

IMAGING SEISMIC REFLECTIONS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties,
in het openbaar te verdedigen
op woensdag 6 april 2011 om 12:45 uur

door

Timotheus Johannes Petrus Maria Op 't Root

geboren op 8 november 1976
te Nederweert.

Dit proefschrift is goedgekeurd door promotor
Prof. dr. ir. E.W.C. van Groesen
en assistent-promotor
Dr. C.C. Stolk.

© 2011 Tim Op 't Root
Wöhrmann Print Service
Kaftontwerp: Sjoukje Schoustra

ISBN 978-90-365-3150-4
DOI 10.3990/1.9789036531504

Samenstelling promotiecommissie:

Voorzitter en secretaris

Prof. dr. ir. A. J. Mouthaan Universiteit Twente

Promotor

Prof. dr. ir. E.W.C. van Groesen Universiteit Twente

Assistent-promotor

Dr. C. C. Stolk Universiteit van Amsterdam

Leden

Prof. dr. S. A. van Gils Universiteit Twente

Prof. dr. ir. A. de Boer Universiteit Twente

Prof. dr. M. V. de Hoop Purdue University

Dr. ir. D. J. Verschuur Technische Universiteit Delft

Dr. R. G. Hanea Technische Universiteit Delft

Het in dit proefschrift beschreven onderzoek werd uitgevoerd bij de groep Toegepaste Analyse en Mathematische Fysica binnen de afdeling Toegepaste Wiskunde van de Universiteit Twente.

Het onderzoek is gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek middels de VIDI-subsidie 639.032.509.

Contents

1	Looking into the Earth	1
2	Reflection seismic imaging	5
2.1	Introduction	5
2.2	Seismic inverse problem	6
2.3	Wave modeling	10
2.4	Seismic imaging	13
2.4.1	Imaging with reflections	13
2.4.2	Reverse-time migration	14
2.4.3	Our contribution	15
2.5	Microlocal analysis	17
2.5.1	Wave front set	17
2.5.2	Propagation of singularities	20
2.6	Outline	22
3	One-way wave equation with symmetric square root	25
3.1	Introduction	26
3.2	One-way wave equation	30
3.2.1	One-way wave decomposition	31
3.2.2	Pseudo-differential operators	32
3.2.3	Square root operator and symmetric quantization	34
3.2.4	Normalization	36
3.3	Numerical implementation	39
3.3.1	Pseudo-spectral interpolation method	40
3.3.2	Absorbing boundaries and stability	43
3.3.3	Symmetric finite-difference one-way	45
3.4	Numerical results	47
3.5	Discussion	51

4	Inverse scattering based on reverse-time migration	55
4.1	Introduction	55
4.2	Asymptotic solutions of initial value problem	58
4.2.1	WKB approximation with plane-wave initial values	58
4.2.2	The phase function on characteristics	60
4.2.3	The amplitude function	61
4.2.4	Solution operator as a FIO	63
4.2.5	Solution operators and decoupling	64
4.2.6	The source field	66
4.3	Forward scattering problem	68
4.3.1	Continued scattered wave field	68
4.3.2	Scattering operator as FIO	71
4.4	Reverse time continuation from the boundary	76
4.5	Inverse scattering	84
4.5.1	Constant background velocity	84
4.5.2	Imaging condition	87
4.6	Numerical examples	94
4.7	Discussion	95
5	Conclusion	99
	Bibliography	101
	Acknowledgments	107
	Summary	108
	Samenvatting	109
	About the author	110

Chapter 1

Looking into the Earth

When we talk about seeing, we usually think about seeing with the eyes. Our eyes are sensitive to light, a wave phenomenon. Another wave phenomenon with which we can ‘see’, is sound. Sound is a sequence of pressure or velocity disturbances of a compressible media such as air that propagates through the medium. Besides air these disturbances can also propagate through a solid medium like the Earth, in which case one speaks of *seismic waves*. And these seismic waves are being used to look at the inside of the Earth. ‘Seeing by listening,’ so to say.

Before we dive into imaging techniques to observe the Earth, we consider another example of the application of sound waves to create an image. Ultrasound is a technique that uses sound waves that propagate through the human body, which makes it possible to visualize the inside. So they can, for example, gain insight into the size, structure and pathology of deviations. It is also being used to visualize an unborn child. Figure 1.1 shows an ultrasound image of an arm and hand of a nineteen week old fetus. The technique relies on the fact that a portion of the waves is reflected at the interfaces between soft and harder tissues. These reflections are detected by an instrument and translated into an image.

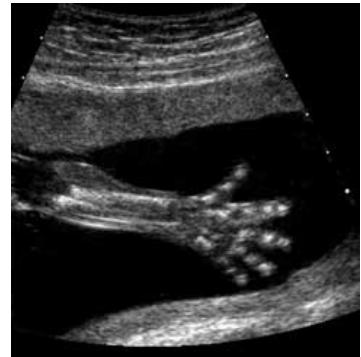


Figure 1.1: Ultrasound image.

We will use the principle of Huygens to explain the wave phenomenon. Huygens’ principle is a notion that goes back to the classical *Traité de la Lumière* by Christiaan Huygens, published in 1690. It gives a geometric description of the propagation of light. The principle states that the wave front of a propagating

wave at any instant is formed by the envelope of spherical waves that emanate from every point on the wave front at a prior instant. This is illustrated in figure 1.2. All disturbances propagate with the same velocity that is characteristic for the medium and constant for light through air. For seismic waves the velocity depends on the position in the Earth. We will see that this is a challenge for the application of seismic imaging.

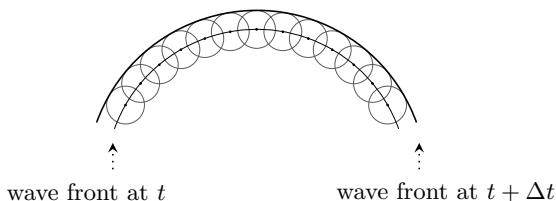


Figure 1.2: Huygens' principle. The wave front at time $t + \Delta t$ is the envelope of the spherical wave fronts emanating from every point of the wave front at time t .

Reflection seismic imaging is the activity of making images of the Earth's interior based on surface measurements of seismic reflections [14, 7, 66]. It is an important tool used by the oil and gas industry to map petroleum deposits in the Earth's upper crust. The method is also being used in shallow depths for engineering, ground water, and environmental studies. In oil exploration the seismic reflection method provides images of formations in consolidated sedimentary rocks at depths in the range from 0.1 to 10 km, using frequencies ranging from 10 to 100 Hz. They involve basically the same type of measurements as in earthquake seismology. However, the energy sources are controlled and movable, and the distance between source and receiver is relatively small. Surveys are performed both on land and on sea.

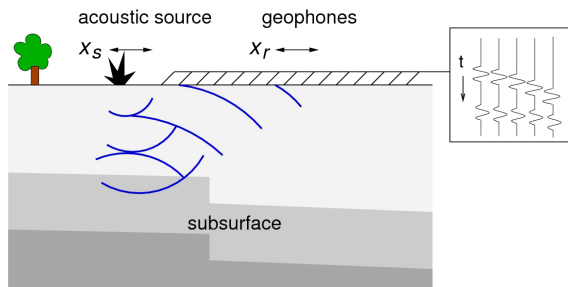


Figure 1.3: Seismic reflection experiment. Data consist of the signals recorded by the geophones. The object is to make an image of the subsurface reflecting features.

Basically, the technique consists of generating seismic waves on or near the Earth's surface and measuring waves that arrive at the surface after being reflected. Those reflections occur at positions where the gradient of physical properties is large, for instance at interfaces between different formations. Figure 1.3 depicts a typical reflection experiment. The seismic waves are usually generated by an explosion, a mechanical impact, or a vibration. The reflections are recorded by detecting instruments, the receivers, which respond to ground motion or pressure. Figure 1.4 shows an example of recorded signals in which the time variable on the vertical axis is presented as a depth coordinate. In reflection surveys the receivers are placed in the vicinity of the point of generation, the source, at distances less than the depth of the assumed reflectors. By noting the time it takes for a reflection to arrive at a receiver, the travel time, it is possible to estimate the depth of the feature that generated the reflection. And the strength of a reflected wave says something about the contrast of the reflection feature in the subsurface.

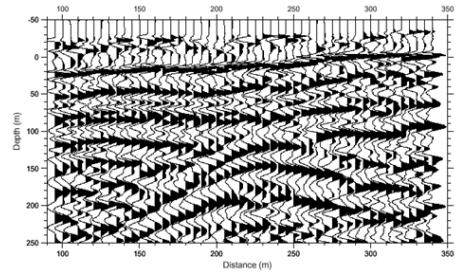


Figure 1.4: Example of data.

The research presented in this thesis concerns techniques from reflection seismic imaging. There has been a great development of techniques that create structural images. Such an image shows the correct position and orientation of underground layer boundaries. In figure 1.5 an example of a seismic image of a salt dome is shown. One can clearly discriminate the subsurface layers in the surrounding of the salt dome. However, is still a challenge to image the amplitude of the reflecting features. By amplitude is meant the size of the subsurface heterogeneities that caused the reflection. This information is valuable for the geological interpretation of the image. Our work provides a mathematical contribution to the development of imaging techniques. In particular, it offers improvements to the amplitude behaviour of the images.

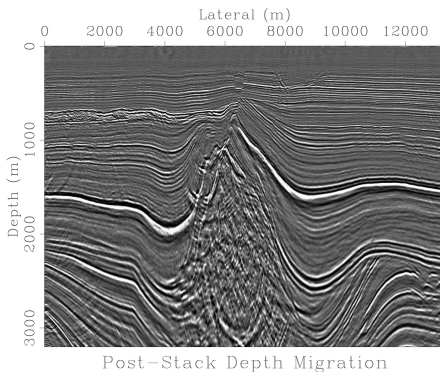


Figure 1.5: Image of a salt dome.

provides a mathematical contribution to the development of imaging techniques. In particular, it offers improvements to the amplitude behaviour of the images.

Chapter 2

Reflection seismic imaging

2.1 Introduction

Migration is the practice of exploration seismology that translates surface measurements of reflected seismic waves into an image of the subsurface. In seismic imaging two schools can be distinguished. From the practical point of view geophysicists are in the first place interested in ways to reveal structures of the Earth's subsurface. Many migration methods developed in exploration seismology were originally presented as procedures to make an image of the subsurface. Such images display where reflections have occurred without identifying the value of the physical quantity. Many of these methods have shown to be empirically evident, although they are often based on a heuristic derivation.

Roughly since the eighties, the migration problem has been formulated as a mathematical inverse problem. These theoretical approaches lead to a direct estimation of the medium parameter perturbations. We mention a few that are relevant to our work. Beylkin (1985) treated migration as a linearized inverse problem and used the generalized inverse Radon transform [4]. It relies on the assumption that the medium parameter, the velocity in this case, is a sum of a long-scale background velocity and a short-scale perturbation. Not much later Rakesh investigated the linearized inverse problem, in particular the continuity, uniqueness and invertibility [54]. Nolan and Symes (1997) examined the conditions that are required to reconstruct the short-scale velocity perturbation and addressed the issue of imaging when the conditions are not satisfied [50]. Also Ten Kroode et al. should be mentioned, who studied the generalization of Beylkin's work [44]. Most of these studies, however, do not offer a migration algorithm so there is still a demand for.

A problem that is still the subject of many studies concerns the amplitude of the image [52, 77, 43]. In the early days, migration methods were capable of

determining the location and attitude of interfaces that give rise to reflections [14, 58]. In this context one speaks of a structural image. For a correct geological interpretation however, the image should give an estimate of the medium parameter perturbation. That is meant when we refer to the amplitude.

The primary objective of the thesis is to improve two methods or techniques of reflection seismic imaging in order to describe the amplitude more accurately. We consider the problem as a linearized inverse problem. Though the approach is formal, the improvements have practical applicable implications. In section 2.2 we will first describe the mathematical inverse problem. Two subproblems can be identified, namely the modeling of seismic waves and the imaging of velocity perturbations. In the wave modeling problem, the amplitude refers to the magnitude of the acoustic wave. In the imaging problem, the amplitude refers to the size of the velocity perturbation.

The *one-way wave equation* is a 1st-order equation that can be derived from the wave equation and describes wave propagation in a predetermined direction. We will derive the one-way wave equation with a symmetric square root operator and show that it correctly describes the wave amplitude. This will be explained and motivated in section 2.3. The main results can be found in chapter 3.

Reverse-time migration is an imaging method that uses simulations of the source and receiver wave fields. The receiver wave field is an in reverse time propagated field that matches the measurements. An *imaging condition* subsequently transforms both fields to an image of the subsurface. Besides the fact that the waves should be calculated with correct amplitude, a persistent difficulty lies in the imaging condition. We present an imaging condition of which we prove that it gives a reconstruction of the velocity perturbation with correct amplitude, the main theorem of chapter 4. We will further explain and motivate this in section 2.4.

The one-way wave equation is an application of *pseudo-differential operators* (PsDO) and the study of reverse-time migration extensively uses *Fourier integral operators* (FIO). These operators are developed to formulate solutions of partial differential equations (PDE) and belong to, what is called, *microlocal analysis*. The very basics of PsDOs and FIOs is included in the chapters where they are used, i.e. 3 and 4, and a more intuitive introduction is included in section 2.3 and 2.4. An important object and a result from microlocal analysis are the *wave front set* and the propagation of singularities. These are not covered in mentioned sections and will hence be explained in section 2.5. The intention is to give some extra preparation, not to be exhaustive.

2.2 Seismic inverse problem

Seismic wave propagation, the physical phenomenon underlying the reflection experiment, is modeled with reasonable accuracy as small-amplitude displacement of

a continuum, the regime of elastodynamics. And many current imaging methods rely on a yet even further simplification, namely the acoustic equation with constant density [11]. Scalar acoustics, the theory of sound propagation through a liquid or gas, can be seen as a special case of elastodynamics. This will be shown at the end of the section, with equation (2.9) as result.

We first describe the model that is used for the reflection experiment. The spatial domain is \mathbb{R}^n with $n = 2, 3$ in which x_n is the depth coordinate. The subsurface is hence represented by $x_n > 0$, the surface by \mathbf{x}' . The medium velocity is the material property related with the wave propagation phenomenon and appears as the spatially varying coefficient $c(\mathbf{x})$ in the acoustic wave equation, a time-dependent hyperbolic PDE:

$$\left[\frac{1}{c(\mathbf{x})^2} \partial_t^2 - \Delta \right] u(\mathbf{x}, t) = f(\mathbf{x}, t). \quad (2.1)$$

The field is initially at rest, so $u(\cdot, 0)$ and $\partial_t u(\cdot, 0)$ are both zero. In accordance with our study in chapter 4, we model a single-shot experiment in which the source f is a sharply defined pulse in spacetime, positioned at the surface. Data are samples of time series of the pressure u or a related quantity collected at points corresponding to locations of the receivers. In the model they are assumed to continuously cover the acquisition domain and so impose a condition on the solution of (2.1):

$$u(\mathbf{x}', t) = d(\mathbf{x}', t), \quad (2.2)$$

in which \mathbf{x}' is in the acquisition domain.

The problem of reflection seismology can hence be stated as an inverse problem, in which one aims to find the coefficient function $c(\mathbf{x})$ such that the solution of the wave equation (2.1) satisfies the condition of the data (2.2). This is a difficult nonlinear problem, and like Beylkin, Rakesh and many others, we will work with a linearized problem. The velocity is considered as a sum of a large-scale background velocity and a short-scale perturbation [4, 54, 50]. The idea behind the scale-separation is that only the short-scale velocity gives rise to reflections. This models the single scattered waves, i.e. waves that have reflected once, as multiples do not depend linearly on the medium perturbation. The slowly varying velocity, in practice determined by what is called velocity analysis, is given in our work.

From practical point of view it is favorable to model the relative perturbation of the velocity. Denoting the slowly varying velocity by c and the 1st-order relative perturbation by r , the velocity becomes $(1 + r(\mathbf{x}))c(\mathbf{x})$. The support of the *reflectivity* r is in the subsurface and assumed to be at some distance from the surface. We model the source by a delta pulse in the origin and reuse equation (2.1) with $f = \delta(\mathbf{x}, t)$ to formulate the unperturbed problem. The reflected wave field u_r is the 1st-order perturbation and is governed by the initial value problem (IVP):

$$\begin{aligned} [\partial_t^2 - c(\mathbf{x})^2 \Delta] u_r(\mathbf{x}, t) &= 2r(\mathbf{x}) \partial_t^2 u(\mathbf{x}, t) \\ u_r(\cdot, 0) &= 0, \quad \partial_t u_r(\cdot, 0) = 0. \end{aligned} \quad (2.3)$$

The objective of the linearized inverse problem is to find the velocity perturbation r such that the solution of problem (2.3), the reflected wave u_r , satisfies the condition imposed by the data:

$$u_r(\mathbf{x}', t) = \tilde{d}(\mathbf{x}', t). \quad (2.4)$$

The data of the linearized problem \tilde{d} do not contain the unperturbed or direct waves, neither do they contain multiples. This is not trivial, as the distinction between direct and reflected waves is a priori theoretical. However, there exist several techniques for this, for instance, a family of methods that discriminate on the basis of travel time.

As pointed out in the introduction, two subproblems can be identified. The inverse problem involves the modeling of acoustic waves through slowly varying inhomogeneous media. This is often done with the one-way wave equation, which is the subject of section 2.3. The second problem concerns the inversion method, how to make an image of the small-scale perturbation. This is covered in section 2.4.

Acoustic wave equation

In reflection seismology the Earth is modeled as a mechanical continuum. The propagation of seismic waves is modeled by linear elastodynamics [66]. And often the acoustic wave equation is used, like we do. Here we present a short review of this equation as a special case of linear elastodynamics and address the underlying assumptions.

In the theory of elasticity internal forces are described by the stress tensor σ [46, 58]. Stress is defined as force per unit area. If the force is perpendicular to the area, the stress is said to be normal stress, or pressure. When the force is tangential to the area, the stress is a shearing stress. Any stress can be resolved into a normal and a shearing component. Subjected to stress a body will change in shape and dimension. The deformation is described by the strain tensor ε . Strain is the relative change in a dimension or shape of a body element [58].

In linear elasticity the displacement \mathbf{u} is small and the relation between the stress and strain tensors, the constitutive relation, is linear. We assume the continuum to be isotropic, i.e. when properties do not depend upon direction. In that case the relation can be expressed as [46, 58]¹

$$\sigma_{ik} = \lambda \text{tr}(\varepsilon) \delta_{ik} + 2\mu \varepsilon_{ik}, \quad (2.5)$$

in which $i, k = 1, 2, 3$. The quantities λ and μ are known as the Lamé constants. The trace $\text{tr}(\varepsilon)$ is the change in volume per unit volume. If it is zero the deformation

¹We use the definition of strain tensor from Landau and Lifshitz [46], i.e. $\varepsilon_{ik} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right)$ if u_i is the displacement. The definition of Sheriff and Geldart [58] differs a factor 2 for the off-diagonals.

is called shear strain. Strain can be decomposed in a normal and a shear component by writing the constitutive relation as the sum

$$\sigma_{ik} = (\lambda + \frac{2}{3}\mu) \text{tr}(\varepsilon)\delta_{ik} + 2\mu(\varepsilon_{ik} - \frac{1}{3} \text{tr}(\varepsilon)\delta_{ik}). \quad (2.6)$$

The first term describes hydrostatic compression, the second term shear strain. In hydrostatic compression an element does not change shape, only in dimension. The modulus of compression, defined as $\kappa = \lambda + \frac{2}{3}\mu$, can be interpreted as the resistance against hydrostatic compression. Shear modulus μ is resistance against shear strain.

Dynamic equations describe the displacement within the linear elastic regime changing in time. If the density of the continuum is denoted by ρ then the application of the second law of Newton yields

$$\rho \partial_t^2 \mathbf{u} = \nabla \cdot \boldsymbol{\sigma}. \quad (2.7)$$

The right hand side is the unbalanced force per unit volume [46, 58]. Equations (2.6) and (2.7) together describe the dynamics of an isotropic elastic medium. Such media support wave phenomena, and it can be shown that there exists two types of waves, compression and shear waves. In a shear wave the displacement is perpendicular to the direction of motion of the wave. Shear waves hence are transverse waves. In a compression wave the disturbance is a compression of the medium. These waves are longitudinal waves.

Earthquakes can produce both kinds of waves. Each kind moves through the interior of the Earth to be detected by seismographs on the other side of the Earth. However, the compression waves travel faster than the shear waves. Since the compression waves outdistance the shear waves and arrive first at distant earthquake detectors, geologists call them primary waves or p-waves. The shear waves are called secondary waves or s-waves [58].

We will show how to obtain the wave equation for compression waves under the simplifying assumption that the continuum does not support shear stress, i.e. $\mu = 0$, the regime of linear acoustics [66]. The continuum then behaves like a fluid or a gas. The stress tensor becomes scalar, $\sigma_{ik} = -p\delta_{ik}$, with p being the pressure. The relative change in volume is $\text{tr}(\varepsilon) = \nabla \cdot \mathbf{u}$ by definition. Using the particle velocity $\mathbf{v} = \partial_t \mathbf{u}$, which is common practice, the equation of momentum (2.7) and the constitutive relation (2.6) respectively collapse to

$$\begin{aligned} \rho \partial_t \mathbf{v} &= -\nabla p + \mathbf{f} \\ -\partial_t p &= \kappa \nabla \cdot \mathbf{v}. \end{aligned} \quad (2.8)$$

We added \mathbf{f} , a force per unit volume that represents external energy input from a source. These equations can also be found in Chapman's book about seismic wave propagation [11].

The modulus of compression κ and density ρ are material parameters that vary in space, though we assume the density to be constant. Supporting argumentation can be found in measurements of material properties made in a borehole, so called wel logs [66]. These showed that the wave velocity, defined as $c = \sqrt{\frac{\kappa}{\rho}}$, much stronger varies with depth than the density. The physical reason for this is that gravitational compression tends to increase the stiffness of the rock with depth, but it is not sufficient to overcome the intermolecular forces to change the density substantially [66].

The external force is assumed to be applied as a pressure, which implies that it is irrotational. From the first equation of (2.8) then follows that $\nabla \times \mathbf{v}$ is constant in time, and hence zero as we assume the medium initially at rest. Then, there exists a potential function ϕ such that $\mathbf{v} = -\nabla\phi$. The system (2.8) therefore reads two 1st-order equations in two scalar quantities, i.e. p and ϕ . By taking the divergence of the first equation and the time derivative of the second, then under elimination of $\nabla \cdot \mathbf{v}$ one derives the 2nd-order equation for the pressure

$$\left[\frac{1}{c(\mathbf{x})^2} \partial_t^2 - \Delta\right] p(\mathbf{x}, t) = f(\mathbf{x}, t). \quad (2.9)$$

We wrote the scalar source term as $f = -\nabla \cdot \mathbf{f}$. This is the wave equation that we will use to model seismic wave propagation through inhomogeneous media. Seismic experiments use transient energy sources, i.e. short-duration signals, that are localized in space. As the time extent of the seismic experiment is limited, the domain for the wave equation (2.9) is essentially unbounded [66]. In our work we will adopt an unbounded domain $\mathbf{x} \in \mathbb{R}^n$ and assume the pressure field initially at rest, i.e. $p(\cdot, t_0) = \partial_t p(\cdot, t_0) = 0$ for some time t_0 preceding the experiment.

The reason why we, and many others within seismic imaging, use the acoustic equation is that its wave propagation shows a strong resemblance with the propagation of p-waves in an elastic medium. Though there are differences, the acoustic approach is widely used in seismic processing and has proven to be very useful.

2.3 Wave modeling

One of the underlying problems in the thesis deals with the propagation of waves through slowly varying inhomogeneous media. Slowly varying means that the velocity only mildly changes in space. In such a medium solutions of the wave equation can be approximated by oscillatory functions, e.g.

$$u(\mathbf{x}, t) = a(\mathbf{x}, t) e^{i\psi(\mathbf{x}, t)}, \quad (2.10)$$

with phase function ψ . Two physical phenomena that affect the amplitude a are geometrical spreading and refraction. Refraction is the change in direction, and

as a result a change of the amplitude, of a wave due to spatial variations of the medium. Hence it is reasonable to assume that the amplitude only mildly varies as function of space, and time, as the propagation velocity is finite.

The assumption has the implication that a differentiation approximately becomes a multiplication, e.g.

$$\partial_{\mathbf{x}}u \approx i(\partial_{\mathbf{x}}\psi)u,$$

in which we used that $\partial_{\mathbf{x}}a \ll (\partial_{\mathbf{x}}\psi)a$. As the frequency spectrum of the waves under study is narrow, the value of $\partial_{\mathbf{x}}\psi$ can be estimated by the wave vector $\boldsymbol{\xi}$. The assumption hence means that $\frac{1}{a}\partial_{\mathbf{x}}a \ll \boldsymbol{\xi}$, the relative change of the amplitude is small compared to the absolute value of the wave vector. The wave vector is related to the frequency ω by the dispersion relation, i.e. $c|\boldsymbol{\xi}| = |\omega|$. The slowly varying medium assumption can therefore be rephrased in a high frequency wave assumption. In accordance with the theory of pseudo-differential and Fourier integral operators we will use oscillatory functions with a high frequency.

For the ansatz $a e^{i\psi}$ to be a solution of the wave equation (2.1), the phase and amplitude must solve the eikonal and first transport equation respectively. The phase determines the geometry of the propagation, the regime of geometrical optics. The level set of the phase function, e.g.

$$\{(\mathbf{x}, t) \mid \psi(\mathbf{x}, t) = \psi_0\}, \quad (2.11)$$

is a way to describe the wave front of a propagating disturbance. A ray is an imaginary curve, perpendicular to the wave front, that shows the direction of propagation. The first transport equation gives a principal order approximation of the wave amplitude, by which is meant that the relative error is of the order of $\frac{1}{\omega}$. This is the leading principle in our work about the one-way wave equation, chapter 3, and in the modeling of waves as part of reverse-time migration in chapter 4.

The one-way wave equation

One particular technique from seismic wave modeling that will be subjected to a detailed analysis is the one-way wave equation. Though the medium inhomogeneity is essential, for the moment we assume that the velocity c is constant and define $A = -\frac{1}{c^2}\partial_t^2 + \partial_x^2$. We consider \mathbb{R}^2 with lateral coordinate x and depth z . Assume that we have a *square root operator* B , of the operator A . Then the wave equation (2.1) can be written as

$$[iB - \partial_z][iB + \partial_z]u = f. \quad (2.12)$$

Now observe that if u is a solution of the so called *one-way wave equation*

$$\partial_z u = -iB u, \quad (2.13)$$

then u also solves equation (2.12), albeit with $f = 0$. The square root operator is not unique, and in section 3.2 we argue that iB must be real, anti-symmetric and such that it becomes $\frac{1}{c}\partial_t$ in the special case of $\partial_x u = 0$. Equation (2.13) then describes wave propagation in the positive z -direction.

The one-way wave equation is widely used in seismic imaging for various reasons. One reason is that it has efficient numerical implementations [28]. Also, in wave field extrapolation it requires only one boundary condition, as it is a 1st-order equation, and it provides a way to propagate in a predetermined direction [74]. One-way wave equations were first used in geophysical imaging by Claerbout [14]. They were used to describe travel time and were not intended to describe wave amplitudes. Since roughly 1980 there has been an growing interest in techniques that not only provide structural images but also provide amplitudes. One of the problems of the one-way wave equation, regarding the amplitudes, is the formulation of the square root operator in case of inhomogeneous media.

To get an idea of how the square root can be constructed, we use the Fourier transform with respect to x and t , i.e. $\hat{u}(\xi, z, \omega) = \int e^{-i(\xi x + \omega t)} u(x, z, t) dx dt$. Differentiations become multiplications, $i\xi$ for ∂_x and $i\omega$ for ∂_t , and A becomes $\frac{\omega^2}{c^2} - \xi^2$. If we carry on with this idea, the one-way wave equation (2.13) comes to be

$$\partial_z \hat{u} = -i \frac{\omega}{c} \sqrt{1 - \frac{c^2 \xi^2}{\omega^2}} \hat{u}. \quad (2.14)$$

Such function of ξ and ω , and possibly x and t , is called the symbol of an operator. Denoting the wave vector by (ξ, ζ) , the dispersion relation is $c^2(\xi^2 + \zeta^2) = \omega^2$. Propagated modes hence satisfy $c^2 \xi^2 < \omega^2$, so the root singularity can be avoided. If we consider, for example, lateral invariant solutions, i.e. when ξ is zero, then the transport equation $[\partial_z + \frac{1}{c}\partial_t]u = 0$ appears. Rigid translations, i.e. functions of $z - ct$, are hence solutions of the one-way wave equation.

The problem starts when the velocity is non-constant. The $\partial_z c$ hampers the factorization that is used in (2.12) as the operators iB and ∂_z are no longer commutative. This leads to a numerically unattractive correction term in the one-way wave equation. Also the $\partial_x c$ puts a question mark above the definition of the square root in (2.14). This is where pseudo-differential operators (PsDO) offer a solution. A PsDO is an extension of the concept of differential operator, which we use to formalize the principal order approximation. The theory provides a way to associate a well-defined operator with the square root symbol in the right hand side of equation (2.14) allowing for $c = c(x, z)$. The thus obtained one-way wave equation corresponds to the WKB approximation.

The outline of the chapter about the one-way waves, chapter 3, can be found in section 2.6. We will first introduce the imaging problem and our approach to it.

2.4 Seismic imaging

We first discuss an easy problem that shows how seismic reflections can be used for the purpose of imaging. Then reverse-time migration, a widely used imaging technique, will be introduced. We conclude the section with an intuitive introduction of Fourier integral operators and explain our contribution to reverse-time migration.

2.4.1 Imaging with reflections

The subsurface inhomogeneity can be quite complex. For illustration however, we sketch the case of a simple reflection problem in \mathbb{R}^3 , with lateral coordinates (x, y) and depth z . The surface is $z = 0$, and two homogeneous layers meet at a mildly sloping interface at depth $z = h(x, y)$. The velocity of the upper layer, c_1 , is given. A delta source is at the surface. See figure 2.1. What will happen? The source wave, described by IVP (2.1) with $f = \delta$, propagates through the upper layer and ‘hits’ the interface. The collision gives rise to the emission of a scattered wave, described by the IVP (2.3). A moment later the scattered wave can be detected at the surface. Using these samples, what can now be concluded about the interface h ? And what about the velocity c_2 of the lower layer?

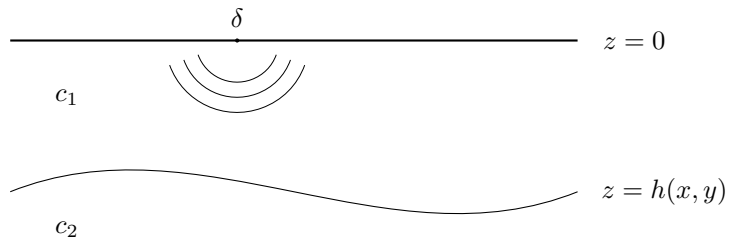


Figure 2.1: Reflection problem with two-layer velocity model.

We describe the source wave by an oscillatory function like (2.10), and first approach the problem geometrically. The source wave front, in the sense of level set (2.11), is spherical and grows in time. Each point of the interface h is reached at a certain time. In this example, we assume that the scattered wave can also be modeled by one oscillatory function. In general this is only true if the interface is ‘flat enough’. Then in a similar way, the wave front of the scattered wave reaches a point of the surface, say (x, y) , at a certain time, the *travel time* $\tau(x, y)$. Now it is sensible to use the travel time τ to reconstruct the interface depth h , as both are functions of two variables. The geometry of the waves hence captures the location of a subsurface reflector.

Besides the travel time of a detected wave, one can identify its amplitude. It is related to the size of the discontinuity at the interface, and hence in this example, to the velocity of the lower layer. More general, the measurements are a function of x , y and t , the data $\tilde{d}(\mathbf{x}', t)$, and the ‘subsurface reflectors’ are a function of x , y and z , namely the reflectivity $r(\mathbf{x})$. To be precise, the reflectors are represented by the *singularities* of the reflectivity. In section 2.5 we show how this is made formal by the so called *wave front set*. Still, the idea is to extend the geometric approach by an amplitude, sometimes called the geometric amplitude, to be able to recover the size of a discontinuity as much as possible.

2.4.2 Reverse-time migration

Geophysicists have developed a variety of migration methods to approximate solutions of the seismic reflection problem. As pointed out in section 2.1, many early developed methods can be interpreted as procedures that convert seismic data into an image of the Earth’s subsurface. Those methods are mainly based on travel time measurements and yield structural images showing reflecting interfaces, but whose amplitudes are unreliable [14, 58]. For example, Claerbout developed one-way wave equation based migration that properly describes travel times but whose geometric spreading effects were inaccurate [12]. During the last three decades, many approaches have been developed to derive images of the velocity perturbation with correct amplitudes [43]. Migration algorithms that preserve amplitudes are also called ‘true-amplitude’ migration [76].

The subject of our study in chapter 4 is *reverse-time migration*. It can be placed among three important migration methods that can be distinguished on the basis of how waves are modeled and how an image is formed. Reverse-time migration, like *Kirchhoff migration*, uses an evolution in time to model wave propagation through the slowly varying background. The latter uses an integral solution of the wave equation, based on Green’s function theory. The image is then expressed as a surface integral of seismic observations [7]. Reverse-time migration has a different approach and uses wave field simulations based on the 2nd-order wave equation [3, 71, 48, 14]. The source wave is propagated through the background in forward time. The data, assuming that the reflections have arrived at the surface as upward traveling waves, form the boundary condition of a receiver wave that is propagated in reverse time. An *imaging condition* thereafter transforms the source and receiver wave fields to an image of the subsurface. The imaging condition, which we further explain in the next paragraph, is central in chapter 4. The principles of *downward continuation migration*, the third method, are similar to those of reverse-time migration. The main difference is that downward continuation migration is based on the *one-way wave equation*, in which the wave fields are extrapolated along the depth axis instead of an evolution in time [7].

The imaging condition is a map that takes the source and receiver wave fields as input, and maps those to an image of the subsurface reflectors. It is based on Claerbout’s imaging principle [13]: A subsurface reflector exists where the source and receiver wave fields coincide in space and time. One imaging condition according to this principle, for example, is the zero-lag cross-correlation of the fields [12],

$$\text{image}(\mathbf{x}) = \sum_{\text{time}} u_s(\mathbf{x}, t) u_r(\mathbf{x}, t). \quad (2.15)$$

Here are u_s and u_r respectively the simulated source and receiver waves. Though the image correctly locates the reflectors, it does not provide a reconstruction of the velocity perturbation. More about this can be found in section 4.1. Many improved imaging conditions have been proposed, for instance the source-normalized cross-correlation condition [40, 12]. It normalizes the image (2.15) by the square of the source illumination strength, $\sum_{\text{time}} u_s(\mathbf{x}, t) u_s(\mathbf{x}, t)$, to improve the amplitudes. Kiyashchenko et al. provided a ‘true-amplitude’ map of the velocity perturbation [43]. In the next paragraphs we will present our approach to the imaging problem and explain how it differs from the approach used in these papers.

2.4.3 Our contribution

We start with the linearized inverse problem, stated in section 2.2, and use the procedure of reverse-time migration to reconstruct the perturbation of the velocity, in the sense of the reflectivity. This means that we mathematically model the source wave field, the scattering event, and the scattered wave field. The scattering event is modeled by an operator, the scattering operator, that maps the reflectivity to the scattered wave field. We propose an approximate inverse of the scattering operator, the imaging operator, and derive an imaging condition based on this operator. Our approach in chapter 4 can be compared with the approaches of the mathematically rigorous studies of the linearized scattering operator, e.g. Nolan and Symes [50]. In this sense it differs from the strategy of Kiyashchenko et al., who took the source-normalized cross-correlation imaging as a starting point.

As mentioned in section 2.1 we use the theory of Fourier integral operators in our study of reverse-time migration. The scattering operator, for instance, is a FIO. Also, the solution of the wave equation IVP can be expressed as a FIO acting on the initial values. A FIO F is an operator that can be written as an oscillatory integral. Acting on u , it has the general expression:

$$Fu(\mathbf{x}) = \frac{1}{(2\pi)^n} \iint e^{i\varphi(\mathbf{x}, \mathbf{y}, \boldsymbol{\xi})} a_r(\mathbf{x}, \mathbf{y}, \boldsymbol{\xi}) u(\mathbf{y}) dy d\boldsymbol{\xi}. \quad (2.16)$$

Variables \mathbf{x} and \mathbf{y} here may denote space or spacetime coordinates, depending on the application. The phase function φ determines to a large extent the properties

of the operator. For example, if we take $\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{y})$ then F becomes a PsDO, which can be written as:

$$Fu(\mathbf{x}) = \frac{1}{(2\pi)^n} \int e^{i\boldsymbol{\xi} \cdot \mathbf{x}} a_F(\mathbf{x}, \boldsymbol{\xi}) \widehat{u}(\boldsymbol{\xi}) d\boldsymbol{\xi}. \quad (2.17)$$

We took the amplitude $a_F(\mathbf{x}, \boldsymbol{\xi})$ independent of \mathbf{y} to avoid technical details. The hat $\widehat{}$ denotes the Fourier transform with respect to \mathbf{y} . In the application of the one-way wave equation $a_F(\mathbf{x}, \boldsymbol{\xi})$ is the square root symbol, i.e. the expression in the right hand side of (2.14). Hence, a PsDO is a special case of a FIO.

To understand why we want to involve more general phase functions we refer to the ansatz $a(\mathbf{x}, t) e^{i\psi(\mathbf{x}, t)}$ introduced in section 2.3 to construct solutions of the wave equation. The solutions that arise are oscillatory integral expressions like (2.16). The derivative $\partial_{\mathbf{x}}\varphi$ can be interpreted as the local wave vector and varies in space, especially in case of varying media. Hence, the phase $\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{y})$ will just not do. But also in case of constant media it is necessary to involve more general phases. This is illustrated in the following example of the IVP of the wave equation:

$$\begin{aligned} [\partial_t^2 - c^2\Delta]u(\mathbf{x}, t) &= 0 \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \quad \partial_t u(\mathbf{x}, 0) = 0. \end{aligned} \quad (2.18)$$

Application of the Fourier transform with respect to the space variables leads to a solution in the form of an integral of plane waves [22]:

$$u(\mathbf{x}, t) = \frac{1}{(2\pi)^n} \iint e^{i(\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{y}) \mp |\boldsymbol{\xi}|ct)} \frac{1}{2} u_0(\mathbf{y}) d\mathbf{y} d\boldsymbol{\xi}, \quad (2.19)$$

where \mp means taking the sum of two terms, one with a $-$ sign and the other with a $+$ sign. The first term has the phase $\varphi(\mathbf{x}, t, \mathbf{y}, \boldsymbol{\xi}) = \boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{y}) - |\boldsymbol{\xi}|ct$. Note that the variables in (2.16) differ in that space and time are taken together.

The theory of FIOs is built on the method of stationary phase, a basic principle of asymptotic analysis. It studies the limit behaviour of oscillatory integrals like (2.16) when the frequency goes to infinity. The idea of the method relies on the addition of many sinusoids with rapidly varying phases. They add constructively only if the phases coincide and remain the same as the frequency increases. The formal statement of the principle is that the asymptotic behaviour depends on the critical points of the phase, i.e. $\partial_{\boldsymbol{\xi}}\varphi = 0$. If we set $\varphi_0 = \varphi(\mathbf{x}, \mathbf{y}, \boldsymbol{\xi}_0)$ and linearize the phase around a critical point $\boldsymbol{\xi}_0$ then the linear term vanishes, so

$$\varphi(\mathbf{x}, \mathbf{y}, \boldsymbol{\xi}) = \varphi_0 + \mathcal{Q}. \quad (2.20)$$

Apart from the remainder \mathcal{Q} , which is quadratic in $\boldsymbol{\xi} - \boldsymbol{\xi}_0$, the phase is constant in a neighborhood of a critical point. In section 2.3 we introduced the level set (2.11) as a way to define the wave front of a propagating wave. Critical points are hence

expected to somehow describe the wave front. To make this explicit we determine the critical points of the first term of example (2.19). Condition $\partial_{\xi}\varphi = 0$ yields:

$$\mathbf{x} = \mathbf{y} + \frac{\xi}{|\xi|}ct. \quad (2.21)$$

This can be interpreted as the movement of the wave fronts, by describing the rays, which in this example are straight lines directed along ξ . If \mathbf{y} is a point of the wave front at initial time then it will be moved to \mathbf{x} at time t .

Dwelling upon the phase function, we would almost forget that the novelty of our work lies in the amplitudes. Besides regularity conditions, e.g. integrability of the oscillatory integrant, the amplitude function $a_{\mathbb{F}}$ of the FIO (2.16) provides a way to describe the geometric amplitude. By that is meant the amplitude of waves that propagate along rays. In our study of reverse-time migration we make explicit calculations of the amplitudes. And the imaging condition that we propose in section 4.5 gives a principal order reconstruction of the reflectivity. This means that, within the aperture of the imaging technique, the relative error of the image is of the order of $\frac{1}{\omega}$. The aperture indicates to what extent the reflectivity can be reconstructed, given the source position and the acquisition domain, of an otherwise ideal experiment. This can be understood as the set of positions and orientations of subsurface reflectors that can be ‘seen’ by the reflection experiment.

The outline of chapter 4, the chapter about reverse-time migration, can be found in section 2.6. We will first have a glance at some results from microlocal analysis in preparation for the theory of the core chapters.

2.5 Microlocal analysis

Fourier integral operators are developed by Hörmander et al. to represent solutions of partial differential equations [37]. A FIO has singularity mapping properties, which are studied by investigating the limit behavior of an oscillatory integral representation when the frequency goes to infinity [22]. Rather than considering the oscillatory integral, in section 2.5.1 we will examine the singularities themselves and show that they are described by, what is called, the *wave front set*. In section 2.5.2 we apply the wave front set to solutions of partial- or pseudo-differential equations, which leads to the notion of propagating singularities.

2.5.1 Wave front set

We use an assumption, applied in many seismic imaging techniques, that can be referred to as the *separation of scales*. In the formulation of the linearized inverse problem, section 2.2, we assume that the medium velocity can be modeled by a sum of a slowly varying background c and a small scale perturbation $\delta c = rc$. And

for the modeling of wave propagation through the slowly varying background we use oscillatory functions like $a e^{i\varphi}$ (2.10). The amplitude a is assumed to vary on a large scale. The high frequency of the phase φ represents the small scale.

The physical distinction between the scales depends on the ratio of the heterogeneity size, i.e. length scale over which the velocity varies, to the seismic wave length and the distance the wave travels through the heterogeneous region [59]. In the mathematical theory of linearized scattering however, the distinction is more abstract. ‘Slowly varying’ is translated into ‘smooth differentiable’ and the ‘small scale perturbation’ is translated into the ‘singularities of a distribution’.

The *wave front set* $\text{WF}(u)$ is used to analyze the location and orientation of singularities of a distribution u . When u is the solution of the wave equation the wave front set can be interpreted literally, in the sense that it describes the singularities that constitute the wave front. If we apply it to the velocity perturbation then $\text{WF}(\delta c)$ describes the reflectors of the subsurface. To explain the wave front set we will apply it to a function in \mathbb{R}^n .

Let $\mathbf{x}' \in \mathbb{R}^{n-1}$ denote the lateral coordinates and x_n the depth. Given are two smooth functions $u_b, u_t : \mathbb{R}^n \rightarrow \mathbb{R}$ and a function $h : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$, also smooth. Then $x_n = h(\mathbf{x}')$ describes a surface. We define function u by

$$u(\mathbf{x}) = \begin{cases} u_t(\mathbf{x}) & \text{if } x_n < h(\mathbf{x}') \\ u_b(\mathbf{x}) & \text{if } x_n \geq h(\mathbf{x}') \end{cases}$$

Figure 2.2 presents an illustration of the function. We assume that $u_b > u_t > 0$, implying a jump discontinuity at h . We will see that the singularities are on the surface, and that their orientations are normal to the surface.

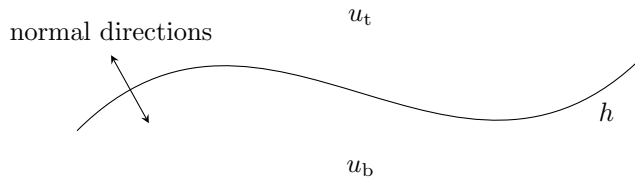


Figure 2.2: Model function u is smooth everywhere except on the surface h .

Singularities occur in points where the distribution is not *microlocally smooth*. A distribution u is microlocally smooth at $(\mathbf{x}_0, \boldsymbol{\xi}_0) \in \mathbb{R}^n \times \mathbb{R}^n \setminus 0$ if the following holds: There exists a test function $\varphi \in C_0^\infty$ supported in the neighborhood of \mathbf{x}_0 with $\varphi(\mathbf{x}_0) \neq 0$ such that φu is C^∞ , i.e.

$$|\widehat{\varphi u}(\boldsymbol{\xi})| \leq C_k (1 + |\boldsymbol{\xi}|)^{-k} \quad (2.22)$$

for all k . The hat $\widehat{}$ denotes the Fourier transform. This implies that u is locally smooth at \mathbf{x}_0 . But microlocal smoothness also involves a sense of direction. This means that there exists an open cone Γ containing $\boldsymbol{\xi}_0$ and such that the inequality (2.22) holds for all $\boldsymbol{\xi} \in \Gamma$. A cone is a set that contains $\lambda \boldsymbol{\xi}$ for all $\lambda > 0$ if it contains $\boldsymbol{\xi}$. For the correct order of quantifiers [64, 22, 1]:

Definition 1. *Distribution u is microlocally smooth at $(\mathbf{x}_0, \boldsymbol{\xi}_0)$ if there exists a $\varphi \in C_0^\infty$ with $\varphi(\mathbf{x}_0) \neq 0$ and an open cone Γ containing $\boldsymbol{\xi}_0$ such that for all k there exists a C_k such that the inequality (2.22) holds for all $\boldsymbol{\xi} \in \Gamma$. The wave front set of u , denoted by $\text{WF}(u)$, is defined as the complement in $\mathbb{R}^n \times \mathbb{R}^n \setminus 0$ of all points that are microlocally smooth.*

There is one consistency condition that should be checked. Once we have a φ that works in showing that u is microlocally smooth at $(\mathbf{x}_0, \boldsymbol{\xi}_0)$, it should also work if we take a cutoff function with smaller support. This is true. See [64], section 8.6.

We simplify the model function u in two steps and apply the definition to the simplification. First we define $\tilde{u} = \frac{u - u_t}{u_b - u_t}$. It is ‘digitized’ in the sense that it only takes the values 0 and 1. It can be shown that $\text{WF}(u) = \text{WF}(\tilde{u})$. Although intuitive, the proof is not fully trivial. As a second simplification we straighten the surface h by a diffeomorphism g on \mathbb{R}^n . Let g be such that the plane $x_n = 0$ is mapped to the surface $x_n = h(\mathbf{x}')$, and such that the half space $x_n > 0$ is mapped to $x_n > h(\mathbf{x}')$. The flattened model is defined as $\bar{u} = \tilde{u} \circ g$. We will use the knowledge that the wave front set of a distribution v transforms like [64, 22]:

$$(\mathbf{x}, \boldsymbol{\xi}) \in \text{WF}(v \circ g) \Leftrightarrow (g(\mathbf{x}), [\partial_{\mathbf{x}} g^{-1}]^T \boldsymbol{\xi}) \in \text{WF}(v). \quad (2.23)$$

Note that \bar{u} is the indicator function of the set $x_n \geq 0$. We will analyse the wave front set of \bar{u} , and use this to identify the wave front set of u itself.

The wave front set $\text{WF}(\bar{u})$ consists of the pairs $(\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^n \times \mathbb{R}^n \setminus 0$ for which hold $\boldsymbol{\xi}' = 0$, and obviously $x_n = 0$. We will argue by the definition 1 why the covector $\boldsymbol{\xi}$ is normal to the plane $x_n = 0$. Let $\boldsymbol{\xi} = \lambda \boldsymbol{\nu}$ with $\|\boldsymbol{\nu}\| = 1$ and $\lambda > 0$. We choose a test function of the form $\varphi(\mathbf{x}) = \varphi_1(\mathbf{x}') \varphi_2(x_n)$. Note that $\bar{u} = \bar{u}(x_n)$. The question now is how

$$\widehat{\varphi \bar{u}}(\lambda \boldsymbol{\nu}) = \widehat{\varphi_1}(\lambda \boldsymbol{\nu}') \widehat{\varphi_2 \bar{u}}(\lambda \nu_n)$$

decays if λ goes to infinity. As long as $\boldsymbol{\nu}' \neq 0$, i.e. $\boldsymbol{\nu}$ is not normal to the plane, for all k holds that $|\widehat{\varphi_1}(\lambda \boldsymbol{\nu}')| \leq C_k (1 + \lambda)^{-k}$. Since φ_2 has compact support one has $|\widehat{\varphi_2 \bar{u}}(\lambda \nu_n)| \leq C (1 + \lambda)^K$ for some K , which is enough. This proves microlocal smoothness of \bar{u} in $(\mathbf{x}, \boldsymbol{\xi})$ if $\boldsymbol{\xi}' \neq 0$ and $x_n = 0$. The covectors of the wave front set $\text{WF}(\bar{u})$ are therefore normal to the plane.

The wave front set of \tilde{u} and hence of the model function u can be concluded from the transformation formula (2.23). The formula shows that the variable $\boldsymbol{\xi}$ transforms like a *cotangent* vector. The perpendicularity of a tangent and a cotangent

vector is preserved under the transformation by g .² Tangent here means tangent to the plane $x_n = 0$ in model \bar{u} , and, tangent to the surface h in model u . Supported by the illustration in figure 2.2, the conclusion is

$$\text{WF}(u) = \{(\mathbf{x}, \boldsymbol{\xi}) \in \mathbb{R}^n \times \mathbb{R}^n \setminus 0 \mid x_n = h(\mathbf{x}') \text{ and } \boldsymbol{\xi} \text{ normal to the surface}\}.$$

The wave front set is the mathematical formalization of localized disturbances. As mentioned in the first paragraphs of this section, it is used to describe wave fronts in the sense of solutions of the wave equation, and discontinuities in case of the subsurface. An example of the latter is the velocity model in figure 2.1, in which the wave front set then would describe the discontinuity at surface h .

2.5.2 Propagation of singularities

We will apply the concept of the wave front set to solutions of partial differential equations, the wave equation in particular. The first observation is that the wave front set of a distribution u is not increased by the application of a differential operator with smooth coefficients P ,

$$\text{WF}(Pu) \subset \text{WF}(u). \quad (2.24)$$

An operator with this property is a microlocal operator [68]. It means that if u is microlocally smooth at $(\mathbf{x}, \boldsymbol{\xi})$ then so is Pu . Pseudo-differential operators are microlocal, and in the following we allow for P to be a PsDO.

Before continuing we say something about the symbol of an operator and its principal symbol. In section 2.3 and 2.4 we saw that the symbol of a PsDO is a function of spacetime variables \mathbf{x} and Fourier associates $\boldsymbol{\xi}$. For example, the symbol of a differential operator of order m is a polynomial, $a(\mathbf{x}, \boldsymbol{\xi}) = \sum_{|\alpha|=0}^m a_\alpha(\mathbf{x})(i\boldsymbol{\xi})^\alpha$. A non-trivial PsDO does not have a polynomial symbol. We will work with classical symbols. Such a symbol has an asymptotic expansion of homogeneous symbols,

$$p(\mathbf{x}, \boldsymbol{\xi}) = \sum_{k=0}^{\infty} p_{m-k}(\mathbf{x}, \boldsymbol{\xi}) \quad (2.25)$$

with $p_j(\mathbf{x}, \lambda\boldsymbol{\xi}) = \lambda^j p_j(\mathbf{x}, \boldsymbol{\xi})$ for large $\boldsymbol{\xi}$ and positive λ . Here, j is the order of p_j . The principal symbol, this is $p_m(\mathbf{x}, \boldsymbol{\xi})$, dominates the expansion for large $\boldsymbol{\xi}$. Applied to our polynomial example the sum of all terms with $|\alpha| = m$ comprise the principal symbol. In section 3.2 one can find more details. In the following we assume that P is a PsDO with classical symbol (2.25) and principal symbol p_m .

We say that $\boldsymbol{\xi}$ is characteristic for P at \mathbf{x} if $p_m(\mathbf{x}, \boldsymbol{\xi}) = 0$. The set of all $(\mathbf{x}, \boldsymbol{\xi})$ where $p_m = 0$ is denoted by $\text{char}(P)$. To answer the question whether an operator

²A tangent vector \mathbf{v} maps to $[\partial_{\mathbf{x}}g]\mathbf{v}$ and a cotangent vector $\boldsymbol{\xi}$ maps to $[\partial_{\mathbf{x}}g^{-1}]^T\boldsymbol{\xi}$. The dot product $\mathbf{v} \cdot \boldsymbol{\xi}$ is hence invariant under g , and so is the perpendicularity.

can decrease the wave front set we first consider elliptic operators. For elliptic operators, i.e. when $p_m(\mathbf{x}, \boldsymbol{\xi}) = 0$ implies $\boldsymbol{\xi} = 0$, the answer is ‘no’. This property is reflected by a general theorem that says that P has an approximate inverse Q , called a *parametrix*, for which holds $QP = PQ = I + R$ [29]. The ‘error’ of the approximation is an operator R that maps every distribution to a smooth function, called a regularizing operator. So for elliptic operators $\text{WF}(Pu) = \text{WF}(u)$. Whether P is elliptic or not, we have the following theorem.

Theorem 1. $\text{WF}(u) \subset \text{WF}(Pu) \cup \text{char}(P)$.

It says something about where the singularities of u can be found if those of Pu are known. In words the theorem says that if Pu is microlocally smooth at $(\mathbf{x}, \boldsymbol{\xi})$, and $\boldsymbol{\xi}$ is non-characteristic at \mathbf{x} , then u is microlocally smooth at $(\mathbf{x}, \boldsymbol{\xi})$. Suppose u is a solution of $Pu = f$ with f smooth or zero. Then one has $\text{WF}(u) \subset \text{char}(P)$. This does not say that much until we analyse $\text{char}(P)$ and show that it is a disjoint union of curves called *bicharacteristics*. And there is an all-or-nothing dichotomy: Either $\text{WF}(u)$ contains the whole bicharacteristic or no points of it. This is meant by the phrase ‘singularities propagate along bicharacteristics’. To examine this we take into consideration operators of principal type. This intuitively means that the principal symbol dominates over the lower order terms [70]. It is a rather broad class of operators that includes both elliptic and strictly hyperbolic operators, like the wave equation. And we assume that p_m is real.

A bicharacteristic is a trajectory $s \mapsto (\mathbf{x}(s), \boldsymbol{\xi}(s))$ of the Hamiltonian system:

$$\begin{aligned} \frac{d}{ds} \mathbf{x} &= \partial_{\boldsymbol{\xi}} p_m(\mathbf{x}, \boldsymbol{\xi}) \\ \frac{d}{ds} \boldsymbol{\xi} &= -\partial_{\mathbf{x}} p_m(\mathbf{x}, \boldsymbol{\xi}). \end{aligned} \tag{2.26}$$

This is a system of 1st-order ordinary differential equations. The existence and uniqueness theorems imply that there is a unique solution for every initial value of $(\mathbf{x}, \boldsymbol{\xi})$. The assumption that P is of principal type implies that $\frac{d}{ds} \mathbf{x}$ is non-zero along a bicharacteristic. A solution hence does not degenerate to a point. One can easily verify that p_m is constant along a bicharacteristic. The set $\text{char}(P)$ splits into a union of bicharacteristics for which holds $p_m = 0$, sometimes called null bicharacteristics. We omit the adjective ‘null’ in the following theorem.

Theorem 2. *Let P be an operator of real principal type and γ a bicharacteristic. If $Pu \in C^\infty$ then either $\gamma \subset \text{WF}(u)$ or $\gamma \cap \text{WF}(u) = \emptyset$.*

We apply the theorem to the wave equation. Then $p_m = -\omega^2 + c(\mathbf{x})^2 |\boldsymbol{\xi}|^2$ and a bicharacteristic, parameterized by t , is a curve $t \mapsto (\mathbf{x}, t; \boldsymbol{\xi}, \omega)$ that solves the Hamiltonian system. Note that in (2.26) the space and time variables are taken together. There are two independent solutions for each $\boldsymbol{\xi}$ and they are distinguished by the frequency $\omega = \mp c(\mathbf{x}) |\boldsymbol{\xi}|$, which is constant on each bicharacteristic. The first

equation of (2.26) becomes $\frac{d}{dt}\mathbf{x} = \pm c(\mathbf{x})\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}$ and shows that a singularity propagates with velocity $c(\mathbf{x})$ in the $\boldsymbol{\xi}$ -direction. And the second equation, $\frac{d}{dt}\boldsymbol{\xi} = \mp(\partial_{\mathbf{x}}c)|\boldsymbol{\xi}|$, implies that the direction of propagation varies if the gradient $\partial_{\mathbf{x}}c$ is non-zero. An example of this is the upward bending of horizontal rays in the Earth due to the fact that the velocity increases with depth, which is often the case. And, as expected, in a constant medium singularities propagate along straight lines.

The system of equations (2.26) shows up in section 4.2.2 when we solve the eikonal equation by means of the method of characteristics, a general method from the theory of PDEs [23]. There the eikonal equation governs the phase function α of an oscillatory function that, by construction, solves the wave equation. In section 4.2.4 we subsequently derive a FIO that gives solutions of the wave equation. The phase function α then determines the geometrical aspects of the operator, i.e. how the singularities propagate.

2.6 Outline

The main content of the thesis is provided by two articles that are presented in chapter 3 and 4. We give an outline for each chapter.

Chapter 3 presents the *one-wave wave equation with symmetric square root*. In section 3.2 we derive the one-way wave equation and identify the steps that are related to the wave amplitude. The mentioned correction term can be avoided by choosing a new field variable, which is related to the original variable u by a PsDO that we call the normalization operator. The details are in section 3.2. We propose a symmetric³ square root operator B , a PsDO, and show that the one-way wave equation provides a principal order approximation of the amplitude. The idea to use a *symmetric* square root is generally applicable. We will apply this to an existing method based on the one-way wave equation, the so called 60 degree Padé finite-difference implementation [14], and show that the amplitude significantly improves. We propose in section 3.3 a new numerical method to implement the normalization and the square root operator. Our amplitude claims are numerically verified in section 3.4.

Chapter 4 treats the *inverse scattering based on reverse-time migration*. In section 4.2 we begin with the construction of solutions of the wave equation, starting by an application of the WKB approximation to a plane wave IVP. We analyse the scattering problem in section 4.3 and formulate the operator that maps the reflectivity to the scattered wave field. In this respect we distinguish between the continued scattered field u_h and the reverse time continued field u_r . The former is the result of a perfect continuation, in reverse time, of the scattered field from its Cauchy values at some time after the scattering has taken place. The latter

³Then iB is anti-symmetric.

is a model for the receiver wave field, i.e. the in reverse time continued field that matches the data. In section section 4.4 we show that u_r can be written as the result of a PsDO, the revert operator, working on u_h . The operator that maps the reflectivity to u_r is referred to as the scattering operator and we obtain an explicit expression that is locally valid and a global characterization as a FIO. The inversion is presented in section 4.5. We first analyse a simplified case in which the velocity is constant. Then we introduce a novel imaging condition and show that it yields a principal order reconstruction of the reflectivity. The imaging condition is composed of the excitation time imaging condition, i.e. the restriction in space and time to the coordinates of the source wave front, and a PsDO that restores the amplitude of the image. In section 4.6 we present a number of numerical tests, which confirm our claims.

Conclusions and suggestions for further research are displayed in chapter 5.

Chapter 3

One-way wave equation with symmetric square root

Abstract.¹ The one-way wave equation is widely used in seismic migration. Equipped with wave amplitudes, migration can be provided with the reconstruction of the strength of reflectivity. We derive the one-way wave equation with geometrical amplitude by using a symmetric square root operator and a wave field normalization. The symbol of the square root operator, $\omega\sqrt{\frac{1}{c(x,z)^2} - \frac{\xi^2}{\omega^2}}$, is a function of space-time variables and frequency ω and horizontal wavenumber ξ . Only by matter of quantization it becomes an operator, and because quantization is subjected to choices it should be made explicit. If one uses a naive asymmetric quantization an extra operator term will appear in the one-way wave equation, proportional to $\partial_x c$. We propose a symmetric quantization, which maps the symbol to a symmetric square root operator. This provides geometrical amplitude without calculating the lower order term. The advantage of the symmetry argument is its general applicability to numerical methods. We apply the argument to two numerical methods. We propose a new pseudo-spectral method, and we adapt the 60 degree Padé type finite-difference method such that it becomes symmetrical at the expense of almost no extra cost. The simulations show in both cases a significant correction to the amplitude. With the symmetric square root operator the amplitudes are correct. The z -dependency of the velocity lead to another numerically unattractive operator term in the one-way wave equation. We show that a suitably chosen normalization of the wave field prevents the appearance of this term. We apply the pseudo-spectral method to the normalization and confirm by a numerical simulation that it yields the correct amplitude.

¹This chapter is exactly the article One-way wave propagation with amplitude based on pseudo-differential operators by T.J.P.M. Op 't Root and C.C. Stolk, published in Wave Motion, 47(2):67-84, 2010. <http://dx.doi.org/10.1016/j.wavemoti.2009.08.001>

3.1 Introduction

One-way wave equations allow us to separate solutions of the wave equation into down- and upward propagating waves. They are frequently used to model wave propagation in the application of depth migration [39, 42, 60, 76, 77, 47, 27, 5, 55]. They are also used in other fields such as underwater acoustics and integrated optics. One of the reasons to use these equations is that they provide wave field extrapolation in a certain direction [74]. Another reason to use these equations is the fact that they can be implemented cost efficiently [28].

One-way wave equations have first been used in geophysical imaging by Claerbout [14]. They were used to describe travel time and were not intended to describe wave amplitudes. In principle, this restricts the migration process to the reconstruction of the locations of the velocity heterogeneities. Roughly since 1980 the development is to also reconstruct the relative strength of the velocity heterogeneities and to estimate petrophysical parameters [17, 4]. For the one-way wave equation to describe amplitudes it needs to be refined. One of the problems is the formulation of the *square root operator* in case the velocity is a spatial function.

The finite-difference one-way methods form a large class of one-way methods [14, 35]. One attempt to tackle the amplitude problem can be found in the work of Zhang et. al [76, 77]. They studied the Padé family of finite-difference one-way methods and modified it such that it provides accurate leading order WKB amplitudes. The Padé family of methods includes the familiar 15, 45, 60 degree equations that are reported in Claerbout's book [14].

Another important class of methods is formed by mixed domain methods. Here calculations are performed both in the lateral space, i.e. horizontal in case of downward continuation, and the lateral wavenumber domain. The fast Fourier transform is used to go forth and back. A mixed domain method can be seen as an adaptation of the phase-shift method for lateral varying media [7]. Phase shift plus interpolation (PSPi), for example, uses several reference velocities and an interpolation technique to adapt the phase-shift for laterally varying velocity [27]. Nonstationary phase shift (NSPS) uses nonstationary filter theory to make this adaptation [47]. Although these methods allows for lateral dependence, they do not preserve the wave amplitude. Among the mixed domain methods one also finds the phase-screen method and its offshoot methods like pseudo-screen and generalized screen [39, 72, 20]. Then the medium is considered as a series of thin slabs or diffraction 'screens', stretched out in lateral direction. It is particularly suited to model propagation through media where the raypaths do not deviate substantially from a given predominant direction.

From mathematical point of view, the one-way wave equation can be found in the rigorous framework of pseudo-differential operator theory. Stolk worked on this and investigated the damping term of the symbol [60]. He compared singularities

in the mathematical sense of the wave front set and showed that singularities are described by one-way wave equations.

In this paper we investigate amplitude aspects of the one-way wave equation from theoretical and practical point of view. We identify the essential ingredients that provide the one-way wave equation with the *geometrical amplitude*, using pseudo-differential operator theory. It is comprised of an investigation of the square root operator and the *normalization operator*, which will be discussed in more detail below, and yield implementations for both. Because we focus on amplitude aspects of wave extrapolation we will show simulations of wave propagation based on one-way methods.

Although the work of Zhang et al. provided an interpretation of the square root and normalization operator through some basic ideas from the theory of pseudo-differential operators, they didn't make a strict distinction between an operator and its *symbol* [1, 37]. We find this aspect of the theory to be fundamental for understanding the true amplitude theory. A symbol is a function of space-time variables and their Fourier associates, e.g. $\omega\sqrt{\frac{1}{c(x)^2} - \frac{\xi^2}{\omega^2}}$, where ξ denotes the horizontal wavenumber. It can be transformed into an operator by matter of *quantization*. Because quantization is subjected to choices it should be made explicit. We will derive the one-way wave equation with amplitude description and explicitly use the theory of pseudo-differential operators. It gives insight in the equation and its modification to involve wave amplitudes.

The depth dependence of the medium leads to an extra operator term in the one-way wave equation which significantly effects the wave amplitude. It is known that by a suitably chosen normalization of the wave field, the one-way wave equation in the new variable does not contain this extra term [20, 60, 77]. Although numerical implementations are well-known for the square root operator [14, 35, 77], these are not reported for the normalization. Zhang et al., for example, used the normalization operator but lacked being explicit about its implementation. We will introduce the pseudo-spectral interpolation method, by which both the square root and the normalization operator can be implemented

To confine the complexity, wave propagation through the earth is modeled by the heterogeneous acoustic wave equation. The heterogeneity is captured by a velocity function, $c = c(x, z)$, which is smooth by assumption. The reader might wonder how this can be applicable to seismic waves, as everybody knows that the Earth is not smooth. Some comments about this follow at the end of the introduction. The pressure $u(t, x, z)$ of an acoustic wave originating at the source $f(t, x, z)$ is governed by:

$$\left(\frac{1}{c(x, z)^2} \partial_t^2 - \partial_x^2 - \partial_z^2\right) u(t, x, z) = f(t, x, z) \quad (3.1)$$

in infinite space-time, i.e. $(t, x, z) \in \mathbb{R} \times \mathbb{R}^2$. The support of the source, i.e. the set

where it is nonzero, is assumed to be bounded:

$$\text{supp } f \subset [t_{s0}, t_{s1}] \times [x_{s0}, x_{s1}] \times [z_{s0}, z_{s1}].$$

Furthermore, the wave field is zero at initial time, i.e. $u(t, x, z) = 0$ for $t < t_{s0}$.

The splitting into down and up going waves relies on the assumption that the wavelength is small compared with the length scale of the heterogeneity of the medium. This means that we assume the medium velocity to vary slowly over space, and, the frequency to be high. The idea can be shown by first assuming a constant velocity. Using the Fourier transform with respect to t and x , the wave operator can be written in two factors:

$$\left(i\omega\sqrt{\frac{1}{c^2} - \frac{\xi^2}{\omega^2}} + \partial_z\right)\left(i\omega\sqrt{\frac{1}{c^2} - \frac{\xi^2}{\omega^2}} - \partial_z\right)\hat{u}(\omega, \xi, z) = 0, \quad (3.2)$$

with Fourier variables ω and ξ . The factors are written in arbitrary order. Restricting the solution space to $\frac{1}{c^2} - \frac{\xi^2}{\omega^2} > 0$, the kernel of each factor consists of uni-directional waves. A numerical method based on the one-way wave equation with constant velocity can be found in [26].

When $\partial_x c \neq 0$ the theory of pseudo-differential operators is used to define the square root operator properly. It is one of our goals to show how this can be done. In case $\partial_z c \neq 0$ the noncommutativity of the square root and ∂_z introduces extra terms. These terms can be removed by a suitable normalization of wave field.

We make a comprehensible derivation of the one-way wave equation and show how the square root operator has to be defined such that the equation includes the wave amplitude. The symbol of the square root operator turns out to be the sum of two terms, i.e. the principal and the subprincipal. The former provides travel time. The terms together describe travel time and the geometric amplitude.

Quantization refers to the procedure of mapping a symbol to an operator. The theory of pseudo-differential operators involves a standard quantization, which produces asymmetric operators. We introduce an alternative quantization that maps real symbols to symmetric operators and show that it maps the principal symbol of the square root to the correct operator. By correct, we mean that it properly describes the wave amplitude. The idea to use a symmetric square root is applicable to any method. We will illustrate this by two numerical methods.

We propose a new method, the pseudo-spectral interpolation method, to implement both the square root and the normalization operator. The method is closely related to some of the mixed domain methods in that it is implemented as a sum in which each term involves a multiplication and a convolution. We will verify numerically the claimed amplitude effect of the symmetric square root operator and the normalization operator by simulating wave propagation through a medium with variable velocity. These simulations also give an illustrative example

of the improvements that can be expected. The results will be compared with a simulation of the full wave equation, which acts as a reference. The comparison shows a great improvement of the wave amplitude due to the symmetric square root and the normalization operator.

As said, the outcome can be applied to other one-way wave methods. We show that the 60 degree Padé type finite-difference one-way implementation can be made true amplitude by using a symmetric implementation for the square root operator. This modification entails almost no extra cost. Again, numerical simulation shows an improved wave amplitude.

The content of the remaining sections of this paper is as follows. In section 3.2 we derive the one-way wave equation for smoothly varying media in the rigorous framework of pseudo-differential operator theory. The numerical implementation of our one-way wave equation is presented in section 3.3, including the pseudo-spectral method to implement pseudo-differential operators. The results of the simulation and a comparison with the simulation of the full wave equation are shown in section 3.4. A discussion is finally presented in section 3.5.

We end this introduction by a few remarks on the application of one-way wave equations in reflection seismic imaging. This is based on the *single scattering assumption* that is generally used in seismic imaging. It involves a geometrical view of wave propagation and assumes that waves present in the data have reflected once. Multiple reflected waves are treated as noise, and to a certain degree they can be suppressed.

The single scattering data can be viewed as the result of linearization of the wave equation with respect to the coefficients, i.e. the medium velocity. Additionally, the reference velocity $c(x)$ is assumed to be smooth, while the perturbation $\delta c(x)$ is oscillatory. In practice this means that c and δc contain different wavelengths, e.g. up to 2-5 Hz for c , and starting somewhere from 5-10 Hz for the reflectivity. These values are converted to the temporal frequency domain using a typical velocity. The unperturbed wave originating from the source f_{src} is the incoming field u_{inc} . Its perturbation is the scattering field u_{sca} . They are governed by the reference and linearized PDE,

$$\frac{1}{c^2} \partial_t^2 u_{\text{inc}} - \Delta u_{\text{inc}} = f_{\text{src}}, \quad \text{and} \quad \frac{1}{c^2} \partial_t^2 u_{\text{sca}} - \Delta u_{\text{sca}} = \frac{2\delta c}{c^3} \partial_t^2 u_{\text{inc}}, \quad (3.3)$$

both supplemented with zero initial and appropriate boundary conditions. The data is assumed to be given by u_{sca} at measurement positions spread over the surface during some time interval.

A typical imaging algorithm [14] involves the incoming field u_{inc} and the back-propagated receiver field, denoted by u_{bp} , which is obtained by solving a wave equation backward in time with the data as source or boundary condition. These

fields are convolved in time

$$I(x, z) = \int_0^{T_{\text{acq}}} u_{\text{inc}}(t, x, z) u_{\text{bp}}(t, x, z) dt,$$

where $[0, T_{\text{acq}}]$ is the time interval for the acquisition assuming the source is set off at $t = 0$. Clearly, solving the wave equation is the most costly part of the imaging algorithm. With present-day data it is to be repeated tens of thousands of times. One-way methods constitute practical tools for modelling wave propagation through smooth media.

3.2 One-way wave equation

In this section we will derive the true amplitude one-way wave equation for a normalized wave field. We will present the essential steps in a self-contained way, without requiring knowledge of pseudo-differential operator theory in advance. For the reader primarily interested in the results we point out that the one-way wave equation is in (3.22), the normalization in (3.23), and the symmetric square root operator in (3.19) and (3.8).

The approach to split the wave field followed in this section slightly differs from the factorization in (3.2). We will tackle the problem by writing the wave equation as a system of equations, which makes it easy to include the source. Diagonalization of the system leads to our first result, the one-way wave equation for wave propagation in the positive z -direction through a medium of which the velocity only depends on x . The decomposition relies on the square root operator. This operator will be subject of further investigation. The normalization of the wave field is used to derive the one-way wave equation for velocities that may also depend on z .

As shown in (3.2), the square root operator is a well-defined convolution when $\partial_x c = 0$. In case $\partial_x c \neq 0$ it becomes a nontrivial pseudo-differential operator. We will review the basic ingredients of a pseudo-differential operator, i.e. its symbol and quantization. A symbol is written as an asymptotic expansion of which the first two terms are called the principal and the subprincipal part. We apply a composition theorem to calculate the principal and subprincipal symbols of the square root operator. Referring to the standard quantization, both are to be taken into account to provide the one-way wave equation with the geometrical amplitude.

From practical point of view it is worthwhile to investigate the quantization. We introduce a symmetric quantization and formulate the symmetric square root, i.e. the operator found by symmetric quantization of the principal symbol only. We show that it provides the one-way wave equation with the geometrical amplitude too. The benefit of the symmetry argument is its applicability to other implementations. For example, the square root of the 60 degree finite-difference one-way

method can be made symmetric and therefore true amplitude, as will be shown in section 3.3.3.

3.2.1 One-way wave decomposition

The object is to derive one-way wave equations for wave propagation along or in opposite direction of the z -axis through inhomogeneous media, i.e. when $c = c(x, z)$.

To start with, we write the vectors $\mathbf{u} = \begin{pmatrix} u \\ \partial_z u \end{pmatrix}$ and $\mathbf{f} = \begin{pmatrix} 0 \\ f \end{pmatrix}$, the operator $A = -\frac{1}{c(x,z)^2} \partial_t^2 + \partial_x^2$ and matrix operator

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -A & 0 \end{pmatrix}.$$

The wave equation in (3.1) can now be rewritten as matrix differential equation:

$$\partial_z \mathbf{u}(t, x, z) = \mathbf{A} \mathbf{u}(t, x, z) - \mathbf{f}(t, x, z). \quad (3.4)$$

Note that it is a 1st-order equation with respect to ∂_z .

Suppose that operators \sqrt{A} and $A^{-\frac{1}{2}}$ exist, whose squares are A and A^{-1} respectively. The construction of operators having approximately these properties will be discuss below. We define the following matrix operators:

$$\mathbf{V} = \begin{pmatrix} 1 & 1 \\ i\sqrt{A} & -i\sqrt{A} \end{pmatrix}, \quad \mathbf{\Lambda} = \frac{1}{2} \begin{pmatrix} 1 & -iA^{-\frac{1}{2}} \\ 1 & iA^{-\frac{1}{2}} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} i\sqrt{A} & 0 \\ 0 & -i\sqrt{A} \end{pmatrix}.$$

The second matrix is the inverse of the first. These matrices yield an eigenvalue decomposition of matrix differential operator \mathbf{A} given by

$$\mathbf{A} = \mathbf{V} \mathbf{B} \mathbf{\Lambda}. \quad (3.5)$$

A change of variables is defined by $\mathbf{u} = \mathbf{V} \mathbf{v}$. The system of differential equations in (3.4) then transforms into

$$\mathbf{V} \partial_z \mathbf{v} = (\mathbf{A} \mathbf{V} - (\partial_z \mathbf{V})) \mathbf{v} - \mathbf{f}.$$

in which the identity $\partial_z \mathbf{V} \mathbf{v} = (\partial_z \mathbf{V}) \mathbf{v} + \mathbf{V} \partial_z \mathbf{v}$ is used. Left multiplication with $\mathbf{\Lambda}$ yields

$$\partial_z \mathbf{v} = (\mathbf{B} - \mathbf{\Lambda}(\partial_z \mathbf{V})) \mathbf{v} - \mathbf{\Lambda} \mathbf{f}, \quad (3.6)$$

in which the decomposition (3.5) is used. Working out the differentiation gives

$$\mathbf{\Lambda}(\partial_z \mathbf{V}) = \frac{1}{4} A^{-1} (\partial_z A) \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

In case the medium is invariant with depth, operator term $\mathbf{\Lambda}(\partial_z \mathbf{V})$ is zero. With the sign choices for \sqrt{A} as in (3.2), i.e. the sign of ω , equation (3.6) then describes two decoupled one-way waves. Writing v for the second component of \mathbf{v} , the second equation describes wave propagation in the positive z -direction. The one-way wave equation for depth invariant media then is:

$$\partial_z v(t, x, z) = -i\sqrt{A} v(t, x, z) - \frac{i}{2} A^{-\frac{1}{2}} f(t, x, z), \quad (3.7)$$

Note that v denotes the wave field, *not the velocity*. In the following, we will review the definition of pseudo-differential operators and find explicit approximations for \sqrt{A} and $A^{-\frac{1}{2}}$. To generalize this result to depth dependent media, we will show in the end of this section how to deal with matrix operator $\mathbf{\Lambda}(\partial_z \mathbf{V})$. A well chosen normalization of the wave field will cancel the diagonals terms. The off-diagonals terms will shown to be negligible by a subtle argument.

3.2.2 Pseudo-differential operators

In case the medium is lateral invariant, i.e. $c = c(z)$, the square root of operator A can be found by using the Fourier transform with respect to t and x . It is a multiplication operator with respect to the Fourier variables and given by

$$b_0(\omega, \xi, z) = \omega \sqrt{\frac{1}{c(z)^2} - \frac{\xi^2}{\omega^2}}. \quad (3.8)$$

The square root is not unique due to sign choices. This form leads to unidirectional waves solutions.

When the medium laterally varies, the square root also depends on x . Although $b_0(x, \omega, \xi, z)$ is well-defined as a function, it doesn't define an operator yet. This function of space-time variables and Fourier associates is called symbol. It transforms into an operator by matter of quantization.

We will show that different operators can be found for one symbol. Subsequently, a short introduction of pseudo-differential operators is given, including an important theorem about composition. In this section y is an arbitrary space-time variable with η its Fourier associate. In the context of our application one reads $y = (t, x)$ and $\eta = (\omega, \xi)$. And we use the following convention for the Fourier transform: $(\mathcal{F}\varphi)(\eta) = \int \varphi(y) e^{-i\eta y} dy$.

Quantization refers to the procedure of mapping a function of space-time variables and their Fourier associates to an operator. Given the function $p = p(y, \eta)$, there are two obvious quantizations. These are the *left* quantization

$$\text{Op}_L(p)\varphi(y) = \frac{1}{(2\pi)^2} \int p(y, \eta) (\mathcal{F}\varphi)(\eta) e^{i\eta y} d\eta, \quad (3.9)$$

and the *right* quantization

$$\text{Op}_R(p)\varphi(y) = (\mathcal{F}^{-1} \int p(\tilde{y}, \eta) \varphi(\tilde{y}) e^{-i\eta\tilde{y}} d\tilde{y})(y). \quad (3.10)$$

The difference becomes clear if we, for example, take $p(y, \eta) = ih(y)\eta$. Then it follows that

$$\begin{aligned} \text{Op}_L(p)\varphi(y) &= h(y) \partial_y \varphi(y) \\ \text{Op}_R(p)\varphi(y) &= \partial_y (h(y) \varphi(y)) = \text{Op}_L(p)\varphi(y) + (\partial_y h) \varphi(y). \end{aligned} \quad (3.11)$$

It shows that the order of multiplication with $h(y)$ and differentiation ∂_y is interchanged. The difference involves derivative $\partial_y h$, as the operators do not commute. Applying this to function b_0 , one concludes that $\text{Op}_L(b_0)$ and $\text{Op}_R(b_0)$ are different operators when the velocity depends on x .

A pseudo-differential operator is the operator found by left quantization of a symbol [37, 22, 1, 73]. A symbol p of order $m \in \mathbb{R}$ is a smooth function of a space-time variable and its Fourier associate, say y and η , that satisfies the following technical mathematical condition. For all multi-indices α, β there exists a constant C such that

$$|\partial_y^\alpha \partial_\eta^\beta p(y, \eta)| \leq C(1 + |\eta|)^{m-|\beta|}. \quad (3.12)$$

This for example implies that the function $\partial_y h$ in (3.11) is a symbol also.

In the context of one-way wave equations it is possible to work with a smaller class, namely the *polyhomogeneous* symbols of order m . A polyhomogeneous or classical symbol can be written as an asymptotic sum

$$p(y, \eta) \sim \sum_{j=0}^{\infty} p_j(y, \eta) \quad (3.13)$$

of homogeneous functions $p_j(y, \eta)$. The homogeneity means that

$$p_j(y, \lambda\eta) = \lambda^{m-j} p_j(y, \eta)$$

for large η and positive λ [37, 1]. The function $p_j(y, \eta)$ is thus a symbol of order $m - j$ itself and $p_0(y, \eta)$ is called the *principal* symbol. We are mainly interested in the principal symbol, which contains the greatest contribution. In some cases the subprincipal symbol, i.e. p_1 , is also considered.

To distinguish the operator P from its symbol p and simultaneously emphasize their relation, the operator will also be denoted by $p(y, D_y)$ or $\text{Op}(p) = \text{Op}_L(p)$. We have written the differential operator as $D_y = -i\partial_y$ and shall also write $D_x = -i\partial_x$ and $D_t = -i\partial_t$. By phrasing *operator symbol* or *symbol of operator*, this always

refers to the standard quantization, which is left. For example, operator A is the 2^{nd} -order pseudo-differential operator given by

$$A = a(x, D_t, D_x, z) \quad \text{with} \quad a(x, \omega, \xi, z) = \frac{\omega^2}{c(x, z)^2} - \xi^2, \quad (3.14)$$

in which z is just a parameter.

An important result is that the composition of two pseudo-differential operators, $P = \text{Op}(p)$ of order m and $Q = \text{Op}(q)$ of order n , is again a pseudo-differential operator and has order $m + n$. The symbol of PQ is denoted by $p\#q$ and its asymptotic expansion can be found by a fundamental theorem of the symbolic calculus [37, 1]. Because our symbols are time independent, we present this theorem with variables x and ξ :

$$\begin{aligned} p\#q(x, \xi) &\sim \sum_{k=0}^{\infty} \frac{1}{k!} \partial_{\xi}^k p(x, \xi) D_x^k q(x, \xi) \\ &\sim p(x, \xi) q(x, \xi) + \partial_{\xi} p(x, \xi) D_x q(x, \xi) + r(x, \xi), \end{aligned} \quad (3.15)$$

in which $r(x, \xi)$ is a symbol of order $m + n - 2$. Clearly, the principal symbol of PQ is the product of the principal symbols of P and Q . An interesting consequence is that the commutator $[P, Q]$ is a pseudo-differential operator of order $m + n - 1$.

3.2.3 Square root operator and symmetric quantization

The principal symbol of the square root of operator A is given by the square root of the symbol of operator A , i.e. $b_0(x, \omega, \xi, z)$ in equation (3.8) with $c = c(x, z)$, which has order 1. For the moment, we ignore the singularity, i.e. when $\frac{\omega^2}{c(x, z)^2} - \xi^2 = 0$.

The theory of one-way wave equations is based on a high frequency assumption, and the order of a term in the asymptotic expansion is an important parameter. The principal symbol b_0 determines the rays of the wave field [60]. Hence, $B_0 = \text{Op}(b_0)$ is our first candidate for the square root operator. The error $A - B_0^2$ is an operator with a 1^{st} -order symbol. This follows from the application of equation (3.15).

To also describe the wave amplitudes, the subprincipal symbol b_1 is needed, as will be argued in the last subsection. Hence, we introduce $B = \text{Op}(b)$ with $b = b_0 + b_1$. The symbol $b_1(x, \omega, \xi, z)$ follows from the symbolic calculus, i.e. (3.15), and the observation that the error $A - B^2$ must have a 0^{th} -order symbol. Writing

the order below each term, we get

$$\begin{aligned}
(b_0 + b_1)\#(b_0 + b_1) &= \sum_{k=0}^{\infty} \frac{1}{k!} \left[\partial_{\xi}^k b_0 \underset{2-k}{D_x^k} b_0 + \partial_{\xi}^k b_0 \underset{1-k}{D_x^k} b_1 + \partial_{\xi}^k b_1 \underset{1-k}{D_x^k} b_0 + \partial_{\xi}^k b_1 \underset{-k}{D_x^k} b_1 \right] \\
&= \underset{2}{b_0^2} + \underset{1}{\partial_{\xi} b_0 D_x b_0} + \underset{1}{b_0 b_1} + \underset{1}{b_1 b_0} + \underset{0}{r'}, \\
&= a + \partial_{\xi} b_0 D_x b_0 + 2b_0 b_1 + r',
\end{aligned}$$

in which r' is a symbol of order 0. To cancel the 1st-order terms, the subprincipal symbol must be

$$b_1 = -\frac{1}{2b_0} \partial_{\xi} b_0 D_x b_0 = i \frac{\xi \partial_x c}{2\omega} \left(1 - \frac{c(x,z)^2 \xi^2}{\omega^2} \right)^{-\frac{3}{2}}, \quad (3.16)$$

wherein the equality $\partial_{\xi} b_0 D_x b_0 = \frac{-i\omega^2 \xi \partial_x c}{a c(x,z)^3}$ is used. Observe that $-ib_1$ is real for propagated waves, which confirms its effect on the amplitude.

With principal and subprincipal parts, i.e. b_0 and b_1 , defined in (3.8) and (3.16), we have found the pseudo-differential operator $B = b(x, D_t, D_x, z)$ such that the error $B^2 - A$ is a pseudo-differential operator of order 0. It is possible to find higher order accurate square roots by repeatedly applying analogous steps. This yields a well-defined infinite asymptotic expansion $\sum_{j=0}^{\infty} b_j$. The order of the error operator $[\sum_{j=0}^n b_j(x, D_t, D_x, z)]^2 - A$ is $1 - n$ and can be taken arbitrary low [1, 22, 37]. This is a theoretical proposition we will not use.

Symmetric quantization

Operator $-\partial_t^2$ and multiplication with $\frac{1}{c(x,z)^2}$ are commutative positive operators. Because we restricted the domain of operator A to the propagation region, i.e. when $\xi^2 < \frac{\omega^2}{c^2}$, A is positive symmetric. Hence, the square root of A is expected to be symmetric, i.e. $(\sqrt{A})^* = \sqrt{A}$. Neglecting regularity conditions, we use symmetry and self-adjointness interchangeably.

In fact, operator $B_0 = \text{Op}(b_0)$ is asymmetric due to the standard left quantization. We therefore define the *symmetric* quantization by

$$\text{Op}_s(p) = \frac{1}{2} [\text{Op}_L(p) + \text{Op}_R(p)]. \quad (3.17)$$

The right quantization is closely related to the adjoint operator, which can be expressed by $\text{Op}(p)^* = \text{Op}_R(\bar{p})$. If p is real then $\text{Op}_s(p) = \frac{1}{2} [\text{Op}(p) + \text{Op}(p)^*]$ obviously is a symmetric operator. Because b_0 is real, operator $\text{Op}_s(b_0)$ is symmetric, considering the propagation region only. The claim is that the principal and subprincipal symbols of $\text{Op}_s(b_0)$ are identical to b_0 and b_1 , respectively. Note that p and the symbol of $\text{Op}_s(p)$ are not identical.

A general theorem says that the adjoint P^* of a pseudo-differential operator $P = \text{Op}(p)$ of order m is a pseudo-differential operator of the same order. The symbol of the adjoint is denoted by p^* and its asymptotic expansion can be found by a fundamental theorem of the symbolic calculus [37, 1]. Again, we present this theorem with variables x and ξ :

$$\begin{aligned} p^*(x, \xi) &\sim \sum_{k=0}^{\infty} \frac{1}{k!} \partial_{\xi}^k D_x^k \bar{p}(x, \xi) \\ &\sim \bar{p}(x, \xi) + \partial_{\xi} D_x \bar{p}(x, \xi) + r(x, \xi), \end{aligned} \quad (3.18)$$

in which $r(x, \xi)$ is a symbol of order $m - 2$.

By applying result (3.18), it can be found that

$$\text{Op}_S(b_0) = \frac{1}{2} \text{Op}(b_0 + b_0^*) = \text{Op}(b_0 + \frac{1}{2} \partial_{\xi} D_x b_0 + r'),$$

in which r' is a symbol of order -1 . By the observation that $b_0 D_x b_0 = \frac{\omega^2}{2} D_x \frac{1}{c(x,z)^2}$ does not depend on ξ , one concludes that

$$\partial_{\xi} D_x b_0 = -\frac{1}{b_0} \partial_{\xi} b_0 D_x b_0.$$

Looking back at (3.16), it can be concluded that the principal and the subprincipal symbols of $\text{Op}_S(b_0)$ are b_0 and b_1 , respectively.

To conclude with, the square root operator with symmetric quantization is therefore given by

$$\text{Op}_S(b_0) = \frac{1}{2} [\text{Op}_L(b_0) + \text{Op}_R(b_0)]. \quad (3.19)$$

The practical value of the symmetric quantization lies in the fact that it describes amplitudes without the need for lower order term b_1 . This term involves c and $\partial_x c$ and appears to be more complicated to compute than the principal part. The symmetry argument is general applicable to numerical schemes. In section 3.3 we will give two examples of symmetric implementations.

3.2.4 Normalization

We started this section with the derivation of the one-way wave equation for inhomogeneous media. The z -dependency of the velocity gives rise to matrix operator $\mathbf{\Lambda}(\partial_z \mathbf{V})$ in (3.6). The operators on the diagonal of $\mathbf{\Lambda}(\partial_z \mathbf{V})$ will appear as extra terms in the one-way wave equation, making it less attractive from the numerical point of view. We therefore incorporate a normalization factor in the change of variables that will cancel the diagonals [60].

Matrix differential equation (3.4) is the starting point. Fractional powers of A are to be read as pseudo-differential operators with polyhomogeneous symbols.

Normalization operators $A^{\frac{1}{4}}$ and $A^{-\frac{1}{4}}$ are real positive by definition.² We redefine matrix operators \mathbf{V} and $\mathbf{\Lambda}$ by including the normalization:

$$\mathbf{V} = A^{-\frac{1}{4}} \begin{pmatrix} 1 & 1 \\ i\sqrt{A} & -i\sqrt{A} \end{pmatrix} \quad \text{and} \quad \mathbf{\Lambda} = \frac{1}{2}A^{\frac{1}{4}} \begin{pmatrix} 1 & -iA^{-\frac{1}{2}} \\ 1 & iA^{-\frac{1}{2}} \end{pmatrix}. \quad (3.20)$$

This again yields an eigenvalue decomposition of the matrix operator \mathbf{A} given by $\mathbf{A} = \mathbf{V}\mathbf{B}\mathbf{\Lambda}$. Using these matrices, the change of variables is defined by $\mathbf{u} = \mathbf{V}\mathbf{v}$. The same procedure as in the begin of this section yields the equation

$$\partial_z \mathbf{v} = (\mathbf{B} - \mathbf{E}) \mathbf{v} - \mathbf{\Lambda} \mathbf{f}, \quad (3.21)$$

in which we used the definition $\mathbf{E} = \mathbf{\Lambda}(\partial_z \mathbf{V})$. Working out the details now gives

$$\mathbf{E} = -\frac{1}{4}A^{-1}(\partial_z A) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Due to the normalization, the diagonal terms are zero.

The off-diagonal entries of matrix operator \mathbf{E} are still of concern because they couple the up and down going waves in (3.21). Actually, there is a subtle argument that says that these terms can be neglected. We will show this in the next subsection. Again writing v for the second component of \mathbf{v} , *the one-way wave equation for propagation in the positive z -direction* is

$$\partial_z v(t, x, z) = -iB v(t, x, z) + \frac{1}{2}HA^{-\frac{1}{4}}f(t, x, z), \quad (3.22)$$

in which B can be either $\text{Op}_L(b_0 + b_1)$ or $\text{Op}_S(b_0)$. We have used the Hilbert transform $H := -i \text{Op}(\text{sgn}(\omega))$. Assuming the first component of \mathbf{v} to be zero, i.e. no up going wave, the wave field in terms of the original variable is given by

$$u(t, x, z) = A^{-\frac{1}{4}} v(t, x, z). \quad (3.23)$$

Comparing (3.22) with (3.7), the factor in the source term is changed due to the normalization.

High frequency decoupling

Following a procedure based on the work of Taylor [67], we show that the off-diagonal entries of \mathbf{E} can be neglected. The argument relies on the high frequency assumption of the theory.

This assumption is made concrete by stating that the one-way wave equation is *invariant* under the transformation of dependent variables given by $\mathbf{v} \leftrightarrow (\mathbf{I} + \mathbf{K})\mathbf{v}$,

²By setting their principal symbols to $a^{\frac{1}{4}}$ and $a^{-\frac{1}{4}}$ respectively. A consequence is that $A^{\frac{1}{4}}A^{\frac{1}{4}} = \text{Op}(\text{sgn}(\omega))A^{\frac{1}{2}}$.

in which all four entries of matrix operator \mathbf{K} are arbitrary pseudo-differential operators of order -1 . The pseudo-differential operators in equation (3.21) have polyhomogeneous symbols. Terms of their expansion that do not change under this set of transformations will be called *invariant*. A term that can be made zero is not invariant, of course.

We will show that there exists a transformation like mentioned above, such that the 0^{th} -order off-diagonals terms of \mathbf{E} become zero. By some knowledge in advance we define

$$\mathbf{K} = \frac{1}{8}A^{-\frac{3}{2}}(\partial_z A) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

which has operator entries of order -1 . The change of variables is defined by $\mathbf{v} = (\mathbf{I} + \mathbf{K})\tilde{\mathbf{v}}$. Inserting this into (3.21) and subtracting $(\partial_z \mathbf{K})\tilde{\mathbf{v}}$, yields

$$(\mathbf{I} + \mathbf{K})\partial_z \tilde{\mathbf{v}} = ((\mathbf{B} - \mathbf{E})(\mathbf{I} + \mathbf{K}) - (\partial_z \mathbf{K}))\tilde{\mathbf{v}} - \mathbf{A}\mathbf{f}.$$

Let \mathbf{R} be a matrix pseudo-differential operator such that $\mathbf{I} - \mathbf{K} + \mathbf{R}$ is a parametrix of $\mathbf{I} + \mathbf{K}$. A parametrix is an approximate inverse in the sense that the difference between the identity and the composition of an operator with its parametrix is an operator of arbitrary low order [37, 1]. It follows that the highest order of the operator entries of \mathbf{R} is -2 . By left multiplication with $\mathbf{I} - \mathbf{K} + \mathbf{R}$ we find

$$\partial_z \tilde{\mathbf{v}} = \begin{pmatrix} \mathbf{I} & & & \\ & -\mathbf{K} & & \\ & & \mathbf{R} & \\ & & & \mathbf{I} \end{pmatrix} ((\mathbf{B} - \mathbf{E})(\mathbf{I} + \mathbf{K}) - (\partial_z \mathbf{K}))\tilde{\mathbf{v}} - (\mathbf{I} - \mathbf{K} + \mathbf{R})\mathbf{A}\mathbf{f}.$$

The highest orders of the operator entries of the matrices in the homogeneous part of the equation are written below each term. Leaving out operator terms of order less than or equal to -1 , yields

$$\partial_z \tilde{\mathbf{v}} = (\mathbf{B} - \mathbf{E} + \mathbf{B}\mathbf{K} - \mathbf{K}\mathbf{B})\tilde{\mathbf{v}} - (\mathbf{I} - \mathbf{K} + \mathbf{R})\mathbf{A}\mathbf{f}. \quad (3.24)$$

Neglecting the commutator $[A^{-\frac{3}{2}}(\partial_z A), A^{\frac{1}{2}}]$ because its leading order is -1 , it can be found by straight forward calculation that

$$\mathbf{B}\mathbf{K} - \mathbf{K}\mathbf{B} = -\frac{1}{4}A^{-1}(\partial_z A) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This expression equals \mathbf{E} , and hence, they cancel each other in the homogeneous part of equation (3.24).

The change of variables formally shows that the 0^{th} -order off-diagonal terms in the homogeneous part of equation (3.21) can be neglected. The highest order terms of $\tilde{\mathbf{v}}$ and \mathbf{v} are identical, so to say. Likewise, the factor $(\mathbf{I} - \mathbf{K} + \mathbf{R})$ does not change the leading order of the source term and can be left out. Hence, the system of equations becomes decoupled:

$$\partial_z \mathbf{v} = \mathbf{B}\mathbf{v} - \mathbf{A}\mathbf{f}. \quad (3.25)$$

The subprincipal part of operator \mathbf{B} describes the wave amplitudes. Therefore, the 0^{th} -order terms of \mathbf{B} in (3.25) are expected to be invariant. It is very easy to confirm this. Let K_1, \dots, K_4 be arbitrary pseudo-differential operators of order -1 and set

$$\mathbf{K} = \begin{pmatrix} K_1 & K_2 \\ K_3 & K_4 \end{pmatrix}$$

to define a formal change of variables by $\mathbf{v} \leftrightarrow (\mathbf{I} + \mathbf{K})\mathbf{v}$. By analogous steps to find (3.24), equation (3.25) transforms into

$$\partial_z \mathbf{v} = (\mathbf{B} + \mathbf{BK} - \mathbf{KB})\mathbf{v} - \mathbf{\Lambda}f.$$

The highest order of the operators on the diagonal of commutator $\mathbf{BK} - \mathbf{KB}$ is -1 . This shows that the 1^{st} and 0^{th} -order terms of \mathbf{B} are invariant. Only the principal symbols of the operator entries of \mathbf{V} and $\mathbf{\Lambda}$ are invariant. The pseudo-differential operator $A^{-\frac{1}{4}}$ in the source term of (3.22) can therefore be replaced by $\text{Op}(a^{-\frac{1}{4}})$.

Pseudo-differential operator on WKB solution

A requirement for true-amplitude one-way wave equations is that they yield the same travel time and amplitude in a WKB expansion as the original wave equation. Fortunately, WKB theory has been generalized to hyperbolic *pseudo*-differential equations. If $p(y, D_y)$ is a pseudo-differential operator of order m with polyhomogeneous symbol, then there is an asymptotic series for $p(y, D_y)\alpha(y)e^{i\tau\theta(y)}$, i.e.

$$p(y, D_y)\alpha(y)e^{i\tau\theta(y)} = \sum_{j=0}^{\infty} \beta_j(y, \tau)e^{i\tau\theta(y)}, \quad (3.26)$$

such that $\beta_j(y, \tau) = O(\tau^{m-j})$, assuming $\alpha(y)$ and $\theta(y)$ are C^∞ -functions. See Duistermaat's book [22], equation (4.3.3). Although a detailed treatment falls outside the scope of this paper, we note that it follows that the first two terms in (3.26) give exactly the required eikonal and transport equations, and therefore, the same WKB solution.

3.3 Numerical implementation

The one-way wave equation (3.22) is an ODE with respect to depth z . The right hand side involves two pseudo-differential operators, namely square root B and $\text{Op}(a^{-\frac{1}{4}})$. We introduce a method to discretize pseudo-differential operators, sufficient general that both operators can be implemented. This method uses interpolation to rephrase the dependency on the velocity such that the operator can be written as a sum of multiplications and convolutions. It will be presented in the

following subsection. Other methods can be found in [47, 27, 45]. We solved the ODE by using the classical 4th-order Runge-Kutta scheme [41].

The (ω, ξ) -domain is split up in two regions. The inequality $\xi^2 < s^2\omega^2$ defines the *propagation* region and $\xi^2 > s^2\omega^2$ defines the *evanescent* region. By a *single-phase* wave we mean an uni-directional wave solution for a single (ω, ξ) -pair. Let $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ be the direction of propagation of a single-phase wave with respect to the z -axis. Then it holds that

$$\sin \phi = \frac{-\xi}{s(x,z)\omega},$$

which shows that the angle depends on the spatial coordinate. This also shows that the singularity of b_0 , i.e. when $\xi^2 = s^2\omega^2$ holds, coincides with horizontally propagating waves.

We are interested in propagating wave solutions. And in the numerical computation, the angle of propagation is bounded. With cutoff angle ϕ_c and maximal angle ϕ_m as parameters, we introduce the angular cutoff function $\psi : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [0, 1]$. If $|\phi| \leq \phi_c$ then $\psi = 1$, if $|\phi| \geq \phi_m$ then $\psi = 0$ and in between the function makes a continuously differentiable transition using a scaled version of $[0, 1] \ni \nu \mapsto \frac{1 - \cos(\pi\nu)}{2}$. Single-phase waves with $\phi \in [-\phi_c, \phi_c]$ are thus undisturbed.

3.3.1 Pseudo-spectral interpolation method

A pseudo-differential operator with constant coefficients, for example equation (3.8), can efficiently be implemented as a multiplication operator in the Fourier domain, using the fast Fourier transform. The integral (3.9) shows that, in general, the implementation of a pseudo-differential operator involves an inverse Fourier transform for each value of the space-time variable.

The numerical costs can be reduced by writing the operator as a discrete sum of multiplication and convolution operators. This idea is generally used. Bao and Symes, for example, also used this formula to implement pseudo-differential operators [2]. In terms of symbols, it is expressed as

$$p(y, \eta) \approx \sum_{k=1}^K g_k(y) h_k(\eta). \quad (3.27)$$

The implementation can be done termwise. The order of application of multiplication $\text{Op}(g_k)$ and convolution $\text{Op}(h_k)$ in each term is prescribed by the quantization. In left quantization the convolution comes first. The implementation of $\text{Op}_L(p)\varphi(y)$ will therefore be

$$\sum_{k=1}^K g_k(y) \mathcal{F}^{-1}(h_k(\eta)\widehat{\varphi}(\eta)), \quad (3.28)$$

in which $\widehat{\varphi} = \mathcal{F}\varphi$. This involves $1 + K$ (inverse) Fourier transforms. To implement the right quantization, one only has to interchange this order. Then $\text{Op}_R(p)\varphi(y)$ becomes

$$\mathcal{F}^{-1}\left(\sum_{k=1}^K h_k(\eta)\widehat{g_k\varphi}(\eta)\right). \quad (3.29)$$

The symmetric quantization (3.17) can therefore be implemented by using $2 + 2K$ (inverse) Fourier transforms. This is cost efficient for small K . In next paragraphs, we will discuss how to find functions g_k and h_k . Remember that we used the notation $y = (t, x)$ and $\eta = (\omega, \xi)$.

To apply this to operators B and $\text{Op}(a^{-\frac{1}{4}})$ we will use the fact that their symbols do not explicitly depend on x and z , and they do not at all depend on t . The symbols are functions of the spatially dependent velocity, which is slowly varying. To express this dependency, we will use the slowness $s(x, z) = \frac{1}{c(x, z)}$ rather than the velocity itself for a reason that will become clear.

A discrete set of slowness values is denoted by s_k with $k = 1, \dots, K$. We take the liberty of writing $g_k(s) = g_k(s(x, z))$ and the symbol of B as the function $b(s, \omega, \xi) = b(s(x, z), \omega, \xi)$. For each fixed s_k operator B ‘freezes’ to a convolution operator because its symbol, $b(s_k, \omega, \xi)$, is independent of x . To reformulate the symbol dependency on s , we use interpolation with respect to s and approximate the symbol by an interpolant of these frozen symbols:

$$b(s, \omega, \xi) \approx \sum_{k=1}^K g_k(s) b(s_k, \omega, \xi). \quad (3.30)$$

This formula requires that the interpolant is a linear function of the ordinates, i.e. the frozen symbols. We used linear interpolation.

As the weight functions g_k do not depend on ω and ξ , their definition can best be explained by thinking of ω and ξ being fixed. Let $\Pi_j(s)$ be the 1^{st} -order polynomial in s through the points $\{(s_i, b(s_i, \omega, \xi))\}_{i=j, j+1}$ with $j \in \{1, \dots, k-1\}$. The weight functions are such that the interpolant, i.e. the right hand side of (3.30), equals $\Pi_j(s)$ if $s_j \leq s \leq s_{j+1}$. This means that for an arbitrary value of the slowness, the interpolant is a weighted sum of the frozen symbols at the two neighboring discrete slowness values.

Various other interpolation methods can be used also like Hermite interpolation, Taylor series extrapolation and spline. Besides linear interpolation we tried piecewise Lagrange interpolation with cubical polynomials [41]. Let $\Pi_j(s)$ be the unique cubic polynomial in s through the points $\{(s_i, b(s_i, \omega, \xi))\}_{i=j, \dots, j+3}$ with $j \in \{1, \dots, k-3\}$. Then the weight functions are such that the interpolant equals $\Pi_j(s)$ if $s_{j+1} \leq s \leq s_{j+2}$. No improvement was found in the simulation.

Symmetric square root and normalization

We will apply the pseudo-spectral interpolation method to implement the symmetric square root operator (3.19). To avoid spurious oscillations in the neighborhood of the singularity, we regularize the principal symbol (3.8) by multiplying s^2 with $1 - i\epsilon(1 - \psi)$, using the angular cutoff function ψ . Here we use $\phi_c = 60^\circ$ and $\phi_m = 75^\circ$. Subsequently writing the slowness in front, this yields

$$b_{0,\epsilon}(s, \omega, \xi) = s\omega\sqrt{1 - \frac{\xi^2}{s^2\omega^2} - i\epsilon(1 - \psi)}. \quad (3.31)$$

The constant $\epsilon = 0.04$ is an empirically optimized parameter. Waves with propagation angle larger than the cutoff angle will thus be damped out. To conclude with, the square root operator is given by the symmetric quantization of the symbol:

$$\sum_{k=1}^K g_k(s(x, z)) b_{0,\epsilon}(s_k, \omega, \xi). \quad (3.32)$$

Note that the angular cutoff function inside the term $b_{0,\epsilon}(s_k, \omega, \xi)$ uses s_k to determine the angle.

For low values of ξ , the principal symbol (3.8) approaches a linear function with respect to the slowness and a hyperbola with respect to the velocity. Using the slowness in the interpolation, instead of the velocity, the error is expected to be lower. Because the slowness is slowly varying the number K can be kept low. The set of slowness values, i.e. $\{s_k\}_{k=1,\dots,K}$, are chosen on a logarithmic grid with increment factor $\gamma = 1.08$. For the same K , a logarithmic grid yields a smaller discretization error in (3.32) than a linear grid.

Figure 3.1 shows the symbol of the square root (3.32) in a dimensionless form. When $s = s_k$ for some k , the discretization error is absent. In that case the deviation of (3.32) from the single square root, i.e. $\nu \mapsto \sqrt{1 - \nu^2}$ [35], is caused by the regularization only. To visualize the worse case scenario, a large set of curves is shown.

Once the one-way wave equation is solved, the solution is transformed to the original variable by the pseudo-differential equation $u(t, x, z) = \text{Op}(a^{-\frac{1}{4}})v(t, x, z)$. This is implemented by the pseudo-spectral method too. Using the angular cutoff function ψ with $\phi_c = 50^\circ$ and $\phi_m = 90^\circ$ for regularization, the symbol of the operator becomes

$$\sum_{k=1}^K g_k(s(x, z)) \psi(s_k^2\omega^2 - \xi^2)^{-\frac{1}{4}}. \quad (3.33)$$

The angular cutoff function uses s_k to determine the angle. It suppresses single-phase waves with propagation angle greater than ϕ_c and evanescent waves. The weight functions $g_k(s)$ are the same as in (3.30), since they are determined by

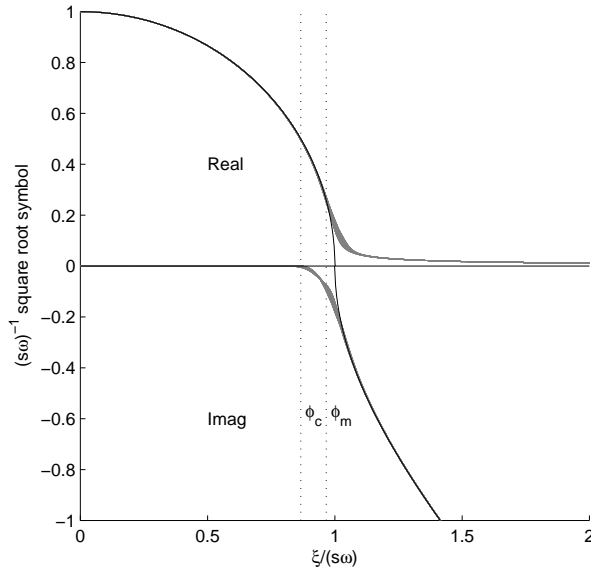


Figure 3.1: Real and imaginary part of the square root symbol (3.32) with $\phi_c = 60^\circ$, $\phi_m = 75^\circ$, $\gamma = 1.08$ and $\epsilon = 0.04$. The graph shows 50 curves, each for a specific slowness s that is randomly chosen between $2 \cdot 10^{-4}$ and $7 \cdot 10^{-4}$ s/m.

the set $\{s_k\}_{k=1,\dots,K}$. In accordance with the left quantization, the convolution is applied first. Figure 3.2 shows the symbol of the normalization operator (3.33) in the same way as for the square root.

3.3.2 Absorbing boundaries and stability

The domain of computation is bounded in the x -coordinate. Due to the discrete Fourier transform it becomes periodic. To suppress the periodicity, absorbing boundaries are obtained by two damping layers with thickness X_d . This is implemented by cutoff function $\chi_d : [-X_d, X + X_d] \rightarrow [0, 1]$. For $x \in [0, X]$ we set $\chi_d(x) = 1$. On $[-X_d, 0]$ and on $[X, X + X_d]$ the function makes a continuously differentiable transition using a scaled version of $[0, 1] \ni \nu \mapsto \frac{1 - \cos(\pi\nu)}{2}$. The damping will be applied simultaneously with solving the ODE. Because the damping should not depend on step size Δz , we apply the root

$$\chi_d(x) \frac{\Delta z}{Z_d} \quad (3.34)$$

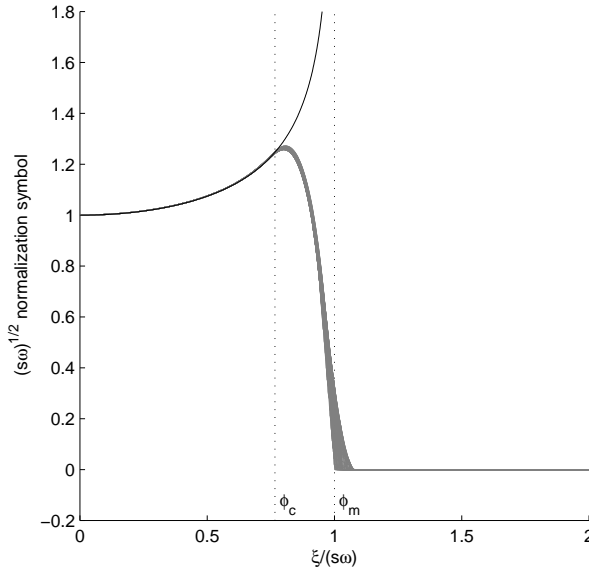


Figure 3.2: Real part of the normalization symbol (3.33) with $\phi_c = 50^\circ$, $\phi_m = 90^\circ$ and $\gamma = 1.08$. The graph shows 50 curves, each for a specific slowness s that is randomly chosen between $2 \cdot 10^{-4}$ and $7 \cdot 10^{-4}$ s/m.

as multiplication operator on the solution after each depth step of the ODE solver. Here, Z_d is the length scale of the damping, measured along the z -axis. This comprises a well-defined, i.e. Δz -independent, damping in the z -direction. For having comprehensible definitions for X_d and Z_d , we relate them to the maximal occurring wavelength λ_{\max} and the maximal propagation angle ϕ_m . We used the empirical lower bounds: $Z_d \geq \lambda_{\max}$ and $X_d \geq Z_d \tan \phi_m$ with $\phi_m = 75^\circ$.

Large negative eigenvalues give rise to stiffness problems [33, 34]. We approximate the eigenvalues of operator $-iB$ by its principal symbol $-ib_0$, using (3.8). With Δz the step size of the z -grid, the dimensionless number $-ib_0 \Delta z$ must be in the stability domain of the ODE solver. Using the 4th-order Runge-Kutta scheme, this yields the restrictive condition $\xi_{\max} \Delta z \leq 2.7853$. In our simulations holds $\xi_{\max} \Delta z = 3.1416$. Because we are interested in propagated waves, the stability problems are avoided by smoothly chopping off the negative real values of the square root symbol. Only the real part of $-ib_0$ that negatively exceeds a threshold value is modified. For the threshold we took $-\frac{\omega_{\max}}{c_{\min}} = -0.1571$.

3.3.3 Symmetric finite-difference one-way

One-way wave equations can be made true amplitude by using a symmetric square root operator and the wave field normalization specified in (3.23). In this section we will show how to construct a symmetric finite-difference one-way scheme in 2-D by giving an example. The modification concerns the order of application of multiplication operators and differentiation. The extra numerical cost is therefore small. We present a symmetric implementation of the 60 degree equation [14, 35]. Other approximations can be treated similarly.

The square root symbol is written as $\frac{\omega}{c}\sqrt{1-s^2}$ with operator symbol $s = \frac{c\xi}{\omega}$. This usage of letter s is restricted to this subsection. Elsewhere it refers to the slowness. To get a symmetric operator, we propose the following quantization. Split the factor $\frac{1}{c}$ into two by defining multiplication operator $M = c(x, z)^{-\frac{1}{2}}$. The square root operator is now written as the composition ωMSM , in which operator S is a symmetric quantization of the symbol $\sqrt{1-s^2}$, or an approximation of this symbol. Note that factors ω can be manipulated easily, as the calculation is done in the frequency domain.

Finite-difference methods rely on a rational approximation of $\sqrt{1-s^2}$ in terms of s^2 . The 60 degree approximation is

$$\sqrt{1-s^2} \approx \frac{1-s^2 + \frac{1}{8}s^4}{1 - \frac{1}{2}s^2} = 1 - \frac{1}{4}s^2 - \frac{\frac{1}{4}s^2}{1 - \frac{1}{2}s^2}. \quad (3.35)$$

Because it is a rational function of s^2 with constant coefficients, we only have to map s^2 into a symmetric operator. An obvious choice is to interpret s^2 as operator $Q = -\frac{1}{\omega^2}\partial_x c(x, z)^2 \partial_x$. The operator is to be interpreted asymptotically, i.e. for large ω , and because numerator and denominator commute, the quotient in approximation (3.35) is well-defined.

Application of our proposal leads to the following symmetric form of the one-way wave equation:

$$\partial_z u = i\omega M \left(-1 + \frac{1}{4}Q + \frac{\frac{1}{4}Q}{1 - \frac{1}{2}Q} \right) M u. \quad (3.36)$$

We will explain how to calculate $u(z + \Delta z)$ from $u(z)$. The one-way wave equation involves three operator terms, dividing the implementation into three sub steps. Each operator term is calculated separately, using operator splitting.

The second and third terms are implemented with a slightly modified Crank-Nicolson method that can be framed in the following scheme with temporary variables \tilde{u} , N and D :

$$\partial_z \tilde{u} = M \frac{N}{D} M \tilde{u}(z) \rightarrow \frac{\tilde{u}(z + \Delta z) - \tilde{u}(z)}{\Delta z} = M \frac{N}{D} M \frac{\tilde{u}(z + \Delta z) + \tilde{u}(z)}{2}. \quad (3.37)$$

Operator $M\frac{N}{D}M$ is, unlike Crank-Nicolson, evaluated at an intermediate grid point $z + \frac{1}{2}\Delta z$. Due to the fact that the velocity is slowly varying, the associated discretization error is negligible. The lateral discretization is done with a central difference. This will be shown in a moment. Rewriting discretization formula (3.37) and left multiplying with DM^{-1} yield the result:

$$(DM^{-1} - \frac{1}{2}\Delta z NM) \tilde{u}(z + \Delta z) = (DM^{-1} + \frac{1}{2}\Delta z NM) \tilde{u}(z). \quad (3.38)$$

We first apply this to the third operator in the expression of the one-way wave equation (3.36), i.e. with $N = i\omega\frac{1}{4}Q$ and $D = 1 - \frac{1}{2}Q$. The first intermediate result u_1 is found by solving:

$$([1 - \frac{1}{2}Q]M^{-1} - \frac{1}{8}\Delta z i\omega QM) u_1 = ([1 - \frac{1}{2}Q]M^{-1} + \frac{1}{8}\Delta z i\omega QM) u(z). \quad (3.39)$$

Then the second operator is to be calculated. We set $N = i\omega\frac{1}{4}Q$ and $D = 1$. The second intermediate result u_2 is given by the solution of:

$$(M^{-1} - \frac{1}{8}\Delta z i\omega QM) u_2 = (M^{-1} + \frac{1}{8}\Delta z i\omega QM) u_1. \quad (3.40)$$

The final step is the calculation of the first operator, which gives a phase shift. The result is our numerical approximation of $u(z + \Delta z)$:

$$u(z + \Delta z) = e^{-i\frac{\omega}{c}\Delta z} u_2, \quad (3.41)$$

in which the velocity c is conformably evaluated at $z + \frac{1}{2}\Delta z$.

We used the central difference formula to discretize M and Q laterally. As the calculation is done in the frequency domain, we may equally discuss how to implement $-\omega^2 Q = \partial_x c(x, z)^2 \partial_x$. By the numerical scheme in (3.37) the velocity is evaluated at $z + \frac{1}{2}\Delta z$, so we omit the depth argument now. Written in temporary variable \tilde{u} , the discrete form of $\partial_x c(x)^2 \partial_x \tilde{u}(x)$ is

$$\frac{c(x + \frac{1}{2}\Delta x)^2 [\tilde{u}(x + \Delta x) - \tilde{u}(x)] - c(x - \frac{1}{2}\Delta x)^2 [\tilde{u}(x) - \tilde{u}(x - \Delta x)]}{\Delta x^2}. \quad (3.42)$$

The value of the velocity at an intermediate grid point is approximated by the average of the values at the two closest grid points. In formula this comes to be $c(x + \frac{1}{2}\Delta x)^2 \approx \frac{1}{2}[c(x)^2 + c(x + \Delta x)^2]$. Because c is smooth the error is small.

Together with (3.42) the numerical formulae in (3.39) and (3.40) are tridiagonal linear systems that are solved successively. The exponential in (3.41) then gives the wave field at next depth. The system is closed with homogeneous Dirichlet boundary conditions at $x = -X_d$ and $x = X + X_d$. The effect of these hard walls is abolished by the absorbing boundaries discussed above.

It is interesting to make a comparison with the work of Zhang et al. in [76, 77]. They interpreted s^2 as the asymmetric operator $-\frac{1}{\omega^2}(c\partial_x)^2$, see [76] appendix A.

However, they did not split the $\frac{1}{c}$ multiplication in two, but kept it on the left side. Careful examination reveals that their square root operator equals ωMSM if we interpret s^2 as $-\frac{1}{\omega^2} c^{\frac{1}{2}} \partial_x c \partial_x c^{\frac{1}{2}}$, which is an symmetric alternative. Their true-amplitude scheme can hence be derived from our symmetrization procedure.

The symmetric implementation will be used in wave propagation simulations. We make a comparison with an asymmetric implementation of the 60 degree equation. In the asymmetric variant we write coefficients on the left of the derivatives, replacing formulae (3.39) and (3.40) with

$$\left(1 + \frac{1}{2} \frac{c^2}{\omega^2} \partial_x^2 + \frac{1}{8} \Delta z i \frac{c}{\omega} \partial_x^2\right) u_1 = \left(1 + \frac{1}{2} \frac{c^2}{\omega^2} \partial_x^2 - \frac{1}{8} \Delta z i \frac{c}{\omega} \partial_x^2\right) u(z) \quad (3.43)$$

and

$$\left(1 + \frac{1}{8} \Delta z i \frac{c}{\omega} \partial_x^2\right) u_2 = \left(1 - \frac{1}{8} \Delta z i \frac{c}{\omega} \partial_x^2\right) u_1. \quad (3.44)$$

Lateral operators are straightforwardly discretized with a central difference method.

3.4 Numerical results

As was theoretically shown in section 3.2, the one-way wave equation describes unidirectional waves with correct amplitude if the symmetric square root operator $\text{Op}_s(b_0)$ is used and the normalization of the wave field $u = A^{-\frac{1}{4}} v$. We numerically verify this by simulating wave propagation through media with a linearly varying velocity. We first use the implementation of the pseudo-spectral method, both with symmetric and asymmetric square root and compare the results. A similar experiment is done with and without normalization, using respectively (3.7) and (3.22). We subsequently repeat the first simulations, now using the finite-difference one-way scheme to implement the symmetric and asymmetric square root operator. The solution of the full wave equation act as a reference. It is implemented by a finite-difference time domain solver. Finally, a simulation with a more complex velocity function is done.

The source term in our model is an approximated delta function given by

$$f(t, x, z) = \delta(t - t_s) \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2} \frac{(x-x_s)^2 + (z-z_s)^2}{\sigma^2}}.$$

This source is used in all simulations. It is located at $(t_s, x_s, z_s) = (0\text{ s}, 1000\text{ m}, 0\text{ m})$ in time and space. The radius of the circular shaped cross section at half height is $\sigma\sqrt{2\ln 2}$ with $\sigma = 25$ m. To regularize the intrinsic singularity with respect to time, a convolution operator is used to cancel high and low frequencies. The frequency band 20 to 30 Hz is undisturbed. From 20 to 10 Hz and from 30 to 50 Hz the source is gradually suppressed to zero.

As was mentioned before, operator $\text{Op}(a^{-\frac{1}{4}})$ is also involved in the source term of the one-way wave equation (3.22). The source we used in the modeling has

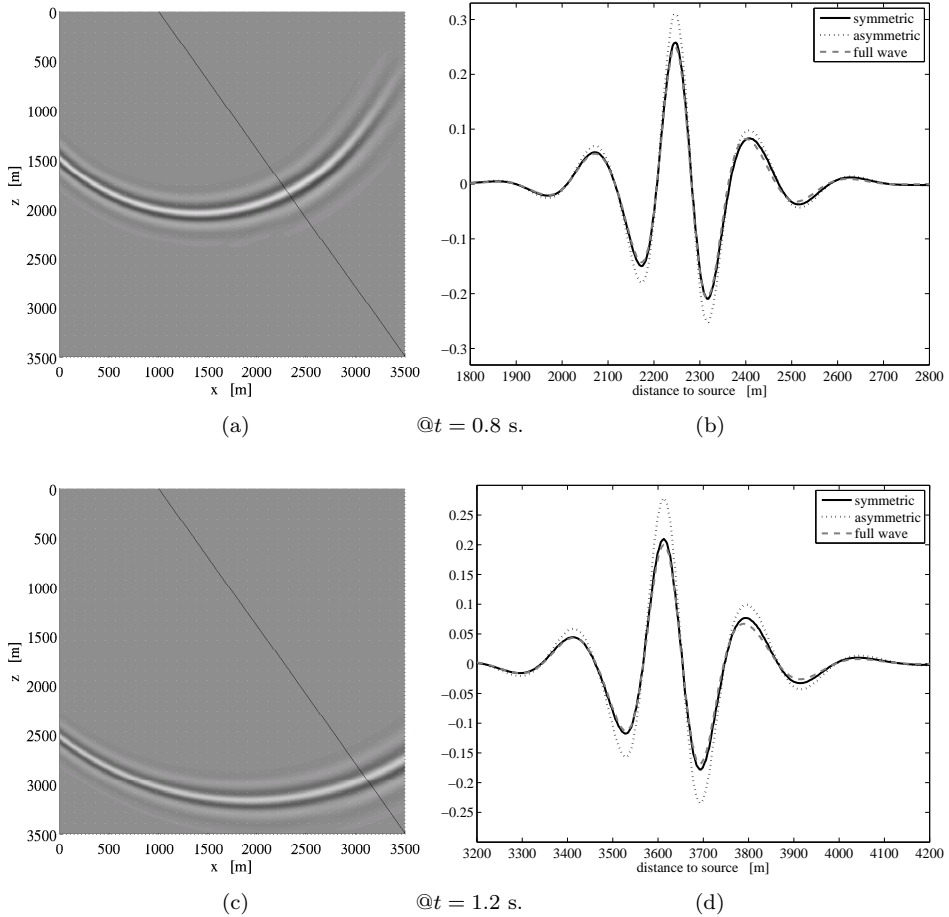


Figure 3.3: Left: Snapshots of wave field $v(t, x, z)$ simulated with pseudo-spectral method. The black line indicates where the cross section is taken. Right: Cross section of snapshot. Solid line: symmetric square root. Dotted line: asymmetric square root. Dashed line: full wave equation. The velocity is $c = 2000 + 0.5x$.

a small support in x and z . In the implementation of the operator we therefore approximate the slowness by a constant, i.e. the average value within the support.

The effect of the symmetric square root is verified in a simulation with the lateral varying velocity function $c = 2000 + 0.5x$. Figure 3.3(a) and (c) show snapshots of the simulated wave field $v(t, x, z)$ at times $t = 0.8$ and 1.2 s. These are simulated with the one-way wave equation and the symmetric square root

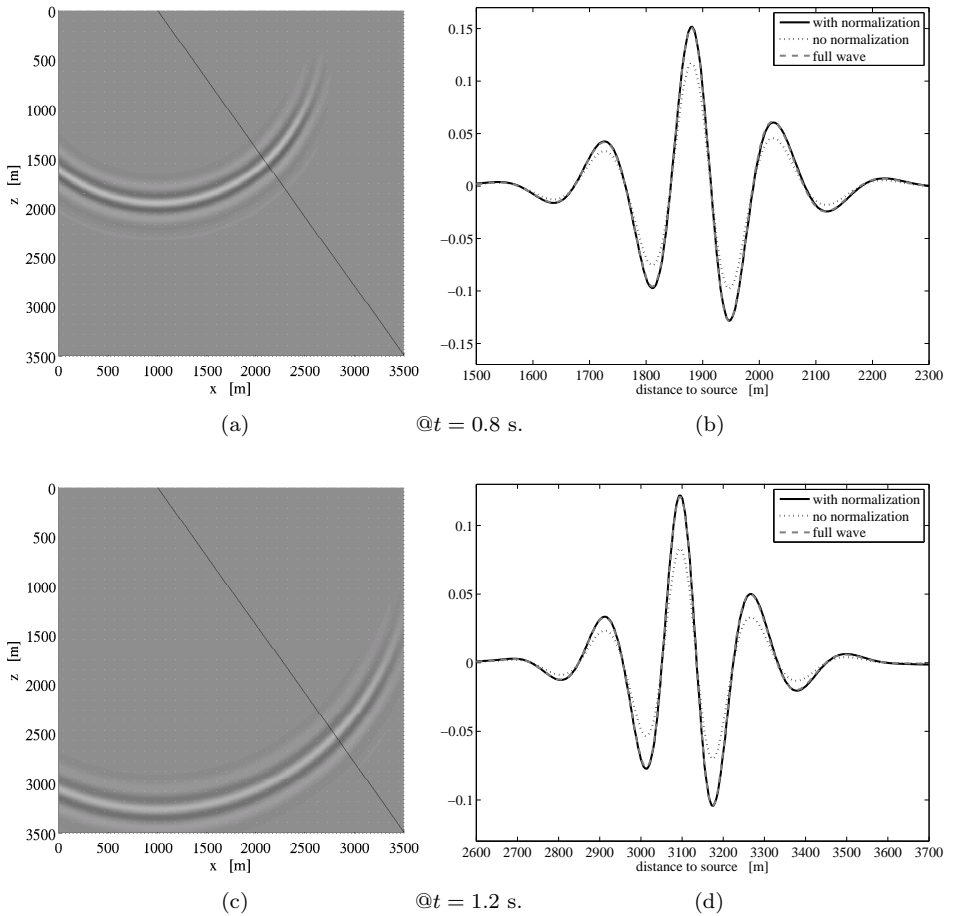


Figure 3.4: Left: Snapshots of wave field $v(t, x, z)$ simulated with pseudo-spectral method. The black line indicates where the cross section is taken. Right: Cross section of snapshot. Solid line: with normalization. Dotted line: without normalization. Dashed line: full wave equation. The velocity is $c = 2000 + 0.5z$.

operator, i.e. equation (3.22) with (3.19). The straight black lines indicate where a cross section is taken.

Figure 3.3(b) and (d) show the cross section of the snapshot as function of the distance to the source, zoomed in on the wave front. Besides the cross section plotted with a solid line, it also shows the analogous curves from other numerical simulations. The dotted line is calculated with the one-way wave equation and

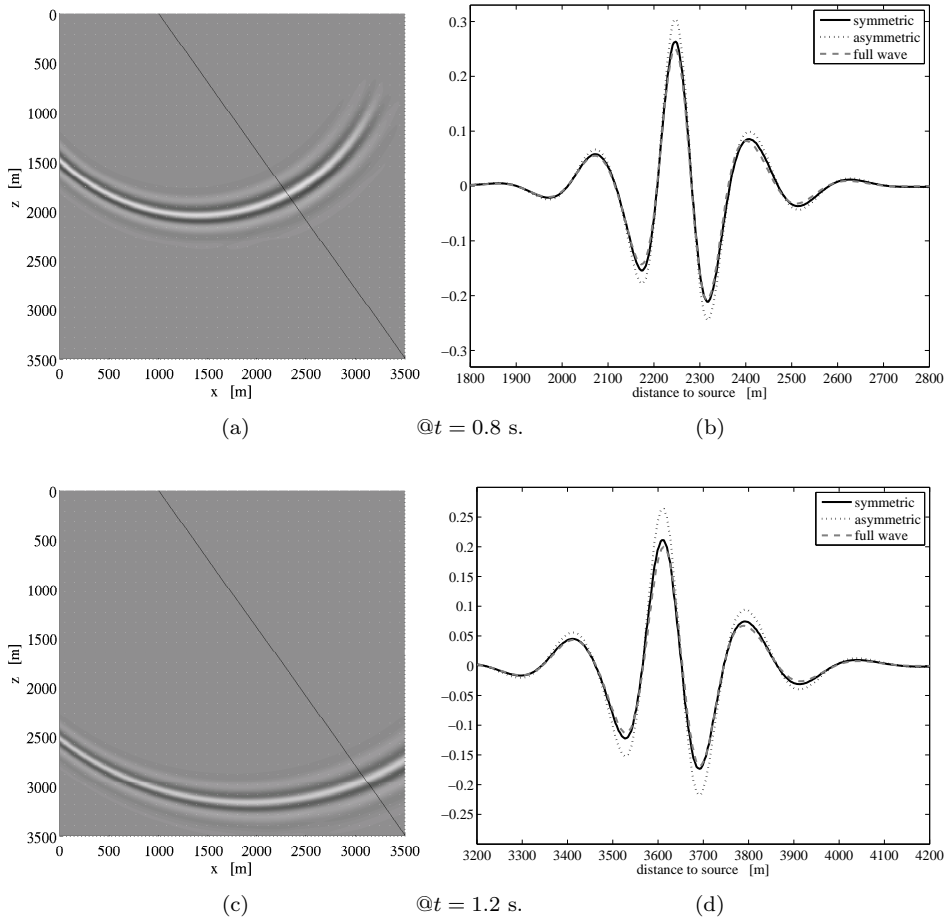


Figure 3.5: Left: Snapshots of wave field $v(t, x, z)$ simulated with finite-difference one-way as depth stepper. The black line indicates where the cross section is taken. Right: Cross section of snapshot. Solid line: symmetric implementation square root. Dotted line: asymmetric square root. Dashed line: full wave equation. The velocity is $c = 2000 + 0.5x$.

the asymmetric square root operator, i.e. $B_0 = \text{Op}(b_0)$. The dashed line is from a simulation of the full wave equation. With respect to the full wave simulation at $t = 1.2$ s, the amplitude from the asymmetric implementation is 39% larger. With the symmetric implementation this is reduced to 5%.

The effect of the normalization is verified in simulations with velocity function $c = 2000 + 0.5z$. Figure 3.4(a) and (c) show snapshots of the simulated wave field

$v(t, x, z)$ at times $t = 0.8$ and 1.2 s. These are simulated with the one-way wave equation with normalization (3.22). Again, straight black lines are indications of cross sections.

Figure 3.4(b) and (d) show the cross section of the snapshot as function of the distance to the source, analogous to the previous figure. Besides the cross section plotted with a solid line, it also shows other numerical results. The dotted line is calculated with the one-way wave equation without normalization, equation (3.7). In both cases, the symmetric square root operator (3.19) is used. The dashed line is again from a simulation of the full wave equation. Again compared with the full wave simulation at $t = 1.2$ s, the amplitude from the implementation without normalization is 31% smaller. With normalization, the amplitude is 1% larger with respect to the result of the full wave simulation. The cross section graphs coincide almost completely.

We derived a symmetric and an asymmetric version of the 60 degree finite-difference equation. With these, we simulated wave propagation with velocity function $c = 2000 + 0.5x$. The normalization operator is still implemented by the pseudo-spectral method. Figure 3.5(a) and (c) show snapshots of the simulated wave field $v(t, x, z)$ at times $t = 0.8$ and 1.2 s. Figure 3.5(b) and (d) show cross sections of several simulations. The dotted line is calculated with the asymmetric square root, the solid line with the symmetric one. The dashed line is from the full wave simulation. With respect to the full wave simulation at $t = 1.2$ s, the amplitude from the asymmetric implementation is 32% larger. With the symmetric implementation this is again reduced to 5%.

Our last simulation has a more complex velocity function, consisting of a circular smooth blob of increased velocity superimposed on a velocity that linearly increases with depth, see figure 3.6. We compare the pseudo-spectral method with the finite-difference implementation of the full wave equation. Figure 3.7 shows the wave fields $v(t, x, z)$ at times $t = 0.9$ s and $t = 1.3$ s for both methods and the cross sections on the indicated diagonal line. The solid black line represents the one-way wave, the dashed gray the full wave. One clearly notices the absence of wide angle wave propagation in the one-way wave simulation. This is of course a consequence of using a one-way method. Apart from this the similarity is quite good, although some small differences remain as can be seen in the bottom graphs of figure 3.7.

3.5 Discussion

The above examples show that the symmetrization of the square root operator and the normalization of the wave field together yield a one-way wave method with correct amplitudes. The theoretical basis is that symmetrization and normalization

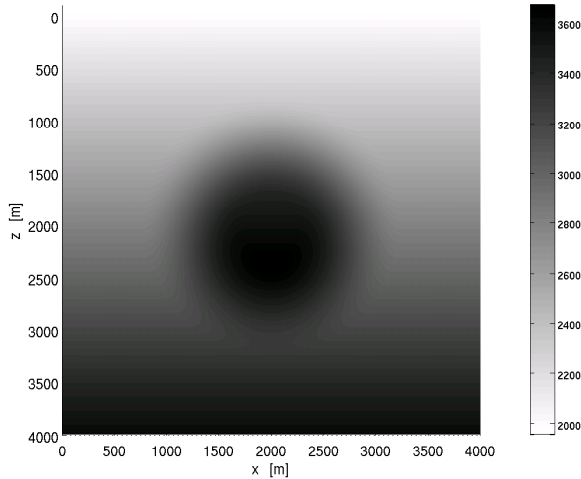


Figure 3.6: Velocity model $c = c(x, z)$ in m/s.

respectively remove the following two correction terms from the equation:

$$-ib_1 = \frac{\xi \partial_x c}{2\omega} \left(1 - \frac{c(x,z)^2 \xi^2}{\omega^2}\right)^{-\frac{3}{2}} \quad (3.45)$$

correcting the asymmetric square root, and

$$-\frac{1}{4}A^{-1}(\partial_z A) = \frac{\partial_z c}{2c} \left(1 - \frac{c(x,z)^2 \xi^2}{\omega^2}\right)^{-1} \quad (3.46)$$

changing the amplitude in depth dependent media. As an alternative we could have calculated these terms explicitly.

The factor $(\cos \phi)^{-3}$ in correction term (3.45) has a strong singularity making it difficult to implement. The symmetry argument on the other hand, can easily be applied in numerical schemes by properly choosing the order of multiplications and derivatives. And the symmetrization of the finite-difference one-way method leads to almost no extra cost. The implementation of correction term (3.46) has the same objection, and because it is to be applied at every depth step, this gives an accumulation of errors. The normalization operator on the other hand, has the advantage that it is to be applied only at source and at depths where output is written, not in every depth step.

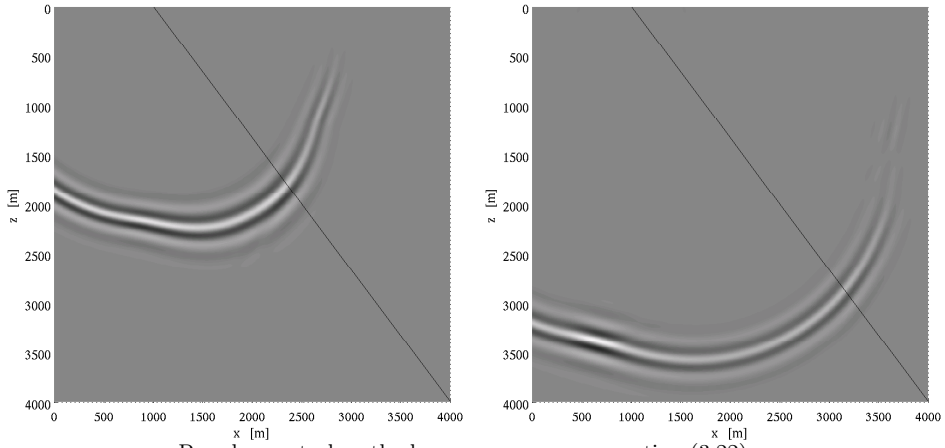
We did not make a detailed analysis of the computational cost of the pseudo-spectral interpolation method. For depth stepping it appears to be expensive due

to the relatively small step size taken in the explicit Runge-Kutta solver. It clearly approximates the pseudo-differential operator and has the advantage to be applicable for both the square root and the normalization operator. Further study could be valuable, for example to find a fast implementation of the normalization.

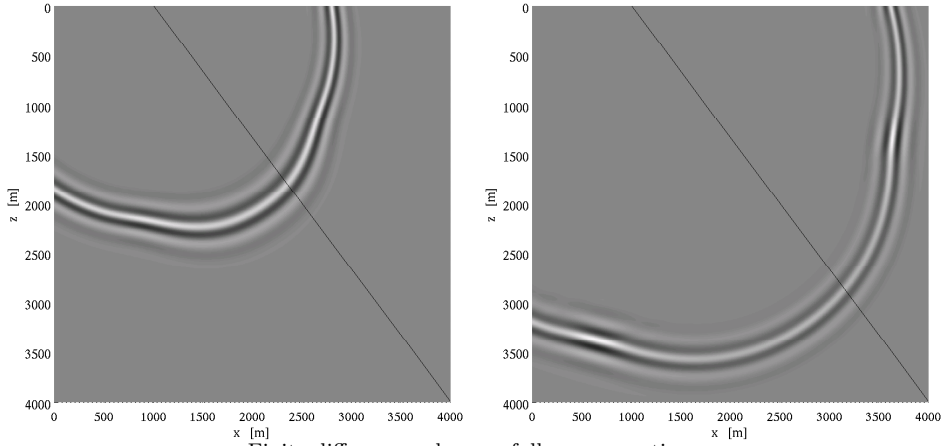
We found that the normalization and the angular cutoff are sensitive to details of the implementation, for example the choice of cutoff angle ϕ_c . Poor parameter choices may lead to artifacts. A better understanding could be an interesting motivation for further study.

Acknowledgement

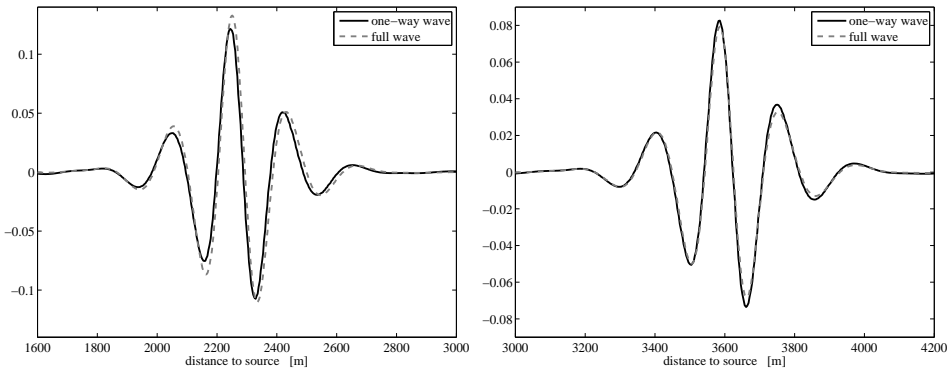
This research was supported by the Netherlands Organisation for Scientific Research through VIDI grant 639.032.509.



Pseudo-spectral method on one-way wave equation (3.22).



Finite-difference solver on full wave equation.



Cross sections of both.

Figure 3.7: Snapshots of simulated wave field $v(t, x, z)$. Left column: @ $t = 0.9$ s. Right column: @ $t = 1.3$ s. The black line in the first four graphs indicates where the cross section is taken. The velocity is given in figure 3.6.

Chapter 4

Inverse scattering based on reverse-time migration

Abstract.¹ In this paper we study the linearized inverse problem associated with imaging of reflection seismic data. We introduce an inverse scattering transform derived from reverse-time migration (RTM). In the process, the explicit evaluation of the so-called normal operator is avoided, while other differential and pseudo-differential operator factors are introduced. We prove that, under certain conditions, the transform yields a partial inverse, and support this with numerical simulations. In addition, we explain the recently discussed ‘low-frequency artifacts’ in RTM, which are naturally removed by the new method.

4.1 Introduction

In reflection seismology one places point sources and point receivers on the earth’s surface. A source generates acoustic waves in the subsurface, which are reflected where the medium properties vary discontinuously. In seismic imaging, one aims to reconstruct the properties of the subsurface from the reflected waves that are observed at the surface [14, 7, 66]. There are various approaches to seismic imaging, each based on a different mathematical model for seismic reflection data with underlying assumptions. In general, seismic scattering and inverse scattering have been formulated in the form of a linearized inverse problem for the medium coefficient in the acoustic wave equation. The linearization is around a smoothly varying

¹This chapter is almost exactly the article Linearized inverse scattering based on seismic Reverse Time Migration by T.J.P.M. Op ’t Root, C.C. Stolk, and M.V. de Hoop, published in ArXiv e-prints, December 2010. <http://arxiv.org/abs/1012.4278v2>

background, called the velocity model, which is a priori also unknown. However, in the inverse scattering setting considered here, we assume the background model to be known. The linearization defines a single scattering operator mapping the model contrast (with respect to the background) to the data, that consists of the restriction to the acquisition set of the scattered field. The adjoint of this map defines the process of imaging in general. The composition of the imaging operator with the single scattering operator yields the so-called normal operator, the properties of which play a central role in developing an inverse scattering procedure.

There are different types of seismic imaging methods. And one can distinguish methods associated with the evolution of waves and data in time from those associated with the evolution in depth, or another principal spatial direction. The first category contains approaches known under the collective names of Kirchhoff migration [8] or generalized Radon transform inversion, and reverse-time migration (RTM) [57, 71, 48, 3, 65]. The second category comprises the downward continuation approach [15, 14, 6, 53, 38] possibly applied in curvilinear coordinates. The analysis pertaining to inverse scattering in the second category can be found in Stolk and De Hoop [62, 63]. The subject of the present paper is an analysis of RTM-based inverse scattering in the first category, with a view to studying the reconstruction of singularities in the contrast. As was done in the analysis of Kirchhoff methods [4, 54, 44, 61], we make use of techniques and concepts from microlocal analysis, and Fourier integral operators (FIOs); see, e.g., [22] for background information on these concepts. As through an appropriate formulation of the wave field continuation approach, we arrive at a representation of RTM in terms of a FIO associated with a canonical graph. Over the past few years, there has been a revived interest in reverse-time migration (RTM), partly because their application has become computationally feasible. RTM is attractive as an imaging procedure because it avoids approximations derived from asymptotic expansions or one-way wave propagation.

The study of the above mentioned normal operator takes into account the available source-receiver acquisition geometry. To avoid the generation of artifacts, one has to invoke the Bolker condition [30], essentially ensuring that the normal operator is a pseudo-differential operator. (In reflection seismology this condition is sometimes referred to as traveltime injectivity condition [50].) RTM is based on a common source geometry, in which case the Bolker condition requires the absence of ‘source caustics’, that is, caustics are not allowed to occur between the source and the image points under consideration [50]. We shall refer to the assumption of absence of source caustics as the source wave multipath exclusion (SME). Additionally, we require that there are no rays connecting the source with a receiver position, which we refer to as the direct source wave exclusion (DSE), and we exclude grazing rays that originate in the subsurface. These conditions can be satisfied by removing the corresponding part of the wavefield by means of

pseudo-differential cutoffs.

In this paper we revisit the original reverse-time imaging procedure. We do this, also, in the context of the integral formulation of Schneider [56] and the inverse scattering integral equation of Bojarski [9]. An RTM migration algorithm consists of three main parts: The modeling of source wave propagation in forward time, the modeling of receiver or reflected wave propagation in reverse time, and the application of the so called imaging condition [14, 7]. The imaging condition is a map that takes as input the source wave field and the backpropagated receiver wave field, and maps these to an image. The imaging condition is based on Claerbout's [13] imaging principle: Reflectors exist in those points in the subsurface where the source and receiver wave fields both have a large contribution at coincident times.

Various imaging conditions have been developed over the past 25 years. The excitation time imaging condition identifies the time that the source field passes an image point, for example, using its maximum amplitude, and evaluates the receiver field at that time. The image can be normalized by dividing by the source amplitude. Alternatively, the image can be computed in the temporal frequency domain by dividing the receiver field by the source field and integrating over frequency, the ratio imaging condition. To avoid division by small values of the source field, regularization techniques have been applied. An alternative is the crosscorrelation imaging condition, in which the product of the fields is integrated over time. Later other variants have been proposed, see e.g. [10, 12, 43]. The authors of [43] use the spatial derivatives of the fields, similarly to what we find in this work.

We introduce a parametrix for the linearized scattering problem on which RTM is based. The explicit evaluation of the normal operator is avoided, at the cost of introducing other pseudo-differential operator factors in the procedure, which is, thus, different from Least-Squares migration-based approaches [52]. The method involves a new variant of the ratio imaging condition that involves time derivatives of the fields and their spatial gradients. The ratio imaging condition, albeit a new variant, is hence finally provided with a mathematical proof. The result is summarized in Theorem 6. As an intermediate result, we also obtain a new variant of the so called excitation time imaging condition in Theorem 7. Moreover, we also address the relation with RTM 'artifacts' [75, 49, 25, 74, 31], as well as certain simplifications that occur when dual sensor streamer data are available.

The seismic waves are governed by the acoustic wave equation with constant density on the spatial domain \mathbb{R}^n with $n = 1, 2, 3$, given by

$$\left[\frac{1}{c(\mathbf{x})^2} \partial_t^2 - \Delta \right] u(\mathbf{x}, t) = f(\mathbf{x}, t). \quad (4.1)$$

Although the subsurface is represented by the half space $\mathbb{R}^{n-1} \times [0, \infty)$, we carry out our analysis in the full space, \mathbb{R}^n . The acquisition domain is a subset of the surface $\mathbb{R}^{n-1} \times \{0\}$. The slowly varying velocity is a given smooth function $c(\mathbf{x})$. The existence, uniqueness and regularity of solutions can be found in [37]. We use the

following Fourier transform: $\mathcal{F}u(\boldsymbol{\xi}, \omega) = \iint e^{-i(\boldsymbol{\xi} \cdot \mathbf{x} + \omega t)} u(\mathbf{x}, t) d\mathbf{x} dt$, and sometimes write \hat{u} for $\mathcal{F}u$.

The outline of the paper is as follows. In section 4.2, solutions of the wave equation are discussed, starting from the WKB approximation with plane-wave initial values. The (forward) scattering problem is analyzed in section 4.3. We focus on the map from the contrast, or ‘reflectivity’, to what we refer to as the continued scattered field, which is the result from a perfect backpropagation of the scattered field from its Cauchy values at some time after the scattering has taken place. We obtain an explicit expression which is locally valid, and a global characterization as a Fourier integral operator. In section 4.4 we study the revert operator, which describes the backpropagation of the receiver field. The relation with the continued scattered field is established. The inversion, that is, parametrix construction, is presented in section 4.5. We first carry out a brief analysis of the case of a constant velocity. Then we introduce a novel version of the excitation time imaging condition and show that it yields an inversion. Following that, we present an imaging condition expressed entirely in terms of the source and backpropagated receiver fields, providing the RTM based linearized inversion. In section 4.6 we show some numerical tests. We end the paper with a short discussion.

4.2 Asymptotic solutions of initial value problem

In this section, we study solutions of the wave equation with smooth coefficients. We introduce explicit expressions for the solution operator for wave propagation over small times. In subsection 4.2.1 we construct an approximate solution of the IVP of the homogeneous wave equation. Using the WKB approximation we introduce phase and amplitude functions, which are solved by the method of characteristics in subsections 4.2.2 and 4.2.3. The asymptotic solution is finally written as a FIO in subsection 4.2.4. Subsection 4.2.5 presents the decoupling of the wave equation and general solution operators. Subsection 4.2.6 deals with the source field problem of RTM.

4.2.1 WKB approximation with plane-wave initial values

Before solving (4.1) we normalize the wave field and use the variable $c^{-1}u$ instead. We consider the equivalent wave equation

$$[\partial_t^2 - c\Delta c](c^{-1}u) = 0. \quad (4.2)$$

In the later analysis it will be advantageous that $c\Delta c$ is a symmetric operator. To find a solution we invoke the WKB ansatz

$$c^{-1}u(\mathbf{x}, t) = a(\mathbf{x}, t)e^{i\lambda\alpha(\mathbf{x}, t)}. \quad (4.3)$$

A straightforward calculation yields

$$\begin{aligned} e^{-i\lambda\alpha}[\partial_t^2 - c\Delta c] a e^{i\lambda\alpha} &= -\lambda^2 a [(\partial_t\alpha)^2 - c^2|\nabla\alpha|^2] \\ &\quad + i\lambda[2(\partial_t a)\partial_t\alpha + a\partial_t^2\alpha - 2c\nabla(ca) \cdot \nabla\alpha - c^2 a\Delta\alpha] \\ &\quad + \partial_t^2 a - c\Delta(ca). \end{aligned} \quad (4.4)$$

One obtains an approximate solution of the form (4.3) by requiring first that the $O(\lambda^2)$ -term vanishes, resulting in an eikonal equation for α , and secondly that the $O(\lambda)$ -term also vanishes, resulting in a transport equation for a . We present the equations respectively in (4.7) and (4.8). In section 4.2.4 we comment on the lower order terms, $O(\lambda^j)$ for $j \leq 0$.

We solve (4.2) with the plane-wave initial values:

$$u(\mathbf{x}, 0) = 0, \quad c(\mathbf{x})^{-1}\partial_t u(\mathbf{x}, 0) = e^{i\mathbf{x}\cdot\xi}. \quad (4.5)$$

The role of λ is here played by $|\xi|$. The WKB type solution of the initial value problem will contain two terms, i.e., the ansatz becomes

$$c^{-1}u(\mathbf{x}, t) = a(\mathbf{x}, t; \xi) e^{i\alpha(\mathbf{x}, t; \xi)} + b(\mathbf{x}, t; \xi) e^{i\beta(\mathbf{x}, t; \xi)}. \quad (4.6)$$

The reason is that there is a sign choice in the equation for the phase, leading to two eikonal equations

$$\partial_t\alpha + c|\nabla\alpha| = 0 \quad \text{and} \quad \partial_t\beta - c|\nabla\beta| = 0. \quad (4.7)$$

Here, α covers the negative frequencies and β the positive ones. The transport equations can concisely be written in terms of a^2 and b^2 . They are

$$\partial_t(a^2\partial_t\alpha) - \nabla \cdot (a^2c^2\nabla\alpha) = 0 \quad \text{and} \quad \partial_t(b^2\partial_t\beta) - \nabla \cdot (b^2c^2\nabla\beta) = 0. \quad (4.8)$$

The WKB ansatz (4.6) can be inserted into the initial conditions (4.5). This straightforwardly yields initial conditions for α and β :

$$\alpha(\mathbf{x}, 0; \xi) = \beta(\mathbf{x}, 0; \xi) = \xi \cdot \mathbf{x}. \quad (4.9)$$

The initial conditions for a and b can be given in the form of a matrix equation:

$$\begin{pmatrix} 1 & 1 \\ -ic(\mathbf{x})|\xi| & ic(\mathbf{x})|\xi| \end{pmatrix} \begin{pmatrix} a(\mathbf{x}, 0; \xi) \\ b(\mathbf{x}, 0; \xi) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (4.10)$$

The two terms in (4.6) are not independent. The initial value problem for α can be transformed into the initial value problem for β by replacing ξ with $-\xi$ and setting $\beta(\mathbf{x}, t; \xi) = -\alpha(\mathbf{x}, t; -\xi)$. Further analysis shows that $b(\mathbf{x}, t; \xi) e^{i\beta(\mathbf{x}, t; \xi)}$ in (4.6) is in fact the complex conjugate of $a(\mathbf{x}, t; -\xi) e^{i\alpha(\mathbf{x}, t; -\xi)}$.

4.2.2 The phase function on characteristics

The method of characteristics [23, section 3.2] will be used to solve the eikonal and transport equations. We will first solve the initial value problem for the phase function $\alpha(\mathbf{y}, t; \boldsymbol{\xi})$ as given in the form:

$$\begin{aligned} \partial_t \alpha(\mathbf{y}, t; \boldsymbol{\xi}) + c(\mathbf{y}) |\nabla \alpha(\mathbf{y}, t; \boldsymbol{\xi})| &= 0 \\ \alpha(\mathbf{y}, 0; \boldsymbol{\xi}) &= \boldsymbol{\xi} \cdot \mathbf{y}. \end{aligned} \quad (4.11)$$

First, we assume that α is actually a smooth solution already. Let $s \mapsto \tilde{\mathbf{r}}(s)$ describe the (\mathbf{y}, t) coordinates of a curve, to be called a ray, parameterized by s . Later space variable \mathbf{y} will be used to denote the map $\mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by following the rays $\tilde{\mathbf{r}}(s)$ from initial to end points. Additionally, we define the functions

$$\begin{aligned} q(s) &= \alpha(\tilde{\mathbf{r}}(s); \boldsymbol{\xi}) \\ \tilde{\mathbf{p}}(s) &= \begin{pmatrix} \nabla \alpha(\tilde{\mathbf{r}}(s); \boldsymbol{\xi}) \\ \partial_t \alpha(\tilde{\mathbf{r}}(s); \boldsymbol{\xi}) \end{pmatrix}. \end{aligned} \quad (4.12)$$

The shorthand notation $(\tilde{\mathbf{p}}_0, q_0, \tilde{\mathbf{r}}_0)$ refers to the value of $(\tilde{\mathbf{p}}, q, \tilde{\mathbf{r}})$ at $s = 0$. To find the *characteristic ODE*, we apply (11) on page 98 of [23] to our problem, giving

$$\begin{aligned} \frac{d}{ds} \tilde{\mathbf{p}}(s) &= \begin{pmatrix} -(\nabla c) |\nabla \alpha| \\ 0 \end{pmatrix} \\ \frac{d}{ds} q(s) &= \begin{pmatrix} \frac{c \nabla \alpha}{|\nabla \alpha|} \\ 1 \end{pmatrix} \cdot \begin{pmatrix} \nabla \alpha \\ \partial_t \alpha \end{pmatrix} = c |\nabla \alpha| + \partial_t \alpha = 0 \\ \frac{d}{ds} \tilde{\mathbf{r}}(s) &= \begin{pmatrix} \frac{c \nabla \alpha}{|\nabla \alpha|} \\ 1 \end{pmatrix}. \end{aligned} \quad (4.13)$$

The right hand side is consequently evaluated on the ray $\tilde{\mathbf{r}}(s)$. Before writing this as a proper ODE, we will simplify the equations by following observations. The first and second equation of (4.13) respectively imply that $[\tilde{\mathbf{p}}(s)]_{n+1} = [\tilde{\mathbf{p}}_0]_{n+1}$ and $q(s) = q_0$ for all s . Both are thus constant on the ray. In virtue of the last equation we may identify $t = s$. Therefore, we can restrict the equations to its spatial part, and write $\mathbf{p}(s) = [\tilde{\mathbf{p}}(s)]_{1:n}$ and $\mathbf{r}(s) = [\tilde{\mathbf{r}}(s)]_{1:n}$. Trivially then, $\mathbf{p}(s) = \nabla \alpha$ on the ray. The characteristic ODE reduces to

$$\begin{aligned} \frac{d}{ds} \mathbf{p}(s) &= -|\mathbf{p}| \nabla c(\mathbf{r}) \\ \frac{d}{ds} \mathbf{r}(s) &= c(\mathbf{r}) \frac{\mathbf{p}}{|\mathbf{p}|}. \end{aligned} \quad (4.14)$$

A triple $(\tilde{\mathbf{p}}_0, q_0, \tilde{\mathbf{r}}_0)$ is an admissible initial value of the characteristic ODE if it is consistent with the initial data of the problem (4.11). In our case that means:

$$\begin{aligned} q_0 &= \boldsymbol{\xi} \cdot \mathbf{r}_0 \\ \tilde{\mathbf{p}}_0 &= \begin{pmatrix} \boldsymbol{\xi} \\ -c(\mathbf{r}_0) |\boldsymbol{\xi}| \end{pmatrix}. \end{aligned} \quad (4.15)$$

The solution of the ODE (4.14) for a given $(\mathbf{r}_0, \mathbf{p}_0)$ is called a *bicharacteristic*. By restriction to the position domain, i.e. $(\mathbf{r}, \mathbf{p}) \mapsto \mathbf{r}$, a bicharacteristic is projected on its ray. The constant $[\tilde{\mathbf{p}}_0]_{n+1} = -c(\mathbf{r}_0)|\boldsymbol{\xi}|$ is interpreted as the frequency of a wave package travelling on the ray. It is negative for the α -term of v .

An admissible triplet is *noncharacteristic* if one can find admissible triplets for nearby initial points, i.e. in the neighborhood of $\mathbf{r}_0 \in \mathbb{R}^n$ [23]. In virtue of lemma 1 of [23] and the observation that $\partial_{\tilde{p}_{n+1}}$ of the eikonal equation is 1, every admissible triplet is noncharacteristic. Hence, for every $\mathbf{x} \in \mathbb{R}^n$ a bicharacteristic with initial value $(\mathbf{r}_0, \mathbf{p}_0) = (\mathbf{x}, \boldsymbol{\xi})$ originates. Given some $s \in \mathbb{R}$, it arrives at $(\mathbf{y}, \boldsymbol{\eta}) = (\mathbf{r}(s), \mathbf{p}(s))$. By lemma 2 of [23] the map $\mathbf{x} \mapsto \mathbf{y}$ is locally invertible: There exist an open interval $I \subset \mathbb{R}$ containing 0 and two neighborhoods U and V of \mathbf{r}_0 in \mathbb{R}^n , such that for each $s \in I$ the restriction to $U \rightarrow V$ is a smooth bijection. In addition, the map $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{y} \in \mathbb{R}^n$ can shown to be globally invertible.

As said, α is only a solution by assumption. We will use the bicharacteristics to construct a solution of the eikonal initial value problem (4.11). Now defining $\tilde{\alpha}(\mathbf{y}, s; \boldsymbol{\xi}) := q(s) = \boldsymbol{\xi} \cdot \mathbf{x}$, by theorem 2 of [23] function $\tilde{\alpha}$ is smooth and solves the eikonal problem. In succeeding text, we simply write α instead of $\tilde{\alpha}$. We observe that $\partial_{\mathbf{y}}\mathbf{x}$ is nonsingular for each $s \in I$, which can be drawn from the proof of lemma 2 of [23]. To examine the $\boldsymbol{\xi}$ -dependence of the constructed solution α , consider (4.15) to observe that the triplet $(\tilde{\mathbf{p}}_0, q_0, \tilde{\mathbf{r}}_0)$ depends smoothly on $\boldsymbol{\xi}$. Therefore, the bicharacteristics depend smoothly on $\boldsymbol{\xi}$, and so does α . Moreover, if we replace $\boldsymbol{\xi}$ by $\lambda\boldsymbol{\xi}$ with $\lambda > 0$, then (\mathbf{r}, \mathbf{p}) becomes $(\mathbf{r}, \lambda\mathbf{p})$, and consequently α becomes $\lambda\alpha$. Hence, function $\alpha(\mathbf{y}, t; \boldsymbol{\xi})$ is positive-homogeneous with respect to $\boldsymbol{\xi}$ of degree one. If (\mathbf{y}, t) is the space-time coordinate of a bicharacteristic with initial value $(\mathbf{x}, \boldsymbol{\xi})$ at $t = 0$ then it can be shown that

$$\partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t; \boldsymbol{\xi}) = \mathbf{x}. \quad (4.16)$$

The same procedure applied to β gives the same *set* of bicharacteristics, as should. Nonetheless, the initial spatial direction for positive time is $\boldsymbol{\xi}$ for α -rays, but $-\boldsymbol{\xi}$ for β -rays. Therefore, an α -bicharacteristic with initial value $(\mathbf{x}, \boldsymbol{\xi})$ is identical with with a β -bicharacteristic with initial value $(\mathbf{x}, -\boldsymbol{\xi})$.

4.2.3 The amplitude function

In this section we solve the IVP of the transport equation (4.8). The initial value, given by (4.10), is $a(\mathbf{x}, 0; \boldsymbol{\xi}) = \frac{i}{2c(\mathbf{x})|\boldsymbol{\xi}|}$. To solve it we examine the energy quantity $a(\mathbf{x}, t; \boldsymbol{\xi})^2$ evaluated on a ray. The result depends on the Jacobian of the flow and is given by (4.21), which means that $\det(\partial_{\mathbf{x}}\mathbf{y})a^2$ is constant along the ray.

The transport equation (4.8) can be rewritten into

$$\begin{aligned} 0 &= \partial_t(a^2\partial_t\alpha) - \nabla \cdot (a^2(\partial_t\alpha)\frac{c^2\nabla\alpha}{\partial_t\alpha}) \\ &= (\partial_t\alpha)[\partial_t - \frac{c^2\nabla\alpha}{\partial_t\alpha} \cdot \nabla]a^2 + a^2\left([\partial_t - \frac{c^2\nabla\alpha}{\partial_t\alpha} \cdot \nabla]\partial_t\alpha - (\partial_t\alpha)\nabla \cdot \frac{c^2\nabla\alpha}{\partial_t\alpha}\right). \end{aligned} \quad (4.17)$$

This expression will be evaluated on a ray. The s -parameterized ray with initial value $(\mathbf{x}, \boldsymbol{\xi})$ is denoted by $\mathbf{y} = \mathbf{y}(\mathbf{x}, s; \boldsymbol{\xi})$. Using subscript e for energy we define

$$q_e(s) = a(\mathbf{y}, s; \boldsymbol{\xi})^2. \quad (4.18)$$

Using the characteristic ODE (4.14), we observe that operator $[\partial_t - \frac{c^2\nabla\alpha}{\partial_t\alpha} \cdot \nabla]$ is the derivative $\frac{d}{ds}$ taken along the ray.² As we have seen in the previous section, the frequency is constant on a ray, so $\frac{d}{ds}\partial_t\alpha = 0$. Equation (4.17) therefore yields the initial value problem for q_e on a ray, i.e.

$$\begin{aligned} \frac{d}{ds}q_e(s) &= \left(\nabla \cdot \frac{c(\mathbf{y})^2\nabla\alpha(\mathbf{y}, s; \boldsymbol{\xi})}{\partial_t\alpha(\mathbf{y}, s; \boldsymbol{\xi})}\right)q_e(s) \\ q_e(0) &= \frac{-1}{4c(\mathbf{x})^2|\boldsymbol{\xi}|^2}. \end{aligned} \quad (4.19)$$

The initial value directly follows from $a(\mathbf{x}, 0; \boldsymbol{\xi})$. Noticing that $-q_e(s)$ is positive, the amplitude now follows from $a(\mathbf{y}, s; \boldsymbol{\xi}) = i\sqrt{-q_e(s)}$.

The IVP (4.19) shows that the amplitude is, in accordance with the principle of energy conservation, determined by the divergence of the rays. We claim that

$$\frac{d}{ds}(\det(\partial_{\mathbf{x}}\mathbf{y})q_e(s)) = 0. \quad (4.20)$$

The amplitude evaluated on the ray $\mathbf{y}(\mathbf{x}, s; \boldsymbol{\xi})$ follows from the claim and the initial value of (4.19). It is given by

$$a(\mathbf{y}, s; \boldsymbol{\xi}) = \frac{i}{2c(\mathbf{x})|\boldsymbol{\xi}|} \sqrt{\det(\partial_{\mathbf{x}}\mathbf{y})^{-1}}. \quad (4.21)$$

To prove (4.20) we use two identities, namely $\det(A)^{-1}\frac{d}{ds}\det(A) = \text{tr}(A^{-1}\frac{d}{ds}A)$, which is Jacobi's formula for an s -parameterized invertible matrix A , and, the identity $\text{tr}(ABC) = \text{tr}(CAB)$ for square matrices. Using the fact that $\mathbf{x} \mapsto \mathbf{y}$ is a diffeomorphism, observe that $[\partial_{\mathbf{x}}\mathbf{y}]^{-1}[\frac{d}{ds}\partial_{\mathbf{x}}\mathbf{y}] = [\partial_{\mathbf{x}}\mathbf{y}]^{-1}[\partial_{\mathbf{x}}\frac{d}{ds}\mathbf{y}] = [\partial_{\mathbf{x}}\mathbf{y}]^{-1}[\partial_{\mathbf{y}}\frac{d}{ds}\mathbf{y}][\partial_{\mathbf{x}}\mathbf{y}]$, in which each pair of brackets denotes a matrix. This yields

$$\det(\partial_{\mathbf{x}}\mathbf{y})^{-1}\frac{d}{ds}\det(\partial_{\mathbf{x}}\mathbf{y}) = \text{tr}(\partial_{\mathbf{y}}\frac{d}{ds}\mathbf{y}) = \nabla \cdot \frac{d}{ds}\mathbf{y}.$$

Using the characteristic ODE (4.14) to rewrite the right hand side, and subsequently multiplying with the determinant, gives

$$\frac{d}{ds}\det(\partial_{\mathbf{x}}\mathbf{y}) = \left(\nabla \cdot \frac{c(\mathbf{y})^2\nabla\alpha(\mathbf{y}, s; \boldsymbol{\xi})}{-\partial_t\alpha(\mathbf{y}, s; \boldsymbol{\xi})}\right)\det(\partial_{\mathbf{x}}\mathbf{y}).$$

Together with amplitude IVP (4.19) this proves the energy conservation claim (4.20).

²This is also the convective derivative with respect to the velocity field $\frac{c^2\nabla\alpha}{-\partial_t\alpha}$.

4.2.4 Solution operator as a FIO

In this subsection we consider more general initial values than plane-waves (4.5) by considering linear combinations of the WKB terms (4.6). This results in an approximate solution operator in the form of a Fourier integral operator (FIO) [22, 69, 70, 29] and we will review some of its properties. Our solutions so far involve only the highest order WKB approximation and they are limited to some small but finite time.

We consider the original wave equation (4.1) with $f = 0$ and the following initial conditions

$$u(\mathbf{x}, 0) = 0, \quad \partial_t u(\mathbf{x}, 0) = h_2(\mathbf{x}). \quad (4.22)$$

Following equation (4.6), the WKB solution for time $t \in I$, which we for the moment write as $S_{12}(t)h_2(\mathbf{y})$, is given by a sum of two terms. These terms will be denoted by $S_{12}(t)h_2(\mathbf{y}) = c(\mathbf{y})(S_{a2}(t)h_2(\mathbf{y}) + S_{b2}(t)h_2(\mathbf{y}))$ with

$$S_{a2}(t)h_2(\mathbf{y}) = \frac{1}{(2\pi)^n} \iint e^{i\alpha(\mathbf{y}, t; \boldsymbol{\xi}) - i\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a(\mathbf{y}, t; \boldsymbol{\xi})}{c(\mathbf{x})} h_2(\mathbf{x}) d\mathbf{x} d\boldsymbol{\xi}. \quad (4.23)$$

Here the subscript ‘a’ refers to the negative frequencies, i.e. phase and amplitude functions α and a . Then S_{b2} is defined similarly, using β and b , and refers to positive frequencies. We recall that the symmetry relations of subsection 4.2.1 imply that $S_{b2}(t)h_2 = \overline{S_{a2}(t)h_2}$. The construction is such that t can be negative.

To argue that S_{a2} is a FIO, we will take a closer look at its phase function, i.e.

$$\varphi(\mathbf{y}, t, \mathbf{x}, \boldsymbol{\xi}) = \alpha(\mathbf{y}, t; \boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \mathbf{x}, \quad (4.24)$$

and observe that it is positive homogeneous with respect to $\boldsymbol{\xi}$ of degree one, as it should. The stationary point set is given by

$$\Gamma_t = \{(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in Y \times X \times \mathbb{R}^n \setminus 0 \mid \mathbf{x} = \partial_{\boldsymbol{\xi}} \alpha(\mathbf{y}, t; \boldsymbol{\xi})\}. \quad (4.25)$$

For Γ_t to be a closed smooth submanifold of $Y \times X \times \mathbb{R}^n \setminus 0$, the matrix,

$$\begin{pmatrix} \partial_{\mathbf{y}} \partial_{\boldsymbol{\xi}} \varphi \\ \partial_{\mathbf{x}} \partial_{\boldsymbol{\xi}} \varphi \\ \partial_{\boldsymbol{\xi}} \partial_{\boldsymbol{\xi}} \varphi \end{pmatrix} = \begin{pmatrix} \partial_{\mathbf{y}} \partial_{\boldsymbol{\xi}} \alpha \\ -\mathbf{I}_n \\ \partial_{\boldsymbol{\xi}} \partial_{\boldsymbol{\xi}} \alpha \end{pmatrix},$$

needs to have maximal rank on Γ_t , which is obviously the case [70, chapter VI, (4.22)]. The stationary point set Γ_t is hence a $2n$ -dimensional manifold with $(\mathbf{y}, \boldsymbol{\xi})$ as coordinates.

The stationary point set can be understood in terms of the bicharacteristics. Definition (4.25) allows us to express \mathbf{x} on Γ_t as a function $\mathbf{x}_{\Gamma}(\mathbf{y}, t, \boldsymbol{\xi}) = \partial_{\boldsymbol{\xi}} \alpha(\mathbf{y}, t; \boldsymbol{\xi})$. Equation (4.16) implies that $(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in \Gamma_t$ if and only if a bicharacteristic initiates at $(\mathbf{x}, \boldsymbol{\xi})$ and passes through $(\mathbf{y}, \boldsymbol{\eta})$ at time t where $\boldsymbol{\eta}$ must be given by

$\boldsymbol{\eta} = \partial_{\mathbf{y}}\alpha(\mathbf{y}, t; \boldsymbol{\xi})$. If $(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in Y \times X \times \mathbb{R}^n \setminus 0$ and $t \in \mathbb{R}$ are such that $(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in \Gamma_t$ then one has $\partial_t \alpha(\mathbf{y}, t; \boldsymbol{\xi}) = -c(\mathbf{x})|\boldsymbol{\xi}|$, since the frequency $\partial_t \alpha$ is constant on a ray.

The propagation of singularities of S_{a2} is described by its canonical relation,

$$\Pi_t = \{((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\xi})) \in T^*Y \setminus 0 \times T^*X \setminus 0 \mid \mathbf{x} = \mathbf{x}_\Gamma(\mathbf{y}, t, \boldsymbol{\xi}), \boldsymbol{\eta} = \partial_{\mathbf{y}}\alpha(\mathbf{y}, t; \boldsymbol{\xi})\}. \quad (4.26)$$

Clearly, Π_t is the image of Γ_t under the map $(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \mapsto ((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\xi}))$. It follows from the characteristic ODE that the map from $(\mathbf{x}, \boldsymbol{\xi})$ to $(\mathbf{y}, \boldsymbol{\eta})$ is a bijection, say $\Phi_t : T^*X \setminus 0 \rightarrow T^*Y \setminus 0$. The canonical relation is hence the graph of an invertible function. Therefore, each pair $(\mathbf{y}, \boldsymbol{\xi})$, $(\mathbf{x}, \boldsymbol{\xi})$ and $(\mathbf{y}, \boldsymbol{\eta})$ can act as coordinates on Γ_t , and on Π_t . We observe that Φ_t depends smoothly on t .

The effect of the FIO S_{a2} working on a distribution v can be explained in terms of the wave front set. If $v \in \mathcal{E}'(X)$, then the wave front set $\text{WF}(v)$ of distribution v is a closed conic subset that describes the *locations* and *directions* of the singularities of v . Operator S_{a2} affects a distribution v by propagating its wave front set by composition with the canonical relation [22, 36, 69, 70]. From the above description of Π_t it follows that

$$\text{WF}(S_{a2}(t)v) \subset \Phi_t(\text{WF}(v)). \quad (4.27)$$

The pair $(\mathbf{x}, -\partial_{\mathbf{x}}\varphi)$ are referred to as the *ingoing variable* and *covariable*, and $(\mathbf{y}, \partial_{\mathbf{y}}\varphi)$ as the *outgoing variable* and *covariable*. The idea behind the names is that S_{a2} , by Φ_t , carries over $(\mathbf{x}, \boldsymbol{\xi})$ of the ingoing wave front set into $(\mathbf{y}, \boldsymbol{\eta})$ of the outgoing wave front set [70, p. 334].

So far the highest order WKB approximation was used. For the amplitudes a and b the notion of symbol classes is needed to properly include lower order terms. By replacing a with an asymptotic sum $a(\mathbf{x}, t; \boldsymbol{\xi}) = \sum_{j=0}^{\infty} a_{m-j}(\mathbf{x}, t; \boldsymbol{\xi})$, with a_k homogeneous of order k in $\boldsymbol{\xi}$ for $|\boldsymbol{\xi}| > 1$, the error in the WKB solution (4.6) can be made to decay as $|\boldsymbol{\xi}|^{-N}$ for any N . In other words, it becomes C^∞ and the approximate solution operator becomes a parametrrix. Moreover, the exact solution operator can be written in the form of $c(S_{a2} + S_{b2})$ by the addition to a and b of certain symbols in $S^{-\infty}$, which in particular decay faster than any power $|\boldsymbol{\xi}|^{-N}$. We note that the latter additions cannot be computed with ray theory.

Solution operators for longer times have been constructed using more general phase functions. For us those explicit expressions are of no interest, but we note that the FIO property, with canonical relation characterized by Φ_t , remains valid, as can be seen by applying the calculus of FIO's [22, theorem 2.4.1] to the product of several short time solution operators.

4.2.5 Solution operators and decoupling

In subsection 4.2.1 we assumed that the functions $ae^{i\alpha}$ and $be^{i\beta}$ propagate independently as solutions of the wave equation. In fact, this is the result of a rather

general procedure to decouple the wave equation [67]. Because the results of the decoupling will be used explicitly in section 4.4 we give a short review of it here. And we will examine its relation to the solution operator S_{a2} .

We write the wave field as the vector $(u_1(\mathbf{x}, t), u_2(\mathbf{x}, t))^T = (u, \partial_t u)^T$. The homogeneous wave equation can now be written as the following system, 1st-order with respect to time.

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & I \\ c(\mathbf{x})^2 \Delta & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (4.28)$$

The solution can be given as a matrix operator that maps the Cauchy data at $t = 0$, say $(u_{0,1}(\mathbf{x}), u_{0,2}(\mathbf{x}))^T$, to the field vector at t .

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = S(t) \begin{pmatrix} u_{0,1} \\ u_{0,2} \end{pmatrix} \quad \text{with} \quad S(t) = \begin{pmatrix} S_{11}(t) & S_{12}(t) \\ S_{21}(t) & S_{22}(t) \end{pmatrix}. \quad (4.29)$$

Naturally it satisfies the group property $S(t)S(s) = S(t + s)$. It is invertible by time reversal.

To decouple the system, we define several pseudo-differential operators. Let operator B be a symmetric approximation of $\sqrt{-c(\mathbf{x})\Delta c(\mathbf{x})}$ with its approximate inverse B^{-1} such that $B^2 + c\Delta c$, $B^{-1}B - I$, and $BB^{-1} - I$ are regularizing operators, i.e. pseudo-differential operators of order $-\infty$. Although the square root does not necessarily have to be symmetric, being symmetric has the advantage that it yields a unitary solution operator, as we will see. Neglecting regularity conditions, we use symmetry and self-adjointness interchangeably. The principal symbols of B and B^{-1} are $c(\mathbf{x})|\xi|$ and $\frac{1}{c(\mathbf{x})|\xi|}$ respectively. The existence of such operators is a well known result in pseudo-differential operator theory, see e.g. [51]. We now have the ingredients to define two matrix pseudo-differential operators Λ and V by

$$V = c(\mathbf{x}) \begin{pmatrix} 1 & 1 \\ -iB & iB \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} 1 & iB^{-1} \\ 1 & -iB^{-1} \end{pmatrix} \frac{1}{2c(\mathbf{x})}, \quad (4.30)$$

which are each others inverses modulo regularizing operators. We finally define the following two fields $(u_a(\mathbf{x}, t), u_b(\mathbf{x}, t))^T = \Lambda(u_1, u_2)^T$. Note that the Cauchy data can be represented by a time evaluation of $(u_a, u_b)^T$. We will use the phrase ‘Cauchy data’ in this way also. Omitting the regularizing error operators, the system (4.28) transforms into a decoupled system for $(u_a, u_b)^T$ of which the first equation, together with its initial value, is

$$\partial_t u_a = -iB u_a \quad \text{and} \quad u_a(\mathbf{x}, 0) = u_{0,a}(\mathbf{x}). \quad (4.31)$$

By removing the minus sign it becomes the equation for u_b . Let S_a and S_b be solution operators of the IVPs, i.e. $u_a(\mathbf{x}, t) = S_a(t)u_{0,a}(\mathbf{x})$ and similar for S_b .

Therefore, modulo regularizing operators

$$S(t) = V \begin{pmatrix} S_a(t) & 0 \\ 0 & S_b(t) \end{pmatrix} \Lambda, \quad (4.32)$$

which means that the original IVP (4.28) and the decoupled system (4.31) have identical solutions disregarding a smooth error. Because B is self-adjoint operators S_a and S_b are unitary, which follows from Stone's Theorem [18]. It can be shown that $S_a(t)$ and $S_b(t)$ with $t \in \mathbb{R}$ are FIOs [67].

We turn to the relation of this matrix formalism and $S_{12} = c(S_{a2} + S_{b2})$, from which we derive a local expression of $\text{p.s.}(S_a)$, the principal symbol of S_a . The amplitude of S_{a2} is a homogeneous symbol, which implies that it coincides with its principal symbol, and from its definition (4.23) can thereafter be concluded that $S_{a2} = \text{p.s.}(S_a \Lambda_{12})$. The principal symbol of a composition is the product of the principal symbols of its factors [22, 70], and hence $\frac{a(\mathbf{y}, t; \boldsymbol{\xi})}{c(\mathbf{x})} = \text{p.s.}(S_a) \frac{i}{2c(\mathbf{x})^2 |\boldsymbol{\xi}|}$. Using the solution of the transport equation (4.21), one concludes that

$$\text{p.s.}(S_a)(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) = \sqrt{\det(\partial_{\mathbf{y}} \mathbf{x}(\mathbf{y}, t; \boldsymbol{\xi}))}. \quad (4.33)$$

The principal symbol of S_b follows from $S_b = \overline{S_a}$.

4.2.6 The source field

In this subsection we discuss the source problem. The unperturbed velocity is a smooth function $c(\mathbf{x})$. The source wave is the fundamental solution of a delta function located at $(\mathbf{x}_s, 0)$ in space-time.

$$\begin{aligned} [c(\mathbf{x})^{-2} \partial_t^2 - \Delta] g(\mathbf{x}, t) &= \delta(\mathbf{x} - \mathbf{x}_s) \delta(t) \\ g(\mathbf{x}, 0) &= 0, \quad \partial_t g(\mathbf{x}, 0) = 0. \end{aligned} \quad (4.34)$$

An important assumption is that

$$\text{the source wave does not exhibit multipathing (SME)}. \quad (4.35)$$

The fundamental solution can therefore be approximated by an asymptotic expansion with a single phase function. This can in principle be found by an application of section 4.2.4 and using a change of phase function [22, section 2.3]. One can show that, if $|\mathbf{x} - \mathbf{x}_s| > \varepsilon$ for an $\varepsilon > 0$ and t bounded, the fundamental solution can be written as the Fourier integral [4]

$$g(\mathbf{x}, \mathbf{x}_s, t) = \frac{1}{2\pi} \int A(\mathbf{x}, \mathbf{x}_s, \omega) e^{i\omega(t - T(\mathbf{x}, \mathbf{x}_s))} d\omega, \quad (4.36)$$

with $A(\mathbf{x}, \mathbf{x}_s, \omega) \in S^{\frac{n-3}{2}}$ and $A(\mathbf{x}, \mathbf{x}_s, \omega) = \sum_{k=0}^{\infty} A_k(\mathbf{x}, \mathbf{x}_s, \omega)$. Each term is homogeneous, i.e. one has $A_k(\mathbf{x}, \mathbf{x}_s, \lambda\omega) = \lambda^{\frac{n-3}{2}-k} A_k(\mathbf{x}, \mathbf{x}_s, \omega)$ for $\lambda > 1$ and $|\omega| > 1$. This holds for $n = 1, 2, 3$. The sum means that for each $N \in \mathbb{N}$ there exists a $C_N > 0$ such that

$$|A(\mathbf{x}, \mathbf{x}_s, \omega) - \sum_{k=0}^{N-1} A_k(\mathbf{x}, \mathbf{x}_s, \omega)| \leq C_N (1 + |\omega|)^{\frac{n-3}{2}-N}. \quad (4.37)$$

The source is real, implying that $\overline{A_k(\mathbf{x}, \mathbf{x}_s, \omega)} = A_k(\mathbf{x}, \mathbf{x}_s, -\omega)$ for all k . In (4.36) one can also view the separate contributions of positive and negative frequencies.

In part of the further analysis we will use the highest order term of the source field. There exist an amplitude $A_s(\mathbf{x})$ and a cutoff $\sigma(\omega)$, both real and such that $A_0(\mathbf{x}, \mathbf{x}_s, \omega) = A_s(\mathbf{x}) \sigma(\omega) (i\omega)^{\frac{n-3}{2}}$ on the support of σ . Function σ is smooth and has value 1 except for a neighborhood of the origin where it is 0. We also abbreviate $T_s(\mathbf{x}) = T(\mathbf{x}, \mathbf{x}_s)$. The principal term of the expansion can now be written as

$$g(\mathbf{x}, t) = A_s(\mathbf{x}) \partial_t^{\frac{n-3}{2}} \delta(t - T_s(\mathbf{x})). \quad (4.38)$$

Functions $A_s(\mathbf{x})$ and $T_s(\mathbf{x})$ will be referred to as the source wave amplitude and travelttime respectively. Operator $\partial_t^{\frac{n-3}{2}}$ denotes the pseudo-differential operator with symbol $\omega \mapsto \sigma(\omega)(i\omega)^{\frac{n-3}{2}}$. The approximation $g(\mathbf{x}, t)$ matches the exact solution in case $\nabla c = 0$ in the limit of $\omega \rightarrow \infty$. In that case one would have $T_s(\mathbf{x}) = \frac{|\mathbf{x}-\mathbf{x}_s|}{c}$ and $A_s(\mathbf{x}) = \frac{c}{2}, \sqrt{\frac{c}{8\pi|\mathbf{x}-\mathbf{x}_s|}}, \frac{1}{4\pi|\mathbf{x}-\mathbf{x}_s|}$ for respectively $n = 1, 2, 3$ [4]. We define the *source wave direction vector*

$$\mathbf{n}_s(\mathbf{x}) = c(\mathbf{x})\partial_{\mathbf{x}}T_s(\mathbf{x}). \quad (4.39)$$

This vector will, for example, be used to provide insight in the microlocal interpretation of the scattering event.

Source waves that arrive at the acquisition set are in the context of the inversion called direct waves. The negative frequency part of the wave front set of the source field is given by

$$\begin{aligned} \Xi_s = \{(\mathbf{x}, t, \boldsymbol{\xi}, \omega) \in T^*(X \times \mathbb{R}) \setminus 0 \mid \dots \\ (\mathbf{x}, \boldsymbol{\xi}) = \Phi_t(\mathbf{x}_s, \boldsymbol{\xi}_s), \boldsymbol{\xi}_s \in \mathbb{R}^n \setminus 0, \omega = -c(\mathbf{x})|\boldsymbol{\xi}|\}. \end{aligned} \quad (4.40)$$

It contains all bicharacteristics that go through $(\mathbf{x}_s, 0)$ in spacetime. In the region where the Fourier integral (4.36) is valid, direct rays are also described by the equations $t = T_s(\mathbf{x})$ and $\boldsymbol{\xi} = |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x})$. The restriction to time t_c is denoted by

$$\Xi_{s,t_c} = \{(\mathbf{x}, \boldsymbol{\xi}) \in T^*X \setminus 0 \mid (\mathbf{x}, t_c, \boldsymbol{\xi}, \omega) \in \Xi_s\}. \quad (4.41)$$

This will be used to describe the direct waves in the Cauchy data of the continued scattered field.

4.3 Forward scattering problem

We consider the scattering problem and formulate the *continued scattered wave field* as the result of the *scattering operator* acting on the *reflectivity*, i.e. the medium perturbation. We start with a description of the scattering model, essentially a linearization of the source problem. In section 4.3.1 we derive an explicit expression for the mentioned operator. It will be used in section 4.3.2 to define the global scattering operator, of which we show in theorem 4 that it is a FIO under the conditions of the DSE and the SME.

4.3.1 Continued scattered wave field

Here, we introduce the *scattered wave field* and the *continued scattered wave field*. Loosely stated, the latter is the reverse time continuation of the former. We introduce the *scattering operator* that maps the medium perturbation to the continued scattered wave field. Theorem 3 shows that a local representation of the operator can be written as an oscillatory integral.

The medium perturbation is modeled by the *reflectivity function* $r(\mathbf{x})$. The non-smooth character of the perturbation gives rise to a scattered or reflected wave. We assume that

$$\text{supp}(r) \subset D \text{ for a compact } D \subset \mathbb{R}^{n-1} \times [\epsilon, \infty) \text{ and some } \epsilon > 0. \quad (4.42)$$

The last component of \mathbf{x} describes the depth. Because the source is at the surface, i.e. $x_{s,n} = 0$, the reflectivity is zero in a neighborhood of the source. Following the Born approximation, the scattering problem is obtained by linearization of the source problem (4.34) with $(1 + r(\mathbf{x}))c(\mathbf{x})$ as the velocity. To find the linearization it is advantageous to first multiply (4.34) with $c(\mathbf{x})^2$. The result is

$$\begin{aligned} [\partial_t^2 - c(\mathbf{x})^2 \Delta]u(\mathbf{x}, t) &= r(\mathbf{x})2A_s(\mathbf{x})\partial_t^{\frac{n+1}{2}} \delta(t - T_s(\mathbf{x})) \\ u(\mathbf{x}, 0) &= 0, \quad \partial_t u(\mathbf{x}, 0) = 0. \end{aligned} \quad (4.43)$$

The *scattered wave field* $u(\mathbf{x}, t)$ is defined as the solution of the scattering problem (4.43). We have used the fact that the source wave field does not exhibit multipathing (SME) and can therefore be formulated as the asymptotic expansion (4.36). In the forward modeling we will use the principal term to approximate the source, i.e. (4.38). The subprincipal source terms do not contribute to the principal symbol of the scattering operator [54].

The *continued scattered wave field* u_h is defined as the solution of a final value problem of the homogeneous wave equation such that the Cauchy data at $t = T_1$ are identical with the Cauchy data of the scattered field u :

$$\begin{aligned} [\partial_t^2 - c(\mathbf{x})^2 \Delta]u_h(\mathbf{x}, t) &= 0, \\ u_h(\mathbf{x}, T_1) &= u(\mathbf{x}, T_1), \quad \partial_t u_h(\mathbf{x}, T_1) = \partial_t u(\mathbf{x}, T_1). \end{aligned} \quad (4.44)$$

The contributions to the scattered field entirely come to pass within the interval $[T_0, T_1]$, i.e. T_0 and T_1 are chosen such that $T_s(\text{supp}(r)) \subset [T_0, T_1]$. For $t \geq T_1$ one has $u_h(\mathbf{x}, t) = u(\mathbf{x}, t)$ but as u_h does and u does not solve the homogeneous wave equation, they differ for $t < T_1$. We also use the decoupled wave fields $(u_{h,a}, u_{h,b})^T = \Lambda(u_h, \partial_t u_h)^T$, with Λ defined in (4.30).

The continued scattered wave field models the receiver wave field in an idealized experiment. Idealized here means that all scattered rays are present, even rays that do not intersect the acquisition set. It hence represents the scattered field by being its continuation in reverse time. The *reverse time continued wave field*, to be defined in section 4.4, models the receiver wave field.

The *scattering operator* F by definition maps r to $(u_h, \partial_t u_h)^T$, and we let F_a and F_b map the reflectivity r to the decoupled components of the continued scattered wave field $u_{h,a}$ and $u_{h,b}$. To show that F_a is a FIO we derive an explicit formulation valid for a small time interval around a localized scattering event. Let $\{\rho_i\}_{i \in \mathcal{I}}$ be a finite smooth partition on D such that $\sum_{i \in \mathcal{I}} \rho_i = 1$ on D . Using ρ_i as a multiplication operator then

$$F_a(t) = \sum_{i \in \mathcal{I}} S_a(t - t_{1i}) F_a(t_{1i}) \rho_i, \quad (4.45)$$

and F_b likewise. S_a is the solution operator (4.32). The i^{th} local scattering event is delimited by $[t_{0i}, t_{1i}]$, so $T_s(\text{supp}(\rho_i)) \subset [t_{0i}, t_{1i}]$. The partition is chosen fine enough such that $[t_{0i}, t_{1i}]$ falls within an interval of definition of (4.23), i.e. the local expression of solution operator S_{a2} .

We write ρ for an arbitrary member of $\{\rho_i\}_{i \in \mathcal{I}}$ and $[t_0, t_1]$ for its delimiting interval, and derive a local expression of the scattering operator evaluated at t_1 . We will prove the following

Theorem 3. *The local scattering operator $F_a(t_1)\rho$ can be written as an oscillatory integral. It maps the reflectivity r to the continued scattered wave field, that is, $u_{h,a}(\mathbf{y}, t_1) = F_a(t_1)\rho r(\mathbf{y})$ and*

$$u_{h,a}(\mathbf{y}, t_1) = \frac{1}{(2\pi)^n} \iint e^{i\varphi_T(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})} A_F(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) d\boldsymbol{\xi} \rho r(\mathbf{x}) d\mathbf{x}, \quad (4.46)$$

in which the phase and amplitude function are respectively defined as

$$\begin{aligned} \varphi_T(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) &= \alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \mathbf{x}, \\ A_F(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) &= (i\partial_t \alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi}))^{\frac{n+1}{2}} \frac{a(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi})}{c(\mathbf{x})} 2A_s(\mathbf{x}). \end{aligned} \quad (4.47)$$

Here (4.46) is only the contribution of ρr . There is a similar statement for $u_{h,b}$, which satisfies $u_{h,b}(\mathbf{y}, t) = \overline{u_{h,a}(\mathbf{y}, t)}$.

Proof. To solve the scattering problem (4.43) it will first be transformed into a τ -parameterized family of IVP's. Duhamel's principle states that the solution, i.e. the scattered wave field, is then given by

$$u(\mathbf{x}, t) := \int_0^t \tilde{u}(\mathbf{x}, t; \tau) d\tau, \quad (4.48)$$

in which for each τ function $\tilde{u}(\mathbf{x}, t; \tau)$ is the solution of the homogeneous wave equation with prescribed Cauchy data on $t = \tau$ [23, §2.4.2]:

$$\begin{aligned} [\partial_t^2 - c(\mathbf{x})^2 \Delta] \tilde{u}(\mathbf{x}, t; \tau) &= 0 \quad \text{with } t \in \mathbb{R} \\ \tilde{u}(\mathbf{x}, \tau; \tau) &= 0 \\ \partial_t \tilde{u}(\mathbf{x}, \tau; \tau) &= r(\mathbf{x}) 2A_s(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta(\tau - T_s(\mathbf{x})). \end{aligned} \quad (4.49)$$

The continued scattered wave field is defined as the solution of the final value problem (4.44). Using the observation that $r(\mathbf{x}) 2A_s(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta(\tau - T_s(\mathbf{x})) = 0$ if $\tau \notin [T_0, T_1]$, it can be found by

$$u_h(\mathbf{x}, t) := \int_{T_0}^{T_1} \tilde{u}(\mathbf{x}, t; \tau) d\tau \quad \text{with } t \in \mathbb{R}. \quad (4.50)$$

Time integration is now over the fixed interval $[T_0, T_1]$, by which u_h solves the homogeneous wave equation. For $t \geq T_1$ the wave fields u and u_h coincide. Therefore, this solves (4.44).

To derive the local expression we solve the τ -parameterized homogeneous IVP of (4.49) with r replaced by ρr and subsequently evaluate the solution at t_1 . If we define $(\tilde{u}_a, \tilde{u}_b)^T = \Lambda(\tilde{u}, \partial_t \tilde{u})^T$ then $\tilde{u} = c(\tilde{u}_a + \tilde{u}_b)$. We apply solution operator S_{a2} with initial state at time τ . This gives

$$\tilde{u}_a(\mathbf{y}, t_1; \tau) = S_{a2}[\rho r(\mathbf{x}) 2A_s(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta(\tau - T_s(\mathbf{x}))](\mathbf{y}, t_1 - \tau). \quad (4.51)$$

Note that S_{a2} involves a *relative* time, i.e. the difference $t_1 - \tau$, which is allowed because the medium velocity does not change in time. Then, time is as much as *absolute* when it agrees with the source time reference.

Consider $u_{h,a}(\mathbf{y}, t_1)$, i.e. integral (4.50) with \tilde{u} replaced by $\tilde{u}_a(\mathbf{y}, t_1; \tau)$ in (4.51). We will eliminate τ by integration and write the field as an oscillatory integral. With the expression (4.23) of S_{a2} and the application of $T_s(\text{supp}(\rho r)) \subset [t_0, t_1]$ one derives the following integral

$$\begin{aligned} u_{h,a}(\mathbf{y}, t_1) &= \frac{1}{(2\pi)^n} \iiint e^{i\alpha(\mathbf{y}, t_1 - \tau; \boldsymbol{\xi}) - i\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a(\mathbf{y}, t_1 - \tau; \boldsymbol{\xi})}{c(\mathbf{x})} \dots \\ &\quad \rho r(\mathbf{x}) 2A_s(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta(\tau - T_s(\mathbf{x})) d\mathbf{x} d\boldsymbol{\xi} d\tau. \end{aligned}$$

We recognize two convolutions, the integral over τ and operator $\partial_t^{\frac{n+1}{2}}$, the operator $\partial_t^{\frac{n+1}{2}}$ can be commuted to act on $e^{i\alpha - i\xi \cdot \mathbf{x} \frac{a}{c}}$. Restricting to the highest order term, one writes $\partial_t^{\frac{n+1}{2}} [e^{i\alpha - i\xi \cdot \mathbf{x} \frac{a}{c}}] = (i\partial_t \alpha)^{\frac{n+1}{2}} e^{i\alpha - i\xi \cdot \mathbf{x} \frac{a}{c}}$, which is an application of a general result of FIO theory [22, 70]. Cutoff σ is omitted to shorten the expression. This yields

$$u_{h,a}(\mathbf{y}, t_1) = \frac{1}{(2\pi)^n} \iiint \left[e^{i\alpha - i\xi \cdot \mathbf{x}} (i\partial_t \alpha)^{\frac{n+1}{2}} \frac{a}{c(\mathbf{x})} \right]_{(\mathbf{y}, t_1 - \tau; \xi)} \cdots \\ \rho r(\mathbf{x}) 2A_s(\mathbf{x}) \delta(\tau - T_s(\mathbf{x})) d\mathbf{x} d\xi d\tau.$$

Notation $[\dots]_{\text{arg}}$ means that α , $\partial_t \alpha$ and a within the square brackets are evaluated in given argument. Explicit integration finally gives the oscillatory integral in (4.46) and (4.47). \square

4.3.2 Scattering operator as FIO

Here we establish that $F_a(t_1)\rho$ is a FIO if the direct waves are excluded (DSE). We define the global scattering operator πF and show that it is a FIO with an injective canonical relation, i.e. theorem 4.

Before we proceed with the theoretical aspects of the operator, we will explain what it does. The stationary points of $F_a(t_1)\rho$ are given by $\partial_\xi \varphi_T = 0$, i.e. $\partial_\xi \alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \xi) - \mathbf{x} = 0$. A stationary point $(\mathbf{y}, \mathbf{x}, \xi)$ has the following interpretation. The source wave front excites the reflectivity at $(\mathbf{x}, T_s(\mathbf{x}))$ in space-time, causing a scattering event. The event emits a scattered ray from $(\mathbf{x}, T_s(\mathbf{x}))$ with initial covariable ξ , which arrives at (\mathbf{y}, t_1) with covariable $\eta = \partial_\mathbf{y} \varphi_T(\mathbf{y}, t_1, \mathbf{x}, \xi)$. Operator $F_a(t_1)\rho$ so describes the scattering event and the propagation of the scattered wave over a small distance. The distance will be extended by application of the solution operator, see (4.45). Using the terminology introduced at the end of subsection 4.2.4, the ingoing variable and covariable are (\mathbf{x}, ζ) with $\zeta = -\partial_\mathbf{x} \varphi_T(\mathbf{y}, t_1, \mathbf{x}, \xi)$. The outgoing variable and covariable are (\mathbf{y}, η) . This means that $F_a(t_1)\rho$ carries over $(\mathbf{x}, \zeta) \in \text{WF}(r)$ into $(\mathbf{y}, \eta) \in \text{WF}(u_{h,a}(\cdot, t_1))$.

For the ingoing covariable one has

$$\zeta = -\partial_\mathbf{x} \varphi_T = \partial_t \alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \xi) \partial_\mathbf{x} T_s(\mathbf{x}) + \xi.$$

Using the source wave direction vector $\mathbf{n}_s(\mathbf{x}) = c(\mathbf{x}) \partial_\mathbf{x} T_s(\mathbf{x})$ and the identity for the frequency $\partial_t \alpha = -c(\mathbf{x}) |\xi|$, this yields the relation between ζ and ξ ,

$$\zeta = \xi - |\xi| \mathbf{n}_s(\mathbf{x}), \quad (4.52)$$

reflecting Snell's law. Figure 4.1 shows the microlocal picture of the scattering event and the scattered ray. Equation (4.52) also implies that $\zeta \cdot \mathbf{n}_s(\mathbf{x}) < 0$ holds

everywhere. This is a result of the geometry of the reflection event with one source. Note that (4.52) only applies for negative frequencies. For positive frequencies, i.e. considering F_b , one gets $\zeta' = \xi + |\xi| \mathbf{n}_s(\mathbf{x})$ instead. In that case $\zeta' \cdot \mathbf{n}_s(\mathbf{x}) > 0$ holds everywhere.

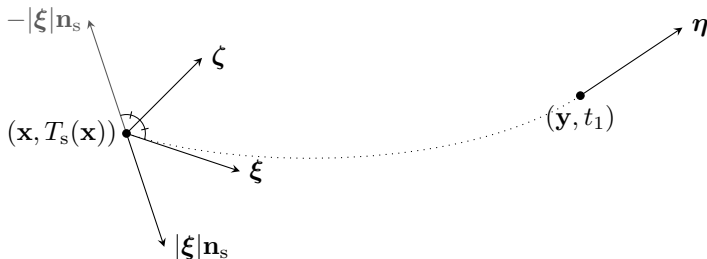


Figure 4.1: Propagation of singularities at $(\mathbf{x}, T_s(\mathbf{x}))$ in space-time. Also see equation (4.52). The dotted line represents the ray. The endpoint of the ray at (\mathbf{y}, t_1) contributes to the scattered field. Here, $|\xi| \mathbf{n}_s$ and ξ can respectively be interpreted as the wave numbers of the initial and reflected waves, and ζ a normal vector that can be associated with a reflector at \mathbf{x} .

If (\mathbf{x}, ξ) is associated with a source ray, i.e. $\xi = |\xi| \mathbf{n}_s(\mathbf{x})$, then $\zeta = 0$ by (4.52). In that case there is no reflection. We show that away from the source rays the scattering operator F is a FIO with an injective canonical relation, which will be made more precise. The practical implication is that source wave arrivals are excluded from the data before the receiver wave field is calculated.

The *direct source wave exclusion* (DSE) is the removal of the source singularities contained in Ξ_s from the wave front set of the continued scattered wave field. Mathematically it will be applied by t -families of pseudo-differential operators $\pi_a(t)$ and $\pi_b(t)$ that act on the Cauchy data $(u_{h,a}(\cdot, t), u_{h,b}(\cdot, t))^T$. The symbol of $\pi_a(t_c)$ is, for some fixed t_c , a smooth cutoff function on $T^*Y \setminus 0$, being 0 on a narrow conic neighborhood of Ξ_{s,t_c} (cf. (4.41)) and 1 outside a slightly larger conic neighborhood. Furthermore, we assume that π_a satisfies

$$\pi_a(t) = S_a(t - t_c) \pi_a(t_c) S_a(t_c - t), \quad (4.53)$$

which implies that the field $\pi_a u_{h,a}$ still satisfies a homogeneous wave equation. The symbol π_b satisfies $\pi_b(t; \mathbf{x}, \xi) = \pi_a(t; \mathbf{x}, -\xi)$.

Since, in the absence of multipathing, rays define paths of shortest traveltime between two arbitrary points, we have the following property. Let $\mathbf{x}, \tilde{\mathbf{x}} \in D$ be not

identical, then

$$\begin{aligned} & \text{if } \boldsymbol{\xi}, \tilde{\boldsymbol{\xi}} \in \mathbb{R}^n \setminus 0 \text{ and } t_i > 0 \text{ such that } (\tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) = \Phi_{t_i}(\mathbf{x}, \boldsymbol{\xi}) \\ & \text{then } |T_s(\tilde{\mathbf{x}}) - T_s(\mathbf{x})| < t_i \text{ or } (\mathbf{x}, \boldsymbol{\xi}) \in \Xi_{s, T_s(\mathbf{x})}. \end{aligned} \quad (4.54)$$

If \mathbf{x} and $\tilde{\mathbf{x}}$ lay on the same source ray then $|T_s(\tilde{\mathbf{x}}) - T_s(\mathbf{x})| = t_i$.

The central result is the theorem that the composition $\pi_a F_a$ is a FIO of which the canonical relation is the graph of an injective function. We write $V_{s,t} \subset T^*Y \setminus 0$ for the zero set of $\pi_a(t)$, a conic neighborhood of $\Xi_{s,t}$. With $\pi_b F_b = \overline{\pi_a F_a}$ we present the following

Theorem 4. *Operator $\pi_a F_a$ defined above, is a FIO. Its canonical relation is*

$$\begin{aligned} \Lambda = \{ & ((\mathbf{y}, t, \boldsymbol{\eta}, \omega), (\mathbf{x}, \boldsymbol{\zeta})) \mid (\mathbf{y}, \boldsymbol{\eta}) \in (T^*Y \setminus 0) \setminus V_{s,t}, t \in \mathbb{R}, \omega = -c(\mathbf{y})|\boldsymbol{\eta}|, \\ & (\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-t}(\mathbf{y}, \boldsymbol{\eta}), \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}| \mathbf{n}_s(\mathbf{x}), \mathbf{x} \in D\}. \end{aligned} \quad (4.55)$$

The projection of Λ to its outgoing variables, i.e. $(\mathbf{y}, t, \boldsymbol{\eta}, \omega)$, is injective.

We will first show that the composition $\pi_a(t_1)F_a(t_1)\rho$ is a FIO. The composition $\pi_a F_a$ is subsequently defined as the sum of local contributions, like in (4.45), and will also be called the ‘scattering operator’. The canonical relation becomes the union of the local relations. A part of the proof is put in lemma 1. The operator can alternatively be defined by means of the bicharacteristics of the wave equation. The papers [54, 50] show how this can be done, although their scattering operator does not fully coincide with ours.

Proof. Because $\pi_a(t)S_a(t-t_{1i}) = S_a(t-t_{1i})\pi_a(t_{1i})$ the scattering operator can be written as

$$\pi_a(t)F_a(t) = \sum_{i \in \mathcal{I}} S_a(t-t_{1i})\pi_a(t_{1i})F_a(t_{1i})\rho_i. \quad (4.56)$$

Again omitting subscript i to denote an arbitrary member of \mathcal{I} we will argue that the local scattering operator $\pi_a(t_1)F_a(t_1)\rho$ is a FIO. Then $\pi_a F_a$ becomes a sum of compositions of FIOs.

The local scattering operator is the oscillatory integral (4.46) in which the amplitude A_F (4.47) is replaced by $\pi_a(t_1; \mathbf{y}, \partial_{\mathbf{y}}\alpha) A_F$. This follows from the application of pseudo-differential operator $\pi_a(t_1)$, its symbol denoted by $\pi_a(t_1; \cdot, \cdot)$, on the integral [22, 70]. To be able to omit the zero set of $\pi_a(t_1)$ from the analysis of the phase φ_T we define the conic set

$$W_{s,t_1} = \{(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in Y \times X \times \mathbb{R}^n \setminus 0 \mid (\mathbf{y}, \boldsymbol{\eta}) \in V_{s,t_1}, (\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-t_1}(\mathbf{y}, \boldsymbol{\eta}), \mathbf{x} \in D\}.$$

The stationary point set of the phase function, by definition $\partial_{\boldsymbol{\xi}}\varphi_T = 0$, is given by

$$\begin{aligned} \Sigma_{t_1} = \{ & (\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in (Y \times X \times \mathbb{R}^n \setminus 0) \setminus W_{s,t_1} \mid \dots \\ & \mathbf{x} = \partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi}), \mathbf{x} \in \text{supp}(\rho)\}. \end{aligned} \quad (4.57)$$

We observe that $|\partial_{\mathbf{x}}\partial_{\boldsymbol{\xi}}\varphi_{\text{T}}| = |\partial_{\boldsymbol{\xi}}\partial_{\mathbf{x}}\varphi_{\text{T}}| = |\partial_{\boldsymbol{\xi}}\boldsymbol{\zeta}|$ by definition of $\boldsymbol{\zeta}$ (4.52). Moreover

$$|\partial_{\boldsymbol{\xi}}\boldsymbol{\zeta}| = |\mathbb{I}_n - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \otimes \mathbf{n}_s(\mathbf{x})| = 1 - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \cdot \mathbf{n}_s(\mathbf{x}). \quad (4.58)$$

By the DSE, applied as the omission of W_{s,t_1} in (4.57), the condition $\boldsymbol{\xi} \parallel \mathbf{n}_s(\mathbf{x})$ is never met, from which follows that the Jacobian $|\partial_{\boldsymbol{\xi}}\boldsymbol{\zeta}|$ is nonsingular. This implies that the derivative $\partial_{(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi})}\partial_{\boldsymbol{\xi}}\varphi_{\text{T}}$ has maximal rank, making Σ_{t_1} a closed smooth $2n$ -dimensional submanifold. The canonical relation relates ingoing (co)variables $(\mathbf{x}, \boldsymbol{\zeta})$ with outgoing (co)variables $(\mathbf{y}, \boldsymbol{\eta})$ and is given by

$$\{((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\zeta})) \mid (\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in \Sigma_{t_1}, \boldsymbol{\eta} = \partial_{\mathbf{y}}\varphi_{\text{T}}, \boldsymbol{\zeta} = -\partial_{\mathbf{x}}\varphi_{\text{T}}\}. \quad (4.59)$$

The relation is the graph of a diffeomorphism. We postpone the proof until after the construction of the global scattering operator $\pi_a(t)F_a(t)$ as the local and the global arguments are basically the same. Therefore the local scattering operator is a FIO with a bijective canonical relation.

The local operator will be composed with the solution operator. This gives a seamless extension because both operators are build on the same flow. It becomes $S_a(t-t_1)\pi_a(t_1)F_a(t_1)\rho$, which is a FIO. The canonical relation is determined by the composition of relations [22, 70]. The global scattering operator $\pi_a(t)F_a(t)$ is subsequently defined as the sum (4.56) of the extended local operators, of which the canonical relation Λ_t is the union of the local relations (4.59).

We will argue that Λ_t is the graph of an injection that is a diffeomorphism onto its image. It will be denoted by $\Theta_t : (T^*D \setminus 0) \setminus U_s \rightarrow (T^*Y \setminus 0) \setminus V_{s,t}$. We used the zero set of $\pi_a(t)$ expressed in the domain of Θ_t :

$$U_s = \{(\mathbf{x}, \boldsymbol{\zeta}) \in T^*X \setminus 0 \mid (\mathbf{x}, \boldsymbol{\xi}) \in V_{s, T_s(\mathbf{x})}, \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}), \mathbf{x} \in D\}. \quad (4.60)$$

The injection implies that Λ_t , for a fixed t , can be parameterized by \mathbf{y} and $\boldsymbol{\eta}$, so

$$\Lambda_t = \{((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\zeta})) \mid (\mathbf{y}, \boldsymbol{\eta}) \in (T^*Y \setminus 0) \setminus V_{s,t}, \dots \\ (\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-t}(\mathbf{y}, \boldsymbol{\eta}), \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}), \mathbf{x} \in D\}. \quad (4.61)$$

We now prove the existence and injectivity of Θ_t . Without loss of generality we assume that t denotes a moment after the scattering event.

Let $(\mathbf{x}, \boldsymbol{\zeta}) \in (T^*D \setminus 0) \setminus U_s$ be given. It can be shown that the transformation $\boldsymbol{\xi} \mapsto \boldsymbol{\zeta}$ given in (4.52) is injective on the complement of U_s and thus determines a unique $(\mathbf{x}, \boldsymbol{\xi})$. By ray tracing over $t - T_s(\mathbf{x})$, i.e. mapping by $\Phi_{t-T_s(\mathbf{x})}$, one finds the pair $(\mathbf{y}, \boldsymbol{\eta})$.

Let $(\mathbf{y}, \boldsymbol{\eta}) \in (T^*Y \setminus 0) \setminus V_{s,t}$ be given. This uniquely determines a bicharacteristic. By ray tracing backwards, i.e. by $\Phi_{T_s(\mathbf{x})-t}$ with $T_s(\mathbf{x}) - t < 0$, the ray goes through $(\mathbf{x}, T_s(\mathbf{x}))$ in space-time. If a second point $(\tilde{\mathbf{x}}, T_s(\tilde{\mathbf{x}}))$ is met, property (4.54) (SME)

implies that the bicharacteristic coincides with one from the source. The condition $(\mathbf{y}, \boldsymbol{\eta}) \notin V_{s,t}$ (DSE) rules out this possibility, leading to the conclusion that \mathbf{x} is unique. The covariable $\boldsymbol{\xi}$ uniquely follows from the ray tracing, and is mapped to $\boldsymbol{\zeta}$ by (4.52). The transformation Θ_t is therefore one-to-one.

To prove the smoothness we analyse the scattering event around a fixed point $(\mathbf{x}_0, \boldsymbol{\xi}_0)$, of which $\mathbf{x}_0 \in \text{supp}(\rho)$, and define $\tau_0 = T_s(\mathbf{x}_0)$. Now Θ_t can be factorized in the following way

$$(\mathbf{x}, \boldsymbol{\zeta}) \xrightarrow{(4.52)} (\mathbf{x}, \boldsymbol{\xi}) \xrightarrow{\Phi_{\tau_0 - T_s(\mathbf{x})}} (\check{\mathbf{x}}, \check{\boldsymbol{\xi}}) \xrightarrow{\Phi_{t - \tau_0}} (\mathbf{y}, \boldsymbol{\eta}).$$

The Jacobian of Θ_t becomes the product of three Jacobians, namely

$$\left| \frac{\partial(\mathbf{y}, \boldsymbol{\eta})}{\partial(\mathbf{x}, \boldsymbol{\zeta})} \right| = \left| \frac{\partial(\mathbf{y}, \boldsymbol{\eta})}{\partial(\check{\mathbf{x}}, \check{\boldsymbol{\xi}})} \right| \left| \frac{\partial(\check{\mathbf{x}}, \check{\boldsymbol{\xi}})}{\partial(\mathbf{x}, \boldsymbol{\xi})} \right| \left| \frac{\partial(\mathbf{x}, \boldsymbol{\xi})}{\partial(\mathbf{x}, \boldsymbol{\zeta})} \right|. \quad (4.62)$$

The leftmost factor in the right hand side is nonsingular because $\Phi_{t - \tau_0}$ is a diffeomorphism. The rightmost factor in the right hand side is nonsingular because the map $\boldsymbol{\xi} \mapsto \boldsymbol{\zeta}$ has a positive Jacobian (4.58). The transformation $\Phi_{\tau_0 - T_s(\mathbf{x})}$ is the least obvious one. We will show in lemma 1 below that it is a smooth bijection. Therefore Θ_t is a diffeomorphism onto its image.

So far t was held fixed to simplify the presentation. Time dependence is determined by the flow Φ_t . This allows t to be included in the canonical relation Λ of the scattering operator $\pi_a F_a$, which is a map to spacetime distributions. Parameterized by \mathbf{y} , $\boldsymbol{\eta}$ and t , Λ becomes (4.55). The injectivity follows from the parameterization. \square

Lemma 1. *Let $\tau_0 = T_s(\mathbf{x}_0)$ and $s(\mathbf{x}) = \tau_0 - T_s(\mathbf{x})$. If $J(\mathbf{x}, \boldsymbol{\xi}) = \Phi_s(\mathbf{x}, \boldsymbol{\xi})$ then J is a smooth bijection that maps $(\mathbf{x}_0, \boldsymbol{\xi}_0)$ onto itself. Its Jacobian is*

$$\det \partial_{(\mathbf{x}, \boldsymbol{\xi})} J(\mathbf{x}_0, \boldsymbol{\xi}_0) = 1 - \frac{\boldsymbol{\xi}_0}{|\boldsymbol{\xi}_0|} \cdot \mathbf{n}_s(\mathbf{x}_0), \quad (4.63)$$

which is nonsingular by the DSE.

Proof. For \mathbf{x} in the neighborhood of \mathbf{x}_0 one has $s(\mathbf{x}) \in I$, so Φ_s is defined. The smoothness of J follows directly from the smoothness of $\mathbf{x} \mapsto T_s(\mathbf{x})$ and Φ_s in its arguments including s . The Jacobian results from the straight forward calculation

$$\begin{aligned} \partial_{(\mathbf{x}, \boldsymbol{\xi})} J(\mathbf{x}_0, \boldsymbol{\xi}_0) &= \partial_{(\mathbf{x}, \boldsymbol{\xi})} \Phi_0(\mathbf{x}_0, \boldsymbol{\xi}_0) + \partial_s \Phi_0(\mathbf{x}_0, \boldsymbol{\xi}_0) \otimes \partial_{(\mathbf{x}, \boldsymbol{\xi})} s(\mathbf{x}_0) \\ &= \begin{pmatrix} \mathbf{I}_n & 0 \\ 0 & \mathbf{I}_n \end{pmatrix} + \begin{pmatrix} c(\mathbf{x}_0) \frac{\boldsymbol{\xi}_0}{|\boldsymbol{\xi}_0|} \\ -|\boldsymbol{\xi}_0| \partial_{\mathbf{x}} c(\mathbf{x}_0) \end{pmatrix} \otimes \begin{pmatrix} -\partial_{\mathbf{x}} T_s(\mathbf{x}_0) & 0 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{I}_n - \frac{\boldsymbol{\xi}_0}{|\boldsymbol{\xi}_0|} \otimes \mathbf{n}_s(\mathbf{x}_0) & 0 \\ |\boldsymbol{\xi}_0| \partial_{\mathbf{x}} c(\mathbf{x}_0) \otimes \partial_{\mathbf{x}} T_s(\mathbf{x}_0) & \mathbf{I}_n \end{pmatrix}. \end{aligned}$$

Herein we substitute the right-hand side of the characteristic ODE (4.14) for $\partial_s \Phi_s$. \square

4.4 Reverse time continuation from the boundary

The receiver wave field is modeled by the reverse time continued wave u_r . In this section, we show that u_r is the result of a pseudo-differential operator of order zero acting on the continued scattered wave u_h . We refer to it as the *revert operator* P .

The processes that are modeled by P are the propagation of the scattered wave field from a certain time, say $t = t_c$, to the surface at $x_n = 0$, the restriction of the wave field to the acquisition domain, the data processing, and eventually the continuation in reverse time. The revert operator suppresses the part of the scattered wave field that cannot be recovered because the contributing waves do not reach the acquisition domain. The data processing comprises a spatial smooth cutoff on the acquisition domain, the removal of direct source waves and the removal of waves reaching the surface following grazing rays. The final reconstruction represents a field related to bicharacteristics that intersect the acquisition domain M only once, and in the upgoing direction.

Let u be the solution to the homogeneous wave equation. When we apply the result of this section to develop the inverse scattering, we set $u = u_h$. Let M be a bounded open subset of $\{(\mathbf{x}, t) \in \mathbb{R}^{n+1} \mid x_n = 0\}$ and let $T_M u$ denote the restriction of u to M . We denote $\mathbf{x}' = (x_1, \dots, x_{n-1})$, so (\mathbf{x}', t) are coordinates on M . The field u_r is an anticausal solution to

$$[c(\mathbf{x})^{-2} \partial_t^2 - \Delta] u_r(\mathbf{x}, t) = \delta(x_n) F_M T_M u(\mathbf{x}', t), \quad (4.64)$$

where F_M is defined as follows.

The boundary operator F_M consists of two types of factors. A pseudo-differential operator accounts for the fact that the boundary data for the backpropagation enters as a source and not as a boundary condition. This operator is given by

$$-2i D_t c^{-1} \sqrt{1 - c^2 D_t^{-2} D_{\mathbf{x}'}^2}, \quad D_t = i^{-1} \partial_t, \quad D_{\mathbf{x}'} = i^{-1} \partial_{\mathbf{x}'}. \quad (4.65)$$

The singularity in the square root is avoided by the cutoff for grazing rays, see below. The second type of factor is composed of three cutoffs:

- (i) The multiplication by a cutoff function that smoothly goes to zero near the boundary of the acquisition domain. The distance over which it goes from 1 to 0 in practice depends on the wavelengths present in the data.
- (ii) The second cutoff is a pseudo-differential operator which removes waves that reach the surface along tangentially incoming rays. Its symbol is zero around $(\mathbf{x}', t, \boldsymbol{\xi}', \omega)$ such that

$$c(\mathbf{y}', 0) |(\boldsymbol{\eta}', 0)| = \pm \omega,$$

and 1 some distance away from this set. If, given the velocity and the support of δc , there are no tangent rays, this cutoff is not needed.

- (iii) The third cutoff suppresses direct rays. Since the velocity model is assumed to be known, these can be identified.

We write $\Psi_M(\mathbf{x}', t, \boldsymbol{\xi}', \omega)$ for the symbol of the composition of the three pseudo-differential cutoffs. The principal symbol of F_M is then

$$-2i\omega c^{-1} \sqrt{1 - c^2 \omega^{-2} \boldsymbol{\xi}'^2} \Psi_M(\mathbf{x}', t, \boldsymbol{\xi}', \omega). \quad (4.66)$$

The decoupling procedure presented above yields two fields u_a and u_b , associated respectively with the negative and positive frequencies in u . We will show that $u_{r,a}$ and $u_{r,b}$ depend locally on u_a and u_b in the following fashion,

$$\begin{aligned} \chi u_{r,a}(\cdot, t) &= \chi [P_a(t)u_a(\cdot, t) + R_1(t)u_b(\cdot, t)] \quad \text{and} \\ \chi u_{r,b}(\cdot, t) &= \chi [P_b(t)u_b(\cdot, t) + R_2(t)u_a(\cdot, t)]. \end{aligned} \quad (4.67)$$

Here, $P_a(t)$ and $P_b(t)$ are pseudo-differential operators described below and $R_1(t)$ and $R_2(t)$ are regularizing operators, and χ is a cutoff because the source in equation (4.64) causes waves in both sides of $x_n = 0$. Note that the decoupling, which so far was mostly a technical procedure, turns out to be essential to characterize the reverse time continued field. The revert operator in matrix form will be defined as the t -family of pseudo-differential operators

$$P(t) = V \begin{pmatrix} P_a(t) & R_1(t) \\ R_2(t) & P_b(t) \end{pmatrix} \Lambda. \quad (4.68)$$

Waves are assumed to hit the set M coming from $x_n > 0$. We assume $\text{supp}(\chi)$ to be compact and contained in the set $x_n > 0$, and we invoke the assumption that

$$\begin{aligned} &\text{bicharacteristics through } M \text{ and } \text{supp}(\chi) \\ &\text{intersect } M \text{ only once and with } dx_n/dt < 0. \end{aligned} \quad (4.69)$$

The operators P_a and P_b depend on F_M and on the bicharacteristic flow in space-time between the hyperplanes $t = 0$ and $x_n = 0$. Let X_s denote the set $\mathbb{R}^n \times \{s\} \subset \mathbb{R}_x^n \times \mathbb{R}_t$. The bicharacteristic flow provides a map³

$$(\mathbf{x}, \boldsymbol{\xi}) \mapsto (\mathbf{y}'_a(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}'_a(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|),$$

from T^*X_0 to T^*M . The principal symbols of P_a and P_b , which we will denote by p_a and p_b , are then given by the following transported versions of Ψ_M :

$$\Psi_{X_{0,a}}(\mathbf{x}, \boldsymbol{\xi}) = \begin{cases} \Psi_M(\mathbf{y}'_a(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}'_a(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|) & \dots \\ & \text{when } \exists t \text{ with } \mathbf{y}_a(\mathbf{x}, \boldsymbol{\xi}, t) \in M \\ 0 & \text{otherwise,} \end{cases} \quad (4.70)$$

³The spatial function $\mathbf{y}_a(\mathbf{x}, \boldsymbol{\xi}, t)$ is identical to $\mathbf{y}(\mathbf{x}, t; \boldsymbol{\xi})$ introduced in section 4.2.2 and 4.2.3.

and $\Psi_{X_{0,b}}$ is defined similarly using the $(\mathbf{y}_b, \boldsymbol{\eta}_b)$ -flow, associated with the negative frequencies. We can now state and prove the following

Theorem 5. *Let $u_{r,a}$, $u_{r,b}$, u_a , u_b , χ and M be as just defined. Equation (4.67) holds, in which P_a and P_b are pseudo-differential operators in $\text{Op} S^0(\mathbb{R}^n)$. Their principal symbols are respectively given by*

$$p_a(t; \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_{0,a}}(\mathbf{x}, \boldsymbol{\xi}) \quad \text{and} \quad p_b(t; \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_{0,b}}(\mathbf{x}, \boldsymbol{\xi}). \quad (4.71)$$

The operators P_a and P_b satisfy property (4.53) as far as they are uniquely determined considering the cutoff χ in (4.67).

The proof will be presented in the remainder of this section. If we take Cauchy values at $t = t_c$, then for small $|t - t_c|$, $T_M u(\mathbf{x}', t)$ can be described by the local FIO representation of the solution operator. This representation can also be used for the description of the map from $T_M u$ to $u_r(\cdot, t_c)$. The result can then be proven by an explicit use of the method of stationary phase. For longer times we apply a partition of unity in time to $T_M u(\mathbf{x}', t)$, so that for each contribution the length of the time interval is small enough to apply the local FIO representation. Egorov's theorem will be used to reduce to the short time case. Alternatively one could consider one-way wave theory as a method of proof.

Proof. We prove (4.67) for some given t . Without loss of generality we may assume that $t = 0$. The field u by assumption solves the homogeneous wave equation and is determined by the Cauchy values $u_{0,a} = u_a(\cdot, 0)$ and $u_{0,b} = u_b(\cdot, 0)$, modulo a smooth contribution. Consider the equation $\partial_t u_a = -iB u_a$. In this proof we write $S_a(t, s)$ instead of $S_a(t - s)$ for the operator that maps initial values at time s to the values of the solution at time t . We write $S_a(\cdot, s)$ for the operator that maps an initial value at time s to the solution as a function of (\mathbf{x}, t) with $t > s$. We will write $S_a(t, \cdot)$ for the operator that gives the anticausal solution to $(\partial_t + iB)u_a = f_a$,

$$S_a(t, \cdot) f_a = - \int_t^\infty S_a(t, s) f_a(\cdot, s) ds.$$

Note that $S_a(t, \cdot)$ maps a function of (\mathbf{x}, t) to a function of \mathbf{x} . The restriction operator T_M introduced above maps $C^\infty(\mathbb{R}^n \times \mathbb{R}) \rightarrow C^\infty(M)$ and is given by

$$T_M u(\mathbf{x}', t) = u(\mathbf{x}', 0, t), \quad (\mathbf{x}', t) \in M.$$

With auxiliary function f , the adjoint of the operator is given by

$$T_M^* f(\mathbf{x}, t) = \delta(x_n) f(\mathbf{x}', t).$$

These operators are well defined on suitable sets of distributions. We will use the notation (cf. (4.64)):

$$f_M(\mathbf{x}', t) = F_M T_M u(\mathbf{x}', t)$$

and study the map $(u_{0,a}, u_{0,b}) \mapsto f_M$. It follows from the results about the decoupling that

$$f_M = F_M T_M (cS_a u_{0,a} + cS_b u_{0,b}), \quad (4.72)$$

modulo a smooth error. Following this decoupling, we will analyze the following map $u_{0,a} \mapsto F_M T_M cS_a u_{0,a}$.

To begin with, there exists a pseudo-differential operator \tilde{F}_M such that

$$F_M T_M u = T_M \tilde{F}_M u$$

modulo a smooth function. This holds for a distribution u that satisfies $|\boldsymbol{\xi}| \leq C|\omega|$ in $\text{WF}(u)$ for some C , like the solution of the homogeneous wave equation. Naturally, $\tilde{F}_M(\mathbf{x}, t, \boldsymbol{\xi}, \omega) \neq F_M(\mathbf{x}', t, \boldsymbol{\xi}', \omega)$, because then the symbol property would not be satisfied around the line $(\boldsymbol{\xi}', \omega) = 0$, $\xi_n \neq 0$, but in the neighborhood of this line the symbol can be modified without affecting the singularities since $|\boldsymbol{\xi}| \leq C|\omega|$ in $\text{WF}(u)$. Thus the first term in (4.72) is given by

$$T_M \tilde{F}_M cS_a(\cdot, 0) \quad (4.73)$$

acting on $u_{0,a}$, which is a product of Fourier integral operators. The operator $S_a(\cdot, 0)$ has canonical relation

$$\{((\mathbf{y}_a(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}_a(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|), (\mathbf{x}, \boldsymbol{\xi}))\}. \quad (4.74)$$

The operator \tilde{F}_M removes singularities propagating on rays that are tangent or close to tangent to the plane $x_n = 0$, and the restriction operator to $x_n = 0$ has canonical relation

$$\{((\mathbf{y}', t, \boldsymbol{\eta}', \omega), (\mathbf{y}', 0, t, \boldsymbol{\eta}', \eta_n, \omega))\}. \quad (4.75)$$

As tangent rays are removed, the composition of the canonical relations (4.75) and (4.74) is transversal. Therefore, (4.73) is a Fourier integral operator. Moreover, from assumption (4.69) it follows that the canonical relation is the graph of an invertible map, given by

$$\{((\mathbf{y}'_a(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}'_a(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|), (\mathbf{x}, \boldsymbol{\xi})) \mid t \text{ s.t. } \mathbf{y}_a(\mathbf{x}, \boldsymbol{\xi}, t) \in M\},$$

or more precisely a subset of it, taking into account the essential support of F_M .

Next, we consider the map $f_M \mapsto \chi u_{r,a}$. We insert a pseudo-differential cutoff $\Xi(\mathbf{x}', t, D_{\mathbf{x}'}, D_t)$. It cuts out tangent rays and is defined such that $\Xi F_M = F_M$. Using the decoupling procedure of section 4.2.5, the source (f_a, f_b) for the inhomogeneous wave equation are given by $(f_a, f_b)^T = \Lambda(0, cf)^T$, hence $\chi u_{r,a}$ satisfies

$$\chi u_{r,a}(\cdot, 0) = \chi S_a(0, \cdot) (\frac{1}{2} B^{-1} c) T_M^* \Xi f_M.$$

There exists an operator $\tilde{\Xi}$ such that $T_M^* \Xi f = \tilde{\Xi} T_M^* f$ at least microlocally on the set $|\boldsymbol{\xi}| \leq C|\omega|$ for large C . Then $\chi u_{r,a}(\cdot, 0)$ is given by the operator

$$\chi S_a(0, \cdot) (\frac{1}{2} B^{-1} c) \tilde{\Xi} T_M^*$$

acting on f_M , modulo a smoothing operator.

The operator $\chi S_a(0, \cdot) (\frac{1}{2} B^{-1} c) \tilde{\Xi}$ is a Fourier integral operator with the following canonical relation

$$\{((\mathbf{x}, \boldsymbol{\xi}), (\mathbf{y}(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|)) \mid |\eta_n(\mathbf{x}, \boldsymbol{\xi}, t)| \geq \epsilon, \epsilon > 0\}.$$

For an element $(\mathbf{y}', 0, \boldsymbol{\eta}', \omega)$ with $|\omega| > c|\boldsymbol{\eta}'|$ there are two rays associated, namely with $\eta_n = \pm \sqrt{c^{-2}\omega^2 - |\boldsymbol{\eta}'|^2}$. The $+$ sign propagates into $x_n < 0$ for decreasing time, the $-$ sign points into $x_n > 0$. The contributions are well separated because of the cutoff for tangent rays present in F_M . Because of the assumption (4.69) and the cutoff χ , the contributions with $+$ sign can be ignored. We write $S_a^{(-)}(0, \cdot) (\frac{1}{2} B^{-1} c) \tilde{\Xi} T_M^*$ for the Fourier integral operator that propagates only the singularities from M into the $x_n > 0$ region for decreasing time. By a similar reasoning as above, this is a Fourier integral operator with canonical relation contained in

$$\{((\mathbf{x}, \boldsymbol{\xi}), (\mathbf{y}'(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}'(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|)) \mid y_n(\mathbf{x}, \boldsymbol{\xi}, t) = 0\}.$$

Again this is an invertible canonical relation.

In the next step we consider the composition of the maps $(u_{0,a}, u_{0,b}) \mapsto f_M$ and $f_M \mapsto (u_{r,a}(\cdot, 0), u_{r,b}(\cdot, 0))$. As both maps are Fourier integral operators with a canonical relation that is the graph of an invertible map, the composition is a (sum of) well defined Fourier integral operators. The fields u_a and $u_{r,a}$ are associated with negative ω , u_b and $u_{r,b}$ with positive ω . One can verify that the ‘cross terms’ $u_{0,a} \mapsto u_{r,b}(\cdot, 0)$ and $u_{0,b} \mapsto u_{r,a}(\cdot, 0)$ are smoothing operators. The maps $u_{0,a} \mapsto u_{r,a}(\cdot, 0)$ and $u_{0,b} \mapsto u_{r,b}(\cdot, 0)$ are pseudo-differential operators. The principal symbol $p_a(0; \mathbf{x}, \boldsymbol{\xi})$ is the product of Ψ_{X_0} and another factor.

We proceed under the assumption that $\Psi_M(\mathbf{x}', t, \boldsymbol{\xi}', t)$ is supported in the region $0 < t < t_1$, with t_1 sufficiently small such that the explicit form of the Fourier integral operator can be used. This assumption will be lifted at the end of the proof. We treat only the map $u_{0,a} \mapsto u_{r,a}(\cdot, 0)$, the map $u_{0,b} \mapsto u_{r,b}(\cdot, 0)$ can be done in a similar way. The map $u_{0,a} \mapsto f_M$ can then be written in the form

$$f_{r,a}(\mathbf{y}', 0) = \frac{1}{(2\pi)^n} \iint a^{(\text{fwd})}(\mathbf{y}', t, \mathbf{x}, \boldsymbol{\xi}) e^{i(\alpha(\mathbf{y}', 0, t, \boldsymbol{\xi}) - \mathbf{x} \cdot \boldsymbol{\xi})} u_{0,a}(\mathbf{x}) d\boldsymbol{\xi} d\mathbf{x},$$

where the amplitude satisfies

$$a^{(\text{fwd})}(\mathbf{y}', t, \mathbf{x}, \boldsymbol{\xi}) = -2i\chi(x_n)\omega\sqrt{1 - \frac{c(\mathbf{y}', 0)^2\boldsymbol{\eta}'^2}{\omega^2}}\sqrt{\det(\partial_{\mathbf{y}}\mathbf{x})}\Psi_M(\mathbf{y}, t, \boldsymbol{\eta}', \omega) \pmod{S^0(\mathbb{R}^{2n} \times \mathbb{R}^n)}, \quad (4.76)$$

where $\omega = \partial_t\alpha = -c(\mathbf{x})|\boldsymbol{\xi}|$, $\boldsymbol{\eta} = \partial_{\mathbf{y}}\alpha$ and $\det(\partial_{\mathbf{y}}\mathbf{x})$ is the Jacobian of the ray flow as explained earlier. The *adjoint* of the map $f_M \mapsto \chi u_{r,a}(\cdot, 0)$ is given by $\Xi^*T_M c \frac{1}{2}B^{-1}S_a(\cdot, 0)\chi$, and is a Fourier integral operator with the same phase function $\alpha(\mathbf{y}', 0, t, \boldsymbol{\xi}) - \mathbf{x} \cdot \boldsymbol{\xi}$ and amplitude

$$a^{(\text{bkd})}(\mathbf{y}', t, \mathbf{z}, \boldsymbol{\zeta}) = \frac{1}{2}\chi(z_n)(-\omega^{-1})c(\mathbf{y})\sqrt{\det(\partial_{\mathbf{y}}\mathbf{x}(\mathbf{y}', t, \boldsymbol{\zeta}))}\Xi \pmod{S^{-2}(\mathbb{R}^{2n} \times \mathbb{R}^n)}. \quad (4.77)$$

The map $f_M \mapsto \chi u_{r,a}(\cdot, 0)$ is therefore given by, with notation \mathbf{z} instead of $\mathbf{x} \in \mathbb{R}^n$,

$$u_{r,a}(\mathbf{z}, 0) = \frac{1}{(2\pi)^n} \iiint \overline{a^{(\text{bkd})}(\mathbf{y}', t, \mathbf{z}, \boldsymbol{\zeta})} e^{i(-\alpha(\mathbf{y}', 0, t, \boldsymbol{\zeta}) + \mathbf{z} \cdot \boldsymbol{\zeta})} f_M(\mathbf{y}', t) d\boldsymbol{\zeta} d\mathbf{y}' dt.$$

Therefore, the map $u_{0,a} \mapsto \chi(z_n)u_{r,a}(\cdot, 0)$ has distribution kernel $K(\mathbf{z}, \mathbf{x})$ given by

$$\frac{1}{(2\pi)^{2n}} \iiint \overline{a^{(\text{bkd})}(\mathbf{y}', t, \mathbf{z}, \boldsymbol{\zeta})} a^{(\text{fwd})}(\mathbf{y}', t, \mathbf{x}, \boldsymbol{\xi}) \dots e^{i(-\alpha(\mathbf{y}', 0, t, \boldsymbol{\zeta}) + \alpha(\mathbf{y}', 0, t, \boldsymbol{\xi}) + \mathbf{z} \cdot \boldsymbol{\zeta} - \mathbf{x} \cdot \boldsymbol{\xi})} d\mathbf{y}' dt d\boldsymbol{\xi} d\boldsymbol{\zeta}. \quad (4.78)$$

Using a smooth cutoff the $(\boldsymbol{\xi}, \boldsymbol{\zeta})$ integration domain can be divided into three parts, one with $|\boldsymbol{\zeta}| \leq 2|\boldsymbol{\xi}|$, one with $|\boldsymbol{\zeta}| \geq \frac{4}{3}|\boldsymbol{\xi}|$, and a third part containing $(\boldsymbol{\zeta}, \boldsymbol{\xi}) = (\mathbf{0}, \mathbf{0})$. In the first part, the method of stationary phase can be applied to the integral over $(\mathbf{y}', t, \boldsymbol{\zeta})$ using $|\boldsymbol{\xi}|$ as large parameter. We show that there is a function $g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ for which holds that

$$\frac{1}{(2\pi)^n} \iiint \overline{a^{(\text{bkd})}} a^{(\text{fwd})} e^{i(-\alpha(\mathbf{y}', 0, t, \boldsymbol{\zeta}) + \alpha(\mathbf{y}', 0, t, \boldsymbol{\xi}) + \mathbf{z} \cdot \boldsymbol{\zeta} - \mathbf{x} \cdot \boldsymbol{\xi})} d\mathbf{y}' dt d\boldsymbol{\zeta} = g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) e^{i(\mathbf{z} - \mathbf{x}) \cdot \boldsymbol{\xi}}, \quad (4.79)$$

and such that $g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ is a symbol that has an asymptotic series expansion with leading order term satisfying $g(\mathbf{x}, \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_0}(\mathbf{x}, \boldsymbol{\xi})$.

The first step in this computation is to determine the stationary points of the phase function

$$\Phi : (\mathbf{y}', t, \boldsymbol{\zeta}) \mapsto -\alpha(\mathbf{y}', 0, t, \boldsymbol{\zeta}) + \alpha(\mathbf{y}', 0, t, \boldsymbol{\xi}) + \mathbf{z} \cdot \boldsymbol{\zeta} - \mathbf{x} \cdot \boldsymbol{\xi}.$$

By the properties of α , $\frac{\partial}{\partial(\mathbf{y}', t)}\Phi = 0$ if and only if $\frac{\partial\alpha}{\partial(\mathbf{y}', t)}(\mathbf{y}', 0, \zeta) = \frac{\partial\alpha}{\partial(\mathbf{y}', t)}(\mathbf{y}', 0, \xi)$ if and only if $(\mathbf{y}', 0, t, \zeta)$ and $(\mathbf{y}', 0, t, \xi)$ are associated with the same bicharacteristic and hence $\zeta = \xi$. Requiring that the derivative with respect to ζ is 0 gives that

$$-\partial_{\xi}\alpha(\mathbf{y}, t, \xi) + \mathbf{z} = 0.$$

Therefore, the bicharacteristic determined by (\mathbf{z}, ξ) must be the same as the bicharacteristic determined by (\mathbf{y}', t, ζ) . Let $\psi(\mathbf{y}', t, \zeta; \mathbf{x}, \xi)$ be a C^{∞} cutoff function that is one for a small neighborhood of (\mathbf{y}', t, ζ) around the stationary value, and zero outside a slightly larger neighborhood. From the lemma of nonstationary phase one can derive that the contribution to g from the region away from the stationary point set is in $S^{-\infty}$.

At this point, observe that the second part, with $|\zeta| \geq \frac{4}{3}|\xi|$, can be treated similarly with the role of ζ and ξ interchanged. In this case the stationary set is in the region where the amplitude is zero, and its contribution is of the form (4.79), but with g in $S^{-\infty}(\mathbb{R}^{2n} \times \mathbb{R}^n)$. The third part, ζ, ξ around zero, also yields such a contribution with $g \in S^{-\infty}(\mathbb{R}^{2n} \times \mathbb{R}^n)$.

To treat the case (\mathbf{y}', t, ζ) around the stationary point set, we apply a change of variables in the phase function. Setting $y_n = 0$, it can be written as

$$\begin{aligned} \alpha(\mathbf{y}, t, \xi) - \alpha(\mathbf{y}, t, \zeta) &= \int_0^1 \frac{\partial}{\partial s} \alpha(\mathbf{y}, t, \zeta + s(\xi - \zeta)) ds \\ &= (\xi - \zeta) \cdot \int_0^1 \frac{\partial\alpha}{\partial\xi}(\mathbf{y}, t, \zeta + s(\xi - \zeta)) ds \\ &\stackrel{\text{def}}{=} (\xi - \zeta) \cdot X(\mathbf{y}', t, \zeta, \xi). \end{aligned}$$

The goal is to rewrite the distribution kernel (4.78) by a change of variables into

$$\frac{1}{(2\pi)^n} \iiint \psi \overline{a^{(\text{bkd})}} a^{(\text{fwd})} e^{i((\xi-\zeta) \cdot X + \mathbf{z} \cdot \zeta - \mathbf{x} \cdot \xi)} \left| \frac{\partial X}{\partial(\mathbf{y}', t)} \right|^{-1} d\zeta dX. \quad (4.80)$$

So the next step is to prove that $\frac{\partial X}{\partial(\mathbf{y}', t)}$ is an invertible matrix at the stationary points. It is clear that, with $\xi = \zeta$ and $y_n = 0$, one has

$$X(\mathbf{y}', t, \xi, \xi) = \frac{\partial\alpha}{\partial\xi}(\mathbf{y}, t, \xi) = \mathbf{x}(\mathbf{y}, t, \xi),$$

where $(\mathbf{y}, t, \xi) \mapsto \mathbf{x}(\mathbf{y}, t, \xi)$ was discussed in subsection 4.2.4. The matrix $\frac{\partial\mathbf{x}}{\partial\mathbf{y}}$ is non-degenerate. Then we apply the implicit function theorem to the map $\mathbf{x} \mapsto (\mathbf{y}', t)$ obtained by setting $\mathbf{y}' = \mathbf{y}'(\mathbf{x}, \tilde{t})$ in which \tilde{t} is such that $y_n(\mathbf{x}, \tilde{t}) = 0$, and use that there are no tangent rays, to obtain that the matrix $\frac{\partial\mathbf{x}}{\partial(\mathbf{y}', t)}$ has maximal rank at

the stationary points, while the Jacobian satisfies

$$\left| \frac{\partial \mathbf{x}}{\partial(\mathbf{y}', t)} \right| = \left| \frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right| \left| \frac{\partial y_n}{\partial t} \right|.$$

The integral (4.80) has a quadratic phase function $\zeta \cdot (X - \mathbf{z})$, and can be performed as usual in the method of stationary phase [22, lemma 1.2.4]. This shows that $g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ satisfies the symbol property. Using (4.76) and (4.77) it follows that

$$g(\mathbf{x}, \mathbf{x}, \boldsymbol{\xi}) = -2i(-c|\boldsymbol{\xi}|) \sqrt{1 - \frac{c^2 \boldsymbol{\eta}'^2}{\omega^2}} \left| \frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right|^{\frac{1}{2}} \Psi_M \quad \dots$$

$$\left(-\frac{i}{2} \right) c(\mathbf{x}) (c|\boldsymbol{\xi}|)^{-1} c(\mathbf{y}) \left| \frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right|^{\frac{1}{2}} \left| \frac{\partial \mathbf{x}}{\partial \mathbf{y}} \right|^{-1} \left| \frac{\partial y_n}{\partial t} \right|^{-1}$$

Two terms need to be worked out, namely $\sqrt{1 - c^2 \omega^{-2} \boldsymbol{\eta}'^2} = \cos(\theta_M)$, in which θ_M is the angle of incidence of a ray at M , and $\left| \frac{\partial y_n}{\partial t} \right| = c(\mathbf{y}) \cos(\theta_M)$. Therefore indeed we have

$$g(\mathbf{x}, \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_0} \quad \text{mod } S^{-1}(\mathbb{R}^{2n} \times \mathbb{R}^n).$$

This concludes the proof of the small time result.

Next we extend this to the result for longer times. By a partition of unity we can write Ψ_M as a sum of terms with $t \in [s, s + t_1]$ for some s . It is sufficient to prove the result for each term, and we may therefore assume $t \in [s, s + t_1]$ in the support of Ψ_M . By a change of variable t to $t - s$, it follows that

$$P_a(s) \stackrel{\text{def}}{=} S_a^{(-)}(s, \cdot) \left(-\frac{i}{2} B^{-1} c \right) T_M^* F_M T_M S_a(\cdot, s) \in \text{Op } S^0(\mathbb{R}^n)$$

with principal symbol

$$\Psi_{X_s}(\mathbf{x}, \boldsymbol{\xi}) = \Psi_M(\mathbf{y}'(\mathbf{x}, \boldsymbol{\xi}, t - s), t, \boldsymbol{\eta}'(\mathbf{x}, \boldsymbol{\xi}, t - s), -c(\mathbf{x})|\boldsymbol{\xi}|),$$

in which t is such that $y_n(\mathbf{x}, \boldsymbol{\xi}, t - s) = 0$. From the group property of the operator $S_a(t, s)$ it follows that

$$P_a(0) = \chi S_a(0, s) P_a(s) S_a(s, 0).$$

The evolutions operators $S_a(0, s)$ and $S_a(s, 0)$ are each others inverses. According to the Egorov theorem [67, section 8.1] the operator $P_a(0)$ is a pseudo-differential operator. For the symbol we find that it is given by $(\mathbf{x}, \boldsymbol{\xi}) \mapsto \Psi_{X_s}(\mathbf{y}(\mathbf{x}, \boldsymbol{\xi}, s), \boldsymbol{\eta}(\mathbf{x}, \boldsymbol{\xi}, s))$, i.e. by Ψ_{X_0} . This completes the proof. \square

4.5 Inverse scattering

This section deals with the inverse scattering problem. The diagram in figure 4.2 shows how we theoretically approach RTM. The forward modeling is given by $r \rightarrow u \rightarrow d$ in the diagram. The *reflectivity* function r causes a scattered wave field u , giving the *data* d by restriction to the surface $x_n = 0$. Recall $\mathbf{x}' = [\mathbf{x}]_{1:n-1}$. The bottom line of the diagram shows the inverse modeling. Data d is propagated in reverse time to the *reverse time continued* wave field u_r . This wave field is mapped by the *imaging operator* G to the *image* i . The *resolution operator* R is the map from the reflectivity to the image as result of the forward modeling and the inversion. The scattering operator F maps the reflectivity to the continued scattered wave u_h . As explained, this field can be seen as the receiver wave field in an idealized experiment. It contains all rays that are present in the scattered wave, regardless whether they can be reconstructed by RTM. The *revert operator* P removes parts that are not present in the receiver wave field. The field u_h , central to the analysis, is not actually computed.

We obtain the main result, the *imaging condition* (4.93), in two steps. We propose the *imaging operator* G and show in theorem 7 and its proof that it is a FIO that maps the *reverse time continued wave field* to an image of the reflectivity. Hence it is an approximate inverse of the scattering operator. From this operator we subsequently derive an imaging condition in terms of solutions of partial differential equations, g and u_r . We first discuss a simplified case with constant coefficient.

For the RTM based inversion we assume that the following holds, which replaces the condition (4.69):

$$\begin{aligned} &\text{Bicharacteristics that enter the region } x_n < 0 \\ &\text{do not return to the region } x_n \geq 0. \end{aligned} \tag{4.81}$$

This will ensure that u_r is properly defined for the purpose of linearized inversion. We also recall the assumption that there is no source wave field multipathing, formalized as the property (4.54). The assumption that there are no direct rays from the source to a receivers is incorporated in P , i.e. by means of Ψ_M , cf. (4.66).

4.5.1 Constant background velocity

In this subsection we consider the case of constant background velocity c with a planar incoming wave, propagating in the positive x_3 -direction. The scattered field will be described by

$$[c^{-2}\partial_t^2 - \Delta] u(\mathbf{x}, t) = A \delta(t - c^{-1}x_3)r(\mathbf{x}), \tag{4.82}$$

which is a slight simplification of the IVP (4.43). For simplicity the analysis will be 3-dimensional, but it applies to other dimensions as well.

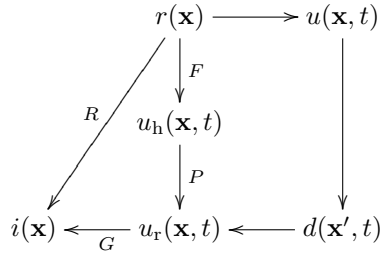


Figure 4.2: Diagram showing the theoretical approach to RTM.

The solution of the PDE (4.82) is given in the (ξ, t) domain by

$$\widehat{u}(\xi, t) = \int_0^t \left(e^{ic|\xi|(t-s)} - e^{-ic|\xi|(t-s)} \right) \frac{c^2}{2ic|\xi|} \widehat{f}(\xi, s) ds, \quad (4.83)$$

where, for now, we denote by f the right hand side of (4.82). The Fourier transform of f is hence needed. Let $\tilde{r}(\xi_1, \xi_2, x_3)$ be the Fourier transform of r with respect to (x_1, x_2) but not x_3 . The Fourier transform of $A\delta(t - \frac{x_3}{c})r(\mathbf{x})$ is given by

$$\int e^{-ix_3\xi_3} A\delta(t - \frac{x_3}{c})\tilde{r}(\xi_1, \xi_2, x_3) dx_3 = cAe^{-i\xi_3 ct}\tilde{r}(\xi_1, \xi_2, ct). \quad (4.84)$$

Next we use (4.83) and (4.84) to solve (4.82), and we subsequently make a change of variable $cs = \tilde{z}$. This yields the integral formula

$$\widehat{u}(\xi, t) = \int_0^{tc} \left(e^{i|\xi|(ct-\tilde{z})} - e^{-i|\xi|(ct-\tilde{z})} \right) \frac{c^2}{2ic|\xi|} Ae^{-i\xi_3\tilde{z}}\tilde{r}(\xi_1, \xi_2, \tilde{z}) d\tilde{z}.$$

We can recognize in this formula a Fourier transformation with respect to \tilde{z} . However, the Fourier transform of r is not evaluated at ξ_3 , but at $\xi_3 \pm |\xi|$, because \tilde{z} occurs at several places in the complex exponents. We assume that the support of r is contained in $0 < x_3 < ct$. In words this means that the field at time t is such that the incoming wave front has completely passed the support of the reflectivity. The integral formula then equals

$$\widehat{u}(\xi, t) = e^{i|\xi|ct} \frac{c^2 A}{2ic|\xi|} \widehat{r}(\xi + (0, 0, |\xi|)) - e^{-i|\xi|ct} \frac{c^2 A}{2ic|\xi|} \widehat{r}(\xi - (0, 0, |\xi|)). \quad (4.85)$$

The field in position coordinates is given by the inverse Fourier transform of the

the formula. Working this out leads to

$$u(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \dots \left[e^{i|\boldsymbol{\xi}|ct} \frac{c^2 A}{2ic|\boldsymbol{\xi}|} \widehat{r}(\boldsymbol{\xi} + (0, 0, |\boldsymbol{\xi}|)) - e^{-i|\boldsymbol{\xi}|ct} \frac{c^2 A}{2ic|\boldsymbol{\xi}|} \widehat{r}(\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)) \right] e^{i\mathbf{x} \cdot \boldsymbol{\xi}} d\boldsymbol{\xi} \quad (4.86)$$

The two terms yield complex conjugate contributions after integration. To see this, change the integration variables in the second term to $-\boldsymbol{\xi}$, and use the property that $r(\mathbf{x})$ is real, which is equivalent to $\widehat{r}(\boldsymbol{\xi}) = \overline{\widehat{r}(-\boldsymbol{\xi})}$ for all $\boldsymbol{\xi}$. Therefore

$$u(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \operatorname{Re} \int_{\mathbb{R}^3} e^{-i|\boldsymbol{\xi}|ct + i\mathbf{x} \cdot \boldsymbol{\xi}} \frac{icA}{|\boldsymbol{\xi}|} \widehat{r}(\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)) d\boldsymbol{\xi}. \quad (4.87)$$

There are three wave vectors in (4.87): $\boldsymbol{\xi}$ is the wave vector of the outgoing reflected wave, $(0, 0, |\boldsymbol{\xi}|)$ can be interpreted as the wave vector of the incoming wave, and $\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)$ is the wave vector of the reflectivity. The last one can be interpreted as the normal direction in case of a reflecting interface.

In this simplified analysis we assume that the reverse time continued wave field u_r solves the homogeneous wave equation and adopts the Cauchy values of u pinpointed at some time after the scattering event, like u_h in (4.44). This means that it results from an idealized experiment as explained in section 4.3.1. Then u_r is also given by (4.87), except that this formula is now valid for all t .

The basic idea of imaging is to time-correlate the source field with the receiver field. Approximating the source field by $A\delta(t - x_3/c)$ this becomes evaluating the receiver field at the arrival time of the incoming wave and multiplication by amplitude A . Hence, a first guess for the image would be $I_0 = Au(\mathbf{x}, x_3/c)$. This, however will not yield an inverse. Using some advance knowledge we will define our image instead as

$$I(\mathbf{x}) = \frac{2}{c^2 A} (\partial_t + c\partial_{x_3})u(\mathbf{x}, x_3/c). \quad (4.88)$$

We derive from equation (4.87)

$$\begin{aligned} \frac{2}{c^2 A} (\partial_t + c\partial_{x_3})u(\mathbf{x}, t) = \dots \\ \frac{2}{(2\pi)^3} \operatorname{Re} \int_{\mathbb{R}^3} \left(1 - \frac{\xi_3}{|\boldsymbol{\xi}|} \right) e^{-i|\boldsymbol{\xi}|ct + i\mathbf{x} \cdot \boldsymbol{\xi}} \widehat{r}(\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)) d\boldsymbol{\xi}. \end{aligned} \quad (4.89)$$

And setting $t = x_3/c$ yields

$$I(\mathbf{x}) = \frac{2}{(2\pi)^3} \operatorname{Re} \int_{\mathbb{R}^3} \left(1 - \frac{\xi_3}{|\boldsymbol{\xi}|} \right) e^{i\mathbf{x} \cdot (\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|))} \widehat{r}(\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)) d\boldsymbol{\xi}. \quad (4.90)$$

We carry out the following coordinate transformation

$$\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|), \quad \left| \frac{\partial \tilde{\boldsymbol{\xi}}}{\partial \boldsymbol{\xi}} \right| = 1 - \frac{\xi_3}{|\boldsymbol{\xi}|}. \quad (4.91)$$

The codomain of this transformation is the halfplane $\tilde{\xi}_3 < 0$. The Jacobian is as given in (4.91) and can be recognized in the integrant of (4.90), which results from the derivative operator $\partial_t + c\partial_{x_3}$. Therefore by a change of variables (4.90) equals $\frac{1}{(2\pi)^3} \operatorname{Re} \int_{\tilde{\xi}_3 < 0} e^{i\mathbf{x} \cdot \tilde{\boldsymbol{\xi}}} \hat{r}(\tilde{\boldsymbol{\xi}}) d\tilde{\boldsymbol{\xi}}$. This can be rewritten into

$$I(\mathbf{x}) = \frac{1}{(2\pi)^3} \int_{\tilde{\xi}_3 \neq 0} e^{i\mathbf{x} \cdot \tilde{\boldsymbol{\xi}}} \hat{r}(\tilde{\boldsymbol{\xi}}) d\tilde{\boldsymbol{\xi}}. \quad (4.92)$$

The right-hand side is almost the inverse Fourier transform, except for the exclusion of the set $\tilde{\xi}_3 = 0$ from the integration domain. This expresses the difficulty of direct waves with respect to the inversion. This simple calculation gives the motivation for the imaging condition (4.93) that we will present below, in particular, for the term involving the gradient $\partial_{\mathbf{x}} \hat{u}_r(\mathbf{x}, \omega)$.

4.5.2 Imaging condition

We now present the main result of the paper. The *imaging condition* is a formula that maps the source wave field $g(\mathbf{x}, t)$ and the reverse time continued wave field $u_r(\mathbf{x}, t)$ to an image $i(\mathbf{x})$ of the reflectivity. We propose the following definition

$$i(\mathbf{x}) = \frac{1}{2\pi} \int \frac{\Omega(\omega)}{i\omega |\hat{g}(\mathbf{x}, \omega)|^2} \left(\overline{\hat{g}(\mathbf{x}, \omega)} \hat{u}_r(\mathbf{x}, \omega) - \frac{c(\mathbf{x})^2}{\omega^2} \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x}, \omega)} \cdot \partial_{\mathbf{x}} \hat{u}_r(\mathbf{x}, \omega) \right) d\omega \quad (4.93)$$

and show that it yields a partial inverse. Here $\Omega(\omega)$ is a smooth function, valued 0 on a bounded neighborhood of the origin, and 1 outside a slightly larger neighborhood. These neighborhoods are obtained in the proof of the theorem.

To characterize $i(\mathbf{x})$, the relation (4.52) between $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ is important. We observe that the inverse function $\boldsymbol{\xi}(\boldsymbol{\zeta})$ of (4.52) is defined on the halfspace

$$\{\boldsymbol{\zeta} \in \mathbb{R}^n \setminus 0 \mid \boldsymbol{\zeta} \cdot \mathbf{n}_s(\mathbf{z}) < 0\}. \quad (4.94)$$

The function $p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta}))$, p_a the principal symbol of the revert operator, is in principle defined only on (4.94). However, due to the DSE, it is zero for $\boldsymbol{\zeta}$ near the boundary of this halfspace and we will consider it as a function on $\mathbb{R}^n \setminus 0$ that is zero outside (4.94). With this definition, the function $(\mathbf{z}, \boldsymbol{\zeta}) \mapsto p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta}))$ is an order 0 symbol.

Theorem 6. Let image $i(\mathbf{x})$ be defined by the formula (4.93), and assume (4.35), (4.42) and (4.81). Define operator R by the map from the reflectivity r to the image, $Rr(\mathbf{x}) = i(\mathbf{x})$. Then R is a pseudo-differential operator of order zero, and its principal symbol satisfies

$$p.s.(R)(\mathbf{z}, \zeta) = p_a(T_s(\mathbf{z}); \mathbf{z}, \xi(\zeta)) + p_a(T_s(\mathbf{z}); \mathbf{z}, \xi(-\zeta)), \quad (4.95)$$

where the map $(\mathbf{z}, \zeta) \mapsto p_a(T_s(\mathbf{z}); \mathbf{z}, \xi(\zeta))$ is as just described.

Operator R will be referred to as the *resolution operator*. From the proof of the result it can be seen that the first contribution on the right-hand side of (4.95) corresponds to the negative frequencies and the second contribution to the positive frequencies. As the supports, i.e. (4.94) for the first, of these two terms are disjoint, (4.95) defines a symbol that is 1 on a subset of $\mathbb{R}^n \times \mathbb{R}^n \setminus 0$. Hence, the map $d \mapsto i$ given by formula (4.93) can rightfully be called a partial inverse.

The imaging condition (4.93) is based on the actual source field g . Before proving theorem 6, we derive an intermediate result with an imaging condition based on the source wave traveltime $T_s(\mathbf{x})$, and the highest order contribution to the amplitude $A_s(\mathbf{x})$. Let $w \in \mathcal{E}'(Y \times \mathbb{R})$ be an auxiliary distribution and let operators H and K be defined by

$$\begin{aligned} Hw(\mathbf{y}, t) &= \frac{1}{A_s(\mathbf{y})} \partial_t^{-\frac{n+1}{2}} [\partial_t + c(\mathbf{y})\mathbf{n}_s(\mathbf{y}) \cdot \partial_{\mathbf{y}}] w(\mathbf{y}, t), \\ Kw(\mathbf{z}) &= w(\mathbf{z}, T_s(\mathbf{z})). \end{aligned} \quad (4.96)$$

Operator K is a restriction to a hypersurface in \mathbb{R}^{n+1} . Operator H is a pseudo-differential operator. Operator $\partial_t^{-\frac{n+1}{2}}$ is to be read as the pseudo-differential operator with symbol $\omega \mapsto \tilde{\sigma}(\omega)(i\omega)^{-\frac{n+1}{2}}$ in which $\tilde{\sigma}$ is a smooth function, valued 1 except for the origin where it is 0. Because P and F are defined as matrix operators we define $V_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}$, which projects out the first component of a two-vector. We define the *imaging operator* $G = KH$.

Theorem 7. If assumptions (4.35), (4.42) and (4.81) hold and \tilde{R} is given by

$$\tilde{R}r = Gu_r = GV_1PFr, \quad (4.97)$$

then \tilde{R} is a pseudo-differential operator of order zero with principal symbol given by equation (4.95).

Proof. We first work out the details for the negative frequencies, which leads to a characterization of $\tilde{R}_a = KHC P_a F_a$. We then consider the positive frequencies and add the contributions, $\tilde{R} = \tilde{R}_a + \tilde{R}_b$.

(i) We show that the composition $\tilde{R}_a = KHcP_aF_a$ is a FIO and that it is microlocal, i.e. has canonical relation that is a subset of the identity. The kernel of operator K is an oscillatory integral,

$$Kw(\mathbf{z}) = (2\pi)^{-n-1} \iint e^{i\boldsymbol{\eta} \cdot (\mathbf{z}-\mathbf{y}) + i\omega(T_s(\mathbf{z})-t)} w(\mathbf{y}, t) d\mathbf{y}, t) d\boldsymbol{\eta} d\omega \quad (4.98)$$

with canonical relation

$$\Upsilon = \{((\mathbf{z}, \boldsymbol{\theta}), (\mathbf{y}, t, \boldsymbol{\eta}, \omega)) \mid (\mathbf{y}, \boldsymbol{\eta}) \in T^*\mathcal{Y} \setminus 0, t = T_s(\mathbf{y}), \omega \in \mathbb{R} \setminus 0, \dots \\ \mathbf{z} = \mathbf{y}, \boldsymbol{\theta} = \boldsymbol{\eta} + c(\mathbf{y})^{-1}\omega \mathbf{n}_s(\mathbf{y})\}. \quad (4.99)$$

First consider $K\pi_aF_a$, which is the composition of K and π_aF_a with canonical relations given respectively by Υ (4.99) and Λ (4.55). We consider the composition of the Fourier integrals K and π_aF_a , using the composition theorem based on the canonical relations, see [22, Theorem 2.4.1] or [70]. Let $((\mathbf{z}, \boldsymbol{\theta}), (\mathbf{x}, \boldsymbol{\zeta})) \in \Upsilon \circ \Lambda$ then there exist a $(\mathbf{y}, \boldsymbol{\eta}) \in T^*\mathcal{Y} \setminus 0$ that is not in $V_{s,t}$, time $t = T_s(\mathbf{y})$ and $\omega = -c(\mathbf{y})|\boldsymbol{\eta}|$ such that $((\mathbf{z}, \boldsymbol{\theta}), (\mathbf{y}, t, \boldsymbol{\eta}, \omega)) \in \Upsilon$ and $((\mathbf{y}, t, \boldsymbol{\eta}, \omega), (\mathbf{x}, \boldsymbol{\zeta})) \in \Lambda$. As a result one has $(\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-T_s(\mathbf{y})}(\mathbf{y}, \boldsymbol{\eta})$, which means that \mathbf{x} and \mathbf{y} are on the same ray and separated in time by $T_s(\mathbf{y}) - T_s(\mathbf{x})$. Condition (4.54) (SME) now implies that this ray must coincide with a source ray. As source rays are excluded, i.e. $(\mathbf{y}, \boldsymbol{\eta}) \notin V_{s,t}$, the only possibility is that $\mathbf{x} = \mathbf{y}$. The conclusion is that $(\mathbf{z}, \boldsymbol{\theta}) = (\mathbf{x}, \boldsymbol{\zeta})$.

It is straightforward to establish that the composition of canonical relations is transversal, and that the additional conditions of the composition theorem of FIOs are satisfied. Hence $K\pi_aF_a$ is a FIO with canonical relation contained in the identity. The operators H and P_a are pseudo-differential operators, and π_a and P_a can be constructed such that $\text{WF}(P_a w) \subset \text{WF}(\pi_a w)$ for all w . The conclusion is that $\tilde{R}_a = KHcP_aF_a$ is a FIO with identity canonical relation, and hence a pseudo-differential operator.

(ii) We show that $\tilde{R} = KHV_1PF$ is a pseudo-differential operator that can be written as the integral (4.113) below. For F we use the local expressions (4.46). Because P is a t -family of pseudo-differential operators and $F_a\rho$ is a t -family of FIOs, the composition $P_a(t)F_a(t)\rho$ is a FIO with phase inherited from $F_a(t)\rho$, i.e. φ_T . The highest order contribution to its amplitude is $p_a(t; \mathbf{y}, \partial_{\mathbf{y}}\varphi_T)A_F$. The composition with H can be done similarly, because $P_a(t)F_a(t)\rho$ can also be viewed as a FIOs with output variables (\mathbf{y}, t) . In this proof we will denote the highest order contribution to the amplitude of $HcP_a(t)F_a(t)\rho$ by $A_{\text{HPF}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})$. It can be written in the form

$$A_{\text{HPF}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) = \left(1 + \frac{c(\mathbf{y})\mathbf{n}_s(\mathbf{y}) \cdot \partial_{\mathbf{y}}\alpha}{\partial_t\alpha}\right) \frac{2ic(\mathbf{y})p_a(t_1; \mathbf{y}, \partial_{\mathbf{y}}\alpha)aA_s(\mathbf{x})\partial_t\alpha}{c(\mathbf{x})A_s(\mathbf{y})}. \quad (4.100)$$

For all occurrences of α and a the arguments are $(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi})$.

Next we consider the application of the restriction operator K . We have already argued that \tilde{R}_a is a FIO with canonical relation contained in the identity. This implies that, to prove the theorem, it is sufficient to do a local analysis using (4.46). The local analysis shows again that \tilde{R}_a is a pseudo-differential operator, but also gives the required explicit formula for the amplitude.

The local phase function of $KHcP_a(t)F_a(t)\rho$ will be denoted by $\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$. Applying K to φ_T , i.e. setting $t = T_s(\mathbf{z})$, yields

$$\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \alpha(\mathbf{z}, T_s(\mathbf{z}) - T_s(\mathbf{x}); \boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \mathbf{x}. \quad (4.101)$$

The stationary set of ψ , denoted by Ψ , is given by the triplets $(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ that solve

$$\partial_{\boldsymbol{\xi}}\alpha(\mathbf{z}, T_s(\mathbf{z}) - T_s(\mathbf{x}); \boldsymbol{\xi}) = \mathbf{x}. \quad (4.102)$$

The interpretation of $(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \in \Psi$ is that a ray with initial condition $(\mathbf{x}, \boldsymbol{\xi})$ arrives at \mathbf{z} after time lapse $T_s(\mathbf{z}) - T_s(\mathbf{x})$. Application of the SME and the DSE now implies that $\mathbf{z} = \mathbf{x}$.

Below we will define a transformation of covariables. To prepare for this, we introduce a smooth cutoff function $\chi : Z \times X \times \mathbb{R}^n \setminus 0 \rightarrow \mathbb{R}$ accordingly. A Fourier integral may be restricted to a neighborhood of the stationary point set at the expense of a regularizing operator. Therefore, $\chi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ is set to 1 in the neighborhood of Ψ and 0 elsewhere. This means that \mathbf{x} is close to \mathbf{z} in $\text{supp}(\chi)$. The second issue is related to the DSE, which is required for the definition of the transformation. The cutoff χ is assumed to also remove singularities on a neighborhood of the direct rays. We set $\chi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ to 0 if $\boldsymbol{\xi}$ lies within a narrow conic set with solid angle $\Omega(\mathbf{z})$ around the principal direction $\mathbf{n}_s(\mathbf{z})$. The solid angle $\Omega(\mathbf{z})$ will be discussed later. We can hence write

$$\tilde{R}_a r(\mathbf{z}) = (2\pi)^{-n} \iint e^{i\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})} \chi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) A_{\text{HPF}}(\mathbf{z}, T_s(\mathbf{z}), \mathbf{x}, \boldsymbol{\xi}) d\boldsymbol{\xi} r(\mathbf{x}) d\mathbf{x}. \quad (4.103)$$

in which, of course, the integration domain is implicitly restricted to $\text{supp}(\chi)$.

Next we introduce covariable $\boldsymbol{\theta}$ to transform phase ψ into the form $\boldsymbol{\theta} \cdot (\mathbf{z} - \mathbf{x})$. By definition $\boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = -\int_0^1 \partial_x \psi(\mathbf{z}, \tilde{\mathbf{x}}(\mu), \boldsymbol{\xi}) d\mu$ in which $\tilde{\mathbf{x}}(\mu) = \mathbf{z} + \mu(\mathbf{x} - \mathbf{z})$. The phase function now transforms into

$$\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \psi(\mathbf{z}, \mathbf{z}, \boldsymbol{\xi}) + \int_0^1 \partial_\mu [\psi(\mathbf{z}, \tilde{\mathbf{x}}(\mu), \boldsymbol{\xi})] d\mu = \boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \cdot (\mathbf{z} - \mathbf{x}). \quad (4.104)$$

To better understand the transformation and to determine the new domain of integration, i.e. $\boldsymbol{\theta}(\text{supp}(\chi))$, and the Jacobian we apply the chain rule to the definition of ψ . This leads to

$$\boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \boldsymbol{\xi} + \int_0^1 \partial_t \alpha(\mathbf{z}, T_s(\mathbf{z}) - T_s(\tilde{\mathbf{x}}); \boldsymbol{\xi}) \partial_x T_s(\tilde{\mathbf{x}}) d\mu.$$

There exists an $\tilde{\mathbf{x}}$ such that $(\mathbf{z}, \tilde{\mathbf{x}}, \boldsymbol{\xi}) \in \Gamma_{T_s(\mathbf{z})-T_s(\tilde{\mathbf{x}})}$, i.e. $\tilde{\mathbf{x}}$ and \mathbf{z} are connected by a ray. Note that $\tilde{\mathbf{x}} = \mathbf{x}_\Gamma(\mathbf{z}, T_s(\mathbf{z}) - T_s(\tilde{\mathbf{x}}), \boldsymbol{\xi})$ will do, see section 4.2.4 for notation \mathbf{x}_Γ . By using the identities $\partial_t \alpha = -c(\tilde{\mathbf{x}})|\boldsymbol{\xi}|$ and $c(\tilde{\mathbf{x}})\partial_{\tilde{\mathbf{x}}}T_s(\tilde{\mathbf{x}}) = \mathbf{n}_s(\tilde{\mathbf{x}})$, one gets

$$\boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \boldsymbol{\xi} - |\boldsymbol{\xi}| \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \quad \text{with} \quad \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \int_0^1 \frac{c(\tilde{\mathbf{x}})}{c(\tilde{\mathbf{x}})} \mathbf{n}_s(\tilde{\mathbf{x}}) d\mu. \quad (4.105)$$

The Jacobian follows from this result. By an easily verified calculation, one finds

$$|\partial_{\boldsymbol{\xi}} \boldsymbol{\theta}| = \left| \det \left(\mathbf{I}_n - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \otimes \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \right) \right| = \left| 1 - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \cdot \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \right|. \quad (4.106)$$

With these formulae at hand a sensible choice can be made for solid angle $\Omega(\mathbf{z})$. The angle must be large enough to meet the following inequality for all elements of the support $\text{supp}(\chi)$:

$$|\boldsymbol{\xi} \cdot \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})| < |\boldsymbol{\xi}| \min\{1, |\mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})|^2\}. \quad (4.107)$$

We will now give the motivation. For $\boldsymbol{\xi} \mapsto \boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ to be injective, given (\mathbf{z}, \mathbf{x}) , the Jacobian must be nonzero. This is true due to the inequality, which is nontrivial if $|\mathbf{n}| > 1$. This affirms the local invertibility, and an easy exercise proves its injectivity. A second argument concerns the domain of integration $\boldsymbol{\theta}(\text{supp}(\chi))$. The inequality guarantees that $\boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \cdot \mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) < 0$ for all points in $\text{supp}(\chi)$, which is nontrivial if $|\mathbf{n}| < 1$. This fact will play a role in gluing $\tilde{R}_a r$ and $\tilde{R}_b r$ together, which will be done in following paragraphs. Because \mathbf{x} is in the neighborhood of \mathbf{z} , so are $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}$. This implies that $\mathbf{n}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ is close to $\mathbf{n}_s(\mathbf{z})$, and $|\mathbf{n}| \approx 1$. This is as close as needed by narrowing the spatial part of the cutoff function χ around the diagonal of $Z \times X$.

By using the new variable $\tilde{R}_a r$'s integral expression (4.103) transforms into

$$\tilde{R}_a r(\mathbf{z}) = (2\pi)^{-n} \iint_{\boldsymbol{\theta}(\text{supp}(\chi))} A_{\tilde{R}}(\mathbf{z}, \mathbf{x}, \boldsymbol{\theta}) e^{i\boldsymbol{\theta} \cdot (\mathbf{z} - \mathbf{x})} d\boldsymbol{\theta} r(\mathbf{x}) d\mathbf{x}, \quad (4.108)$$

where we define

$$A_{\tilde{R}}(\mathbf{z}, \mathbf{x}, \boldsymbol{\theta}) = |\partial_{\boldsymbol{\xi}} \boldsymbol{\theta}|^{-1} \chi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) A_{\text{HPF}}(\mathbf{z}, T_s(\mathbf{z}), \mathbf{x}, \boldsymbol{\xi}). \quad (4.109)$$

Concerning the integration domain it can be observed that, for a given (\mathbf{z}, \mathbf{x}) the set $\boldsymbol{\theta}(\text{supp}(\chi))$ is contained in the halfspace $\{\boldsymbol{\theta} \in \mathbb{R}^n \setminus \{0 \mid \boldsymbol{\theta} \cdot \mathbf{n} < 0\}\}$.

(iv) While expression (4.108) defines a pseudo-differential operator of order 0, it is given in a non-standard form. It differs from a regular pseudo-differential operator, because the the amplitude $A_{\tilde{R}}(\mathbf{z}, \mathbf{x}, \boldsymbol{\theta})$ depends on $(\mathbf{z}, \mathbf{x}, \boldsymbol{\theta})$ and not only on $(\mathbf{z}, \boldsymbol{\theta})$. Another amplitude that does not depend on \mathbf{x} can be found by

$$A_{\tilde{R}}(\mathbf{z}, \mathbf{z}, \boldsymbol{\theta}) + \sum_{k=1}^n \int_0^1 D_{\theta_k} \partial_{x_k} A_{\tilde{R}}(\mathbf{z}, \mathbf{z} + \mu(\mathbf{x} - \mathbf{z}), \boldsymbol{\theta}) d\mu, \quad (4.110)$$

which is an application of formulae (4.8)-(4.10) of Treves [69]. The first term is the principal symbol of \tilde{R}_a , which has order 0. The second term in (4.110) does not contribute to the principal part, it corresponds to a pseudo-differential operator of order -1 . We will denote by $A_{\tilde{R}}(\mathbf{z}, \boldsymbol{\theta})$, with two arguments, the symbol of \tilde{R} .

To evaluate of A_{HPF} (4.100) on the diagonal one applies (4.9), the relation $\partial_t \alpha(\mathbf{z}, 0; \boldsymbol{\xi}) = -c(\mathbf{z})|\boldsymbol{\xi}|$ for the phase and the result $a(\mathbf{z}, 0; \boldsymbol{\xi}) = \frac{i}{2c(\mathbf{z})|\boldsymbol{\xi}|}$ for the amplitude. This yields

$$\begin{aligned} A_{\text{HPF}}(\mathbf{z}, T_s(\mathbf{z}), \mathbf{z}, \boldsymbol{\xi}) &= \left(1 - \frac{\mathbf{n}_s(\mathbf{z}) \cdot \boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right) p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}) \\ &= |\partial_{\boldsymbol{\xi}} \boldsymbol{\theta}| p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}). \end{aligned} \quad (4.111)$$

See also (4.106). In view of (4.109)-(4.111) conclude that for the principal symbol one has p.s. $(A_{\tilde{R}})(\mathbf{z}, \boldsymbol{\theta}) = \chi(\mathbf{z}, \mathbf{z}, \boldsymbol{\xi}) p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi})$. Note that $\boldsymbol{\eta} = \partial_y \alpha = \boldsymbol{\xi}$ holds on the diagonal, and that $\boldsymbol{\xi} = \boldsymbol{\xi}(\boldsymbol{\theta})$.

We now come back to the formal role of cutoff function χ . By requiring $\chi(\mathbf{z}, \mathbf{z}, \boldsymbol{\xi}) = 1$ on $\text{supp}(p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}))$ the cutoff function can be left out. This requirement is allowed because $\Omega(\mathbf{z})$ in the construction of χ can be chosen arbitrarily tight by narrowing the spatial support of χ around the diagonal. Therefore

$$\text{p.s.}(A_{\tilde{R}})(\mathbf{z}, \boldsymbol{\theta}) = p_a(T_s(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\theta})). \quad (4.112)$$

(v) A key step is the inclusion of both negative and positive frequencies. In section 4.2 we saw that $a(\mathbf{x}, t; \boldsymbol{\xi}) e^{i\lambda\alpha(\mathbf{x}, t; \boldsymbol{\xi})}$ and $b(\mathbf{x}, t; -\boldsymbol{\xi}) e^{i\lambda\beta(\mathbf{x}, t; -\boldsymbol{\xi})}$ have the symmetry relation that says they are complex conjugates of each other, taking into account the $-$ signs. The consequences of this property can be traced throughout the proof. We find that $\tilde{R}_b r(\mathbf{z}) = (2\pi)^{-n} \iint_{-\boldsymbol{\theta}(\text{supp}(\chi))} \overline{A_{\tilde{R}}(\mathbf{z}, -\boldsymbol{\theta})} e^{i\boldsymbol{\theta} \cdot (\mathbf{z} - \mathbf{x})} d\boldsymbol{\theta} r(\mathbf{x}) d\mathbf{x}$, and consequently, modulo a regularizing contribution,

$$\tilde{R}r(\mathbf{z}) = (2\pi)^{-n} \iint \left[A_{\tilde{R}}(\mathbf{z}, \boldsymbol{\theta}) + \overline{A_{\tilde{R}}(\mathbf{z}, -\boldsymbol{\theta})} \right] e^{i\boldsymbol{\theta} \cdot (\mathbf{z} - \mathbf{x})} d\boldsymbol{\theta} r(\mathbf{x}) d\mathbf{x}. \quad (4.113)$$

The $\boldsymbol{\theta}$ -integration is over the full space because the definition of $A_{\tilde{R}}(\mathbf{z}, \boldsymbol{\theta})$ can be smoothly extended such that it is zero outside the domain $\boldsymbol{\theta}(\text{supp}(\chi))$. In view of equation (4.112) this proves the claim. \square

Proof of theorem 6. The first step in deriving the imaging condition is to rewrite operators H , K and G (4.96). Let $w(\mathbf{x}, t)$ again be an auxiliary distribution. In this section $\hat{w}(\mathbf{x}, \omega)$ will denote its temporal Fourier transform. Because the inverse transform is $w(\mathbf{x}, t) = \frac{1}{2\pi} \int e^{i\omega t} \hat{w}(\mathbf{x}, \omega) d\omega$, one has

$$\begin{aligned} \widehat{H}w(\mathbf{x}, \omega) &= \frac{\tilde{\sigma}(\omega)}{A_s(\mathbf{x})} (i\omega)^{-\frac{n+1}{2}} [i\omega + c(\mathbf{x})\mathbf{n}_s(\mathbf{x}) \cdot \partial_{\mathbf{x}}] \hat{w}(\mathbf{x}, \omega) \\ Kw(\mathbf{x}) &= \frac{1}{2\pi} \int e^{i\omega T_s(\mathbf{x})} \hat{w}(\mathbf{x}, \omega) d\omega. \end{aligned} \quad (4.114)$$

Applied to the reverse time continued wave field $u_r(\mathbf{x}, \omega)$, equation (4.97) becomes

$$\tilde{R}r(\mathbf{x}) = \frac{1}{2\pi} \int e^{i\omega T_s(\mathbf{x})} \frac{\tilde{\sigma}(\omega)}{A_s(\mathbf{x})} (i\omega)^{-\frac{n+1}{2}} [i\omega + c(\mathbf{x})\mathbf{n}_s(\mathbf{x}) \cdot \partial_{\mathbf{x}}] \hat{u}_r(\mathbf{x}, \omega) d\omega. \quad (4.115)$$

The next step is to eliminate $T_s(\mathbf{x})$, $A_s(\mathbf{x})$ and $\mathbf{n}_s(\mathbf{x})$ by expressing them in terms of the source field explicitly. The principal term of the geometrical optics approximation of the source (4.38) is

$$\hat{g}(\mathbf{x}, \omega) = A_s(\mathbf{x})\sigma(\omega)(i\omega)^{\frac{n-3}{2}} e^{-i\omega T_s(\mathbf{x})}.$$

Function σ , introduced in (4.38), is smooth and has value 1 except for a small neighborhood of the origin where it is 0. Later we will examine the effect of the subprincipal terms of the source and the division by its amplitude. One naively derives the following identities

$$\begin{aligned} e^{i\omega T_s(\mathbf{x})} \frac{1}{A_s(\mathbf{x})} (i\omega)^{-\frac{n+1}{2}} &= \frac{\sigma(\omega)}{(i\omega)^2 \hat{g}(\mathbf{x}, \omega)} \\ c(\mathbf{x})\mathbf{n}_s(\mathbf{x}) &= \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \hat{g}(\mathbf{x}, \omega)}{-i\omega \hat{g}(\mathbf{x}, \omega)} = \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x}, \omega)}}{i\omega \overline{\hat{g}(\mathbf{x}, \omega)}}, \end{aligned} \quad (4.116)$$

in which it is used that the second equation is real-valued. Substitution of involved factors occurring in the integral (4.115) yields

$$\tilde{R}r(\mathbf{x}) = \frac{1}{2\pi} \int \frac{\tilde{\sigma}(\omega)\sigma(\omega)}{i\omega \hat{g}(\mathbf{x}, \omega)} \left[1 + \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x}, \omega)} \cdot \partial_{\mathbf{x}}}{(i\omega)^2 \hat{g}(\mathbf{x}, \omega)} \right] \hat{u}_r(\mathbf{x}, \omega) d\omega. \quad (4.117)$$

We will finally argue that the division by the source amplitude is well-defined and that the subprincipal terms in the expansion for $\hat{g}(\mathbf{x}, \omega)$ do not affect the expression for the principal symbol (4.95). The source wave field is free of caustics by assumption. The transport equation yields that, on a compact domain in spacetime, there exists a lower bound $L > 0$ for the principal amplitude, thus $|A_0(\mathbf{x}, \mathbf{x}_s, \omega)| \geq L$. Division by A_0 is therefore well-defined, and from its homogeneity and the inequality (4.37) it can be deduced that there exists a constant $C > 0$ such that $\left| \frac{A(\mathbf{x}, \mathbf{x}_s, \omega)}{A_0(\mathbf{x}, \mathbf{x}_s, \omega)} - 1 \right| \leq \frac{C}{1+|\omega|}$. For $|\omega|$ sufficiently high, division by A is therefore well-defined. We choose $1 - \Omega$ wide enough such that all $\omega \in \text{supp}(\Omega)$ are high and satisfy $\tilde{\sigma}(\omega)\sigma(\omega) = 1$. The difference between $\frac{1}{A_0}$ and $\frac{1}{A}$ is of lower order in ω . By construction it holds that $A_0(\mathbf{x}, \mathbf{x}_s, \omega) = A_s(\mathbf{x}) (i\omega)^{\frac{n-3}{2}}$ on $\text{supp}(\Omega)$. Taking (4.117) we replace $\tilde{\sigma}\sigma$ with Ω to define the imaging condition (4.93). \square

4.6 Numerical examples

In this section we give numerical examples to support our theorems. The general setup of the examples was as follows. First a model was chosen, consisting of a background medium c , a medium perturbation (contrast) $\delta c = cr$, a domain of interest and a computational domain. The latter was larger than the domain of interest and included absorbing boundaries. Data were generated by solving the inhomogeneous wave equation with velocity $c + \delta c$, and a Ricker wavelet source signature at position $\mathbf{x}_s = (0, 0)$, using an order (2,4) finite difference scheme [16]. The direct wave was eliminated. The operator (4.65) could be applied in the Fourier domain since in the examples c was constant at the surface. The backpropagated field was then computed using the finite difference method, and the same for the source field. Finally the imaging condition (4.93) was applied to obtain an approximate reconstruction of δc .

As we mentioned, only a partial reconstruction of δc is possible in realistic situations. Relation (4.52) and the wave propagation restrict the directions of ζ where inversion is possible. The frequency range present in the data also restricts the length of ζ , according to (4.52) and using that $|\xi| = c^{-1}|\omega|$. To be able to compare the original and reconstructed reflectivity we used bandlimited functions for δc , which were obtained by multiplying a plane wave with a window function. Such functions are localized in position, by the support of the window, and in wave vector by the plane-wave.

Our first example concerns a constant gradient medium with velocity function $c(x_1, x_2) = 2.0 + 0.001x_2$ with c in km/s and x_2 in meters. Our model region was the square with x_1 and x_2 between 0 and 2000 meters. The purpose was to show a successful reconstruction of velocity perturbations at different positions and with different orientations in the model. We therefore chose for δc a linear combination of three wave packets at different locations, with central wave vector well within in the inversion aperture. We included one with large dip, as one of the interesting abilities of RTM is imaging of large dips. The results of the above procedure are shown in figures 4.3 and 4.4. The reconstruction of the phase is excellent. However, the reconstructed amplitude is around 8-10 % smaller than the original amplitude. Possible explanations for this are inaccuracies related to the linearization and to a limited aperture.

Our second example concerns a bandlimited continuous reflector, illustrated in figure 4.6(a). For a continuous reflector one might expect less loss in amplitude when compared to the localized velocity perturbations. One of the strengths of RTM and wave equation migration in general is that multipathing is easily incorporated, where in our case of single source RTM, multipathing is only allowed between the reflector and the receiver point. To see this in an example we included in our background model a low velocity lens at (800, 1200) m. The background

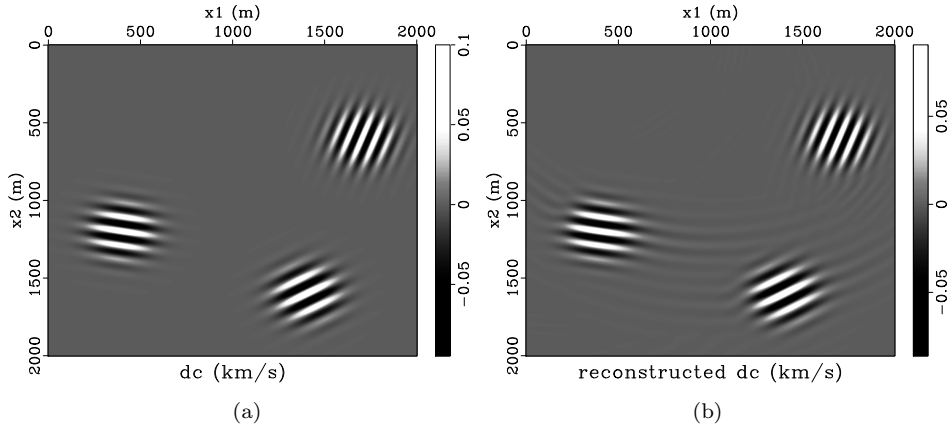


Figure 4.3: Example 1. (a) Velocity perturbation and (b) reconstructed velocity perturbation. The background medium is a gradient $c = 2.0 + 0.001x_2$, with x_2 in meters and c in km/s.

medium including some rays is plotted in figure 4.5(a), and some data are plotted in figure 4.5(b). The velocity perturbation was located at $x_2 = 1600$ m. The result of the inversion for this example is given in figure 4.6(b). The reconstruction of the phase is again excellent. The amplitude varies somewhat depending on location, being about 0-10 % too low. The smooth tapering which was applied has diminished smiles and amplitude variations, but not fully eliminated them. The multipathing leads to singularities in the inverse of the source field \hat{u}_{inc}^{-1} , around $(x_1, x_2) = (1900, 1000)$ m, which leads to the two artifacts that can be seen there.

4.7 Discussion

We have presented a comprehensive analysis of RTM-based imaging, and introduced an imaging condition involving only local (with respect to data point and image point) operators which yields a parametrix for the single scattering problem for a given point source.

We make the following observations concerning our inverse scattering procedure. (i) The symbol of the normal operator associated with a single point source contains a singularity which has been observed in the form of ‘low-frequency’ artifacts [75, 49, 25, 74, 31]. Our imaging condition yields a parametrix and naturally avoids this singularity. (ii) The square-root operator (4.65), a factor of F_M introduced in section 4.4, can be removed if dual sensor (streamer) data is used. In that case the surface-normal derivative of the wave field is measured. We note that F_M

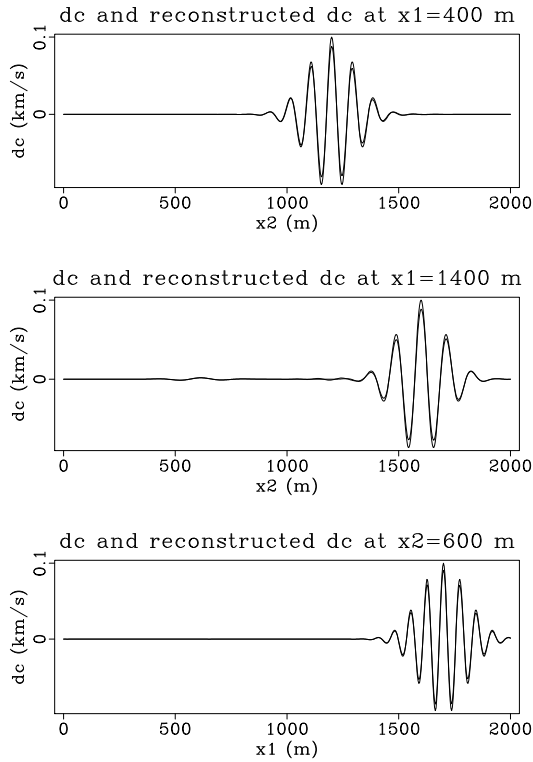


Figure 4.4: Example 1. Comparison of some traces from figure 4.3(a) and 4.3(b) at $x_1 = 400$ m, $x_1 = 1400$ m and $x_2 = 600$ m.

is available only microlocally. (iii) Division by the source field, in frequency, can lead to poor results when its amplitude is small. There are two main reasons why this can occur. First, a realistic source signature can yield very small values for particular frequencies in its amplitude spectrum. Moving averaging in frequency typically resolves this situation [32, 12, 19]. Secondly, the illumination due to propagation in a velocity model of high complexity may result in small values. Spatial averaging over small neighborhoods of the image points may then be beneficial. In this context we note that the cross-correlation imaging has been adapted by normalization with the source wave field energy at the imaging points as a proxy to inverse scattering [13, 7].

The acquisition aperture, and associated illumination, is intimately connected to the resolution operator R . This operator is pseudo-differential and the support of its symbol expresses which parts of the contrast or reflectivity can be recovered

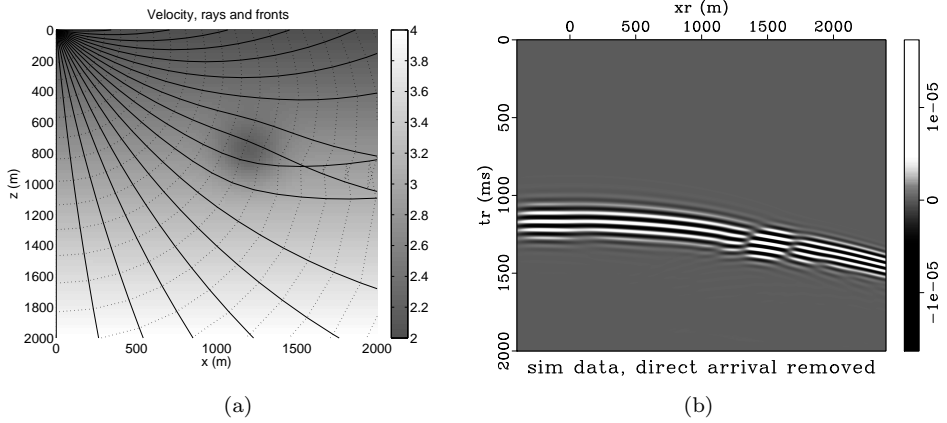


Figure 4.5: Example 2. (a) A velocity model with some rays and (b) simulated data, in which the direct arrivals are removed.

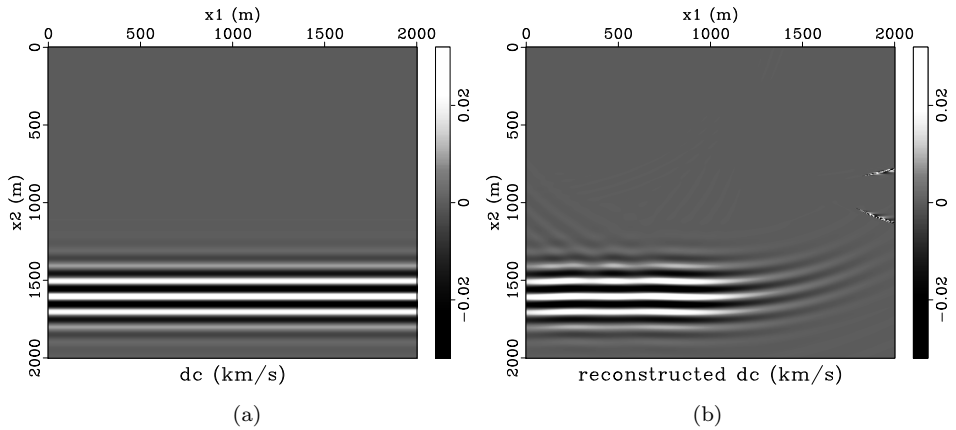


Figure 4.6: Example 2. (a) Velocity perturbation and (b) partial reconstruction of the velocity perturbation.

from the available data. Partial reconstruction is optimally formulated in terms of curvelets or wave packets. A detailed procedure, making use of the fact that the single scattering or imaging operator is associated with a canonical graph, can be found in [21]; see also [24].

We have addressed the single-source acquisition geometry, which arises naturally in RTM. One can anticipate an immediate extension of our reconstruction

to multi-source data, but a major challenge arises because the single source reconstructions are only partial. Because each of the single source images result in reconstructions at different sets of points and orientations, in general, which are not identified within the RTM algorithm, averaging must be excluded. However, techniques from microlocal analysis can be invoked to properly exploit the discrete multi-source acquisition geometry. We note that in the case of open sets of sources the generation of source caustics will be allowed.

Acknowledgement

CCS and TJPMOtR were supported by the Netherlands Organisation for Scientific Research through VIDI grant 639.032.509. MVdH was supported in part by the members of the Geo-Mathematical Imaging Group at Purdue University.

Chapter 5

Conclusion

The ultimate goal of reflection seismic imaging is to make detailed images of the Earth subsurface. Besides the position and orientation of the discontinuities, the subsurface reflecting interfaces, it is still a challenge to recover the size or amplitude of the discontinuities as much as possible.

We have investigated two methods or techniques of reflection seismic imaging in order to improve the amplitude properties. In case of the one-way wave equation, chapter 3, the amplitude refers to the amplitude of the wave solutions. In the case of our study of reverse-time migration, chapter 4, the amplitude refers to the amplitude of the reflectivity, i.e. the medium perturbation.

We used a mathematically rigorous approach and extensively used the theory of pseudo-differential and Fourier integral operators. Though the approach is formal, we support our theoretical claims with numerical simulations. Moreover, our results have practical implications for seismic imaging algorithms, which improve or correct the amplitude behaviour.

Symmetric square root operator

We have derived the one-way wave equation (3.22) and identified the steps that determine the wave amplitude. We introduced a symmetric square root operator (3.19) and a wave field normalization operator (3.23) and showed that these provide the one-way wave equation with a geometrical amplitude. The advantage of the symmetry argument lies in its general applicability to numerical methods. For instance, we proposed an adaptation (3.36) of the 60 degree Padé type finite-difference method such that it becomes symmetrical. This can be done at the expense of almost no extra cost. We have shown by wave field simulations that this significantly improves the amplitudes.

We have introduced in section 3.3 a numerical method, which we call the pseudo-spectral method. This method was used to implement the square root and normalization operator. However, it was in the first place designed to verify and illustrate the amplitude claims by wave field simulations. Though there are alternative implementations for the square root, for the normalization operator they do not exist, to our opinion. We found that the implementation is sensitive to some parameter choices, for example the cutoff angle ϕ_c used in the implementation of the normalization operator (3.33). Therefore, it could be worthwhile to devote further research to the method and numerically optimize it.

A novel imaging condition

We have studied the linearized inverse problem associated with reverse-time migration and presented a comprehensive analysis of the RTM-based imaging procedure. An inverse scattering transform (4.96) is introduced of which we proved that it yields a partial inverse, i.e. it gives a principal order reconstruction within the aperture. From the inverse transform we derived an imaging condition (4.93) of which we proved that it maps the source and receiver wave fields to a principal order reconstruction, within the aperture, of the reflectivity, the image. The imaging condition consists of the excitation time imaging condition together with a PsDO factor that works on the receiver wave field and provides a correct image amplitude. The claims are confirmed and illustrated by numerical simulations.

In many articles about the linearized inverse problem, e.g. Rakesh [54], one considers the adjoint of the scattering operator as a ‘first approach’ of the inverse scattering operator. The so called normal operator, i.e. composition of the scattering operator together with its adjoint, then is a PsDO under certain conditions and is used to derive a partial inverse. In our work we did not use the normal operator at the expense of introducing other differential and pseudo-differential operators.

The aperture of a single-source experiment is the set of subsurface positions and orientations that can be recovered from the data within the acquisition domain. And it depends on the source position. To extend our results, which assume a single source, to multi-source data, one should determine the aperture for each source in order to be able to combine the contributions of different sources. This is still a major challenge because the aperture is not identified within contemporary RTM algorithms. For this purpose the techniques of microlocal analysis can be used to exploit the discrete multi-source acquisition geometry. In this context we note that resolution operator R is intimately related with the aperture of the single-source acquisition geometry.

Bibliography

- [1] S. Alinhac and P. Gérard. *Pseudo-differential Operators and the Nash-Moser Theorem*. American Mathematical Society, 2007.
- [2] G. Bao. and W. W. Symes. Computation of pseudo-differential operators. *SIAM J. Sci. Comput.*, 17(2):416–429, 1996.
- [3] K. Baysal, D.D. Kosloff, and J.W.C. Sherwood. Reverse time migration. *Geophysics*, 48:1514–1524, 1983.
- [4] G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *J. Math. Phys.*, 26(1):99–108, 1985.
- [5] B. Biondi. Stable wide-angle fourier finite-difference downward extrapolation of 3-d wavefields. *Geophysics*, 67(03):872–882, 2002.
- [6] B. Biondi and G. Palacharla. 3-d prestack migration of common-azimuth data. *Geophysics*, 61(6):1822–1832, 1996.
- [7] B. L. Biondi. *3D seismic imaging*. Society of Exploration Geophysicists, 2006.
- [8] N. Bleistein, J. K. Cohen, and Jr. J. W. Stockwell. *Mathematics of multi-dimensional seismic imaging, migration, and inversion*. Springer-Verlag New York, Inc., 2001.
- [9] N. N. Bojarski. A survey of the near-field far-field inverse scattering inverse source integral equation. *Inst. Electr. Electron. Eng. Trans. Antennas and Propagation*, AP-30:975–979, 1982.
- [10] W. F. Chang and G. A. McMechan. Reverse-time migration of offset vertical seismic profiling data using the excitation-time imaging condition. *Geophysics*, 51(1):67–84, 1986.

- [11] C. H. Chapman. *Fundamentals of Seismic Wave Propagation*. Cambridge University Press, 2004.
- [12] S. Chattopadhyay and G. A. McMechan. Imaging conditions for prestack reverse-time migration. *Geophysics*, 73(3):S81–S89, 2008.
- [13] J. F. Claerbout. Toward a unified theory of reflector mapping. *Geophysics*, 36(3):467–481, 1971.
- [14] J. F. Claerbout. *Imaging the Earth's Interior*. Blackwell Scientific Publications, Inc., 1985.
- [15] R. Clayton. Common midpoint migration: Stanford Expl. Proj. Rep. no. 14, Stanford University, 1978.
- [16] G. C. Cohen. *Higher-order numerical methods for transient wave equations*. Scientific Computation. Springer-Verlag, Berlin, 2002. With a foreword by R. Glowinski.
- [17] J. K. Cohen and N. Bleistein. Velocity inversion procedure for acoustic waves. *Geophysics*, 44(6):1077–1087, 1979.
- [18] J. B. Conway. *A course in functional analysis, Second Edition*. Springer, 1990.
- [19] J. C. Costa, F. A. Silva Neto, M. R. M. Alcântara, J. Schleicher, and A. Novais. Obliquity-correction imaging condition for reverse time migration. *Geophysics*, 74(3):S57–S66, 2009.
- [20] M. V. de Hoop, J. H. Le Rousseau, and R. S. Wu. Generalization of the phase-screen approximation for the scattering of acoustic waves. *Wave Motion*, 31:43–70, 2000.
- [21] M. V. de Hoop, H. Smith, G. Uhlmann, and R. D. van der Hilst. Seismic imaging with the generalized Radon transform: a curvelet transform perspective. *Inverse Problems*, 25(2):025005, 21, 2009.
- [22] J. J. Duistermaat. *Fourier Integral Operators*. Birkhäuser, 1996.
- [23] L. C. Evans. *Partial differential equations*. American Mathematical Society, 1998.
- [24] P. P. Moghaddam F. J. Herrmann and C. C. Stolk. Sparsity- and continuity-promoting seismic image recovery with curvelet frames. *Appl. Comput. Harmon. Anal.*, 24(2):150–173, 2008.

-
- [25] R. F. Fletcher, P. Fowler, P. Kitchenside, and U. Albertin. Suppressing artifacts in prestack reverse time migration. In *Expanded Abstracts*, pages 2049–2052. Society of Exploration Geophysicists, 2005.
- [26] J. Gazdag. Wave-equation migration with the phase-shift method. *Geophysics*, 43(7):1342–1351, 1978.
- [27] J. Gazdag and P. Sguazzero. Migration of seismic data by phase shift plus interpolation. *Geophysics*, 49(2):124–131, 1984.
- [28] R. W. Graves and R. W. Clayton. Modeling acoustic waves with paraxial extrapolators. *Geophysics*, 55(3):306–319, 1990.
- [29] A. Grigis and J. Sjöstrand. *Microlocal Analysis for Differential Operators, An Introduction*. Cambridge University Press, 1994.
- [30] V. Guillemin. On some results of Gelfand in integral geometry. In *Pseudodifferential operators and applications*, volume 43 of *Proc. Sympos. Pure Math.*, pages 149–155. Amer. Math. Soc., Providence, RI, 1985.
- [31] A. Guitton, B. Kaelin, and B. Biondi. Least-square attenuation of reverse-time migration artifacts. *Geophysics*, 72:S19–S23, 2007.
- [32] A. Guitton, A. Valenciano, and D. Bevc. Robust imaging condition for shot-profile migration. *SEG Technical Program Expanded Abstracts*, 25(1):2519–2523, 2006.
- [33] E. Hairer, S. P. Norsett, and G. Wanner. *Solving ordinary differential equations I, Nonstiff Problems*. Springer, 1993.
- [34] E. Hairer and G. Wanner. *Solving ordinary differential equations II, Stiff and Differential-Algebraic Problems*. Springer, 1996.
- [35] L. Halpern and L. N. Trefethen. Wide-angle one-way wave equations. *J. Acoust. Soc. Am.*, 84(4):1397–1404, 1988.
- [36] L. Hörmander. *The Analysis of Linear Partial Differential Operatros I*. Springer-Verlag, 1990.
- [37] L. Hörmander. *The Analysis of Linear Partial Differential Operatros III*. Springer-Verlag, 1994.
- [38] S. Jin, C. C. Mosher, and R. S. Wu. Offset-domain pseudoscreen prestack depth migration. *Geophysics*, 67(6):1895–1902, 2002.
- [39] S. Jin, R. S. Wu, and C. Peng. Seismic depth migration with pseudo-screen propagator. *Computational Geosciences*, 3:321–335, 1999.

- [40] B. Kaelin and A. Guitton. Imaging condition for reverse time migration. *SEG Technical Program Expanded Abstracts*, 25(1):2594–2598, 2006.
- [41] D. Kincaid and W. Cheney. *Numerical Analysis: Mathematics of Scientific Computing*. Brooks/Cole, Thomson Learning, Inc., 2002.
- [42] D. Kiyashchenko, R. E. Plessix, B. Kashtan, and V. Troyan. Improved amplitude multi-one-way modeling method. *Wave Motion*, 43:99–115, 2005.
- [43] D. Kiyashchenko, R. E. Plessix, B. Kashtan, and V. Troyan. A modified imaging principle for true-amplitude wave-equation migration. *Geophys. J. Int.*, 168(3):1093–1104, 2007.
- [44] A. P. E. Ten Kroode, D. J. Smit, and A. R. Verdel. A microlocal analysis of migration. *Wave Motion*, 28:149–172, 1998.
- [45] M. P. Lamoureux and G. F. Margrave. An introduction to numerical methods of pseudodifferential operators. *Pseudo-Differential Operators*, 1949:79–133, 2008.
- [46] L. D. Landau, E. M. Lifshitz, A. M. Kosevich, and L. P. Pitaevskii. *Theory of Elasticity, 3rd edition*. Butterworth-Heinemann, Elsevier, 1986.
- [47] G. F. Margrave and R. J. Ferguson. Wavefield extrapolation by nonstationary phase shift. *Geophysics*, 64(4):1067–1078, 1999.
- [48] G. A. McMechan. Migration by extrapolation of time-dependent boundary values. *Geophys. Prosp.*, 31:413–420, 1983.
- [49] W. A. Mulder and R. E. Plessix. A comparison between one-way and two-way wave-equation migration. *Geophysics*, 69:1491–1504, 2004.
- [50] C. J. Nolan and W. W. Symes. Global solution of a linearized inverse problem for the wave equation. *Comm. PDE*, 22(5):127–149, 1997.
- [51] T. J. P. M. Op ’t Root and C. C. Stolk. One-way wave propagation with amplitude based on pseudo-differential operators. *Wave Motion*, 47(2):67 – 84, 2010.
- [52] R. E. Plessix and W. A. Mulder. Frequency-domain finite-difference amplitude-preserving migration. *Geophys. J. Int.*, 157:975–987, 2004.
- [53] A. M. Popovici. Prestack migration by split-step dsr. *Geophysics*, 61(5):1412–1416, 1996.
- [54] Rakesh. A linearised inverse problem for the wave equation. *Comm. PDE*, 13(5):573–601, 1988.

-
- [55] D. Ristow and T. Ruhl. Fourier finite-difference migration. *Geophysics*, 59(12):1882–1893, 1994.
- [56] W. A. Schneider. Integral formulation for migration in two and three dimensions. *Geophysics*, 43:49–76, 1978.
- [57] P. S. Schultz and J. W. C. Sherwood. Depth migration before stack. *Geophysics*, 45(3):376–393, 1980.
- [58] R. E. Sheriff and L. P. Geldart. *Exploration Seismology*, 2nd ed. Cambridge University Press, 1995.
- [59] S. Stein and M. Wyssession. *An introduction to seismology, earthquakes, and earth structure*. Wiley-Blackwell, 2003.
- [60] C. C. Stolk. A pseudodifferential equation with damping for one-way wave propagation in inhomogeneous acoustic media. *Wave Motion*, 40:111–121, 2004.
- [61] C. C. Stolk and M. V. De Hoop. Microlocal analysis of seismic inverse scattering in anisotropic elastic media. *Communications on Pure and Applied Mathematics*, 55:261–301, 2002.
- [62] C. C. Stolk and M. V. De Hoop. Modeling of seismic data in the downward continuation approach. *SIAM Journal on Applied Mathematics*, 65:1388–1406, 2005.
- [63] C. C. Stolk and M. V. De Hoop. Seismic inverse scattering in the downward continuation approach. *Wave Motion*, 43:579–598, 2006.
- [64] R. S. Strichartz. *A guide to distribution theory and Fourier transforms*. World Scientific, 2003.
- [65] R. Sun and G. A. McMechan. Scalar reverse-time depth migration of prestack elastic seismic data. *Geophysics*, 66:1519–1527, 2001.
- [66] W. W. Symes. Topical review: The seismic reflection inverse problem. *Inverse Problems*, 25(12):123008, dec 2009.
- [67] M. E. Taylor. *Pseudodifferential Operators*. Princeton University Press, 1981.
- [68] M. E. Taylor. *Pseudodifferential Operators and Nonlinear PDE*. Birkhuser, 1991.
- [69] F. Trèves. *Introduction to Pseudodifferential and Fourier Integral Operators, Volume I*. Plenum Press, New York, 1980.

- [70] F. Trèves. *Introduction to Pseudodifferential and Fourier Integral Operators, Volume II*. Plenum Press, New York, 1980.
- [71] D. Whitmore. Iterative depth migration by backward time propagation. In *Expanded Abstracts*, pages 382–385. Society of Exploration Geophysicists, 1983.
- [72] A. J. Wild, R. W. Hobbs, and L. Frenje. Modelling complex media: an introduction to the phase-screen method. *Physics of the Earth and Planetary Interiors*, 120:219–225, 2000.
- [73] M. W. Wong. *An introduction to pseudo-differential operators*. World Scientific, 1999.
- [74] X. B. Xie and R. S. Wu. A depth migration method based on the full-wave reverse-time calculation and local one-way propagation. *SEG Technical Program Expanded Abstracts*, 25(1):2333–2337, 2006.
- [75] K. Yoon, K. J. Marfurt, and W. Starr. Challenges in reverse-time migration. In *Expanded Abstracts*, pages 1057–1060. Society of Exploration Geophysicists, 2004.
- [76] Y. Zhang, G. Zhang, and N. Bleistein. True amplitude wave equation migration arising from true amplitude one-way wave equations. *Inverse Problems*, 19:1113–1138, 2003.
- [77] Y. Zhang, G. Zhang, and N. Bleistein. Theory of true-amplitude one-way wave equations and true-amplitude common-shot migration. *Geophysics*, 70(4):E1–E10, 2005.

Acknowledgments

The research presented in this thesis has been carried out over the past four and a half year in the Applied Analysis and Mathematical Physics (AAMP) group at the University of Twente. I want to express my gratitude to a number of people.

I want to thank Brenny van Groesen and Chris Stolk for the opportunity to do a PhD and for their attention for me and my work. The advice and support of Stephan van Gils was much appreciated. I want to thank Hil Meijer for the contributions to the ‘proofs behind the scene’. And Sander Rhebergen was more than helpful in giving practical support. I would like to thank Alyona Ivanova for the ergonomic keyboard that I got from her. Finally, I want to thank Sjoukje Schoustra for reading and correcting these acknowledgements and other parts of my thesis where necessary.

Summary

The goal of reflection seismic imaging is making images of the Earth subsurface using surface measurements of reflected seismic waves. Besides the position and orientation of subsurface reflecting interfaces it is a challenge to recover the size or amplitude of the discontinuities. We investigate two methods or techniques of reflection seismic imaging in order to improve the amplitude. First we study the *one-way wave equation* with attention to the amplitude of the waves. In our study of *reverse-time migration* the amplitude refers to the amplitude of the image. Though the approach in the thesis is formal, our results have practical implications for seismic imaging algorithms, which improve or correct the amplitudes.

The one-way wave equation is a 1st-order equation that describes wave propagation in a predetermined direction. We derive the equation and identify the steps that determine the wave amplitude. We introduce a *symmetric square root operator* and a wave field normalization operator and show that they provide the correct amplitude. The idea to use a symmetric square root is generally applicable. Our amplitude claims are numerically verified. The one-way wave equation is an application of *pseudo-differential operators*.

Reverse-time migration (RTM) is an imaging method that uses simulations of the source and receiver wave fields through a slowly varying estimate of the subsurface medium. The receiver wave is an in reverse time continued field that matches the measurements. An *imaging condition* transforms the fields into an image of the small scale medium contrast.

We investigate the linearized inverse problem and use the RTM procedure to reconstruct the perturbation of the medium. We model the scattering event by the scattering operator, which maps the medium perturbation to the scattered wave. We propose an approximate inverse of the scattering operator, derive a novel imaging condition and show that it yields a reconstruction of the perturbation with correct amplitude. The study extensively uses the theory of *Fourier integral operators* (FIO). Besides that we globally characterize the scattering operator as a FIO, we also obtain a local expression to calculate the amplitude explicitly. The claims are confirmed and illustrated by numerical simulations.

Samenvatting

Reflection seismic imaging is een beeldvormende techniek die het inwendige van de aarde afbeeldt op basis van aan het oppervlak gemeten gereflecteerde seismische golven. Naast de diepte en oriëntatie van ondergrondse reflecterende grenslagen is het een uitdaging om de grootte of amplitude van de discontinuïteiten te reconstrueren. Wij onderzoeken twee methoden of technieken van reflection seismic imaging met als doel de amplitude te verbeteren. We bestuderen de *one-way wave equation* en richten ons daarbij op de amplitude van de golven. In onze studie van *reverse-time migration* staat de amplitude van het beeld centraal. Hoewel het proefschrift een vrij formele benadering heeft, kunnen onze resultaten toegepast worden in de praktijk van de seismiek.

De one-way wave equation is een 1^{ste}-orde vergelijking die golfvoortplanting in een vooraf bepaalde richting beschrijft. Wij leiden de vergelijking af en identificeren de stappen die de golfamplitude bepalen. We introduceren een *symmetric square root operator* en een golfveld normalisatie operator en laten zien dat de amplitude correct beschreven wordt. Het idee om een symmetrische wortel te gebruiken, is algemeen toepasbaar. De amplitude claims zijn numeriek geverifieerd. De one-way wave equation is een toepassing van *pseudo-differentiaal operatoren*.

Reverse-time migration (RTM) maakt gebruik van simulaties van bron- en ontvangergolfvelden door een glad model van het medium. De ontvangergolf is een in omgekeerde tijd voortgezet veld dat de metingen voorspelt. Een *imaging condition* transformeert de velden tot een beeld van het kleine-schaal mediumcontrast.

Wij onderzoeken het gelineariseerde inverse probleem en gebruiken de RTM procedure om de kleine-schaal mediumverstoring te reconstrueren. We modelleren het scatteringsproces met de scatteringsoperator, de afbeelding van de mediumverstoring naar de gescatterde golf. We formuleren een benaderende inverse van de scatteringsoperator, introduceren een nieuwe imaging condition en laten zien dat deze de mediumverstoring met juiste amplitude reconstrueert. De studie maakt intensief gebruik van de theorie van de *Fourier integraal operatoren* (FIO). Naast het karakteriseren van de scatteringsoperator als een globale FIO, formuleren we een lokale uitdrukking voor de berekening van de amplitude. Het resultaat is bevestigd en geïllustreerd door numerieke simulaties.

About the author

Tim Op 't Root was born in 1976 in Nederweert, the Netherlands. He studied Applied Physics at the Eindhoven University of Technology. His master thesis project was carried out in 2002 at Océ Technologies in Venlo. The project involved the theory and practice of the ultrasonic rheometer, an experimental instrument to dynamically measure the rheological properties of solidifying ink. After completing his master he combined a part time job, in research and development of wheel suspension testers at Altesco, with studying Mathematics at the Radboud University in Nijmegen. In between he has done two technical programming assignments. One of them involved the calculation of strength and stiffness of concrete floors for Rootsoft. Tim started his PhD at the University of Twente in 2006. He presented a part of this thesis at the *Waves* conference in Pau, France, in 2009. In his spare time he designs, builds and fine-tunes electronic amplifiers for audio. In addition to the technical and rather mental activities, he enthusiastically started running.