=20$ at the top to $l_{\text{max}}=4$ at the bottom ($T=40\text{s}$, $\epsilon=1.0$), and the corresponding scalelength (Λ) is shown. The input wave into the numerical model remains fixed ($\lambda \approx 2050 \text{ km}$), and the correlation worsens with increasing l_{max} . Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays, green = 5 rays. See Section 7.5.

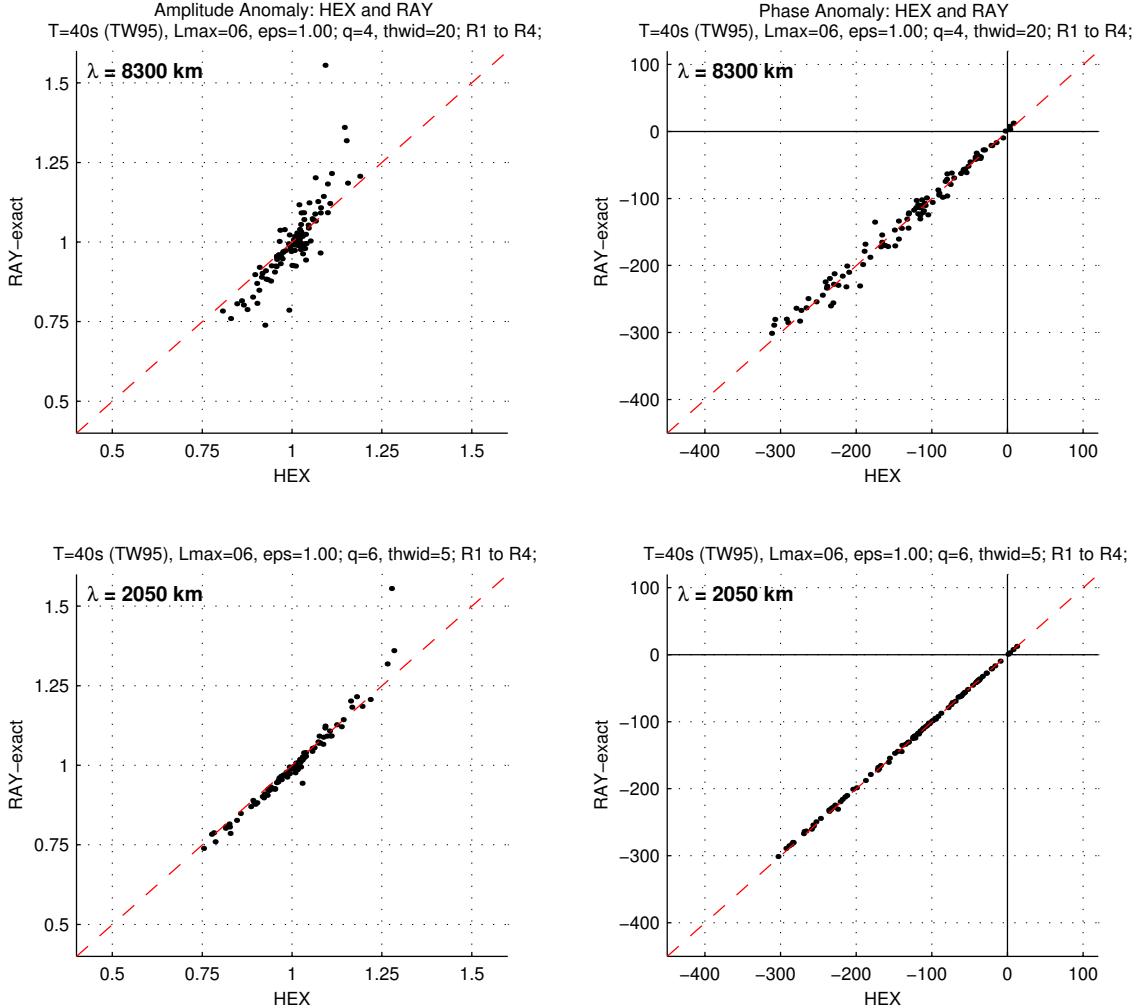


Figure 7.11: Vary λ . Comparison of amplitude anomalies (left column) and phase anomalies (right column), as predicted from the numerical model (HEX) and from exact ray tracing (RAY-exact). The input wavelength into the numerical model is $\lambda \approx 8300 \text{ km}$ at the top (order $q=4$ numerical grids) and $\lambda \approx 2050 \text{ km}$ at the bottom (order $q=6$ grids). The phase velocity map is the same: the $l_{\text{max}}=6$ map ($T=40\text{s}$, $\epsilon=1.0$), which contains a minimum scalelength of $\Lambda \approx 6158 \text{ km}$. As we go to finer grids and use shorter wavelength waves in the numerical model, the correlation between exact ray theory and the numerical results increases. No multipathing occurs, due to the low $l_{\text{max}}=6$ of the map.

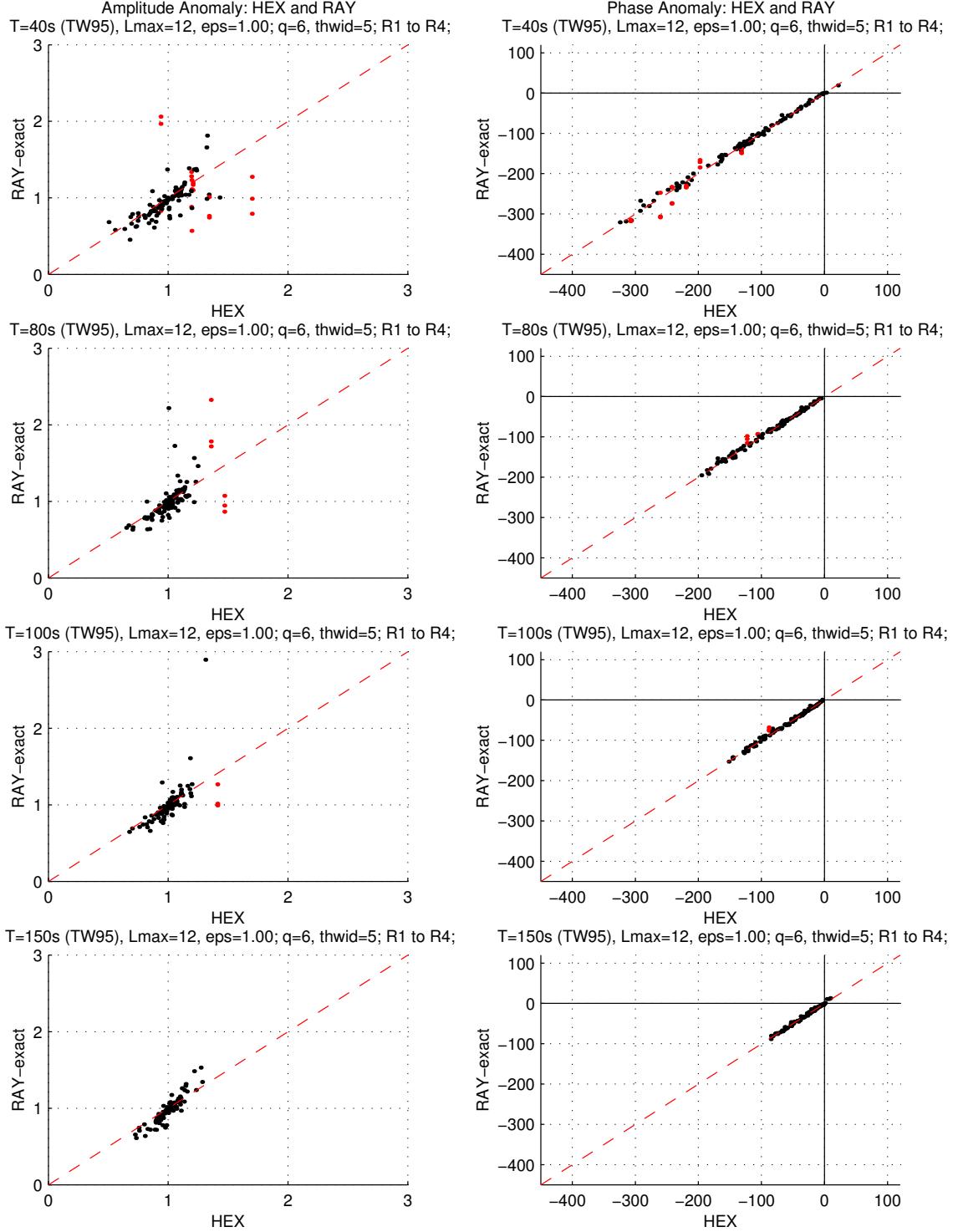


Figure 7.12: Vary T . Comparison of amplitude anomalies (left column) and phase anomalies (right column), as predicted from the numerical model (HEX) and from exact ray tracing (RAY-exact). The *period* of each phase velocity map varies from $T=40\text{s}$ at the top to $T=150\text{s}$ at the bottom ($l_{\max}=12$, $\epsilon=1.0$). The input wave into the numerical model remains fixed ($\lambda \approx 2050 \text{ km}$), and the correlation worsens with decreasing T . Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays. See Section 7.5.

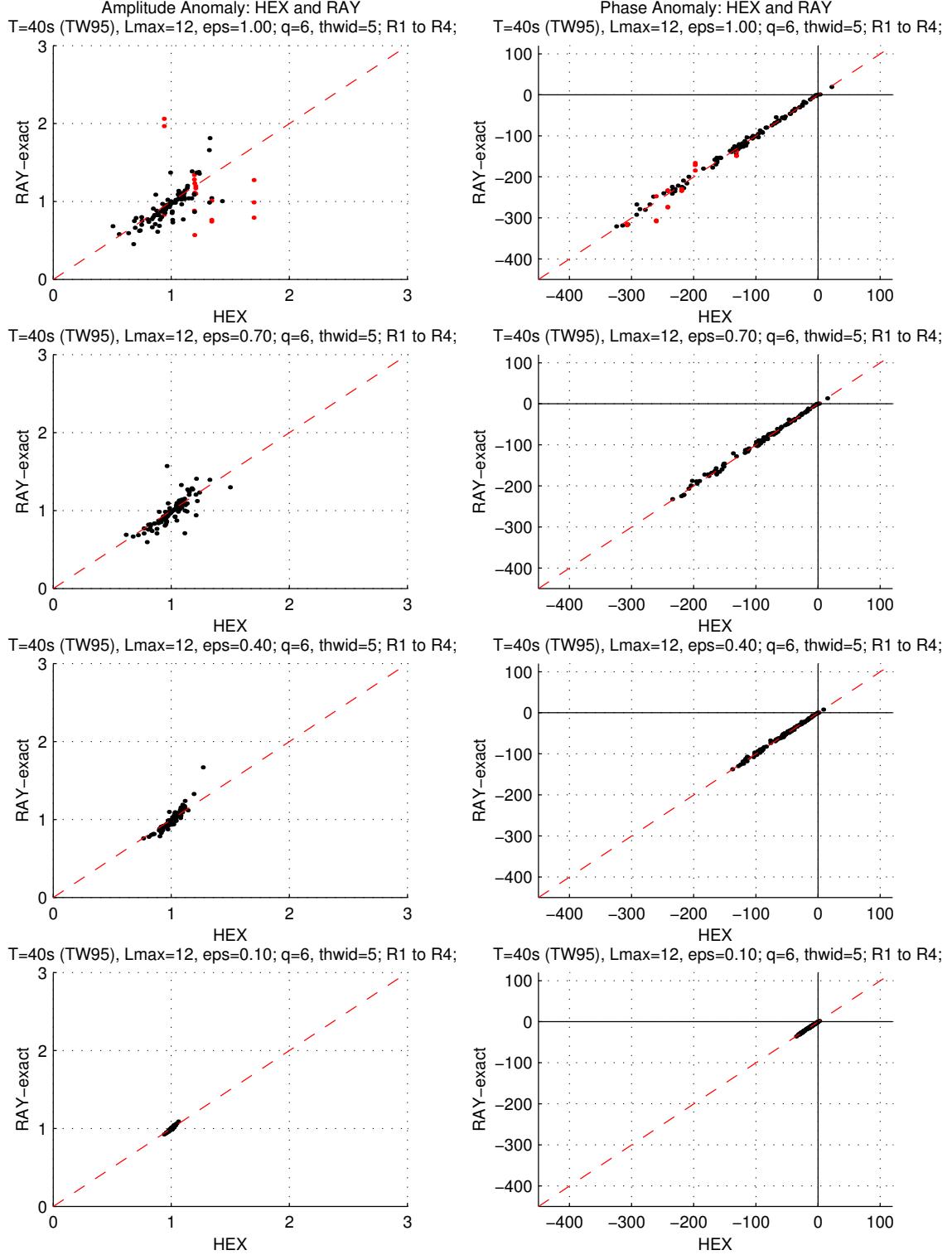


Figure 7.13: Vary ϵ . Comparison of amplitude anomalies (left column) and phase anomalies (right column), as predicted from the numerical model (HEX) and from exact ray tracing (RAY-exact). The ϵ value of each phase velocity map varies from $\epsilon=1.0$ at the top to $\epsilon=0.1$ at the bottom ($l_{\text{max}}=12$, $T=40\text{s}$). The input wave into the numerical model remains fixed ($\lambda \approx 2050 \text{ km}$), and the correlation worsens with increasing ϵ . Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays. See Section 7.5.

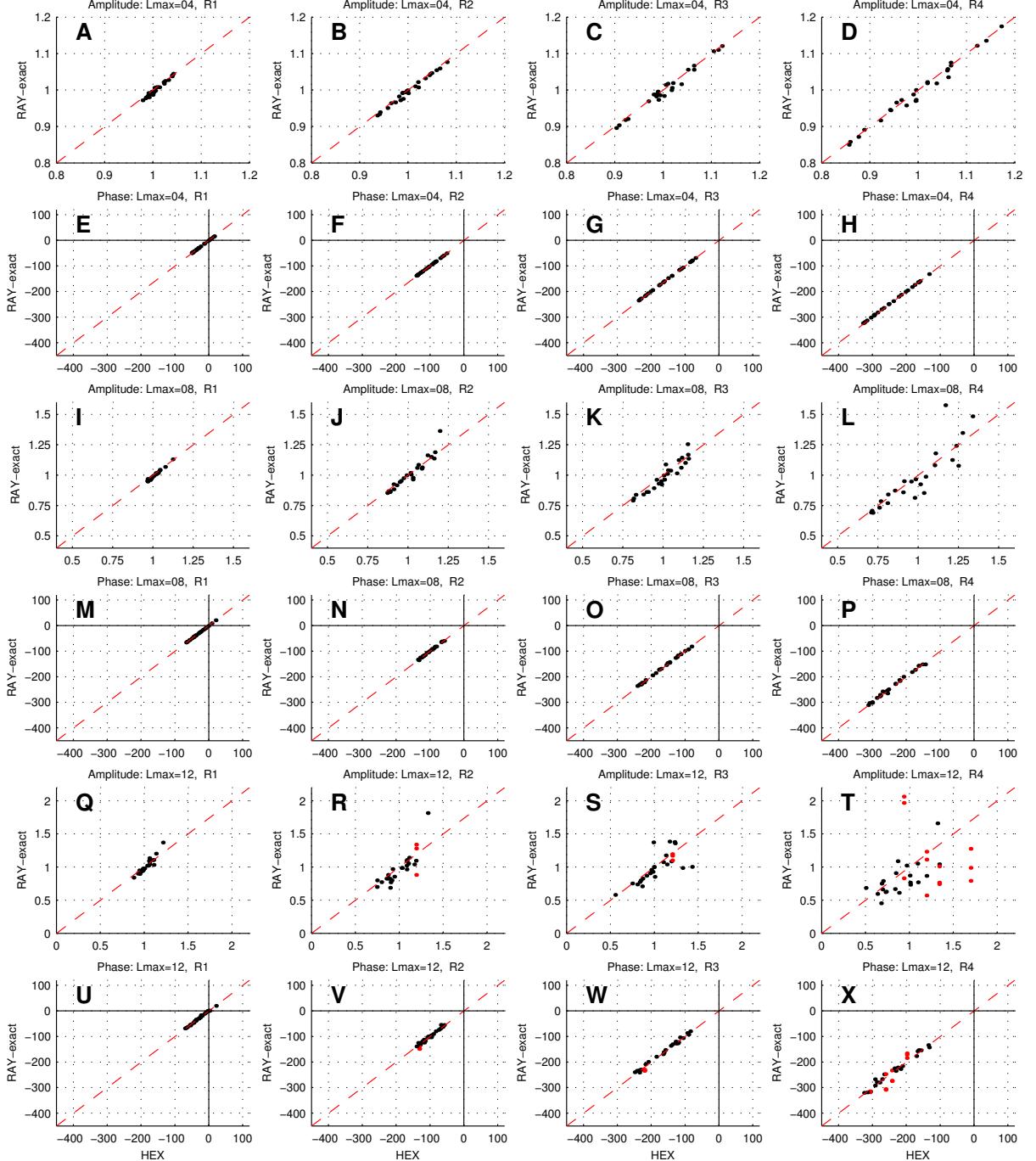


Figure 7.14: Vary orbit (and l_{max}). Comparison of amplitude anomalies (rows 1, 3, 5) and phase anomalies (rows 2, 4, 6), as predicted from the numerical model (HEX) and from exact ray tracing (RAY-exact). Each column shows the results for a particular orbit, R1 (left column) to R4. Each phase velocity map is $T=40\text{s}$ ($\epsilon=1.0$) for three l_{max} values (4, 8, 12). The input wave into the numerical model remains fixed ($\lambda \approx 2050 \text{ km}$). Note the multipath rays for the exact ray tracing: black = 1 ray, red = 3 rays. **A-H.** Variation with orbit for $l_{\text{max}}=4$ for amplitude anomalies (A-D) and for phase anomalies (E-H). **I-P.** Same as A-H, only for $l_{\text{max}}=8$. **Q-X.** Same as I-P, only for $l_{\text{max}}=12$. See Section 7.5.

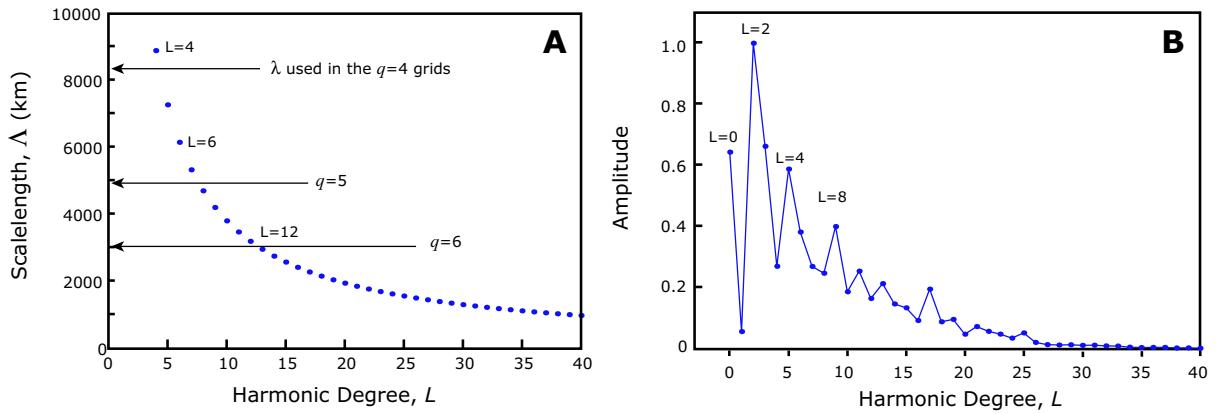


Figure 7.15: The relationship between the detail of a phase velocity map and the wavelength of the probing surface waves. **A.** Scalelength Λ as a function of spherical harmonic degree l : $\Lambda = 2\pi a/(l + \frac{1}{2})$, a is the earth's radius (Appendix F). A phase velocity map of degree $l_{max}=40$ may have structure with scalelength as short as 1000 km; a map of degree $l_{max}=4$ may have structure with scalelength as short as 9000 km (Table 5.2). Also shown are the approximate wavelengths of the waves in our numerical model, where q indicates the order of the spherical grid (Figure 7.7). This is shown in order to facilitate a discussion in terms of the ray theory validity condition, $\lambda \ll \Lambda$ (Sections 7.4 and 7.5). **B.** Amplitude spectrum of the $T=40$ s phase velocity map in Figure 5.2A ($l_{max}=40$, $\epsilon=1.0$). The power at each degree l is the most important factor, not simply the l_{max} truncation of the map. In this map, most of the structure is within the range $l < 20$.

Appendix A

Functions on the sphere

In this appendix we examine what it means to be a well-defined function on the sphere. When deriving the analytical solutions to the wave equation on the sphere (Chapter 3), we must ensure that all of our functions on the sphere satisfy the invariance conditions presented in Equations (A.11)–(A.14). Spherical harmonic functions, the subject of Appendix B, all satisfy the invariance conditions for the sphere.

A well-defined function on the sphere

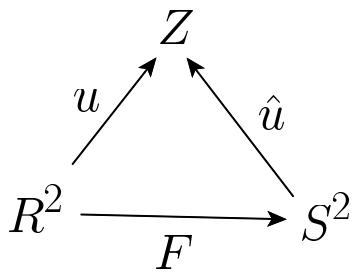


Figure A.1: Function mappings defined in text. S^2 is the set of all points on the unit sphere. R^2 is the set of all points in the ϕ - θ plane. Z is a set of real numbers.

Let $F : R^2 \rightarrow S^2$ be the spherical coordinate mapping, i.e.,

$$F(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (\text{A.1})$$

where θ is the colatitude (polar angle) and ϕ is the azimuthal angle.

DEFINITION. Let Z be any set and let $u : R^2 \rightarrow Z$. Then u is *well-defined on the sphere S^2* if there is a function $\hat{u} : S^2 \rightarrow Z$ such that $u = \hat{u} \circ F$. (It is really

\hat{u} that we care about.) These mappings are shown schematically in Figure A.1.

THEOREM. u is well-defined on S^2 iff

$$F(x_1) = F(x_2) \text{ implies } u(x_1) = u(x_2) \quad (x_1, x_2 \in R^2). \quad (\text{A.2})$$

IDEA: Try to define \hat{u} . Let $y \in S^2$. Then $y = F(x_1)$ for some $x_1 \in R^2$. “Define” $\hat{u}(y) = u(x_1)$. But what if also $y = F(x_2)$? If \hat{u} is to be a function, then you need $u(x_1) = u(x_2)$. In other words, Equation (A.2) is exactly the condition that you need.

Basic lemma from trigonometry:

$$r_1 \cos \phi_1 = r_2 \cos \phi_2 \quad (\text{A.3})$$

$$r_1 \sin \phi_1 = r_2 \sin \phi_2 \quad (\text{A.4})$$

if and only if one of the three conditions is met:

$$(1) \quad r_1 = r_2 = 0 \quad (\text{A.5})$$

$$(2) \quad r_1 = r_2 \quad \text{and} \quad \phi_2 = \phi_1 + 2n\pi \quad (\text{A.6})$$

$$(3) \quad r_1 = -r_2 \quad \text{and} \quad \phi_2 = \phi_1 + \pi + 2n\pi . \quad (\text{A.7})$$

Now it is just a matter of using this lemma to see when $F(x_1) = F(x_2)$. Let us suppose that $F(\theta_1, \phi_1) = F(\theta_2, \phi_2)$. Using Equation (A.1), we have

$$\sin \theta_1 \cos \phi_1 = \sin \theta_2 \cos \phi_2 \quad (\text{A.8})$$

$$\sin \theta_1 \sin \phi_1 = \sin \theta_2 \sin \phi_2 \quad (\text{A.9})$$

$$\cos \theta_1 = \cos \theta_2 . \quad (\text{A.10})$$

This is our basic lemma (Eqs. A.3–A.4), with $r_1 = \sin \theta_1$ and $r_2 = \sin \theta_2$. It is true if and only if at least one of the three conditions (Eqs. A.5–A.7) is met.

Equation (A.5) gives us

$$(1) \quad \sin \theta_1 = \sin \theta_2 = 0 \quad (\text{and} \quad \cos \theta_1 = \cos \theta_2)$$

$$\Rightarrow \theta_1 = 2n\pi \quad \text{and} \quad \theta_2 = 2m\pi$$

$$\text{or} \quad \theta_1 = \pi + 2n\pi \quad \text{and} \quad \theta_2 = \pi + 2m\pi .$$

Equation (A.6) gives us

$$(2) \quad \sin \theta_1 = \sin \theta_2 \quad \text{and} \quad \phi_2 = \phi_1 + 2n\pi$$

$$\Rightarrow \theta_2 = \theta_1 + 2m\pi .$$

Equation (A.7) gives us

$$\begin{aligned}
 (3) \quad & \sin \theta_1 = -\sin \theta_2 \quad \text{and} \quad \phi_2 = \phi_1 + \pi + 2n\pi \\
 \rightarrow \quad & \sin \theta_1 = \sin(-\theta_2) \\
 \rightarrow \quad & \cos \theta_1 = \cos \theta_2 = \cos(-\theta_2) \\
 \rightarrow \quad & \theta_1 = -\theta_2 + 2m\pi \\
 \Rightarrow \quad & \theta_2 = -\theta_1 + 2m\pi .
 \end{aligned}$$

Putting this all together, we can state the following:

A function u is well-defined on the sphere if and only if

$$(1) \quad u(-\theta, \phi + \pi) = u(\theta, \phi) \tag{A.11}$$

$$(2) \quad u(\theta + 2\pi, \phi) = u(\theta, \phi) \tag{A.12}$$

$$(3) \quad u(0, \phi) = \text{const}_1 \equiv u_{\text{NP}} \tag{A.13}$$

$$(4) \quad u(\pi, \phi) = \text{const}_2 \equiv u_{\text{SP}} . \tag{A.14}$$

Equations (A.11)–(A.14) are the four invariance conditions.

Extending a function to the ϕ - θ plane

Let us now say we have a function $\tilde{g}(\theta, \phi)$ defined only in the domain $R^2 \in D : 0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$, whose function values there are $Z = \tilde{g}(D)$. If $\tilde{g}(\theta, \phi)$ meets the four invariance conditions¹ within D (Eqs. A.11–A.14), then we can extend this function to all ϕ - θ points (R^2) by applying two fundamental operations γ and τ , which can be seen upon examination of Figure A.2:

$$\begin{aligned}
 \gamma(\theta, \phi) &= (-\theta, \phi + \pi) \Rightarrow \gamma(-\theta, \phi + \pi) = (\theta, \phi + 2\pi) \\
 \tau(\theta, \phi) &= (\theta + 2\pi, \phi) ,
 \end{aligned} \tag{A.15}$$

where γ represents a glide reflection and τ represents a translation. In the ϕ -direction, we have a glide reflection of $\phi = \pi$, where π is the translational part. Two successive applications of the glide reflection (' \Rightarrow ') equates to one simple translation of $\phi = 2\pi$ in the ϕ -direction. In

¹If we are within D , we can speak of the invariance conditions as the boundary conditions.

the θ -direction, we have a simple translation of $\theta = 2\pi$. Figures A.3 and A.4 show examples of extending a function from D to the entire ϕ - θ plane. In Figure A.3 the function does *not* meet the invariance conditions, and it is only shown to help illustrate the mapping. In Figure A.4 the function does meet the invariance condition.

Here is a basic algorithm for extending a function $\tilde{g}(\theta, \phi)$ to the ϕ - θ plane:

1. Begin with a function defined on $D : 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$.
2. Ensure that the function meets the boundary conditions on D , which can be determined from the invariance conditions (Eqs. A.11–A.14).
3. Extend the function to the ϕ - θ plane using the two operations τ and γ (Eq. A.15).
4. The new function $g(\theta, \phi)$, defined for the entire ϕ - θ plane, will satisfy the invariance conditions, and thus will be a well-defined function on the sphere.

If we begin with a function $\tilde{g}(\theta, \phi)$ and obtain $g(\theta, \phi)$ in this manner, we cannot be sure that the derivatives of $g(\theta, \phi)$ are continuous at the boundaries between the sub-domains of the ϕ - θ plane. In order to meet this requirement, we could set boundary conditions on the derivatives in step 2. However, the simplest thing is to start with a function $g(\theta, \phi)$ that is defined in the entire ϕ - θ plane and that meets the invariance conditions.

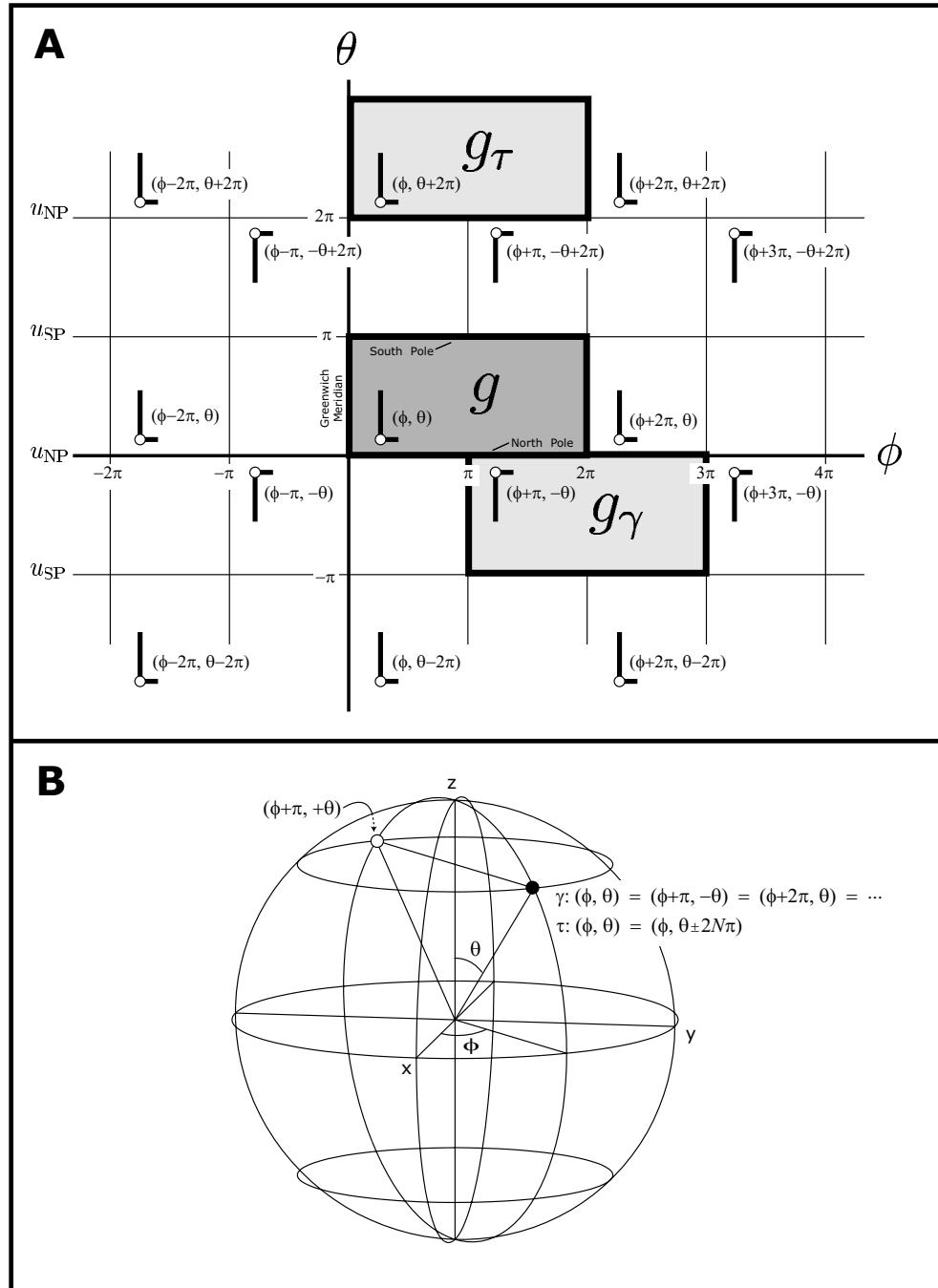


Figure A.2: **A.** Extending the function $\tilde{g}(\theta, \phi)$ to the entire ϕ - θ plane (R^2). The two basic operations are given by Eq. A.15. The first operation for each case gives g_γ and g_τ ; repeating these operations will cover the entire plane. The L-shape is shown to assist in illustrating the transformations. The ϕ -axis can be thought of as the North Pole; the θ -axis can be thought of as the Greenwich Meridian. u_{SP} and u_{NP} are the values of u at the poles; they are constant for a fixed time t . **B.** Spherical coordinate system (S^2), illustrating the glide reflection $\gamma(\theta, \phi) = (-\theta, \phi + \pi)$ and the translation $\tau(\theta, \phi) = (\theta \pm 2N\pi, \phi)$.

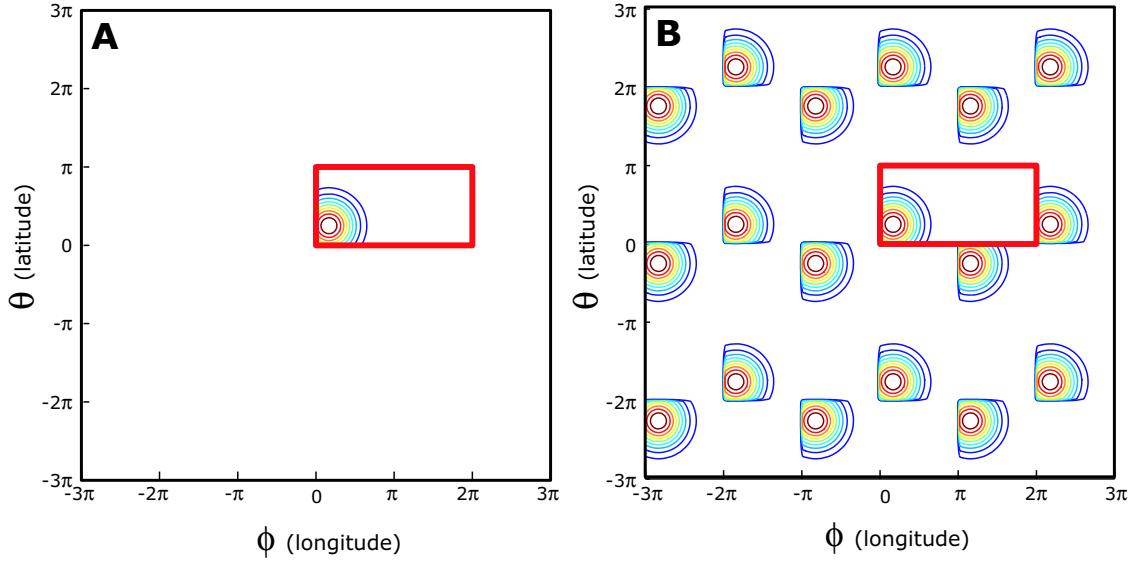


Figure A.3: **A.** Example function $\tilde{g}(\theta, \phi) = e^{-((\phi-\phi_0)^2+(\theta-\theta_0)^2)/\mu^2}$, defined in domain D of the ϕ - θ plane. **B.** Our new function $g(\theta, \phi)$, found by extending $\tilde{g}(\theta, \phi)$ to the ϕ - θ plane using the method in Figure A.2A. In this example, $\tilde{g}(\theta, \phi)$ does not satisfy the invariance conditions given by Eqs. A.11–A.14: here the North Pole and South Pole are multiple-valued.

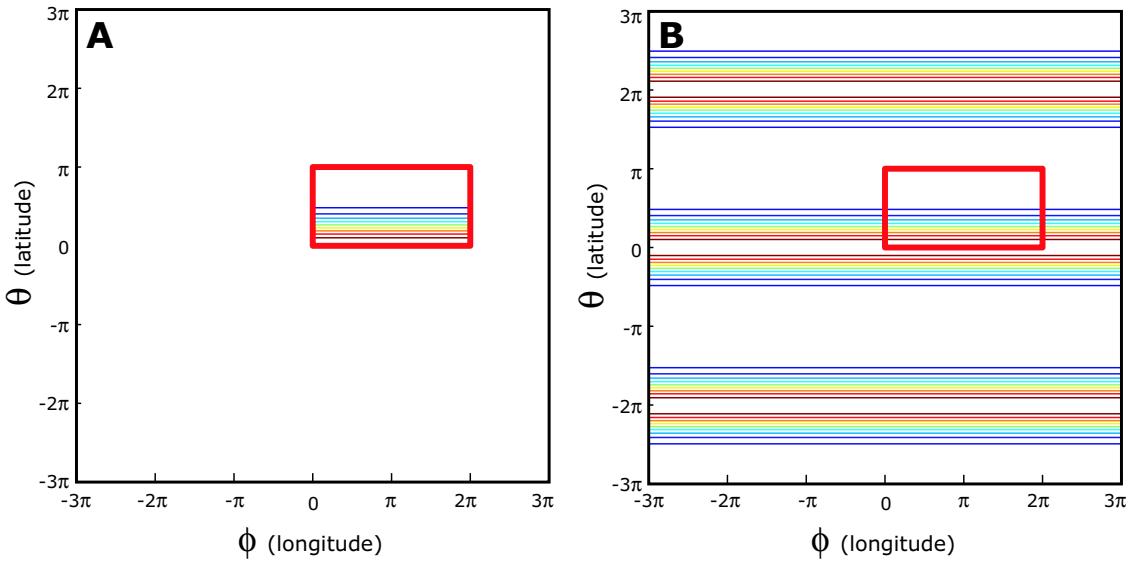


Figure A.4: **A.** The function $\tilde{g}(\theta, \phi) = e^{-\theta^2/\mu^2}$, defined in domain D of the ϕ - θ plane. **B.** Our new function $g(\theta, \phi)$, after extending $u(\theta, \phi)$ to the ϕ - θ plane using the method in Figure A.2A. Here $\tilde{g}(\theta, \phi)$ satisfies the invariance conditions given by Eqs. A.11–A.14.

Appendix B

Spherical harmonic functions

Spherical harmonic functions, $Y_{lm}(\theta, \phi)$, are smooth, well-defined functions on the sphere (see Appendix A). Spherical harmonic conventions have been presented and discussed in *Edmonds* (1960); *Phinney and Burridge* (1973); *Gilbert and Dziewonski* (1975); *Dziewonski and Woodhouse* (1983); *Dahlen and Tromp* (1998), among others. The purpose of this appendix is to present the standard conventions used in seismology — those of *Gilbert and Dziewonski* (1975) — and then to show how Matlab’s built-in functions can be modified to match the conventions used in seismology. Then we look at expanding functions in terms of spherical harmonic functions.

B.1 Spherical harmonic conventions

Adopting the convention of *Gilbert and Dziewonski* (1975), we define *complex surface spherical harmonics* of degree l and order m as

$$Y_{lm}(\theta, \phi) = X_{lm}(\theta) e^{im\phi} \quad (\text{B.1})$$

$$= X_{lm}(\theta) \cos m\phi + i X_{lm}(\theta) \sin m\phi , \quad (\text{B.2})$$

where $X_{lm}(\theta)$ are the *generalized Legendre functions*, defined as

$$X_{lm}(\theta) = \begin{cases} (-1)^m \left(\frac{2l+1}{4\pi}\right)^{1/2} \left[\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_{lm}(\cos \theta) & \text{if } 0 \leq m \leq l \\ X_{l(-m)}(\theta) & \text{if } -l \leq m \leq 0 , \end{cases} \quad (\text{B.3})$$

or, equivalently,

$$X_{lm}(\theta) = (-1)^{|m|} \left(\frac{2l+1}{4\pi} \right)^{1/2} \left[\frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} P_{l|m|}(\cos \theta) \quad (-l \leq m \leq l) . \quad (\text{B.4})$$

where we have used the definition of *Dziewonski and Woodhouse* (1983) to explicitly show the case for $m < 0$. For the special case of the spherically-symmetric function with $l=0$ and $m=0$, we have $X_{00}(\theta) = (4\pi)^{-1/2}$.

The *associated Legendre functions* of degree l and order m are defined as

$$P_{lm}(\mu) = \frac{1}{2^l l!} (1 - \mu^2)^{m/2} \left(\frac{d}{d\mu} \right)^{l+m} (\mu^2 - 1)^l \quad (m \geq 0) \quad (\text{B.5})$$

where $\mu = \cos \theta$. (With $\mu = \cos \theta$, we always meet the required condition $-1 \leq \mu \leq 1$.) The *Legendre polynomial of degree l* is given by the associated Legendre function of order $m = 0$:

$$P_l(\mu) = \frac{1}{2^l l!} \left(\frac{d}{d\mu} \right)^l (\mu^2 - 1)^l . \quad (\text{B.6})$$

The associated Legendre functions (Eq. B.5) of positive order, $m > 0$, can be written in terms of the Legendre polynomials $P_l(\mu)$:

$$P_{lm}(\mu) = (1 - \mu^2)^{m/2} \left(\frac{d}{d\mu} \right)^m P_l(\mu) \quad (m > 0) \quad (\text{B.7})$$

Dahlen and Tromp (1998) define *real surface spherical harmonics* by

$$\mathcal{Y}_{lm}(\theta, \phi) = \begin{cases} \sqrt{2} X_{l|m|}(\theta) \cos m\phi & \text{if } -l \leq m < 0 \\ X_{l0}(\theta) & \text{if } m = 0 \\ \sqrt{2} X_{lm}(\theta) \sin m\phi & \text{if } 0 < m \leq l . \end{cases} \quad (\text{B.8})$$

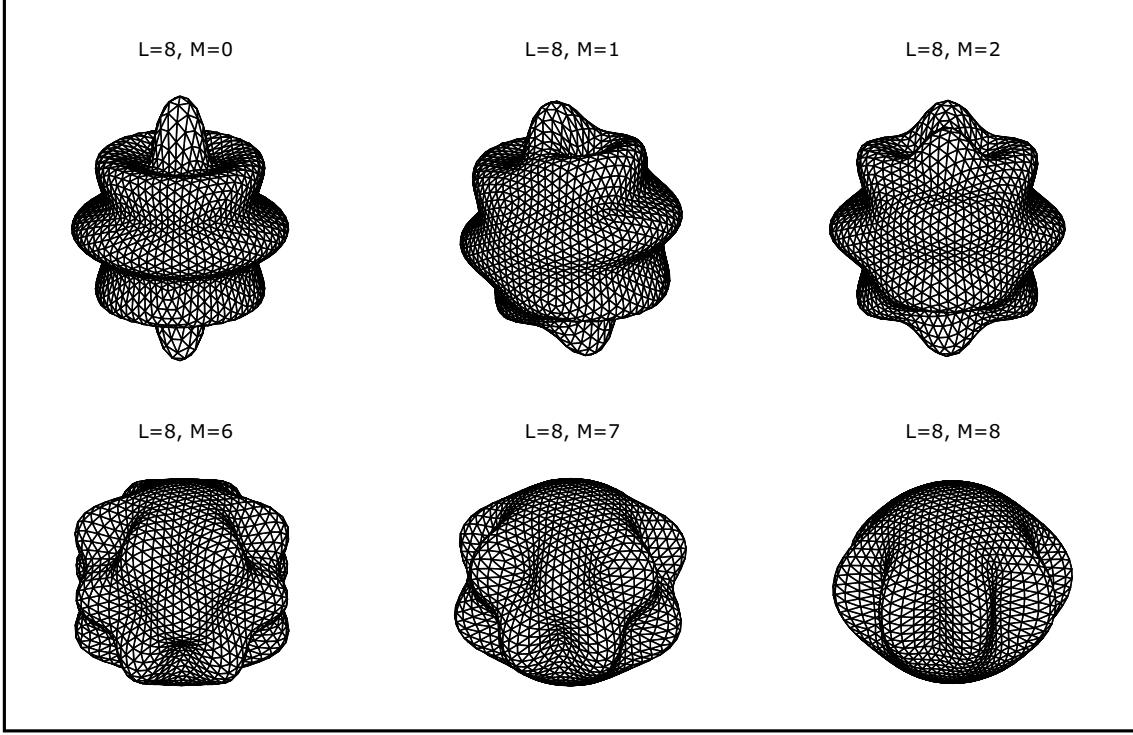


Figure B.1: Sample functions of $Y_{lm}(\theta, \phi)$. The functions are plots of Eq. B.14 for $l = 8$ with varying m . The functions are plotted on the order $q=3$ triangular grids.

B.2 Spherical harmonic conventions in Matlab

Matlab has the built-in function `legendre`. Two types of commands are possible. First, $P_{lm}(\mu) = \text{legendre}(l, \mu)$ are the associated Legendre functions of degree l and order $m = 0, 1, \dots, l$, defined as

$$P_{lm}(\mu) = (-1)^m (1 - \mu^2)^{m/2} \left(\frac{d}{d\mu} \right)^m P_l(\mu) \quad (m \geq 0) \quad (\text{B.9})$$

$$= (-1)^m P_{lm}(\mu) \quad (m \geq 0), \quad (\text{B.10})$$

where $P_{lm}(\mu)$ is given by Equation (B.7). The input $\mu = \cos \theta$ is a vector of N values, and the output is a $(l+1)$ -by- N matrix, with the m th row corresponding to the associated Legendre function of order $m-1$ (and degree l). For example say $l = 5$ and the input vector is 100 elements of $\theta = [0, \pi]$ and $\mu = \cos \theta$. Then the Matlab command `legendre(l, mu)` gives a 6-by-100 matrix of the associated Legendre functions of degree $l = 5$, orders $m = 0$ (first row) to $m = 5$. Second, $X_{lm}(\theta) = \text{legendre}(l, \mu, 'sch')$ is the *Schmidt semi-normalized associated*

Legendre functions, defined as

$$X_{lm}(\theta) = \begin{cases} P_{lm}(\cos \theta) & \text{if } m = 0 \\ (-1)^m \sqrt{2} \left[\frac{(l-m)!}{(l+m)!} \right]^{1/2} P_{lm}(\cos \theta) & \text{if } m > 0 . \end{cases} \quad (\text{B.11})$$

Substituting Equation (B.10) (with $\mu = \cos \theta$) gives

$$X_{lm}(\theta) = \begin{cases} P_{lm}(\cos \theta) & \text{if } m = 0 \\ \sqrt{2} \left[\frac{(l-m)!}{(l+m)!} \right]^{1/2} P_{lm}(\cos \theta) & \text{if } m > 0 , \end{cases} \quad (\text{B.12})$$

where $P_{lm}(\mu)$ is given by Equation (B.7). In order to adopt the convention of *Dziewonski and Woodhouse* (1983), we have to alter the Matlab function, changing Equations (B.12) and (B.4) for $m \geq 0$. We alter the Matlab output for $X_{lm}(\theta)$ by (i) multiplying all values by $A = \sqrt{(2l+1)/(4\pi)}$, (ii) changing the sign of every other order (multiply by $(-1)^m$), and (iii) dividing all non- $(m=0)$ values by $\sqrt{2}$.

$$X_{lm}(\theta) = \begin{cases} A X_{l0}(\theta) & \text{if } m = 0 \\ (-1)^m \frac{A}{\sqrt{2}} X_{lm}(\theta) & \text{if } m \geq 0 . \end{cases} \quad (\text{B.13})$$

Eq. B.13 gives the expression for converting the Matlab function $X_{lm}(\theta) = \text{legendre}(l, \mu, 'sch')$ into the $X_{lm}(\theta)$ functions used by *Dziewonski and Woodhouse* (1983). We use real surface spherical harmonics $Y_{lm}(\theta, \phi)$ to use as test functions for testing the accuracy of our numerical approximation for the Laplacian operator (Section 4.5). We define these functions as

$$Y_{lm}(\theta, \phi) = \begin{cases} X_{l0}(\theta) & \text{if } m = 0 \\ X_{lm}(\theta) \sin m\phi & \text{if } 0 < m \leq l , \end{cases} \quad (\text{B.14})$$

where $X_{lm}(\theta)$ is given by Equation (B.4). Note that for each m there exists a ‘cos $m\phi$ ’ \mathcal{Y}_{lm} function and a ‘sin $m\phi$ ’ \mathcal{Y}_{lm} function (Eq. B.8). We chose to use the test functions only for the case with $m > 0$. Plots of Equation (B.14) for $l = 8$ are shown in Figure B.1.

B.3 Spherical harmonic expansions

Functions on the sphere can be expanded in terms of the spherical harmonic functions, $Y_{lm}(\theta, \phi)$. Phase velocity maps are generally expressed as a set of spherical harmonic coefficients. The spherical harmonic expansion for a phase velocity map $c(\theta, \phi)$ ¹ is written as

$$c(\theta, \phi) = \sum_{l=0}^{l_{max}} \sum_{m=0}^l (A_{lm} \cos m\phi + B_{lm} \sin m\phi) X_{lm}(\theta), \quad (\text{B.15})$$

where A_{lm} and B_{lm} are the spherical harmonic coefficients up to maximum degree $l = l_{max}$, and $X_{lm}(\theta)$ is given by Equation (B.4). In the ray-tracing calculations it is necessary to know the derivatives of $c(\theta, \phi)$:

$$\frac{dc(\theta, \phi)}{d\phi} \equiv d_\phi c = \sum_{l=0}^{l_{max}} \sum_{m=0}^l (-m A_{lm} \sin m\phi + m B_{lm} \cos m\phi) X_{lm}(\theta) \quad (\text{B.16})$$

$$\frac{dc(\theta, \phi)}{d\theta} \equiv d_\theta c = \sum_{l=0}^{l_{max}} \sum_{m=0}^l (A_{lm} \cos m\phi + B_{lm} \sin m\phi) d_\theta [X_{lm}(\theta)] \quad (\text{B.17})$$

$$\frac{d^2 c(\theta, \phi)}{d\theta d\phi} \equiv d_{\theta\phi} c = \sum_{l=0}^{l_{max}} \sum_{m=0}^l (-m A_{lm} \sin m\phi + m B_{lm} \cos m\phi) d_\theta [X_{lm}(\theta)] \quad (\text{B.18})$$

$$\frac{d^2 c(\theta, \phi)}{d\theta^2} \equiv d_{\theta\theta} c = \sum_{l=0}^{l_{max}} \sum_{m=0}^l (A_{lm} \cos m\phi + B_{lm} \sin m\phi) d_{\theta\theta} [X_{lm}(\theta)]. \quad (\text{B.19})$$

Using Equation (B.4) with $m \geq 0$ and

$$k_{lm} = (-1)^m \left(\frac{2l+1}{4\pi} \right)^{1/2} \left[\frac{(l-m)!}{(l+m)!} \right]^{1/2}, \quad (\text{B.20})$$

we find:

$$X_{lm}(\theta) = k_{lm} P_{lm}(\cos \theta) \quad (\text{B.21})$$

$$d_\theta [X_{lm}(\theta)] = k_{lm} d_\theta [P_{lm}(\cos \theta)] \quad (\text{B.22})$$

$$d_{\theta\theta} [X_{lm}(\theta)] = k_{lm} d_{\theta\theta} [P_{lm}(\cos \theta)]. \quad (\text{B.23})$$

What is needed are the derivatives $d_\theta P_{lm}(\cos \theta)$ and $d_{\theta\theta} P_{lm}(\cos \theta)$. The first derivatives

¹Typically the maps are written as $\delta c(\theta, \phi)/c_0$, not $c(\theta, \phi)$.

are calculated using the recursion relation (*Arfken and Weber*, 1995, Eq. 12.87)

$$2(1-x^2)^{1/2} \frac{d}{dx} [P_{lm}(x)] = P_{l(m+1)}(x) - (l+m)(l-m+1)P_{l(m-1)}(x), \quad (\text{B.24})$$

where $x = \cos \theta$. (Eq. B.24 shows why the ray-tracing calculations do not work *exactly* at the poles, where $\theta = 0$ and $\theta = \pi$, and thus $1 - x^2 = 0$.) Equation (B.24) can be written in matrix form:

$$d_x P_{lm}(x) = \frac{1}{2(1-x^2)^{1/2}} \begin{bmatrix} 1 \\ -(l+m)(l-m+1) \end{bmatrix} \begin{bmatrix} P_{l(m+1)}(x) \\ P_{l(m-1)}(x) \end{bmatrix}^\top. \quad (\text{B.25})$$

Making the substitution $x = \cos \theta$, we get:

$$d_x P_{lm}(x) = \frac{dP_{lm}(\cos \theta)}{d(\cos \theta)} = \frac{dP_{lm}(\cos \theta)}{d\theta} \frac{d\theta}{d\cos \theta} = d_\theta P_{lm}(\cos \theta) \left(\frac{1}{-\sin \theta} \right) \quad (\text{B.26})$$

$$d_\theta P_{lm}(\cos \theta) = -\sin \theta d_x P_{lm}(x). \quad (\text{B.27})$$

Using Equation (B.25) with $x = \cos \theta$, we can write Equation (B.27) as:

$$d_\theta P_{lm}(\cos \theta) = -\frac{1}{2} \begin{bmatrix} 1 \\ -(l+m)(l-m+1) \end{bmatrix} \begin{bmatrix} P_{l(m+1)}(\cos \theta) \\ P_{l(m-1)}(\cos \theta) \end{bmatrix}^\top. \quad (\text{B.28})$$

When $m = 0$ for $l \geq 1$ in Equation (B.25), we have to use the relationship (*Arfken and Weber*, 1995, Eq. 12.81a):

$$P_{l(-m)}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_{lm}(x). \quad (\text{B.29})$$

When $m = 0$, Equations (B.25), (B.29), and (B.28) give:

$$d_x P_{l0}(x) = \frac{1}{2(1-x^2)^{1/2}} [P_{l(1)}(x) + P_{l(1)}(x)] = \frac{P_{l(1)}(x)}{2(1-x^2)^{1/2}} \quad (\text{B.30})$$

$$d_\theta P_{l0}(\cos \theta) = \frac{1}{2} [P_{l(m+1)}(\cos \theta) + P_{l(1)}(\cos \theta)] = -P_{l(1)}(\cos \theta). \quad (\text{B.31})$$

The associated Legendre equation is given by (*Arfken and Weber*, 1995, Eqs. 12.71 and 12.72):

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left[\sin \theta \frac{dP_{lm}(\cos \theta)}{d\theta} \right] + \left[l(l+1) - \frac{m^2}{\sin^2 \theta} \right] P_{lm}(\cos \theta) = 0 \quad (\text{B.32})$$

$$(1-x^2) \frac{d^2}{dx^2} P_{lm}(x) - 2x \frac{d^2}{dx^2} P_{lm}(x) + \left[l(l+1) - \frac{m^2}{1-x^2} \right] = 0. \quad (\text{B.33})$$

Equation (B.32) can be solved for the second derivatives $d_{\theta\theta} P_{lm}(\cos \theta)$:

$$\begin{aligned} \frac{1}{\sin \theta} [\cos \theta d_\theta P_{lm}(\cos \theta) + \sin \theta d_{\theta\theta} P_{lm}(\cos \theta)] &+ \left[l(l+1) - \frac{m^2}{\sin^2 \theta} \right] P_{lm}(\cos \theta) = 0 \\ d_{\theta\theta} P_{lm}(\cos \theta) &= -\frac{\cos \theta}{\sin \theta} d_\theta P_{lm}(\cos \theta) - \left[l(l+1) - \frac{m^2}{\sin^2 \theta} \right] P_{lm}(\cos \theta). \end{aligned} \quad (\text{B.34})$$

Equations (B.28) and (B.34) are used to calculate the first and second derivatives of the Legendre functions, which are used in the X_{lm} functions and derivatives (Equations (B.21)–(B.23)), which are in turn used in the ray-tracing calculations.

An important property of the spherical harmonic functions is that the $Y_{lm}(\theta, \phi)$ functions are eigenfunctions of the Laplacian operator:

$$\nabla^2 Y_{lm} = \left(\frac{-l(l+1)}{a^2} \right) Y_{lm}, \quad (\text{B.35})$$

where the $Y_{lm}(\theta, \phi)$ functions are plotted on a sphere with radius a . This was the reason for choosing the $Y_{lm}(\theta, \phi)$ functions as our test functions for our numerical approximation for the Laplacian (Section 4.5).

Appendix C

Euler angles and rotations

Rotations in three dimensions were used throughout this study. Any rotation of a rigid body (e.g., a sphere) may be represented by three real numbers. The *Euler angles* (α, β, γ) represent a useful convention for describing a rotation in 3D (Edmonds, 1960). Using the Euler-angle convention, each 3D rotation consists of three separate rotations. Each separate rotation is pictured in Figure C.1 and has an associated matrix, shown in Equation (C.1).

The *Euler rotation* $D(\alpha, \beta, \gamma)$ on the point (x, y, z) consists of three successive rotations and is given by:

$$\begin{pmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (\text{C.1})$$

The reverse of the operation $D(\alpha, \beta, \gamma)$ is given by

$$D^{-1}(\alpha, \beta, \gamma) = D(\pi - \gamma, \beta, \pi - \alpha). \quad (\text{C.2})$$

Figure C.1 illustrates the forward rotations, and the reverse operation is shown in Figure C.2. The key concept is that the Euler angles specify rotation angles for the axes, while the rotation matrix D is applied to the points (Edmonds, 1960). Thus, D is a 3-by-3 matrix whose entries are determined by (α, β, γ) via Equation (C.1).

Here we will illustrate an example of how we use the Euler rotations. Suppose we wish to compare the phase anomalies and amplitude anomalies calculated using the numerical model (i.e., spherical grids) with those calculated using ray tracing. Here are the four steps in selecting the receivers.

1. Specify a target source (lat, lon) and a set of target receivers (lats, lons). The locations are arbitrary — if we were comparing our results (numerical or ray theory) with actual data, then we would use the actual source and receiver locations.
2. Determine the Euler angles (α, β, γ) that rotate the target source to the North Pole, which is the position of the source in the reference frame of the numerical calculations. Rotate the target source and receivers by $D(\alpha, \beta, \gamma)$.
3. Find the nearest gridpoint to each target receiver. These are your receiver points (lats, lons) in the reference frame for the numerical calculations. Carry out the numerical calculations on the spherical grids and obtain the HEX phase and amplitude anomalies.
4. Rotate these receivers (and the source) back into the original reference frame, by using the reverse of the initial rotation, $D^{-1}(\alpha, \beta, \gamma)$. The source ends up back at the target source, and the receivers end up close to the target receivers. These are the “actual” lat-lon points of the receivers that you will use in the ray-tracing calculations.

For the ray-tracing calculations, the Euler angles are specified such that the source is on the equator at $(0, 0)$ and the receiver is at $(0, \Delta)$. Thus, each source-receiver pair has a different set of Euler angles to rotate them into the ray-tracing reference frame. Once a source-receiver pair is rotated to the equator, then we calculate the phase and amplitude anomalies associated with the ray path.

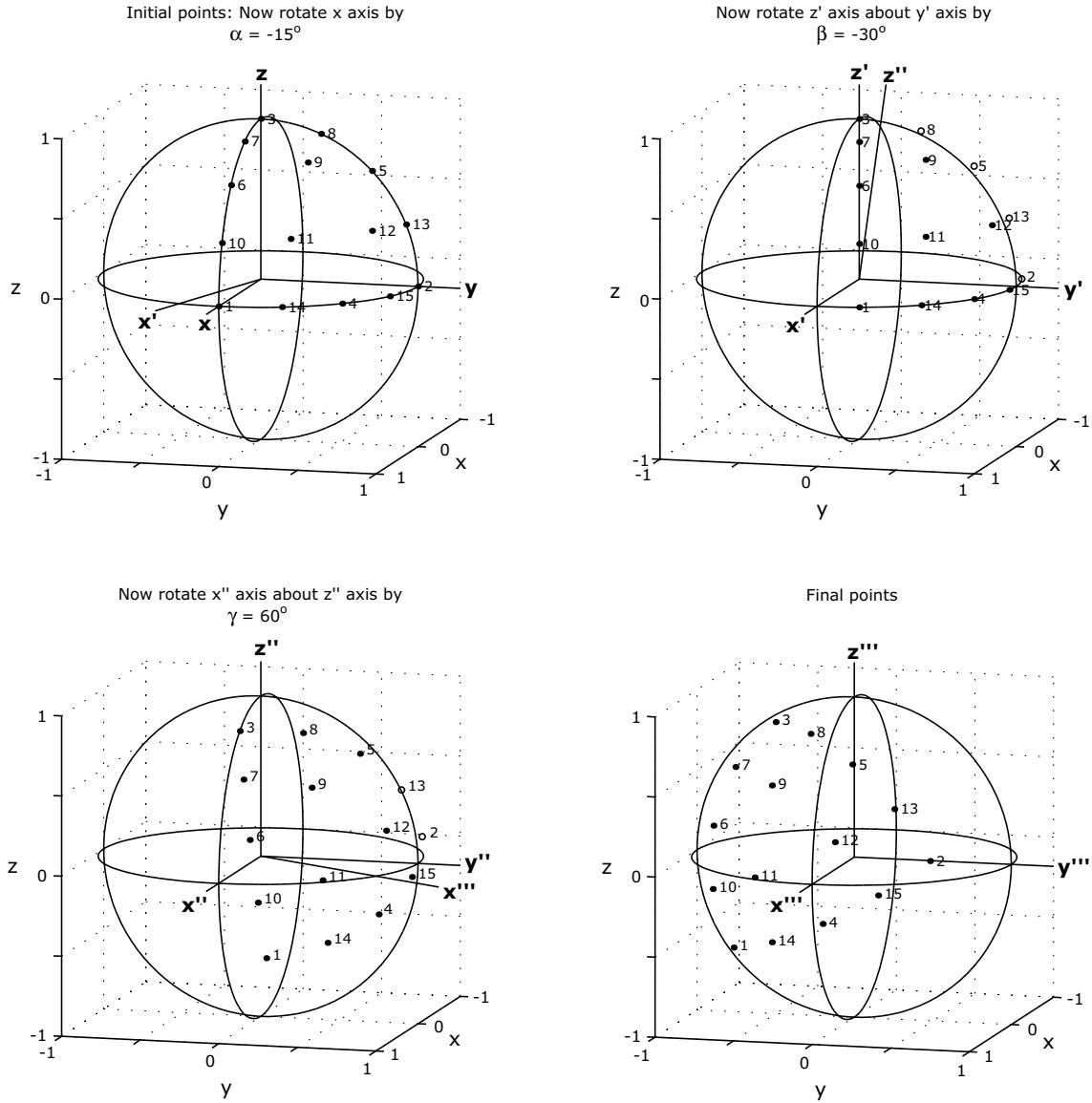


Figure C.1: Euler angle rotation. The set of sample points consists of 15 points in the quadrant $x, y, z > 0$. Each Euler angle (α, β, γ) specifies a rotation of the coordinate axes. The corresponding rotation of the sample points is given by one of the matrices in Eq. C.1. The α -rotation is a rotation in the xy -plane. The β -rotation is a rotation in the $x'z'$ -plane. The γ -rotation is a rotation in the $x''z''$ -plane. The final reference frame S''' is the same as the initial reference frame S .

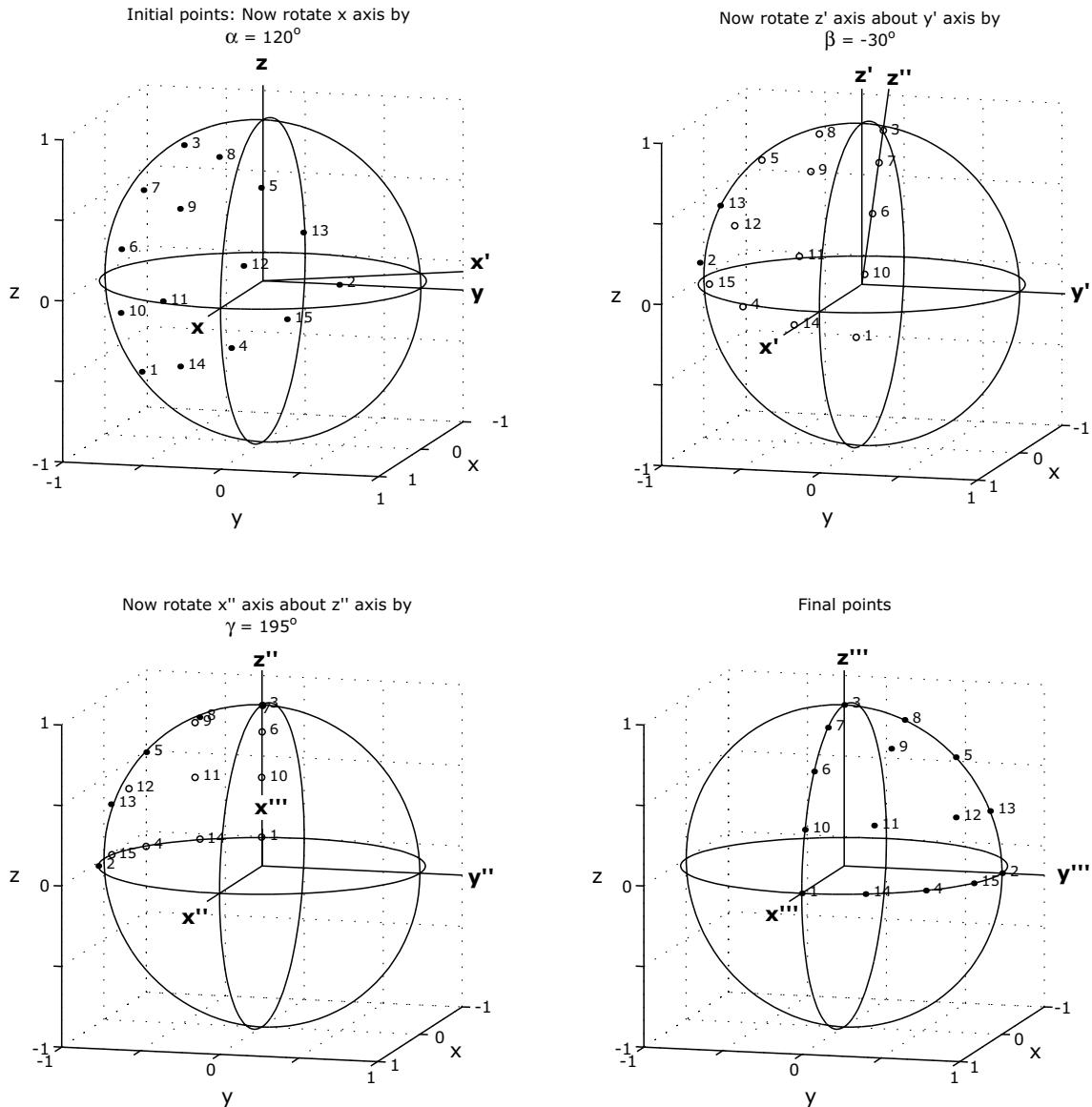


Figure C.2: Euler angle rotation (in reverse). Here we reverse the operations in Figure C.1; the final points here are the initial points in Figure C.1. The net rotation matrix is given by Eq. C.2.

Appendix D

Derivations for selected equations

In this appendix we include derivations for several of the equations presented in Chapter 3. Section D.4 below contains various definitions and properties used in the derivations.

D.1 Equation (3.12)

The method of Green's functions in obtaining Equation (3.12) is that the solution to Equation (3.10) is the convolution of the source-time function $f_{lm}(t)$ with the Green's function, i.e., the solution to Equation (3.12) for an impulsive force. Thus, using the definitions in Section D.4, the solution to Equation (3.12) is given by

$$u_{lm}(t) = G(t) * c^2 f_{lm}(t) = c^2 \int_{-\infty}^t G(t - \tau) f_{lm}(\tau) d\tau , \quad (\text{D.1})$$

where

$$G(t) = \frac{\sin \omega_l t}{\omega_l} . \quad (\text{D.2})$$

D.2 Equations (3.27) and (3.33)

The derivation of Equation (3.27) is as follows. Using the identity

$$\begin{aligned}\sin(A - B) &= \sin A \cos B - \cos A \sin B \\ \sin(\omega_l t - \omega_l \tau) &= \sin \omega_l t \cos \omega_l \tau - \cos \omega_l t \sin \omega_l \tau,\end{aligned}$$

we expand the left side of Equation (3.27):

$$\int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\sin \omega_l t \cos \omega_l \tau - \cos \omega_l t \sin \omega_l \tau}{\omega_l} d\tau \quad (\text{D.3})$$

$$= \int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\sin \omega_l t}{\omega_l} \cos \omega_l \tau d\tau - \int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\cos \omega_l t}{\omega_l} \sin \omega_l \tau d\tau \quad (\text{D.4})$$

$$= \frac{\sin \omega_l t}{\omega_l} \int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \cos \omega_l \tau d\tau - \frac{\cos \omega_l t}{\omega_l} \int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \sin \omega_l \tau d\tau \quad (\text{D.5})$$

$$= \frac{\sin \omega_l t}{\omega_l} \text{R} \left[\int_{-\infty}^{\infty} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) e^{-i \omega_l \tau} d\tau \right] - \frac{\cos \omega_l t}{\omega_l} \cdot 0 \quad (\text{D.6})$$

$$= \frac{\sin \omega_l t}{\omega_l} \text{R} \left[\mathcal{F} \left(\frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \right] \quad (\text{D.21}) \quad (\text{D.7})$$

$$= \frac{\sin \omega_l t}{\omega_l} \text{R} \left[e^{-\omega_l^2 \sigma^2 / 2} \right] \quad (\text{D.23}) \quad (\text{D.8})$$

$$= \frac{\sin \omega_l t}{\omega_l} e^{-\omega_l^2 \sigma^2 / 2}, \quad (\text{D.9})$$

where in obtaining Equation (D.6) we have used $\int_{-\infty}^{\infty} H_{even}(\tau) H_{odd}(\tau) d\tau = 0$.

The derivation of Equation (3.33) is similar to that above, where instead of Equation (D.3) we have

$$\int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\sin \omega_l t \cos \omega_l \tau - \cos \omega_l t \sin \omega_l \tau}{\omega_l} d\tau \quad (\text{D.10})$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\sin \omega_l t}{\omega_l t} \cos \omega_l \tau d\tau \\ &\quad - \int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \frac{\cos \omega_l t}{\omega_l t} \sin \omega_l \tau d\tau \end{aligned} \quad (\text{D.11})$$

$$\begin{aligned} &= \frac{\sin \omega_l t}{\omega_l} \int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) \cos \omega_l \tau d\tau \\ &\quad + \frac{\cos \omega_l t}{\omega_l} \int_{-\infty}^{\infty} \left(\frac{-\tau}{\sigma^2} \frac{e^{-\tau^2/2\sigma^2}}{\sqrt{2\pi} \sigma} \right) (-\sin \omega_l \tau) d\tau \end{aligned} \quad (\text{D.12})$$

$$= \frac{\sin \omega_l t}{\omega_l} \cdot 0 + \frac{\cos \omega_l t}{\omega_l} I \left[\int_{-\infty}^{\infty} h_2(\tau) e^{-i\omega_l \tau} d\tau \right] \quad (\text{D.13})$$

$$= \frac{\cos \omega_l t}{\omega_l} I \left[\mathcal{F} \left(\frac{d}{d\tau} (h_1(\tau)) \right) \right] \quad (\text{D.21}), (\text{3.31}) \quad (\text{D.14})$$

$$= \frac{\cos \omega_l t}{\omega_l} I \left[i \omega_l e^{-\omega_l^2 \sigma^2 / 2} \right] \quad (\text{D.24}) \quad (\text{D.15})$$

$$= \frac{\cos \omega_l t}{\omega_l} \omega_l e^{-\omega_l^2 \sigma^2 / 2} \quad (\text{D.16})$$

$$= \cos \omega_l t e^{-\omega_l^2 \sigma^2 / 2}, \quad (\text{D.17})$$

where in obtaining Equation (D.13) we have used $\int_{-\infty}^{\infty} H_{odd}(\tau) H_{even}(\tau) H_{even}(\tau) d\tau = 0$.

D.3 Equations (3.20) and (3.62)

For Equation (3.20), we want to simplify the following expression:

$$\begin{aligned}
 \int_0^\infty \left(\frac{e^{-\theta^2/2\mu^2}}{\mu^2} \right) \theta \, d\theta &= \int_0^\infty \frac{e^{w/2\mu^2}}{\mu^2} \frac{dw}{-2} \\
 &= \frac{-1}{2\mu^2} \int_0^{-\infty} e^{w/2\mu^2} dw \\
 &= -e^{w/2\mu^2} \Big|_0^{-\infty} = -(0 - 1) = 1,
 \end{aligned} \tag{D.18}$$

where we have made the substitution $w \equiv -\theta^2$. This also shows how we selected the normalizations on $g(\theta)$.

For Equation (3.62), we begin with

$$I_l(0) = \lim_{\mu \rightarrow 0} \int_0^\pi P_l(\cos \theta) \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \sin \theta \, d\theta. \tag{D.19}$$

In the limit as $\mu \rightarrow 0$, the function in the integral becomes a delta function $\delta(\theta)$ centered at $\theta = 0$. Thus, as $\mu \rightarrow 0$, $\theta \rightarrow 0$ (a spike at the North Pole), and we can use the first-order approximation $\sin \theta \approx \theta$. Furthermore, $\cos \theta \rightarrow 1$ and $P_l(1) = 1$. Putting this together we have

$$I_l(0) = P_l(1) \int_0^\pi \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \theta \, d\theta \approx P_l(1) \int_0^\infty \frac{e^{-\theta^2/2\mu^2}}{\mu^2} \theta \, d\theta = 1. \tag{D.20}$$

The upper integration limit of π can be replaced by ∞ , because $g(\theta) \approx 0$ for $\theta \in [\pi, \infty]$. The last step follows from the results in Equation (D.18). (This result can also be seen graphically in Figure 3.3.)

D.4 Some definitions and properties

Here we list some equations used in the Chapter 3. Spherical harmonic functions are discussed in Appendix B.

Fourier transforms

We define forward and inverse Fourier transforms:

$$\mathcal{F}[h(t)] = \tilde{h}(\omega) = \int_{-\infty}^{\infty} h(t) e^{-i\omega t} dt \quad (\text{D.21})$$

$$\mathcal{F}^{-1}[\tilde{h}(\omega)] = h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{h}(\omega) e^{i\omega t} d\omega. \quad (\text{D.22})$$

Some properties of Fourier transforms that we will use are as follows:

$$\mathcal{F}\left[k e^{-a^2 t^2}\right] = \left(\int_{-\infty}^{\infty} k e^{-a^2 t^2} dt\right) e^{-\frac{\omega^2}{4a^2}} \quad (\text{D.23})$$

$$\mathcal{F}[(f'(t))] = i\omega \mathcal{F}(f(t)) \quad (\text{D.24})$$

For example, using Equation (D.23), the Fourier transform of the Gaussian function $h_1(t)$ in Equation (3.17) is, with $k = 1/(\sqrt{2\pi}\sigma)$ and $a = 1/(\sqrt{2}\sigma)$:

$$\mathcal{F}[h_1(t)] = e^{-\omega^2 \sigma^2 / 2}. \quad (\text{D.25})$$

Convolution

The *convolution of $f(t)$ with $g(t)$* is given by the formula¹

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(t - \tau) g(\tau) d\tau. \quad (\text{D.26})$$

Properties of convolutions that we will use are as follows:

$$f(t) * g(t) = g(t) * f(t); \quad (\text{D.27})$$

¹Arfken and Weber (1995, p. 863) have slightly different definition of convolution.

Let's say we have $u = u(x, t)$ and $h = h(t)$. Then the following is true:

$$F_a(t, x) = u * h = \int_{-\infty}^{\infty} u(t - \tau, x) h(\tau) d\tau \quad (\text{D.28})$$

$$F_b(t, x) = h * u = \int_{-\infty}^{\infty} h(t - \tau) u(\tau, x) d\tau \quad (\text{D.29})$$

$$\frac{\partial F_a}{\partial t} = \frac{\partial}{\partial t}(u * h) = \int_{-\infty}^{\infty} \dot{u}(t - \tau, x) h(\tau) d\tau = \dot{u} * h \quad (\text{D.30})$$

$$\frac{\partial F_b}{\partial t} = \frac{\partial}{\partial t}(h * u) = \int_{-\infty}^{\infty} \dot{h}(t - \tau) u(\tau, x) d\tau = \dot{h} * u \quad (\text{D.31})$$

Green's functions

The solution $x(t)$ to the second-order ODE

$$\frac{d^2x}{dt^2} + k \frac{dx}{dt} + \omega_0^2 x = h(t) \quad (\text{D.32})$$

is given by

$$x(t) = G(t) * h(t) = \int_{-\infty}^{\infty} G(t - \tau) h(\tau) d\tau, \quad (\text{D.33})$$

where $G(t)$ is Green's function and satisfies Equation (D.32) for an impulsive source, i.e.,

$$\frac{d^2G}{dt^2} + k \frac{dG}{dt} + \omega_0^2 x = \delta(t). \quad (\text{D.34})$$

By definition, $G(t) = 0$ for $t < 0$, which implies that $G(t - \tau) = 0$ for $\tau > t$. We can split $x(t)$ into two integrals and simplify:

$$\begin{aligned} x(t) &= \int_{-\infty}^t G(t - \tau) h(\tau) d\tau + \int_t^{\infty} G(t - \tau) h(\tau) d\tau \\ &= \int_{-\infty}^t G(t - \tau) h(\tau) d\tau. \end{aligned} \quad (\text{D.35})$$

Thus, this expression is true for $t > \tau$, where τ is the time value within Equation (D.32). This places the condition that the function $x(t)$ and its derivatives must vanish at (sufficiently large) negative times.

Appendix E

Ray-tracing equations

E.1 Receiver at a fixed longitude

The equations for path, amplitude, and phase anomaly in *Woodhouse and Wong* (1986) are specified for a source and receiver situated on the equator. Here we show a slightly more generalized version of the equations, where only the source is on the equator and the receiver is at a particular longitude ϕ_f but not necessarily on the equator. Thus, the source is on the Greenwich meridian at $(\theta_0, \phi_0) = (\frac{\pi}{2}, 0)$, and the receiver is at (θ_f, ϕ_f) . Figure E.1A shows the setup for the ray-tracing.

Variables and notation

Variable definitions follow from *Woodhouse and Wong* (1986):

$$\begin{aligned}\gamma &= \cot \theta \\ \nu &= -\frac{d\gamma}{d\phi} = -\frac{d \cot \theta}{d\theta} \frac{d\theta}{d\phi} = \csc^2 \theta \frac{d\theta}{d\phi} \\ \gamma' &= \frac{\partial \gamma}{\partial \nu_0} \\ \nu' &= \frac{\partial \nu}{\partial \nu_0} = -\frac{\partial \gamma'}{\partial \phi}.\end{aligned}$$

Referring to Figure E.1, we define the azimuth along the ray, ξ , as follows:

$$\tan \xi = \frac{h_1 d\theta}{h_2 d\phi} = \frac{-1}{\sin \theta} \frac{d\theta}{d\phi}, \quad (\text{E.1})$$

or, in terms of ν ,

$$\tan \xi = \frac{-1}{\sin \theta} \frac{d\theta}{d\gamma} \frac{d\gamma}{d\phi} = \frac{-1}{\sin \theta} \left(\frac{1}{-\csc^2 \theta} \right) (-\nu) = -\nu \sin \theta. \quad (\text{E.2})$$

Note that the azimuth is measured in a direction counter-clockwise from eastward at a given point on the ray. The initial azimuth is the take-off angle,

$$\alpha \equiv \xi_0, \quad (\text{E.3})$$

and the final azimuth is the off-azimuth arrival angle, ξ_f , which can be expressed as

$$\tan \xi_f = -\nu_f \sin \theta_f, \quad (\text{E.4})$$

where θ_f is the finishing latitude of the ray at the finishing longitude ϕ_f .

Initial conditions are (see Figure E.1):

$$\theta_0 = \frac{\pi}{2} \quad (\text{E.5})$$

$$\phi_0 = 0 \quad (\text{E.6})$$

$$\tan \alpha = \tan \xi_0 = \frac{-1}{\sin \theta_0} \left. \frac{d\theta}{d\phi} \right|_{\phi=\phi_0} = -\left. \frac{d\theta}{d\phi} \right|_{\phi=0}. \quad (\text{E.7})$$

These lead to the following expressions:

$$\gamma_0 = \cot \theta_0 = 0 \quad (\text{E.8})$$

$$\nu_0 = \csc^2 \theta_0 \left. \frac{d\theta}{d\phi} \right|_{\phi=\phi_0} = -\tan \alpha. \quad (\text{E.9})$$

In the remainder of this section, we develop equations for the homogeneous ray paths. Then we derive an expression for the amplitude anomaly due to a heterogeneous earth, for a receiver at a fixed longitude ϕ_f .

Homogeneous ray paths

The solutions to Equation (6.3) have the form

$$\gamma(\phi) = A \cos \phi + B \sin \phi. \quad (\text{E.10})$$

Differentiating Equation (E.10) we have, in two different forms,

$$\begin{aligned} \frac{d\gamma}{d\phi} &= -A \sin \phi + B \cos \phi \\ \nu &= A \sin \phi - B \cos \phi. \end{aligned}$$

Our initial conditions specify the initial position of the source and the take-off angle α :

$$\gamma(0) = 0 \quad (\text{E.11})$$

$$\left. \frac{d\gamma}{d\phi} \right|_{\phi=0} \equiv -\nu_0 = \tan \alpha \quad (\text{E.12})$$

Using the initial conditions, we have

$$\gamma(0) = A = 0 \quad (\text{E.13})$$

$$\left. \frac{d\gamma}{d\phi} \right|_{\phi=0} = B = -\nu_0 = \tan \alpha. \quad (\text{E.14})$$

Our solution can be expressed in several different forms:

$$\gamma(\phi) = -\nu_0 \sin \phi \quad (\text{E.15})$$

$$\nu(\phi) = \nu_0 \cos \phi \quad (\text{E.16})$$

$$\cot \theta = \tan \alpha \sin \phi \quad (\text{E.17})$$

$$\theta(\phi) = \cot^{-1}(\tan \alpha \sin \phi) \quad (\text{E.18})$$

$$\bar{\theta}(\phi) = \tan^{-1}(\tan \alpha \sin \phi), \quad (\text{E.19})$$

where θ is colatitude and $\bar{\theta}$ is latitude ($\bar{\theta} = \frac{\pi}{2} - \theta$). Plots of Equation (E.19) are shown in Figure 6.2.

Amplitude anomaly

The amplitude anomaly is defined as the square root of the ratio of geometrical spreading in the homogeneous case to the geometrical spreading in the heterogeneous case:

$$A = \left| \frac{\frac{d\sigma}{d\alpha}|_{\text{hom}}}{\frac{d\sigma}{d\alpha}|_{\text{het}}} \right|^{1/2}. \quad (\text{E.20})$$

Figure E.1B shows the starting point for the derivation. We consider two rays emanating from the source with an infinitesimal take-off angle difference, $\delta\alpha$. The two rays form the ray tube, whose width is measured by $d\sigma$.

HOMOGENEOUS CASE. With the homogeneous earth, the ray paths are great-circle paths, and the geometrical spreading on the sphere in the homogeneous case is simply given by

$$\frac{d\sigma}{d\alpha} \Big|_{\text{hom}} = a \sin \Delta, \quad (\text{E.21})$$

where Δ is the arc distance between the source and the receiver and a is the radius of the earth (*Woodhouse and Wong*, 1986, Eq. 40b). Physically, this means that the geometrical spreading is a minimum at the source ($\Delta = 0$) and the antipode ($\Delta = \pi$), where all the great-circle paths converge, and it is a maximum at a distance $\Delta = \frac{\pi}{2}$ from the source (and antipode), where the rays are most spread out.

In our case, the source is at (θ_0, ϕ_0) and the receiver is at (θ_f, ϕ_f) . The dot-product gives us an expression for the arc distance, which we show here for a source at $(\frac{\pi}{2}, 0)$:

$$\begin{aligned} \cos \Delta &= \mathbf{r}_s \cdot \mathbf{r}_r \\ &= \sin \theta_0 \cos \phi_0 \sin \theta_f \cos \phi_f + \sin \theta_0 \sin \phi_0 \sin \theta_f \sin \phi_f + \cos \theta_0 \cos \theta_f \\ &= \sin \theta_f \cos \phi_f. \end{aligned} \quad (\text{E.22})$$

Equation (E.21) then can be written as

$$\frac{d\sigma}{d\alpha} \Big|_{\text{hom}} = a (1 - \cos^2 \Delta)^{1/2} = a (1 - \sin^2 \theta_f \cos^2 \phi_f)^{1/2}, \quad (\text{E.23})$$

If our receiver is on the equator ($\theta_f = \frac{\pi}{2}$), then $\phi_f = \Delta$, and Equation (E.23) reduces to Equation (E.21).

HETEROGENEOUS CASE. The geometrical spreading in the heterogeneous case can be expanded using the relationships above (Eq. E.9):

$$\frac{d\sigma}{d\alpha} \Big|_{\text{het}} = \frac{d\sigma}{d\tan\alpha} \frac{d\tan\alpha}{d\alpha} = -\frac{d\sigma}{d\nu_0} \sec^2\alpha. \quad (\text{E.24})$$

From Figure E.1B we see that we have the relationship $\cos\xi = d\sigma/d\theta$, or $d\sigma = d\theta \cos\xi$. Differentiating $\gamma = \cot\theta$, we have $d\gamma = -\csc^2\theta d\theta$, or $d\theta = -d\gamma \sin^2\theta$. These two expressions can be used in Equation (E.24):

$$\begin{aligned} \frac{d\sigma}{d\alpha} \Big|_{\text{het}} &= -\frac{d\sigma}{d\nu_0} \sec^2\alpha \\ &= -\frac{d\theta \cos\xi}{d\nu_0} \sec^2\alpha \\ &= \frac{d\gamma \sin^2\theta \cos\xi}{d\nu_0} \sec^2\alpha \\ &= \frac{d\gamma}{d\nu_0} \cos\xi \sec^2\alpha \sin^2\theta. \end{aligned} \quad (\text{E.25})$$

Using trigonometric substitutions, we have

$$\frac{d\sigma}{d\alpha} \Big|_{\text{het}} = \frac{d\gamma}{d\nu_0} \left(\frac{1}{\sqrt{1 + \tan^2\xi_f}} \right) (1 + \tan^2\alpha) \sin^2\theta_f, \quad (\text{E.26})$$

where we have entered the f -subscripts to indicate that we are at the finishing longitude ϕ_f .

Substituting Equations (E.4) and (E.9) into Equation (E.26), we get

$$\frac{d\sigma}{d\alpha} \Big|_{\text{het}} = a \gamma'_f (1 + \nu_f^2 \sin^2\theta_f)^{-1/2} (1 + \nu_0^2) \sin^2\theta_f, \quad (\text{E.27})$$

where we have also substituted $\gamma' \equiv d\gamma/d\nu_0$ and inserted the earth radius, a , which is simply a scaling factor for the geometric spreading. Note that Equation (E.27) reduces to Equation 36 of *Woodhouse and Wong (1986)* when the receiver is placed on the equator ($\theta_f = \frac{\pi}{2}$ and $\phi_f = \Delta$).

We now substitute Equations (E.27) and (E.23) into Equation (E.20) to get an expression for the amplitude anomaly:

$$\begin{aligned}
A &= \left| \frac{\frac{d\sigma}{d\alpha}|_{\text{hom}}}{\frac{d\sigma}{d\alpha}|_{\text{het}}} \right|^{1/2} \\
&= \left| \frac{a (1 - \sin^2 \theta_f \cos^2 \phi_f)^{1/2}}{a \sin^2 \theta_f (1 + \nu_0^2) (1 + \nu_f^2 \sin^2 \theta_f)^{-1/2} \gamma'_f} \right|^{1/2} \\
&= \left| \frac{1}{\sin^2 \theta_f (1 + \nu_0^2) \gamma'_f} \right|^{1/2} (1 - \sin^2 \theta_f \cos^2 \phi_f)^{1/4} (1 + \nu_f^2 \sin^2 \theta_f)^{1/4} \\
&= \frac{(1 - \sin^2 \theta_f \cos^2 \phi_f)^{1/4} (1 + \nu_f^2 \sin^2 \theta_f)^{1/4}}{|\sin \theta_f| (\gamma'_f)^{1/2} (1 + \nu_0^2)^{1/2}}. \tag{E.28}
\end{aligned}$$

Note the inclusion of the factor θ_f in three places, which results from allowing the receiver to be situated at non-zero latitudes. For the case when the receiver is placed on the equator ($\theta_f = \frac{\pi}{2}$ and $\phi_f = \Delta$), Equation (E.28) reduces to

$$A = \left| \frac{\sin \Delta}{\gamma'_f (1 + \nu_0^2)} \right|^{1/2} [1 + \nu^2(\Delta)]^{1/4}, \tag{E.29}$$

which is exactly Equation 41 of *Woodhouse and Wong (1986)*. Plots of Equation (E.28) are shown in Figure 6.10, which showing the variation of amplitude anomaly as a function of take-off angle α along a finishing longitude ϕ_f .

E.2 Equation (6.19) — $ds/d\phi$

Beginning with Eq. 12 of *Woodhouse and Wong* (1986) (*Dahlen and Tromp*, 1998, Eq. A.100), we have the expression for the infinitesimal distance ds between a point (x^1, x^2) and $(x^1 + dx^1, x^2 + dx^2)$:

$$\begin{aligned} ds &= (g_{\sigma\nu} \dot{x}^\sigma \dot{x}^\nu)^{1/2} dt \\ &= (g_{\sigma\nu} dx^\sigma dx^\nu)^{1/2}, \end{aligned} \quad (\text{E.30})$$

where $x^1 = \gamma \equiv \cot \theta$, $x^2 = \phi$, and $g_{\sigma\nu}$ are the covariant components of the metric tensor, given by

$$\begin{aligned} g_{11} &= \frac{a^2}{(1 + \gamma^2)^2} & g_{12} &= 0 \\ g_{21} &= 0 & g_{22} &= \frac{a^2}{(1 + \gamma^2)}, \end{aligned} \quad (\text{E.31})$$

and the contravariant components (*Woodhouse and Wong*, 1986, Eq. 32) are found via

$$\begin{aligned} [g^{\sigma\nu}] &= [g_{\sigma\nu}]^{-1} : \\ g^{11} &= a^{-2} (1 + \gamma^2)^2 & g^{12} &= 0 \\ g^{21} &= 0 & g^{22} &= a^{-2} (1 + \gamma^2). \end{aligned} \quad (\text{E.32})$$

(Note that with $\gamma = \cot \theta$, we have $1/(1 + \gamma^2) = \sin^2 \theta$.) Substituting Equation (E.31) into Equation (E.30), we have

$$\begin{aligned} ds &= [g_{\sigma\nu} dx^\sigma dx^\nu]^{1/2} \\ &= [g_{11}(dx^1)^2 + g_{22}(dx^2)^2]^{1/2} \\ &= [g_{11} d\gamma^2 + g_{22} d\phi^2]^{1/2} \\ &= [g_{11} \nu^2 + g_{22}]^{1/2} d\phi \\ &= \left[\frac{a^2}{(1 + \gamma^2)^2} \nu^2 + \frac{a^2}{(1 + \gamma^2)} \right]^{1/2} d\phi \end{aligned} \quad (\text{E.33})$$

$$\frac{ds}{d\phi} = a \left[\frac{\nu^2}{(1 + \gamma^2)^2} + \frac{1}{(1 + \gamma^2)} \right]^{1/2}, \quad (\text{E.34})$$

where we have used $x^1 = \gamma$, $x^2 = \phi$, and $\nu = -d\gamma/d\phi$. From the latter follows $d\gamma^2 = \nu^2 d\phi^2$.

E.3 Equation (6.11)

We use Newton's method to help find the take-off angles of the ray(s) that converges on a particular receiver for the ray-shooting. We expand $\nu_0(\gamma_n)$ as a first-order Taylor series, where $\gamma_n = \gamma(\Delta)$ is the finishing γ value for the *next* iteration and γ_f is the finishing γ value for the *current* iteration:

$$\begin{aligned}
\nu_0(\gamma_n) &= \nu_0(\gamma_f) + (\gamma_n - \gamma_f) \frac{\partial \nu_0}{\partial \gamma}(\gamma_f) + \dots \\
&\approx \nu_0(\gamma_f) + (0 - \gamma_f) \frac{\partial \nu_0}{\partial \gamma}(\gamma_f) \\
&= \nu_0(\gamma_f) - \frac{\gamma_f}{\frac{\partial \nu_0}{\partial \gamma}(\gamma_f)} \\
&= \nu_0(\gamma_f) - \frac{\gamma_f}{\gamma'_f},
\end{aligned} \tag{E.35}$$

where in the second step we set $\gamma_n = 0$ and in the last step we substituted Equation (6.6). Thus, Equation (E.35) allows a method of picking the ν_0 's so that the rays converge on the receiver. See Section 6.3.

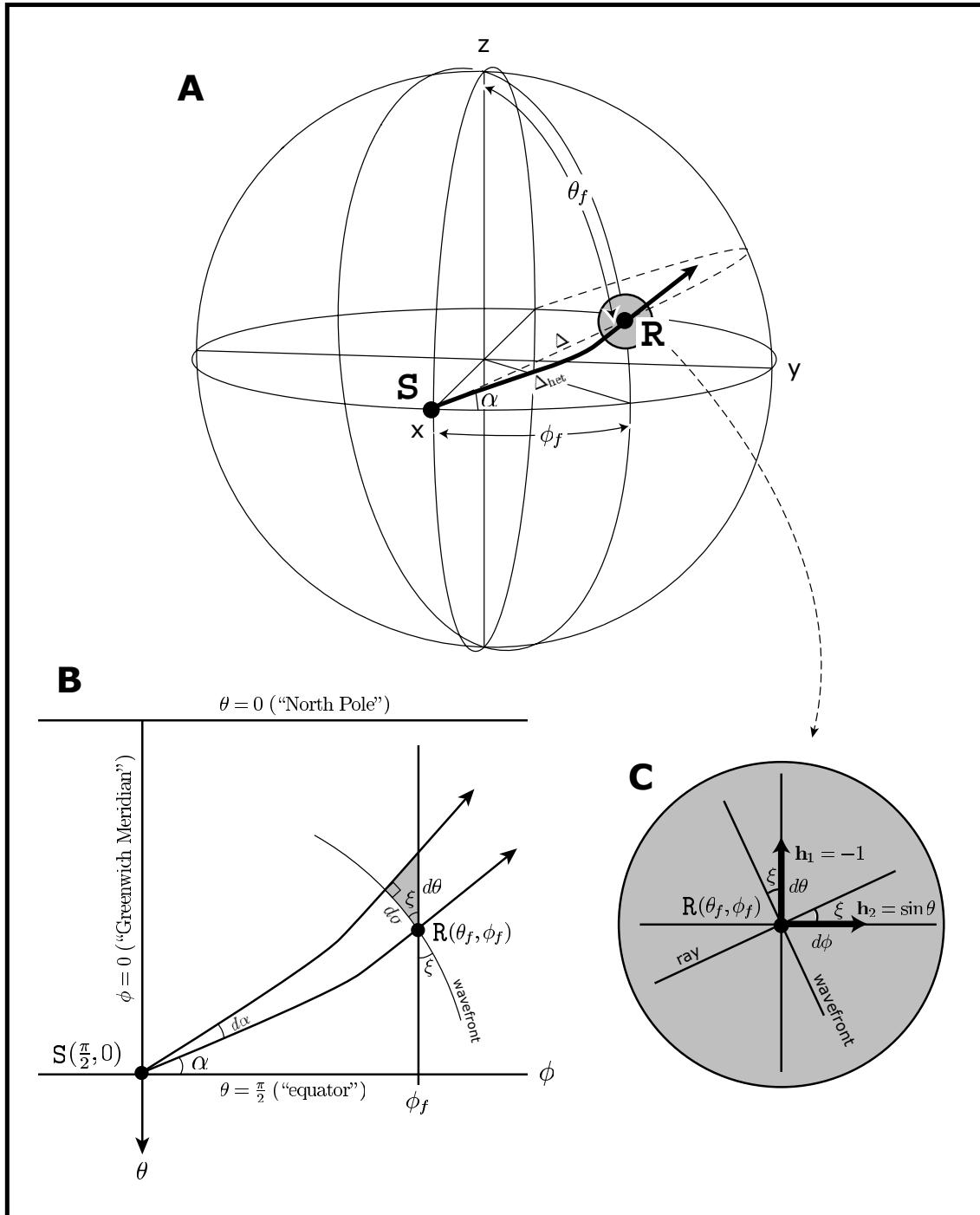


Figure E.1: Variables and setup for the ray-tracing equations on the sphere (Section E.1). The source is situated on the equator at $(\theta_0, \phi_0) = (\frac{\pi}{2}, 0)$, and the receiver is at (θ_f, ϕ_f) . The take-off angle is $\alpha \equiv \xi_0$. **A.** The homogeneous ray path is a great-circle ray path (dashed, Δ); the heterogeneous ray path (bold, Δ_{het}) deviates from the great-circle path ($\Delta_{\text{het}} > \Delta$). **B.** Exaggerated view of the ray in the vicinity of the receiver, showing the relationship $\cos \xi = d\sigma/d\theta$, where ξ is the azimuth of the ray. The geometrical spreading is defined as $d\sigma/d\alpha$. The quote-marks highlight the fact that the earth model is rotated into the ray-tracing reference frame (Appendix C), so that the actual North Pole, etc., are not where we have labeled them (except in the case where the earth model is unrotated). **C.** Coordinate system in the vicinity of the receiver.

Appendix F

Wave parameters and dispersion

Table F.1: Wave parameters and some related equations. The equations in the last column incorporate l and a and are approximate for $l \gg 1$ (see *Henson and Dahlen*, 1986; *Woodhouse*, 1996, Eq. 2.110).

earth radius	a	6371 km	
earth circumference	$2\pi a$	40030 km	
harmonic degree	l		
period	T	$T = \frac{1}{f} = \frac{2\pi}{\omega} = \frac{\lambda}{c}$	$T = \frac{2\pi a}{c(l+\frac{1}{2})}$
frequency	f	$f = \frac{\omega}{2\pi} = \frac{1}{T} = \frac{c}{\lambda}$	$f = \frac{c(l+\frac{1}{2})}{2\pi a}$
wavelength	λ	$\lambda = \frac{c}{f} = c T = \frac{2\pi}{k}$	$\lambda = \frac{2\pi a}{l+\frac{1}{2}}$
wavenumber	k	$k = \frac{\omega}{c} = \frac{2\pi}{\lambda} = \frac{2\pi f}{c}$	$k = \frac{l+\frac{1}{2}}{a}$
angular frequency	ω	$\omega = \frac{2\pi}{T} = c k$	$\omega = \frac{c(l+\frac{1}{2})}{a}$
velocity	c	$c = \frac{\lambda}{T} = f \lambda = \frac{\omega}{k}$	$c = \frac{2\pi a}{T(l+\frac{1}{2})}$

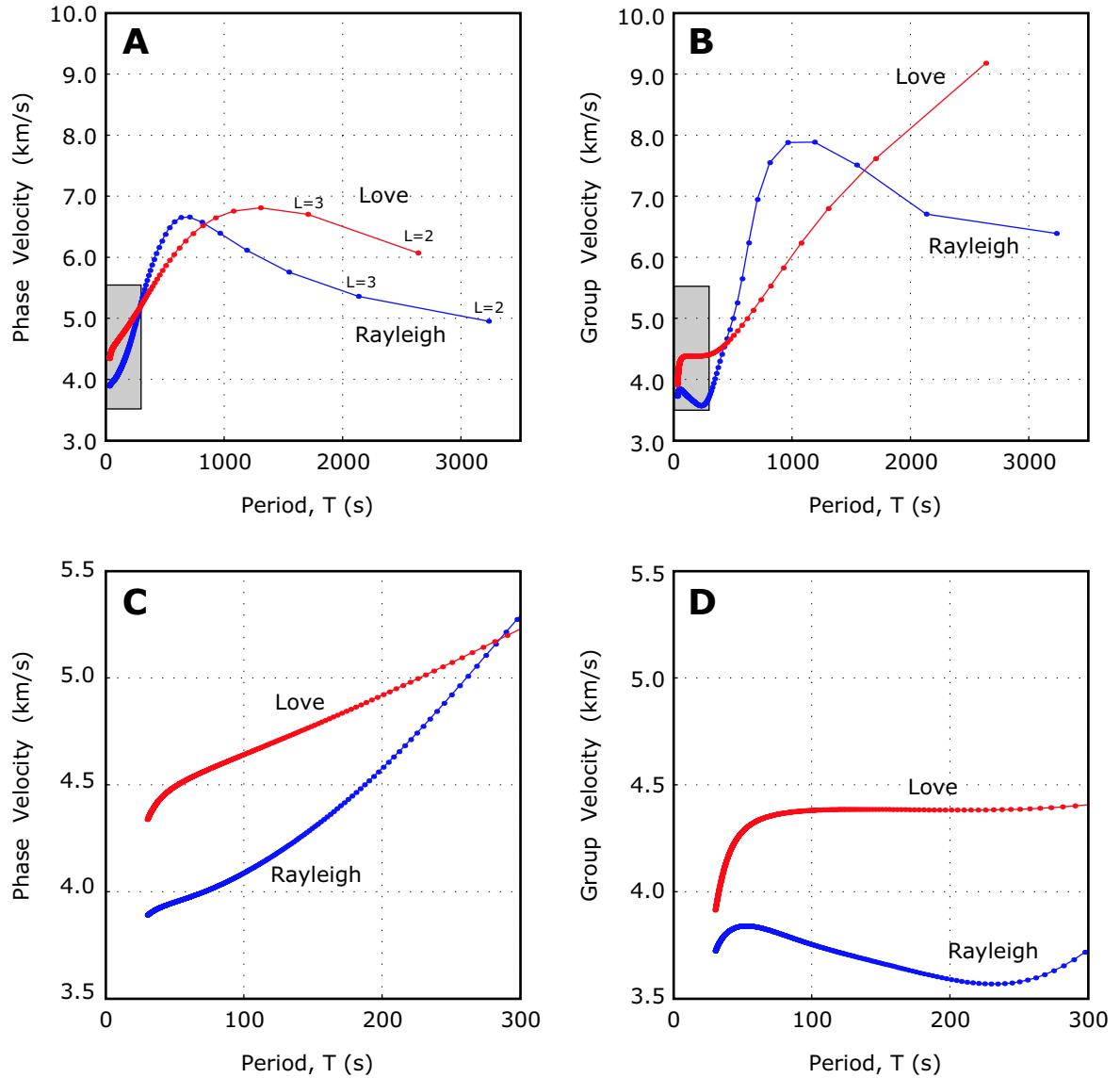


Figure F.1: Dispersion of surface waves, plotted as phase velocity $c(T)$ and group velocity $U(T)$. The curves are calculated using normal-modes summation with the 1D earth model PREM (Dziewonski and Anderson, 1981); each point corresponds to a normal mode with harmonic degree l . A couple points are labeled in A. These curves are for the fundamental spheroidal (Rayleigh) and toroidal (Love) modes ($n = 0$); dispersion curves of overtone branches ($n \geq 1$) can be seen in Figures 11.5 and 11.6 of Dahlen and Tromp (1998). **A.** Phase velocity dispersion $c(T)$. Gray box shows region of C. **B.** Group velocity dispersion $U(T)$. Gray box shows region of D. **C-D.** Same as A-B, only with different axes ranges. With periods $T < 300$ s, Love waves travel faster than Rayleigh waves. Calculations courtesy of Ana Ferreira.

Appendix G

Flow chart of computer programs

Figure 7.1 shows a very schematic flow chart of the method to generate scatterplot comparisons among the three different methods: numerical model, exact ray theory, linearized ray theory. Figure G.1 shows a more accurate depiction of the actual method, including the naming convention for the data files.

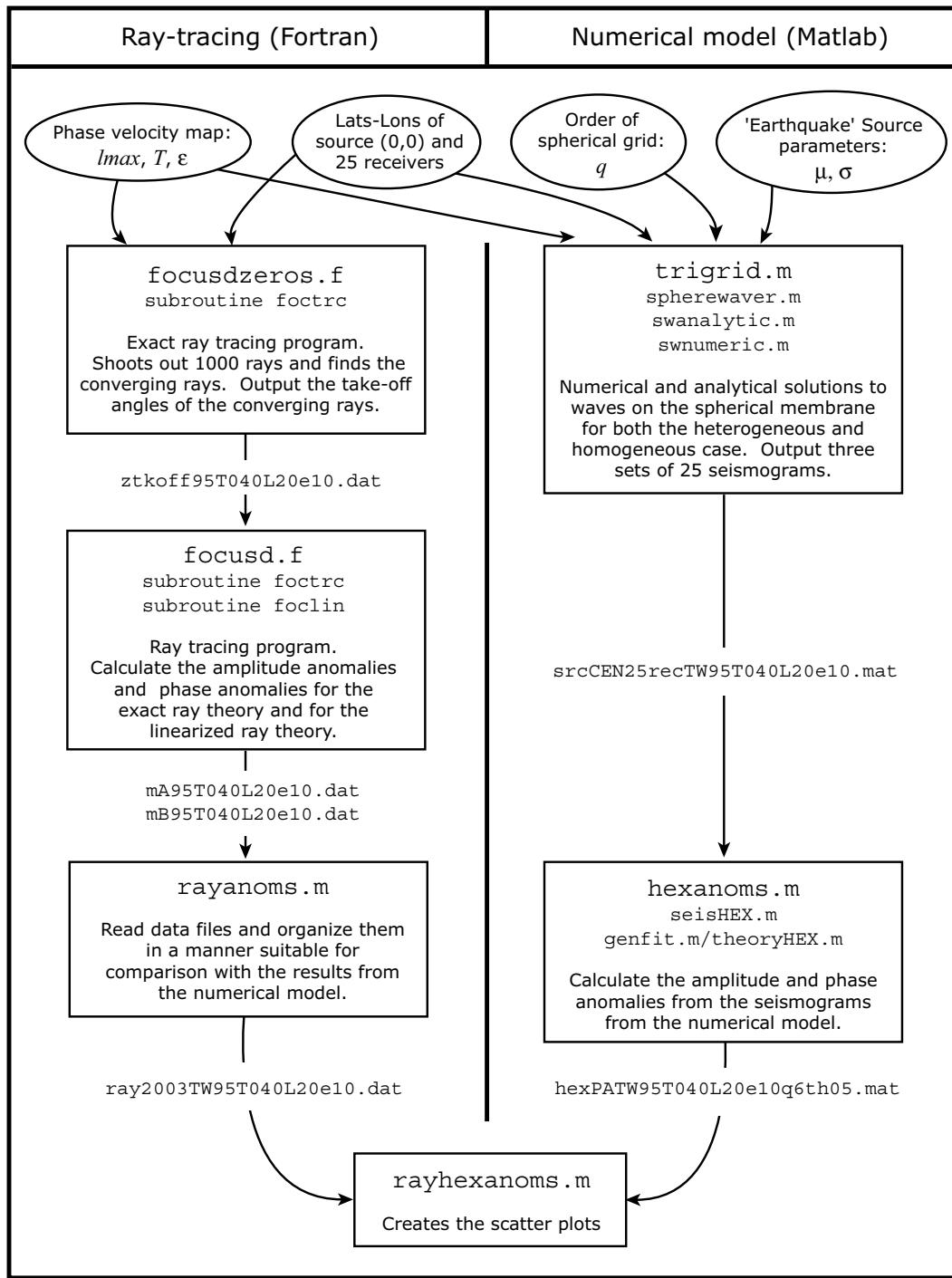


Figure G.1: Flowchart of computer programs used in calculating the phase and amplitude anomalies from ray theory (left) and from the numerical model (right). See Figure 7.1 for a more conceptual view of the process.

Bibliography

- Albarede, F., and R. D. van der Hilst, Zoned mantle convection, *Phil. Trans. R. Soc. A.*, **360**, 2569–2592, 2002.
- Arfken, G. B., and H. J. Weber, *Mathematical Methods for Physicists*, 4 ed., Academic Press, San Diego, Calif., 1995.
- Augenbaum, J. M., and C. S. Peskin, On the construction of the Voronoi mesh on a sphere, *J. Comp. Phys.*, **14**, 177–192, 1985.
- Baig, A. M., F. A. Dahlen, and S.-H. Hung, traveltimes of waves in three-dimensional random media, *Geophys. J. Int.*, **153**, 467–482, 2003.
- Bassin, C., G. Laske, and G. Masters, The current limits of resolution for surface wave tomography in North America, in *EOS Trans. Am. Geophys. Union*, vol. F897, p. 81, 2000.
- Baumgardner, J. R., and P. O. Frederickson, Icosahedral discretization of the two-sphere, *SIAM J. Num. Anal.*, **22**, 1107–1115, 1985.
- Boschi, L., and G. Ekström, New images of the Earth’s upper mantle from measurements of surface-wave phase velocity anomalies, *J. Geophys. Res.*, **107(B4)**, 2059, 10.1029/2000JB000,059, 2002.
- Capdeville, Y., E. Chaljub, J. P. Vilotte, and J. P. Montagner, Coupling the spectral element method with a modal solution for elastic wave propagation in global earth models, *Geophys. J. Int.*, **152**, 34–67, 2003.
- Clayton, R., and B. Engquist, Absorbing boundary conditions for acoustic and elastic wave equations, *Bull. Seismol. Soc. Am.*, **67**, 1529–1540, 1977.
- Cui, J., and W. Freeden, Equidistribution on the sphere, *SIAM J. Sci. Comp.*, **18**(2), 595–609, 1997.
- Cullen, M. J. P., Integration of the primitive equations on a sphere using the finite element method, *Q. J. R. Meteorol. Soc.*, **100**, 555–562, 1974.
- Dahlen, F. A., and J. Tromp, *Theoretical Global Seismology*, Princeton U. Press, New Jersey, 1998.
- Dziewonski, A., and D. Anderson, Preliminary reference earth model, *Phys. Earth Planet. Inter.*, **25**, 297–356, 1981.

- Dziewonski, A., and J. H. Woodhouse, Studies of the seismic source using normal-mode theory, in *Earthquakes: observation, theory and interpretation: notes from the International School of Physics “Enrico Fermi” (1982: Varenna, Italy)*, vol. LXXXV, edited by H. Kanamori and E. Boschi, pp. 45–137, North-Holland Pub. Co., Amsterdam, 1983.
- Edmonds, A. R., *Angular Momentum in Quantum Mechanics*, 2 ed., Princeton U. Press, Princeton, New Jersey, 1960.
- Ekström, G., J. Tromp, and E. W. F. Larson, Measurements and global models of surface wave propagation, *J. Geophys. Res.*, **102**, 8137–8157, 1997.
- Gilbert, F., and A. M. Dziewonski, An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, *Phil. Trans. R. Soc. Lond.*, **278**, 187–269, 1975.
- Grand, S. P., Mantle shear-wave tomography and the fate of subducted slabs, *Phil. Trans. R. Soc. A.*, **360**, 2545–2567, 2002.
- Heikes, R., and D. A. Randall, Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part I: Basic design and results of tests, *Monthly Weather Review*, **123**, 1862–1880, 1995a.
- Heikes, R., and D. A. Randall, Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part II: A detailed description of the grid and an analysis of numerical accuracy, *Monthly Weather Review*, **123**, 1881–1887, 1995b.
- Henson, I. H., and F. A. Dahlen, Asymptotic normal modes of a laterally heterogeneous earth 2. Further results, *J. Geophys. Res.*, **91**, 12,467–12,481, 1986.
- Komatitsch, D., and J. Tromp, Introduction to the spectral element method for three-dimensional seismic wave propagation, *Geophys. J. Int.*, **139**, 806–822, 1999.
- Komatitsch, D., and J. Tromp, Spectral-element simulations of global seismic wave propagation—I. Validation, *Geophys. J. Int.*, **149**, 390–412, 2002a.
- Komatitsch, D., and J. Tromp, Spectral-element simulations of global seismic wave propagation—II. Three-dimensional models, oceans, rotation and self-gravitation, *Geophys. J. Int.*, **150**, 308–318, 2002b.
- Kreyszig, E., *Advanced Engineering Mathematics*, Wiley, New York, 1988.
- Laske, G., Global observation of off-great-circle propagation of long-period surface waves, *Geophys. J. Int.*, **123**, 245–259, 1995.
- Laske, G., and G. Masters, Constraints on global phase velocity maps from long-period polarization data, *J. Geophys. Res.*, **101**, 16,059–16,075, 1996.
- Laske, G., G. Masters, and W. Zürn, Frequency-dependent polarization measurements of long-period surface waves and their implications for global phase-velocity maps, *Phys. Earth Planet. Inter.*, **84**, 111–137, 1994.

- Masuda, Y., and H. Ohnishi, An integration scheme of the primitive equation model with an icosahedral-hexagonal grid system and its application to the shallow-water equations, in *Short- and Medium-Range Numerical Weather Prediction*, edited by T. Matsumo, pp. 317–326, Japan Meteorological Society, 1987.
- Mooney, W. D., G. Laske, and G. Masters, CRUST-5.1: A global crustal model at $5^\circ \times 5^\circ$, *J. Geophys. Res.*, 103, 727–747, 1998.
- Phinney, R. A., and R. Burridge, Representation of the elastic-gravitational excitation of a spherical earth model by generalized spherical harmonics, *Geophys. J. R. Astron. Soc.*, 34, 451–487, 1973.
- Rancic, M., R. J. Purser, and F. Mesinger, A global shallow-water model using an expanded spherical cube: Gnomic versus conformal coordinates, *Q. J. R. Meteorol. Soc.*, 122, 959–982, 1996.
- Randall, D. A., T. D. Ringler, R. P. Heikes, P. Jones, and J. Baumgardner, Climate modeling with spherical geodesic grids, *Comp. Sci. Eng.*, 3, 32–41, 2002.
- Ringler, T. D., and D. A. Randall, A potential enstrophy and energy conserving numerical scheme for solution of the shallow-water equations on a geodesic grid, *Monthly Weather Review*, 130, 1397–1410, 2002a.
- Ringler, T. D., and D. A. Randall, The ZM grid: An alternative to the Z grid, *Monthly Weather Review*, 130, 1411–1422, 2002b.
- Ringler, T. D., R. P. Heikes, and D. A. Randall, Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores, *Monthly Weather Review*, 128, 2471–2489, 2000.
- Ritsema, J., H. J. van Heijst, and J. H. Woodhouse, Complex shear velocity structure imaged beneath Africa and Iceland, *Science*, 286, 1925–1928, 1999.
- Ronchi, C., R. Iacono, and P. S. Paolucci, The “cubed sphere”: A new method for the solution of partial differential equations in spherical geometry, *J. Comp. Phys.*, 124, 93–114, 1996.
- Sadourny, R., Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids, *Monthly Weather Review*, 100, 136–144, 1972.
- Sadourny, R., and P. Morel, A finite-difference approximation of the primitive equations for a hexagonal grid on a plane, *Monthly Weather Review*, 100, 439–445, 1969.
- Sadourny, R., A. Arakawa, and Y. Mintz, Integration of the non-divergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere, *Monthly Weather Review*, 96, 351–356, 1968.
- Saff, E. B., and A. B. J. Kuijlaars, Distributing many points on a sphere, *Math. Intelligencer*, 19(1), 5–11, 1997.
- Shearer, P. M., *Introduction to Seismology*, Cambridge U. Press, Cambridge, 1999.
- Stuhne, G. R., and W. R. Peltier, Vortex erosion and amalgamation in a new model of large scale flow on the sphere, *J. Comp. Phys.*, 128, 58–81, 1996.

- Stuhne, G. R., and W. R. Peltier, New icosahedral grid-point discretizations of the shallow water equations on the sphere, *J. Comp. Phys.*, 148, 23–58, 1999.
- Swarztrauber, P. N., D. L. Williamson, and J. B. Drake, The Cartesian method for solving partial differential equations in spherical geometry, *Dynamics Atmos. Oceans*, 27, 679–706, 1997.
- Sword, C. H., J. F. Claerbout, and N. H. Sleep, Finite-element propagation of acoustic waves on a spherical shell, in *Stanford Exploration Project (SEP)*, vol. 50, pp. 43–78, 1986a.
- Sword, C. H., J. F. Claerbout, and N. H. Sleep, Modeling global surface waves by a finite-element method, in *EOS Trans. Am. Geophys. Union*, vol. 67, pp. 1094–1095, 1986b.
- Tanimoto, T., Modelling curved surface wave paths: membrane surface wave synthetics, *Geophys. J. Int.*, 102, 89–100, 1990.
- Thuburn, J., A PV-based shallow-water model on a hexagonal-icosahedral grid, *Monthly Weather Review*, 125, 2328–2347, 1997.
- Trampert, J., and R. Snieder, Model estimations biased by truncated expansions: Possible artifacts in seismic tomography, *Science*, 271, 1257–1260, 1996.
- Trampert, J., and J. H. Woodhouse, Global phase velocity maps of Love and Rayleigh waves between 40 and 150 seconds, *Geophys. J. Int.*, 122, 675–690, 1995.
- Trampert, J., and J. H. Woodhouse, High resolution global phase velocity distribution, *Geophys. Res. Lett.*, 23, 21–24, 1996.
- Trampert, J., and J. H. Woodhouse, Assessment of global phase velocity models, *Geophys. J. Int.*, 144, 165–174, 2001.
- Trefethen, L. N., Group velocity in finite difference schemes, *SIAM Review*, 24, 113–136, 1982.
- Trefethen, L. N., Finite difference and spectral methods for ordinary and partial differential equations, available at <http://web.comlab.ox.ac.uk/oucl/work/nick.trefethen/pdetext.html>, 1996.
- Trefethen, L. N., *Spectral Methods in Matlab*, SIAM, Philadelphia, Penn., 2000.
- Tromp, J., and F. A. Dahlen, Variational principles for surface wave propagation on a laterally heterogeneous Earth—I. Time-domain JWKB theory, *Geophys. J. Int.*, 109, 581–598, 1992a.
- Tromp, J., and F. A. Dahlen, Variational principles for surface wave propagation on a laterally heterogeneous Earth—II. Frequency-domain JWKB theory, *Geophys. J. Int.*, 109, 599–619, 1992b.
- Um, J., and F. A. Dahlen, Source phase and amplitude anomalies of long-period surface waves, *Geophys. Res. Lett.*, 19, 1575–1578, 1992.
- Wang, Z., and F. A. Dahlen, JKWB surface-wave seismograms on a laterally heterogeneous earth, *Geophys. J. Int.*, 119, 381–401, 1994.

- Wang, Z., and F. A. Dahlen, Spherical-spline parameterization of three-dimensional Earth, *Geophys. Res. Lett.*, 22, 3099–3102, 1995a.
- Wang, Z., and F. A. Dahlen, Validity of surface-wave ray theory on a laterally heterogeneous Earth, *Geophys. J. Int.*, 123, 757–773, 1995b.
- Wang, Z., F. A. Dahlen, and J. Tromp, Surface wave caustics, *Geophys. J. Int.*, 114, 311–324, 1993.
- Wang, Z., J. Tromp, and G. Ekström, Global and regional surface-wave inversion: A spherical-spline parameterization, *Geophys. Res. Lett.*, 25, 207–210, 1998.
- Williamson, D. L., Integration of the barotropic vorticity equation on a spherical geodesic grid, *Tellus*, 20, 642–653, 1968.
- Williamson, D. L., Numerical integration of fluid flow over triangular grids, *Monthly Weather Review*, 97, 885–895, 1969.
- Williamson, D. L., Integration of the primitive barotropic model over a spherical geodesic grid, *Monthly Weather Review*, 98, 512–520, 1970.
- Williamson, D. L., J. B. Drake, J. J. Hack, R. Jakob, and P. N. Swarztrauber, A standard test for numerical approximations to the shallow water equations in spherical geometry, *J. Comp. Phys.*, 102, 211–224, 1992.
- Wong, Y. K., Upper Mantle Heterogeneity from Phase and Amplitude Data of Mantle Waves, Ph.D. thesis, Harvard University, 1989.
- Woodhouse, J. H., Surface waves in a laterally varying layered structure, *Geophys. J. R. Astron. Soc.*, 37, 461–490, 1974.
- Woodhouse, J. H., Long period seismology and the Earth's free oscillations, in *Seismic Modelling of Earth Structure*, edited by E. Boschi, G. Ekström, and A. Morelli, pp. 31–80, Editrice Compositori, Bologna, Italy, 1996.
- Woodhouse, J. H., and A. M. Dziewonski, Mapping the upper mantle: Three-dimensional modeling of Earth structure by inversion of seismic waveforms, *J. Geophys. Res.*, 89, 5953–5986, 1984.
- Woodhouse, J. H., and Y. K. Wong, Amplitude, phase and path anomalies of mantle waves, *Geophys. J. R. Astron. Soc.*, 87, 753–773, 1986.