Available online at www.sciencedirect.com

SciVerse ScienceDirect

J. Math. Pures Appl. 98 (2012) 211–238

JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES

www.elsevier.com/locate/matpur

Linearized inverse scattering based on seismic reverse time migration ☆

Tim J.P.M. Op 't Root [a], Christiaan C. Stolk [b,∗], Maarten V. de Hoop [c]

[a] Department of Applied Mathematics, University of Twente, Drienerlolaan 5, 7522 NB, Enschede, Netherlands
[b] Korteweg–de Vries Institute for Mathematics, University of Amsterdam, PO Box 94248, 1090 GE Amsterdam, Netherlands
[c] Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47907-2067, United States

Received 9 February 2011

Available online 28 February 2012

Abstract

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 pseudodifferential 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.

Résumé

Dans cet article, on étudie le problème inverse linéarisé associé à l'imagerie sismique par réflexion. On propose une transformation de diffusion inverse dérivée de la migration à temps inverse. On démontre que, sous certaines conditions, cette transformation donne un inverse microlocal partiel. De plus, le résultat est verifié par des simulations numériques.

Keywords: Inverse scattering; Microlocal analysis; Reflection seismology

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 [1–3]. 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 background, called the velocity model, which is also unknown in general. However, in the inverse scattering

☆ First presented at the Conference on Applied Inverse Problems, June 28, 2007.

∗ Corresponding author.
E-mail addresses: t.j.p.m.optroot@math.utwente.nl (T.J.P.M. Op 't Root), C.C.Stolk@uva.nl (C.C. Stolk).

0021-7824/$ – see front matter
http://dx.doi.org/10.1016/j.matpur.2012.02.009

setting considered here, the background model is assumed to be known. The linearization defines a map from the model contrast (that is, the perturbation with respect to the background) to the data, that consists of the restriction to the acquisition set of the scattered field (that is, the perturbation of the wave field).

There are different types of seismic imaging methods. 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 [4] or generalized Radon transform (GRT) inversion, and reverse time migration (RTM) [5–9]; the second category comprises the downward continuation approach [10,1,11–13] possibly applied in curvilinear coordinates. For Kirchhoff/GRT methods sufficient conditions on the background medium have been derived under which the methods reconstruct the singularities in the contrast [14–18]. They are characterized as inverse scattering methods. The analysis pertaining to inverse scattering in the second category can be found in Stolk and De Hoop [19,20].

The subject of the present paper is an RTM-based method for inverse scattering. We demonstrate that a modified RTM method reconstructs the singularities in the contrast. 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 from one-way wave propagation.

The main condition on the background is the Bolker condition [21]. (In reflection seismology this condition is sometimes referred to as traveltime injectivity condition [22].) 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 [22]. 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 parts of the wave field using pseudodifferential cutoffs.

As said, we consider a modified RTM method. An RTM procedure consists of three parts: The modeling of the source wave propagation in forward time, the modeling of the receiver or reflected wave propagation in reverse time, and the application of the so-called imaging condition [1,2]. 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 [23] imaging principle: Reflectors exist in those points in the subsurface where the source and receiver wave fields both have a significant 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 cross-correlation imaging condition, in which the product of the fields is integrated over time. Later other variants have been proposed, see e.g. [24–26]. The authors of [26] use the spatial derivatives of the fields, similarly to what we find in this work.

In this paper we consider a modified ratio imaging condition that involves time derivatives of the fields and their spatial gradients. It contains some elements of the integral formulation of Schneider [27] and the inverse scattering integral equation of Bojarski [28]. The ratio imaging condition, albeit a new variant, is hence finally provided with a mathematical proof. The result is summarized in the next section and formulated precisely in Theorem 4. Moreover, we address the relation with RTM "artifacts" [29–33], see Section 8.

The outline of the paper is as follows. In the next section, we present our inverse scattering transform and state the main result. In Section 3, we discuss various aspects of the parametrix construction of the wave equation; we start from the WKB approximation with plane-wave initial values. The (forward) scattering problem is analyzed in Section 4. 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 5 we study the revert operator, which describes the backpropagation of the receiver field. The relation with the continued scattered field is established. We develop the inverse scattering in Section 6. 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 7 we show numerical experiments. We end the paper with a short discussion.

2. Statement of the main result

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

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

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 [34]. We use the Fourier transform: $\mathcal{F}u(\boldsymbol{\xi}, \omega) = \iint \mathrm{e}^{-\mathrm{i}(\boldsymbol{\xi}\cdot\mathbf{x}+\omega t)}u(\mathbf{x}, t)\,d\mathbf{x}\,dt$.

We consider a Dirac source, which generates the source wave, that is, the fundamental solution of the wave equation,

$$\begin{aligned}\left[c(\mathbf{x})^{-2}\partial_t^2 - \Delta\right]g(\mathbf{x}, t) &= \delta(\mathbf{x} - \mathbf{x}_\mathrm{s})\delta(t),\\ g(\mathbf{x}, 0) = 0, \qquad \partial_t g(\mathbf{x}, 0) &= 0.\end{aligned} \tag{2}$$

Here, c is a smooth function. We impose the assumption:

the propagation of singularities by the source wave does not exhibit multipathing (*SME*). (3)

We introduce coordinates, $\mathbf{x} = (x_1, \dots, x_n)$, such that x_n signifies depth. The acquisition surface, where the source is located and the observations are made, is given by $x_n = 0$. 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

$$\operatorname{supp}(r) \subset D \quad \text{for a compact } D \subset \mathbb{R}^{n-1} \times [\epsilon, \infty) \text{ and some } \epsilon > 0. \tag{4}$$

Because the source is located at the surface, the reflectivity is zero in a neighborhood of the source. Following the Born approximation, the scattering problem is obtained by linearization of (2) with $(1 + r(\mathbf{x}))c(\mathbf{x})$ as the velocity. We multiply (2) with $c(\mathbf{x})^2$ and find

$$\begin{aligned}\left[\partial_t^2 - c(\mathbf{x})^2\Delta\right]u(\mathbf{x}, t) &= r(\mathbf{x})2\partial_t^2 g(\mathbf{x}, t),\\ u(\mathbf{x}, 0) = 0, \qquad \partial_t u(\mathbf{x}, 0) &= 0.\end{aligned} \tag{5}$$

The *scattered wave field* $u(\mathbf{x}, t)$ is defined as the solution of this problem. 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 operator $\mathcal{F}_M$ which maps r to $T_M u$ models the data, d say.

We introduce the *reverse time continued field*, u_r, as the anticausal solution to

$$\left[c(\mathbf{x})^{-2}\partial_t^2 - \Delta\right]u_\mathrm{r}(\mathbf{x}, t) = \delta(x_n)F_M d\left(\mathbf{x}', t\right), \tag{6}$$

here F_M consists of a composition of appropriately chosen pseudodifferential cutoffs and the operator

$$-2\mathrm{i}D_t c^{-1}\sqrt{1 - c^2 D_t^{-2} D_{\mathbf{x}'}^2}, \quad D_t = \mathrm{i}^{-1}\partial_t,\ D_{\mathbf{x}'} = \mathrm{i}^{-1}\partial_{\mathbf{x}'}. \tag{7}$$

We define the inverse scattering transform, $\mathcal{H}_M$, as

$$(\mathcal{H}_M d)(\mathbf{x}) = \frac{1}{2\pi}\int \frac{\Omega(\omega)}{\mathrm{i}\omega|\hat{g}(\mathbf{x}, \omega)|^2}\left(\overline{\hat{g}(\mathbf{x}, \omega)}\hat{u}_\mathrm{r}(\mathbf{x}, \omega) - \frac{c(\mathbf{x})^2}{\omega^2}\,\partial_\mathbf{x}\overline{\hat{g}(\mathbf{x}, \omega)} \cdot \partial_\mathbf{x}\hat{u}_\mathrm{r}(\mathbf{x}, \omega)\right)d\omega. \tag{8}$$

Here, ˆ denotes the Fourier transform with respect to time, and $\Omega(\omega)$ is a smooth function, valued 0 on a bounded neighborhood of the origin, and 1 outside a slightly larger neighborhood. Theorem 4, the main result of this paper, essentially states that with assumption SME, $\mathcal{H}_M$ is microlocally the inverse of $\mathcal{F}_M$.

3. Asymptotic solutions of the 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 Section 3.1 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 Sections 3.2 and 3.3. The asymptotic solution is finally written as an FIO in Section 3.4. Section 3.5 presents the decoupling of the wave equation and general solution operators. Section 3.6 deals with the source field problem of RTM.

3.1. WKB approximation with plane-wave initial values

Instead of solving (1) directly, we solve for $c^{-1}u$, and consider the equivalent wave equation,

$$\left[\partial_t^2 - c\Delta c\right]\left(c^{-1}u\right) = 0. \tag{9}$$

In the later analysis it will be advantageous that $c\Delta c$ is a symmetric operator. We invoke the WKB ansatz,

$$c^{-1}u(\mathbf{x},t) = a(\mathbf{x},t)\mathrm{e}^{\mathrm{i}\lambda\alpha(\mathbf{x},t)}. \tag{10}$$

A straightforward calculation yields

$$\begin{aligned}\mathrm{e}^{-\mathrm{i}\lambda\alpha}\left[\partial_t^2 - c\Delta c\right]a\mathrm{e}^{\mathrm{i}\lambda\alpha} &= -\lambda^2 a\left[(\partial_t\alpha)^2 - c^2|\nabla\alpha|^2\right] \\ &\quad + \mathrm{i}\lambda\left[2(\partial_t a)\partial_t\alpha + a\partial_t^2\alpha - 2c\nabla(ca)\cdot\nabla\alpha - c^2 a\Delta\alpha\right] + \partial_t^2 a - c\Delta(ca).\end{aligned} \tag{11}$$

An approximate solution of the form (10) is obtained by requiring first that the term $O(\lambda^2)$ vanishes, resulting in an eikonal equation for α, and secondly that the term $O(\lambda)$ also vanishes, resulting in a transport equation for a. We will give these equations momentarily, and comment below on the vanishing of terms $O(\lambda^j)$ for $j \leqslant 0$.

We solve (9) with plane-wave initial values:

$$u(\mathbf{x},0) = 0, \qquad c(\mathbf{x})^{-1}\partial_t u(\mathbf{x},0) = \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}}. \tag{12}$$

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

$$c^{-1}u(\mathbf{x},t) = a(\mathbf{x},t;\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\alpha(\mathbf{x},t;\boldsymbol{\xi})} + b(\mathbf{x},t;\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\beta(\mathbf{x},t;\boldsymbol{\xi})}. \tag{13}$$

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

$$\partial_t\alpha + c|\nabla\alpha| = 0 \quad \text{and} \quad \partial_t\beta - c|\nabla\beta| = 0. \tag{14}$$

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

$$\partial_t\left(a^2\partial_t\alpha\right) - \nabla\cdot\left(a^2c^2\nabla\alpha\right) = 0 \quad \text{and} \quad \partial_t\left(b^2\partial_t\beta\right) - \nabla\cdot\left(b^2c^2\nabla\beta\right) = 0. \tag{15}$$

The WKB ansatz (13) can be inserted into the initial conditions (12). This straightforwardly yields initial conditions for α, β:

$$\alpha(\mathbf{x},0;\boldsymbol{\xi}) = \beta(\mathbf{x},0;\boldsymbol{\xi}) = \boldsymbol{\xi}\cdot\mathbf{x}. \tag{16}$$

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

$$\begin{pmatrix} 1 & 1 \\ -\mathrm{i}c(\mathbf{x})|\boldsymbol{\xi}| & \mathrm{i}c(\mathbf{x})|\boldsymbol{\xi}| \end{pmatrix}\begin{pmatrix} a(\mathbf{x},0;\boldsymbol{\xi}) \\ b(\mathbf{x},0;\boldsymbol{\xi}) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The two terms in (13) are not independent. The initial value problem for α can be transformed into the initial value problem for β by replacing $\boldsymbol{\xi}$ with $-\boldsymbol{\xi}$ and setting $\beta(\mathbf{x},t;\boldsymbol{\xi}) = -\alpha(\mathbf{x},t;-\boldsymbol{\xi})$. Further analysis shows that $b(\mathbf{x},t;\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\beta(\mathbf{x},t;\boldsymbol{\xi})}$ in (13) is in fact the complex conjugate of $a(\mathbf{x},t;-\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\alpha(\mathbf{x},t;-\boldsymbol{\xi})}$.

3.2. The phase function on characteristics

The method of characteristics [35, Section 3.2] will be used to solve the eikonal and transport equations, as usual. We first solve the initial value problem for $\alpha(\mathbf{y}, t; \boldsymbol{\xi})$, cf. (14) and (16). The same procedure can be applied to β.

The characteristic equations are formulated in terms of $(\mathbf{y}, t)$, $(\mathbf{p}, \omega)$ associated with $(\nabla\alpha, \partial_t\alpha)$, and a variable q associated with α. The eikonal equation is hence given by

$$F(\mathbf{y}, t, \nabla\alpha, \partial_t\alpha, \alpha) = 0, \qquad F(\mathbf{y}, t, \mathbf{p}, \omega, q) = \omega + c(\mathbf{y})|\mathbf{p}|. \tag{17}$$

The characteristic equations are then

$$\frac{d}{ds}\begin{pmatrix} \mathbf{y} \\ t \end{pmatrix} = \begin{pmatrix} \frac{c\mathbf{p}}{|\mathbf{p}|} \\ 1 \end{pmatrix}, \qquad \frac{d}{ds}\begin{pmatrix} \mathbf{p} \\ \omega \end{pmatrix} = \begin{pmatrix} -(\nabla c)|\mathbf{p}| \\ 0 \end{pmatrix}, \qquad \frac{dq}{ds} = 0. \tag{18}$$

The only nontrivial equations are those for $\mathbf{y}$ and $\mathbf{p}$. By $(\mathbf{y}(\mathbf{x}, t; \boldsymbol{\xi}), \mathbf{p}(\mathbf{x}, t; \boldsymbol{\xi}))$ we denote a solution with $(\mathbf{y}(0), \mathbf{p}(0)) = (\mathbf{x}, \boldsymbol{\xi})$.

When α is a solution to (14), (16) on some open set $U \subset \mathbb{R}^{n+1}$, and $(\mathbf{y}(\cdot), t(\cdot))$ is a solution to the first equation of (18), where $(\mathbf{p}(\cdot), \omega(\cdot), q(\cdot)) = (\nabla_{\mathbf{y}}\alpha(\mathbf{y}(\cdot), t(\cdot)), \partial_t\alpha(\mathbf{y}(\cdot), t(\cdot)), \alpha(\mathbf{y}(\cdot), t(\cdot)))$, then $(\mathbf{p}(\cdot), \omega(\cdot), q(\cdot))$ solve the other equations of (18), and in particular $\alpha(\mathbf{y}(\mathbf{x}, s; \boldsymbol{\xi}), s; \boldsymbol{\xi}) = \alpha(\mathbf{y}(\mathbf{x}, 0; \boldsymbol{\xi}), 0; \boldsymbol{\xi})$. Differentiating this identity, and using the identity $(\partial\alpha/\partial\mathbf{y}) \cdot (\partial\mathbf{y}/\partial\boldsymbol{\xi}) = 0$, which is a consequence of the linearization of (18), it follows that

$$\text{if } \mathbf{y} = \mathbf{y}(\mathbf{x}, t; \boldsymbol{\xi}) \quad \text{then } \partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t; \boldsymbol{\xi}) = \mathbf{x}. \tag{19}$$

To verify the local existence of solutions of (14), (16), one must derive the initial conditions for (18) from (14) and (16) for each point $\mathbf{y}$, and verify that these initial conditions are noncharacteristic, i.e. $\partial_\omega F \neq 0$. The latter is trivially the case. It follows therefore from [35] that solutions exist up to some finite time locally, when $\partial_{\mathbf{y}}\mathbf{x}$ becomes singular.

To examine the $\boldsymbol{\xi}$-dependence of the constructed solution α, we note that the initial conditions for (18) depend in a smooth fashion on $\boldsymbol{\xi}$. Consequently, so does α. Furthermore, a short calculation shows that the function $\alpha(\mathbf{y}, t; \boldsymbol{\xi})$ is positive homogeneous with respect to $\boldsymbol{\xi}$ of degree one.

3.3. The amplitude function

In this subsection, we solve for the amplitude in terms of a Jacobian of the flow of the rays. The result in Eqs. (23) and (24) is a manifestation of the energy conservation property. The first step is to carefully write Eq. (15) into the form

$$0 = \left(\partial_t - \frac{c^2\nabla\alpha}{\partial_t\alpha} \cdot \nabla - \left(\nabla \cdot \frac{c^2\nabla\alpha}{\partial_t\alpha}\right)\right)a^2 = \big(\partial_t + \mathbf{v} \cdot \nabla + (\nabla \cdot \mathbf{v})\big)a^2, \tag{20}$$

where we define $\mathbf{v} = -\frac{c^2\nabla\alpha}{\partial_t\alpha}$. We used that $(\partial_t + \mathbf{v} \cdot \nabla)\partial_t\alpha = 0$, i.e. the frequency is constant on a ray. The field $\mathbf{v}$ is associated with the rays, which satisfy

$$\frac{d\mathbf{y}}{dt}(t; \mathbf{x}) = \mathbf{v}\big(\mathbf{y}(t; \mathbf{x}), t\big). \tag{21}$$

We have $\frac{d}{dt}\frac{\partial\mathbf{y}}{\partial\mathbf{x}} = \frac{\partial\mathbf{v}}{\partial\mathbf{y}}\frac{\partial\mathbf{y}}{\partial\mathbf{x}}$. The derivative $\frac{d}{dt}|\frac{\partial\mathbf{y}}{\partial\mathbf{x}}|$ is hence related to $\nabla \cdot \mathbf{v}$ as

$$\frac{d}{dt}\det\left(\frac{\partial\mathbf{y}}{\partial\mathbf{x}}\right) = \det\left(\frac{\partial\mathbf{y}}{\partial\mathbf{x}}\right)\operatorname{tr}\left(\left(\frac{\partial\mathbf{y}}{\partial\mathbf{x}}\right)^{-1}\frac{\partial}{\partial t}\frac{\partial\mathbf{y}}{\partial\mathbf{x}}\right) = (\nabla \cdot \mathbf{v})\det\left(\frac{\partial\mathbf{y}}{\partial\mathbf{x}}\right). \tag{22}$$

This implies that

$$\det(\partial_{\mathbf{x}}\mathbf{y})a^2 \text{ is constant along the ray.} \tag{23}$$

Indeed, (23) is easily established by computing the derivative $\frac{d}{dt}[\det(\partial_{\mathbf{x}}\mathbf{y}(t; \mathbf{x}))a(\mathbf{y}(t; \mathbf{x}), t)^2]$ and using (20). From (23) it follows that $a(\mathbf{y}(t; \mathbf{x}), t; \cdot) = \sqrt{\det(\partial_{\mathbf{x}}\mathbf{y}(t; \mathbf{x})^{-1})}a(\mathbf{x}, 0; \cdot)$. Inserting the $\boldsymbol{\xi}$-dependence back into the notation, and using that the map $\mathbf{x} \mapsto \mathbf{y}(\mathbf{x}, t; \boldsymbol{\xi})$ is invertible results in

$$a(\mathbf{y}, t; \boldsymbol{\xi}) = \frac{\mathrm{i}}{2c(\mathbf{x}(\mathbf{y}, t; \boldsymbol{\xi}))|\boldsymbol{\xi}|}\sqrt{\det\big(\partial_{\mathbf{y}}\mathbf{x}(\mathbf{y}, t; \boldsymbol{\xi})\big)}. \tag{24}$$

3.4. Solution operator as an FIO

In this subsection we consider more general initial values than (12) by considering linear combinations of the terms in (13). This results in an approximate solution operator in the form of a Fourier integral operator (FIO) [36–39], and we will review some of its properties. Our solutions so far involve only the highest order WKB terms and are limited to some small but finite time.

We consider the original wave equation (1) with $f = 0$ and the initial conditions

$$u(\mathbf{x}, 0) = 0, \qquad \partial_t u(\mathbf{x}, 0) = h_2(\mathbf{x}). \tag{25}$$

Following (13), its WKB solution for time $t \in I$, which we will denote for the moment by $S_{12}(t)h_2(\mathbf{y})$ is given by a sum of two terms $S_{12}(t)h_2(\mathbf{y}) = c(\mathbf{y})(S_{a2}(t)h_2(\mathbf{y}) + S_{b2}(t)h_2(\mathbf{y}))$, with

$$S_{a2}(t)h_2(\mathbf{y}) = \frac{1}{(2\pi)^n} \iint e^{i\alpha(\mathbf{y},t;\boldsymbol{\xi}) - i\boldsymbol{\xi}\cdot\mathbf{x}} \frac{a(\mathbf{y}, t; \boldsymbol{\xi})}{c(\mathbf{x})} h_2(\mathbf{x})\, d\mathbf{x}\, d\boldsymbol{\xi}. \tag{26}$$

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 Section 3.1 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 an FIO, we will take a closer look at its phase function, i.e.,

$$\varphi(\mathbf{y}, t, \mathbf{x}, \boldsymbol{\xi}) = \alpha(\mathbf{y}, t; \boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \mathbf{x}, \tag{27}$$

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

$$\Gamma_t = \bigl\{(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in Y \times X \times \mathbb{R}^n \setminus \{0\} \bigm| \mathbf{x} = \partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t; \boldsymbol{\xi})\bigr\}. \tag{28}$$

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

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

needs to have maximal rank on Γ_t, which is obviously the case [38, Chapter VI, (4.22)]. The stationary point set Γ_t is hence a $2n$-dimensional manifold with coordinates $(\mathbf{y}, \boldsymbol{\xi})$.

The stationary point set can be understood in terms of the bicharacterstics. Definition (28) allows us to express $\mathbf{x}$ on Γ_t as a function $\mathbf{x}_\Gamma(\mathbf{y}, t, \boldsymbol{\xi}) = \partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t; \boldsymbol{\xi})$. Eq. (19) 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,

$$\Pi_t = \bigl\{\bigl((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\xi})\bigr) \in T^*Y \setminus 0 \times T^*X \setminus 0 \bigm| \mathbf{x} = \mathbf{x}_\Gamma(\mathbf{y}, t, \boldsymbol{\xi}),\ \boldsymbol{\eta} = \partial_{\mathbf{y}}\alpha(\mathbf{y}, t; \boldsymbol{\xi})\bigr\}. \tag{29}$$

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, $\Phi_t : T^*X \setminus 0 \to T^*Y \setminus 0$ say. 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 $\mathrm{WF}(v)$ of 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 [36,40,37,38]. From the above description of Π_t it follows that

$$\mathrm{WF}\bigl(S_{a2}(t)v\bigr) \subset \Phi_t\bigl(\mathrm{WF}(v)\bigr). \tag{30}$$

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 [38, p. 334].

So far the highest order WKB approximation was used. The notion of symbol classes for a, b is needed to properly include lower order terms. By replacing a by 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 (13) 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 parametrix. Moreover, the *exact* solution operator can be written in the form of $c(S_{\mathrm{a}2} + S_{\mathrm{b}2})$ by the addition to a and b of certain symbols in $S^{-\infty}$, which in particular decay faster than any power $|\boldsymbol{\xi}|^{-N}$ (unsurprisingly, 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 [36, Theorem 2.4.1] to the product of several short time solution operators.

3.5. Decoupling

In Section 3.1 we assumed that the functions $a\mathrm{e}^{\mathrm{i}\alpha}$ and $b\mathrm{e}^{\mathrm{i}\beta}$ propagate independently as solutions of the wave equation. In fact, this is the result of a rather general procedure to decouple the wave equation [41]. Because the results of the decoupling will be used explicitly in Section 5 we give a short review of it here; we will examine its relation to the solution operator $S_{\mathrm{a}2}$.

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:

$$\partial_t \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & I \\ c(\mathbf{x})^2 \Delta & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{31}$$

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:

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = S(t) \begin{pmatrix} u_{0,1} \\ u_{0,2} \end{pmatrix} \quad \text{with } S(t) = \begin{pmatrix} S_{11}(t) & S_{12}(t) \\ S_{21}(t) & S_{22}(t) \end{pmatrix}. \tag{32}$$

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 pseudodifferential 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. pseudodifferential 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 $c(\mathbf{x})|\boldsymbol{\xi}|$ and $\frac{1}{c(\mathbf{x})|\boldsymbol{\xi}|}$ respectively. The existence of such operators is a well-known result in pseudodifferential operator theory, see e.g. [42]. We now have the ingredients to define two matrix pseudodifferential operators Λ and V by

$$V = c(\mathbf{x}) \begin{pmatrix} 1 & 1 \\ -\mathrm{i}B & \mathrm{i}B \end{pmatrix} \quad \text{and} \quad \Lambda = \begin{pmatrix} 1 & \mathrm{i}B^{-1} \\ 1 & -\mathrm{i}B^{-1} \end{pmatrix} \frac{1}{2c(\mathbf{x})}, \tag{33}$$

which are each others inverses modulo regularizing operators. We finally define the following two fields $(u_{\mathrm{a}}(\mathbf{x}, t), u_{\mathrm{b}}(\mathbf{x}, t))^T = \Lambda(u_1, u_2)^T$. Note that the Cauchy data can be represented by a time evaluation of $(u_{\mathrm{a}}, u_{\mathrm{b}})^T$. We will use the phrase 'Cauchy data' in this way also. Omitting the regularizing error operators, the system (31) transforms into a decoupled system for $(u_{\mathrm{a}}, u_{\mathrm{b}})^T$ of which the first equation, together with its initial value, is

$$\partial_t u_{\mathrm{a}} = -\mathrm{i}B u_{\mathrm{a}} \quad \text{and} \quad u_{\mathrm{a}}(\mathbf{x}, 0) = u_{0,\mathrm{a}}(\mathbf{x}). \tag{34}$$

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. $u_{\mathrm{a}}(\mathbf{x}, t) = S_{\mathrm{a}}(t) u_{0,\mathrm{a}}(\mathbf{x})$ and similar for S_{b}. Therefore, modulo regularizing operators

$$S(t) = V \begin{pmatrix} S_{\mathrm{a}}(t) & 0 \\ 0 & S_{\mathrm{b}}(t) \end{pmatrix} \Lambda, \tag{35}$$

which means that the original IVP (31) and the decoupled system (34) 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 [43]. It can be shown that $S_{\mathrm{a}}(t)$ and $S_{\mathrm{b}}(t)$ with $t \in \mathbb{R}$ are FIOs [41].

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 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 (26) 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 [36,38], and hence $\frac{a(\mathbf{y},t;\boldsymbol{\xi})}{c(\mathbf{x})} = \text{p.s.}(S_a)\frac{\mathrm{i}}{2c(\mathbf{x})^2|\boldsymbol{\xi}|}$. Using the solution of the transport equation (24), one concludes that

$$\text{p.s.}(S_a)(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) = \sqrt{\det\big(\partial_{\mathbf{y}}\mathbf{x}(\mathbf{y}, t; \boldsymbol{\xi})\big)}. \tag{36}$$

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

3.6. The absence of caustics: 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 given by the fundamental solution, $g(\mathbf{x}, t)$, cf. (2). Subject to assumption SME, 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 3.4 and using a change of phase function [36, 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 [14]

$$g(\mathbf{x}, \mathbf{x}_s, t) = \frac{1}{2\pi}\int A(\mathbf{x}, \mathbf{x}_s, \omega)\mathrm{e}^{\mathrm{i}\omega(t - T(\mathbf{x},\mathbf{x}_s))}\, d\omega, \tag{37}$$

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

$$\left| A(\mathbf{x}, \mathbf{x}_s, \omega) - \sum_{k=0}^{N-1} A_k(\mathbf{x}, \mathbf{x}_s, \omega) \right| \leqslant C_N\big(1 + |\omega|\big)^{\frac{n-3}{2}-N}. \tag{38}$$

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 (37) 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 $A_0(\mathbf{x}, \mathbf{x}_s, \omega) = A_s(\mathbf{x})\sigma(\omega)(\mathrm{i}\omega)^{\frac{n-3}{2}}$ 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

$$g(\mathbf{x}, t) = A_s(\mathbf{x})\partial_t^{\frac{n-3}{2}}\delta\big(t - T_s(\mathbf{x})\big). \tag{39}$$

Functions $A_s(\mathbf{x})$ and $T_s(\mathbf{x})$ will be referred to as the source wave amplitude and traveltime respectively. Operator $\partial_t^{\frac{n-3}{2}}$ denotes the pseudodifferential operator with symbol $\omega \mapsto \sigma(\omega)(\mathrm{i}\omega)^{\frac{n-3}{2}}$. The approximation $g(\mathbf{x}, t)$ matches the exact solution in case $\nabla c = 0$ in the limit of $\omega \to \infty$. 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 $n = 1, 2, 3$ [14]. We define the *source wave direction* vector

$$\mathbf{n}_s(\mathbf{x}) = c(\mathbf{x})\partial_{\mathbf{x}} T_s(\mathbf{x}). \tag{40}$$

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

$$\Xi_s = \big\{(\mathbf{x}, t, \boldsymbol{\xi}, \omega) \in T^*(X \times \mathbb{R}) \setminus 0 \,\big|\, (\mathbf{x}, \boldsymbol{\xi}) = \Phi_t(\mathbf{x}_s, \boldsymbol{\xi}_s),\ \boldsymbol{\xi}_s \in \mathbb{R}^n \setminus \{0\},\ \omega = -c(\mathbf{x})|\boldsymbol{\xi}|\big\}. \tag{41}$$

It contains all bicharacteristics that go through $(\mathbf{x}_s, 0)$ in space–time. In the region where the Fourier integral (37) 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

$$\Xi_{s,t_c} = \big\{(\mathbf{x}, \boldsymbol{\xi}) \in T^*X \setminus 0 \,\big|\, (\mathbf{x}, t_c, \boldsymbol{\xi}, \omega) \in \Xi_s\big\}. \tag{42}$$

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

4. 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 4.1 we derive an explicit expression for the mentioned operator. It will be used in Section 4.2 to define the global scattering operator, of which we show in Theorem 2 that it is an FIO under the conditions of the DSE and the SME.

4.1. 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 1 shows that a local representation of the operator can be written as an oscillatory integral.

The *scattered wave field* $u(\mathbf{x}, t)$ is defined as the solution of the scattering problem (5). We use that the source wave field does not exhibit multipathing (SME) and can therefore be formulated as the asymptotic expansion (37). In the forward modeling we will use the principal term to approximate the source, i.e. (39). (The subprincipal source terms do not contribute to the principal symbol of the scattering operator [15].) We obtain

$$\begin{gathered}\left[\partial_t^2 - c(\mathbf{x})^2\Delta\right]u(\mathbf{x}, t) = r(\mathbf{x})2A_{\mathrm{s}}(\mathbf{x})\partial_t^{\frac{n+1}{2}}\delta\big(t - T_{\mathrm{s}}(\mathbf{x})\big),\\ u(\mathbf{x}, 0) = 0, \qquad \partial_t u(\mathbf{x}, 0) = 0.\end{gathered} \tag{43}$$

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:

$$\begin{gathered}\left[\partial_t^2 - c(\mathbf{x})^2\Delta\right]u_{\mathrm{h}}(\mathbf{x}, t) = 0,\\ u_{\mathrm{h}}(\mathbf{x}, T_1) = u(\mathbf{x}, T_1), \qquad \partial_t u_{\mathrm{h}}(\mathbf{x}, T_1) = \partial_t u(\mathbf{x}, T_1).\end{gathered} \tag{44}$$

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 $T_{\mathrm{s}}(\mathrm{supp}(r)) \subset [T_0, T_1]$. For $t \geqslant T_1$ one has $u_{\mathrm{h}}(\mathbf{x}, t) = u(\mathbf{x}, t)$ 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 $(u_{\mathrm{h},a}, u_{\mathrm{h},b})^T = \Lambda(u_{\mathrm{h}}, \partial_t u_{\mathrm{h}})^T$, with Λ defined in (33).

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 5, models the receiver wave field.

The *scattering operator* F by definition maps r to $(u_{\mathrm{h}}, \partial_t u_{\mathrm{h}})^T$, and we let F_{a} and F_{b} map the reflectivity r to the decoupled components of the continued scattered wave fields $u_{\mathrm{h,a}}$ and $u_{\mathrm{h,b}}$. To show that F_{a} is an 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 multiplication operator then

$$F_{\mathrm{a}}(t) = \sum_{i\in\mathcal{I}} S_{\mathrm{a}}(t - t_{1i})F_{\mathrm{a}}(t_{1i})\rho_i, \tag{45}$$

and F_{b} likewise. S_{a} is the solution operator (35). The ith local scattering event is delimited by $[t_{0i}, t_{1i}]$, so $T_{\mathrm{s}}(\mathrm{supp}(\rho_i)) \subset [t_{0i}, t_{1i}]$. The partition is chosen fine enough such that $[t_{0i}, t_{1i}]$ falls within an interval of definition of (26), i.e. the local expression of solution operator $S_{\mathrm{a}2}$.

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

$$u_{\mathrm{h,a}}(\mathbf{y}, t_1) = \frac{1}{(2\pi)^n}\iint e^{i\varphi_{\mathrm{T}}(\mathbf{y},t_1,\mathbf{x},\boldsymbol{\xi})}A_{\mathrm{F}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})\, d\boldsymbol{\xi}\, \rho r(\mathbf{x})\, d\mathbf{x}, \tag{46}$$

in which the phase and amplitude functions are respectively defined as

$$\varphi_{\mathrm{T}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) = \alpha\big(\mathbf{y}, t_1 - T_{\mathrm{s}}(\mathbf{x}); \boldsymbol{\xi}\big) - \boldsymbol{\xi} \cdot \mathbf{x},$$
$$\mathrm{A}_{\mathrm{F}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) = \big(\mathrm{i}\partial_t \alpha\big(\mathbf{y}, t_1 - T_{\mathrm{s}}(\mathbf{x}); \boldsymbol{\xi}\big)\big)^{\frac{n+1}{2}} \frac{a(\mathbf{y}, t_1 - T_{\mathrm{s}}(\mathbf{x}); \boldsymbol{\xi})}{c(\mathbf{x})} 2A_{\mathrm{s}}(\mathbf{x}). \tag{47}$$

Here (46) is only the contribution of ρr. There is a similar statement for $u_{\mathrm{h},b}$, which satisfies $u_{\mathrm{h},b}(\mathbf{y}, t) = \overline{u_{\mathrm{h},a}(\mathbf{y}, t)}$.

Proof of Theorem 1. To solve the scattering problem (43) it will be transformed into a τ-parameterized family of IVP's. Duhamel's principle states that the solution, i.e. the scattered wave field, is given by

$$u(\mathbf{x}, t) := \int_0^t \tilde{u}(\mathbf{x}, t; \tau)\, d\tau, \tag{48}$$

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 $t = \tau$ [35, §2.4.2]:

$$\big[\partial_t^2 - c(\mathbf{x})^2 \Delta\big]\tilde{u}(\mathbf{x}, t; \tau) = 0 \quad \text{with } t \in \mathbb{R},$$
$$\tilde{u}(\mathbf{x}, \tau; \tau) = 0,$$
$$\partial_t \tilde{u}(\mathbf{x}, \tau; \tau) = r(\mathbf{x}) 2A_{\mathrm{s}}(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta\big(\tau - T_{\mathrm{s}}(\mathbf{x})\big). \tag{49}$$

The continued scattered wave field is the solution of the final value problem (44). Using the observation that $r(\mathbf{x}) 2A_{\mathrm{s}}(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta(\tau - T_{\mathrm{s}}(\mathbf{x})) = 0$ if $\tau \notin [T_0, T_1]$, it can be found by

$$u_{\mathrm{h}}(\mathbf{x}, t) := \int_{T_0}^{T_1} \tilde{u}(\mathbf{x}, t; \tau)\, d\tau \quad \text{with } t \in \mathbb{R}. \tag{50}$$

Time integration is now over the fixed interval $[T_0, T_1]$, by which u_{h} solves the homogeneous wave equation. For $t \geqslant T_1$ the wave fields u and u_{h} coincide. Therefore, this solves (44).

To derive the local expression we solve the τ-parameterized homogeneous IVP (49) with r replaced by ρr and evaluate the solution at t_1. Let $(\tilde{u}_{\mathrm{a}}, \tilde{u}_{\mathrm{b}})^T = \Lambda(\tilde{u}, \partial_t \tilde{u})^T$, then $\tilde{u} = c(\tilde{u}_{\mathrm{a}} + \tilde{u}_{\mathrm{b}})$. We apply solution operator $S_{\mathrm{a}2}$ with initial state at time τ. This gives

$$\tilde{u}_{\mathrm{a}}(\mathbf{y}, t_1; \tau) = S_{\mathrm{a}2}\big[\rho r(\mathbf{x}) 2A_{\mathrm{s}}(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta\big(\tau - T_{\mathrm{s}}(\mathbf{x})\big)\big](\mathbf{y}, t_1 - \tau). \tag{51}$$

Note that $S_{\mathrm{a}2}$ 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 $u_{\mathrm{h},a}(\mathbf{y}, t_1)$, i.e. integral (50) with $\tilde{u}$ replaced by $\tilde{u}_{\mathrm{a}}(\mathbf{y}, t_1; \tau)$ in (51). We will eliminate τ by integration and write the field as an oscillatory integral. With the expression (26) of $S_{\mathrm{a}2}$ and the application of $T_{\mathrm{s}}(\operatorname{supp}(\rho r)) \subset [t_0, t_1]$ one derives the following integral

$$u_{\mathrm{h},a}(\mathbf{y}, t_1) = \frac{1}{(2\pi)^n} \iiint \mathrm{e}^{\mathrm{i}\alpha(\mathbf{y}, t_1 - \tau; \boldsymbol{\xi}) - \mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a(\mathbf{y}, t_1 - \tau; \boldsymbol{\xi})}{c(\mathbf{x})} \rho r(\mathbf{x}) 2A_{\mathrm{s}}(\mathbf{x}) \partial_t^{\frac{n+1}{2}} \delta\big(\tau - T_{\mathrm{s}}(\mathbf{x})\big)\, d\mathbf{x}\, d\boldsymbol{\xi}\, d\tau.$$

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 $\mathrm{e}^{\mathrm{i}\alpha - \mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a}{c}$. Restricting to the highest order term, one writes $\partial_t^{\frac{n+1}{2}} [\mathrm{e}^{\mathrm{i}\alpha - \mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a}{c}] = (\mathrm{i}\partial_t \alpha)^{\frac{n+1}{2}} \mathrm{e}^{\mathrm{i}\alpha - \mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}} \frac{a}{c}$, which is an application of a general result of FIO theory [36,38]. Cutoff σ is omitted to shorten the expression. This yields

$$u_{\mathrm{h},a}(\mathbf{y}, t_1) = \frac{1}{(2\pi)^n} \iiint \bigg[\mathrm{e}^{\mathrm{i}\alpha - \mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}} (\mathrm{i}\partial_t \alpha)^{\frac{n+1}{2}} \frac{a}{c(\mathbf{x})}\bigg]_{(\mathbf{y}, t_1 - \tau; \boldsymbol{\xi})} \rho r(\mathbf{x}) 2A_{\mathrm{s}}(\mathbf{x}) \delta\big(\tau - T_{\mathrm{s}}(\mathbf{x})\big)\, d\mathbf{x}\, d\boldsymbol{\xi}\, d\tau.$$

Notation $[\ldots]_{\mathrm{arg}}$ 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 (46), (47). □

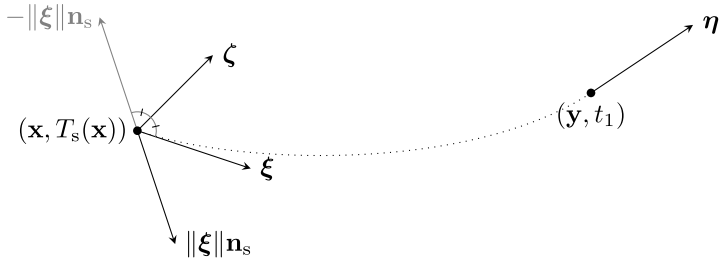

Fig. 1. Propagation of singularities at $(\mathbf{x}, T_s(\mathbf{x}))$ in space–time. See Eq. (52). The dotted line represents the ray. The endpoint of the ray at $(\mathbf{y}, t_1)$ contributes to the scattered field. Here, $|\boldsymbol{\xi}|\mathbf{n}_s$ and $\boldsymbol{\xi}$ can respectively be interpreted as the wave numbers of the initial and reflected waves, and $\boldsymbol{\zeta}$ a normal vector that can be associated with a reflector at $\mathbf{x}$.

4.2. *Scattering operator as an FIO*

Here we establish that $F_a(t_1)\rho$ is an FIO if the direct waves are excluded (DSE). We define the global scattering operator πF and show that it is an FIO with an injective canonical relation, i.e. Theorem 2.

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 $\partial_{\boldsymbol{\xi}}\varphi_T = 0$, i.e. $\partial_{\boldsymbol{\xi}}\alpha(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi}) - \mathbf{x} = 0$. A stationary point $(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi})$ 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 $\boldsymbol{\xi}$, which arrives at $(\mathbf{y}, t_1)$ with covariable $\boldsymbol{\eta} = \partial_{\mathbf{y}}\varphi_T(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})$. 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 (45). Using the terminology introduced at the end of Section 3.4, the ingoing variable and covariable are $(\mathbf{x}, \boldsymbol{\zeta})$ with $\boldsymbol{\zeta} = -\partial_{\mathbf{x}}\varphi_T(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})$. The outgoing variable and covariable are $(\mathbf{y}, \boldsymbol{\eta})$. This means that $F_a(t_1)\rho$ carries over $(\mathbf{x}, \boldsymbol{\zeta}) \in \mathrm{WF}(r)$ into $(\mathbf{y}, \boldsymbol{\eta}) \in \mathrm{WF}(u_{h,a}(., t_1))$.

We have

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

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

$$\boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}), \tag{52}$$

reflecting Snell's law. Fig. 1 shows the microlocal picture of the scattering event and the scattered ray. Eq. (52) also implies that $\boldsymbol{\zeta} \cdot \mathbf{n}_s(\mathbf{x}) < 0$ everywhere. This is a result of the geometry of the reflection event with one source. Note that (52) only holds for negative frequencies. For positive frequencies, i.e. considering F_b, one gets $\boldsymbol{\zeta}' = \boldsymbol{\xi} + |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x})$ instead. In that case $\boldsymbol{\zeta}' \cdot \mathbf{n}_s(\mathbf{x}) > 0$ everywhere.

If $(\mathbf{x}, \boldsymbol{\xi})$ is associated with a source ray, i.e. $\boldsymbol{\xi} = |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x})$, then $\boldsymbol{\zeta} = 0$ by (52). In that case there is no reflection. We show that away from the source rays the scattering operator F is an 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 pseudodifferential 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. (42)) and 1 outside a slightly larger conic neighborhood. Furthermore, we assume that π_a satisfies

$$\pi_a(t) = S_a(t - t_c)\pi_a(t_c)S_a(t_c - t), \tag{53}$$

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

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

$$\text{if } \boldsymbol{\xi}, \tilde{\boldsymbol{\xi}} \in \mathbb{R}^n \setminus \{0\} \quad \text{and} \quad t_i > 0 \quad \text{such that} \quad (\tilde{\mathbf{x}}, \tilde{\boldsymbol{\xi}}) = \Phi_{t_i}(\mathbf{x}, \boldsymbol{\xi}) \quad \text{then}$$
$$\left|T_s(\tilde{\mathbf{x}}) - T_s(\mathbf{x})\right| < t_i \quad \text{or} \quad (\mathbf{x}, \boldsymbol{\xi}) \in \Xi_{s, T_s(\mathbf{x})}. \tag{54}$$

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 an FIO of which the canonical relation is the graph of an injective function. Let $V_{s,t} \subset T^*Y \setminus 0$ be 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 2. *Operator $\pi_a F_a$ defined above, is an FIO. Its canonical relation is*

$$\Lambda = \big\{\big((\mathbf{y}, t, \boldsymbol{\eta}, \omega), (\mathbf{x}, \boldsymbol{\zeta})\big) \;\big|\; (\mathbf{y}, \boldsymbol{\eta}) \in \big(T^*Y \setminus 0\big) \setminus V_{s,t},\ t \in \mathbb{R},\ \omega = -c(\mathbf{y})|\boldsymbol{\eta}|,$$
$$(\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-t}(\mathbf{y}, \boldsymbol{\eta}),\ \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}),\ \mathbf{x} \in D\big\}. \tag{55}$$

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 an FIO. Composition $\pi_a F_a$ is subsequently defined as the sum of local contributions, like in (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 [15,22] show how this can be done, although their scattering operator does not fully coincide with ours.

Proof of Theorem 2. 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

$$\pi_a(t)F_a(t) = \sum_{i \in \mathcal{I}} S_a(t - t_{1i})\pi_a(t_{1i})F_a(t_{1i})\rho_i. \tag{56}$$

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 an FIO. Then $\pi_a F_a$ becomes a sum of compositions of FIOs.

The local scattering operator is the oscillatory integral (46) in which the amplitude A_F (47) is replaced by $\pi_a(t_1; \mathbf{y}, \partial_{\mathbf{y}}\alpha)A_F$. This follows from the application of pseudodifferential operator $\pi_a(t_1)$, its symbol denoted by $\pi_a(t_1; \cdot, \cdot)$, on the integral [36,38]. 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

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

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

$$\Sigma_{t_1} = \big\{(\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in \big(Y \times X \times \mathbb{R}^n \setminus \{0\}\big) \setminus W_{s,t_1} \;\big|\; \mathbf{x} = \partial_{\boldsymbol{\xi}}\alpha\big(\mathbf{y}, t_1 - T_s(\mathbf{x}); \boldsymbol{\xi}\big),\ \mathbf{x} \in \operatorname{supp}(\rho)\big\}. \tag{57}$$

We observe that $|\partial_{\mathbf{x}}\partial_{\boldsymbol{\xi}}\varphi_T| = |\partial_{\boldsymbol{\xi}}\partial_{\mathbf{x}}\varphi_T| = |\partial_{\boldsymbol{\xi}}\boldsymbol{\zeta}|$ by definition of $\boldsymbol{\zeta}$ (52). Moreover

$$|\partial_{\boldsymbol{\xi}}\boldsymbol{\zeta}| = \left|I_n - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \otimes \mathbf{n}_s(\mathbf{x})\right| = 1 - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} \cdot \mathbf{n}_s(\mathbf{x}). \tag{58}$$

By the DSE, applied as the omission of W_{s,t_1} in (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 $\partial_{(\mathbf{y},\mathbf{x},\boldsymbol{\xi})}\partial_{\boldsymbol{\xi}}\varphi_T$ 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

$$\big\{\big((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\zeta})\big) \;\big|\; (\mathbf{y}, \mathbf{x}, \boldsymbol{\xi}) \in \Sigma_{t_1},\ \boldsymbol{\eta} = \partial_{\mathbf{y}}\varphi_T,\ \boldsymbol{\zeta} = -\partial_{\mathbf{x}}\varphi_T\big\}. \tag{59}$$

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 an 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 an FIO. The canonical relation is determined by the composition of relations [36,38]. The global scattering operator $\pi_a(t)F_a(t)$ is subsequently defined as the sum (56) of the extended local operators, of which the canonical relation Λ_t is the union of the local relations (59).

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

$$U_s = \{(\mathbf{x}, \boldsymbol{\zeta}) \in T^*X \setminus 0 \,|\, (\mathbf{x}, \boldsymbol{\xi}) \in V_{s,T_s(\mathbf{x})},\ \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}),\ \mathbf{x} \in D\}. \tag{60}$$

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

$$\Lambda_t = \{((\mathbf{y}, \boldsymbol{\eta}), (\mathbf{x}, \boldsymbol{\zeta})) \,|\, (\mathbf{y}, \boldsymbol{\eta}) \in (T^*Y \setminus 0) \setminus V_{s,t},\ (\mathbf{x}, \boldsymbol{\xi}) = \Phi_{T_s(\mathbf{x})-t}(\mathbf{y}, \boldsymbol{\eta}),\ \boldsymbol{\zeta} = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}_s(\mathbf{x}),\ \mathbf{x} \in D\}. \tag{61}$$

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 (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 $(\mathbf{y}, \boldsymbol{\eta})$.

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 (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 (52). The transformation Θ_t is therefore one-to-one.

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

$$(\mathbf{x}, \boldsymbol{\zeta}) \xrightarrow{(52)} (\mathbf{x}, \boldsymbol{\xi}) \xrightarrow{\Phi_{\tau_0 - T_s(\mathbf{x})}} (\check{\mathbf{x}}, \check{\boldsymbol{\xi}}) \xrightarrow{\Phi_{t-\tau_0}} (\mathbf{y}, \boldsymbol{\eta}).$$

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

$$\left|\frac{\partial(\mathbf{y}, \boldsymbol{\eta})}{\partial(\mathbf{x}, \boldsymbol{\zeta})}\right| = \left|\frac{\partial(\mathbf{y}, \boldsymbol{\eta})}{\partial(\check{\mathbf{x}}, \check{\boldsymbol{\xi}})}\right| \left|\frac{\partial(\check{\mathbf{x}}, \check{\boldsymbol{\xi}})}{\partial(\mathbf{x}, \boldsymbol{\xi})}\right| \left|\frac{\partial(\mathbf{x}, \boldsymbol{\xi})}{\partial(\mathbf{x}, \boldsymbol{\zeta})}\right|. \tag{62}$$

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 (58). The transformation $\Phi_{\tau_0 - T_s(\mathbf{x})}$ is the least obvious one. We will show in Lemma 1 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 space–time distributions. Parameterized by $\mathbf{y}$, $\boldsymbol{\eta}$ and t, Λ becomes (55). The injectivity follows from the parameterization. □

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*

$$\det \partial_{(\mathbf{x},\boldsymbol{\xi})} J(\mathbf{x}_0, \boldsymbol{\xi}_0) = 1 - \frac{\boldsymbol{\xi}_0}{|\boldsymbol{\xi}_0|} \cdot \mathbf{n}_s(\mathbf{x}_0), \tag{63}$$

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 straightforward calculation

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

Herein we substitute the right-hand side of the characteristic ODE (18) for $\partial_s \Phi_s$. □

5. 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 pseudodifferential 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 $t = t_{\mathrm{c}}$, 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 will set $u = u_{\mathrm{h}}$. We let $T_M u$ denote the restriction of u to M as before. The field u_{r} is an anticausal solution to

$$\left[c(\mathbf{x})^{-2}\partial_t^2 - \Delta\right]u_{\mathrm{r}}(\mathbf{x}, t) = \delta(x_n) F_M T_M u\left(\mathbf{x}', t\right); \tag{64}$$

here, F_M is a boundary operator consisting of two types of factors. The first factor is the pseudodifferential operator given in (7) which accounts for the fact that the boundary data for the backpropagation enters as a source and not as a boundary condition. The singularity in the square root is avoided by the cutoff for grazing rays 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 pseudodifferential operator which removes waves that reach the surface along tangently incoming rays. Its symbol is zero around $(\mathbf{x}', t, \boldsymbol{\xi}', \omega)$ such that

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

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 these pseudodifferential cutoffs. The principal symbol of F_M is then

$$-2\mathrm{i}\omega c^{-1}\sqrt{1 - c^2\omega^{-2}\boldsymbol{\xi}'^2}\,\Psi_M\left(\mathbf{x}', t, \boldsymbol{\xi}', \omega\right). \tag{65}$$

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 $u_{\mathrm{r,a}}$ and $u_{\mathrm{r,b}}$ depend locally on u_{a} and u_{b} in the following fashion,

$$\begin{aligned} \chi u_{\mathrm{r,a}}(\cdot, t) &= \chi\left[P_{\mathrm{a}}(t)u_{\mathrm{a}}(\cdot, t) + R_1(t)u_{\mathrm{b}}(\cdot, t)\right], \quad \text{and} \\ \chi u_{\mathrm{r,b}}(\cdot, t) &= \chi\left[P_{\mathrm{b}}(t)u_{\mathrm{b}}(\cdot, t) + R_2(t)u_{\mathrm{a}}(\cdot, t)\right]. \end{aligned} \tag{66}$$

Here, $P_{\mathrm{a}}(t)$ and $P_{\mathrm{b}}(t)$ are pseudodifferential operators described below and $R_1(t)$ and $R_2(t)$ are regularizing operators, and χ is a cutoff because the source in Eq. (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 pseudodifferential operators

$$P(t) = V\begin{pmatrix} P_{\mathrm{a}}(t) & R_1(t) \\ R_2(t) & P_{\mathrm{b}}(t) \end{pmatrix}\Lambda. \tag{67}$$

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

$$\text{bicharacteristics through } M \text{ and } \operatorname{supp}(\chi) \text{ intersect } M \text{ only once and with } dx_n/dt < 0. \tag{68}$$

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 $\mathbb{R}^n \times \{s\} \subset \mathbb{R}^n_{\mathbf{x}} \times \mathbb{R}_t$. The bicharacteristic flow provides a map

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

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

$$\Psi_{X_s,a}(\mathbf{x}, \boldsymbol{\xi}) = \begin{cases} \Psi_M(\mathbf{y}'_a(\mathbf{x}, \boldsymbol{\xi}, t-s), t, \boldsymbol{\eta}'_a(\mathbf{x}, \boldsymbol{\xi}, t-s), -c(\mathbf{x})|\boldsymbol{\xi}|) & \text{when } \exists t \text{ with } \mathbf{y}_a(\mathbf{x}, \boldsymbol{\xi}, t-s) \in M, \\ 0 & \text{otherwise,} \end{cases} \tag{69}$$

and $\Psi_{X_0,b}$ is defined similarly using the $(\mathbf{y}_b, \boldsymbol{\eta}_b)$ flow. We can now state and prove the following:

Theorem 3. *Let $u_{r,a}$, $u_{r,b}$ and u_a, u_b, χ and M be as just defined. Eq. (66) holds, in which P_a and P_b are pseudodifferential operators in* $\operatorname{Op} S^0(\mathbb{R}^n)$*, the principal symbols of which are given by*

$$p_a(t; \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_t,a}(\mathbf{x}, \boldsymbol{\xi}) \quad \text{and} \quad p_b(t; \mathbf{x}, \boldsymbol{\xi}) = \Psi_{X_t,b}(\mathbf{x}, \boldsymbol{\xi}), \tag{70}$$

respectively. The operators P_a, P_b satisfy property (53) as far as they are uniquely determined considering the cutoff χ in (66).

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 of Theorem 3. We prove (66) 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 (possibly modulo a smooth contribution) by the Cauchy values $u_{0,a} = u_a(\cdot, 0)$ and $u_{0,b} = u_b(\cdot, 0)$. Consider the equation $\partial_t u_a = -\mathrm{i}Bu_a$. 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 $(\mathbf{x}, t)$, $t > s$. We will write $S_a(t, \cdot)$ for the operator that gives the anticausal solution to $(\partial_t + \mathrm{i}Bu_a)u_a = f_a$,

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

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

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

The adjoint of this operator is given by the following. With auxiliary function f it is:

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

These operators are well-defined on suitable sets of distributions. We use the notation (cf. (64))

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

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

$$f_M = F_M T_M (c S_a u_{0,a} + c S_b u_{0,b}), \tag{71}$$

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

To begin with, there exists a pseudodifferential 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 $|\boldsymbol{\xi}| \leqslant C|\omega|$ in $\mathrm{WF}(u)$ 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 $|\boldsymbol{\xi}| \leqslant C|\omega|$ in $\mathrm{WF}(u)$. Thus the first term in (71) is given by

$$T_M \tilde{F}_M c S_{\mathrm{a}}(\cdot, 0) \tag{72}$$

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

$$\big\{\big(\big(\mathbf{y}_{\mathrm{a}}(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}_{\mathrm{a}}(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|\big), (\mathbf{x}, \boldsymbol{\xi})\big)\big\}. \tag{73}$$

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

$$\big\{\big(\big(\mathbf{y}', t, \boldsymbol{\eta}', \omega\big), \big(\mathbf{y}', 0, t, \boldsymbol{\eta}', \eta_n, \omega\big)\big)\big\}. \tag{74}$$

As tangent rays are removed, the composition of canonical relations (74) and (73) is transversal. Therefore, (72) is a Fourier integral operator. Moreover, from assumption (68) it follows that the canonical relation is the graph of an invertible map, given by

$$\big\{\big(\big(\mathbf{y}_{\mathrm{a}}'(\mathbf{x}, \boldsymbol{\xi}, t), t, \boldsymbol{\eta}_{\mathrm{a}}'(\mathbf{x}, \boldsymbol{\xi}, t), -c(\mathbf{x})|\boldsymbol{\xi}|\big), (\mathbf{x}, \boldsymbol{\xi})\big) \;\big|\; t \text{ s.t. } \mathbf{y}_{\mathrm{a}}(\mathbf{x}, \boldsymbol{\xi}, t) \in M\big\},$$

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

Next, we consider the map $f_M \mapsto \chi u_{\mathrm{r,a}}$. We insert a pseudodifferential 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 3.5, the source $(f_{\mathrm{a}}, f_{\mathrm{b}})$ for the inhomogeneous wave equation is given by $(f_{\mathrm{a}}, f_{\mathrm{b}})^T = \Lambda(0, cf)^T$, hence $\chi u_{\mathrm{r,a}}$ satisfies

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

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

$$\chi S_{\mathrm{a}}(0, \cdot)\left(\frac{\mathrm{i}}{2} B^{-1} c\right) \tilde{\Xi} T_M^*$$

acting on f_M, modulo a smoothing operator.

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

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

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 assumption (68) and the cutoff χ, the contributions with $+$ sign can be ignored. We write $S_{\mathrm{a}}^{(-)}(0, \cdot)(-\frac{\mathrm{i}}{2} B^{-1} c)\tilde{\Xi} T_M^*$ 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

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

Again this is an invertible canonical relation.

The next step is the composition of the maps $(u_{0,\mathrm{a}}, u_{0,\mathrm{b}}) \mapsto f_M$ and $f_M \mapsto (u_{\mathrm{r,a}}(\cdot, 0), u_{\mathrm{r,b}}(\cdot, 0))$. As both maps are Fourier integral operators with canonical relations that are the graph of an invertible map, the composition is a (sum of) well-defined Fourier integral operators. The fields u_{a} and $u_{\mathrm{r,a}}$ are associated with negative ω, u_{b} and $u_{\mathrm{r,b}}$ with positive ω. One can verify that the "cross terms" $u_{0,\mathrm{a}} \mapsto u_{\mathrm{r,b}}(\cdot, 0)$ and $u_{0,\mathrm{b}} \mapsto u_{\mathrm{r,a}}(\cdot, 0)$ are smoothing operators. The maps $u_{0,\mathrm{a}} \mapsto u_{\mathrm{r,a}}(\cdot, 0)$ and $u_{0,\mathrm{b}} \mapsto u_{\mathrm{r,b}}(\cdot, 0)$ are pseudodifferential operators. The principal symbol $p_{\mathrm{a}}(0; \mathbf{x}, \boldsymbol{\xi})$ 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 $u_{0,\mathrm{a}} \mapsto u_{\mathrm{r,a}}(\cdot, 0)$, the map $u_{0,\mathrm{b}} \mapsto u_{\mathrm{r,b}}(\cdot, 0)$ can be done in a similar way. The map $u_{0,\mathrm{a}} \mapsto f_M$ can then be written in the form

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

where the amplitude satisfies

$$a^{(\mathrm{fwd})}\big(\mathbf{y}', t, \mathbf{x}, \boldsymbol{\xi}\big) = -2\mathrm{i}\chi(x_n)\omega\sqrt{1 - c\big(\mathbf{y}', 0\big)^2\omega^{-2}\boldsymbol{\eta}'^2}\sqrt{\det(\partial_{\mathbf{y}}\mathbf{x})}\Psi_M\big(\mathbf{y}, t, \boldsymbol{\eta}', \omega\big) \mod S^0\big(\mathbb{R}^{2n} \times \mathbb{R}^n\big), \tag{75}$$

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 $f_M \mapsto \chi u_{\mathrm{r,a}}(\cdot, 0)$ is given by $\Xi^* T_M c \frac{\mathrm{i}}{2} B^{-1} S_{\mathrm{a}}(\cdot, 0)\chi$, 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

$$a^{(\mathrm{bkd})}\big(\mathbf{y}', t, \mathbf{z}, \boldsymbol{\zeta}\big) = \frac{\mathrm{i}}{2}\chi(z_n)\big(-\omega^{-1}\big)c(\mathbf{y})\sqrt{\det\big(\partial_{\mathbf{y}}\mathbf{x}\big(\mathbf{y}', t, \boldsymbol{\zeta}\big)\big)}\,\Xi \mod S^{-2}\big(\mathbb{R}^{2n} \times \mathbb{R}^n\big). \tag{76}$$

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

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

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

$$\frac{1}{(2\pi)^{2n}} \iiiint \overline{a^{(\mathrm{bkd})}\big(\mathbf{y}', t, \mathbf{z}, \boldsymbol{\zeta}\big)} a^{(\mathrm{fwd})}\big(\mathbf{y}', t, \mathbf{x}, \boldsymbol{\xi}\big)\mathrm{e}^{\mathrm{i}(-\alpha(\mathbf{y}',0,t,\boldsymbol{\zeta}) + \alpha(\mathbf{y}',0,t,\boldsymbol{\xi}) + \mathbf{z}\cdot\boldsymbol{\zeta} - \mathbf{x}\cdot\boldsymbol{\xi})}\, d\mathbf{y}'\, dt\, d\boldsymbol{\xi}\, d\boldsymbol{\zeta}. \tag{77}$$

Using a smooth cutoff the $(\boldsymbol{\xi}, \boldsymbol{\zeta})$ integration domain can be divided into three parts, one with $|\boldsymbol{\zeta}| \leqslant 2|\boldsymbol{\xi}|$, one with $|\boldsymbol{\zeta}| \geqslant \frac{4}{3}|\boldsymbol{\xi}|$, 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 $g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})$ such that

$$\frac{1}{(2\pi)^n} \iiint \overline{a^{(\mathrm{bkd})}} a^{(\mathrm{fwd})}\mathrm{e}^{\mathrm{i}(-\alpha(\mathbf{y}',0,t,\boldsymbol{\zeta}) + \alpha(\mathbf{y}',0,t,\boldsymbol{\xi}) + \mathbf{z}\cdot\boldsymbol{\zeta} - \mathbf{x}\cdot\boldsymbol{\xi})}\, d\mathbf{y}'\, dt\, d\boldsymbol{\zeta} = g(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi})\mathrm{e}^{\mathrm{i}(\mathbf{z}-\mathbf{x})\cdot\boldsymbol{\xi}}, \tag{78}$$

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 map

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

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

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

Therefore, the bicharacteristic determined by $(\mathbf{z}, \boldsymbol{\xi})$ must be the same as the bicharacteristic determined by $(\mathbf{y}', t, \boldsymbol{\zeta})$. Let $\psi(\mathbf{y}', t, \boldsymbol{\zeta}; \mathbf{x}, \boldsymbol{\xi})$ be a C^∞ cutoff function that is one for a small neighborhood of $(\mathbf{y}', t, \boldsymbol{\zeta})$ 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 $|\boldsymbol{\zeta}| \geqslant \frac{4}{3}|\boldsymbol{\xi}|$, can be treated similarly, with the role of $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ interchanged. In this case the stationary point set is in the region where the amplitude is zero, and its contribution is of the form (78), but with g in $S^{-\infty}(\mathbb{R}^{2n} \times \mathbb{R}^n)$. The third part, $\boldsymbol{\zeta}, \boldsymbol{\xi}$ around zero, also yields such a contribution with $g \in S^{-\infty}(\mathbb{R}^{2n} \times \mathbb{R}^n)$.

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

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

The goal is to rewrite (77) by change of variables into

$$\frac{1}{(2\pi)^n}\iiint \psi\overline{a^{(\mathrm{bkd})}}a^{(\mathrm{fwd})}\mathrm{e}^{\mathrm{i}((\boldsymbol{\xi}-\boldsymbol{\zeta})\cdot X+\mathbf{z}\cdot\boldsymbol{\zeta}-\mathbf{x}\cdot\boldsymbol{\xi})}\left|\frac{\partial X}{\partial(\mathbf{y}',t)}\right|^{-1}d\boldsymbol{\zeta}\,dX, \tag{79}$$

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 $\boldsymbol{\xi}=\boldsymbol{\zeta}$ and $y_n=0$,

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

where $(\mathbf{y},t,\boldsymbol{\xi})\mapsto\mathbf{x}(\mathbf{y},t,\boldsymbol{\xi})$ was discussed in Section 3.4. 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 (79) has a quadratic phase function $\boldsymbol{\zeta}\cdot(X-\mathbf{z})$, and can be performed as usual in the method of stationary phase [36, Lemma 1.2.4]. This shows that $g(\mathbf{z},\mathbf{x},\boldsymbol{\xi})$ satisfies the symbol property. Using (75) and (76) it follows that

$$g(\mathbf{x},\mathbf{x},\boldsymbol{\xi})=-2\mathrm{i}\big(-c|\boldsymbol{\xi}|\big)\sqrt{1-c^2\omega^{-2}\boldsymbol{\eta}'^2}\left|\frac{\partial\mathbf{x}}{\partial\mathbf{y}}\right|^{\frac{1}{2}}\Psi_M\left(-\frac{\mathrm{i}}{2}\right)c(\mathbf{x})\big(c|\boldsymbol{\xi}|\big)^{-1}c(\mathbf{y})\left|\frac{\partial\mathbf{x}}{\partial\mathbf{y}}\right|^{\frac{1}{2}}\left|\frac{\partial\mathbf{x}}{\partial\mathbf{y}}\right|^{-1}\left|\frac{\partial y_n}{\partial t}\right|^{-1}.$$

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 $|\frac{\partial y_n}{\partial t}|=c(\mathbf{y})\cos(\theta_M)$. Therefore indeed we have

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

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

$$P_{\mathrm{a}}(s)\stackrel{\mathrm{def}}{=}S_{\mathrm{a}}^{(-)}(s,\cdot)\left(-\frac{\mathrm{i}}{2}B^{-1}c\right)T_M^*F_MT_MS_{\mathrm{a}}(\cdot,s)\in\mathrm{Op}\,S^0\big(\mathbb{R}^n\big)$$

with principal symbol

$$\Psi_{X_s}(\mathbf{x},\boldsymbol{\xi})=\Psi_M\big(\mathbf{y}'(\mathbf{x},\boldsymbol{\xi},t-s),t,\boldsymbol{\eta}'(\mathbf{x},\boldsymbol{\xi},t-s),-c(\mathbf{x})|\boldsymbol{\xi}|\big),\quad\text{with } t \text{ s.t. } y_n(\mathbf{x},\boldsymbol{\xi},t-s)=0.$$

From the group property of the $S_{\mathrm{a}}(t,s)$ it follows that

$$P_{\mathrm{a}}(0)=\chi S_{\mathrm{a}}(0,s)P_{\mathrm{a}}(s)S_{\mathrm{a}}(s,0).$$

The evolution operators $S_{\mathrm{a}}(0,s)$, $S_{\mathrm{a}}(s,0)$ are each others inverses. According to the Egorov theorem [41, Section 8.1] the operator $P_{\mathrm{a}}(0)$ is a pseudodifferential 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. □

6. Inverse scattering

This section deals with the inverse scattering problem. The diagram in Fig. 2 shows how we theoretically approach RTM. The forward modeling is given by $r\to u\to d$ 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

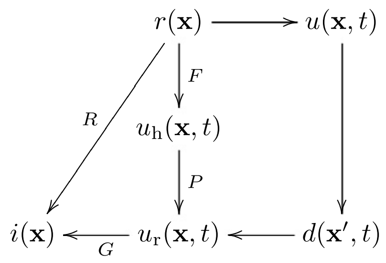

Fig. 2. Diagram showing the theoretical approach to RTM.

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* (92), in two steps. We propose the *imaging operator* G and show in Theorem 5 and its proof that it is an 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.

Instead of condition (68) we have the following condition for the RTM-based inversion

$$\text{bicharacteristics that enter the region } x_n < 0 \text{ do not return to the region } x_n \geqslant 0. \tag{80}$$

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 (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. (65).

6.1. 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

$$\big[c^{-2}\partial_t^2 - \Delta\big]u(\mathbf{x},t) = A\delta\big(t - c^{-1}x_3\big)r(\mathbf{x}), \tag{81}$$

which is a slight simplification of (43). For simplicity the analysis will be 3-dimensional, but it applies to other dimensions as well.

The solution of the PDE (81) is given in the $(\boldsymbol{\xi}, t)$ domain by

$$\hat{u}(\boldsymbol{\xi},t) = \int_0^t \big(\mathrm{e}^{\mathrm{i}c|\boldsymbol{\xi}|(t-s)} - \mathrm{e}^{-\mathrm{i}c|\boldsymbol{\xi}|(t-s)}\big)\frac{c^2}{2\mathrm{i}c|\boldsymbol{\xi}|}\hat{f}(\boldsymbol{\xi},s)\,ds, \tag{82}$$

where, for now, we denote by f the right-hand side of (81). 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

$$\int \mathrm{e}^{-\mathrm{i}x_3\xi_3}A\delta\left(t-\frac{x_3}{c}\right)\tilde{r}(\xi_1,\xi_2,x_3)\,dx_3 = cA\mathrm{e}^{-\mathrm{i}\xi_3 ct}\tilde{r}(\xi_1,\xi_2,ct). \tag{83}$$

Next we use (82) and (83) to solve (81), and we make a change of variable $cs = \tilde{z}$. This yields

$$\hat{u}(\boldsymbol{\xi},t) = \int_0^{tc} \big(\mathrm{e}^{\mathrm{i}|\boldsymbol{\xi}|(ct-\tilde{z})} - \mathrm{e}^{-\mathrm{i}|\boldsymbol{\xi}|(ct-\tilde{z})}\big)\frac{c^2}{2\mathrm{i}c|\boldsymbol{\xi}|}A\mathrm{e}^{-\mathrm{i}\xi_3\tilde{z}}\tilde{r}(\xi_1,\xi_2,\tilde{z})\,d\tilde{z}.$$

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 $\xi_3 \pm |\boldsymbol{\xi}|$, because $\tilde{z}$ occurs at several places in the complex exponents. Under the assumption that the support of r is contained in $0 < x_3 < ct$ (in other words, that we consider the field at time t such that the incoming wave front has completely passed the support of the reflectivity), the formula equals

$$\hat{u}(\boldsymbol{\xi}, t) = \mathrm{e}^{\mathrm{i}|\boldsymbol{\xi}|ct} \frac{c^2 A}{2\mathrm{i}c|\boldsymbol{\xi}|} \hat{r}\big(\boldsymbol{\xi} + \big(0, 0, |\boldsymbol{\xi}|\big)\big) - \mathrm{e}^{-\mathrm{i}|\boldsymbol{\xi}|ct} \frac{c^2 A}{2\mathrm{i}c|\boldsymbol{\xi}|} \hat{r}\big(\boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big)\big). \tag{84}$$

The field in position coordinates is given by the inverse Fourier transform of this, i.e. by

$$u(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \int\limits_{\mathbb{R}^3} \left[\mathrm{e}^{\mathrm{i}|\boldsymbol{\xi}|ct} \frac{c^2 A}{2\mathrm{i}c|\boldsymbol{\xi}|} \hat{r}\big(\boldsymbol{\xi} + \big(0, 0, |\boldsymbol{\xi}|\big)\big) - \mathrm{e}^{-\mathrm{i}|\boldsymbol{\xi}|ct} \frac{c^2 A}{2\mathrm{i}c|\boldsymbol{\xi}|} \hat{r}\big(\boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big)\big) \right] \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}}\, d\boldsymbol{\xi}. \tag{85}$$

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 that the property that $r(\mathbf{x})$ is real for all $\mathbf{x}$ is equivalent to $\hat{r}(\boldsymbol{\xi}) = \overline{\hat{r}(-\boldsymbol{\xi})}$ for all $\boldsymbol{\xi}$. Therefore

$$u(\mathbf{x}, t) = \frac{1}{(2\pi)^3} \operatorname{Re} \int\limits_{\mathbb{R}^3} \mathrm{e}^{-\mathrm{i}|\boldsymbol{\xi}|ct + \mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}} \frac{\mathrm{i}cA}{|\boldsymbol{\xi}|} \hat{r}\big(\boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big)\big)\, d\boldsymbol{\xi}. \tag{86}$$

There are three wave vectors in (86), $\boldsymbol{\xi}$ is the wave vector of the outgoing reflected wave, $(0, 0, |\boldsymbol{\xi}|)$ can be interpreted as the wave vector of the incoming wave, while $\boldsymbol{\xi} - (0, 0, |\boldsymbol{\xi}|)$ can be interpreted as the reflectivity wave number, which, for a conormal singularity for example, would be normal to the reflector.

In this simplified analysis we assume that the reverse time continued receiver field u_{r} satisfies a homogeneous wave equation with equal final values (after the scattering) as u, like u_{h} in (44), i.e. it results from an idealized experiment as explained in Section 4.1. This means that u_{r} is also given by (86), 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 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 instead as our image

$$I(\mathbf{x}) = \frac{2}{c^2 A} (\partial_t + c\partial_{x_3}) u(\mathbf{x}, x_3/c). \tag{87}$$

We have from (86)

$$\frac{2}{c^2 A} (\partial_t + c\partial_{x_3}) u(\mathbf{x}, t) = \frac{2}{(2\pi)^3} \operatorname{Re} \int\limits_{\mathbb{R}^3} \left(1 - \frac{\xi_3}{|\boldsymbol{\xi}|}\right) \mathrm{e}^{-\mathrm{i}|\boldsymbol{\xi}|ct + \mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}} \hat{r}\big(\boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big)\big)\, d\boldsymbol{\xi}. \tag{88}$$

Setting $t = x_3/c$ we find

$$I(\mathbf{x}) = \frac{2}{(2\pi)^3} \operatorname{Re} \int\limits_{\mathbb{R}^3} \left(1 - \frac{\xi_3}{|\boldsymbol{\xi}|}\right) \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot(\boldsymbol{\xi} - (0,0,|\boldsymbol{\xi}|))} \hat{r}\big(\boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big)\big)\, d\boldsymbol{\xi}. \tag{89}$$

We carry out a coordinate transformation,

$$\tilde{\boldsymbol{\xi}} = \boldsymbol{\xi} - \big(0, 0, |\boldsymbol{\xi}|\big), \qquad \left|\frac{\partial \tilde{\boldsymbol{\xi}}}{\partial \boldsymbol{\xi}}\right| = 1 - \frac{\xi_3}{|\boldsymbol{\xi}|}. \tag{90}$$

The image of this transformation is the half-plane $\tilde{\xi}_3 < 0$, while the Jacobian is as given in (90), and exactly equals the factor $1 - \frac{\xi_3}{|\boldsymbol{\xi}|}$ from the derivative operator $\partial_t + c\partial_{x_3}$. Therefore by a change of variables (89) equals $\frac{1}{(2\pi)^3} \operatorname{Re} \int_{\tilde{\xi}_3 < 0} \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\tilde{\boldsymbol{\xi}}} \hat{r}(\tilde{\boldsymbol{\xi}})\, d\tilde{\boldsymbol{\xi}}$. This can be rewritten as

$$I(\mathbf{x}) = \frac{1}{(2\pi)^3} \int\limits_{\tilde{\xi}_3 \neq 0} \mathrm{e}^{\mathrm{i}\mathbf{x}\cdot\tilde{\boldsymbol{\xi}}} \hat{r}(\tilde{\boldsymbol{\xi}})\, d\tilde{\boldsymbol{\xi}}. \tag{91}$$

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 with inverting from direct waves. This simple calculation gives the motivation for the imaging condition (92) below, in particular, for the term involving the gradient $\partial_{\mathbf{x}} \hat{u}_{\mathrm{r}}(\mathbf{x}, \omega)$.

6.2. Variable background velocity

The *imaging condition* yields a mapping of the source wave $g(\mathbf{x}, t)$ and the reverse time continued wave $u_{\mathrm{r}}(\mathbf{x}, t)$ to an image $i(\mathbf{x})$ of the reflectivity. We will show that the imaging condition, $i(\mathbf{x}) = \mathcal{H}_M d(\mathbf{x})$, that is,

$$i(\mathbf{x}) = \frac{1}{2\pi} \int \frac{\Omega(\omega)}{\mathrm{i}\omega|\hat{g}(\mathbf{x},\omega)|^2} \left(\overline{\hat{g}(\mathbf{x},\omega)} \hat{u}_{\mathrm{r}}(\mathbf{x},\omega) - \frac{c(\mathbf{x})^2}{\omega^2} \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x},\omega)} \cdot \partial_{\mathbf{x}} \hat{u}_{\mathrm{r}}(\mathbf{x},\omega) \right) d\omega, \tag{92}$$

cf. (8), yields a partial inverse. The support of Ω is obtained in the proof of the theorem. To characterize $i(\mathbf{x})$, the relation (52) between $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ is important. We observe that the inverse function $\boldsymbol{\xi}(\boldsymbol{\zeta})$ of (52) is defined on the half-space

$$\left\{ \boldsymbol{\zeta} \in \mathbb{R}^n \setminus 0 \;\middle|\; \boldsymbol{\zeta} \cdot \mathbf{n}_{\mathrm{s}}(\mathbf{z}) < 0 \right\}. \tag{93}$$

The function $p_{\mathrm{a}}(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta}))$, p_{a} the principal symbol of the revert operator, is in principle defined only on (93). However, due to the DSE, it is zero for $\boldsymbol{\zeta}$ near the boundary of this half-space and we will consider it as a function on $\mathbb{R}^n \setminus 0$ that is zero outside (93). With this definition, the function $(\mathbf{z}, \boldsymbol{\zeta}) \mapsto p_{\mathrm{a}}(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta}))$ is an order 0 symbol.

Theorem 4. *Let image $i(\mathbf{x})$ be defined by* (92)*, and assume* (3)*,* (4) *and* (80)*. Define operator R by the map from the reflectivity r to the image, $Rr(\mathbf{x}) = i(\mathbf{x})$. Then R is a pseudodifferential operator of order zero, and its principal symbol satisfies*

$$\text{p.s.}(R)(\mathbf{z}, \boldsymbol{\zeta}) = p_{\mathrm{a}}\big(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta})\big) + p_{\mathrm{a}}\big(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(-\boldsymbol{\zeta})\big), \tag{94}$$

where the map $(\mathbf{z}, \boldsymbol{\zeta}) \mapsto p_{\mathrm{a}}(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\zeta}))$ 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 (94) corresponds to the negative frequencies and the second contribution to the positive frequencies. As the supports, i.e. (93) for the first, of these two terms are disjoint, (94) defines a symbol that is one on a subset of $\mathbb{R}^n \times \mathbb{R}^n \setminus 0$. Hence, the map $d \mapsto i$ given by (92) can rightfully be called a partial inverse.

The imaging condition (92) is based on the actual source field g. Before proving Theorem 4, we derive an intermediate result with an imaging condition based on the source wave traveltime $T_{\mathrm{s}}(\mathbf{x})$, and the highest order contribution to the amplitude $A_{\mathrm{s}}(\mathbf{x})$. Let $w \in \mathcal{E}'(Y \times \mathbb{R})$ be an auxiliary distribution. Let operators H and K be defined by

$$\begin{aligned} Hw(\mathbf{y}, t) &= \frac{1}{A_{\mathrm{s}}(\mathbf{y})} \partial_t^{-\frac{n+1}{2}} \big[\partial_t + c(\mathbf{y}) \mathbf{n}_{\mathrm{s}}(\mathbf{y}) \cdot \partial_{\mathbf{y}}\big] w(\mathbf{y}, t), \\ Kw(\mathbf{z}) &= w\big(\mathbf{z}, T_{\mathrm{s}}(\mathbf{z})\big). \end{aligned} \tag{95}$$

Operator K is a restriction to a hypersurface in $\mathbb{R}^{n+1}$. Operator H is a pseudodifferential operator. Operator $\partial_t^{-\frac{n+1}{2}}$ is to be read as the pseudodifferential operator with symbol $\omega \mapsto \tilde{\sigma}(\omega)(\mathrm{i}\omega)^{-\frac{n+1}{2}}$ 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 $V_1 = (1\ \ 0)$ which projects out the first component of a two-vector. We define the *imaging operator* $G = KH$.

Theorem 5. *If* (3)*,* (4) *and* (80) *are satisfied and $\tilde{R}$ is given by*

$$\tilde{R}r = Gu_{\mathrm{r}} = GV_1 PFr, \tag{96}$$

then $\tilde{R}$ is a pseudodifferential operator of order zero with principal symbol given by (94)*.*

Proof. We first work out the details for the negative frequencies, leading to a characterization of $\tilde{R}_{\mathrm{a}} = KHcP_{\mathrm{a}}F_{\mathrm{a}}$. We then consider the positive, and add the contributions, $\tilde{R} = \tilde{R}_{\mathrm{a}} + \tilde{R}_{\mathrm{b}}$.

(i) We show that the composition $\tilde{R}_{\mathrm{a}} = KHcP_{\mathrm{a}}F_{\mathrm{a}}$ is an 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,

$$Kw(\mathbf{z}) = (2\pi)^{-n-1} \iint \mathrm{e}^{\mathrm{i}\boldsymbol{\eta}\cdot(\mathbf{z}-\mathbf{y}) + \mathrm{i}\omega(T_{\mathrm{s}}(\mathbf{z}) - t)} w(\mathbf{y}, t)\, d(\mathbf{y}, t)\, d\boldsymbol{\eta}\, d\omega \tag{97}$$

with canonical relation

$$\Upsilon = \{((\mathbf{z}, \boldsymbol{\theta}), (\mathbf{y}, t, \boldsymbol{\eta}, \omega)) \mid (\mathbf{y}, \boldsymbol{\eta}) \in T^*Y \setminus 0,\ t = T_s(\mathbf{y}),\ \omega \in \mathbb{R} \setminus \{0\},\ \mathbf{z} = \mathbf{y},\ \boldsymbol{\theta} = \boldsymbol{\eta} + c(\mathbf{y})^{-1}\omega \mathbf{n}_s(\mathbf{y})\}. \tag{98}$$

First consider $K\pi_a F_a$, which is the composition of K and $\pi_a F_a$ with canonical relations given respectively by Υ (98) and Λ (55). We consider the composition of the Fourier integrals K and $\pi_a F_a$, using the composition theorem based on the canonical relations, see [36, Theorem 2.4.1] or [38]. Let $((\mathbf{z}, \boldsymbol{\theta}), (\mathbf{x}, \boldsymbol{\zeta})) \in \Upsilon \circ \Lambda$ then there exist a $(\mathbf{y}, \boldsymbol{\eta}) \in T^*Y \setminus 0$ 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 (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 $K\pi_a F_a$ is an FIO with canonical relation contained in the identity. The operators H and P_a are pseudodifferential operators, and π_a and P_a can be constructed such that $\mathrm{WF}(P_a w) \subset \mathrm{WF}(\pi_a w)$ for all w. The conclusion is that $\tilde{R}_a = KHcP_aF_a$ is an FIO with identity canonical relation, and hence a pseudodifferential operator.

(ii) We show that $\tilde{R} = KHV_1PF$ is a pseudodifferential operator that can be written as the integral (112) below. For F we use the local expressions (46). Because P is a t-family of pseudodifferential operators and $F_a\rho$ is a t-family of FIOs, the composition $P_a(t)F_a(t)\rho$ is an FIO with phase inherited from $F_a(t)\rho$, i.e. φ_T. The highest order contribution to its amplitude is $p_a(t; \mathbf{y}, \partial_{\mathbf{y}}\varphi_T)\mathrm{A}_F$. The composition with H can be done similarly, because $P_a(t)F_a(t)\rho$ can also be viewed as an 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 $\mathrm{A}_{\mathrm{HPF}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi})$. It can be written in the form

$$\mathrm{A}_{\mathrm{HPF}}(\mathbf{y}, t_1, \mathbf{x}, \boldsymbol{\xi}) = \left(1 + \frac{c(\mathbf{y})\mathbf{n}_s(\mathbf{y}) \cdot \partial_{\mathbf{y}}\alpha}{\partial_t \alpha}\right) \frac{2\mathrm{i}c(\mathbf{y})p_a(t_1; \mathbf{y}, \partial_{\mathbf{y}}\alpha)aA_s(\mathbf{x})\partial_t\alpha}{c(\mathbf{x})A_s(\mathbf{y})}. \tag{99}$$

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 an FIO with canonical relation contained in the identity. This implies that, to prove the theorem, it is sufficient to do a local analysis using (46). The local analysis shows again that $\tilde{R}_a$ is a pseudodifferential 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

$$\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \alpha(\mathbf{z}, T_s(\mathbf{z}) - T_s(\mathbf{x}); \boldsymbol{\xi}) - \boldsymbol{\xi} \cdot \mathbf{x}. \tag{100}$$

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

$$\partial_{\boldsymbol{\xi}}\alpha(\mathbf{z}, T_s(\mathbf{z}) - T_s(\mathbf{x}); \boldsymbol{\xi}) = \mathbf{x}. \tag{101}$$

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 $\chi : Z \times X \times \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$ 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 $\mathrm{supp}(\chi)$. 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

$$\tilde{R}_a r(\mathbf{z}) = (2\pi)^{-n} \iint \mathrm{e}^{\mathrm{i}\psi(\mathbf{z},\mathbf{x},\boldsymbol{\xi})} \chi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \mathrm{A}_{\mathrm{HPF}}(\mathbf{z}, T_s(\mathbf{z}), \mathbf{x}, \boldsymbol{\xi})\, d\boldsymbol{\xi}\, r(\mathbf{x})\, d\mathbf{x}, \tag{102}$$

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

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

$$\psi(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) = \psi(\mathbf{z}, \mathbf{z}, \boldsymbol{\xi}) + \int_0^1 \partial_\mu[\psi(\mathbf{z}, \tilde{\mathbf{x}}(\mu), \boldsymbol{\xi})]\, d\mu = \boldsymbol{\theta}(\mathbf{z}, \mathbf{x}, \boldsymbol{\xi}) \cdot (\mathbf{z} - \mathbf{x}). \tag{103}$$

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

$$\boldsymbol{\theta}(\mathbf{z},\mathbf{x},\boldsymbol{\xi}) = \boldsymbol{\xi} + \int_0^1 \partial_t\alpha\big(\mathbf{z}, T_\mathrm{s}(\mathbf{z}) - T_\mathrm{s}(\tilde{\mathbf{x}}); \boldsymbol{\xi}\big)\partial_\mathbf{x} T_\mathrm{s}(\tilde{\mathbf{x}})\,d\mu.$$

There exists an $\check{\mathbf{x}}$ such that $(\mathbf{z},\check{\mathbf{x}},\boldsymbol{\xi}) \in \Gamma_{T_\mathrm{s}(\mathbf{z})-T_\mathrm{s}(\tilde{\mathbf{x}})}$, i.e. $\check{\mathbf{x}}$ and $\mathbf{z}$ are connected by a ray. Note that $\check{\mathbf{x}} = \mathbf{x}_\Gamma(\mathbf{z}, T_\mathrm{s}(\mathbf{z}) - T_\mathrm{s}(\tilde{\mathbf{x}}), \boldsymbol{\xi})$ will do, see Section 3.4 for notation $\mathbf{x}_\Gamma$. By using the identities $\partial_t\alpha = -c(\check{\mathbf{x}})|\boldsymbol{\xi}|$ and $c(\tilde{\mathbf{x}})\partial_\mathbf{x} T_\mathrm{s}(\tilde{\mathbf{x}}) = \mathbf{n}_\mathrm{s}(\tilde{\mathbf{x}})$, one gets

$$\boldsymbol{\theta}(\mathbf{z},\mathbf{x},\boldsymbol{\xi}) = \boldsymbol{\xi} - |\boldsymbol{\xi}|\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi}) \quad \text{with } \mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi}) = \int_0^1 \frac{c(\check{\mathbf{x}})}{c(\tilde{\mathbf{x}})}\mathbf{n}_\mathrm{s}(\tilde{\mathbf{x}})\,d\mu. \tag{104}$$

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

$$|\partial_{\boldsymbol{\xi}}\boldsymbol{\theta}| = \left|\det\left(\mathrm{I}_n - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\otimes\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi})\right)\right| = \left|1 - \frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\cdot\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi})\right|. \tag{105}$$

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

$$\big|\boldsymbol{\xi}\cdot\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi})\big| < |\boldsymbol{\xi}|\min\big\{1, \big|\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi})\big|^2\big\}. \tag{106}$$

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 proofs its injectivity. A second argument concerns the domain of integration $\boldsymbol{\theta}(\mathrm{supp}(\chi))$. 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 $\mathrm{supp}(\chi)$, which is nontrivial if $|\mathbf{n}| < 1$. This fact will play a role in gluing $\tilde{R}_\mathrm{a}r$ and $\tilde{R}_\mathrm{b}r$ together, which will be done in the following paragraphs. Because $\mathbf{x}$ is in the neighborhood of $\mathbf{z}$, so are $\tilde{\mathbf{x}}$ and $\check{\mathbf{x}}$. This implies that $\mathbf{n}(\mathbf{z},\mathbf{x},\boldsymbol{\xi})$ is close to $\mathbf{n}_\mathrm{s}(\mathbf{z})$, 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 $\tilde{R}_\mathrm{a}r$'s integral expression (102) transforms into

$$\tilde{R}_\mathrm{a}r(\mathbf{z}) = (2\pi)^{-n}\iint_{\boldsymbol{\theta}(\mathrm{supp}(\chi))} \mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z},\mathbf{x},\boldsymbol{\theta})\mathrm{e}^{\mathrm{i}\boldsymbol{\theta}\cdot(\mathbf{z}-\mathbf{x})}\,d\boldsymbol{\theta}\,r(\mathbf{x})\,d\mathbf{x}, \tag{107}$$

where we define

$$\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = |\partial_{\boldsymbol{\xi}}\boldsymbol{\theta}|^{-1}\chi(\mathbf{z},\mathbf{x},\boldsymbol{\xi})\mathrm{A}_\mathrm{HPF}\big(\mathbf{z}, T_\mathrm{s}(\mathbf{z}), \mathbf{x}, \boldsymbol{\xi}\big). \tag{108}$$

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

(iii) While the expression (107) defines a pseudodifferential operator of order 0, it is given in a non-standard from. It differs from a regular pseudodifferential operator, because the amplitude $\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z},\mathbf{x},\boldsymbol{\theta})$ 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

$$\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z},\mathbf{z},\boldsymbol{\theta}) + \sum_{k=1}^n\int_0^1 D_{\theta_k}\partial_{x_k}\mathrm{A}_{\tilde{\mathrm{R}}}\big(\mathbf{z},\mathbf{z}+\mu(\mathbf{x}-\mathbf{z}),\boldsymbol{\theta}\big)\,d\mu, \tag{109}$$

which is an application of formulae (4.8)–(4.10) of Treves [37]. The first term is the principal symbol of $\tilde{R}_\mathrm{a}$, which has order 0. The second term in (109) does not contribute to the principal part, it corresponds to a pseudodifferential operator of order -1. We will denote by $\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z},\boldsymbol{\theta})$ (with two arguments) the symbol of $\tilde{R}$.

To evaluate of A_HPF (99) on the diagonal one applies (16), the relation $\partial_t\alpha(\mathbf{z},0;\boldsymbol{\xi}) = -c(\mathbf{z})|\boldsymbol{\xi}|$ for the phase and the result $a(\mathbf{z},0;\boldsymbol{\xi}) = \frac{\mathrm{i}}{2c(\mathbf{z})|\boldsymbol{\xi}|}$ for the amplitude. This yields

$$\mathrm{A}_\mathrm{HPF}\big(\mathbf{z}, T_\mathrm{s}(\mathbf{z}), \mathbf{z}, \boldsymbol{\xi}\big) = \left(1 - \frac{\mathbf{n}_\mathrm{s}(\mathbf{z})\cdot\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\right)p_\mathrm{a}\big(T_\mathrm{s}(\mathbf{z});\mathbf{z},\boldsymbol{\xi}\big) = |\partial_{\boldsymbol{\xi}}\boldsymbol{\theta}|\,p_\mathrm{a}\big(T_\mathrm{s}(\mathbf{z});\mathbf{z},\boldsymbol{\xi}\big), \tag{110}$$

see also (105). In view of (108)–(110), we have $\text{p.s.}(\mathrm{A}_{\tilde{\mathrm{R}}})(\mathbf{z}, \boldsymbol{\theta}) = \chi(\mathbf{z}, \mathbf{z}, \boldsymbol{\xi}) p_{\mathrm{a}}(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi})$. Note that $\boldsymbol{\eta} = \partial_{\mathbf{y}}\alpha = \boldsymbol{\xi}$ 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 $\text{supp}(p_{\mathrm{a}}(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}))$ 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

$$\text{p.s.}(\mathrm{A}_{\tilde{\mathrm{R}}})(\mathbf{z}, \boldsymbol{\theta}) = p_{\mathrm{a}}\big(T_{\mathrm{s}}(\mathbf{z}); \mathbf{z}, \boldsymbol{\xi}(\boldsymbol{\theta})\big). \tag{111}$$

(iv) A key step is the inclusion of both negative and positive frequencies. In Section 3 we saw that $a(\mathbf{x}, t; \boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\lambda\alpha(\mathbf{x},t;\boldsymbol{\xi})}$ and $b(\mathbf{x}, t; -\boldsymbol{\xi})\mathrm{e}^{\mathrm{i}\lambda\beta(\mathbf{x},t;-\boldsymbol{\xi})}$ have a symmetry relation: They yield complex conjugate contributions (note the $-$ signs). The consequences of this property can be traced through this proof. We find that $\tilde{R}_{\mathrm{b}} r(\mathbf{z}) = (2\pi)^{-n} \iint_{-\boldsymbol{\theta}(\text{supp}(\chi))} \overline{\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z}, -\boldsymbol{\theta})} \mathrm{e}^{\mathrm{i}\boldsymbol{\theta}\cdot(\mathbf{z}-\mathbf{x})}\, d\boldsymbol{\theta}\, r(\mathbf{x})\, d\mathbf{x}$, and consequently, modulo a regularizing contribution,

$$\tilde{R} r(\mathbf{z}) = (2\pi)^{-n} \iint \big[\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z}, \boldsymbol{\theta}) + \overline{\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z}, -\boldsymbol{\theta})}\big] \mathrm{e}^{\mathrm{i}\boldsymbol{\theta}\cdot(\mathbf{z}-\mathbf{x})}\, d\boldsymbol{\theta}\, r(\mathbf{x})\, d\mathbf{x}. \tag{112}$$

The $\boldsymbol{\theta}$-integration is over the full space because the definition of $\mathrm{A}_{\tilde{\mathrm{R}}}(\mathbf{z}, \boldsymbol{\theta})$ can be smoothly extended such that it is zero outside the domain $\boldsymbol{\theta}(\text{supp}(\chi))$. In view of (111) this proves the claim. □

Proof of Theorem 4. The first step in deriving the imaging condition is to rewrite operators H, K and G (95). 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 $w(\mathbf{x}, t) = \frac{1}{2\pi} \int \mathrm{e}^{\mathrm{i}\omega t} \hat{w}(\mathbf{x}, \omega)\, d\omega$, one has

$$\begin{aligned} \widehat{Hw}(\mathbf{x}, \omega) &= \frac{\tilde{\sigma}(\omega)}{A_{\mathrm{s}}(\mathbf{x})} (\mathrm{i}\omega)^{-\frac{n+1}{2}} \big[\mathrm{i}\omega + c(\mathbf{x})\mathbf{n}_{\mathrm{s}}(\mathbf{x}) \cdot \partial_{\mathbf{x}}\big] \hat{w}(\mathbf{x}, \omega), \\ Kw(\mathbf{x}) &= \frac{1}{2\pi} \int \mathrm{e}^{\mathrm{i}\omega T_{\mathrm{s}}(\mathbf{x})} \hat{w}(\mathbf{x}, \omega)\, d\omega. \end{aligned} \tag{113}$$

Applied to the reverse time continued wave field $u_{\mathrm{r}}(\mathbf{x}, \omega)$, Eq. (96) becomes

$$\tilde{R} r(\mathbf{x}) = \frac{1}{2\pi} \int \mathrm{e}^{\mathrm{i}\omega T_{\mathrm{s}}(\mathbf{x})} \frac{\tilde{\sigma}(\omega)}{A_{\mathrm{s}}(\mathbf{x})} (\mathrm{i}\omega)^{-\frac{n+1}{2}} \big[\mathrm{i}\omega + c(\mathbf{x})\mathbf{n}_{\mathrm{s}}(\mathbf{x}) \cdot \partial_{\mathbf{x}}\big] \hat{u}_{\mathrm{r}}(\mathbf{x}, \omega)\, d\omega. \tag{114}$$

The next step is to eliminate $T_{\mathrm{s}}(\mathbf{x})$, $A_{\mathrm{s}}(\mathbf{x})$ and $\mathbf{n}_{\mathrm{s}}(\mathbf{x})$ by expressing them in terms of the source field explicitly. The principal term of the geometrical optics approximation of the source (39) is

$$\hat{g}(\mathbf{x}, \omega) = A_{\mathrm{s}}(\mathbf{x}) \sigma(\omega) (\mathrm{i}\omega)^{\frac{n-3}{2}} \mathrm{e}^{-\mathrm{i}\omega T_{\mathrm{s}}(\mathbf{x})}.$$

Function σ, introduced in (39), 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:

$$\begin{aligned} \mathrm{e}^{\mathrm{i}\omega T_{\mathrm{s}}(\mathbf{x})} \frac{1}{A_{\mathrm{s}}(\mathbf{x})} (\mathrm{i}\omega)^{-\frac{n+1}{2}} &= \frac{\sigma(\omega)}{(\mathrm{i}\omega)^2 \hat{g}(\mathbf{x}, \omega)}, \\ c(\mathbf{x})\mathbf{n}_{\mathrm{s}}(\mathbf{x}) = \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \hat{g}(\mathbf{x}, \omega)}{-\mathrm{i}\omega \hat{g}(\mathbf{x}, \omega)} &= \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x}, \omega)}}{\mathrm{i}\omega \overline{\hat{g}(\mathbf{x}, \omega)}}, \end{aligned} \tag{115}$$

in which it is used that the second equation is real-valued. Substitution of involved factors occurring in the integral (114) yields

$$\tilde{R} r(\mathbf{x}) = \frac{1}{2\pi} \int \frac{\tilde{\sigma}(\omega)\sigma(\omega)}{\mathrm{i}\omega \hat{g}(\mathbf{x}, \omega)} \left[1 + \frac{c(\mathbf{x})^2 \partial_{\mathbf{x}} \overline{\hat{g}(\mathbf{x}, \omega)} \cdot \partial_{\mathbf{x}}}{(\mathrm{i}\omega)^2 \overline{\hat{g}(\mathbf{x}, \omega)}}\right] \hat{u}_{\mathrm{r}}(\mathbf{x}, \omega)\, d\omega. \tag{116}$$

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 (94). The source wave field is free of caustics by assumption. The transport equation yields that, on a compact domain in space–time, there exists a lower bound $L > 0$ for the principal amplitude, thus $|A_0(\mathbf{x}, \mathbf{x}_{\mathrm{s}}, \omega)| \geqslant L$. Division by A_0 is therefore well-defined, and from its homogeneity and the inequality (38) it can be deduced that there exists a constant $C > 0$ such that

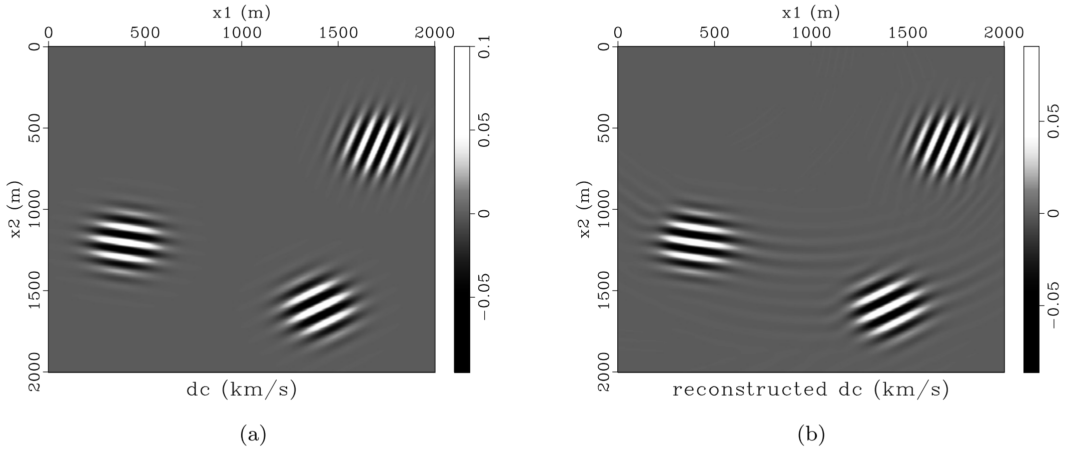

Fig. 3. Example 1: Velocity perturbation and reconstructed velocity perturbation. The background medium is a gradient $c = 2.0 + 0.001x_2$, with x_2 in meters and c in km/s.

$|\frac{A(\mathbf{x},\mathbf{x}_s,\omega)}{A_0(\mathbf{x},\mathbf{x}_s,\omega)} - 1| \leqslant \frac{C}{1+|\omega|}$. For $|\omega|$ sufficiently large, division by A is therefore well-defined. We choose $1 - \Omega$ wide enough such that all $\omega \in \operatorname{supp}(\Omega)$ 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 $\operatorname{supp}(\Omega)$. Taking (116) we replace $\tilde{\sigma}\sigma$ with Ω to define the imaging condition (92). □

7. 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 [44]. The direct wave was eliminated. The operator (7) 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 (92) was applied to obtain an approximate reconstruction of δc.

As we mentioned, only a partial reconstruction of δc is possible in realistic situations. Relation (52) and the wave propagation restrict the directions of $\boldsymbol{\zeta}$ where inversion is possible. The frequency range present in the data also restricts the length of $\boldsymbol{\zeta}$, according to (52) and using that $|\boldsymbol{\xi}| = c^{-1}|\omega|$. 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 gradient type medium with $c(x_1, x_2) = 2.0 + 0.001x_2$ 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 Figs. 3 and 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. 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 medium including some rays, as well

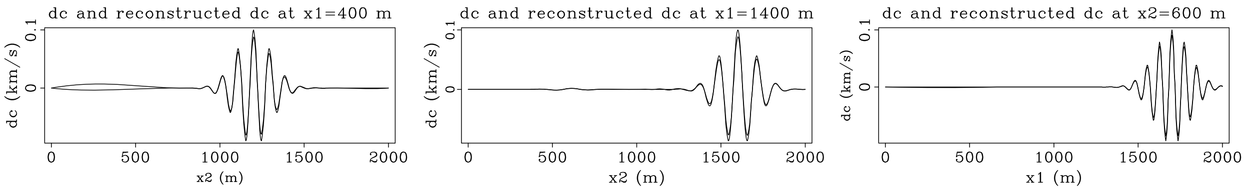

Fig. 4. Example 1: Comparison of some traces from Fig. 3 at $x_1 = 400$ m, $x_1 = 1400$ m and $x_2 = 600$ m.

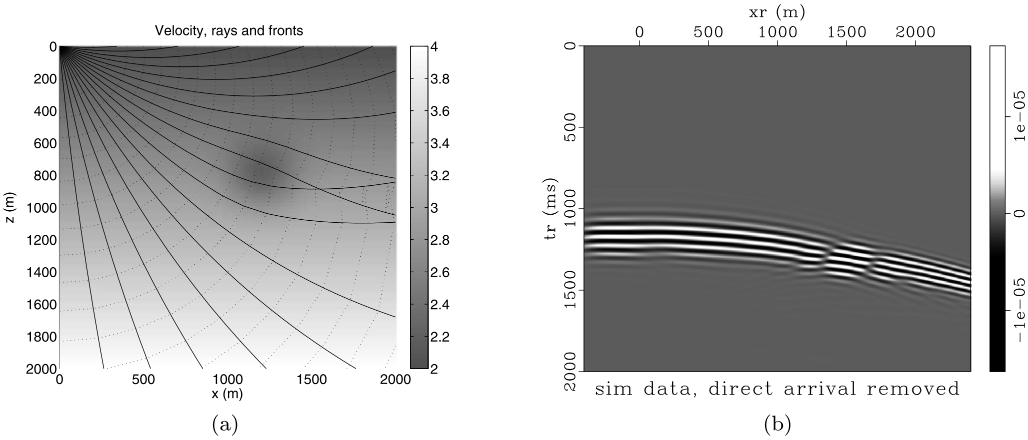

Fig. 5. Example 2: (a) A velocity model with some rays; (b) Simulated data, with direct arrival removed.

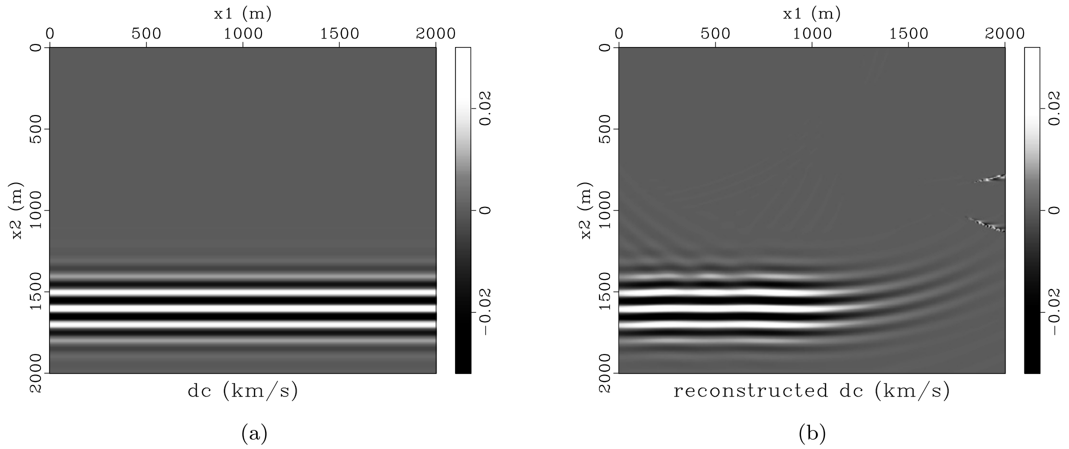

Fig. 6. Example 2: (a) Velocity perturbation; (b) Partial reconstruction of the velocity perturbation.

as some data are plotted in Fig. 5. The velocity perturbation was located at $x_2 = 1600$ m. The results of this example are given in Fig. 6. 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 $\hat{u}_{\text{inc}}^{-1}$, around $(x_1, x_2) = (1900, 1000)$ m, which leads to the two artifacts that can be seen there.

8. Discussion

We presented a comprehensive analysis of RTM-based imaging, and introduced an imaging condition involving only local (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 [29–33]. Our imaging condition yields a parametrix and naturally avoids this singularity. (ii) The square-root operator (7), a factor of F_M introduced in Section 5, can be removed with dual sensor (streamer) data, that is, if the surface-normal derivative of the wave field is measured. We note that F_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 [45,25,46]. 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 be beneficial. (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 [23,2].)

The acquisition aperture, and associated illumination, is intimately connected to the resolution operator R. This operator is pseudodifferential and the support of its symbol expresses which parts of the contrast or reflectivity can be recovered 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 [47]; see also [48].

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 results in reconstructions at different sets of points and orientations, in general, which are not identified within the RTM algorithm, averaging must be avoided. However, techniques from microlocal analysis can be invoked to properly exploit the discrete multi-source acquisition geometry. A sampling type theorem in this context has yet to be developed. (We note that in the case of open sets of sources the generation of source caustics will be allowed.)

Acknowledgements

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.

References

[1] J.F. Claerbout, Imaging the Earth's Interior, Blackwell Scientific Publications, Inc., 1985.
[2] B.L. Biondi, 3D Seismic Imaging, Society of Exploration Geophysicists, 2006.
[3] W.W. Symes, Topical review: The seismic reflection inverse problem, Inverse Problems 25 (12) (2009) 123008.
[4] N. Bleistein, J.K. Cohen, J.J.W. Stockwell, Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion, Springer-Verlag, New York, 2001.
[5] P.S. Schultz, J.W.C. Sherwood, Depth migration before stack, Geophysics 45 (3) (1980) 376–393, http://dx.doi.org/10.1190/1.1441088, http://link.aip.org/link/?GPY/45/376/1.
[6] D. Whitmore, Iterative depth migration by backward time propagation, in: Expanded Abstracts, Society of Exploration Geophysicists, 1983, pp. 382–385.
[7] G.A. McMechan, Migration by extrapolation of time-dependent boundary values, Geophys. Prosp. 31 (1983) 413–420.
[8] K. Baysal, D.D. Kosloff, J.W.C. Sherwood, Reverse time migration, Geophysics 48 (1983) 1514–1524.
[9] R. Sun, G.A. McMechan, Scalar reverse-time depth migration of prestack elastic seismic data, Geophysics 66 (2001) 1519–1527.
[10] R. Clayton, Common midpoint migration: Stanford Expl. Proj., Rep. No. 14, Stanford University, 1978.
[11] B. Biondi, G. Palacharla, 3-d prestack migration of common-azimuth data, Geophysics 61 (6) (1996) 1822–1832, http://dx.doi.org/10.1190/1.1444098, http://link.aip.org/link/?GPY/61/1822/1.
[12] A.M. Popovici, Prestack migration by split-step dsr, Geophysics 61 (5) (1996) 1412–1416, http://dx.doi.org/10.1190/1.1444065, http://link.aip.org/link/?GPY/61/1412/1.
[13] S. Jin, C.C. Mosher, R.-S. Wu, Offset-domain pseudoscreen prestack depth migration, Geophysics 67 (6) (2002) 1895–1902, http://dx.doi.org/10.1190/1.1527089, http://link.aip.org/link/?GPY/67/1895/1.
[14] G. Beylkin, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, J. Math. Phys. 26 (1) (1985) 99–108.
[15] Rakesh, A linearised inverse problem for the wave equation, Comm. Partial Differential Equations 13 (5) (1988) 573–601.
[16] A.P.E. Ten Kroode, D.J. Smit, A.R. Verdel, A microlocal analysis of migration, Wave Motion 28 (1998) 149–172.

[17] C.C. Stolk, Microlocal analysis of a seismic linearized inverse problem, Wave Motion 32 (2000) 267–290.
[18] C.C. Stolk, M.V. De Hoop, Microlocal analysis of seismic inverse scattering in anisotropic elastic media, Comm. Pure Appl. Math. 55 (2002) 261–301.
[19] C.C. Stolk, M.V. De Hoop, Modeling of seismic data in the downward continuation approach, SIAM J. Appl. Math. 65 (2005) 1388–1406.
[20] C.C. Stolk, M.V. De Hoop, Seismic inverse scattering in the downward continuation approach, Wave Motion 43 (2006) 579–598.
[21] V. Guillemin, On some results of Gelfand integral geometry, in: Pseudodifferential Operators and Applications, in: Proc. Sympos. Pure Math., vol. 43, Amer. Math. Soc., Providence, RI, 1985, pp. 149–155.
[22] C.J. Nolan, W.W. Symes, Global solution of a linearized inverse problem for the wave equation, Comm. Partial Differential Equations 22 (5) (1997) 127–149.
[23] J.F. Claerbout, Toward a unified theory of reflector mapping, Geophysics 36 (3) (1971) 467–481.
[24] W.F. Chang, G.A. McMechan, Reverse-time migration of offset vertical seismic profiling data using the excitation-time imaging condition, Geophysics 51 (1) (1986) 67–84.
[25] S. Chattopadhyay, G.A. McMechan, Imaging conditions for prestack reverse-time migration, Geophysics 73 (3) (2008) S81–S89.
[26] D. Kiyashchenko, R.-E. Plessix, B. Kashtan, V. Troyan, A modified imaging principle for true-amplitude wave-equation migration, Geophys. J. Int. 168 (3) (2007) 1093–1104.
[27] W.A. Schneider, Integral formulation for migration in two and three dimensions, Geophysics 43 (1978) 49–76.
[28] N.N. Bojarski, A survey of the near-field far-field inverse scattering inverse source integral equation, IEEE Trans. Antennas and Propagation AP-30 (1982) 975–979.
[29] K. Yoon, K. Marfurt, W. Starr, Challenges in reverse-time migration, in: Expanded Abstracts, Society of Exploration Geophysicists, 2004, pp. 1057–1060.
[30] W.A. Mulder, R.E. Plessix, A comparison between one-way and two-way wave-equation migration, Geophysics 69 (2004) 1491–1504.
[31] R.F. Fletcher, P. Fowler, P. Kitchenside, U. Albertin, Suppressing artifacts in prestack reverse time migration, in: Expanded Abstracts, Society of Exploration Geophysicists, 2005, pp. 2049–2052.
[32] X.-B. Xie, R.-S. Wu, A depth migration method based on the full-wave reverse-time calculation and local one-way propagation, SEG Techn. Progr. Exp. Abstr. 25 (1) (2006) 2333–2337.
[33] A. Guitton, B. Kaelin, B. Biondi, Least-square attenuation of reverse-time migration artifacts, Geophysics 72 (2007) S19–S23.
[34] L. Hörmander, The Analysis of Linear Partial Differential Operators III, Springer-Verlag, 1994.
[35] L.C. Evans, Partial Differential Equations, Amer. Math. Soc., 1998.
[36] J.J. Duistermaat, Fourier Integral Operators, Birkhäuser, 1996.
[37] F. Treves, Introduction to Pseudodifferential and Fourier Integral Operators, vol. I, Plenum Press, New York, 1980.
[38] F. Treves, Introduction to Pseudodifferential and Fourier Integral Operators, vol. II, Plenum Press, New York, 1980.
[39] A. Grigis, J. Sjöstrand, Microlocal Analysis for Differential Operators. An Introduction, Cambridge University Press, 1994.
[40] L. Hörmander, The Analysis of Linear Partial Differential Operators I, Springer-Verlag, 1990.
[41] M.E. Taylor, Pseudodifferential Operators, Princeton University Press, 1981.
[42] T.J.P.M. Op 't Root, C.C. Stolk, One-way wave propagation with amplitude based on pseudo-differential operators, Wave Motion 47 (2) (2010) 67–84, http://dx.doi.org/10.1016/j.wavemoti.2009.08.001.
[43] J.B. Conway, A Course in Functional Analysis, second ed., Springer-Verlag, 1990.
[44] G.C. Cohen, Higher-Order Numerical Methods for Transient Wave Equations, Sci. Comput., Springer-Verlag, Berlin, 2002, with a foreword by R. Glowinski.
[45] A. Guitton, A. Valenciano, D. Bevc, Robust imaging condition for shot-profile migration, SEG Techn. Progr. Exp. Abstr. 25 (1) (2006) 2519–2523.
[46] J.C. Costa, F.A.S. Neto, M.R.M. Alcântara, J. Schleicher, A. Novais, Obliquity-correction imaging condition for reverse time migration, Geophysics 74 (3) (2009) S57–S66.
[47] M.V. de Hoop, H. Smith, G. Uhlmann, R.D. van der Hilst, Seismic imaging with the generalized Radon transform: A curvelet transform perspective, Inverse Problems 25 (2) (2009) 025005, http://dx.doi.org/10.1088/0266-5611/25/2/025005, 21 pp.
[48] F.J. Herrmann, P.P. Moghaddam, C.C. Stolk, Sparsity- and continuity-promoting seismic image recovery with curvelet frames, Appl. Comput. Harmon. Anal. 24 (2) (2008) 150–173, http://dx.doi.org/10.1016/j.acha.2007.06.007.