On mathematical modelling and parameter estimation of seismic media

by

© **Theodore Stanoev**

A dissertation submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Department of Earth Sciences

Memorial University of Newfoundland

October 2024

St. John's, Newfoundland and Labrador, Canada

Abstract

Within the multidisciplinary field of geophysics, seismology is a branch that accounts for the generation and propagation of vibrations within the Earth, otherwise referred to as seismic waves. To study the mechanical properties of these waves, seismologists model the Earth's subsurface as a continuous elastic medium. Such an approximation facilitates the interpretation of physical observations within a mathematical framework, which can be approached either as forward or inverse problem. To that end, this dissertation is comprised of two forward problems and one inverse problem.

From a forward perspective, theoretical models are proposed based on *a priori* assumptions of mechanical properties of the subsurface, which can be quantified as changes in velocity with location (inhomogeneity) or direction (anisotropy). Two chapters of this dissertation reside within this context and are applied to homogeneous anisotropic media. In the first of these chapter, we determine the conditions for elliptical roots of the Christoffel equation in media that are the result of the Backus average. Within these conditions, we demonstrate that the slowness surfaces are nondetached. In the second chapter, we present a novel formulation for the purpose of forward modelling traveltimes. Through the Taylor expansion along vertical rays in a horizontally stratified Earth model, we obtain a homogeneous transversely isotropic medium within which the traveltimes are similar to the Fermat traveltimes of its constituent layers.

From an inverse perspective, the parameters of the theoretical models are estimated so as to provide an agreement with physical observations. In this dissertation, one chapter resides within this context, where we perform an inversion on traveltime measurements acquired from a vertical seismic profile, otherwise referred to as field data. We implement a derivative-free approach to minimize the residual sum of

squares between the measurements and the model. Since field data are not necessarily complete and can be subject to measurement errors, we conduct a simulation study on synthetically generated traveltimes to assess the accuracy of our estimates. Then, we apply our approach to the field data to estimate the inhomogeneity and anisotropy of the subsurface.

To my parents, George and Pavlina,
who taught me that education can open
many doors in life. Here's to finding out
what doors a Ph.D. can open . . .

Preface

We've all heard it before: "So, when do you finish your Ph.D.?" It's a benign question that well-intentioned friends and family members pathologically pose to their resident doctoral candidates. If by some circumstance you've managed to set eyes on this preface, it's quite likely that you have already been through these trials, obtained your doctorate, and are now reminiscing on those days when you were on the receiving end of this question.[1] You might recall that your response varied from week to week, often depending on: how many times you've been recently asked, how much sleep you've been getting, whether or not you've had any exposure to sunlight as of late (the list goes ever on)... In all seriousness, however, your response ultimately depended on how close you felt you were to finishing that dissertation.

Admittedly, anyone that has embarked down this path and completed the task would know that you never *actually* finish. We — curious, perfectionist, academic types — are afflicted by an internal conviction (or, at times, compulsion) to continually improve upon our work and, in so doing, we believe that we can promote self-growth. This affliction forces us to grapple with our inner dialogue: "Maybe you should clear up that one iffy explanation with a bit more elaboration... Do you *really* think the examiners *won't* ask you that on defence day?"; "How about touching up just a little bit more that last figure you finished... What's another hour at this point?"; "Oh, and do you remember that one assumption we made a year ago? Yeah, maybe you should redo all of your calculations for the tenth time, just in case there's a glitch..."

Thankfully, our doctoral supervisor is often there to remind us that it's not only okay to quiet that voice, but it's entirely necessary. In the words of my supervisor, Dr.

[1] If you're that friend or family member who decided to follow up and want to give reading this document the ol' college try, well, kudos to you — I hope you have a sweet spot for mathematical modelling (or me)!

Michael A. Slawinski, "at some point, you have to let it go" (and hope it'll be enough). Well, for those of you that have been with me as I've walked along this path, rest assured: I've finally decided to let it go — truly, I hope it will be enough.

Theodore Stanoev
St. John's, NL, Canada
October 2024

General summary

The focus of this dissertation is on mathematical modelling and parameter estimation, using numerical optimization, applied to a geophysical setting within the domain of seismology. In particular, we assume that the Earth's subsurface can be modelled as a series of thin horizontally stratified layers. The properties of each layer affect the macroscopic behaviour of a seismic wave travelling therein. With the aid of a mathematical model, we replace large portions of the subsurface that have similar properties with homogenized sections, whose averaged macroscopic properties are determined by their constituent layers. Within this context, we study the mathematical properties of the homogenized media, the traveltimes along the rays of the seismic waves therein, and the estimation of parameter values to ensure an agreement between a model and data measurements.

The main chapters of the dissertation are based on three research projects. In the first, we conduct a theoretical investigation to determine the necessary conditions for a specific type of wavefront to arise in a homogenized medium that is the result of a well-known seismological average. In the second, we formulate a novel mathematical model for which the traveltimes in the homogenized medium are similar to those of its constituents. In the third, we estimate the parameter values using numerical optimization to account for measured traveltimes acquired from a vertical seismic profile.

Acknowledgements

Throughout the years of my graduate experience, I have been fortunate to meet and work alongside many wonderful people! I hope the following finds you well.

In this extended format, I am pleased to express my gratitude and respect for my doctoral supervisor, Dr. Michael A. Slawinski. In 2015, I told Dr. Slawinski that I hoped he would not only be my supervisor but he could also become my mentor. Since then, Dr. Slawinski has gone above and beyond at every stage of my academic development. From providing me with the resources to conduct my studies to facilitating my international collaborations and conference presentations to giving me the freedom to pursue my independent research, Dr. Slawinski has always tried to provide me opportunities to grow. From trusting my editorial capacity in the context of his textbooks to facilitating my attendance at the courses he has taught internationally to providing a historical context for our discussions across a wide range of topics, Dr. Slawinski has always tried to broaden my horizons. From demonstrating how to handle adverse situations as an academic (as well as a gentleman) to showing me that the best way to stay atop your game is to never get out of those mathematical trenches to embracing the notion that "being comfortable with confusion" is a necessary component to academic success, Dr. Slawinski has always lead by example and inspired me to do the same. In short, Dr. Slawinski has exceeded my expectations as a mentor and supervisor. I look forward, with great enthusiasm, to our continued collaboration and friendship.

Alongside my supervisor, I would like to thank two of my collaborators, with whom I have conducted research that appears in this dissertation. To Dr. Len Bos, thank you for being my academic champion and role model; you have played an essential role in my development by showing me your mathematical rigour, which I continually strive to emulate. To Mr. Ayiaz Kaderali, thank you for your attention to detail, for

helping me to learn the importance of a simulation study, and for our many Zoom meetings that have helped to form our friendship.

Now, to my supervisory committee, Dr. Mikhail V. Kotchetov and Dr. Alison Malcolm, thank you for your encouragement and helpful revisions at the final stages of the dissertation. Your guidance, patience, and support during the revision process have been invaluable and greatly appreciated.

Next, I would be remiss to not thank the following professors that have left their mark on my development. I would like to thank: Dr. Tomasz Danek, for demonstrating the value of scientific computing; Dr. Ivan Booth, for inspiring my interest in differential geometry; Dr. Kristin M. Poduska, for helping me to find direction toward the end of my undergraduate; Dr. Robert Sarracino, for meeting with me to discuss the "Scientific method in brief"; Dr. Ran Bachrach, for his wise words during my attendance at the 18th International Workshop on Seismic Anisotropy (18IWSA); Dr. Piotr Stachura, for demonstrating the versatility of a mathematical structure in conducting interdisciplinary research; Dr. Karem Azmy, for helping me to set this entire graduate experience in motion by introducing me to Dr. Slawinski.

Also, I would like to thank collaborators that have not yet been mentioned: Mr. Md Abu Sayed, Dr. Filip P. Adamus, Mrs. Izabela Adamus, Dr. David R. Dalton, Dr. Bartosz Gierlach, Dr. Raphaël A. Slawinski, and Dr. Maurizio Vianello. In addition, I would like to thank: Dr. Maria Kotsi, for her reliable advice and friendship; Dr. Zvi Koren, for being a gracious host following my presentation at 18IWSA; and Mrs. Elena Patarini Slawinski, for her graphical support, but, most importantly, for her passion for culture, languages, and everything in between.

Beyond these personal declarations, I would like to acknowledge the financial support provided to me by Memorial University of Newfoundland in the form of the five-year School of Graduate Studies Fellowship, the Chevron Canada Limited Rising Star Award that I won twice, and The Geomechanics Project supported by Husky Energy and NSERC grants.

Lastly, to my family and friends... You, above all, have been the ones that have dealt with the consequences of the sacrifices that I made toward the completion of this dissertation. With deep respect and sincerity, I thank you for sticking with me throughout these years.

Co-authorship statements

This dissertation is based on three research collaborations and focuses on mathematical modelling and parameter estimation, using numerical optimization, in the context of seismic media. In view of the interdisciplinary nature of Earth Sciences, the collaborations have been with members of Dr. Michael A. Slawinski's research group, The Geomechanics Project: the first two were conducted with professors Dr. Len Bos (*Universitá degli studi di Verona*, Italy) and Dr. Slawinski (my doctoral supervisor); the third with Mr. Ayiaz Kaderali, who is another Ph.D. candidate within the research group.

For the collaborative projects, the established norm of the research group is to engage in alphabetical authorship lists, which is a common practice in mathematical journals. Motivated by the spirit of collaboration, we do not invoke the notion of a principal author. Likewise, in an effort to promote parity among colleagues, each aspect of the research, analysis, and manuscript preparation has been conducted jointly — as such, it is difficult to separate distinct contributions.

With that being said, I acknowledge that, for the purpose of fulfilling the requirements of a doctoral degree, a clear statement should be provided concerning my contribution to each portion of the dissertation that is the result of a collaborative effort. Following the statement of the School of Graduate Studies regarding co-authorship, and for the benefit of my examiners, I shall proceed to address my contributions made to each collaborative effort.

Collaborative research

On Christoffel roots for nondetached slowness surfaces

Chapter 3 is based on Bos et al. (2019a), which is a peer-reviewed publication, "On Christoffel roots for nondetached slowness surfaces", by Len Bos, Michael A. Slawinski, and Theodore Stanoev. The project was conducted between the Fall of 2018 and the Spring of 2019.

Following my presentation, "On traveltimes in Backus media", at the 18th International Workshop on Seismic Anisotropy (18IWSA), Dr. Slawinski suggested that we launch our investigation into the roots of the Christoffel equation in Backus media, with particular attention to the conditions that lead to nondetachment. Soon after, Dr. Bos joined the research project and, by February of 2019, we had obtained our results: the presented Lemma 1 was devised by Dr. Bos whereas Proposition 2 was borne out of my geophysical conversations with Dr. Slawinski. As the sole student collaborator, I prepared a first draft of our research, from which point we collaboratively finalized the manuscript. In March of 2019, I posted our manuscript to arXiv and submitted to Geophysical Prospecting. The article appears in the November 2019 Special Issue: Advances in Seismic Anisotropy, which includes the proceedings of 18IWSA.

For the purpose of this traditional-format dissertation, I have enhanced the manuscript with supplementary derivations and explanatory statements for this dissertation.

On forward modelling of traveltimes in approximate media

Chapter 4 is based on a collaborative research project with Dr. Len Bos and Dr. Slawinski, conducted between July of 2019 and April of 2021.

Following the May 2019 acceptance of Bos et al. (2019a), I still sought to understand better the traveltimes in Backus media. On the recommendation of Dr. Slawinski, I organized a research trip to visit Dr. Bos in Verona during July of 2019, where had the insight to leverage the Taylor expansion of vertical traveltimes as a means to answer my questions. During the trip, I worked with Dr. Bos to derive the mathematical expressions of the approximate media. After my stay, and for the majority of 2020, I continued our collaboration remotely to devise the numerical examples. As with Bos et al. (2019a), I prepared the first draft and then, collaboratively, we finalized the

manuscript.

For the purpose of this traditional-format dissertation, I have adjusted the manuscript to flow within this single, unified document.

On inverse modelling of traveltimes

Chapter 5 is based on a collaborative research project with Mr. Ayiaz Kaderali. The research project was conducted collaboratively between the Fall of 2019 and the Fall of 2022, and independently between November 2023 and May 2024.

In 2019, I sought to gain practical research experience with seismic traveltimes. To that end, Dr. Slawinski recommended that I collaborate with Mr. Kaderali, who has extensive practical experience as an expert industry geophysicist and uses vertical seismic profile (VSP) traveltimes in his doctoral research. From that point, every decision concerning the project's design and implementation has been made jointly by Mr. Kaderali and me. Since we consider ourselves as equal contributors, we have chosen to include the research results in our respective dissertations. To remain in accordance with Memorial University of Newfoundland guidelines, I have converted my dissertation from a manuscript to traditional format and written Chapter 5 independently of Mr. Kaderali.

Regarding these contributions, Dr. Slawinski recommended the research topic whereas the details of a VSP and the notion of a simulation study were brought forth entirely by Mr. Kaderali. I spearheaded the theoretical efforts to adjust the raytracing expression for a multilayer setting, to learn numerical optimization for the parameter estimation, and to devise penalization strategies to constrain model-parameter values and ensure model feasibility. The validation of the optimization programs were done in a joint manner of equal contribution. The analysis of the optimization results and interpretation with regard to their correspondence to the Earth's subsurface were done in a joint manner.

In 2020, we consolidated our research up to that point and posted a preprint to arXiv (Kaderali and Stanoev, 2020). Thereafter, we continued our research efforts into the Fall of 2022, the results of which were used in my original submission in the Fall of 2023. Between November 2023 and May 2024, I have made substantial revisions to the research project in response to examiner critiques, which were done independently.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

Throughout this dissertation, we focus on the mathematical modelling and parameter estimation of seismic media. In Section 1.1, we provide a general overview of the dissertation, with a description of the modelling set up and techniques used for homogenization and parameter estimation. In Section 1.2, we provide a chapter-by-chapter outline of the dissertation; we include a separate literature review for each chapter.

1.1 General overview

This dissertation is built upon physical concepts, expressed in a mathematical language, and is applied, through computational techniques, to seismology. A key motivation for the advancement of theoretical seismology is to bridge the gap between mathematical analogies that allow us to study physical phenomena and their material properties. In other words, we use mathematical analogies to model responses of the Earth expressed by seismic data.

To that effect, we use mathematical models to study the homogenization of inhomogeneous media and use numerical optimization to estimate the parameter values of these models to account for measured traveltimes along seismic rays therein. The advantages of homogenization are — at least — two-fold. On the one hand, with regard to forward modelling, homogenization can provide closed-form expressions that permit simplified raytracing and traveltime expressions related to the seismic phenomena. On the other, with regard to inverse modelling, the reduced

number of model parameters required to adequately model the physical phenomena lends itself to investigations regarding parameter estimation.

The process of replacing portions of heterogenous media with homogeneous sections is known as homogenization or seismic upscaling (e.g., Gold et al., 2000) — throughout this dissertation, we refer to the process by the former term. To facilitate homogenization, we model inhomogeneous media by a series of thin horizontally stratified Earth layers. The elastic properties of each thin layer, denoted by homogeneous Hookean solids of varying symmetry classes, determine the behaviour of seismic wave propagation within the model. We restrict our attention to down-going long-wavelength seismic waves and assume that the layer properties vary in a gradual manner. As a consequence, many thin layers are encompassed within the seismic wave and their microscopic properties affect wave propagation at a macroscopic level. In other words, within the context of long wavelengths, we may average a section of stratified Earth layers by a single equivalent medium.

In general, not all homogenizations are the same and depend greatly upon the underlying assumptions of the formulation. Herein, we consider three mathematical formulations of homogenizations. The first is the so-called Backus average (Backus, 1962), which replaces a series of layers by a single (approximately) equivalent medium that is said to be "long-wave equivalent" to the original. The second is a novel Taylor series approximation of traveltimes about the vertical axis that replaces a series of layers by a single approximate medium, wherein the traveltimes are similar to the Fermat traveltimes within its constituent layers. The third is the traveltime model by Slawinski et al. (2004) that accounts for a linear increase in velocity with depth and anisotropy that is the result of an elliptical velocity dependence. We refer to the media obtained from these homogenizations as *Backus*, *approximate*, and *ab*χ media, respectively. They are the background models used in this dissertation. Within the context of these media, we use the homogenizations as background models through which we investigate aspects of the subsurface. We do so from a forward and inverse perspective, with each perspective constituting one of the two parts that comprise this dissertation.

Residing within the framework of a mathematical model, on the one hand, we derive the conditions to observe a specific phenomenon to occur within a given model and, on the other, design a new model subject to specific constraints. In both cases, we

analyze the outputs of a model based on specified inputs, hence, a forward perspective. For instance, in Chapter 3, we investigate slowness surfaces in transversely isotropic (TI) media. The motivation of this chapter is to assess the claim that in TI media that are the result of the Backus average applied to isotropic layers, which is the first homogenization that we consider, the qP wavefront is never ellipsoidal (Helbig, 1983). While this statement likely holds in most practical settings that pertain to seismology, we revisit the so-called Christoffel equations, which govern the existence of seismic waves within elastic media, to not only derive the necessary conditions for a counterexample, but also to provide a numerical example as a demonstration. In Chapter 4, we formulate the novel approximate medium, which is the second homogenization that we consider, for the purpose of forward modelling Fermat traveltimes. The motivation of this chapter is to devise a model that is consistent with a larger mathematical framework while subjected to a constraint. As it pertains to seismology, the homogenization is a tool through which the traveltimes of Backus media can be cross examined. Likewise, it can replace sections of the subsurface within a larger multilayer structure.

Residing within a given set of data, on the one hand, we estimate the parameters of a given model that improve the agreement between data and model and, on the other, choose the most appropriate to represent the data given a set of model candidates. In both cases, we analyze a model and its agreement with the data, hence, an inverse perspective. In Chapter 5, we extend the formulation of Slawinski et al. (2004) to account for a multilayer application of the $ab\chi$ traveltime model to a set of vertical seismic profile (VSP) traveltime measurements. The motivation of this chapter is to estimate the model-parameter values in order to make inferences of the properties of the subsurface. To that effect, we restrict our attention to the Nelder-Mead algorithm, which is an unconstrained, derivative-free method that only uses function evaluations (Nelder and Mead, 1965). We perform the optimization through the reduction of the residual sum of squares (RSS), which ensures an agreement between the measured and modelled traveltimes. First, we devise a simulation study to assess the extent to which measurement error affects the reliability of the estimation. For that purpose, we use simulated traveltimes, generated with known parameter values, and several noise profiles. Then, with an understanding of the accuracy of the approach, we apply the optimization to the actual VSP measurements.

In summary, this dissertation is focused on using mathematical modelling and parameter estimation, using numerical optimization, to study homogenizations of seismic media and the traveltimes therein. We rely on the framework of mathematics to answer questions of: i) forward modelling, concerning slowness surfaces and approximate-traveltime homogenizations, and ii) inverse modelling, concerning parameter estimation and model selection as it pertains to a given data set. Together the two modelling perspectives permit investigations that may be conducted in seismic media.

1.2 Dissertation outline

This dissertation is written in a traditional format and is based on the results of four research projects. In accordance with Memorial University of Newfoundland regulations, the presentation of the research resembles that of a "proto-book" that includes this introduction, four chapters, and a conclusion; the remaining pages constitute the front and back matter, the latter of which contains two appendices. Each portion of this dissertation is written in a consistent writing style and is focused on the common theme of mathematical modelling and parameter estimation, using numerical optimization, in the context of seismic media. Now, let us outline the four chapters of the main text.

In Chapter 2, we provide the essential background for the mathematical theories that we use within this dissertation. We introduce balance equations that are based on the first principles of continuum mechanics. Then, we demonstrate aspects of the derivation of seismic ray theory, with a particular focus on the assumptions of the linearized stress-strain relationship, plane-wave propagation, and high-frequency approximation.

In Chapter 3, we consider roots of the Christoffel equation for seismic waves in TI media. We determine the conditions for elliptical roots within Backus media and demonstrate our findings with a numerical example. Also, as a result of our investigation, we provide a qualifier for the ellipticity condition of Thomsen (1986) and provide illustrative examples to that effect.

In Chapter 4, we formulate a novel approximate medium within which the traveltimes are similar to the Fermat traveltimes of its constituent layers. We

perform a Taylor series expansion of the Fermat traveltimes about the vertical axis and derive formulæ for the elasticity parameters to ensure their similarity with their counterparts in the homogenized medium. We present numerical examples within isotropic and TI approximate media along with a comparison of the traveltimes therein to other homogenized media, including the Backus medium.

In Chapter 5, we apply the traveltime model of Slawinski et al. (2004) to accommodate VSP traveltime measurements in a multilayer setting. To ensure an agreement between measured and modelled traveltimes, we reduce the RSS with numerical optimization. We conduct a simulation study to assess the reliability of the parameter estimation for simulated traveltimes, generated with known model-parameter values. We apply the optimization, using a multistart approach, to various velocity models under the presence of simulated noise profiles. With an understanding of the accuracy of the approach, we proceed with the estimation for the measured VSP traveltimes and interpret our results.

Additionally, we include an appendix that contains supplementary material to the main text.

Chapter 2

Theoretical background

In this dissertation, our purpose is to gain insight into the structure of the Earth's subsurface through the propagation of seismic waves. To facilitate this purpose, we require a mathematical framework to characterize the subsurface and quantify waves propagating therein. This framework is based on several foundational principles and simplifying approximations, which we overview as follows.

From the onset, we require continuum mechanics for the fundamental balance principles that apply to any material and must be satisfied at all times. To apply these principles to a seismological context, we apply constitutive equations, based on Hooke's Law, that govern the material properties of a linearized elastic medium. Using this medium as the background model for the subsurface, we derive the elastodynamic equation, which quantifies the propagation of a seismic disturbance in such a medium. We proceed with a ray theory approach that is based on a high-frequency approximation of the elastodynamic equation. Finally, the resulting mathematical equations permit the computation of geometric raypaths and the associated traveltime along them.

In this chapter, we begin with a high-level seismological overview as well as provide commentary on the mathematical notation used in this dissertation. Then, we proceed to discuss the details of the aforementioned concepts. We provide references to the literature where applicable.

2.1 Seismological setting

Throughout this dissertation, we use a common background model that consists of series of horizontally stratified Earth layers. The elastic properties of each layer are determined by homogeneous Hookean solids. These solids are mathematical analogies for physical materials that obey Hooke's law, which amounts to a linearized stress-strain relationship. The elasticity constants of these solids can belong to a variety of symmetry classes and they influence the propagation of seismic waves, which is determined by the elastic wave equation, otherwise referred to as the elastodynamic equation.

An important aspect of seismic wave propagation is the velocity of the wave, which can vary depending on the direction of its travel. This variation is referred to as seismic anisotropy and it can be produced by several factors within the subsurface: in particular, oriented cracks or pores, layering, or preferential alignment of constituent materials. In many cases, all three sources simultaneously contribute to the development of anisotropy (Cholach and Schmitt, 2003).

In Chapters 3 and 4, we consider homogenizations that belong to the symmetry class of transverse isotropy (TI), which involves elastic properties that are the same in any direction perpendicular to a symmetry axis (Sheriff, 2002, p. 365). The purpose of these homogenizations is to replace a series of layers within the background model, wherein the anisotropy is induced by layering. The constituent layers belong to the symmetry class of isotropy, which involves elastic properties that are the same in all directions.

In Chapter 5, we do not specify the symmetry classes of constituent layers but, rather, consider the macroscopic behaviour as a result of their compounding anisotropic effects. We focus on media that are vertically inhomogeneous and elliptically anisotropic. The former specifies that the seismic velocity depends only on the depth coordinate, which is an apposite modelling assumption as velocity can increase with depth due to overburden compaction. The latter specifies that the horizontal velocity is a scalar multiple of the vertical, which is a suitable manner in which to model the preferential alignment of materials. Since the velocity in this type of media depends only on one coordinate, it is sometimes referred to as one-dimensional (1D).

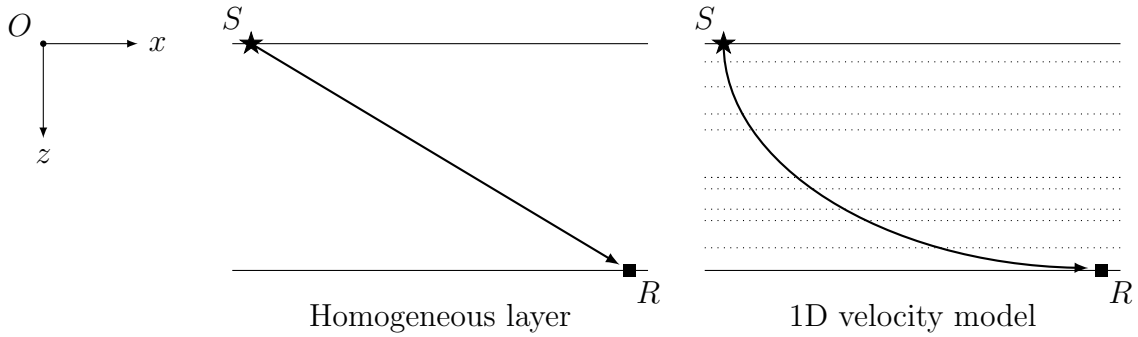


Figure 2.1: Rays connecting point sources, S, and receivers, R, at specified depths, denoted by stars and squares, respectively. Within the cross section, the rays travel a horizontal distance, x, and a vertical distance, z. On the left, the ray in the homogeneous layer is a straight line. On the right, the ray bends as a result of increasing velocities through the series of homogeneous layers (1D velocity model), whose interfaces are denoted by dotted lines.

In each of these chapters, we approximate the propagation of the wave using seismic ray theory, which is a well-established procedure based on an asymptotic solution of the elastodynamic equation (e.g., Červený, 2001, p. 1). This approximation leads to useful expressions related to plane waves that permit the computation of seismic rays and the traveltimes associated with them. In particular, these expressions apply to a plane wave that is generated from a point source, S, from which is possible to trace a path — known as the raypath — that connects to the location of a known receiver, R. The ray at an interface obeys Snell's Law, which governs the angles of reflection and refraction of with respect to an incident wave. The raypath between any two points follows Fermat's principle, which refers to a stationary traveltime between the points. We present an example of such properties in Figure 2.1.

2.2 Mathematical notation

At times in this dissertation, we adopt the so-called Einstein summation convention (e.g., Holzapfel, 2000, Section 1), which is useful for representing dot, cross, and tensor products. For example, any vector $\mathbf{u}$ in the three-dimensional Cartesian coordinate system is represented as

$$\mathbf{u} = \sum_{i=1}^{3} u_i \, \mathbf{e}_i = u_1 \, \mathbf{e}_1 + u_2 \, \mathbf{e}_2 + u_3 \, \mathbf{e}_3, \tag{2.1}$$

where each term is the product of a component of vector $\mathbf{u}$ along the given direction of the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$. By the summation convention, the notation of vector (2.1) is shortened as

$$\mathbf{u} = u_i \, \mathbf{e}_i, \tag{2.2}$$

where the repeated indices i indicate a summation. The dot product of basis vectors is expressed by the Kronecker delta, where

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij} := \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} . \tag{2.3}$$

Thus, using notation (2.3), the dot product of two arbitrary vectors is

$$\mathbf{u} \cdot \mathbf{v} = (u_i \, \mathbf{e}_i) \cdot (v_j \, \mathbf{e}_j) = u_i v_j (\mathbf{e}_i \cdot \mathbf{e}_j) = u_i v_j \delta_{ij} = u_i v_i. \tag{2.4}$$

For the cross product of basis vectors,

$$\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk} \, \mathbf{e}_k, \tag{2.5}$$

we require the so-called permutation symbol

$$\varepsilon_{ijk} := \begin{cases} 1, & \text{for even permutations of indices, i.e., (123, 231, 312)} \\ -1, & \text{for odd permutations, i.e., (132, 321, 213)} \\ 0, & \text{for repeated indices} \end{cases} . \tag{2.6}$$

Using notations (2.5) and (2.6), the cross product of two arbitrary vectors is

$$\mathbf{u} \times \mathbf{v} = (u_i \, \mathbf{e}_i) \times (v_j \, \mathbf{e}_j) = u_i v_j (\mathbf{e}_i \times \mathbf{e}_j) = \varepsilon_{ijk} u_i v_j \, \mathbf{e}_k. \tag{2.7}$$

For the tensor product of vectors $\mathbf{u}$ and $\mathbf{v}$, it is defined as a second order tensor that transforms a third vector $\mathbf{w}$ into a vector in the direction of $\mathbf{u}$, whereby

$$(\mathbf{u} \otimes \mathbf{v}) \, \mathbf{w} := (\mathbf{v} \cdot \mathbf{w}) \, \mathbf{u} = \mathbf{u} \, (\mathbf{v} \cdot \mathbf{w}) . \tag{2.8}$$

Another purpose of the tensor product is to represent a tensor $\mathbf{A}$, where, for example, a second-order tensor is

$$\mathbf{A} = A_{ij} \, \mathbf{e}_i \otimes \mathbf{e}_j \tag{2.9}$$

and the components are represented as

$$A_{ij} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}.$$

Along with the Einstein summation convention, we require several identities (e.g., Holzapfel (2000, p. 51) or Gurtin et al. (2010, p. 46)). Specifically, we consider relationships among the gradient and divergence operators, an arbitrary scalar quantity, φ, and arbitrary vector quantities, $\mathbf{u}$, $\mathbf{v}$:

$$\operatorname{grad}(\varphi\mathbf{u}) = \mathbf{u} \otimes \operatorname{grad}\varphi + \varphi \operatorname{grad}\mathbf{u}, \tag{2.10}$$

$$\operatorname{div}(\varphi\mathbf{u}) = \varphi \operatorname{div}\mathbf{u} + \mathbf{u} \cdot \operatorname{grad}\varphi, \tag{2.11}$$

$$\operatorname{div}(\mathbf{u} \otimes \mathbf{v}) = (\operatorname{grad}\mathbf{u})\mathbf{v} + \mathbf{u} \operatorname{div}\mathbf{v}. \tag{2.12}$$

In identities (2.10)–(2.12), the scalar and vector functions assign a scalar and vector, respectively, for each point $\boldsymbol{x} \in \mathbb{R}^3$ within some region of the coordinate system. As such, it is useful to refer to these functions as scalar, $\varphi(\boldsymbol{x})$, and vector, $\mathbf{u}(\boldsymbol{x})$, fields; this applies to tensor fields, $\mathbf{A}(\boldsymbol{x})$, as in expression (2.9).

To express the gradient and divergence of the scalar, vector, and tensor fields, we use the Nabla operator, ∇, and follow the conventions of e.g. Holzapfel (2000, Section 1.8). For a scalar field, $\varphi(\boldsymbol{x})$, the components of the gradient are

$$\operatorname{grad}\varphi := \nabla\varphi = \frac{\partial\varphi}{\partial x_j}\,\mathbf{e}_j.$$

However, for an arbitrary vector or tensor field, the gradient and divergence rely on the dot and tensor products of ∇ with the given field. Using the tensor product, the gradient of a vector field, $\mathbf{u}(\boldsymbol{x})$, is

$$\operatorname{grad}\mathbf{u} := \nabla \otimes \mathbf{u} = \frac{\partial u_i}{\partial x_j}\,\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.13}$$

and the gradient of a tensor field, $\mathbf{A}(\boldsymbol{x})$, is

$$\operatorname{grad}\mathbf{A} := \nabla \otimes \mathbf{A} = \frac{\partial A_{ij}}{\partial x_k}\,\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k.$$

Using the dot product, the divergence of a vector field is

$$\operatorname{div} \mathbf{u} := \nabla \cdot \mathbf{u} = \frac{\partial\, u_i}{\partial x_j}\, \mathbf{e}_i \cdot \mathbf{e}_j = \frac{\partial\, u_i}{\partial x_i}$$

and the divergence of a tensor field is

$$\operatorname{div} \mathbf{A} := \nabla \cdot \mathbf{A} = \frac{\partial\, A_{ij}}{\partial x_k}\, (\mathbf{e}_i \otimes \mathbf{e}_j) \cdot \mathbf{e}_k = \frac{\partial\, A_{ij}}{\partial x_k} \delta_{jk}\, \mathbf{e}_i = \frac{\partial\, A_{ij}}{\partial x_j}\, \mathbf{e}_i.$$

Also, for the sake of brevity, at times we use a comma notation for partial derivatives, akin to Misner et al. (1973, Section 2.7),

$$(\circ)_{,i} := \frac{\partial\, (\circ)}{\partial x_i} \quad \text{and} \quad (\circ)_{,ji} := \frac{\partial^2 (\circ)}{\partial x_j \partial x_i}. \tag{2.14}$$

2.3 Continuum mechanics

To quantify the propagation of seismic waves in the Earth, we rely on the framework of continuum mechanics as an analogy for the subsurface. By virtue of its nomenclature, continuum mechanics incorporates the language of mathematics and the notion of a continuous medium to study physical phenomena. The former depends upon the mathematical machinery of calculus, linear algebra, and tensors. The latter assumes a smooth material, whose disregard of discrete, microscopic particles evaluates the macroscopic response of a given input.

The crux of a continuous medium is that properties averaged over a small element within the medium are assumed to vary continuously with position. This idealization is permissible because the lengths characterizing the microscopic structure are generally much smaller than any lengths arising in the deformation of the medium (Achenbach, 1973, p. 3). Under such an assumption, continuum mechanics establishes the governing laws of balance of mass, linear and angular momentum, which hold for any material. The mechanical behaviour of particular materials depends on constitutive relations, which complement the conservation laws and describe the mechanical response to an applied load.

Within continuum mechanics, the configuration of a continuum body refers to an infinite set of points that are assigned to unique positions embedded in a

three-dimensional coordinate system. There are two configurations: reference and current; the former corresponds to an initial time $t = 0$; the latter corresponds to the configuration of the body under the action of a motion, otherwise known as a change in shape, position, or orientation, at a subsequent time $t > 0$. During the motion, each point in the reference configuration has a one-to-one correspondence in the current configuration. The associated coordinate systems for the two configurations are known as material and spatial, respectively.

2.3.1 Master balance principle

At any given instant in time t, the status of the configuration of a continuum body, occupying an arbitrary fixed region Ω with a surface boundary $\partial\Omega$, is summarized by the so-called master balance principle (Holzapfel, 2000, Section 4.9)

$$\frac{\mathrm{D}}{\mathrm{D}t}\int_{\Omega} f(\boldsymbol{x},t)\,\mathrm{d}v = \int_{\partial\Omega} \boldsymbol{\phi}(\boldsymbol{x},t,\mathbf{n})\,\mathrm{d}s + \int_{\Omega} \boldsymbol{\Sigma}(\boldsymbol{x},t)\,\mathrm{d}v, \tag{2.15}$$

where $\boldsymbol{x}$ is a position within the region, t is an instant in time, and $\mathrm{d}s$ and $\mathrm{d}v$ refer to an infinitesimal element on the surface boundary or within the volume of the region, respectively. In the following paragraphs, wherein we define quantities[1] f, $\boldsymbol{\phi}$, and $\boldsymbol{\Sigma}$, we use expression (2.15) to provide the fundamental statements of the balance of mass, balance of linear momentum, and balance of angular momentum, which apply to any continuous medium. With these balances, we obtain the so-called Cauchy equations of motion, which are crucial for the quantification of our seismological studies.

To understand expression (2.15), let us address each term therein. The left-hand side represents the rate of change per unit volume, $\mathrm{d}v$, of a physical quantity, f, which is a function of both position, $\boldsymbol{x}$, and time, t, and could characterize density, linear and angular momentum, as well as thermodynamical quantities. The differential operator is the material-time derivative, whose expression, by way of the chain rule, is

$$\frac{\mathrm{D}\,f(\boldsymbol{x},t)}{\mathrm{D}t} = \frac{\partial\,f(\boldsymbol{x},t)}{\partial t} + \frac{\partial\,f(\boldsymbol{x},t)}{\partial \boldsymbol{x}}\cdot\frac{\partial\,\boldsymbol{x}}{\partial t} = \frac{\partial\,f(\boldsymbol{x},t)}{\partial t} + \operatorname{grad} f(\boldsymbol{x},t)\cdot\mathbf{v}(\boldsymbol{x},t), \tag{2.16}$$

where the dot-product term refers to the change of quantity f associated with a change

[1] As it pertains to the fundamental statements, quantity f will be defined as a scalar in the case of the balance of mass, but, in the case of the balances of momentum, it will be defined as a vector quantity; quantities $\boldsymbol{\phi}$ and $\boldsymbol{\Sigma}$ are vector quantities in all three cases.

of position $\mathbf{v}(\boldsymbol{x}, t) = \partial \boldsymbol{x}/\partial t$ of the surrounding points, i.e., a spatial velocity. Also, let us remark that if f is a vector quantity, Holzapfel (2000, p. 67) indicates that the dot product should be removed, i.e., the last term in equation (2.16) should be replaced with $\operatorname{grad} \boldsymbol{f}(\boldsymbol{x}, t)\mathbf{v}(\boldsymbol{x}, t)$.

It is possible to commute the differential and integral operators using Reynolds' transport theorem,

$$\frac{\mathrm{D}}{\mathrm{D}t} \int_{\Omega} f(\boldsymbol{x}, t) \, \mathrm{d}v = \int_{\Omega} \frac{\partial \, f(\boldsymbol{x}, t)}{\partial t} \, \mathrm{d}v + \int_{\partial\Omega} f(\boldsymbol{x}, t)\mathbf{v}(\boldsymbol{x}, t) \cdot \mathbf{n} \, \mathrm{d}s, \tag{2.17}$$

where the first term is the rate of change of f within Ω and the second term is the flux of $f(\boldsymbol{x}, t)\mathbf{v}(\boldsymbol{x}, t)$ per unit area, $\mathrm{d}s$, along the direction of the outward normal, $\mathbf{n}$, to the surface at $\boldsymbol{x}$; a proof of relation (2.17) may be found in Holzapfel (2000, Section 4.2).

Turning to the right-hand side of expression (2.15), the terms describe the action of the environment upon the region Ω. In particular, the first term represents the surface density, $\boldsymbol{\phi}$, per unit area, $\mathrm{d}s$, which is a function of position, time, and the outward normal to the surface, whereas the second term represents the volume density, $\boldsymbol{\Sigma}$, consisting of forces, either external or internal, that apply to the entirety of the body.

An important requirement in the development of the balance laws is the relation between the surface and volume integrals, which is facilitated by the divergence theorem. Consulting Gurtin et al. (2010, Section 4.1), for a bounded region Ω with boundary $\partial\Omega$, the divergence theorem states

$$\left.\begin{aligned} \int_{\partial\Omega} \varphi \, \mathbf{n} \, \mathrm{d}s &= \int_{\Omega} \operatorname{grad} \varphi \, \mathrm{d}v \\ \int_{\partial\Omega} \mathbf{v} \cdot \mathbf{n} \, \mathrm{d}s &= \int_{\Omega} \operatorname{div} \mathbf{v} \, \mathrm{d}v \\ \int_{\partial\Omega} \mathbf{T} \, \mathbf{n} \, \mathrm{d}s &= \int_{\Omega} \operatorname{div} \mathbf{T} \, \mathrm{d}v \end{aligned}\right\}, \tag{2.18}$$

where φ is a scalar field, $\mathbf{v}$ is a vector field, $\mathbf{T}$ is a tensor field, and $\mathbf{n}$ is the outward normal to the boundary $\partial\Omega$ of region Ω. Note that expression $\mathbf{T}\,\mathbf{n}$ indicates a matrix multiplication. Also, we recall from Section 2.2 that the gradient of a scalar field is $\operatorname{grad} \phi$ and the divergence of a vector, or tensor, field is $\operatorname{div} \mathbf{v}$ or $\operatorname{div} \mathbf{T}$, respectively.

Now, using the divergence theorem, the second term in relation (2.17) becomes[2]

$$\int_{\partial\Omega} f(\boldsymbol{x},t)\mathbf{v}(\boldsymbol{x},t)\cdot\mathbf{n}\,\mathrm{d}s = \int_{\Omega} \operatorname{div}\left(f(\boldsymbol{x},t)\mathbf{v}(\boldsymbol{x},t)\right)\mathrm{d}v. \tag{2.19}$$

However, to apply the divergence theorem to the surface integral in balance (2.15), we require the so-called Cauchy stress theorem. According to Gurtin et al. (2010, Section 19.5), a deep result central to all of continuum mechanics is Cauchy's Theorem: a consequence of balance (2.15) is that there exists a spatial tensor field, $\boldsymbol{\Phi}$, called the Cauchy stress tensor, such that

$$\boldsymbol{\phi}(\boldsymbol{x}\,,t\,,\mathbf{n}) = \boldsymbol{\Phi}(\boldsymbol{x}\,,t)\mathbf{n}, \tag{2.20}$$

where we evaluate $\boldsymbol{\Phi}(\boldsymbol{x}\,,t)\mathbf{n}$ by matrix multiplication. As such, using the divergence theorem, expression (2.18) for tensors, along with expression (2.20), the surface integral in balance (2.15) becomes

$$\int_{\partial\Omega} \boldsymbol{\phi}(\boldsymbol{x}\,,t\,,\mathbf{n})\,\mathrm{d}s = \int_{\partial\Omega} \boldsymbol{\Phi}(\boldsymbol{x}\,,t)\mathbf{n}\,\mathrm{d}s = \int_{\Omega} \operatorname{div}\boldsymbol{\Phi}(\boldsymbol{x}\,,t)\,\mathrm{d}v. \tag{2.21}$$

Thus, in view of Reynolds' transport theorem and volume integrals (2.19) and (2.21) — as a result of the divergence theorem — the master balance principle (2.15) becomes

$$\int_{\Omega} \frac{\partial\, f(\boldsymbol{x},t)}{\partial t}\,\mathrm{d}v + \int_{\Omega} \operatorname{div}\left(f(\boldsymbol{x},t)\mathbf{v}(\boldsymbol{x},t)\right)\mathrm{d}v = \int_{\Omega} \operatorname{div}\boldsymbol{\Phi}(\boldsymbol{x}\,,t)\,\mathrm{d}v + \int_{\Omega} \boldsymbol{\Sigma}(\boldsymbol{x},t)\,\mathrm{d}v,$$

which can be written as

$$\int_{\Omega} \left(\frac{\partial\, f(\boldsymbol{x},t)}{\partial t} + \operatorname{div}\left(f(\boldsymbol{x},t)\mathbf{v}(\boldsymbol{x},t)\right) - \operatorname{div}\boldsymbol{\Phi}(\boldsymbol{x}\,,t) - \boldsymbol{\Sigma}(\boldsymbol{x},t) \right)\mathrm{d}v = 0. \tag{2.22}$$

For equation (2.22) to hold for an arbitrary fixed volume, the integrand must equal zero. As such, the master balance principle becomes

$$\frac{\partial\, f(\boldsymbol{x},t)}{\partial t} + \operatorname{div}\left(f(\boldsymbol{x},t)\mathbf{v}(\boldsymbol{x},t)\right) = \operatorname{div}\boldsymbol{\Phi}(\boldsymbol{x}\,,t) + \boldsymbol{\Sigma}(\boldsymbol{x},t). \tag{2.23}$$

For the remainder of the section, we shall demonstrate the procedures by which we

[2]Let us remark that if f is a scalar quantity, $f\,\mathbf{v}$ is a scalar product. However, should f represent a vector quantity, the product becomes a tensor product $\boldsymbol{f}\otimes\mathbf{v}$ (Holzapfel, 2000, p. 176).

adjust expression (2.23) to obtain the balances of mass, linear and angular momentum.

2.3.2 Balance of mass

To obtain the balance of mass, we set $f = \rho = \rho(\boldsymbol{x}, t)$, $\boldsymbol{\phi} = \mathbf{0}$, and $\boldsymbol{\Sigma} = \mathbf{0}$. By Cauchy's theorem, in the absence of a surface density, we require $\boldsymbol{\Phi} = \mathbf{0}$ to satisfy expression (2.20). As such, balance (2.23) reduces to the so-called balance of mass

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) = 0, \tag{2.24}$$

where, for the sake of brevity, we omit function arguments from numbered expressions for the remainder of the section. Balance (2.24) states that the change in the amount of mass in a volume at any instant is balanced by the mass flowing through the surface that encloses this volume (Slawinski, 2020a, p. 48). Using identity (2.11), we expand balance (2.24) so that

$$\frac{\partial \rho}{\partial t} + \rho(\operatorname{div} \mathbf{v}) + \mathbf{v} \cdot (\operatorname{grad} \rho) = 0.$$

Then, by the material-time derivative (2.16), we obtain an alternative expression for the balance of mass,

$$\frac{\mathrm{D} \rho}{\mathrm{D} t} + \rho(\operatorname{div} \mathbf{v}) = 0. \tag{2.25}$$

2.3.3 Balance of linear momentum

Let us recall that the momentum of an object is the product of its mass, m, and its velocity, $\mathbf{v}$. Likewise, we recall from Newton's Laws that (Feynman et al., 2011a, Section 9-1) the time-rate-of-change of the momentum is equal to the force

$$\mathcal{F} = \frac{\mathrm{d}}{\mathrm{d} t}(m \mathbf{v}). \tag{2.26}$$

In continuum mechanics, the force in question is comprised of traction, $\mathbf{t} = \mathbf{t}(\boldsymbol{x}, t, \mathbf{n})$, which is a force applied to an oriented surface element, $\mathrm{d}s$, and a (density-scaled) body force, $\mathbf{b} = \mathbf{b}(\boldsymbol{x}, t)$, which applies to the entirety of the fixed volume. For the balance of linear momentum, we set $\boldsymbol{f} = \rho \mathbf{v}$, $\boldsymbol{\phi} = \mathbf{t}$, and $\boldsymbol{\Sigma} = \rho \mathbf{b}$. By Cauchy's theorem, under the effect of traction, it is common to set $\boldsymbol{\Phi} = \boldsymbol{\sigma} = \boldsymbol{\sigma}(\boldsymbol{x}, t)$, which is known as Cauchy's stress tensor, in order to satisfy expression (2.20). As such, we

write balance (2.23) as

$$\frac{\partial\,(\rho\,\mathbf{v})}{\partial t} + \operatorname{div}\left((\rho\,\mathbf{v})\otimes\mathbf{v}\right) - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b} = 0, \tag{2.27}$$

where, in accordance with footnote 2 on page 14, we use the tensor product. Using identities (2.12), (2.10), and then (2.8), the second term in balance (2.27) becomes

$$\operatorname{div}\left((\rho\,\mathbf{v})\otimes\mathbf{v}\right) = (\mathbf{v}\cdot(\operatorname{grad}\rho))\mathbf{v} + (\rho\,(\operatorname{grad}\mathbf{v}))\mathbf{v} + (\rho\,(\operatorname{div}\mathbf{v}))\,\mathbf{v}. \tag{2.28}$$

With expression (2.28) along with the product rule on the first term of balance (2.27), we obtain

$$\frac{\partial\,\rho}{\partial t}\,\mathbf{v} + \rho\,\frac{\partial\,\mathbf{v}}{\partial t} + (\mathbf{v}\cdot(\operatorname{grad}\rho))\,\mathbf{v} + (\rho\,(\operatorname{grad}\mathbf{v}))\,\mathbf{v} + (\rho\,(\operatorname{div}\mathbf{v}))\,\mathbf{v} - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b} = 0.$$

Collecting like terms of ρ and $\mathbf{v}$,

$$\left(\frac{\partial\,\rho}{\partial t} + (\operatorname{grad}\rho)\cdot\mathbf{v} + \rho\,(\operatorname{div}\mathbf{v})\right)\mathbf{v} + \rho\left(\frac{\partial\,\mathbf{v}}{\partial t} + (\operatorname{grad}\mathbf{v})\mathbf{v}\right) - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b} = 0. \tag{2.29}$$

Then, by the material-time derivative (2.16) of ρ, as well as of $\mathbf{v}$, we write balance (2.29) as

$$\left(\frac{\mathrm{D}\,\rho}{\mathrm{D}t} + \rho(\operatorname{div}\mathbf{v})\right)\mathbf{v} + \rho\frac{\mathrm{D}\,\mathbf{v}}{\mathrm{D}t} - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b} = 0. \tag{2.30}$$

However, by the balance of mass (2.25), the first term in expression (2.30) is zero. Thus, expression (2.30) reduces to

$$\rho\frac{\mathrm{D}\,\mathbf{v}}{\mathrm{D}t} = \operatorname{div}\boldsymbol{\sigma} + \rho\mathbf{b}, \tag{2.31}$$

which is also known as Cauchy's first equation of motion or, in seismological contexts, the elastodynamic equation; henceforth, we adopt the latter nomenclature. These equations state that the acceleration of an element within a continuum results from the application of surface and body forces (Slawinski, 2020a, p. 65).

2.3.4 Balance of angular momentum

The counterpart to the linear momentum in expression (2.26) is angular momentum

$$\boldsymbol{\tau} := \boldsymbol{r} \times \mathcal{F} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\boldsymbol{r} \times (m\mathbf{v}) \right), \tag{2.32}$$

where $\boldsymbol{\tau}$ is torque and $\boldsymbol{r} = \boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{x}_0$ is a position vector relative to a reference coordinate $\boldsymbol{x}_0$. In continuum mechanics, for the balance of angular momentum, we include the cross product within $\boldsymbol{f} = \boldsymbol{r} \times (\rho\, \mathbf{v})$, $\boldsymbol{\phi} = \boldsymbol{r} \times \mathbf{t}$, and $\boldsymbol{\Sigma} = \boldsymbol{r} \times (\rho \mathbf{b})$.

Now, in the presence of a cross product, the conversion of the surface integral in balance (2.15), by means of the divergence theorem, to a volume integral is not given by conversion (2.21). By contrast, Holzapfel (2000, p. 54) indicates that the divergence theorem for the cross product $\mathbf{u} \times \mathbf{An}$ is

$$\int_{\partial\Omega} \mathbf{u} \times (\mathbf{An})\, \mathrm{d}s = \int_{\Omega} \left(\mathbf{u} \times (\operatorname{div} \mathbf{A}) + \boldsymbol{\mathcal{E}} : \mathbf{A}^T \right) \mathrm{d}v\,, \tag{2.33}$$

where $\mathbf{u}$ and $\mathbf{A}$ are a vector and tensor defined in region Ω and on its surface $\partial\Omega$, with $\mathbf{n}$ as the outward normal to the surface, and $\mathbf{An}$ is evaluated by matrix multiplication. The second term in the volume integrand of expression (2.33) includes the permutation tensor,

$$\boldsymbol{\mathcal{E}} := \varepsilon_{ijk}\, \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k, \tag{2.34}$$

where ε_{ijk} is the permutation symbol (2.6). The colon represents a double contraction, akin to the dot product, which indicates a summation over two sets of indices. Thus, using conversion (2.33) in place of (2.21), the master balance principle (2.23) is

$$\frac{\partial \left(\boldsymbol{r} \times (\rho \mathbf{v}) \right)}{\partial t} + \operatorname{div} \left((\boldsymbol{r} \times (\rho \mathbf{v})) \otimes \mathbf{v} \right) = \boldsymbol{r} \times (\operatorname{div} \boldsymbol{\sigma}) + \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T + \boldsymbol{r} \times (\rho \mathbf{b}), \tag{2.35}$$

where, by Cauchy's theorem $\mathbf{t} = \boldsymbol{\sigma}\mathbf{n}$, cross product $\boldsymbol{\phi} = \boldsymbol{r} \times \mathbf{t} = \boldsymbol{r} \times \boldsymbol{\sigma}\mathbf{n}$ replaces $\mathbf{u} \times \mathbf{An}$ in conversion (2.33).

In contrast to the derivation of Cauchy's first equation of motion, the introduction of the cross product leads to a more drawn out derivation of the second equation of motion. For that reason, we relegate the majority of the derivation to Appendix A.1. As is shown therein, it is possible to reduce master balance principle (2.35) to

$$\mathbf{0} = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T. \tag{2.36}$$

Recalling the permutation tensor (2.34) along with $\boldsymbol{\sigma} = \sigma_{mn}\,\mathbf{e}_m \otimes \mathbf{e}_n$, the double contraction of expression (2.34) results in

$$\begin{aligned}\mathbf{0} &= (\varepsilon_{ijk}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k) : (\sigma_{nm}\,\mathbf{e}_m \otimes \mathbf{e}_n) \\ &= \varepsilon_{ijk}\sigma_{nm}(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k) : (\mathbf{e}_m \otimes \mathbf{e}_n) \\ &= \varepsilon_{ijk}\sigma_{nm}\delta_{jm}\delta_{kn}\,\mathbf{e}_i \\ &= \varepsilon_{ijk}\sigma_{kj}\,\mathbf{e}_i \end{aligned} \tag{2.37}$$

To satisfy (2.37), we require $\varepsilon_{ijk}\sigma_{kj} = 0$ for $i = 1, 2, 3$, which, by permutation symbol (2.6), yields

$$\sigma_{32} - \sigma_{23} = 0, \quad \sigma_{31} - \sigma_{13} = 0, \quad \sigma_{21} - \sigma_{12} = 0. \tag{2.38}$$

Requirements (2.38) are tantamount to the symmetry of the Cauchy stress tensor,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^T, \tag{2.39}$$

which is otherwise known as Cauchy's second equation of motion.

2.4 Linearization

In Section 2.3, we presented the master balance principle of continuum mechanics, which can be adjusted for the statements of the balances of mass, linear momentum, and angular momentum. These fundamental balances are used to derive Cauchy's equations of motion: the first is the elastodynamic equation, expression (2.31), which is nonlinear partial differential equation that relates an element's acceleration within a continuum to the application of surface and body forces to the region it occupies; the second, expression (2.39), is the symmetry condition of the Cauchy stress tensor. Thus far, the balances and equations of motion can apply to any continuous material. However, to apply these principles to seismology, we require so-called constitutive equations, which are mathematical relations for the macroscopic stress-strain responses to the application of a load.

For seismology, under the assumptions of a continuous material, it is common to consider a linear stress-strain relationship. This linearization pertains to the deformation of the material, which is quantified by the displacement of a collection

of points between the aforementioned reference and current configurations. In other words, the constitutive equation for seismology is based on Hooke's law, which states that the force in a body that tries to restore the body to its original condition when it is deformed is proportional to the deformation (Feynman et al., 2011a, Section 12-2). Hooke's law is valid for such forces that do not exceed a so-called proportional limit, beyond which stress is no longer linearly related to strain (e.g., Beer et al., 2011, Section 9.4). In this dissertation, we restrict our attention to materials that adhere to Hooke's law and we refer to them as Hookean.

In this section, we detail the effect of linearization on the stress-strain relation and the elastodynamic equation.

2.4.1 Infinitesimal strain tensor

Broadly speaking, the linearization of a differentiable single-variable scalar function, $f = f(x)$, refers to the tangent line approximation of f when x is near some x_0. More precisely, we obtain this approximation using the nth-degree Taylor series of f centered at x_0 (e.g., Stewart et al., 2021, Section 11.10),

$$f(x) \approx \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i, \tag{2.40}$$

where $f^{(i)}(x_0)$ refers to the ith derivative of f evaluated at $x = x_0$. Hence, for a linearization, we use Taylor series (2.40) with $n = 1$ to obtain

$$f(x) \approx f(x_0) + f'(x_0)\,(x - x_0). \tag{2.41}$$

Now, let us consider the linearization for the displacement of two points within a region of the continuum, $\boldsymbol{x}$ and $\boldsymbol{y} = \boldsymbol{x} + \mathrm{d}\boldsymbol{x}$. The points are separated by an infinitesimal vector, $\mathrm{d}\boldsymbol{x} = \mathrm{d}x_i\,\mathbf{e}_i$, and the squared distance between them is

$$\mathrm{d}s^2 = \mathrm{d}\boldsymbol{x} \cdot \mathrm{d}\boldsymbol{x} = \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2. \tag{2.42}$$

Next, suppose a deformation is applied to the continuum, resulting in a displacement, $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x})$, of the two points such that the new positions are

$$\widetilde{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}) \quad \text{and} \quad \widetilde{\boldsymbol{y}} = \boldsymbol{y} + \boldsymbol{u}(\boldsymbol{y}). \tag{2.43}$$

Using approximation (2.41), we apply the linearization to $\boldsymbol{u}$ at $\boldsymbol{y}$, for the kth component, which yields

$$\boldsymbol{u}(\boldsymbol{y}) = \boldsymbol{u}(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}) \approx \left(u_k + \sum_{\ell=k}^{3} u_{k,\ell}\, \mathrm{d}x_\ell \right) \mathbf{e}_k = \boldsymbol{u}(\boldsymbol{x}) + (\operatorname{grad} \boldsymbol{u}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}. \tag{2.44}$$

With linearization (2.44), the vector from $\widetilde{\boldsymbol{x}}$ to $\widetilde{\boldsymbol{y}}$ is

$$\begin{aligned} \widetilde{\boldsymbol{y}} - \widetilde{\boldsymbol{x}} &= \boldsymbol{y} + \boldsymbol{u}(\boldsymbol{y}) - \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}) \\ &= \boldsymbol{x} + \mathrm{d}\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}) - \boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}) \\ &\approx \mathrm{d}\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}) + (\operatorname{grad} \boldsymbol{u}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x} - \boldsymbol{u}(\boldsymbol{x}) \\ &\approx \mathrm{d}\boldsymbol{x} + (\operatorname{grad} \boldsymbol{u}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}, \end{aligned} \tag{2.45}$$

whose squared length is

$$\widetilde{\mathrm{d}s}^2 \approx (\mathrm{d}\boldsymbol{x} + (\operatorname{grad} \boldsymbol{u}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}) \cdot (\mathrm{d}\boldsymbol{x} + (\operatorname{grad} \boldsymbol{u}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}) = \sum_{k=1}^{3} \left(\mathrm{d}x_k + \sum_{\ell=k}^{3} u_{k,\ell}\, \mathrm{d}x_\ell \right)^2. \tag{2.46}$$

Expanding summation (2.46) and neglecting terms with the product of two infinitesimal derivatives of displacement, expression (2.46) reduces to

$$\begin{aligned} \widetilde{\mathrm{d}s}^2 \approx\ & \mathrm{d}x_1^2 + \mathrm{d}x_2^2 + \mathrm{d}x_3^2 + 2 \big(u_{1,1}\mathrm{d}x_1^2 + u_{2,2}\mathrm{d}x_2^2 + u_{3,3}\mathrm{d}x_3^2 \\ & + (u_{1,2} + u_{2,1})\mathrm{d}x_1\mathrm{d}x_2 + (u_{1,3} + u_{3,1})\mathrm{d}x_1\mathrm{d}x_3 + (u_{2,3} + u_{3,2})\mathrm{d}x_2\mathrm{d}x_3 \big). \end{aligned} \tag{2.47}$$

Substituting distance (2.42) in place of the first three terms of the right-hand side, rearranging, and expressing the remaining terms in summation form, we obtain

$$\widetilde{\mathrm{d}s}^2 - \mathrm{d}s^2 \approx \sum_{k=1}^{3} \sum_{\ell=1}^{3} (u_{k,\ell} + u_{\ell,k})\, \mathrm{d}x_k \mathrm{d}x_\ell. \tag{2.48}$$

The parenthetical term in distance (2.48) is the definition of the strain tensor, $\boldsymbol{\varepsilon} = \varepsilon_{k\ell}\, \mathbf{e}_k \otimes \mathbf{e}_\ell$, for infinitesimal displacements (e.g., Slawinski, 2020a, Section 1.4), where components

$$\varepsilon_{k\ell} = \frac{1}{2} \left(\frac{\partial\, u_k}{\partial x_\ell} + \frac{\partial\, u_\ell}{\partial x_k} \right) \tag{2.49}$$

indicate that the strain tensor is symmetric.

2.4.2 Modified elastodynamic equation

For the purposes of seismic wave propagation, we can modify the elastodynamic equation (2.31) by reducing the material-time derivative and neglecting body forces. To demonstrate, let us recall that the left-hand side of expression (2.31) is the product of mass density and the material-time derivative (2.16) of the velocity associated with a displacement. We expand the latter as

$$\frac{\mathrm{D}\mathbf{v}(\boldsymbol{x},t)}{\mathrm{D}t} = \frac{\partial \mathbf{v}(\boldsymbol{x},t)}{\partial t} + (\operatorname{grad}\mathbf{v}(\boldsymbol{x},t))\,\mathbf{v}(\boldsymbol{x},t). \tag{2.50}$$

For infinitesimal strain, we assume that both the spatial velocity, $\mathbf{v}(\boldsymbol{x},t) = \partial\boldsymbol{x}/\partial t$, and its gradient are infinitesimal and, as such, the second term in expression (2.50), which is their matrix product, can be neglected. In view of expression (2.43), the temporal derivative of the spatial velocity is equivalent to the second temporal derivative of displacement. Thus, in the context of infinitesimal strain, material-time derivative (2.50) reduces to

$$\frac{\mathrm{D}\mathbf{v}(\boldsymbol{x},t)}{\mathrm{D}t} \approx \frac{\partial \mathbf{v}(\boldsymbol{x},t)}{\partial t} = \frac{\partial^2 \boldsymbol{u}(\boldsymbol{x},t)}{\partial t^2}. \tag{2.51}$$

Likewise, as a consequence of infinitesimal strain, the deformations considered are small and there is no distinction between the aforementioned reference and current configurations. Hence, the mass density of the continuum no longer depends on time and becomes solely a function of position, $\rho = \rho(\boldsymbol{x})$.

Let us recall that the right-hand side of the elastodynamic equation (2.31) refers to the application of surface and body forces to the boundary of a region within the continuum. Under the effect of large displacements, the body force can refer to an earthquake source. In this dissertation, we restrict our attention to infinitesimal strains and, as such, we do not model wave propagation in the vicinity of a source. Beyond this vicinity, the only body force that applies to the seismic wave is that of gravitation, which depends on the period of the wave. It has been shown (e.g., Udías, 1999, Section 3.5) that the effect of gravity on wave propagation is only applicable for waves with very long periods. As we discuss in Section 2.5, we model wave propagation under the assumption of the so-called high-frequency approximation. Therefore, we

omit body forces from further consideration.

In view of these simplifications, the elastodynamic equation (2.31) reduces to

$$\rho(\boldsymbol{x})\frac{\partial^2 \boldsymbol{u}(\boldsymbol{x},t)}{\partial t^2} = \operatorname{div}\boldsymbol{\sigma}(\boldsymbol{x},t). \tag{2.52}$$

2.4.3 Stress-strain relationship

Within the context of the aforementioned linearization, the second-order Cauchy stress tensor, $\boldsymbol{\sigma}$, is related to the second-order strain tensor, $\boldsymbol{\varepsilon}$, by the so-called elasticity tensor

$$\boldsymbol{c} = c_{ijk\ell}\,\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell, \tag{2.53}$$

which is of order four. Since each vector has three components, the elasticity tensor has $3^4 = 81$ components. The linear relastionship, which is valid at every point within the continuum, is given by the double contraction

$$\boldsymbol{\sigma}(\boldsymbol{x},t) = \boldsymbol{c}(\boldsymbol{x}) : \boldsymbol{\varepsilon}(\boldsymbol{x},t). \tag{2.54}$$

In a similar manner to mass density, the elasticity tensor, within the context of infinitesimal strains, is solely a function of position, $\boldsymbol{c} = \boldsymbol{c}(\boldsymbol{x})$. The components of expression (2.54) are (e.g., Slawinski, 2020a, Section 3.2.2)

$$\sigma_{ij} = \sum_{k=1}^{3}\sum_{\ell=1}^{3} c_{ijk\ell}\,\varepsilon_{k\ell}\,, \qquad i,j = 1,2,3\,. \tag{2.55}$$

The elasticity tensor has what are said to be minor and major index symmetries. The minor symmetries refer to the invariance of order within pairs of indices i,j and k,ℓ whereas the major symmetries refer to the invariance of order of the pairs themselves. In other words, the index symmetries of the elasticity tensor are

$$c_{ijk\ell} = c_{jik\ell} = c_{ij\ell k} = c_{k\ell ij}. \tag{2.56}$$

The minor index symmetries, relating to the first and second pairs of indices, are justified by the symmetries of the stress and strain tensors, respectively, which are given by expressions (2.39) and (2.49).

The major index symmetry is justified by the existence of the so-called strain-energy function, W. This function is related to the balance of energy, which is not explicitly discussed in this dissertation, but can be derived from the master balance principle (2.15) (Holzapfel, 2000, Table 4.1). Within this context, Feynman et al. (2011b, Section 39-2) indicates that this function is the work that goes into each unit volume of distorted material,

$$W = \frac{1}{2} \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{k=1}^{3} \sum_{\ell=1}^{3} c_{ijk\ell} \varepsilon_{ij} \varepsilon_{k\ell}. \tag{2.57}$$

Scalar function (2.57) can be differentiated with respect to strain. Akin to the manner in which a conservative force is obtained by differentiating a scalar potential, differentiating (2.57) yields

$$\frac{\partial W}{\partial \varepsilon_{ij}} = \sum_{k=1}^{3} \sum_{\ell=1}^{3} c_{ijk\ell} \varepsilon_{k\ell} = \sigma_{ij}, \qquad i, j = 1, 2, 3, \tag{2.58}$$

which is stress-strain relationship (2.55). Taking the second derivative, we obtain

$$\frac{\partial^2 W}{\partial \varepsilon_{k\ell} \partial \varepsilon_{ij}} = c_{ijk\ell}, \qquad i, j, k, \ell = 1, 2, 3. \tag{2.59}$$

However, by the equality of mixed second-order partial derivatives (e.g., Colley, 2012, Theorem 4.3), the order of differentiation is interchangeable, which justifies the major symmetry. With regard to the number of independent elasticity tensor components, the minor symmetries reduce the number from 81 to 36 whereas the major symmetry further reduces the number to 21.

As a consequence of index symmetries (2.56), stress-strain relationship (2.55) may be rewritten in matrix form

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}, \tag{2.60}$$

with components

$$\begin{bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{13} \\ \sqrt{2}\sigma_{12} \end{bmatrix} = \begin{bmatrix} c_{1111} & c_{1122} & c_{1133} & \sqrt{2}\,c_{1123} & \sqrt{2}\,c_{1113} & \sqrt{2}\,c_{1112} \\ c_{1122} & c_{2222} & c_{2233} & \sqrt{2}\,c_{2223} & \sqrt{2}\,c_{2213} & \sqrt{2}\,c_{2212} \\ c_{1133} & c_{2233} & c_{3333} & \sqrt{2}\,c_{3323} & \sqrt{2}\,c_{3313} & \sqrt{2}\,c_{3312} \\ \sqrt{2}\,c_{1123} & \sqrt{2}\,c_{2223} & \sqrt{2}\,c_{3323} & 2\,c_{2323} & 2\,c_{2313} & 2\,c_{2312} \\ \sqrt{2}\,c_{1113} & \sqrt{2}\,c_{2213} & \sqrt{2}\,c_{3313} & 2\,c_{2313} & 2\,c_{1313} & 2\,c_{1312} \\ \sqrt{2}\,c_{1112} & \sqrt{2}\,c_{2212} & \sqrt{2}\,c_{3312} & 2\,c_{2312} & 2\,c_{1312} & 2\,c_{1212} \end{bmatrix} \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ \sqrt{2}\,\varepsilon_{23} \\ \sqrt{2}\,\varepsilon_{13} \\ \sqrt{2}\,\varepsilon_{12} \end{bmatrix}, \tag{2.61}$$

where the $\sqrt{2}$ and 2 ensure that the basis of the stress and strain components remain the same by requiring the Frobenius norm of each basis element to be equal to unity (Slawinski, 2020b, Section 3.2.4.4).

The existence of strain-energy function (2.57) imposes a physical constraint on the continuum. Specifically, an element of the continuum can only be deformed if energy is expended. In the absence of deformation, the continuum remains in its undeformed state. However, since energy is a positive quantity, the strain energy of an undeformed continuum is zero. Thus, the strain-energy function must be a positive quantity that vanishes only in the undeformed state (Slawinski, 2020a, Section 4.3.1). Using matrix form (2.60), this constraint is

$$W = \frac{1}{2}(\boldsymbol{C}\boldsymbol{\varepsilon}) \cdot \boldsymbol{\varepsilon} > 0, \quad \text{for} \quad \boldsymbol{\varepsilon} \neq \mathbf{0}, \tag{2.62}$$

which is the statement of the positive definite property of matrix $\boldsymbol{C}$ (e.g., Kolman and Hill, 2008, p. 311).

2.5 Ray theory

Broadly speaking, the elastodynamic equation (2.52) relates the acceleration of an element due to nonlinear displacements within a material to the application of forces to the boundary of the region it occupies. This is a complicated equation to solve as the material properties can be anisotropic and/or inhomogeneous. The former refers to material properties that vary with respect to direction at a given position within the material whereas the latter refers to material properties varying with position. According to Červený (2001, p. 1), the most common approaches to the investigation of seismic wavefields in such complex structures are methods based on

direct numerical solutions of the elastodynamic equation or methods based on its high-frequency approximation. In this dissertation, we use the latter method, which gives rise to the so-called seismic ray theory.

Seismic ray theory is an idealized approximation to modelling wave propagation within media whose material properties vary smoothly and slowly in comparison to its wavelength. The theory is based on high-frequency approximation to elastodynamic equation (2.52) within the context of plane waves. We use this approximation to obtain the so-called eikonal equation, which we use to quantify the kinematic features of seismic rays such as their trajectories and traveltimes.

2.5.1 Plane waves

A displacement modelled by a plane wave is

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{A}(\boldsymbol{x}) e^{-i\omega(t-\psi(\boldsymbol{x}))}, \tag{2.63}$$

where $\boldsymbol{A}(\boldsymbol{x})$ is the amplitude of the displacement that varies with respect to position, i is an imaginary number, whose square is -1, ω is the angular frequency of the wave, t is the temporal variable, and $\psi(\boldsymbol{x})$ is the so-called eikonal function, and $t - \psi(\boldsymbol{x})$ is the phase of the wave. We refer to expression (2.63) as plane wave because its phase satisfies the equation of a plane,

$$\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0, \tag{2.64}$$

where $\boldsymbol{x}$ and $\boldsymbol{x}_0$ are two points on a plane with unit normal $\boldsymbol{n}$. Within this structure, it is common to parameterize phase expression (2.64) as

$$\boldsymbol{n} \cdot \boldsymbol{x} = v\, t, \tag{2.65}$$

where v is the phase velocity, which corresponds to the velocity of the plane. We can obtain an alternative parameterization of expression (2.65) using the so-called slowness vector,

$$\boldsymbol{p} = \frac{\boldsymbol{n}}{v}, \tag{2.66}$$

which is oriented in the direction of the unit normal and is scaled by the phase velocity. The squared magnitude of expression (2.66) is

$$p^2 = \boldsymbol{p} \cdot \boldsymbol{p} = \left(\frac{\boldsymbol{n}}{v}\right) \cdot \left(\frac{\boldsymbol{n}}{v}\right) = \frac{\boldsymbol{n} \cdot \boldsymbol{n}}{v^2} = \frac{1}{v^2}, \tag{2.67}$$

which states that the squared slowness is the reciprocal of the squared phase velocity. In view of slowness (2.66), expression (2.65) becomes

$$\boldsymbol{p} \cdot \boldsymbol{x} = t, \tag{2.68}$$

Setting the left-hand side of expression (2.68) to be the eikonal function, we have

$$\psi(\boldsymbol{x}) = \boldsymbol{p} \cdot \boldsymbol{x} \tag{2.69}$$

and the argument of plane wave (2.63) describes a moving planar wavefront for some t.

Let us acknowledge that for a wave's wavefront to be planar, it is understood that the wave's source is at an infinite distance and — as a result — is not included in the elastodynamic equation. Strictly speaking, this is a nonphysical situation. However, if we are interested only in wave propagation and can assume the source is at a large distance, plane waves are a good simplifying approximation (Udías, 1999, Section 3.6).

In view of expression (2.69), the gradient of the eikonal function yields a useful relationship for the slowness vector. Using components, we have

$$\operatorname{grad} \psi(\boldsymbol{x}) = \frac{\partial}{\partial x_k} \left(p_i \, \mathbf{e}_i \cdot x_j \, \mathbf{e}_j\right) \mathbf{e}_k = \frac{\partial}{\partial x_k} \left(p_i x_j \delta_{ij}\right) \mathbf{e}_k = p_i \frac{\partial \, x_i}{\partial x_k} \mathbf{e}_k = p_i \delta_{ik} \mathbf{e}_k = p_i \mathbf{e}_i = \boldsymbol{p}. \tag{2.70}$$

In other words, expression (2.70) states that the gradient of the eikonal function is the slowness vector, which is expected due to the orthogonality of in-plane vectors and the planar normal.

2.5.2 High-frequency approximation

Let us recall the elastodynamic equation (2.52), whose components are

$$\rho(\boldsymbol{x}) \frac{\partial^2 u_i(\boldsymbol{x}, t)}{\partial t^2} = \sum_{j=1}^{3} \frac{\partial \, \sigma_{ij}(\boldsymbol{x}, t)}{\partial x_j}, \qquad i = 1, 2, 3; \tag{2.71}$$

Substituting in the linearized stress-strain (2.55) relation, with strain components (2.49), expression (2.71) becomes

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^{3}\frac{\partial}{\partial x_j}\left(\sum_{k=1}^{3}\sum_{\ell=1}^{3}\frac{c_{ijk\ell}}{2}\left(\frac{\partial\, u_k}{\partial x_\ell}+\frac{\partial\, u_\ell}{\partial x_k}\right)\right), \qquad i=1,2,3,$$

where, for brevity, we have dropped the function arguments. Distributing the derivative, we obtain

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{\ell=1}^{3}\left(\frac{\partial\, c_{ijk\ell}}{\partial x_j}\left(\frac{\partial\, u_k}{\partial x_\ell}+\frac{\partial\, u_\ell}{\partial x_k}\right)+c_{ijk\ell}\left(\frac{\partial^2 u_k}{\partial x_j\partial x_\ell}+\frac{\partial^2 u_\ell}{\partial x_j\partial x_k}\right)\right), \quad i=1,2,3. \tag{2.72}$$

To evaluate the derivatives in expression (2.72), we use a plane wave (2.63) to model the displacements. For the temporal derivatives, it is immediate that

$$\frac{\partial \boldsymbol{u}}{\partial t} = -i\omega\boldsymbol{u} \quad \text{and} \quad \frac{\partial^2 \boldsymbol{u}}{\partial t^2} = -\omega^2\boldsymbol{u}. \tag{2.73}$$

By contrast, the spatial derivatives are more complicated. For the sake of brevity, we take up again the comma notation for partial derivatives (2.14) for expressions with many partial derivatives as well as omit function arguments where applicable. Also, we combine the multiple summations in expression (2.72) as a single summation with multiple indices, where the index order is preserved.

Using components for arbitrary indices k and ℓ, the first derivative is

$$\frac{\partial\, u_k}{\partial x_\ell} = A_{k,\ell}\, e^{-i\omega(t-\psi)} + i\omega A_k e^{-i\omega(t-\psi)}\psi_{,\ell} = e^{-i\omega(t-\psi)}\left(A_{k,\ell}+i\omega A_k\psi_{,\ell}\right) \tag{2.74}$$

and the second derivative, denoted by index j, is

$$\begin{aligned}\frac{\partial}{\partial x_j}\left(\frac{\partial\, u_k}{\partial x_\ell}\right) &= i\omega e^{-i\omega(t-\psi)}\psi_{,j}\left(A_{k,\ell}+i\omega A_k\psi_{,\ell}\right)+e^{-i\omega(t-\psi)}\left(A_{k,j\ell}+i\omega A_{k,j}\psi_{,\ell}+i\omega\psi_{,j\ell}\right)\\ &= e^{-i\omega(t-\psi)}\left(A_{k,j\ell}+i\omega\left(A_{k,j}\psi_{,\ell}+A_{k,\ell}\psi_{,j}+A_k\psi_{,j\ell}\right)-\omega^2 A_k\psi_{,j}\psi_{,\ell}\right).\end{aligned} \tag{2.75}$$

Analogous expressions for $u_{\ell,k}$ and $u_{\ell,jk}$ can be obtained by swapping indices k and ℓ

in expressions (2.74) and (2.75), respectively.

Substituting temporal derivative (2.73) and spatial derivatives (2.74) and (2.75) in expression (2.72), then cancelling the exponential term from both sides, we obtain

$$-\omega^2\rho\,A_i = \frac{1}{2}\sum_{j,k,\ell}\Bigg(c_{ijk\ell,j}\left(A_{k,\ell} + i\omega A_k\psi_{,\ell} + A_{\ell,k} + i\omega A_\ell\psi_{,k}\right) + c_{ijk\ell}\Bigg(A_{k,j\ell} + i\omega\left(A_{k,j}\psi_{,\ell} + A_{k,\ell}\psi_{,j} + A_k\psi_{,j\ell}\right) - \omega^2 A_k\psi_{,j}\psi_{,\ell} + A_{\ell,jk} + i\omega\left(A_{\ell,j}\psi_{,k} + A_{\ell,k}\psi_{,j} + A_\ell\psi_{,jk}\right) - \omega^2 A_\ell\psi_{,j}\psi_{,k}\Bigg)\Bigg), \quad i = 1,2,3. \tag{2.76}$$

Expression (2.76) can be separated into real- and imaginary-valued terms. We use the real terms to derive the eikonal equation, which we use to derive equations for the raypaths and traveltimes of individual seismic waves. The imaginary terms are used to derive the transport equation, which governs the amplitude along a given ray.

In this dissertation, we restrict our attention to eikonal equation only. As such, considering the real part of expression (2.76), we divide both sides by ω^2 to obtain

$$-\rho A_i = \frac{1}{2}\sum_{j,k,\ell}\Bigg(\frac{c_{ijk\ell,j}}{\omega^2}\left(A_{k,\ell} + A_{\ell,k}\right) + c_{ijk\ell}\left(\frac{1}{\omega^2}\left(A_{k,j\ell} + A_{\ell,jk}\right) - \left(\psi_{,j}\psi_{,\ell}A_k + \psi_{,j}\psi_{,k}A_\ell\right)\right)\Bigg), \quad i = 1,2,3. \tag{2.77}$$

Then, in the limit of frequency, $\omega \to \infty$, the terms divided by ω tend to zero. Rearranging the remaining terms in expression (2.77) and undoing the multiple summation combination, we have

$$\frac{1}{2}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{\ell=1}^{3}\left(c_{ijk\ell}\left(\psi_{,j}\psi_{,\ell}A_k + \psi_{,j}\psi_{,k}A_\ell\right)\right) - \rho A_i = 0, \quad i = 1,2,3. \tag{2.78}$$

To simplify expression (2.78), we recall the elasticity-tensor index symmetries (2.56),

in particular $c_{ijk\ell} = c_{ij\ell k}$, which yields

$$\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{\ell=1}^{3}\left(c_{ijk\ell}\psi_{,j}\psi_{,\ell}A_k\right) - \rho A_i = 0, \quad i = 1,2,3. \tag{2.79}$$

Expanding the kth summation for each i, then contracting by the Kronecker delta, δ_{ik}, we are able to reduce expression (2.79) to

$$\left(\sum_{j=1}^{3}\sum_{\ell=1}^{3} c_{ijk\ell}(\boldsymbol{x})\frac{\partial\,\psi(\boldsymbol{x})}{\partial x_j}\frac{\partial\,\psi(\boldsymbol{x})}{\partial x_\ell} - \rho(\boldsymbol{x})\delta_{ik}\right) A_i(\boldsymbol{x}) = 0, \quad i,k = 1,2,3, \tag{2.80}$$

where we have undone the comma notation for partial derivatives and reinstated the function arguments. Expression (2.80) is known as the Christoffel equation and it constitutes a system of three equations for plane-wave displacement amplitudes $\boldsymbol{A}(\boldsymbol{x})$.

Let us present the Christoffel equations in matrix form. By gradient (2.70), we can replace the derivatives of the eikonal function by the slowness vector. Then, we normalize the components of the matrix using the slownesses, $p^2 = \boldsymbol{p}\cdot\boldsymbol{p}$. With these adjustments, Christoffel equations (2.80) are

$$\left(\boldsymbol{\Gamma}(\boldsymbol{x}) - p^{-2}\boldsymbol{I}\right)\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{0}, \tag{2.81}$$

where $\boldsymbol{\Gamma}(\boldsymbol{x}) = \Gamma_{ik}\,\mathbf{e}_i\otimes\mathbf{e}_k$ is the so-called Christoffel matrix, with components

$$\Gamma_{ik} = \sum_{j=1}^{3}\sum_{\ell=1}^{3}\frac{c_{ijk\ell}(\boldsymbol{x})}{\rho(\boldsymbol{x})}\frac{p_j p_\ell}{p^2}, \tag{2.82}$$

and $\boldsymbol{I} = \delta_{ik}\,\mathbf{e}_i\otimes\mathbf{e}_k$ is an identity matrix. The Christoffel matrix has several important properties. First, components (2.82) are symmetric, $\Gamma_{ik} = \Gamma_{ki}$, as a consequence of the elasticity tensor index symmetries (2.56). Second, the Christoffel matrix is positive definite as a consequence of the positive definiteness of the elasticity tensor (e.g., Červený, 2001, p. 23). Third, the Christoffel equations (2.81) are in the form of an eigenvalue problem, whose nontrivial solution exists if and only if

$$\det\left(\boldsymbol{\Gamma}(\boldsymbol{x}) - v(\boldsymbol{x},\boldsymbol{p})^2\boldsymbol{I}\right) = 0,$$

where we recall $p^{-2} = v^2$ from expression (2.67). Since the Christoffel matrix is

symmetric, the characteristic equation has only real roots (Anton, 1984, p. 289) and, since it is positive definite, its eigenvalues are positive, which is verified directly by taking the dot product with the eigenvalue problem statement. Thus, the factored form of this polynomial is

$$\left(p^2 - \frac{1}{v_1(\boldsymbol{x},\boldsymbol{p})^2}\right)\left(p^2 - \frac{1}{v_2(\boldsymbol{x},\boldsymbol{p})^2}\right)\left(p^2 - \frac{1}{v_3(\boldsymbol{x},\boldsymbol{p})^2}\right) = 0, \tag{2.83}$$

where $1/v_i(\boldsymbol{x},\boldsymbol{p})^2$ are the roots of the characteristic polynomial (Slawinski, 2020a, p. 328) and, in view of Christoffel matrix (2.82), are necessarily a function of position and the slowness vectors.

Since polynomial (2.83) has three roots, there are three distinct phase velocities that propagate in an anisotropic medium. The greatest of these corresponds to the so-called quasi-compressional (quasi-P or qP) wave. The other two roots correspond to the quasi-shear (quasi-S_1 and quasi-S_2 or qS_1 and qS_2) waves; alternatively, quasi-shear can be written as qSV and SH, where V corresponds to a vertical shear and H to a horizontal shear.

Regarding the use of P and S, in isotropic media, characteristic polynomial (2.83) only has two roots, which gives rise to the P and S waves. The displacement of the P wave is in the direction of propagation whereas it is perpendicular to them for S waves (Červený, 2001, Section 2.2.9). This leads to the adjectival distinctions of compressional and shear, which refers to the displacements of the respective waves. Since these displacements are not entirely aligned in anisotropic media, it is common to refer to these three waves with the adjective "quasi". In Chapter 3, we discuss the phase velocity of the qP, qSV, and SH waves within the symmetry class of transverse isotropy.

2.5.3 Eikonal equation

In expression (2.70), we demonstrated that the slowness vector is the gradient of the so-called eikonal function, namely, $\boldsymbol{p} = \operatorname{grad}\psi(\boldsymbol{x})$. Thus, with respect to $\psi(\boldsymbol{x})$, each root of characteristic polynomial (2.83) is a nonlinear first-order partial differential

equation. In view of expression (2.67), the roots can be written as

$$\sum_{k=1}^{3} p_k^2 = \frac{1}{v_i(\boldsymbol{x}, \boldsymbol{p})^2}, \quad i = 1, 2, 3, \tag{2.84}$$

where p_k are the components of the slowness vector. Expression (2.84) is known as the eikonal equation and it applies to each of the three aforementioned waves.

To solve the eikonal equation, we use the method of characteristics. The purpose of this method is to recast the partial differential equation into a set of ordinary differential equation (ODE) through a change of variables (e.g., Farlow, 1982, Lesson 27). To demonstrate, we can abbreviate the eikonal equation with the so-called Hamiltonian function,

$$\mathcal{H} = \mathcal{H}(\boldsymbol{x}, \boldsymbol{p}) = \text{constant}, \tag{2.85}$$

whose total differential is

$$\mathrm{d}\mathcal{H} = \sum_{j=1}^{3} \left(\frac{\partial \mathcal{H}}{\partial p_j} \mathrm{d}p_j + \frac{\partial \mathcal{H}}{\partial x_j} \mathrm{d}x_j \right) = 0. \tag{2.86}$$

Notice that expression (2.86) can only be satisfied if

$$-\frac{\mathrm{d}p_j}{\left(\dfrac{\partial \mathcal{H}}{\partial x_j}\right)} = \frac{\mathrm{d}x_j}{\left(\dfrac{\partial \mathcal{H}}{\partial p_j}\right)} = \mathrm{d}u, \tag{2.87}$$

where $\mathrm{d}u$ is an arbitrary differential parameter. Expression (2.87) can be separated into two derivatives,

$$\frac{\mathrm{d}p_j}{\mathrm{d}u} = -\frac{\partial \mathcal{H}}{\partial x_j} \quad \text{and} \quad \frac{\mathrm{d}x_j}{\mathrm{d}u} = \frac{\partial \mathcal{H}}{\partial p_j}. \tag{2.88}$$

The utility of $\mathrm{d}u$ is clear in the context of the derivative of the eikonal function, $\psi(\boldsymbol{x})$, where, by the chain rule and expressions (2.70) and (2.87), we have

$$\frac{\mathrm{d}\psi}{\mathrm{d}u} = \sum_{j=1}^{3} \frac{\partial \psi}{\partial x_j} \frac{\mathrm{d}x_j}{\mathrm{d}u} = \sum_{j=1}^{3} p_j \frac{\partial \mathcal{H}}{\partial p_j}. \tag{2.89}$$

Together, derivatives (2.88) and (2.89) comprise the so-called system of characteristic

equations to partial differential equation (2.85). Since these derivatives depend on $\mathcal{H}$, we can only specify the physical meaning of u — as, for example, a parameter of time or arclength — upon choosing a parameterization.

In this dissertation, we focus on two parameterizations, $\mathcal{H} = 0$ and $\mathcal{H} = 1/2$, both of which can be used to obtain raytracing equations in media with different velocity profiles. For the former, we consider a general parameterization suggested by Červený (2001, expression (3.1.6)),

$$\mathcal{H} = \frac{1}{n}\left(\left(\sum_{k=1}^{3} p_k^2\right)^{\frac{n}{2}} - \frac{1}{v(\boldsymbol{x})^n}\right) = 0, \tag{2.90}$$

which is valid for $n \neq 0$, applies to inhomogeneous isotropic media, and $v(\boldsymbol{x})$ applies to any of the three waves. It can be demonstrated that the associated characteristic system is

$$\frac{\mathrm{d}x_j}{\mathrm{d}u} = -\left(\sum_{k=1}^{3} p_k^2\right)^{\frac{n}{2}-1} p_j, \quad \frac{\mathrm{d}p_j}{\mathrm{d}u} = \frac{1}{n}\frac{\partial}{\partial x_j}\left(\frac{1}{v(\boldsymbol{x})^n}\right), \quad \frac{\mathrm{d}\psi}{\mathrm{d}u} = \frac{1}{v(\boldsymbol{x})^n}. \tag{2.91}$$

In Chapter 4, we develop an approximate traveltime model based on raytracing expressions (2.93) within homogeneous media, for which we use $n = 1$ and set remove the positional dependence of velocity, i.e., $v(\boldsymbol{x}) = v_0$. Characteristic system (2.91) becomes

$$\frac{\mathrm{d}x_j}{\mathrm{d}u} = -v_0 p_j, \quad \frac{\mathrm{d}p_j}{\mathrm{d}u} = 0, \quad \frac{\mathrm{d}\psi}{\mathrm{d}u} = \frac{1}{v_0}. \tag{2.92}$$

Since the eikonal function is related to time and $\mathrm{d}\psi/\mathrm{d}u$ results in a reciprocal measure of velocity, it stands to reason that u is a unit of arclength, s. We can solve ODEs (2.92) directly through integration, where, using the third as an example, we obtain

$$\int_{s_0}^{s} \frac{\mathrm{d}\psi}{\mathrm{d}\sigma}\,\mathrm{d}\sigma = -\int_{s_0}^{s} \frac{1}{v_0}\,\mathrm{d}\sigma \implies \psi(s) - \psi(s_0) = \frac{s - s_0}{v_0},$$

which yields traveltime, ψ, as a function of position, s, with a linear dependence on velocity, v_0. In this manner, we obtain solutions for the characteristic system,

$$x_j(s) = x_0 + v_i p_j (s - s_0), \quad p_j(s) = (p_i)_0, \quad \psi(s) = \psi_0 + \frac{s - s_0}{v}, \tag{2.93}$$

where $x_j(s_0) = x_0$, $p_j(s_0) = (p_i)_0$, and $\psi(s_0) = \psi_0$ are initial values. Thus, solutions (2.93) indicate that, in homogeneous isotropic media, raypaths are straight lines and the slowness vector, otherwise referred to as a ray parameter, is constant.

For the latter parameterization, we adjust the eikonal equation (2.84) such that the right-hand side is equal to unity and then set

$$\mathcal{H} = \frac{1}{2} = \frac{\left(\sum_{k=1}^{3} p_k^2\right) v_i(\boldsymbol{x}, \boldsymbol{p})^2}{2}. \tag{2.94}$$

Obtaining the characteristic system for Hamiltonian (2.94) is a more complicated procedure than for Hamiltonian (2.90) due to the phase velocity dependence on both position and the slowness vector. This case applies to wave propagation in an inhomogeneous anisotropic medium, which requires taking into account the elasticity symmetry class of the medium and obtaining raytracing equations within the context the Christoffel matrix (e.g., Červený, 2001, Section 3.6). By contrast, we use the approach of Rogister and Slawinski (2005), which, through several simplifying approximations, yields closed-form raytracing equations for a velocity profile that increases linearly with depth and depends elliptically on direction. In Chapter 5, we develop a model for traveltime inversion based on these raytracing equations.

Chapter 3

On Christoffel roots for nondetached slowness surfaces

Author note

The majority of this chapter appears elsewhere in the scientific literature as Bos et al. (2019a), "On Christoffel roots for nondetached slowness surfaces", which is a peer-reviewed article published in Geophysical Prospecting by Len Bos, Michael A. Slawinski, and Theodore Stanoev. Additionally, the preprint of the accepted article is Bos et al. (2019b). For further details regarding co-authorship, we refer the reader to the statement provided on page xi.

3.1 Introductory remarks

Since the studies of Rudzki (1911)[1], characterizing the shapes of wavefronts in anisotropic media has been of interest to seismologists. A wavefront represents two-dimensional curve (or surface in three-dimensions) connecting points of equal phase (e.g., Serway and Jewett, 2013, Section 24.3). In equation (2.63), we consider wavefronts that correspond to a plane wave, whose phase is a function of traveltime. As such, we consider a wavefront to represent a locus of equal traveltime (Sheriff, 2002, p. 388). From a geometric perspective, a wavefront propagating through such

[1]Rudzki (1911) was presented to the Academy of Sciences at Cracow in 1911. His work has been translated with comments by Klaus Helbig and Michael A. Slawinski; it appears as Rudzki (2003).

a medium will change shape as a result of changes in velocity in different directions within the medium. The practical aspect of understanding the shapes of wavefront is that if we select a point on a wavefront and, as it propagates, follow the changes in the direction of its normal, we have the trajectory of the raypath corresponding to that point of the wavefront (Udías, 1999, Section 6.2).

As we discuss in Section 2.5.2, the existence of waves in such media are governed by the Christoffel equation (2.80), which can be used to obtain the phase velocities of such waves. To facilitate the characterization of the wavefront shapes, seismologists often consider the wave's slowness surface, which is a constant-velocity surface specified by the reciprocals of the phase velocity (Sheriff, 2002, p. 323). Seismologists consider the slowness surface because it is related the wavefront by polar reciprocity. In particular, two curves are polar reciprocals if for each point on the first curve the radius vector is parallel to the normal of the second, and vice versa (Helbig, 1994, Section 2A.4).

In anisotropic media, a slowness surface is comprised of three sheets, one for each of the roots of characteristic polynomial (2.83), which is the factorized form of the solubility condition of the Christoffel equation. Each sheet traces out a closed continuous surface that is symmetric about the origin. The innermost sheet corresponds to qP waves as its phase velocity is the greatest and it is always convex whereas the other two sheets that correspond to the qSV and SH waves can have regions of concavity in media with strong anisotropy (Červený, 2001, Section 2.2). Typically, the innermost sheet is completely separated from the outer two; we refer to this scenario as *detached* whereas we use *nondetached* if there is no separation.

From a historical perspective within the foundations of seismology, it has been assumed that the innermost sheet is never ellipsoidal. Let us provide some context to this statement. Postma (1955) derived a condition for elliptical velocity dependence in homogeneous transversely isotropic media that is equivalent to alternating isotropic layers. This condition was generalized by Berryman (1979) for "any horizontally stratified, homogeneous material whose constituent layers are isotropic." The proof for nonexistence of ellipticity of qP wavefronts in media resulting from lamellation came from Helbig (1979), in response to Levin (1978). Shortly thereafter, Helbig (1983, p. 826) stated the following:

(1) the wavefront of qP waves is never an ellipsoid;

(2) the wavefront of qSV waves is never an ellipsoid;

(3) the wavefront of SH waves is always an oblate ellipsoid.

Lamellation, which is described by Helbig (1979, 1983) as fine layering on a scale small compared with the wavelength, is tantamount to using the Backus (1962) average; throughout this chapter, we use the methodology of the latter.

The purpose of this chapter is to demonstrate that the commonly accepted propositions of Helbig (1983) are incomplete and have restrictions that bear on their validity. By specifying the conditions under which the propositions do not hold, we contribute to the foundational understanding of seismology. We present several examples to demonstrate our findings and, in so doing, confirm the importance of questioning theoretical issues.

To proceed with our demonstration, we consider the three roots of the solubility condition of the Christoffel equation, which we refer to as *Christoffel roots*. These roots correspond to the wavefront-slowness surfaces of the three waves that propagate in an anisotropic Hookean solid. Herein, we examine transversely isotropic media resulting from the Backus average of isotropic layers to which we refer as *Backus media*. We derive the conditions under which the spherical-coordinate plots of the three roots are ellipsoidal; we refer to such roots as *elliptical*. In accordance with polar reciprocity, the ellipticity of slownesses is equivalent to ellipticity of wavefronts; for a verification of this statement, we refer the reader to Appendix A.2.

As it turns out, a necessary condition for the ellipticity of roots in Backus media is the nondetachment of the qP slowness surface. Although the Hookean solids that represent most materials encountered in seismology exhibit a detached qP slowness surface, the existence of both detached and nondetached slowness surfaces is, indeed, permissible within the stability condition of the elasticity tensor (Bucataru and Slawinski, 2009). Mathematically, this condition is the positive definiteness of the elasticity tensor.

3.2 Christoffel equation in Backus media

The existence of waves in anisotropic media is governed by the Christoffel equation (2.80); its solubility condition is (e.g., Slawinski, 2015, Section 7.3)

$$\det\left[\sum_{j=1}^{3}\sum_{\ell=1}^{3} c_{ijk\ell}\, p_j p_\ell - \delta_{ik}\right] = 0\,, \qquad i,k = 1\,,\, 2\,,\, 3\,, \tag{3.1}$$

where $c_{ijk\ell}$ is a density-scaled elasticity tensor and $\boldsymbol{p}$ is the wavefront-slowness vector. The three roots of this cubic equation in p^2 can be stated as the expressions for the wavefront speeds of the qP, qSV and SH waves, which we demonstrate in expression (2.83); herein, $p^2 = \boldsymbol{p}\cdot\boldsymbol{p}$ is the squared magnitude of the slowness vector.

3.2.1 Backus average

Within the context of the Backus average (Backus, 1962), a homogeneous transversely isotropic medium is long-wave-equivalent to a stack of thin isotropic or transversely isotropic layers. This low-frequency wave propagates with normal (or quasi-normal) incidence to the stack of thin layers, whose layering is consistently parallel but the thicknesses of layers need not be uniform. The averaging process uses the function $f(x_3)$ of "width" ℓ' is the moving average given by

$$\overline{f}(x_3) := \int_{-\infty}^{\infty} w(\zeta - x_3) f(\zeta)\,\mathrm{d}\zeta\,,$$

where the weight function, $w(x_3)$, has the following properties:

$$w(x_3) \geqslant 0, \quad w(\pm\infty) = 0,$$

$$\int_{-\infty}^{\infty} w(x_3)\,\mathrm{d}x_3 = 1\,, \quad \int_{-\infty}^{\infty} x_3 w(x_3)\,\mathrm{d}x_3 = 0\,, \quad \int_{-\infty}^{\infty} x_3^2 w(x_3)\,\mathrm{d}x_3 = (\ell')^2\,.$$

These properties define $w(x_3)$ as a probability-density function with mean 0 and standard deviation ℓ', explaining the use of the term "width" for ℓ'. The average itself is a moving average with a specifically set size of window. The size of the window might be determined by the wavelength of the propagating wave.

The medium through which the Backus average propagates is considered to be in static equilibrium, i.e. disturbances created by the Backus average are infinitesimal. The stress and strain of the medium, along the x_3-axis are constant along the interfaces of layers. Components in the $x_1\,x_2$-plane can vary significantly along the x_3-axis due to distinct properties of different layers. Furthermore, the components of the displacement vector, u_1, u_2, u_3, are continuous. In view of the lateral homogeneity of the medium, $\partial u_i/\partial x_1$ and $\partial u_i/\partial x_2$ are smoothly varying functions of x_3, but $\partial u_i/\partial x_3$ might not be, due to the vertical inhomogeneity of layers. Hence, the medium satisfies the dynamic and kinematic boundary conditions, wherein the former refers to equality of forces along an interface and the latter refers to a welded contact between layers (Slawinski, 2015, Section 10.2.1).

For the case of isotropic layers, elasticity parameters of each layer are c_{1111} and c_{2323}, and the corresponding parameters of the transversely isotropic medium are

$$c_{1111}^{\overline{\text{TI}}} = \overline{\left(\frac{c_{1111}-2\,c_{2323}}{c_{1111}}\right)^2}\;\overline{\left(\frac{1}{c_{1111}}\right)}^{-1} + \overline{\left(\frac{4(c_{1111}-c_{2323})c_{2323}}{c_{1111}}\right)}, \tag{3.2a}$$

$$c_{1122}^{\overline{\text{TI}}} = \overline{\left(\frac{c_{1111}-2\,c_{2323}}{c_{1111}}\right)^2}\,\overline{\left(\frac{1}{c_{1111}}\right)}^{-1} + \overline{\left(\frac{2(c_{1111}-2\,c_{2323})c_{2323}}{c_{1111}}\right)}, \tag{3.2b}$$

$$c_{1133}^{\overline{\text{TI}}} = \overline{\left(\frac{c_{1111}-2\,c_{2323}}{c_{1111}}\right)}\;\overline{\left(\frac{1}{c_{1111}}\right)}^{-1}, \tag{3.2c}$$

$$c_{1212}^{\overline{\text{TI}}} = \overline{c_{2323}}\,, \tag{3.2d}$$

$$c_{2323}^{\overline{\text{TI}}} = \overline{\left(\frac{1}{c_{2323}}\right)}^{-1}, \tag{3.2e}$$

$$c_{3333}^{\overline{\text{TI}}} = \overline{\left(\frac{1}{c_{1111}}\right)}^{-1}. \tag{3.2f}$$

Herein, superscript $^{\overline{\text{TI}}}$ indicates transverse isotropy resulting from the Backus average. The Backus average of a stack of different isotropic layers results in a transversely isotropic medium; a stack of layers exhibiting another symmetry results in the medium of that symmetry; a stack of layers of various symmetries results in a medium of the lowest among these symmetries, as is exemplified by Bos et al. (2017) (details of derivations to other symmetry classes are also discussed therein).

Now, regarding the possible invertibility of the Backus average, let us remark that

the precise information concerning the individual layer properties is lost during the averaging process. As such, inferences of such properties from their averages lack uniqueness (Danek and Slawinski, 2016, Section 4.3). With that being said, it is known that the properties of the constituent layers affect the anisotropy of the equivalent medium, where, for example, should the rigidity (c_{2323}) be constant throughout the layers, the Backus average results in an isotropic equivalent medium (Backus, 1962, Section 6). Adamus et al. (2020) investigate further the effects of inhomogeneity on anisotropy to find that the strength of anisotropy in the transversely isotropic equivalent medium is solely a measure of inhomogeneity among the isotropic layers. Therein, they show that the anisotropy of the Backus average is a consequence of the difference in rigidity among layers, not in compressibility (c_{1111}). Such a scenario would correspond to that of a porous rock of constant rigidity, whose compressibility varies depending on the amount of liquid within its pores. Therefore, while inverting the Backus average for its unique constituent layers might not be possible, it is possible to obtain a qualitative measure of the stack's inhomogeneity through the strength of anisotropy of the equivalent medium.

3.2.2 Backus media

Let us consider a homogeneous transversely isotropic medium, whose elasticity parameters are

$$c^{\overline{\mathrm{TI}}} = \begin{bmatrix} c^{\overline{\mathrm{TI}}}_{1111} & c^{\overline{\mathrm{TI}}}_{1122} & c^{\overline{\mathrm{TI}}}_{1133} & 0 & 0 & 0 \\ c^{\overline{\mathrm{TI}}}_{1122} & c^{\overline{\mathrm{TI}}}_{1111} & c^{\overline{\mathrm{TI}}}_{1133} & 0 & 0 & 0 \\ c^{\overline{\mathrm{TI}}}_{1133} & c^{\overline{\mathrm{TI}}}_{1133} & c^{\overline{\mathrm{TI}}}_{3333} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\,c^{\overline{\mathrm{TI}}}_{2323} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\,c^{\overline{\mathrm{TI}}}_{2323} & 0 \\ 0 & 0 & 0 & 0 & 0 & c^{\overline{\mathrm{TI}}}_{1111} - c^{\overline{\mathrm{TI}}}_{1122} \end{bmatrix} . \tag{3.3}$$

If we consider a stack of n isotropic layers, whose elasticity parameters are

$$c_{1111} = \{(c_{1111})_1\,,\,\ldots\,,\,(c_{1111})_n\} \qquad \text{and} \qquad c_{2323} = \{(c_{2323})_1\,,\,\ldots\,,\,(c_{2323})_n\}\,, \tag{3.4}$$

the stability condition for each layer is (e.g., Slawinski, 2018, Exercise 5.3)

$$(c_{1111})_i > \frac{4}{3}\,(c_{2323})_i > 0\,, \qquad i = 1\,,\,\ldots\,,\,n\,. \tag{3.5}$$

For notational convenience, we write the Backus-average elasticity parameters of expressions (3.2) as

$$c_{1111}^{\overline{\mathrm{TI}}} = \left(1 - \frac{2}{n}\,Y\right)^2 \left(n\,W^{-1}\right) + \frac{4}{n}\,(U - Z)\;, \tag{3.6a}$$

$$c_{1122}^{\overline{\mathrm{TI}}} = \left(1 - \frac{2}{n}\,Y\right)^2 \left(n\,W^{-1}\right) + \frac{2}{n}\,(U - 2\,Z)\;, \tag{3.6b}$$

$$c_{3333}^{\overline{\mathrm{TI}}} = n\,W^{-1}\,, \tag{3.6c}$$

$$c_{1133}^{\overline{\mathrm{TI}}} = \left(1 - \frac{2}{n}\,Y\right) \left(n\,W^{-1}\right)\;, \tag{3.6d}$$

$$c_{2323}^{\overline{\mathrm{TI}}} = n\,V^{-1}\,, \tag{3.6e}$$

$$c_{1212}^{\overline{\mathrm{TI}}} = n^{-1}\,U\,, \tag{3.6f}$$

where

$$\begin{gathered} U := \sum_{i=1}^{n} \left(c_{2323}\right)_i\;, \quad V := \sum_{i=1}^{n} \frac{1}{\left(c_{2323}\right)_i}\,, \quad W := \sum_{i=1}^{n} \frac{1}{\left(c_{1111}\right)_i}\,, \\ Y := \sum_{i=1}^{n} \frac{\left(c_{2323}\right)_i}{\left(c_{1111}\right)_i}\,, \quad Z := \sum_{i=1}^{n} \frac{\left[\left(c_{2323}\right)_i\right]^2}{\left(c_{1111}\right)_i}\,. \end{gathered} \tag{3.7}$$

A standard form of these parameters is given by, for example, Slawinski (2018, Section 4.2.2); the expressions, therein, and those of parameterizations (3.6), are equivalent to A, B, C, F, L, M of Backus (1962, equations (13)), respectively; these expressions are analogous to expressions (3.2). The stability of the Backus average is inherited from the stability of the layers (Slawinski, 2018, Proposition 4.1); in other words, if the layers are stable, so is the average.

3.3 Christoffel roots

Returning to Christoffel equations (3.1), the three roots within transversely isotropic media are (e.g., Slawinski, 2015, equation (9.2.19), (9.2.20))

$$v_{qP,qSV}(\vartheta) = \sqrt{\frac{\left(c_{3333}^{\overline{\mathrm{TI}}} - c_{1111}^{\overline{\mathrm{TI}}}\right)(1 - n_3) + c_{1111}^{\overline{\mathrm{TI}}} + c_{2323}^{\overline{\mathrm{TI}}} \pm \sqrt{\Delta(\vartheta)}}{2\,\rho}} \tag{3.8}$$

and

$$v_{SH}(\vartheta) = \sqrt{\frac{c_{1212}^{\overline{\mathrm{TI}}}\, n_3 + c_{2323}^{\overline{\mathrm{TI}}}\, (1 - n_3)}{\rho}}\,. \tag{3.9}$$

Herein, we use $n_3 = \sin^2\vartheta$ to express the wavefront orientation and we parameterize

$$\Delta(\vartheta) = a\,(n_3)^2 + b\,n_3 + c,$$

where

$$a = \left(c_{1111}^{\overline{\mathrm{TI}}} + 2\,c_{1133}^{\overline{\mathrm{TI}}} + c_{3333}^{\overline{\mathrm{TI}}}\right)\left(c_{1111}^{\overline{\mathrm{TI}}} - 2\,c_{1133}^{\overline{\mathrm{TI}}} - 4\,c_{2323}^{\overline{\mathrm{TI}}} + c_{3333}^{\overline{\mathrm{TI}}}\right), \tag{3.10a}$$

$$b = 2\,c_{1111}^{\overline{\mathrm{TI}}}\,c_{2323}^{\overline{\mathrm{TI}}} - 2\,c_{1111}^{\overline{\mathrm{TI}}}\,c_{3333}^{\overline{\mathrm{TI}}} + 4\left(c_{1133}^{\overline{\mathrm{TI}}}\right)^2 + 8\,c_{1133}^{\overline{\mathrm{TI}}}\,c_{2323}^{\overline{\mathrm{TI}}} + 6\,c_{2323}^{\overline{\mathrm{TI}}}\,c_{3333}^{\overline{\mathrm{TI}}} - 2\left(c_{3333}^{\overline{\mathrm{TI}}}\right)^2, \tag{3.10b}$$

$$c = \left(c_{2323}^{\overline{\mathrm{TI}}} - c_{3333}^{\overline{\mathrm{TI}}}\right)^2. \tag{3.10c}$$

Next, for elliptical roots, we require the reciprocals of the roots in the form of $\sqrt{d + f\sin^2\vartheta}$, where d and f are nonzero real constants. To demonstrate, we begin with the standard form of an ellipse,

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1,$$

where x and y are Cartesian coordinates and a and b are positive real constants. Using the relationship between Cartesian and polar coordinates,

$$x = r\cos\vartheta \quad \text{and} \quad y = r\sin\vartheta,$$

for a positive radius, r, and central angle, ϑ, the standard form is

$$\left(\frac{r\cos\vartheta}{a}\right)^2 + \left(\frac{r\sin\vartheta}{b}\right)^2 = 1.$$

Solving for the squared radius, and availing of the Pythagorean trigonometric identity, we have

$$r^2 = \frac{a^2b^2}{b^2(1 - \sin^2\vartheta) + a^2\sin^2\vartheta}.$$

Taking the square root of both sides and simplifying yields

$$r = \frac{ab}{\sqrt{b^2 + (a^2 - b^2)\sin^2\vartheta}},$$

where we omit the negative solution since $r > 0$. Thus, setting $f := a^2 - b^2$, substituting $b = 1/a$, and setting $d := 1/a^2$, we obtain

$$r = \frac{1}{\sqrt{d + f\sin^2\vartheta}},$$

as required.

Returning to expression (3.9), we observe that the reciprocal of v_{SH} is already in the form of an ellipse. However, for reciprocals of roots (3.8) are elliptical if and only if Δ is a perfect square. To demonstrate, we complete the square to obtain

$$\Delta = a\left(n_3 + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right).$$

For a perfect square, we require the latter term to equal zero, which imposes the condition

$$\mathrm{Disc}(\Delta) := b^2 - 4ac = 0,$$

or, in other words, if and only if the discriminant of Δ, which we denote by $\mathrm{Disc}\,(\Delta)$, is zero. If the condition is satisfied, expression (3.8) is in the form of $\sqrt{d + f\sin^2\vartheta}$, where d and f are nonzero real constants. In view of expressions (3.10a)–(3.10c),

$$\mathrm{Disc}\,(\Delta) := 16\left(c_{1133}^{\overline{\mathrm{TI}}} + c_{2323}^{\overline{\mathrm{TI}}}\right)^2\left[c_{1111}^{\overline{\mathrm{TI}}}\left(c_{2323}^{\overline{\mathrm{TI}}} - c_{3333}^{\overline{\mathrm{TI}}}\right) + \left(c_{1133}^{\overline{\mathrm{TI}}}\right)^2 + 2\,c_{1133}^{\overline{\mathrm{TI}}}\,c_{2323}^{\overline{\mathrm{TI}}} + c_{2323}^{\overline{\mathrm{TI}}}\,c_{3333}^{\overline{\mathrm{TI}}}\right] = 0\,, \tag{3.11}$$

for which the solutions are

$$c_{2323}^{\overline{\mathrm{TI}}} = -c_{1133}^{\overline{\mathrm{TI}}}\,, \tag{3.12a}$$

$$c_{2323}^{\overline{\mathrm{TI}}} = -c_{1133}^{\overline{\mathrm{TI}}} \quad \text{and} \quad c_{1111}^{\overline{\mathrm{TI}}} = c_{2323}^{\overline{\mathrm{TI}}}\,, \tag{3.12b}$$

$$c_{3333}^{\overline{\mathrm{TI}}} = \frac{\left(c_{1133}^{\overline{\mathrm{TI}}}\right)^2 + c_{1111}^{\overline{\mathrm{TI}}}\,c_{2323}^{\overline{\mathrm{TI}}} + 2\,c_{1133}^{\overline{\mathrm{TI}}}\,c_{2323}^{\overline{\mathrm{TI}}}}{c_{1111}^{\overline{\mathrm{TI}}} - c_{2323}^{\overline{\mathrm{TI}}}}\,. \tag{3.12c}$$

Discriminant (3.11), and its solutions (3.12), can be written in terms of parameterizations (3.6) as

$$\frac{64\,(n\,(V+W)-2\,V\,Y)^2\,((n-Y)^2+(V-W)\,(Z-U))}{V^3\,W^3}=0\,, \tag{3.13}$$

whose solutions are

$$Y=\frac{n}{2}\left(1+W\,V^{-1}\right)\,, \tag{3.14a}$$

$$W=V\quad\text{and}\quad Y=n\,, \tag{3.14b}$$

$$Z=U-\frac{(n-Y)^2}{V-W}\,. \tag{3.14c}$$

Let us discuss solutions (3.14). With regard to solution (3.14b), we recognize that it is a special case of solution (3.14a), but we keep both for convenient referencing, below. Within the context of the Backus average, solution (3.14b) cannot be satisfied by the stability condition as it would require

$$W=V\implies\sum_{i=1}^{n}\frac{1}{(c_{1111})_i}=\sum_{i=1}^{n}\frac{1}{(c_{2323})_i},$$

which is not allowed.

For solution (3.12c), the equality can be satisfied if and only if c_{2323} is the same for all layers. To demonstrate, we set $(c_{2323})_i=c_{2323}$ in parameterizations (3.7), which, by summation arithmetic, results in

$$Z=(c_{2323})^2W,\quad Y=c_{2323}W,\quad V=\frac{n}{c_{2323}},\quad\text{and}\quad U=n\,c_{2323}.$$

Substituting these terms in solution (3.14c) and simplifying, we obtain

$$\begin{aligned}(c_{2323})^2W&=n\,c_{2323}-\frac{(n-c_{2323}W)^2}{\frac{n}{c_{2323}}-W}\\&=n\,c_{2323}-c_{2323}\frac{(n-c_{2323}W)^2}{n-c_{2323}W}\\&=n\,c_{2323}-c_{2323}(n-c_{2323}W)\\&=(c_{2323})^2W,\end{aligned}$$

as required. However, as indicated by Backus (1962, Section 6), if a layered isotropic medium has constant c_{2323} in all layers, the resulting average is isotropic. Since our focus is a transversely isotropic average, we omit solution (3.12c) from our consideration.

As such, the only viable condition that remains is solution (3.14a). In Section 3.3.1, we proceed to prove the existence of solution (3.14a) for $n \geqslant 4$ layers, followed by a numerical example to illustrate the result.

3.3.1 Nondetachment

Within the constraints of the stability, both detachment and nondetachment are permitted. The nondetachment of the qP slowness surface occurs if and only if $c_{2323}^{\overline{\text{TI}}} = -c_{1133}^{\overline{\text{TI}}}$. In Figure 3.2, we present an illustration of detached slowness surfaces becoming nondetached as $c_{2323}^{\overline{\text{TI}}} \to -c_{1133}^{\overline{\text{TI}}}$. From solution (3.12a), and its parameterization (3.14a), it follows that roots (3.8) are elliptical. Hence, we have the following lemma.

Lemma 1. *There exists a Backus average of at least four isotropic layers for which the Christoffel roots are elliptical.*

Proof. We fix $a > \frac{4}{3}$ and let $x \in (0\,,1]$ so that

$$c_{1111} = \left(\frac{a}{x}\,,\frac{a}{x}\,,\frac{a}{x}\,,\frac{a}{x}\right) \qquad \text{and} \qquad c_{2323} = \left(x\,,\frac{1}{x}\,,\frac{1}{x}\,,\frac{1}{x}\right). \tag{3.15}$$

Following definitions (3.7), variables Y, V, and W of solution (3.14a) become

$$Y(x) = \frac{x^2}{a} + \sum_{i=2}^{n} \frac{(c_{2323})_i}{(c_{1111})_i} = \frac{x^2}{a} + \sum_{i=2}^{n} \frac{1/x}{a/x} = \frac{x^2}{a} + \frac{n-1}{a}\,, \tag{3.16a}$$

$$V(x) = \frac{1}{x} + \sum_{i=2}^{n} \frac{1}{(c_{2323})_i} = \frac{1}{x} + \sum_{i=2}^{n} \frac{1}{1/x} = \frac{1+(n-1)\,x^2}{x}\,, \tag{3.16b}$$

$$W(x) = \sum_{i=1}^{n} \frac{1}{(c_{1111})_i} = \sum_{i=1}^{n} \frac{1}{a/x} = \frac{n\,x}{a}\,. \tag{3.16c}$$

We define $X(x) := \frac{n}{2}\left(1 + W(x)\,V(x)^{-1}\right)$, which results in

$$X(x) = \frac{n}{2}\left\{1 + \left(\frac{n\,x}{a}\right)\left[\frac{1+(n-1)\,x^2}{x}\right]^{-1}\right\} = \frac{n}{2}\left\{1 + \frac{1}{a}\left[\frac{n\,x^2}{1+(n-1)\,x^2}\right]\right\}. \tag{3.17}$$

As $x \to 0^+$, we have

$$Y(0^+) \to \frac{n-1}{a} \approx \frac{3}{4}(n-1), \quad X(0^+) \to \frac{n}{2}, \tag{3.18}$$

and, hence, $Y(0^+) > X(0^+)$ for a close to $\frac{4}{3}$. Furthermore, as $x \to 1^-$,

$$Y(1^-) \to \frac{1}{a} + \frac{n-1}{a} \quad \text{and} \quad X(1^-) \to \frac{n}{2}\left(1 + \frac{1}{a}\right). \tag{3.19}$$

Thus, for $n \geqslant 4$, and a close to, but greater than, $\frac{4}{3}$, $Y(1^-) < X(1^-)$.

It follows form the Intermediate Value Theorem that there exists an $x \in (0,1)$ for which $Y(x) = X(x)$, which completes the proof. □

3.3.2 Numerical example

Let us consider a numerical example, where

$$x = 0.2299, \quad a = 1.3440, \quad \text{and} \quad Y = X = 2.2714; \tag{3.20}$$

indeed, there exists a Backus average of at least four isotropic layers, whose Christoffel roots are elliptical.

Using results (3.20) with parameters (3.15), we obtain

$$\begin{aligned} c_{1111} &= \{5.8472\,,\, 5.8472\,,\, 5.8472\,,\, 5.8472\} \quad \text{and} \\ c_{2323} &= \{0.2299\,,\, 4.3506\,,\, 4.3506\,,\, 4.3506\}\,; \end{aligned}$$

the Backus-average parameters, following expressions (3.6), are

$$\begin{aligned} c_{1111}^{\overline{\text{TI}}} &= 3.6692\,, \quad c_{1122}^{\overline{\text{TI}}} = -2.9717\,, \quad c_{3333}^{\overline{\text{TI}}} = 5.8472\,, \\ c_{1133}^{\overline{\text{TI}}} &= -0.7936\,, \quad c_{2323}^{\overline{\text{TI}}} = 0.7936\,, \quad c_{1212}^{\overline{\text{TI}}} = 3.3204\,. \end{aligned} \tag{3.21}$$

The eigenvalues of tensor (3.3) with values (3.21) are $\lambda_1 = \lambda_2 = 6.6409$, $\lambda_3 = 6.0812$, $\lambda_4 = \lambda_5 = 1.5873$, $\lambda_6 = 0.4635$, which belong to a transversely isotropic tensor (Bóna et al., 2007a); since they are positive, the stability condition of the average are satisfied. Also, $\mathrm{Disc}\,(\Delta) = 1.9883 \times 10^{-12}$, which can be considered zero, as required. Consequently, equation (3.8) becomes

$$v_{qP,qSV}(\vartheta) = \sqrt{\frac{4.4628 + 2.1781\ (1 - n_3) \pm \sqrt{62.8718\ (0.6373 - n_3)^2}}{2\,\rho}}\,. \tag{3.22}$$

Recalling that $n_3 = \sin^2\vartheta$, and since ρ must be a positive scalar quantity, we see that equation (3.22) is $\sqrt{d + f\sin^2\vartheta}$, as required. Letting $\rho = 1$, the three roots are

$$v_{qP}(\vartheta) = \sqrt{\frac{\left(c^{\overline{\mathrm{TI}}}_{3333} - c^{\overline{\mathrm{TI}}}_{1111}\right)(1 - n_3) + c^{\overline{\mathrm{TI}}}_{1111} + c^{\overline{\mathrm{TI}}}_{2323} + \sqrt{\Delta}}{2\,\rho}} = \sqrt{5.8472 + 5.0536\ \sin^2\vartheta} \tag{3.23a}$$

$$v_{qSV}(\vartheta) = \sqrt{\frac{\left(c^{\overline{\mathrm{TI}}}_{3333} - c^{\overline{\mathrm{TI}}}_{1111}\right)(1 - n_3) + c^{\overline{\mathrm{TI}}}_{1111} + c^{\overline{\mathrm{TI}}}_{2323} - \sqrt{\Delta}}{2\,\rho}} = \sqrt{0.7936 - 2.8756\ \sin^2\vartheta}\,, \tag{3.23b}$$

$$v_{SH}(\vartheta) = \sqrt{\frac{c^{\overline{\mathrm{TI}}}_{1212}\,n_3 + c^{\overline{\mathrm{TI}}}_{2323}\ (1 - n_3)}{\rho}} = \sqrt{0.7936 + 2.5268\ \sin^2\vartheta}\,. \tag{3.23c}$$

The reciprocals of expressions (3.23) are ellipses, illustrated in Figure 3.1. Therein, the green curve represents $1/v_{SH}(\vartheta)$; the blue curve represents $1/v_{qSV}(\vartheta)$; and the red curve represents $1/v_{qP}(\vartheta)$.

3.3.3 Interpretation

Values (3.21) do not represent typical Hookean solids used in seismology. To quantify this statement, let us invoke Poisson's ratio,

$$\nu_{ji} := -\frac{\epsilon_{ii}}{\epsilon_{jj}}, \tag{3.24}$$

which is defined as the negative ratio of the lateral strain to the axial strain, in a uniaxial strain state, i.e., $\sigma_{jj} \neq 0$ and $\sigma_{ik} = 0$ for $i \neq j$ and $k = 1, 2, 3$ (e.g., Mavko

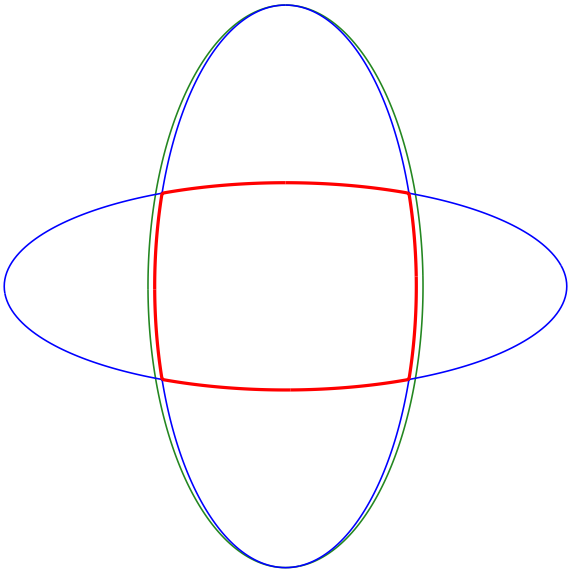
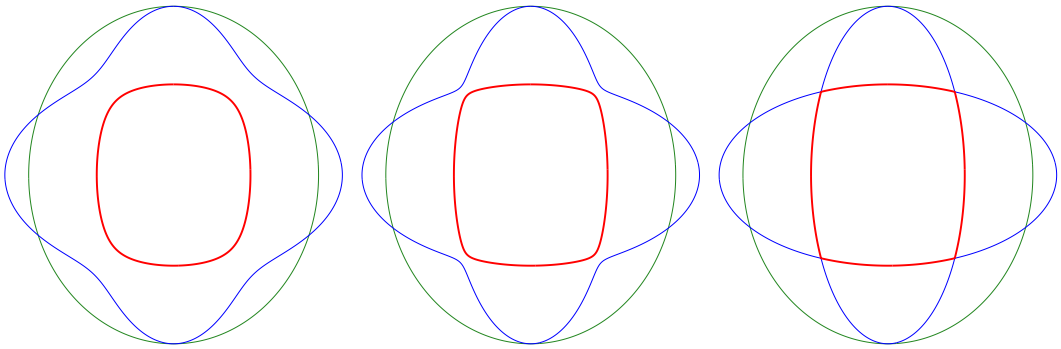


Figure 3.1: Three Christoffel roots resulting in three slowness curves



(a) (b) (c)

Figure 3.2: Slowness curves, $1/v_{qP}(\vartheta)$, $1/v_{qSV}(\vartheta)$, $1/v_{SH}(\vartheta)$, for modified values of elasticity parameters for Green-River shale. As $c^{\mathrm{TI}}_{1133} \to -c^{\mathrm{TI}}_{2323}$, the innermost curve ceases to be detached and smooth.

et al., 2009, p. 22). For transverse isotropy (e.g., Mavko et al., 2009, pp. 33–35), along with values (3.21),

$$\nu_{31} = -\frac{\epsilon_{11}}{\epsilon_{33}} = \frac{c^{\overline{\text{TI}}}_{1133}}{c^{\overline{\text{TI}}}_{1111} + c^{\overline{\text{TI}}}_{1122}} = -1.1378. \tag{3.25}$$

In contrast to value (3.25), Poisson's ratio for sedimentary rocks, such as limestone and sandstone, is typically positive — Ji et al. (2018, Table 3) present such values in a statistical analysis of typical rock and ore samples at various pressures. For shale, Gereck (2007, Figure 4) presents typical ranges for shale of $\nu \in (0.05, 0.32)$. However, Lakes (2017, Section 3.7) indicates that it is reasonable to obtain large negative Poisson's ratios in highly anisotropic materials: in arsenic crystals, the ratio can exceed $\frac{1}{2}$ or can be less than -1, which are the isotropic bounds for ν (e.g., Slawinski, 2020a, Exercise 5.18). With that being said, let us emphasize that the only restriction on the elasticity parameters is the positive definiteness of tensor (3.3). Since parameter values (3.21) satisfy this condition, they are acceptable and, hence, suffice as a counterexample to the propositions of Helbig (1983). Further justification of these values — concerning the likelihood of their observation in typical seismological settings — is beyond the scope of this analysis.

To gain an insight into the appearance of slowness curves in Figure 3.1, let us examine an example using the density-scaled elasticity parameters for Green-River shale (e.g., Slawinski (2015, Exercise 9.3), Thomsen (1986, Table 1)),

$$c^{\text{TI}}_{1111} = 13.55\,, \quad c^{\text{TI}}_{1133} = 1.47\,, \quad c^{\text{TI}}_{3333} = 9.74\,, \quad c^{\text{TI}}_{2323} = 2.81\,, \quad c^{\text{TI}}_{1212} = 3.81\,, \tag{3.26}$$

where each parameter is scaled by 10^6; herein, superscript $^{\text{TI}}$ refers to an intrinsically transversely isotropic medium, as opposed to $^{\overline{\text{TI}}}$, which is the Backus medium.

The curves of $1/v_{qP}(\vartheta)$, $1/v_{qSV}(\vartheta)$ and $1/v_{SH}(\vartheta)$ with values (3.26) are illustrated in Figure 3.2(a); such curves are typical in seismology. For the wavefronts of these slowness curves, which are obtained using polar reciprocity, we refer the reader to Appendix A.2.3. The progression of Figures 3.2(a)–(c), however, illustrates an important property. For detached qP slowness surfaces, the expressions for the qP, qSV and SH wavefront speeds, indeed, correspond to distinct smooth wavefronts. However, if $c^{\text{TI}}_{1133} = -c^{\text{TI}}_{2323}$, the qP and qSV slowness surfaces lose their smoothness. Also, their expressions — not only their curves — become connected with one another; neither root corresponds to a distinct slowness curve nor does a given curve

result from a single root.

For each slowness surface, the criterion of belonging to a particular wave on either side of an intersection is not its belonging to a single root but the orientation of the corresponding eigenvectors, which are the displacement vectors of a given wave (Bóna et al., 2007b). In Figure 3.2, the qSV and SH slownesses are equal along the rotation-symmetry axis. As such, the innermost surface corresponds to the qP wave.

3.4 Ellipticity condition

According to Thomsen (1986), the ellipticity condition is $\varepsilon = \delta$, where

$$\varepsilon = \frac{c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{3333}}{2\,c^{\mathrm{TI}}_{3333}} \qquad \text{and} \qquad \delta = \frac{\left(c^{\mathrm{TI}}_{1133} + c^{\mathrm{TI}}_{2323}\right)^2 - \left(c^{\mathrm{TI}}_{3333} - c^{\mathrm{TI}}_{2323}\right)^2}{2\,c^{\mathrm{TI}}_{3333}\left(c^{\mathrm{TI}}_{3333} - c^{\mathrm{TI}}_{2323}\right)}\,, \tag{3.27}$$

for either $^{\mathrm{TI}}$ or $^{\overline{\mathrm{TI}}}$. However, equations (3.23a)–(3.23c) lead to ellipsoidal forms, even though, therein, $\varepsilon = -0.2363 \neq -0.4445 = \delta$. To avoid this discrepancy, we state the following proposition with a qualifier.

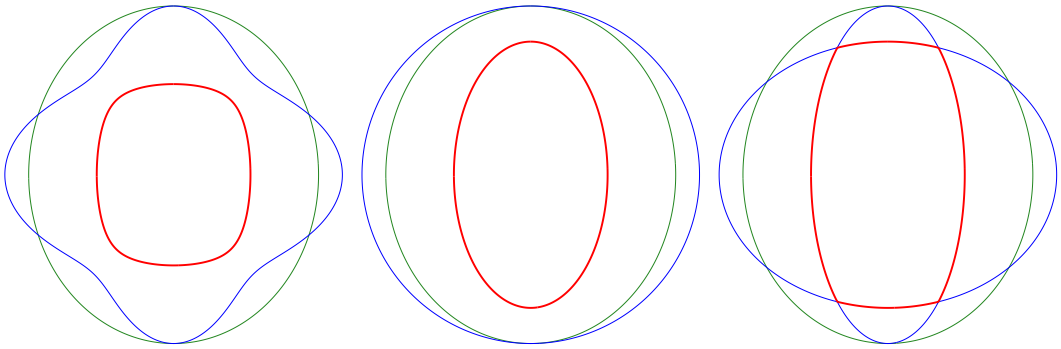
Proposition 2. *The detached qP slowness surface is ellipsoidal if and only if $\varepsilon = \delta$.*

Proof. Following expressions (3.27), $\varepsilon = \delta$ if and only if

$$\begin{aligned} &c^{\mathrm{TI}}_{2323} = -c^{\mathrm{TI}}_{1133} \quad \text{and} \quad c^{\mathrm{TI}}_{1111} = c^{\mathrm{TI}}_{2323} \qquad \text{or} \\ &c^{\mathrm{TI}}_{3333} = \frac{\left(c^{\mathrm{TI}}_{1133}\right)^2 + c^{\mathrm{TI}}_{1111}\,c^{\mathrm{TI}}_{2323} + 2\,c^{\mathrm{TI}}_{1133}\,c^{\mathrm{TI}}_{2323}}{c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{2323}}\,, \end{aligned}$$

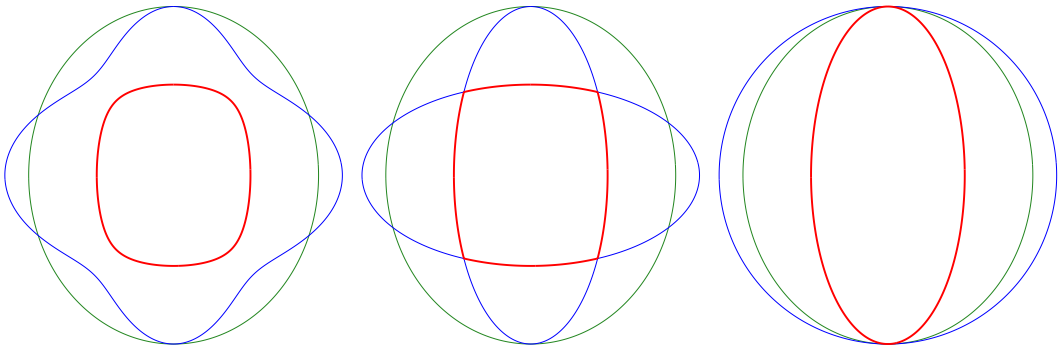
which are solutions (3.12b) and (3.12c), respectively. Solution (3.12a), $c^{\mathrm{TI}}_{2323} = -c^{\mathrm{TI}}_{1133}$, is — in general — the condition for nondetachment, and — for the Backus average — is the condition for elliptical roots. For the Backus average, solution (3.12b) is not allowed within the stability condition and solution (3.12c) results in an isotropic average, hence, circular roots. Thus, the ellipticity condition, $\varepsilon = \delta$, is valid for detached qP slowness surfaces only. □

To gain an insight into Proposition 2, we modify parameters for Green-River shale — illustrated in plot 3.3(a) — by applying expression (3.12c) to obtain an ellipsoidal qP slowness surface illustrated in plot 3.3(b). Applying subsequently expression (3.12a),



(a) (b) (c)

Figure 3.3: Slowness curves, $1/v_{qP}(\vartheta)$, $1/v_{qSV}(\vartheta)$, $1/v_{SH}(\vartheta)$, for modified values of elasticity parameters for Green-River shale. Plot (b) corresponds to expression (3.12c); plot (c) corresponds to expression (3.12c) followed by expression (3.12a).



(a) (b) (c)

Figure 3.4: Slowness curves, $1/v_{qP}(\vartheta)$, $1/v_{qSV}(\vartheta)$, $1/v_{SH}(\vartheta)$, for modified values of elasticity parameters for Green-River shale. Plot (b) corresponds to modifications by expression (3.12a); plot (c) is modified by expression (3.12a) followed by expression (3.12c).

we obtain three elliptical Christoffel roots, where $\varepsilon \neq \delta$, as expected in view of expressions (3.23a) and (3.23b). The innermost slowness surface is neither detached nor elliptical, as illustrated in plot 3.3(c).

Let us apply these expressions in the opposite order. Using expression (3.12a), we obtain the result illustrated in plot 3.4(b). Applying subsequently expression (3.12c), we obtain three elliptical slowness surfaces, illustrated in plot 3.4(c). The stability conditions are satisfied but δ is indeterminate. However, if we let $c^{\mathrm{TI}}_{2323} \approx -c^{\mathrm{TI}}_{1133}$, as opposed to $c^{\mathrm{TI}}_{2323} = -c^{\mathrm{TI}}_{1133}$, we obtain $\delta = \varepsilon$, as a consequence of detachment.

In contrast to these results, the Backus average can only satisfy the stability condition with expression (3.12a), which results in nondetachment.

3.5 Concluding remarks

The only restriction on the values of the elasticity parameters is the stability condition, which — mathematically — is tantamount to positive definiteness of the elasticity tensor. Within this condition, we examine properties of the Christoffel roots for nondetached slowness surfaces in transversely isotropic media.

Our analysis focuses on the propositions of Helbig (1983, p. 826), among which it is stated that the wavefront of qP waves is never an ellipsoid. The condition for this statement is a detached innermost slowness surface and, according to Musgrave (1970, p. 92), any inner detached sheet must be wholly convex. However, Bucataru and Slawinski (2009) extend the theorem of Musgrave by showing that the condition of detachment is unnecessary. Thus, within that context, our analysis yields a counterexample to Helbig that demonstrates elliptical nondetached slowness curves in Backus media.

To summarize, the qP slowness surface is detached if and only if $c^{\mathrm{TI}}_{2323} \neq -c^{\mathrm{TI}}_{1133}$. Under such a condition, each root corresponds to a distinct smooth wavefront. The qP slowness surface is nondetached if and only if $c^{\mathrm{TI}}_{2323} = -c^{\mathrm{TI}}_{1133}$. Under such conditions, the roots are elliptical but do not correspond to distinct wavefronts; also, the qP and qSV slowness surfaces are not smooth. For transversely isotropic media generated by the Backus average, the qP roots are elliptical only if the slowness surface is nondetached.

Chapter 4

On forward modelling of traveltimes in approximate media

Author note

This chapter is based on a collaborative research project with Dr. Len Bos and Dr. Michael A. Slawinski. For further details regarding co-authorship, we refer the reader to the statement provided on pages xi–xii.

4.1 Introductory remarks

One of the most important techniques for extracting information on the Earth's properties is traveltime inversion (e.g., Aki and Richards, 2002, Section 9.4). However, even a stack of horizontal layers is computationally complicated in both modelling and data processing, and, hence, it might be replaced by a single medium with similar traveltimes. In this chapter, we formulate an approach to reproduce traveltimes for near-vertical wave propagation, based on a Taylor series matching of the traveltimes in the stack of layers within a single medium.

For this purpose, we consider a number of parallel isotropic layers and construct two homogeneous media, the first single isotropic (iso) and the second single transversely isotropic (TI), for which the Fermat traveltimes are nearly equal to the Fermat traveltime within the layers, at least for propagation angles near the vertical.

Indeed, the isotropic medium will be constructed so that its Fermat traveltime agrees with the Fermat traveltime in the layered medium to the first order as a function of the propagation angle, while the transversely isotropic medium so that its Fermat traveltime will agree with the Fermat traveltime in the layered medium to the second order as a function of the propagation angle.

To achieve this formulation, noting that isotropy is a subclass of transverse isotropy, we require an expression of a TI tensor whose components are parameterized by properties of its constituent layers. Furthermore, we require two types of analytic expressions for traveltime. The first is the traveltime for rays of P and S waves in layered media as a function of their takeoff angle, ϕ. The second is the traveltime for rays of qP, qSV, and SH waves in the TI medium as a function of their takeoff angle, θ.

We require that waves propagate, in both media, from the same source and arrive at the same receiver, even though, except for vertical raypaths, $\phi \neq \theta$. Thus, we express the traveltimes of the waves propagating along the vertical axis as a second-order Taylor series approximation. To obtain the elasticity parameters of an approximate medium, we solve the system of equations that arise by equating coefficients of the Taylor approximations in both media; to the first order for an isotropic medium and to the second order for a transversely isotropic medium.

Upon completing the formulation, we obtain a homogeneous isotropic medium, which we refer to as the *approximate isotropic medium*, and a homogeneous transversely isotropic medium, which we refer to as the *approximate transversely isotropic medium*, for which the Fermat traveltimes have the approximation properties stated above.

We use a numerical example to demonstrate the validity of the traveltime approximations in the approximate media. In particular, we use a stack of ten isotropic layers and validate the traveltimes of rays through the layers for increasing receiver offsets. Then, we evaluate the traveltimes through the approximate media for the same source-receiver combinations. We find that the approximation leads to empirically adequate results for not only near-vertical rays, but even for rays whose takeoff angles approach 30° in the case of the TI medium.

Now, let us remark that our motivation is not only to approximate Fermat traveltimes, but also to develop a homogenized medium determined by its elasticity parameters. As such, our approach differs from other traveltime approximations, which rely on the

longstanding method of performing a Taylor approximation of a traveltime squared expression as a function of offset (e.g., Taner and Koehler, 1969). In recent years, this method has been used to approximate traveltimes by, for example, Sripanich and Fomel (2015) in TI and orthotropic media, Ravve and Koren (2017) as well as Abedi et al. (2021) in layered TI media, and Farra and Pšenčík (2020) in weakly anisotropic layered media of arbitrary symmetry classes.

By contrast, our alternative approach, which is based on Taylor approximations as well, yields closed-form expressions of elasticity parameters for the novel approximate medium. Using these expressions, we compare traveltimes directly with those in a TI medium that is the result of the Backus average of the same constituent layers, i.e., the Backus medium.

We begin this chapter by formulating traveltime expressions for both the layered and approximate media. Subsequently, we examine the Taylor expansions — developed about the vertical direction — of the traveltime expressions. We proceed to obtain the parameters of the approximate medium and make a comparison with traveltimes, for the same source-receiver combinations, in the Backus medium.

4.2 Traveltime expressions

4.2.1 Hooke's law

For both media, the stress-strain relationship is expressed by Hooke's law,

$$\sigma_{ij} = \sum_{k=1}^{3} \sum_{\ell=1}^{3} c_{ijk\ell}\, \varepsilon_{k\ell}\,, \qquad i,j = 1,2,3\,.$$

Herein, the stress tensor, σ_{ij}, is linearly related to the strain tensor,

$$\varepsilon_{k\ell} := \frac{1}{2}\left(\frac{\partial\, u_k}{\partial x_\ell} + \frac{\partial\, u_\ell}{\partial x_k}\right)\,, \qquad k,\ell = 1,2,3\,,$$

where u and x are displacement and position vectors, respectively. The elasticity tensor, $c_{ijk\ell}$, whose components can be expressed as a 6×6 matrix (e.g., Chapman, 2004, p. 92), is positive-definite. In general, there are three types of waves in anisotropic media but only two types in isotropic ones. The velocity expressions for waves in such media arise from the solubility conditions of the Christoffel

equation (e.g., Slawinski, 2015, Section 9.2.3). However, examining this equation is beyond the scope of this chapter and, as such, we only state, and use, the expressions for isotropic and TI media. To examine either medium, we use the x_1x_3-plane, where x_1 and x_3 are positive rightward and downward, respectively. For more details on the stress-strain relationship, Hooke's law, and the application toward ray theory, we refer the reader to Sections 2.4 and 2.5.

4.2.2 Layered medium

The layered medium is comprised of n parallel layers that are homogeneous and isotropic. For any given layer, its elastic properties are given by

$$c^{\text{iso}} = \begin{bmatrix} c^{\text{iso}}_{1111} & c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & 0 & 0 & 0 \\ c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & c^{\text{iso}}_{1111} & c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & 0 & 0 & 0 \\ c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & c^{\text{iso}}_{1111} - 2\,c^{\text{iso}}_{2323} & c^{\text{iso}}_{1111} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\,c^{\text{iso}}_{2323} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\,c^{\text{iso}}_{2323} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\,c^{\text{iso}}_{2323} \end{bmatrix} .$$

The thickness of the medium is

$$h = \sum_{i=1}^{n} (I_{i+1} - I_i) = I_{n+1} - I_1 \,, \tag{4.1}$$

where I_i is the interface depth of the ith layer. The receiver coordinates of a ray that connects a source at the origin to the bottom of the medium are $(x(\phi), h)$, where $x(\phi)$ is the horizontal position as a function of the takeoff angle ϕ. Thus, for transmitting P and S waves with the same takeoff angle,

$$x(\phi) = \sum_{i=1}^{n} (I_{i+1} - I_i) \tan\phi_i \quad \text{or} \quad x(\phi) = \sum_{i=1}^{n} (I_{i+1} - I_i) \tan\delta_i \,, \tag{4.2}$$

where ϕ_i and δ_i are, respectively, the propagation angles of P and S waves in the ith layer and the takeoff in the first layer is $\phi_1 = \delta_1$. The angle at which the ray is refracted is governed by the ray parameter

$$p = \frac{\sin\phi_1}{\alpha_1} = \frac{\sin\delta_1}{\beta_1} = \frac{\sin\phi_2}{\alpha_2} = \frac{\sin\delta_2}{\beta_2} = \cdots = \frac{\sin\phi_n}{\alpha_n} = \frac{\sin\delta_n}{\beta_n} \,.$$

Herein, α and β correspond to the P and S layer speeds. For the angle in the second layer,

$$\phi_2 = \arcsin\left(\frac{\alpha_2}{\alpha_1}\sin\phi_1\right) \quad \text{and} \quad \delta_2 = \arcsin\left(\frac{\beta_2}{\beta_1}\sin\phi_1\right).$$

For the third layer,

$$\phi_3 = \arcsin\left(\frac{\alpha_3}{\alpha_2}\sin\phi_2\right) = \arcsin\left(\frac{\alpha_3}{\alpha_2}\sin\left(\arcsin\left(\frac{\alpha_2}{\alpha_1}\sin\phi_1\right)\right)\right) = \arcsin\left(\frac{\alpha_3}{\alpha_1}\sin\phi_1\right),$$
$$\delta_3 = \arcsin\left(\frac{\beta_3}{\beta_2}\sin\delta_2\right) = \arcsin\left(\frac{\beta_3}{\beta_2}\sin\left(\arcsin\left(\frac{\beta_2}{\beta_1}\sin\phi_1\right)\right)\right) = \arcsin\left(\frac{\beta_3}{\beta_1}\sin\phi_1\right).$$

Thus, for the ith layer, where $\phi = \phi_1$, we have

$$\phi_i = \arcsin\left(\frac{\alpha_i}{\alpha_1}\sin\phi\right) \quad \text{and} \quad \delta_i = \arcsin\left(\frac{\beta_i}{\beta_1}\sin\phi\right).$$

To avoid duplication, we substitute ν for both α and β as a notational convenience to denote a generic wavespeed in a layer. As such, equation (4.2) becomes

$$\begin{aligned} x(\phi) &= \sum_{i=1}^{n} (I_{i+1} - I_i) \tan\left(\arcsin\left(\frac{\nu_i}{\nu_1}\sin\phi\right)\right) \\ &= \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\dfrac{\nu_i}{\nu_1}\sin\phi}{\sqrt{1 - \left(\dfrac{\nu_i}{\nu_1}\sin\phi\right)^2}} \\ &= \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\nu_i\,\sin\phi}{\sqrt{{\nu_1}^2 - {\nu_i}^2\,\sin^2\phi}}, \end{aligned} \tag{4.3}$$

where ν_i is the speed of a wave in the ith layer, to be replaced, respectively, by α_i or β_i for a P or S wave. The values for the wavespeeds in the ith layer are given by

$$\alpha_i = \sqrt{\frac{c^{\text{iso}}_{1111,i}}{\rho_i}} \qquad \text{and} \qquad \beta_i = \sqrt{\frac{c^{\text{iso}}_{2323,i}}{\rho_i}}, \qquad i = 1\,,\,\ldots\,,\,n\,, \tag{4.4a,b}$$

where ρ_i is the mass density of the ith layer. The corresponding traveltime for the ray of either the P or S wave is

$$t(\phi) = \sum_{i=1}^{n} \frac{1}{\nu_i} \sqrt{(I_{i+1} - I_i)^2 + \left(\frac{(I_{i+1} - I_i)\, \nu_i \, \sin\phi}{\sqrt{{\nu_1}^2 - {\nu_i}^2 \, \sin^2\phi}} \right)^2} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\, \nu_1}{\nu_i \sqrt{{\nu_1}^2 - {\nu_i}^2 \, \sin^2\phi}} \,. \tag{4.5}$$

4.2.3 Approximate medium

Noting again that isotropy is a subclass of transverse isotropy, we assume the symmetry class of the approximate medium to be TI. Although it is possible to consider other symmetry classes, the well-known and frequently employed Backus average of a layered medium (Backus, 1962) also results in a composite TI medium, chosen, however, to provide a more general physical approximation than just for traveltimes. In Section 4.6.3, we provide numerical comparisons between our approximate media and the Backus average. Hence, the elasticity tensor of the approximate medium is

$$c^{\mathrm{TI}} = \begin{bmatrix} c^{\mathrm{TI}}_{1111} & c^{\mathrm{TI}}_{1122} & c^{\mathrm{TI}}_{1133} & 0 & 0 & 0 \\ c^{\mathrm{TI}}_{1122} & c^{\mathrm{TI}}_{1111} & c^{\mathrm{TI}}_{1133} & 0 & 0 & 0 \\ c^{\mathrm{TI}}_{1133} & c^{\mathrm{TI}}_{1133} & c^{\mathrm{TI}}_{3333} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\, c^{\mathrm{TI}}_{2323} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\, c^{\mathrm{TI}}_{2323} & 0 \\ 0 & 0 & 0 & 0 & 0 & c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{1122} \end{bmatrix} . \tag{4.6}$$

In an anisotropic medium, the relationship between the ray and wavefront angles is (e.g., Slawinski, 2015, equation (8.4.12))

$$\tan\theta = \frac{\tan\vartheta + \dfrac{1}{v(\vartheta)} \dfrac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}}{1 - \dfrac{\tan\vartheta}{v(\vartheta)} \dfrac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}} \,, \tag{4.7}$$

where $v(\vartheta)$ stands for any of the three wavefront velocities, $v_{qP}(\vartheta)\,, v_{qSV}(\vartheta)\,, v_{SH}(\vartheta)\,,$ which are given in Section 4.4.4. In a homogeneous TI medium, the ray angle for a

raypath from the origin to coordinates $(x(\phi), h)$ is

$$\theta(\phi) = \arctan\left(\frac{x(\phi)}{h}\right), \tag{4.8}$$

where $x(\phi)$ is given by equation (4.3). Since the medium is homogeneous, rays are straight lines. Using trigonometry, the distance travelled for a takeoff angle $\theta(\phi)$ is

$$L(\phi) = \sqrt{x(\phi)^2 + h^2}\,. \tag{4.9}$$

The ray velocity of waves propagating through the anisotropic medium is given in terms of the wavefront velocity (e.g., Slawinski, 2015, equation (8.4.9)), where

$$V(\vartheta) = \sqrt{v(\vartheta)^2 + \left(\frac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}\,. \tag{4.10}$$

To evaluate equation (4.10), we must supply a value for θ in relation (4.7) and solve for its corresponding ϑ. However, relation (4.7) can only be solved analytically if and only if v is quadratic in the components of a vector that specifies the orientation of the wavefront, which corresponds to an elliptical wavefront (Slawinski, 2015, p. 355). Otherwise, we solve relation (4.7) numerically. Since the final horizontal position of a ray in the layered medium is dependent on the medium itself, we cannot obtain a particular value for θ to use in relation (4.7). Nevertheless, in general, the traveltime along a ray in the approximate medium, for the same source-receiver combination as the layered medium, whose receiver coordinates are $(x(\phi), h)$, is

$$T(\theta(\phi)) = \frac{L(\phi)}{V(\vartheta(\theta(\phi)))}\,. \tag{4.11}$$

The receiver coordinates may be expressed as a function of the takeoff angle in the layered medium, ϕ. Since the formulation uses the same source-receiver combinations, we express the ray angle in the approximate medium as a function of ϕ. Since the wavefront angle, ϑ, is a function of θ, which, in turn, is a function of ϕ, with an abuse of notation, we write

$$T(\phi) := T(\theta(\phi))\,, \quad L := L(\phi)\,, \quad V := V(\phi) := V(\vartheta(\theta(\phi)))\,. \tag{4.12}$$

Using notation (4.12), we write equation (4.11) as $T(\phi) = L/V$.

4.2.4 Vertical rays

A trivial solution of equation (4.7) is $\theta = \vartheta = 0$, which corresponds to vertical rays. To demonstrate this, we substitute $\theta = \vartheta = 0$ in relation (4.7) to obtain

$$\tan(0) = \frac{\tan(0) + \dfrac{1}{v(0)} \left.\dfrac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0}}{1 - \dfrac{\tan(0)}{v(0)} \left.\dfrac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0}} \,. \tag{4.13}$$

Since $v(0) > 0$, a solution can exists only if

$$\left.\frac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0} = 0 \,. \tag{4.14}$$

Thus, for vertical rays, equation (4.10) becomes

$$V(0) = v(0) \,; \tag{4.15}$$

in other words, the ray and wavefront velocities are equal to one another. Also, using equation (4.8),

$$0 = \arctan\left(\frac{x(\phi)}{h}\right) \implies x(\phi) = 0 \,, \quad \text{for} \quad \theta = \vartheta = 0 \,.$$

Recalling equation (4.3), $x(\phi) = 0 \implies \phi = 0$. Hence, for vertical rays,

$$\phi = 0 \iff \theta = \vartheta = 0 \,. \tag{4.16}$$

In other words, the ray and wavefront angles are equal to one another. Using property (4.16), we may evaluate the traveltimes along vertical rays in the layered and approximate media, $t(\phi)$ and $T(\phi)$, respectively. Setting them to be equal to one another, so that the traveltimes along vertical rays are the same in both media, imposes certain conditions on the elasticity parameters of the approximate medium, but does not completely determine them. To determine them completely, we need to consider their Taylor series up to the second order,

$$f(\zeta) \approx \sum_{m=0}^{2} \frac{f^{(m)}(\zeta)\big|_{\zeta=\zeta_0}}{m!} \left(\zeta - \zeta_0\right)^m \,,$$

where ζ is the argument of the function, ζ_0 is the base point of the expansion, and m is the order of the derivative.

Thus, for the layered medium,

$$t(\phi) = t(0) + \left.\frac{\mathrm{d}\, t(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \phi + \frac{1}{2} \left.\frac{\mathrm{d}^2 t(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} \phi + \mathcal{O}\left(\phi^3\right) , \tag{4.17}$$

and, for the approximate medium,

$$T(\phi) = T(0) + \left.\frac{\mathrm{d}\, T(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \phi + \frac{1}{2} \left.\frac{\mathrm{d}^2 T(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} \phi^2 + \mathcal{O}\left(\phi^3\right) , \tag{4.18}$$

where $\mathcal{O}$ is the big O Landau symbol. These series expansions result in an approximation of traveltime in the neighbourhood of $\phi = 0$.

4.3 Zeroth-order approximations

To complete the traveltime approximation, we must equate the zeroth, first, and second-order coefficients of approximations (4.17) and (4.18). Let us consider the ray of a wave traveling downward along the vertical symmetry axis. For the layered medium, we substitute $\phi = 0$ in equations (4.3) and (4.5) to obtain

$$x(0) = 0\,, \qquad t(0) = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)}{\nu_i}\,. \tag{4.19a,b}$$

Substituting equations (4.4a) and (4.4b) for ν, we obtain two forms of equation (4.19a)

$$x_P(0) = 0\,, \qquad x_S(0) = 0\,, \tag{4.20a,b}$$

whereas, for equation (4.19b),

$$t_P(0) = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)}{\sqrt{\dfrac{c^{\mathrm{iso}}_{1111,i}}{\rho_i}}}\,, \qquad t_S(0) = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)}{\sqrt{\dfrac{c^{\mathrm{iso}}_{2323,i}}{\rho_i}}}\,.$$

For the approximate medium, the vertical direction corresponds to the rotation symmetry axis. As a result of equation (4.14), the magnitude of its ray velocity is equal to the wavefront velocity, which is shown in equation (4.15). Thus, evaluating equations (4.39) for $\theta = \vartheta = 0$, we obtain

$$V_{qP}(0) = \sqrt{\frac{c^{\mathrm{TI}}_{3333}}{\rho^{\mathrm{TI}}}}\,, \quad V_{qSV}(0) = \sqrt{\frac{c^{\mathrm{TI}}_{2323}}{\rho^{\mathrm{TI}}}}\,, \quad V_{SH}(0) = \sqrt{\frac{c^{\mathrm{TI}}_{2323}}{\rho^{\mathrm{TI}}}}\,. \tag{4.22a,b,c}$$

Recalling equation (4.9), we calculate the distance travelled in the approximate medium for vertical rays. Using equations (4.1) and (4.19a),

$$L(0) = \sqrt{0 + h^2} = h = \sum_{i=1}^{n} I_{i+1} - I_i = I_{n+1} - I_1\,. \tag{4.23}$$

The corresponding traveltimes are

$$T_{qP}(0) = \frac{L(0)}{V_{qP}(0)}\,, \quad T_{qSV}(0) = \frac{L(0)}{V_{qSV}(0)}\,, \quad T_{SH}(0) = \frac{L(0)}{V_{SH}(0)}\,. \tag{4.24a,b,c}$$

Substituting expressions (4.22) in equations (4.24), and recognizing that the latter $T_{qSV}(0) = T_{SH}(0)$, we obtain

$$c^{\mathrm{TI}}_{3333} = \left(\frac{L(0)}{T_{qP}(0)}\right)^2 \rho^{\mathrm{TI}}\,, \quad c^{\mathrm{TI}}_{2323} = \left(\frac{L(0)}{T_{qSV}(0)}\right)^2 \rho^{\mathrm{TI}}\,. \tag{4.25a,b}$$

Thus, for $t_P(0) = T_{qP}(0)$ and $t_S(0) = T_{qSV}(0)$, we substitute equations (4.20a,b) in equations (4.25a,b), respectively, to obtain

$$c^{\mathrm{TI}}_{3333} = \left(\frac{\sum_{i=1}^{n} (I_{i+1} - I_i)}{\sum_{i=1}^{n} \dfrac{I_{i+1} - I_i}{\sqrt{c^{\mathrm{iso}}_{1111,i}/\rho_i}}}\right)^2 \rho^{\mathrm{TI}} =: H^2 \left(\sqrt{\frac{c^{\mathrm{iso}}_{1111,i}}{\rho_i}}\right) \rho^{\mathrm{TI}} \tag{4.26a}$$

and

$$c_{2323}^{\mathrm{TI}} = \left(\frac{\sum_{i=1}^{n} (I_{i+1} - I_i)}{\sum_{i=1}^{n} \frac{I_{i+1} - I_i}{\sqrt{c_{2323,i}^{\mathrm{iso}}/\rho_i}}} \right)^2 \rho^{\mathrm{TI}} =: H^2 \left(\sqrt{\frac{c_{2323,i}^{\mathrm{iso}}}{\rho_i}} \right) \rho^{\mathrm{TI}} . \tag{4.26b}$$

Here,

$$H(a_i) := \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} \frac{w_i}{a_i}} , \quad w_i := I_{i+1} - I_i$$

is the weighted harmonic mean of the a_i with weights w_i , where $I_{i+1} - I_i$, the layer thicknesses. In other words, equations (4.26) are the products of the squared harmonic mean of their corresponding wavespeeds, weighted by layer thickness, and the mass density of the transversely isotropic medium, which is given by the arithmetic mean of the mass densities of the constituent layers, weighted by layer thickness,

$$\rho^{\mathrm{TI}} = \frac{\sum_{i=1}^{n} (I_{i+1} - I_i)\, \rho_i}{\sum_{i=1}^{n} (I_{i+1} - I_i)} .$$

The results of the zeroth-order approximation are tantamount to the elasticity parameters of pure-mode P and S waves, which propagate along the rotation symmetry axis of the approximate medium. Hence, in agreement with expressions (4.22b,c), $V_{qSV}(0) = V_{SH}(0)$.

Thus, elasticity parameters (4.26) comprise the approximate isotropic medium, whose elasticity tensor is

$$c^{\mathrm{ISO}} := \begin{bmatrix} c_{1111}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & 0 & 0 & 0 \\ c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & 0 & 0 & 0 \\ c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} - 2\, c_{2323}^{\mathrm{ISO}} & c_{1111}^{\mathrm{ISO}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\, c_{2323}^{\mathrm{ISO}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\, c_{2323}^{\mathrm{ISO}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\, c_{2323}^{\mathrm{ISO}} \end{bmatrix} , \tag{4.27}$$

where $c_{1111}^{\mathrm{ISO}} := c_{3333}^{\mathrm{TI}}$ and $c_{2323}^{\mathrm{ISO}} := c_{2323}^{\mathrm{TI}}$.

Even though derived for vertical rays in TI media, while examining the approximate isotropic medium, we consider the wavefront velocity to be constant for all wavefront angles. Consequently, both the ray and wavefront angles and velocities are equivalent to one another. Thus, the magnitudes of the ray velocities of the P and S waves are tantamount to equations (4.22), where

$$V_P := \sqrt{\frac{c_{1111}^{\mathrm{ISO}}}{\rho^{\mathrm{TI}}}} = V_{qP}(0)\,, \quad V_S := \sqrt{\frac{c_{2323}^{\mathrm{ISO}}}{\rho^{\mathrm{TI}}}} = V_{qSV}(0) = V_{SH}(0)\,. \tag{4.28a,b}$$

Using ray velocities (4.28), the traveltimes, for any receiver offset, $x(\phi)$, are

$$T_P(\phi) := \frac{L(x(\phi))}{V_P}\,, \quad T_S(\phi) := \frac{L(x(\phi))}{V_S}\,. \tag{4.29a,b}$$

Note that by (4.26a) and (4.26b), it follows easily that the velocities in the approximate ISO medium are given by

$$V_\nu = H(\nu_i) \tag{4.30}$$

where ν_i is the wave speed in the ith layer and H is the weighted harmonic mean.

The isotropic elasticity parameters determined by matching the traveltimes for a vertical ray (propagation angle 0; zeroth order approximation) results in an (isotropic) medium for which the traveltimes also agree with the layered medium traveltimes to *first* order. Relaxing the symmetry class to TI requires solving for the remaining TI elasticity parameters, c_{1111}^{TI}, c_{1122}^{TI}, c_{1133}^{TI}, which are determined by matching the second-order coefficients of the Taylor approximations (4.17) and (4.18).

4.4 First- and second-order derivatives

4.4.1 Layered medium

Let us perform the first and second derivatives of equations (4.3) and (4.5) with respect to takeoff angle ϕ. They are

$$\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\nu_i\,\sin\phi}{\sqrt{{\nu_1}^2 - {\nu_i}^2\,\sin^2\phi}}\right) = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,{\nu_1}^2\,\nu_i\,\cos\phi}{\left({\nu_1}^2 - {\nu_i}^2\,\sin^2\phi\right)^{3/2}}\,, \tag{4.31}$$

$$\frac{\mathrm{d}\,t(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\nu_1}{\nu_i\sqrt{{\nu_1}^2-{\nu_i}^2\,\sin^2\phi}}\right) = \sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\nu_1\,\nu_i\,\cos\phi\,\sin\phi}{\left({\nu_1}^2-{\nu_i}^2\,\sin^2\phi\right)^{3/2}}\,, \quad (4.32)$$

$$\frac{\mathrm{d}^2x(\phi)}{\mathrm{d}\phi^2} = \sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,{\nu_1}^2\,\nu_i\,\sin\phi\,((2+\cos(2\,\phi))\,{\nu_i}^2-{\nu_1}^2)}{\left({\nu_1}^2-{\nu_i}^2\,\sin^2\phi\right)^{5/2}}\,, \quad (4.33)$$

$$\frac{\mathrm{d}^2t(\phi)}{\mathrm{d}\phi^2} = \sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\nu_1\,\nu_i\left(2\,\cos(2\,\phi)\,{\nu_1}^2+(3+\cos(2\,\phi))\,{\nu_i}^2\,\sin^2\phi\right)}{2\left({\nu_1}^2-{\nu_i}^2\,\sin^2\phi\right)^{5/2}}\,. \quad (4.34)$$

4.4.2 Ray angle

Recalling expression (4.8), the ray angle for a raypath in a homogeneous TI medium from the origin to coordinates $(x(\phi), h)$ is

$$\theta(\phi) = \arctan\left(\frac{x(\phi)}{h}\right)\,.$$

The first derivative of expression (4.8) is

$$\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\arctan\left(\frac{x(\phi)}{h}\right)\right) = \frac{h}{x(\phi)^2+h^2}\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\,. \quad (4.35)$$

The second derivative of expression (4.8) is

$$\begin{aligned}\frac{\mathrm{d}^2\theta(\phi)}{\mathrm{d}\phi^2} &= \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{h}{x(\phi)^2+h^2}\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)\\ &= \frac{h}{\left(x(\phi)^2+h^2\right)^2}\left(\left(x(\phi)^2+h^2\right)\frac{\mathrm{d}^2x(\phi)}{\mathrm{d}\phi^2}-2\,x(\phi)\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right)\,. \quad (4.36)\end{aligned}$$

4.4.3 Distance travelled

Recalling expression (4.9), the distance travelled, for a takeoff angle (4.8) to coordinates $(x(\phi), h)$, is

$$L(\phi) = \sqrt{x(\phi)^2+h^2}\,.$$

The first derivative of expression (4.9) is

$$\frac{\mathrm{d}\,L(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\sqrt{x(\phi)^2+h^2}\right) = \frac{x(\phi)}{\sqrt{x(\phi)^2+h^2}}\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\,. \quad (4.37)$$

The second derivative of expression (4.9) is

$$\begin{aligned}\frac{\mathrm{d}^2 L(\phi)}{\mathrm{d}\phi^2} &= \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{x(\phi)}{\sqrt{x(\phi)^2+h^2}}\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)\\ &= \frac{1}{\left(x(\phi)^2+h^2\right)^{3/2}}\left(x(\phi)\left(x(\phi)^2+h^2\right)\frac{\mathrm{d}^2 x(\phi)}{\mathrm{d}\phi^2}+h^2\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right). \end{aligned} \tag{4.38}$$

4.4.4 Ray velocity

Recalling expression (4.10), the magnitude of the ray velocity is

$$V(\vartheta) = \sqrt{v(\vartheta)^2+\left(\frac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2},$$

where $v(\vartheta)$ is tantamount to any of $v_{qP}(\vartheta), v_{qSV}(\vartheta), v_{SH}(\vartheta)$, which are the three wavefront velocities in a TI medium, and ϑ is the wavefront angle. The existence and properties of the three waves that propagate in a Hookean solid are a consequence of the Christoffel equation (e.g., Slawinski, 2015, Chapter 9). In Chapter 3, we discuss the roots of the Christoffel equation within the context of a transversely isotropic medium that is the result of Backus average. However, for this chapter, we consider the roots within a generic transversely isotropic medium, whose expressions are

$$v_{qP}(\vartheta) = \sqrt{\frac{\left(c^{\mathrm{TI}}_{3333}-c^{\mathrm{TI}}_{1111}\right){n_3}^2+c^{\mathrm{TI}}_{1111}+c^{\mathrm{TI}}_{2323}+\sqrt{\Delta}}{2\,\rho^{\mathrm{TI}}}}, \tag{4.39a}$$

$$v_{qSV}(\vartheta) = \sqrt{\frac{\left(c^{\mathrm{TI}}_{3333}-c^{\mathrm{TI}}_{1111}\right){n_3}^2+c^{\mathrm{TI}}_{1111}+c^{\mathrm{TI}}_{2323}-\sqrt{\Delta}}{2\,\rho^{\mathrm{TI}}}}, \tag{4.39b}$$

$$v_{SH}(\vartheta) = \sqrt{\frac{\frac{1}{2}\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{1122}\right)\left(1-{n_3}^2\right)+c^{\mathrm{TI}}_{2323}\,{n_3}^2}{\rho^{\mathrm{TI}}}}. \tag{4.39c}$$

where

$$\begin{aligned}\Delta = &\left(\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{3333}\right){n_3}^2-c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{2323}\right)^2-4\Big(c^{\mathrm{TI}}_{3333}\,c^{\mathrm{TI}}_{2323}\,{n_3}^4-\\ &\left(2\,c^{\mathrm{TI}}_{1133}\,c^{\mathrm{TI}}_{2323}-c^{\mathrm{TI}}_{1111}\,c^{\mathrm{TI}}_{3333}+{c^{\mathrm{TI}}_{1133}}^2\right){n_3}^2\left(1-{n_3}^2\right)+c^{\mathrm{TI}}_{1111}\,c^{\mathrm{TI}}_{2323}\left(1-{n_3}^2\right)^2\Big),\\ n_3 = &\cos\vartheta,\end{aligned}$$

and $^{\mathrm{TI}}$ denotes a quantity of a transversely isotropic medium. As the displacement vectors of these waves are neither parallel nor perpendicular to the direction of propagation, the notation q refers to "quasi" and represents quasi-P and quasi-S waves. As a consequence of the addition and subtraction of the radical term in the numerators of equations (4.39a) and (4.39b), respectively, the qP wave is, in general, faster than the qSV wave, except in singular cases (Bucataru and Slawinski, 2009; Bos et al., 2019a). The generality of the previous sentence is linked to the convexity of the innermost sheet of the wavefront-slowness surface, which is described by Musgrave (1970, p. 91–92) and Slawinski (2015, Theorem 9.3.1). Notably, P stands for *primary* and S for *secondary*. For a discussion on wavefront velocities and their displacement directions in TI media, we refer the reader to Slawinski (2015, Chapter 9.2.3). As they are required in the following sections, the first derivatives of equations (4.39) are

$$\frac{\mathrm{d}\,v_{qP}}{\mathrm{d}\vartheta} = \frac{2\,\sin(2\,\vartheta)\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{3333}\right)\left(\sqrt{2}\,c^{\mathrm{TI}}_{1111}-2\,\sqrt{2}\,c^{\mathrm{TI}}_{2323}+\sqrt{2}\,c^{\mathrm{TI}}_{3333}+\sqrt{\xi_1}\right)-\sqrt{2}\,\xi_2}{16\,\sqrt{\rho^{\mathrm{TI}}\,\Delta}\,\sqrt{(c^{\mathrm{TI}}_{3333}-c^{\mathrm{TI}}_{1111})\cos^2\vartheta+c^{\mathrm{TI}}_{1111}+c^{\mathrm{TI}}_{2323}+\sqrt{\Delta}}}\,, \tag{4.40}$$

$$\frac{\mathrm{d}\,v_{qSV}}{\mathrm{d}\vartheta} = \frac{-2\,\sin(2\,\vartheta)\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{3333}\right)\left(\sqrt{2}\,c^{\mathrm{TI}}_{1111}-2\,\sqrt{2}\,c^{\mathrm{TI}}_{2323}+\sqrt{2}\,c^{\mathrm{TI}}_{3333}-\sqrt{\xi_1}\right)+\sqrt{2}\,\xi_2}{16\,\sqrt{\rho^{\mathrm{TI}}\,\Delta}\,\sqrt{(c^{\mathrm{TI}}_{3333}-c^{\mathrm{TI}}_{1111})\cos^2\vartheta+c^{\mathrm{TI}}_{1111}+c^{\mathrm{TI}}_{2323}-\sqrt{\Delta}}}\,, \tag{4.41}$$

$$\frac{\mathrm{d}\,v_{SH}}{\mathrm{d}\vartheta} = \frac{\sin\vartheta\,\cos\vartheta\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{1122}-2\,c^{\mathrm{TI}}_{2323}\right)}{\sqrt{\rho^{\mathrm{TI}}\left(\cos(2\,\vartheta)\left(-c^{\mathrm{TI}}_{1111}+c^{\mathrm{TI}}_{1122}+2\,c^{\mathrm{TI}}_{2323}\right)+c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{1122}+2\,c^{\mathrm{TI}}_{2323}\right)}}\,, \tag{4.42}$$

where

$$\begin{aligned}
\xi_1 := 8\,\Delta = {}& 3\,{c^{\mathrm{TI}}_{1111}}^2+\cos(4\,\vartheta)\left(c^{\mathrm{TI}}_{1111}+2\,c^{\mathrm{TI}}_{1133}+c^{\mathrm{TI}}_{3333}\right)\left(c^{\mathrm{TI}}_{1111}-2\,c^{\mathrm{TI}}_{1133}-4\,c^{\mathrm{TI}}_{2323}+c^{\mathrm{TI}}_{3333}\right)\\
&-4\left(c^{\mathrm{TI}}_{1111}-c^{\mathrm{TI}}_{3333}\right)\cos(2\,\vartheta)\left(c^{\mathrm{TI}}_{1111}-2\,c^{\mathrm{TI}}_{2323}+c^{\mathrm{TI}}_{3333}\right)-2\,c^{\mathrm{TI}}_{1111}\left(2\,c^{\mathrm{TI}}_{2323}+c^{\mathrm{TI}}_{3333}\right)\\
&+4\,{c^{\mathrm{TI}}_{1133}}^2+8\,c^{\mathrm{TI}}_{2323}\left(c^{\mathrm{TI}}_{1133}+c^{\mathrm{TI}}_{2323}\right)-4\,c^{\mathrm{TI}}_{2323}\,c^{\mathrm{TI}}_{3333}+3\,{c^{\mathrm{TI}}_{3333}}^2\,,\\
\xi_2 := {}& \sin(4\,\vartheta)\left(c^{\mathrm{TI}}_{1111}+2\,c^{\mathrm{TI}}_{1133}+c^{\mathrm{TI}}_{3333}\right)\left(c^{\mathrm{TI}}_{1111}-2\,c^{\mathrm{TI}}_{1133}-4\,c^{\mathrm{TI}}_{2323}+c^{\mathrm{TI}}_{3333}\right)\,.
\end{aligned}$$

The second derivatives of equation (4.39a) and (4.39b) extend further than a page and, as such, we do not present them here. However, they may be derived using a symbolic mathematical software; their expressions for vertical rays, i.e., $\theta=\vartheta=0\,$,

are given in Section 4.5.2. As for the second derivative of equation (4.39c), it is

$$\frac{\mathrm{d}^2 v_{SH}}{\mathrm{d}\vartheta^2} = -\frac{\xi_3 \left(\xi_3 \cos^4\vartheta + \left(c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{1122}\right)\sin^2\vartheta + \left(c^{\mathrm{TI}}_{1122} - c^{\mathrm{TI}}_{1111}\right)\cos^2\vartheta\right)}{\sqrt{2\,\rho^{\mathrm{TI}}}\left(\cos^2\vartheta\left(-c^{\mathrm{TI}}_{1111} + c^{\mathrm{TI}}_{1122} + 2\,c^{\mathrm{TI}}_{2323}\right) + c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{1122}\right)^{3/2}}\,, \tag{4.43}$$

where

$$\xi_3 := c^{\mathrm{TI}}_{1111} - c^{\mathrm{TI}}_{1122} - 2\,c^{\mathrm{TI}}_{2323}\,.$$

First derivative

For the remainder of this section, we do not explicitly state the first and second derivatives of wavefront velocities (4.39). Instead, we retain the unspecified forms, $v(\vartheta)$ and $\mathrm{d}\,v(\vartheta)/\mathrm{d}\vartheta$.

Let us proceed to the first derivative of equation (4.10). We recognize that $\vartheta = \vartheta(\theta(\phi))$ and, as such,

$$\frac{\mathrm{d}\,V(\vartheta(\theta(\phi)))}{\mathrm{d}\phi} =: \frac{\mathrm{d}\,V(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\,\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\,\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\,, \tag{4.44}$$

where

$$\frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta} = \frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\sqrt{v(\vartheta)^2 + \left(\frac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}\right) = \frac{\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\left(v(\vartheta) + \dfrac{\mathrm{d}^2 v(\vartheta)}{\mathrm{d}\vartheta^2}\right)}{\sqrt{v(\vartheta)^2 + \left(\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}}\,. \tag{4.45}$$

Differentiating the ray and wavefront angle relationship (4.7), with respect to θ, where we use the chain rule on the right-hand side, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\left(\tan\theta\right) = \frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\frac{\tan\vartheta + \dfrac{1}{v(\vartheta)}\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}}{1 - \dfrac{\tan\vartheta}{v(\vartheta)}\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}}\right)$$

which results in

$$\sec^2\theta = \frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\left(\frac{v(\vartheta)\left(v(\vartheta) + \dfrac{\mathrm{d}^2 v(\vartheta)}{\mathrm{d}\vartheta^2}\right)}{\left(v(\vartheta)\,\cos\vartheta - \dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\,\sin\vartheta\right)^2}\right)\,.$$

Solving for $\mathrm{d}\,\vartheta(\theta)/\mathrm{d}\theta$,

$$\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta} = \frac{\sec^2\theta\left(\cos\vartheta\,v(\vartheta) - \sin\vartheta\,\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}{v(\vartheta)\left(v(\vartheta) + \dfrac{\mathrm{d}^2 v(\vartheta)}{\mathrm{d}\vartheta^2}\right)}\,. \tag{4.46}$$

Using equations (4.35), (4.45), and (4.46), equation (4.44) becomes

$$\frac{\mathrm{d}\,V(\vartheta(\theta(\phi)))}{\mathrm{d}\phi} =: \frac{\mathrm{d}\,V(\phi)}{\mathrm{d}\phi} = \frac{h\,(\sec^2\theta)\,\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\left((\cos\vartheta)\,v(\vartheta) - \sin\vartheta\left(\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)\right)^2\dfrac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}}{v(\vartheta)\left(x(\phi)^2 + h^2\right)\sqrt{v(\vartheta)^2 + \left(\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}}\,. \tag{4.47}$$

Second derivative

Recalling that $\vartheta = \vartheta(\theta(\phi))$, the second derivative of equation (4.10) is

$$\begin{aligned}
\frac{\mathrm{d}^2 V(\phi)}{\mathrm{d}\phi^2} &:= \frac{\mathrm{d}^2 V(\vartheta(\theta(\phi)))}{\mathrm{d}\phi^2} \\
&= \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\,\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\,\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\right) \\
&= \frac{\mathrm{d}^2 V(\vartheta)}{\mathrm{d}\vartheta^2}\left(\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\right)^2\left(\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\right)^2 + \frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\frac{\mathrm{d}^2\vartheta(\theta)}{\mathrm{d}\theta^2}\left(\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\right)^2 \\
&\qquad\qquad + \frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\,\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\,\frac{\mathrm{d}^2\theta(\phi)}{\mathrm{d}\phi^2} \\
&= \left(\frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\right)^2\left(\frac{\mathrm{d}^2 V(\vartheta)}{\mathrm{d}\vartheta^2}\left(\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\right)^2 + \frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\frac{\mathrm{d}^2\vartheta(\theta)}{\mathrm{d}\theta^2}\right) + \frac{\mathrm{d}\,V(\vartheta)}{\mathrm{d}\vartheta}\,\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\,\frac{\mathrm{d}^2\theta(\phi)}{\mathrm{d}\phi^2}\,.
\end{aligned} \tag{4.48}$$

To complete equation (4.48), we require $\mathrm{d}^2 V(\vartheta)/\mathrm{d}\vartheta^2$ and $\mathrm{d}^2\vartheta(\theta)/\mathrm{d}\theta^2$. For the former, using equation (4.45), we obtain

$$\frac{\mathrm{d}^2 V(\vartheta)}{\mathrm{d}\vartheta^2} = \frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\frac{\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\left(v(\vartheta) + \dfrac{\mathrm{d}^2 v(\vartheta)}{\mathrm{d}\vartheta^2}\right)}{\sqrt{v(\vartheta)^2 + \left(\dfrac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right)^2}}\right)\,.$$

Writing $v = v(\vartheta)$, $\vartheta = \vartheta(\theta)$, the above equation simplifies to

$$\frac{\mathrm{d}^2 V(\vartheta)}{\mathrm{d}\vartheta^2} = \frac{\left(\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\right)^4 - v\left(\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\right)^2 \dfrac{\mathrm{d}^2 v}{\mathrm{d}\vartheta^2} + v^2 \dfrac{\mathrm{d}^2 v}{\mathrm{d}\vartheta^2}\left(v + \dfrac{\mathrm{d}^2 v}{\mathrm{d}\vartheta^2}\right) + v^2 \dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\left(\dfrac{\mathrm{d}^3 v}{\mathrm{d}\vartheta^3}\right) + \left(\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\right)^3 \dfrac{\mathrm{d}^3 v}{\mathrm{d}\vartheta^3}}{\left(v^2 + \left(\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\right)^2\right)^{3/2}} \,. \tag{4.49}$$

For $\mathrm{d}^2\vartheta/\mathrm{d}\theta^2$, we make use of the second derivative of relationship (4.7). For the right-hand side, we use the chain and product rules, where

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}\theta^2}(\circ) &= \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\mathrm{d}\,\vartheta}{\mathrm{d}\theta} \cdot \frac{\mathrm{d}}{\mathrm{d}\vartheta}(\circ)\right) \\ &= \frac{\mathrm{d}^2\vartheta}{\mathrm{d}\theta^2} \cdot \frac{\mathrm{d}}{\mathrm{d}\vartheta}(\circ) + \frac{\mathrm{d}\vartheta}{\mathrm{d}\theta} \cdot \frac{\mathrm{d}}{\mathrm{d}\theta}\left(\frac{\mathrm{d}}{\mathrm{d}\vartheta}(\circ)\right) \\ &= \frac{\mathrm{d}^2\vartheta}{\mathrm{d}\theta^2} \cdot \frac{\mathrm{d}}{\mathrm{d}\vartheta}(\circ) + \left(\frac{\mathrm{d}\vartheta}{\mathrm{d}\theta}\right)^2 \cdot \frac{\mathrm{d}^2}{\mathrm{d}\vartheta^2}(\circ) \,. \end{aligned}$$

Substituting the right-hand side of relationship (4.7) for $(\circ)$ and solving for $\mathrm{d}^2\vartheta/\mathrm{d}\theta^2$, we obtain

$$\frac{\mathrm{d}^2\vartheta(\theta)}{\mathrm{d}\theta^2} = \frac{\dfrac{\mathrm{d}^2 \tan\theta}{\mathrm{d}\theta^2} - \left(\dfrac{\mathrm{d}\,\vartheta}{\mathrm{d}\theta}\right)^2 \cdot \dfrac{\mathrm{d}^2}{\mathrm{d}\vartheta^2}\left(\dfrac{\tan\vartheta + \dfrac{1}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}\right)}{\dfrac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\dfrac{\tan\vartheta + \dfrac{1}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}\right)} \,, \tag{4.50}$$

where $\mathrm{d}\,\vartheta/\mathrm{d}\theta$ is given by equation (4.46) and the remaining derivatives are obtained below. Recalling $\theta(\phi)$ from equation (4.8), we obtain

$$\frac{\mathrm{d}^2 \tan\theta}{\mathrm{d}\theta^2} = 2\, \sec^2\theta\, \tan\theta \,. \tag{4.51a}$$

Next,

$$\frac{\mathrm{d}}{\mathrm{d}\vartheta}\left(\frac{\tan\vartheta + \dfrac{1}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}{1 - \dfrac{\tan\vartheta}{v}\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}}\right) = \frac{v\left(v + \dfrac{\mathrm{d}^2 v}{\mathrm{d}\vartheta^2}\right)}{\left(v\,\cos\vartheta - \sin\vartheta\left(\dfrac{\mathrm{d}\, v}{\mathrm{d}\vartheta}\right)\right)^2} \tag{4.51b}$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}\vartheta^2}\left(\frac{\tan\vartheta+\dfrac{1}{v}\dfrac{\mathrm{d}\,v}{\mathrm{d}\vartheta}}{1-\dfrac{\tan\vartheta}{v}\dfrac{\mathrm{d}\,v}{\mathrm{d}\vartheta}}\right)=\frac{-\sin\vartheta\left(\dfrac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\dfrac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}+\xi_4+\xi_5+2\,v^3\,\sin\vartheta}{\left(v\,\cos\vartheta-\sin\vartheta\left(\dfrac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)\right)^3}\,, \tag{4.51c}$$

where

$$\xi_4 := v^2\left(\frac{\mathrm{d}^3v}{\mathrm{d}\vartheta^3}\,\cos\vartheta+4\,\sin\vartheta\,\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}+2\,\cos\vartheta\,\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right),$$
$$\xi_5 := v\left(2\,\sin\vartheta\,\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)^2-2\,\sin\vartheta\,\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2+\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(\cos\vartheta\,\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}-\frac{\mathrm{d}^3v}{\mathrm{d}\vartheta^3}\,\sin\vartheta\right)\right).$$

Using equations (4.46) and (4.51), equation (4.50) becomes

$$\frac{\mathrm{d}^2\vartheta(\theta)}{\mathrm{d}\theta^2}=\frac{\xi_6\,\sec^4\theta\left(v\cos\vartheta-\sin\vartheta\,\dfrac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2}{2\,v^3\left(v+\dfrac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)^3}\,, \tag{4.52}$$

where

$$\begin{aligned}\xi_6 :=\ & 2\,v^4\,\xi_7\\ & -2\,\sin^2\vartheta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^3\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)\\ & -2\,v^3\left(2\,\cos(2\,\vartheta)\,\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}-2\,\xi_7\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)+\cos^2\vartheta\left(\frac{\mathrm{d}^3v}{\mathrm{d}\vartheta^3}\right)\right)\\ & -2\,(\sin\vartheta)\,v\,\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(2\,\sin\vartheta\left(\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2-\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)^2\right)+\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(-2\,\cos\vartheta\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)+\sin\vartheta\left(\frac{\mathrm{d}^3v}{\mathrm{d}\vartheta^3}\right)\right)\right)\\ & +v^2\left(4\,\sin(2\,\vartheta)\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2+2\,\xi_7\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)^2+\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left((3-5\,\cos(2\,\vartheta))\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)+2\,\sin(2\,\vartheta)\left(\frac{\mathrm{d}^3v}{\mathrm{d}\vartheta^3}\right)\right)\right),\end{aligned}$$
$$\xi_7 := \sin(2\,\theta)-\sin(2\,\vartheta)\ .$$

Returning to equation (4.48), and using equations (4.35) for $\mathrm{d}\,\theta(\phi)/\mathrm{d}\phi$, (4.36) for $\mathrm{d}^2\theta(\phi)/\mathrm{d}\phi^2$, (4.45) for $\mathrm{d}\,V(\vartheta)/\mathrm{d}\vartheta$, (4.49) for $\mathrm{d}^2V(\vartheta)/\mathrm{d}\vartheta^2$, (4.46) for $\mathrm{d}\,\vartheta(\theta)/\mathrm{d}\theta$,

and (4.52) for $\mathrm{d}^2\vartheta(\theta)/\mathrm{d}\theta^2$, we obtain

$$\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2} = \frac{h\,\sec^4\theta\left((\cos\vartheta)\,v - \sin\vartheta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)\right)^2(\xi_8+\xi_9+\xi_{10}+\xi_{11}+\xi_{14})}{v^3\left(h^2+x(\phi)^2\right)^2\left(v^2+\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\right)^{3/2}\left(v+\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)}\,, \tag{4.53}$$

where

$$\xi_8 := v^3\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(\xi_{12}+\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2(h\,(\sin(2\,\vartheta)+\sin(2\,\theta))-\xi_{13})\right)+3\,h\left(\sin^2\vartheta\right)\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right),$$

$$\xi_9 := v^2\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^3\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\left(\xi_{12}+\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2(h\,(\sin(2\,\theta)-\sin(2\,\vartheta))-\xi_{13})\right)-h\left(\cos^2\vartheta\right)\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right),$$

$$\xi_{10} := v^4\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\left(\xi_{12}+\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2(h\,(\sin(2\,\theta)-2\,\sin(2\,\vartheta))-\xi_{13})\right)-2\,h\,(\cos(2\,\vartheta))\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right),$$

$$\xi_{11} := v^5\left(\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)\left(\xi_{12}+\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2(h\,(\sin(2\,\theta)-\sin(2\,\vartheta))-\xi_{13})\right)+h\left(\cos^2\vartheta\right)\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\right),$$

$$\xi_{12} := \cos^2\theta\left(x(\phi)^2+h^2\right)\left(\frac{\mathrm{d}^2x(\phi)}{\mathrm{d}\phi^2}\right),$$

$$\xi_{13} := 2\,x(\phi)\,\cos^2\theta\,,$$

$$\xi_{14} := -h\,\sin^2\vartheta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^6\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2+2\,h\,v\,\sin\vartheta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^4\left(\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right)^2\left(\sin\vartheta\left(\frac{\mathrm{d}^2v}{\mathrm{d}\vartheta^2}\right)+\cos\vartheta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)\right).$$

4.4.5 Traveltime

Recalling expression (4.11), the traveltime along a ray in the approximate medium, for the same source-receiver combination as the layered medium, whose receiver coordinates are $(x(\phi),h)$, is

$$T(\theta(\phi)) = \frac{L(\phi)}{V(\vartheta(\theta(\phi)))}\,.$$

As stated in Section 4.2.3, the receiver coordinates may be expressed as a function of the takeoff angle in the layered medium, ϕ. Since the formulation uses the same source-receiver combinations, we express the ray angle in the approximate medium as a function of ϕ. Since the wavefront angle ϑ is a function of θ, which, in turn, is a function of ϕ, we recall notation (4.12) to write

$$T(\phi) := T(\theta(\phi))\,,\quad L := L(\phi)\,,\quad V := V(\phi) = V(\vartheta(\theta(\phi)))\,.$$

Using notation (4.12), we write equation (4.11) as $T(\phi) = L/V$. From the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}\phi} = \frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\frac{\mathrm{d}}{\mathrm{d}\theta} = \frac{\mathrm{d}\,\theta(\phi)}{\mathrm{d}\phi}\,\frac{\mathrm{d}\,\vartheta(\theta)}{\mathrm{d}\theta}\,\frac{\mathrm{d}}{\mathrm{d}\vartheta}\,,$$

the first derivative of equation (4.11) is

$$\frac{\mathrm{d}\,T(\phi)}{\mathrm{d}\phi} = \frac{\mathrm{d}}{\mathrm{d}\phi}\left(\frac{L}{V}\right) = \frac{1}{V^2}\left(V\,\frac{\mathrm{d}\,L}{\mathrm{d}\,\phi} - L\,\frac{\mathrm{d}\,V}{\mathrm{d}\,\phi}\right)\,. \tag{4.54}$$

Using equation (4.9) for L , (4.37) for $\mathrm{d}\,L/\mathrm{d}\phi$, (4.10) for V , and (4.47) for $\mathrm{d}\,V/\mathrm{d}\phi$, equation (4.54) becomes

$$\frac{\mathrm{d}\,T(\phi)}{\mathrm{d}\phi} = \frac{\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\left(v\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\left(h\,\sin(2\,\vartheta)\,\sec^2\theta + x(\phi)\right) - h\,\sec^2\theta\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)\left(v^2\,\cos^2\vartheta + \sin^2\vartheta\,\left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\right) + v^3\,x(\phi)\right)}{v\,\sqrt{x(\phi)^2 + h^2}\left(v^2 + \left(\frac{\mathrm{d}\,v}{\mathrm{d}\vartheta}\right)^2\right)^{3/2}}\,. \tag{4.55}$$

The second derivative of equation (4.11) is

$$\begin{aligned}\frac{\mathrm{d}^2T(\phi)}{\mathrm{d}\phi^2} &= \frac{\mathrm{d}}{\mathrm{d}\phi}\left(V^{-1}\,\frac{\mathrm{d}\,L}{\mathrm{d}\,\phi} - L\,V^{-2}\,\frac{\mathrm{d}\,V}{\mathrm{d}\,\phi}\right)\\ &= \left(-\frac{1}{V^2}\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\,\frac{\mathrm{d}\,L}{\mathrm{d}\phi} + \frac{1}{V}\frac{\mathrm{d}^2L}{\mathrm{d}\phi^2}\right) - \left(\frac{\mathrm{d}\,L}{\mathrm{d}\phi}\,V^{-2}\,\frac{\mathrm{d}\,V}{\mathrm{d}\phi} - 2\,\frac{L}{V^3}\,\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\frac{\mathrm{d}\,V}{\mathrm{d}\phi} + L\,V^{-2}\,\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2}\right)\\ &= \frac{1}{V^3}\left(-2\,V\,\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\,\frac{\mathrm{d}\,L}{\mathrm{d}\phi} + V^2\frac{\mathrm{d}^2L}{\mathrm{d}\phi^2} + 2\,L\,\left(\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\right)^2 - L\,V\,\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2}\right)\,.\end{aligned} \tag{4.56}$$

Since the expression for equation (4.56) exceeds a full page when fully simplified, we suppress the details.

4.5 First- and second-order approximations

4.5.1 First order

Let us examine the first-order coefficients of Taylor approximations (4.17) and (4.18). For the layered medium, we evaluate the first derivatives of equations (4.3) and (4.5),

which are given in (4.31) and (4.32), for $\phi = 0$ to obtain

$$\left.\frac{\mathrm{d}\, x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\nu_i}{\nu_1}\,, \qquad \left.\frac{\mathrm{d}\, t(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = 0\,.$$

Recalling equations (4.4a) and (4.4b) for ν, we obtain

$$\left.\frac{\mathrm{d}\, x_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\alpha_i}{\alpha_1}\,, \qquad \left.\frac{\mathrm{d}\, x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\beta_i}{\beta_1}\,, \quad (4.57a,b)$$

and

$$\left.\frac{\mathrm{d}\, t_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = 0\,, \quad \left.\frac{\mathrm{d}\, t_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = 0\,. \qquad (4.58a,b)$$

For the approximate medium, we evaluate the first derivative of equation (4.11), which is equation (4.55), to obtain

$$\begin{aligned}
&\left.\frac{\mathrm{d}\, T(\phi)}{\mathrm{d}\phi}\right|_{\phi,\theta,\vartheta=0} \\
&\quad = \frac{\left.\frac{\mathrm{d}\, x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \left(v(0)\, x(0) \left(\left.\frac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0} \right)^2 - h \left(\left.\frac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0} \right) v(0)^2 + v(0)^3\, x(0) \right)}{v(0) \sqrt{x(0)^2 + h^2} \left(v(0)^2 + \left(\left.\frac{\mathrm{d}\, v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0} \right)^2 \right)^{3/2}}\,.
\end{aligned} \qquad (4.59)$$

Recalling equations (4.14) and (4.19a), the right-hand side of equation (4.59) reduces to 0. Consequently, for any of $T_{qP}(\phi)$, $T_{qSV}(\phi)$, $T_{SH}(\phi)$, we have

$$\left.\frac{\mathrm{d}\, T_{qP}(\phi)}{\mathrm{d}\phi}\right|_{\phi,\theta,\vartheta=0} = 0\,, \qquad \left.\frac{\mathrm{d}\, T_{qSV}(\phi)}{\mathrm{d}\phi}\right|_{\phi,\theta,\vartheta=0} = 0\,, \qquad \left.\frac{\mathrm{d}\, T_{SH}(\phi)}{\mathrm{d}\phi}\right|_{\phi,\theta,\vartheta=0} = 0\,. \qquad (4.60a,b,c)$$

Thus, in view of layer values (4.58) and approximate-medium values (4.60), the first-order coefficients of Taylor approximations (4.17) and (4.18) are in agreement for rays of P and S waves in the layered medium and rays of $qP\,, qSV\,, SH$ waves in the approximate medium. We remark that this is a consequence of the fact that traveltimes along vertical rays are a minimum in both media. Hence, to consider

expressions for the remaining TI elasticity parameters, c_{1111}^{TI} , c_{1122}^{TI} , c_{1133}^{TI} , we proceed to the second-order coefficients.

4.5.2 Second order

Let us examine the second-order coefficients of Taylor approximations (4.17) and (4.18). For the layered medium, we evaluate the second derivatives of equations (4.3) and (4.5), which are equations (4.33) and (4.34), for $\phi = 0$ to obtain

$$\left.\frac{\mathrm{d}^2 x(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = 0\,, \qquad \left.\frac{\mathrm{d}^2 t(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\nu_i}{{\nu_1}^2}\,.$$

Recalling equations (4.4a) and (4.4b) for ν , we obtain

$$\left.\frac{\mathrm{d}^2 x_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = 0\,, \quad \left.\frac{\mathrm{d}^2 x_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = 0\,,$$

and

$$\left.\frac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\alpha_i}{{\alpha_1}^2}\,, \quad \left.\frac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \sum_{i=1}^{n} \frac{(I_{i+1} - I_i)\,\beta_i}{{\beta_1}^2}\,. \qquad (4.61\text{a,b})$$

For the approximate medium, we refer to the second derivative of equation (4.11), which is equation (4.56) and we restate as

$$\frac{\mathrm{d}^2 T(\phi)}{\mathrm{d}\phi^2} = \frac{1}{V^3}\left(-2\,V\,\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\,\frac{\mathrm{d}\,L}{\mathrm{d}\phi} + V^2\frac{\mathrm{d}^2 L}{\mathrm{d}\phi^2} + 2\,L\,\left(\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\right)^2 - L\,V\,\frac{\mathrm{d}^2 V}{\mathrm{d}\phi^2}\right)\,. \qquad (4.62)$$

In order to evaluate equation (4.62) for vertical rays, let us address the zeroth-, first-, and second-order derivatives of L and V . First, we recall equations (4.15) and (4.23), respectively, and find that the zeroth-order derivatives have been evaluated. Second, we address the first derivatives. For the former, we recall equation (4.19a) and evaluate equation (4.37) to obtain

$$\left.\frac{\mathrm{d}\,L}{\mathrm{d}\phi}\right|_{\phi=0} = \frac{x(0)}{\sqrt{x(0)^2 + (I_{n+1} - I_1)^2}}\,\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} = 0\,. \qquad (4.63)$$

For the latter, we recall equation (4.14) and evaluate (4.47) to find that

$$\left.\frac{\mathrm{d}\,V}{\mathrm{d}\phi}\right|_{\phi,\theta,\vartheta=0} = \frac{(I_{n+1}-I_1)\left(\left.\frac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0}\right)v(0)^2\left(\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)}{v(0)\left(x(0)^2+(I_{n+1}-I_1)^2\right)\sqrt{v(0)^2+\left(\left.\frac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0}\right)^2}} = 0\,. \tag{4.64}$$

Finally, we address second derivatives. For the former, we recall equation (4.19a) and evaluate equation (4.38) to obtain

$$\begin{aligned}\left.\frac{\mathrm{d}^2L}{\mathrm{d}\phi^2}\right|_{\phi=0} &= \frac{(I_{n+1}-I_1)^2\left(\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2 + x(0)\left((I_{n+1}-I_1)^2+x(0)^2\right)\left.\frac{\mathrm{d}^2x(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}}{\left(x(0)^2+(I_{n+1}-I_1)^2\right)^{3/2}} \\ &= \frac{\left(\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)}\,. \end{aligned} \tag{4.65}$$

For the latter, we evaluate (4.53) to obtain

$$\begin{aligned}&\left.\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} \\ &\quad= \frac{(I_{n+1}-I_1)^2\,v(0)^4\left(\left.\frac{\mathrm{d}^2v(\vartheta)}{\mathrm{d}\vartheta^2}\right|_{\vartheta=0}\right)\left(\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{\left(x(0)^2+(I_{n+1}-I_1)^2\right)^2\left(v(0)^2+\left(\left.\frac{\mathrm{d}\,v(\vartheta)}{\mathrm{d}\vartheta}\right|_{\vartheta=0}\right)^2\right)^{3/2}\left(v(0)+\left.\frac{\mathrm{d}^2v(\vartheta)}{\mathrm{d}\vartheta^2}\right|_{\vartheta=0}\right)}\,.\end{aligned}$$

Recalling equations (4.14) and (4.19a),

$$\left.\frac{\mathrm{d}^2V}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} = \frac{v(0)\left(\left.\frac{\mathrm{d}^2v(\vartheta)}{\mathrm{d}\vartheta^2}\right|_{\vartheta=0}\right)\left(\left.\frac{\mathrm{d}\,x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)^2\left(v(0)+\left.\frac{\mathrm{d}^2v(\vartheta)}{\mathrm{d}\vartheta^2}\right|_{\vartheta=0}\right)}\,. \tag{4.66}$$

Thus, in view of equation (4.23) for $L(0)$, (4.15) for $V(0)$, (4.63) for $\mathrm{d}\,L/\mathrm{d}\phi|_{\phi=0}$, (4.64) for $\mathrm{d}\,V/\mathrm{d}\phi|_{\phi,\theta,\vartheta=0}$, (4.65) for $\mathrm{d}^2L/\mathrm{d}\phi^2|_{\phi=0}$, and (4.66) for $\mathrm{d}^2V/\mathrm{d}\phi^2|_{\phi,\theta,\vartheta=0}$,

equation (4.62) becomes

$$\left.\frac{\mathrm{d}^2 T(\phi)}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} = \frac{\left(\left.\frac{\mathrm{d}\, x(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)\left(v(0)+\left.\frac{\mathrm{d}^2 v(\vartheta)}{\mathrm{d}\vartheta^2}\right|_{\vartheta=0}\right)}\,. \tag{4.67}$$

For the rays of qP, qSV, SH waves, let us recall equations (4.39) for $v(0)$ and (4.57a,b) for $\mathrm{d}\, x(\phi)/\mathrm{d}\phi|_{\phi=0}$. For $\mathrm{d}^2 v(\vartheta)/\mathrm{d}\vartheta^2|_{\vartheta=0}$, we recall equation (4.43) for SH waves but rely on a symbolic software to perform, and evaluate, the second derivative of wavefront velocities of qP and qSV waves. Thus, for $T(\phi) = T_{qP}(\phi)$,

$$\left.\frac{\mathrm{d}^2 T_{qP}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=\theta=\vartheta=0} = \frac{\left(\sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\alpha_i}{\alpha_1}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\frac{c_{3333}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}}+\frac{\left(c_{1133}^{\mathrm{TI}}+c_{3333}^{\mathrm{TI}}\right)\left(c_{1133}^{\mathrm{TI}}+2\,c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}}\right)}{\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{3333}^{\mathrm{TI}}}\,(c_{3333}^{\mathrm{TI}}-c_{2323}^{\mathrm{TI}})}\right)}\,; \tag{4.68a}$$

for $T(\phi) = T_{qSV}(\phi)$,

$$\left.\frac{\mathrm{d}^2 T_{qSV}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=\theta=\vartheta=0} = \frac{\left(\sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\beta_i}{\beta_1}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\frac{c_{2323}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}}+\frac{c_{1111}^{\mathrm{TI}}\left(c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}}\right)+{c_{1133}^{\mathrm{TI}}}^2+2\,c_{1133}^{\mathrm{TI}}\,c_{2323}^{\mathrm{TI}}+c_{2323}^{\mathrm{TI}}\,c_{3333}^{\mathrm{TI}}}{\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{2323}^{\mathrm{TI}}}\,(c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}})}\right)}\,; \tag{4.68b}$$

and for $T(\phi) = T_{SH}(\phi)$,

$$\left.\frac{\mathrm{d}^2 T_{SH}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=\theta=\vartheta=0} = \frac{\left(\sum_{i=1}^{n}\frac{(I_{i+1}-I_i)\,\beta_i}{\beta_1}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\frac{c_{2323}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}}+\frac{c_{1111}^{\mathrm{TI}}-c_{1122}^{\mathrm{TI}}-2\,c_{2323}^{\mathrm{TI}}}{2\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{2323}^{\mathrm{TI}}}}\right)}\,. \tag{4.68c}$$

4.5.3 Elasticity parameters

To obtain expressions for the remaining elasticity parameters, c_{1111}^{TI}, c_{1122}^{TI}, c_{1133}^{TI}, we use equations (4.61) and (4.68) to solve a system of three equations, which is

$$\begin{cases} \left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \left.\dfrac{\mathrm{d}^2 T_{qP}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} \\ \left.\dfrac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \left.\dfrac{\mathrm{d}^2 T_{qSV}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} \\ \left.\dfrac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \left.\dfrac{\mathrm{d}^2 T_{SH}(\phi)}{\mathrm{d}\phi^2}\right|_{\phi,\theta,\vartheta=0} \end{cases} . \tag{4.69}$$

In solving system (4.69), we do not specify $\mathrm{d}^2 t_P(\phi)/\mathrm{d}\phi^2|_{\phi=0}$, $\mathrm{d}^2 t_S(\phi)/\mathrm{d}\phi^2|_{\phi=0}$ or $\mathrm{d}\,x_P(\phi)/\mathrm{d}\phi|_{\phi=0}$, $\mathrm{d}\,x_S(\phi)/\mathrm{d}\phi|_{\phi=0}$, as they are not functions of the approximate-medium elasticity parameters. Thus, using modified forms of equations (4.68a), (4.68b), (4.68c), system (4.69) becomes

$$\begin{cases} \left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \dfrac{\left(\left.\dfrac{\mathrm{d}\,x_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\dfrac{c_{3333}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}} + \dfrac{\left(c_{1133}^{\mathrm{TI}}+c_{3333}^{\mathrm{TI}}\right)\left(c_{1133}^{\mathrm{TI}}+2\,c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}}\right)}{\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{3333}^{\mathrm{TI}}}\left(c_{3333}^{\mathrm{TI}}-c_{2323}^{\mathrm{TI}}\right)}\right)} \\ \left.\dfrac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} \\ \qquad = \dfrac{\left(\left.\dfrac{\mathrm{d}\,x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\dfrac{c_{2323}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}} + \dfrac{c_{1111}^{\mathrm{TI}}\left(c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}}\right)+{c_{1133}^{\mathrm{TI}}}^2+2\,c_{1133}^{\mathrm{TI}}\,c_{2323}^{\mathrm{TI}}+c_{2323}^{\mathrm{TI}}\,c_{3333}^{\mathrm{TI}}}{\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{2323}^{\mathrm{TI}}}\left(c_{2323}^{\mathrm{TI}}-c_{3333}^{\mathrm{TI}}\right)}\right)} \\ \left.\dfrac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} = \dfrac{\left(\left.\dfrac{\mathrm{d}\,x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{(I_{n+1}-I_1)\left(\sqrt{\dfrac{c_{2323}^{\mathrm{TI}}}{\rho^{\mathrm{TI}}}} + \dfrac{c_{1111}^{\mathrm{TI}}-c_{1122}^{\mathrm{TI}}-2\,c_{2323}^{\mathrm{TI}}}{2\sqrt{\rho^{\mathrm{TI}}}\sqrt{c_{2323}^{\mathrm{TI}}}}\right)} \end{cases} .$$

Solving for c_{1111}^{TI}, c_{1122}^{TI}, c_{1133}^{TI}, we obtain

$$c_{1111}^{\mathrm{TI}} = -c_{2323}^{\mathrm{TI}} + \frac{1}{(I_{n+1}-I_1)}\left(\frac{\sqrt{c_{3333}^{\mathrm{TI}}\,\rho^{\mathrm{TI}}}\left(\left.\dfrac{\mathrm{d}\,x_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{\left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}} + \frac{\sqrt{c_{2323}^{\mathrm{TI}}\,\rho^{\mathrm{TI}}}\left(\left.\dfrac{\mathrm{d}\,x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}\right)^2}{\left.\dfrac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}}\right), \tag{4.70a}$$

$$c^{\mathrm{TI}}_{1122} = -c^{\mathrm{TI}}_{2323} + \frac{1}{(I_{n+1}-I_1)} \left(\frac{\sqrt{c^{\mathrm{TI}}_{3333}\,\rho^{\mathrm{TI}}} \left(\left.\frac{\mathrm{d}\,x_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \right)^2}{\left.\frac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}} - \frac{\sqrt{c^{\mathrm{TI}}_{2323}\,\rho^{\mathrm{TI}}} \left(\left.\frac{\mathrm{d}\,x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \right)^2}{\left.\frac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}} \right), \tag{4.70b}$$

$$c^{\mathrm{TI}}_{1133} = -c^{\mathrm{TI}}_{2323} \tag{4.70c}$$

$$+ \frac{\sqrt{\dfrac{\left(c^{\mathrm{TI}}_{2323} - c^{\mathrm{TI}}_{3333}\right) \left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} \left((I_{n+1}-I_1)\, c^{\mathrm{TI}}_{2323} \left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} - \sqrt{c^{\mathrm{TI}}_{3333}\,\rho^{\mathrm{TI}}} \left(\left.\dfrac{\mathrm{d}^2 x_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0} \right)^2 \right)}{(I_{n+1}-I_1)}}}{\left.\dfrac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}}. \tag{4.70d}$$

There exists another set of solutions where the second term in equation (4.70c) is subtracted. However, we reject this solution since, upon further numerical inspection, the corresponding elasticity tensor is not positive definite.

4.5.4 Formulae

Let us substitute the expressions for c^{TI}_{3333} and c^{TI}_{2323}, given in equations (4.26a) and (4.26b), along with equations (4.57a,b) and (4.61a,b), namely,

$$\left.\frac{\mathrm{d}\,x_P(\phi)}{\mathrm{d}\phi}\right|_{\phi=0}, \quad \left.\frac{\mathrm{d}\,x_S(\phi)}{\mathrm{d}\phi}\right|_{\phi=0} \quad \text{and} \quad \left.\frac{\mathrm{d}^2 t_P(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0}, \quad \left.\frac{\mathrm{d}^2 t_S(\phi)}{\mathrm{d}\phi^2}\right|_{\phi=0},$$

in to equations (4.70a), (4.70b), (4.70c) to obtain the approximate-medium elasticity parameters in terms of its constituent layer parameters. For conciseness, we adopt the following notation

$$\tau_1 := \sum_{i=1}^{n} \frac{I_{i+1}-I_i}{\sqrt{\dfrac{c^{\mathrm{iso}}_{1111,i}}{\rho_i}}}, \qquad \tau_2 := \sum_{i=1}^{n} \frac{I_{i+1}-I_i}{\sqrt{\dfrac{c^{\mathrm{iso}}_{2323,i}}{\rho_i}}}, \tag{4.71a,b}$$

$$\tau_3 := \sum_{i=1}^{n} \frac{\rho_1\,(I_{i+1}-I_i)}{c^{\mathrm{iso}}_{1111,1}} \sqrt{\frac{c^{\mathrm{iso}}_{1111,i}}{\rho_i}}, \qquad \tau_4 := \sum_{i=1}^{n} \frac{\rho_1\,(I_{i+1}-I_i)}{c^{\mathrm{iso}}_{2323,1}} \sqrt{\frac{c^{\mathrm{iso}}_{2323,i}}{\rho_i}}, \tag{4.71c,d}$$

$$\tau_5 := \sum_{i=1}^{n} (I_{i+1}-I_i) \sqrt{\frac{\rho_1\, c^{\mathrm{iso}}_{1111,i}}{\rho_i\, c^{\mathrm{iso}}_{1111,1}}}, \qquad \tau_6 := \sum_{i=1}^{n} (I_{i+1}-I_i) \sqrt{\frac{\rho_1\, c^{\mathrm{iso}}_{2323,i}}{\rho_i\, c^{\mathrm{iso}}_{2323,1}}}. \tag{4.71e,f}$$

Using notation (4.71), the approximate-medium elasticity parameters are

$$c_{1111}^{\mathrm{TI}} = \rho^{\mathrm{TI}} \left(\frac{{\tau_5}^2}{\tau_1\,\tau_3} + \frac{\dfrac{\tau_2\,{\tau_6}^2}{\tau_4} - \left(I_{n+1} - I_1\right)^2}{{\tau_2}^2} \right) , \tag{4.72a}$$

$$c_{1122}^{\mathrm{TI}} = \rho^{\mathrm{TI}} \left(\frac{{\tau_5}^2}{\tau_1\,\tau_3} - \frac{\dfrac{\tau_2\,{\tau_6}^2}{\tau_4} + \left(I_{n+1} - I_1\right)^2}{\tau_2^2} \right) , \tag{4.72b}$$

$$c_{1133}^{\mathrm{TI}} = \frac{\rho^{\mathrm{TI}}\left(I_{n+1} - I_1\right)}{{\tau_2}^2\,\tau_3} \left(\frac{\sqrt{\left({\tau_1}^2 - {\tau_2}^2\right)\tau_3\left(\left(I_{n+1} - I_1\right)^2\tau_1\,\tau_3 - {\tau_2}^2\,{\tau_5}^2\right)}}{{\tau_1}^{3/2}} - \left(I_{n+1} - I_1\right)\tau_3 \right) , \tag{4.72c}$$

$$c_{3333}^{\mathrm{TI}} = \frac{\rho^{\mathrm{TI}}\left(I_{n+1} - I_1\right)^2}{\tau_1^2} , \tag{4.72d}$$

$$c_{2323}^{\mathrm{TI}} = \frac{\rho^{\mathrm{TI}}\left(I_{n+1} - I_1\right)^2}{\tau_2^2} . \tag{4.72e}$$

4.6 Numerical examples

4.6.1 Approximate medium

Let us consider a stack of ten layers, whose density-normalized elasticity properties are given in Table 4.1a, whereas traveltimes through the stack at increasing offsets are given in Table 4.1b.

We begin by calculating the values for equations (4.71), where

$$\tau_1 = 0.2295\,, \quad \tau_2 = 0.5340\,, \quad \tau_3 = 0.4257\,, \quad \tau_4 = 0.9667\,,$$
$$\tau_5 = 1383.2206\,, \quad \tau_6 = 1374.0038\,, \quad \rho^{\mathrm{TI}} = 1\,.$$

Using these values, we evaluate parameters (4.72) to obtain

$$\begin{aligned} &c_{1111}^{\mathrm{TI}} = 19738951.6410\,, \quad c_{1122}^{\mathrm{TI}} = 12425456.1922\,, \quad c_{1133}^{\mathrm{TI}} = 12274442.1940\,, \\ &c_{3333}^{\mathrm{TI}} = 18991558.3336\,, \quad c_{2323}^{\mathrm{TI}} = 3506410.3613\,, \end{aligned} \tag{4.73}$$

whose units are $\mathrm{m}^2/\mathrm{s}^2$. Using elasticity parameters (4.73) in (4.6), the elasticity tensor

Layer	c^{iso}_{1111}	c^{iso}_{2323}	α	β
1	10.56	2.02	3.25	1.42
2	20.52	4.45	4.53	2.11
3	31.14	2.89	5.58	1.70
4	14.82	2.62	3.85	1.62
5	32.15	2.92	5.67	1.71
6	16.00	2.56	4.00	1.60
7	16.40	6.35	4.05	2.52
8	18.06	4.33	4.25	2.08
9	31.47	8.01	5.61	2.83
10	17.31	3.76	4.16	1.94

(a)

Receiver	Offset	P	S
1	0	229.5	534.0
2	100	230.6	536.6
3	200	233.9	544.2
4	300	239.2	556.5
5	400	246.5	573.2
6	500	255.6	593.9
7	600	266.1	618.0
8	700	278.0	644.9
9	800	291.0	674.2
10	900	304.9	705.3
11	1000	319.5	737.7

(b)

Table 4.1: Density-scaled elasticity parameters and corresponding traveltimes (adapted from Tables 4.1 and 4.2 (p. 175 and p. 185) of Slawinski (2018)). Table on the left tabulates density-scaled elasticity parameters, whose units are $10^6\,\mathrm{m^2 s^{-2}}$, for a stack of isotropic layers, and the corresponding P- and S-wave velocities in $\mathrm{km\,s^{-1}}$. Table on the right tabulates traveltimes, in milliseconds, through a stack of isotropic layers, whose properties are given in the left table and whose thicknesses are one hundred metres. The source is directly opposite the first receiver, and other receivers are spaced in hundred-metre intervals. The P-wave and S-wave traveltimes correspond to the layered medium and are obtained by invoking Fermat's principle.

for the approximate TI medium is

$$c^{\mathrm{TI}} = \begin{bmatrix} 19738951.6410 & 12425456.1922 & 12274442.1940 & 0 & 0 & 0 \\ 12425456.1922 & 19738951.6410 & 12274442.1940 & 0 & 0 & 0 \\ 12274442.1940 & 12274442.1940 & 18991558.3336 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\,(3506410.3613) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\,(3506410.3613) & 0 \\ 0 & 0 & 0 & 0 & 0 & 7313495.4488 \end{bmatrix}, \tag{4.74}$$

whose eigenvalues are

$$\lambda_1 = 44144211.9233\,, \quad \lambda_2 = \lambda_3 = 7313495.4488\,,$$

$$\lambda_4 = \lambda_5 = 7012820.7225\,, \quad \lambda_6 = 7011754.2436\,,$$

which belong to a transversely isotropic tensor (Bóna et al., 2007a), as expected.

Remark 3. *Let us exemplify the solution to system* (4.69) *that breaks positive definiteness. Specifically, in solution* (4.70c)*, the second term may be negative, which results in* $c^{\mathrm{TI}}_{1133} = -19287262.9165$*. Replacing this value in tensor* (4.74) *leads to eigenvalues that are*

$$\lambda_1 = 53638239.8999\,, \quad \lambda_2 = \lambda_3 = 7313495.4488\,,$$

$$\lambda_4 = \lambda_5 = 7012820.7225\,, \quad \lambda_6 = -2482273.7330\,.$$

Since $\lambda_6 < 0$ *the matrix is not positive definite.*

Using c^{TI}_{3333} and c^{TI}_{2323} of values (4.73) in (4.27), the elasticity tensor of the approximate isotropic medium is

$$c^{\mathrm{ISO}} = \begin{bmatrix} 18991558.3336 & 11978737.6111 & 11978737.6111 & 0 & 0 & 0 \\ 11978737.6111 & 18991558.3336 & 11978737.6111 & 0 & 0 & 0 \\ 11978737.6111 & 11978737.6111 & 18991558.3336 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\,(3506410.3613) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\,(3506410.3613) & 0 \\ 0 & 0 & 0 & 0 & 0 & 2\,(3506410.3613) \end{bmatrix}, \tag{4.75}$$

whose eigenvalues are

$$\lambda_1 = 42949033.5559\,, \quad \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = \lambda_6 = 7012820.7225\,,$$

which belong to an isotropic tensor (Bóna et al., 2007a), as expected.

Now, let us compare traveltimes through the layered and approximate media. For the layered medium, using equation (4.3), we calculate ϕ_P and ϕ_S for rays of P and S waves for receiver locations spaced in hundred-metre intervals. Then, for each receiver location, we substitute ϕ_P and ϕ_S in equation (4.5) to calculate t_P and t_S, respectively; these results are tabulated in the columns three and five of Table 4.2.

For the approximate isotropic medium, we calculate the distances traveled along the rays, L_P and L_S, for each receiver location using equation (4.9). Using equations (4.29), we calculate the traveltimes along the two rays, T_P and T_S; these results are tabulated in the last two columns of Table 4.2. In Figures 4.2 and 4.3 of Section 4.6.2, we plot the approximate isotropic P and S traveltimes and compare with traveltimes of the root-mean-square medium.

For the approximate TI medium, we calculate the distances traveled along the rays, L_{qP}, L_{qSV} and L_{SH}, for each receiver location using equation (4.9), and, using equation (4.8), we calculate the ray angles, θ_{qP}, θ_{qSV} and θ_{SH}. Using the ray angles, we solve relation (4.7) for the associated wavefront angles, ϑ_{qP}, ϑ_{qSV} and ϑ_{SH}, which are, in turn, used in equation (4.10) to calculate the ray velocities, V_{qP}, V_{qSV} and V_{SH}. Substituting the respective distances traveled and ray velocities in equation (4.11), we calculate the traveltimes along the three rays, T_{qP}, T_{qSV} and T_{SH}; these results are tabulated in Table 4.3. In Figures 4.4 and 4.5 of Section 4.6.3, we plot the approximate TI qP, qSV and SH traveltimes and compare with traveltimes of the Backus medium.

Let us elaborate on the raytracing for the layered and approximate media. On the left-hand side of Figure 4.1, we have the layered medium, which is comprised of ten, one-hundred-meter thick, horizontally stratified, homogeneous and isotropic layers, with the layer interfaces indicated by dotted lines and layer properties specified in Table 4.1(a). Therein, the source is located at the origin, the first receiver (square) are located at a depth of 1000 m directly opposite to the source, and the remaining receivers are spaced in one-hundred-metre intervals to the right. For each

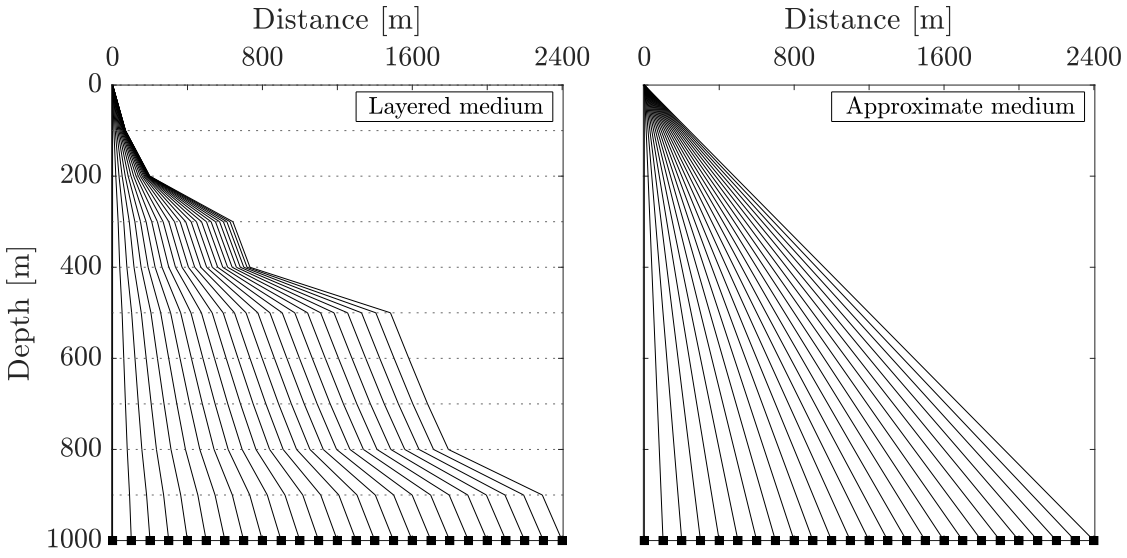


Figure 4.1: Selected raytracing for P waves in the layered medium (left) and for qP waves in the approximate medium (right). In both media, the source is located at the origin directly opposite the first receiver, which is located at a depth of 1000 m, and other receivers are spaced by one-hundred-metre intervals.

source-receiver combination, we use MATLAB's `fzero` function to calculate the takeoff angle in equation (4.3) that results in the required horizontal distance travelled. In Table 4.2, we present the takeoff angles for selected source-receiver combinations, for both P and S waves.

On the right-hand side of Figure 4.1, we have the approximate medium, which is homogeneous and either isotropic or transversely isotropic. For the former, the medium can represent either approximate ISO and root-mean-square medium, which is discussed in Section 4.6.2. For the latter, the medium can represent the approximate TI or the Backus medium, which we discuss in Section 4.6.3. For each source-receiver combination in the transversely isotropic media, we use MATLAB's `fzero` function with equation (4.7) to calculate the wavefront angle for the given ray angle, then calculate the ray velocity along the ray using equation (4.10). In Tables 4.3 and 4.4, we present the wavefront angles for selected source-receiver combinations, for qP, qSV, and SH waves.

Turning our attention to traveltimes, we observe in Figures 4.2 – 4.5 that the approximate TI traveltimes (second order) are always closer to the layered traveltimes than the isotropic (first order) traveltimes, as expected. The traveltime

x	ϕ_P	t_P	ϕ_S	t_S	θ	$T_{P,\mathrm{rms}}$	$T_{S,\mathrm{rms}}$	T_P	T_S
0	0.0000	229.4667	0.0000	534.0337	0.0000	225.9426	522.9408	229.4667	534.0337
200	8.1351	233.8683	8.1814	544.1598	11.3099	230.4172	533.2970	234.0111	544.6096
400	15.4507	246.5387	15.4857	573.2453	21.8014	243.3476	563.2244	247.1432	575.1719
600	21.4338	266.1239	21.3286	617.9779	30.9638	263.4921	609.8485	267.6019	622.7849
800	25.9486	290.9649	25.4767	674.1813	38.6598	289.3477	669.6909	293.8608	683.8968
1000	29.1251	319.5119	27.9682	737.7364	45.0000	319.5311	739.5499	324.5149	755.2376
1200	31.2288	350.5187	29.1599	805.1996	50.1944	352.9336	816.8596	358.4385	834.1872
1400	32.5634	383.0849	29.6504	874.3673	54.4623	388.7264	899.7013	394.7895	918.7862
1600	33.3957	416.6158	29.8608	944.2354	57.9946	426.3077	986.6827	432.9569	1007.6130
1800	33.9183	450.7433	29.9633	1014.4200	60.9454	465.2443	1076.8010	472.5009	1099.6430
2000	34.2527	485.2460	30.0195	1084.7670	63.4350	505.2230	1169.3310	513.1032	1194.1360
2200	34.4714	519.9904	30.0533	1155.2070	65.5561	546.0150	1263.7430	554.5314	1290.5510
2400	34.6170	554.8937	30.0751	1225.7040	67.3801	587.4508	1359.6460	596.6135	1388.4880

Table 4.2: Traveltimes along rays, for increasing horizontal offset, of P and S waves through layered medium, t_P and t_S, through the rms medium, $T_{P,\mathrm{rms}}$ and $T_{S,\mathrm{rms}}$, through approximate isotropic medium, T_P and T_S; layer properties are given in Table 4.1, whereas elastic properties of approximate isotropic medium are given by $c^{\mathrm{ISO}}_{1111} := c^{\mathrm{TI}}_{3333}$ and $c^{\mathrm{ISO}}_{2323} := c^{\mathrm{TI}}_{2323}$ of parameters (4.73). The receiver offsets x are in metres while the traveltimes in milliseconds. The ray angles in the layered medium, ϕ_P and ϕ_S, as well as in the rms and approximate isotropic media, θ, are in degrees.

curves of the approximate TI medium remain within 1% error of those in the layered medium up to 1,100 metres of offset, and, for the approximate isotropic medium, up to 800 metres.

4.6.2 Root-mean-square medium

As noted in (4.30) the wave velocity in the approximate ISO medium is the weighted harmonic mean of the velocities in the individual layers. However, other averaging procedures have been proposed in the literature. In particular, Dix (1955, expression (10)) proposes the root-mean-square (rms) velocity,

$$V_{\nu,\mathrm{rms}} = \sqrt{\frac{\sum_{i=1}^{n} {\nu_i}^2\, t_i}{\sum_{i=1}^{n} t_i}}\,, \tag{4.76}$$

where ν_i is the speed of a wave in the ith layer, to be replaced for a P or S wave by either equation (4.4a) or (4.4b), respectively, and t_i is the traveltime along a vertical ray of that wave in the ith layer. We refer to this medium as the *rms medium*. For a detailed derivation of expression (4.76), we refer the reader to Taner and Koehler (1969). Therein, the rms velocity is a coefficient of a power series expansion within the

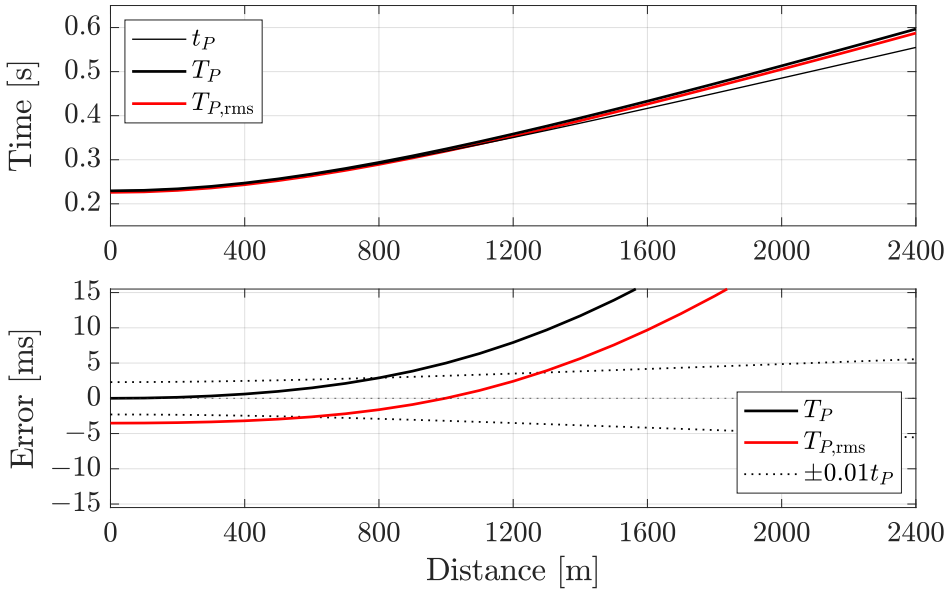


Figure 4.2: Illustration of P-wave traveltimes and error for increasing horizontal offset. The traveltimes in the approximate isotropic medium remain within 1% of the layered traveltimes for up to 800 metres; traveltimes in the root-mean-square medium has no such properties.

context of the so-called X^2-T^2 method, which is beyond the scope of this dissertation.

Now, we may calculate, using the fact that $t_i = (I_{i+1} - I_i)/\nu_i$,

$$V^2_{\nu,\mathrm{rms}} = \frac{\sum_{i=1}^{n} {\nu_i}^2 \, t_i}{\sum_{i=1}^{n} t_i} = \frac{\sum_{i=1}^{n} \nu_i (I_{i+1} - I_i)}{\sum_{i=1}^{n} \frac{I_{i+1} - I_i}{\nu_i}} = A(\nu_i)\, H(\nu_i)$$

where

$$A(\nu_i) := \frac{\sum_{i=1}^{n} \nu_i (I_{i+1} - I_i)}{\sum_{i=1}^{n} (I_{i+1} - I_i)}$$

is the weighted *arithmetic* mean of the velocities, and H is the weighted harmonic mean. In other words, the rms mean of the velocities is the geometric mean of $A(\nu_i)$ and $H(\nu_i)$.

In case the layer velocities are all very similar then it is to be expected that the rms

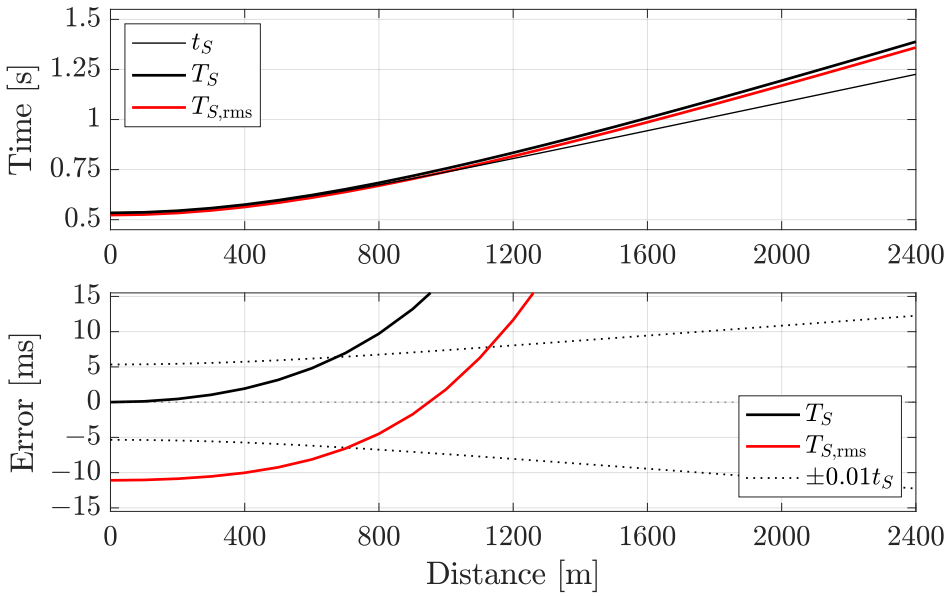


Figure 4.3: Illustration of S-wave traveltimes and error for increasing horizontal offset. The traveltimes in the approximate isotropic medium remain within 1% of the layered traveltimes for up to 600 metres; traveltimes in the root-mean-square medium has no such properties.

x	θ	ϑ_{qP}	T_{qP}	ϑ_{qSV}	T_{qSV}	ϑ_{SH}	T_{SH}
0	0.0000	0.0000	229.4667	0.0000	534.0337	0.0000	534.0337
200	11.3099	10.9675	233.8726	10.8886	544.1942	10.8562	544.1789
400	21.8014	21.1568	246.6060	21.1998	573.7510	20.9846	573.5387
600	30.9638	30.0884	266.4517	30.4604	620.2535	29.9132	619.3869
800	38.6598	37.6353	291.9438	38.4066	680.4970	37.4922	678.3884
1000	45.0000	43.8982	321.7342	45.0147	751.3071	43.7977	747.4349
1200	50.1944	49.0681	354.7393	50.4273	830.0106	49.0073	824.0052
1400	54.4623	53.3458	390.1446	54.8496	914.5577	53.3172	906.1939
1600	57.9946	56.9083	427.3554	58.4817	1003.4500	56.9038	992.6065
1800	60.9454	59.9001	465.9406	61.4920	1095.6100	59.9131	1082.2320
2000	63.4350	62.4355	505.5866	64.0132	1190.2660	62.4608	1174.3340
2200	65.5561	64.6037	546.0633	66.1474	1286.8550	64.6375	1268.3740
2400	67.3801	66.4741	587.1993	67.9724	1384.9680	66.5136	1363.9510

Table 4.3: Traveltimes along rays, for increasing horizontal offset, of qP , qSV and SH waves through the approximate TI medium, T_{qP} , T_{qSV} , T_{SH} ; elastic properties of medium are given by parameters (4.73). The receiver offsets x are in metres while the traveltimes in milliseconds. The ray angle, θ , along with wavefront angles for the qP , qSV and SH waves, ϑ_{qP} , ϑ_{qSV} , ϑ_{SH} , are in degrees.

(and indeed *any* average procedure) will produce a velocity nearly equal to the harmonic mean. Consider the following numerical example as a demonstration. We calculate the rms velocities for P and S waves, which are $V_{P,\mathrm{rms}} = 4425.9027\,\mathrm{m/s}$ and $V_{S,\mathrm{rms}} = 1912.2625\,\mathrm{m/s}$. Then, we calculate the traveltimes along the rays of the P and S waves propagating through the rms medium for source-receiver combinations that are consistent with Section 4.6.1; we tabulate the results in columns seven and eight of Table 4.2. Also, we illustrate the rms-velocity traveltime curves of the P and S waves in Figures 4.2 and 4.3, respectively.

More generally, if the layer velocities differ significantly, then it can be expected that the rms mean and the harmonic mean differ substantially. However, the harmonic mean for our proposed ISO medium is exact for vertical rays, agrees with the layered velocity to order 1 (in the propagation angle), and hence is guaranteed to be a good approximation for small propagation angles — the rms medium has no such properties. At the cost of a slightly more complicated formulation, the approximate TI medium will provide an even better approximation (for small propagation angles).

4.6.3 Backus medium

To gain insight on the traveltimes in the approximate TI medium, let us compare them with traveltime along the same raypaths within a TI medium that is the result of the Backus average[1]. The criterion of our homogenization is an approximate equality of traveltimes between a stack of layers and a single medium. In contrast, the criterion of Backus's homogenization is derived based on mechanical considerations by assuming a load applied to a horizontal plane of a stack of layers in a static equilibrium (e.g., Slawinski, 2018, Section 4.2.2.2). Hence, the stress-tensor components of the applied load are σ_{13}, σ_{23} and σ_{33}. In view of the static equilibrium, these stress-tensor components are constant throughout.

To compare, we perform the Backus average on the layers in Table 4.1(a) to obtain an equivalent homogenous TI medium, which we refer to as the *Backus medium*. Its elasticity tensor is

$$c^{\overline{\mathrm{TI}}} = \begin{bmatrix} 18835691.9900 & 10853691.9900 & 10959063.4413 & 0 & 0 & 0 \\ 10853691.9900 & 18835691.9900 & 10959063.4413 & 0 & 0 & 0 \\ 10959063.4413 & 10959063.4413 & 18432619.1430 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\,(3378620.2475) & 0 & 0 \\ 0 & 0 & 0 & 0 & 2\,(3378620.2475) & 0 \\ 0 & 0 & 0 & 0 & 0 & 7982000.0000 \end{bmatrix} \tag{4.77}$$

and its eigenvalues are

$$\begin{aligned} \lambda_1 &= 40549810.8475\,, \quad & \lambda_2 &= \lambda_3 = 7982000.0000\,, \\ \lambda_4 &= 7572192.2754\,, \quad & \lambda_5 &= \lambda_6 = 6757240.4950\,; \end{aligned}$$

herein, $^{\overline{\mathrm{TI}}}$ denotes a quantity of the Backus medium. We calculate the traveltimes along the rays of qP, qSV, SH waves propagating through the Backus medium for source-receiver combinations that are consistent with Section 4.6.1; the results are tabulated in Table 4.4.

We recall that the Backus medium is said to be "long-wave equivalent" to the original, layered medium. Thus, the limitation to long waves removes high-frequency wavelengths and, as such, the first arrivals of nearly vertical waves in the layered medium arrive earlier than in the Backus medium (Backus, 1962, p. 4432). As well,

[1]Readers interested in the details of the Backus average may refer to the original article by Backus (1962) as well as investigations by Bos et al. (2017, 2018, 2019c).

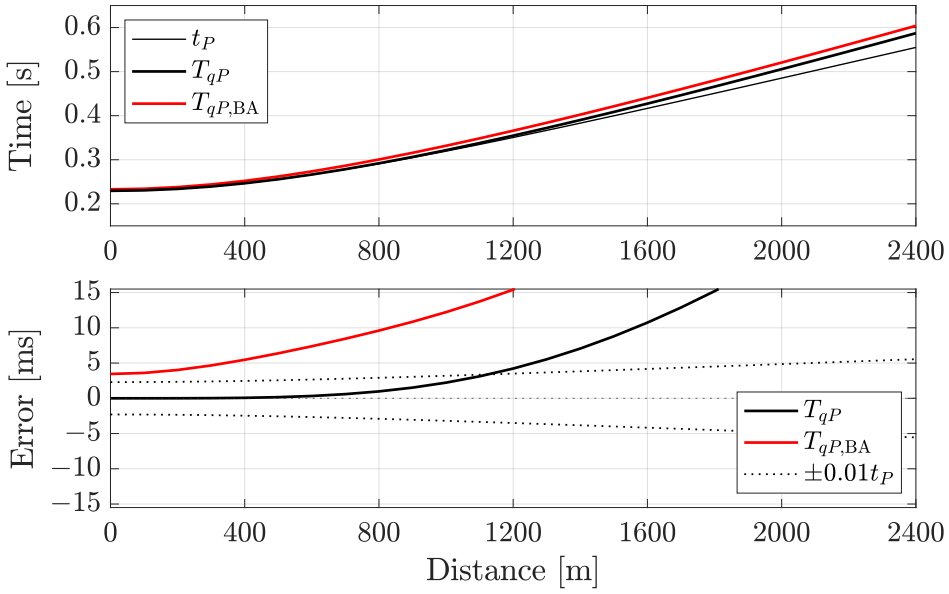


Figure 4.4: Illustration of P-wave traveltimes and error for increasing horizontal offset. The traveltimes in the approximate TI medium remain within 1% of the layered traveltimes for up to 1,100 metres; traveltimes in the Backus medium have no such properties.

as a consequence of the above formulation, this statement holds for the approximate TI medium.

Let us compare the traveltimes within the layered medium, which are tabulated in columns three and five of Table 4.2, against the approximate TI and Backus media, which are tabulated in Tables 4.3 and 4.4. Therein, nearly vertical traveltimes in the layered medium, and the approximate TI medium, are faster than those in the Backus medium; the traveltimes of the layered and Backus media are illustrated in Figures 4.4 and 4.5. However, this is not the case for qSV and SH waves. Comparing traveltimes, we observe that, for this example,

$$T_{qSV} < T_{qSV,BA} \quad \text{for} \quad \theta \leq 20.4135^\circ\,, \qquad T_{qSV} < T_{SH,BA} \quad \text{for} \quad \theta \leq 32.6853^\circ\,,$$
$$T_{SH} < T_{qSV,BA} \quad \text{for} \quad \theta \leq 20.5859^\circ\,, \qquad T_{SH} < T_{SH,BA} \quad \text{for} \quad \theta \leq 34.4625^\circ\,.$$

Traveltimes in Figures 4.4 and 4.5 demonstrate that the Backus medium does not provide an accurate approximation of traveltime within its constituent layers, which — we must emphasize — is not the criterion of the Backus average. However,

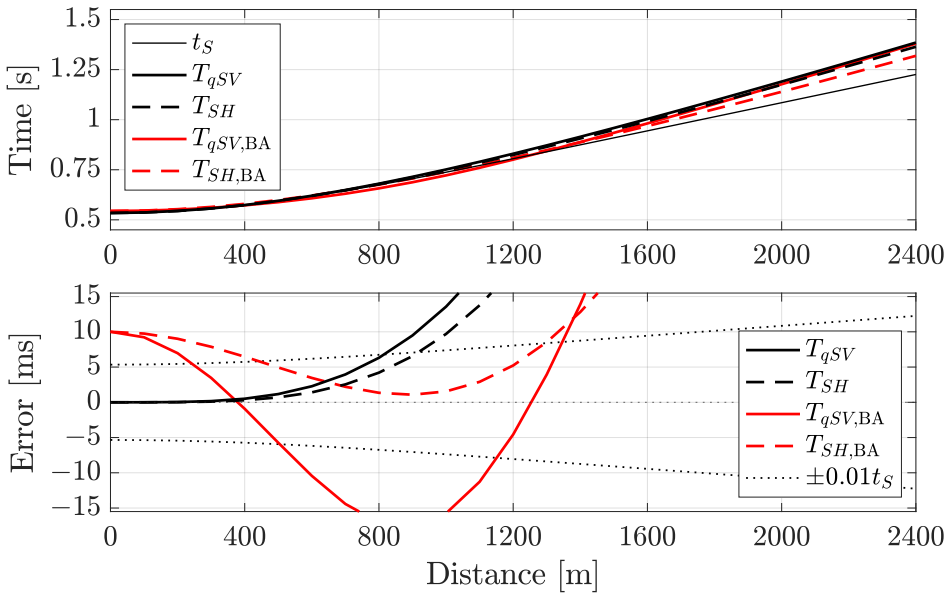


Figure 4.5: Illustration of S-wave traveltimes and error for increasing horizontal offset. The traveltimes in the approximate isotropic medium remain within 1% of the layered traveltimes for up to 800 metres; traveltimes in the Backus medium have no such properties.

in contrast to Dalton and Slawinski (2016), who "[...] postulate that [the] validity of the Backus (1962) average, whose weights are layer thicknesses, is limited to waves whose incidence is nearly vertical. The accuracy of this average decreases with the increase of the source-receiver offset," the Backus medium cannot be validated by a comparison with Fermat traveltimes through its constituent layers. The Backus average simulates the behaviour of actual physical phenomena; it produces an estimate of physical behaviour as opposed to an approximation of traveltimes. The traveltime is a result of this simulation.

x	θ	$\vartheta_{qP,BA}$	$T_{qP,BA}$	$\vartheta_{qSV,BA}$	$T_{qSV,BA}$	$\vartheta_{SH,BA}$	$T_{SH,BA}$
0	0.0000	0.0000	232.9198	0.0000	544.0393	0.0000	544.0393
200	11.3099	12.1112	237.8863	7.5128	551.1282	9.6097	553.1739
400	21.8014	22.7944	252.0031	15.2873	572.3210	18.7073	579.7146
600	30.9638	31.5758	273.5057	23.7440	607.5431	26.9277	621.4352
800	38.6598	38.6596	300.5706	33.6522	657.1353	34.1077	675.5291
1000	45.0000	44.4062	331.7358	45.1896	721.9872	40.2499	739.2852
1200	50.1944	49.1266	365.9307	54.6983	800.6333	45.4511	810.4262
1400	54.4623	53.0544	402.3906	60.8739	888.2660	49.8440	887.1774
1600	57.9946	56.3611	440.5689	65.0791	981.3334	53.5622	968.2056
1800	60.9454	59.1741	480.0709	68.1541	1077.9280	56.7251	1052.5230
2000	63.4350	61.5893	520.6076	70.5189	1176.9640	59.4328	1139.4010
2200	65.5561	63.6806	561.9645	72.4032	1277.7660	61.7672	1228.2950
2400	67.3801	65.5051	603.9801	73.9447	1379.8920	63.7941	1318.7970

Table 4.4: Traveltimes along rays, for increasing horizontal offset, of qP , qSV and SH waves through the Backus medium, $T_{qP,\mathrm{BA}}$, $T_{qSV,\mathrm{BA}}$, $T_{SH,\mathrm{BA}}$; elastic properties of medium are given in tensor (4.77). The receiver offsets x are in metres while the traveltimes in milliseconds. The ray angle, θ , along with wavefront angles for the qP , qSV and SH waves, $\vartheta_{qP,\mathrm{BA}}$, $\vartheta_{qSV,\mathrm{BA}}$, $\vartheta_{SH,\mathrm{BA}}$, are in degrees.

4.7 Concluding remarks

The Taylor series expansion of traveltime expressions, about the vertical axis, lead to an approximate homogeneous medium, wherein traveltimes are similar to Fermat traveltimes within its constituent layers. A numerical example suggests that the approximation leads to empirically adequate results for not only near-vertical rays, but even for rays whose takeoff angles approach 30° . Even though a source-receiver offset of a homogenized medium is not restricted by a critical angle, a Fermat ray in a stack of layers might be restricted. Since these angles can result in sufficiently large offsets, our ability to examine greater offsets are limited.

In summary, we have developed a formulation for approximate TI media that has similarities to the commonly used root-mean-square velocity. For a stack of isotropic layers, there is an approximate equality between a layered and single medium (e.g., Sheriff, 2002, p. 377, NMO velocity). The approximate medium can be used to examine the accuracy of using the hyperbolic curve for a horizontally stratified medium as well as of using the root-mean-square velocity in TI media.

Finally, the resulting approximation could be considered if modelling traveltimes for transversely isotropic media, where the anisotropy results from a stack of layers. In particular, the approximate medium could serve as an alternative to the equivalent medium that is the result of the Backus average applied to the same constituent layers.

Chapter 5

On inverse modelling of traveltimes

Author note

The content of this chapter has partial overlap with Kaderali and Stanoev (2020), which is a collaborative research project with Mr. Ayiaz Kaderali. For further details regarding co-authorship, we refer the reader to the statement provided on page xii.

5.1 Introductory remarks

In this chapter, we deviate from the homogeneous anisotropic representations of a horizontally stratified subsurface, which has been the focus of Chapters 3 and 4. As opposed to considering homogenizations of seismic media for forward modelling based on elasticity theory, we turn our attention to the practical application of estimation of anisotropy and inhomogeneity properties within the layers of the subsurface. In other words, this chapter focuses on traveltime tomography, otherwise known as inversion, which is a procedure that allows us to invert observed seismic traveltimes to estimate the subsurface velocity structure (e.g., Lines and Newrick, 2004, Chapter 17).

The chapter proceeds as follows. We begin the chapter with a brief review of traveltime inversion in Section 5.2 and how it applies to our problem. Next, in Section 5.3, we discuss the modelling considerations for the inversion. Specifically, we provide an overview of the vertical seismic profile (VSP), establish the background model, and specify raytracing and traveltime expressions for multilayer

model, wherein the inhomogeneity and anisotropy of each layer can vary. Then, in Section 5.4, we detail the numerical optimization procedure that we use to estimate the parameter values through the process of minimizing the residual sum of squares with an arbitrary set of traveltimes. We adapt the derivative-free algorithm to account for constraints on model parameters and ensure model feasibility. After, in Section 5.5, to establish a sense of the multilayer model's accuracy in parameter estimation using the optimization procedure, we devise a simulation study. In particular, we use identical source-receiver combinations as the VSP and, using a set of predetermined model-parameter values, we generate simulated traveltimes that resemble those that are measured. We apply the optimization procedure to assess its reliability in retrieving the predetermined, otherwise referred to as known, parameter values. Finally, in Section 5.6, we apply the procedure to the measured traveltimes and, within an understanding of the procedure's limitations, we discuss the results.

5.2 Traveltime inversion

Traveltime inversion relies on mathematical formulations to locate a path between a source and receiver within an anisotropic and/or inhomogeneous medium. In a recent review, Rawlinson et al. (2008) discuss two approaches that are used to locate such paths: raytracing and grid-based schemes. The former corresponds to trajectories of paths between points in a medium that are normal to the wavefront and has been covered in a comprehensive manner (e.g. Červený, 2001; Chapman, 2004; Slawinski, 2020a).

The latter corresponds to the finite-difference solution of the eikonal equation within a discretized velocity grid that is useful in modelling complicated subsurface structures. This approach has become popular in recent studies. For example, Bóna et al. (2009) propose a raytracing method that determines minimal-time rays within a discretized grid based on the concept of simulated annealing. Zhang et al. (2017) present an elliptically anisotropic raytracing method that models with irregular subsurface structures by applying flexible triangular grids. Within the context of traveltime inversion, Lelièvre et al. (2011) develop a method to invert traveltimes without ray tracing using a finite-difference forward solution of the eikonal equation, which is implemented on unstructured grids that discretize the subsurface velocity. Giroux and Gloaguen (2012) invert cross-hole traveltime data using ray tracing in a

discretized elliptically anisotropic model, which is based on an assumption that the horizontal velocity is a scalar multiple of the vertical velocity (e.g., Helbig, 1983). Meléndez et al. (2019) perform a traveltime inversion to compute the velocities observed by rays under the weak VTI anisotropy formulation of Thomsen (1986).

By contrast, the subsurface in question is reported to have a horizontally stratified structure comprised of shale formations, as discussed in Section 5.3.1. Thus, it is reasonable to model the subsurface as a series of horizontal layers. Likewise, to account for these properties in each layer, we assume that the vertical velocity increases linearly with depth, phase velocity exhibits an elliptical velocity dependence, and each layer is laterally homogeneous. Since the traveltime accuracy within grid-based schemes is generally not as high as raytracing (Rawlinson et al., 2008, p. 3), we proceed with the latter.

Under the assumptions of the previous paragraph, the eikonal equation may be solved explicitly by the method of characteristics, which yields closed-form raytracing and traveltime expressions (Rogister and Slawinski, 2005). As such, the velocity in any given layer depends on only three model parameters and the raypaths and traveltimes are calculated using closed-form expressions. Using numerical optimization, we adjust the model parameters so as to minimize the residual sum of squares (RSS) between the traveltimes.

The traveltime tomography that we perform in this chapter depends heavily on ray theory, which we discuss in greater detail in Section 2.5. In particular, we calculate the traveltimes along a specified ray path that is obtained from the eikonal equation. However, it is possible to perform traveltime tomography based on the wave equation, which is elastodynamic equation (2.52).

In contrast to the ray-based approach, wave-equation-based tomography can account for not only traveltime information but also amplitude, attenuation, and even full waveform content to image subsurface structures (Tong, 2021). One advantage of wave-equation methods is that they do not require the high-frequency assumption, within the context of the eikonal equation, and can account for different types of waves simultaneously as well as their reflections and resulting multiples. A popular inversion approach that uses the wave equation is the so-called full-waveform inversion (FWI), wherein many types of waves are involved for the least-squares optimization scheme to iteratively improve the model of the subsurface (Virieux and Operto, 2009).

The FWI approach, originally developed by Tarantola (1984), has been used to create heterogeneous velocity models in a variety of settings. For example, in recent years, Yang et al. (2016) perform a time-lapse inversion on the Valhall field in the North Sea using ocean-bottom-cable data, Beller et al. (2018) use a multiparameter 3-D inversion to determine the lithospheric structure of the South-Western Alps, Li et al. (2022) recover the P-wave velocity with first-arrival signals of land data in western Canada, and Dhabaria and Singh (2024) estimate elasticity parameters with Monte Carlo methods using ocean-bottom-seismometer data from the Atlantic Ocean. Even beyond geophysical applications, Guasch et al. (2020) have applied FWI to generate accurate three-dimensional images of the brain with sub-millimetre resolution.

While FWI can be more accurate at reconstructing a velocity model of the Earth than those methods based on raytracing, it is said to be at least an order of magnitude more costly because the wave equation must be numerically solved for each source (Fu and Hanafy, 2017). Likewise, given the increased dimensionality of this type of inversion, matching the whole waveform can lead to many local minima. As such, one of the key issues of FWI is obtaining an accurate initial model from which the local optimization can start. In many of the above examples, the initial model is provided by raytracing methods. Therefore, one of the applications of our results is in its use as an initial model for more-sophisticated imaging techniques like FWI.

5.3 Modelling considerations

5.3.1 VSP overview

We base the experimental setup on the VSP acquisition described by Kaderali (2009), which corresponds to a deviated survey well in offshore Newfoundland. These data consist of VSPs (primary arrivals, P-wave traveltimes) with zero-offset and walkaway source-receiver combinations — a walkaround acquisition is provided as well, but falls beyond the scope of this dissertation. In an earlier work by Zhou and Kaderali (2006, p. 2), the subsurface is confirmed to have a horizontally stratified structure comprised of shale formations up to a depth boundary of approximately 2300 metres. In Figure 1 therein, the subsurface is partitioned into five macro layers based on the zero-offset VSP, with interfaces at approximately 450 m, 1300 m, 1700 m, 2100 m, and 2300 m.

The well in question is deviated from the vertical, as detailed by its well

trajectory (Kaderali, 2009, Figure 2.2). The zero-offset VSP (ZVSP) consists of 93 measurements — provided by Kaderali (2009, Table A.1) — between depths of 418.60 metres, which corresponds to the sea floor, and approximately 2950 metres. The walkaway VSP (WVSP) consists of 200 source locations centred over the receiver array and are spaced at approximately 25-metre intervals, with a maximum offset of approximately 4000 metres in one direction and 1000 metres in the other. Respectively, the measurements in either direction are denoted as long- or short-side traveltimes. The source is placed at six metres below sea level and the signal is recorded by a five-receiver array placed above the 2300-metre boundary. The receiver depths, referenced from mean sea level (MSL), are reported as

$$\boldsymbol{Z} = [1979.923, 1989.809, 1999.669, 2009.758, 2019.927]^T. \tag{5.1}$$

Depths (5.1), which can be found in Zhou and Kaderali (2006, Table 2), refer to the first of three so-called shot lines, with the remaining lines pertaining to deeper receiver placements. Note that the indicated receiver depths in the tables therein indicate the vertical depth without the MSL correction. Further details of the acquisition may be found in Kaderali (2009, Chapters 2 and 3).

For the ZVSP traveltimes, we only consider measurements up to 2100 metres as the WVSP receivers are placed above the indicated macro-layer interface. However, between the depths of 418.60 and 1343.40 metres, the 13 traveltimes are recorded at non-regular depth intervals, whereas the remaining 39 are recorded at approximately 20-metre intervals. Consequently, there is a scarcity of ZVSP data for nearly 65% of the subsurface: an unavailability of measurements for approximately the shallowest 20% and non-regular measurements for the next 45%. For the WVSP traveltimes, we focus on the first shot line as the measurements are provided explicitly in Kaderali (2009, Appendix C). Of the WVSP traveltimes, we omit from consideration the final 21 measurements that correspond to receivers 1 and 2 due to measurement error that has resulted in observable scatter. Thus, of this dataset, we use 52 of the ZVSP traveltimes, which correspond to a depth of up to 2100 metres, and 958 of the WVSP traveltimes; in total, we consider $M = 1010$ data points.

5.3.2 Background model

To perform the traveltime inversion on the VSP data, we consider a background model comprised of N horizontal layers. The background model is embedded in a left-handed three-dimensional coordinate system, where the z-axis increases downward.

We make several simplifying assumptions about the layers. First, we assume that they are bounded by planar interfaces in welded contact. The thicknesses of the layers are calculated by taking the differences of successive interface depths, $\boldsymbol{\xi} = [\xi_0, \ldots, \xi_{N-1}]^T$, where, for example, the thickness of the ith layer is $\xi_i - \xi_{i-1}$; consequently, the Nth layer is a halfspace. Second, we assume that the layers are laterally homogenous, which means that the properties vary only with respect to depth. Third, we assume that the first layer, which represents the water column, is homogeneous and isotropic. Therein, the velocity of the P wave is constant in all directions. Finally, we assume that the subsequent $N-1$ layers are subject to a linear increase in velocity with respect to depth, z, and exhibit an elliptical velocity dependence with respect to the wavefront angle, ϑ.

We model the propagation of the P wave using seismic ray theory, which we discuss in Section 2.5. Since the VSP data contain traveltimes that correspond to vertical and offset source-receiver combinations, we require a framework for rays that are entirely vertical as well as those that are offset by a horizontal distance. We use ray theory to specify the raypath — and quantify the traveltime along it — that connects a given source, $\boldsymbol{S} \in \mathbb{R}^3$, and receiver, $\boldsymbol{R} \in \mathbb{R}^3$. Within the model, we consider M source-receiver combinations. As such, the jth ray is traced from

$$\boldsymbol{S}_j := \left[X_S^{(j)}, Y_S^{(j)}, Z_S^{(j)}\right]^T \quad \text{to} \quad \boldsymbol{R}_j := \left[X_R^{(j)}, Y_R^{(j)}, Z_R^{(j)}\right]^T, \quad \text{where} \quad j = 1, \ldots, M.$$

Now, let us consider Figure 5.1, which illustrates the background model and the jth source-receiver combination. Since the model consists of horizontal and parallel layer interfaces, the jth ray is traced within a unique plane — we refer to this plane as the rayplane — that contains the unit source-receiver vector, $\mathbf{d}_j = (\boldsymbol{R}_j - \boldsymbol{S}_j)/\|(\boldsymbol{R}_j - \boldsymbol{S}_j)\|$, and the z-direction basis vector, $\mathbf{e}_z = [0, 0, 1]^T$. For offset rays, we identify the jth rayplane by its unit normal vector, $\mathbf{n}_j = (\mathbf{d}_j \times \mathbf{e}_z)/\|(\mathbf{d}_j \times \mathbf{e}_z)\|$. For vertical rays, we do not use the rayplane terminology because the normal is indeterminate, as $\mathbf{d}_j \cdot \mathbf{e}_z = 1$; as such, rayplane refers only to offset source-receiver combinations. Within the jth

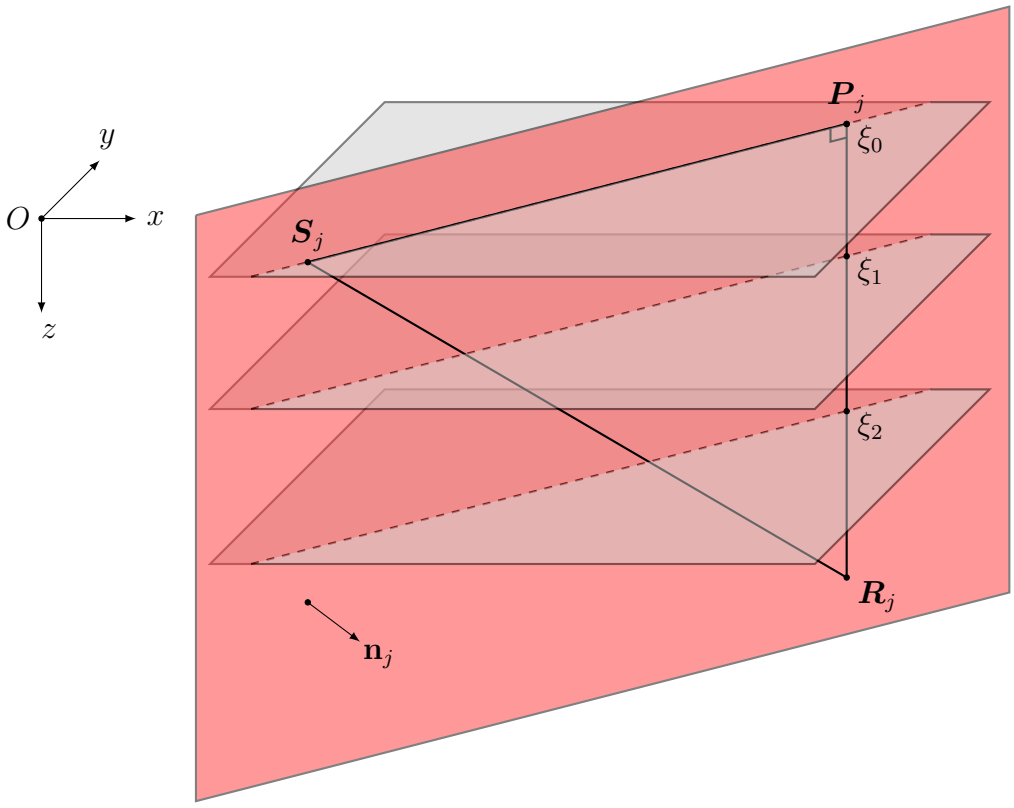


Figure 5.1: Rayplane within background model comprised of $N = 3$ layers. The left-handed coordinate system is specified by the coordinate axes, which are centred on the origin O. The planar layer interfaces (gray) are horizontal and parallel, with depths $\boldsymbol{\xi} = [\xi_0, \xi_1, \xi_2]^T$. For the jth source-receiver combination, the ray travels within the unique rayplane (red) with unit normal $\mathbf{n}_j$; herein, the receiver is in layer $\ell_j = 3$. The horizontal and vertical distances travelled by the ray are specified by right triangle $\triangle \boldsymbol{S}_j \boldsymbol{P}_j \boldsymbol{R}_j$.

rayplane, we specify a right triangle $\triangle \boldsymbol{S}_j \boldsymbol{P}_j \boldsymbol{R}_j$, where

$$\boldsymbol{P}_j := \boldsymbol{R}_j - (\boldsymbol{R}_j - \boldsymbol{S}_j) \cdot \mathbf{e}_z.$$

The vertical distance of the triangle is $(\boldsymbol{R}_j - \boldsymbol{S}_j) \cdot \mathbf{e}_z = Z_R^{(j)} - Z_S^{(j)}$ and the signed horizontal distance is $\text{sgn}(\mathbf{d}_j \cdot \mathbf{e}_x) \|\boldsymbol{P}_j - \boldsymbol{S}_j\|$, where $\mathbf{e}_x = [1, 0, 0]^T$ is the x-direction basis vector. We identify the layer within which the jth receiver resides by

$$\ell_j := \sum_{i=1}^{N-1} \left(Z_R^{(j)} - \xi_i \right).$$

5.3.3 Raytracing and traveltime expressions

In this section, we provide the raytracing and traveltime expressions in media with either a constant velocity or those that exhibit a linear increase velocity with depth and an elliptical velocity dependence. Since the source and receiver coordinates are known, we parameterize the expressions with respect to depth. For this section, we omit layer, i, and ray, j, subscripts, which means that the expressions, herein, apply to a single ray within either type of medium.

Constant velocity

In a homogeneous medium with constant velocity, the characteristic system of the eikonal equation results in a constant ray-parameter value and straight-line solutions for the raypaths, as is demonstrated Section 2.5.3. Within a given rayplane, we can adapt raytracing expressions (2.93) to quantify the horizontal, x, distance covered by the ray with respect to depth, z, by

$$x(z) = x_0 + \frac{pv(z - z_0)}{\sqrt{1 - p^2 v^2}}, \tag{5.2}$$

where p is the ray parameter, v is the constant velocity, and x_0 and z_0 are initial conditions for horizontal and vertical distances, respectively. In such a medium, the traveltime along the ray is

$$t(z) = t_0 + \frac{z - z_0}{v\sqrt{1 - p^2 v^2}}, \tag{5.3}$$

where t_0 is an initial time condition.

We convert raytracing expressions (2.93) from an arclength to a depth parameterization using the Pythagorean Theorem, where $(s - s_0)^2 = (x - x_0)^2 + (z - z_0)^2$ and s_0 is an initial arclength condition. In terms of the ray parameter, $p = \sin\theta/v$, the relationships are

$$s - s_0 = 1, \quad x - x_0 = pv, \quad z - z_0 = \sqrt{1 - p^2 v^2}.$$

For the traveltime inversion, we use expression (5.2) to trace the ray through the water column and expression (5.3) to quantify the traveltime. Therein, we fix $v = v_w := 1500\,\mathrm{m/s}$.

Linear inhomogeneity and elliptical anisotropy

Let us consider a medium with a linear increase in velocity with depth and an elliptical velocity dependence. We refer to the former as linear inhomogeneity and the latter as elliptical anisotropy. In such a medium, the phase velocity is

$$v(\vartheta) := \sqrt{{v_x}^2 \sin^2\vartheta + {v_z}^2 \cos^2\vartheta}\,, \tag{5.4}$$

where v_x and v_z are the magnitudes of the wavefront velocity along the x-axis and z-axis, respectively, and ϑ is measured from the z-axis. The following anisotropy parameter,

$$\chi = \frac{{v_x}^2 - {v_z}^2}{2\,{v_z}^2},$$

quantifies the difference between horizontal and vertical velocities; if the velocities are equal, $\chi = 0$, which is tantamount to an isotropic medium. Then, the wavefront velocity (5.4) becomes

$$v(\vartheta) = v_z\sqrt{(1 + 2\,\chi)\sin^2\vartheta + \cos^2\vartheta}\,. \tag{5.5}$$

Subjecting the wave to a linear increase of velocity along the z-axis, wavefront velocity (5.5) becomes

$$v(\vartheta, z) = (a + b\,(z - z_0))\sqrt{(1 + 2\,\chi)\sin^2\vartheta + \cos^2\vartheta}\,, \tag{5.6}$$

where a and b are constants with units of velocity and reciprocal of time, respectively, and z_0 is an initial depth. For brevity, henceforth, we refer to linearly inhomogeneous and elliptically anisotropic media as $ab\chi$ media.

To obtain the characteristic raytracing system for the $ab\chi$ medium, Rogister and Slawinski (2005) substitute phase velocity (5.6) in Hamiltonian (2.94). Let us summarize their result. The corresponding raytracing system is

$$\begin{cases} \dfrac{\mathrm{d}\,x}{\mathrm{d}t} = \dfrac{\partial\,\mathcal{H}}{\partial p_x} = \left(a + b\,(z - z_0)\right)^2 (1 + 2\,\chi)\,p_x\,, \\ \dfrac{\mathrm{d}\,z}{\mathrm{d}t} = \dfrac{\partial\,\mathcal{H}}{\partial p_z} = \left(a + b\,(z - z_0)\right)^2 p_z\,, \\ \dfrac{\mathrm{d}\,p_x}{\mathrm{d}t} = -\dfrac{\partial\,\mathcal{H}}{\partial x} = 0\,, \\ \dfrac{\mathrm{d}\,p_z}{\mathrm{d}t} = -\dfrac{\partial\,\mathcal{H}}{\partial z} = -b\,(a + b\,(z - z_0))\left((1 + 2\,\chi)\,{p_x}^2 + {p_z}^2\right)\,. \end{cases} \tag{5.7}$$

The solution to system (5.7), which is detailed further by Slawinski (2020a, Section 8.5), is

$$\begin{cases} x(t) = x_0 \\ \qquad + \dfrac{1}{b\,p}\left(\tanh\left(b\,t - \operatorname{arctanh}\left(\sqrt{1 - p^2a^2\,(1 + 2\,\chi)}\right)\right) + \sqrt{1 - p^2a^2\,(1 + 2\,\chi)}\right), \\ z(t) = z_0 + \dfrac{a}{b}\left(\dfrac{1}{p\,a\,\sqrt{1 + 2\,\chi}\,\cosh\left(\operatorname{arctanh}\left(\sqrt{1 - p^2a^2\,(1 + 2\,\chi)}\right) - b\,t\right)} - 1\right), \\ p_x(t) = p\,, \\ p_z(t) = \mathfrak{p}\sqrt{1 + 2\,\chi}\,\sinh\left(\operatorname{arctanh}\left(\sqrt{1 - p^2a^2\,(1 + 2\,\chi)}\right) - b\,t\right)\,. \end{cases} \tag{5.8}$$

The first two expressions within solution (5.8) can be used to provide a geometrical interpretation of the shape of the rays. Specifically, these expressions can be manipulated to obtain

$$\frac{\left((x - x_0) - \dfrac{\sqrt{1 - p^2a^2\,(1 + 2\,\chi)}}{p\,b}\right)^2}{\left(\dfrac{1}{pb}\right)^2} + \frac{\left((z - z_0) + \dfrac{a}{b}\right)^2}{\left(\dfrac{1}{p\,b\sqrt{1 + 2\,\chi}}\right)^2} = 1, \tag{5.9}$$

which is the equation of an ellipse. Solving ellipse (5.9) for the horizontal distance travelled with respect to depth (e.g., Slawinski, 2020a, expression (14.3.9)) yields

$$x - x_0 = \frac{\sqrt{1-A} - \sqrt{1-B}}{pb}, \tag{5.10}$$

where we define

$$A := p^2a^2(1+2\chi) \quad \text{and} \quad B := p^2(a+b(z-z_0))^2(1+2\chi).$$

For the traveltime along the ray, solving the first equation of solution (5.8) yields

$$t = \frac{\operatorname{arctanh}\left(p\,b\,(x-x_0) - \sqrt{1-A}\,\right) + \operatorname{arctanh}\left(\sqrt{1-A}\,\right)}{b}. \tag{5.11}$$

We can obtain an alternative expression to traveltime (5.11) by substituting expression (5.10) for the horizontal distance travelled. Since inverse hyperbolic tangent is odd, $\operatorname{arctanh}(-x) = y \implies \operatorname{arctanh}(x) = -y$, for $x \in (-1,1)$ and $y \in \mathbb{R}$, traveltime (5.11) is tantamount to

$$t = \frac{\operatorname{arctanh}\sqrt{1-A} - \operatorname{arctanh}\sqrt{1-B}}{b}. \tag{5.12}$$

Vertical rays

To compare the model traveltimes along vertical rays with the ZVSP data, we require traveltime expressions in both constant-velocity and $ab\chi$ media. Since vertical rays correspond to $p = 0$, expressions (5.2) and (5.3) for the constant-velocity medium simplify trivially to

$$x(z) = x_0 \quad \text{and} \quad t(z) = t_0 + \frac{z - z_0}{v_w}. \tag{5.13}$$

However, for $p = 0$ in the $ab\chi$ medium, the raytracing expression (5.10) leads to an indeterminate form, $\frac{0}{0}$ and the numerator traveltime expression (5.12) becomes undefined, $\infty - \infty$. Therefore, to evaluate these expressions, we use l'Hôpital's rule to evaluate their limits as $p \to 0$.

Now, let us address raytracing expression (5.10). To evaluate its limit, we require the

derivative of the numerator, which simplifies to

$$\frac{\partial}{\partial p}\left(\sqrt{1-A}-\sqrt{1-B}\right)=-p\left(\frac{a^2(1+2\chi)}{\sqrt{1-A}}-\frac{(a+b(z-z_0))^2(1+2\chi)}{\sqrt{1-B}}\right).$$

Then, by l'Hôpital's rule,

$$\lim_{p\to 0}(x-x_0)=\lim_{p\to 0}-\frac{p}{b}\left(\frac{a^2(1+2\chi)}{\sqrt{1-A}}-\frac{(a+b(z-z_0))^2(1+2\chi)}{\sqrt{1-B}}\right)=0. \tag{5.14}$$

As expected, $x=x_0$ for vertical rays in the $ab\chi$ medium.

For traveltime expression (5.12), we require the logarithmic form of the inverse hyperbolic tangent,

$$\operatorname{arctanh} x=\frac{1}{2}\ln\frac{1+x}{1-x}. \tag{5.15}$$

Using expression (5.15), and by the properties of logarithms, traveltime (5.12) can be written as

$$t=\frac{1}{2b}\left(\ln\frac{1+\sqrt{1-A}}{1+\sqrt{1-B}}+\ln\frac{1-\sqrt{1-B}}{1-\sqrt{1-A}}\right). \tag{5.16}$$

Since $\ln 1=0$, the limit of the first term is zero. Then, by properties of the limit, we obtain

$$\lim_{p\to 0}t=\lim_{p\to 0}\frac{1}{2b}\left(\ln\frac{1+\sqrt{1-A}}{1+\sqrt{1-B}}+\ln\frac{1-\sqrt{1-B}}{1-\sqrt{1-A}}\right)=\frac{1}{2b}\ln\left(\lim_{p\to 0}\frac{1-\sqrt{1-B}}{1-\sqrt{1-A}}\right).$$

Applying l'Hôpital's rule again,

$$\lim_{p\to 0}\frac{1-\sqrt{1-B}}{1-\sqrt{1-A}}=\lim_{p\to 0}\frac{-\frac{1}{2}(1-B)^{-1/2}(-2p(a+b(z-z_0))^2(1+2\chi))}{-\frac{1}{2}(1-A)^{-1/2}(-2pa^2(1+2\chi))}=\frac{(a+b(z-z_0))^2}{a^2}.$$

Finally, by properties of the logarithm,

$$t=\frac{1}{b}\ln\left(\frac{a+b(z-z_0)}{a}\right) \tag{5.17}$$

for vertical rays in the $ab\chi$ medium.

5.3.4 Multilayer raytracing

As discussed in Section 5.3.1, the VSP dataset is comprised of traveltimes that correspond to $j = 1, \ldots, M$ source-receiver combinations. To model these traveltimes, we consider the multilayer background model comprised of $i = 1, \ldots, N$ layers, which we established in Section 5.3.2.

In this section, we adapt the expressions of Section 5.3.3 to obtain the raytracing constraint for offset source-receiver combinations in a multilayer setting. Since the horizontal distance travelled by vertical rays is zero — as demonstrated in expressions (5.13) and (5.14) — the raytracing is satisfied trivially. However, for offset rays, we rely on numerical techniques, such as the Newton-Raphson method, to obtain an approximate solution for the raytracing constraint. In particular, we discuss the method's the convergence conditions as well as provide a strategy to improve its initial guess.

Raytracing constraint

Within the jth rayplane, we parameterize the horizontal distance travelled along the ray for a given depth. In view of the lateral homogeneity of the layer structure, the ray parameter, p_j, is constant along the jth ray. Therefore, for the ray to be traced from the jth source to receiver, we require the following constraint to be satisfied,

$$X_j(p_j) := \frac{p_j v_w z_{j1}}{\sqrt{1 - p_j^2 v_w^2}} + \frac{1}{p_j} \sum_{i=2}^{\ell_j} \frac{\sqrt{1 - A_{ji}} - \sqrt{1 - B_{ji}}}{b_i} - \operatorname{sgn}(\mathbf{d}_j \cdot \mathbf{e}_x) \|\boldsymbol{P}_j - \boldsymbol{S}_j\| = 0, \tag{5.18}$$

where

$$A_{ji} := A_{ji}(p_j) = p_j^2 a_i^2 (1 + 2\chi_i) \quad \text{and} \quad B_{ji} := B_{ji}(p_j) = p_j^2 (a_i + b_i z_{ji})^2 (1 + 2\chi_i) \tag{5.19}$$

are the variables that contain the anisotropy and inhomogeneity velocity parameters,

$$z_{ji} := \xi_i - \xi_{i-1} + \delta_{i(\ell_j)} \left(Z_R^{(j)} - \xi_i \right) - \delta_{i1} \left(Z_S^{(j)} - \xi_0 \right) \tag{5.20}$$

is the vertical distance travelled in each layer, and $\delta_{i(\ell_j)}$ and δ_{i1} are Kronecker deltas. With regard to constraint (5.18): the first term corresponds to the straight-line ray within the first layer; the second term corresponds to the summation of the elliptical

raypaths within the remaining $\ell_j - 1$ layers; and, the third term is the signed horizontal distance travelled within the rayplane.

To demonstrate the functionality of expression (5.20), let us return to Figure 5.1. Suppose that the jth receiver is in layer $\ell_j = 3$. Then, the respective vertical distances travelled by the ray in layers $i = 1, 2, 3$ are

$$z_{j1} = \xi_1 - Z_S^{(j)}, \quad z_{j2} = \xi_2 - \xi_1, \quad z_{j3} = Z_R^{(j)} - \xi_2.$$

For our consideration, the vertical P-wave velocity and its linear increase are strictly positive, and the anisotropy parameter is nonnegative, i.e.,

$$a_i > 0, \quad b_i > 0, \quad \text{and} \quad \chi_i \geq 0, \quad \text{for} \quad i = 1, \ldots, N. \tag{5.21}$$

We discuss the physical meaning of restrictions (5.21) in Section 5.4.3. The depths of the consecutive interfaces are strictly increasing, i.e., $\xi_i > \xi_{i-1}$. As such, $z_{ji} > 0$ and, consequently,

$$B_{ji} > A_{ji} > 0. \tag{5.22}$$

To obtain the ray parameter that satisfies raytracing constraint (5.18), we rely on the Newton-Raphson method for approximating the roots of the nonlinear equation. This iterative method produces a sequence of approximations, $p_j^{(k)}$, based on the linearization of $X_j(p_j^{(k)})$, where k is the iteration parameter. The method converges to a solution provided that $X_j(p_j)$ and $X_j'(p_j)$ are continuous, $X_j'(p_j) \neq 0$, and $p_j^{(0)}$ is sufficiently close to the solution. The formula for the $(k+1)$th approximation is

$$p_j^{(k+1)} = p_j^{(k)} - \frac{X_j(p_j^{(k)})}{X_j'(p_j^{(k)})}.$$

Newton-Raphson method: continuity conditions

Let us analyze the continuity of constraint (5.18), as well as its derivative, in regard to the convergence properties of the Newton-Raphson method. To analyze the continuity, we consider the domain for which $X_j(p_j) : \mathbb{R}^M \to \mathbb{R}^M$. For the function to return real

values, we have the following restrictions,

$$p_j \neq 0, \quad 1 - p_j^2 v_w^2 > 0, \quad 1 - A_{ji} \geq 0, \quad 1 - B_{ji} \geq 0.$$

The first restriction means that we do not consider vertical rays with raytracing constraint (5.18). The remaining restrictions impose bounds on p_j, which, by expressions (5.19), are

$$p_j^2 < \frac{1}{v_w^2}, \quad p_j^2 \leq \frac{1}{a_i^2(1+2\chi_i)}, \quad p_j^2 \leq \frac{1}{(a_i + b_i z_{ji})^2(1+2\chi_i)}.$$

In view of these inequalities, the maximum value of the jth ray parameter is

$$p_j^{\max} := \sqrt{\min\left(\left[\frac{1}{v_w^2}, \left\{\frac{1}{a_i^2(1+2\chi_i)}\right\}_{i=1}^{N}, \left\{\frac{1}{(a_i + b_i z_{ji})^2(1+2\chi_i)}\right\}_{i=1}^{N}\right]\right)},$$

the minimum is $p_j^{\min} := -p_j^{\max}$, and the domain within which the constraint is continuous is

$$p_j \in \left(p_j^{\min}, p_j^{\max}\right) \setminus \{0\}. \tag{5.23}$$

Within this domain, the requirement that a solution exists is

$$\frac{p_j^{\max} v_w z_{j1}}{\sqrt{1 - (p_j^{\max})^2 v_w^2}} + \frac{1}{p_j^{\max}} \sum_{i=2}^{\ell_j} \frac{\sqrt{1 - A_{ji}(p_j^{\max})} - \sqrt{1 - B_{ji}(p_j^{\max})}}{b_i} \geq \|\boldsymbol{P}_j - \boldsymbol{S}_j\|. \tag{5.24}$$

If this condition is not met, the Newton-Raphson method will continue to produce approximations $p_j^{(k)}$ without approaching a solution. For that reason, we include a stopping criterion that terminates the method after a specified number of iterations.

Now, let us consider the continuity of $X_j'(p_j)$. The first derivative of the first term in constraint (5.18) is

$$\frac{\partial}{\partial p_j}\left(\frac{p_j v_w z_{j1}}{\sqrt{1 - p_j^2 v_w^2}}\right) = v_w z_{j1}\left(\frac{1}{\sqrt{1 - p_j^2 v_w^2}} - \frac{p_j(-2p_j v_w^2)}{2\left(1 - p_j^2 v_w^2\right)^{3/2}}\right) = \frac{v_w z_{j1}}{\left(1 - p_j^2 v_w^2\right)^{3/2}}. \tag{5.25}$$

Prior to addressing the first derivative of the ith summand of the second term, let us

consider

$$\frac{\partial}{\partial p_j}\left(\sqrt{1-A_{ji}}-\sqrt{1-B_{ji}}\right)=\frac{-2p_ja_i^2(1+2\chi_i)}{2\sqrt{1-A_{ji}}}+\frac{2p_j(a_i+b_iz_{ji})^2(1+2\chi_i)}{2\sqrt{1-B_{ji}}}$$
$$=\frac{1}{p_j}\left(\frac{A_{ji}}{\sqrt{1-A_{ji}}}-\frac{A_{ji}}{\sqrt{1-B_{ji}}}\right).$$

Then, it can be shown that

$$\frac{\partial}{\partial p_j}\left(\frac{\sqrt{1-A_{ji}}-\sqrt{1-B_{ji}}}{p_jb_i}\right)=-\frac{1}{p_j^2b_i}\left(\frac{1}{\sqrt{1-A_{ji}}}-\frac{1}{\sqrt{1-B_{ji}}}\right). \tag{5.26}$$

As such, the first derivative of constraint (5.18) is

$$X_j'(p_j)=\frac{v_wz_{j1}}{\left(1-p_j^2v_w^2\right)^{3/2}}-\frac{1}{p_j^2}\sum_{i=2}^{\ell_j}\frac{1}{b_i}\left(\frac{1}{\sqrt{1-A_{ji}}}-\frac{1}{\sqrt{1-B_{ji}}}\right).$$

In view of inequalities (5.22) and ray parameter domain (5.23), the layer to which $p_j^{\max}$ pertains results in $1-B_{ji}\to 0$ as $p_j\to\pm p_j^{\max}$. As such,

$$\lim_{p_j\to p_j^{\max}}X_j'(p_j)=\infty\quad\text{and}\quad\lim_{p_j\to p_j^{\min}}X_j'(p_j)=\infty. \tag{5.27}$$

For the limit of $X_j'(p_j)$ as $p_j\to 0$, while the first term yields v_wz_{j1} trivially, the summation leads to an indeterminate form, $\frac{0}{0}$. To proceed with the limit, we must apply l'Hôpital's rule to the summation, for which we consider

$$\frac{\partial}{\partial p_j}\left(\frac{1}{\sqrt{1-A_{ji}}}-\frac{1}{\sqrt{1-B_{ji}}}\right)=p_j\left(\frac{a_i^2(1+2\chi_i)}{(1-A_{ji})^{3/2}}-\frac{(a_i+b_iz_{ji})^2(1+2\chi_i)}{(1-B_{ji})^{3/2}}\right).$$

Then, by l'Hôpital's rule,

$$\lim_{p_j\to 0}X_j'(p_j)=v_wz_{j1}-\lim_{p_j\to 0}\frac{\sum_{i=2}^{\ell_j}\frac{p_j(1+2\chi_i)}{b_i}\left(\frac{a_i^2}{(1-A_{ji})^{3/2}}-\frac{(a_i+b_iz_{ji})^2}{(1-B_{ji})^{3/2}}\right)}{2p_j}.$$

However, by expressions (5.19), $\lim_{p_j \to 0} A_{ji} = \lim_{p_j \to 0} B_{ji} = 0$ and, hence,

$$\lim_{p_j \to 0} X_j'(p_j) = v_w z_{j1} + \frac{1}{2} \sum_{i=2}^{\ell_j} \frac{(1 + 2\chi_i)}{b_i} \left((a_i + b_i z_{ji})^2 - a_i^2 \right).$$

Since the inhomogeneity parameters are strictly positive and the anisotropy parameter is nonnegative, as detailed by inequalities (5.21), we find that

$$\lim_{p_j \to 0} X_j'(p_j) > 0. \tag{5.28}$$

Therefore, in view of inequalities (5.27) and (5.28), $X_j(p_j)$ does not have extrema as it is increasing throughout the domain. Through this analysis, we find that, indeed, $X_j(p_j)$ and $X_j'(p_j)$ are continuous on the ray parameter domain as well as $X_j'(p_j) \neq 0$.

Finally, let us also consider the concavity of $X_j(p_j)$, for which we require its second derivative. Using expression (5.25), the second derivative of the first term in constraint (5.18) is

$$\frac{\partial^2}{\partial p_j^2} \left(\frac{p_j v_w z_{j1}}{\sqrt{1 - p_j^2 v_w^2}} \right) = \frac{p_j v_w^3 z_{j1}}{\left(1 - p_j^2 v_w^2\right)^{5/2}}. \tag{5.29}$$

Using expression (5.26), it can be shown that the second derivative of the summand in constraint (5.18) is

$$\frac{\partial^2}{\partial p_j^2} \left(\frac{\sqrt{1 - A_{ji}} - \sqrt{1 - B_{ji}}}{p_j b_i} \right) = -\frac{1}{p_j^3 b_i} \left(\frac{2 - 3A_{ji}}{(1 - A_{ji})^{3/2}} - \frac{2 - 3B_{ji}}{(1 - B_{ji})^{3/2}} \right). \tag{5.30}$$

Hence, using derivatives (5.29) and (5.30), the second derivative of constraint (5.18) is

$$X_j''(p_j) = \frac{p_j v_w^3 z_{j1}}{\left(1 - p_j^2 v_w^2\right)^{5/2}} + \frac{1}{p_j^3} \sum_{i=2}^{\ell_j} \frac{1}{b_i} \left(\frac{2 - 3A_{ji}}{(1 - A_{ji})^{3/2}} - \frac{2 - 3B_{ji}}{(1 - B_{ji})^{3/2}} \right). \tag{5.31}$$

Since v_w and z_{j1} are strictly positive, the sign of the first term in expression (5.31) is $\operatorname{sgn}(p_j)$. For the second term, since $B_{ji} > A_{ji} > 0$ as a consequence of inequalities (5.21), let us suppose that $A_{ji} = u$ and $B_{ji} = \alpha u$, where $u > 0$ and

$\alpha > 0$. Then, the difference of terms in the ith summand can be simplified to

$$\frac{2-3u}{(1-u)^{3/2}} - \frac{2-3\alpha u}{(1-\alpha u)^{3/2}} = \frac{(2-3u)(1-\alpha u)^{3/2} - (2-3\alpha u)(1-u)^{3/2}}{(1-u)^{3/2}(1-\alpha u)^{3/2}},$$

which has restrictions $0 < u < 1$ and $0 < \alpha u < 1$. However, since $\alpha > 1$, $\alpha u > u$ and, so, $0 < \alpha u < 1$ is the dominating restriction. Let us consider the limits at either end of this restriction. On the left, as $\alpha u \to 0$ so does $x \to 0$, and so we have

$$\lim_{\alpha u \to 0} \frac{(2-3u)(1-\alpha u)^{3/2} - (2-3\alpha u)(1-u)^{3/2}}{(1-u)^{3/2}(1-\alpha u)^{3/2}} = 0.$$

On the right,

$$\lim_{\alpha u \to 1} \frac{(2-3u)(1-\alpha u)^{3/2} - (2-3\alpha u)(1-u)^{3/2}}{(1-u)^{3/2}(1-\alpha u)^{3/2}} = \infty.$$

By these limits, we find that the difference in the ith summand is nonnegative, which means the entire summation is nonnegative. Hence, the reciprocal of the ray parameter that scales the summation determines the sign, i.e., $\text{sgn}(p_j^3)$. As such, we find that $X_j(p_j)$ is concave up for $0 < p_j < p_j^{\max}$ and concave down for $p_j^{\min} < p_j < 0$, with the inflection point at $p_j \to 0$.

Newton-Raphson method: initial guess

The Newton-Raphson method requires that the initial ray parameter value, $p_j^{(0)}$, is sufficiently close to the solution. Since raytracing constraint (5.18) does not have extrema, choosing any $p_j^{(0)}$ that is in the domain will lead to convergence. Be that as it may, the features of $X_j(p_j)$ resemble that of the inverse hyperbolic tangent function. To demonstrate, let $f(x) = \text{arctanh}(x)$. Rearranging and implicitly differentiating, we obtain

$$\tanh(f(x)) = x \implies \text{sech}^2(f(x))f'(x) = 1.$$

Solving for $f'(x)$, and simplifying using standard trigonometric identities, we obtain

$$f'(x) = \frac{1}{1-\text{sech}^2(f(x))} = \frac{1}{1-\tanh^2(\text{arctanh}(x))} = \frac{1}{1-x^2}.$$

In an analogous manner to $X_j'(p_j) > 0$ for all $p_j \in (p_j^{\min}, p_j^{\max}) \setminus \{0\}$, we find that $f'(x) > 0$ for all $x \in (-1, 1)$. Regarding concavity, we find

$$f''(x) = \frac{2x}{(1-x^2)^2},$$

which is implies that $f(x)$ is concave up for $x \in (0, 1)$ and concave down for $x \in (-1, 0)$; this is analogous to the concavity of $X_j(p_j)$ within the two branches of its domain. As such, we can use $f(x)$ to approximate $X_j(p_j)$, which improves the quality of the initial guess for the Newton-Raphson method.

The approximating procedure is as follows. Recalling expression (5.15), the logarithmic form of the inverse hyperbolic tangent is

$$f(x) = \operatorname{arctanh}(x) = \frac{1}{2} \ln \frac{1+x}{1-x}.$$

Now, consider its algebraic extension

$$f(x) = \frac{1}{2} \log_p \frac{t+x}{t-x} + h, \tag{5.32}$$

where p corresponds to an arbitrary base, $x = \pm t$ are the vertical asymptotes, and h is a vertical translation. To apply $f(x)$ to $X_j(p_j)$, we center the former upon the latter by setting $h = -\operatorname{sgn}(\mathbf{d}_j \cdot \mathbf{e}_x) \| \boldsymbol{P}_j - \boldsymbol{S}_j \|$, which is the signed horizontal distance between the source and receiver within the jth rayplane. We match the domains of the two functions by setting $t = p_j^{\max}$. For the base, we require that $f(x)$ passes through coordinates (q, Q), where $Q = X_j(q)$. Substituting these coordinates in expression (5.32) and rearranging, we obtain

$$2(Q-h) = \log_p \frac{t+q}{t-q}.$$

Converting to exponential form and then exponentiating both sides with respect to $\frac{1}{2(Q-h)}$, we obtain

$$p = \left(\frac{t+q}{t-q} \right)^{\frac{1}{2(Q-h)}}. \tag{5.33}$$

Using base (5.33), we can calculate the solution to $f(x^*) = 0$. Through algebraic

manipulation, the solution is

$$x^* = t\left(\frac{p^{-2h}-1}{p^{-2h}+1}\right). \tag{5.34}$$

Since $f(x)$ is constructed to approximate $X_j(p_j)$, we use solution (5.34) as an initial guess to solve $X_j(p_j) = 0$ with the Newton-Raphson method.

5.3.5 Multilayer traveltimes

Let us adapt the expressions of Section 5.3.3 to calculate the traveltime along a ray in a multilayer setting. The traveltime expressions depend on whether the source-receiver combination is vertical or offset. In both cases, we assume that the jth receiver is within in the $ab\chi$ layers, not the water column, i.e., $\ell_j \geq 2$.

If the jth source-receiver combination is vertical, the raytracing is satisfied trivially. As such, the traveltime along the vertical ray is the sum of traveltimes in the constant velocity layer and the remaining $(\ell_j - 1)$ $ab\chi$ layers, which are given by expressions (5.13) and (5.17), respectively. Hence, the jth traveltime along a vertical ray is

$$t_j^v = \frac{z_{j1}}{v_w} + \sum_{i=2}^{\ell_j} \frac{1}{b_i} \ln\left(\frac{a_i + b_i z_{ji}}{a_i}\right). \tag{5.35}$$

If the jth source-receiver combination is offset, we require that raytracing constraint (5.18) is satisfied. We use the Newton-Raphson method to approximate the ray parameter, p_j, that satisfies $X_j(p_j) = 0$. Then, the traveltime along the offset ray is the sum of traveltimes in the constant velocity layer and the remaining $(\ell_j - 1)$ $ab\chi$ layers, which are given by expressions (5.3) and (5.12), respectively. Hence, the jth traveltime along an offset ray is

$$t_j^o = \frac{z_{j1}}{v_w\sqrt{1-p_j^2 v_w^2}} + \sum_{i=2}^{\ell_j} \frac{\operatorname{arctanh}\sqrt{1-A_{ji}} - \operatorname{arctanh}\sqrt{1-B_{ji}}}{b_i}. \tag{5.36}$$

5.4 Optimization procedure

In this section, we outline the optimization procedure for the inversion of VSP traveltimes. To achieve the inversion, we use numerical techniques to estimate the

model-parameter values such that the difference between measured and modelled traveltimes is at a minimum.

5.4.1 Residual sum of squares

Within an N-layer background model, comprised of a constant-velocity layer followed by $N-1$ $ab\chi$ layers, there are $2(2N-1)$ model parameters: one constant velocity, v_w; $3(N-1)$ linear inhomogeneity and elliptical anisotropy parameters, $\{a_i\}_{i=1}^{N-1}$, $\{b_i\}_{i=1}^{N-1}$, $\{\chi_i\}_{i=1}^{N-1}$; and N interface depths, $\{\xi_i\}_{i=0}^{N-1}$. Of these parameters, we fix the constant velocity as $v_w = 1500\,\mathrm{m/s}$, which corresponds to the speed of sound in water, as well as interface depths $\xi_0 = 0\,\mathrm{m}$ and $\xi_1 = 130.8\,\mathrm{m}$, the latter of which corresponds to the depth of the sea floor. Thus, we consider the remaining $4N-5$ model parameters,

$$\boldsymbol{\beta} := [\{a_i\}_{i=1}^{N-1}, \{b_i\}_{i=1}^{N-1}, \{\chi_i\}_{i=1}^{N-1}, \{\xi\}_{i=2}^{N-1}]^T, \tag{5.37}$$

as variables to be estimated within the optimization.

For the purpose of obtaining an agreement between the measured and modelled traveltimes, we seek a set of optimal model-parameter values, $\boldsymbol{\beta}^*$, that yield a minimum residual sum of squares, which we define as

$$\mathrm{RSS}(\boldsymbol{\beta}) := \sum_{j=1}^{M} r_j(\boldsymbol{\beta})^2, \tag{5.38}$$

which is a scalar $\mathrm{RSS} : \mathbb{R}^{4N-5} \to \mathbb{R}^1$. Herein, for the jth source-receiver combination, we use the measured VSP traveltime, T_j, and modelled traveltime, t_j, to the determine the jth residual,

$$r_j(\boldsymbol{\beta}) = T_j - t_j(\boldsymbol{\beta}), \tag{5.39}$$

where $t_j(\boldsymbol{\beta})$ corresponds to t_j^v for vertical rays or t_j^o for offset rays. From the point of view of optimization theory (Nocedal and Wright, 2006, p. 12), the optimal solution is said to be a global minimizer if

$$\mathrm{RSS}(\boldsymbol{\beta}^*) \leq \mathrm{RSS}(\boldsymbol{\beta}), \quad \text{for all} \quad \boldsymbol{\beta} \in \mathbb{R}^{4N-5}.$$

In general, due to the nonlinearity of RSS and the possible restrictions to which it might be subjected, obtaining $\boldsymbol{\beta}^*$ is a difficult task. However, within a smaller region

of the domain, $\boldsymbol{\beta}^*$ is said to be a local minimizer if

$$\text{RSS}(\boldsymbol{\beta}^*) \leq \text{RSS}(\boldsymbol{\beta}), \quad \text{for} \quad \boldsymbol{\beta} \in \mathcal{N}, \tag{5.40}$$

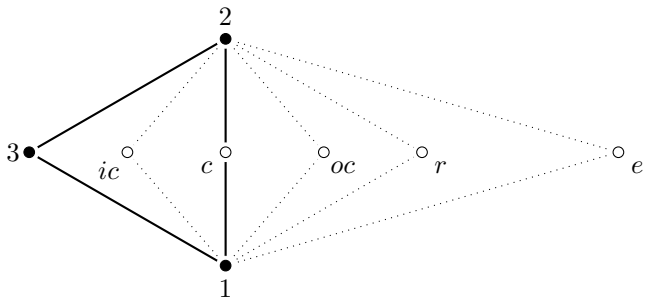
where $\mathcal{N}$ is an open set that we denote as a neighbourhood and for which $\boldsymbol{\beta}^* \in \mathcal{N} \subset \mathbb{R}^{4N-5}$. The point is classified as a strict local minimizer if the inequality is strictly $\text{RSS}(\boldsymbol{\beta}^*) < \text{RSS}(\boldsymbol{\beta})$ within the neighbourhood.

5.4.2 Nelder-Mead algorithm

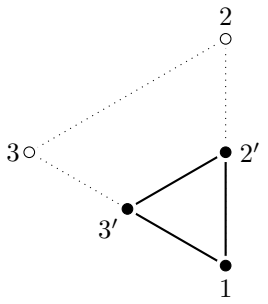
To obtain $\boldsymbol{\beta}^*$, we proceed with a derivative-free optimization approach, which is useful in scenarios where obtaining derivative information could be prohibitive. In our case, obtaining derivative information is prohibitive due to the complicated terms in the raytracing constraint for offset source-receiver combinations. In particular, taking first- and second-order derivatives of the inverse hyperbolic tangent function with radical arguments leads to increasingly complicated expressions. Likewise, in view of the multilayer setting, these expressions are repeated for each layer with different a_i, b_i, and χ_i values, which results in cumbersome simplifications. Moreover, even to evaluate the traveltime expression, we require to solve a subproblem to obtain the ray parameter value, p_j, that specifies the raypath for each source-receiver combination. In view of these considerations, we choose to proceed with derivative-free optimization.

The method of our choice is the Nelder-Mead algorithm (Nelder and Mead, 1965), which is a simplex approach that minimizes the residual sum of squares using only function evaluations. In particular, this algorithm is said to minimize a function of n variables by comparing the function values at the $n + 1$ vertices of a simplex. Then, within an iterative structure, the vertex with the highest function value is replaced using one of four mechanisms: reflect, expand, contract (inside or outside), and shrink. In Figure 5.2, we illustrate these mechanisms for a function of $n = 2$ variables.

The Nelder-Mead algorithm provides local convergence only. To ensure adequate coverage, we use a multistart strategy, which means that we repeat the optimization with different initial simplexes whose values are randomly sampled within the domain of the model parameters. Following the formalism of expression (5.40), we denote the final solution as the $\boldsymbol{\beta}^*$ that has the lowest residual sum of squares out of the set of initial simplexes considered.



(a) inside contract (ic), centroid (c), outside contract (oc), reflect (r), expand (e)



(b) Shrink

Figure 5.2: Mechanisms of the Nelder-Mead algorithm for a function of $n = 2$ variables (adapted from Conn et al., 2009, Figures 8.1 and 8.2).

To reduce the residual sum of squares using the $ab\chi$ traveltime model, we follow the algorithm specifications set by Conn et al. (2009) and Gao and Han (2012). For convenience of the readership, we provide these specifications as Algorithm 1, which is presented in Appendix A.3.

5.4.3 Implementation considerations

Let us discuss several aspects of the algorithm and its implementation. In particular, we consider the choice of scaling parameters, termination conditions, and penalization constraints.

Adaptive scaling parameters

As indicated by Conn et al. (2009, p. 142), Algorithm 1 is the standard implementation of the Nelder-Mead algorithm for modern applications. However, there are modifications that could be made to improve upon its results. In particular, Wright (1996, p. 6) indicates that, in higher dimensional settings, the Nelder-Mead method can take endless iterations while making negligible progress, despite being nowhere close to a minimizer. Gao and Han (2012) refer to this as the effect of dimensionality and propose the so-called adaptive approach wherein, as opposed to the standard scaling constants (A.22), the parameters

$$\alpha = 1, \quad \beta = 1 + \frac{2}{n}, \quad \gamma = \frac{3}{4} - \frac{1}{2n}, \quad \delta = 1 - \frac{1}{n}, \tag{5.41}$$

are a function of the number of variables (Gao and Han, 2012, expression (4.1)) in the optimization problem. This adjustment keeps the simplex open for longer as it diminishes the effect of the contraction and shrink scaling constants for increasing n, thus resulting in more iterations before termination. Note that these parameters reduce to the standard scaling constants for $n = 2$.

Termination conditions

The termination conditions of this algorithm are threefold. The first condition is that we terminate if the absolute difference between the highest and lowest function value of the simplex is less than a user-specified tolerance. Since we perform our computations in MATLAB, we set the tolerance to the system floating-point accuracy

$$\text{eps} = 2^{-52} \approx 2.2204 \times 10^{-16}.$$

In this case, the vertices of any simplex that satisfy this termination condition have effectively the same function value, which justifies termination. The second condition compares the value of the first condition for successive iterations. We terminate if the absolute difference of the value between the current and previous iteration is less than the system floating-point accuracy. In this case, the simplex is effectively unchanged on successive iterations and, as such, the reduction of the function value has stalled, which justifies termination. The third condition is that we terminate if the iteration count of the optimization exceeds a user-specified number.

Note that, in most settings, the tolerances of the first two conditions could be relaxed. However, we impose strict tolerances for two reasons. From the perspective of the $ab\chi$ model, minute changes in the inhomogeneity, b, and anisotropy, χ, parameters can result in significant changes in the function values. The effect is compounded at greater offsets as each ray traces out a longer path in each layer. Since our model parameters correspond to the anisotropy and inhomogeneity of an entire layer, minute changes can have a greater effect on the ray than in grid-based schemes, which would amount to perturbing a single cell only.

Parameter constraints

To ensure physically meaningful optimization results, we have several model-parameter constraints. Specifically, following inequalities (5.21), we require

that the inhomogeneity parameters, a_i and b_i, are strictly positive and the anisotropy parameter, χ_i is nonnegative. Note that strict positivity pertains to a vertical increase in seismic velocity as a result of compaction due to overburden whereas the nonnegativity pertains to the anisotropic properties of shale. Also, $b_i \neq 0$ as it is a non-permissible value of traveltime expressions (5.35) and (5.36).

Along with the velocity-parameter constraints, we consider several restrictions that are a consequence of the ray theory validity conditions, which are usually are only of a qualitative, not quantitative character (Červený, 2001, Section 5.9). Among these conditions, it is said that the wavelength of the wave under consideration must be considerably smaller than any characteristic quantity. These characteristic quantities correspond to the radii of curvature of the interfaces, the scale lengths of the inhomogeneity of the medium, and thicknesses of layers. For the curvature, since our interfaces are planar, the radii of curvature are infinite, which do not impose a restriction. For the inhomogeneity, while we do not attempt to quantify the limiting scale length, we impose maximum values on the ith layer velocity gradient, b_i, and anisotropy, χ_i — we specify these maximum values in Sections 5.5.2. For the thicknesses, we impose a minimum layer thickness of 300 m, which corresponds to approximately 15% of the total depth of the VSP.

Now, let us acknowledge that the Nelder-Mead algorithm is unconstrained and, as it only considers function values, it does not have the requisite mechanisms to impose parameter constraints or ensure model feasibility. As such, we have devised the following penalization strategies.

For parameter constraints, we consider penalties that are defined in a piecewise form

$$p(x) = \begin{cases} -(x - p_\ell), & x < p_\ell \\ 0, & p_\ell \leq x \leq p_u \\ x - p_u, & x > p_u \end{cases}, \tag{5.42}$$

where, for a given model parameter, x, the penalty is zero within the user-specified domain $x \in [p_\ell, p_u]$, but returns linearly increasing values beyond that domain. Herein, p_ℓ is the lower bound on x and p_u is the upper bound, as illustrated in Figure 5.3.

In practice, within a MATLAB environment, we express the piecewise penalties as

$$p(x) = (p_\ell - x) \cdot (x < p_\ell) + (x - p_u) \cdot (p_u < x), \tag{5.43}$$

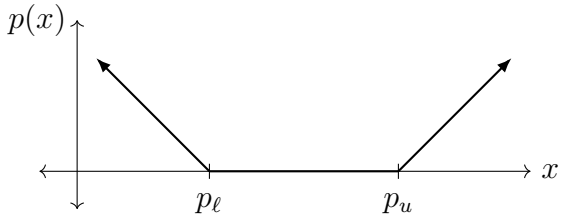


Figure 5.3: Visualization of piecewise penalty function (5.42) for arbitrary lower, p_ℓ, and upper, p_u, bounds on model parameter x

We apply the penalties to every inhomogeneity and anisotropy parameter, and we add the sum of the parameter penalties to the residual sum of squares.

For model feasibility, our model is feasible if there exists a ray-parameter value, p_j, for every source-receiver combination that satisfies raytracing expression (5.18). In other words, we consider any such combination that satisfies inequality (5.24) as a traceable combination and, otherwise, it is considered untraceable. At every iteration of the optimization, we identify the number of untraceable combinations and add the sum of the inequality errors to the residual sum of squares. We provide further details in Section 5.5.2.

5.5 Simulation study

The motivation of the simulation study to assess the $ab\chi$ model's accuracy of parameter estimation using the Nelder-Mead optimization procedure discussed in Section 5.4. We outline the procedure in the following three steps.

First, we consider three velocity profiles of the subsurface, whose parameter values are known. We generate simulated traveltimes based on these profiles as well as three sets of simulated measurement errors, otherwise referred to as additive noise. The nine datasets serve as the input data for the simulation study. We provide details of the velocity profiles, their traveltimes, and noise in Section 5.5.1.

Second, we apply the optimization procedure to each of the datasets. Given the increased dimensionality of the of the traveltime model, there are many combinations of model-parameter values that yield similar traveltime residuals. We counter this nonuniqueness by performing two passes of optimization. On the first pass, we use randomly generated initial simplexes that are sampled from a large domain for each

parameter, which serves to identify the regions of the parameter space that minimize the traveltime residual. On the second pass, we generate new initial simplexes based on the identified regions of the first pass and repeat the optimization, hence, improving upon the quality of our parameter estimate of the first pass. For each pass, we use a multistart strategy to improve the reliability of our optimization results through redundancy. We discuss details of this implementation strategy in Section 5.5.2.

Third, we compare the optimization results to the known parameter values. By taking into account the optimization outputs from both passes, we obtain an improved estimate of the optimal model-parameter values for each dataset. Also, we acquire insight into the extent to which the model parameters may vary for those combinations that are near the minimum traveltime residual. We illustrate these results in Section 5.5.3.

5.5.1 Simulated data

Let us discuss the considerations that have been made for the generation of simulated traveltime data. Following the VSP acquisition specifications of Section 5.3.1, we require traveltimes that resemble those in the walkaway VSP, with respect to source-receiver combinations. For this purpose, we use the multilayer $ab\chi$ model of Section 5.3.5 to generate the simulated traveltimes with known model-parameter values. Likewise, we generate simulated noise profiles that are added to the simulated traveltimes in order to simulate measurement errors obtained in the VSP acquisition.

Simulated model-parameter values

In view of the macro layers of Zhou and Kaderali (2006, Figure 1(b)), we choose a $(N = 4)$-layer background model, which, by expression (5.37), yields $4N - 5 = 11$ model parameters to be estimated. To assess the optimization in various scenarios, we consider three sets of simulated model parameters, which we present in Table 5.1. Therein, model A constitutes a continuous vertical velocity profile, with increasing homogeneity with depth and low, constant anisotropy throughout. Model B constitutes a discontinuous vertical velocity profile, with varying inhomogeneity and increasing anisotropy with depth. Model C has a discontinuous vertical velocity profile, but with constant inhomogeneity and a buried anisotropic layer.

	Simulated model-parameter values		
$\boldsymbol{\beta}$	A	B	C
a_1	1800	1900	2000
a_2	2478	2300	2300
a_3	2728	2500	2600
b_1	0.7	0.5	0.3
b_2	0.5	1.25	0.3
b_3	0.3	0.75	0.3
χ_1	0.01	0.05	0.01
χ_2	0.01	0.1	0.1
χ_3	0.01	0.15	0.01
ξ_2	1100	1000	800
ξ_3	1600	1500	1400

Table 5.1: Simulated model-parameter values for a ($N = 4$)-layer model. Each column represents a set of parameters that are substituted in expression (5.37), to which we refer as model A, B, or C. Along with each set, we fix $v_w = 1500\,\mathrm{m/s}$, $\xi_0 = 0\,\mathrm{m}$, and $\xi_1 = 130.8\,\mathrm{m}$. Note: parameter (units): a (m/s); b (s^{-1}); χ (dimensionless); ξ (m).

Noise

To generate simulated noise, we consider a noise range up to a fixed percentage of the measured traveltimes. In particular, since the traveltimes are approximately two seconds at the maximum acquisition offset, we consider noise up to 0.1%, which results in a traveltime error on the order of milliseconds (e.g., Slawinski et al., 2003). Likewise, as source offset grows, the error in traveltime increases (e.g., Brown et al., 2000). In view of these considerations, we calculate $\pm 0.1\%$ of the simulated traveltimes at each source-receiver combination to set the boundary of the noise values. Then, within these boundaries, we draw samples from the uniform distribution using MATLAB's `rand` function. Applying this procedure yields WVSP traveltime errors for all five receivers. We apply the same procedure to the ZVSP traveltimes.

For the simulation study, we repeat the process three times, resulting in the three noise profiles presented in Figures A.4–A.7 in Appendix A.4.1.

5.5.2 Implementation

Within a ($N = 4$)-layer model, there are eleven model parameters to estimate, which results in many combinations of parameter values that produce similar traveltimes.

We contend with the issue of nonuniqueness by performing two passes of optimization. For each pass, we use a multistart strategy to improve the reliability of our optimization results through redundancy.

First pass

On the first pass, we generate random initial simplexes that are sampled from a large domain of each parameter using the uniform distribution. As indicated by Gao and Han (2012, Section 3), starting with larger initial simplexes can help improve the performance of the Nelder-Mead method in high dimensions. To that end, we consider 250 random initial simplexes with the following specifications. For the vertical velocity at the top of the ith layer, we choose

$$a_i \sim \mathcal{U}(1800\,\mathrm{m/s}, 3000\,\mathrm{m/s}) \quad \text{such that} \quad a_{i-1} < a_i,$$

which means that the velocities increase with depth. To ensure a wide range of inhomogeneity and anisotropy values, we choose

$$b_i \sim \mathcal{U}(0.1\,\mathrm{s}^{-1}, 1.5\,\mathrm{s}^{-1}) \quad \text{and} \quad \chi_i \sim \mathcal{U}(0.01, 0.3).$$

Also, we consider interface depths that are strictly increasing with depth,

$$\xi_i \sim \mathcal{U}(500\,\mathrm{m}, 1900\,\mathrm{m}) \quad \text{such that} \quad \xi_{i-1} < \xi_i.$$

This method of initial simplex generation promotes random combinations of model parameters, which serves to sample the parameter space adequately. As such, the initial residual sums of squares are guaranteed to vary and, hence, promote large initial simplexes.

During the first pass, we use parameter constraints to keep the model-parameter values within specified domains using penalty functions. In practice, we add a penalty term for each model parameter to the residual sum of squares, such as penalty function (5.43). The bounds for the velocity parameters, (p_ℓ, p_u), are

$$a_i \in (1600\,\mathrm{m/s}, 4000\,\mathrm{m/s}), \quad b_i \in (10^{-3}\,\mathrm{s}^{-1}, 2\,\mathrm{s}^{-1}), \quad \chi_i \in (10^{-3},\, 0.5). \tag{5.44}$$

However, for the interface depths, we adjust the penalty function such that

$$p_\xi(\{\xi_2,\ldots,\xi_{N-1}\}) = (500-\xi_2)\cdot(\xi_2<500) + (\xi_{N-1}-1900)\cdot(\xi_{N-1}>1900) + \sum_{i=3}^{N-1}(\xi_{i-1}+300-\xi_i)\cdot(\xi_i<\xi_{i-1}+300), \tag{5.45}$$

which means that a penalty is applied if ξ_2 is less than 500m, ξ_{N-1} is greater than 1900m, or if subsequent interface depths are less than 300 m from one another. Through the mechanisms of the Nelder-Mead algorithm, any adjusted vertex that takes on parameter values beyond the bounds receives a linearly increasing penalty. A visualization of this penalization for arbitrary bounds is presented in Figure 5.3.

Along with the model-parameter constraints, we require that each vertex results in a traceable raypath for every source-receiver combination. However, given its unconstrained nature, it is possible for the Nelder-Mead algorithm to produce combinations of $ab\chi$ values that do not satisfy inequality (5.24). To ensure model feasibility, we flag any unfeasible combination and set its ray parameter value to $p_j^{\max}$. Then, we add the amount by which inequality (5.24) is not satisfied to the RSS for each untraceable combination, which adversely affects the overall RSS.

Let us emphasize that these optimization strategies have been devised to take effect only if the parameter constraints are broken and/or model feasibility is not ensured. Thus, the total quantity that the Nelder-Mead algorithm seeks to minimize is

$$\text{(residual sum of squares)}+\sum\text{(parameter penalties)}+\sum\text{(untraceable combinations)}. \tag{5.46}$$

In this manner, we maintain the use of the unconstrained Nelder-Mead algorithm while accounting for model-parameter restrictions and ray-tracing feasibility through the use of penalization strategies.

Second pass

At the end of the first pass of optimization, we sort the output parameters in terms of RSS, identify those outputs that are within 1% of the minimum, and compute the lower and upper bounds of the parameter values. For the second pass of optimization, we generate another 250 random initial simplexes, but within these bounds. Likewise, we adjust the velocity bounds (5.44) and interface bounds (5.45) to correspond to

these minimum and maximum parameter values. Then, we redeploy the Nelder-Mead algorithm to minimize expression (5.46) within this narrower parameter domain.

5.5.3 Results

Let us apply the optimization procedure to the noisy traveltimes. We repeat the process for the three sets of model parameters and the three noise profiles. For consistency of comparison, we use the same 250 initial simplexes for the multistart strategy within the first pass of optimization. However, we use the bounds obtained from the first pass determine the 250 initial simplexes for the second pass. Due to the similarity of results, we discuss only those results that correspond to noise profile #1. However, we include the results for noise profiles #2 and #3 in Appendix A.4.2.

Regarding the number of iterations, we find that the Nelder-Mead algorithm terminates typically within 6 000 iterations on the first pass and 4 000 iterations on the second pass. The specific number of iterations varies among the nine datasets and can be reduced by relaxing the strict termination conditions, which we discuss in Section 5.4.3. To reduce overall computation time, we perform the optimizations using MATLAB's Parallel Computing Toolbox.

Broadly speaking, the optimization procedure yields accurate estimates of the true model-parameter values. To quantify this statement, we turn our attention to the optimization results presented in Table 5.2.

For the three velocity models, we find that the bounds of the first-pass-optimization outputs — that have RSS values within 1% of the minimum RSS — can vary. This is an expected result as the initial simplexes for the first pass are generated for the purpose of sampling the parameter space. We find that the first pass is successful in identifying suitable parameter intervals that correspond to the true parameters.

On the second pass of optimization, using initial simplexes generated from the first-pass bounds, we find that the estimates are accurate. While the optimal model parameters, $\boldsymbol{\beta}^*$, vary with respect to noise profile, the sample mean, μ, of the second-pass outputs is stable for each of the velocity models. Furthermore, the optimal and true model parameters are within one (or nearly one) sample standard deviation, σ, from their respective means. The stability of results, with respect to different velocity models, suggest that the optimization procedure yields reliable and precise model-

parameter estimates. This conclusion is also borne out for noise profiles #2 and #3 in Tables A.1 and A.2, respectively, wherein the same observations can be made.

Let us acknowledge that some parameters are estimated better than others. We observe that the most-variable estimates pertain to velocity models with low inhomogeneity and/or anisotropy. This is an expected result as smaller-valued b and χ parameters have less of an effect on the traveltime within a given layer and, as such, such parameters would be difficult to estimate within a traveltime optimization. However, for those layers with larger inhomogeneities and anisotropies, their effect on traveltime is more significant and, consequently, these parameters are estimated with less variance.

We quantify the variation using the coefficient of variation, σ/μ, which is said to be the most widely used measure of the extent of trait variation (Botta-Dukát, 2023). We use this relative statistic to assess the estimate of a given parameter across different noise profiles for a specified velocity model, or vice versa. Across the three models, the vertical speeds, a_i, have the lowest coefficients of variance, which suggests that the estimates are suitable. Also, we find that the coefficients — across all parameters — are lowest for velocity model B, which has the greatest inhomogeneity and anisotropy parameters, as expected.

However, we find that the model parameters that pertain to the middle layer for all three velocity models have some of the highest variance in comparison to the other layers. We suspect this variance is due to the scarcity of ZVSP data up to the depth of 1343.40 m, as is discussed in Section 5.3.1. Likewise, we suspect that the scarcity of data contributes to the higher variance for ξ_2 as opposed to ξ_3 — across all velocity models — because the former has been set to depths that are shallower than 1343.40 m, whereas the latter is below that depth. Also, we find that the variance is high for layers with low inhomogeneity and anisotropy values, as expected.

Finally, we can assess the quality of the parameter estimates by the plot of traveltime residuals. In a residual plot, we should observe a random scatter of residuals around zero, without an observable trend in the scatter of the residuals. Such a trend usually indicates that something is missing in the model structure (Wood, 2006, Section 1.1.2).

Turning our attention to Figure 5.4, we present such a plot for the simulated traveltimes corresponding to model A with noise profile #1. We calculate the residuals for the specified source-receiver combinations using the optimal parameters

Model A		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1800	1750	1911	1804	1808	19	0.0103
a_2	2478	2073	2642	2543	2346	143	0.0608
a_3	2728	2468	2870	2703	2663	64	0.0240
b_1	0.7	0.0496	1.0072	0.6908	0.6664	0.1118	0.1677
b_2	0.5	0.0184	1.3232	0.8599	0.5818	0.2178	0.3743
b_3	0.3	0.0014	0.9486	0.4302	0.3898	0.0974	0.2498
χ_1	0.01	0.0010	0.0298	0.0091	0.0082	0.0053	0.6583
χ_2	0.01	0.0015	0.0547	0.0228	0.0181	0.0091	0.5050
χ_3	0.01	0.0010	0.0104	0.0030	0.0063	0.0024	0.3704
ξ_2	1100	500	1393	1315	910	224	0.2465
ξ_3	1600	996	1900	1615	1471	183	0.1242
Model B		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1900	1872	1914	1902	1901	4	0.0020
a_2	2300	2103	2746	2551	2387	158	0.0664
a_3	2500	2457	2532	2489	2489	9	0.0038
b_1	0.5	0.4550	0.6338	0.4969	0.5016	0.0151	0.0300
b_2	1.25	0.0010	1.7811	0.9608	1.0735	0.2285	0.2128
b_3	0.75	0.6707	0.9932	0.8181	0.8106	0.0347	0.0428
χ_1	0.05	0.0235	0.0631	0.0573	0.0504	0.0072	0.1437
χ_2	0.1	0.0867	0.1315	0.1162	0.1119	0.0060	0.0539
χ_3	0.15	0.1280	0.1557	0.1422	0.1429	0.0033	0.0231
ξ_2	1000	752	1226	1161	1043	112	0.1070
ξ_3	1500	1472	1535	1510	1508	7	0.0044
Model C		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	2000	1916	2033	1999	1995	8	0.0042
a_2	2300	2090	2434	2296	2300	48	0.0208
a_3	2600	2452	2648	2579	2572	20	0.0076
b_1	0.3	0.1297	0.8021	0.3142	0.3285	0.0427	0.1299
b_2	0.3	0.0010	0.5605	0.3161	0.2957	0.0946	0.3198
b_3	0.3	0.3030	0.4309	0.3470	0.3547	0.0197	0.0554
χ_1	0.01	0.0010	0.0373	0.0107	0.0123	0.0064	0.5173
χ_2	0.1	0.0671	0.1709	0.1085	0.1124	0.0124	0.1104
χ_3	0.01	0.0010	0.0142	0.0058	0.0055	0.0024	0.4374
ξ_2	800	513	1082	811	820	70	0.0849
ξ_3	1400	1172	1571	1388	1377	39	0.0282

Table 5.2: Optimization outputs for simulated traveltimes corresponding to models A, B, C with noise profile #1. Model corresponds to the true model-parameter values that are used to generate traveltimes. First-pass information contains the bounds of the optimization outputs that have RSS values within 1% of the minimum RSS. Second-pass information contains the optimal parameters, $\boldsymbol{\beta}^*$, the sample mean, μ, and standard deviation, σ, of outputs with RSS within 1% of the minimum, as well as the coefficient of variation, σ/μ.

$\boldsymbol{\beta}$	First-pass bounds		Second-pass statistics			
	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1621	1664	1609	1622	17	0.0105
a_2	2066	2249	2227	2160	64	0.0297
a_3	2617	2620	2620	2618	3	0.0010
b_1	0.4275	0.7136	0.7622	0.6836	0.0937	0.1371
b_2	0.3618	0.6118	0.2659	0.3528	0.1111	0.3150
b_3	0.0010	0.0041	0.0015	0.0074	0.0065	0.8771
χ_1	0.0010	0.0017	0.0010	0.0013	0.0003	0.2584
χ_2	0.0010	0.0042	0.0010	0.0010	0.0001	0.0907
χ_3	0.2377	0.2402	0.2377	0.2385	0.0018	0.0076
ξ_2	586	1229	940	830	208	0.2512
ξ_3	1524	1529	1522	1524	3	0.0020

Table 5.3: Optimization outputs for field data traveltimes. First-pass information contains the bounds of the optimization outputs that have RSS values within 1% of the minimum RSS. Second-pass information contains the optimal parameters, $\boldsymbol{\beta}^*$, the sample mean, μ, and standard deviation, σ, of outputs with RSS within 1% of the minimum, as well as the coefficient of variation, σ/μ.

presented in Table 5.2, along with the traveltime expressions in Section 5.3.5. As expected, the residuals are centred around zero without an observable trend. Moreover, the residuals agree with the input noise profile, whose values are plotted in Figures A.4 and A.5, which is indicative of an accurate fit. Similar residual plots can be produced in this manner for the remaining velocity models and noise profiles.

In view of these explanations, we have demonstrated that the optimization procedure yields consistent and reliable results. Within the scope of the simulation study, we can successfully estimate model parameters for three different velocity models under the effect of three different noise profiles. Thus, we deem the optimization procedure viable for parameter estimation using the field data.

5.6 Field-data application

Let us apply the optimization procedure to the field data obtained from the VSP acquisition, which are discussed in Section 5.3.1. In Table 5.3 of this section, we tabulate the optimization outputs for the first- and second-passes of optimization. Also, we present the traveltime residuals in Figure 5.5. Now, let us proceed with a discussion of the results.

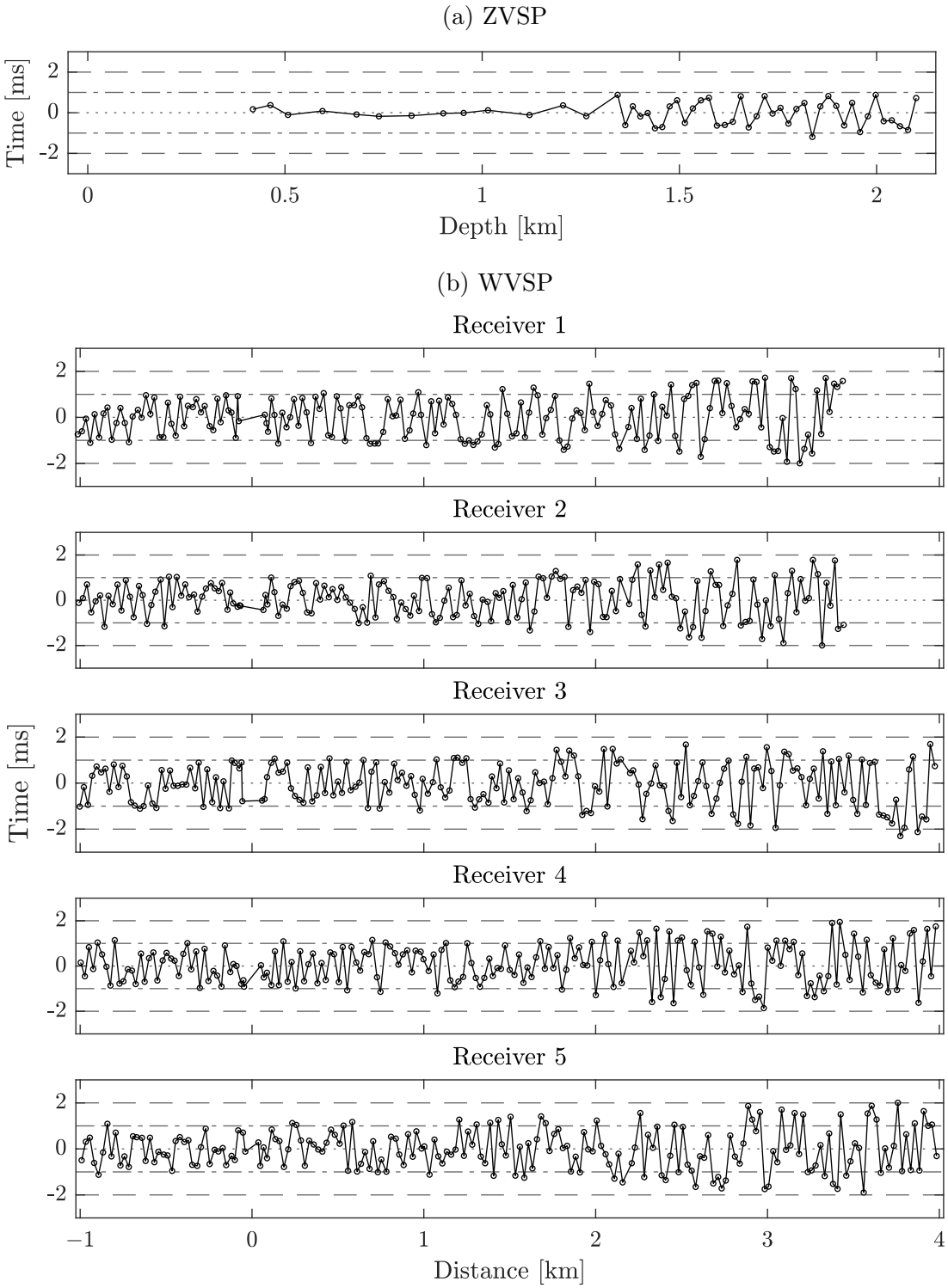


Figure 5.4: Traveltime residuals with optimal parameters for simulated traveltimes corresponding to model A with noise profile #1. Subfigure (a) corresponds to the ZVSP traveltimes whereas subfigure (b) to the WVSP.

As a first observation, we find that the bounds of the first-pass-optimization outputs — that have RSS values within 1% of the minimum RSS — do not vary greatly for most parameters. The bounds in Table 5.3 suggest that the variability of model-parameter values near the minimizer is low. The lower variability would imply that the field data have less noise than our simulated traveltimes. Qualitative support of this implication is readily obtained through direct comparison of the simulated and field-data residuals of Figures 5.4 and 5.5, respectively. The relative scatter of the former is much larger than the latter, especially for the near offset traveltimes.

Turning our attention to the second-pass-optimization outputs, we find that the optimal values are within one standard deviation of the mean. This is a similar result to that of the simulation study, wherein we found that an agreement between the mean and optimal parameter values corresponded to an agreement with the true model parameters. As such, we consider the optimal values to be adequate estimates of the subsurface.

Let us assess each of the optimal velocity model, $\boldsymbol{\beta}^*$. In Table 5.3, we find that the vertical velocities, a_i, increase with depth, which is an expected outcome in view of compaction due to overburden. For the inhomogeneities, we find that b_i decreases with depth, which implies that the Earth material — within the region to which the VSP pertains — is of a similar composition at depth. We suspect that the shallowest layer has the greatest vertical-velocity gradient due to these differing material compositions. For the anisotropies, we find that the deepest layer is the only anisotropic layer. We consider this to be reasonable in view of the increasing anisotropic behaviour of shale with increasing compaction due to overburden. Likewise, the most anisotropic layer is also the least inhomogeneous, which would suggest that it is comprised of a similar material composition.

Regarding the variability of the estimates, we observe that the greatest variability, as measured by the coefficient of variation, occurs for parameters with low values. This is a consequence of low values of inhomogeneity and/or anisotropy being more difficult to estimate within a traveltime optimization. However, of the estimates with larger values, we find that b_2 and ξ_2 have the most variability, which is an expected result due to the scarcity of ZVSP traveltimes shallower than approximately 1300 m. These conclusions are supported by having already encountered similar scenarios during the simulation study, as discussed in Section 5.5.3.

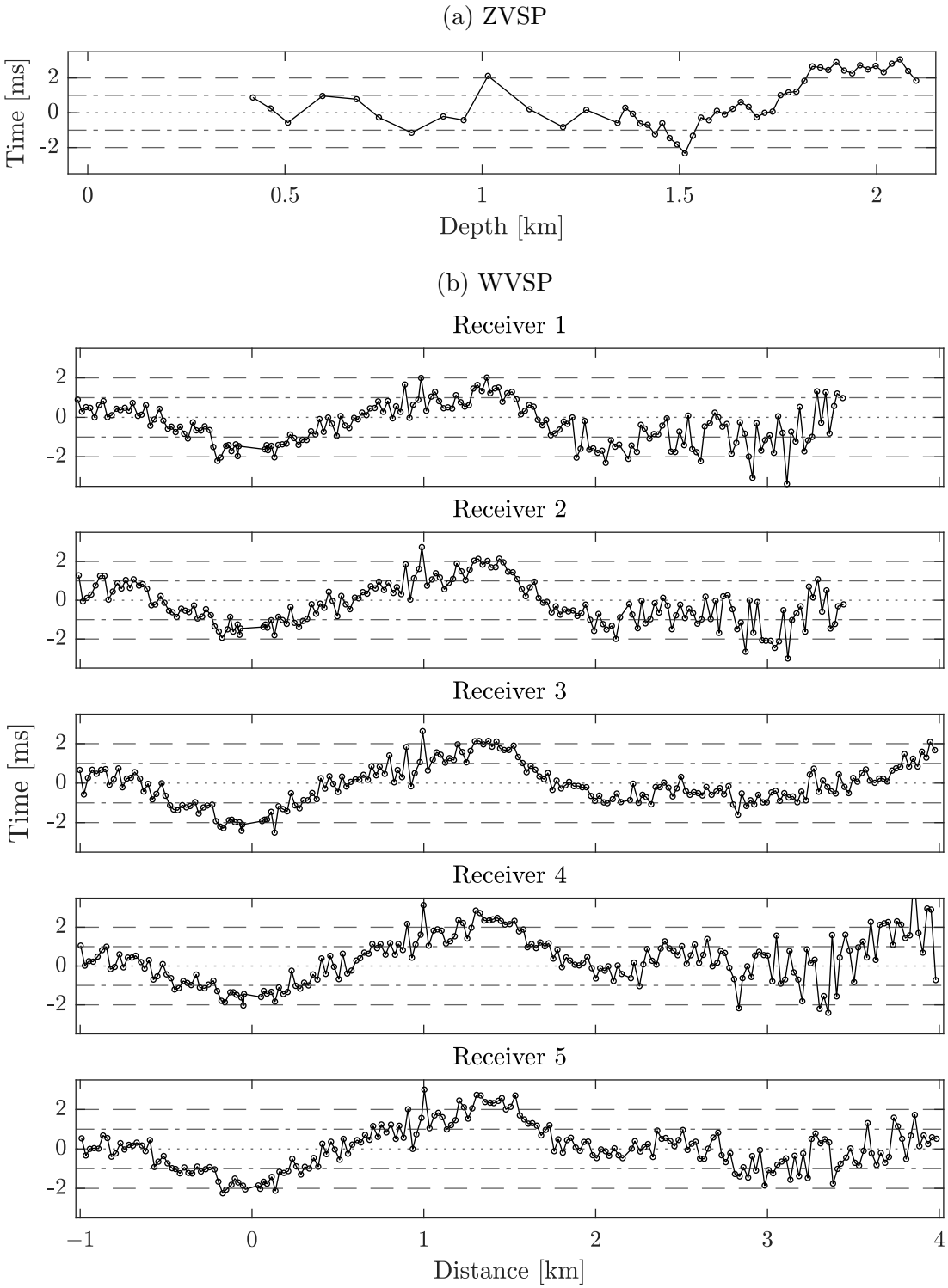


Figure 5.5: Traveltime residuals with optimal parameters for field data traveltimes. Subfigure (a) corresponds to the ZVSP traveltimes whereas subfigure (b) to the WVSP.

Finally, let us examine the residual plots in Figure 5.5. We find that the residuals obtained using the optimal velocity model are within two milliseconds for the majority of the ZVSP and WVSP traveltimes. Such an agreement indicates that the velocity model models the traveltime data adequately.

With that being said, we observe that the residual plots are not randomly scattered about the mean of zero, which is the case for the residuals in the simulation study. Specifically, the residuals in Figure 5.5 exhibit an undulating pattern, which is indicative of a form of model inadequacy. For example, such an inadequacy would arise if one were to perform a polynomial regression using a polynomial of lesser degree than that which was used to generate the data. In other words, the residual plot for a simple linear regression upon data generated from a quadratic model would undulate in a similar manner to the plots in Figure 5.5.

The undulating residual pattern suggests that our traveltime model is not able to not capture the properties of the subsurface entirely. In particular, we suspect that the model inadequacy stems from our assumption of lateral homogeneity. In a field-data setting, this assumption can be broken either by the presence of a lateral velocity gradient or non-planar intermediate interfaces, both of which are reasonable to assume in the Earth's subsurface. However, accommodating such a model inadequacy is beyond the scope of this work.

5.7 Alternative modelling considerations

In view of the results of Section 5.6, let us discuss several alternative modelling considerations that could prove fruitful for future researchers.

Homogeneous layers

Within the optimizations of field-data traveltimes, we observe that some outputs resulted in nearly homogeneous layers. To gain a better understanding, it would be beneficial to repeat the simulation study with a homogeneous layer as a part of the true simulated parameters. However, since setting b to zero is nonpermissible, as it would lead to indeterminate raytracing and traveltime expressions (5.10) and (5.12), respectively, we require an adjustment to the model.

To that end, setting $b = 0$ in velocity expression (5.6) would form the foundations of

a so-called $a\chi$-model. In such a case, eikonal equation is

$$a^2\left((1+2\chi)p_x^2+p_z^2\right)=1.$$

Establishing and solving the Hamiltonian system of equations yields

$$\begin{cases} x(t)=a^2t(1+2\chi)p+x(0) \\ z(t)=at\sqrt{1-p^2a^2(1+2\chi)}+z(0) \\ p_x(t)=p \\ p_z(t)=\frac{1}{a}\sqrt{1-p^2a^2(1+2\chi)} \end{cases}. \tag{5.47}$$

In contrast to expression (5.9), it can be shown that

$$\frac{(x(t)-x(0))^2}{p^2a^2(1+2\chi)^2}-(z(t)-z(0))^2=1, \tag{5.48}$$

which indicates that the shape of a ray in a homogeneous and elliptically anisotropic layer is a hyperbola, not an ellipse. We obtain the raytracing expression by solving (5.48) for

$$x(t)=x(0)+pa\,(z(t)-z(0))(1+2\chi)\sqrt{\frac{1}{1-\mathfrak{p}^2a^2(1+2\chi)}}. \tag{5.49}$$

We obtain the traveltime expression by solving the second equation in system (5.47) for

$$t=\frac{z(t)-z(0)}{a\sqrt{1-p^2a^2(1+2\chi)}}. \tag{5.50}$$

Using expressions (5.49) and (5.50) for homogeneous layers reduces the number of parameters in the multilayer model, which might be useful in further simulation studies.

Generalized elliptical anisotropy

Another manipulation to the model could be to consider a so-called generalized elliptical anisotropy. Throughout this dissertation, we consider the standard elliptical anisotropy established by Slawinski et al. (2004), wherein

$$v_x^2=(1+2\chi)v_z^2 \quad \text{and} \quad v_z=a+bz$$

are used in velocity model (5.6). As opposed to restricting v_x to be a scalar multiple of v_z, it is possible to relax this assumption and establish a velocity model using

$$v_x = c + dz \quad \text{and} \quad v_z = a + bz,$$

where both velocities increase with depth. This model is the topic of a forthcoming article by Diner and Beyaz (Diner, personal communication, May 2023). The generalized model allows for the ellipticity to vary at each point of the medium. In particular, the analytical solution for raytracing presented therein, along with the traveltime expressions, could prove useful toward quantifying the anisotropy within the VSP.

Lateral heterogeneity

Recently, two research groups have put forth studies toward accounting for the effects of lateral heterogeneity. In particular, Sripanich et al. (2019) propose a general framework for evaluating the one-way traveltime derivatives in layered anisotropic media in the presence of weak lateral heterogeneity from curved reflectors and lateral velocity variations. Most recently, Grechka et al. (2024) present a method to identify the presence of weak lateral heterogeneity from VSP data and then developed the procedure to remove its influence on estimated model parameters. These promising research results indicate that future work could be done to re-estimate the model parameters within our VSP data set, while accounting for the effect of lateral heterogeneity.

5.8 Concluding remarks

The linearly inhomogeneous and elliptically anisotropic traveltime model of Slawinski et al. (2004) is applied to a multilayer setting for the purpose of traveltime inversion obtained from VSP field data. Relying on a numerical optimization procedure, we conduct a simulation study to assess the model's accuracy for the estimation of model parameters, using synthetic traveltime data, generated from known model parameters, under the effect of simulated measurement error as noise. We demonstrate that the procedure yields accurate estimates of model parameters for three different velocity models and noise profiles, which supports its application on the VSP traveltimes.

In the field-data case, we obtain parameter estimates that model the VSP traveltimes to within ± 2 milliseconds. However, we observe a non-random pattern within the residual plots, which suggests that some features within the subsurface, namely lateral heterogeneity, are not adequately modelled. To that end, we conclude by providing alternative modelling considerations for future researchers that might work with these VSP field data.

Chapter 6

Conclusion

This dissertation focuses on mathematical modelling and parameter estimation, using numerical optimization, in the context of seismic media. As a point of consistency, the projects described in Chapters 3–5 use a horizontally stratified background model of the Earth's subsurface. We conduct our investigations at the macroscopic level within homogenized media whose properties are determined by the constituent layers of the background model. In these chapters, we use mathematical models to i) determine the conditions for elliptical roots of the Christoffel equation in transversely isotropic media, ii) perform Taylor expansions of traveltimes along the vertical axis of the background model to design a novel approximate-traveltime homogenization, and iii) estimate the parameter values of a traveltime model by applying numerical optimization to VSP traveltime measurements.

Now, for the purpose of a unified conclusion for this traditional-format dissertation, what follows is a discussion regarding the achievements, limitations, and recommendations (for future researchers) regarding the original developments of Chapters 3–5.

On Christoffel roots for nondetached slowness surfaces

Achievement In Chapter 3, we demonstrate that elliptical roots of the Christoffel equation require nondetached slowness surfaces. In particular, we determine this condition within the framework of mathematical proof as well as a numerical example. Our result can be used to gain insight on what is permissible within a mathematical

model, despite the outcome being contrary to one's expectations.

Limitation Since our motivation is simply to demonstrate the existence of elliptical roots, our result is a theoretical finding and an arbitrary numerical example. Likewise, our result is limited to transversely isotropic media that is the result of the Backus average on isotropic layers. Since the Backus average can be applied to layers of any symmetry class, our result lacks generality.

Recommendation Much like the work of Bos et al. (2017), which focuses on the Backus average for generally anisotropic layers, determining the conditions for elliptical Christoffel roots in Backus media of different symmetry classes could prove an interesting research direction. In that vein, perhaps the numerical examples in such media could correspond better to the layer properties of seismic media.

On forward modelling of traveltimes in approximate media

Achievement In Chapter 4, we develop a novel model of a homogenized approximate-traveltime medium. In particular, we use a Taylor expansion of traveltimes about the vertical axis to obtain the elasticity parameters that result in equal vertical traveltimes in the approximate medium. Our model can be used to compare VSP traveltimes with those that are a result of well log interval velocities.

Limitation Our forward model has two limitations. First, as in Chapter 3, our result applies only to isotropic constituent layers. Consequently, it is likely to have limited applicability toward modelling subsurface regions that are anisotropic. Second, our formulation is not restricted to a critical angle, which is the case for the Fermat traveltimes. Hence, our modelling approach has limited applicability in examining greater offsets.

Recommendation The formulation could benefit from the inclusion of anisotropy in the input layers. While considering TI constituents would lead to complicated raytracing, introducing a scalar difference between the horizontal and vertical velocities, e.g., Slawinski et al. (2004), leads to closed-form raytracing expressions. The resulting approximate medium could increase the accuracy of traveltimes at even further offsets than the current limitation of 30°.

On inverse modelling of traveltimes

Achievement In Chapter 5, we estimate the inhomogeneity and anisotropy parameters of a multilayer Earth model using VSP traveltime measurements. In particular, we conduct a simulation study to assess the reliability of our results and introduce penalization strategies to impose parameter constraints within an unconstrained, derivative-free optimization algorithm. We demonstrate that our approach obtains accurate estimates for various velocity models and noise profiles.

Limitation Our study is limited by the availability of first-break traveltime information only from the VSP and the *a priori* assumption of lateral homogeneity. The first limits our inversion to P-wave traveltime tomography, which precludes the use of more-sophisticated techniques such as FWI. The second is an assumption that can be broken if the subsurface has non-planar layer interfaces and/or lateral heterogeneity within its layers, which limits our model to less geologically complex subsurface structures. While our ray-based approach leads to efficient computation of seismic ray trajectories and satisfactory inversion if the subsurface velocity model is known, the approach is limited to simpler geological structures. As such, the utility of our field-data estimates is limited to an initial velocity model that can be used for other inversion techniques.

Recommendation To account for the stated limitations, we would revisit the traveltime modelling using a wave-equation-based techniques, without the *a priori* assumptions of horizontally stratified and laterally homogeneous layers. While relaxing such assumptions would forgo our use of closed-form raytracing and traveltime expressions, we expect that the improved imaging resolution of the subsurface velocity justifies the increased computational costs. Recent studies demonstrate the viability of this approach through the use of full-traveltime inversion, which is used to invert for a kinematically accurate velocity model from traveltime information only (Luo et al., 2016).

Closing remarks

In closing, this dissertation represents our efforts to study seismic media through the analogies put forth by mathematical models as well as their agreement with data measurements through the practice of parameter estimation. We acknowledge that

the approaches used herein, along with those that have been developed, are not exhaustive — likely, they constitute mere drops in the proverbial bucket of seismology.

Bibliography

Abedi, M. M., Pardo, D., and Stovas, A. (2021). Approximations for traveltime, slope, curvature, and geometric spreading of elastic waves in layered transversely isotropic media. *Geophysics*, 86(6):1ND–Z3.

Achenbach, J. D. (1973). *Wave propagation in elastic solids*, volume 16 of *Applied Mathematics and Mechanics*. North-Holland Publishing Company.

Adamus, F. P., Kaderali, A., Slawinski, M. A., and Stanoev, T. (2020). On effects of inhomogeneity on anisotropy in Backus average. *arXiv*, 1802.04075 [physics.geo-ph].

Aki, K. and Richards, P. G. (2002). *Quantitative seismology*. University Science Books, 2nd edition.

Anton, H. (1984). *Elementary linear algebra*. John Wiley & Sons, Inc., 4th edition.

Backus, G. E. (1962). Long-wave elastic anisotropy produced by horizontal layering. *Journal of Geophysical Research*, 67(11):4427–4440.

Beer, F. P., Johnston Jr., E. R., DeWolf, J. T., and Mazurek, D. F. (2011). *Statics and mechanics of materials*. McGraw-Hill Companies, Inc.Hill.

Beller, S., Montellier, V., Operto, S., Nolet, G., Paul, A., and Zhao, L. (2018). Lithospheric architecture of the South-Western Alps revealed by multiparameter teleseismic full-waveform inversion. *Geophysical Journal International*, 212(2):1369–1388.

Berryman, J. G. (1979). Long-wave elastic anisotropy in transversely isotropic media. *Geophysics*, 44(5):896–917.

Bóna, A., Bucataru, I., and Slawinski, M. A. (2007a). Coordinate-free characterization of the symmetry classes of elasticity tensors. *Journal of Elasticity*, 87(2):109–132.

Bóna, A., Bucataru, I., and Slawinski, M. A. (2007b). Material symmetries versus wavefront symmetries. *The Quarterly Journal of Mechanics and Applied Mathematics*, 60(2):73–84.

Bóna, A., Slawinski, M. A., and Smith, P. (2009). Ray tracing by simulated annealing: Bending method. *Geophysics*, 74(2):T25–T32.

Bos, L., Dalton, D. R., Slawinski, M. A., and Stanoev, T. (2017). On Backus average for generally anisotropic layers. *Journal of Elasticity*, 127(2):179–196.

Bos, L., Danek, T., Slawinski, M. A., and Stanoev, T. (2018). Statistical and numerical considerations of Backus-average product approximation. *Journal of Elasticity*, 132(1):141–159.

Bos, L., Slawinski, M. A., and Stanoev, T. (2019a). On Christoffel roots for nondetached slowness surfaces. *Geophysical Prospecting*, 67(9):2280–2286.

Bos, L., Slawinski, M. A., and Stanoev, T. (2019b). On Christoffel roots for nondetached slowness surfaces. *arXiv*, 1903.02514 [physics.geo-ph].

Bos, L., Slawinski, M. A., and Stanoev, T. (2019c). On the Backus average of a layered medium with elasticity tensors in random orientations. *Zeitschrift für Angewandte Mathematik und Physik*, 70(84).

Botta-Dukát, Z. (2023). Quartile coefficient of variation is more robust than CV for traits calculated as a ratio. *Nature*, 13(4671).

Brown, R. J., Lamoureux, M. P., Slawinski, M. A., and Slawinski, R. A. (2000). Direct traveltime inversion of VSP data for elliptical anisotropy in layered media. *CREWES Research Report*, 12(3):1–16.

Bucataru, I. and Slawinski, M. A. (2009). On convexity and detachment of innermost wavefront-slowness sheet. *Geophysics*, 74(5):WB63—WB66.

Červený, V. (2001). *Seismic ray theory*. Cambridge University Press.

Chapman, C. (2004). *Fundemantals of seismic wave propagation*. Cambridge University Press.

Cholach, P. Y. and Schmitt, D. R. (2003). Seismic anisotropy of shales. *CSEG Recorder*, 28(7).

Colley, S. J. (2012). *Vector calculus*. Pearson Education, Inc., 4th edition.

Conn, A. R., Scheinberg, K., and Vincente, L. N. (2009). *Introduction to derivative-free optimization*. Society for Industrial and Applied Mathematics/Mathematical Programming Society, Philadelphia.

Dalton, D. R. and Slawinski, M. A. (2016). On Backus average for oblique incidence. *arXiv*, 1601.02966 [physics.geo-ph].

Danek, T. and Slawinski, M. A. (2016). Backus average under random perturbations of layered media. *SIAM Journal on Applied Mathematics*, 76(4):1239–1249.

Dhabaria, N. and Singh, S. C. (2024). Hamiltonian Monte Carlo based elastic full-waveform inversion of wide-angle seismic data. *Geophysical Journal International*, 237:1384–1399.

Diner, C. and Beyaz, A. (Forthcoming). Ray-tracing equations for generalized elliptical anisotropy.

Dix, C. H. (1955). Seismic velocities from surface measurements. *Geophysics*, 20(1):68–86.

Farlow, S. J. (1982). *Partial Differential Equations for Scientists & Engineers*. John Wiley & Sons, Ltd.

Farra, V. and Pšenčík, I. (2020). P-wave reflection-moveout appoximation for horizontally layered media of tilted weak-to-moderate orthorhombic symmetry. *Journal of Applied Geophysics*, 178(104057).

Feynman, R. P., Leighton, R., and Sands, M. (2011a). *The Feynman Lectures on Physics, Vol. I: The New Millennium Edition: Mainly Mechanics, Radiation, and Heat*. The Feynman Lectures on Physics. Basic Books, new millenium edition edition.

Feynman, R. P., Leighton, R., and Sands, M. (2011b). *The Feynman Lectures on Physics, Vol. II: The New Millennium Edition: Mainly Electromagnetism and Matter*, volume II of *The Feynman Lectures on Physics*. Basic Books, new millenium edition edition.

Fu, L. and Hanafy, S. M. (2017). Ray-tracing traveltime tomography versus wave-

equation traveltime inversion for near-surface seismic land data. *Interpretation*, 5(3):1A–T449.

Gao, F. and Han, L. (2012). Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, 51:259–277.

Gereck, H. (2007). Poisson's ratio values for rocks. *International Journal of Rock Mechanics and Mining Sciences*, 44(1):1–13.

Giroux, B. and Gloaguen, E. (2012). Geostatical traveltime tomography in elliptically anisotropic media. *Geophysical Prospecting*, 60(6):1133–1149.

Gold, N., Shapiro, S. A., Bojinski, S., and Müller, T. M. (2000). An approach to upscaling for seismic waves in statistically isotropic heterogeneous elastic media. *Geophysics*, 65:1837–1850.

Grechka, V., Tsvankin, I., and Contreras, P. (2024). Inversion of walkaway VSP data in the presence of lateral velocity heterogeneity. *Energy Geoscience*, 5(100255).

Guasch, L., Agudo, O. C., Tang, M., Nanchev, P., and Warner, M. (2020). Full-waveform inversion imaging of the human brain. *npj Digital Medicine*, 3(28).

Gurtin, M. E., Fried, E., and Anand, L. (2010). *The Mechanics and Thermodynamics of Continua*. Cambridge University Press.

Helbig, K. (1979). On: "The reflection, refraction, and diffraction of waves in media with elliptical velocity dependence", by Franklyn K. Levin (Geophysics, April 1978, p. 528–537). *Geophysics*, 44(5):987–990.

Helbig, K. (1983). Elliptical anisotropy—Its significance and meaning. *Geophysics*, 48(7):825–832.

Helbig, K. (1994). Foundations of anisotropy for exploration seismics. In *Handbook of geophysical exploration: Seismic exploration*, volume 22. Elsevier Science Serials.

Holzapfel, G. A. (2000). *Nonlinear solid mechanics: A continuum approach for engineering*. John Wiley & Sons, Ltd.

Ji, S., Li, L., Motra, H. B., Wuttke, F., Sun, S., Michibayashi, K., and Salisbury, M. H. (2018). Poisson's ratio and auxetic properties of natural rocks. *Journal of Geophysical Research: Solid Earth*, 123:1161–1185.

Kaderali, A. (2009). Investigating anisotropy and inhomogeneity using tomographic inversion of VSP traveltimes: Validation of analytic expressions for linearly inhomogeneous elliptically anisotropic models. Master's thesis, Memorial University of Newfoundland.

Kaderali, A. and Stanoev, T. (2020). On anisotropy and inhomogeneity parameter estimation using traveltimes. *arXiv*, 2012.03393 [physics.geo-ph].

Kolman, B. and Hill, D. R. (2008). *Elementary linear algebra with applications*. Pearson Education, Inc., 9th edition.

Lakes, R. S. (2017). Negative-poisson's-ratio materials: Auxetic solids. *Annual Review of Materials Research*, 47(1):63–81.

Lelièvre, P. G., Farquharson, C., and Hurich, C. A. (2011). Inversion of first-arrival seismic traveltimes without rays, implemented on unstructured grids. *Geophysical Journal International*, 185(2):749–763.

Levin, F. K. (1978). The reflection, refraction, and diffraction of waves in media with an elliptical velocity dependence. *Geophysics*, 43(3):528–537.

Lewintan, P., Müller, S., and Neff, P. (2021). Korn inequalities for incompatible tensor fields in three space dimensions with conformally invariant dislocation energy. *Calculus of Variations*, 60(150).

Li, X., Yao, G., Niu, F., and Liu, N. (2022). Waveform inversion of seismic first arrivals acquired on irregular surface. *Geophysics*, 87(3):R291–R304.

Lines, L. R. and Newrick, R. T. (2004). *Fundamentals of geophysical interpretation*. Society of Exploration Geophysicists.

Lipschutz, M. M. (1969). *Schaum's Outline of Differential Geometry*. McGraw Hill Professional.

Luo, Y., Ma, Y., Wu, Y., Liu, H., and Cao, L. (2016). Full-traveltime inversion. *Geophysics*, 81(5):R261–R274.

Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The rock physics handbook: Tools for seismic analysis of porous media*. Cambridge University Press, Cambridge, 2nd edition.

Meléndez, A., Jiménez, C. E., Sallaraès, V., and Ranero, C. R. (2019). Anisotropic

P-wave travel-time tomography implementing Thomsen's weak approximation in TOMO3D. *Solid Earth*, 10(6):1857–1876.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*. W. H. Freeman and Company.

Musgrave, M. J. P. (1970). *Crystal acoustics: Introduction to the study of elastic waves and vibrations in crystals*. Holden-Day.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

Nocedal, J. and Wright, S. J. (2006). *Numerical optimization*. Springer series in operations research and financial engineering. Springer,, 2 edition.

Postma, G. W. (1955). Wave propagation in a stratified medium. *Geophysics*, 20(4):780–806.

Ravve, I. and Koren, Z. (2017). Traveltime approximation in vertical transversely isotropic layered media. *Geophysical Prospecting*, 65(6):1559–1581.

Rawlinson, N., Hauser, J., and Sambridge, M. (2008). Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in Geophysics*, 49:203–273.

Rogister, Y. and Slawinski, M. A. (2005). Analytic solution of ray-tracing equations for a linearly inhomogeneous and elliptically anisotropic velocity model. *Geophysics*, 70(5):D37–D41.

Rudzki, M. P. (1911). Parametrische Darstellung der elastischen Welle in anisotropen Medien. *Anzeiger der Akademie der Wissenshaften Krakau*, pages 503–536.

Rudzki, M. P. (2003). Parametric representation of the elastic wave in anisotropic media. *Journal of Applied Geophysics*, 54(3):165–183.

Serway, R. A. and Jewett, J. W. (2013). *Principles of physics: A calculus-based text, hybrid*. Brooks/Cole.

Sheriff, R. E. (2002). *Encyclopedic dictionary of applied geophysics*, volume 13 of *Geophysical references*. Society of Explorational Geophysicists, 4 edition.

Slawinski, M. A. (2015). *Waves and rays in elastic continua*. World Scientific, 3rd edition.

Slawinski, M. A. (2018). *Waves and rays in seismology: Answers to unasked questions*. World Scientific, 2nd edition.

Slawinski, M. A. (2020a). *Waves and rays in elastic continua*. World Scientific, 4 edition.

Slawinski, M. A. (2020b). *Waves and rays in seismology: Answers to unasked questions*. World Scientific, 3 edition.

Slawinski, M. A., Lamoureux, M. P., Slawinski, R. A., and Brown, R. J. (2003). VSP traveltime inversion for anisotropy in a buried layer. *Geophysical Prospecting*, 51:131–139.

Slawinski, M. A., Wheaton, C. J., and Powojowski, M. (2004). VSP traveltime inversion for linear inhomogeneity and elliptical anisotropy. *Geophysics*, 69(2):373–377.

Sripanich, Y. and Fomel, S. (2015). On anelliptic approximations for qP velocities in transversely isotropic and orthorhombic media. *Geophysics*, 80(5):C89–C105.

Sripanich, Y., Fomel, S., and Stovas, A. (2019). Effects of lateral heterogeneity on time-domain processing parameters. *Geophysical Journal International*, 219:1181–1201.

Stewart, J., Clegg, D. K., and Watson, S. (2021). *Calculus: Early transcendentals*. Cengage Learning, Inc., 9th edition.

Taner, M. T. and Koehler, F. (1969). Velocity spectra—Digital computer derivation and applications of velocity functions. *Geophysics*, 34(6):859–881.

Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.

Thomsen, L. (1986). Weak elastic aniostropy. *Geophysics*, 51(10):1954–1966.

Tong, P. (2021). Adjoint-state traveltime tomography: Eikonal equation-based methods and application to the Anza area in Southern California. *Journal of Geophysical Research: Solid Earth*, 126(5).

Udías, A. (1999). *Principles of seismology*. Cambridge University Press.

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74(6).

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC.

Wright, M. H. (1996). Direct search methods: Once scorned, now respectable. In Griffiths, D. F. and Watson, G. A., editors, *Numerical Analysis*, Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis, pages 191–208. Addison Wesley Longman, Harlow, United Kingdom.

Yang, D., Liu, F., Morton, S., Malcolm, A., and Fehler, M. (2016). Time-lapse full-waveform inversion with ocean-bottom-cable data: Application on Valhall field. *Geophysics*, 81(4):R225–R235.

Zhang, M., Xu, T., Bai, Z., Liu, Y., Hou, J., and Yu, G. (2017). Ray tracing of turning wave in elliptically anisotropic media with an irregular surface. *Earthquake Science*, 30:219–228.

Zhou, R. and Kaderali, A. (2006). Anisotropy evaluation using an array walk-away VSP. Paper presented at the Offshore Technology Conference, Houston, Texas, USA, May 2006. OTC-18177-MS.

Appendix A

Supplementary material

A.1 Cauchy's second equation of motion

In Section 2.3.4, we discuss the derivation of Cauchy's second equation of motion. However, for the sake of readability, we have relegated the majority of the derivation to this appendix. We pick up that derivation with master balance principle (2.35), which is

$$\frac{\partial\left(\boldsymbol{r}\times(\rho\mathbf{v})\right)}{\partial t}+\operatorname{div}\left(\left(\boldsymbol{r}\times(\rho\mathbf{v})\right)\otimes\mathbf{v}\right)=\boldsymbol{r}\times(\operatorname{div}\boldsymbol{\sigma})+\boldsymbol{\mathcal{E}}:\boldsymbol{\sigma}^{T}+\boldsymbol{r}\times(\rho\mathbf{b}).$$

By Cauchy's theorem $\mathbf{t}=\boldsymbol{\sigma}\mathbf{n}$, cross product $\boldsymbol{\phi}=\boldsymbol{r}\times\mathbf{t}=\boldsymbol{r}\times\boldsymbol{\sigma}\mathbf{n}$ replaces $\mathbf{u}\times\mathbf{An}$ in conversion (2.33). To simplify the first term in balance (A.1), we recognize that the partial derivative of $\boldsymbol{r}(\boldsymbol{x})=\boldsymbol{x}-\boldsymbol{x}_0$ with respect to time is $\mathbf{v}$. However, since the cross product of a vector with itself is zero, it follows that $\mathbf{v}\times(\rho\mathbf{v})=\mathbf{0}$. Therefore, we have

$$\begin{aligned}\frac{\partial\left(\boldsymbol{r}\times(\rho\mathbf{v})\right)}{\partial t}&=\frac{\partial\boldsymbol{r}}{\partial t}\times(\rho\mathbf{v})+\boldsymbol{r}\times\left(\frac{\partial(\rho\mathbf{v})}{\partial t}\right)\\&=\mathbf{v}\times(\rho\mathbf{v})+\boldsymbol{r}\times\left(\frac{\partial\rho}{\partial\mathbf{t}}\mathbf{v}+\rho\frac{\partial\mathbf{v}}{\partial\mathbf{t}}\right)\\&=\boldsymbol{r}\times\left(\frac{\partial\rho}{\partial t}\mathbf{v}+\rho\frac{\partial\mathbf{v}}{\partial t}\right).\end{aligned}\tag{A.1}$$

For the second term in balance (2.35), since the cross product of two vectors is a vector, we factor out the ρ and simplify the divergence of the tensor product by

identity (2.12),

$$\begin{aligned}\operatorname{div}\left((\boldsymbol{r}\times(\rho\mathbf{v}))\otimes\mathbf{v}\right) &= \operatorname{div}\left((\rho(\boldsymbol{r}\times\mathbf{v}))\otimes\mathbf{v}\right)\\ &= \left(\operatorname{grad}\left(\rho(\boldsymbol{r}\times\mathbf{v})\right)\right)\mathbf{v}+\rho(\boldsymbol{r}\times\mathbf{v})(\operatorname{div}\mathbf{v}).\end{aligned}\tag{A.2}$$

The first term in expression (A.2) contains the gradient of the product of a scalar and a vector, which, by identity (2.10), we simplify as

$$\left(\operatorname{grad}(\rho(\boldsymbol{r}\times\mathbf{v}))\right)\mathbf{v} = \left((\boldsymbol{r}\times\mathbf{v})\otimes(\operatorname{grad}\rho)\right)\mathbf{v}+\left(\rho(\operatorname{grad}(\boldsymbol{r}\times\mathbf{v}))\right)\mathbf{v}.\tag{A.3}$$

We rearrange the first term in expression (A.3) using the definition of the tensor product (2.8) and, for the second term, we distribute the gradient to obtain

$$\left(\operatorname{grad}\left(\rho(\boldsymbol{r}\times\mathbf{v})\right)\right)\mathbf{v} = \left(\mathbf{v}\cdot(\operatorname{grad}\rho)\right)(\boldsymbol{r}\times\mathbf{v})+\rho\left((\operatorname{grad}\boldsymbol{r})\times\mathbf{v}\right)\mathbf{v}+\rho\left(\boldsymbol{r}\times\operatorname{grad}\mathbf{v}\right)\mathbf{v}.\tag{A.4}$$

To evaluate the second and third terms in expression (A.4), let us consider an alternative representation of the cross product. For arbitrary vectors $\mathbf{a},\mathbf{b}\in\mathbb{R}^3$, there is a skew-symmetric matrix representation of the cross product such that

$$\mathbf{a}\times\mathbf{b} =: [\mathbf{a}]_\times\mathbf{b}, \qquad \text{where} \qquad [\mathbf{a}]_\times = \begin{bmatrix} 0 & -a_3 & a_2\\ a_3 & 0 & -a_1\\ -a_2 & a_1 & 0\end{bmatrix}.\tag{A.5}$$

According to Lewintan et al. (2021, Section 2.1), this representation is useful as it allows for the generalization of a cross product in $\mathbb{R}^3$ to a cross product of a vector $\mathbf{a}\in\mathbb{R}^3$ and a matrix $\mathbf{M}\in\mathbb{R}^{3\times3}$ from the left and from the right,

$$\mathbf{a}\times\mathbf{M} := [\mathbf{a}]_\times\mathbf{M} \qquad \text{and} \qquad \mathbf{M}\times\mathbf{a} := \mathbf{M}[\mathbf{a}]_\times,\tag{A.6}$$

which is given by column- and row-wise cross products, respectively. To demonstrate the use of representation (A.5), let us consider the cross product in the third term of (A.4), which is

$$[\boldsymbol{r}]_\times(\operatorname{grad}\mathbf{v}) = \begin{bmatrix} 0 & -r_3 & r_2\\ r_3 & 0 & -r_1\\ -r_2 & r_1 & 0\end{bmatrix}\begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3}\\ v_{2,1} & v_{2,2} & v_{2,3}\\ v_{3,1} & v_{3,2} & v_{3,3}\end{bmatrix}$$

$$= \begin{bmatrix} -r_3 v_{2,1} + r_2 v_{3,1} & -r_3 v_{2,2} + r_2 v_{3,2} & -r_3 v_{2,3} + r_2 v_{3,3} \\ r_3 v_{1,1} - r_1 v_{3,1} & r_3 v_{1,2} - r_1 v_{3,2} & r_3 v_{1,3} - r_1 v_{3,3} \\ -r_2 v_{1,1} + r_1 v_{2,1} & -r_2 v_{1,2} + r_1 v_{2,2} & -r_2 v_{1,3} + r_1 v_{2,3} \end{bmatrix},$$

where the components of the gradient of the vector are given by expression (2.13), which is

$$\operatorname{grad} \mathbf{v} = \nabla \otimes \mathbf{v} = \frac{\partial v_i}{\partial x_j} \, \mathbf{e}_i \otimes \mathbf{e}_j.$$

Then, by matrix multiplication with $\mathbf{v}$, using comma notation (2.14), we can rearrange the resulting vector components such that

$$([\boldsymbol{r}]_\times (\operatorname{grad} \mathbf{v}))\, \mathbf{v} = \begin{bmatrix} -r_3 \left(v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3\right) + r_2 \left(v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3\right) \\ r_3 \left(v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3\right) - r_1 \left(v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3\right) \\ -r_2 \left(v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3\right) + r_1 \left(v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3\right) \end{bmatrix}. \tag{A.7}$$

By comparison, using matrix multiplication

$$(\operatorname{grad} \mathbf{v})\, \mathbf{v} = \begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3} \\ v_{2,1} & v_{2,2} & v_{2,3} \\ v_{3,1} & v_{3,2} & v_{3,3} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3 \\ v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3 \\ v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3 \end{bmatrix},$$

we find that

$$\begin{aligned} [\boldsymbol{r}]_\times \left((\operatorname{grad} \mathbf{v})\mathbf{v}\right) &= \begin{bmatrix} 0 & -r_3 & r_2 \\ r_3 & 0 & -r_1 \\ -r_2 & r_1 & 0 \end{bmatrix} \begin{bmatrix} v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3 \\ v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3 \\ v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3 \end{bmatrix} \\ &= \begin{bmatrix} -r_3 \left(v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3\right) + r_2 \left(v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3\right) \\ r_3 \left(v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3\right) - r_1 \left(v_{3,1} v_1 + v_{3,2} v_2 + v_{3,3} v_3\right) \\ -r_2 \left(v_{1,1} v_1 + v_{1,2} v_2 + v_{1,3} v_3\right) + r_1 \left(v_{2,1} v_1 + v_{2,2} v_2 + v_{2,3} v_3\right) \end{bmatrix}. \end{aligned} \tag{A.8}$$

Hence, by the equality of matrices (A.7) and (A.8), we have the identity

$$(\boldsymbol{r} \times (\operatorname{grad} \mathbf{v}))\, \mathbf{v} = ([\boldsymbol{r}]_\times (\operatorname{grad} \mathbf{v}))\, \mathbf{v} = [\boldsymbol{r}]_\times \left((\operatorname{grad} \mathbf{v})\mathbf{v}\right) = \boldsymbol{r} \times \left((\operatorname{grad} \mathbf{v})\mathbf{v}\right). \tag{A.9}$$

For the cross product in the second term of (A.4), we recognize that the gradient of the position vector, $\operatorname{grad} \boldsymbol{r}$, is an identity matrix, I. Then, using representations (A.6),

it follows that

$$(\operatorname{grad} \boldsymbol{r} \times \mathbf{v})\mathbf{v} = (I \times \mathbf{v})\mathbf{v} = (I[\mathbf{v}]_{\times})\mathbf{v} = [\mathbf{v}]_{\times}\mathbf{v} = \mathbf{v} \times \mathbf{v} = \mathbf{0}. \tag{A.10}$$

In view of identity (A.9) and result (A.10), expression (A.4) simplifies to

$$\begin{aligned}(\operatorname{grad}(\rho(\boldsymbol{r} \times \mathbf{v})))\,\mathbf{v} &= (\mathbf{v} \cdot (\operatorname{grad}\rho))\,(\boldsymbol{r} \times \mathbf{v}) + \rho\,((\operatorname{grad}\boldsymbol{r}) \times \mathbf{v})\,\mathbf{v} + \boldsymbol{r} \times ((\operatorname{grad}\mathbf{v})\mathbf{v}) \\ &= (\mathbf{v} \cdot (\operatorname{grad}\rho))\,(\boldsymbol{r} \times \mathbf{v}) + \mathbf{0} + \rho\,(\boldsymbol{r} \times ((\operatorname{grad}\mathbf{v})\mathbf{v})) \\ &= (\mathbf{v} \cdot (\operatorname{grad}\rho))\,(\boldsymbol{r} \times \mathbf{v}) + \boldsymbol{r} \times (\rho(\operatorname{grad}\mathbf{v})\mathbf{v})\,. \end{aligned} \tag{A.11}$$

Returning to the divergence of the tensor product of expression (A.2), along with expression (A.11), the second term in balance (2.35) becomes

$$\operatorname{div}((\boldsymbol{r} \times (\rho\mathbf{v})) \otimes \mathbf{v}) = (\mathbf{v} \cdot (\operatorname{grad}\rho))\,(\boldsymbol{r} \times \mathbf{v}) + \boldsymbol{r} \times (\rho\,(\operatorname{grad}\mathbf{v})\mathbf{v}) + (\boldsymbol{r} \times \mathbf{v})(\rho(\operatorname{div}\mathbf{v})). \tag{A.12}$$

Therefore, using derivative (A.1) and relation (A.12), along with algebraic manipulation, the master balance principle (2.35) takes on the form

$$\begin{aligned}\boldsymbol{r} \times \left(\frac{\partial\,\rho}{\partial t}\mathbf{v} + \rho\frac{\partial\,\mathbf{v}}{\partial t} + \rho(\operatorname{grad}\mathbf{v})\mathbf{v} - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b}\right) \\ + (\boldsymbol{r} \times \mathbf{v})\,(\rho(\operatorname{div}\mathbf{v}) + \mathbf{v} \cdot (\operatorname{grad}\rho)) = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T.\end{aligned} \tag{A.13}$$

We can simplify the first cross product in balance (A.13) by recognizing that the second and third terms combine to form the material-time derivative (2.16) of vector $\mathbf{v}$, whereby

$$\boldsymbol{r} \times \left(\frac{\partial\,\rho}{\partial t}\mathbf{v} + \rho\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} - \operatorname{div}\boldsymbol{\sigma} - \rho\mathbf{b}\right) + (\boldsymbol{r} \times \mathbf{v})\,(\rho(\operatorname{div}\mathbf{v}) + \mathbf{v} \cdot (\operatorname{grad}\rho)) = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T. \tag{A.14}$$

Then, we recognize that the latter terms in the first cross product of balance (A.14) sum to zero by virtue of the elastodynamic equation (2.31). Commuting the time derivative of ρ with the cross product, we collect like terms and rearrange balance (A.14) such that

$$(\boldsymbol{r} \times \mathbf{v})\left(\frac{\partial\,\rho}{\partial t} + (\operatorname{grad}\rho) \cdot \mathbf{v} + \rho(\operatorname{div}\mathbf{v})\right) = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T. \tag{A.15}$$

We simplify balance (A.15) further by the material-time derivative (2.16) of ρ to obtain

$$(\boldsymbol{r} \times \mathbf{v}) \left(\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho(\operatorname{div} \mathbf{v}) \right) = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T. \tag{A.16}$$

However, by the balance of mass (2.25), the left-hand side of balance (A.16) is zero and, as such,

$$\mathbf{0} = \boldsymbol{\mathcal{E}} : \boldsymbol{\sigma}^T. \tag{A.17}$$

We continue the derivation of Cauchy's second equation of motion in expression (2.36) of Section 2.3.4.

A.2 Polar reciprocity

In Section 3.1, we state that, in accordance with polar reciprocity, the ellipticity of slownesses is equivalent to ellipticity of wavefronts. To verify this statement in a seismological context, we present the principle of polar reciprocity, as discussed by Helbig (1994), derive the mathematical formulæ that quantify the principle, and conclude with an illustration of the principle. Also, we provide an illustration of polar reciprocity in the context of slownesses and wavefronts in transversely isotropic media.

A.2.1 Plane curve formulation

We begin with Helbig (1994, Section 2A.4), who states that two curves are polar reciprocals if for each point of the first curve there is one point on the second curve such that the position vector of the first curve is parallel to the normal of the second and the position vector of the second is parallel to the normal of the first. Also, he states that polar reciprocity of two curves is tantamount to the dot product of all corresponding position vectors being constant.

To assess these statements, we consider a simple and closed plane curve that is centred on the origin. We define the curve in parametric form as the locus of all points (x, y) that are determined by position vector, relative to the origin,

$$\boldsymbol{r}(\vartheta) := \begin{bmatrix} f(\vartheta) \\ g(\vartheta) \end{bmatrix}, \tag{A.18}$$

where $f(\vartheta)$ and $g(\vartheta)$ are continuous (and at least once differentiable) functions and ϑ is the parameter. In this context, simple means that the curve never intersects itself and closed implies that the initial and final coordinates of the curve are equal. Associated with the position vector are the unit tangent and unit normal vectors (e.g., Lipschutz, 1969, Chapter 4),

$$\boldsymbol{T}(\vartheta) := \frac{\boldsymbol{r}'(\vartheta)}{\|\boldsymbol{r}'(\vartheta)\|} \quad \text{and} \quad \boldsymbol{N}(\vartheta) := \frac{\boldsymbol{T}'(\vartheta)}{\|\boldsymbol{T}'(\vartheta)\|},$$

respectively, where $(\circ)' := \mathrm{d}(\circ)/\mathrm{d}\vartheta$ and $\|(\circ)\| := \sqrt{(\circ)\cdot(\circ)}$ is the Euclidean norm. As such, taking the derivative of the position vector and normalizing, the unit tangent

vector is

$$\boldsymbol{T}(\vartheta) = \frac{\boldsymbol{r}'(\vartheta)}{\|\boldsymbol{r}'(\vartheta)\|} = \frac{1}{\sqrt{f'(\vartheta)^2 + g'(\vartheta)^2}} \begin{bmatrix} f'(\vartheta) \\ g'(\vartheta) \end{bmatrix}.$$

For the unit normal vector, we differentiate $\boldsymbol{T}(\vartheta)$, using the chain rule, to obtain

$$\boldsymbol{T}'(\vartheta) = \frac{1}{\left(f'(\vartheta)^2 + g'(\vartheta)^2\right)^{3/2}} \begin{bmatrix} g'(\vartheta)\left(g'(\vartheta)f''(\vartheta) - f'(\vartheta)g''(\vartheta)\right) \\ f'(\vartheta)\left(f'(\vartheta)g''(\vartheta) - g'(\vartheta)f''(\vartheta)\right) \end{bmatrix},$$

from which we manipulate algebraically to obtain its squared norm

$$\|\boldsymbol{T}'(\vartheta)\|^2 = \frac{\left(g'(\vartheta)f''(\vartheta) - f'(\vartheta)g''(\vartheta)\right)^2}{\left(f'(\vartheta)^2 + g'(\vartheta)^2\right)^2}.$$

Therefore, the unit normal vector simplifies to

$$\boldsymbol{N}(\vartheta) = \frac{\boldsymbol{T}'(\vartheta)}{\|\boldsymbol{T}'(\vartheta)\|} = \frac{1}{\sqrt{f'(\vartheta)^2 + g'(\vartheta)^2}} \begin{bmatrix} g'(\vartheta) \\ -f'(\vartheta) \end{bmatrix}.$$

Now, consider a second curve whose position vector is parallel to normal of the first, $\boldsymbol{R}(\vartheta) := \rho(\vartheta)\left[g'(\vartheta), -f'(\vartheta)\right]^T$, where $\rho(\vartheta) > 0$. Given these two curves, the polar reciprocity condition requires that the dot product of their position vectors remains constant for all ϑ, i.e., $\boldsymbol{r}(\vartheta) \cdot \boldsymbol{R}(\vartheta) = \alpha^2$, where α is a constant nonzero scalar value. We satisfy this condition with $\rho(\vartheta) = \alpha^2 / \left(f(\vartheta)g'(\vartheta) - g(\vartheta)f'(\vartheta)\right)$. As such, the polar reciprocal curve to $\boldsymbol{r}(\vartheta)$ is

$$\boldsymbol{R}(\vartheta) = \frac{\alpha^2}{f(\vartheta)g'(\vartheta) - g(\vartheta)f'(\vartheta)} \begin{bmatrix} g'(\vartheta) \\ -f'(\vartheta) \end{bmatrix}. \tag{A.19}$$

A.2.2 Ellipse

Suppose that the first curve is an ellipse centred on the origin. To define such an ellipse, we consider its semi-major, a, and semi-minor, b, axes such that $a \geq b > 0$. The coordinates of the ellipse are $(x, y) = (f(\vartheta), g(\vartheta))$ with $f(\vartheta) = a\cos\vartheta$, $g(\vartheta) = b\sin\vartheta$, and $\vartheta \in [0, 2\pi)$. Consequently, position vector (A.18) becomes

$$\boldsymbol{r}(\vartheta) = \begin{bmatrix} a\cos\vartheta \\ b\sin\vartheta \end{bmatrix}. \tag{A.20}$$

To specify the second curve, we substitute $f(\vartheta)$ and $g(\vartheta)$ — along with their derivatives $f'(\vartheta) = -a\sin\vartheta$ and $g'(\vartheta) = b\cos\vartheta$ — in position vector (A.19) to obtain

$$\boldsymbol{R}(\vartheta) = \frac{\alpha^2}{(a\cos\vartheta)(b\cos\vartheta) - (b\sin\vartheta)(-a\sin\vartheta)} \begin{bmatrix} b\cos\vartheta \\ a\sin\vartheta \end{bmatrix} = \frac{\alpha^2}{ab} \begin{bmatrix} b\cos\vartheta \\ a\sin\vartheta \end{bmatrix}. \quad \text{(A.21)}$$

As such, it is clear that the second curve is an ellipse as well, albeit with different semi-major, α^2/a, and semi-minor, α^2/b, axes. We can confirm the polar reciprocity of the two curves by substituting position vectors (A.20) and (A.21) in the polar reciprocity condition to obtain $\boldsymbol{r}(\vartheta) \cdot \boldsymbol{R}(\vartheta) = (\alpha^2/ab)\left((a\cos\vartheta)(b\cos\vartheta) + (b\sin\vartheta)(a\sin\vartheta)\right) = \alpha^2$, which is a constant, as expected.

With this result, we confirm that the polar reciprocal of an ellipse centred on the origin is another ellipse centred on the origin. The implication of this result is that, in a seismological context, an elliptical slowness curve is equivalent to an elliptical wavefront. As an illustration of this result, we present a plot of the two ellipses in Figure A.1 for parameter values specified in the caption. Therein, we have set $\alpha = 2$ for scaling purposes so that $\boldsymbol{r}(\vartheta)$ and $\boldsymbol{R}(\vartheta)$ intersect on the y-axis; note that, in general, these curves may or may not intersect.

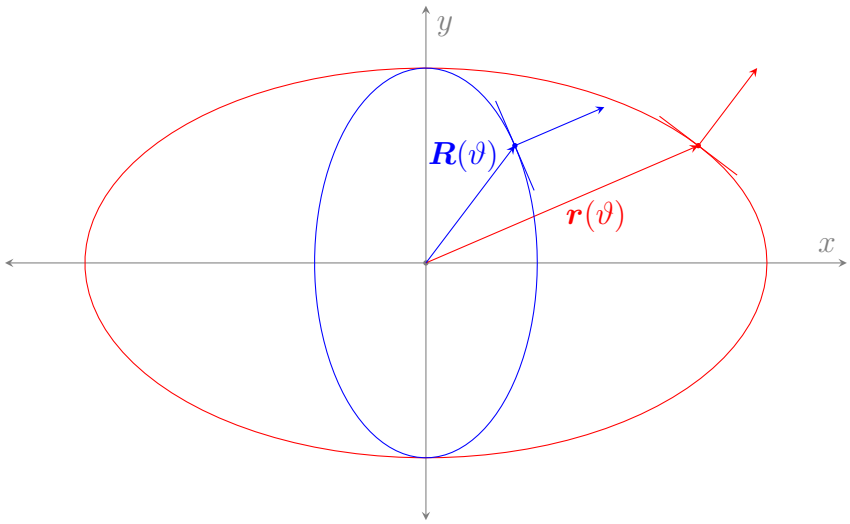


Figure A.1: The polar reciprocal of an ellipse is an ellipse. The first curve (red ellipse) is specified by position vector (A.20), where $a = 5/2$ and $b = 2$. The second curve (blue ellipse) is specified by position vector (A.21), where $\alpha = 2$. For both curves, we draw the position vectors $\boldsymbol{r}(\vartheta)$ and $\boldsymbol{R}(\vartheta)$, specified at $\vartheta = 17°$, along with a segment of a tangent line and the unit normal vector.

A.2.3 Transverse isotropy

In Section 3.3, we present the slowness curves for the qP, qSV, and SH waves in several transversely isotropic media, specified by different elasticity-parameter values. In a similar manner to the ellipse, we can plot the wavefronts that correspond to the slowness curves using polar reciprocity and position vector (A.19).

Let us recall that the slowness curve for a given wave is generated using the inverse of its phase velocity, which are given in expression (3.8) for the qP and qSV wave and expression (3.9) for the SH wave. For any of these waves, we specify the coordinates of the slowness curve using position vector (A.18), with $f(\vartheta) = \cos\vartheta/v(\vartheta)$ and $g(\vartheta) = \sin\vartheta/v(\vartheta)$, where $v(\vartheta)$ can be replaced $v_{qP}(\vartheta)$, $v_{qSV}(\vartheta)$, or $v_{SH}(\vartheta)$. By polar reciprocity, the coordinates of wavefronts, corresponding to each of the slowness curves, are specified using position vector (A.19), with

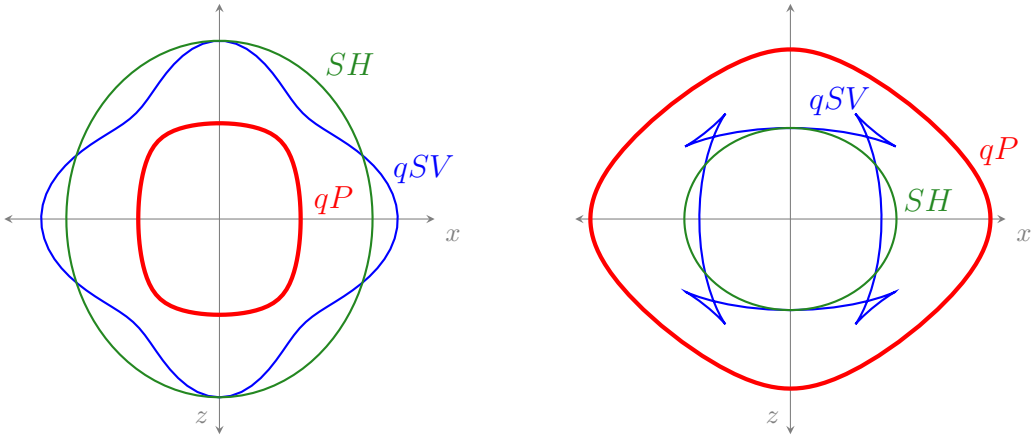
$$f'(\vartheta) = -\frac{v(\vartheta)\sin\vartheta + v'(\vartheta)\cos\vartheta}{v(\vartheta)^2} \quad \text{and} \quad g'(\vartheta) = \frac{v(\vartheta)\cos\vartheta - v'(\vartheta)\sin\vartheta}{v(\vartheta)^2}\,;$$

the derivatives $v'_{qP}(\vartheta)$, $v'_{qSV}(\vartheta)$, and $v'_{SH}(\vartheta)$ are given in expressions (4.40)–(4.42).

Now, let us turn our attention now to the Green-River shale medium. We consider the slowness curves in Figure 3.2(a), which are generated using elasticity parameters (3.26), and in Figure 3.2(c), which are generated with the same elasticity parameters, but subjected to nondetachment condition (3.12a), i.e., $c^{\mathrm{TI}}_{1133} = -c^{\mathrm{TI}}_{2323}$.

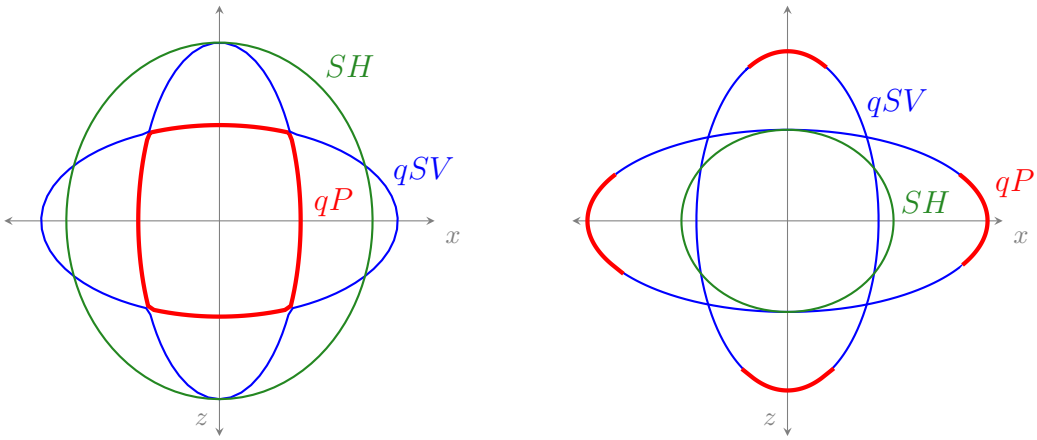
For the former, we use position vector (A.19) to generate the wavefronts in Figure A.2(b). Therein, we observe that magnitude of the qP wavefront is the greatest among the three waves, which is consistent with its slowness having the least magnitude, that the eight cusps of the qSV wavefront correspond to the eight points of inflection of its slowness surface, and that the SH wavefront is an ellipse as a consequence of its elliptical slowness.

For the latter, we generate the wavefronts in Figure A.3(b) subject to $c^{\mathrm{TI}}_{1133} = -c^{\mathrm{TI}}_{2323}$, As is discussed in Chapter 3, while the innermost curve corresponds to the qP wave, the slowness curves of the qP and qSV waves become connected with one another. Moreover, the curves are no longer smooth at the point of their intersection. Therefore, while continuity is maintained, the derivative of the phase velocity is not defined at



(a) Slownesses (b) Wavefronts

Figure A.2: Slownesses and wavefronts in Green-River shale, generated using elasticity parameters (3.26). The red curves correspond to the qP wave, blue curves to the qSV wave, and green curves to the SH wave. In subplot (b), $\alpha = \sqrt{\frac{2}{11}}$ for scaling purposes so that the wavefronts are comparable to the slownesses.



(a) Slownesses (b) Wavefronts

Figure A.3: Slownesses and wavefronts in Green-River shale, generated using elasticity parameters (3.26), but subjected to $c^{\mathrm{TI}}_{1133} = -c^{\mathrm{TI}}_{2323}$. In subplot (b), $\alpha = \sqrt{\frac{2}{11}}$.

these locations. As a result, although the outer curve is continuous — and despite segments of ellipses being present in the wavefront — the qP and qSV wavefronts are discontinuous. This result can be verified by matching the red and blue wavefront segments in Figure A.3(b) to the corresponding segments of the slowness curve in Figure A.3(a) using polar reciprocity. Beyond these aspects, it remains the case that the SH wavefront is an ellipse as a consequence of its elliptical slowness.

A.3 Nelder-Mead algorithm

In Section 5.4.2, we discuss the Nelder-Mead algorithm and its specifications for the traveltime inversion. For convenience of the readership, we provide these specifications within Algorithm 1. Note that, for consistency with the notation in this dissertation, we have made several notational adjustments from the cited sources.

Algorithm 1 Nelder-Mead algorithm (adapted from Conn et al. (2009, p. 143) and Gao and Han (2012, Section 1))

Initialization: Choose an initial simplex of vertices $\boldsymbol{\beta}_0 = \{\boldsymbol{\beta}_0^1, \ldots, \boldsymbol{\beta}_0^{n+1}\}$. Evaluate f at the points in $\boldsymbol{\beta}_0$. Choose scaling constants:

$$\alpha = 1, \quad \beta = 2, \quad \gamma = \tfrac{1}{2}, \quad \delta = \tfrac{1}{2}. \tag{A.22}$$

for $k = 0, 1, 2, \ldots$ **do**

0. Set $\boldsymbol{\beta} = \boldsymbol{\beta}_k$.
1. **Order**: Order the $n+1$ vertices of $\boldsymbol{\beta} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^{n+1}\}$ so that
$$f^1 = f(\boldsymbol{\beta}^1) \leq \cdots \leq f^N = f(\boldsymbol{\beta}^N).$$
2. **Reflect**: Reflect the worst vertex $\boldsymbol{\beta}^{n+1}$ over the centroid $\boldsymbol{\beta}^c = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\beta}^i$ of the remaining n vertices:
$$\boldsymbol{\beta}^r = \boldsymbol{\beta}^c + \alpha\left(\boldsymbol{\beta}^c - \boldsymbol{\beta}^{n+1}\right).$$
Evaluate $f^r = f(\boldsymbol{\beta}^r)$. If $f^1 \leq f^r < f^n$, then replace $\boldsymbol{\beta}^{n+1}$ by the reflected point $\boldsymbol{\beta}^r$ and terminate the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^n, \boldsymbol{\beta}^r\}$.
3. **Expand**: If $f^r < f^1$, then calculate the expansion point
$$\boldsymbol{\beta}^e = \boldsymbol{\beta}^c + \beta\left(\boldsymbol{\beta}^r - \boldsymbol{\beta}^c\right)$$
and evaluate $f^e = f(\boldsymbol{\beta}^e)$. If $f^e \leq f^r$, replace $\boldsymbol{\beta}^{n+1}$ by the expansion point $\boldsymbol{\beta}^e$ and terminate the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^n, \boldsymbol{\beta}^e\}$. Otherwise, replace $\boldsymbol{\beta}^n$ by the reflected point $\boldsymbol{\beta}^r$ and terminate the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^n, \boldsymbol{\beta}^r\}$.
4. **Contract**: If $f^r \geq f^n$, then a contraction is performed between the best of $\boldsymbol{\beta}^r$ and $\boldsymbol{\beta}^{n+1}$.
   (a) **Outside contraction**: If $f^r < f^{n+1}$, perform an outside contraction
   $$\boldsymbol{\beta}^{oc} = \boldsymbol{\beta}^c + \gamma\left(\boldsymbol{\beta}^r - \boldsymbol{\beta}^c\right)$$
   and evaluate $f^{oc} = f(\boldsymbol{\beta}^{oc})$. If $f^{oc} \leq f^r$, replace $\boldsymbol{\beta}^{n+1}$ by the outside contraction point $\boldsymbol{\beta}^{oc}$ and terminate the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^n, \boldsymbol{\beta}^{oc}\}$. Otherwise, perform a shrink.
   (b) **Inside contraction**: If $f^r \geq f^{N+1}$, perform an inside contraction
   $$\boldsymbol{\beta}^{ic} = \boldsymbol{\beta}^c - \gamma\left(\boldsymbol{\beta}^r - \boldsymbol{\beta}^c\right)$$
   and evaluate $f^{ic} = f(\boldsymbol{\beta}^{ic})$. If $f^{ic} \leq f^{n+1}$, replace $\boldsymbol{\beta}^{n+1}$ by the inside contraction point $\boldsymbol{\beta}^{ic}$ and terminate the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1, \ldots, \boldsymbol{\beta}^n, \boldsymbol{\beta}^{ic}\}$. Otherwise, perform a shrink.
5. **Shrink**: Evaluate f at the n points $\boldsymbol{\beta}^1 - \delta\left(\boldsymbol{\beta}^i - \boldsymbol{\beta}^1\right)$, for $i = 2, \ldots, n+1$, and replace $\boldsymbol{\beta}^2, \ldots, \boldsymbol{\beta}^{n+1}$ by these points, terminating the iteration: $\boldsymbol{\beta}_{k+1} = \{\boldsymbol{\beta}^1 - \delta\left(\boldsymbol{\beta}^i - \boldsymbol{\beta}^1\right)\}$.

end for

A.4 Simulation study

A.4.1 Noise

In Section 5.5.1, we discuss the generation of simulated measurement errors, otherwise referred to as noise. For the purposes of the simulation study, we generate three sets of noise for both the simulated ZVSP and WVSP traveltimes, which we present herein.

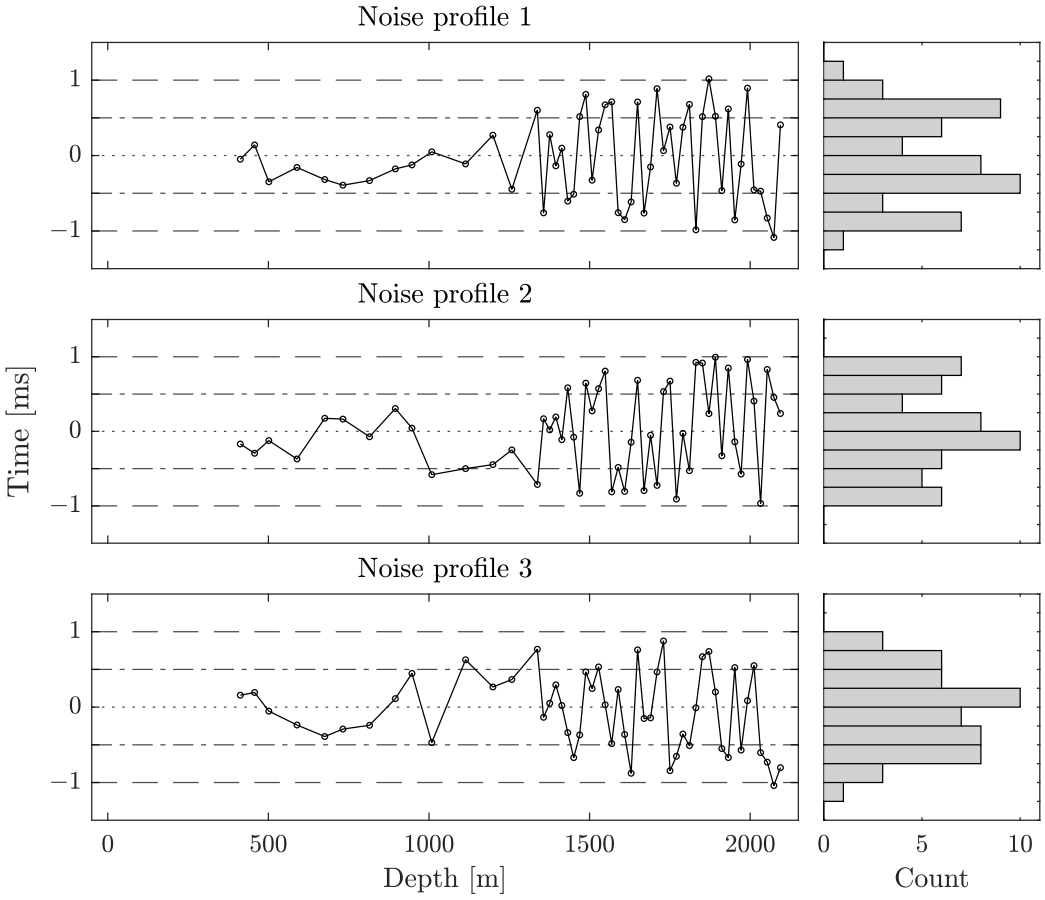


Figure A.4: ZVSP noise profiles. Left-hand column: residuals with respect to receiver depth; horizontal lines spaced at 0.5-millisecond intervals. Right-hand column: residuals in histogram format; binning width of 0.25 milliseconds.

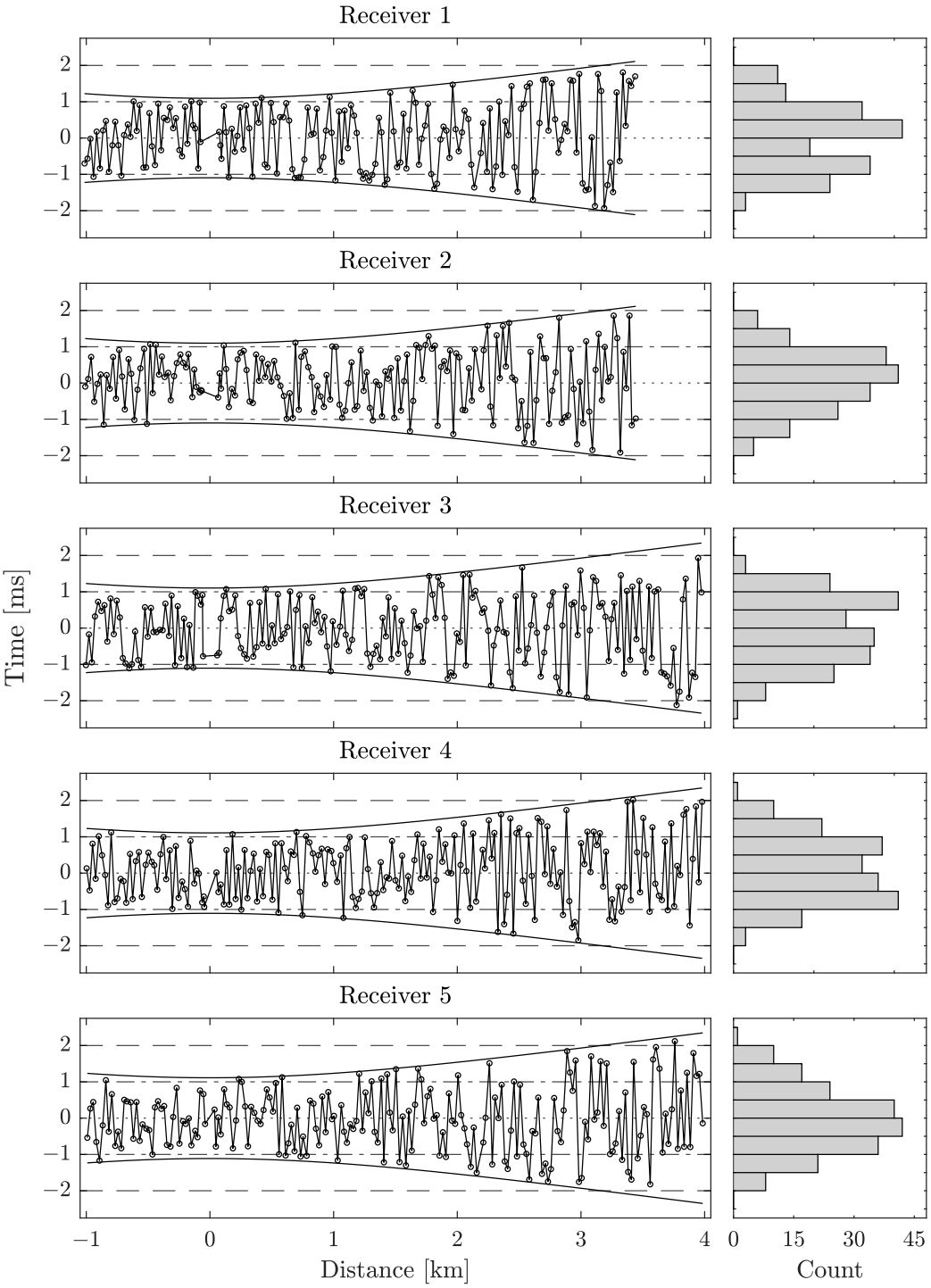


Figure A.5: WVSP noise profile #1. Left-hand column: residuals with respect to source offset; horizontal lines spaced at one-millisecond intervals. Right-hand column: residuals in histogram format; binning width of 0.5 milliseconds.

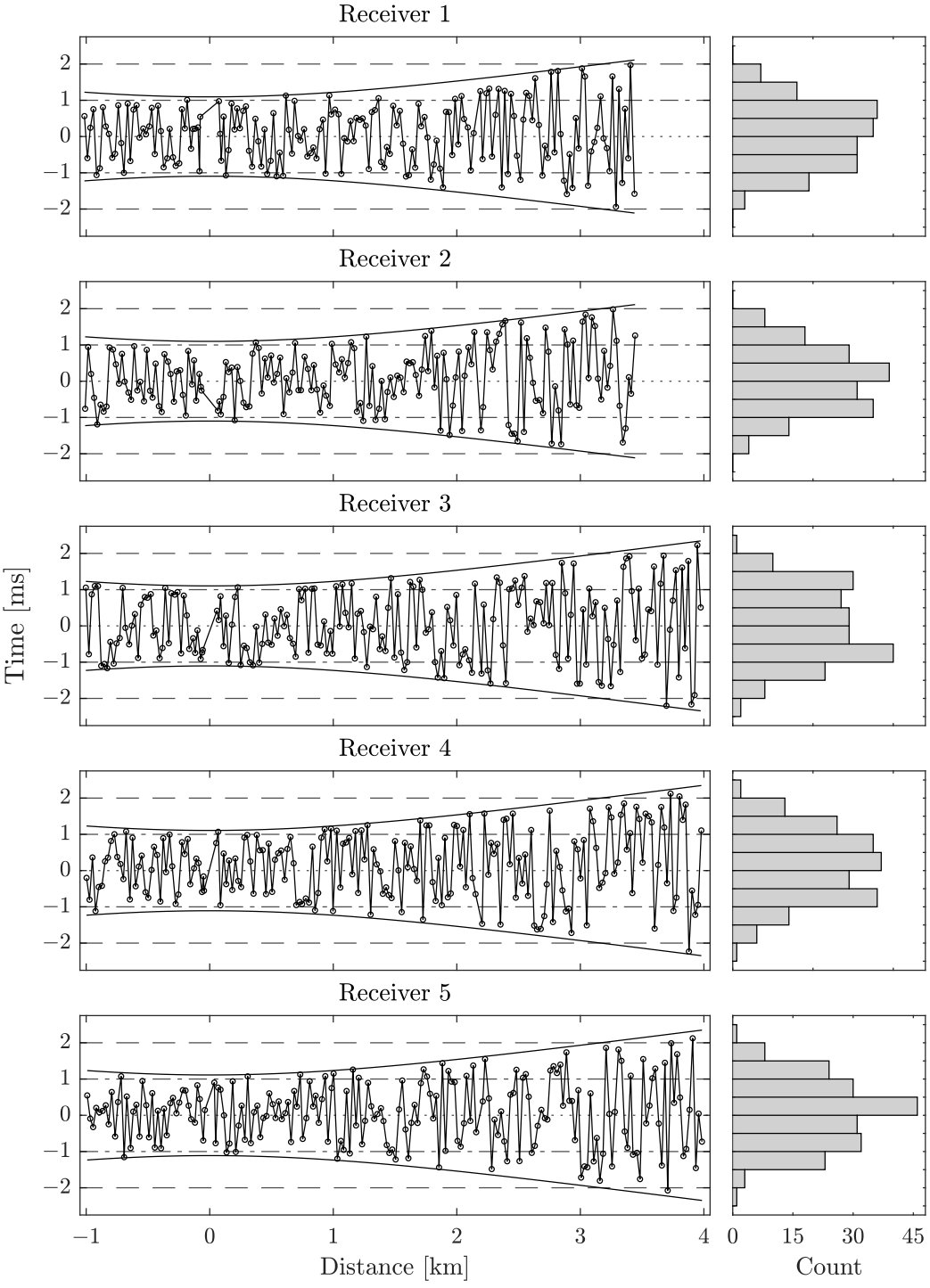


Figure A.6: WVSP noise profile #2. Figure formatting consistent with Figure A.5.

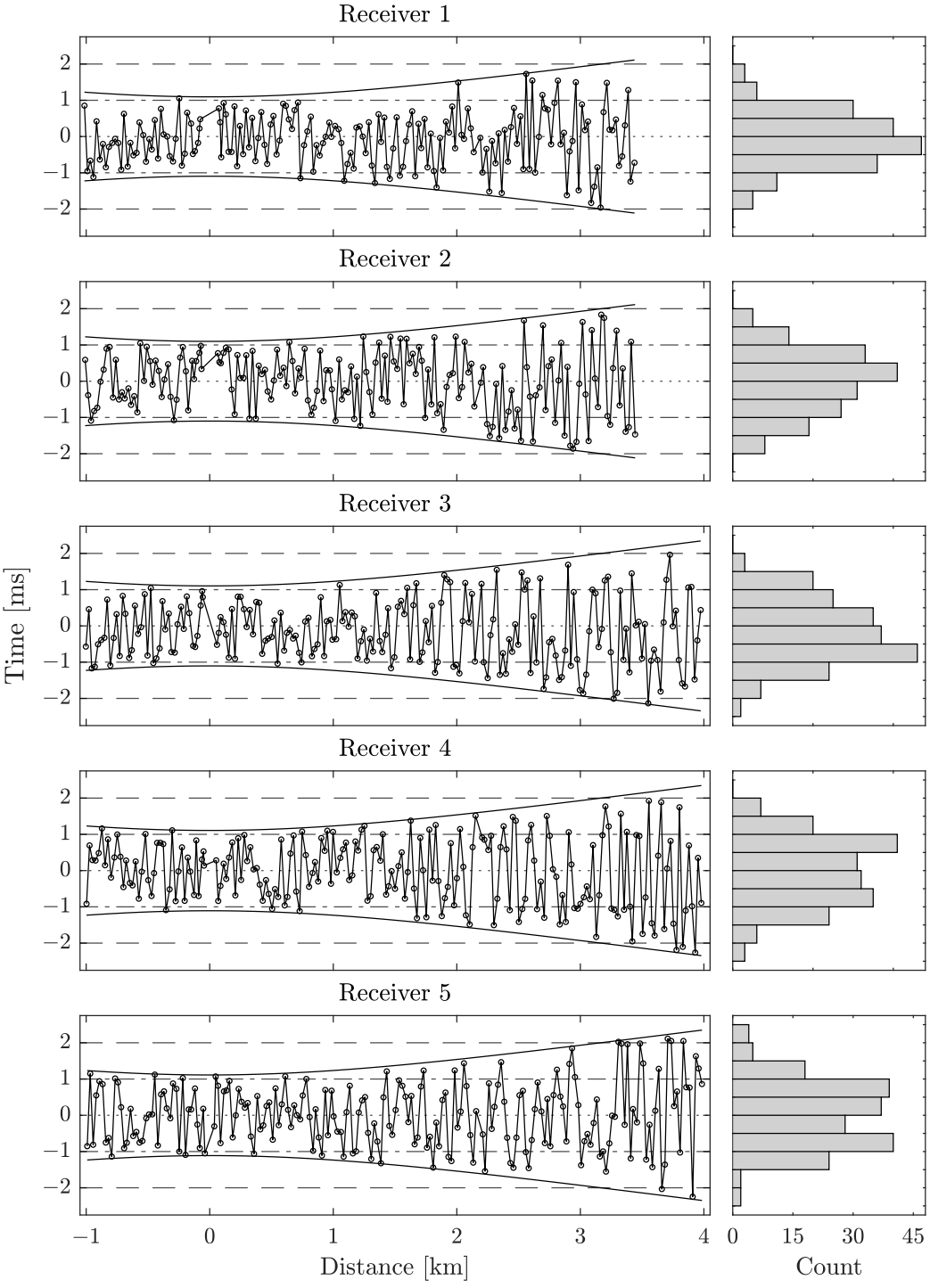


Figure A.7: WVSP noise profile #3. Figure formatting consistent with Figure A.5.

A.4.2 Optimization outputs

In Section 5.5.3, we discuss the results of the simulation study optimization. Due to the similarity of results across noise profiles, we present only those results that correspond to noise profile #1 in Table 5.2, therein. Herein, for completeness, we include the results for noise profiles #2 and #3 in Tables A.1 and A.2.

Noise profile #2

Model A		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1800	1794	1837	1805	1810	11	0.0059
a_2	2478	2064	2530	2396	2329	109	0.0469
a_3	2728	2507	2793	2677	2686	50	0.0187
b_1	0.7	0.4774	0.7170	0.6811	0.6550	0.0525	0.0802
b_2	0.5	0.4212	1.0602	0.4212	0.5690	0.1308	0.2299
b_3	0.3	0.2091	0.7986	0.2991	0.3436	0.0856	0.2492
χ_1	0.01	0.0010	0.0331	0.0137	0.0111	0.0056	0.5089
χ_2	0.01	0.0010	0.0360	0.0049	0.0105	0.0072	0.6804
χ_3	0.01	0.0010	0.0199	0.0112	0.0100	0.0027	0.2684
ξ_2	1100	500	1295	907	862	192	0.2233
ξ_3	1600	1070	1866	1438	1510	154	0.1023
Model B		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1900	1855	1909	1901	1899	6	0.0031
a_2	2300	2045	2742	2335	2347	132	0.0563
a_3	2500	2489	2541	2509	2512	9	0.0037
b_1	0.5	0.4680	0.7620	0.4978	0.5089	0.0284	0.0559
b_2	1.25	0.0468	1.6245	1.1512	1.0561	0.1924	0.1822
b_3	0.75	0.6196	0.8786	0.7224	0.7221	0.0375	0.0520
χ_1	0.05	0.0120	0.0644	0.0508	0.0484	0.0096	0.1975
χ_2	0.1	0.0658	0.1272	0.0975	0.0984	0.0077	0.0786
χ_3	0.15	0.1414	0.1648	0.1536	0.1537	0.0033	0.0215
ξ_2	1000	629	1226	1015	1000	124	0.1243
ξ_3	1500	1479	1542	1503	1506	10	0.0065
Model C		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	2000	1974	2052	2009	2008	7	0.0035
a_2	2300	2099	2418	2325	2322	34	0.0148
a_3	2600	2566	2658	2610	2611	9	0.0033
b_1	0.3	0.0179	0.4396	0.2667	0.2694	0.0343	0.1272
b_2	0.3	0.0010	0.4829	0.2232	0.2293	0.0581	0.2535
b_3	0.3	0.2264	0.3258	0.2757	0.2753	0.0132	0.0478
χ_1	0.01	0.0010	0.0300	0.0138	0.0139	0.0036	0.2590
χ_2	0.1	0.0645	0.1123	0.0923	0.0919	0.0054	0.0582
χ_3	0.01	0.0085	0.0188	0.0132	0.0132	0.0013	0.0981
ξ_2	800	500	1008	800	798	48	0.0598
ξ_3	1400	1310	1532	1405	1408	17	0.0119

Table A.1: Optimization outputs for simulated traveltimes corresponding to models A, B, C and noise profile #2. Formatting consistent with Table 5.2.

Noise profile #3

Model A		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1800	1709	1907	1807	1803	31	0.0171
a_2	2478	2071	2682	2638	2311	198	0.0857
a_3	2728	2530	2814	2814	2680	68	0.0253
b_1	0.7	0.0377	1.2216	0.6784	0.6824	0.1846	0.2705
b_2	0.5	0.2331	1.2625	0.3664	0.5962	0.2341	0.3927
b_3	0.3	0.2335	1.0648	0.3029	0.3612	0.1772	0.4907
χ_1	0.01	0.0010	0.0394	0.0097	0.0101	0.0072	0.7139
χ_2	0.01	0.0015	0.0405	0.0122	0.0140	0.0089	0.6388
χ_3	0.01	0.0010	0.0126	0.0044	0.0080	0.0032	0.4033
ξ_2	1100	500	1452	1386	866	312	0.3602
ξ_3	1600	1083	1900	1881	1487	220	0.1482
Model B		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	1900	1881	1906	1902	1900	4	0.0020
a_2	2300	2113	2778	2414	2375	130	0.0546
a_3	2500	2472	2538	2495	2499	10	0.0041
b_1	0.5	0.4816	0.5928	0.4951	0.5024	0.0133	0.0265
b_2	1.25	0.0010	1.9069	1.5617	1.4120	0.2670	0.1891
b_3	0.75	0.5925	0.8321	0.7373	0.7312	0.0374	0.0511
χ_1	0.05	0.0302	0.0659	0.0594	0.0562	0.0046	0.0826
χ_2	0.1	0.0700	0.1284	0.1020	0.1012	0.0075	0.0745
χ_3	0.15	0.1366	0.1628	0.1442	0.1460	0.0046	0.0312
ξ_2	1000	790	1226	1120	1076	81	0.0752
ξ_3	1500	1471	1526	1486	1490	8	0.0055
Model C		First-pass bounds		Second-pass statistics			
$\boldsymbol{\beta}$	true	lower	upper	$\boldsymbol{\beta}^*$	μ	σ	σ/μ
a_1	2000	1874	2023	1993	1989	13	0.0063
a_2	2300	2082	2442	2331	2290	65	0.0285
a_3	2600	2530	2683	2571	2601	25	0.0095
b_1	0.3	0.2138	1.0404	0.3371	0.3565	0.0698	0.1958
b_2	0.3	0.0099	0.5945	0.2338	0.3412	0.1252	0.3669
b_3	0.3	0.1594	0.3602	0.3243	0.2944	0.0268	0.0912
χ_1	0.01	0.0010	0.0387	0.0136	0.0129	0.0081	0.6257
χ_2	0.1	0.0651	0.1556	0.1330	0.1121	0.0177	0.1581
χ_3	0.01	0.0010	0.0156	0.0010	0.0044	0.0033	0.7545
ξ_2	800	503	1107	866	833	99	0.1191
ξ_3	1400	1274	1601	1344	1402	59	0.0417

Table A.2: Optimization outputs for simulated traveltimes corresponding to model A, B, C and noise profile #3. Formatting consistent with Table 5.2.