UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE ENGENHARIA MECÂNICA
E INSTITUTO DE GEOCIÊNCIAS

CARLOS ANDRÉ MARTINS DE ASSIS

DEVELOPMENT OF INTEGRAL REPRESENTATIONS FOR THE COUPLED ONE-WAY WAVE EQUATIONS AND STRATEGIES FOR PARAMETER ESTIMATION BASED ON JOINT MIGRATION INVERSION

DESENVOLVIMENTO DAS REPRESENTAÇÕES INTEGRAIS PARA AS EQUAÇÕES UNIDIRECIONAIS DE ONDA ACOPLADAS E ESTRATÉGIAS PARA A ESTIMATIVA DE PARÂMETROS BASEADAS NA MIGRAÇÃO E INVERSÃO CONJUNTAS

CAMPINAS
2020

CARLOS ANDRÉ MARTINS DE ASSIS

DEVELOPMENT OF INTEGRAL REPRESENTATIONS FOR THE COUPLED ONE-WAY WAVE EQUATIONS AND STRATEGIES FOR PARAMETER ESTIMATION BASED ON JOINT MIGRATION INVERSION

DESENVOLVIMENTO DAS REPRESENTAÇÕES INTEGRAIS PARA AS EQUAÇÕES UNIDIRECIONAIS DE ONDA ACOPLADAS E ESTRATÉGIAS PARA A ESTIMATIVA DE PARÂMETROS BASEADAS NA MIGRAÇÃO E INVERSÃO CONJUNTAS

Thesis presented to the School of Mechanical Engineering and Institute of Geosciences of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Petroleum Sciences and Enginnering in the area of Reservoirs and Management.

Tese apresentada à Faculdade de Engenharia Mecânica e Instituto de Geociências da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências e Engenharia de Petróleo na área de Reservatórios e Gestão.

Orientador: Prof. Dr. Joerg Dietrich Wilhelm Schleicher

Este exemplar corresponde à versão final da Tese defendida pelo aluno Carlos André Martins de Assis e orientada pelo Prof. Dr. Joerg Dietrich Wilhelm Schleicher.

CAMPINAS

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Luciana Pietrosanto Milla - CRB 8/8129

As76d

Assis, Carlos André Martins de, 1988-
Development of integral representations for the coupled one-way wave equations and strategies for parameter estimation based on joint migration inversion / Carlos André Martins de Assis. – Campinas, SP : [s.n.], 2020.

Orientador: Joerg Dietrich Wilhelm Schleicher.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Mecânica.

1. Geofísica. 2. Teorema de Green. 3. Inversão (Geofísica). 4. Método sísmico de reflexão. 5. Propagação de ondas. I. Schleicher, Joerg Dietrich Wilhelm, 1964-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Mecânica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Desenvolvimento das representações integrais para as equações unidirecionais de onda acopladas e estratégias para a estimativa de parâmetros baseadas na migração e inversão conjuntas
Palavras-chave em inglês:
Geophysics
Green's theorem
Inversion (Geophysics)
Seismic reflection method
Wave propagation
Área de concentração: Reservatórios e Gestão
Titulação: Doutor em Ciências e Engenharia de Petróleo
Banca examinadora:
Joerg Dietrich Wilhelm Schleicher
Jessé Carvalho Costa
Paulo José da Silva e Silva
Reynam da Cruz Pestana
Sérgio Adriano Moura Oliveira
Data de defesa: 07-02-2020
Programa de Pós-Graduação: Ciências e Engenharia de Petróleo

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-1433-9609
- Currículo Lattes do autor: http://lattes.cnpq.br/4657522685033307

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA
E INSTITUTO DE GEOCIÊNCIAS

TESE DE DOUTORADO

DEVELOPMENT OF INTEGRAL REPRESENTATIONS FOR THE COUPLED ONE-WAY WAVE EQUATIONS AND STRATEGIES FOR PARAMETER ESTIMATION BASED ON JOINT MIGRATION INVERSION

DESENVOLVIMENTO DAS REPRESENTAÇÕES INTEGRAIS PARA AS EQUAÇÕES UNIDIRECIONAIS DE ONDA ACOPLADAS E ESTRATÉGIAS PARA A ESTIMATIVA DE PARÂMETROS BASEADAS NA MIGRAÇÃO E INVERSÃO CONJUNTAS

Autor: Carlos André Martins de Assis
Orientador: Prof. Dr. Joerg Dietrich Wilhelm Schleicher

A Banca Examinadora composta pelos membros abaixo aprovou esta Tese:

Prof. Dr. Joerg Dietrich Wilhelm Schleicher, Presidente
DMA/IMECC/UNICAMP

Prof. Dr. Paulo José da Silva e Silva
DMA/IMECC/UNICAMP

Prof. Dr. Jessé Carvalho Costa
IGEOG/UFPA

Prof. Dr. Reynam da Cruz Pestana
DGN/IGEO/UFBA

Prof. Dr. Sérgio Adriano Moura Oliveira
LENEP/CCT/UENF

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no processo de vida acadêmica do aluno.

Campinas, 7 de Fevereiro de 2020.

Para o meu pai Carlos Roberto de Assis e minha mãe Vera Silveira Martins de Assis

Acknowledgements

I am grateful to my parents Vera S. M. de Assis and Carlos R. de Assis, my sister Amanda M. de Assis and my wife Leidiane P. Helmer. My supervisor Jörg Schleicher, all the professors, friends and staff from University of Campinas. My friends Henrique B. Santos and Edwin H. Fagua for sharing their experiences and being supportive.

I also thank Capes/CNPq/DAAD, Karlsruhe Institute of Technology group and ION Geophysical for providing and being part of the special year of 2018. During the exchange program, we also met Michael and Denise Neuerburg that were like family for me and my wife. It was an amazing time. Naturally, special thanks also goes to our friend Bill! Now let me write some words in my mother language.

Trabalhar nesta tese de doutorado foi um conjunto de experiências importantes para mim e sou grato por todas as oportunidades durante a caminhada. De certa forma, poder estudar algum assunto em profundidade sempre foi o que busquei.

Na minha trajetória, meus pais Vera S. M. de Assis e Carlos R. de Assis e minha irmã Amanda M. de Assis são fundamentais. Minha esposa Leidiane P. Helmer e minha sogra Maria P. Merscher também tem um papel de grande importância. Naturalmente, estendo os agradecimentos a todos os meus familiares.

O grupo de professores no Laboratório de Engenharia e Exploração de Petróleo (LENEP) e na Universidade Estadual de Campinas (UNICAMP) também foram de grande importância para a minha formação como um todo. Todo o trabalho desenvolvido desde a graduação no LENEP tem impacto nesta tese de doutorado.

Agradeço ao meu orientador Jörg Schleicher pela compreensão, apoio e paciência no desenvolvimento deste trabalho. Também agradeço aos velhos e novos amigos pela parceria durante esses anos. Durante o desenvolvimento deste trabalho recebi apoio financeiro do PRH-PB 230 no Brasil. E da Capes/CNPq/DAAD processo nº{88887.161391/2017-00} no programa de doutorado sanduíche.

Abstract

We investigate the coupled one-way wave equations and develop strategies for the estimation of model parameters of the subsurface from seismic reflection data acquired near the Earth's surface. In the first part, we develop in detail the one-way wave equations from the equation of motion and the temporal derivative of Hooke's law. Then, we define one-way Green's functions consistently with the two-way counterpart. Next, we derive in an original manner the integral representations of the one-way wave equations. These representations form the conceptual basis of the adopted modeling algorithm. In the second part, we develop strategies to estimate the seismic image and to improve an initial velocity model. All the methodologies for parameter estimation are built upon the nonlinear least-squares method for data fitting. Considering that the known velocity model is sufficiently accurate for migration, we investigate the parameterization of this problem as a function of the reflection coefficients. The expression obtained for the misfit-function gradient suggests a new imaging condition. Still in the context of a known precise velocity model, we propose the impedance parameterization of the migration problem. Then, given an initial homogeneous impedance section, we estimate the relative acoustic impedance. Finally, we include updates of the velocity model in the inversion procedure, which characterizes the joint migration inversion methodology that inspired this work. We propose two regularizing functions based on the image such that the updates of the velocity model can benefit from the high spatial-frequency content in the image. The numerical tests indicate that the investigated methodologies are promising.

Keywords: Geophysics; Green's theorem; Inversion (Geophysics); Seismic reflection method; Wave propagation.

Resumo

Nós investigamos as equações de onda unidirecionais acopladas e estratégias para a estimativa de parâmetros da subsuperfície a partir de dados sísmicos de reflexão medidos próximo à superfície terrestre. Na primeira parte, desenvolvemos em detalhe as equações de onda unidirecionais a partir da equação de movimento e da derivada temporal da lei de Hooke. Então, definimos as funções de Green unidirecionais de maneira consistente com a função equivalente para o caso completo. Na sequência, deduzimos originalmente as representações integrais das equações de onda unidirecionais. As expressões integrais formam a base conceitual do algoritmo de modelagem direta adotado. Na segunda parte, desenvolvemos e discutimos em detalhe estratégias para obter uma estimativa da imagem sísmica e o aprimoramento de um modelo de velocidades inicial, sempre aplicando o método de quadrados mínimos não-linear para o ajuste de dados. Considerando que o modelo de velocidades tem precisão suficiente para a migração, nós revisitamos a parametrização deste problema em função dos coeficientes de reflexão e obtemos uma expressão para o gradiente da função objetivo que sugere uma nova condição de imagem. Além disso, ainda neste contexto de modelo de velocidades suficiente para a migração, nós propomos a parametrização do problema de migração em profundidade em função da impedância. Então, a partir de uma seção de impedância homogênea inicial, estimamos a impedância acústica relativa. Finalmente, incluímos atualizações do modelo de velocidades no procedimento de inversão, o que caracteriza o método de migração e inversão conjuntas que inspirou este trabalho. Nós propomos duas regularizações baseadas na imagem de forma que as atualizações do modelo de velocidades possam se beneficiar da informação de alta-frequência espacial contida na imagem. Os testes numéricos indicam que as metodologias investigadas são promissoras.

Palavras-chaves: Geofísica; Teorema de Green; Inversão (Geofísica); Método sísmico de reflexão; Propagação de ondas.

List of figures

List of tables

List of symbols

Here, we define the symbols, variables and the mathematical conventions adopted. Given a function $f = f(\mathbf{x}, t)$, its Fourier transform of the time variable will be represented by a correspondent uppercase letter $F(\mathbf{x}, \omega)$. Moreover, bold letters represent vectors and matrices. See Appendix A for the Fourier transform convention adopted.

Math symbols

$*$	Denotes complex conjugate
$\dagger$	Denotes complex conjugate and transpose, i.e., the adjoint. Moreover, we use this notation for operators and matrices
$\approx$	Approximately equal
$\|.\|_2$	L^2 norm
$\mathbb{C}$	Set of complex numbers
$\mathbb{R}$	Set of real numbers
t	As a superscript denotes transpose

Basic symbols

$\mathbf{x}$	Spatial position (x, y, z). Depending on the context it can be the two-dimensional case (x, z)
$\mathbf{x}'$	Spatial position with superscript prime indicate a different coordinate
ω	Angular frequency
$\hat{\mathbf{x}}$	Lateral coordinates (x, y)
i	Imaginary unit
k_x	Horizontal wavenumber related to x

k_y Horizontal wavenumber related to y

k_z Vertical wavenumber

t Time

x Horizontal coordinate

y Horizontal coordinate. It is omitted in the two-dimensional case

z Depth coordinate pointing downward

Model parameters

$\kappa(\mathbf{x})$ Compressibility

$\rho(\mathbf{x})$ Mass density

$c(\mathbf{x})$ Compressional wavespeed, or equivalently, velocity model

$R^{\pm}(\mathbf{x})$ Angle-independent reflection coefficient

$R_c^{\pm}(\mathbf{x})$ Angle-independent reflectivity in a continuous model

$T^{\pm}(\mathbf{x})$ Angle-independent transmission coefficient

$T_c^{\pm}(\mathbf{x})$ Angle-independent transmissivity in a continuous model

$Z(\mathbf{x})$ Acoustic impedance

Operators

$\hat{\mathcal{G}}^{\pm}(\mathbf{x}', \omega; \mathbf{x})$ Downgoing $\hat{\mathcal{G}}^{+}$ and upgoing $\hat{\mathcal{G}}^{-}$ Green's functions. They perform extrapolation from a boundary

$\mathcal{G}^{\pm}(\mathbf{x}', \omega; \mathbf{x})$ Downgoing $\mathcal{G}^{+}$ and upgoing $\mathcal{G}^{-}$ Green's functions. They perform integration over the $\mathbb{R}^3$

$\hat{}$ The hat superscript indicates that the operator acts on the lateral coordinate(s)

$\hat{\mathcal{H}}_1(\mathbf{x}, \omega; \hat{\mathbf{x}}')$ Generalized vertical wavenumber or, for short, square-root (of $\hat{\mathcal{H}}_2$) operator

$\hat{\mathcal{H}}_2(\mathbf{x}, \omega; \hat{\mathbf{x}}')$ Transversal Helmholtz operator

$\hat{\boldsymbol{\mathcal{H}}}(\mathbf{x}, \omega; \hat{\mathbf{x}}')$ Matrix of eigenvalues of $\hat{\boldsymbol{\mathcal{A}}}$

$\hat{\boldsymbol{\mathcal{R}}}(\mathbf{x}, \omega; \hat{\mathbf{x}}')$ Matrix of scattering operators in a continuous model

$\hat{\boldsymbol{\mathcal{A}}}(\mathbf{x},\omega;\hat{\mathbf{x}}')$ Two-way matrix

$\hat{\mathcal{R}}^{\pm}(\mathbf{x},\omega;\mathbf{x}')$ Reflection operators defined in a discontinuous model

$\hat{\mathcal{R}}_c^{\pm}(\mathbf{x},\omega;\mathbf{x}')$ Reflectivity operators defined in a continuous model

$\hat{\mathcal{T}}^{\pm}(\mathbf{x},\omega;\mathbf{x}')$ Transmission operators defined in a discontinuous model

$\hat{\mathcal{T}}_c^{\pm}(\mathbf{x},\omega;\mathbf{x}')$ Transmissivity operators defined in a continuous model

$\hat{\boldsymbol{\mathcal{Z}}}(\mathbf{x},\omega;\hat{\mathbf{x}}')$ Matrix of operators that decomposes the two-way vector $\mathbf{U}$ into the one-way vector $\mathbf{P}$

$\hat{\boldsymbol{\mathcal{Z}}}^{-1}(\mathbf{x},\omega;\hat{\mathbf{x}}')$ The inverse of $\hat{\boldsymbol{\mathcal{Z}}}$

Wavefields

$\Lambda^{+}(\mathbf{x},\omega)$ Adjoint downgoing wavefield

$\Lambda^{-}(\mathbf{x},\omega)$ Adjoint upgoing wavefield

$\mathbf{F}(\mathbf{x},\omega)$ Injected force density

$\mathbf{P}$ Column vector with the downgoing and upgoing wavefields, i.e., $[P^{+}\ \ P^{-}]^{t}$

$S^{+}(\mathbf{x},\omega)$ Downgoing one-way source

$\mathbf{U}$ Column vector with the total pressure and vertical particle-displacement velocity, i.e., $[P\ \ V_z]^{t}$

$P(\mathbf{x},\omega)$ Two-way pressure wavefield

$P^{+}(\mathbf{x},\omega)$ Downgoing pressure wavefield

$P^{-}(\mathbf{x},\omega)$ Upgoing pressure wavefield

$Q(\mathbf{x},\omega)$ Injected volume density

$V_z(\mathbf{x},\omega)$ Vertical particle-displacement velocity

Contents

1 Introduction

Reflection seismic data are of great importance for the petroleum exploration industry. After processing, these data provide an approximate image of the subsurface, from which geologists try to infer possible petroleum reservoirs. Recently, seismic data have also been applied to the monitoring of reservoirs and have assisted in the decision making during petroleum production (Blangy et al., 2014).

The main stages in an investigation of the subsurface are the data acquisition, followed by processing and interpretation (Yilmaz, 2001). The seismic reflection method is based on artificial sources, positioned near the Earth surface, generating elastic waves that propagate through the Earth's interior. Then, after reflection in the subsurface, these waves return to the surface and are recorded by some apparatus, e.g., hydrophones in marine data acquisition.

One simple way of classifying the reflection phenomena is by considering the number of occurrences during propagation at those subsurface positions where there is a contrast in the impedance of adjacent rock layers. If only one reflection occurs along the path between the source, the subsurface and the recording instruments, we denominate the corresponding recorded event a primary reflection. On the other hand, if the propagating waves reflect more than once, the recorded event is denominated a multiple reflection. Figure 1.1 sketches this discussion for one seismic experiment.

Commonly, multiple reflections are treated as unwanted data during the processing stage, i.e., noise, (Foster and Mosher, 1992; Verschuur et al., 1992; Innanen, 2017; da Costa Filho et al., 2017). This approach is of practical importance, because assuming that the seismic data is composed of primary reflections only makes the task of building imaging algorithms easier. This is a consequence of the linearity involved in imaging primary reflections, while taking into account multiple reflections is a nonlinear problem.

In practice, recorded multiple reflections are a big challenge, being they classified as noise or useful data (Weglein, 2016). Considering multiple reflections as useful data has the potential of improving seismic imaging results (Davydenko and Verschuur, 2018), because they may have been reflected at positions of the subsurface not reached by primary reflections, Figure 1.1. Another possible beneficial outcome from a methodology that tries

to image multiple reflections, is that at least this information might be removed from the final image, in this way not generating false reflector images.

The use of multiple reflections for seismic imaging and velocity model estimation are an active research topic. Berkhout (2016) indicates that handling multiple reflections will be an important component of the seismic imaging algorithms of the future. It is important to highlight that to build an image of the subsurface, in general, it is necessary to have some knowledge of the velocity model. Hence, in order to include multiple reflections in the imaging stage and at the same time improving an initial velocity model, the joint migration inversion (JMI) methodology was conceived (Staal and Verschuur, 2012b,a; Berkhout, 2012; Verschuur et al., 2016). In this manner, in principle, JMI has the potential to be nearly a complete processing methodology of seismic reflection data.

In JMI, the migration part is a formulation of the least-squares migration problem in the data domain (see, e.g., Nemeth et al., 1999). The inversion part is similar to the reflection waveform inversion (Xu et al., 2012). Furthermore, JMI is entirely based on the separation of the seismic waves into their down- and upgoing components (see, e.g., Wapenaar, 1996). It is noteworthy to emphasize that the down- and upgoing waves are coupled by the reflection phenomenon.

Most of the published works on this topic describe the methodology entirely in the discrete case. We took a few steps back to look into the continuous modeling equations behind JMI, these are the coupled one-way wave equations. Hence, we develop many concepts related to the one-way wave equations already established in the literature (see, e.g., Wapenaar and Berkhout, 1989; Wapenaar, 1996). Furthermore, similarly to Berkhout and Wapenaar (1989), we also use Green's theorem to build a bridge between the up- and downgoing wavefields and the acoustic wave-equation. Naturally, all this investigation led us to a deeper understanding of JMI and new contributions to the methodology.

1.1 Seismic imaging

Seismic imaging methods can be seen as a broad collection of techniques applied in the processing of seismic data. Some methods are: depth migration, time migration and demigration (for a detailed discussion see Schleicher et al., 2007). Here, the nomenclature seismic imaging is used in the context of methods related to depth migration.

In this manner, seismic imaging aims at building an approximate image of the subsurface from the signals recorded at the measuring apparatus as a function of space and time (Robinson, 1986). Thus, considering that the acoustic wave-equation represents the propagation of seismic waves and given a velocity model of the subsurface, the common procedure is to extrapolate the seismic data from the recording positions back to deeper

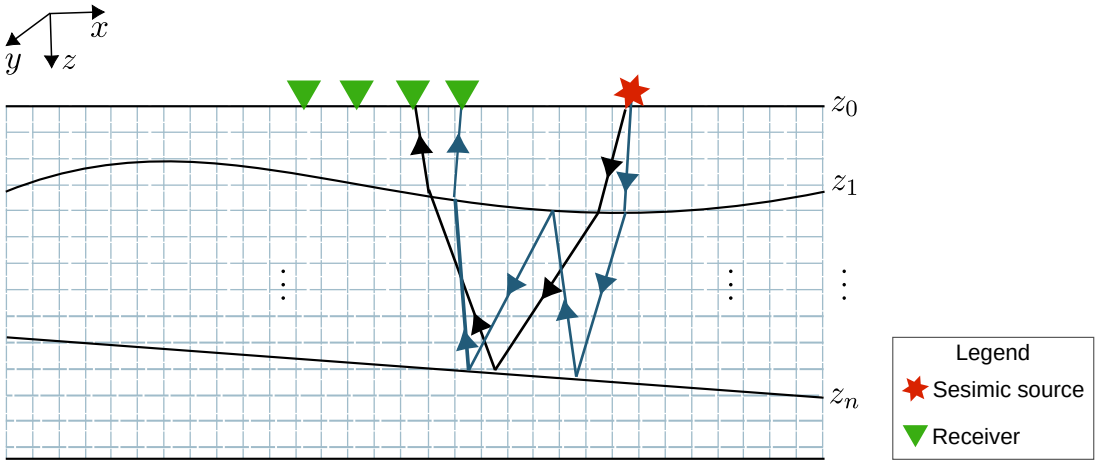


Figure 1.1: Representation of a seismic experiment, i.e., a collection of measurements at receivers that are originated at the seismic source. The black arrows sketch a primary reflection and the blue arrows sketch an internal multiple. The black continuous curves mark the occurrence of discontinuities in the wavespeed and mass density.

positions in the model and check if an image occurs at the output spatial position by means of some kind of image condition (Claerbout, 1971).

Sava and Hill (2009) classify the seismic imaging methods into two big groups, one based on the integral formulation and the other based on the differential formulation. For example, imaging methods based on the Kirchhoff integral are widely used in the petroleum exploration. Commonly, they are implemented via the seismic ray theory, which is a high-frequency approximation to the wave propagation phenomena and provides efficient computational algorithms.

The differential methods largely rely on the acoustic wave-equation, for example, reverse time migration (RTM) (see, e.g., Baysal et al., 1983). Given a sufficiently accurate velocity model, the RTM methodology is one of the most precise approaches to the seismic imaging problem. It has no limitations concerning the velocity model inhomogeneities and steeply dipping geological structures. The wave extrapolation is carried out in time steps, which implies that, for complicated velocity models, it is not straightforward to predict the position of the extrapolation output in the model.

The separation of the wave equation into its down- and upgoing components forms the basis of the one-way wave equations. These equations extrapolate the wavefields in the depth coordinate. As a consequence, it is not necessary to sweep the entire velocity model at each extrapolation step. Besides that, it is possible to discretize the velocity model in a coarser grid than in the RTM methodology. This way, imaging methods based on the one-way wave equations, also called wave-equation-migration (WEM) methods, are usually less computationally demanding than RTM methods. The drawbacks of the methodology are related to the degree of inhomogeneities and maximum layer dip that it

can accurately handle. WEM methods can be implemented in differential form, in integral form or even as a mix of both approaches (see, e.g., Claerbout, 1985; Amazonas et al., 2007).

One common assumption made in the development of most seismic imaging methods is that the data is composed exclusively of primary reflections, i.e., they rely on the first term of the Born series (see, e.g., Bleistein et al., 2001). Hence, in this context, multiple reflections and refracted waves must be removed in a pre-processing stage. Otherwise, they may mask the image provided by the primary reflections or even provide false information in the image.

1.2 Velocity estimation

Seismic imaging depends on an accurate smooth velocity model for the compressional wave. Commonly, the workflow is to estimate the velocity and then apply the imaging, this approach agrees with theoretical observations about the information content of the seismic data. Jannane et al. (1989) demonstrated that there is a gap in the wavenumber spectrum due to the seismic data limited temporal and spatial-frequency content. In practice, this gap translates into a decoupling of the high spatial-frequency features from low spatial-frequency information, which justifies the common practice involved in seismic processing.

Therefore, the scale separation is intrinsic to the recorded seismic data and it is largely explored by migration velocity analysis (MVA) methods. Basically, these methods looks for the velocity model that best optimize some pre-defined image feature. Thus, MVA methods makes use of migrated images to obtain velocity model updates (Sava and Biondi, 2004).

Inversion methods, like full waveform inversion (FWI), are more ambitious. For example, in the acoustic case, the model parameters to be estimated can be set as the wavespeed and mass density. This way, the FWI methodology, in its most basic form and upon linearization of the inverse problem, tries to estimate iteratively updates to these models from the residuals between the nonlinearly modeled data and the observed data (Virieux and Operto, 2009).

Hence, methods based on data fitting pave the way to estimates of the model parameters that contain both the low and high spatial-frequency content. While MVA methods have the information split into models with different resolutions, e.g., the migration velocity model and the migrated image. However, the MVA problem is easier to solve than the data fitting counterpart. For example, the MVA approach allows the starting model to be less accurate (Sava and Biondi, 2004). As a consequence, MVA techniques can

assist in the construction of an initial velocity model for FWI. One approach that tries to combine the best of MVA and data fitting methods is the joint migration inversion (JMI) methodology (Verschuur et al., 2016).

Note that we did not discuss ray-based velocity estimation methods, but they play a key role in the velocity model building workflow (see, e.g., Jones, 2010).

1.3 Joint migration inversion

JMI is based on the separation of the seismic waves into their down- and upgoing components. This approach, intrinsically decouples dynamics (amplitude) from kinematics (events positioning). That is, scattering operators represents the high spatial-frequency information, being responsible for reflection and transmission effects. The velocity model can contain small spatial details, but its main purpose is to propagate the waves. Thus, a smooth and kinematically accurate velocity model, is enough in the one-way framework (Wapenaar and Grimbergen, 1996).

Thus, it is straightforward to realize that JMI is also built upon spatial-scale separation. Besides that, the methodology deals mainly with reflections below the critical angle because the wavefield separation into down- and upgoing components has difficulties in representing horizontally propagating waves (Ursin et al., 2012). In this manner, an approximate analogy with reflection full waveform inversion (RFWI) can be noted. Therefore, accordingly, the migration part of JMI is related to seismic imaging. In its most simple form this part provides approximately the angle-independent reflection coefficients, and the inversion part aims at estimating velocity model updates.

The main weakness of JMI is the fact that it relies on one-way propagators. These propagators are limited in reproducing wave propagation in geological scenarios with steep dips. Moreover, is it not straightforward to implement properly angle-dependent scattering in the data domain. The main advantages of JMI are the ability to handle intrabed multiples even in a smooth velocity model, the correction of transmission effects during imaging and lower computational cost than waveform inversion methods based on finite-differences solution of the wave equation.

It is also noteworthy that velocity estimation methods based on reflected waves, e.g., JMI and RFWI, are able to update deeper parts of the velocity model without much restriction on the source-receive offset (Zhou et al., 2015). Besides that, due to the reflected waves acting as secondary sources, they reduce the problem of measuring data only along a limited portion of the Earth surface, which improves the capability of the inversion part of estimating the velocity model (Mora, 1989).

The modeling procedure involved in JMI is based on the integral representation

of the coupled one-way wave equations. These equations can be expanded in a series, similarly to the Born expansion applied to the acoustic wave-equation, with each term accounting for one scattering order (see, e.g., Wapenaar and Grimbergen, 1996; Berkhout, 2014a). Another feature of the obtained series is that it has the form of a recurrence. For instance, to model seismic reflections, given a downgoing source, the incident downgoing wavefield is modeled. Next, at each model position, the downgoing wavefield is reflected and the upgoing wavefield is calculated. Then, given the source wavefield and the upgoing wavefield, the recurrence starts again with the computation of the downgoing wavefield due to the reflection of the previous upgoing wavefield. In this way, the procedure accounts for one additional scattering order at a time.

This recurrence involving one-way wavefields is closely related to the work of Bremmer (1951). In the unidimensional case, Bremmer started from the zero-order WKBJ solution of the wave equation and, upon consideration that down- and upgoing waves, coupled by reflections, occur in a layered model, showed that the obtained recursive series is a solution of the acoustic wave-equation.

1.4 This work

In the first part, our objective is to review the coupled one-way wave equations for an acoustic medium and interpret its components. They are coupled equations because the reflected upgoing wavefield acts as a secondary source for the downgoing wavefield and vice-versa. Most of the time we will be developing the equations in the three-dimensional case and with variable mass-density.

In preparation for an integral implementation of the modeling equations, we derive the integral representations of the coupled down- and upgoing wavefields from scratch. The objective is to reconcile Green's theorem, commonly applied to wavefields described by the acoustic wave-equation, and the one-way wavefields. In order to do that, we demonstrate that the one-way wavefields obey the acoustic wave-equation. Moreover, in the application of Green's theorem, the one-way wave equations are used to fulfill the required boundary conditions in the surface integral of the volume under consideration.

During the development of integral equations, it is common practice in the seismic literature to impose homogeneous boundary conditions or to apply the Kirchhoff approximation. Hence, the way we build the one-way integral representations is a new approach to this problem. Although the algebraic developments may be cumbersome, they provide plenty of insights about the integral modeling equations commonly used in the JMI method.

After discussing the modeling equations in great detail, our focus turns to the model-

parameters-estimation problem by iterative least-square data fitting. In the implementation of the forward modeling and inversion procedures, we consider the model parameters to be two-dimensional, the mass density is set constant and we assume that scattering is angle-independent. Note that we are using an acoustic formulation, thus, shear and locally converted waves are considered noise.

Seismic inversion is a large-scale problem. Therefore, local optimization methods that rely mainly on gradient information are the common methodology to the model-parameter estimation. Using the Lagrangian multipliers, we formulate the inverse problem in terms of variations of the forward down- and upgoing wavefields, the corresponding adjoint down- and upgoing wavefields, the scattering operators completely parameterized in terms of the reflectivity operator and the velocity model.

Considering that the velocity model is accurate enough for imaging, we reformulate the angle-independent migration part of JMI. The expression obtained for the misfit-function gradient suggests a new imaging condition. We also investigate the imaging problem parameterized as a function of the acoustic impedance. The methodology developed, may replace the common practice of performing seismic imaging followed by the image conversion to impedance. Finally, we propose and verify the effectiveness of two image-based regularizing functions for the velocity estimation part of JMI. The goal is to make use of the high spatial-frequency information contained in the image during the estimation of the velocity model. Although one special feature of JMI is it capability of handling internal multiples, we do not explore this aspect on the numerical examples.

1.5 Thesis outline

This thesis is organized such that the main mathematical and physical ideas are in the body of the text. We did our best to confine the algebraic developments to the appendices, but in some cases it seemed beneficial for the exposition to develop the equations in the chapters. Additionally, each chapter is practically self-contained. Below, we summarize the subject of each chapter and appendix.

Chapter 2 presents the forward modeling equations. Starting from the equation of motion and the constitutive relation for fluids, we develop the coupled one-way wave equations in the differential form. Next, we define the associated Green's functions and, in an original manner, we develop the integral representations. The series expansion of the integral equations provides the expressions that form the basis of the modeling algorithm behind the JMI methodology. We close this chapter presenting the implementation of the forward modeling equations and discussing numerical examples.

Chapter 3 discusses the implementation of the migration part of JMI. The objec-

tive is to estimate the angle-independent reflection coefficient using the Gauss–Newton method. Besides that, the gradient of the misfit function suggests a new imaging condition. We evaluate this new imaging condition in a simple model and demonstrate the proposed least-squares migration. Note that in this Chapter we used interchangeably the terms migration and imaging.

Chapter 4 proposes the parameterization of the migration problem as a function of the acoustic impedance. Additionally, we develop the misfit-function gradient after discretization. Thus, it is possible to appreciate the inverse problem in a different from from the one provided by the Lagrangian multipliers in their continuous form. We also develop an approximation to the Gauss–Newton method, that turns out to be analogous to the concept of the deconvolution imaging condition, but adapted to the impedance parameterization of the migration problem.

Chapter 5 presents two regularization functions to the velocity estimation problem, i.e., the inversion part of JMI. The objective is to make use of the high spatial-frequency information in the image during the velocity estimation.

Chapter 6 summarizes the developments made in this thesis, elaborates on the conclusions, and provides suggestions for future work.

Appendix A presents the convention adopted for the Fourier transform.

Appendix B demonstrates the diagonalization of the two-way matrix operator. We also discuss the decomposition components, the integral representation of the involved operators, and the relationship between one-way and two-way wavefields. We close this appendix demonstrating the two-way wave operator's action on the one-way wavefields and defining the source terms of the one-way Green's functions consistently with the two-way case.

Appendix C defines the scattering operators in a discontinuous model and derives the continuous scattering operators from the first-order Taylor expansion of the corresponding operators in the discontinuous case.

Appendix D defines the Ricker wavelet.

Appendix E demonstrates the reciprocity between the transmitted down- and upgoing Green's functions.

Appendix F discusses the complex Padé Fourier finite-difference method and its implementation.

Appendix G shows the application of the Lagrangian multipliers in the derivation of the adjoint modeling equations and the misfit-function partial derivatives with respect to the model parameters.

2 Forward modeling equations

In this chapter, we develop the forward modeling equations, starting from the equation of motion and the constitutive relation for an acoustic medium.

First, we present the quantities related to the two-way wave propagation, in which there is no distinction of propagation direction. Then, we develop the coupled one-way wave equations, that differentiate between down- and upgoing waves. We give special attention to the physical interpretations of the coupled-equations components.

Next, we develop the corresponding integral representations and wavefield expansions. These integral equations form the basis of the forward modeling algorithm used in this thesis. Furthermore, the way we develop these representations is to the best of our knowledge, a new contribution to the topic under investigation.

Then, we approximate the generalized vertical-wavenumber, here also called for short square-root operator, with the complex Padé Fourier finite-difference method and the related extrapolation procedure is presented. Finally, we define the modeling algorithm and show two applications comparing the one-way modeling method with the more accurate finite-difference solution of the two-way wave equation.

2.1 Acoustic wave-equation

In an acoustic medium, the pressure $p(\mathbf{x}, t)$ and the particle-displacement velocity $\mathbf{v}(\mathbf{x}, t)$ are related by the equation of motion,

$$\nabla p + \rho \frac{\partial \mathbf{v}}{\partial t} = -\mathbf{f}\,, \tag{2.1}$$

and the constitutive relation, i.e., temporal derivative of Hooke's law for fluids, (see, e.g., Rosa, 2018),

$$\nabla \cdot \mathbf{v} + \kappa \frac{\partial p}{\partial t} = -\frac{\partial q}{\partial t}\,, \tag{2.2}$$

where κ is the compressibility, i.e., inverse of the bulk modulus, ρ is the mass density, q is the injected volume density and $\mathbf{f}$ is the injected force density. Applying the Fourier

transform in the time variable, we obtain

$$\nabla P + i\omega\rho\mathbf{V} = -\mathbf{F}\,, \tag{2.3}$$

$$\nabla \cdot \mathbf{V} + i\omega\kappa P = -i\omega Q\,, \tag{2.4}$$

where the capital letters P, Q and $\mathbf{V}$ correspond to the temporal Fourier transform of p, q and $\mathbf{v}$ (for more details about the Fourier transform see Appendix A).

We define the z coordinate as the preferential direction and isolate its derivatives. Moreover, we group the remaining derivatives under the nabla operator $\hat{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ and obtain

$$\hat{\nabla} P = -i\omega\rho\hat{\mathbf{V}}\,, \tag{2.5}$$

$$\frac{\partial P}{\partial z} = -i\omega\rho V_z - F\,, \tag{2.6}$$

$$\frac{\partial V_z}{\partial z} = -i\omega\kappa P - \hat{\nabla}\cdot\hat{\mathbf{V}} - i\omega Q\,, \tag{2.7}$$

where $\hat{\mathbf{V}} = (V_x, V_y)$ is the two-dimensional vector of the horizontal components of $\boldsymbol{V}$ and we considered $\mathbf{F} = (0, 0, F)$.

We isolate $\hat{\mathbf{V}}$ in equation 2.5, substitute in equation 2.7 and rearrange the terms to obtain

$$\frac{\partial P}{\partial z} = -i\omega\rho V_z - F\,, \tag{2.8}$$

$$\frac{\partial V_z}{\partial z} = \frac{1}{i\omega\rho}\left[\frac{\omega^2}{c^2}P + \rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}P\right)\right] - i\omega Q\,, \tag{2.9}$$

where we have substituted $\kappa = 1/\rho c^2$ with c denoting the compressional wavespeed. These equations can be arranged into the matrix-vector product form

$$\frac{\partial}{\partial z}\begin{bmatrix} P \\ V_z \end{bmatrix} = \begin{bmatrix} 0 & -i\omega\rho \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & 0 \end{bmatrix}\begin{bmatrix} P \\ V_z \end{bmatrix} - \begin{bmatrix} F \\ i\omega Q \end{bmatrix}\,, \tag{2.10}$$

where $\hat{\mathcal{H}}_2$ is the transversal Helmholtz operator,

$$\boxed{\hat{\mathcal{H}}_2 = \rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}\right) + \frac{\omega^2}{c^2}\,.} \tag{2.11}$$

We use compact notation to rewrite equation 2.10 as

$$\boxed{\frac{\partial \mathbf{U}}{\partial z} = \hat{\boldsymbol{\mathcal{A}}}\mathbf{U} - \mathbf{Q}\,.} \tag{2.12}$$

In which the hat denotes action over the lateral coordinates and the calligraphic font-style indicates operators. Here, the wave vector $\mathbf{U}$, the two-way matrix operator $\hat{\boldsymbol{\mathcal{A}}}$, and the two-way source vector $\mathbf{Q}$ are given by

$$\mathbf{U} = \begin{bmatrix} P \\ V_z \end{bmatrix}, \tag{2.13}$$

$$\mathbf{Q} = \begin{bmatrix} F \\ i\omega Q \end{bmatrix}, \tag{2.14}$$

$$\hat{\boldsymbol{\mathcal{A}}} = \begin{bmatrix} 0 & -i\omega\rho \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & 0 \end{bmatrix}. \tag{2.15}$$

In this section, we have rearranged the equation of motion and the constitutive equation in such a way that the vertical derivatives of the total pressure and vertical particle-displacement velocity were grouped into a vector. Note that equation 2.12 is nothing but a vectorial representation of the acoustic two-way wave equation.

2.2 Up/downgoing wavefields

Now, we are ready to decompose the two-way wave equation 2.12 into two coupled one-way wave equations for down- and upgoing waves. For this purpose, we substitute the two-way matrix $\hat{\boldsymbol{\mathcal{A}}}$ in equation 2.12 by its diagonalized form, see Appendix B. Then, we present the physical interpretations of the resulting equations and recognize that, for homogeneous media, we have the down- and upgoing waves. At the end of this section, we briefly discuss the downgoing source.

2.2.1 Coupled equations

In Appendix B, aligned with the work of Wapenaar and Grimbergen (1996), we develop the diagonalization of the two-way matrix operator, equation 2.15, in the following form

$$\hat{\boldsymbol{\mathcal{A}}} = i\hat{\boldsymbol{\mathcal{Z}}}^{-1}\hat{\boldsymbol{\mathcal{H}}}\hat{\boldsymbol{\mathcal{Z}}}, \tag{2.16}$$

where $\hat{\boldsymbol{\mathcal{Z}}}$ is a matrix with the eigenvectors of $\hat{\boldsymbol{\mathcal{A}}}$ and $i\hat{\boldsymbol{\mathcal{H}}}$ is a diagonal matrix of eigenvalues. All the elements in these matrices are operators, with the exception of the element $-i\omega\rho$ in matrix $\hat{\boldsymbol{\mathcal{A}}}$, equation 2.15.

These matrices of operators are given by

$$\hat{\boldsymbol{\mathcal{Z}}} = \frac{1}{2}\begin{bmatrix} \mathcal{I} & \hat{\mathcal{Z}} \\ \mathcal{I} & -\hat{\mathcal{Z}} \end{bmatrix}, \tag{2.17}$$

$$\hat{\boldsymbol{\mathcal{Z}}}^{-1} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{Z}}^{-1} & -\hat{\mathcal{Z}}^{-1} \end{bmatrix}, \tag{2.18}$$

$$\hat{\boldsymbol{\mathcal{H}}} = \begin{bmatrix} -\hat{\mathcal{H}}_1 & 0 \\ 0 & \hat{\mathcal{H}}_1 \end{bmatrix}, \tag{2.19}$$

where $\mathcal{I}$ is the identity operator and $\hat{\mathcal{Z}}$ is the impedance operator. Moreover, $\hat{\mathcal{H}}_1$ is the generalized vertical-wavenumber, here also called square-root operator for short, which is related to the transversal Helmholtz operator, equation 2.11, via twofold application, i.e.,

$$\boxed{\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1\,.} \tag{2.20}$$

The operators $i\hat{\mathcal{H}}_1$ and $-i\hat{\mathcal{H}}_1$ are the eigenvalues of the two-way matrix $\hat{\boldsymbol{\mathcal{A}}}$. They differ only by a minus sign which is a consequence of the trace of $\hat{\boldsymbol{\mathcal{A}}}$ being zero. Operator $\hat{\mathcal{H}}_1$ acts on a wavefield through an integral, see Appendix B, Section B.3.

In matrices 2.17 and 2.18 the impedance operator $\hat{\mathcal{Z}}$ and its inverse are defined as

$$\hat{\mathcal{Z}} = \hat{\mathcal{H}}_1^{-1}\omega\rho\,, \tag{2.21}$$

$$\hat{\mathcal{Z}}^{-1} = \frac{1}{\omega\rho}\hat{\mathcal{H}}_1\,. \tag{2.22}$$

We note that in a laterally invariant model and for waves propagating with wavefront perpendicular to the depth axis, $\hat{\mathcal{Z}}$ reduces to the acoustic impedance. In the next section, we discuss this physical interpretation in more detail.

As shown in Appendix B, Section B.4, applying the matrix $\hat{\boldsymbol{\mathcal{Z}}}$ to the wavefield vector $\mathbf{U}$ defines a modified wavefield vector $\mathbf{P}$ by

$$\mathbf{P} = \hat{\boldsymbol{\mathcal{Z}}}\mathbf{U} \qquad \text{or} \qquad \hat{\boldsymbol{\mathcal{Z}}}^{-1}\mathbf{P} = \mathbf{U}\,, \tag{2.23}$$

where $\mathbf{U}$ is given by equation 2.13. At this moment, we know the vector $\mathbf{U}$ and the vector $\mathbf{P}$ results from a change of variables. In Section 2.2.2, we give an interpretation to the components of $\mathbf{P}$.

Substituting the diagonalization of matrix $\hat{\boldsymbol{\mathcal{A}}}$, equation 2.16, in the two-way wave equation 2.12, we obtain

$$\frac{\partial\hat{\boldsymbol{\mathcal{Z}}}^{-1}\mathbf{P}}{\partial z} = i\hat{\boldsymbol{\mathcal{Z}}}^{-1}\hat{\boldsymbol{\mathcal{H}}}\hat{\boldsymbol{\mathcal{Z}}}\hat{\boldsymbol{\mathcal{Z}}}^{-1}\mathbf{P} - \mathbf{Q}\,. \tag{2.24}$$

Applying the product rule to the left-hand side and simplifying the right-hand side, we obtain

$$\frac{\partial \hat{\boldsymbol{\mathcal{Z}}}_m^{-1}}{\partial z}\mathbf{P} + \hat{\boldsymbol{\mathcal{Z}}}^{-1}\frac{\partial \mathbf{P}}{\partial z} = i\hat{\boldsymbol{\mathcal{Z}}}^{-1}\hat{\boldsymbol{\mathcal{H}}}\mathbf{P} - \mathbf{Q}\,, \tag{2.25}$$

where we have used the identity $\hat{\boldsymbol{\mathcal{Z}}}\hat{\boldsymbol{\mathcal{Z}}}^{-1} = \mathcal{I}$. Multiplying by $\hat{\boldsymbol{\mathcal{Z}}}$ and rearranging the terms, we arrive at

$$\boxed{\frac{\partial \mathbf{P}}{\partial z} = i\hat{\boldsymbol{\mathcal{H}}}\mathbf{P} + \hat{\boldsymbol{\mathcal{R}}}\mathbf{P} + \mathbf{S}\,.} \tag{2.26}$$

where the matrices $\mathbf{S}$ and $\hat{\boldsymbol{\mathcal{R}}}$ are given by

$$\mathbf{S} = -\hat{\boldsymbol{\mathcal{Z}}}\mathbf{Q} = \frac{1}{2}\begin{bmatrix} F + i\omega\hat{\mathcal{Z}}Q \\ F - i\omega\hat{\mathcal{Z}}Q \end{bmatrix}\,, \tag{2.27}$$

and

$$\hat{\boldsymbol{\mathcal{R}}} = -\hat{\boldsymbol{\mathcal{Z}}}\frac{\partial \hat{\boldsymbol{\mathcal{Z}}}_m^{-1}}{\partial z} = \frac{1}{2}\begin{bmatrix} -\hat{\mathcal{Z}}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} & \hat{\mathcal{Z}}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} \\ \hat{\mathcal{Z}}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} & -\hat{\mathcal{Z}}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} \end{bmatrix}\,. \tag{2.28}$$

Equation 2.26 is the main result of this section. According to the physical interpretation in the next section, it represents a system of coupled one-way wave equations for the down- and upgoing wavefields.

2.2.2 Physical interpretations

In this section, we interpret the components of equation 2.26. First, we discuss the operator $\hat{\mathcal{H}}_1$ in a laterally invariant model and recognize that in the Fourier domain it is the vertical wavenumber. Then, considering a homogeneous medium, we split the two-way wave equation into two equations and recognize the down- and upgoing waves. Next, using this interpretation and the additional investigation of the elements in the matrix of equation 2.28, we arrive at the coupled down- and upgoing wave equations in the space-frequency domain. Finally, we apply a first-order perturbation to the down- and upgoing wavefields in order to gain more insight into the modeling equations.

Laterally invariant medium

In a laterally invariant model, we can develop the equations of Section 2.1 with the Fourier transform applied over the lateral coordinates and the time variable. As a consequence, the transversal-Helmholtz operator defined equation 2.11, but in the domain

(k_x, k_y, z, ω) is

$$\widetilde{\hat{\mathcal{H}}_2} = \frac{\omega^2}{c^2} - \left(k_x^2 + k_y^2\right) . \tag{2.29}$$

In this equation, we recognize the square of the vertical wavenumber, k_z^2. Thus, as a consequence of the relation $\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1$, the analogous quantity to the square-root operator $\hat{\mathcal{H}}_1$, in a model without lateral variations and in the wavenumber-frequency domain, is

$$\widetilde{\hat{\mathcal{H}}_1} = k_z = \frac{\omega}{c}\sqrt{1 - \frac{c^2}{\omega^2}\left(k_x^2 + k_y^2\right)} , \tag{2.30}$$

where k_z is the vertical wavenumber. Here, we have defined the sign of k_z to equal that of ω, so that $\pm k_z$ describe unique propagation directions. We observe that dividing by $\omega\rho$, we obtain for purely vertical propagation ($k_x = k_y = 0$)

$$\widetilde{\hat{\mathcal{H}}_1}/\omega\rho = k_z/\omega\rho = \frac{1}{\rho c} , \tag{2.31}$$

where on the left-hand side we recognize an analogous quantity to $\hat{\mathcal{Z}}^{-1}$ of equation 2.22. Thus, under the assumptions made in this section, the operator $\hat{\mathcal{Z}}$ reduces to the acoustic impedance. Therefore, we conclude that $\hat{\mathcal{Z}}$, defined in equation 2.21, can be understood as a generalized acoustic-impedance.

Homogeneous medium

The equation of motion 2.3 in the frequency domain can be rearranged into

$$\mathbf{V} = -\frac{1}{i\omega\rho}\nabla P - \frac{1}{i\omega\rho}\mathbf{F} . \tag{2.32}$$

Substituting this expression for $\mathbf{V}$ in the constitutive relation in the frequency domain, equation 2.4, we obtain

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla P\right) + \frac{\omega^2}{c^2}P = -\omega^2\rho Q - \rho\nabla\cdot\left(\frac{1}{\rho}\mathbf{F}\right) . \tag{2.33}$$

Considering the medium to be homogeneous, discarding the source term and applying the Fourier transform in the lateral coordinates, we obtain

$$\frac{\partial^2\widetilde{P}}{\partial z^2} + k_z^2\widetilde{P} = 0 , \tag{2.34}$$

where $k_z^2 = \omega^2/c^2 - k_x^2 - k_y^2$. Since the medium under consideration is homogeneous, we can split this equation according to

$$\left[\frac{\partial}{\partial z} - ik_z\right]\left[\frac{\partial}{\partial z} + ik_z\right]\widetilde{P} = 0\,, \tag{2.35}$$

where the order of the two operators is arbitrary. As a consequence, we see that any solution to one of the equations

$$\frac{\partial \widetilde{P}}{\partial z} = -ik_z\widetilde{P}\,, \tag{2.36}$$

$$\frac{\partial \widetilde{P}}{\partial z} = ik_z\widetilde{P}\,, \tag{2.37}$$

is a solution of equation 2.35.

We include the boundary condition $\widetilde{P}(x, y, z_0, \omega) = \widetilde{P}_0$, integrate over z and apply the inverse temporal Fourier transform, defined in Appendix A, to obtain the solution of both equations as

$$\widetilde{p} = \frac{1}{2\pi}\int_{\mathbb{R}} \widetilde{P}_0 e^{i(\mp k_z(z-z_0)+\omega t)} d\omega\,. \tag{2.38}$$

Considering a wavefront, where the phase must be constant as the wavefields advance in time. The positive sign multiplying k_z must be related to a decrease in z and the negative sign must be related to an increase in z. This way, we reinterpret equations 2.36 and 2.37 as follows

$$\frac{\partial \widetilde{P}^+}{\partial z} = -ik_z\widetilde{P}^+\,, \tag{2.39}$$

$$\frac{\partial \widetilde{P}^-}{\partial z} = ik_z\widetilde{P}^-\,, \tag{2.40}$$

where $\widetilde{P}^+$ is a downgoing wavefield and $\widetilde{P}^-$ is an upgoing wavefield.

Hence, we use the fact that in a laterally invariant model $\widetilde{\hat{\mathcal{H}}_1} = k_z$, as discussed at the beginning of this section, it is clear that these conventions for the extrapolation direction are kept in the application of $\pm i\hat{\mathcal{H}}_1$, embedded in the matrix $i\hat{\boldsymbol{\mathcal{H}}}$ of equation 2.26. And it is also expected that they hold in an inhomogeneous medium. Furthermore, these results indicate that the factors $\pm i\hat{\mathcal{H}}_1$ are mainly related to the kinematics of the wave equations defined in the compact notation of 2.26.

Therefore, in the space-frequency domain, the vector $\mathbf{P}$ in equation 2.26 can be interpreted as

$$\mathbf{P} = \begin{bmatrix} P^+ \\ P^- \end{bmatrix}\,, \tag{2.41}$$

where P^+ is the downgoing wavefield and P^- is the upgoing wavefield. The relationship with the two-way wavefields, equation 2.13, is given by the decomposition 2.23 (see also Section B.4 of Appendix B).

Scattering matrix and coupled one-way equations

From the results of the Appendix C, Section C.2, we interpret $\hat{\boldsymbol{\mathcal{R}}}$, equation 2.28, as a scattering matrix with the elements given by

$$\hat{\boldsymbol{\mathcal{R}}} = \begin{bmatrix} \hat{\mathcal{T}}_c^+ & \hat{\mathcal{R}}_c^- \\ -\hat{\mathcal{R}}_c^+ & -\hat{\mathcal{T}}_c^- \end{bmatrix}, \tag{2.42}$$

where $\hat{\mathcal{R}}_c^\pm$ are reflectivities and $\hat{\mathcal{T}}_c^\pm$ are transmissivities (see also Wapenaar, 1996). These operators have dimension of inverse distance and are defined in a continuous model as

$$\hat{\mathcal{R}}_c^\pm = \hat{\mathcal{T}}_c^\pm = \mp \frac{\hat{\mathcal{H}}_1^{-1}}{2} \rho \frac{\partial}{\partial z} \left(\frac{1}{\rho} \hat{\mathcal{H}}_1 \right). \tag{2.43}$$

The superscript $+$ is related to the downgoing wavefield that impinges on a model position from above, and the superscript $-$ is related to the upgoing waves that impinges on a model position from below.

Linearization of the scattering operators $\hat{\mathcal{R}}^\pm$ and $\hat{\mathcal{T}}^\pm$, see Appendix C, Section C.2, for small displacements $\pm\Delta z/2$ in the depth yields the following relationships

$$\hat{\mathcal{R}}_c^\pm(z) \approx \frac{\hat{\mathcal{R}}^\pm(z + \Delta z/2, z - \Delta z/2)}{\Delta z}, \tag{2.44}$$

$$\hat{\mathcal{T}}_c^\pm(z) \approx \frac{\hat{\mathcal{T}}^\pm(z + \Delta z/2, z - \Delta z/2) - \mathcal{I}}{\Delta z}, \tag{2.45}$$

where $\mathcal{I}$ is the identity operator.

Finally, with the interpretations of this section, the compact notation in equation 2.26 can be explicitly rewritten as

$$\frac{\partial P^+}{\partial z} = -i\hat{\mathcal{H}}_1 P^+ + \underbrace{\hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^-}_{downgoing\ \ secondary\ \ source} + S^+, \tag{2.46}$$

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \underbrace{\hat{\mathcal{T}}_c^- P^- - \hat{\mathcal{R}}_c^+ P^+}_{upgoing\ \ secondary\ \ source} + S^-, \tag{2.47}$$

where $S^\pm$ are the one-way physical sources given by the components of $\mathbf{S}$ in equation 2.27. Note again that $\hat{\mathcal{H}}_1$ and the scattering operators act on the wavefields via integration as discussed in the Appendix B, Section B.3. Moreover, it is important to observe that the

sum $P^+ + P^-$ provides the total pressure wavefield P, see Appendix B, Section B.4.

Equations 2.46 and 2.47 are the partial differential equations that describe the problem under investigation. Equation 2.46 represents the downgoing wave equation and equation 2.47 the upgoing counterpart. Note that we are working in the space-frequency domain $(\mathbf{x} - \omega)$.

An important difference between these equations and the two-way wave equation, besides the directional separation of the wavefields, is the intrinsic decoupling of the propagation effects, represented by $i\hat{\mathcal{H}}_1$, from the scattering effects, represented by $\hat{\mathcal{R}}_c^{\pm}$ and $\hat{\mathcal{T}}_c^{\pm}$. Furthermore, we emphasize that these expressions are coupled one-way wave equations due to the presence of the reflectivities $\hat{\mathcal{R}}_c^{\mp}$ in equations 2.46 and 2.47.

Depth discretization

At first sight, the signs in the right-hand side of equations 2.46 and 2.47 seem to be inconsistent. But in fact, if we consider a first order variation of each wavefield, we obtain

$$\Delta P^+(z) = P^+(z + \Delta z) - P^+(z) \approx \frac{\partial P^+}{\partial z}\Delta z\,, \tag{2.48}$$

$$\Delta P^-(z) = P^-(z - \Delta z) - P^-(z) \approx -\frac{\partial P^-}{\partial z}\Delta z\,, \tag{2.49}$$

where we exhibit only the depth dependence for simplicity. The complete notation is $P^+(z) = P^+(\mathbf{x}, \omega)$. We substitute the vertical derivatives by equations 2.46 and 2.47 and obtain

$$\begin{aligned} P^+(z + \Delta z) \approx\; & P^+(z) - i\hat{\mathcal{H}}_1 P^+(z)\Delta z + \hat{\mathcal{T}}_c^+ \Delta z P^+(z) \\ & + \hat{\mathcal{R}}_c^- \Delta z P^-(z) + S^+(z)\Delta z\,, \end{aligned} \tag{2.50}$$

$$\begin{aligned} P^-(z - \Delta z) \approx\; & P^-(z) - i\hat{\mathcal{H}}_1 P^-(z)\Delta z + \hat{\mathcal{T}}_c^- \Delta z P^-(z) \\ & + \hat{\mathcal{R}}_c^+ \Delta z P^+(z) - S^-(z)\Delta z\,. \end{aligned} \tag{2.51}$$

In the right-hand side of these equations, we recognize the linearization of the reflection and transmission operators, equations 2.44 and 2.45 (see also Section C.2 of Appendix C), that provides

$$P^+(z + \Delta z) \approx -i\hat{\mathcal{H}}_1 P^+(z)\Delta z + \hat{\mathcal{T}}^+ P^+(z) + \hat{\mathcal{R}}^- P^-(z) + S^+(z)\Delta z\,, \tag{2.52}$$

$$P^-(z - \Delta z) \approx -i\hat{\mathcal{H}}_1 P^-(z)\Delta z + \hat{\mathcal{T}}^- P^-(z) + \hat{\mathcal{R}}^+ P^+(z) - S^-(z)\Delta z\,. \tag{2.53}$$

Note that we used the relations

$$\hat{\mathcal{T}}^{+}P^{+} \approx P^{+} + \hat{\mathcal{T}}_{c}^{+}\Delta z P^{+}\,, \tag{2.54}$$

$$\hat{\mathcal{T}}^{-}P^{-} \approx P^{-} + \hat{\mathcal{T}}_{c}^{-}\Delta z P^{-}\,. \tag{2.55}$$

Equations 2.52 and 2.53 provide some interesting insights, but this is not how we implement them. In practice, we use the integral formulation discussed below in Section 2.4.3. We observe that after discretization of the vertical derivatives, we recover the reflection operators and the complete transmission operators. Furthermore, the signals of the right-hand side terms are consistent in both equations 2.52 and 2.53. The difference in the sign of the source terms occurs due to the definition of S^{+} that encapsulates a negative sign. The source terms will be discussed in the next section.

2.2.3 One-way sources

From equation 2.27 and the interpretation in Section 2.2.2, the one-way sources are given by

$$\begin{bmatrix} S^{+} \\ S^{-} \end{bmatrix} = \begin{bmatrix} F - \frac{i\omega^2}{2}\hat{\mathcal{H}}_1^{-1}\rho Q \\ F + \frac{i\omega^2}{2}\hat{\mathcal{H}}_1^{-1}\rho Q \end{bmatrix}\,. \tag{2.56}$$

In this thesis, we will work only with downgoing physical sources. This way, we impose $S^{-} = 0$ and obtain

$$F = -\frac{i\omega^2}{2}\hat{\mathcal{H}}_1^{-1}\rho Q\,. \tag{2.57}$$

The substitution of this result in the definition of S^{+} in equation 2.56 provides

$$S^{+}(\mathbf{x},\omega) = -i\omega^2\hat{\mathcal{H}}_1^{-1}\rho Q(\mathbf{x},\omega)\,. \tag{2.58}$$

Moreover, we define the source term Q as a point source according to

$$Q(\mathbf{x},\omega;\mathbf{x}_s) = \frac{1}{2\Delta z\omega^2\rho}\delta(\mathbf{x}-\mathbf{x}_s)C(\omega)\,, \tag{2.59}$$

where the factor $C(\omega)$ describes the source wavelet spectrum, e.g. the Ricker wavelet defined in Appendix D. Moreover,we included the factor $2\Delta z\omega^2\rho$ to make the downgoing source consistent with the two-way source, as discussed in the Appendix B, Section B.5.1.

The substitution of the definition of Q in equation 2.58 provides

$$\boxed{S^{+}(\mathbf{x},\omega;\mathbf{x}_s) = -\frac{i}{2\Delta z}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}_s)C(\omega)\,.} \tag{2.60}$$

The weighting factor multiplying $C(\omega)$ in equation 2.60 compensates the factor $i\hat{\mathcal{H}}_1$ related

to extrapolation in equation 2.46, which provides a monopole source (Wapenaar, 1990).

We performed a simple experiment to verify the amplitude behavior of a wavefront due to equation 2.58, in which we considered $Q = \delta(\mathbf{x} - \mathbf{x}_s)C(\omega)/\rho$. We observed that in a homogeneous medium, the ω^2 factor can be omitted without compromising much the wavefront amplitude behavior and this choice may avoid bursting high-frequency noise. Figure 2.1 shows the amplitude pattern of the monopole, monopole without ω^2 and dipole, i.e., $S^+ = \delta(\mathbf{x} - \mathbf{x}_s)C(\omega)$. The amplitude of the dipole source presents greater focus than the monopole and the monopole without the factor ω^2 is very close to the exact monopole response.

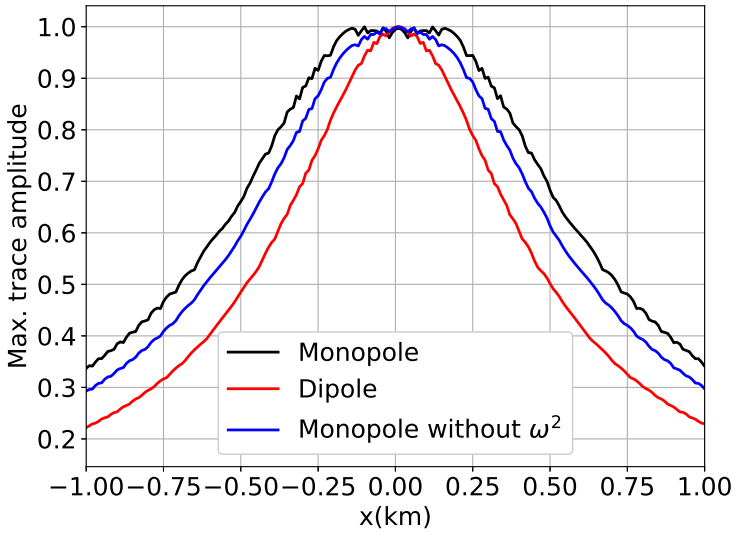


Figure 2.1: Maximum amplitude of each trace for a punctual source extrapolated 500 m in a homogeneous medium with wavespeed equal to 1.0 km/s. The wavelet used was a 10 Hz Ricker.

2.3 Integral representations

In this section, we develop the integral representations of the downgoing P^+ and upgoing P^- wavefields. First, we define the decoupled one-way Green's functions consistent with their two-way counterparts. Then, we proceed to the integral representation in an unbounded space. Next, we develop integral equations for extrapolation from a boundary. These results form the basis of the modeling algorithm used in this thesis.

2.3.1 Decoupled Green's functions

We neglect the reflectivity terms in the coupled one-way wave equations 2.46 and 2.47, and obtain the decoupled downgoing and upgoing equations as follows

$$\frac{\partial P^+}{\partial z} = -i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{T}}_c^+ P^+ + S^+ \,, \tag{2.61}$$

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \hat{\mathcal{T}}_c^- P^- \,, \tag{2.62}$$

where the source term S^+ is given by equation 2.60.

These equations are very similar to the results of Zhang et al. (2005) and Amazonas et al. (2010). They used zero-order ray theory approximation to correct the amplitude behavior of the decoupled one-way wave equations. Their correction term is very similar to the transmissivities $\hat{\mathcal{T}}_c^{\pm}$ above.

In the Appendix B, Section B.5, we define the decoupled Green's functions G_0^+ and G_0^- such that their sum provides the two-way Green's function G_0. They must obey the following equations

$$\frac{\partial G_0^+}{\partial z} = -i\hat{\mathcal{H}}_1 G_0^+ + \hat{\mathcal{T}}_c^+ G_0^+ - \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,, \tag{2.63}$$

$$\frac{\partial G_0^-}{\partial z} = i\hat{\mathcal{H}}_1 G_0^- - \hat{\mathcal{T}}_c^- G_0^- + \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{2.64}$$

The associated one-way wave operators are

$$\mathcal{L}^{\pm} = \frac{\partial}{\partial z} \pm i\hat{\mathcal{H}}_1 \mp \hat{\mathcal{T}}_c^{\pm} \,. \tag{2.65}$$

It is also beneficial to define the Green's functions that are the scaled inverses of the one-way wave operators as follows

$$\mathcal{L}^+ G^+ = \Delta z \delta(\mathbf{x}-\mathbf{x}') \,, \tag{2.66}$$

$$\mathcal{L}^- G^- = -\Delta z \delta(\mathbf{x}-\mathbf{x}') \,. \tag{2.67}$$

In Section 2.2.2, we have shown that the operator $\hat{\mathcal{H}}_1$ is related to the vertical wavenumber. For this reason, we introduced the constant Δz, that represents an infinitesimal variation in the vertical direction, into equations 2.66 and 2.67 to maintain the dimensional consistency between $G^{\pm}$ and $G_0^{\pm}$.

Thus, from the definitions of $G^{\pm}$ and $G_0^{\pm}$ we deduce their relationship as

$$2i\hat{\mathcal{H}}_1 G_0^{\pm} = -\frac{1}{\Delta z} G^{\pm} \,, \tag{2.68}$$

where we considered that the operators $\mathcal{L}^{\pm}$ commute with the scaling factors of each Green's function source due to the fact that the wave operators are acting on the $\mathbf{x}$ coordinate and the scaling factor of the source terms act on $\mathbf{x}'$.

2.3.2 Representation in an unbounded space

Now we are looking for an integral representation of the down- and upgoing coupled partial differential equations 2.46 and 2.47. Using the one-way operator, equation 2.65, we have

$$\mathcal{L}^{+}G^{+} = \Delta z \delta(\mathbf{x}-\mathbf{x}')\,, \tag{2.69}$$

$$\mathcal{L}^{+}P^{+} = \left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x}',\omega) + S^{+}(\mathbf{x}',\omega;\mathbf{x}_s)\,. \tag{2.70}$$

We multiply equation 2.69 by the source term of 2.70, integrate over the $\mathbb{R}^3$ and obtain

$$\begin{aligned}\mathcal{L}^{+}\int_{\mathbb{R}^3} G^{+}(\mathbf{x},\omega;\mathbf{x}')\left[\left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x}',\omega) + S^{+}(\mathbf{x}',\omega;\mathbf{x}_s)\right]dV' \\ = \Delta z\int_{\mathbb{R}^3}\delta(\mathbf{x}-\mathbf{x}')\left[\left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x}',\omega) + S^{+}(\mathbf{x}',\omega;\mathbf{x}_s)\right]dV'\,.\end{aligned} \tag{2.71}$$

Using the sifting property of the delta distribution, we obtain

$$\begin{aligned}\mathcal{L}^{+}\int_{\mathbb{R}^3} G^{+}(\mathbf{x},\omega;\mathbf{x}')\left[\left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x}',\omega) + S^{+}(\mathbf{x}',\omega;\mathbf{x}_s)\right]dV' = \\ \Delta z\left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x},\omega) + \Delta z S^{+}(\mathbf{x},\omega;\mathbf{x}_s)\,.\end{aligned} \tag{2.72}$$

We compare this result with equation 2.70 and conclude that

$$P^{+}(\mathbf{x},\omega) = \frac{1}{\Delta z}\int_{\mathbb{R}^3} G^{+}(\mathbf{x},\omega;\mathbf{x}')\left[\left(\hat{\mathcal{R}}_c^{-}P^{-}\right)(\mathbf{x}',\omega) + S^{+}(\mathbf{x}',\omega;\mathbf{x}_s)\right]dV'\,. \tag{2.73}$$

This is the integral representation of the downgoing wavefield in an unbounded space. In order to represent equation 2.73 compactly, we introduce the Green's function operator $\mathcal{G}^{+}$ as follows

$$\boxed{P^{+} = \mathcal{G}^{+}\left[\hat{\mathcal{R}}_c^{-}P^{-} + S^{+}\right]\,.} \tag{2.74}$$

An analogous result is obtained for the upgoing wavefield and given by

$$\boxed{P^{-} = \mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}P^{+}\,.} \tag{2.75}$$

In which the upgoing source term is absent as discussed in Section 2.2.3.

2.3.3 Representation in a half-space

Considering two wavefields A and B, that obey the two-way wave equation in two states, Green's theorem (see, e.g., Schleicher et al., 2007) states that

$$\int_{\mathcal{V}} A\nabla\cdot\left(\frac{1}{\rho}\nabla B\right) - B\nabla\cdot\left(\frac{1}{\rho}\nabla A\right) dV = \int_{\partial\mathcal{V}} \frac{1}{\rho}\left[A\nabla B - B\nabla A\right]\cdot\mathbf{n}dS\,, \tag{2.76}$$

where $\mathcal{V}$ is an arbitrary volume, $\partial\mathcal{V}$ is the volume's surface and $\mathbf{n}$ is a unitary vector pointing outward the surface.

We want to relate a downgoing wavefield P^+ and an upgoing Green's function G_0^- in a volume $\mathcal{V}$ in which the mass density and wavespeed are vertically homogeneous. The region outside $\mathcal{V}$ may have any degree of inhomogeneity.

In the Appendix B, Section B.5, we demonstrate that the one-way wavefields obey the two-way wave operator in an inhomogeneous medium. Therefore, applying the two-way wave operator to the one-way wavefields P^+ and G_0^-, we obtain

$$\nabla\cdot\left(\frac{1}{\rho}\nabla P^+\right) + \frac{\omega^2}{\rho c^2}P^+ = \frac{1}{\rho}\frac{\partial S^+(\mathbf{x}^{out},\omega)}{\partial z} - \frac{1}{\rho}i\hat{\mathcal{H}}_1 S^+(\mathbf{x}^{out},\omega)\,, \tag{2.77}$$

$$\nabla\cdot\left(\frac{1}{\rho}\nabla G_0^-\right) + \frac{\omega^2}{\rho c^2}G_0^- = \frac{i}{2\rho}\hat{\mathcal{H}}_1^{-1}\frac{\partial\delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2\rho}\delta(\mathbf{x}-\mathbf{x}')\,, \tag{2.78}$$

in which we considered that the model parameters are vertically homogeneous in a small region around the downgoing source. The upgoing Green's function source term is deduced in Appendix B, Section B.5.1. It is defined such that

$$\mathcal{L}G_0^+ + \mathcal{L}G_0^- = -\delta(\mathbf{x}-\mathbf{x}')\,, \tag{2.79}$$

in which the sum $G_0^+ + G_0^-$ provides the total Green's function G_0 and $\mathcal{L}$ is the two-way wave operator. The same relation holds for the pressure wavefields, i.e., $P^+ + P^- = P$, see also Appendix B.

Furthermore, we design a configuration with the upgoing impulsive source at $\mathbf{x}'$ inside the integration volume and it originates $G_0^-(\mathbf{x},\omega;\mathbf{x}')$ in equation 2.78. The source of the downgoing wavefield P^+ in equation 2.77 is positioned above the integration volume.

Figure 2.2 sketches this configuration. It is similar to the Kirchhoff integral representation of transmitted two-way wavefields (see, e.g., Schleicher et al., 2007), but using one-way wavefields and also taking into consideration the possible reflection, at depth z'', of an upgoing wavefield. We highlight that the combination of the downgoing wavefield that arrives at z'' and a downgoing Green's function G_0^+ with source at $\mathbf{x}'$ would not work, because G_0^+ would vanish at z''.

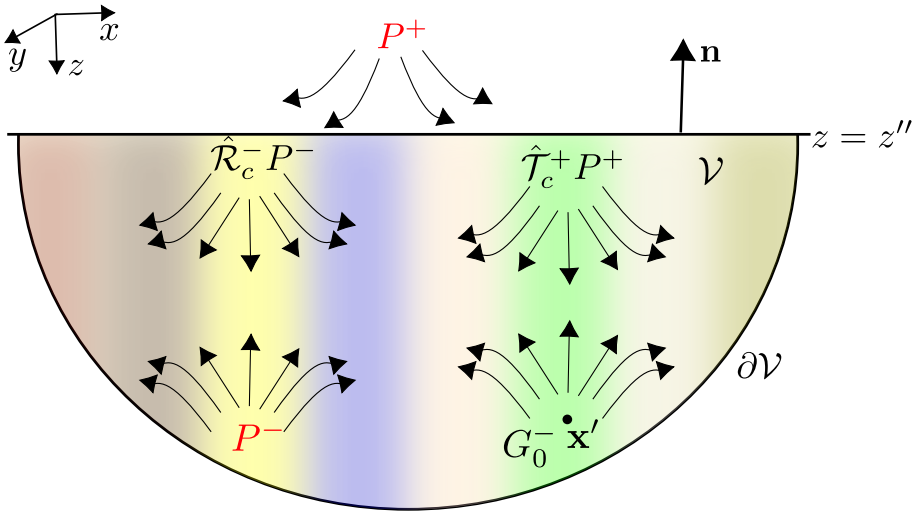


Figure 2.2: Closed surface $\partial\mathcal{V}$ for the application of Green's theorem, in which $\mathcal{V}$ is the integration volume, $\mathbf{x}'$ is the spatial position of the upgoing impulsive source and P^- is an upgoing wavefield. In the developments, we use an approximate reciprocity between the up- and downgoing Green's functions to turn the position $\mathbf{x}'$ into the observation point. At depth z'' a reflector may occurs, where the transmissivity acts on the downgoing wavefield $\hat{\mathcal{T}}_c^+ P^+$ and the reflectivity acts on the upgoing wavefield $\hat{\mathcal{R}}_c^- P^-$.

In this manner, we substitute $A = P^+$ and $B = G_0^-$ into equation 2.76 and, accordingly, the volume integral is given by

$$I_V = \int_{\mathcal{V}} -P^+(\mathbf{x},\omega)\frac{1}{2\rho(\mathbf{x})}\delta(\mathbf{x}-\mathbf{x}') + P^+(\mathbf{x},\omega)\frac{i\hat{\mathcal{H}}_1^{-1}}{2\rho(\mathbf{x})}\frac{\partial\delta(\mathbf{x}-\mathbf{x}')}{\partial z}dV\,, \tag{2.80}$$

where the wavefield G_0^- is absent because the downgoing source is outside the integration volume. Applying the sifting property of the delta distribution, we obtain

$$I_V = -\frac{P^+(\mathbf{x}',\omega)}{2\rho(\mathbf{x}')} - \frac{i}{2}\frac{\partial}{\partial z}\left(\frac{\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x})}P^+(\mathbf{x},\omega)\right)\Bigg|_{\mathbf{x}=\mathbf{x}'}. \tag{2.81}$$

or equivalently, we have

$$I_V = -\frac{P^+(\mathbf{x}',\omega)}{2\rho(\mathbf{x}')} - \frac{i}{2}\frac{\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x}')}\frac{\partial P^+(\mathbf{x}',\omega)}{\partial z} - \frac{i}{2}\frac{\partial}{\partial z}\left(\frac{\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x})}\right)\Bigg|_{\mathbf{x}=\mathbf{x}'} P^+(\mathbf{x}',\omega)\,. \tag{2.82}$$

The substitution of the vertical derivative of the downgoing wavefield by the corresponding

one-way wave equation 2.46 provides

$$I_V = -\frac{P^+(\mathbf{x}',\omega)}{2\rho(\mathbf{x}')} - \frac{i\hat{\mathcal{H}}_1^{-1}}{2\rho(\mathbf{x}')}\left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^-\right]_{\mathbf{x}=\mathbf{x}'} - \frac{i}{2}\frac{\partial}{\partial z}\left(\frac{\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x})}\right)\Bigg|_{\mathbf{x}=\mathbf{x}'} P^+(\mathbf{x}',\omega)\,, \tag{2.83}$$

where the source term of the downgoing wavefield was not included because it is outside the integration volume. The transmissivity and reflectivity are composed of vertical derivatives of ρ and $\hat{\mathcal{H}}_1$, see Appendix C, Section C.2. Besides that, the vertical derivatives are evaluated at position $\mathbf{x}'$ that is inside the volume, and there we considered the medium without vertical variations. In this way, simplifying the vertical derivatives and rearranging terms, we obtain

$$I_V = -\frac{P^+(\mathbf{x}',\omega)}{\rho(\mathbf{x}')}\,. \tag{2.84}$$

This is the left-hand side of equation 2.76. The complete equation is

$$-\frac{P^+(\mathbf{x}',\omega)}{\rho(\mathbf{x}')} = \int_{\partial\mathcal{V}} \frac{1}{\rho(\mathbf{x})}\left[P^+(\mathbf{x},\omega)\nabla G_0^-(\mathbf{x},\omega;\mathbf{x}') - G_0^-(\mathbf{x},\omega;\mathbf{x}')\nabla P^+(\mathbf{x},\omega)\right]\cdot\mathbf{n}dxdy\,. \tag{2.85}$$

Until now, the shape of the integration volume $\mathcal{V}$ was arbitrary. At this point, we consider the particular geometry of Figure 2.2. We extend the lower semicircle to the infinity and apply Sommerfeld's radiation conditions (see, e.g., Schleicher et al., 2007). Moreover, we change the direction of the vector $\mathbf{n}$ and use the identity $\nabla\cdot\mathbf{n} = \partial/\partial z$ at $\partial\mathcal{V}_0$. In this manner, we obtain

$$\boxed{-\frac{P^+(\mathbf{x}',\omega)}{\rho(\mathbf{x}')} = \int_{\partial\mathcal{V}_0}\frac{1}{\rho}\left[P^+\frac{\partial G_0^-}{\partial z} - G_0^-\frac{\partial P^+}{\partial z}\right]dxdy\,,} \tag{2.86}$$

where $\partial\mathcal{V}_0$ is the $\mathbb{R}^2$ plane at z''. Rearranging the mass-density factor, we obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} P^+\frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})}\frac{\partial G_0^-}{\partial z} - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})}G_0^-\frac{\partial P^+}{\partial z}dxdy\,. \tag{2.87}$$

This is an important point of the development. We need to define the boundary conditions for our problem in order to proceed. Commonly, the options for the boundary conditions are related to the requirement of homogeneity of the wavefield or its derivative at the boundary (see, e.g., Bleistein, 1984). Another option is to use the Kirchhoff approximation (see, e.g., Schleicher et al., 2007).

Here, we follow a different path. We use the one-way wave equations 2.46 and 2.64 to define the inhomogeneous boundary conditions of our problem. We replace the vertical derivatives by the correspondent one-way wave equations and obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} P^+ \left[i\hat{\mathcal{H}}_1 G_0^- - \hat{\mathcal{T}}_c^- G_0^- + \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \right] - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^- \right] dxdy\,, \tag{2.88}$$

where $\hat{\mathcal{T}}_c^{\pm}$ are the transmissivity operators acting on the wavefields and the term $\hat{\mathcal{R}}_c^- P^-$ is an upgoing wavefield inside the integration volume that reflected at the boundary at z''. The support of the delta distribution is inside the volume at position $\mathbf{x}'$, therefore it does not contribute to the surface integral at z'' and we obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} P^+ \left[i\hat{\mathcal{H}}_1 G_0^- - \hat{\mathcal{T}}_c^- G_0^- \right] - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^- \right] dxdy\,. \tag{2.89}$$

We use the identity $\hat{\mathcal{T}}_c^+ = -\hat{\mathcal{T}}_c^-$ and the transpose of the transmissivities, see Appendix G, Section G.2.2, and obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} P^+ i\hat{\mathcal{H}}_1 G_0^- - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} \left(\hat{\mathcal{T}}_c^-\right)^t P^+ G_0^- - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{R}}_c^- P^- \right] + \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} \left(\hat{\mathcal{T}}_c^-\right)^t G_0^- P^+ dxdy\,. \tag{2.90}$$

The simplification of the terms involving the transmissivity, yields

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} P^+ i\hat{\mathcal{H}}_1 G_0^- - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{R}}_c^- P^- \right] dxdy\,. \tag{2.91}$$

The square-root operator is symmetric (Wapenaar and Grimbergen, 1996), i.e. $\hat{\mathcal{H}}_1^t = \hat{\mathcal{H}}_1$. Applying this property to the first term, we obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- i\hat{\mathcal{H}}_1 P^+ - \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[-i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{R}}_c^- P^- \right] dxdy\,. \tag{2.92}$$

The sum of the first and second term provides

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^- \left[2i\hat{\mathcal{H}}_1 P^+ - \hat{\mathcal{R}}_c^- P^-\right] dxdy\,. \tag{2.93}$$

Restoring the independent variables, we obtain

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} \frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^-(\mathbf{x},\omega;\mathbf{x}') \left(2i\hat{\mathcal{H}}_1 P^+ - \hat{\mathcal{R}}_c^- P^-\right)(\mathbf{x},\omega) dxdy\,. \tag{2.94}$$

In the Appendix E, we demonstrate the reciprocity between the decoupled Green's functions given by

$$\frac{\rho(\mathbf{x}')}{\rho(\mathbf{x})} G_0^-(\mathbf{x},\omega;\mathbf{x}') = G_0^+(\mathbf{x}',\omega;\mathbf{x})\,. \tag{2.95}$$

The substitution of this reciprocity relation in equation 2.94 provides

$$P^+(\mathbf{x}',\omega) = -\int_{\partial\mathcal{V}_0} G_0^+(\mathbf{x}',\omega;\mathbf{x}) \left(2i\hat{\mathcal{H}}_1 P^+ - \hat{\mathcal{R}}_c^- P^-\right)(\mathbf{x},\omega) dxdy\,, \tag{2.96}$$

or equivalently, we have

$$\begin{aligned} P^+(\mathbf{x}',\omega) = &-\int_{\partial\mathcal{V}_0} 2i\hat{\mathcal{H}}_1 G_0^+(\mathbf{x}',\omega;\mathbf{x}) P^+(\mathbf{x},\omega) dxdy \\ &+\int_{\partial\mathcal{V}_0} G_0^+(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{R}}_c^- P^-\right)(\mathbf{x},\omega) dxdy\,. \end{aligned} \tag{2.97}$$

From equation 2.68, we have the following identity

$$2i\hat{\mathcal{H}}_1 G_0^+ = -\frac{1}{\Delta z} G^+\,, \tag{2.98}$$

where G^+ is the one-way Green's function that is the scaled inverse of the corresponding one-way wave operator, see equation 2.69. Substituting this relation in the first term of equation 2.97, we obtain

$$\begin{aligned} P^+(\mathbf{x}',\omega) = &\int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^+(\mathbf{x}',\omega;\mathbf{x}) P^+(\mathbf{x},\omega) dxdy \\ &+\int_{\partial\mathcal{V}_0} G_0^+(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{R}}_c^- P^-\right)(\mathbf{x},\omega) dxdy\,. \end{aligned} \tag{2.99}$$

In compact notation, we have

$$\boxed{P^+ = \hat{\mathcal{G}}^+ P^+ + \hat{\mathcal{G}}_0^+ \hat{\mathcal{R}}_c^- P^-\,.} \tag{2.100}$$

This result is insightful. It shows that the total downgoing wavefield observed at position $\mathbf{x}'$ is the sum of the extrapolated downgoing wavefield from depth z'' plus the

extrapolation of the reflected downgoing wavefield. Moreover, the difference between G^+ and G_0^+ is in the source radiation pattern, see Section 2.3.1, and both account for the transmissivity effect. The fact that after reflection, the downgoing wavefield $\hat{\mathcal{R}}_c^- P^-$ is extrapolated by the Green's function that corrects for the radiation pattern, similarly to the impulsive source term, reinforces the idea that reflected wavefields are in fact secondary sources.

Analogously, for the upgoing wavefield, we obtain

$$P^-(\mathbf{x}',\omega) = \int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^-(\mathbf{x}',\omega;\mathbf{x}) P^-(\mathbf{x},\omega) dxdy + \int_{\partial\mathcal{V}_0} G_0^-(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{R}}_c^+ P^+\right)(\mathbf{x},\omega) dxdy\,, \tag{2.101}$$

which, in compact notation can be represented as

$$\boxed{P^- = \hat{\mathcal{G}}^- P^- + \hat{\mathcal{G}}_0^- \hat{\mathcal{R}}_c^+ P^+\,.} \tag{2.102}$$

We observe that to extrapolate the wavefields from (x, y, z'') to (x', y', z'), the integration over the horizontal plane $x - y$ needs to be performed at depth z'', equations 2.99 and 2.101. In the two-dimensional case, the extrapolation is performed via integration over the horizontal line x.

2.3.4 Specification to extrapolation

Now we specify the decoupled Green's functions for the case of pure extrapolation. Accordingly, the Green's functions are defined in a model without vertical wavespeed contrasts at the integration boundary. Thus, the transmissivity operators $\hat{\mathcal{T}}_c^{\pm}$ can be neglected, allowing us to recast the associated decoupled one-way equations 2.63 and 2.64 into the form

$$\frac{\partial G_0^+}{\partial z} = -i\hat{\mathcal{H}}_1 G_0^+ - \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}')\,, \tag{2.103}$$

$$\frac{\partial G_0^-}{\partial z} = i\hat{\mathcal{H}}_1 G_0^- + \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}')\,. \tag{2.104}$$

The associated Green's functions G^+ and G^- are again defined with the delta distribution weighted by Δz, see Section 2.3.1. Therefore, starting the development from the boundary integral in equation 2.87 and following the same steps used to obtain the representations

in equations 2.99 and 2.101, we obtain

$$P^+(\mathbf{x}',\omega) = \int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^+(\mathbf{x}',\omega;\mathbf{x}) P^+(\mathbf{x},\omega) dxdy + \int_{\partial\mathcal{V}_0} G_0^+(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^-\right)(\mathbf{x},\omega) dxdy\,, \tag{2.105}$$

$$P^-(\mathbf{x}',\omega) = \int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^-(\mathbf{x}',\omega;\mathbf{x}) P^-(\mathbf{x},\omega) dxdy + \int_{\partial\mathcal{V}_0} G_0^-(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}_c^- P^- + \hat{\mathcal{R}}_c^+ P^+\right)(\mathbf{x},\omega) dxdy\,. \tag{2.106}$$

In order to account for possible discontinuities in the wavespeed model at the integration boundary, we use the approximations to the scattering operators developed in Appendix C, Section C.2, given by

$$\hat{\mathcal{R}}_c^{\pm}(z) \approx \frac{\hat{\mathcal{R}}^{\pm}(z+\Delta z/2, z-\Delta z/2)}{\Delta z}\,, \tag{2.107}$$

$$\hat{\mathcal{T}}_c^{\pm}(z) \approx \frac{\hat{\mathcal{T}}^{\pm}(z+\Delta z/2, z-\Delta z/2) - \mathcal{I}}{\Delta z} = \frac{\hat{\mathcal{R}}^{\pm}(z+\Delta z/2, z-\Delta z/2)}{\Delta z}\,. \tag{2.108}$$

The substitution of these expressions in equations 2.105 and 2.106 provides

$$P^+(\mathbf{x}',\omega) = \int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^+(\mathbf{x}',\omega;\mathbf{x}) P^+(\mathbf{x},\omega) dxdy + \frac{1}{\Delta z}\int_{\partial\mathcal{V}_0} G_0^+(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}^+ P^+ - P^+ + \hat{\mathcal{R}}^- P^-\right)(\mathbf{x},\omega) dxdy\,, \tag{2.109}$$

$$P^-(\mathbf{x}',\omega) = \int_{\partial\mathcal{V}_0} \frac{1}{\Delta z} G^-(\mathbf{x}',\omega;\mathbf{x}) P^-(\mathbf{x},\omega) dxdy + \frac{1}{\Delta z}\int_{\partial\mathcal{V}_0} G_0^-(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}^- P^- - P^- + \hat{\mathcal{R}}^+ P^+\right)(\mathbf{x},\omega) dxdy\,. \tag{2.110}$$

We define Green's operators that represent these integrals and obtain the following compact notation

$$P^+(\mathbf{x}',\omega) = \hat{\mathcal{G}}^+(\mathbf{x}',\omega;\mathbf{x}) P^+(\mathbf{x},\omega) + \hat{\mathcal{G}}_0^+(\mathbf{x}';\mathbf{x}) \left(\hat{\mathcal{T}}^+ P^+ - P^+ + \hat{\mathcal{R}}^- P^-\right)(\mathbf{x},\omega)\,, \tag{2.111}$$

$$P^-(\mathbf{x}',\omega) = \hat{\mathcal{G}}^-(\mathbf{x}',\omega;\mathbf{x}) P^-(\mathbf{x},\omega) + \hat{\mathcal{G}}_0^-(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}^- P^- - P^- + \hat{\mathcal{R}}^+ P^+\right)(\mathbf{x},\omega)\,. \tag{2.112}$$

In analogy to the integral representation commonly developed for scattered two-way wavefields (see, e.g., Schleicher et al., 2007), we recognize that the terms $\hat{\mathcal{G}}^{\pm} P^{\pm}$ are similar to the idea of an incident wavefield. That is, they are extrapolated wavefields without

interaction with scatterers. The terms extrapolated with $\hat{\mathcal{G}}_0^{\pm}$ are related to the interaction of the wavefields with possible scatterers at the depth level of the boundary integration.

Again, the difference between $\hat{\mathcal{G}}^{\pm}$ and $\hat{\mathcal{G}}_0^{\pm}$ is in the source radiation pattern. If the source terms of the Green's operators $\hat{\mathcal{G}}_0^{\pm}$ already account for the monopole radiation pattern, we may adopt the approximation

$$\hat{\mathcal{G}}_0^{\pm} \approx \hat{\mathcal{G}}^{\pm} \,. \tag{2.113}$$

Under this simplifying assumption, we rewrite equations 2.111 and 2.112 as

$$P^{+}(\mathbf{x}',\omega) = \hat{\mathcal{G}}^{+}(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}^{+}P^{+} + \hat{\mathcal{R}}^{-}P^{-}\right)(\mathbf{x},\omega)\,, \tag{2.114}$$

$$P^{-}(\mathbf{x}',\omega) = \hat{\mathcal{G}}^{-}(\mathbf{x}',\omega;\mathbf{x}) \left(\hat{\mathcal{T}}^{-}P^{-} + \hat{\mathcal{R}}^{+}P^{+}\right)(\mathbf{x},\omega)\,. \tag{2.115}$$

The integral representations given by equations 2.114 and 2.115 form the basis of the extrapolation scheme in this thesis. They are also the components of the so-called full-wavefield modeling proposed by Berkhout (2014a). The only missing detail is the separation of the scattering order. It is developed in the next section.

2.4 Separating scattering orders

From the integral representations in an unbounded space, equations 2.74 and 2.75, we separate the scattering orders using a series expansion. First, we define the incident and scattered wavefields. Then, the series expansion is developed. Finally, from the results for boundary extrapolation, Section 2.3.4, and the series expansion discussed in the present section, we define the general recursive modeling equations that allows discontinuities in the wavespeed and mass density.

2.4.1 Scattered wavefields

The total downgoing wavefield is defined as

$$P^{+} = \mathcal{G}^{+}\left(\hat{\mathcal{R}}_c^{-}P^{-} + S^{+}\right)\,, \tag{2.116}$$

where the Green's operator realizes a volume integral, equation 2.73.

We define the incident downgoing wavefield P_0^{+} as the term decoupled from the upgoing wavefield and obtain

$$\boxed{P_0^{+} = \mathcal{G}^{+}S^{+}\,.} \tag{2.117}$$

The subtracting of equation 2.117 from 2.116 provides the scattered downgoing wavefield

δP^+,

$$\boxed{\delta P^+ = \mathcal{G}^+ \hat{\mathcal{R}}_c^- P^- \,.} \tag{2.118}$$

This definition is consistent with the one used in the development of the boundary extrapolation, equation 2.99. Moreover, the relation between the downgoing wavefields is

$$P^+ = P_0^+ + \delta P^+ \,. \tag{2.119}$$

Similarly, the total upgoing wavefield is given by

$$P^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ P^+ \,. \tag{2.120}$$

Consider a seismic reflection experiment with sources and receivers distributed near the Earth's surface. The only way of creating an upgoing incident wavefield is by means of a secondary upgoing source. Thus, we define the incident upgoing wavefield from the reflection of the incident downgoing wavefield as

$$\boxed{P_0^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ P_0^+ \,.} \tag{2.121}$$

The subtraction of this result from equation 2.120, results in the associated scattered upgoing wavefield

$$\boxed{\delta P^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ \delta P^+ \,,} \tag{2.122}$$

where

$$P^- = P_0^- + \delta P^- \,. \tag{2.123}$$

The incident $P_0^\pm$ and scattered wavefields $\delta P^\pm$ form the basis of the series expansion in the discussion that follows.

2.4.2 Series expansion

The total down- and upgoing wavefields, equations 2.116 and 2.120, account for full scattering. Now, we want to obtain a series expansion of the wavefields $P^\pm$ that recursively accounts for higher scattering orders. Starting with the scattered downgoing and the total upgoing wavefield, we have

$$\delta P^+ = \mathcal{G}^+ \hat{\mathcal{R}}_c^- P^- \,, \tag{2.124}$$

$$P^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ P^+ \,. \tag{2.125}$$

We substitute the second equation into the first one and obtain

$$\delta P^{+} = \mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}P^{+}\,. \tag{2.126}$$

The separation of the total downgoing wavefield as an incident part plus a scattered part, i.e., $P^{+} = P_0^{+} + \delta P^{+}$, and the reorganization of the previous result provides

$$\left(\mathcal{I} - \mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}\right)\delta P^{+} = \mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}P_0^{+}\,. \tag{2.127}$$

Using the inverse of the left-hand side coefficient to isolate δP^{+} and recognizing the definition of $P_0^{-} = \mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}P_0^{+}$, equation 2.121, we obtain

$$\delta P^{+} = \left[\mathcal{I} - \mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}\right]^{-1}\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}P_0^{-}\,. \tag{2.128}$$

The expansion of the inverse coefficient in a power series provides

$$\delta P^{+} = \sum_{j'=0}^{\infty}\left[\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}\right]^{j'}\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}P_0^{-}\,. \tag{2.129}$$

Kennett (1974) denoted this expansion as the ray expansion. In terms of the total downgoing wavefield in equation 2.119, we have

$$P^{+} = P_0^{+} + \sum_{j'=0}^{\infty}\left[\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}\right]^{j'}\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}P_0^{-}\,. \tag{2.130}$$

The infinite sum is not realizable in practice, therefore we truncate it at some term $j+1$, i.e.,

$$\boxed{P_{j+1}^{+} \approx P_0^{+} + \sum_{j'=0}^{j+1}\left[\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}\mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}\right]^{j'}\mathcal{G}^{+}\hat{\mathcal{R}}_c^{-}P_0^{-}\,,} \tag{2.131}$$

where j starts from zero.

We have already observed that the incident downgoing wavefield P_0^{+} after reflection turns into the incident upgoing wavefield P_0^{-}, equation 2.121, and that the scattered downgoing wavefield δP^{+} after reflection provides the scattered upgoing wavefield δP^{-}, equation 2.122.

Hence, for any value j of the downgoing wavefield, and using the relation between the total wavefields, equation 2.120, the series representation of the upgoing wavefield can be defined as

$$\boxed{P_{j+1}^{-} = \mathcal{G}^{-}\hat{\mathcal{R}}_c^{+}P_{j+1}^{+}\,.} \tag{2.132}$$

For example, considering scattering up to second order, $j = 0$, the down and upgoing

wavefields are given by

$$P_1^+ = P_0^+ + \mathcal{G}^+ \hat{\mathcal{R}}_c^- P_0^- \,, \tag{2.133}$$

$$P_1^- = P_0^- + \mathcal{G}^- \hat{\mathcal{R}}_c^+ P_1^+ \,, \tag{2.134}$$

with the incident wavefields given by

$$P_0^+ = \mathcal{G}^+ S^+ \,, \tag{2.135}$$

$$P_0^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ P_0^+ \,, \tag{2.136}$$

where P_0^- models primary reflections.

Generalizing these results, we arrive at the following recursive formulas

$$\boxed{P_{j+1}^+ = \mathcal{G}^+ \hat{\mathcal{R}}_c^- P_j^- + P_0^+ \,,} \tag{2.137}$$

$$\boxed{P_{j+1}^- = \mathcal{G}^- \hat{\mathcal{R}}_c^+ P_{j+1}^+ \,,} \tag{2.138}$$

where we extended j to start from -1, with the associated initial condition that at the beginning of each iteration $P_{j+1}^+ = 0$.

2.4.3 Recursive modeling

Finally, we use the specification of the Green's operators to extrapolation of Section 2.3.4, adapt the recursive volume integrals in equations 2.137 and 2.138 to a sequence of boundary extrapolations and obtain

$$P_{j+1}^+(\hat{\mathbf{x}}', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(\hat{\mathbf{x}}', z_{n+1}, \omega; \hat{\mathbf{x}}, z_n) \left(\hat{\mathcal{R}}^- P_j^- + \hat{\mathcal{T}}^+ P_{j+1}^+ + S^+ \Delta z \right) (\hat{\mathbf{x}}, z_n, \omega) \,, \tag{2.139}$$

$$P_{j+1}^-(\hat{\mathbf{x}}', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(\hat{\mathbf{x}}', z_{n-1}, \omega; \hat{\mathbf{x}}, z_n) \left(\hat{\mathcal{R}}^+ P_{j+1}^+ + \hat{\mathcal{T}}^- P_{j+1}^- \right) (\hat{\mathbf{x}}, z_n, \omega) \,, \tag{2.140}$$

where j is related to the scattering order, $\hat{\mathbf{x}} = (x, y)$ are the lateral coordinates, z_n is the position of the nth boundary, $z_{n\pm1} = z_n \pm \Delta z$ and $n = 0, 1, 2, 3, \ldots, N$ is the boundary number. The source term is discussed in Section 2.2.3. Note that in these equations we used the transmission and reflection operators due to possible discontinuities in the model.

The action of the Green's operators $\hat{\mathcal{G}}^+$ and $\hat{\mathcal{G}}^-$ is defined in more detail in Section 2.3.4. Besides that, these Green's operators are based on the wavefields of equations 2.63 and 2.64, with the additional assumption that between extrapolation levels the wavespeed

and mass-density models are vertically homogeneous and there are no contrasts at the boundaries. Also observe that these expressions are similar to the depth discretized partial differential equations describing down- and upgoing wavefields, equations 2.52 and 2.53.

Figure 2.3 sketches the quantities in the modeling equations. At some depth levels the discretization may not coincide with discontinuities in the model parameters. In these cases, the action of the transmission operator will be given by the identity operator and the reflection operator will vanish (for more details about the scattering operators see Appendix C, Section C.1).

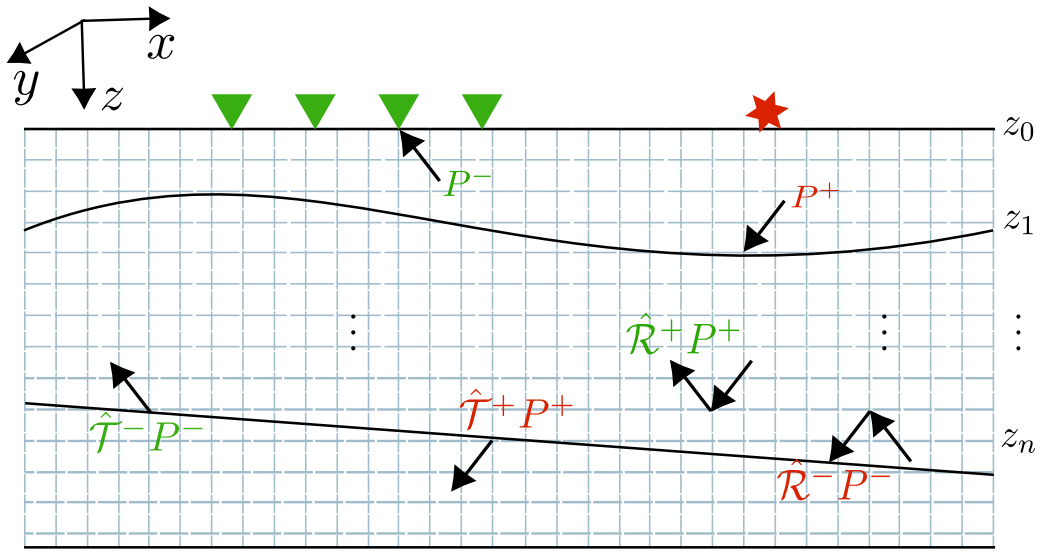


Figure 2.3: Representation of the wavefields in the forward modeling integral equations. The continuous lines represent boundaries between layers with different model parameters, i.e., mass density and velocity. The dashed lines represent the model cells related to the discretization. In red we have downgoing quantities and in green the upgoing counterpart. The triangles indicate receiver positions and the star is the source position.

The recursion starts from the downgoing wavefield, with $j = -1$ and $P^-_{-1} = 0$, which provides

$$P^+_0(\hat{\mathbf{x}}', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(\hat{\mathbf{x}}', z_{n+1}, \omega; \hat{\mathbf{x}}, z_n)\left(\hat{\mathcal{T}}^+ P^+_0 + S^+\Delta z\right)(\hat{\mathbf{x}}, z_n, \omega)\ . \tag{2.141}$$

Initially, P^+_0 is zero and starts to be different from zero at the source position. After calculating P^+_0, the upgoing wavefield is calculated by means of

$$P^-_0(\hat{\mathbf{x}}', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(\hat{\mathbf{x}}', z_{n-1}, \omega; \hat{\mathbf{x}}, z_n)\left(\hat{\mathcal{R}}^+ P^+_0 + \hat{\mathcal{T}}^- P^-_0\right)(\hat{\mathbf{x}}, z_n, \omega)\,. \tag{2.142}$$

Then, the scattering index is updated to $j = 0$ and the procedure returned to the downgoing wavefield, equation 2.139. It is only necessary to keep the upgoing wavefield from $j = -1$, i.e., P^-_0. Higher scattering orders are simulated by repeating this procedure and

keeping the upgoing wavefield the same between iterations.

2.5 Practical aspects of the two-dimensional square-root operator

The transversal Helmholtz operator, equation 2.11, for constant mass density and in a two-dimensional model is given by

$$\hat{\mathcal{H}}_2 = \frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2}\,. \tag{2.143}$$

By definition, it can be replaced by twofold application of the square-root operator $\hat{\mathcal{H}}_1$, i.e.,

$$\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1\,. \tag{2.144}$$

Hence, we can formally write $\hat{\mathcal{H}}_1$ as

$$\hat{\mathcal{H}}_1 = \left[\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2}\right]^{1/2}\,. \tag{2.145}$$

In practice, the square root of the nabla operator makes the task of evaluating this equation very difficult (Claerbout, 1985). Therefore, in this thesis we use the complex Padé Fourier finite-difference (CPFFD) method to approximate equation 2.145. This method provides wide-angle propagation accuracy in the one-way framework and deals adequately with evanescent waves (Amazonas et al., 2007), making it a stable one-way propagator.

The CPFFD approximation of the square-root operator, equation 2.145, is the following

$$\hat{\mathcal{H}}_1 \approx \underbrace{\frac{\omega}{c_r}\sqrt{1+p^2X^2}}_{phase-shift} + \underbrace{\frac{\omega}{c_r}C_0\,(p-1)}_{split-step} + \underbrace{\frac{\omega}{c_r}\sum_{n=1}^{N}\frac{p(1-p)A_nX^2}{1+\sigma B_nX^2}}_{CPFD}\,, \tag{2.146}$$

where C_0, A_n and B_n are the complex Padé coefficients, $p = c_r/c$ with c_r being a reference wavespeed, and $X^2 = (c/\omega)^2\partial^2/\partial x^2$ (for more details see Appendix F). Moreover, we approximate σ as

$$\sigma \approx 1 + p - p^3\,. \tag{2.147}$$

The three terms in equation 2.146 are: the phase-shift term, which can be applied in the wavenumber domain; and the split-step; and finite-difference terms, which must be applied in the space domain. These terms are named after the solution of the partial differential equation associated with the one-way propagation, discussed in the next section. Here, the CPFFD approximation is used only in the extrapolation part of the down- and

upgoing wavefields.

2.5.1 Wavefield propagation scheme

Consider, as an example, the partial differential equation related to the downgoing wavefield in a vertically homogeneous medium, i.e.,

$$\frac{\partial P^{+}(x,z,\omega)}{\partial z} = -i\hat{\mathcal{H}}_1 P^{+}(x,z,\omega)\,. \tag{2.148}$$

The wave extrapolation is made by cascading the solutions corresponding to each term in equation 2.146. This is possible due to the exponential form of the solution. First, we apply the phase-shift and split-step solutions

$$P^{+}_{ps}(k_x,z+\Delta z,\omega) = P^{+}(k_x,z,\omega)\exp\left\{-i\frac{\omega}{c_r}\sqrt{1-\frac{c_r^2k_x^2}{\omega^2}}\Delta z\right\}\,, \tag{2.149}$$

$$P^{+}_{ps,ss}(x,z+\Delta z,\omega) = P^{+}_{ps}(x,z+\Delta z,\omega)\exp\left\{-i\frac{\omega}{c_r}C_0(p-1)\Delta z\right\}\,, \tag{2.150}$$

where Δz is the depth extrapolation step. Next, the finite-difference term is obtained by solving the difference equation

$$\begin{aligned}&\left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n\ +i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P^{+}_{i,cpffd}(x,z+\Delta z,\omega) = \\ &\left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n - i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P^{+}_{i,ps,ss}(x,z+\Delta z,\omega)\,,\end{aligned} \tag{2.151}$$

where the subscript i indicates the x-axis grid and D_x is the centered FD approximation of the second derivative in the x coordinate. For $n=1$ the wavefield on the right-hand side is obtained from 2.150.

We will consider three terms in the finite-difference part, i.e., $N=3$ in equation 2.146. In this way, each time equation 2.151 is solved, the new solution composes the right-hand side in subsequent terms, $n=2$ and $n=3$. Moreover, the linear system provided by equation 2.151 is solved efficiently with a tridiagonal solver.

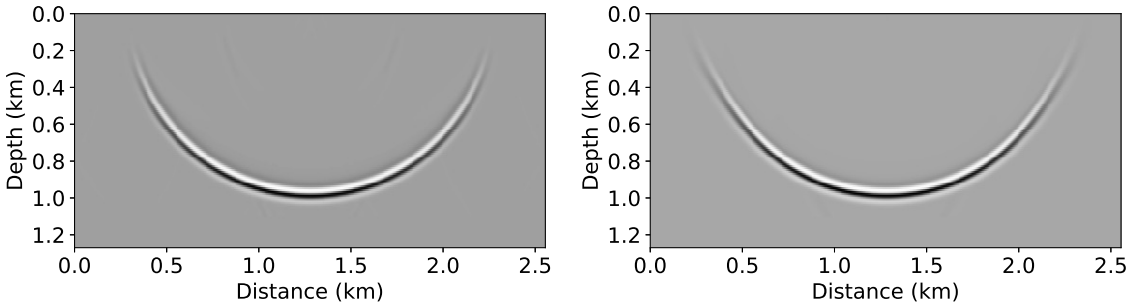
For more details about the CPFFD method and its implementation see Appendix F.

2.5.2 Propagation impulse response

Now, we evaluate the impulse response of the CPFFD method. The experiment was performed in a homogeneous medium with wavespeed of 2.0 km/s. Figure 2.4(a) shows the phase-shift result, that is considered the best case scenario. Figure 2.4(b) shows the CPFFD result with c_r =1.0 km/s, i.e., $p=0.5$, that simulates a strong wavespeed

variation.

The CPFFD result is very close to the phase-shift impulse response. This observation is confirmed qualitatively by visual inspection of the wavefront position shown in Figure 2.5(a). And quantitatively by the absolute error below 5% in most part of the wavefront as shown in Figure 2.5(b). The large errors at both ends are related to high propagation angles and may be acceptable in practical applications considering layers with mild dipping-angle. Note that waves propagating at angles higher than approximately 70° are not described by the one-way propagators.



(a) (b)

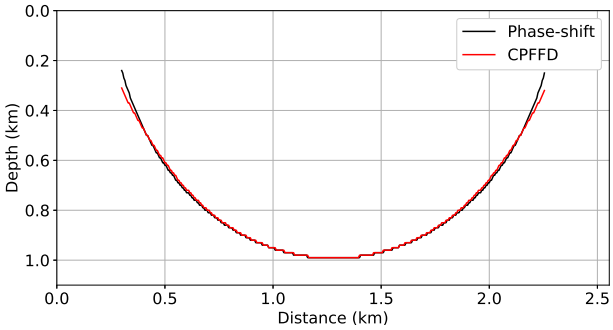
Figure 2.4: Impulse response. The initial wavefield was built with a Ricker wavelet at 25 Hz and positioned at $(1.25, 0.0)$. The wavespeed was set constant and equal to 2.0 km/s. (a) Phase-shift; (b) CPFFD with c_r =1.0 km/s.

2.6 Modeling algorithm

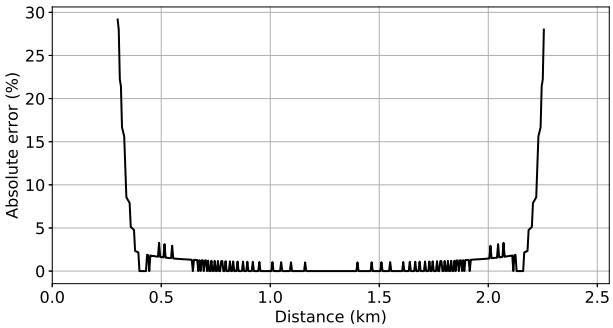
2.6.1 Further assumptions and algorithm

Although most part of the presentation until Section 2.5 was made in three dimensions, the final algorithm implemented in this thesis is two dimensional. The additional simplifying assumptions behind the implemented methodology are

- Horizontal propagation and evanescent waves are not modeled due to the directional decoupling and the square-root operator approximation, Section 2.5;
- Mass density is sufficiently well-approximated by a constant. Except in Chapter 4, in which we investigate the impedance parameterization of the migration problem;
- Scattering is angle-independent. Therefore, after reflection or transmission, the wavefield is scaled by the corresponding coefficient from normal incidence;
- Layer dips can be ignored. Note that dipping layers can be easily included in the reflection and transmission coefficients if available.



(a)



(b)

Figure 2.5: Same experiment as in Figure 2.4. (a) Maximum amplitude of each trace in the impulse responses of the phase-shift and CPFFD methods; (b) Absolute percentage error considering phase-shit result as the reference.

The scattering operators in a constant density medium are given by

$$\hat{\mathcal{R}}_c^{\pm} = \hat{\mathcal{T}}_c^{\pm} = \mp\frac{1}{2}\hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\,. \tag{2.152}$$

In terms of the reflection and transmission operators, see Appendix C, Section C.2, we have

$$\hat{\mathcal{R}}^{\pm} \approx \hat{\mathcal{R}}_c^{\pm}\Delta z \qquad \text{and} \qquad \hat{\mathcal{T}}^{\pm} \approx \hat{\mathcal{T}}_c^{\pm}\Delta z + \mathcal{I}\,, \tag{2.153}$$

where $\mathcal{I}$ is the identity operator.

Similarly to the approximation of the vertical wavenumber in equation 2.31, the angle-dependence of the scattering operators is neglected, i.e., the square-root operator is approximated as

$$\hat{\mathcal{H}}_1 \approx \frac{\omega}{c}\,. \tag{2.154}$$

Therefore, substituting this approximation in equation 2.152 and using the relations

in 2.153, the reflection and transmission coefficients are given by

$$R^{+} = \frac{c_l - c_u}{c_l + c_u} \qquad \text{and} \qquad T^{+} = \frac{2c_l}{c_l + c_u}\,, \tag{2.155}$$

$$R^{-} = \frac{c_u - c_l}{c_l + c_u} \qquad \text{and} \qquad T^{-} = \frac{2c_u}{c_l + c_u}\,, \tag{2.156}$$

where $c_u(x,z) = c(x, z - \Delta z/2)$ and $c_l(x,z) = c(x, z + \Delta z/2)$. In order to apply these equations in an inhomogeneous medium, we calculate them during modeling for each grid point. Hence, using these definitions of scattering operators and the discrete form of the modeling equations 2.139 and 2.140, we have

$$P^{+}_{j+1}(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^{+}(x', z_{n+1}, \omega; x, z_n)\left(R^{-}P^{-}_{j} + T^{+}P^{+}_{j+1} + S^{+}\Delta z\right)(x, z_n, \omega)\,, \tag{2.157}$$

$$P^{-}_{j+1}(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^{-}(x', z_{n-1}, \omega; x, z_n)\left(R^{+}P^{+}_{j+1} + T^{-}P^{-}_{j+1}\right)(x, z_n, \omega)\,, \tag{2.158}$$

where $\hat{\mathcal{G}}^{\pm}$ represent the Green's functions acting as wave extrapolation operators using the CPFFD method, Section 2.5. S^{+} represents the seismic monopole source, defined in equation 2.60.

Algorithm 1 sketches the modeling procedure for the downgoing wavefield discussed above, the upgoing wavefield implementation is analogous. Each time the down- and upgoing algorithms are called, one order of scattering is added. Furthermore, between iterations it is only necessary to keep the downgoing wavefield P^{-}. See Section 2.4.3, for a review of the concepts behind the algorithm from a continuous to a discontinuous model.

2.6.2 Practical comparison with finite difference

To get a better understanding of the one-way modeling algorithm, discussed in Section 2.6.1, we compared it to the finite difference (FD) method applied to the two-way wave equation. In the following discussions, the FD data are considered the reference result and any differences are regarded as the inaccuracies of the one-way method. Two models are used to illustrate the methods, one is a Lens-shaped model (Masaya and Verschuur, 2018) and the other is the modified Marmousi2 (Pan et al., 2018). In both experiments, the data were modeled with a Ricker wavelet with peak frequency of 20 Hz.

Lens model

Figure 2.6 shows the exact wavespeed model used in our first numerical test. It is composed of a background wavespeed of 2.0 km/s, the lens with wavespeed of 2.5 km/s and fine-layering at the bottom. Figure 2.7 shows the amplitude versus angle (AVA) behavior of the lens-top. It is a strong amplitude increase.

We modeled one common shot-gather in the middle of the model using FD and the

Algorithm 1: Downgoing wavefield modeling procedure.

Data: $C(\omega)$, $P_j^-(z,\omega,x)$, $R^+(z,x)$, $c(z,x)$
Result: $P_{j+1}^+(z,\omega,x)$
$S^+[\omega,x] = -\frac{i\hat{\mathcal{H}}_1^{-1}(\omega,x)}{2\Delta z}\delta(x-x_s)\delta(z-z_s)C[\omega]$; ▷ Monopole eq.2.60
$S^+[\omega,x] = HalfIntgeral(S^+[\omega,x])$; ▷ Correct 2D phase distortion
$P_{j+1}^+ = zeros((nz,nf,nx))$; ▷ P_{j+1}^+[z-axis,frequency,x-axis]
$Q^+ = zeros((nf,nx))$; ▷ Secondary source
$T^+ = 1 + R^+$; ▷ Transmission operator
$iz = iz_s$; ▷ Depth counter
while $iz < nz$ **do**
  $ifr = 0$; ▷ Frequency counter
  while $ifr < nf$ **do**
    $ix = 0$; ▷ Lateral position counter
    while $ix < nx$ **do**
      $Q^+[ifr,ix] = R^-[iz,ix]P_j^-[iz,ifr,ix] + T^+[iz,ix]P_{j+1}^+[iz,ifr,ix]$;
      $ix \mathrel{+}= 1$;
    $ifr \mathrel{+}= 1$;
  if $iz == iz_s$ **then**
    $Q^+ \mathrel{+}= \Delta z S^+$;
  $P_{j+1}^+[iz+1,:,:] \mathrel{+}= Extrapolate(Q^+, c[iz,:], \Delta z)$; ▷ CPFFD
  $iz \mathrel{+}= 1$;

one-way algorithm studied in this thesis. The data were recorded at the depth level of 37.5 m at all lateral grid points, Figure 2.8. The AVA effect for the first event is very clear in the FD data, Figure 2.8(a). The one-way data do not display the same amplitude behavior, as expected, but have similar kinematic shape, Figure 2.8(b).

Figure 2.9 shows traces at offsets corresponding to zero, 512 m and 987 m. The zero-offset trace is very similar in both modeling methods, Figure 2.9(a). Small amplitude differences occur probably due to the stronger diffractions in the FD data, Figure 2.8(a). The amplitude differences increase with the offset, Figures 2.9(b) and 2.9(c). The amplitude differences were expected due to the normal incidence transmission and reflection coefficients used in the one-way approach.

The kinematic errors are associated with large propagation angles. Note that from the acquisition surface to the lens-top, the distance is approximately 0.4 km. Therefore, for the offset of 512 m the incidence angle is approximately 51 degrees. And at the offset of 987 m it is approximately 67 degrees, that can be considered a high incidence-angle.

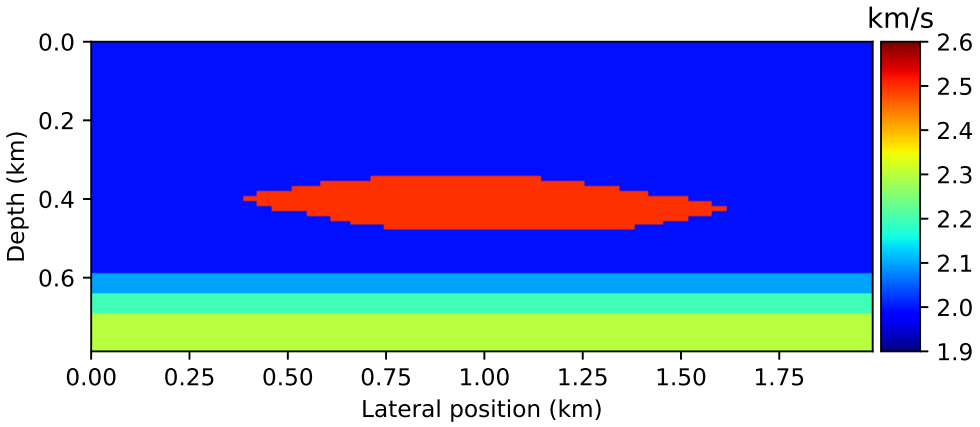


Figure 2.6: Lens-shaped model exact wavespeed.

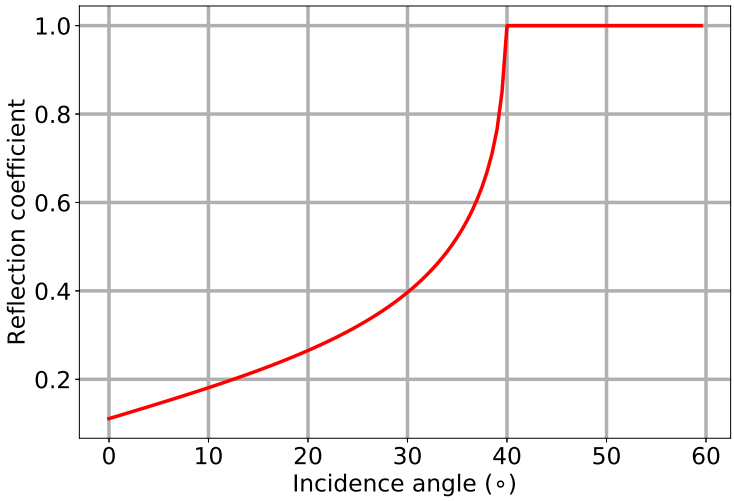


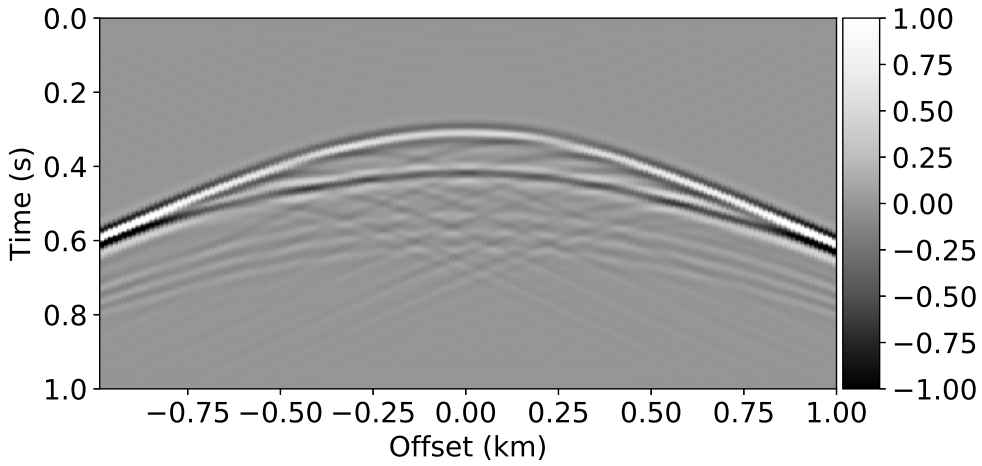
Figure 2.7: Lens-top amplitude versus angle curve.

Marmousi2

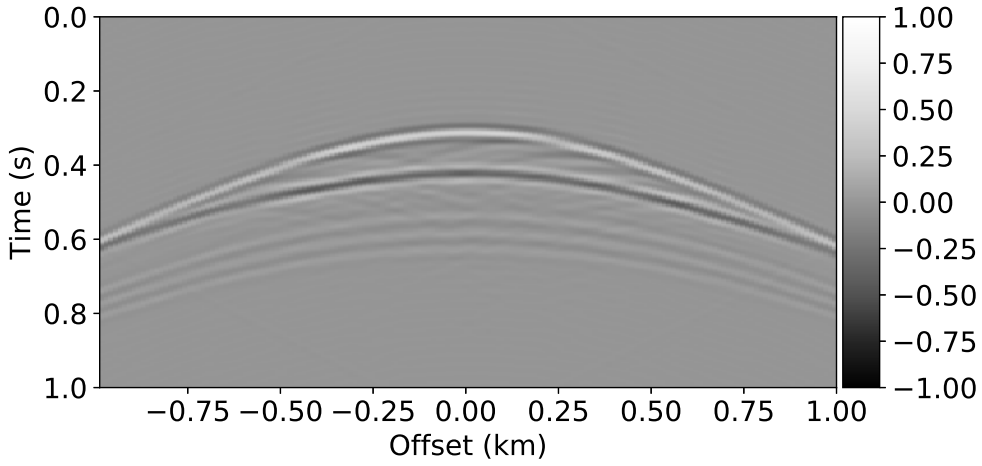
Figure 2.10 shows the exact Marmousi2 wavespeed model used for the second numerical test. We compare the one-way and FD modeling methods for one shot gather at position 0.5 km and the other at position 1.7 km. We chose these shot gathers because one is near a layered area without much geological complexity, while the other is located in a more complex region.

Figure 2.11 shows the shot gather at position 0.5 km. From the offset -0.5 km until 1.0 km the events are mainly related to specular reflections, this way, the events are similar in both results, Figures 2.11(a) and 2.11(b). After one second, stronger scattered energy occurs in the FD data, Figure 2.11(a). Offsets higher than 1.0 km are dominated by head waves and high incidence-angle reflections, that are practically absent in the one-way data, Figure 2.11(b).

Figure 2.13 shows the shot gather at position 1.7 km. The overall amplitudes are very different between the methods, Figures 2.13(a) and 2.13(b). It was more difficult for



(a)



(b)

Figure 2.8: FD and one-way modeling in the lens-shaped model. Common shot-gather in the middle of the model. (a) FD; (b) One-way.

the one-way method to reproduce the zero-offset FD data, Figure 2.14, than in the shot acquired at 0.5 km, Figure 2.12. However, the shape and positioning of many events in the one-way data are very similar to the FD result.

2.7 Summary

In this chapter, we have derived the coupled one-way wave equations from the equation of motion and the constitutive relation for fluids. They are coupled equations due to the presence of the secondary-source term, in each equation, involving the reflection of the wavefield that propagates in the opposite direction. Using the assumption of a laterally invariant medium, we verified that the square-root operator divided by the product between the angular frequency and the mass density is related to the inverse of the acoustic impedance, equation 2.31. The source separation into its down- and upgoing

components provided a monopole response, equation 2.60, which we defined consistently with the corresponding source in the two-way wave equation, see Appendix B, Section B.5.1.

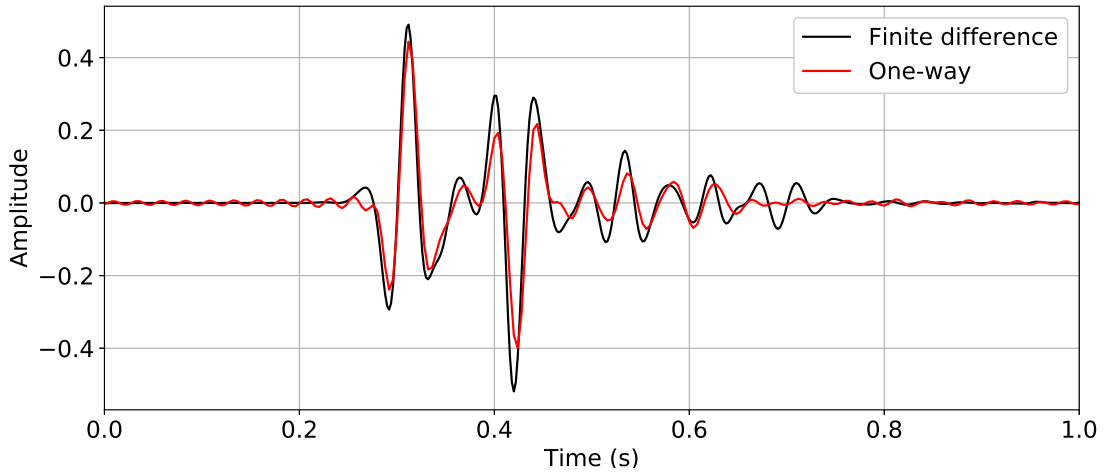
In Section 2.3, we defined the decoupled Green's functions that account for transmission effects besides propagation. Then, from the relation between the one-way and two-way wave operators developed in Appendix B, Section B.5, we applied Green's theorem directly to the one-way wavefields to obtain the associated integral representations for extrapolation from a boundary. During algebraic manipulation of the surface integral in the Green's theorem, we chose to use the one-way wave equations as boundary conditions.

In section 2.4, the integral representations in an unbounded space provided a framework to make explicit the contribution of each scattering order using a wavefield expansion of the total downgoing wavefield. The upgoing counterpart was obtained by applying a reflection to the downgoing series representation.

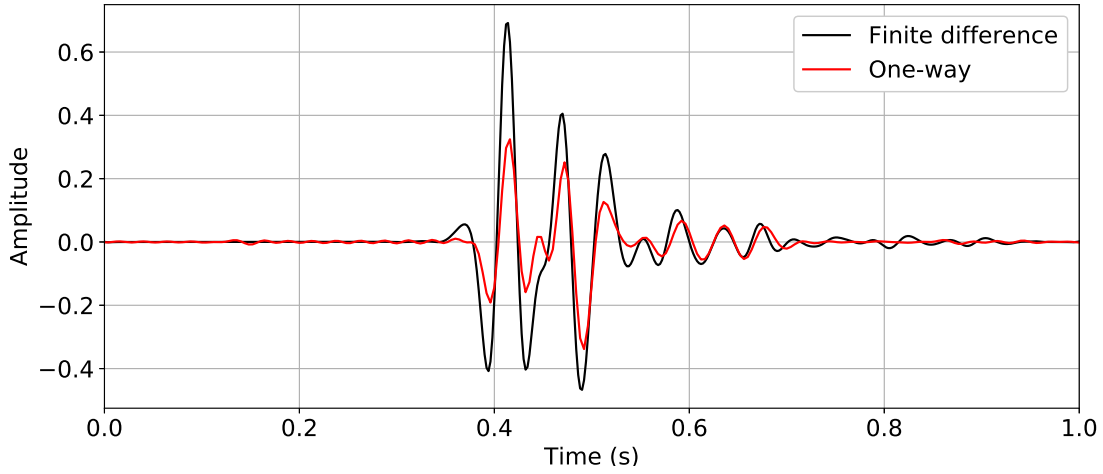
We also reviewed the complex Padé Fourier finite-difference approximation to the square-root operator in a two-dimensional model with constant mass-density, followed by the definition of the one-way extrapolation scheme, Section 2.5, see also the Appendix F. Finally, we discussed the one-way modeling algorithm and presented two applications comparing the one-way modeling with the more accurate FD method, Section 2.6. The main simplifications adopted in the implementation were: two-dimensional acoustic medium, only specular reflections are modeled, constant mass-density, angle-independent scattering and scattering only due to vertical variations. The pseudo-algorithm for downgoing extrapolation was sketched in Algorithm 1, with the upgoing procedure performed similarly.

Numerical tests provided insights into the behavior of the one-way modeling algorithm in two models with different complexity. The seismic response of the simple Lens model exhibited similar kinematic behavior to the FD data, but with difficulties to reproduce the AVA effect. In the Marmousi2 model, the shot gather acquired near the layered region showed that, for simple geology, the one-way data compare well with the FD data. But for a shot gather closer to a region with complicated geology, e.g. containing faults, the one-way data had difficulties to represent some events and the overall correct amplitudes. However, the one-way data still predicted a fair amount of events with the correct shape.

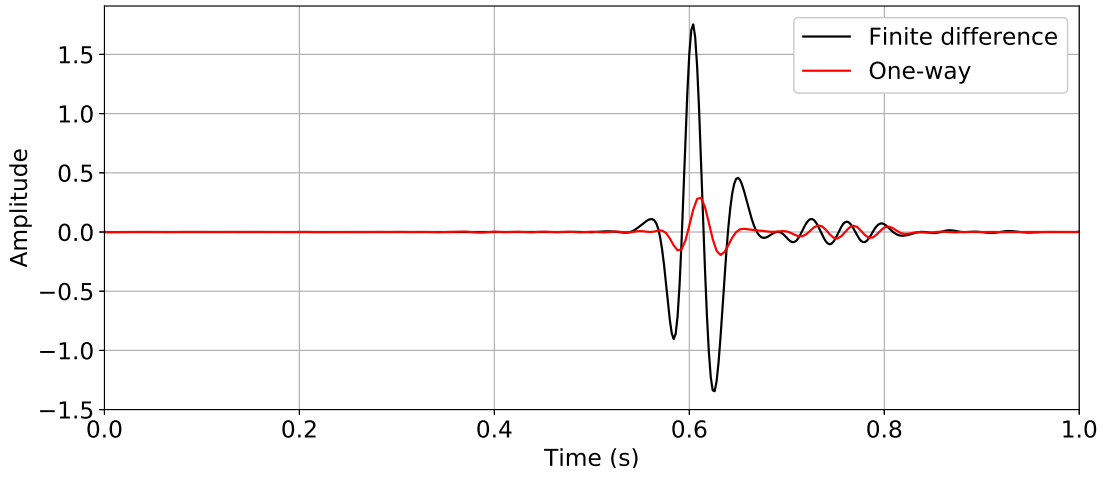
Moreover, even in this worst case scenario, near offset traces showed some similarity with FD data, at least similar kinematic effects. This indicates that in regions with complicated geology, an inversion based on phase information may be still effective. Another possibility is to improve the algorithm with the inclusion of the angle effect on scattering.



(a)



(b)



(c)

Figure 2.9: FD and one-way modeling in the lens-shaped model. Note the different amplitude scales. Offset: (a) zero ; (b) 512 m; (c) 987 m.

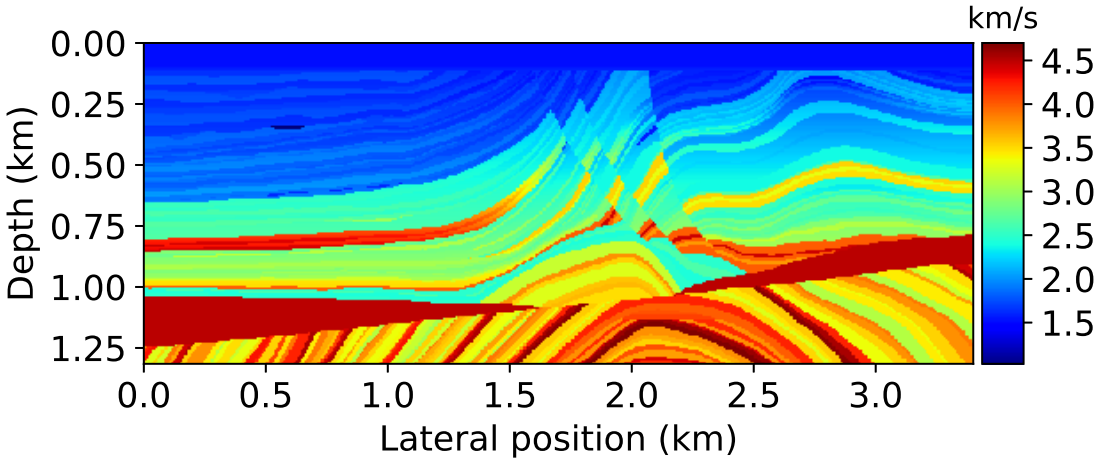
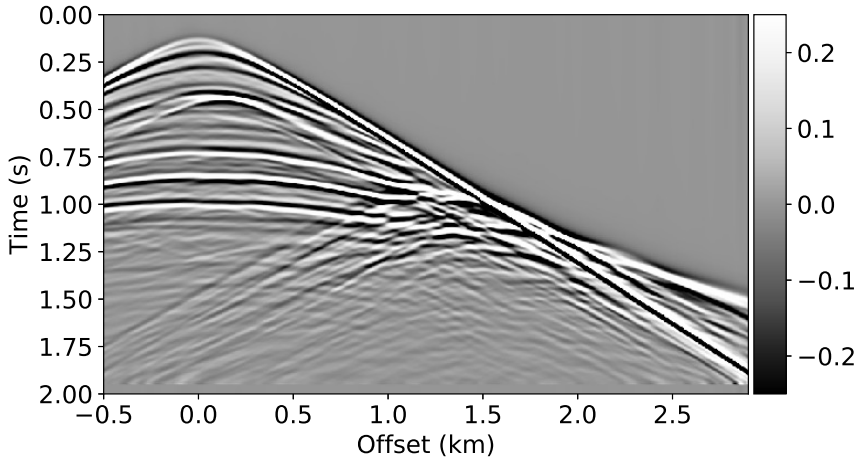
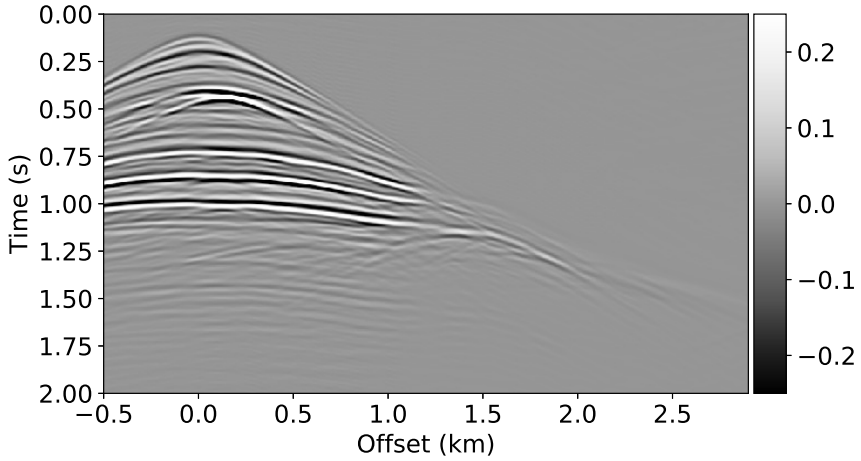


Figure 2.10: Marmousi2 exact wavespeed.

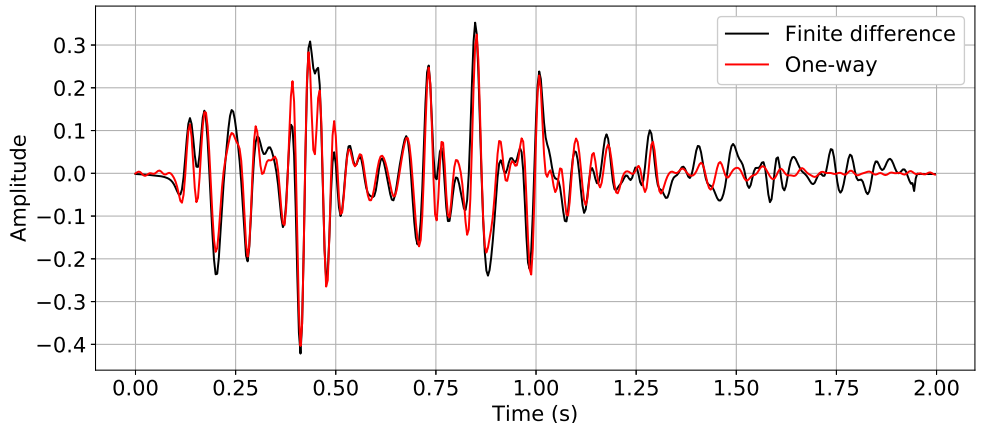


(a)



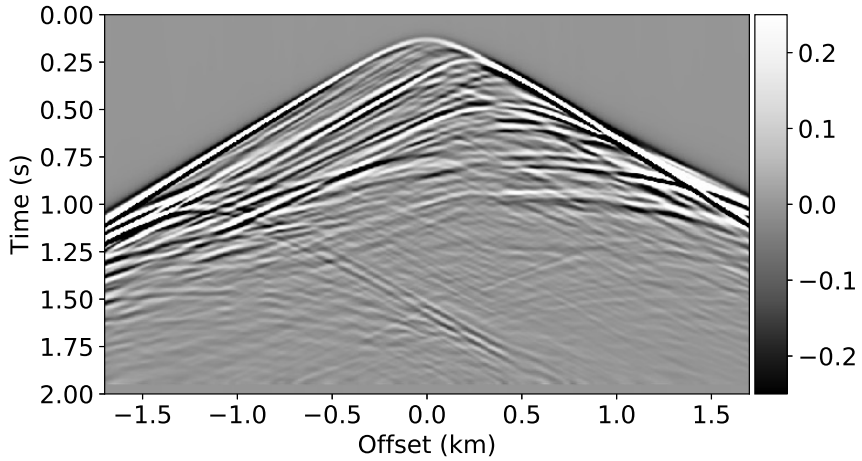
(b)

Figure 2.11: FD and one-way modeling in the Marmousi2 model. Common shot-gather at position 0.5 km. (a) FD; (b) One-way.

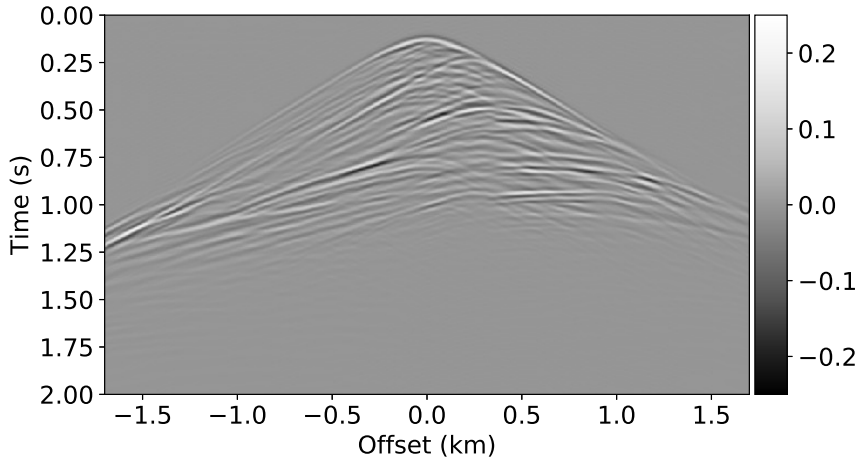


(a)

Figure 2.12: Zero-offset trace of the shot gather at position 0.5 km.



(a)



(b)

Figure 2.13: FD and one-way modeling in the Marmousi2 model. Common shot-gather at position 1.7 km. (a) FD; (b) One-way.

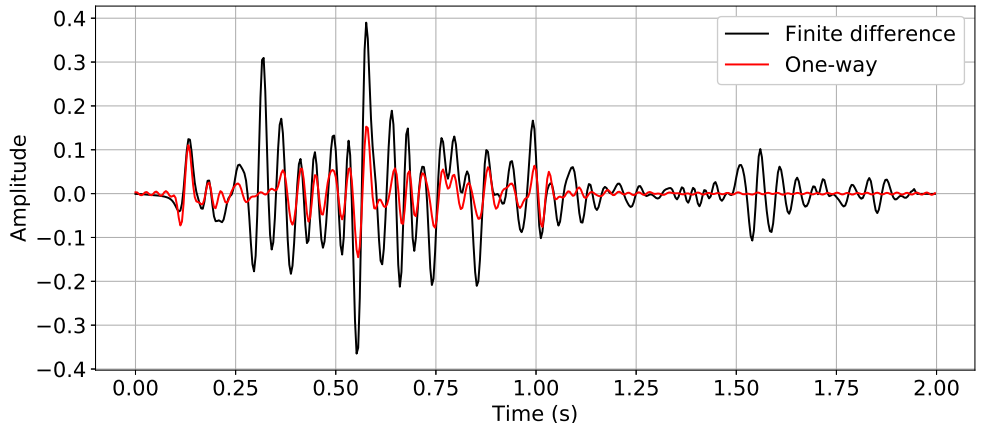


Figure 2.14: Zero-offset trace for the shot gather at position 1.7 km.

3 Least-squares migration: A parameterization study

We investigate two parameterizations of the least-squares migration problem in the acoustic case and for angle-independent scattering. Under the considered assumptions, the objective is to recover true-amplitude seismic images using full wavefields simulated with the coupled one-way wave equations. Here, the concept of full wavefields represents the ability of the modeling algorithm to produce multiple reflections.

From the continuous description of the coupled one-way wave equations, we obtain a general expression to the misfit-function gradient. This result suggests a modification of the wavefields used on the conventional imaging condition. Simplifying the general formulation, we obtain an expression similar to the conventional imaging condition, but still with an additional term that comes from the relation between the transmission and reflection coefficients.

The inverse problem is solved iteratively using the inexact Gauss–Newton method. During the development of the inversion methodology, we believe that it is reasonable to substitute the adjoint of the transmission coefficient by its inverse and, as a consequence, we obtain backpropagation equations that may balance amplitudes of deep reflectors relatively to the shallow ones.

We evaluate the imaging conditions on a simple model and test the inversion methodologies on two models with increasing complexity.

3.1 Modeling equations

In this section, we present the modeling equations. We start from the differential formulation to provide clarity. Then, we discuss the scattering operators and the adopted assumptions. Finally, we present the modeling equations in their integral form, that is how we implemented them.

3.1.1 Differential formulation

We investigate the two-dimensional acoustic migration problem using the coupled downgoing $P^+(\mathbf{x},\omega)$ and upgoing $P^-(\mathbf{x},\omega)$ wavefields. We use pressure-normalization, i.e., the sum $P^+ + P^-$ provides the total pressure wavefield, and the coupled one-way wave equations that describe this problem are

$$\frac{\partial P^+}{\partial z} = -i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{R}}_c^- P^- + \hat{\mathcal{T}}_c^+ P^+ + S^+ \,, \tag{3.1}$$

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \hat{\mathcal{R}}_c^+ P^+ - \hat{\mathcal{T}}_c^- P^- \,. \tag{3.2}$$

Here, $\mathbf{x} = (x, z)$, x represents the horizontal axis and z the depth axis pointing downward, ω is the angular frequency, the hat over the operators indicates integration over the lateral coordinate x, $\hat{\mathcal{R}}_c$ is the reflectivity operator, $\hat{\mathcal{T}}_c$ is the transmissivity operator, the superscript $+$ in the scattering operators denotes incidence from above on model position and the superscript $-$ indicates incidence from below, S^+ is a downgoing seismic source. See Chapter 2 for a detailed discussion of these modeling equations.

In a discontinuous model, the reflectivity and transmissivity operators are related to the first-order approximation of the corresponding reflection and transmission operators $\hat{\mathcal{R}}^{\pm}$ and $\hat{\mathcal{T}}^{\pm}$, see Appendix C. In the angle-independent case, the scattering operators are the acoustic reflection and transmission coefficients, with their action on the wavefields performed by simple multiplication.

The two-dimensional constant mass-density generalized vertical wavenumber, called for short square-root operator, $\hat{\mathcal{H}}_1$ is

$$\hat{\mathcal{H}}_1 = \left[\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c(x,z)^2}\right]^{1/2} . \tag{3.3}$$

It is related to the lateral Helmholtz operator $\hat{\mathcal{H}}_2$ via the twofold application

$$\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1 \hat{\mathcal{H}}_1 \,. \tag{3.4}$$

In a laterally invariant model and in the domain (k_x, z, ω), we have that $\hat{\mathcal{H}}_1$ is equivalent to the vertical wavenumber.

3.1.2 Scattering operators and coefficients

For an acoustic and constant mass-density continuous model, the reflectivities and transmissivities in equations 3.1 and 3.2 are

$$\hat{\mathcal{R}}_c^{\pm} = \hat{\mathcal{T}}_c^{\pm} = \mp\frac{1}{2}\hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\,, \tag{3.5}$$

where the reflectivities and transmissivities have dimension of distance inverse. The action of these operators on a wavefield is performed via integration over the lateral axis, see Appendix B, Section B.3.

The associated transmission and reflection operators for a discontinuous model are

$$\hat{\mathcal{R}}^{+} = \left[\hat{\mathcal{H}}_{1,l} + \hat{\mathcal{H}}_{1,u}\right]^{-1}\left[\hat{\mathcal{H}}_{1,u} - \hat{\mathcal{H}}_{1,l}\right]\,, \tag{3.6}$$

$$\hat{\mathcal{T}}^{+} = \mathcal{I} + \hat{\mathcal{R}}^{+}\,, \tag{3.7}$$

$$\hat{\mathcal{R}}^{-} = -\hat{\mathcal{R}}^{+}\,, \tag{3.8}$$

$$\hat{\mathcal{T}}^{-} = \mathcal{I} - \hat{\mathcal{R}}^{+}\,, \tag{3.9}$$

where $\mathcal{I}$ is the identity operator, $\hat{\mathcal{H}}_{1,l} = \hat{\mathcal{H}}_1(x, z+\Delta z)$ and $\hat{\mathcal{H}}_{1,u} = \hat{\mathcal{H}}_1(x, z-\Delta z)$.

According to the procedure and nomenclature outlined by Foster (1975). The linearization of these equations provides the reflectivities and transmissivities (for more details see Section C.2 of Appendix C). For example,

$$\hat{\mathcal{R}}^{+}(z+\Delta z/2, z-\Delta z/2) \approx \hat{\mathcal{R}}_c^{+}(z)\Delta z\,, \tag{3.10}$$

$$\hat{\mathcal{T}}^{+}(z+\Delta z/2, z-\Delta z/2) - \mathcal{I} \approx \hat{\mathcal{T}}_c^{+}(z)\Delta z\,. \tag{3.11}$$

Assuming angle-independent scattering, we approximate the square-root operator in the reflectivity, equation 3.5, as

$$\hat{\mathcal{H}}_1^{scat} \approx \frac{\omega}{c}\,, \tag{3.12}$$

where in connection with the vertical wavenumber, this approximation is equivalent in the wavenumber domain to impose $k_x = 0$, which makes this operator angle-independent, see also the discussion in Section 2.2.2. We included the superscript "*scat*" to indicate that this approximation is applied only to the scattering operators.

The substitution of the approximation in equation 3.12 in 3.5 provides

$$\hat{\mathcal{R}}_c^{+} \approx R_c^{+} = \frac{1}{2c}\frac{\partial c}{\partial z}\,, \tag{3.13}$$

where the angle-independent reflectivity R_c^{+} is not an integral operator. Moreover, its

action on a wavefield is performed by simple multiplication and it is independent of the angular frequency. In this expression, we recognize the first-order approximation to the angle-independent reflection coefficient (Berteussen and Ursin, 1983). In a discontinuous model it is given by

$$R^{+}(x,z) = \frac{c(x,z+\Delta z/2) - c(x,z-\Delta z/2)}{c(x,z+\Delta z/2) + c(x,z-\Delta z/2)}\,. \tag{3.14}$$

This result will be used to parameterize the forward modeling equations 3.1 and 3.2 during the least-squares migration. The other scattering coefficients can be obtained by using equations 3.7, 3.8 and 3.9 with the identity operator substituted by a unitary constant.

In the context of the modeling equations, we highlight that the approximation made for the reflectivity operator may look inconsistent. Because we will keep the first term of the modeling equations 3.1 and 3.2 intact in order to preserve wide-angle propagation effects. In practice, this combination may work well in scenarios of low contrasts in the velocity model. Furthermore, it is consistent with the angle-independent approximation commonly used in the joint migration inversion literature (see, e.g., Masaya and Verschuur, 2018) and adopted here.

3.1.3 Discrete integral representation

Now, we define the Green's functions associated with equations 3.1 and 3.2. Considering that the velocity model is vertically homogeneous between adjacent depth positions, we have

$$\frac{\partial G^{+}}{\partial z} = -i\hat{\mathcal{H}}_1^{extrap} G^{+} - \Delta z\,\delta(x-x')\delta(z-z_{n+1})\,, \tag{3.15}$$

$$\frac{\partial G^{-}}{\partial z} = i\hat{\mathcal{H}}_1^{extrap} G^{-} + \Delta z\,\delta(x-x')\delta(z-z_{n-1})\,, \tag{3.16}$$

where the superscript indicates that the approximation of $\hat{\mathcal{H}}_1$ made for scattering in equation 3.12 is not applied to these equations to preserve wide-angle extrapolation.

We use these Green's functions to build integral representations for extrapolation from a boundary, Section 2.3.3. Additionally, we apply a procedure similar to the Born expansion to detach scattering orders, Section 2.4, and obtain the integral recursive modeling equations

$$P_{j+1}^{+}(x',z_{n+1},\omega) = \hat{\mathcal{G}}^{+}(x',z_{n+1},\omega;x,z_n)\left(R^{-}P_j^{-} + T^{+}P_{j+1}^{+} + S^{+}\Delta z\right)(x,z_n,\omega)\,, \tag{3.17}$$

$$P_{j+1}^{-}(x',z_{n-1},\omega) = \hat{\mathcal{G}}^{-}(x',z_{n-1},\omega;x,z_n)\left(R^{+}P_{j+1}^{+} + T^{-}P_{j+1}^{-}\right)(x,z_n,\omega)\,, \tag{3.18}$$

where j is related to the scattering order, z_n is the nth boundary, $n = 0, 1, 2, 3, \ldots, N$ and $z_{n\pm1} = z_n \pm \Delta z$. In these equations, we already considered the angle-independent reflection and transmission coefficients. Moreover, the hat over the Green's functions indicates that the extrapolation is performed via integration over the x coordinate, as usual in two-dimensional boundary extrapolation.

Note that equations 3.17 and 3.18 feed one another in a recursive scheme. Starting with P_0^+, given a known downgoing source wavefield S^+.

For $j = -1$, we have primary reflections modeled as

$$P_0^+(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}; x, z_n)\left(T^+ P_0^+ + S^+ \Delta z\right)(x, z_n, \omega)\,, \tag{3.19}$$

$$P_0^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left(R^+ P_0^+ + T^- P_0^-\right)(x, z_n, \omega)\,. \tag{3.20}$$

Hence, equations 3.17 and 3.18 form the basis of the forward modeling algorithm used in the inversion methodology. The downgoing wavefield is calculated from the model top to bottom and the upgoing counterpart from the model bottom to top. In this recursive modeling scheme, it is necessary to keep only the downgoing wavefield P^- in memory between iterations. See also Section 2.4.3 for a detailed explanation of these modeling equations.

3.1.4 Implementation details

In this section, we describe the implementation details of the extrapolation scheme and the source term.

Green's function

We implemented the Green's function using the complex Padé Fourier finite-difference (CPFFD) approximation (Amazonas et al., 2007). Consistently with this methodology, the square-root operator was approximated by

$$\hat{\mathcal{H}}_1 \approx \underbrace{\frac{\omega}{c_r}\sqrt{1 + p^2 X^2}}_{phase-shift} + \underbrace{\frac{\omega}{c_r} C_0 (p - 1)}_{split-step} + \underbrace{\frac{\omega}{c_r} \sum_{n=1}^{N} \frac{p(1-p) A_n X^2}{1 + \sigma B_n X^2}}_{CPFD}\,, \tag{3.21}$$

where C_0, A_n and B_n are the complex Padé coefficients, $p = c_r/c$ with c_r a reference velocity and c the actual velocity model, $\sigma = 1 + p - p^3$, and $X^2 = (c/\omega)^2 \partial^2/\partial x^2$. See Appendix F for more details.

The three terms are: the phase-shift, which can be applied in the wavenumber domain; the split-step applied in the space domain; the finite-difference term that must

be applied in the space domain. These terms are named after the solution of the partial differential equation associated with the one-way propagation (Amazonas et al., 2007).

Source term

The source term was implemented as a monopole (Wapenaar, 1990) given by

$$S^{+}(x,z,\omega;x_s,z_s) = -\frac{i}{2\Delta z}\hat{\mathcal{H}}_1^{-1}\delta(x-x_s)\delta(z-z_s)C(\omega)\,, \tag{3.22}$$

where $C(\omega)$ describes the source wavelet spectrum, in our experiments it is set as the Ricker wavelet, more details in Appendix D. See Section 2.2.3 for a detailed discussion about the source term.

3.2 Inverse problem

We investigate the estimation of the angle-independent reflection coefficient vector $\boldsymbol{R}^{+}$ using a nonlinear least-squares methodology. Figure 3.1 sketches the inversion methodology. Given an actual estimate of the reflection coefficients $\boldsymbol{R}_k^{+}$, the down- and upgoing wavefields are calculated. Then, the inexact Gauss–Newton method is applied to estimate the update direction $\Delta\boldsymbol{R}_k^{+}$. Next, the vector of model parameters is updated using the update direction scaled by a step-length α_k. If the misfit function is smaller than 1.01 times the misfit value in the last iteration, the inversion procedure continues, otherwise, the algorithm is stopped.

We develop the discrete inverse problem in a grid of model parameters with M elements, e.g. $\boldsymbol{R}^{+}$ is a column vector with M elements. For the continuous equations related to the partial derivatives, see Appendix G. Although in practice we implement the discretized equations, the continuous equations are insightful.

Here, we omit the subscript related to the scattering order to simplify notation. In practice, as the model parameters estimation iterations progresses, the upgoing wavefield is kept in memory to account for possible internal multiple reflections in the recursive modeling equations 3.17 and 3.18. This way, the recursion involved in the modeling equations is performed only once at the beginning of each iteration of the inverse problem.

3.2.1 Gauss–Newton method

The Gauss–Newton method is an approximate approach to account for the effect of the inverse Hessian on the gradient. Some benefits of the method are its capability to focus the gradient by reducing band-limiting effects from data acquisition and, consequently, it may speed up the convergence of the inversion process (Pratt et al., 1998).

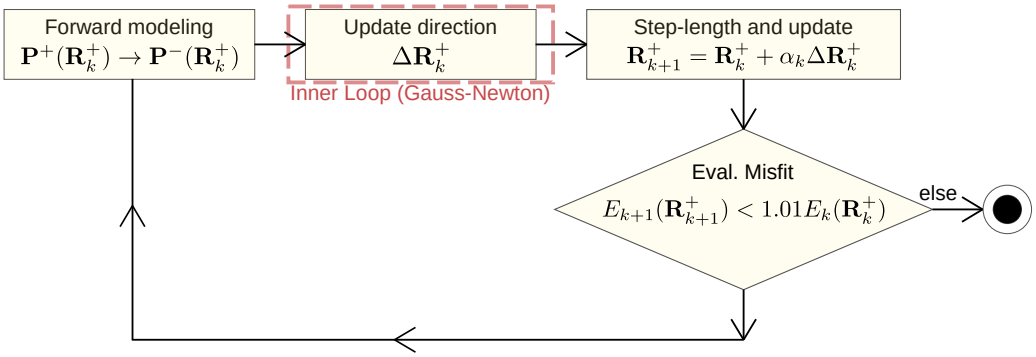


Figure 3.1: Seismic migration workflow. First the seismic data is modeled, then, an update direction is estimated with the Gauss–Newton method. Next, the model parameters are updated and the misfit function evaluated. If the misfit reduction criterion is satisfied the inversion procedure continues, otherwise, it is finished.

Adapting the notation from Métivier et al. (2017) to our application, we want to solve the least-squares problem

$$E(\boldsymbol{R}^+) = \frac{1}{2}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\sum_{m=1}^{M}\left|D_{slm}^- - \left(\mathcal{S}_s\boldsymbol{P}_{sl}^-(\boldsymbol{R}^+)\right)_m\right|^2 , \tag{3.23}$$

or in compact notation, we have

$$E(\boldsymbol{R}^+) = \frac{1}{2}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\left\|\boldsymbol{D}_{sl}^- - \mathcal{S}_s\boldsymbol{P}_{sl}^-(\boldsymbol{R}^+)\right\|_2^2 , \tag{3.24}$$

where s indicates one shot, N_s is the number of shots, l indicates one frequency, N_ω is the number of frequencies, m indicates one model grid point and M is the number of model grid points. Moreover, $\boldsymbol{D}_{sl}^-$ is a column vector with the upgoing observed data, $\boldsymbol{P}_{sl}^-$ is a column vector with the modeled upgoing wavefield at all M model spatial positions, $\mathcal{S}_s$ is an operator that samples the wavefield at the receivers position and may vary with shot position, $\|.\|_2^2$ denotes the L^2 norm squared over the receivers coordinate. The definition of the reflection coefficients vector is

$$\boldsymbol{R}^+ = \begin{bmatrix} R_1^+ \\ R_2^+ \\ \vdots \\ R_m^+ \\ \vdots \\ R_M^+ \end{bmatrix} , \tag{3.25}$$

in which the model parameters have N_x horizontal positions, N_z depth levels such that

the total number of element is $M = N_x \times N_z$ and $\boldsymbol{R}^+ \in \mathbb{R}^M$. The same disposition applies to $\boldsymbol{D}^-_{sl}$ and $\boldsymbol{P}^-_{sl}$, but they belongs to $\mathbb{C}^M$.

In order to estimate $\boldsymbol{R}^+$ that minimizes the misfit function $E(\boldsymbol{R}^+)$, the inexact Gauss–Newton method is applied to find update directions $\Delta\boldsymbol{R}^+_k$ by solving approximately the subproblem

$$\min_{\Delta\boldsymbol{R}^+_k} \frac{1}{2} \sum_{s=1}^{N_s} \sum_{l=1}^{N_\omega} \left\| \mathcal{S}_s \frac{\partial \boldsymbol{P}^-_{sl}}{\partial \boldsymbol{R}^+} \bigg|_{\boldsymbol{R}^+_k} \Delta\boldsymbol{R}^+_k - \left(\boldsymbol{D}^-_{sl} - \mathcal{S}_s \boldsymbol{P}^-_{sl}(\boldsymbol{R}^+_k)\right) \right\|_2^2 , \tag{3.26}$$

where k is the iteration number. In a model with M grid points, for one shot and one frequency, the wavefield Jacobian matrix is represented as

$$\frac{\partial \boldsymbol{P}^-}{\partial \boldsymbol{R}^+} = \begin{bmatrix} \dfrac{\partial P^-_1}{\partial R^+_1} & \dfrac{\partial P^-_1}{\partial R^+_2} & \cdots \dfrac{\partial P^-_1}{\partial R^+_M} \\ \dfrac{\partial P^-_2}{\partial R^+_1} & \dfrac{\partial P^-_2}{\partial R^+_2} & \cdots \dfrac{\partial P^-_2}{\partial R^+_M} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial P^-_M}{\partial R^+_1} & \dfrac{\partial P^-_M}{\partial R^+_2} & \cdots \dfrac{\partial P^-_M}{\partial R^+_M} \end{bmatrix} . \tag{3.27}$$

The update of the model parameters $\Delta\boldsymbol{R}^+_k$ is estimated approximately by solving equation 3.26 with the LSMR algorithm (Fong and Saunders, 2011). In this manner, the wavefield Jacobian matrix is not formed explicitly, but it is necessary only to implement its action on a vector, i.e., the so-called matrix-free approach. The solution satisfies approximately the normal equations

$$\left[\left(\mathcal{S}_s \frac{\partial \boldsymbol{P}^-_{sl}}{\partial \boldsymbol{R}^+} \right)^\dagger \mathcal{S}_s \frac{\partial \boldsymbol{P}^-_{sl}}{\partial \boldsymbol{R}^+} \right] \bigg|_{\boldsymbol{R}^+_k} \Delta\boldsymbol{R}^+_k = \left(\mathcal{S}_s \frac{\partial \boldsymbol{P}^-_{sl}}{\partial \boldsymbol{R}^+} \right)^\dagger \bigg|_{\boldsymbol{R}^+_k} \left(\boldsymbol{D}^-_{sl} - \mathcal{S}_s \boldsymbol{P}^-_{sl}(\boldsymbol{R}^+_k)\right) , \tag{3.28}$$

where $\dagger$ denotes transpose and complex conjugate, i.e., the adjoint, and the adjoint of sampling operator $\mathcal{S}^\dagger_s$ injects the data in the model grid at the receivers position. Moreover, we omitted the sum over shots and frequencies and the right-hand side is equivalent to the negative of the misfit-function gradient. After estimating the update direction, the model parameters are updated according to

$$\boldsymbol{R}^+_{k+1} = \boldsymbol{R}^+_k + \alpha_k \Delta\boldsymbol{R}^+_k , \tag{3.29}$$

where the step-length α_k is calculated using a parabolic fit (Bohlen et al., 2009).

Note that the upgoing-wavefield Jacobian-matrix action on the vector of model parameters $\Delta\boldsymbol{R}^+_k$ can be interpreted as the first-order perturbed wavefield (see, e.g., Macedo,

2014; Camargo, 2019) given by

$$\Delta \boldsymbol{P}^{-}(\boldsymbol{R}_k^{+}) = \boldsymbol{P}^{-}(\boldsymbol{R}_k^{+} + \Delta \boldsymbol{R}_k^{+}) - \boldsymbol{P}^{-}(\boldsymbol{R}_k^{+}) \approx \left.\frac{\partial \boldsymbol{P}^{-}}{\partial \boldsymbol{R}^{+}}\right|_{\boldsymbol{R}_k^{+}} \Delta \boldsymbol{R}_k^{+} \,. \tag{3.30}$$

One element of the perturbed wavefield vector is given by

$$\Delta P_m^{-}(\boldsymbol{R}_k^{+}) = P_m^{-}(\boldsymbol{R}_k^{+} + \Delta \boldsymbol{R}_k^{+}) - P_m^{-}(\boldsymbol{R}_k^{+}) \approx \left.\frac{\partial P_m^{-}}{\partial \boldsymbol{R}^{+}}\right|_{\boldsymbol{R}_k^{+}} \Delta \boldsymbol{R}_k^{+} \,. \tag{3.31}$$

where $P_m^{-} = P^{-}(x, z, \omega, \boldsymbol{R}_k^{+})$ is the upgoing wavefield at one model grid point and it is calculated with equation 3.18. The scalar P_m^{-} depends on the vector of reflection coefficients because, if primary reflections are being modeled, one observation point (x, z) can be influenced by all the transmission effects in shallower positions and by the reflections that occurred in deeper positions until the model bottom. If internal multiples occur, the same observation point may also be influenced by reflections that occurred on shallower positions. Analogous notation and argumentation hold for the downgoing wavefield. Figure 3.2 sketches this discussion.

We will investigate two methodologies for the wavefield partial derivative with respect to $\boldsymbol{R}^{+}$ required by equation 3.26.

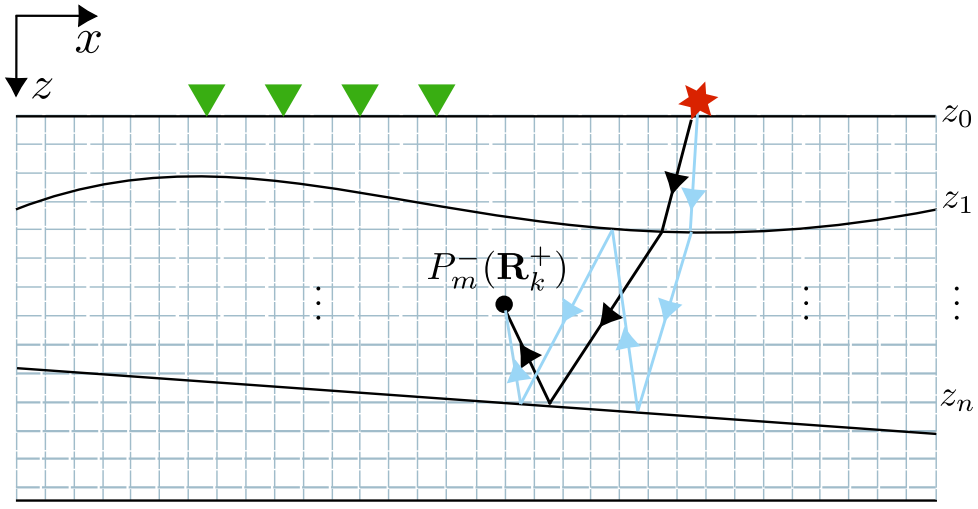


Figure 3.2: Representation of an upgoing wavefield P_m^{-} at one model grid point. It is composed of a wavefield that is transmitted and reflected (black arrows) and by another wavefield that transmitted and was reflected three times (blue arrows).

3.2.2 Adjoint wavefields

Before introducing the partial derivatives of the upgoing wavefield, we present the adjoint wavefields.

As demonstrated in the Appendix G, Section G.3, the continuous description of the adjoint wavefields is

$$\frac{\partial \Lambda_{sl}^{-}}{\partial z} = -i\hat{\mathcal{H}}_1 \Lambda_{sl}^{-} + \left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger} \Lambda_{sl}^{+} + \left(\hat{\mathcal{T}}_c^{-}\right)^{\dagger} \Lambda_{sl}^{-} - \mathcal{S}_s^{\dagger}\left(\boldsymbol{D}_{sl}^{-} - \mathcal{S}_s \boldsymbol{P}_{sl}^{-}\right) , \tag{3.32}$$

$$\frac{\partial \Lambda_{sl}^{+}}{\partial z} = i\hat{\mathcal{H}}_1 \Lambda_{sl}^{+} - \left(\hat{\mathcal{T}}_c^{+}\right)^{\dagger} \Lambda_{sl}^{+} - \left(\hat{\mathcal{R}}_c^{-}\right)^{\dagger} \Lambda_{sl}^{-} , \tag{3.33}$$

where Λ_{sl}^{+} is the adjoint of the downgoing wavefield and Λ_{sl}^{-} is the adjoint of the upgoing wavefield. Moreover, we adapted the sampling procedure from the delta distribution in the continuous case to the operator $\mathcal{S}_s$.

The right-most term in equation 3.32 is a boundary condition, indicating that at the receivers position the adjoint wavefield Λ_{sl}^{-} is the difference between the measured and the calculated data. The transposed and complex-conjugated operator $\mathcal{S}_s^{\dagger}$ simply positions the data in the model grid with the spatial dimensions of the simulation.

In view of the linearized inverse problem, it is desirable that the adjoint formulation behaves as close as possible of an inverse operation, if it was exactly the case, the inversion could be performed in one step. In order to achieve this goal approximately, in the discrete integral-formulation correspondent to equations 3.32 and 3.33, we substitute the adjoint of the transmissivity operator by its inverse. Therefore, analogously to the forward modeling equations 3.19 and 3.20, the zero-order adjoint-modeling equations are

$$\Lambda_{sl}^{-}(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^{+}(x', z_{n+1}, \omega; x, z_n)\left[\frac{1}{T^{-}}\Lambda_{sl}^{-} - \Delta z \mathcal{S}_s^{\dagger}\left(\boldsymbol{D}_{sl}^{-} - \mathcal{S}_s \boldsymbol{P}_{sl}^{-}\right)\right](x, z_n, \omega) , \tag{3.34}$$

$$\Lambda_{sl}^{+}(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^{-}(x', z_{n-1}, \omega; x, z_n)\left[\frac{1}{T^{+}}\Lambda_{sl}^{+} + R^{-}\Lambda_{sl}^{-}\right](x, z_n, \omega) , \tag{3.35}$$

where $\Lambda_{sl}^{\pm}$ are the adjoint wavefields at one grid point, $T^{\pm}$ are the angle-independent transmission coefficients and R^{-} is the angle-independent reflection coefficient. Note that the adjoint wavefields propagate in the opposite direction of the associated forward wavefields, that is why we call their propagation as backpropagation.

Expressions 3.34 and 3.35 for the computation of the adjoint wavefields are also physically reasonable. Besides the residuals backpropagation, the transmission losses are being compensated. We believe that this simple modification has the potential to balance the amplitude of deeper reflectors relatively to the shallow ones. Furthermore, these equations are implemented similarly to the forward modeling equations, but first the adjoint wavefield Λ_{sl}^{-} is calculated from the model top to bottom, then, the wavefield Λ_{sl}^{+} is calculated from the model bottom to top. Moreover, the adjoint upgoing wavefield is a secondary source for the adjoint downgoing wavefield.

3.2.3 Derivatives of the complete parameterization

Misfit-function derivative

In Appendix G, Section G.3.1, we applied the Lagrangian multipliers to the least-squares misfit function, equation 3.24, constrained by the continuous form of the forward modeling equations 3.1 and 3.2. Furthermore, using the relationship between the scattering operators, we fully parameterized the forward modeling equations in terms of the reflectivity $\hat{\mathcal{R}}_c^+$. Then, we obtained an expression for the misfit function derivative with respect to the reflectivity operator $\hat{\mathcal{R}}_c^+$.

We use the calculus chain-rule and the relation between reflectivity and reflection coefficient, equation 3.10, to obtain the expression for the derivative with respect to the angle-independent reflection coefficient $\boldsymbol{R}^+$

$$-\left(\frac{\partial E}{\partial \hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{c}}^{\,+}}\frac{\partial \hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{c}}^{\,+}}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+}\right)_m = \frac{1}{\Delta z}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\boldsymbol{Re}\left\{\left(P_{slm}^+ - P_{slm}^-\right)^*\left(\Lambda_{slm}^+ - \Lambda_{slm}^-\right)\right\}, \tag{3.36}$$

where m indicates one grid point in a model grid with M elements, we multiplied this expression by -1 to obtain a descent direction, $\boldsymbol{Re}\{.\}$ is the real-part operator and the asterisk denotes complex conjugate. Moreover, $\boldsymbol{P}_{sl}^+$ is the downgoing wavefield, $\boldsymbol{P}_{sl}^-$ is the upgoing wavefield, $\boldsymbol{\Lambda}_{sl}^+$ the adjoint downgoing wavefield and $\boldsymbol{\Lambda}_{sl}^-$ the adjoint upgoing wavefield. Note that in the angle-independent case considered here, we have the zero-lag cross-correlation between the involved wavefields (for the general case, see Appendix G, Section G.3).

Equation 3.36 also tells us how to calculate the action of the adjoint of the partial derivative of the upgoing wavefield on some perturbed wavefield. Taking the negative derivative of the misfit function in equation 3.24 and substituting in equation 3.36, we obtain the identity

$$\left(\frac{\partial P_{slm}^-}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+}\right)^* \mathcal{S}_s^\dagger\left(\boldsymbol{D}_{sl}^- - \mathcal{S}_s\boldsymbol{P}_{sl}^-\right) = \frac{1}{\Delta z}\,\boldsymbol{Re}\left\{\left(P_{slm}^+ - P_{slm}^-\right)^*\left(\Lambda_{slm}^+ - \Lambda_{slm}^-\right)\right\}. \tag{3.37}$$

In the general case of an upgoing residual wavefield ΔP^- other than the data residual in the least-squares misfit function. Equation 3.37 can be calculated by substituting the data residual in the adjoint upgoing wavefield, equation 3.34, by the quantity ΔP^-.

Wavefields partial derivative

Consider the first-order perturbation of the down- and upgoing modeling equations 3.17 and 3.18 fully parameterized in terms of R^+, for one shot and one frequency. This way,

one element of the perturbed-wavefield vector obtained from the action of the Jacobian matrix on a vector $\Delta \boldsymbol{R}_k^+$ is defined as

$$\left(\frac{\partial P_m^+}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+} \Delta \boldsymbol{R}_k^+\right)(x', z_{n+1}, \omega) \approx \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)$$
$$\left[\left(P^+ - P^-\right)\Delta R_k^+ + T_k^+ \Delta P^+ + R_k^- \Delta P^-\right](x, z_n, \omega)\,, \tag{3.38}$$
$$\left(\frac{\partial P_m^-}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+} \Delta \boldsymbol{R}_k^+\right)(x', z_{n-1}, \omega) \approx \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)$$
$$\left[\left(P^+ - P^-\right)\Delta R_k^+ + T_k^- \Delta P^- + R_k^+ \Delta P^+\right](x, z_n, \omega)\,, \tag{3.39}$$

where in the right-hand side the model parameters multiply the wavefields at the corresponding position. Observe that these equations are very similar to the forward modeling equations 3.17 and 3.18, except for the perturbed model parameter and the perturbed wavefields. In fact, this similarity is reasonable, since the modeling equations 3.17 and 3.18 form a recursive scheme, in which the first iteration is the linear approximation (see Section 2.4 for a detailed discussion).

It also noteworthy that, although, equations 3.38 and 3.39 are the result of a linearization, the last term in the right-hand side is related to multiple reflections if these equations are applied recursively. Therefore, in the one-way framework, it is possible to apply the Gauss–Newton method using linearized equations that accounts for multiple reflections, or more generally, nonlinear scattering effects. If these higher-order effects are required, then, the adjoint modeling equations 3.34 and 3.35 must also be applied recursively. These observations goes against the common sense, at least for the two-way wave equations, that the Gauss–Newton method is related only to primaries reflections (see, e.g., Pratt et al., 1998; Métivier et al., 2017). In this work, the zero-order terms are sufficient for the considered numerical examples.

Hence, the right-hand side of equation 3.28 is calculated using expression 3.36. And the upgoing-wavefield-derivative action on a vector of model parameters, required in the left-hand side, is given by the first-order perturbed wavefields, equations 3.38 and 3.39. The adjoint of the upgoing-wavefield-derivative action on some wavefield ΔP^- is calculated with equation 3.36, upon substitution of the data residual by the perturbed wavefield ΔP^-.

3.2.4 Derivatives of the quasi-conventional parameterization

Misfit-function derivative

The negligence of variations in the downgoing wavefield in the Lagrangian formulation of Appendix G, Section G.3.1, naturally, discards the adjoint wavefield Λ^+. In the same manner, the misfit partial-derivative in equation 3.36, is adapted to

$$-\left(\frac{\partial E}{\partial \hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{c}}^{\,+}}\frac{\partial \hat{\boldsymbol{\mathcal{R}}}_{\boldsymbol{c}}^{\,+}}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+}\right)_m = -\frac{1}{\Delta z}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\boldsymbol{Re}\left\{\left(P_{slm}^+ - P_{slm}^-\right)^*\Lambda_{slm}^-\right\}. \tag{3.40}$$

This expression is almost the conventional imaging condition. It is not exactly the conventional result due to the presence of the upgoing wavefield. This additional term arises from the parameterization of the transmissivity as a function of the reflectivity in the upgoing modeling equation.

Analogously to the development of the adjoint equation 3.37, the action of the adjoint partial-derivative of the upgoing wavefield on the data residual is given by

$$\left(\frac{\partial P_{slm}^-}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+}\right)^* \mathcal{S}_s^\dagger\left(\boldsymbol{D}_{sl}^- - \mathcal{S}_s\boldsymbol{P}_{sl}^-\right) = -\frac{1}{\Delta z}\,\boldsymbol{Re}\left\{\left(P_{slm}^+ - P_{slm}^-\right)^*\Lambda_{slm}^-\right\}. \tag{3.41}$$

Upgoing-wavefield derivative

Consistently with the misfit-function derivative in equation 3.40. We consider perturbations only in the upgoing wavefield and obtain the following expression for the correspondent action of the partial derivative of the upgoing wavefield on a vector

$$\left(\frac{\partial P_m^-}{\partial \boldsymbol{R}^+}\bigg|_{\boldsymbol{R}_k^+}\Delta\boldsymbol{R}_k^+\right)(x', z_{n-1}, \omega) \approx \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)$$
$$\left[\left(P^+ - P^-\right)\Delta R_k^+ + T_k^-\Delta P^-\right](x, z_n, \omega), \tag{3.42}$$

where we omitted the shot and frequency indexes due to the reintroduction of the independent variables. Moreover, this expression describes one experiment, i.e., one shot gather at one angular frequency.

Hence, equation 3.40 is a second option to the right-hand side of the Gauss–Newton subproblem, equation 3.28. And equation 3.42 is the required upgoing wavefield derivative action on the vector $\Delta\boldsymbol{R}_k^+$.

3.2.5 First nonlinear iteration

Initially, in the numerical experiments, the image will be set zero and, as a consequence, the forward upgoing and the adjoint downgoing wavefields will be also zero. Thus, in the first iteration, independently of the chosen parameterization, the negative of the misfit-function gradient will be

$$-\left(\left.\frac{\partial E}{\partial \boldsymbol{R}^{+}}\right|_{\boldsymbol{R}_0^{+}}\right)_m = -\frac{1}{\Delta z}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\boldsymbol{Re}\left\{\left(P_{slm}^{+}\right)^{*}\Lambda_{slm}^{-}\right\}. \tag{3.43}$$

Accordingly, the wavefields partial derivatives simplifies to

$$\left(\left.\frac{\partial P_m^{+}}{\partial \boldsymbol{R}^{+}}\right|_{\boldsymbol{R}_0^{+}}\Delta\boldsymbol{R}_0^{+}\right)(x', z_{n+1}, \omega) \approx \hat{\mathcal{G}}^{+}(x', z_{n+1}, \omega; x, z_n)\left(P^{+}\Delta R_0^{+} + \Delta P^{+}\right)(x, z_n, \omega), \tag{3.44}$$

$$\left(\left.\frac{\partial P_m^{-}}{\partial \boldsymbol{R}^{+}}\right|_{\boldsymbol{R}_0^{+}}\Delta\boldsymbol{R}_0^{+}\right)(x', z_{n-1}, \omega) \approx \hat{\mathcal{G}}^{-}(x', z_{n-1}, \omega; x, z_n)\left(P^{+}\Delta R_0^{+} + \Delta P^{-}\right)(x, z_n, \omega). \tag{3.45}$$

3.2.6 Relationship with conventional migration

In this section, we discuss the relationship between the least-squares solution to the migration problem and the conventional migration approach. The concept of imaging condition is fundamental in the development of migration algorithms. Put in simple terms, first a migration algorithm performs forward and adjoint modeling, then the obtained wavefields are used as input for an imaging condition. Finally, a result different from zero is a necessary condition for an image to occur.

Conventional seismic migration can be understood as the negative of the gradient of the regular least-squares misfit function multiplied by a normalization factor (see, e.g., Plessix and Mulder, 2004). This is a consequence of, upon negligence of multiple reflections, the relationship between the upgoing wavefield P^{-} and the reflection coefficients R^{+} being practically linear in equation 3.18, except for the additive transmitted wavefield. This relation translates into an approximate linear problem for estimation of the coefficients R^{+}.

That is why, starting from a image set as zero, the direct computation of the misfit function gradient already provides a reasonable result. The subsequent iterations account, for example, for proper amplitude scaling and source wavelet deconvolution (see, e.g., Nemeth et al., 1999).

From this line of thought, in the complete parameterization of the migration problem

in equation 3.36, we recognize the new imaging condition

$$\boxed{\left(IC_{new}\right)_m = \frac{1}{\Delta z}\frac{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\boldsymbol{Re}\left\{\left(P^+_{slm}-P^-_{slm}\right)^*\left(\Lambda^+_{slm}-\Lambda^-_{slm}\right)\right\}}{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\left(P^+_{slm}-P^-_{slm}\right)^*\left(P^+_{slm}-P^-_{slm}\right)}\,,} \tag{3.46}$$

where the normalization factor is a rough approximation to the diagonal of the Gauss–Newton Hessian. In Chapter 4, we will derive this normalization in the context of the impedance estimation. In equation 3.46, the terms

$$\left(P^+_{slm}\right)^*\Lambda^+_{slm} \qquad \text{and} \qquad \left(P^-_{slm}\right)^*\Lambda^-_{slm}\,, \tag{3.47}$$

are originated from writing the transmissivities in the forward downgoing and upgoing equations as a function of the reflectivity, thus, they can be related to forward scattering. The terms

$$\left(P^+_{slm}\right)^*\Lambda^-_{slm} \qquad \text{and} \qquad \left(P^-_{slm}\right)^*\Lambda^+_{slm}\,, \tag{3.48}$$

come directly from the reflectivity, thus, they are related to backscattering (for more details see Appendix G, Section G.3).

Commonly, the imaging condition is implemented without the forward upgoing wavefield and the adjoint downgoing wavefield. The first simplification is equivalent to neglecting the upgoing transmission term. The second is a consequence of neglecting variations in the forward downgoing wavefield such that the associated adjoint wavefield is zero. These simplifications are represented by

$$\Lambda^+_{slm} = 0 \qquad \text{and} \qquad \left(P^-_{slm}\right)^*\Lambda^-_{slm} = 0\,. \tag{3.49}$$

From these assumptions, the conventional imaging condition (see, e.g., Claerbout, 1971) is recovered and it is given by

$$\boxed{\left(IC_{conv.}\right)_m = -\frac{1}{\Delta z}\frac{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\boldsymbol{Re}\left\{\left(P^+_{slm}\right)^*\Lambda^-_{slm}\right\}}{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\left(P^+_{slm}\right)^*\left(P^+_{slm}\right)}\,,} \tag{3.50}$$

in which the negative sign balances the negative sign of the source term of the adjoint wavefield in equation 3.32.

We also observe that the additional terms on the new imaging condition in equation 3.46 do not impact the first iteration of regular migration algorithms based on the one-way wave equations. The reason is that the procedure starts at a zero image, i.e., at $R^+_0 = 0$. Thus, initially, there is no forward upgoing wavefield and the adjoint downgoing wavefield is also zero. Under these circumstances, the new and the conventional imaging conditions

are equivalent.

However, if the migration problem is solved by some iterative scheme that requires more than one nonlinear iteration to estimate the image, the additional terms in the new imaging condition can be taken into account. Another possibility occurs in the context of migration using the two-way wave equation and a velocity model that produces reflections. In this case, it may be also possible to apply the new imaging condition in equation 3.46.

We illustrate this discussion on imaging conditions with Figure 3.3. We observe that the forward scattering terms $(P^-)^* \Lambda^-$ and $(P^+)^* \Lambda^+$ are different from zero at the reflector on position z_1 and may extend the imaged portion in addition to the term related to reflection $(P^+)^* \Lambda^-$. But these terms also provide information along the path of wave propagation that lies outside the reflectors. Note that the term $(P^-)^* \Lambda^+$ was left out of the discussion, because it is related to the reflections from below the reflectors.

We highlight that the forward scattering terms are also related to the velocity estimation problem (see Appendix G, Section G.4). Moreover, in migration methodologies that makes use of the two-way wavefields, the forward scattering terms are well known for the low-frequency noise in the final image. Therefore, the effectiveness of expression 3.46 should be verified with great care.

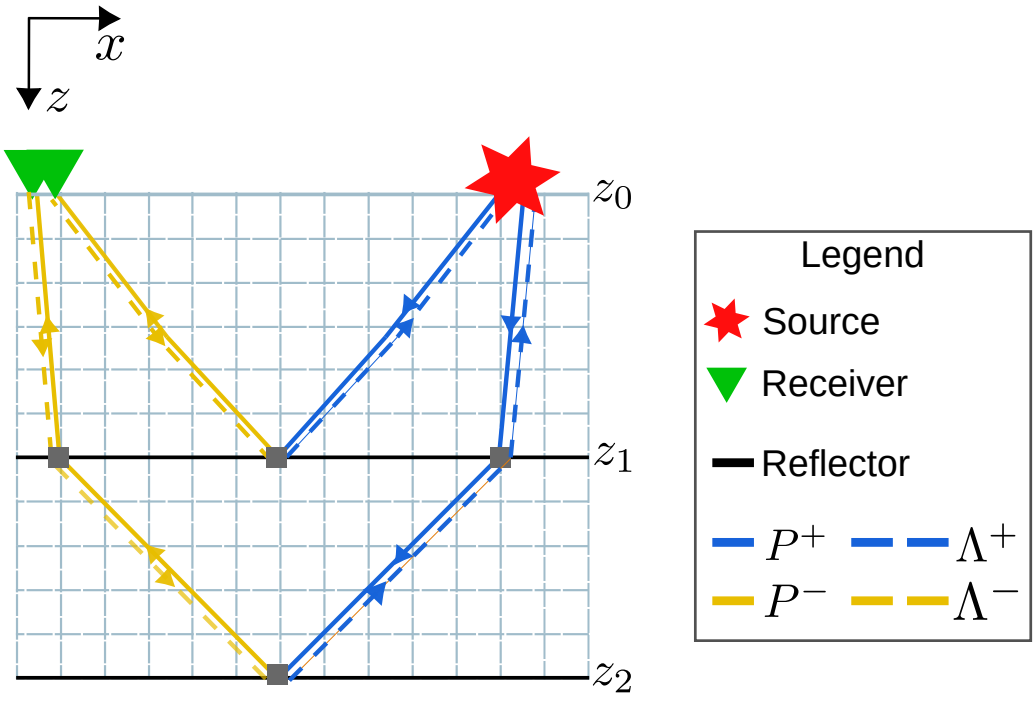


Figure 3.3: Representation of the seismic experiment for one seismic source and two receivers. Additionally, we sketch the forward ($P^{\pm}$) and adjoint ($\Lambda^{\pm}$) wavefields. Only primary reflections were considered and the reflectors mark positions in which occurs contrasts in the velocity.

3.3 Numerical tests

In this section, we test the proposed migration methodologies. First, we use a simple model with three layers to test the imaging conditions. Next, we present the results of the least-squares migration in two models with increasing complexity. In the discussions, we use the terms image and reflection coefficients interchangeably.

3.3.1 Evaluation of the first iteration

To simplify the analysis, we use a model with three layers, Figure 3.4. The reflection coefficient at the top of the middle layer is 0.05 and at the bottom it is -0.05. We acquire 20 shot gathers in the middle of the model with spacing of 10 m. The source wavelet is a Ricker with peak frequency of 20 Hz.

In order to produce the upgoing wavefield in the modeled data, we set the initial image as the exact reflection coefficients divided by 2. Thus, the exact image update should have magnitudes of ± 0.025 at the corresponding reflectors.

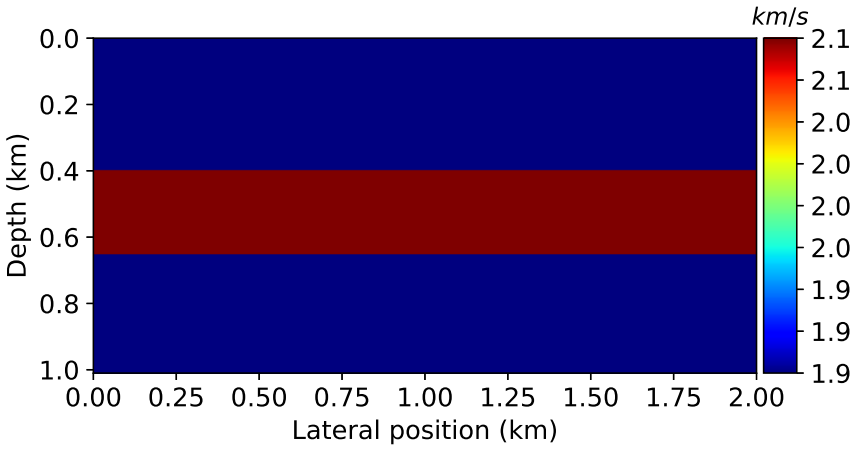
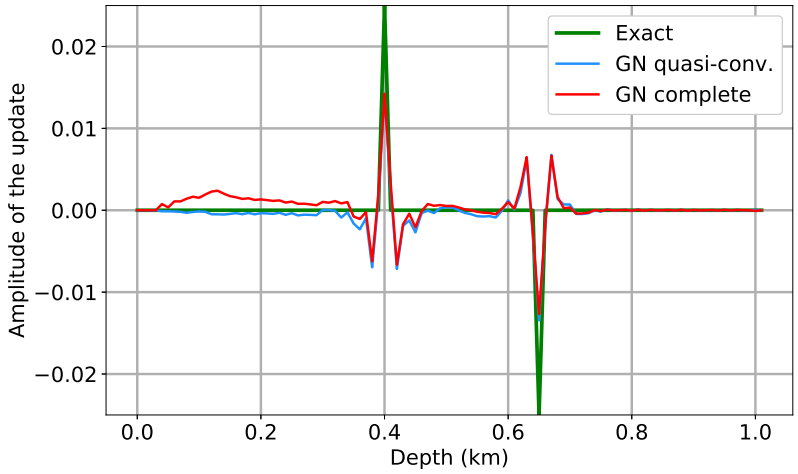


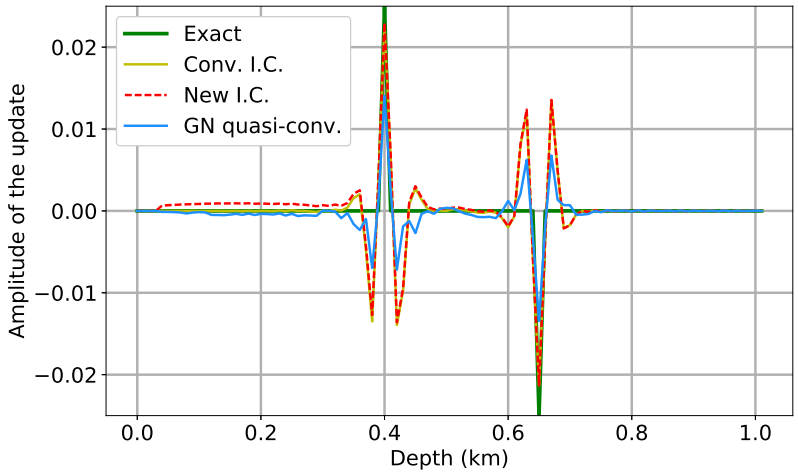
Figure 3.4: Velocity model with three layers.

We evaluate the new and the conventional imaging conditions defined in equations 3.46 and 3.50. That can be seen as the first iteration of the steepest-descent solution of the least-squares problem with step-length set to unity and without updating the initial image. For comparison purposes, we also include the first nonlinear iteration of the least-squares solution via Gauss–Newton considering the complete and the quasi-conventional parameterizations.

All the approaches produced results with smaller magnitudes than the exact update, Figure 3.5. The Gauss–Newton solutions were nearly the same, except that the complete parameterization exhibited more noise from depth position 0.1 to 0.3 km, Figure 3.5(a).



(a)



(b)

Figure 3.5: Results of the migration at the middle of the model with three layers. (a) Gauss–Newton results; (b) New and conventional imaging conditions.

The results of the new and conventional imaging conditions were also good, Figure 3.5(b). In comparison to the Gauss–Newton solutions, the imaging conditions produced a result with magnitude closer to the exact update at the position of 0.4 km, Figure 3.5(b). On the other hand, the Gauss–Newton solutions had a better performance on the wavelet removal, but at the cost of 4 inner iterations. Additionally, if a line-search for the step-length was applied, the Gauss–Newton updates would obtain step-lengths nearly unitary, while the results from the imaging conditions would obtain step-lengths smaller than the unity due to the poor wavelet removal.

For the simple model considered here, the new imaging condition in equation 3.3 did not provide any considerable improvement on the result of the shallowest reflector, Figure 3.6. Maybe the model is too simple, or the extra information from the transmission terms $(P^-)^*\Lambda^-$ and $(P^+)^*\Lambda^+$ are already in the reflection term $(P^+)^*\Lambda^-$. However, the noise before the first reflector at 0.4 km is similar in the result of the complete parameter-

ization of the Gauss–Newton and in the new imaging condition, Figure 3.5. It seems to be a result of the additional terms $(P^{\pm})^{*}\,\Lambda^{\pm}$ as sketched in Figure 3.3. The high amplitude errors at the edges of the top reflector, Figure 3.6, occurred due to the reduced number of shot gathers and their concentration in the middle of the model.

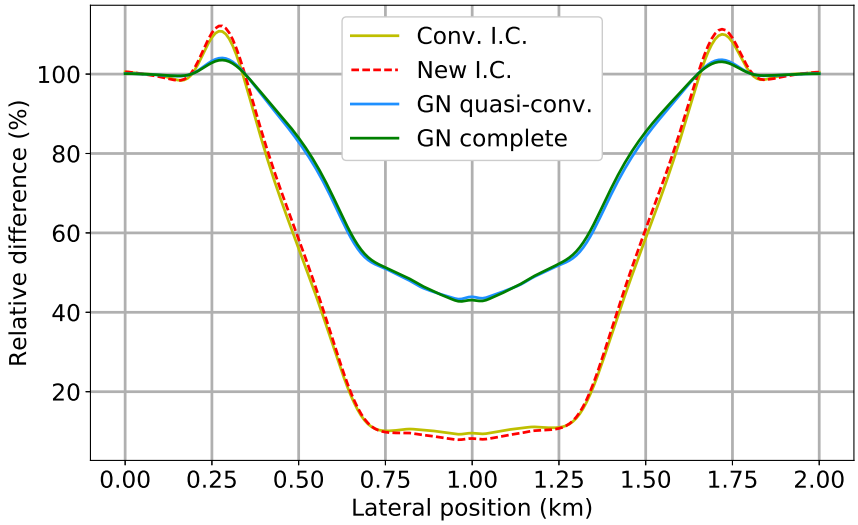


Figure 3.6: Relative differences between the exact and the estimated updates along the top reflector (position 0.4 km in Figure 3.4).

3.3.2 Least-squares solution

We test the parameterizations presented in Section 3.2.3 with the Lens model, (see, e.g., Masaya and Verschuur, 2018) and the modified Marmousi2 (Pan et al., 2018). All the frequency content and the exact velocity model are considered. The least-squares migration procedure is iterative, it is composed of an outer loop related to the nonlinear least-squares method and the inner loop related to the update direction estimated with the inexact Gauss–Newton method. Therefore, in the analysis of the results, we simply refer to the nonlinear iteration/outer and inner loop. The stopping criterion is a misfit function reduction smaller than 1% or an increase bigger than 1.01% relatively to the preceding iteration. The initial image is set zero for both tests and the maximum number of outer iterations is set to 10. We include the results of the conventional imaging condition given by equation 3.50 as a reference.

Lens model

Figure 3.7 shows the exact velocity model, it is composed of a background at 2.0 km/s, the lens at 2.5 km/s and fine-layering at the bottom. Table 3.1 shows the velocity model dimensions and the data acquisition parameters.

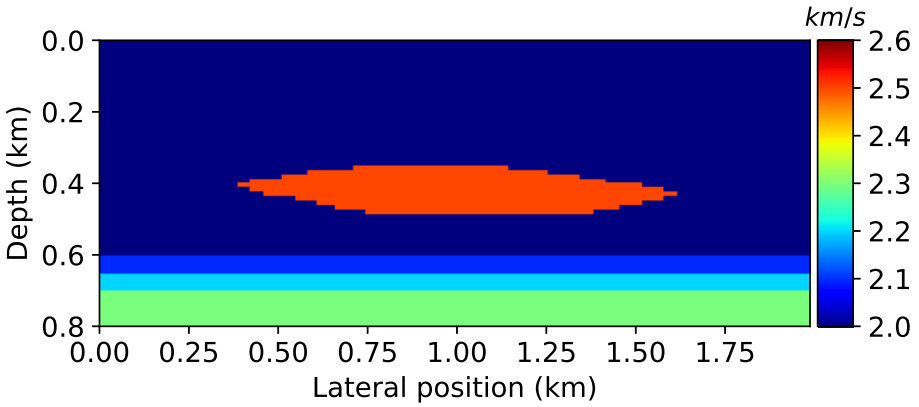


Figure 3.7: Lens model exact velocity.

Imaging results

The image profiles indicate that both approaches based on the least-squares solution provided similar results, Figure 3.8. The conventional imaging condition provided a result very close to the least-squares solutions, but still there are considerable side lobes related to a poor deconvolution of the source wavelet.

Table 3.1: Lens model dimensions and data acquisition parameters.

nx	160	Sources depth (m)	37.5
nz	64	Source spacing (m)	50
nt	306	First source (m)	0
dx (m)	12.5	Number of sources	40
dz (m)	12.5	Receivers depth (m)	37.5
dt (ms)	4	Number of receivers	160
Min. freq. (Hz)	0	Wavelet	Ricker ($f_p = 20$ Hz)
Max. freq. (Hz)	40	-	-

The least-squares solutions are very similar as already discussed for the profiles in Figure 3.8. Thus, we exhibit only one of the images together with the conventional imaging condition, Figure 3.9. Naturally, due to the good performance on the source wavelet removal, the least-squares solution, Figure 3.9(b), exhibits higher resolution than the result of the conventional imaging condition, Figure 3.9(a).

We observe that some artifacts similar to reverberations occur around the lens structure, Figure 3.9(b). Looking again into the profiles, Figure 3.8, we speculate that probably this is a sharp boundary effect related to the source wavelet deconvolution. It may be mitigated by regularizing the misfit function, e.g., adding the L^2 norm of the reflection coefficients scaled by a constant. Another possibility is to let the inner loop iterate more to obtain a more precise update direction.

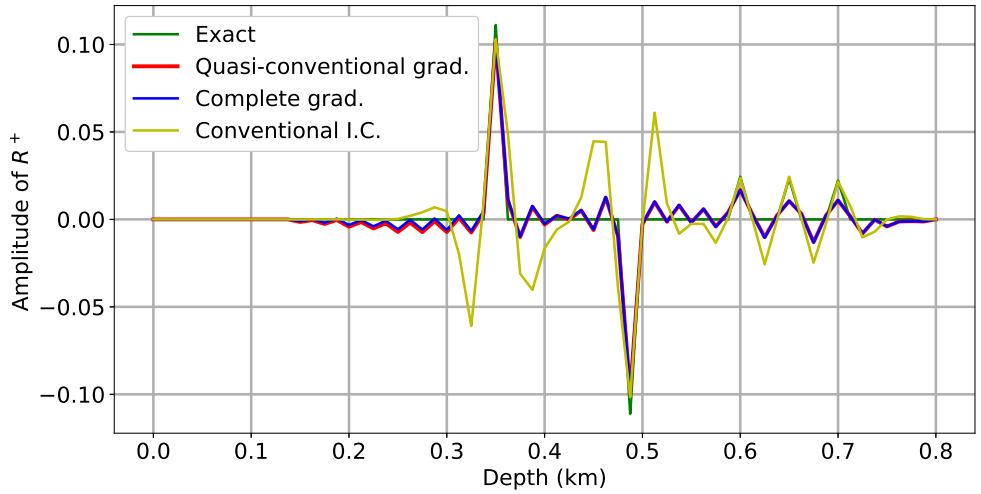


Figure 3.8: Profile in the middle of the migrated image of the Lens model.

Number of iterations, residuals and step-lengths

The complete gradient parameterization performed 10 outer iterations and on the average 2.8 inner iterations, Table 3.2. The quasi-conventional parameterization performed 10 outer iteration and on the average 2.7 inner iterations. Apparently, the complete and quasi-conventional parameterizations require approximately the same gradient computational power. But, the forward and adjoint modeling equations of the complete parameterization perform two times the number of operations of the quasi-conventional case due to the downgoing wavefield perturbation being taken into account. Therefore, the complete parameterization had higher computational cost on the Lens model.

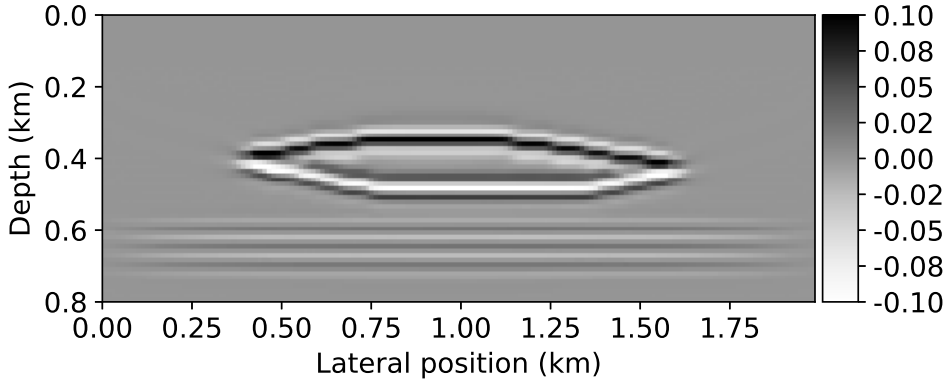
Starting at the end of the first iteration, Figure 3.13(a) shows the misfit function reduction. The complete parameterization was more effective in the reduction of the data residuals over the iterations.

It is also beneficial to verify the change of the model residuals along the iterations. We measured the change of the model residuals with the following expression

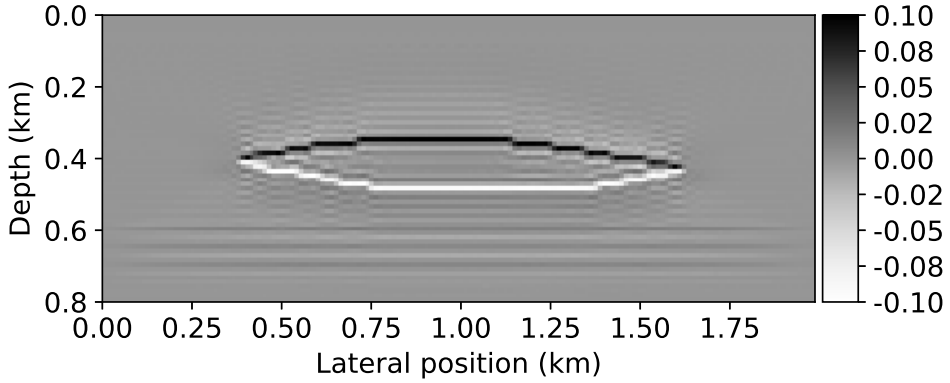
$$\text{Model residuals change} = 100 \times \frac{\|\boldsymbol{R}^+_{est} - \boldsymbol{R}^+_{ex}\|_2}{\|\boldsymbol{R}_{ex}\|_2}\,, \tag{3.51}$$

where $\boldsymbol{R}^+_{est}$ is the estimated image and $\boldsymbol{R}^+_{ex}$ is the exact image. Again, as we observed for the data residuals, the complete parameterization provided a higher reduction of the model residuals over the outer iterations.

The first step-length was approximately unitary for both approaches, Figure 3.13(c). We recall that we applied the inexact Gauss–Newton method to obtain the update directions. This way, we did not expect unitary step-lengths for all iterations.



(a)



(b)

Figure 3.9: Lens model: (a) Conventional imaging condition; (b) Least-squares solution.

Table 3.2: Results of the Lens model. Number of iterations in the outer (nonlinear least-squares) and inner loops (Gauss–Newton method) for different gradient parameterizations.

	Nonlinear least-squares	Gauss–Newton	Total
Complete	10	2.8	28
Quasi-conventional	10	2.7	27

Marmousi2

Figure 3.10 shows the exact velocity model of the Marmousi2. The velocity model dimensions and the acquisition configuration are described on Table 3.3.

Imaging results

The image profile follows very well the general aspects of the exact reflection coefficients, Figure 3.11(a). But the frequency range used in the acquisition was not enough to resolve adjacent reflection coefficients and the magnitudes are far from being matched with the exact response. Using a source wavelet with a wider frequency content may improve these details. The result of the conventional imaging condition in various positions has amplitudes smaller than the least-squares solutions and also exhibits less details,

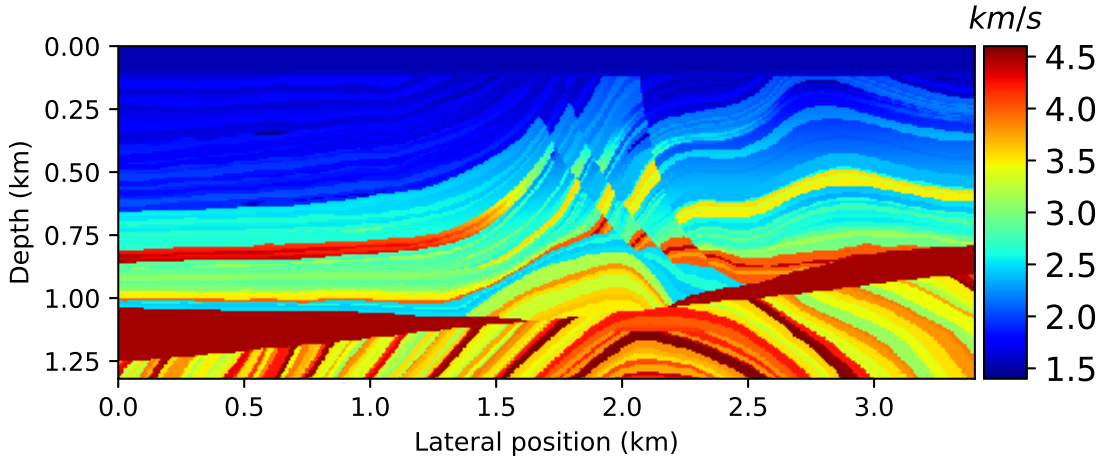


Figure 3.10: Exact velocity model of the Marmousi2.

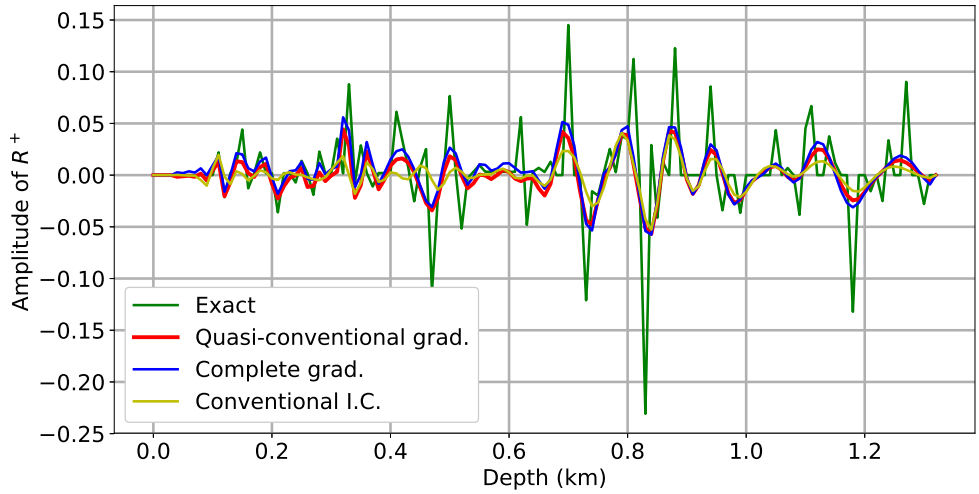
Figure 3.11(a). The first nonlinear iteration compares well with the final results of both parameterizations, Figure 3.11(b). Moreover, we observe small differences between the final results for both parameterizations of the Gauss–Newton method.

Table 3.3: Marmousi2 dimensions and data acquisition parameters.

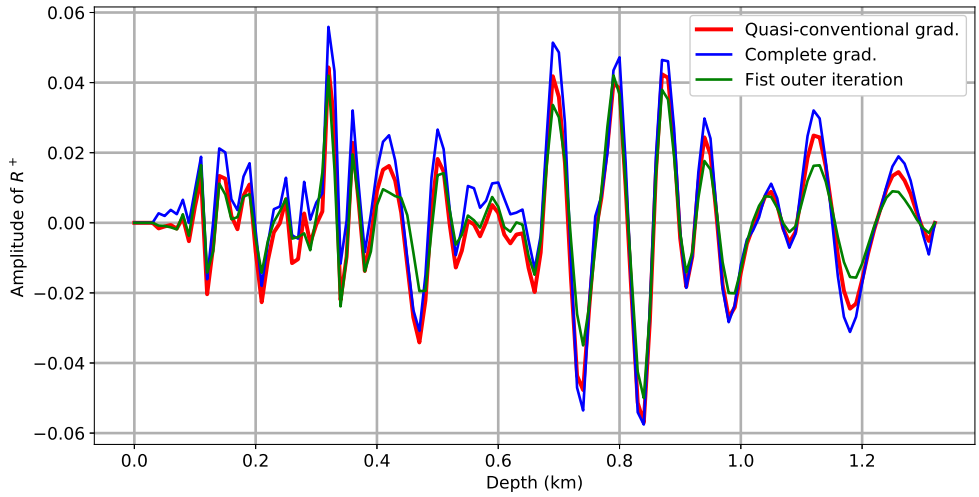
nx	341	Sources depth (m)	10.0
nz	132	Source spacing (m)	70
nt	256	First source (m)	0
dx (m)	10.0	Number of sources	48
dz (m)	10.0	Receivers depth (m)	10.0
dt (ms)	8	Number of receivers	341
Min. freq. (Hz)	0	Wavelet	Ricker ($f_p = 20$ Hz)
Max. freq. (Hz)	40	-	-

Figure 3.12 shows the images estimated with the conventional imaging condition and the least-squares method with the complete gradient. As the profiles already indicated, Figure 3.11(a), the least-squares solution provided an image with much more details than the conventional imaging condition, Figure 3.12. The first nonlinear iteration result exhibits greater resolution than the conventional imaging condition and has a quality near the final least-squares solution, Figure 3.12. We observe that the final least-squares solution has slightly more details and the amplitudes are better balanced than the result of the first nonlinear iteration.

We also note that after the depth level of 1.0 km, the dipping layers are not very well resolved on the least-squares solution, Figure 3.12(c). Probably, this is a consequence of the limited acquisition range that stopped at the lateral edges of the model. Besides that, maybe these dips are outside the application range of the one-way propagator scheme. In order to verify the limitation of the one-way extrapolation adopted here, the same



(a)



(b)

Figure 3.11: Image profiles in the middle of Marmousi2. (a) Exact image and estimated results; (b) Gauss–Newton results and first nonlinear iteration.

experiment should be performed with the two-way wave equation.

Number of iterations, residuals and step-lengths

The complete gradient performed 10 outer iterations and on the average 2.7 inner iterations. The quasi-conventional parameterization performed 2 outer iterations and on the average 4 inner iterations, Table 3.4. The quasi-conventional parameterization stopped after the second outer iteration due to an increase of the misfit function higher than the threshold of 1.01% after the third outer iteration.

Both approaches to the Gauss–Newton method are equivalent in the first iteration and this is verified by the data same residual, Figure 3.14(a). Although the quasi-conventional parameterization performed only two outer iterations, it is possible to observe that the complete parameterization converged faster, Figure 3.14(a). Similar behavior is

observed in the model residuals, Figure 3.14(b).

Despite the improved convergence of the complete parameterization. We highlight that the quasi-conventional parameterization provided similar results as indicated in the profiles of Figure 3.11(a). Moreover, the first nonlinear iteration already provides considerable improvements in comparison to the conventional imaging condition, Figure 3.12(b).

Table 3.4: Marmousi2 results. Number of iterations in the outer (nonlinear least-squares) and inner loops (Gauss–Newton method) for different gradient parameterizations.

	Nonlinear least-squares	Gauss–Newton	Total
Complete	10	2.7	27
Quasi-conventional	2	4	8

3.3.3 Summary

We have reviewed the migration part of JMI given the exact velocity model. The inversion methodology described is related to the conventional least-squares migration and we applied the Gauss–Newton method to obtain the update directions. The main objective was to estimate the angle-independent reflection coefficients.

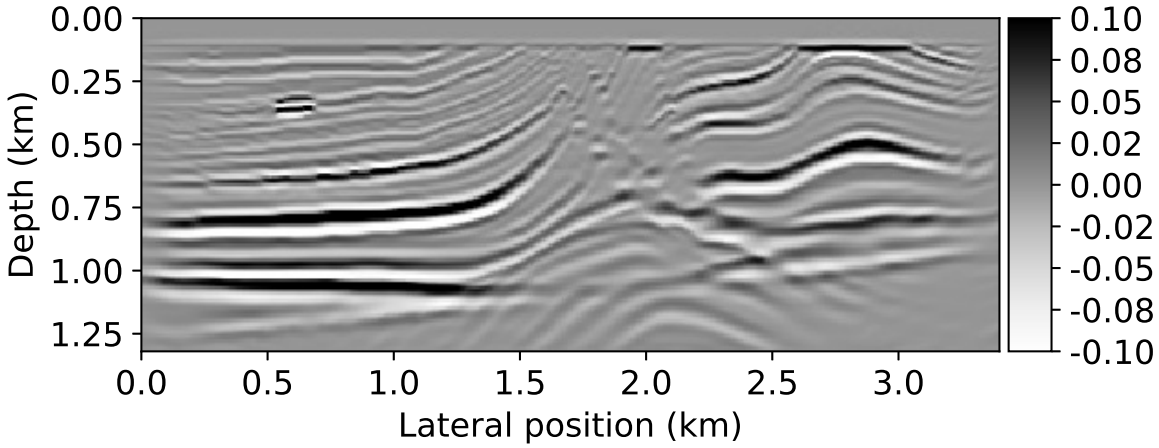
In the discussion of the forward and adjoint equations required by the Gauss–Newton method, we recognized that in the one-way framework it is possible to account for nonlinear effects by recursively applying the linearized equations. In the numerical tests, it was sufficient to consider only zero-order terms.

The expression derived for the gradient of the misfit function considering the forward and adjoint equations completely parameterized in terms of the reflectivity/reflection coefficient provided a new imaging condition. We analyzed this new imaging condition and realized that the terms related to forward scattering have the potential to complement the image formed by the terms related to backscattering. However, for the simple model considered, we could not observe any improvement over the conventional imaging condition. Additional investigations are required before disregarding this new expression. Maybe it can be useful in the reverse time migration methodology.

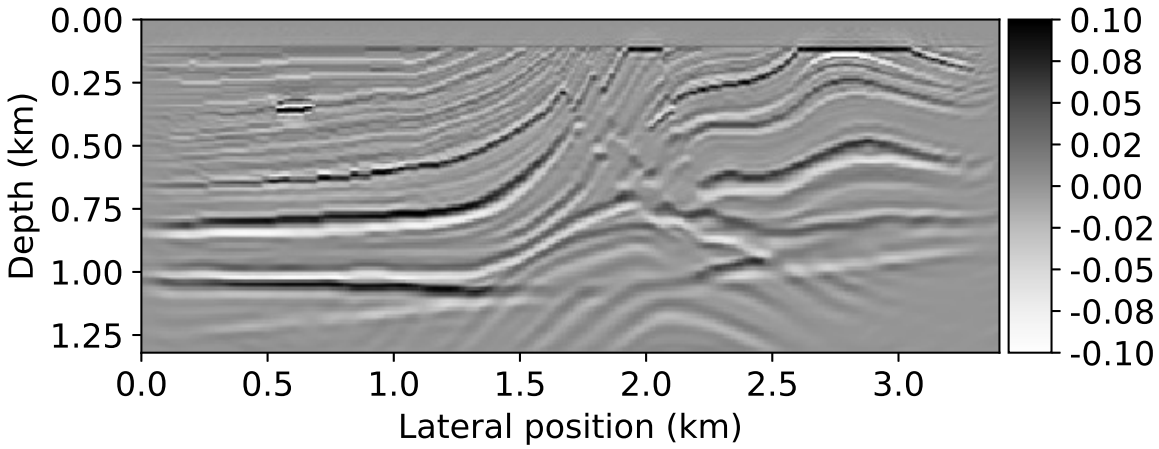
The numerical results indicated that the complete parameterization of the misfit function gradient, that take into account down- and upgoing wavefields variations, can provide better convergence than the quasi-conventional parameterization, that neglects variations in the downgoing wavefield. However, the final image was very similar for both parameterizations. Moreover, the first nonlinear iteration already provided an improved result in comparison to the image obtained solely by the conventional imaging condition.

This fact may assist in the compensation of the assumption of angle-independent scattering in the forward modeling equations. Because in the first nonlinear/outer itera-

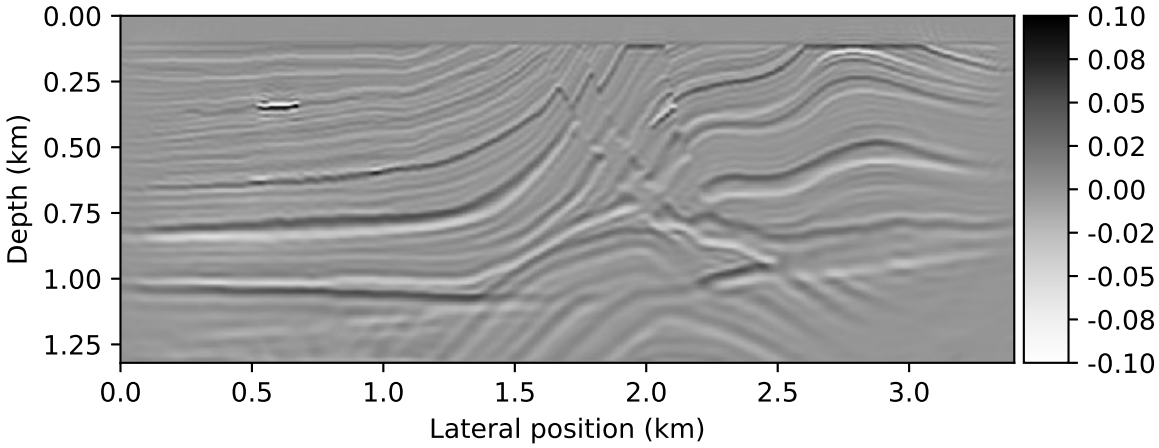
tion, the data residuals are simply the observed data. In this manner, inaccuracies due to the assumptions made in the modeling equations are less harmful to the estimated solution. Probably, in seismic data with considerable imprint of intrabed multiple reflections, more nonlinear iterations can be of greater importance.



(a)

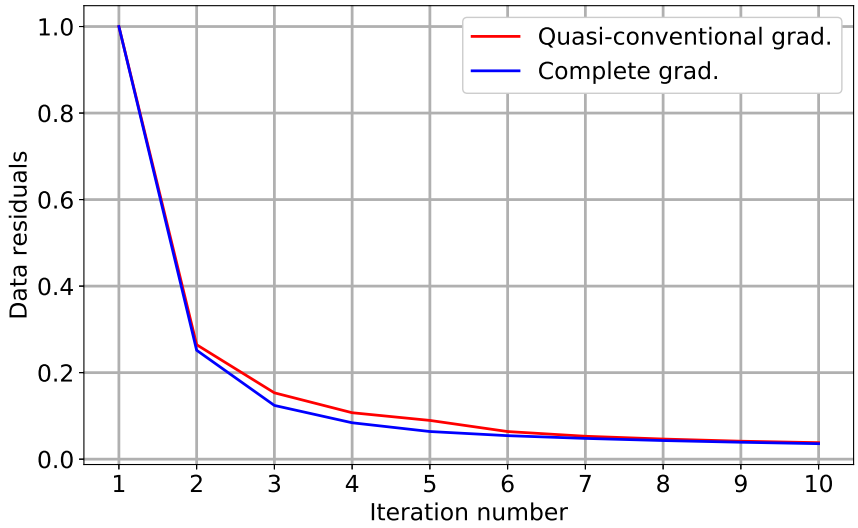


(b)

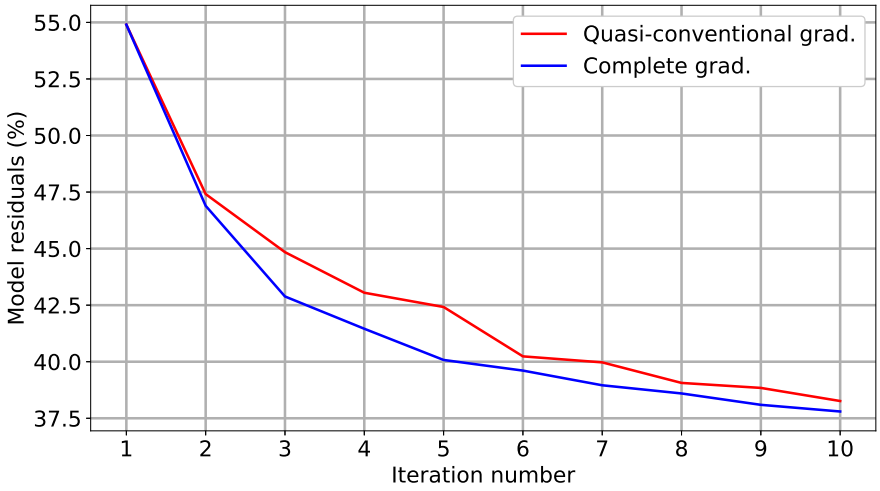


(c)

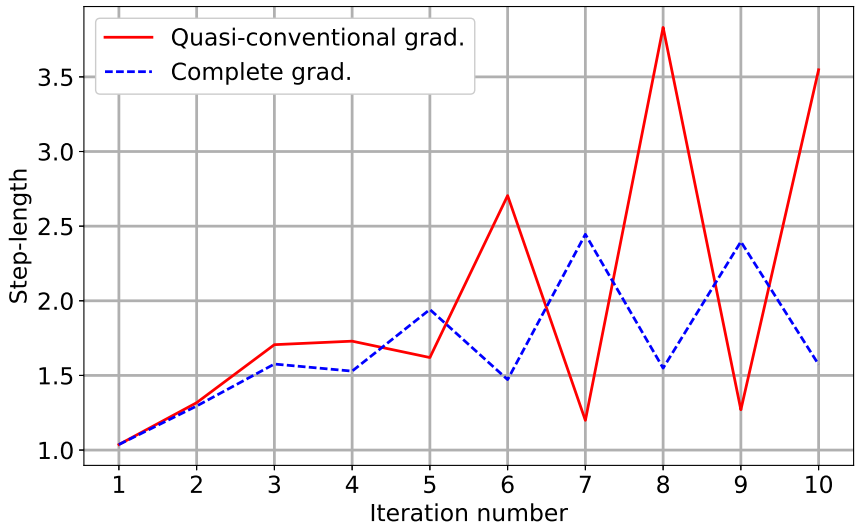
Figure 3.12: Marmousi2 results: (a) Conventional imaging condition; (b) First nonlinear iteration; (c) Least-squares solution.



(a)

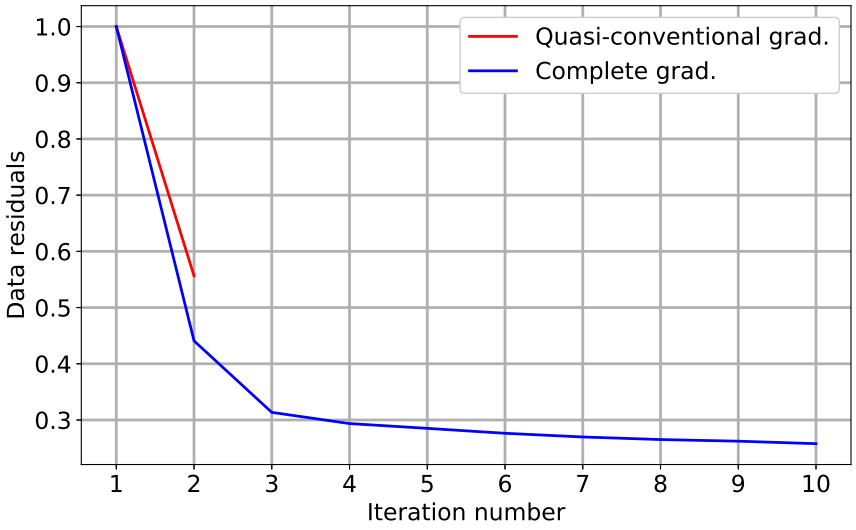


(b)

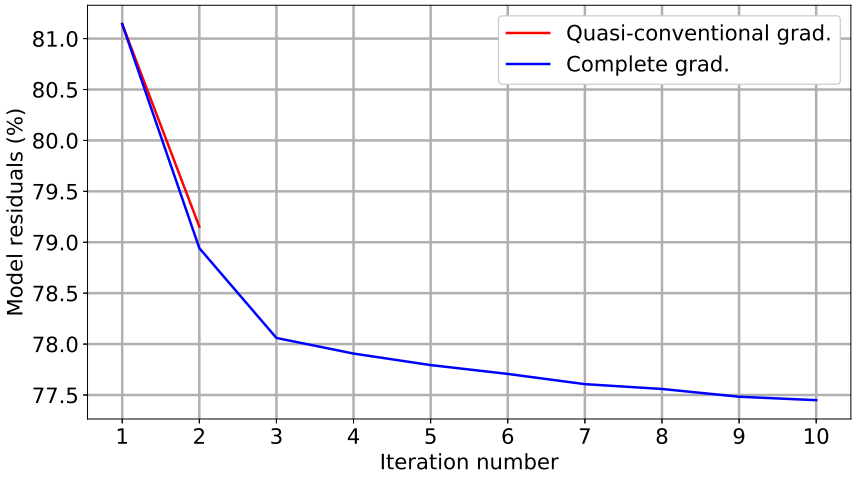


(c)

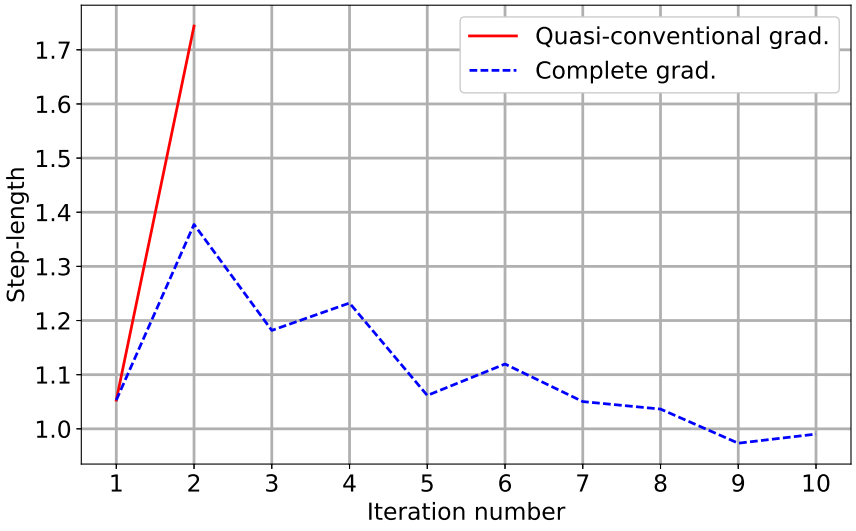
Figure 3.13: Lens results: (a) Data residuals; (b) Model residuals; (c) Step-length.



(a)



(b)



(c)

Figure 3.14: Marmousi2 results: (a) Data residuals; (b) Model residuals; (c) Step-length.

4 Least-squares wave-equation migration to impedance

Our aim is to estimate the acoustic impedance directly from shot gathers using the coupled one-way wave equations. Commonly, the impedance is estimated in two steps, first the shot gathers are migrated to estimate the reflections coefficients, then, some procedure, involving the image conversion from the depth domain to the time domain, is applied to transform the seismic image to impedance (see, e.g., Jones et al., 2018). This approach is justified by the fact that the seismic wavelet is approximately constant in the time domain, while in the depth domain it is distorted by variations in the velocity and in the layers dip (Tygel et al., 1994).

However, once a sufficient velocity model for depth migration is available, the application of a depth migration technique that directly provides at least information of the impedance variations, i.e., the relative acoustic impedance, is highly desirable. In this direction, Fletcher et al. (2012) proposed a methodology based on reverse-time migration and applied point spread functions in the inversion to impedance. Moreover, they used synthetic models to demonstrate that in complicated geological settings, the estimation of the impedance directly in the depth domain can outperform the conventional procedure that involves one-dimensional inversion and the convolutional model.

Here, we develop in detail a migration methodology for acoustic impedance in the depth domain. We start from the first-order approximation of the reflection coefficients. Adding to that, from the discrete form of the modeling equations, the least-squares solution is obtained with a gradient-based method. We also develop an approximate normalization factor to the gradient, that akes to the diagonal of the Gauss–Newton Hessian. Besides that, it is analogous to the conventional deconvolution imaging condition, but adapted to the impedance parameterization of the migration problem and considering the relation between reflection and transmission coefficients.

Note that, differently from the other chapters, here we work on the inverse problem in the discrete form from scratch. The objective is to gain more insight into the model parameters estimation procedure. In this manner, we confront the final expression for the

update direction of the model parameters with the result obtained from the Lagrangian multipliers.

4.1 Impedance parameterization

In Chapter 2, we presented in detail the forward modeling equations and the scattering operators. A short version was presented in Section 3.1. Here, we focus on the impedance parameterization of these modeling equations.

The reflectivity operator in a media with inhomogeneous wavespeed and mass density is

$$\hat{\mathcal{R}}_c^+ = -\frac{1}{2}\hat{\mathcal{H}}_1^{-1}\rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\hat{\mathcal{H}}_1\right) , \tag{4.1}$$

where ρ is the mass density and $\hat{\mathcal{H}}_1$ is the generalized vertical wavenumber, called for short square-root operator, defined as

$$\hat{\mathcal{H}}_1 = \left[\rho\frac{\partial}{\partial x}\left(\frac{1}{\rho}\frac{\partial}{\partial x}\right) + \frac{\omega^2}{c^2}\right]^{1/2} , \tag{4.2}$$

where c is the compressional wavespeed and we are considering a two-dimensional model. Moreover, the hat over $\hat{\mathcal{H}}_1$ indicates it is an integral operator that acts on the lateral coordinate x (for more details see Appendix B, Section B.3). From equations 4.1 and 4.2, we observe that the square-root operator accounts for lateral variations in the mass density.

In order to simplify these equations. We neglect lateral variations in the mass density and consider that scattering is angle-independent. This way, we approximate $\hat{\mathcal{H}}_1$ by

$$\hat{\mathcal{H}}_1^{scat} \approx \frac{\omega}{c} . \tag{4.3}$$

Due to the connection of $\hat{\mathcal{H}}_1$ with the vertical wavenumber in a laterally invariant model, see Section 2.2.2, this approximation is equivalent to set $k_x = 0$ in the vertical wavenumber expression. It is important to highlight that $\hat{\mathcal{H}}_1$ is also related to the wavefield extrapolation. Therefore, we preserve the wide-angle extrapolation accuracy by representing this operator as

$$\hat{\mathcal{H}}_1^{extrap} = \left[\frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2}\right]^{1/2} , \tag{4.4}$$

where the lateral variations in the mass density were neglected in order to have some consistency with the approximation in $\hat{\mathcal{H}}_1^{scat}$. Besides that, it is not expected that the mass density plays an important role in the wavefield extrapolation. We also highlight that

although the approximations in the operators $\hat{\mathcal{H}}_1^{scat}$ and $\hat{\mathcal{H}}_1^{extrap}$ may seem inconsistent. In practice, this approach may work if the contrasts in the wavespeed and mass-density model are small, due to the angle-independent approximation, and for model variations mainly in the vertical.

Thus, we substitute the approximation $\hat{\mathcal{H}}_1^{scat}$, equation 4.3, in equation 4.1 and obtain

$$\hat{\mathcal{R}}_c^+ \approx R_c^+ = -\frac{1}{2} Z \frac{\partial}{\partial z}\left(\frac{1}{Z}\right) , \tag{4.5}$$

where $Z = \rho c$ is the acoustic impedance. Note that after the approximation, the reflectivity becomes frequency-independent and it is not an integral operator anymore, see also the discussion in Section 3.1.2.

Equation 4.5 can also be written as

$$R_c^+ = \frac{1}{2} \frac{\partial \ln\left(Z/Z_0\right)}{\partial z} , \tag{4.6}$$

where Z_0 is a constant. The substitution of this result into the continuous upgoing modeling equation 3.2 and the relation $\hat{\mathcal{T}}_c^- = -\hat{\mathcal{R}}_c^+$ provides

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1^{extrap} P^- - \frac{1}{2} \frac{\partial \ln\left(Z/Z_0\right)}{\partial z} \left(P^+ - P^-\right) . \tag{4.7}$$

In a discontinuous and discrete model, the angle-independent reflectivity, equation 4.6, can be written as

$$R^+(x, z + \Delta z/2, z - \Delta z/2) = \frac{Z(x, z + \Delta z/2) - Z(x, z - \Delta z/2)}{Z(x, z + \Delta z/2) + Z(x, z - \Delta z/2)} , \tag{4.8}$$

where R^+ is the angle-independent reflection coefficient. We expand equation 4.8 in a first-order Taylor series and obtain

$$R^+ \approx \left.\frac{\partial R^+}{\partial z}\right|_{\Delta z/2=0} (\Delta z/2) + \left.\frac{\partial R^+}{\partial z}\right|_{-\Delta z/2=0} (-\Delta z/2) , \tag{4.9}$$

where we already considered that $R^+|_{\Delta z/2=0} = 0$. Developing the partial derivatives, we obtain

$$R^+ \approx \left[\frac{1}{Z_2 + Z_1}\frac{\partial Z_2}{\partial z} - \frac{Z_2 - Z_1}{\left(Z_2 + Z_1\right)^2}\frac{\partial Z_2}{\partial z}\right]_{\Delta z/2=0} (\Delta z/2) \tag{4.10}$$

$$+ \left[\frac{-1}{Z_2 + Z_1}\frac{\partial Z_1}{\partial z} - \frac{Z_2 - Z_1}{\left(Z_2 + Z_1\right)^2}\frac{\partial Z_1}{\partial z}\right]_{-\Delta z/2=0} (-\Delta z/2) , \tag{4.11}$$

where $Z_2 = Z(x, z + \Delta z/2)$ and $Z_1 = Z(x, z - \Delta z/2)$. The evaluation of the coefficients

inside the brackets provides

$$R^+ \approx \frac{1}{2Z}\frac{\partial Z}{\partial z}\left(\Delta z/2\right) - \frac{1}{2Z}\frac{\partial Z}{\partial z}\left(-\Delta z/2\right) , \tag{4.12}$$

where $Z = Z(x, z)$. Or equivalently, we have

$$\boxed{R^+ \approx \frac{1}{2Z}\frac{\partial Z}{\partial z}\Delta z = R_c^+ \Delta z\,.} \tag{4.13}$$

This result confirms the connection between the reflection coefficient R^+ and the angle-independent reflectivity R_c^+ define in equation 4.6. It will be important in the derivation of the upgoing wavefield partial derivative with respect to the acoustic impedance.

In an acoustic medium, the relationship between the scattering coefficients (see,e.g., Berkhout, 2014a) are given by

$$T^+ = 1 + R^+ , \tag{4.14}$$
$$R^- = -R^+ , \tag{4.15}$$
$$T^- = 1 - R^+ , \tag{4.16}$$

where $T^{\pm}$ are the transmission coefficients, the superscript $+$ is related to the scattering of the downgoing wavefield and the superscript $-$ is related to the scattering of the upgoing wavefield. Figure 2.3 sketches the relation between the scattering operators and the associated wavefields.

4.2 Integral modeling equations

In a vertical portion of the wavespeed model that is homogeneous. The decoupled one-way Green's functions are

$$\frac{\partial G^+}{\partial z} = -i\hat{\mathcal{H}}_1^{extrap} G^+ - \Delta z \delta(x - x')\delta(z - z_{n+1})\,, \tag{4.17}$$
$$\frac{\partial G^-}{\partial z} = i\hat{\mathcal{H}}_1^{extrap} G^- + \Delta z \delta(x - x')\delta(z - z_{n-1})\,. \tag{4.18}$$

These Green's functions are used to build integral representations for extrapolation from a boundary, see Section 2.3.3. After that, applying a procedure similar to the Born expansion to detach scattering orders, Section 2.4, we obtain

$$P_{j+1}^+(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)\left(R^- P_j^- + T^+ P_{j+1}^+ + S^+\Delta z\right)(x, z_n, \omega)\,, \tag{4.19}$$
$$P_{j+1}^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left(R^+ P_{j+1}^+ + T^- P_{j+1}^-\right)(x, z_n, \omega)\,, \tag{4.20}$$

where j is related to scattering order, z_n denotes the nth boundary, $z_{n\pm1} = z_n \pm \Delta z$ and $n = 0, 1, 2, 3, \ldots, N$. In these equations, we considered the angle-independent reflection and transmission coefficients. This way, the action of these coefficients on the wavefields is performed by simple multiplication at the corresponding grid point.

Note that equation 4.19 should be read as the downgoing wavefield calculated at depth level z_{n+1} for one frequency and one lateral position x', therefore, we have a scalar. The same understanding holds for the upgoing wavefield. Additionally, the Green's operators are implemented with the complex Padé Fourier finite-difference (for more details see Appendix F).

The downgoing source is

$$S^+(x', z_n, \omega; x_s, z_s) = \begin{cases} -\frac{i}{2\Delta z}\hat{\mathcal{H}}_1^{-1}\delta(x' - x_s)C(\omega), z_n = z_s\,, \\ 0, z_n \neq z_s\,, \end{cases} \tag{4.21}$$

where $C(\omega)$ is the wavelet in the frequency domain, e.g. the Ricker wavelet.

Also note that equations 4.19 and 4.20 feed one another in a recursive scheme that starts with P_0^+ given a known downgoing source wavefield S^+. For $j = -1$, we have primary reflections modeled by

$$P_0^+(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)\left(T^+P_0^+ + S^+\Delta z\right)(x, z_n, \omega)\,, \tag{4.22}$$

$$P_0^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left(R^+P_0^+ + T^-P_0^-\right)(x, z_n, \omega)\,. \tag{4.23}$$

Therefore, equations 4.19 and 4.20 form the basis of the modeling algorithm used in the inversion methodology. The downgoing wavefield is calculated from the model top to bottom and the upgoing counterpart from the model bottom to top. In this recursive scheme, it is necessary to keep only the downgoing wavefield P^- between iterations. See also Sections 2.4.3 and 3.1.4 for a detailed explanation of these modeling equations.

4.3 Inverse problem

In this section, we develop the discrete inverse problem. We adopt a methodology based on data fitting, that is, the gradient of a misfit measure between the observed and calculated data is used to iteratively update an initial estimate of the acoustic impedance. The procedure is similar to the one sketched in Figure 3.1, but instead of $\boldsymbol{R}^+$, we want to estimate an acoustic impedance vector $\boldsymbol{Z}$. Moreover, we start from a homogeneous impedance model and our objective is to recover the impedance variations, i.e., the relative acoustic impedance (see, e.g., Assis et al., 2019b).

In the formulation, we consider that variations in the impedance model gives rise

only to variations in the upgoing wavefield. The complete formulation should also involve variations in the downgoing wavefield. In the Appendix G, Section G.3.2, we develop the complete formulation using the Lagrangian multipliers and the forward modeling equations in the differential form.

Here, we omit the subscript related to scattering order to simplify notation. In the implementation, as the iteration progresses, the upgoing wavefield is kept the same between iterations such that the recursive modeling equations 4.19 and 4.20 are able to model internal multiples.

4.3.1 Misfit function

We estimate the acoustic impedance by solving the least-squares misfit function,

$$E(\mathbf{Z}) = \frac{1}{2}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\|\mathbf{D}^-_{sl} - \mathcal{S}_s\mathbf{P}^-_{sl}(\mathbf{Z})\|_2^2\,, \tag{4.24}$$

where s is the shot index, N_s is the number of shots, l is the frequency index, N_ω is the number of frequencies, $\mathbf{D}^-_{sl}$ is a column vector with M elements and it is the upgoing observed data, $\mathbf{P}^-_{sl}$ is a column vector with M elements and it is the calculated data at all spatial positions, $\mathcal{S}_s$ samples the wavefield at the receivers position and $\|.\|_2^2$ is the L^2 norm squared over the receivers position. Moreover, $\mathbf{Z}$ is the impedance model vector,

$$\boldsymbol{Z} = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_m \\ \vdots \\ Z_M \end{bmatrix}\,, \tag{4.25}$$

in which the model parameters domain has N_x horizontal positions and N_z vertical levels, this way the total number of elements is $M = N_x \times N_z$. The first N_x elements are related to the shallowest depth level and $\boldsymbol{Z} \in \mathbb{R}^M$. The vectors $\boldsymbol{D}^-_{sl}$ and $\mathcal{S}_s\boldsymbol{P}^-_{sl}$ are arranged in the same form but they belong to $\mathbb{C}^M$. These data vectors are different from zero only at the recording positions.

The misfit function in 4.24, for one shot and one frequency, can be written as

$$E_{sl} = \frac{1}{2}\left(\mathbf{D}^-_{sl} - \mathcal{S}_s\boldsymbol{P}^-_{sl}\right)^t\left(\mathbf{D}^-_{sl} - \mathcal{S}_s\boldsymbol{P}^-_{sl}\right)^*\,, \tag{4.26}$$

where t denotes the transpose and the asterisk denotes the complex conjugate. We defined

the dot product as the sum over the grid points, i.e.,

$$\boldsymbol{A}_{sl}^{t}\boldsymbol{A}_{sl}^{*} = \sum_{m=1}^{M} A_{slm}A_{slm}^{*}\,, \tag{4.27}$$

where $\boldsymbol{A}$ represents the data residuals. After some algebraic manipulations of equation 4.26, we obtain

$$\begin{aligned} E_{sl} = &\frac{1}{2}\left(\mathbf{D}_{sl}^{-}\right)^{t}\left(\mathbf{D}_{sl}^{-}\right)^{*} - \frac{1}{2}\left(\mathbf{D}_{sl}^{-}\right)^{t}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{*} - \frac{1}{2}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{t}\left(\mathbf{D}_{sl}^{-}\right)^{*} \\ &+ \frac{1}{2}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{t}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{*}\,. \end{aligned} \tag{4.28}$$

For a complex scalar a, we have that

$$2\,\boldsymbol{Re}\left\{a\right\} = a + a^{*}\,, \tag{4.29}$$

where $\boldsymbol{Re}\{.\}$ denotes the real-part operator. We recognize this identity in the second and third terms of equation 4.28, rearranging the terms, we obtain

$$E_{sl} = \frac{1}{2}\left(\mathbf{D}_{sl}^{-}\right)^{t}\left(\mathbf{D}_{sl}^{-}\right)^{*} - \boldsymbol{Re}\left\{\left(\mathbf{D}_{sl}^{-}\right)^{t}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{*}\right\} + \frac{1}{2}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{t}\left(\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right)^{*}\,. \tag{4.30}$$

Then, we take the gradient and obtain

$$\nabla E_{sl} = -\,\boldsymbol{Re}\left\{\left(\mathbf{D}_{sl}^{-}\right)^{t}\left(\mathcal{S}_{s}\frac{\partial\boldsymbol{P}_{sl}^{-}}{\partial\boldsymbol{Z}}\right)^{*}\right\} + \boldsymbol{Re}\left\{\left[\left(\mathcal{S}_{s}\frac{\partial\boldsymbol{P}_{sl}^{-}}{\partial\boldsymbol{Z}}\right)^{t}\right]^{*}\mathcal{S}_{s}\boldsymbol{P}_{sl}^{-}\right\}\,, \tag{4.31}$$

where the gradient of the misfit function with respect to impedance is a column vector and should be understood as

$$\nabla E_{sl} = \begin{bmatrix} \dfrac{\partial E_{sl}}{\partial Z_1} \\ \dfrac{\partial E_{sl}}{\partial Z_2} \\ \vdots \\ \dfrac{\partial E_{sl}}{\partial Z_M} \end{bmatrix}\,. \tag{4.32}$$

The Jacobian matrix of the wavefield, for one shot and one frequency, in a model grid

with M elements is defined as

$$\frac{\partial \boldsymbol{P}^{-}}{\partial \boldsymbol{Z}} = \begin{bmatrix} \frac{\partial P_1^-}{\partial Z_1} & \frac{\partial P_1^-}{\partial Z_2} & \cdots \frac{\partial P_1^-}{\partial Z_M} \\ \frac{\partial P_2^-}{\partial Z_1} & \frac{\partial P_2^-}{\partial Z_2} & \cdots \frac{\partial P_2^-}{\partial Z_M} \\ \vdots & \vdots & \vdots \\ \frac{\partial P_M^-}{\partial Z_1} & \frac{\partial P_M^-}{\partial Z_2} & \cdots \frac{\partial P_M^-}{\partial Z_M} \end{bmatrix} . \tag{4.33}$$

Finally, we use the dot product identity $\boldsymbol{A}^t\boldsymbol{B} = \boldsymbol{B}^t\boldsymbol{A}$ in the first term of equation 4.31, reorganize the terms and obtain

$$\boxed{-\nabla E\left(\boldsymbol{Z}\right) = \sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} \boldsymbol{Re}\left\{\left(\mathcal{S}_s\frac{\partial \boldsymbol{P}_{sl}^{-}}{\partial \boldsymbol{Z}}\right)^{\dagger}\left(\mathbf{D}_{sl}^{-} - \mathcal{S}_s\boldsymbol{P}_{sl}^{-}\right)\right\} .} \tag{4.34}$$

where † denotes transpose and complex conjugate, i.e., the adjoint, we reintroduced the sum over shots and frequencies. Besides that, we multiplied by -1 to obtain a descent direction. This expression will be used to estimate the acoustic impedance iteratively. In the next section, we develop the required partial derivative of the upgoing wavefield. It is noteworthy that we never compute explicitly the Jacobian shown in equation 4.33. The implementation of the misfit-function gradient requires only its action.

4.3.2 Wavefield Jacobian matrix

The partial derivative of the upgoing wavefield with respect to impedance is fundamental in the calculation of the misfit function, equation 4.34. In this section, we manipulate the upgoing modeling equation 4.20 with the reflection coefficient parameterized as function of the impedance. The equation obtained makes easier the subsequent development of the action of the upgoing-wavefield Jacobian matrix on a vector. In many steps, we adapt the notation to improve readability of the expressions.

Preparation of the modeling equation

We parameterize the upgoing modeling equation 4.20 totally as a function of R^+ by considering the relationship

$$T^{-} = 1 - R^{+} . \tag{4.35}$$

Together with the substitution of the approximate reflection coefficient in equation 4.13, we obtain

$$P^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n) \cdot \left[\frac{\Delta z}{2}\frac{\partial ln\,(Z/Z_0)}{\partial z}P^+ + P^- - \frac{\Delta z}{2}\frac{\partial \ln (Z/Z_0)}{\partial z}P^-\right](x, z_n, \omega)\,. \quad (4.36)$$

We rearrange the terms and obtain

$$P^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n) \cdot \left[\frac{\Delta z}{2}\frac{\partial \ln (Z/Z_0)}{\partial z}\left(P^+ - P^-\right) + P^-\right](x, z_n, \omega)\,. \quad (4.37)$$

The introduction of an integration over the depth coordinate, yields

$$\int_{z_n}^{z_{n-1}} P^-(x', z_{n-1}, \omega; x, z)dz = \int_{z_n}^{z_{n-1}} \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z) \cdot \left[\frac{\Delta z}{2}\frac{\partial \ln (Z/Z_0)}{\partial z}\left(P^+ - P^-\right) + P^-\right](x, z, \omega)dz\,, \quad (4.38)$$

where the integration interval goes from the actual layer bottom at position z_n to the top at z_{n-1}. In this interval, the wavespeed and mass-density models are vertically homogeneous as discussed in Section 4.2.

The integrating by parts of the first term provides

$$\int_{z_n}^{z_{n-1}} P^- dz = \left[\hat{\mathcal{G}}^-\left(P^+ - P^-\right)\right]\frac{\Delta z}{2}\ln (Z/Z_0)\Big|_{z_n}^{z_{n-1}} + \int_{z_n}^{z_{n-1}} -\frac{\partial}{\partial z}\left[\hat{\mathcal{G}}^-\left(P^+ - P^-\right)\right]\frac{\Delta z}{2}\ln (Z/Z_0) + \hat{\mathcal{G}}^- P^- dz\,, \quad (4.39)$$

where we omitted the arguments to simplify notation. We label the first term in the right-hand side as I_1, evaluate the boundary values and obtain

$$I_1 = \hat{\mathcal{G}}^-(z_{n-1}; z_{n-1})\left(P^+ - P^-\right)(z_{n-1})\frac{\Delta z}{2}\ln (Z(z_{n-1})/Z_0) - \hat{\mathcal{G}}^-(z_{n-1}; z_n)\left(P^+ - P^-\right)(z_n)\frac{\Delta z}{2}\ln (Z(z_n)/Z_0)\,, \quad (4.40)$$

where we exhibit only the depth argument to simplify notation.

Considering that $Z(z_{n-1})$ and $Z(z_n)$ are quantities in the same layer, although near the top and bottom boundaries, they must be equal due to the vertical homogeneity considered in each layer interval. Furthermore, the Green's operator in the first term

does not perform any extrapolation and the one in the second term takes the wavefields to the position z_{n-1}. From these observations, we believe that it is acceptable to assume that

$$I_1 = 0\,. \tag{4.41}$$

We use this result in equation 4.39 and obtain

$$\int_{z_n}^{z_{n-1}} P^- dz = \int_{z_n}^{z_{n-1}} -\frac{\partial}{\partial z}\left[\hat{\mathcal{G}}^-\left(P^+ - P^-\right)\right]\frac{\Delta z}{2}\ln\left(Z/Z_0\right) + \hat{\mathcal{G}}^- P^-\, dz\,. \tag{4.42}$$

We perform the vertical derivative in the right-hand side, impose that the integrands must be equal at $z = z_n$ and obtain

$$\begin{aligned} P^-(x', z_{n-1}, \omega) = &-\frac{\partial\hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)}{\partial z_n}\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(P^+ - P^-\right)\right](x, z_n, \omega) \\ &- \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-}{\partial z_n}\right)\right](x, z_n, \omega) \\ &+ \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)P^-(x, z_n, \omega)\,. \end{aligned} \tag{4.43}$$

The developments in this section provided three options for the upgoing modeling equation. It seems that forward modeling with equation 4.36 or 4.43 is less practical than using the estimated impedance to calculate the reflection coefficient and then applying equation 4.20. However, equation 4.43 makes easier the task of developing an expression for the upgoing wavefield derivative with respect to impedance in the following sections.

Action of the Jacobian matrix on a vector

The action of the upgoing-wavefield Jacobian matrix, for one shot and one frequency, on a vector $\Delta\boldsymbol{Z}$ can be calculated from the first-order Taylor expansion,

$$\Delta\boldsymbol{P}^-(\boldsymbol{Z}) = \boldsymbol{P}^-(\boldsymbol{Z} + \Delta\boldsymbol{Z}) - \boldsymbol{P}^-(\boldsymbol{Z}) \approx \left.\frac{\partial\boldsymbol{P}^-}{\partial\boldsymbol{Z}}\right|_{\Delta\boldsymbol{Z}=0}\Delta\boldsymbol{Z}\,, \tag{4.44}$$

where the wavefield Jacobian matrix is M by M, equation 4.33, the wavefield and impedance vectors are arranged as shown in equation 4.25.

The perturbed wavefield at one grid point is given by

$$\Delta P_m^-(\boldsymbol{Z}) = P_m^-(\boldsymbol{Z} + \Delta\boldsymbol{Z}) - P_m^-(\boldsymbol{Z}) \approx \left.\frac{\partial P_m^-}{\partial\boldsymbol{Z}}\right|_{\Delta Z_k=0}\Delta\boldsymbol{Z}\,, \tag{4.45}$$

where $P_m^-(\boldsymbol{Z}) = P^-(x, z, \omega, \boldsymbol{Z})$ is the upgoing wavefield at one grid point and the deriva-

tive $\partial P_m^-/\partial \boldsymbol{Z}$ is a row vector. For a physical interpretation of $P_m^-(\boldsymbol{Z})$ in the context of migration parameterized as a function of the reflection coefficients, see the discussion that follows equation 3.31 and Figure 3.2.

Repeating equation 4.43 and exhibiting only the dependence on the impedance, we have

$$\begin{aligned} P_m^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right) = &-\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(P^+ - P^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right)\right)\right] \\ &-\hat{\mathcal{G}}^-\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right)}{\partial z_n}\right)\right] + \hat{\mathcal{G}}^- P^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right), \end{aligned} \tag{4.46}$$

where the calculated wavefield in the left-hand side is at one grid point.

In order to linearize the calculated upgoing wavefield around $\Delta\boldsymbol{Z}$, we substitute in the right-hand side the linearized logarithm,

$$\ln\left(Z+\Delta Z\right) - \ln\left(Z_0\right) \approx \ln\left(Z/Z_0\right) + \frac{\Delta Z}{Z}, \tag{4.47}$$

where Z_0 is a constant and Z should be understood as a reference value. Additionally, we substitute in 4.46 the wavefield as a reference value plus a perturbation, i.e.,

$$P_m^-(\boldsymbol{Z}+\Delta\boldsymbol{Z}) = P_m^-(\boldsymbol{Z}) + \Delta P_m^-(\boldsymbol{Z}), \tag{4.48}$$

and obtain

$$\begin{aligned} P_m^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right) \approx &-\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\frac{\Delta z}{2}\left[\left(\ln\left(Z/Z_0\right)+\frac{\Delta Z}{Z}\right)\left(P^+ - P^-\right) - \ln\left(Z/Z_0\right)\Delta P^-\right] \\ &-\hat{\mathcal{G}}^-\frac{\Delta z}{2}\left[\left(\ln\left(Z/Z_0\right)+\frac{\Delta Z}{Z}\right)\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-}{\partial z_n}\right) - \ln\left(Z/Z_0\right)\frac{\partial\Delta P^-}{\partial z_n}\right] \\ &+\hat{\mathcal{G}}^-\left[P^- + \Delta P^-\right], \end{aligned} \tag{4.49}$$

where we neglected the variations higher than first-order, i.e, we considered

$$\frac{\Delta Z}{Z}\Delta P^- \qquad \text{and} \qquad \frac{\Delta Z}{Z}\frac{\partial\Delta P^-}{\partial z_n}. \tag{4.50}$$

The upgoing wavefield calculated at the reference impedance is given by

$$\begin{aligned} P_m^-\left(\boldsymbol{Z}\right) = &-\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(P^+ - P^-\left(\boldsymbol{Z}\right)\right)\right] \\ &-\hat{\mathcal{G}}^-\frac{\Delta z}{2}\left[\ln\left(Z/Z_0\right)\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-\left(\boldsymbol{Z}\right)}{\partial z_n}\right)\right] + \hat{\mathcal{G}}^- P^-\left(\boldsymbol{Z}\right). \end{aligned} \tag{4.51}$$

The subtraction of equation 4.51 from 4.49 provides

$$P_m^-\left(\boldsymbol{Z}+\Delta\boldsymbol{Z}\right)-P_m^-\left(\boldsymbol{Z}\right)\approx-\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(P^+-P^-\right)-\ln\left(Z/Z_0\right)\Delta P^-\right]$$
$$-\hat{\mathcal{G}}^-\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)-\ln\left(Z/Z_0\right)\frac{\partial\Delta P^-}{\partial z_n}-\frac{2}{\Delta z}\Delta P^-\right], \quad (4.52)$$

where in the right-hand side we omitted the upgoing wavefield dependence on $\boldsymbol{Z}$.

Therefore, from the Taylor expansion formula, equation 4.44, one element of the upgoing perturbed wavefield vector is given by

$$\left(\frac{\partial P_m^-}{\partial\boldsymbol{Z}}\Delta\boldsymbol{Z}\right)\left(x',z_{n-1},\omega\right)\approx-\frac{\partial\hat{\mathcal{G}}^-(x',z_{n-1},\omega;x,z_n)}{\partial z_n}$$
$$\cdot\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(P^+-P^-\right)-\ln\left(Z/Z_0\right)\Delta P^-\right](x,z_n,\omega)$$
$$-\hat{\mathcal{G}}^-(x',z_{n-1},\omega;x,z_n)$$
$$\cdot\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)-\ln\left(Z/Z_0\right)\frac{\partial\Delta P^-}{\partial z_n}-\frac{2}{\Delta z}\Delta P^-\right](x,z_n,\omega), \quad (4.53)$$

where in the left-hand side the subscript m and the arguments (x',z_{n-1},ω) indicate in different ways the same information, i.e., we have one element of the upgoing wavefield.

In preparation to define the action of the adjoint of the upgoing-wavefield Jacobian matrix on a vector, consider the dot product

$$\boldsymbol{Re}\left\{\Delta\boldsymbol{Z}^t\left(\frac{\partial\boldsymbol{P}^-}{\partial\boldsymbol{Z}}\right)^\dagger\Delta\boldsymbol{P}^-\right\}=\boldsymbol{Re}\left\{\sum_{m=1}^{M}\Delta\boldsymbol{Z}^t\left(\frac{\partial P_m^-}{\partial\boldsymbol{Z}}\right)^\dagger\Delta P_m^-\right\}, \quad (4.54)$$

where † denotes the transpose and complex conjugate. This is an identity between two real numbers, thus, we have the equivalent form

$$\boldsymbol{Re}\left\{\left(\Delta\boldsymbol{P}^-\right)^\dagger\left(\frac{\partial\boldsymbol{P}^-}{\partial\boldsymbol{Z}}\right)\Delta\boldsymbol{Z}\right\}=\boldsymbol{Re}\left\{\sum_{m=1}^{M}\left(\Delta P_m^-\right)^*\left(\frac{\partial P_m^-}{\partial\boldsymbol{Z}}\Delta\boldsymbol{Z}\right)\right\}. \quad (4.55)$$

It is convenient to work with this expression, since we know how to calculate the action of the wavefield derivative via equation 4.53.

Now, we prepare to calculate one row of the action of the adjoint Jacobian matrix on a vector. In this manner, we observe that the adjoint of Green's operator physically means a change in the direction of extrapolation, that is, the approximate relation holds

$$\left[\hat{\mathcal{G}}^-(x',z_{n-1},\omega;x,z_n)\right]^\dagger\approx\hat{\mathcal{G}}^+(x,z_n,\omega;x',z_{n-1}), \quad (4.56)$$

in which the approximation comes from the negligence of evanescent waves (see, e.g.,

Wapenaar, 1998). This expression is easily verified in a homogeneous media. Consider the Green's functions described by equations 4.17 and 4.18. Applying the Fourier transform over the horizontal coordinate x and solving for depth positions that are out of the support of the impulsive source, we obtain

$$\widetilde{G}^{+}(k_x, z_n + \Delta z, \omega; k_x, z_n) = e^{-ik_z \Delta z}\,, \tag{4.57}$$

$$\widetilde{G}^{-}(k_x, z_n - \Delta z, \omega; k_x, z_n) = e^{ik_z \Delta z}\,, \tag{4.58}$$

where we considered $\widetilde{\hat{\mathcal{H}}_1} = k_z$, see Section 2.2.2 and a unitary boundary condition. This result confirms that a least for homogeneous media the approximate relation in equation 4.56 holds. Also note that the transpose in 4.56 could have been omitted, because, in analogy with the discrete case in the space-frequency domain, the operators $\hat{\mathcal{G}}^{\pm}$ are symmetric, at least for the acoustic and constant mass-density case.

Finally, we gathered all the necessary information to develop an expression for the adjoint operation required in the partial derivative of the misfit function in 4.34. In this manner, we multiply equation 4.53 by $(\Delta P_m^-)^*$ and obtain

$$\begin{aligned}
&\left[\left(\Delta P_m^-\right)^* \frac{\partial P_m^-}{\partial \boldsymbol{Z}} \Delta \boldsymbol{Z}\right](x', z_{n-1}, \omega) \approx -\left(\Delta P^-\right)^*(x', z_{n-1}, \omega) \frac{\partial \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)}{\partial z_n} \\
&\cdot \frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(P^+ - P^-\right) - \ln\left(Z/Z_0\right)\Delta P^-\right](x, z_n, \omega) \\
&- \left(\Delta P^-\right)^*(x', z_{n-1}, \omega)\hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-}{\partial z_n}\right)\right](x, z_n, \omega) \\
&- \left(\Delta P^-\right)^*(x', z_{n-1}, \omega)\hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n) \\
&\cdot \frac{\Delta z}{2}\left[-\ln\left(Z/Z_0\right)\frac{\partial \Delta P^-}{\partial z_n} - \frac{2}{\Delta z}\Delta P^-\right](x, z_n, \omega)\,.
\end{aligned} \tag{4.59}$$

We apply the adjoint on both sides of this equation and obtain

$$\begin{aligned}
&\left[\Delta \boldsymbol{Z}^t\left(\frac{\partial P_m^-}{\partial \boldsymbol{Z}}\right)^{\dagger}\Delta P_m^-\right](x, z_n, \omega) \approx -\left[\frac{\partial \hat{\mathcal{G}}^+(x, z_n, \omega; x', z_{n-1})}{\partial z_n}\Delta P^-(x', z_{n-1}, \omega)\right] \\
&\cdot \frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(P^+ - P^-\right)^* - \ln\left(Z/Z_0\right)\left(\Delta P^-\right)^*\right](x, z_n, \omega) \\
&- \left[\hat{\mathcal{G}}^+(x, z_n, \omega; x', z_{n-1})\Delta P^-(x', z_{n-1}, \omega)\right]\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-}{\partial z_n}\right)^*\right](x, z_n, \omega) \\
&- \left[\hat{\mathcal{G}}^+(x, z_n, \omega; x', z_{n-1})\Delta P^-(x', z_{n-1}, \omega)\right] \\
&\cdot \frac{\Delta z}{2}\left[-\ln\left(Z/Z_0\right)\left(\frac{\partial \Delta P^-}{\partial z_n}\right)^* - \frac{2}{\Delta z}\left(\Delta P^-\right)^*\right](x, z_n, \omega)\,,
\end{aligned} \tag{4.60}$$

where we used the approximate relation between the up- and downgoing Green's functions

in equation 4.56. Then, neglecting variations higher than first-order in the wavefield,

$$\Delta P^{-}\left(\Delta P^{-}\right)^{*} \qquad \text{and} \qquad \Delta P^{-}\left(\frac{\partial \Delta P^{-}}{\partial z_n}\right)^{*}, \tag{4.61}$$

we arrive at

$$\begin{aligned}\left[\Delta \boldsymbol{Z}^{t}\left(\frac{\partial P_m^{-}}{\partial \boldsymbol{Z}}\right)^{\dagger}\Delta P_m^{-}\right](x,z_n,\omega) \approx -\left[\frac{\partial \hat{\mathcal{G}}^{+}(x,z_n,\omega;x',z_{n-1})}{\partial z_n}\Delta P^{-}(x',z_{n-1},\omega)\right] \\ \cdot\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(P^{+}-P^{-}\right)^{*}\right](x,z_n,\omega) - \left[\hat{\mathcal{G}}^{+}(x,z_n,\omega;x',z_{n-1})\Delta P^{-}(x',z_{n-1},\omega)\right] \\ \cdot\frac{\Delta z}{2}\left[\frac{\Delta Z}{Z}\left(\frac{\partial P^{+}}{\partial z_n}-\frac{\partial P^{-}}{\partial z_n}\right)^{*}\right](x,z_n,\omega)\,.\end{aligned} \tag{4.62}$$

We observe that each term in the right-hand side has a factor ΔZ. Thus, we can easily isolate the action of the adjoint partial-derivative of the upgoing wavefield on the perturbed wavefield from the impedance vector to obtain

$$\begin{aligned}\left[\left(\frac{\partial P_m^{-}}{\partial \boldsymbol{Z}}\right)^{\dagger}\Delta P_m^{-}\right](x,z_n,\omega) \approx -\left[\frac{\partial \hat{\mathcal{G}}^{+}(x,z_n,\omega;x',z_{n-1})}{\partial z_n}\Delta P^{-}(x',z_{n-1},\omega)\right] \\ \cdot\frac{\Delta z}{2Z}\left[\left(P^{+}-P^{-}\right)^{*}\right](x,z_n,\omega) - \left[\hat{\mathcal{G}}^{+}(x,z_n,\omega;x',z_{n-1})\Delta P^{-}(x',z_{n-1},\omega)\right] \\ \cdot\frac{\Delta z}{2Z}\left[\left(\frac{\partial P^{+}}{\partial z_n}-\frac{\partial P^{-}}{\partial z_n}\right)^{*}\right](x,z_n,\omega)\,.\end{aligned} \tag{4.63}$$

Thus, this expression provides a vector with M elements in the same domain of the impedance vector, the so-called model domain, due to the perturbed wavefield at one position of the grid. The comparison with equation 4.34, shows that the misfit-function gradient can be interpreted as the sum of the partial derivative of the upgoing wavefield action on each sample of the data residuals. Moreover, the data residuals can be seen as perturbed wavefields generated by the discrepancy between the impedance vector considered in the forward modeling and the impedance that gave rise to the observed data.

Now, we add more one simplification to equation 4.63 to make the implementation practical. In a depth interval that is vertically homogeneous and away from the source term, the downgoing Green's operator is a solution of

$$\frac{\partial \hat{\mathcal{G}}^{+}}{\partial z_n} = -i\hat{\mathcal{H}}_1^{extrap}\hat{\mathcal{G}}^{+}\,. \tag{4.64}$$

Considering that the vertical derivative of the Green's operator in equation 4.63 is a consequence of the impedance parameterization. We approximate this square-root operator

as in the scattering case, equation 4.3, and obtain

$$\frac{\partial \hat{\mathcal{G}}^{+}}{\partial z_n} = -i\frac{\omega}{c}\hat{\mathcal{G}}^{+} \,. \tag{4.65}$$

In this manner, equation 4.63 simplifies to

$$\begin{aligned}
&\left[\left(\frac{\partial P_m^-}{\partial \boldsymbol{Z}}\right)^{\dagger} \Delta P_m^-\right](x, z_n, \omega) \approx \\
&\frac{\Delta z}{2Z}\left[i\frac{\omega}{c}\left(P^+ - P^-\right)^* - \left(\frac{\partial P^+}{\partial z_n} - \frac{\partial P^-}{\partial z_n}\right)^*\right](x, z_n, \omega) \\
&\cdot \hat{\mathcal{G}}^{+}(x, z_n, \omega; x', z_{n-1}) \Delta P^-(x', z_{n-1}, \omega) \,.
\end{aligned} \tag{4.66}$$

Due to its simplicity, we implement this approximation. Additionally, before applying the data residuals backpropagation, we divide them by the angle-independent transmission coefficient. In the discussion that follows, this procedure will be clarified.

4.3.3 Relationship with the continuous case

Equation 4.66 is equivalent to the negative of the misfit-function gradient in equation 4.34, disregarding the sum over sources and frequencies. In the Appendix G, Section G.3.2, a similar result was obtained from the application of the Lagrangian multipliers on the inverse problem we want to solve and it is given by

$$-\frac{dE}{dZ} = -\frac{1}{2Z}\left[\left(\frac{\partial \Lambda^+}{\partial z} - \frac{\partial \Lambda^-}{\partial z}\right)\left(P^+ - P^-\right)^* + \left(\Lambda^+ - \Lambda^-\right)\left(\frac{\partial P^+}{\partial z} - \frac{\partial P^-}{\partial z}\right)^*\right] , \tag{4.67}$$

where $\Lambda^{\pm}$ are the adjoint wavefields and they have dimension of pressure times distance (for more details see Appendix G, Section G.2.2). Note that we omitted the sum over sources and frequencies.

In order to uncover the similarity between equations 4.66 and 4.67, in equation 4.67 we neglect the adjoint downgoing wavefield, i.e., discard variations in the forward downgoing wavefield, and obtain

$$-\frac{dE}{dZ} = \frac{1}{2Z}\left[\frac{\partial \Lambda^-}{\partial z}\left(P^+ - P^-\right)^* + \Lambda^-\left(\frac{\partial P^+}{\partial z} - \frac{\partial P^-}{\partial z}\right)^*\right] . \tag{4.68}$$

The negligence of the secondary sources and boundary conditions in the equation that describes the adjoint upgoing wavefield provides

$$\frac{\partial \Lambda^-}{\partial z} \approx -i\hat{\mathcal{H}}_1 \Lambda^- \,. \tag{4.69}$$

Adding to that, we approximate $\hat{\mathcal{H}}_1$ as in the scattering operator and obtain

$$\frac{\partial \Lambda^-}{\partial z} \approx -i\frac{\omega}{c}\Lambda^- \,. \tag{4.70}$$

The substitution in equation 4.68 and rearrangement of terms provides

$$\boxed{-\frac{dE}{dZ} \approx \frac{1}{2Z}\left[-i\frac{\omega}{c}\left(P^+ - P^-\right)^* + \left(\frac{\partial P^+}{\partial z} - \frac{\partial P^-}{\partial z}\right)^*\right]\Lambda^- \,.} \tag{4.71}$$

From Section 3.2.2, the integral representation of Λ^- for one common source and one angular frequency is

$$\Lambda^-(x, z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x, z_{n+1}, \omega; x', z_n)\left[\frac{1}{T^-}\Lambda^- - \Delta z \mathcal{S}^\dagger\left(D^- - \mathcal{S}_s P^-\right)\right](x', z_n, \omega)\,. \tag{4.72}$$

This equation represents the backpropagation of the data residuals from the receivers position with compensation of amplitude related to transmission effects. Therefore, after the additional assumptions in this subsection, we conclude that the approximate result in equation 4.71 is equivalent to the equation 4.66. The differences in the signs and in the scaling factor Δz are reconciled between these equations by the data residuals in equation 4.72.

4.3.4 Approximate Gauss–Newton update

The second-order Taylor expansion of the misfit function is

$$E(\boldsymbol{Z}_0 + \Delta\boldsymbol{Z}) \approx E(\boldsymbol{Z}_0) + \sum_{l=1}^{M}\frac{\partial E}{\partial Z_l}\bigg|_{\Delta Z_l = 0}\Delta Z_l + \frac{1}{2}\sum_{l=1}^{M}\sum_{m=1}^{M}\Delta Z_m \frac{\partial^2 E}{\partial Z_m \partial Z_l}\bigg|_{\Delta Z_{ml}=0}\Delta Z_l \,, \tag{4.73}$$

where the sum is over all grid points of the impedance model. The subscript in Z_l indicates one spatial position (x, z) as defined in equation 4.25. In compact notation, we have

$$E(\boldsymbol{Z}_0 + \Delta\boldsymbol{Z}) \approx E(\boldsymbol{Z}_0) + (\Delta\boldsymbol{Z})^t \nabla E\bigg|_{\Delta\boldsymbol{Z}=0} + \frac{1}{2}(\Delta\boldsymbol{Z})^t \frac{d^2E}{d\boldsymbol{Z}^2}\bigg|_{\Delta\boldsymbol{Z}=0}\Delta\boldsymbol{Z}\,. \tag{4.74}$$

We define the total impedance as the sum of a reference value and a perturbation, i.e., $\boldsymbol{Z} = \boldsymbol{Z}_0 + \Delta\boldsymbol{Z}$. Thus, we use this definition to derive equation 4.74 with respect to $\boldsymbol{Z}$. Moreover, we impose that this derivative must be zero in order to find the stationary points of E and obtain

$$\frac{d^2E}{d\boldsymbol{Z}^2}\bigg|_{\Delta\boldsymbol{Z}_k=0}\Delta\boldsymbol{Z}_k = -\nabla E\bigg|_{\Delta\boldsymbol{Z}_k=0}\,, \tag{4.75}$$

where we introduced a subscript k that indicates iteration number. This equation is the basis of Newton's method (Vogel, 2002).

We need an expression for the second derivative of the misfit function, i.e., the Hessian. The first derivative of the misfit function is given by equation 4.34, deriving once more, we obtain

$$\frac{d^2E}{d\boldsymbol{Z}^2} = -\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} \boldsymbol{Re}\left\{\left(\mathcal{S}_s\frac{\partial^2\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}^2}\right)^\dagger\left(\mathbf{D}^-_{sl}-\mathcal{S}_s\boldsymbol{P}^-_{sl}\right) - \left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)^\dagger\left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)\right\} . \quad (4.76)$$

If we consider that the data residuals are small, we can keep only the linear term and, consequently, the approximation is obtained

$$\frac{d^2E}{d\boldsymbol{Z}^2} \approx \mathbf{H} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} \boldsymbol{Re}\left\{\left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)^\dagger\left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)\right\} . \quad (4.77)$$

The substitution of this result in equation 4.75 provides the Gauss–Newton method (see, e.g., Pratt et al., 1998).

We specialize the approximate Hessian in 4.77 even more by considering only the diagonal terms

$$\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\left(\mathbf{H_D}\right)_{sl} = \sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} Diag\left\{\left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)^\dagger\left(\mathcal{S}_s\frac{\partial\boldsymbol{P}^-_{sl}}{\partial\boldsymbol{Z}}\right)\right\} , \quad (4.78)$$

where $Diag\{.\}$ extracts the matrix diagonal. In order to obtain an expression for $\mathbf{H_D}$, we apply the adjoint of the partial derivative of the upgoing wavefield, equation 4.63, on its forward derivative, equation 4.53, at the same observation point and obtain

$$\begin{aligned}
&\left[\left(\frac{\partial P^-_m}{\partial\boldsymbol{Z}}\right)^\dagger\left(\frac{\partial P^-_m}{\partial\boldsymbol{Z}}\Delta\boldsymbol{Z}\right)\right](x,z_n,\omega) \approx \\
&\left(\frac{\Delta z}{2Z}\right)^2\left[\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)^*\hat{\mathcal{G}}^+\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\Delta Z\left(P^+-P^-\right)\right. \\
&+\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)^*\hat{\mathcal{G}}^+\hat{\mathcal{G}}^-\Delta Z\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right) \\
&+\left(P^+-P^-\right)^*\frac{\partial\hat{\mathcal{G}}^+}{\partial z_n}\frac{\partial\hat{\mathcal{G}}^-}{\partial z_n}\Delta Z\left(P^+-P^-\right) \\
&\left.+\left(P^+-P^-\right)^*\frac{\partial\hat{\mathcal{G}}^+}{\partial z_n}\hat{\mathcal{G}}^-\Delta Z\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)\right] ,
\end{aligned} \quad (4.79)$$

where we omitted the shot and frequency indexes and neglected terms higher than first-order with the form $\Delta z^2\Delta P^-$. We observe that each factor in the right-hand side has a

factor ΔZ. This way, we separate the product between the wavefield derivatives from the impedance vector and obtain

$$\begin{aligned}
&\left[\left(\frac{\partial P_m^-}{\partial \boldsymbol{Z}}\right)^{\dagger}\left(\frac{\partial P_m^-}{\partial \boldsymbol{Z}}\right)\right](x, z_n, \omega) \approx \\
&\left(\frac{\Delta z}{2Z}\right)^2\left[\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)^{*}\hat{\mathcal{G}}^{+}\frac{\partial \hat{\mathcal{G}}^{-}}{\partial z_n}\left(P^+-P^-\right)\right. \\
&+\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)^{*}\hat{\mathcal{G}}^{+}\hat{\mathcal{G}}^{-}\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right) \\
&+\left(P^+-P^-\right)^{*}\frac{\partial \hat{\mathcal{G}}^{+}}{\partial z_n}\frac{\partial \hat{\mathcal{G}}^{-}}{\partial z_n}\left(P^+-P^-\right) \\
&\left.+\left(P^+-P^-\right)^{*}\frac{\partial \hat{\mathcal{G}}^{+}}{\partial z_n}\hat{\mathcal{G}}^{-}\left(\frac{\partial P^+}{\partial z_n}-\frac{\partial P^-}{\partial z_n}\right)\right]. \qquad (4.80)
\end{aligned}$$

In this equation, after the action of the Green's operators, the right-hand side is only composed by zero-lag correlations between wavefields. This observation suggests that in fact, this expression already provides one element of the diagonal of the Gauss–Newton Hessian.

Now, we further simplify equation 4.80. We recall that the Green's operators corresponds to

$$\hat{\mathcal{G}}^{+}=\hat{\mathcal{G}}^{+}(x, z_n, \omega; x', z_{n-1}) \qquad \text{and} \qquad \hat{\mathcal{G}}^{-}=\hat{\mathcal{G}}^{-}(x', z_{n-1}, \omega; x, z_n)\,. \qquad (4.81)$$

We adopt the approximations

$$\hat{\mathcal{G}}^{+}\hat{\mathcal{G}}^{-} \approx \mathcal{I}\,, \qquad (4.82)$$

$$\frac{\partial \hat{\mathcal{G}}^{+}}{\partial z_n} \approx -i\frac{\omega}{c}\hat{\mathcal{G}}^{+}\,, \qquad (4.83)$$

$$\frac{\partial \hat{\mathcal{G}}^{-}}{\partial z_n} \approx i\frac{\omega}{c}\hat{\mathcal{G}}^{-}\,, \qquad (4.84)$$

in which we considered a vertically homogeneous interval, used the definition of the Green's functions, equations 4.17 and 4.18, and the approximate square-root operator for scattering, equation 4.3. The first approximation, basically states that the downgoing Green's function undo the action of the upgoing counterpart. Moreover, we justify the approximation in the Green's functions vertical derivatives by considering that this factor arises from the impedance parameterization. Thus, the action of the square-root operator is approximated analogously to the angle-independent scattering operators.

Hence, one element of the approximate diagonal of the Gauss–Newton Hessian, for

one shot and one frequency, is

$$\boxed{(\mathbf{H_D})_{sl}(x, z_n, \omega) \approx \left(\frac{\Delta z}{2Z}\right)^2 \left[\left|\frac{\partial P_{sl}^{+}}{\partial z_n} - \frac{\partial P_{sl}^{-}}{\partial z_n}\right|^2 + \frac{\omega^2}{c^2}\left|P_{sl}^{+} - P_{sl}^{-}\right|^2\right](x, z_n, \omega)\,,} \tag{4.85}$$

where $|.|$ denotes the magnitude of a complex and we reintroduced the shot and frequency indexes. It is noteworthy that this equation is evaluated at the current known impedance model $\boldsymbol{Z}_k$. After the sum of $(\mathbf{H_D})_{sl}$ over shots and frequencies, we obtain the required approximate diagonal of the Gauss–Newton Hessian $\mathbf{H_D}$.

Further parameterization and impedance update

To reduce scaling issues in the impedance update, we use the parameterization

$$\boldsymbol{Z}_L = \ln(\boldsymbol{Z}/Z_0) \tag{4.86}$$

where the natural logarithm is applied to each sample of the impedance vector and Z_0 is a constant. Thus, using the approximate Newton's method, equation 4.75, the impedance is updated iteratively via

$$\boldsymbol{Z}_L^{k+1} = \boldsymbol{Z}_L^k + \alpha_k \Delta \boldsymbol{Z}_L^k\,, \tag{4.87}$$

where α_k is the step-length calculated with the parabolic fit (Bohlen et al., 2009). The update direction given by

$$\Delta \boldsymbol{Z}_L = -\left[\frac{1}{\boldsymbol{Z}_k^2 (\mathbf{H_D})_k + \epsilon}\right] \nabla E\Big|_{\Delta \boldsymbol{Z}_k = 0} \boldsymbol{Z}_k\,, \tag{4.88}$$

where ϵ is a small quantity included to stabilize the division. The gradient of the misfit function is defined in equation 4.34, and together with the required adjoint wavefield partial derivative, equation 4.63, it is basically the backpropagation of the data residuals from the receivers position, followed by the zero-lag cross-correlation with the composition of wavefields and its vertical derivatives. In the implementation, we smooth the misfit function gradient to reduce spikes in the estimated impedance.

As discussed in the development, the updating term in equation 4.87 is an approximation to the Gauss–Newton method. Furthermore, in analogy with regular seismic imaging, it is the so-called deconvolution imaging condition adapted to our problem (see, e.g., Schleicher et al., 2008).

4.4 Numerical tests

We apply the gradient-based methodology iteratively in the Lens model, (see, e.g., Masaya and Verschuur, 2018) and in the modified Marmousi2 (Pan et al., 2018). The corresponding model dimensions and data acquisition configurations are shown in Tables 3.1 and 3.3. For the Lens model the mass density is considered constant and equal to 1 $km/s.g/cm^3$. For both models, the initial impedance is a constant array with values of 2 $km/s.g/cm^3$.

4.4.1 Lens model

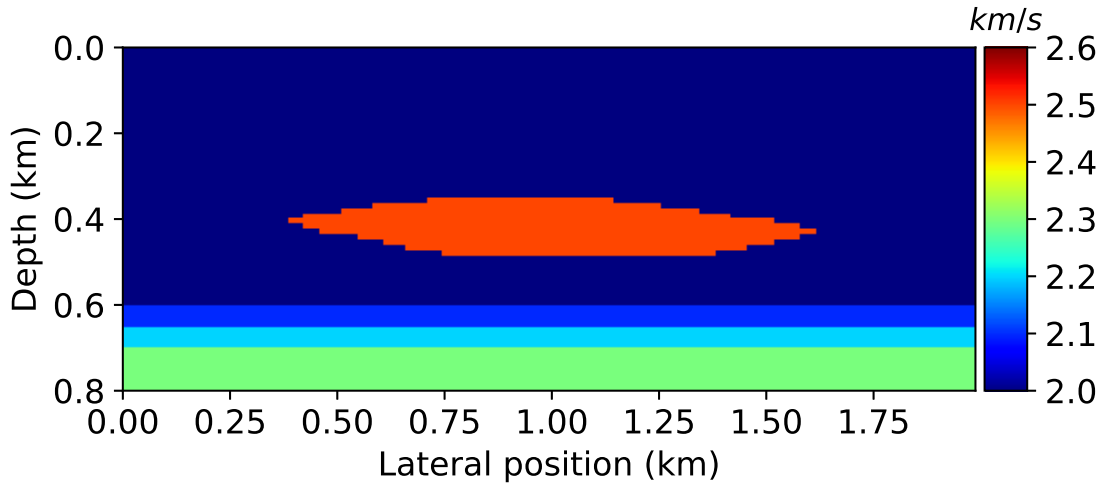
Figure 4.1(a) shows the exact velocity model and Figure 4.1(b) the impedance estimated with the proposed methodology. The lens and the fine layering shapes are well recovered. We use equation 4.8 to transform the estimated acoustic impedance to the reflection coefficients. The obtained image very close to the least-squares result from Chapter 3, Figure 4.2.

Figure 4.2(c) shows profiles in the middle of the impedance and image sections, the impedance profile was scaled and shifted. As expected, the impedance represents interval information, while the reflection coefficients are positioned at the discontinuities of the impedance model. From this result, we also verify that we recovered only impedance variations instead of the absolute values. If the absolute values had been recovered, the magnitude of the fine layering should increase with depth, as in the exact model in Figure 4.1(a). The norm of the data residuals over the iterations indicates that the minimization procedure behaved as expected, Figure 4.3.

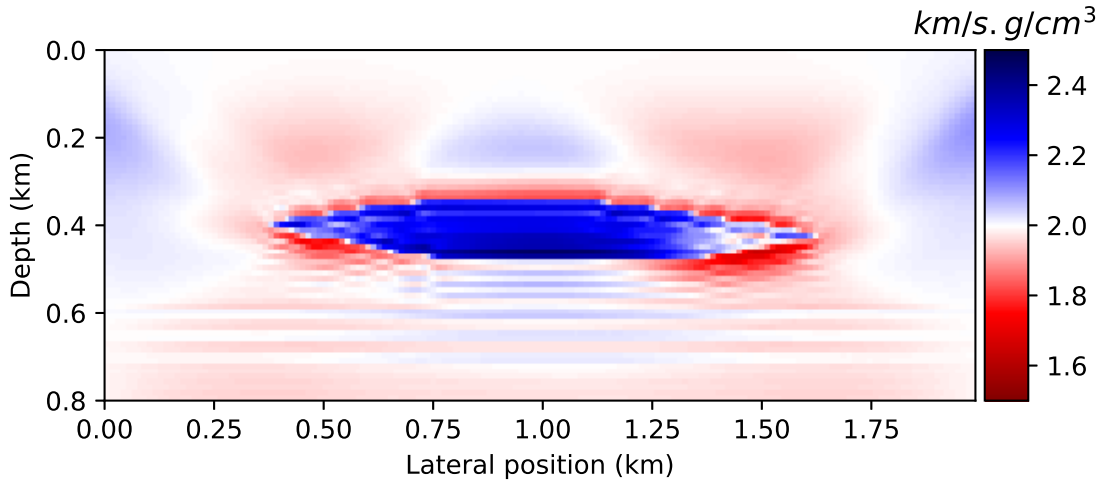
4.4.2 Marmousi2

Figures 4.4(a) and 4.4(b) show the exact and the estimated impedance models. Again, we recovered only the variations of the impedance, Figure 4.4(b). This is the case, because the seismic source has band-limited frequency content. One workaround is to start the inversion from an initial impedance model that contains the low spatial-frequency information, i.e., a smooth initial impedance estimate instead of a constant model as considered here.

The impedance section converted to reflection coefficients is very close to the result of the least-squares imaging, Figures 4.5(a) and 4.5(b). The image obtained from the impedance differs from the least-squares imaging result mainly in the central part, Figure 4.5(c). But in the estimated impedance section, we observe that the correspondent structural information is very well defined, Figure 4.4(b).



(a)



(b)

Figure 4.1: Lens model: (a) Exact velocity model; (b) Estimated acoustic impedance.

Figure 4.6(a) shows the image and impedance profiles, it confirms that the impedance represent interval information, while the image is at the discontinuities of the impedance model. Figure 4.6(b) shows profiles of the exact and estimated impedance at the lateral position 2.83 km. We removed the linear trend, scaled and shifted the exact impedance profile. The local increase and decrease of both profiles are very similar. Again, the norm of the data residuals decreased over iterations, Figure 4.7.

4.5 Summary

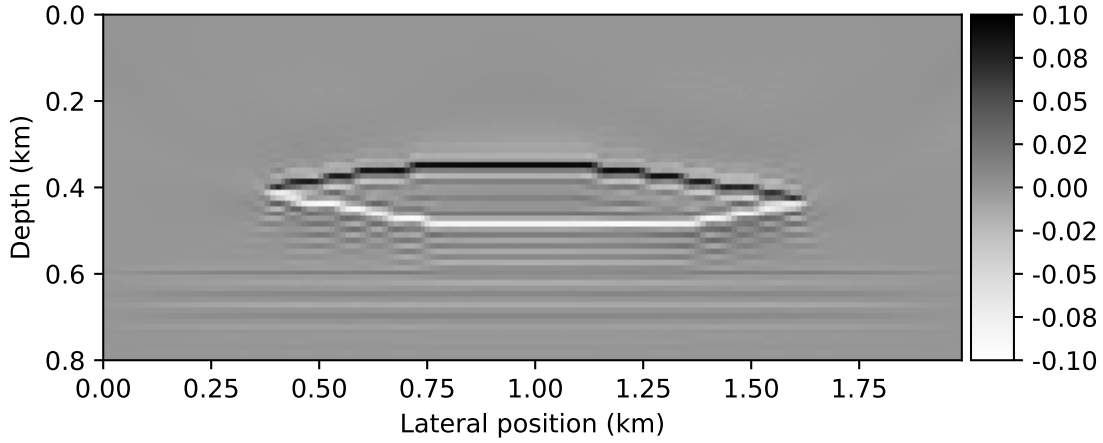
We started this chapter approximating the general expression of the reflectivity operator to the angle-independent case. Moreover, we also neglected the contribution of lateral variations in the mass density to the angle-independent reflectivity. Under these assumptions, we demonstrated that the angle-independent reflectivity is given by

the vertical derivative of the acoustic impedance logarithm. Then, we used this result to parameterize the migration problem as function of the acoustic impedance.

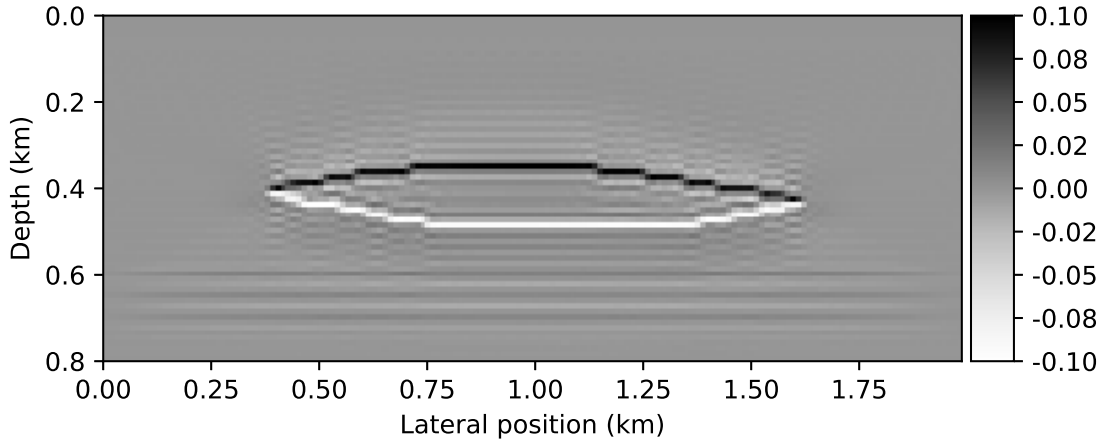
Next, we adapted the upgoing modeling equation in the integral form to the proposed parameterization. Additionally, we developed the partial derivative of the upgoing wavefield with respect to the impedance and also provided the expression of its adjoint action on a vector. The derivation of these equations provided detailed understanding of the misfit-function gradient.

We also derived the approximate diagonal of the Gauss–Newton Hessian for the impedance parameterization. The expression obtained is analogous to the deconvolution imaging condition used in regular seismic imaging, but adapted to the impedance parameterization and taking into consideration the relation between the reflection and transmission coefficients.

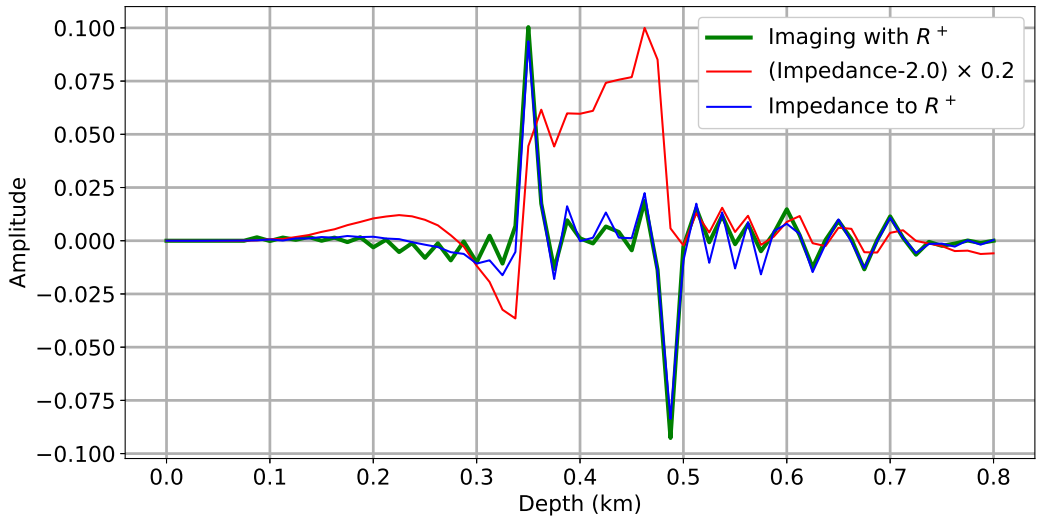
The numerical tests were performed using the exact velocity models. Thus, the detailed information in the velocity may have influenced the estimated impedance sections. However, the results indicate that the impedance parameterization of the imaging problem can provide valuable results that may assist the work of seismic interpreters.



(a)



(b)



(c)

Figure 4.2: Images of the Lens model: (a) Estimated impedance converted to R^+; (b) Direct estimation of R^+; (c) Profiles in the middle of the Lens model.

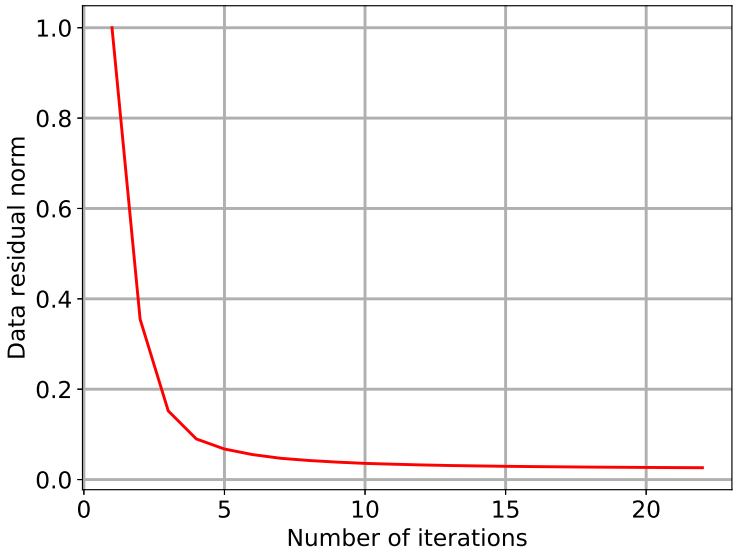
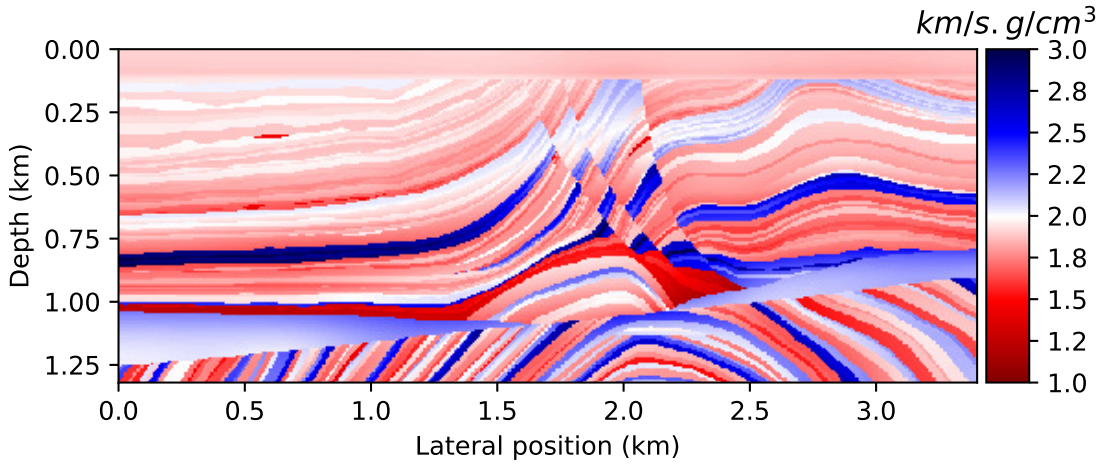
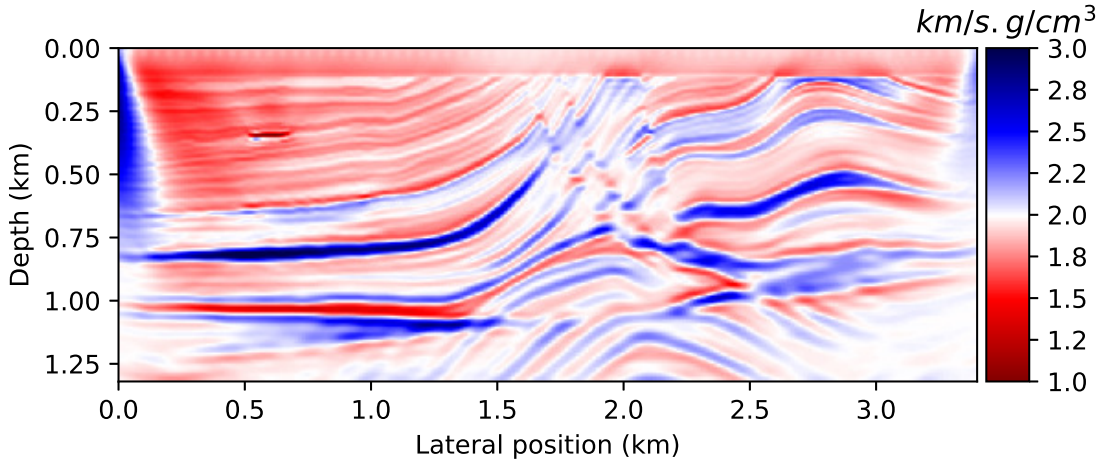


Figure 4.3: Data residuals norm over iterations from the test with the Lens model.

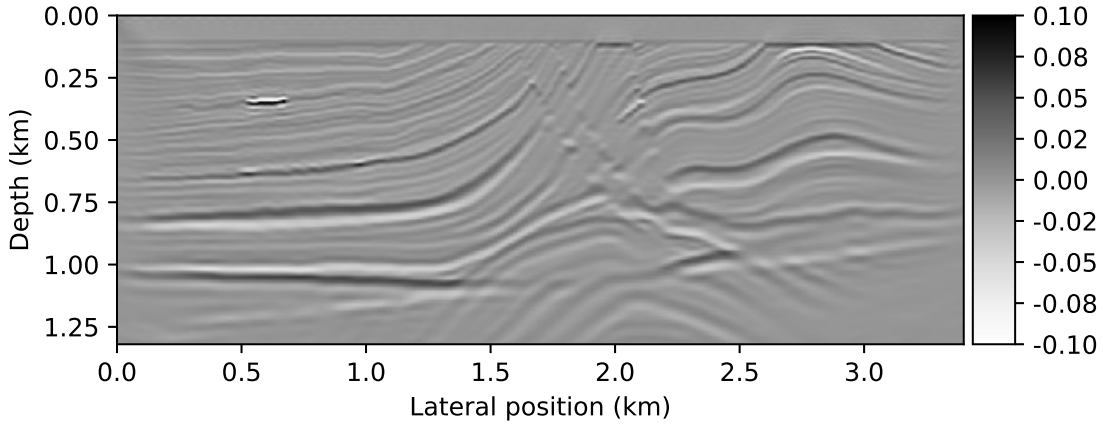


(a)

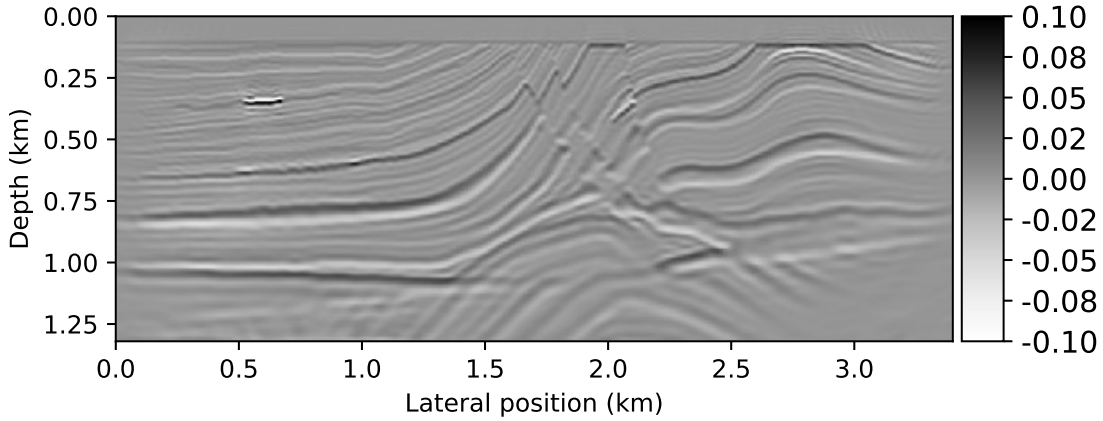


(b)

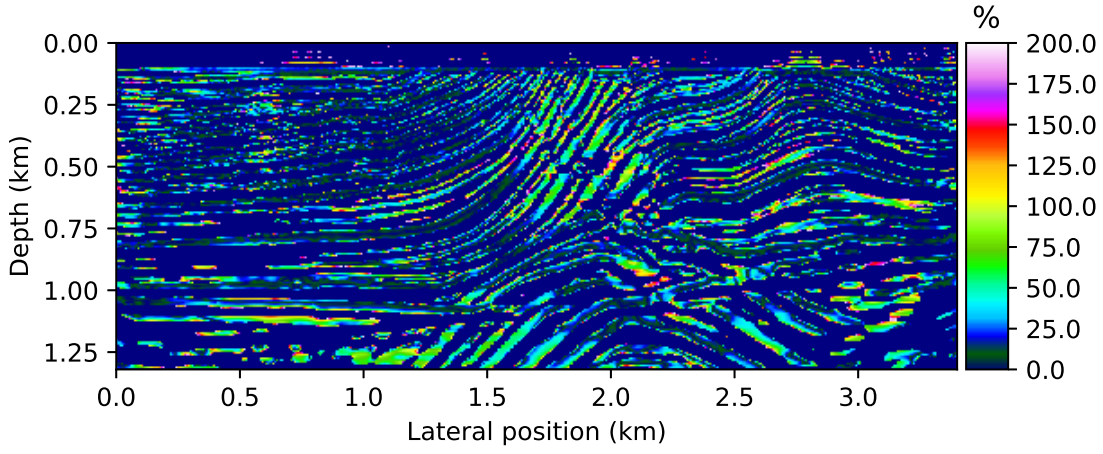
Figure 4.4: Marmousi2 relative acoustic impedance: (a) Exact obtained after $(Exact - trend) \times 0.25 + 1.95$, in which *trend* refers to the smooth part of the exact impedance; (b) Estimated.



(a)

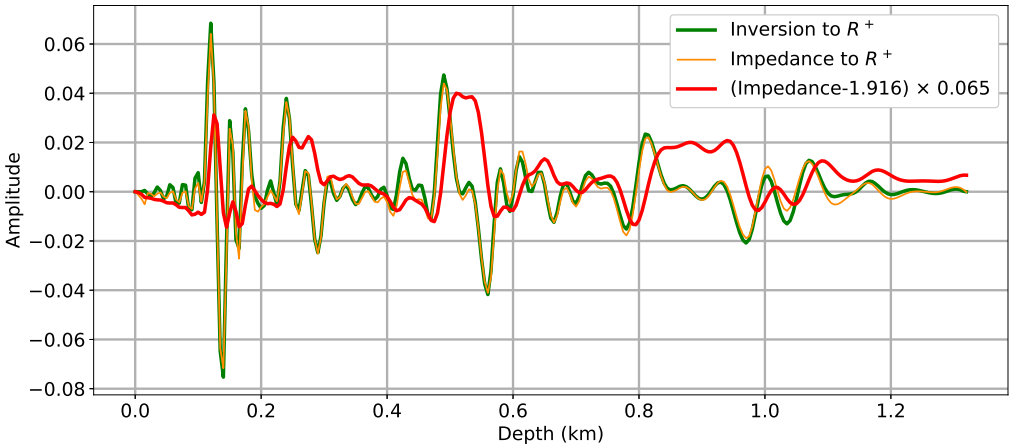


(b)

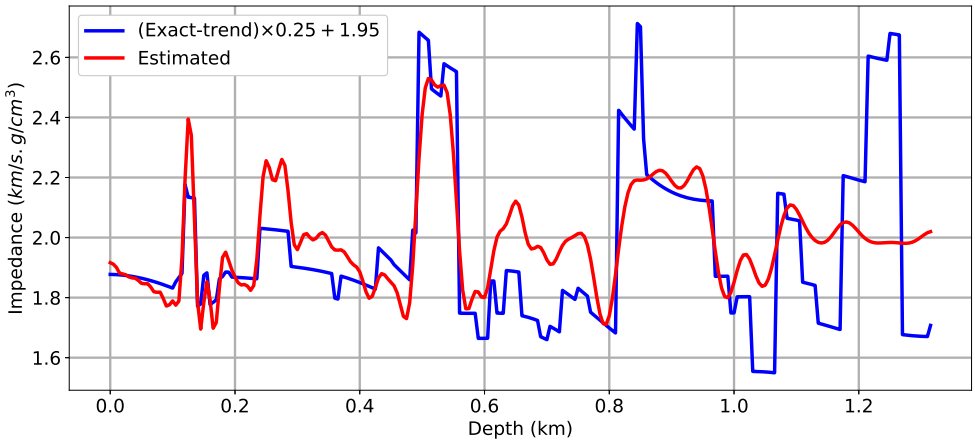


(c)

Figure 4.5: Images from the Marmousi2: (a) Estimated impedance converted to R^+; (b) Direct estimation of R^+; (c) Absolute percentage difference between (a) and (b).



(a)



(b)

Figure 4.6: Marmousi2 profiles at position 2.83 km. (a) Estimated impedance and related images; (b) Estimated and exact acoustic impedance.

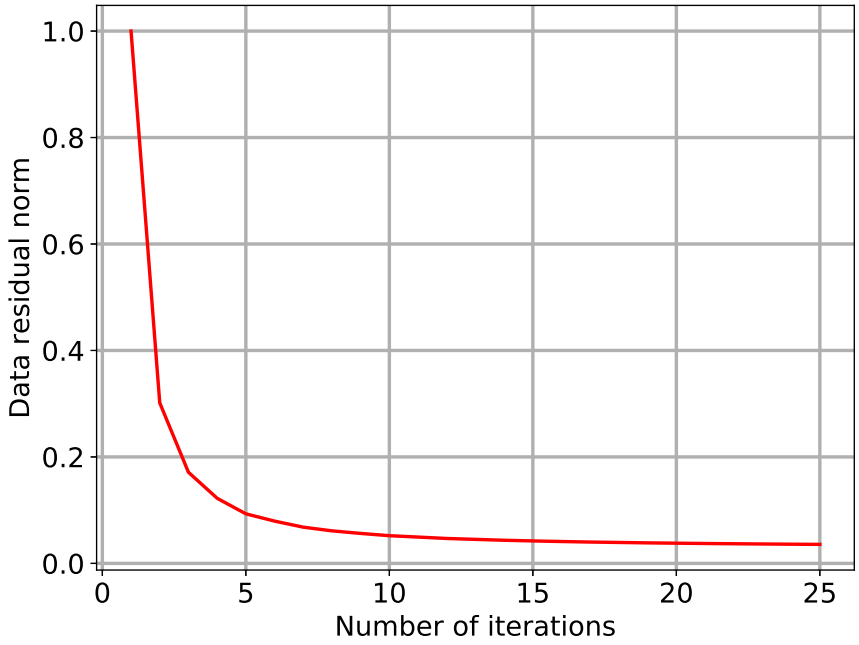


Figure 4.7: Marmousi2 result. Data residual norm over iterations.

5 Velocity estimation with an image-based regularization

Joint Migration Inversion (JMI) is an approach to the seismic inverse problem based on scattering and propagation operators (Berkhout, 2014b). In its most simple form, the reflection scattering operator is given by the stacked seismic image. Primarily, it carries information related to the subsurface impedance contrasts at the boundaries between geologic formations. On the other hand, the estimated velocity model carry information about each layer interval. In this line of thought, given the nature of information in the image and velocity model, it is quite acceptable to think that these quantities are highly uncorrelated. However, with additional analytical relations between these quantities, it is possible to make them approximately correlated. For example, considering the mass-density constant and derivating in the depth axis, the velocity model it is transformed approximately to a reflectivity model. Conversely, a stacked seismic image can be integrated to estimate the acoustic impedance (see, e.g., Assis et al., 2019b).

Relating different classes of model parameters or different data types is very common in geophysical inverse problems. For example, Gallardo and Meju (2003) proposed the cross-gradients constraint, which relates the cross-product of the model parameters gradient to enforce structural similarity in the inversion. Also, imposing structural information from the image to the velocity model has already been proposed using different methodologies. Costa et al. (2008) applied a structurally motivated smoothing constraint in the direction of a potential reflector in slope tomography and showed that this constraint led to geologically more consistent models. In the wave-equation migration velocity analysis, Williamson et al. (2011) solved a differential equation to smooth the updated velocity model consistently with structural information.

In the context of JMI, Maciel et al. (2015) applied morphological operators directly to the updated velocity model. Masaya and Verschuur (2018) proposed a regularization for the inversion part of JMI that relates the image and an approximate image obtained from the velocity model. In this manner, Masaya and Verschuur (2018) improved the structural information in the velocity model gradient and not directly in the velocity like

Williamson et al. (2011) and Maciel et al. (2015).

Regularization of ill-posed inverse problems is fundamental. It has basically two roles, one being to reduce instability and the other to select solution candidates according to desired characteristics of the solution, carried over to the problem by means of the regularization function (Zhdanov, 2015). For example, defining the regularization term as the first derivative of the model parameter penalizes roughness and favors smooth estimates or, in the words of Constable et al. (1987), the simplest model that explains the data.

Inspired by the cross-gradients regularization, we investigate the minimization of the inner-product between the image and the velocity model to improve the structural information in the velocity model. We also analyze the inclusion of each model parameter's gradient in the regularization function. We demonstrate the effectiveness of the methodologies in one synthetic data. The only assumption assumed here is that the information in the image is, to some extent, correlated with velocity contrasts. Also note that the regularization developed here is very similar to the work of Masaya and Verschuur (2018), but the procedure is simpler.

5.1 Regularizing functions

We want to relate the velocity model $c(x, z)$ with the reflection coefficient model $R^+(x, z)$, also interchangeably called image, in which x indicates the horizontal coordinate and z is the depth coordinate. The image R^+ represents the reflection coefficient for incidence from above on a boundary between adjacent model positions. It is assumed to be equivalent to the stacked image provided by most least-squares migration methods.

5.1.1 Function I: Inner product of the model parameters

The first function is defined by the inner product between velocity and image,

$$J_{reg1}(c) = \int_{\Omega} c(\mathbf{x}')R^+(\mathbf{x}')d\mathbf{x}' \,, \tag{5.1}$$

where $\mathbf{x} = (x, z)$ and Ω is as two-dimensional spatial domain in which the model parameters are defined. The kernel of the first derivative with respect to velocity is the image, i.e.,

$$\frac{\partial J_{reg1}}{\partial c} = R^+(\mathbf{x}) \,. \tag{5.2}$$

5.1.2 Function II: Including the gradient of the model parameters

The second function is based on the combination of three relations between the velocity and image using the gradient, i.e.,

$$J_{reg2}(c) = \int_{\Omega} \left[\nabla c(\mathbf{x}') \cdot \nabla R^{+}(\mathbf{x}')\right] d\mathbf{x}' + \frac{1}{2}\int_{\Omega} \left[c(\mathbf{x}')\nabla R^{+}(\mathbf{x}') \cdot \mathbf{n} - R^{+}(\mathbf{x}')\nabla c(\mathbf{x}') \cdot \mathbf{n}\right] d\mathbf{x}' \,, \tag{5.3}$$

where $\mathbf{n} = (a_x, a_z)$, in which a_x and a_z are constants with dimension of distance inverse. After discretization, we set them as

$$a_x = 1/\Delta x \qquad \text{and} \qquad a_z = 1/\Delta z \,. \tag{5.4}$$

Regularization using the gradients aims at penalizing roughness in the model parameters (Constable et al., 1987). Furthermore, we introduced the negative sign in the last term of equation 5.3 with the goal of increasing the correlation between R^{+} and ∇c during the minimization procedure to be presented in the next section. Integration by parts and adoption of homogeneous boundary conditions yields

$$J_{reg2}(c) = \int_{\Omega} c(\mathbf{x}') \left[-\nabla^2 R^{+}(\mathbf{x}') + \nabla R^{+}(\mathbf{x}') \cdot \mathbf{n}\right] d\mathbf{x}' \,. \tag{5.5}$$

Thus, the kernel of the derivative of this regularization function with respect to velocity is

$$\frac{\partial J_{reg2}}{\partial c} = \left[-\nabla^2 R^{+}(\mathbf{x}) + \nabla R^{+}(\mathbf{x}) \cdot \mathbf{n}\right] \,. \tag{5.6}$$

5.1.3 Velocity parameterization

Consider the logarithm of the slowness model

$$\sigma_L = \ln\left(\bar{c}^{init}/c\right) \,, \tag{5.7}$$

where $\bar{c}^{init}$ is the mean of the velocity model. To reduce gradient scaling issues during the velocity update, we propose the parameterization

$$\sigma_{LN} = \frac{\sigma_L - \bar{\sigma}_L^{init}}{\max\left(\sigma_L^{init} - \bar{\sigma}_L^{init}\right)} \,, \tag{5.8}$$

where σ_L^{init} is the logarithm of the initial velocity model and $\bar{\sigma}_L^{init}$ its mean. Therefore, instead of the derivatives of the regularization functions with respect to c, we need those with respect to σ_{LN}. We find that

$$\frac{\partial J_{reg}}{\partial \sigma_{LN}} = \frac{\partial J_{reg}}{\partial c}\frac{\partial c}{\partial \sigma_{LN}} = -M^{init} c \frac{\partial J_{reg}}{\partial c}\,, \tag{5.9}$$

where the derivative of the regularization function is given by equation 5.2 or 5.6. We define the scalar M^{init} as

$$M^{init} = \max\left(\sigma_L^{init} - \bar{\sigma}_L^{init}\right)\,. \tag{5.10}$$

5.2 Inverse problem

In this section, we present the model parameters estimation methodology based on the joint migration inversion (JMI). We apply an iterative least-squares data fitting procedure to estimate the angle-independent reflection coefficient R^+, i.e., the image, and the velocity model c. The forward modeling engine is composed of the integral equations for the down- and upgoing wavefields in an acoustic, two-dimensional and constant mass-density model. In the following sections, we briefly present the modeling equations (for a detailed presentation see Chapter 2).

5.2.1 Overview of the JMI algorithm

Here, we present an overview of the JMI algorithm, Figure 5.1. In the subsequent sections, we discuss the technical details.

First, the algorithm loads the seismic data, the initial velocity model and the initial image is set zero. Additionally, the inversion parameters are defined, e.g., minimum/maximum frequencies, number of iterations and so on. A multiscale approach is adopted, i.e., the minimum frequency is fixed and the maximum frequency is increased after a criteria is satisfied at the end of an iteration. Each frequency range is denominated a stage. Finally, the estimation procedure starts with the imaging workflow represented by the yellow blocks in Figure 5.1.

Next, the inversion sequence is performed to estimate the velocity update. It is sketched by the blue blocks in Figure 5.1. During the inversion, we parameterize the velocity model as a function of the normalized slowness logarithm defined in equation 5.8. This choice improves the scaling of the update direction and, consequently, enhances the effectiveness of the calculated step-length. Note that we update the parameter σ_{LN}, but whenever necessary we transform it back to velocity.

After updating the velocity model, the percentage change of the misfit function

is checked. If the reduction is greater than a tolerance, the algorithm returns to the migration workflow. Otherwise, it is checked if the maximum frequency in the actual iteration is equal to the maximum frequency set to apply JMI, if they agree, the algorithm is finished. On the contrary, the algorithm increments the maximum frequency and returns to the migration. The quantities $\bar{c}^{init}$ and $\bar{\sigma}_L^{init}$ are updated at the beginning of each frequency stage, they are used in equations 5.7, 5.8 and 5.10

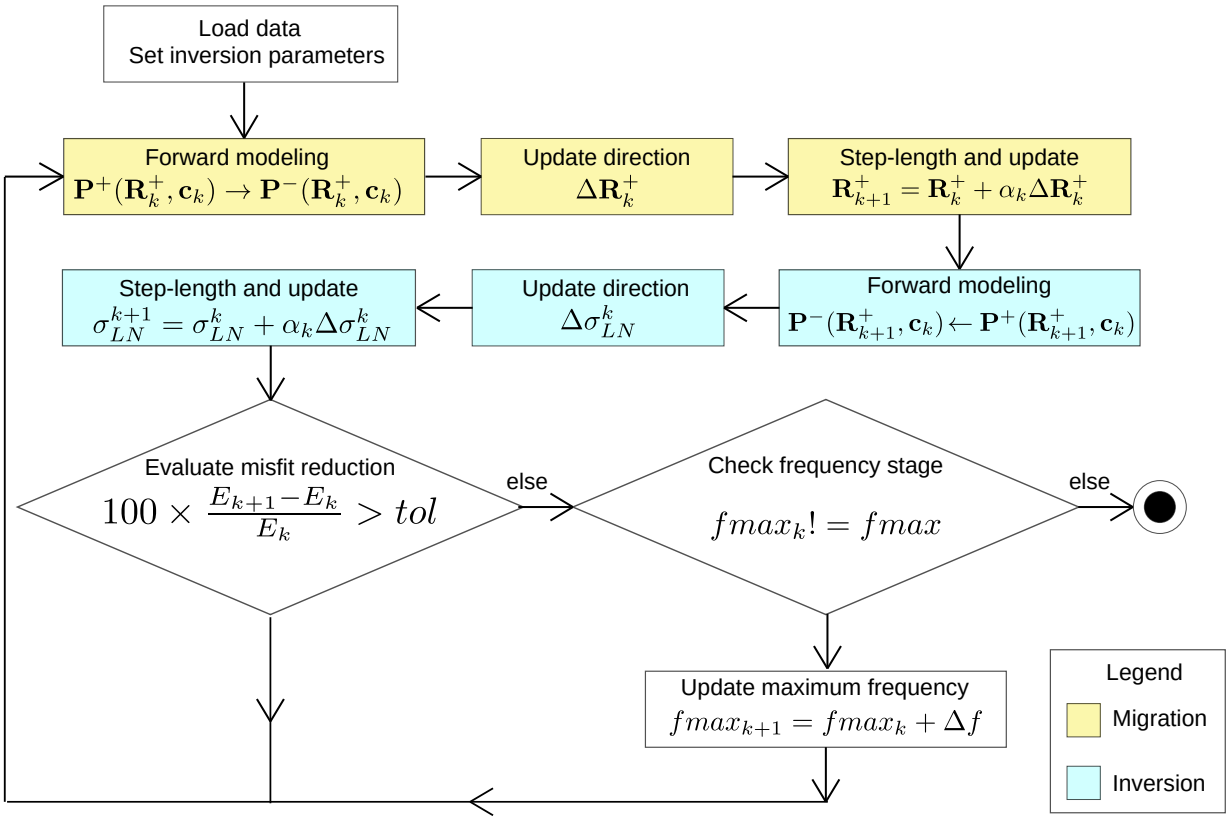


Figure 5.1: Joint migration inversion workflow. The yellow blocks are related to migration and the blue blocks to the inversion.

5.2.2 Statement of the problem

By alternating between imaging and inversion, we apply JMI for the estimation of the reflection coefficient R^+ and the normalized slowness σ_{LN}. This is accomplished by minimizing the least-squares misfit function given by

$$E(\boldsymbol{R}^+, \boldsymbol{\sigma}_{LN}) = \frac{1}{2} \sum_{s=1}^{N_s} \sum_{l=1}^{N_\omega} \| \boldsymbol{D}_{sl}^- - \mathcal{S}_s \boldsymbol{P}_{sl}^-(\boldsymbol{R}^+, \boldsymbol{\sigma}_{LN}) \|_2^2 + \beta J_{reg}(\boldsymbol{\sigma}_{LN}) \,, \tag{5.11}$$

where N_s is the number of shots, s is the shot index, N_ω is the number of frequencies, l is the frequency index, $\boldsymbol{D}_{sl}^-$ is the upgoing measured data in the space-frequency domain (x, z, ω), $\boldsymbol{P}_{sl}^-$ is the upgoing modeled data at all grid points, $\mathcal{S}_s$ samples the upgoing

wavefield at the receiver positions, $\|.\|_2^2$ is the L^2 norm squared over the receivers position, β is a scalar and J_{reg} is one of the regularization functions given by equations 5.1 and 5.5. Consider, for example, the column vector of reflection coefficients

$$\boldsymbol{R}^+ = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \\ \vdots \\ R_M \end{bmatrix}, \tag{5.12}$$

in which m is one grid point, the model parameters domain has N_x horizontal positions and N_z vertical levels, this way the total number of elements is $M = N_x \times N_z$ and the first N_x elements are related to the shallowest depth level and so on. Moreover, $\boldsymbol{R}^+$ belongs to $\mathbb{R}^M$, the same disposition applies to $\boldsymbol{D}_{sl}^-$ and $\boldsymbol{P}_{sl}^-$ that belongs to $\mathbb{C}^M$ and $\boldsymbol{\sigma}_{LN} \in \mathbb{R}^M$.

The recursive integral representations of the modeled downgoing P^+ and upgoing P^- wavefields, at one model grid point, are

$$P_{j+1}^+(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)\left(R^- P_j^- + T^+ P_{j+1}^+ + S^+ \Delta z\right)(x, z_n, \omega), \tag{5.13}$$

$$P_{j+1}^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left(R^+ P_{j+1}^+ + T^- P_{j+1}^-\right)(x, z_n, \omega), \tag{5.14}$$

where j indicates the scattering order, z_n denotes the depth of the nth boundary and $z_{n\pm1} = z_n \pm \Delta z$. Moreover, $\hat{\mathcal{G}}^\pm$ are the Green's operators, $R^\pm$ are the angle-independent reflection coefficients, $T^\pm$ are the angle-independent transmission coefficients and S^+ is the downgoing source. The Green's operators are implemented using the complex Padé Fourier finite-difference (CPFFD) method (Amazonas et al., 2007). See also Appendix F for more details about the CPFFD method.

The zero-order terms model primary reflections and are given by

$$P_0^+(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)\left(T^+ P_0^+ + S^+ \Delta z\right)(x, z_n, \omega), \tag{5.15}$$

$$P_0^-(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left(R^+ P_0^+ + T^- P_0^-\right)(x, z_n, \omega). \tag{5.16}$$

5.2.3 Model parameters update

We estimate the model parameters iteratively using a gradient-based method. Here, we present the discretized partial derivatives. In the Appendix G, we develop the required partial derivatives for the corresponding continuous formulation. We omit the subscript related to scattering order, but as the iterative estimation progress, the same upgoing

wavefield is used between iterations.

Migration

Neglecting variations in the downgoing wavefield, the partial derivative of the misfit function, equation 5.11, with respect to R^+ is

$$-\left(\frac{\partial E}{\partial \boldsymbol{R}^+}\right)_m = -\frac{1}{\Delta z}\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} \boldsymbol{Re}\left\{\left(P^+_{slm} - P^-_{slm}\right)^* \Lambda^-_{slm}\right\}\,, \tag{5.17}$$

where m is one model grid point, $\boldsymbol{Re}\{.\}$ is the real-part operator, P^+_{slm} is the downgoing wavefield at one grid point m, P^-_{slm} is the upgoing wavefield and the asterisk denotes the complex conjugate. Moreover, we multiplied by -1 to obtain a descent direction. For more details about this expression see Chapter 3, Section 3.2.4.

The zero-order terms of the integral representation of the adjoint wavefields are given by

$$\Lambda^-_{sl}(x', z_{n+1}, \omega) = \hat{\mathcal{G}}^+(x', z_{n+1}, \omega; x, z_n)\left[\frac{1}{T^-}\Lambda^-_{sl} - \Delta z \mathcal{S}^\dagger_s\left(\boldsymbol{D}^-_{sl} - \boldsymbol{\mathcal{S}}_s \boldsymbol{P}^-_{sl}\right)\right](x, z_n, \omega)\,, \tag{5.18}$$

$$\Lambda^+_{sl}(x', z_{n-1}, \omega) = \hat{\mathcal{G}}^-(x', z_{n-1}, \omega; x, z_n)\left[\frac{1}{T^+}\Lambda^+_{sl} + R^-\Lambda^-_{sl}\right](x, z_n, \omega)\,, \tag{5.19}$$

where the superscript † denotes transpose and complex conjugate, i.e., the adjoint. Moreover, we believe that the division by the transmission coefficients $T^\pm$ has the potential to balance the amplitude of deep reflectors relatively to the shallow ones.

The image is updated iteratively via

$$\boldsymbol{R}^+_{k+1} = \boldsymbol{R}^+_k - \alpha_k\left[\frac{1}{(\mathbf{H_D})_k + \epsilon}\right]\left.\frac{\partial E}{\partial \boldsymbol{R}^+}\right|_{\Delta \boldsymbol{R}^+_k = 0}\,, \tag{5.20}$$

where $\boldsymbol{R}^+$ is arranged as a column vector of M elements, see equation 5.12. Additionally, k is the iteration number, the step-length α_k is calculated using a parabolic fit (Bohlen et al., 2009), $\mathbf{H_D}$ is an approximation to the diagonal of the Gauss–Newton Hessian and ϵ is a stabilization factor.

The elements of the vector $\mathbf{H_D}$ are composed of the magnitude squared of the factor multiplying the adjoint wavefield in expression 5.17. For a detailed discussion, in the context of the impedance estimation problem, see Section 4.3.4. In the migration part of the stated problem, one element m of the normalization vector is given by

$$(\mathbf{H_D})_m = \sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega}\left(P^+_{slm} - P^-_{slm}\right)^*\left(P^+_{slm} - P^-_{slm}\right)\,, \tag{5.21}$$

that due to the upgoing wavefield term, is a modification of the conventional normalization factor in the deconvolution imaging condition (see, e.g., Schleicher et al., 2008). This modified normalization factor is a consequence of considering variations in the transmission coefficient. See Appendix G, Section G.3 for the development of the migration partial derivative using the Lagrangian multipliers.

Inversion

The partial derivative with respect to $\boldsymbol{\sigma}_{LN}$ is

$$-\left(\frac{\partial E}{\partial \boldsymbol{\sigma}_{LN}}\right)_m = \sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} \boldsymbol{Re}\left\{-iM^{init}\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c_m^2}\left[-P^-_{slm}\left(\Lambda^-_{slm}\right)^* + P^+_{slm}\left(\Lambda^+_{slm}\right)^*\right]\right\} - \beta\left(\frac{\partial J_{reg}}{\partial \boldsymbol{\sigma}_{LN}}\right)_m , \tag{5.22}$$

where again we multiplied by -1 to obtain a descent direction (for a detailed derivation with respect to velocity/wavespeed see Appendix G, Section G.4). One implementation of the generalized vertical wavenumber, called for short square-root operator, inverse action on a wavefield is discussed in Appendix G, Section G.4.1. We observe in this equation that the slowness updates come from the correlation of wavefields along transmission paths (see also the discussion in Chapter 3, Section 3.2.6).

We parameterized the inverse problem as function of $\boldsymbol{\sigma}_{LN}$. Thus, we perform the update via

$$\boldsymbol{\sigma}^{k+1}_{LN} = \boldsymbol{\sigma}^{k}_{LN} + \alpha_k \Delta\boldsymbol{\sigma}^{k}_{LN} , \tag{5.23}$$

where again the step-length α_k is calculated using a parabolic fit. The update direction is given by

$$\left(\Delta\boldsymbol{\sigma}^k_{LN}\right)_m = \frac{1}{H^-_m}\left[\underbrace{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} iM^{init}\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c_m^2}P^-_{slm}\left(\Lambda^-_{slm}\right)^*}_{F^-_I} - \beta^-_k\underbrace{\left(\frac{\partial J_{reg}}{\partial \boldsymbol{\sigma}_{LN}}\right)_m}_{F^-_{II}}\right] + \frac{1}{H^+_m}\left[\underbrace{\sum_{s=1}^{N_s}\sum_{l=1}^{N_\omega} -iM^{init}\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c_m^2}P^+_{slm}\left(\Lambda^+_{slm}\right)^*}_{F^+_I} - \beta^+_k\underbrace{\left(\frac{\partial J_{reg}}{\partial \boldsymbol{\sigma}_{LN}}\right)_m}_{F^+_{II}}\right] . \tag{5.24}$$

The scaling factors $\beta_k^{\pm}$ calculated as follows

$$\beta_k^{\pm} = \beta \frac{\max\left(|F_I^{\pm}|\right)}{\max\left(|F_{II}^{\pm}|\right)} \,. \tag{5.25}$$

This expression was designed to make the dimension of the terms in equation 5.24 consistent. The normalization factors are given by

$$H_m^- = \sum_{s=1}^{N_s} \sum_{l=1}^{N_\omega} \left| iM^{init} \hat{\mathcal{H}}_1^{-1} \Delta z \frac{\omega^2}{c_m^2} P_{slm}^- \right|^2 + \epsilon^- \,, \tag{5.26}$$

$$H_m^+ = \sum_{s=1}^{N_s} \sum_{l=1}^{N_\omega} \left| -iM^{init} \hat{\mathcal{H}}_1^{-1} \Delta z \frac{\omega^2}{c_m^2} P_{slm}^+ \right|^2 + \epsilon^+ \,, \tag{5.27}$$

where $\epsilon^{\pm}$ are scalars used to stabilize the division. After updating $\boldsymbol{\sigma}_{LN}$ with equation 5.23, we recover the velocity model via

$$\boldsymbol{c}^{k+1} = \bar{c}^{init} \exp\left\{ -M^{init} \boldsymbol{\sigma}_{LN}^{k+1} - \bar{\sigma}_L^{init} \right\} \,, \tag{5.28}$$

where the quantities in the right-hand side are defined in Section 5.1.3. This transformation back to velocity is important because the forward and adjoint modeling equations are more naturally applied using the velocity model.

5.3 Numerical examples

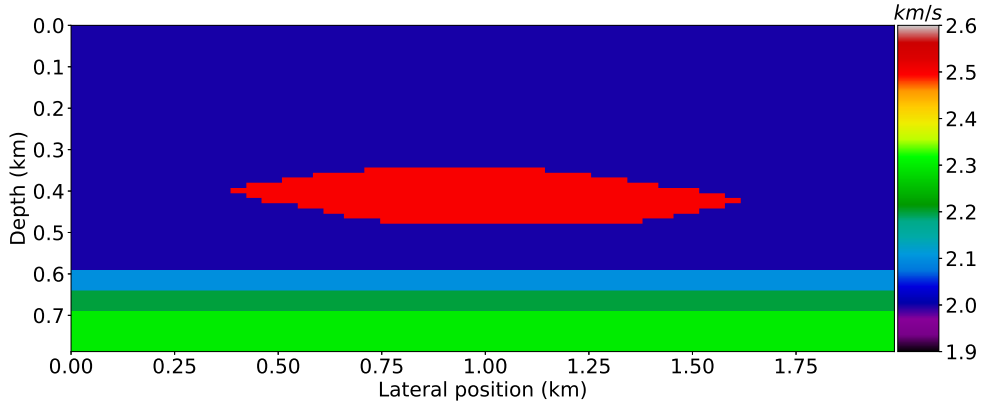
We demonstrate the use of the proposed regularization functions in a model that is composed of a lens-shaped anomaly and fine layering at the bottom, (see, e.g., Masaya and Verschuur, 2018). The same algorithm is used for modeling and inversion. The data are simulated with a 20 Hz Ricker wavelet. For more details about the Lens model and the acquisition configuration see Table 3.1.

The inversion works in frequency stages. Whenever the change of the misfit function is less than a tolerance of 1% or a maximum number of 20 iterations was reached, a new frequency stage is initiated. At the beginning of each new stage, we slightly increase the tolerance such that at the final stage the tolerance is 1.5%. The minimum frequency is fixed at the first sample after 0 Hz and the first maximum frequency is 15 Hz. Then, at each new frequency stage, the maximum frequency is increased 5 Hz and the last maximum frequency set to 40 Hz.

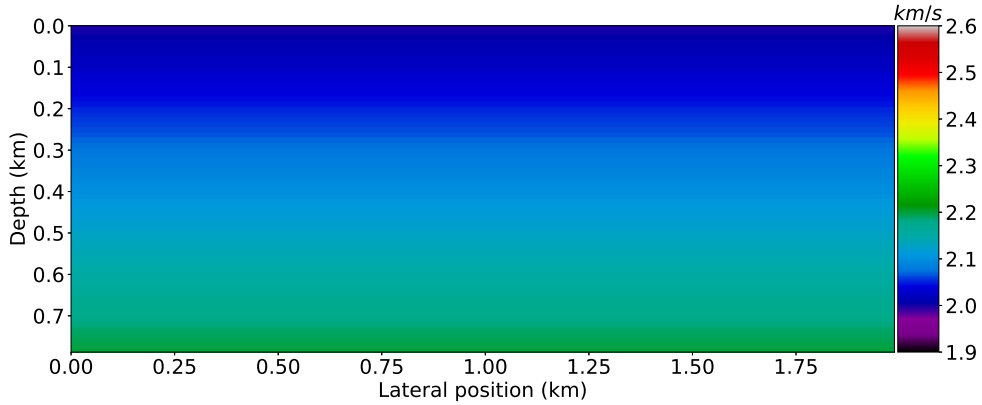
For the regularized tests, we calculate $\beta^{\pm}$ at the beginning of each frequency stage with equation 5.25. We set initially $\beta = 0.01$ and increase it at each new frequency

stage by a factor of 1.38 such that at the final stage we have $\beta \approx 0.05$. We anticipate that regularization I did not perform very well with this approach. Thus, for this regularization we also included a test with $\beta^{\pm}$ calculated at each iteration with equation 5.25.

Moreover, we neglect updates near the sources and receivers. Table 5.1 summarizes the inversion parameters.



(a)



(b)

Figure 5.2: Lens model: (a) Exact; (b) Initial.

5.3.1 Lens model

Figure 5.2(a) shows the exact velocity model. The initial velocity model is composed of a vertical gradient, Figure 5.2(b), and the initial image is composed of zeros.

Estimated model parameters

The velocity profiles in the middle of the model indicate that, at this position, regularization II and the result without regularization are quite similar, Figure 5.3(a). From the depth position of 0.5 km until 0.8 km, regularization II is slightly better than the result without regularization. We also observe that regularization provided a better shape near the top and bottom of the lens anomaly.

Table 5.1: Lens model inversion parameters.

Parameter	**Value**
β in eq. 5.25	$0.01 - 0.05$
min. frequency (Hz)	first sample after 0
max. frequency (Hz)	40
First-stage maximum frequency (Hz)	15
Number of stages	6
Stage stopping criterion (*tol* in Fig 5.1) (%)	$1 - 1.5$
$\epsilon^{\pm}$	$\max\left(0.02 \times \lvert \boldsymbol{H}^{\pm} \rvert\right)$

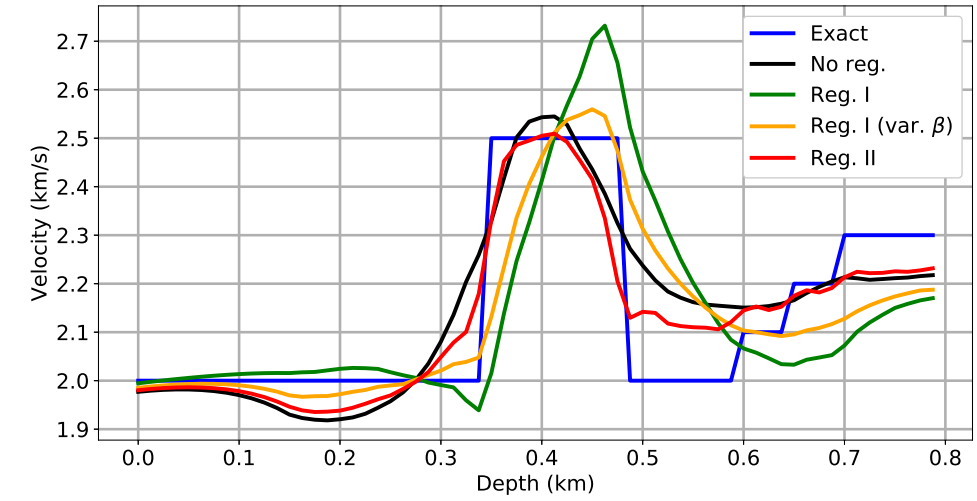
Regularization I marked very well the beginning of the velocity anomaly near the depth position of 0.3 km, but mispositioned the peak of the lens anomaly and exceeded its correct value, green curve in Figure 5.3(a). Regularization I combined with the calculation of the scaling factors $\beta^{\pm}$ at each iteration had a performance between the same regularization I with this factor fixed during each stage and the result from regularization II, orange curve in Figure 5.3(a).

The image profiles in the middle of the model are very similar, Figure 5.3(b). In general, the different tests provided similar results for the image. Thus, in Figure 5.4, we only exhibit the exact image, the result without regularization and the result from regularization II. We observe that the methodology recovered high-resolution images.

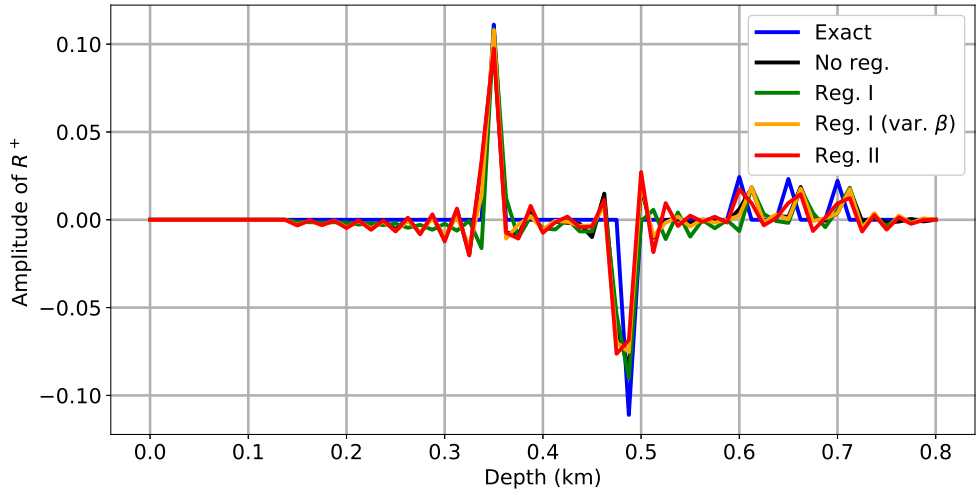
Figure 5.5(a) shows the estimated velocity model without regularization. The Lens shape is not completely recovered, but it is a good result considering that we started from the vertical velocity gradient in Figure 5.2(b). Besides that, the fine layering at the bottom is not well resolved. All the results including regularization improved the resolution of the estimated velocity model, Figure 5.5. Although, in the profiles of Figure 5.3(a) the improvements seems to be small, the regularizations had a considerable impact in the velocity estimation, Figure 5.5.

For the scaling factors $\beta^{\pm}$ fixed at each stage, regularization I improved the structural information in the velocity model, but as the profile of Figure 5.3(a) indicates, it exceeded the value of the velocity inside the lens, Figure 5.5(b). The approach with $\beta^{\pm}$ calculated at each iteration for regularization I improved the estimation of the velocity inside the lens, Figure 5.5(c), but at the cost of less resolution in the lens shape compared to the result of Figure 5.5(b).

Regularization II had the best performance in the recovery of the lens shape, Figure 5.5(d). It also slightly improved the lateral continuity of the velocity at the bottom of the model.



(a)



(b)

Figure 5.3: Profiles in the middle of the Lens model: (a) Exact and estimated velocities; (b) Exact and estimated images.

Data and model residuals

The data residuals decreased over the iterations in the same frequency stage, Figure 5.6(a). The sudden increase at some positions is related to the beginning of a new frequency stage in which additional frequency content is included in the inversion process.

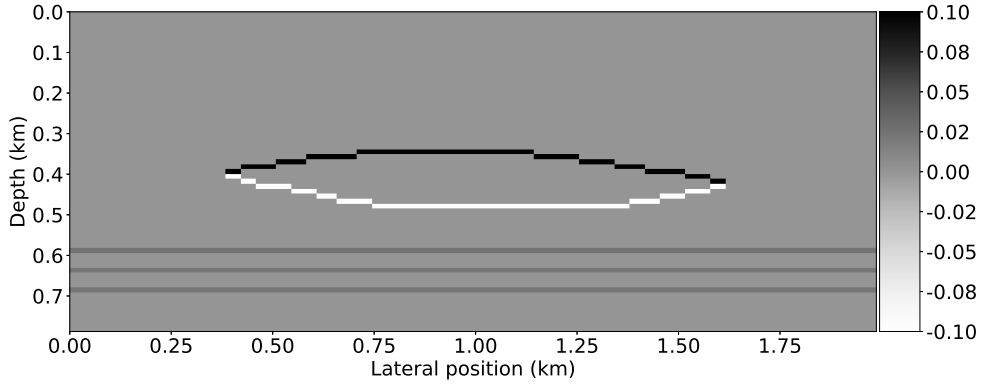
The model residuals indicate that regularization II approached the exact model faster than the other methodologies, Figure 5.6(b). Regularization I with $\beta^{\pm}$ fixed in each stage reduced the model residuals approximately until iteration number 40, then it started to diverge. Regularization I with $\beta^{\pm}$ calculated at each iteration solved this divergence problem, orange curve in Figure 5.6(b).

5.4 Summary

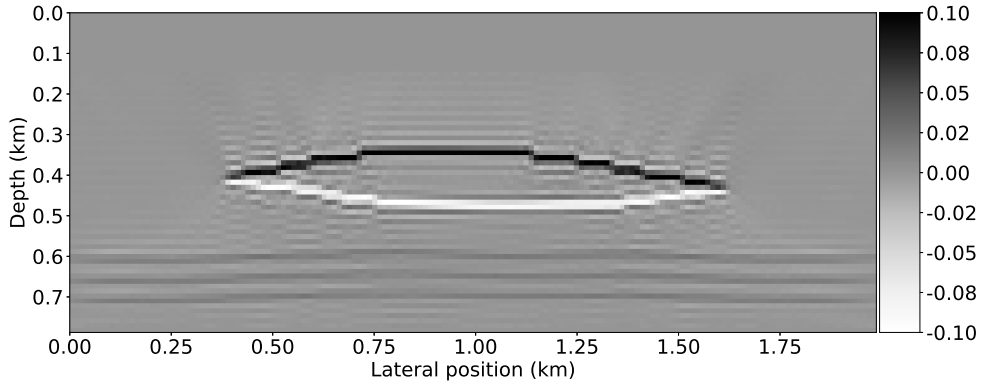
We have proposed two regularization functions that aim at taking advantage of the high spatial-frequency information of the image, estimated by the migration part of JMI, in the calculation of the velocity updates, i.e., inversion part of JMI. Moreover, we have proposed a new parameterization of the inversion part that is based on a normalization of the slowness logarithm. During the preparation of the presented numerical tests, the proposed parameterization seemed to be an important technical detail that contributed to the effectiveness of the regularizations.

The numerical experiments with the Lens model indicated that the methodology proposed is promising. Additionally, regularization II besides improving structural information in the velocity model, it also improved the overall problem convergence. Regularization I with the scaling factors $\beta^{\pm}$ fixed at each stage improved the structural information in the estimated velocity model. But overestimated the value of the velocity inside the lens. Thus, we modified the calculation of $\beta^{\pm}$ and updated them at each iteration. This strategy improved the accuracy of the velocity model estimated by regularization I.

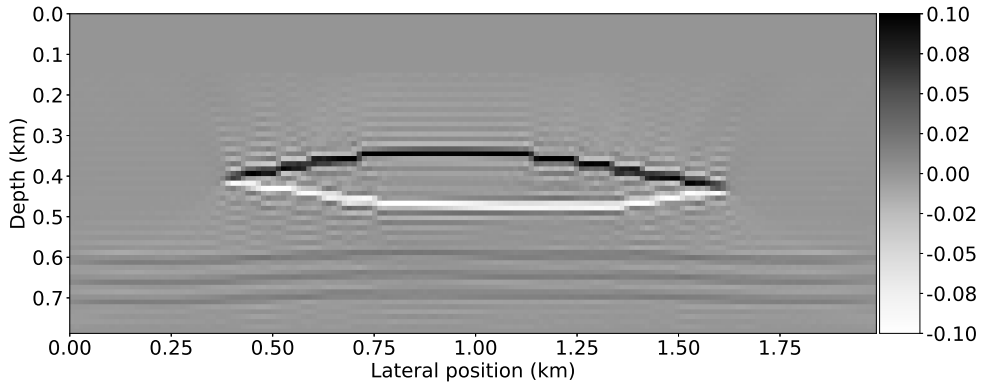
Although the results are satisfactory, we recognize that tests with models of higher complexity must be performed, e.g., the Marmousi2.



(a)

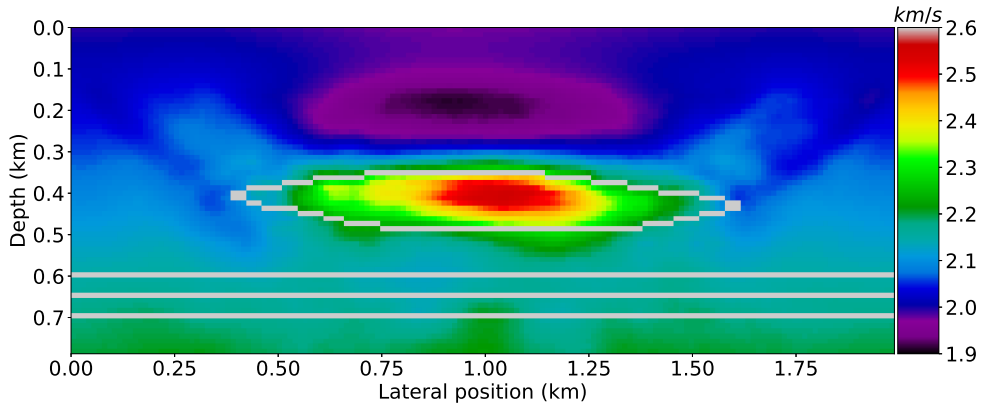


(b)

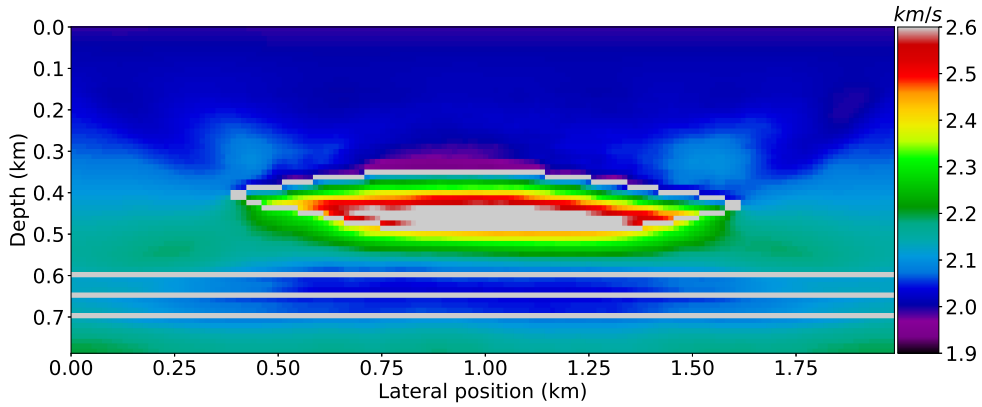


(c)

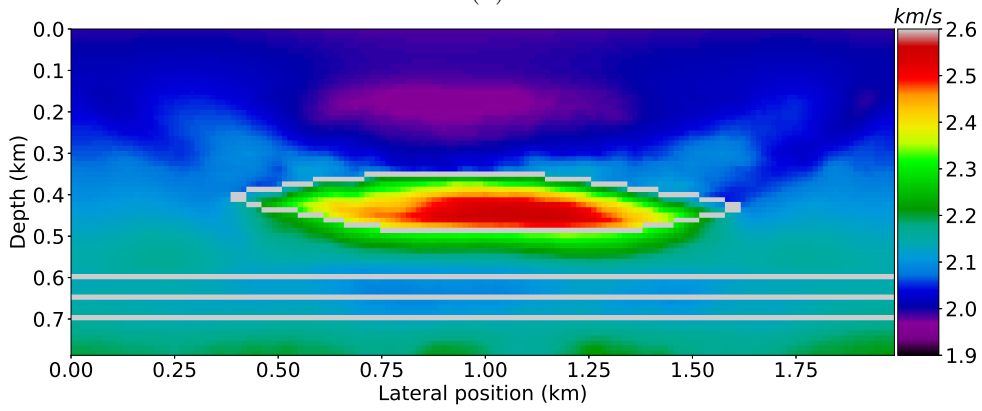
Figure 5.4: Images of the Lens model: (a) Exact; (b) Without regularization; (c) Regularization II.



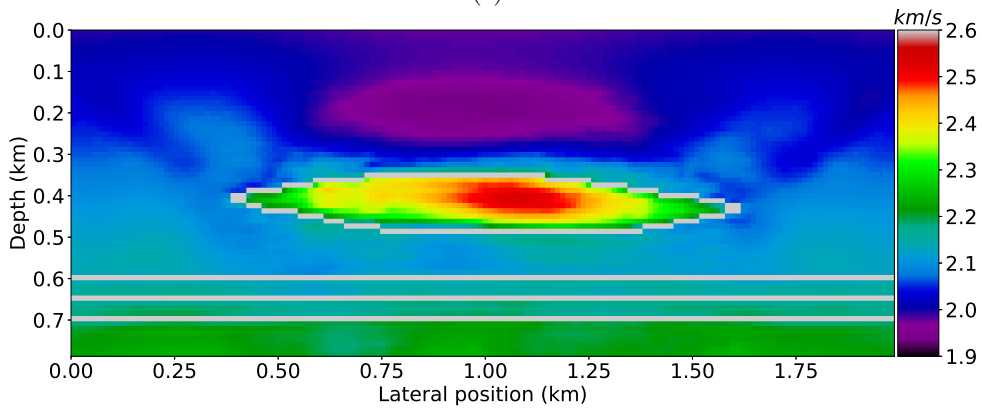
(a)



(b)

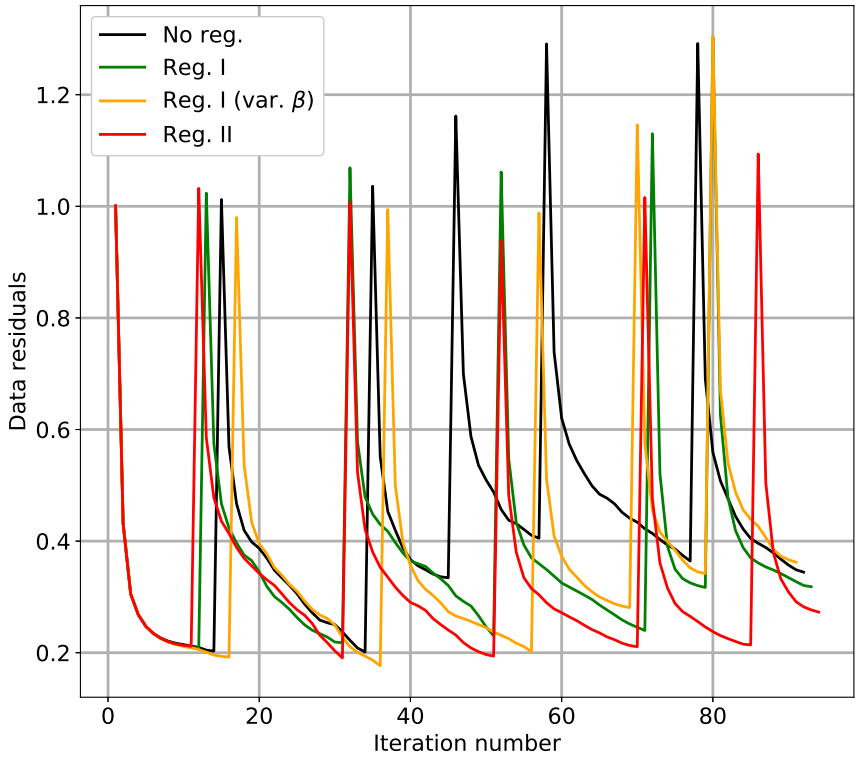


(c)

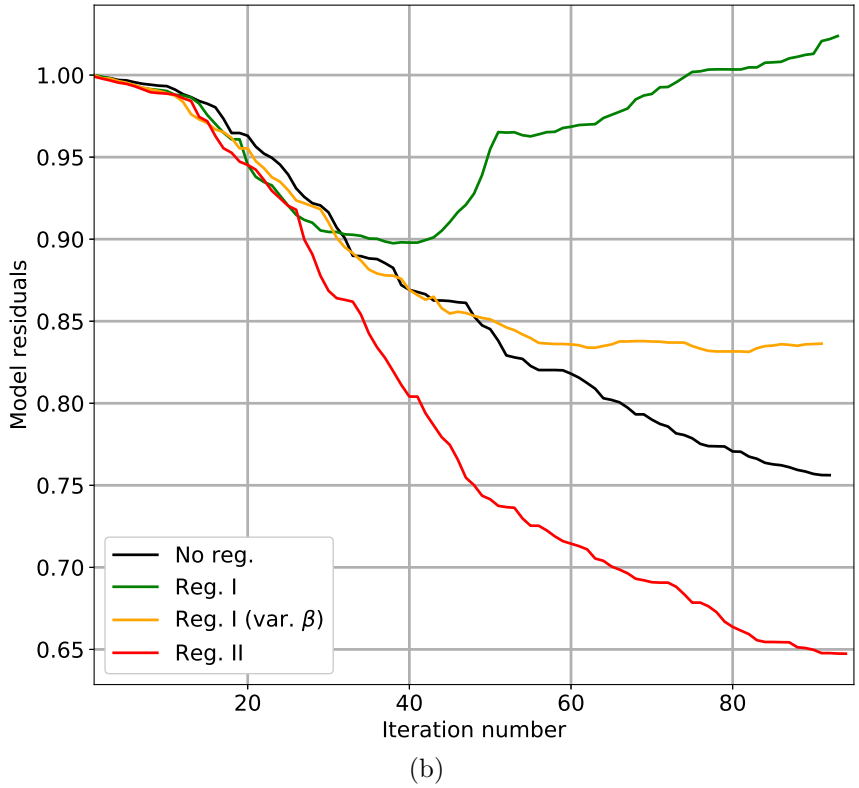


(d)

Figure 5.5: Estimated velocity for the Lens model. The white curves mark the boundaries of the layers. (a) Without regularization; (b) Regularization I; (b) Regularization I with scaling factors $\beta^{\pm}$ varying in each iteration; (d) Regularization II.



(a)



(b)

Figure 5.6: L^2 norm of the residuals over the iterations: (a) Data residuals; (b) Model residuals.

6 Conclusions

Here, we summarize the developments made in this thesis, elaborates on the conclusions, and provides suggestions for future work.

6.1 Overview

This work was motivated by the joint migration inversion (JMI) methodology. JMI is built upon the separation of the seismic waves into their down- and upgoing components. The forward modeling equations that describes these waves intrinsically separates scattering from propagation effects. Thus, this approach takes advantage of the seismic reflection data separation into its dynamic and kinematic information content and, as a consequence, JMI explores the spatial-scale separation.

We started with the equation of motion and the time derivative of the Hooke's law for acoustic media. Then, we separated the vertical derivatives of the pressure and of the vertical particle-displacement velocity and organized the initial equations as a combination of vectors and one matrix. Next, using a diagonalization procedure and changing the problem variables, we derived the coupled one-way wave equations that describe the down- and the upgoing waves. Actually, the change of variables transforms a vector composed of pressure and vertical particle-displacement velocity into a vector composed of the down- and the upgoing pressure-wavefields. Thus, we defined a decomposition of the two-way wavefields such that the sum of the one-way pressure wavefields provides the total counterpart, i.e., the two-way pressure wavefield. This form of decomposition is commonly referred in the literature as pressure normalization.

The derivation of the integral representations of the one-way wavefields is a key contribution of this thesis. First, we demonstrated that the one-way wavefields satisfy the acoustic wave-equation. This result made possible the direct application of Green's theorem in the same manner applied to two-way wavefields. Moreover, the specification of the one-way Green's functions to act as extrapolation operators provided the modeling equations that are the basis of JMI and was very insightful.

For example, one approach to derive these integral equations, is by defining the

decoupled one-way Green's functions in a model without vertical variations and the corresponding wavefields being able to scatter at the boundary of the closed volume under consideration. Then, propagation or backpropagation can be performed as a sequence of boundary extrapolations. Additionally, the series expansion of the integral representations provided the recursive modeling equations used in this work.

After developing the integral representations of the forward modeling equations in acoustic media, we specialized our implementation to the two-dimensional and constant mass-density case. Moreover, the scattering operators were assumed angle-independent.

In the second part of this thesis, we focused on the estimation of the model parameters by nonlinear least-squares data fitting. From the application of the Lagrangian multipliers to the coupled one-way equations, we obtained a parameterization to the estimation of the reflection operator, i.e., the migration stage of JMI, that consider variations in the forward/adjoint down- and upgoing wavefields. Moreover, variations in the transmission operator were also considered by means of the relation between transmission and reflection operators in acoustic media.

The gradient of the least-squares migration problem provided a new imaging condition. Neglecting the normalization factor, the new imaging condition is composed of the correlation between the difference of the forward down- and upgoing wavefields with the difference between the correspondent adjoint wavefields. Neglecting variations in the downgoing wavefield, the conventional imaging condition should be modified at least to the correlation of the backpropagated observed data and the difference between the down- and upgoing wavefields in order to be consistent with our result. Also neglecting variations in the transmission operator provides the conventional imaging condition.

Still considering that a known velocity model is sufficient for migration, we parameterized the forward modeling equation as a function of the acoustic impedance, which made possible the estimation of the impedance directly from shot gathers. Commonly, first depth migration is performed to estimate the reflection coefficients, then some procedure is applied to convert the image to impedance. In a more straightforward approach, the proposed methodology is capable of migrating the shot gathers directly to the relative acoustic impedance.

Finally, considering the migration and velocity estimation problems, i.e, the complete JMI workflow, we proposed two regularization functions that take advantage of the high spatial-frequency content of the image in the estimation of the velocity model updates. The results demonstrated that the proposed methodology is effective in synthetic data.

6.2 Suggestions for future work

6.2.1 Approximations in the modeling equations

We developed the three-dimensional coupled one-way wave equations for acoustic media with variable mass-density. In the implementation, we made many assumptions, they are: two-dimensional model, constant mass-density and angle-independent scattering.

In our opinion, the first assumption to be removed should be the angle-independency of the scattering operators, because the methodology for model parameters estimation was formulated in the data domain and the objective is to recover true-amplitude images. Additionally, envisioning applications on real data, the removal of this assumption also reduces the restrictions on the input observed data. Appendix C and the work of Hammad and Verschuur (2019) are good starting points to the investigation of the scattering operators.

The assumption of constant mass-density may not be critical in the JMI workflow. But in reservoir characterization studies, this model parameter is very important for discrimination of fluid content, due to the fact that, given a velocity model together with the mass density, the elastic parameters can be estimated (see, e.g., Avseth et al., 2005).

It also important to have in mind that most of the modern marine seismic data are three-dimensional. Hence, an efficient three-dimensional implementation of the forward/inverse modeling equations is desirable, if the objective is to approximate JMI of the routinely applied industrial workflow. Otherwise, for research purposes, the two-dimensional extension for elastic media may be more interesting (see, e.g., Wapenaar and Berkhout, 1989).

6.2.2 Parameters from amplitude variations with angle

After implementing forward modeling equations with scattering operators that are angle-dependent. One possible investigation is the parameterization of the inverse problem as a function of the amplitude variation with angle (AVA) parameters, e.g., intercept and gradient (Shuey, 1985). The methodology developed in Chapter 4 for the estimation of the acoustic impedance can be seen as a first step in this direction. In this manner, instead of the common AVA inversion of migrated data that assumes an unidimensional subsurface and relies on the convolutional model, a less restrict method based on the wave equation could be built such that the required AVA parameters are estimated directly from shot gathers. This investigation, may reduce the gap between waveform inversion methods and the common practice adopted in reservoir characterization studies.

6.2.3 Strategies for the estimation of the model parameters

We presented results for the least-squares solution of the migration part of JMI in Chapter 3. We concluded that the inexact Gauss–Newton method is capable of providing a sufficient result in one nonlinear iteration, if multiple reflections are absent. Thus, the residuals between the observed and modeled data are not formed.

In the complete JMI methodology of Chapter 5, we alternated between migration and inversion using a gradient-based method. In this approach, least-squares migration is not performed accurately in one nonlinear iteration due to the poor approximation of the misfit-function Hessian. Consequently, on the subsequent iterations, the data residuals are formed. Thus, this methodology may suffer from spurious image updates for observed data that does not fulfill the assumption of angle-independent scattering, as assumed in the problem formulation. The velocity updates may face the same issues.

One workaround for the complete JMI methodology is to discard the current image at the beginning of each nonlinear iteration and then to solve the migration problem with the inexact Gauss–Newton method as in Chapter 3. The velocity estimation part is a nonlinear problem, thus, the simplest approach is to keep the iterative updates. But maybe instead of using the L^2 norm of the data residuals, misfit functions that works mainly with phase information are good candidates to the mitigation of amplitude innacuracies in the forward modeling equations (see, e.g., Davydenko and Verschuur, 2019).

Now, consider that in the implementation of the JMI methodology discussed here, we have two classes of parameters, i.e., the image and the velocity. It may be also interesting to formulate the inverse problem completely in terms of the Gauss–Newton method or even the Newton method (see, e.g., Métivier et al., 2017). The simultaneous estimation of image and velocity updates using the Gauss–Newton method can improve the resolution of the estimated velocity model (see, e.g., Soubaras and Gratacos, 2017; Assis et al., 2019a).

A Fourier transform

The Fourier transform is fundamental in seismic signal analysis and in the development of the theoretical concepts. We will adopt the following convention for the forward and inverse transforms in the time variable

$$F(\mathbf{x},\omega) = \int_{\mathbb{R}} f(\mathbf{x},t)e^{-i\omega t}dt\,, \tag{A.1}$$

$$f(\mathbf{x},t) = \frac{1}{2\pi}\int_{\mathbb{R}} F(\mathbf{x},\omega)e^{i\omega t}d\omega\,, \tag{A.2}$$

where $\mathbf{x}$ is the position vector, ω is the angular frequency, t is time and i is the imaginary unity. In a three dimensional problem $\mathbf{x} = (x, y, z)$.

We define the Fourier transform in the lateral coordinates x and y as follows

$$\widetilde{F}(k_x,k_y,z,\omega) = \frac{1}{4\pi^2}\int_{\mathbb{R}^2} F(\mathbf{x},\omega)e^{i(k_x x+k_y y)}dxdy\,, \tag{A.3}$$

$$F(\mathbf{x},\omega) = \frac{1}{4\pi^2}\int_{\mathbb{R}^2} \widetilde{F}(k_x,k_y,z,\omega)e^{-i(k_x x+k_y y)}dk_x dk_y\,. \tag{A.4}$$

We also define the short notation for the spatial Fourier transform as

$$\hat{\mathcal{F}}\left\{F\right\} = \widetilde{F}\,, \tag{A.5}$$

$$\hat{\mathcal{F}}^{-1}\left\{\widetilde{F}\right\} = F\,. \tag{A.6}$$

B Directional decomposition operators

In this appendix, we investigate the diagonalization of the two-way matrix. In the sequence, we demonstrate the relation between the one-way and two-way wavefields. Next, we discuss the action of the operators on a wavefield. Following the notation convention adopted in this thesis, calligraphic font-style indicates operators and hat over the operators indicates action over the lateral coordinates. Then, we approximate the two-way wave operator from the composition of the decoupled one-way operators and define the corresponding one-way Green's functions that added provide the two-way counterpart.

B.1 Diagonalization of the two-way matrix operator

Now we develop the diagonalization of the two-way matrix operator $\hat{\boldsymbol{\mathcal{A}}}$ given by

$$\hat{\boldsymbol{\mathcal{A}}} = \begin{bmatrix} 0 & -i\omega\rho \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & 0 \end{bmatrix}. \tag{B.1}$$

We want a decomposition with the form

$$\hat{\boldsymbol{\mathcal{A}}} = \hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}}\hat{\boldsymbol{\mathcal{L}}}^{-1}. \tag{B.2}$$

Given that $\hat{\boldsymbol{\mathcal{A}}}$ is 2 by 2, the columns of $\hat{\boldsymbol{\mathcal{L}}}$ will be built with two independent vectors and $\hat{\boldsymbol{\mathcal{D}}}$ with the corresponding eigenvalues in the diagonals. We expect all these quantities to be operators. Besides that, due to the fact that the trace of $\hat{\boldsymbol{\mathcal{A}}}$ is zero, we have some degree of freedom to define the eigenvectors. Thus, we choose the elements of $\hat{\boldsymbol{\mathcal{L}}}$ as

$$\hat{\boldsymbol{\mathcal{L}}} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{L}}_1 & -\hat{\mathcal{L}}_1 \end{bmatrix}, \tag{B.3}$$

where $\mathcal{I}$ is the identity operator and we also expect $\hat{\mathcal{L}}_1$ to be an operator. Still, as a

consequence of $\hat{\boldsymbol{\mathcal{A}}}$ null trace, the matrix of eigenvalues is

$$\hat{\boldsymbol{\mathcal{D}}} = \begin{bmatrix} -\hat{\lambda} & 0 \\ 0 & \hat{\lambda} \end{bmatrix} . \tag{B.4}$$

We want to define $\hat{\mathcal{L}}_1$ and $\hat{\lambda}$. Rearranging equation B.2, we start the problem in the form

$$\hat{\boldsymbol{\mathcal{A}}}\hat{\boldsymbol{\mathcal{L}}} = \hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}} . \tag{B.5}$$

Explicitly we have

$$\begin{bmatrix} 0 & -i\omega\rho \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & 0 \end{bmatrix} \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{L}}_1 & -\hat{\mathcal{L}}_1 \end{bmatrix} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{L}}_1 & -\hat{\mathcal{L}}_1 \end{bmatrix} \begin{bmatrix} -\hat{\lambda} & 0 \\ 0 & \hat{\lambda} \end{bmatrix} . \tag{B.6}$$

We perform the products and obtain

$$\begin{bmatrix} -i\omega\rho\hat{\mathcal{L}}_1 & i\omega\rho\hat{\mathcal{L}}_1 \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 \end{bmatrix} = \begin{bmatrix} -\hat{\lambda} & \hat{\lambda} \\ -\hat{\mathcal{L}}_1\hat{\lambda} & -\hat{\mathcal{L}}_1\hat{\lambda} \end{bmatrix} . \tag{B.7}$$

The left columns provide

$$i\omega\rho\hat{\mathcal{L}}_1 = \hat{\lambda} , \tag{B.8}$$

$$\frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 = -\hat{\mathcal{L}}_1\hat{\lambda} . \tag{B.9}$$

The substitution of the first equation in the second one provides

$$\frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 = -\frac{1}{i\omega\rho}\hat{\lambda}^2 . \tag{B.10}$$

Considering that the transversal Helmholtz operator $\hat{\mathcal{H}}_2$ can be decomposed into two equal operators, we obtain

$$\boxed{\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1 .} \tag{B.11}$$

Thus, from equation B.10, we define the positive and negative eigenvalues as

$$\boxed{\pm\hat{\lambda} = \pm i\hat{\mathcal{H}}_1} . \tag{B.12}$$

The substitution of the positive eigenvalue in equation B.8 provides

$$\boxed{\hat{\mathcal{L}}_1 = \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 .} \tag{B.13}$$

Gathering the results, the matrices that diagonalize $\hat{\boldsymbol{\mathcal{A}}}$ as proposed in equation B.2 are

$$\hat{\boldsymbol{\mathcal{D}}} = \begin{bmatrix} -i\hat{\mathcal{H}}_1 & 0 \\ 0 & i\hat{\mathcal{H}}_1 \end{bmatrix}, \tag{B.14}$$

$$\hat{\boldsymbol{\mathcal{L}}} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 & -\frac{1}{\omega\rho}\hat{\mathcal{H}}_1 \end{bmatrix}. \tag{B.15}$$

The inverse of $\hat{\boldsymbol{\mathcal{L}}}$ is

$$\hat{\boldsymbol{\mathcal{L}}}^{-1} = \frac{1}{2}\begin{bmatrix} \mathcal{I} & \hat{\mathcal{H}}_1^{-1}\omega\rho \\ \mathcal{I} & -\hat{\mathcal{H}}_1^{-1}\omega\rho \end{bmatrix}. \tag{B.16}$$

Now we verify the diagonalization. We start with the product

$$\hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}} = \begin{bmatrix} -i\hat{\mathcal{H}}_1 & i\hat{\mathcal{H}}_1 \\ -\frac{i}{\omega\rho}\hat{\mathcal{H}}_2 & -\frac{i}{\omega\rho}\hat{\mathcal{H}}_2 \end{bmatrix}, \tag{B.17}$$

where we considered $\hat{\mathcal{H}}_1\hat{\mathcal{H}}_1 = \hat{\mathcal{H}}_2$. Applying this result to the matrix of eigenvectors inverse, we obtain

$$\hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}}\hat{\boldsymbol{\mathcal{L}}}^{-1} = \begin{bmatrix} 0 & -i\omega\rho \\ \frac{1}{i\omega\rho}\hat{\mathcal{H}}_2 & 0 \end{bmatrix}. \tag{B.18}$$

This result confirms that the expressions for $\hat{\boldsymbol{\mathcal{L}}}$, $\hat{\boldsymbol{\mathcal{L}}}^{-1}$ and $\hat{\boldsymbol{\mathcal{D}}}$ realize the decomposition $\hat{\boldsymbol{\mathcal{A}}} = \hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}}\hat{\boldsymbol{\mathcal{L}}}^{-1}$.

B.2 Decomposition interpretation

Now that we developed the decomposition $\hat{\boldsymbol{\mathcal{L}}}\hat{\boldsymbol{\mathcal{D}}}\hat{\boldsymbol{\mathcal{L}}}^{-1}$, it is possible to attach some interpretation to the elements of each matrix of operators.

The elements $-i\hat{\mathcal{H}}_1$ and $i\hat{\mathcal{H}}_1$ of $\hat{\boldsymbol{\mathcal{D}}}$ are associated with downgoing and upgoing propagation, respectively. See Section 2.2.2 for a detailed discussion of this interpretation. In this manner, we rename $\hat{\boldsymbol{\mathcal{D}}}$, equation B.14, as

$$\hat{\boldsymbol{\mathcal{D}}} = i\hat{\boldsymbol{\mathcal{H}}} = i\begin{bmatrix} -\hat{\mathcal{H}}_1 & 0 \\ 0 & \hat{\mathcal{H}}_1 \end{bmatrix}, \tag{B.19}$$

where we expect that this new nomenclature improves readability.

Besides the identity operator, $\hat{\boldsymbol{\mathcal{L}}}$ and $\hat{\boldsymbol{\mathcal{L}}}^{-1}$ are basically composed of the element $\hat{\mathcal{H}}_1^{-1}\omega\rho$ and its inverse. On a laterally invariant model and in the domain (k_x, k_y, z, ω),

we have

$$\widetilde{\hat{\mathcal{H}}_1} = k_z = \frac{\omega}{c}\sqrt{1 - \frac{c^2}{\omega^2}\left(k_x^2 + k_y^2\right)}\,, \tag{B.20}$$

where we have defined the sign of k_z to equal that of ω.

Now, consider a wave in a homogeneous medium that propagates with wavefront perpendicular to the vertical axis and increasing depth, i.e.,

$$k_x = k_y = 0\,, \tag{B.21}$$

we have

$$k_z = \omega/c\,, \tag{B.22}$$

which provides

$$\widetilde{\hat{\mathcal{H}}_1}^{-1}\,\omega\rho = \rho c\,. \tag{B.23}$$

This is the acoustic impedance. Hence, we define the generalized acoustic impedance as

$$\hat{\mathcal{Z}} = \hat{\mathcal{H}}_1^{-1}\omega\rho\,. \tag{B.24}$$

And its inverse

$$\hat{\mathcal{Z}}^{-1} = \frac{1}{\omega\rho}\hat{\mathcal{H}}_1\,. \tag{B.25}$$

Thus, we redefine $\hat{\boldsymbol{\mathcal{L}}}$ and $\hat{\boldsymbol{\mathcal{L}}}^{-1}$, equations B.15 and B.16, in terms of the generalized impedance as follows

$$\hat{\boldsymbol{\mathcal{L}}} = \hat{\boldsymbol{\mathcal{Z}}}^{-1} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{Z}}^{-1} & -\hat{\mathcal{Z}}^{-1} \end{bmatrix}\,, \tag{B.26}$$

$$\hat{\boldsymbol{\mathcal{L}}}^{-1} = \hat{\boldsymbol{\mathcal{Z}}} = \frac{1}{2}\begin{bmatrix} \mathcal{I} & \hat{\mathcal{Z}} \\ \mathcal{I} & -\hat{\mathcal{Z}} \end{bmatrix}\,. \tag{B.27}$$

As as consequence, the diagonalization of the two-way matrix operator $\hat{\boldsymbol{\mathcal{A}}}$ is recast in the form

$$\boxed{\hat{\boldsymbol{\mathcal{A}}} = i\hat{\boldsymbol{\mathcal{Z}}}^{-1}\hat{\boldsymbol{\mathcal{H}}}\hat{\boldsymbol{\mathcal{Z}}}\,.} \tag{B.28}$$

B.3 Compact integral representation

Most of the operators discussed in this appendix perform their action via integration. Let us start our investigation with the transversal Helmholtz operator

$$\hat{\mathcal{H}}_2 = \rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}\right) + \frac{\omega^2}{c^2}\,, \tag{B.29}$$

where the operator $\hat{\nabla}$ is given by

$$\hat{\nabla} = \frac{\partial}{\partial x}\mathbf{i}_x + \frac{\partial}{\partial y}\mathbf{i}_y \,. \tag{B.30}$$

The transversal Helmholtz operator can also be defined upon twofold application of the square-root operator

$$\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1 \,. \tag{B.31}$$

In a model without lateral variations, i.e., at most $c = c(z)$ and $\rho = \rho(z)$, the explicit representation of the square-root operator in the domain (k_x, k_y, z, ω) is

$$\widetilde{\hat{\mathcal{H}}_1} = k_z = \frac{\omega}{c(z)}\sqrt{1 - \frac{c(z)^2}{\omega^2}\left(k_x^2 + k_y^2\right)} \,. \tag{B.32}$$

For example, the action of $-k_z$ on a wavefield P^+ provides

$$k_z\hat{\mathcal{F}}\left\{P^+\right\} = -\frac{\omega}{c(z)}\sqrt{1 - \frac{c(z)^2}{\omega^2}\left(k_x^2 + k_y^2\right)}\hat{\mathcal{F}}\left\{P^+\right\} \,, \tag{B.33}$$

where $\hat{\mathcal{F}}\{.\}$ is the Fourier transform over the lateral coordinates x and y. We apply the inverse Fourier transform in the coordinate (k_x, k_y) and obtain

$$\hat{\mathcal{F}}^{-1}\left\{k_z\hat{\mathcal{F}}\left\{P^+\right\}\right\} = -\int_{\mathbb{R}^2}\left\{\frac{1}{4\pi^2}\int_{\mathbb{R}^2} k_z e^{-ik_x x' - ik_y y'}dk_x dk_y\right\}P^+ dx'dy' \,. \tag{B.34}$$

Note that we have a convolution over x' and y'. Moreover, we recognize inside the right-hand side brackets, the generalized vertical-wavenumber, i.e., the vertical wavenumber in the space domain, therefore we can write

$$\left(\hat{\mathcal{H}}_1 P^+\right)(x, y, z, \omega) = -\int_{\mathbb{R}^2}\hat{\mathcal{H}}_1(x, y, z, \omega; x', y')P^+(x', y', z, \omega)dx'dy' \,. \tag{B.35}$$

We Multiply by the imaginary unit i, recognize the partial differential equation that describes the downgoing wave of Section 2.2.2, without source terms, and obtain

$$\frac{\partial P^+}{\partial z}(x, y, z, \omega) = -\int_{\mathbb{R}^2} i\hat{\mathcal{H}}_1(x, y, z, \omega; x', y')P^+(x', y', z, \omega)dx'dy' \,. \tag{B.36}$$

Now we introduce the compact operator notation

$$\boxed{\frac{\partial P^+}{\partial z}(\mathbf{x}, \omega) = -i\hat{\mathcal{H}}_1(\mathbf{x}, \omega; \hat{\mathbf{x}}')P^+(\hat{\mathbf{x}}', z, \omega) \,,} \tag{B.37}$$

where $\hat{\mathbf{x}} = (x, y)$. This notation with repeated indexes is pretty close to the Einstein

summation convention, here representing integration over the coordinates x' and y'.

The same notation is valid for the application of the square-root vertical derivative

$$\left(\frac{\partial\hat{\mathcal{H}}_1}{\partial z}P^+\right)(\mathbf{x},\omega) = \frac{\partial\hat{\mathcal{H}}_1}{\partial z}(\mathbf{x},\omega;\hat{\mathbf{x}}',z)P^+(\hat{\mathbf{x}}',z,\omega)\,. \tag{B.38}$$

This representation is useful, because the vertical derivative of the square-root operator appears in the continuous scattering operators (for more details see Appendix C).

Therefore, disregarding lateral variations, we developed the intuition of the square-root operator $\hat{\mathcal{H}}_1$ application on a wavefield. For the general inhomogeneous case the forms in equations B.35 and B.36 are still valid.

B.4 Relationship between one-way and two-way wavefields

In this section, we verify that the matrix of operators $\hat{\boldsymbol{\mathcal{Z}}}^{-1}$, equation B.26, applied to the downgoing and upgoing wavefields $\mathbf{P}$, provides the two-way quantities in $\mathbf{U}$, i.e., total pressure and vertical particle velocity. Consider the equality in compact notation

$$\hat{\boldsymbol{\mathcal{Z}}}^{-1}\mathbf{P} = \mathbf{U}\,. \tag{B.39}$$

In which the wavefield vectors are given by

$$\mathbf{P} = \begin{bmatrix} P^+ \\ P^- \end{bmatrix}, \tag{B.40}$$

$$\mathbf{U} = \begin{bmatrix} P \\ V_z \end{bmatrix}. \tag{B.41}$$

Explicitly, we have

$$\begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 & -\frac{1}{\omega\rho}\hat{\mathcal{H}}_1 \end{bmatrix}\begin{bmatrix} P^+ \\ P^- \end{bmatrix} = \begin{bmatrix} P \\ V_z \end{bmatrix}, \tag{B.42}$$

where we recognize that the sum $P^+ + P^-$ provides the total pressure. Therefore, we conclude that P^+ and P^- are pressure wavefields. Proceeding, from the composition in

the left-hand side, we have

$$\begin{aligned}\hat{\boldsymbol{\mathcal{Z}}}^{-1}\mathbf{P} &= \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 & -\frac{1}{\omega\rho}\hat{\mathcal{H}}_1 \end{bmatrix} \begin{bmatrix} P^+ \\ P^- \end{bmatrix}, \\ &= \begin{bmatrix} P^+ + P^- \\ \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 P^+ - \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 P^- \end{bmatrix}. \end{aligned} \tag{B.43}$$

The forward modeling equations without the physical sources are

$$\frac{\partial P^+}{\partial z} = -i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{R}}_c^- P^- + \hat{\mathcal{T}}_c^+ P^+ , \tag{B.44}$$

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \hat{\mathcal{R}}_c^+ P^+ - \hat{\mathcal{T}}_c^- P^- . \tag{B.45}$$

We rearrange the terms and obtain

$$\hat{\mathcal{H}}_1 P^+ = i\frac{\partial P^+}{\partial z} - i\hat{\mathcal{R}}_c^- P^- - i\hat{\mathcal{T}}_c^+ P^+ , \tag{B.46}$$

$$-\hat{\mathcal{H}}_1 P^- = i\frac{\partial P^-}{\partial z} + i\hat{\mathcal{R}}_c^+ P^+ + i\hat{\mathcal{T}}_c^- P^- . \tag{B.47}$$

The substitution of these equations in the last line of B.43 provides

$$i\frac{1}{\omega\rho}\left[\frac{\partial\left(P^+ + P^-\right)}{\partial z} + \left(\hat{\mathcal{R}}_c^+ - \hat{\mathcal{T}}_c^+\right)P^+ + \left(\hat{\mathcal{T}}_c^- - \hat{\mathcal{R}}_c^-\right)P^-\right] . \tag{B.48}$$

In acoustic media, we know that $\hat{\mathcal{R}}_c^{\pm} = \hat{\mathcal{T}}_c^{\pm}$ (for more details see Appendix C, Sections C.2.1 and C.2.2). Thus, the last equation simplifies to

$$i\frac{1}{\omega\rho}\frac{\partial}{\partial z}\left(P^+ + P^-\right) . \tag{B.49}$$

Considering from the first line of the composition definition, equation B.42, that the relation $P^+ + P^- = P$ holds. We multiply and divide, equation B.49, by the imaginary unity i and obtain

$$-\frac{1}{i\omega\rho}\frac{\partial P}{\partial z} . \tag{B.50}$$

This is the definition of V_z, as shown in equation 2.6. Therefore, we confirmed that the operator $\hat{\boldsymbol{\mathcal{Z}}}^{-1}$ represents the composition of the one-way wavefields into pressure and vertical particle displacement velocity. Consequently, $\hat{\boldsymbol{\mathcal{Z}}}$ decomposes the two-way quantities $\mathbf{U}$ into the one-way vector $\mathbf{P}$.

B.5 Relationship between two-way and one-way wave operators

Let us define the decoupled one-way wave operators as

$$\mathcal{L}^{+} = \frac{\partial}{\partial z} + i\hat{\mathcal{H}}_1 - \hat{\mathcal{T}}_c^{+}\,, \tag{B.51}$$

$$\mathcal{L}^{-} = \frac{\partial}{\partial z} - i\hat{\mathcal{H}}_1 + \hat{\mathcal{T}}_c^{-}\,, \tag{B.52}$$

where $\hat{\mathcal{T}}_c^{+}$ and $\hat{\mathcal{T}}_c^{-}$ are the transmissivities, see Appendix C, Section C.2.2.

Consider the coupled downgoing wavefield P^{+}

$$\mathcal{L}^{+}P^{+} = \hat{\mathcal{R}}_c^{-}P^{-} + S^{+}\,. \tag{B.53}$$

The composition of the action of $\mathcal{L}^{-}$ and $\mathcal{L}^{+}$ on P^{+} yields

$$\begin{aligned}\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \frac{\partial^2 P^{+}}{\partial z^2} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} + i\hat{\mathcal{H}}_1\frac{\partial P^{+}}{\partial z} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} - \hat{\mathcal{T}}_c^{+}\frac{\partial P^{+}}{\partial z}\\ - i\hat{\mathcal{H}}_1\frac{\partial P^{+}}{\partial z} + \hat{\mathcal{H}}_2 P^{+} + i\hat{\mathcal{H}}_1\hat{\mathcal{T}}_c^{+}P^{+} + \hat{\mathcal{T}}_c^{-}\frac{\partial P^{+}}{\partial z}\\ + \hat{\mathcal{T}}_c^{-}i\hat{\mathcal{H}}_1 P^{+} - \hat{\mathcal{T}}_c^{-}\hat{\mathcal{T}}_c^{+}P^{+}\,.\end{aligned} \tag{B.54}$$

From equations C.40 and C.42, we have that

$$\hat{\mathcal{T}}_c^{-} = -\hat{\mathcal{T}}_c^{+}\,. \tag{B.55}$$

The substitution in equation B.54 together with the identity $\hat{\mathcal{H}}_1\hat{\mathcal{T}}_c^{+} = \hat{\mathcal{T}}_c^{+}\hat{\mathcal{H}}_1$ and the simplification of the common terms provides

$$\begin{aligned}\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \frac{\partial^2 P^{+}}{\partial z^2} + \hat{\mathcal{H}}_2 P^{+} - 2\hat{\mathcal{T}}_c^{-}\frac{\partial P^{+}}{\partial z} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+}\\ + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,.\end{aligned} \tag{B.56}$$

The transmissivity operator is given by

$$\hat{\mathcal{T}}_c^{+} = -\frac{1}{2}\rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\right) - \frac{1}{2}\hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\,. \tag{B.57}$$

The substitution of this equation in the third term of equation B.56 provides

$$\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \frac{\partial^2 P^{+}}{\partial z^2} + \hat{\mathcal{H}}_2 P^{+} + \rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\right)\frac{\partial P^{+}}{\partial z} + \hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\frac{\partial P^{+}}{\partial z} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,. \tag{B.58}$$

In which we recognize the product rule involving the first and third terms, grouping them, we obtain

$$\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \hat{\mathcal{H}}_2 P^{+} + \rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial P^{+}}{\partial z}\right) + \hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\frac{\partial P^{+}}{\partial z} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,. \tag{B.59}$$

The transversal Helmholtz operator is given by

$$\hat{\mathcal{H}}_2 = \rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}\right) + \frac{\omega^2}{c^2}\,. \tag{B.60}$$

Its substitution in equation B.59 provides

$$\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}P^{+}\right) + \frac{\omega^2}{c^2}P^{+} + \rho\frac{\partial}{\partial z}\left(\frac{1}{\rho}\frac{\partial P^{+}}{\partial z}\right) + \hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\frac{\partial P^{+}}{\partial z} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,. \tag{B.61}$$

We group the first and third terms and obtain

$$\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \rho\nabla\cdot\left(\frac{1}{\rho}\nabla P^{+}\right) + \frac{\omega^2}{c^2}P^{+} + \hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\frac{\partial P^{+}}{\partial z} - i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,. \tag{B.62}$$

The first and second terms form the definition of the two-way wave operator $\mathcal{L}$, the substitution this definition provides

$$\mathcal{L}^{-}\mathcal{L}^{+}P^{+} = \mathcal{L}P^{+} + \hat{\mathcal{H}}_1^{-1}\frac{\partial \hat{\mathcal{H}}_1}{\partial z}\frac{\partial P^{+}}{\partial z} + i\frac{\partial \hat{\mathcal{H}}_1}{\partial z}P^{+} - \frac{\partial \hat{\mathcal{T}}_c^{+}}{\partial z}P^{+} + \hat{\mathcal{T}}_c^{+}\hat{\mathcal{T}}_c^{+}P^{+}\,. \tag{B.63}$$

From equation B.53, we have that

$$\frac{\partial P^{+}}{\partial z} = -i\hat{\mathcal{H}}_1 P^{+} + \hat{\mathcal{R}}_c^{-}P^{-} + \hat{\mathcal{T}}_c^{+}P^{+} + S^{+}\,. \tag{B.64}$$

We substitute in the second term of equation B.63 and obtain

$$\mathcal{L}^-\mathcal{L}^+P^+ = \mathcal{L}P^+ + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\left(-i\hat{\mathcal{H}}_1P^+ + \hat{\mathcal{R}}_c^-P^- + \hat{\mathcal{T}}_c^+P^+ + S^+\right) + i\frac{\partial\hat{\mathcal{H}}_1}{\partial z}P^+ - \frac{\partial\hat{\mathcal{T}}_c^+}{\partial z}P^+ + \hat{\mathcal{T}}_c^+\hat{\mathcal{T}}_c^+P^+ . \tag{B.65}$$

In the second term, using the fact that $\hat{\mathcal{H}}_1^t = \hat{\mathcal{H}}_1$ provides

$$\mathcal{L}^-\mathcal{L}^+P^+ = \mathcal{L}P^+ - i\hat{\mathcal{H}}_1^{-1}\hat{\mathcal{H}}_1\frac{\partial\hat{\mathcal{H}}_1}{\partial z}P^+ + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\hat{\mathcal{R}}_c^-P^- + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\hat{\mathcal{T}}_c^+P^+ + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}S^+ + i\frac{\partial\hat{\mathcal{H}}_1}{\partial z}P^+ - \frac{\partial\hat{\mathcal{T}}_c^+}{\partial z}P^+ + \hat{\mathcal{T}}_c^+\hat{\mathcal{T}}_c^+P^+ . \tag{B.66}$$

In which we observe that the second and sixth terms cancel, simplifying, we obtain

$$\mathcal{L}^-\mathcal{L}^+P^+ = \mathcal{L}P^+ + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\hat{\mathcal{T}}_c^+P^+ + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\hat{\mathcal{R}}_c^-P^- + \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}S^+ + \hat{\mathcal{T}}_c^+\hat{\mathcal{T}}_c^+P^+ . \tag{B.67}$$

Applying the operator $\mathcal{L}^-$ on equation B.53, we have that

$$\mathcal{L}^-\mathcal{L}^+P^+ = \mathcal{L}^-\hat{\mathcal{R}}_c^-P^- + \frac{\partial S^+}{\partial z} - i\hat{\mathcal{H}}_1S^+ - \hat{\mathcal{T}}_c^+S^+ . \tag{B.68}$$

The substitution in the left-hand side of equation B.67 provides

$$\mathcal{L}P^+ = -\hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}\hat{\mathcal{T}}_c^+P^+ + \mathcal{L}^-\hat{\mathcal{R}}_c^-P^- - \hat{\mathcal{H}}_1^{-1}\frac{\partial\hat{\mathcal{H}}_1}{\partial z}S^+ - \hat{\mathcal{T}}_c^+\hat{\mathcal{T}}_c^+P^+ + \frac{\partial S^+}{\partial z} - i\hat{\mathcal{H}}_1S^+ - \hat{\mathcal{T}}_c^+S^+ . \tag{B.69}$$

In which the two-way wave operator is

$$\mathcal{L} = \rho\nabla\cdot\left(\frac{1}{\rho}\nabla\right) + \frac{\omega^2}{c^2} . \tag{B.70}$$

Following the same steps for an upgoing wavefield P^-, a similar expression is obtained. Furthermore, note that the right-hand side of equation B.69 is a source term. It is possible to simplify the right-hand side by considering a small region around the source term vertically homogeneous, this way, we obtain

$$\boxed{\mathcal{L}P^+ = \frac{\partial S^+}{\partial z} - i\hat{\mathcal{H}}_1S^+ .} \tag{B.71}$$

B.5.1 Decoupled Green's functions source term

We define the decoupled Green's functions source terms under the action of the one-way and two-way wave operators. We assume, that in a small region around the source term, the model parameters are vertically homogeneous.

One-way and two-way wave operators source term

We define G_0^+ and G_0^- as the downgoing and upgoing Green's functions that added provide the two-way Green's function G_0. The decoupled Green's function must obey

$$\mathcal{L}^+ G_0^+ = A^+ \,, \tag{B.72}$$

$$\mathcal{L}^- G_0^- = A^- \,, \tag{B.73}$$

in which we want to determine the source terms $A^\pm$. From equation B.69, we have that

$$\mathcal{L} G_0^+ = \frac{\partial A^+}{\partial z} - i\hat{\mathcal{H}}_1 A^+ \,, \tag{B.74}$$

$$\mathcal{L} G_0^- = \frac{\partial A^-}{\partial z} + i\hat{\mathcal{H}}_1 A^- \,, \tag{B.75}$$

where we considered the model parameters vertically homogeneous around the source terms. The two-way Green's function G_0 must obey

$$\mathcal{L} G_0 = -\delta(\mathbf{x} - \mathbf{x}') \,. \tag{B.76}$$

The sum of equations B.74 and B.75 provides

$$\mathcal{L}\left(G_0^+ + G_0^-\right) = \frac{\partial\left(A^+ + A^-\right)}{\partial z} + i\hat{\mathcal{H}}_1\left(A^- - A^+\right) \,. \tag{B.77}$$

We define that $G_0^+ + G_0^- = G_0$ and substitute equation B.76 in equation B.77 to obtain

$$-\delta(\mathbf{x} - \mathbf{x}') = \frac{\partial\left(A^+ + A^-\right)}{\partial z} + i\hat{\mathcal{H}}_1\left(A^- - A^+\right) \,. \tag{B.78}$$

We have one equation and two quantities to be determined. In order to proceed, we impose that one source factor must be the negative of the other, i.e.,

$$A^+ = -A^- \,. \tag{B.79}$$

This choice is inspired in the one-way source terms defined in equation 2.27, but discarding

the injected force density term. The substitution in equation B.78 for A^- provides

$$-\delta(\mathbf{x}-\mathbf{x}') = i\hat{\mathcal{H}}_1 2A^- \,. \tag{B.80}$$

Rearranging the terms, we obtain

$$A^- = \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{B.81}$$

We use the relation $A^+ = -A^-$ and arrive at

$$A^+ = -\frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{B.82}$$

Hence, we substitute these results in equations B.72 and B.73 and conclude that the decoupled Green's are described by

$$\frac{\partial G_0^+}{\partial z} = -i\hat{\mathcal{H}}_1 G_0^+ + \hat{\mathcal{T}}_c^+ G_0^+ - \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,, \tag{B.83}$$

$$\frac{\partial G_0^-}{\partial z} = i\hat{\mathcal{H}}_1 G_0^- - \hat{\mathcal{T}}_c^- G_0^- + \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{B.84}$$

This source terms are consistent with Zhang et al. (2005). They are also similar to the one-way monopole source presented by Wapenaar (1990), except for a $\omega^2\rho$ factor. In Section 2.2.3, we define the downgoing source wavefield consistent with equation B.82.

Substituting the results for $A^\pm$ in equations B.74 and B.75, the action of the two-way wave operator on the decoupled Green's functions is given by

$$\mathcal{L}G_0^+ = -\frac{\partial \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2}\delta(\mathbf{x}-\mathbf{x}') \,, \tag{B.85}$$

$$\mathcal{L}G_0^- = \frac{\partial \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{B.86}$$

We considered the model parameters vertically homogeneous around the source, this way, these expressions simplifies to

$$\mathcal{L}G_0^+ = -\frac{i}{2}\hat{\mathcal{H}}_1^{-1}\frac{\partial\delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2}\delta(\mathbf{x}-\mathbf{x}') \,, \tag{B.87}$$

$$\mathcal{L}G_0^- = \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\frac{\partial\delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2}\delta(\mathbf{x}-\mathbf{x}') \,. \tag{B.88}$$

C Scattering operators

In practical applications, the subsurface models contain discontinuities in the mass density and wavespeed, note that we use the terms wavespeed and velocity interchangeably. This fact leads us to the necessity of defining the reflection and transmission operators for a discontinuous model. Subsequently, after linearization of the reflection and transmission operators, we define the associated continuous reflectivities and transmissivities. The equations developed here are for the general case of a model with lateral variations.

C.1 Reflection and transmission operators

Consider the stack of two horizontal acoustic layers with lateral variations, the upper layer with velocity and mass density (c_u, ρ_u) and correspondingly in the lower layer (c_l, ρ_l). This configuration is illustrated in Figure C.1. The relation between the wavefields represented in Figure C.1 are

$$P_l^+ = \hat{\mathcal{T}}^+ P_u^+ \,, \tag{C.1}$$

$$P_u^- = \hat{\mathcal{R}}^+ P_u^+ \,. \tag{C.2}$$

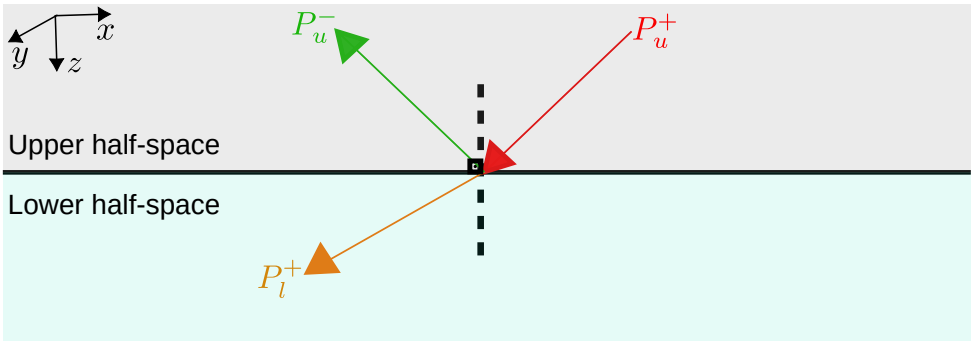


Figure C.1: Scattering of an incident downgoing wave P_u^+ impinging on the boundary between two horizontal layers. P_u^- is the reflected wave traveling upward and P_l^+ is the transmitted wave traveling downward.

In order to proceed with the developments, we impose the continuity of the two-way

quantities, i.e., total pressure P and vertical particle displacement velocity V_z, at the boundary between two homogeneous half spaces as

$$P_u = P_l \,, \tag{C.3}$$

$$V_{z,u} = V_{z,l} \,. \tag{C.4}$$

These relations can be expressed in terms of the one-way wavefields by means of the composition operator as

$$\hat{\boldsymbol{\mathcal{Z}}}_u^{-1}\mathbf{P}_u = \hat{\boldsymbol{\mathcal{Z}}}_l^{-1}\mathbf{P}_l \,. \tag{C.5}$$

The composition operator is given by

$$\hat{\boldsymbol{\mathcal{Z}}}^{-1} = \begin{bmatrix} \mathcal{I} & \mathcal{I} \\ \hat{\mathcal{Z}}^{-1} & -\hat{\mathcal{Z}}^{-1} \end{bmatrix} , \tag{C.6}$$

where $\mathcal{I}$ is the identity operator. For more details about the composition operator, see Appendix B, Section B.2. The elements of the operator $\hat{\boldsymbol{\mathcal{Z}}}^{-1}$ are given by

$$\hat{\mathcal{Z}}^{-1} = \frac{1}{\omega\rho}\hat{\mathcal{H}}_1 \,, \tag{C.7}$$

where $\hat{\mathcal{Z}}$ represents the generalized acoustic impedance, see Section 2.2.2.

For the configuration under consideration, Figure C.1, we have

$$\mathbf{P}_l = \begin{bmatrix} P_l^+ \\ 0 \end{bmatrix} , \tag{C.8}$$

$$\mathbf{P}_u = \begin{bmatrix} P_u^+ \\ P_u^- \end{bmatrix} . \tag{C.9}$$

Thus, the compact notation in equation C.5 represents the relations

$$P_u^+ + P_u^- = P_l^+ \,, \tag{C.10}$$

$$\hat{\mathcal{Z}}_u^{-1}P_u^+ - \hat{\mathcal{Z}}_u^{-1}P_u^- = \hat{\mathcal{Z}}_l^{-1}P_l^+ \,. \tag{C.11}$$

From equations C.1 and C.2, we substitute P_l^+ and P_u^- into equations C.10 and C.11 and obtain

$$P_u^+ + \hat{\mathcal{R}}^+P_u^+ = \hat{\mathcal{T}}^+P_u^+ \,, \tag{C.12}$$

$$\hat{\mathcal{Z}}_u^{-1}P_u^+ - \hat{\mathcal{Z}}_u^{-1}\hat{\mathcal{R}}^+P_u^+ = \hat{\mathcal{Z}}_l^{-1}\hat{\mathcal{T}}^+P_u^+ \,. \tag{C.13}$$

The simplification of the common factor P_u^+ provides

$$\mathcal{I} + \hat{\mathcal{R}}^+ = \hat{\mathcal{T}}^+ , \tag{C.14}$$

$$\hat{\mathcal{Z}}_u^{-1} - \hat{\mathcal{Z}}_u^{-1}\hat{\mathcal{R}}^+ = \hat{\mathcal{Z}}_l^{-1}\hat{\mathcal{T}}^+ . \tag{C.15}$$

We substitute the first equation into the second and obtain

$$\hat{\mathcal{Z}}_u^{-1} - \hat{\mathcal{Z}}_u^{-1}\hat{\mathcal{R}}^+ = \hat{\mathcal{Z}}_l^{-1}\left(\mathcal{I} + \hat{\mathcal{R}}^+\right) . \tag{C.16}$$

Rearranging terms, we obtain

$$\boxed{\hat{\mathcal{R}}^+ = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_u^{-1} - \hat{\mathcal{Z}}_l^{-1}\right] .} \tag{C.17}$$

From equation C.14, we have the relation $\hat{\mathcal{T}}^+ = \hat{\mathcal{R}}^+ + \mathcal{I}$, substituting $\hat{\mathcal{R}}^+$, equation C.17, we obtain the transmission operator

$$\boxed{\hat{\mathcal{T}}^+ = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_u^{-1} - \hat{\mathcal{Z}}_l^{-1}\right] + \mathcal{I} .} \tag{C.18}$$

Similarly for an upgoing wave P_l^- that impinges from below on the boundary between two layers, we have

$$P_u^- = \hat{\mathcal{T}}^- P_l^- , \tag{C.19}$$

$$P_l^+ = \hat{\mathcal{R}}^- P_l^+ . \tag{C.20}$$

Following the procedure outlined above, we obtain

$$\hat{\mathcal{R}}^- = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_l^{-1} - \hat{\mathcal{Z}}_u^{-1}\right] , \tag{C.21}$$

$$\hat{\mathcal{T}}^- = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_l^{-1} - \hat{\mathcal{Z}}_u^{-1}\right] + \mathcal{I} . \tag{C.22}$$

C.2 Recovering the continuous case

In this section, we linearize $\hat{\mathcal{R}}^+$ and $\hat{\mathcal{T}}^+$, equations C.17 and C.18, in order to obtain the associated reflectivity $\hat{\mathcal{R}}_c^+$ and transmissivity $\hat{\mathcal{T}}_c^+$ operators defined in a continuous model and with dimension of distance inverse. Analogous results are obtained for $\hat{\mathcal{R}}^-$ and $\hat{\mathcal{T}}^-$. The linearizations performed in this section were inspired in the work of Foster (1975).

C.2.1 Reflectivity operator

We want to obtain the scattering operator $\hat{\mathcal{R}}^+$, equation C.17, in a continuous model.

Considering that the generalized impedance is continuously dependent of depth coordinate through the model parameters. We define the lower and upper generalized impedances as

$$\hat{\mathcal{Z}}_l = \hat{\mathcal{Z}}(c(z+\Delta z), \rho(z+\Delta z)) = \hat{\mathcal{Z}}(z+\Delta z)\,, \tag{C.23}$$

$$\hat{\mathcal{Z}}_u = \hat{\mathcal{Z}}(c(z-\Delta z), \rho(z-\Delta z)) = \hat{\mathcal{Z}}(z-\Delta z)\,, \tag{C.24}$$

where Δz is a small depth increment. As a consequence, we write the reflection operator as

$$\begin{aligned}\hat{\mathcal{R}}^+(z+\Delta z, z-\Delta z) = &\left[\hat{\mathcal{Z}}^{-1}(z+\Delta z) + \hat{\mathcal{Z}}^{-1}(z-\Delta z)\right]^{-1} \\ &\cdot \left[\hat{\mathcal{Z}}^{-1}(z-\Delta z) - \hat{\mathcal{Z}}^{-1}(z+\Delta z)\right]\,,\end{aligned} \tag{C.25}$$

where we interpret $\hat{\mathcal{R}}^+$ as a function of two variables.

One important result for the following developments, is the derivative of an inverse operator. Consider the composition $\mathcal{P}^{-1}\mathcal{P}$, its derivative is given by

$$\frac{\partial \mathcal{P}^{-1}\mathcal{P}}{\partial z} = \frac{\partial \mathcal{P}^{-1}}{\partial z}\mathcal{P} + \mathcal{P}^{-1}\frac{\partial \mathcal{P}}{\partial z}\,. \tag{C.26}$$

In the left-hand side we recognize the derivative of the identity and it must be zero. Rearranging terms, we obtain

$$\frac{\partial \mathcal{P}^{-1}}{\partial z} = -\mathcal{P}^{-1}\frac{\partial \mathcal{P}}{\partial z}\mathcal{P}^{-1}\,. \tag{C.27}$$

This way, the derivative of $\hat{\mathcal{R}}^+$ with respect to $\hat{\mathcal{Z}}_u^{-1}$ is

$$\frac{\partial \hat{\mathcal{R}}^+}{\partial \hat{\mathcal{Z}}_u^{-1}} = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1} + \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\,. \tag{C.28}$$

And the derivative with respect to $\hat{\mathcal{Z}}_l^{-1}$ is

$$\frac{\partial \hat{\mathcal{R}}^+}{\partial \hat{\mathcal{Z}}_l^{-1}} = \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1} - \left[\hat{\mathcal{Z}}_l^{-1} + \hat{\mathcal{Z}}_u^{-1}\right]^{-1}\,. \tag{C.29}$$

The linearization of equation C.25 around $\Delta z/2$ is

$$\hat{\mathcal{R}}^{+}(z+\Delta z/2, z-\Delta z/2) \approx \hat{\mathcal{R}}^{+}\Big|_{\Delta z/2=0} + \frac{\partial \hat{\mathcal{R}}^{+}}{\partial \hat{\mathcal{Z}}_u^{-1}} \frac{\partial \hat{\mathcal{Z}}_u^{-1}}{\partial z}\Big|_{-\Delta z/2=0} (-\Delta z/2) + \frac{\partial \hat{\mathcal{R}}^{+}}{\partial \hat{\mathcal{Z}}_l^{-1}} \frac{\partial \hat{\mathcal{Z}}_l^{-1}}{\partial z}\Big|_{\Delta z/2=0} \Delta z/2\,, \tag{C.30}$$

where we used the first-order Taylor expansion in two variables. We substitute equations C.28 and C.29 and obtain

$$\hat{\mathcal{R}}^{+}(z+\Delta z/2, z-\Delta z/2) \approx -\frac{\hat{\mathcal{Z}}}{2} \frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} \Delta z\,. \tag{C.31}$$

Thus, we define the reflectivity operator as

$$\boxed{\hat{\mathcal{R}}_c^{+}(z) = -\frac{\hat{\mathcal{Z}}}{2} \frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\,.} \tag{C.32}$$

Following analogous steps for the reflection operator $\hat{\mathcal{R}}^{-}$, we obtain

$$\hat{\mathcal{R}}^{-}(z+\Delta z/2, z-\Delta z/2) \approx \frac{\hat{\mathcal{Z}}}{2} \frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z} \Delta z\,. \tag{C.33}$$

The associated reflectivity is

$$\boxed{\hat{\mathcal{R}}_c^{-}(z) = \frac{\hat{\mathcal{Z}}}{2} \frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\,.} \tag{C.34}$$

C.2.2 Transmissivity operator

The relation between the transmission operators and the reflection operators are

$$\hat{\mathcal{T}}^{+}(z+\Delta z/2, z-\Delta z/2) = \hat{\mathcal{R}}^{+}(z+\Delta z/2, z-\Delta z/2) + \mathcal{I}\,, \tag{C.35}$$

$$\hat{\mathcal{T}}^{-}(z+\Delta z/2, z-\Delta z/2) = \hat{\mathcal{R}}^{-}(z+\Delta z/2, z-\Delta z/2) + \mathcal{I}\,. \tag{C.36}$$

Therefore, the linearization of the transmission operator $\hat{\mathcal{T}}^{+}$ around $\Delta z/2 = 0$ is

$$\hat{\mathcal{T}}^{+}(z+\Delta z/2, z-\Delta z/2) \approx \hat{\mathcal{T}}^{+}\Big|_{\Delta z/2=0} + \frac{\partial \hat{\mathcal{R}}^{+}}{\partial \hat{\mathcal{Z}}_u^{-1}} \frac{\partial \hat{\mathcal{Z}}_u^{-1}}{\partial z}\Big|_{-\Delta z/2=0} (-\Delta z/2) + \frac{\partial \hat{\mathcal{R}}^{+}}{\partial \hat{\mathcal{Z}}_l^{-1}} \frac{\partial \hat{\mathcal{Z}}_l^{-1}}{\partial z}\Big|_{\Delta z/2=0} \Delta z/2\,. \tag{C.37}$$

The substitution of the reflection operators derived in the last section provides

$$\hat{\mathcal{T}}^{+}(z+\Delta z/2, z-\Delta z/2) \approx \mathcal{I} - \frac{\hat{\mathcal{Z}}}{2}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\Delta z\,. \tag{C.38}$$

From which we define the perturbed transmission operator as

$$\delta\hat{\mathcal{T}}^{+} = \hat{\mathcal{T}}^{+} - \mathcal{I} \approx -\frac{\hat{\mathcal{Z}}}{2}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\Delta z\,, \tag{C.39}$$

and the associated transmissivity operator given by

$$\boxed{\hat{\mathcal{T}}_c^{+}(z) = -\frac{\hat{\mathcal{Z}}}{2}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\,.} \tag{C.40}$$

Analogous result is obtained for the transmission with incidence from below a point in the model as

$$\delta\hat{\mathcal{T}}^{-} = \hat{\mathcal{T}}^{-} - \mathcal{I} \approx \frac{\hat{\mathcal{Z}}}{2}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\Delta z\,, \tag{C.41}$$

and the correspondent transmissivity is

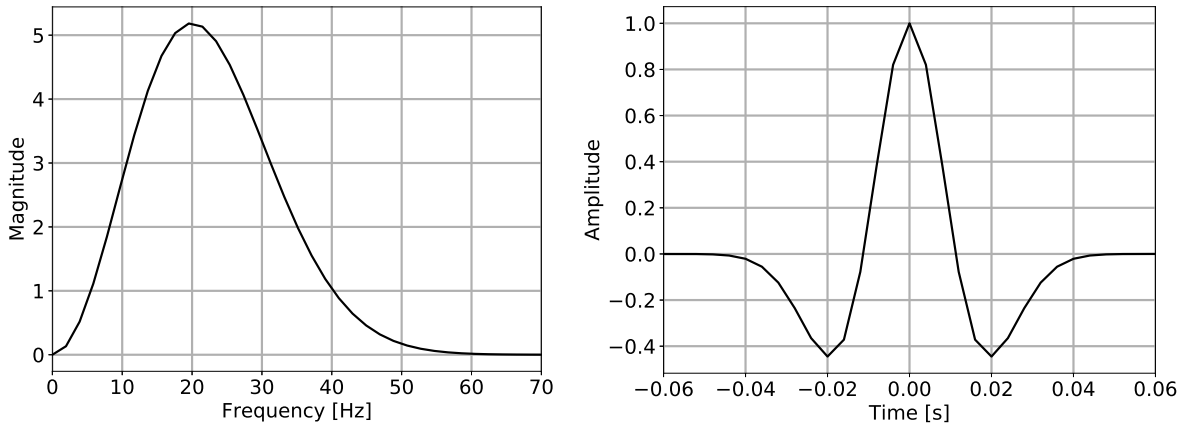
$$\boxed{\hat{\mathcal{T}}_c^{-}(z) = \frac{\hat{\mathcal{Z}}}{2}\frac{\partial \hat{\mathcal{Z}}^{-1}}{\partial z}\,.} \tag{C.42}$$

D Ricker wavelet

In the synthetic tests, we adopted the Ricker wavelet. The algorithms were developed in the space-frequency domain $(\mathbf{x}, \omega)$. Therefore, we used the definition of the Ricker wavelet in the frequency domain as follows

$$W(f, f_p) = \frac{1}{\Delta t} \frac{2f^2}{\sqrt{\pi} f_p^3} e^{-f^2/f_p^2}, \tag{D.1}$$

where f is the frequency in Hz, f_p is the peak frequency and Δt is the time sampling interval in seconds. Considering that we measure only the real-part of the seismic signal, it is only necessary to work with the positive frequencies. Figure D.1(a) shows the magnitude spectrum of the Ricker wavelet. The phase spectrum is zero, which in the time domain translates into the symmetric wavelet of Figure D.1(b).



(a) (b)

Figure D.1: Ricker wavelet with 20 Hz peak frequency. (a) Magnitude spectrum; (b) Time domain.

E Reciprocity

In this section, we develop the reciprocity relation between the decoupled Green's functions. We only consider the transmission configuration.

E.1 Decoupled Green's functions

In the Appendix B, Section B.5, we defined decoupled Green's function G_0^+ and G_0^- that added provides the two-way counterpart G_0, see Appendix B, Section B.5.1. In a vertically homogeneous model in small region around the source, the action of the two-way wave operator on these decoupled Green's functions is given by

$$\nabla \cdot \left(\frac{1}{\rho}\nabla G_0^+\right) + \frac{\omega^2}{c^2} G_0^+ = -\frac{i}{2\rho}\hat{\mathcal{H}}_1^{-1}\frac{\partial \delta(\mathbf{x}-\mathbf{x}')}{\partial z} - \frac{1}{2\rho}\delta(\mathbf{x}-\mathbf{x}')\,, \tag{E.1}$$

$$\nabla \cdot \left(\frac{1}{\rho}\nabla G_0^-\right) + \frac{\omega^2}{c^2} G_0^- = \frac{i}{2\rho}\hat{\mathcal{H}}_1^{-1}\frac{\partial \delta(\mathbf{x}-\mathbf{x}'')}{\partial z} - \frac{1}{2\rho}\delta(\mathbf{x}-\mathbf{x}'')\,, \tag{E.2}$$

where $\delta(\mathbf{x}) = \delta(x)\delta(y)\delta(z)$.

The Green's theorem relating two wavefields A and B is

$$\int_{\mathcal{V}} A\nabla\cdot\left(\frac{1}{\rho}\nabla B\right) - B\nabla\cdot\left(\frac{1}{\rho}\nabla A\right) dV = \int_{\partial\mathcal{V}} \frac{1}{\rho}\left[A\nabla B - B\nabla A\right]\cdot\mathbf{n}dS\,, \tag{E.3}$$

where $\mathcal{V}$ is the integration volume, $\partial\mathcal{V}$ is the volume surface and $\mathbf{n}$ a unitary vector pointing outward $\partial\mathcal{V}$.

We select $A = G_0^+$ and $B = G_0^-$ with source terms inside the integration volume $\mathcal{V}$, see Figure E.1, the volume integral is

$$I_V = \int_{\mathcal{V}} G_0^+\nabla\cdot\left(\frac{1}{\rho}\nabla G_0^-\right) - G_0^-\nabla\cdot\left(\frac{1}{\rho}\nabla G_0^+\right) dV\,. \tag{E.4}$$

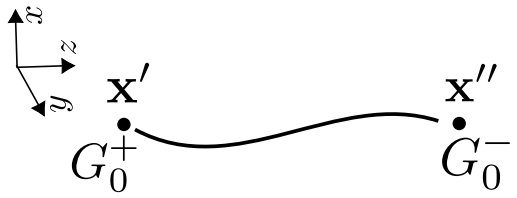


Figure E.1: Representation of the transmission experiment considered in the derivation of the reciprocity relation. G_0^+ and G_0^- are decoupled Green's functions with source term consistent with the two-way counterpart G_0.

The substitution of equations E.1 and E.2 in E.4 provides

$$\begin{aligned} I_V = \int_{\mathcal{V}} & G_0^+ \frac{i}{2\rho} \hat{\mathcal{H}}_1^{-1} \frac{\partial \delta(\mathbf{x}-\mathbf{x}'')}{\partial z} - G_0^+ \frac{1}{2\rho} \delta(\mathbf{x}-\mathbf{x}'') \\ & + G_0^- \frac{i}{2\rho} \hat{\mathcal{H}}_1^{-1} \frac{\partial \delta(\mathbf{x}-\mathbf{x}')}{\partial z} + G_0^- \frac{1}{2\rho} \delta(\mathbf{x}-\mathbf{x}') dV \,, \end{aligned} \tag{E.5}$$

where $G_0^+ = G_0^+(\mathbf{x}, \omega; \mathbf{x}')$ and $G_0^- = G_0^-(\mathbf{x}, \omega; \mathbf{x}'')$. We apply the delta distribution sifting property and obtain

$$\begin{aligned} I_V = & -\frac{i}{2}\frac{\partial}{\partial z}\left(G_0^+ \frac{\hat{\mathcal{H}}_1^{-1}}{\rho} \right)_{\mathbf{x}=\mathbf{x}''} - \frac{1}{2\rho(\mathbf{x}'')} G_0^+(\mathbf{x}'', \omega; \mathbf{x}') \\ & - \frac{i}{2}\frac{\partial}{\partial z}\left(G_0^- \frac{\hat{\mathcal{H}}_1^{-1}}{\rho} \right)_{\mathbf{x}=\mathbf{x}'} + \frac{1}{2\rho(\mathbf{x}')} G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'') \,. \end{aligned} \tag{E.6}$$

The vertical derivative of $\hat{\mathcal{H}}_1$ and ρ vanish at $\mathbf{x}'$ and $\mathbf{x}''$. This way, we obtain

$$\begin{aligned} I_V = & -\frac{i\hat{\mathcal{H}}_1^{-1}}{2\rho(\mathbf{x}'')} \frac{\partial G_0^+(\mathbf{x}'', \omega; \mathbf{x}')}{\partial z} - \frac{1}{2\rho(\mathbf{x}'')} G_0^+(\mathbf{x}'', \omega; \mathbf{x}') \\ & - \frac{i\hat{\mathcal{H}}_1^{-1}}{2\rho(\mathbf{x}')} \frac{\partial G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'')}{\partial z} + \frac{1}{2\rho(\mathbf{x}')} G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'') \,. \end{aligned} \tag{E.7}$$

Considering that the integration volume $\mathcal{V}$ is the whole $\mathbb{R}^3$ and applying Sommerfeld's radiation conditions, the surface integral in Green's theorem must be zero. As a consequence, the volume integral given by equation E.7 is zero and we obtain the reciprocity relation

$$\begin{aligned} & \frac{i\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x}'')} \frac{\partial G_0^+(\mathbf{x}'', \omega; \mathbf{x}')}{\partial z} + \frac{1}{\rho(\mathbf{x}'')} G_0^+(\mathbf{x}'', \omega; \mathbf{x}') = \\ & - \frac{i\hat{\mathcal{H}}_1^{-1}}{\rho(\mathbf{x}')} \frac{\partial G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'')}{\partial z} + \frac{1}{\rho(\mathbf{x}')} G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'') \,. \end{aligned} \tag{E.8}$$

The decoupled Green's functions, see Appendix B, Section B.5.1, for a vertically

homogeneous medium are given by

$$\frac{\partial G_0^+}{\partial z} = -i\hat{\mathcal{H}}_1 G_0^+ - \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}'' - \mathbf{x}')\,, \tag{E.9}$$

$$\frac{\partial G_0^-}{\partial z} = i\hat{\mathcal{H}}_1 G_0^- + \frac{i}{2}\hat{\mathcal{H}}_1^{-1}\delta(\mathbf{x}' - \mathbf{x}'')\,. \tag{E.10}$$

According to the transmission experiment sketched in Figure E.1, the points $\mathbf{x}'$ and $\mathbf{x}''$ are not coincident. As a consequence, the impulsive source terms are zero. This way, substituting the remaining term in each of these equations into equation E.8 and rearranging terms, we obtain

$$\boxed{\frac{2}{\rho(\mathbf{x}'')}G_0^+(\mathbf{x}'', \omega; \mathbf{x}') = \frac{2}{\rho(\mathbf{x}')}G_0^-(\mathbf{x}', \omega; \hat{\mathbf{x}}, \mathbf{x}'')\,.} \tag{E.11}$$

We highlight that these decoupled Green's functions discard any reflections, and that this reciprocity relation was deduced for a vertically homogeneous medium. Hence, it is expected that this is a reasonable result compared to the reciprocity relation of the total Green's function G_0, at least for the transmitted part.

F Complex Padé Fourier finite-difference

The two-dimensional constant mass-density downgoing one-way wave equation is

$$\frac{\partial P^+(x,z,\omega)}{\partial z} = -i\frac{\omega}{c}\sqrt{1+\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}}P^+(x,z,\omega)\,, \tag{F.1}$$

where the velocity has the spatial dependence $c = c(x,z)$ and we neglected transmission and coupling effects. Using a velocity that varies only with depth $c_r = c_r(z)$, the difference d between the square root defined using $c(x,z)$ and using $c_r(z)$ is

$$d = i\frac{\omega}{c}\sqrt{1+\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}} - i\frac{\omega}{c_r}\sqrt{1+\frac{c_r^2}{\omega^2}\frac{\partial^2}{\partial x^2}}\,. \tag{F.2}$$

We will use the complex representation of the Padé expansion to approximate the square roots in equation F.2 (see, e.g., Amazonas et al., 2007). This approximation is given by

$$\sqrt{1+Z} \approx C_0 + \sum_{n=1}^{N}\frac{A_n Z}{1+B_n Z}\,, \tag{F.3}$$

where

$$A_n = \frac{a_n e^{-i\alpha/2}}{\left[1+b_n\left(e^{-i\alpha}-1\right)\right]^2}\,, \tag{F.4}$$

$$B_n = \frac{b_n e^{-i\alpha}}{1+b_n\left(e^{-i\alpha}-1\right)}\,, \tag{F.5}$$

$$a_n = \frac{2}{2N+1}\sin^2\frac{n\pi}{2N+1}\,, \tag{F.6}$$

$$b_n = \cos^2\frac{n\pi}{2N+1}\,, \tag{F.7}$$

and

$$C_0 = e^{i\alpha/2}\left[1+\sum_{n=1}^{N}\frac{a_n\left(e^{-i\alpha}-1\right)}{1+b_n\left(e^{-i\alpha}-1\right)}\right]\,, \tag{F.8}$$

where A_n and B_n are the complex Padé coefficients, and α is the rotation angle of the branch cut of the square root in the complex plane.

Moreover, we define the variables

$$X^2 = \frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}\,, \tag{F.9}$$

$$p = \frac{c_r}{c}\,. \tag{F.10}$$

As a consequence, we have

$$p^2X^2 = \frac{c_r^2}{\omega^2}\frac{\partial^2}{\partial x^2}\,. \tag{F.11}$$

We apply the Padé expansion to equation F.2 and obtain

$$d \approx i\frac{\omega}{c}C_0 + i\frac{\omega}{c}\sum_{n=1}^{N}\frac{A_nX^2}{1+B_nX^2} - i\frac{\omega}{c_r}C_0 - i\frac{\omega}{c_r}\sum_{n=1}^{N}\frac{p^2A_nX^2}{1+p^2B_nX^2}\,. \tag{F.12}$$

Rearranging this equation, we obtain

$$d \approx i\frac{\omega}{c_r}C_0\,(p-1) + i\frac{\omega}{c_r}\left[\sum_{n=1}^{N}\left(\frac{p}{1+B_nX^2} - \frac{p^2}{1+p^2B_nX^2}\right)A_nX^2\right]\,. \tag{F.13}$$

A rational function with the form

$$f(x) = \frac{a}{1+bx}\,, \tag{F.14}$$

has first-order Taylor expansion around $x = 0$ given by

$$\frac{a}{1+bx} \approx a - abx\,. \tag{F.15}$$

We apply this expansion to each fraction in equation F.13 and obtain

$$d \approx i\frac{\omega}{c_r}C_0\,(p-1) + i\frac{\omega}{c_r}\left\{\sum_{n=1}^{N}\left[p - p^2 + \left(-p+p^4\right)B_nX^2\right]A_nX^2\right\}\,, \tag{F.16}$$

$$= i\frac{\omega}{c_r}C_0\,(p-1) + i\frac{\omega}{c_r}\left\{\sum_{n=1}^{N}(p-p^2)\left[1 - \frac{1-p^3}{1-p}B_nX^2\right]A_nX^2\right\}\,. \tag{F.17}$$

We interpret the remaining fraction as the division of two functions dependent of p, i.e., $x(p)$ and $y(p)$. This way, the first-order Taylor expansion for two variables provides

$$\frac{x(p)}{y(p)} \approx \frac{x(p_0)}{y(p_0)} + \frac{1}{y}\bigg|_{p_0}\left[x - x(p_0)\right] - \frac{x}{y^2}\bigg|_{p_0}\left[y - y(p_0)\right]\,. \tag{F.18}$$

This result and the assumption that the reference velocity c_r is much smaller than c,

which implies $p_0 = 0$, provides the following approximation

$$\boxed{\frac{1-p^3}{1-p} \approx 1 + p - p^3 \,.} \tag{F.19}$$

We could have performed the polynomial division exactly. But in practice, we observed that this approximation provides good results during extrapolation. The substitution of this approximation in equation F.17 provides

$$d \approx i\frac{\omega}{c_r} C_0 \left(p - 1\right) + i\frac{\omega}{c_r} \left\{ \sum_{n=1}^{N} (p - p^2) \left[1 - \sigma B_n X^2\right] A_n X^2 \right\} , \tag{F.20}$$

where $\sigma = 1 + p - p^3$. Moreover, we recognize the Taylor expansion given in equation F.15, with $a = 1$ and $b = \sigma$, which provides

$$\boxed{d \approx i\frac{\omega}{c_r} C_0 \left(p - 1\right) + i\frac{\omega}{c_r} \left\{ \sum_{n=1}^{N} \frac{(p - p^2) A_n X^2}{1 + \sigma B_n X^2} \right\} .} \tag{F.21}$$

The substitution of the difference expression defined in equation F.2 into the left-hand side of equation F.21 provides the approximate expression for the square root in a model withvelocity $c(x, z)$ and it is given by

$$i\frac{\omega}{c}\sqrt{1 + X^2} \approx \underbrace{i\frac{\omega}{c_r}\sqrt{1 + p^2 X^2}}_{phase-shift} + \underbrace{i\frac{\omega}{c_r} C_0 \left(p - 1\right)}_{split-step} + \underbrace{i\frac{\omega}{c_r} \sum_{n=1}^{N} \frac{p(1-p) A_n X^2}{1 + \sigma B_n X^2}}_{CPFD} , \tag{F.22}$$

where the factor X^2 is defined in equation F.9. Hence, the complex Padé Fourier finite-difference (CPFFD) approximation to the vertical wavenumber is given by equation F.22.

The CPFFD approximation is composed of three terms, they are: the phase-shift, which can be applied in the $k_x - \omega$ domain and it is related to a model without lateral variations; the split-step correction that must be applied in the $x - \omega$ domain, it is a correction to the wave-propagation near the vertical axis; the complex Padé finite-difference (CPFD) that must be applied in the $x - \omega$ domain.

F.1 CPFFD implementation

The substitution of equation F.22 in the one-way partial differential equation F.1 provides

$$\frac{\partial P^+(x,z,\omega)}{\partial z} = -\left[\underbrace{i\frac{\omega}{c_r}\sqrt{1+p^2X^2}}_{phase-shift} + \underbrace{i\frac{\omega}{c_r}C_0\,(p-1)}_{split-step} + \underbrace{i\frac{\omega}{c_r}\sum_{n=1}^{N}\frac{p(1-p)A_nX^2}{1+\sigma B_nX^2}}_{CPFD}\right]P^+(x,z,\omega)\,. \tag{F.23}$$

The solution of this partial differential equation can be built by cascading the solutions of terms inside the brackets. The phase-shift solution is calculated via

$$P^+_{ps}(k_x,z+\Delta z,\omega) = P^+(k_x,z,\omega)\exp\left\{-i\frac{\omega}{c_r}\sqrt{1-\frac{c_r^2k_x^2}{\omega^2}}\Delta z\right\}. \tag{F.24}$$

Next, we apply the inverse Fourier transform in the horizontal wavenumber k_x. Then, we solve for the split-step term and obtain

$$P^+_{ps,ss}(x,z+\Delta z,\omega) = P^+_{ps}(x,z+\Delta z,\omega)\exp\left\{-i\frac{\omega}{c_r}C_0\,(p-1)\,\Delta z\right\}. \tag{F.25}$$

Note that the terms in the CPFD sum in equation F.23 have the same form, the difference is in the index of the coefficients A_n and B_n. Therefore, considering a generic term of the CPFD part, we have

$$\frac{\partial P^+(x,z,\omega)}{\partial z} = -i\frac{\omega}{c_r}\frac{p(1-p)A_nX^2}{1+\sigma B_nX^2}P^+(x,z,\omega)\,, \tag{F.26}$$

where X^2 is defined in equation F.9. Rearranging the fraction in the right-hand side and substituting X^2, we obtain

$$\left[1+\sigma B_n\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}\right]\frac{\partial P^+(x,z,\omega)}{\partial z} = -i\frac{\omega}{c_r}\left[p(1-p)A_n\frac{c^2}{\omega^2}\frac{\partial^2}{\partial x^2}\right]P^+(x,z,\omega)\,. \tag{F.27}$$

We define a compact notation with $P^+(x,z,\omega) = P_i^{+,j}$ and $P^+(x,z+\Delta z,\omega) = P_i^{+,j+1}$, in which j is related to the z coordinate and i is related to x. Then, we apply the Crank–Nicolson method (see, e.g., Claerbout, 1985) and obtain

$$\left[1+\sigma B_n\frac{c^2}{\omega^2}\frac{D_x^2}{\Delta x^2}\right]\frac{P_i^{+,j+1}-P_i^{+,j}}{\Delta z} = -i\frac{\omega}{c_r}\left[p(1-p)A_n\frac{c^2}{\omega^2}\frac{D_x^2}{\Delta x^2}\right]\frac{P_i^{+,j+1}+P_i^{+,j}}{2}\,, \tag{F.28}$$

where D_x^2 is the centered second-derivative in the x coordinate. The rearrangement of the

terms in the same step of z provides

$$\left[\frac{1}{\Delta z}+\left(\sigma B_n\frac{1}{\Delta z}+i\frac{\omega}{c_r}p(1-p)A_n\frac{1}{2}\right)\frac{c^2}{\omega^2}\frac{D_x^2}{\Delta x^2}\right]P_i^{+,j+1}= \\ \left[\frac{1}{\Delta z}+\left(\sigma B_n\frac{1}{\Delta z}-i\frac{\omega}{c_r}p(1-p)A_n\frac{1}{2}\right)\frac{c^2}{\omega^2}\frac{D_x^2}{\Delta x^2}\right]P_i^{+,j}\,. \tag{F.29}$$

The reorganization of the coefficients involving Δz and Δx provides

$$\left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n\ +i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P_i^{+,j+1}= \\ \left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n-i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P_i^{+,j}\,. \tag{F.30}$$

We are concatenating different solutions of the one-way wave equation. In this manner, we substitute the actual wavefield $P_i^{+,j}$ in the right-hand side of equation F.30 by the extrapolated wavefield after the application of the phase-shift and split-step solutions given by $P^+_{i,ps,ss}$ in equation F.25. Thus, $P_i^{+,j+1}$ is the final wavefield that combines all one-way propagators, we label it as $P^+_{i,cpffd}$, and its is a solution of

$$\left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n\ +i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P^+_{i,cpffd}= \\ \left[\frac{\omega^2\Delta x^2}{c^2}+\left(\sigma B_n-i\frac{\omega\Delta z}{2c_r}p(1-p)A_n\right)D_x^2\right]P^+_{i,ps,ss}\,. \tag{F.31}$$

If more terms are considered in the CPFD sum, the equation F.31 is solved again with the correspondent coefficients A_n and B_n, and the last solution is considered as the new wavefield in the right-hand side.

G Lagrangian multipliers

In this appendix, we apply the Lagrangian multipliers method to compute the misfit-function partial derivative for migration and inversion. We follow the methodology outlined by Plessix (2006), Askan et al. (2007) and Métivier et al. (2017).

G.1 Forward modeling equations

The continuous down- and upgoing modeling equations are

$$\frac{\partial P^+}{\partial z} = -i\hat{\mathcal{H}}_1 P^+ + \hat{\mathcal{T}}_c^+ P^+ + \hat{\mathcal{R}}_c^- P^- + S^+ \,, \tag{G.1}$$

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \hat{\mathcal{T}}_c^- P^- - \hat{\mathcal{R}}_c^+ P^+ \,, \tag{G.2}$$

where $P^+(\mathbf{x},\omega)$ is the downgoing wavefield, $P^-(\mathbf{x},\omega)$ is the upgoing wavefield, $\hat{\mathcal{H}}_1$ is the square-root operator, $\hat{\mathcal{T}}_c^{\pm}$ are the transmissivities, $\hat{\mathcal{R}}_c^{\pm}$ are the reflectivities and S^+ is the downgoing source. See Section 2.2.2 for more details.

It is important to have in mind that the action of these operators on the wavefields is given by

$$\left(\hat{\mathcal{B}}U\right)(\mathbf{x},\omega) = \int_{\mathbb{R}^2} \hat{\mathcal{B}}(\mathbf{x},\omega;x',y')U(x',y',z,\omega)dx'dy' \,, \tag{G.3}$$

where $\hat{\mathcal{B}}$ represents $\hat{\mathcal{H}}_1$, $\hat{\mathcal{T}}_c^{\pm}$ or $\hat{\mathcal{R}}_c^{\pm}$, and U is a wavefield (for more details see Appendix B, Section B.3).

G.2 Lagrangian function

In preparation to build the Lagrangian function, we write the forward modeling equations G.1 and G.2 as

$$F^+(P^+, P^-, \hat{\mathcal{R}}_c^+, c) = \frac{\partial P^+}{\partial z} + i\hat{\mathcal{H}}_1 P^+ - \hat{\mathcal{R}}_c^+ \left(P^+ - P^-\right) , \tag{G.4}$$

$$F^-(P^+, P^-, \hat{\mathcal{R}}_c^+, c) = \frac{\partial P^-}{\partial z} - i\hat{\mathcal{H}}_1 P^- + \hat{\mathcal{R}}_c^+ \left(P^+ - P^-\right) , \tag{G.5}$$

where we separated the dependency of the wavefields on the models parameters and neglected the downgoing source. Moreover, we used the following relationships between scattering operators in acoustic media

$$\hat{\mathcal{R}}_c^- = -\hat{\mathcal{R}}_c^+ , \tag{G.6}$$

$$\hat{\mathcal{T}}_c^- = -\hat{\mathcal{R}}_c^+ . \tag{G.7}$$

We are investigating modeling equations that evolve in the depth coordinate, this way, we define the associated boundary conditions as

$$P^+(x, y, z = z_f, \omega) = 0 , \tag{G.8}$$

$$P^-(x, y, z = 0, \omega) = 0 , \tag{G.9}$$

where z_f is the depth level at the bottom of the model domain.

In order to estimate the model parameters, we minimize the least-squares misfit function,

$$E(P^-) = \frac{1}{2} \sum_{r=1}^{N_r} \int_{\omega_i}^{\omega_f} \int_{\Omega} \left(D^- - P^-\right)^* \left(D^- - P^-\right) \delta(\mathbf{x} - \mathbf{x}_r) d\mathbf{x} d\omega , \tag{G.10}$$

where r is the receiver index, N_r is the number of receivers, ω_i the initial angular frequency, ω_f the final angular frequency, Ω is the model parameters spatial domain, $D^-(\mathbf{x}, \omega)$ is the observed data, $P^-(\mathbf{x}, \omega)$ is the modeled data, $\mathbf{x}_r = (x_r, y_r, z_r)$ is the receiver position and the asterisk denotes the complex conjugate. Moreover, in this equation we considered that the misfit function depends on the upgoing wavefield. As the derivation evolves, we will restore the dependence of the one-way wavefields on the model parameters and, consequently, the misfit function will depend on the same variables.

We pose the optimization problem as the minimization of the misfit function $E(P^-)$, subject to the forward modeling equations $F^{\pm}(P^+, P^-, \hat{\mathcal{R}}_c^+, c)$. Formulating this problem

in terms of a Lagrangian function, yields

$$L(P^+, P^-, \hat{\mathcal{R}}_c^+, c, \Lambda^+, \Lambda^-) = E(P^-) + \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_\Omega \left(\Lambda^+\right)^* F^+(P^+, P^-, \hat{\mathcal{R}}_c^+) d\mathbf{x} d\omega\right\} + \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_\Omega \left(\Lambda^-\right)^* F^-(P^+, P^-, \hat{\mathcal{R}}_c^+) d\mathbf{x} d\omega\right\}, \tag{G.11}$$

where $\Lambda^{\pm}(\mathbf{x}, \omega)$ are Lagrangian multipliers, also called adjoint-state variables, and $\boldsymbol{Re}\{.\}$ denotes the real-part operator.

A necessary condition for a solution to be considered optimum, is that the first variation of the Lagrangian must be stationary. This is accomplished by making

$$\Delta L(P^+, P^-, \hat{\mathcal{R}}_c^+, c, \Lambda^+, \Lambda^-) = \frac{\partial L}{\partial P^+}\Delta P^+ + \frac{\partial L}{\partial P^-}\Delta P^- + \frac{\partial L}{\partial \hat{\mathcal{R}}_c^+}\Delta \hat{\mathcal{R}}_c^+ + \frac{\partial L}{\partial c}\Delta c + \frac{\partial L}{\partial \Lambda^+}\Delta\Lambda^+ + \frac{\partial L}{\partial \Lambda^-}\Delta\Lambda^- = 0\,. \tag{G.12}$$

G.2.1 Forward wavefields

The Lagrangian variation with respect to Λ^+ is

$$\Delta_{\Lambda^+} L = \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_\Omega \left(\Delta\Lambda^+\right)^* F^+ d\mathbf{x} d\omega\right\}. \tag{G.13}$$

We impose $\Delta_{\Lambda^+} L = 0$ and obtain

$$\boxed{F^+(P^+, P^-, \hat{\mathcal{R}}_c^+, c) = 0\,.} \tag{G.14}$$

The Lagrangian variation with respect to Λ^- is

$$\Delta_{\Lambda^-} L = \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_\Omega \left(\Delta\Lambda^-\right)^* F^- d\mathbf{x} d\omega\right\}. \tag{G.15}$$

We impose $\Delta_{\Lambda^-} L = 0$ and obtain

$$\boxed{F^-(P^+, P^-, \hat{\mathcal{R}}_c^+, c) = 0\,.} \tag{G.16}$$

Hence, from the variations with respect to Λ^+ and Λ^-, we recover the forward modeling equations **??** and G.5.

G.2.2 Adjoint wavefields

The variation of L with respect to P^- is more demanding to derive. Let us manipulate the third term in the Lagrangian. The substitution of F^- provides

$$\int_{\omega_i}^{\omega_f}\int_{\mathbb{R}^2}\int_0^{z_f}\left(\Lambda^-\right)^*\left\{\frac{\partial P^-}{\partial z}-i\hat{\mathcal{H}}_1P^-+\hat{\mathcal{R}}_c^+\left(P^+-P^-\right)\right\}d\mathbf{x}d\omega\,. \tag{G.17}$$

The operators $\hat{\mathcal{H}}_1$ and $\hat{\mathcal{R}}_c^+$ are convolutions with the wavefields in the lateral coordinates, see equation G.3. Taking the reflectivity as an example, we have

$$\int_{\mathbb{R}^2}\left(\Lambda^-\right)^*(x,y,z,\omega)\int_{\mathbb{R}^2}\hat{\mathcal{R}}_c^+(x,y,z,\omega;x',y')P^+(x',y',z,\omega)dx'dy'dxdy\,. \tag{G.18}$$

We change the integration order of the horizontal variables and obtain

$$\int_{\mathbb{R}^2}\left\{\int_{\mathbb{R}^2}\hat{\mathcal{R}}_c^+(x,y,z,\omega;x',y')\left(\Lambda^-\right)^*(x,y,z,\omega)dxdy\right\}P^+(x',y',z,\omega)dx'dy'\,, \tag{G.19}$$

where we recognize the correlation between Λ^+ and $\hat{\mathcal{R}}_c^+$ in the horizontal coordinates. This result is equivalent to

$$\int_{\mathbb{R}^2}\left\{\int_{\mathbb{R}^2}\hat{\mathcal{R}}_c^+(x',y',z,\omega;x,y)\left(\Lambda^-\right)^*(x',y',z,\omega)dx'dy'\right\}P^+(\mathbf{x},\omega)dxdy\,, \tag{G.20}$$

where, in analogy with the discrete case represented by matrices, we observe that the change in the order of integration is similar to a matrix transposition (see, e.g., Menke, 2018). Thus, we define the compact notation

$$\begin{aligned}\int_{\mathbb{R}^2}\left(\hat{\mathcal{R}}_c^+\right)^t\left(\Lambda^-\right)^*P^+d\mathbf{x}=\int_{\mathbb{R}^2}\left\{\int_{\mathbb{R}^2}\hat{\mathcal{R}}_c^+(x',y',z,\omega;x,y)\left(\Lambda^-\right)^*(x',y',z,\omega)dx'dy'\right\}\\ P^+(\mathbf{x},\omega)dxdy\,,\end{aligned} \tag{G.21}$$

Now, we integrate the first term of equation G.17 by parts over the z axis and obtain

$$\begin{aligned}\int_{\omega_i}^{\omega_f}\int_{\mathbb{R}^2}\int_0^{z_f}\left(\Lambda^-\right)^*\frac{\partial P^-}{\partial z}d\mathbf{x}d\omega=\int_{\omega_i}^{\omega_f}\int_{\mathbb{R}^2}\left(\Lambda^-\right)^*P^-\Big|_0^{z_f}dxdyd\omega\\ -\int_{\omega_i}^{\omega_f}\int_{\mathbb{R}^2}\int_0^{z_f}\frac{\partial\left(\Lambda^-\right)^*}{\partial z}P^-d\mathbf{x}d\omega\,.\end{aligned} \tag{G.22}$$

We defined that $P^-(x,y,z=0,\omega)=0$, therefore, in order to make the first term in this expression vanish, we also impose the boundary condition

$$\boxed{\Lambda^-(x,y,z=z_f,\omega)=0\,.} \tag{G.23}$$

As a consequence, we have

$$\int_{\omega_i}^{\omega_f}\int_{\Omega}\left(\Lambda^{-}\right)^{*}\frac{\partial P^{-}}{\partial z}d\mathbf{x}d\omega=-\int_{\omega_i}^{\omega_f}\int_{\Omega}\frac{\partial\left(\Lambda^{-}\right)^{*}}{\partial z}P^{-}d\mathbf{x}d\omega\,. \tag{G.24}$$

Wapenaar and Grimbergen (1996) demonstrated that the square-root operator is symmetric, i.e., $\hat{\mathcal{H}}_1^t=\hat{\mathcal{H}}_1$. Hence, using equations G.21, $\hat{\mathcal{H}}_1^t=\hat{\mathcal{H}}_1$ and G.24, the third term in the Lagrangian can be written as

$$\int_{\omega_i}^{\omega_f}\int_{\Omega}\left\{-\frac{\partial\left(\Lambda^{-}\right)^{*}}{\partial z}-i\hat{\mathcal{H}}_1\left(\Lambda^{-}\right)^{*}-\left(\hat{\mathcal{R}}_c^{+}\right)^{t}\left(\Lambda^{-}\right)^{*}\right\}P^{-}+\left(\Lambda^{-}\right)^{*}\hat{\mathcal{R}}_c^{+}P^{+}d\mathbf{x}d\omega\,. \tag{G.25}$$

Or approximately, we have

$$\int_{\omega_i}^{\omega_f}\int_{\Omega}\left\{-\frac{\partial\Lambda^{-}}{\partial z}-i\hat{\mathcal{H}}_1\Lambda^{-}-\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\Lambda^{-}\right\}^{*}P^{-}+\left(\Lambda^{-}\right)^{*}\hat{\mathcal{R}}_c^{+}P^{+}d\mathbf{x}d\omega\,, \tag{G.26}$$

where $\dagger$ denotes transpose and complex conjugate. This expression is approximate because we considered $\hat{\mathcal{H}}_1^{*}\approx\hat{\mathcal{H}}_1$, it holds only for propagating waves (Wapenaar and Grimbergen, 1996).

Now, we are ready to calculate the variation of the Lagrangian with respect to P^{-}. We substitute equation G.26 in the third term of the Lagrangian in G.11 and obtain

$$\begin{aligned}\Delta_{P^{-}}L=&\sum_{r=1}^{N_r}\int_{\omega_i}^{\omega_f}\int_{\Omega}-\left(D^{-}-P^{-}\right)\delta(\mathbf{x}-\mathbf{x}_r)(\mathbf{x}-\mathbf{x}_r)\left(\Delta P^{-}\right)^{*}d\mathbf{x}d\omega\\&+\boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega}\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\Lambda^{+}\left(\Delta P^{-}\right)^{*}d\mathbf{x}d\omega\right\}\\&+\boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega}\left\{-\frac{\partial\Lambda^{-}}{\partial z}-i\hat{\mathcal{H}}_1\Lambda^{-}-\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\Lambda^{-}\right\}\left(\Delta P^{-}\right)^{*}d\mathbf{x}d\omega\right\}.\end{aligned} \tag{G.27}$$

We impose $\Delta_{P^{-}}L=0$ and obtain the adjoint modeling equation

$$\boxed{\frac{\partial\Lambda^{-}}{\partial z}=-i\hat{\mathcal{H}}_1\Lambda^{-}+\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\left(\Lambda^{+}-\Lambda^{-}\right)-\sum_{r=1}^{N_r}\left(D^{-}-P^{-}\right)\delta(\mathbf{x}-\mathbf{x}_r)\,.} \tag{G.28}$$

Following similar steps for the variation of L with respect to P^{+}, we obtain

$$\begin{aligned}\Delta_{P^{+}}L=&\boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega}\left\{-\frac{\partial\Lambda^{+}}{\partial z}+i\hat{\mathcal{H}}_1\Lambda^{+}-\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\Lambda^{+}\right\}\left(\Delta P^{+}\right)^{*}d\mathbf{x}d\omega\right\}\\&+\boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega}\left(\hat{\mathcal{R}}_c^{+}\right)^{\dagger}\Lambda^{-}\left(\Delta P^{+}\right)^{*}d\mathbf{x}d\omega\right\}.\end{aligned} \tag{G.29}$$

From the integration by parts over depth, we obtain the boundary condition

$$\boxed{\Lambda^{+}(x,y,z=0,\omega)=0\,.} \tag{G.30}$$

Then, we impose $\Delta_{P^{+}}L=0$ and obtain the adjoint modeling equation

$$\boxed{\frac{\partial\Lambda^{+}}{\partial z}=i\hat{\mathcal{H}}_{1}\Lambda^{+}-\left(\hat{\mathcal{R}}_{c}^{+}\right)^{\dagger}\left(\Lambda^{+}-\Lambda^{-}\right)\,.} \tag{G.31}$$

Therefore, the variations of the Lagrangian with respect to $\Lambda^{\pm}$ provide the associated adjoint modeling equations G.28 and G.31. We observe that both adjoint-state variables acts as a secondary sources in these equations, similarly to the forward modeling equations G.1 and G.2. Thus, we have coupled adjoint equations. Although the adjoint modeling equation G.31 does not have a physical source term, it has a secondary source related to Λ^{-}.

G.3 Migration partial derivatives

During migration, we are interested in estimating the scattering operators, e.g. reflectivity or reflection operator. In this manner, we consider that at least a background wavespeed model is available.

The variation of the Lagrangian with respect to $\hat{\mathcal{R}}_{c}^{+}$, yields

$$\begin{aligned}\Delta_{\hat{\mathcal{R}}_{c}^{+}}L=&\boldsymbol{Re}\left\{\int_{\omega_{i}}^{\omega_{f}}\int_{\Omega}-\left(\Lambda^{+}\right)^{*}\Delta\hat{\mathcal{R}}_{c}^{+}\left(P^{+}-P^{-}\right)d\mathbf{x}d\omega\right\}\\&+\boldsymbol{Re}\left\{\int_{\omega_{i}}^{\omega_{f}}\int_{\Omega}\left(\Lambda^{-}\right)^{*}\Delta\hat{\mathcal{R}}_{c}^{+}\left(P^{+}-P^{-}\right)d\mathbf{x}d\omega\right\}\,.\end{aligned} \tag{G.32}$$

Considering that we can change the role of the operators and the wavefields and transposing the result as we defined in equation G.21, we obtain

$$\begin{aligned}\Delta_{\hat{\mathcal{R}}_{c}^{+}}L=&\boldsymbol{Re}\left\{\int_{\omega_{i}}^{\omega_{f}}\int_{\Omega}-\left(\hat{\mathcal{P}}^{+}-\hat{\mathcal{P}}^{-}\right)^{t}\left(\Lambda^{+}\right)^{*}\Delta\mathcal{R}_{c}^{+}d\mathbf{x}d\omega\right\}\\&+\boldsymbol{Re}\left\{\int_{\omega_{i}}^{\omega_{f}}\int_{\Omega}\left(\hat{\mathcal{P}}^{+}-\hat{\mathcal{P}}^{-}\right)^{t}\left(\Lambda^{-}\right)^{*}\Delta\mathcal{R}_{c}^{+}d\mathbf{x}d\omega\right\}\,,\end{aligned} \tag{G.33}$$

where $\hat{\mathcal{P}}^{\pm}$ encapsulate the concept of integral operator defined in equation G.3. Then, we impose $\Delta_{\hat{\mathcal{R}}_{c}^{+}}L=0$ and obtain

$$-\left(\hat{\mathcal{P}}^{+}-\hat{\mathcal{P}}^{-}\right)^{t}\left(\Lambda^{+}-\Lambda^{-}\right)^{*}=0\,. \tag{G.34}$$

That is equivalent to

$$\boxed{-\left(\hat{\mathcal{P}}^{+}-\hat{\mathcal{P}}^{-}\right)^{\dagger}\left(\Lambda^{+}-\Lambda^{-}\right)=0\,.} \tag{G.35}$$

G.3.1 Reduced Lagrangian

Considering that in the Lagrangian function the wavefields are dependent of the scattering operators, we have

$$P^{+}=\bar{P}^{+}(\hat{\mathcal{R}}_{c}^{+},c)\,, \tag{G.36}$$

$$P^{-}=\bar{P}^{-}(\hat{\mathcal{R}}_{c}^{+},c)\,. \tag{G.37}$$

This definition yields

$$F^{+}(\bar{P}^{+},\bar{P}^{-},\hat{\mathcal{R}}_{c}^{+},c)=0\,, \tag{G.38}$$

$$F^{-}(\bar{P}^{+},\bar{P}^{-},\hat{\mathcal{R}}_{c}^{+},c)=0\,. \tag{G.39}$$

As a consequence, we obtain the reduced Lagrangian

$$L(\bar{P}^{+},\bar{P}^{-},\hat{\mathcal{R}}_{c}^{+},c,\bar{\Lambda}^{+},\bar{\Lambda}^{-})=E(\bar{P}^{-})\,. \tag{G.40}$$

It is noteworthy that defining $\bar{P}^{\pm}$ such that $F^{\pm}=0$, makes the choice of $\Lambda^{\pm}$ arbitrary in equation G.11. As usual, we define $\Lambda^{\pm}$ though equations G.28 and G.31, to maintain consistency with the general formulation. This way, $\bar{\Lambda}^{\pm}$ are the adjoint wavefields, defined in equations G.28 and G.31, calculated using the wavefields $\bar{P}^{\pm}$.

The derivative of equation G.40 with respect to $\hat{\mathcal{R}}_{c}^{+}$ provides the identity

$$\frac{\partial L(\bar{P}^{+},\bar{P}^{-},\hat{\mathcal{R}}_{c}^{+},c,\bar{\Lambda}^{+},\bar{\Lambda}^{-})}{\partial\hat{\mathcal{R}}_{c}^{+}}=\frac{\partial E(\bar{P}^{-})}{\partial\hat{\mathcal{R}}_{c}^{+}}\,. \tag{G.41}$$

From the variation $\Delta_{\hat{\mathcal{R}}_{c}^{+}}L$ given by equation G.33 and the last identity, we conclude that

$$\frac{\partial E(\bar{P}^{-})}{\partial\hat{\mathcal{R}}_{c}^{+}}=-\left(\hat{\mathcal{P}}^{+}-\hat{\mathcal{P}}^{-}\right)^{\dagger}\left(\bar{\Lambda}^{+}-\bar{\Lambda}^{-}\right)\,. \tag{G.42}$$

In the developments, we considered only one experiment, i.e., one shot gather. In practice, we have many shot-gathers, therefore, we include a summation in the last result and obtain

$$\boxed{\frac{\partial E(\bar{P}^{-})}{\partial\hat{\mathcal{R}}_{c}^{+}}=-\sum_{s=1}^{N_{s}}\left(\hat{\mathcal{P}}_{s}^{+}-\hat{\mathcal{P}}_{s}^{-}\right)^{\dagger}\left(\bar{\Lambda}_{s}^{+}-\bar{\Lambda}_{s}^{-}\right)\,,} \tag{G.43}$$

where N_s is the number of shots and s is the shot index.

In the angle-independent case, see Section 4.1, the reflectivity operator is approximately

$$\hat{\mathcal{R}}_c^+ \approx R_c^+ = \frac{1}{2Z}\frac{\partial Z}{\partial z}\,, \tag{G.44}$$

where $Z = \rho c$ is the acoustic impedance. We observe that in this case, the reflectivity is also independent of the frequency and there is no operation over the lateral coordinates. This way, the misfit-function partial derivative simplifies to

$$\boxed{\frac{\partial E(\bar{P}^-)}{\partial R_c^+} = -\sum_{s=1}^{N_s} \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} \left(\bar{P}_s^+ - \bar{P}_s^-\right)^* \left(\bar{\Lambda}_s^+ - \bar{\Lambda}_s^-\right) d\omega\right\}\,.} \tag{G.45}$$

G.3.2 Impedance parameterization

The reflectivity, in the angle-independent case, can be approximated by

$$\hat{\mathcal{R}}_c^+ \approx R_c^+ = \frac{1}{2}\frac{\partial \ln\left(Z/Z_0\right)}{\partial z}\,, \tag{G.46}$$

where $Z = \rho c$ is the acoustic impedance and Z_0 is a constant. In this approximation, the reflectivity operator acts on the wavefields by direct multiplication, i.e., there is no integration over the lateral coordinates, and it is frequency independent. See Section 4.1 for a detailed discussion about this approximation. The substitution in the forward modeling equations G.4 and G.5 provides

$$F^+(P^+, P^-, Z, c) = \frac{\partial P^+}{\partial z} + i\hat{\mathcal{H}}_1 P^+ - \frac{1}{2}\frac{\partial \ln\left(Z/Z_0\right)}{\partial z}\left(P^+ - P^-\right)\,, \tag{G.47}$$

$$F^-(P^+, P^-, Z, c) = \frac{\partial P^-}{\partial z} - i\hat{\mathcal{H}}_1 P^- + \frac{1}{2}\frac{\partial \ln\left(Z/Z_0\right)}{\partial z}\left(P^+ - P^-\right)\,. \tag{G.48}$$

Therefore, the Lagrangian in equation G.11 is redefined as

$$\begin{aligned} L(P^+, P^-, Z, c, \Lambda^+, \Lambda^-) = E(P^-) &+ \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega} \left(\Lambda^+\right)^* F^+(P^+, P^-, Z, c) d\mathbf{x} d\omega\right\} \\ &+ \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f}\int_{\Omega} \left(\Lambda^-\right)^* F^-(P^+, P^-, Z, c) d\mathbf{x} d\omega\right\}\,, \end{aligned} \tag{G.49}$$

In preparation to calculate the variation of L with respect to Z, we define the first-order expansion of the impedance logarithm,

$$\frac{1}{2}\ln\left(Z + \Delta Z\right) - \frac{1}{2}\ln\left(Z_0\right) \approx \frac{1}{2}\ln\left(Z/Z_0\right) + \frac{1}{2}\frac{\Delta Z}{Z}\,. \tag{G.50}$$

Using this result and integrating by parts over the depth coordinate the term involving the impedance vertical derivative in G.49, the Lagrangian variation with respect

to the impedance is

$$\Delta_Z L = \int_\Omega \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} \frac{1}{2Z}\left[\frac{\partial\left(\Lambda^+\right)^*}{\partial z}\left(P^+ - P^-\right) + \left(\Lambda^+\right)^*\left(\frac{\partial P^+}{\partial z} - \frac{\partial P^-}{\partial z}\right)\right] d\omega\right\} \Delta Z\, d\mathbf{x} + \int_\Omega \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} -\frac{1}{2Z}\left[\frac{\partial\left(\Lambda^-\right)^*}{\partial z}\left(P^+ - P^-\right) + \left(\Lambda^-\right)^*\left(\frac{\partial P^+}{\partial z} - \frac{\partial P^-}{\partial z}\right)\right] d\omega\right\} \Delta Z\, d\mathbf{x}\,. \tag{G.51}$$

Finally, considering the reduced Lagrangian developed in Section G.3.1 and multiple shot gathers, the misfit-function partial derivative with respect to impedance is

$$\frac{\partial E}{\partial Z} = \sum_{s=1}^{N_s} \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} \frac{1}{2Z}\left[\left(\frac{\partial \bar{\Lambda}_s^+}{\partial z} - \frac{\partial \bar{\Lambda}_s^-}{\partial z}\right)\left(\bar{P}_s^+ - \bar{P}_s^-\right)^* + \left(\bar{\Lambda}_s^+ - \bar{\Lambda}_s^-\right)\left(\frac{\partial \bar{P}_s^+}{\partial z} - \frac{\partial \bar{P}_s^-}{\partial z}\right)^*\right] d\omega\right\}, \tag{G.52}$$

where the adjoint state-variables are given by equations G.28 and G.31.

G.4 Inversion partial derivatives

In this section, we consider that an initial known wavespeed model is not accurate enough for migration. Therefore, we use the Lagrangian function in G.11 to derive the partial derivative of the misfit function with respect to wavespeed.

The square-root and the scattering operators depend on the wavespeed. However, keeping the philosophy of separating dynamic from kinematic effects, we consider only the square-root operator as wavespeed dependent. This operator is given by

$$\hat{\mathcal{H}}_1 = \left[\rho\hat{\nabla}\cdot\left(\frac{1}{\rho}\hat{\nabla}\right) + \frac{\omega^2}{c^2}\right]^{1/2}, \tag{G.53}$$

such that two-fold application provides the transversal Helmholtz operator, i.e., $\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1\hat{\mathcal{H}}_1$. The associated Lagrangian function is

$$L(P^+, P^-, \hat{\mathcal{R}}_c^+, c, \Lambda^+, \Lambda^-) = \frac{1}{2}\sum_{r=1}^{N_r}\left\langle D^- - P^-, \left(D^- - P^-\right)\delta(\mathbf{x}-\mathbf{x}_r)\right\rangle + \boldsymbol{Re}\left\{\left\langle \Lambda^+, F^+(P^+, P^-, \hat{\mathcal{R}}_c^+, c)\right\rangle\right\} + \boldsymbol{Re}\left\{\left\langle \Lambda^-, F^-(P^+, P^-, \hat{\mathcal{R}}_c^+, c)\right\rangle\right\}, \tag{G.54}$$

where $\langle .,. \rangle$ denotes inner product. For example, consider two complex quantities $A(\mathbf{x},\omega)$ and $B(\mathbf{x},\omega)$, their inner product is

$$\langle A, B \rangle = \int_{\Omega} \int_{\omega_i}^{\omega_f} A^* B d\omega d\mathbf{x} \,. \tag{G.55}$$

The variation of L, equation G.54, with respect to wavespeed is

$$\Delta_c L = \boldsymbol{Re} \left\{ \left\langle \Lambda^+, \frac{\partial F^+}{\partial c} \Delta c \right\rangle + \left\langle \Lambda^-, \frac{\partial F^-}{\partial c} \Delta c \right\rangle \right\} . \tag{G.56}$$

We consider that the dependency of F^+ and F^- on wavespeed comes only from the square-root operator related to extrapolation, that is the second term in equations G.4 and G.5. This way, the required partial derivatives are

$$\boxed{\frac{\partial F^{\pm}}{\partial c} = \pm i \frac{\partial \hat{\mathcal{H}}_1}{\partial c} P^{\pm} \,.} \tag{G.57}$$

From the relation $\hat{\mathcal{H}}_2 = \hat{\mathcal{H}}_1 \hat{\mathcal{H}}_1$, we have that

$$\frac{\partial \hat{\mathcal{H}}_2}{\partial c} = \frac{\partial \hat{\mathcal{H}}_1}{\partial c} \hat{\mathcal{H}}_1 + \hat{\mathcal{H}}_1 \frac{\partial \hat{\mathcal{H}}_1}{\partial c} \,. \tag{G.58}$$

The symmetry $\hat{\mathcal{H}}_1^t = \hat{\mathcal{H}}_1$ (Wapenaar and Grimbergen, 1996) provides

$$\frac{\partial \hat{\mathcal{H}}_2}{\partial c} = 2 \hat{\mathcal{H}}_1 \frac{\partial \hat{\mathcal{H}}_1}{\partial c} \,. \tag{G.59}$$

That is equivalent to

$$\frac{\partial \hat{\mathcal{H}}_1}{\partial c} = \frac{1}{2} \hat{\mathcal{H}}_1^{-1} \frac{\partial \hat{\mathcal{H}}_2}{\partial c} \,. \tag{G.60}$$

The transversal Helmholtz operator is given by

$$\hat{\mathcal{H}}_2 = \rho \hat{\nabla} \cdot \left(\frac{1}{\rho} \hat{\nabla} \right) + \frac{\omega^2}{c^2} \,. \tag{G.61}$$

Thus, its partial derivative with respect to velocity is

$$\frac{\partial \hat{\mathcal{H}}_2}{\partial c} = -2 \frac{\omega^2}{c^3} \,. \tag{G.62}$$

Finally, the substitution of the partial derivative of the transversal Helmholtz in equation G.60, yields

$$\boxed{\frac{\partial \hat{\mathcal{H}}_1}{\partial c} = -\hat{\mathcal{H}}_1^{-1} \frac{\omega^2}{c^3} \,.} \tag{G.63}$$

Hence, from the results in G.57 and G.63, the variation $\Delta_c L$ in equation G.56 is

$$\Delta_c L = \boldsymbol{Re}\left\{\left\langle \Lambda^+, -i\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c^3}P^+\Delta c\right\rangle\right\} + \boldsymbol{Re}\left\{\left\langle \Lambda^-, i\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c^3}P^-\Delta c\right\rangle\right\} . \tag{G.64}$$

The wavespeed is independent of the angular frequency, thus, we rearrange this expression and obtain

$$\boxed{\Delta_c L = \int_\Omega \boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} i\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c^3}\left[P^-\left(\Lambda^-\right)^* - P^+\left(\Lambda^+\right)^*\right]d\omega\right\}\Delta c\, d\mathbf{x}\,.} \tag{G.65}$$

In the same manner as in the reduced Lagrangian developments, Section G.3.1, we have the identity

$$\frac{\partial L(\bar{P}^+, \bar{P}^-, \hat{\mathcal{R}}_c^+, c, \bar{\Lambda}^+, \bar{\Lambda}^-)}{\partial c} = \frac{\partial E(\bar{P}^-)}{\partial c}\,. \tag{G.66}$$

Finally, considering that more experiments are available, i.e., shot-gathers, from equations G.65 and G.66, we conclude that the misfit-function partial derivative with respect to the wavespeed is

$$\boxed{\frac{\partial E(\bar{P}^-)}{\partial c} = \sum_{s=1}^{N_s}\boldsymbol{Re}\left\{\int_{\omega_i}^{\omega_f} i\hat{\mathcal{H}}_1^{-1}\frac{\omega^2}{c^3}\left[\bar{P}_s^-\left(\bar{\Lambda}_s^-\right)^* - \bar{P}_s^+\left(\bar{\Lambda}_s^+\right)^*\right]d\omega\right\}\,.} \tag{G.67}$$

G.4.1 Implementation of the square-root inverse

The forward modeling equation for an upgoing wavefield P^- is

$$\frac{\partial P^-}{\partial z} = i\hat{\mathcal{H}}_1 P^- - \hat{\mathcal{R}}_c^+\left(P^+ - P^-\right)\,. \tag{G.68}$$

The square-root operator is defined such that a twofold application provides the transversal Helmholtz operator, i.e., $\hat{\mathcal{H}}_1\hat{\mathcal{H}}_1 = \hat{\mathcal{H}}_2$. This way, multiplying equation G.68 by $\hat{\mathcal{H}}_2^{-1}$, we obtain

$$\hat{\mathcal{H}}_2^{-1}\left[\frac{\partial P^-}{\partial z} + \hat{\mathcal{R}}_c^+\left(P^+ - P^-\right)\right] = i\hat{\mathcal{H}}_1^{-1}P^-\,. \tag{G.69}$$

We want to calculate the quantity $i\hat{\mathcal{H}}_1^{-1}P^-$, labeling it as

$$Y = i\hat{\mathcal{H}}_1^{-1}P^-\,. \tag{G.70}$$

Substituting in equation G.69 and rearranging the terms, we obtain

$$\frac{\partial P^-}{\partial z} + \hat{\mathcal{R}}_c^+\left(P^+ - P^-\right) = \hat{\mathcal{H}}_2 Y\,. \tag{G.71}$$

In the two-dimensional and constant mass-density case, the transversal Helmholtz operator is

$$\hat{\mathcal{H}}_2 = \frac{\partial^2}{\partial x^2} + \frac{\omega^2}{c^2}\,. \tag{G.72}$$

Hence, we need to solve equation G.71 for Y to obtain the action of square-root inverse, scaled by the imaginary unit, on the wavefield P^-. Discretizing and considering the angle-independent reflection coefficient, we obtain

$$\left(\frac{D_x^2}{\Delta x^2} + \frac{\omega^2}{c^2}\right) Y = \frac{D_z}{\Delta z} P^- + \frac{R^+}{\Delta z}\left(P^+ - P^-\right)\,, \tag{G.73}$$

where D_x^2 is the centered second-derivative in the x coordinate and D_z the forward derivative in the z coordinate.

References

Amazonas, D., R. Aleixo, G. Melo, J. Schleicher, A. Novais, and J. C. Costa, 2010, Including lateral velocity variations into true-amplitude common-shot wave-equation migration: Geophysics, **75**, S175–S186, doi: 10.1190/1.3481469.

Amazonas, D., J. C. Costa, J. Schleicher, and R. Pestana, 2007, Wide-angle FD and FFD migration using complex padé approximations: Geophysics, **72**, S215–S220, doi: 10.1190/1.2785813.

Askan, A., V. Akcelik, J. Bielak, and O. Ghattas, 2007, Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures: Bulletin of the Seismological Society of America, **97**, 1990–2008, doi: 10.1785/0120070079.

Assis, C. A. M., Y. Pan, J. Schleicher, T. Bohlen, and H. B. Santos, 2019a, Joint migration inversion continuous equations and discretized solution via multiparameter Gauss-Newton method: SEG Technical Program Expanded Abstracts 2019, Society of Exploration Geophysicists, 5115–5119, doi: 10.1190/segam2019-3216644.1.

Assis, C. A. M., H. B. Santos, and J. Schleicher, 2019b, Colored and linear inversions to relative acoustic impedance: Geophysics, **84**, N15–N27, doi: 10.1190/geo2018-0185.1.

Avseth, P., T. Mukerji, and G. Mavko, 2005, Quantitative seismic interpretation: Cambridge University Press.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, doi: 10.1190/1.1441434.

Berkhout, A. G., 2014a, Review paper: An outlook on the future of seismic imaging, Part I: forward and reverse modelling: Geophysical Prospecting, **62**, 911–930, doi: 10.1111/1365-2478.12161.

———, 2014b, Review paper: An outlook on the future of seismic imaging, Part III: Joint migration inversion: Geophysical Prospecting, **62**, 950–971, doi: 10.1111/1365-2478.12158.

———, 2016, Utilization of multiple scattering: the next big step forward in seismic imaging: Geophysical Prospecting, **65**, 106–145, doi: 10.1111/1365-2478.12395.

Berkhout, A. J., 2012, Combining full wavefield migration and full waveform inversion, a glance into the future of seismic imaging: Geophysics, **77**, S43–S50, doi:

10.1190/geo2011-0148.1.

Berkhout, A. J., and C. P. A. Wapenaar, 1989, One-way versions of the Kirchhoff integral: Geophysics, **54**, 460–467, doi: 10.1190/1.1442672.

Berteussen, K. A., and B. Ursin, 1983, Approximate computation of the acoustic impedance from seismic data: Geophysics, **48**, 1351–1358, doi: 10.1190/1.1441415.

Blangy, J. P., C. Schiott, O. Vejbaek, and D. Maguire, 2014, The value of 4D seismic: Has the promise been fulfilled?: SEG Technical Program Expanded Abstracts 2014, Society of Exploration Geophysicists, 2552–2557, doi: 10.1190/segam2014-1118.1.

Bleistein, N., 1984, Mathematical methods for wave phenomena: Elsevier.

Bleistein, N., J. W. Stockwell, and J. K. Cohen, 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer New York.

Bohlen, T., A. Kurzmann, D. Koehn, A. Przebindowska, and N. Nguyen, 2009, 2D Acoustic Full Waveform Tomography – Performance and Optimization: Presented at the 71st EAGE Conference and Exhibition incorporating SPE EUROPEC 2009, EAGE Publications BV, doi: 10.3997/2214-4609.201400012.

Bremmer, H., 1951, The W.K.B. approximation as the first term of a geometric-optical series: Communications on Pure and Applied Mathematics, **4**, 105–115, doi: 10.1002/cpa.3160040111.

Camargo, A. W., 2019, Analysis of the full waveform inversion using the augmented lagrangian method: PhD thesis, University of Campinas, São Paulo, Brazil, url: http://repositorio.unicamp.br/handle/REPOSIP/333384?locale=es.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481, doi: 10.1190/1.1440185.

———, 1985, Imaging the earth's interior: Blackwell Scientific Publications, London.

Constable, S. C., R. L. Parker, and C. G. Constable, 1987, Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data: Geophysics, **52**, 289–300, doi: 10.1190/1.1442303.

Costa, J. C., F. J. C. da Silva, E. N. S. Gomes, J. Schleicher, L. A. V. Melo, and D. Amazonas, 2008, Regularization in slope tomography: Geophysics, **73**, no. 5, VE39–VE47, doi: 10.1190/1.2967499.

da Costa Filho, C. A., G. A. Meles, and A. Curtis, 2017, Elastic internal multiple analysis and attenuation using marchenko and interferometric methods: Geophysics, **82**, Q1–Q12, doi: 10.1190/geo2016-0162.1.

Davydenko, M., and D. J. Verschuur, 2018, Including and using internal multiples in closed-loop imaging — field data examples: Geophysics, **83**, R297–R305, doi: 10.1190/geo2017-0533.1.

———, 2019, Joint imaging of angle-dependent reflectivity and estimation of the mi-

gration velocity model using multiple scattering: Geophysics, **84**, R859–R868, doi: 10.1190/geo2018-0637.1.

Fletcher, R. P., S. Archer, D. Nichols, and W. Mao, 2012, Inversion after depth imaging: Presented at the SEG Technical Program Expanded Abstracts 2012, Society of Exploration Geophysicists, doi: 10.1190/segam2012-0427.1.

Fong, D. C.-L., and M. Saunders, 2011, LSMR: An iterative algorithm for sparse least-squares problems: SIAM Journal on Scientific Computing, **33**, 2950–2971, doi: 10.1137/10079687x.

Foster, D. J., and C. C. Mosher, 1992, Suppression of multiple reflections using the radon transform: Geophysics, **57**, 386–395, doi: 10.1190/1.1443253.

Foster, M., 1975, Transmission Effects in the Continuous One-dimensional Seismic Model: Geophysical Journal of the Royal Astronomical Society, **42**, 519–527, doi: 10.1111/j.1365-246x.1975.tb05875.x.

Gallardo, L. A., and M. A. Meju, 2003, Characterization of heterogeneous near-surface materials by joint 2D inversion of DC resistivity and seismic data: Geophysical Research Letters, **30**, 1658–1661, doi: 10.1029/2003gl017370.

Hammad, H., and D. Verschuur, 2019, Slowness and reflection coefficient curves for laterally heterogeneous media: Presented at the 81st EAGE Conference and Exhibition 2019, EAGE Publications BV, doi: 10.3997/2214-4609.201901538.

Innanen, K. A., 2017, Time- and offset-domain internal multiple prediction with nonstationary parameters: Geophysics, **82**, V105–V116, doi: 10.1190/geo2016-0220.1.

Jannane, M., W. Beydoun, E. Crase, D. Cao, Z. Koren, E. Landa, M. Mendes, A. Pica, M. Noble, G. Roeth, S. Singh, R. Snieder, A. Tarantola, D. Trezeguet, and M. Xie, 1989, Wavelengths of earth structures that can be resolved from seismic reflection data: Geophysics, **54**, 906–910, doi: 10.1190/1.1442719.

Jones, I., 2010, Tutorial: Velocity estimation via ray-based tomography: First Break, **28**, 45–52, doi: 10.3997/1365-2397.2010006.

Jones, I. F., J. Singh, S. Greenwood, J. Chigbo, P. Cox, and C. Hawke, 2018, High-resolution impedance estimation using refraction and reflection FWI constraints: the Fortuna region, offshore Equatorial Guinea: First Break, **36**, 39–44, doi: 10.3997/1365-2397.2018006.

Kennett, B. L. N., 1974, Reflections, rays, and reverberations: Bulletin of the Seismological Society of America, **64**, 1685–1696.

Macedo, D. L., 2014, Scattering-based decomposition of sensitivity kernels of acoustic full waveform inversion: PhD thesis, University of Campinas, São Paulo, Brazil, url: http://repositorio.unicamp.br/jspui/handle/REPOSIP/265785.

Maciel, J., J. Costa, and D. Verschuur, 2015, Enhancing resolution in imaging-based veloc-

ity estimation using morphological operators: SEG Technical Program Expanded Abstracts 2015, Society of Exploration Geophysicists, 5228–5232, doi: 10.1190/segam2015-5900483.1.

Masaya, S., and D. J. E. Verschuur, 2018, Iterative reflectivity-constrained velocity estimation for seismic imaging: Geophysical Journal International, **214**, 1–13, doi: 10.1093/gji/ggy105.

Menke, W., 2018, Geophysical data analysis: Discrete inverse theory: Academic press.

Métivier, L., R. Brossier, S. Operto, and J. Virieux, 2017, Full waveform inversion and the truncated newton method: SIAM Review, **59**, 153–195, doi: 10.1137/16m1093239.

Mora, P., 1989, Inversion= migration+ tomography: Geophysics, **54**, 1575–1586, doi: 10.1190/1.1442625.

Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221, doi: 10.1190/1.1444517.

Pan, W., Y. Geng, and K. A. Innanen, 2018, Interparameter trade-off quantification and reduction in isotropic-elastic full-waveform inversion: synthetic experiments and hussar land data set application: Geophysical Journal International, **213**, 1305–1333, doi: 10.1093/gji/ggy037.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503, doi: 10.1111/j.1365-246x.2006.02978.x.

Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987, doi: 10.1111/j.1365-246x.2004.02282.x.

Pratt, G., C. Shin, and Hicks, 1998, Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362, doi: 10.1046/j.1365-246x.1998.00498.x.

Robinson, E. A., 1986, Migration of seismic data by the WKBJ method: Proceedings of the IEEE, **74**, 428–439, doi: 10.1109/PROC.1986.13484.

Rosa, A. L. R., 2018, The seismic signal and its meaning: Society of Exploration Geophysicists.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis. i. theory: Geophysical Prospecting, **52**, 593–606, doi: 10.1111/j.1365-2478.2004.00447.x.

Sava, P., and S. J. Hill, 2009, Overview and classification of wavefield seismic imaging methods: The Leading Edge, **28**, 170–183, doi: 10.1190/1.3086052.

Schleicher, J., J. C. Costa, and A. Novais, 2008, A comparison of imaging conditions for wave-equation shot-profile migration: Geophysics, **73**, S219–S227, doi: 10.1190/1.2976776.

Schleicher, J., M. Tygel, and P. Hubral, 2007, Seismic true-amplitude imaging: Society of Exploration Geophysicists.

Shuey, R. T., 1985, A simplification of the Zoeppritz equations: Geophysics, **50**, 609–614, doi: 10.1190/1.1441936.

Soubaras, R., and B. Gratacos, 2017, Mitigating the gradient artefacts of migration velocity analysis by gauss-newton update: Presented at the 79th EAGE Conference and Exhibition 2017 - Workshops, EAGE Publications BV, doi: 10.3997/2214-4609.201701718.

Staal, X. R., and D. J. Verschuur, 2012a, Velocity estimation using internal multiples: SEG Technical Program Expanded Abstracts 2012, Society of Exploration Geophysicists, 1–5, doi: 10.1190/segam2012-1580.1.

———, 2012b, Velocity Model Estimation by Full Waveform Tomography of Seismic Reflection Data: Presented at the 74th EAGE Conference and Exhibition incorporating EUROPEC 2012, EAGE Publications, doi: 10.3997/2214-4609.20148590.

Tygel, M., J. Schleicher, and P. Hubral, 1994, Pulse distortion in depth migration: Geophysics, **59**, 1561–1569, doi: 10.1190/1.1443545.

Ursin, B., Ø. Pedersen, and B. Arntsen, 2012, Flux-normalized wavefield decomposition and migration of seismic data: Geophysics, **77**, S83–S92, doi: 10.1190/geo2011-0234.1.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, 1166–1177, doi: 10.1190/1.1443330.

Verschuur, D. J., X. R. Staal, and A. J. Berkhout, 2016, Joint migration inversion: Simultaneous determination of velocity fields and depth images using all orders of scattering: The Leading Edge, **35**, 1037–1046, doi: 10.1190/tle35121037.1.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26, doi: 10.1190/1.3238367.

Vogel, C. R., 2002, Computational methods for inverse problems: Siam, **23**.

Wapenaar, C., and A. J. Berkhout, 1989, Elastic wave field extrapolation: Redatuming of single- and multi-component seismic data (advances in exploration geophysics): Elsevier Science.

Wapenaar, C. P. A., 1990, Representation of seismic sources in the one-way wave equations: Geophysics, **55**, 786–790, doi: 10.1190/1.1442892.

———, 1996, One-way representations of seismic data: Geophysical Journal International, **127**, 178–188, doi: 10.1111/j.1365-246x.1996.tb01543.x.

Wapenaar, C. P. A., and J. L. T. Grimbergen, 1996, Reciprocity theorems for one-way wavefields: Geophysical Journal International, **127**, 169–177, doi: 10.1111/j.1365-246x.1996.tb01542.x.

Wapenaar, K., 1998, Reciprocity properties of one-way propagators: Geophysics, **63**, 1795–1798, doi: 10.1190/1.1444473.

Weglein, A. B., 2016, Multiples: Signal or noise?: Geophysics, **81**, V283–V302, doi: 10.1190/geo2014-0486.1.

Williamson, P., A. Atle, W. Fei, and D. Hale, 2011, Regularization of wave-equation migration velocity analysis by structure-oriented smoothing: SEG Technical Program Expanded Abstracts 2011, Society of Exploration Geophysicists, 3877–3881, doi: 10.1190/1.3628015.

Xu, S., D. Wang, F. Chen, Y. Zhang, and G. Lambare, 2012, Full waveform inversion for reflected seismic data: Presented at the 74th EAGE Conference and Exhibition incorporating EUROPEC 2012, doi: 10.3997/2214-4609.20148725.

Yilmaz, O., 2001, Seismic data analysis: Society of Exploration Geophysicists.

Zhang, Y., G. Zhang, and N. Bleistein, 2005, Theory of true-amplitude one-way wave equations and true-amplitude common-shot migration: Geophysics, **70**, E1–E10, doi: 10.1190/1.1988182.

Zhdanov, M. S., 2015, Ill-posed problems and the methods of their solution, *in* Inverse Theory and Applications in Geophysics: Elsevier, 33–61.

Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2015, Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation: Geophysical Journal International, **202**, 1535–1554, doi: 10.1093/gji/ggv228.