SEISMIC INVERSE SCATTERING IN THE 'WAVE-EQUATION' APPROACH

CHRISTIAAN C. STOLK AND MAARTEN V. DE HOOP

ABSTRACT. In this paper we use methods from partial differential equations, in particular microlocal analysis and theory of Fourier integral operators, to study seismic imaging in the wave-equation approach. Seismic data are commonly modeled by a high-frequency single scattering approximation. This amounts to a linearization in the medium coefficient about a smooth background. The discontinuities are contained in the medium perturbation. The wave solutions in the background medium admit a geometrical optics representation. Here we describe the wave propagation in the background medium by a one-way wave equation. Based on this we derive the double-square-root equation, which is a first order pseudodifferential equation, that describes the continuation of seismic data in depth. We consider the modeling operator, its adjoint and reconstruction based on this equation. If the rays in the background that are associated with the reflections due to the perturbation are nowhere horizontal, the singular part of the data is described by the solution to an inhomogeneous double-square-root equation. We derive a microlocal reconstruction equation. The main result is a characterization of the angle transform that generates the common image point gathers, and a proof that this transform contains no artifacts. Finally, pseudodifferential annihilators based on the double-square-root equation are constructed. The double-square-root equation approach is used in seismic data processing.

Keywords: Seismic inversion, microlocal analysis, double-square-root equation.
AMS Subject Classification: 86A15, 35R30.

1. INTRODUCTION

In reflection seismology one places point sources and point receivers on the earth's surface. The source generates acoustic waves in the subsurface, that are reflected where the medium properties vary discontinuously. The recorded reflections that can be observed in the data are used to reconstruct these discontinuities.

The data are commonly modeled by a high-frequency single scattering approximation. In the sequel, we distinguish the vertical coordinate $z \in \mathbb{R}$ from the horizontal coordinates $x \in \mathbb{R}^{n-1}$ and write $(x, z) \in \mathbb{R}^n$. In these coordinates the scalar acoustic wave equation is given by

$$Pu = f \ , \quad P = c(x,z)^{-2} \frac{\partial}{\partial t}^2 + \sum_{j=1}^{n-1} D_{x_j}^2 + D_z^2 \ , \tag{1}$$

where $D_x = -\mathrm{i}\frac{\partial}{\partial x}, D_z = -\mathrm{i}\frac{\partial}{\partial z}$. The equation is considered for $(x, z) \in \mathbb{R}^n$, t in a time interval $]0, T[$. We do not impose the usual free-surface boundary condition at $z = 0$. It is straightforward to remove such a condition in a microlocal representation of the solution,

The authors thank The Mathematical Sciences Research Institute for partial support, through NSF grant DMS-9810361. They also thank the members of the institute, and in particular Gunther Uhlmann, for providing a very stimulating environment during the Inverse Problems program in Fall 2001.

in our setting. By Duhamel's principle, a causal solution operator for the inhomogeneous equation (1) is given by

$$u(x,z,t) = \int_0^t \int G(x,z,t-t_0,x_0,z_0) f(x_0,z_0,t_0)\,\mathrm{d}x_0\mathrm{d}z_0\mathrm{d}t_0, \tag{2}$$

where (when the coefficient c is C^∞) G is a Fourier integral operator with canonical relation that is essentially a union of bicharacteristics. The source f is a distribution in $\mathcal{E}'(X_0\times]0,T[)$ where X_0 is a bounded open subset of $\mathbb{R}^n$. The kernel of the Fourier integral operator can be written as a sum of contributions

$$G(x,z,t,x_0,z_0) = \sum_{j\in J} \int_{\mathbb{R}^{N(i)}} a^{(j)}(x,z,t,x_0,z_0,\theta)\, \exp[\mathrm{i}\phi^{(j)}(x,z,x_0,z_0,t,\theta)]\,\mathrm{d}\theta, \tag{3}$$

where the $\phi^{(j)}$ are non-degenerate phase functions and the $a^{(j)}$ suitable symbols, see [8, (4.2.1) and chapter 5]; J is a finite set.

We adopt the linearized scattering approximation. The linearization is in the coefficient c around a smooth background c_0, $c = c_0 + \delta c$. The perturbation δc may contain singularities. The perturbation in G at the acquisition surface $z = 0$ is given by

$$\delta G(r,0,t,s,0) = \int_{\mathbb{R}^{n-1}\times\mathbb{R}_+} \int_0^t G(r,0,t-t_0,x_0,z_0)\, 2c_0^{-3}(x_0,z_0)\delta c(x_0,z_0) \\ \partial_{t_0}^2 G(x_0,z_0,t_0,s,0)\,\mathrm{d}t_0\mathrm{d}x_0\mathrm{d}z_0, \tag{4}$$

where both $s, r \in \mathbb{R}^{n-1}$. We assume that the acquisition manifold Y, that contains the set of values of (s,r,t) used in the acquisition, is a bounded open subset of $\mathbb{R}^{2n-2}\times\mathbb{R}_+$. Except for Remark 6.3 we assume that no further restrictions apply to Y; the acquisition geometry is maximal. Let $\mathcal{R}_0$ be the restriction to Y defined by

$$\mathcal{R}_0 \,:\, \mathcal{D}'_\Gamma(\mathbb{R}^{2n+1}) \to \mathcal{D}'(Y)\ ,\ u(x,z,t,x_0,z_0) \mapsto (\mathcal{R}_0 u)(x_0,x,t) = u(x,0,t,x_0,0).$$

The modeling map F is defined as the map from δc to δG restricted to Y

$$F = F[c_0] \,:\, \delta c \mapsto \mathcal{R}_0 \delta G.$$

One can show that this restriction is well defined. Since Y is bounded and the waves propagate with finite speed we may assume that δc is supported in a bounded open subset X of $\mathbb{R}^{n-1}\times\mathbb{R}_+$. We assume that $\overline{X}\cap\{z=0\} = \emptyset$.

In the seismic inverse problem, both the smooth background c_0 and the perturbation δc are unknown and have to be reconstructed. This leads to two related inverse problems. First the linear inverse problem of reconstructing δc given c_0. A solution operator for this problem is used in the determination of c_0, the second problem. This is explained in more detail below. It is noted that, in general, c_0 is not uniquely determined by the singular part of the data. The results presented in this paper are of particular importance to the second problem.

The reconstruction of the perturbation δc given the background c_0 is essentially done by applying the adjoint of the above mentioned linear map (seismic imaging). The modeling operator is a Fourier integral operator [19] (for a general reference of Fourier integral operators see [8]). If the composition of adjoint and modeling operators, the normal operator, is pseudodifferential, then the positions of the singularities of the perturbation are recovered by applying the adjoint to the data. By computing explicitly the symbol of the normal operator, and applying its inverse (or by adding suitable factors to the expression for the adjoint operator) a microlocal reconstruction can be carried out. Under various assumptions on the background, concerning the presence of caustics and the geometry of the rays, results concerning the normal operator have been obtained [1, 10, 16, 14, 20]. These are summarized in section 2. In the Kirchhoff approach an approximation to the adjoint operator is constructed that depends on the background only through travel times and complex amplitudes that appear in the geometrical optics approximation.

The main topic of this paper is the so-called wave-equation approach [4, 3, 17]. In this approach the data are downward continued, leading to data from fictitious experiments below the surface at varying depths. To make this precise we first study the forward map F and show that the singular part of the data can be described by the solution to an inhomogeneous pseudodifferential evolution equation in depth, provided that the rays in the background that are associated with the reflections are nowhere horizontal (section 5).

To form the adjoint a restriction is applied to the downward continued data (imaging condition). Using a one-way wave equation (section 4) for the propagation of waves in the background, the downward continuation is described by the so-called double-square-root equation. In the wave-equation approach to inverse scattering, this equation is solved numerically; it results in a more complete computation of the wave propagation and diffraction in the reconstruction than the one based on the geometrical optics approximation. (However, the downward continuation can also be carried out in the latter approximation; the results of this paper remain valid.) We derive a microlocal reconstruction equation for δc in the wave-equation approach in Theorem 6.1.

To determine the background c_0, the redundancy of the data is exploited. The background model c_0 is acceptable if the data are in the range of the modeling operator $F[c_0]$. There exist invertible Fourier integral operators derived from the modeling operator that map the data to a set of reconstructions that are identical (microlocally), say a p-family of reconstructions where p is in a subset of $\mathbb{R}^{n-1}$ [22]. If $A = A[c_0]$ is such an operator, then $A[c_0]F[c_0]$ is a p-family of pseudodifferential operators, equal to the identity microlocally on a subset of $\mathbb{R}^n \times \mathbb{R}^n \backslash \{0\} \times \mathbb{R}^{n-1}$. The criterion that the data $d \in \mathrm{range}(F[c_0])$ becomes equivalent to

$$(A[c_0]d)(x, z, p) \text{ is independent of } p, \tag{5}$$

microlocally. This is the criterion that is used in the reconstruction of the background (in seismology called migration velocity analysis). The redundancy also leads to the existence of pseudodifferential operators that annihilate the singular part of the data [22].

Under Beylkin's conditions [1], in particular in the absence of caustics, such invertible Fourier integral operators can be constructed by restricting the modeling operator to different subsets of the data. In the presence of caustics the parameter p can be chosen as the angle between in- and out-going rays at the image point. In the Kirchhoff approach it has been proposed to use a generalized Radon transform to generate such a set [2, 27]. However, in the presence of caustics artifacts were observed in numerical examples in [2]. By microlocal analysis of the Kirchhoff approach the presence of artifacts was shown in [21].

In section 3 an approach to suppress these artifacts is discussed, which requires knowledge of the background, however.

The main result of this paper is the introduction of invertible Fourier integral operators of the type described above derived from the modeling operator in the wave equation approach. In Theorem 7.1 and Proposition 7.2 an angle transform $\widetilde{A}_{\mathrm{WE}}$ is introduced for this purpose. For each p a reconstruction equation of the type

$$\Psi_{\mathrm{WE}}(x, z, p, D_x, D_z, 0)\delta c = \widetilde{A}_{\mathrm{WE}} d,$$

where Ψ_{WE} represents a p-family of pseudodifferential operators, is obtained. This angle transform does not generate artifacts in the presence of caustics. Hence with this angle transform migration velocity analysis (i.e. estimating c_0 using (5)) is extended to allow for media with caustics. Annihilators derived from $\widetilde{A}_{\mathrm{WE}}$ and based on the double-square-root equation are constructed in Theorem 8.1.

2. HIGH-FREQUENCY BORN MODELING AND IMAGING

In this section we summarize some results in the literature about the modeling map and the reconstruction of the perturbation δc given the background c_0. The solution operator (3) propagates singularities along bicharacteristics. These are the solutions of a Hamilton system with Hamiltonian given by the principal symbol of P

$$P(x, z, \xi, \zeta, \tau) = -c(x, z)^{-2}\tau^2 + \|\xi\|^2 + \zeta^2.$$

The Hamilton system is given by

$$\frac{\mathrm{d}(x, z, t)}{\mathrm{d}\lambda} = \frac{\partial P}{\partial(\xi, \zeta, \tau)}, \quad \frac{\mathrm{d}(\xi, \zeta, \tau)}{\mathrm{d}\lambda} = -\frac{\partial P}{\partial(x, z, t)}. \tag{6}$$

Its solutions will be parameterized by initial position (x_0, z_0), take-off direction $\alpha \in S^{n-1}$, frequency τ and time t,

$$\boldsymbol{x} = \boldsymbol{x}(x_0, z_0, \alpha, \tau, t)$$

and similarly for $\boldsymbol{z}, \boldsymbol{t}, \boldsymbol{\xi}, \boldsymbol{\zeta}$; τ is invariant along the Hamilton flow. Here the evolution parameter is the time t. In section 4 we will change the evolution parameter to z, and use a similar notation to denote the bicharacteristics parameterized by z.

To ensure that δG defines a continuous map from $\mathcal{E}'(X)$ to $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n \times]0, T[)$, and that the restriction of δG to Y is a Fourier integral operator we invoke the following assumption on c_0

Assumption 1. *There are no rays from $(s, 0)$ to $(r, 0)$ with travel time t such that $(s, r, t) \in Y$. For all ray pairs connecting $(r, 0)$ via some $(x, z) \in X$ to $(s, 0)$ with total time t such that $(s, r, t) \in Y$, the rays intersect the plane $z = 0$ transversally at r and s.*

The first part of the assumption excludes direct rays, i.e. rays that scatter over π; the second part of the assumption excludes rays grazing the plane $z = 0$. Strictly, only caustics grazing the plane $z = 0$ have to be excluded.

Theorem 2.1. [19, 14] *With Assumption 1 the map F is a Fourier integral operator $\mathcal{D}'(X) \to \mathcal{D}'(Y)$ of order $(n-1)/4$ with canonical relation*

(7)
$$\big\{(\boldsymbol{x}(x,z,\beta,\tau,t_{\rm s}),\boldsymbol{x}(x,z,\alpha,\tau,t_{\rm r}),t_{\rm s}+t_{\rm r},\boldsymbol{\xi}(x,z,\beta,\tau,t_{\rm s}),\boldsymbol{\xi}(x,z,\alpha,\tau,t_{\rm r}),\tau;x,z,\xi,\zeta)\mid t_{\rm s},t_{\rm r}>0,\ \boldsymbol{z}(x,z,\beta,\tau,t_{\rm s})=\boldsymbol{z}(x,z,\alpha,\tau,t_{\rm r})=0,\ (\xi,\zeta)=-\tau c_0(x,z)^{-1}(\alpha+\beta),\ (x,z,\alpha,\beta,\tau)\in \textit{ subset of } X\times(S^{n-1})^2\times\mathbb{R}\backslash 0\big\}.$$

Assumption 1 is microlocal. One can identify the set of points $(s,r,t,\sigma,\rho,\tau)\in T^*Y\backslash 0$ where this assumption is violated. If the symbol $\psi=\psi(s,r,t,\sigma,\rho,\tau)$ vanishes on a neighborhood of this set, then the composition ψF of the pseudodifferential cutoff $\psi=\psi(s,r,t,D_s,D_r,D_t)$ with F is a Fourier integral operator as in the theorem.

We assume that ψ is as before and in addition vanishes outside Y. To image the singularities of δc from the data we consider the adjoint $F^*\psi$, which is a Fourier integral operator also. Denoting the data by d we arrive at the reconstruction equation

$$F^*\psi F\delta c = F^*\psi d. \tag{8}$$

To ensure that $F^*\psi$ images singularities of δc, which is guaranteed if $F^*\psi F$ is pseudodifferential, we invoke Guillemin's

Assumption 2. [9] *The projection of the canonical relation (7) on $T^*Y\backslash 0$ is an embedding.*

Since (7) is a canonical relation that projects submersively on the subsurface variables (x,z,ξ,ζ), the projection of (7) on $T^*Y\backslash 0$ is immersive [13, 25.3.6]. Therefore only the injectivity in the assumption needs to be verified [14].

The following theorem [14] describes the reconstruction of singularities of δc. In general by (7) not all singularities in δc are mapped to singularities in the data (the projection of (7) on the first (acquisition) variables is not surjective), so that only part of the singularities in δc will be reconstructed. Whether a singularity is reconstructed is determined by the principal symbol of $F^*\psi F$: it is nonzero at (x,z,ξ,ζ) whenever there is a point $(s,r,t,\sigma,\rho,\tau;x,z,\xi,\zeta)$ in the canonical relation (7) with (s,r,t,σ,ρ,τ) in the support of ψ (i.e. whenever there is illumination or insonification).

Theorem 2.2. *With Assumption 2 the operator $F^*\psi F$ is pseudodifferential of order $n-1$.*

3. Generalized Radon transform in scattering angle

Consider the projection of the canonical relation (7) on the (x,z,s,r,τ) variables. Where this projection is locally diffeomorphic, i.e. away from caustics forming at s or r, the canonical relation (7) can be described by a phase function of the form

$$\tau(T^{(m)}(x,z,s,r)-t)$$

where $T^{(m)}$ is the value of the time variable in (7). There is a set $\{T^{(m)}\}_{m\in M}$, with M a finite set, that describes the canonical relation except for a neighborhood of the subset of the canonical relation where mentioned projection is degenerate. Since the time interval $]0,T[$ is bounded, the set M is finite. Each $T^{(m)}$ is defined on a subset $D^{(m)}$ of $\mathbb{R}^{3n-2}_{(x,z,s,r)}$. We define $F^{(m)}$ to be a contribution to F with phase function given by $\tau(T^{(m)}-t)$, such that on a subset of the canonical relation where the projection is nondegenerate F is given microlocally by $\sum_{m\in M} F^{(m)}$.

We can use $(x, z, \xi, \zeta) \in T^*(\mathbb{R}^{n-1} \times \mathbb{R}_+)\backslash 0$ as the first $2n$ local coordinates on the canonical relation (7). In addition, we need to parameterize the subsets of the canonical relation given by $(x, z, \xi, \zeta) =$ constant; we denote such parameters by e. The canonical relation (7) was parameterized by $(x, z, \alpha, \beta, \tau)$. We relate (x, z, ξ, ζ, e) by a coordinate transformation to $(x, z, \alpha, \beta, \tau)$: A suitable choice when $\alpha \neq \beta$ is the scattering angles given by

$$\boldsymbol{e}(x, z, \alpha, \beta) = \left(\arccos(\alpha \cdot \beta), \frac{-\alpha + \beta}{2\sin(\arccos(\alpha \cdot \beta)/2)}\right) \in]0, \pi[\times S^{n-2}. \tag{9}$$

The migration dip ν is defined as

$$\nu(x, z, \alpha, \beta) = \frac{\alpha + \beta}{\|\alpha + \beta\|} \in S^{n-1}. \tag{10}$$

At (x, z), $(\nu, \boldsymbol{e})$ replace (α, β). The first component of $\boldsymbol{e}$ defines the opening angle between α and β. The second component of $\boldsymbol{e}$ defines, together with ν, the plane spanned by (α, β). Since ν can attain any value in S^{n-1} the opening angle is restricted to $]0, \pi[$. On $D^{(m)}$ there is a map $(x, z, \alpha, \beta) \mapsto (x, z, s, r)$. We define $\boldsymbol{e}^{(m)} = \boldsymbol{e}^{(m)}(x, z, s, r)$ as the composition of $\boldsymbol{e}$ with the inverse of this map. Likewise, we define $\nu^{(m)} = \nu^{(m)}(x, z, s, r)$.

We define the generalized Radon transform in scattering angle or the Kirchhoff angle transform via a restriction in F^* of the mapping $\boldsymbol{e}^{(m)}$ to a prescribed value e, i.e. the distribution kernel of each contribution $F^{(m)*}$ is multiplied by $\delta(e - \boldsymbol{e}^{(m)}(x, z, s, r))$. Its kernel is given by

$$L(x, z, e, r, s, t) = \sum_{m\in M} (2\pi)^{-(n-1)} \int \overline{A^{(m)}(x, z, s, r, \tau)} \mathrm{e}^{\mathrm{i}\Phi^{(m)}(x,z,e,s,r,t,\varepsilon,\tau)} \,\mathrm{d}\tau \mathrm{d}\varepsilon, \tag{11}$$

where $A^{(m)}$ is a symbol for the m-th contribution to F, supported on $D^{(m)}$, and

$$\Phi^{(m)}(x, z, e, s, r, t, \varepsilon, \tau) = \tau(T^{(m)}(x, z, s, r) - t) + \langle \varepsilon, e - \boldsymbol{e}^{(m)}(x, z, s, r)\rangle. \tag{12}$$

Here, ε is the cotangent vector corresponding to e. Let $\psi_L = \psi_L(D_s, D_r, D_t)$ be a pseudo-differential cutoff with symbol $\psi(\sigma, \rho, \tau)$, homogeneous of order 0, such that $\psi(\sigma, \rho, \tau) = 0$ on a closed conic neighborhood of $\{(\sigma, \rho, \tau) \mid \tau = 0,\ (\sigma, \rho) \neq (0, 0)\}$. Then [21] $\psi_L L$ is a Fourier integral operator with canonical relation

$$\begin{aligned} \cup_{m\in M} \{ & (x, z, \boldsymbol{e}^{(m)}(x, z, s, r), \boldsymbol{\xi}^{(m)}(x, z, s, r, \tau, \varepsilon), \boldsymbol{\zeta}^{(m)}(x, z, s, r, \tau, \varepsilon), \varepsilon; \\ & s, r, T^{(m)}(x, z, s, r), \boldsymbol{\rho}^{(m)}(x, z, s, r, \tau, \varepsilon), \boldsymbol{\sigma}^{(m)}(x, z, s, r, \tau, \varepsilon), \tau) \mid \\ & \qquad (x, z, s, r) \in D^{(m)},\ \varepsilon \in \mathbb{R}^{n-1},\ \tau \in \mathbb{R}\backslash 0\} \end{aligned} \tag{13}$$

where

$$\boldsymbol{\xi}^{(m)}(x, z, s, r, \tau, \varepsilon) = \partial_x \Phi^{(m)} = \tau \partial_x T^{(m)}(x, z, s, r) - \langle \varepsilon, \partial_x \boldsymbol{e}^{(m)}(x, z, s, r)\rangle, \tag{14}$$

and likewise expressions for $\boldsymbol{\zeta}^{(m)}$, $\boldsymbol{\sigma}^{(m)}$ and $\boldsymbol{\rho}^{(m)}$.

Let d be the Born modeled data. To reveal any artifacts or false images generated by L, i.e. singularities in Ld at positions not corresponding to an element of $\mathrm{WF}(\delta c)$, we consider the composition $L\,F$. This composition is equal to the sum of a smooth e-family of pseudodifferential operators and, in general, a non-microlocal operator the wavefront set of which contains no elements with $\varepsilon = 0$ [21, theorem 6.1].

We discuss the practical considerations of the suppression of artifacts generated by L by modification of transform (11). The correct way to remove artifacts from Ld is by a Fourier cut-off in $\varepsilon/\|(\xi,\zeta)\|$ about 0. For bandlimited data, this is done approximately by local averaging in e at every (x,z). Such procedure has been applied to both synthetic and real data examples in [2]. The artifacts were not fully suppressed. In addition to this filtering procedure, the kernel of L in (11) was modified in the following way

$$L_\chi(x,z,e,r,s,t) = \sum_{m\in M} (2\pi)^{-(n-1)} \int \overline{A^{(m)}(x,z,s,r,\tau)}\chi(x,z,\nu^{(m)}(x,z,s,r)) \times \mathrm{e}^{\mathrm{i}\Phi^{(m)}(x,z,e,s,r,t,\varepsilon,\tau)}\,\mathrm{d}\tau\mathrm{d}\varepsilon,$$

where $\chi(x,z,\nu)$ is a smooth cutoff function on $\mathbb{R}^n \times S^{n-1}$. Observe that $\nu^{(m)}$ is the direction of $\frac{\partial T^{(m)}}{\partial(x,z)}$. Thus this cutoff filters out isochrons, and differs from the Fourier cutoff in $\varepsilon/\|(\xi,\zeta)\|$.

Remark 3.1. The transform L_χ restricts the wavefront set of operator L. If $\mathrm{WF}(\delta c)$ is contained in $\{(x,z,\lambda(\alpha_x(x,z),\alpha_z(x,z))) \mid (x,z,\lambda) \in \mathbb{R}^{n-1}\times\mathbb{R}_+\times\mathbb{R}\}$, where $(\alpha_x(x,z), \alpha_z(x,z))$ is a smooth covector field on $\mathbb{R}^{n-1}\times\mathbb{R}_+$, there is a χ with a small support in ν such that L_χ generates the true image. Artifacts with singular direction $(\boldsymbol{\xi}^{(m)},\boldsymbol{\zeta}^{(m)})$ such that the direction of $\frac{\partial T^{(m)}}{\partial(x,z)}$ is outside the support of χ, are suppressed.

The main difficulty arises when the background medium is not (accurately) known. Without knowledge of the background medium there is no criterion to distinguish artifacts from the true image. Also, the double-square-root equation approach to be developed below requires knowledge of the background medium. However, this approach does not generate artifacts even in the presence of caustics and hence no cutoff is needed. This is of particular importance in the development of tomographic methods for reconstructing the background medium.

4. The one-way wave or single-square-root equation

In this section we discuss the solution of the wave equation by evolution in one of the space variables (wave field extrapolation). This evolution problem is in general not well posed, but the propagation of the singularities of the solution can be obtained microlocally, when the propagation direction of the corresponding rays stays somewhat away from horizontal. Here we will introduce the pseudodifferential evolution equation as well as a regularized solution operator that yields this part of the solution. The regularized solution operator, composed with a pseudodifferential cutoff that is also introduced, will serve as a model for numerical wave field extrapolation operators.

Consider the wave equation rewritten as a first-order system

$$\frac{\partial}{\partial z}\begin{pmatrix} u \\ \frac{\partial u}{\partial z}\end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A(x,z,D_x,D_t) & 0\end{pmatrix}\begin{pmatrix} u \\ \frac{\partial u}{\partial z}\end{pmatrix} + \begin{pmatrix} 0 \\ f\end{pmatrix}, \tag{15}$$

where $A(x,z,\xi,\tau) = c_0(x,z)^{-2}\tau^2 - \|\xi\|^2$. Microlocally, away from the zeroes of $A(x,z,\xi,\tau)$, this system can be transformed into diagonal form modulo a smoothing operator [24]. There is a family of pseudodifferential operator matrices $\mathsf{Q}(z) = \mathsf{Q}(x,z,D_x,D_t)$ such that, microlocally,

$$\begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \mathsf{Q}(z) \begin{pmatrix} u \\ \frac{\partial u}{\partial z} \end{pmatrix}, \qquad \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \mathsf{Q}(z) \begin{pmatrix} 0 \\ f \end{pmatrix},$$

satisfy the one-way wave or single-square-root (SSR) equations

$$\left(\tfrac{\partial}{\partial z} \pm \mathrm{i} B_\pm(x,z,D_x,D_t)\right) u_\pm = f_\pm, \tag{16}$$

see [24]. The principal symbol b of the $B_\pm$ are given by $b(x,z,\xi,\tau) = \sqrt{A(x,z,\xi,\tau)} = \tau\sqrt{\frac{1}{c_0(x,z)^2} - \tau^{-2}\|\xi\|^2}$. For (x,t,ξ,τ) such that the symbol $B_\pm$ is real, the equation is of hyperbolic type, corresponding (microlocally) to propagating waves. To describe the associated bicharacteristics the Hamiltonian P can be used, as well as the principal symbol of the one-way wave operator in (16) $\zeta \mp b(x,z,\xi,\tau)$. The solution of a one-way wave equation describes the propagation of singularities along rays in intervals where $\pm\frac{\partial z}{\partial t} > 0$. The case where the symbol $B_\pm$ is imaginary corresponds to either evanescent waves, or waves that blow up like in a backward heat equation.

The subprincipal part of the $B_\pm$ depends on the normalization of $\mathsf{Q}(z)$. We choose this normalization such that (16) is selfadjoint microlocally where the symbol is real,

$$\begin{aligned} u &= Q_+^* u_+ + Q_-^* u_-, \\ f_\pm &= \pm\tfrac{1}{2}\mathrm{i} Q_\pm f, \end{aligned}$$

where $Q_\pm = Q_\pm(z) = Q_\pm(x,z,D_x,D_t)$ are z-families of pseudodifferential operators with principal symbols $\tau^{-1/2}\left(\frac{1}{c_0(x,z)^2} - \tau^{-2}\|\xi\|^2\right)^{-1/4}$ which form the first column of $\mathsf{Q}(z)$.

From now on we work with the equation with the $-$ sign, the $+$ sign can be treated similarly. We will also use z as the evolution parameter for bicharacteristics, and denote them as

(17)
$$\begin{aligned} \gamma(z,z_0;x_0,t_0,\xi_0,\tau) = (&\boldsymbol{x}(z,z_0,x_0,\xi_0,\tau), z, \boldsymbol{t}(z,z_0,x_0,\xi_0,\tau) + t_0, \boldsymbol{\xi}(z,z_0,x_0,\xi_0,\tau), \\ &b(\boldsymbol{x}(z,z_0,x_0,\xi_0,\tau), z, \boldsymbol{\xi}(z,z_0,x_0,\xi_0,\tau),\tau),\tau). \end{aligned}$$

At the zeroes of $A(x,z,\xi,\tau)$, given by $\frac{c(x,z)\|\xi\|}{|\tau|} = 1$, the operators $B_-(z), Q_-(z)$ are not yet defined. To define a (regularized) solution operator we modify $B_-(z), Q_-(z)$ on a neighborhood of this set. Suppose that $\epsilon > 0$ is small and that equation (16) must be satisfied on the subset of phase space given by

$$\frac{c(x,z)\|\xi\|}{|\tau|} \le 1 - \epsilon, \tag{18}$$

corresponding to local propagation angles in the interval $[-\arccos(1-\epsilon), \arccos(1-\epsilon)]$. We define regularized pseudodifferential operators B_- and Q_-, that are as given above when $\frac{c(x,z)\|\xi\|}{|\tau|} < 1 - \epsilon_0$, for some $\epsilon_0 < \epsilon$, and smoothly go to zero outside this region. The initial value problem for (16) with the modified operator B_- can be solved and the solution operator is a Fourier integral operator that will be denoted by $G_-(z,z_0)$. With the bicharacteristics is associated a maximal interval $I_1(z_0, x_0, \xi_0, \tau, \epsilon)$, where they stay in the region given by (18). The regularized solution operator G_- correctly describes the singularities on the set V_{ϵ_0} defined by

$$V_{\epsilon_0} = \{\gamma(z, z_0, x_0, t_0, \xi_0, \tau) \,|\, (x_0, t_0, \xi_0, \tau) \in T^*\mathbb{R}^n, z \in I_1(z_0, x_0, \xi_0, \tau, \epsilon_0)\} \tag{19}$$

Outside the set V_{ϵ_0} the singularities of G_- are no longer correct, but instead depend on the regularization. Most of the times we will therefore use G_- together with a pseudo-differential cutoff $\psi_1(z, z_0) = \psi_1(z, z_0, x, D_x, D_t)$. We assume the symbol $\psi_1(z, x, \xi, \tau)$ is 1 on the set V_ϵ, and in $S^{-\infty}$ outside the set V_{ϵ_0}. The operator ψ_1 removes the effect of the modification of B_- on the singularities of the solutions, and $\psi_1 G_-$ only has the singularities of G_- that are correct.

The composition $\psi_1(z, z_0)G_-(z, z_0)$ will serve as a model for numerical solution operators. Indeed, it gives what would be expected from a good solution operator, namely the propagation of singularities of the solution up to some a priori specified angle, and suppression of any incorrectly propagated singularities. Naturally the question arises in how far such properties are achieved by numerical methods. A detailed discussion of the numerical methods falls outside the scope of this paper, but we did find that such a solution operator can be obtained by solving a regularized version of (16) with an additional dissipative pseudodifferential term. We assume that this term is of order δ, $0 < \delta < 1/2$, and that it is non-zero when $\frac{c(x,z)\|\xi\|}{|\tau|} > 1 - \epsilon_1$, for some ϵ_1, $\epsilon_0 < \epsilon_1 < \epsilon$. A further discussion and a proof that this yields the desired pseudodifferential cutoff is given in the appendix. The cost associated with such a regularization, as far as our analysis is concerned, is that the symbol of the solution operator satisfies slightly weaker symbol estimates than a standard symbol. In this case the solution operator is a Fourier integral operator of type $\rho = 1 - \delta$ with symbol in S^μ_ρ, for some μ. It still enjoys the calculus properties that we use to obtain our results. The equation with additional dissipative term is solved only in one direction (depending on the sign of the dissipative term), the adjoint of the solution operator gives a regularized solution in the opposite propagation direction.

By Duhamel's principle a microlocal solution operator for the inhomogeneous equation is given by

$$u_-(\cdot, z) = \int_{-\infty}^{z} G_-(z, z_0) f_-(\cdot, z_0) \, \mathrm{d}z_0. \tag{20}$$

Remark 4.1. The symbols of the operators $B_\pm(z), Q_\pm(z)$ can be written as an infinite sum of elementary symbols, $\sum_{i=1}^{\infty} v_i(x, z, \tau) w_i(z, \xi, \tau)$ say, that is rapidly converging where they are smooth, i.e. away from the set of zeroes of $A(x, z, \xi, \tau)$, in accordance with [25, equation (2.1.11)]. The elementary symbols correspond to multiplications either in horizontal position space (v_i) or in horizontal wave number (Fourier) space (w_i). This is exploited in one type of fast numerical solvers of (16), see [6] and the references therein.

5. MODELING, THE DOUBLE-SQUARE-ROOT EQUATION

We show that the Born modeling operator can be written, modulo smoothing terms, in terms of the solution operator to the double-square-root (DSR) equation (25) below. We assume that the tangent vectors to the rays that connect source and receiver to a scattering point in X stay away some finite distance from horizontal as given in (18)

Assumption 3. *(DSR assumption) If $(x,z) \in X$ and $\alpha, \beta \in S^{n-1}$, $t_{\mathrm{s}}, t_{\mathrm{r}} > 0$ depending on (x,z,α,β) are such that $\boldsymbol{z}(x,z,\beta,\tau,t_{\mathrm{s}}) = \boldsymbol{z}(x,z,\alpha,\tau,t_{\mathrm{r}}) = 0$ and $(\boldsymbol{x}(x,z,\beta,\tau,t_{\mathrm{s}}), \boldsymbol{x}(x,z,\alpha,\tau,t_{\mathrm{r}}), t_{\mathrm{s}}+t_{\mathrm{r}}) \in Y$ (cf. (7)), then*

$$c(x,z)^{-1}\frac{\partial \boldsymbol{z}}{\partial t}(x,z,\beta,\tau,t) < -\sqrt{2\epsilon-\epsilon^2}, t \in [0,t_{\mathrm{s}}],$$
$$c(x,z)^{-1}\frac{\partial \boldsymbol{z}}{\partial t}(x,z,\alpha,\tau,t) < -\sqrt{2\epsilon-\epsilon^2}, t \in [0,t_{\mathrm{r}}],$$

where $\epsilon > 0$ was introduced above (18).

This assumption is stronger than Assumption 1. In general the set of rays violating this assumption is not small, but can contain an open subset of the canonical relation (7), depending on the properties of the background medium. This limits the applicability of the method discussed here, which however is still useful in many cases, as discussed in the introduction. The assumption is microlocal, and, given the background medium, a pseudodifferential cutoff can be applied to the data to remove microlocally the part of the data where Assumption 3 is violated.

Under Assumption 3 and the assumption that $\delta c = 0$ on a neighborhood of $z = 0$, the singular part of the Born data is unchanged when G in (4) is replaced by $-\frac{1}{2}\mathrm{i}Q_-^*(x,z,D_x,D_t)G_-Q_-(x,z,D_x,D_t)$. Define the operator $H(z,z_0), z < z_0$, by

$$(H(z,z_0))(s,r,t,s_0,r_0,t_0) = \int_{\mathbb{R}} (G_-(z,z_0))(s,t-t_0-t',s_0,0)(G_-(z,z_0))(r,t',r_0,0)\,\mathrm{d}t'. \tag{21}$$

Here $(G_-(z,z_0))(r,t',r_0,0)$ denotes the distribution kernel of $G_-(z,z_0)$, and similarly for $H(z,z_0)$. Define the maps $\mathcal{I}_1, \mathcal{I}_2$ by

$$\mathcal{I}_1 : \mathcal{D}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^{2n-1}) : u(x,z) \mapsto \delta(r-s)u(\tfrac{r+s}{2},z), \tag{22}$$
$$\mathcal{I}_2 : \mathcal{D}'(\mathbb{R}^{2n-1}) \to \mathcal{D}'(\mathbb{R}^{2n}) : u(r,s,z) \mapsto \delta(t)u(r,s,z). \tag{23}$$

The operators G, G_-, Q_- all are of convolution type in the time variable. It follows that the singular part of the Born approximated data (4) is given by

$$F\delta c = \int_0^{z_{\max}} Q_{-,s}^*(0)Q_{-,r}^*(0)H(0,z)Q_{-,s}(z)Q_{-,r}(z)\,\tfrac{1}{2}D_t^2(\mathcal{I}_2\mathcal{I}_1\,c_0^{-3}\delta c)(\cdot,z)\,\mathrm{d}z. \tag{24}$$

where $Q_{-,s}(z)$ is short for $Q_-(s,z,D_s,D_t)$ and similarly for $Q_{-,r}(z)$. Here $z_{\max}$ is the maximum depth illuminated from acquisition manifold Y given the background medium c_0.

Define the inhomogeneous double-square-root (DSR) equation by

$$(\tfrac{\partial}{\partial z} - \mathrm{i}B_-(s,z,D_s,D_t) - \mathrm{i}B_-(r,z,D_r,D_t))u = g(s,r,t,z). \tag{25}$$

It follows from the definition of $G_-(z,z_0)$ that the operator $H(z,z_0)$ is a regularized solution operator for the Cauchy initial value problem for the DSR equation. The Born approximated data is modulo smooth functions equal to the solution of the DSR equation at $z = 0$ with inhomogeneous term $g(s,r,t,z) = \frac{1}{2}D_t^2(\mathcal{I}_2\mathcal{I}_1\,c_0^{-3}\delta c)(s,r,t,z)$ in analogy with (2).

The operator $H(z, z_0)$ propagates singularities along the bicharacteristics associated with (25), given by, in the notation of (17),

$$(\boldsymbol{x}(z, z_0, s_0, \sigma_0, \tau), \boldsymbol{x}(z, z_0, r_0, \rho_0, \tau), t_0 + \boldsymbol{t}(z, z_0, s_0, \sigma_0, \tau) + \boldsymbol{t}(z, z_0, r_0, \rho_0, \tau), z, \\ \boldsymbol{\xi}(z, z_0, s_0, \sigma_0, \tau), \boldsymbol{\xi}(z, z_0, r_0, \rho_0, \tau), \tau, \\ \boldsymbol{\Gamma}(\boldsymbol{x}(z, z_0, s_0, \sigma_0, \tau), \boldsymbol{x}(z, z_0, r_0, \rho_0, \tau), \boldsymbol{\xi}(z, z_0, s_0, \sigma_0, \tau), \boldsymbol{\xi}(z, z_0, r_0, \rho_0, \tau), \tau, z)), \tag{26}$$

where

$$\boldsymbol{\Gamma}(s, r, \sigma, \rho, \tau, z) = b(s, z, \sigma, \tau) + b(r, z, \rho, \tau). \tag{27}$$

Here z is in an interval that is the intersection of the maximal intervals associated with the source and receiver bicharacteristics introduced below (18). The lower and upper bounds of this interval will be denoted by $Z_{\min}(z_0, s_0, r_0, \sigma_0, \rho_0, \tau, \epsilon)$, $Z_{\max}(z_0, s_0, r_0, \sigma_0, \rho_0, \tau, \epsilon)$, respectively. We define a pseudodifferential cutoff $\psi_2(z, z_0) = \psi_2(z, z_0, s, r, D_s, D_r, D_t)$ that removes incorrectly propagated singularities from H similar as we defined $\psi_1(z, z_0)$ below (19) (using the DSR bicharacteristics and their maximal interval to define the equivalent of the set V_ϵ.) We assume the symbol of $\psi_2(z, z_0)$ is in $S^0_\rho(\mathbb{R}^{2n-1} \times \mathbb{R}^{2n-1})$, so that the theory of the appendix applies.

6. Double-square-root reconstruction

The reconstruction of the perturbation δc given the background is essentially done by applying the adjoint of the linear modeling map (24) discussed in the previous section to the data d. (The outcome is a so-called seismic image of the perturbation.) Let $N := F^* \psi F$, where $\psi = \psi(s, r, t, D_s, D_r, D_t)$ is a suitable pseudodifferential cutoff as in section 2. By Theorem 2.1, with Assumption 2, $N = N(x, z, D_x, D_z)$ is a pseudodifferential operator. Therefore applying the adjoint $F^* \psi$ to the data yields

$$N(x, z, D_x, D_z)\delta c = F^* \psi(s, r, t, D_s, D_r, D_t) d, \tag{28}$$

and δc can be reconstructed microlocally where the principal symbol of N is non-zero.

In this section present a reconstruction formula similar to (28) based on the DSR Born modeling (24). This leads, again, to reconstruction modulo a pseudodifferential operator for which an explicit expression is given.

Assumption 3 implies the injectivity in Assumption 2 and is hence stronger than Assumption 2 by the remark below Assumption 2. (A point $(s, r, t, \sigma, \rho, \tau)$ in the projection of Λ on $T^*Y \backslash 0$ uniquely determines the initial values of a source and a receiver one-way bicharacteristic. Since under Assumption 3 travel time increases monotonically with depth, travel time injectivity could only be violated within a plane of fixed depth. However, by the same monotonicity, the one-way receiver bicharacteristic intersects each plane of fixed depth in at most one point. The same is true for the one-way source bicharacteristic.) So Assumption 3 is sufficient for the reconstruction. However, since the operator H^* will be used in the imaging formula together with a cutoff as discussed in sections 4 and 5, in fact the following weaker assumption (a modification of the traveltime injectivity condition) will be sufficient.

Assumption 4. *If several elements of (7) project to the same point in T^*Y, then none of the (x, z, α, β) that describe these elements are such that Assumption 3 is satisfied.*

The adjoint operator $H(0,z)^*$ is equal to $H(z,0)$ (the regularization can be chosen such that this is true.) It propagates the data downward and backward in time. We consider $H(0,z)^*d$ as a function of (s,r,t,z). The operator $H(0,z)^*$ is microlocally a Fourier integral operator with canonical relation which follows from (26) with $z_0 = 0$, for z in the interval $[0, Z_{\max}(0, s_0, r_0, \sigma_0, \rho_0, \tau, \epsilon_0)[$

$$\begin{aligned}(29)\quad \big\{(&\boldsymbol{x}(z,0,s_0,\sigma_0,\tau), \boldsymbol{x}(z,0,r_0,\rho_0,\tau), t_0 + \boldsymbol{t}(z,0,s_0,\sigma_0,\tau) + \boldsymbol{t}(z,0,r_0,\rho_0,\tau), z,\\ &\boldsymbol{\xi}(z,0,s_0,\sigma_0,\tau), \boldsymbol{\xi}(z,0,r_0,\rho_0,\tau), \tau, \boldsymbol{\Gamma}(s_0,r_0,\sigma_0,\rho_0,\tau,z); s_0,r_0,t_0,\sigma_0,\rho_0,\tau) \mid \\ &\qquad\qquad (s_0,r_0,t_0,\sigma_0,\rho_0,\tau) \in T^*\mathbb{R}^{2n-1}\backslash 0, z \in \mathbb{R}_+\big\}.\end{aligned}$$

For $z > Z_{\max}(0, s_0, r_0, \sigma_0, \rho_0, \tau, \epsilon_0)$, the bicharacteristics are only defined when a regularization is chosen (we will use a pseudodifferential cutoff below to remove the corresponding singularities). The adjoint of the operator $\mathcal{I}_2$ (cf. (23)) is given by the restriction $\mathcal{R}_2$ defined by

$$g(r,s,t,z) \mapsto (\mathcal{R}_2 g)(r,s,z) = g(r,s,0,z).$$

If $u = u(s,r,z)$ then $Ku = (Ku)(s,r,t)$ will be defined by

$$Ku = \int_{-\infty}^{0} H(0,z)\mathcal{I}_2 u(\cdot,z)\,\mathrm{d}z. \tag{30}$$

The adjoint K^* of K is given by

$$K^* = \mathcal{R}_2 H(0,z)^*.$$

Since $\frac{\partial \boldsymbol{t}}{\partial z} < 0$, $\boldsymbol{t}$ depends strictly monotone on z for each $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ in the region $z \in [0, Z_{\max}(0, s_0, r_0, \sigma_0, \rho_0, \tau, \epsilon_0)]$. For such z, the DSR bicharacteristic with initial values $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ hence intersects the $t=0$ hyperplane at most once, and the intersection is transversal. From this, it follows that microlocally the operator K^* is a Fourier integral operators with a locally invertible canonical relation, that maps points $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ in a subset of the cotangent acquisition space $T^*\mathbb{R}^{2n-1}\backslash 0$ to $(r,s,z,\sigma,\rho,\zeta)$ in a subset of $T^*\mathbb{R}^{2n-1}\backslash 0$. This map converts time to depth.

In our imaging formula we will replace $H(z,0)$ by a version with cutoff $\psi_2(z,0)H(z,0)$, where $\psi_2(z_0,z)$ was introduced at the end of section 5. In addition we introduce a smooth cutoff function $\psi_Y = \psi_Y(s,r,t)$ on Y that is zero near the boundary of Y. The modified version of K^* now reads

$$d \mapsto \mathcal{R}_2 \psi_2(z,0) H(z,0) \psi_Y d. \tag{31}$$

It is not difficult to see that the two cutoffs ψ_2, ψ_Y can in fact be replaced by a single cutoff $\psi(s,r,t,D_s,D_r,D_t)$ acting on d, and we will derive its symbol. Instead of $\psi_2(z,0)H(z,0)$ we can also write $H(z,0)\psi_2(0,z)$. As we already noted, a DSR bicharacteristic with initial value $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ at $z=0$ intersects the $t=0$ plane at some depth $\tilde{z}$, that increases monotonically with t_0 (for z in the interval given above). With $\tilde{z} = \tilde{z}(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ we find for the principal symbol of ψ

$$\psi(s,r,t,\sigma,\rho,\tau) = \psi_2(0,\tilde{z},s,r,\sigma,\rho,\tau)\psi_Y(s,r,t). \tag{32}$$

As mentioned the canonical relation of $K^*\psi$ maps points $(s_0,r_0,t_0,\sigma_0,\rho_0,\tau)$ in a subset of the cotangent acquisition space $T^*\mathbb{R}^{2n-1}\backslash 0$ to $(r,s,z,\sigma,\rho,\zeta)$ in a subset of $T^*\mathbb{R}^{2n-1}\backslash 0$. We define the symbol $\Psi(s,r,z,\sigma,\rho,\zeta) \in S^0_\rho(\mathbb{R}^{2n-1}\times\mathbb{R}^{2n-1})$ as the pull back of $\psi(s_0,r_0,t_0,\sigma_0,$

$\rho_0, \tau)$ under the inverse of this map. Starting from the Born modeling (24), we obtain the following reconstruction result.

Theorem 6.1. *Suppose Assumption 4 is satisfied. Then there are pseudodifferential operators* $\Phi = \Phi(x, z, D_x, D_z)$ *of order* $n-1$, $\Phi \in S_\rho^{n-1}$, *with principal symbol*

$$\Phi(x, z, \xi, \zeta) = \int_{\mathbb{R}^{n-1}} \Psi(x, x, z, \tfrac{1}{2}\xi - \theta, \tfrac{1}{2}\xi + \theta, \zeta)\,\mathrm{d}\theta, \tag{33}$$

and $\Xi(z) = \Xi(s, r, t, D_s, D_r, D_t, z)$ *of order* 0 *with principal symbol*

$$\begin{aligned}\Xi(s, r, t, \sigma, \rho, \tau, z) &= \left|\frac{\partial \boldsymbol{\Gamma}}{\partial \tau}(s, r, \sigma, \rho, \tau, z)\right| \\ &= c_0(s, z)^{-2}(c_0(s, z)^{-2} - \tau^{-2}\|\sigma\|^2)^{-1/2} + c_0(r, z)^{-2}(c_0(r, z)^{-2} - \tau^{-2}\|\rho\|^2)^{-1/2},\end{aligned}$$

such that

$$\begin{aligned}&\Phi(x, z, D_x, D_z)\delta c \\ &\quad = 2c_0^3 \mathcal{R}_1 \mathcal{R}_2 \Xi(z) Q_{-,s}^*(z)^{-1} Q_{-,r}^*(z)^{-1} \psi_2(z, 0) H(z, 0) Q_{-,s}(0)^{-1} Q_{-,r}(0)^{-1} D_t^{-2} \psi_Y d,\end{aligned} \tag{34}$$

where $d = F\delta c$ *is the Born modeled data.*

Proof. Because of Theorem 2.2 and the modeling formula (24) the right hand side of (34) is equal to a pseudodifferential operator acting on δc. It remains therefore to compute the symbol. We calculate microlocally the principal symbol of K^*K. The kernel of the operator $H(0, z)$ has microlocally an oscillatory integral representation with a phase function associated with generating function, $S = S(z, s, r, t, y_{0I}, \eta_{0J})$ say, where $y_0 = (s_0, r_0, t_0)$ and η_0 is the corresponding cotangent vector, and $\{I, J\}$ is a partition of $\{1, \ldots, 2n-1\}$,

$$\begin{aligned}H(0, z)(y_0, s, r, t) = (2\pi)^{-(2n-1+|I|)/2} \\ \times \int A(z, s, r, t, y_0, \eta_{0J}) \mathrm{e}^{\mathrm{i}(S(z,s,r,t,y_{0I},\eta_{0J}) - \langle \eta_{0J}, y_{0J}\rangle)}\,\mathrm{d}\eta_{0J},\end{aligned}$$

The adjoint $H(0, z)^*$ has amplitude $\overline{A(z, s, r, t, y_0, \eta_{0J})}$ and phase $-S(z, s, r, t, y_{0I}, \eta_{0J}) + \langle \eta_{0J}, y_{0J}\rangle$. Hence, the kernel of the composition $H(0, z)^* H(0, z)$ has the oscillatory integral representation

$$\begin{aligned}(2\pi)^{-(2n-1)} \int \overline{A(z, s', r', t', y_0, \eta_{0J})} A(z, s, r, t, y_0, \eta_{0J}) \\ \times \mathrm{e}^{\mathrm{i}[-S(z,s',r',t',y_{0I},\eta_{0J}) + S(z,s,r,t,y_{0I},\eta_{0J})]}\,\mathrm{d}y_{0I}\mathrm{d}\eta_{0J}.\end{aligned}$$

We expand the phase in a Taylor series around $(s', r', t') = (s, r, t)$ and identify the gradient

$$\begin{aligned}-\frac{\partial S}{\partial(s, r, t)}(z, s, r, t, y_{0I}, \eta_{0J}) = \\ (\sigma(z, s, r, t, y_{0I}, \eta_{0J}), \rho(z, s, r, t, y_{0I}, \eta_{0J}), \tau(z, s, r, t, y_{0I}, \eta_{0J})).\end{aligned}$$

Applying a change of variables, $(y_{0I}, \eta_{0J}) \mapsto (\sigma, \rho, \tau)$, the phase takes the form

$$\langle (\sigma, \rho, \tau), (s' - s, r' - r, t' - t)\rangle.$$

Since $H(0,z)^*H(0,z)$ is a pseudodifferential operator with symbol 1, microlocally in the support of the cutoff ψ, we conclude that the principal part a of the amplitude A satisfies

$$|a(z,s,r,t,y_0,\eta_{0J})| = \left|\frac{\partial(\sigma,\rho,\tau)}{\partial(y_{0I},\eta_{0J})}\right|^{1/2}.$$

The kernel of operator K has an oscillatory integral representation similar to the one of $H(0,z)$,

$$K(y_0,s,r,z) = (2\pi)^{-(2n-1+|I|)/2}\int A(z,s,r,0,y_0,\eta_{0J})\mathrm{e}^{\mathrm{i}(S(z,s,r,0,y_{0I},\eta_{0J})-\langle\eta_{0J},y_{0J}\rangle)}\,\mathrm{d}\eta_{0J}$$

(we have applied $\mathcal{I}_2$ at z). It follows that the composition K^*K, carrying out an analysis similar to the one for $H(0,z)^*H(0,z)$, is a pseudodifferential operator, microlocally. Its amplitude has principal part

$$\left|\frac{\partial(\zeta,\sigma,\rho)}{\partial(y_{0I},\eta_{0J})}\right|^{-1}\left|\frac{\partial(\sigma,\rho,\tau)}{\partial(y_{0I},\eta_{0J})}\right|.$$

For fixed (s,r,σ,ρ,z) the map $\tau \mapsto \mathbf{\Gamma}(s,r,\sigma,\rho,\tau,z)$ is invertible on a set given by $|\tau|$ sufficiently large. This map will be denoted by $\mathbf{\Gamma}^{-1} = \mathbf{\Gamma}^{-1}(s,r,z,\sigma,\rho,\zeta)$. It follows that the principal part of K^*K is microlocally given by

$$\left|\frac{\partial\mathbf{\Gamma}}{\partial\tau}(s,r,\sigma,\rho,\mathbf{\Gamma}^{-1}(s,r,z,\sigma,\rho,\zeta),z)\right|^{-1}.$$

We finally consider the composition of operators $\mathcal{R}_1\Psi\mathcal{I}_1$. Its kernel has an oscillatory integral representation,

$$(2\pi)^{-(2n-1)}\int_{\mathbb{R}^{2n-1}}\Psi(x,x,z,\sigma,\rho,\zeta)\,\mathrm{e}^{\mathrm{i}\langle(x,x,z)-(x',x',z'),(\sigma,\rho,\zeta)\rangle}\,\mathrm{d}\rho\mathrm{d}\sigma\mathrm{d}\zeta =$$
$$(2\pi)^{-(2n-1)}\int_{\mathbb{R}^n}\int_{\mathbb{R}^{n-1}}\Psi(x,x,z,\tfrac{1}{2}\xi-\theta,\tfrac{1}{2}\xi+\theta,\zeta)\,\mathrm{d}\theta\,\mathrm{e}^{\mathrm{i}(\langle(x-x'),\xi\rangle+\langle(z-z'),\zeta\rangle)}\,\mathrm{d}\xi\mathrm{d}\zeta,$$

upon changing variables of integration, $\sigma = \frac{1}{2}\xi-\theta$, $\rho = \frac{1}{2}\xi+\theta$. The domain of the θ integral is bounded depending on (ξ,ζ), since Ψ is a cutoff in (s,r,t,σ,ρ,τ). Hence $\mathcal{R}_1\Psi\mathcal{I}_1$ is a pseudodifferential operator of order $n-1$ with principal symbol (33).

Applying the above results to the expression (24) for the Born modeled data leads to the statement of the theorem. □

Remark 6.2. Note that in (34) all the operators on the right hand side, except the downward continuation $\psi_2 H$, act at depth z only. On the contrary, the operator $\Phi(x,z,D_x,D_z)$ on the left hand side depends on the Hamiltonian flow associated with the background medium in the depth interval $[0,z]$. In the usual wave-equation imaging algorithms it is hence straightforward to include the pseudodifferential factors on the right hand side of (34). To account for the operator $\Phi(x,z,D_x,D_z)$ on the left hand side one requires an additional ray computation.

Remark 6.3. Depending on the background medium, the reconstruction can also be done using data on a submanifold Y' of Y. Let $\mathcal{R}'$ be the restriction of a function on Y to Y', so that the forward map for this case is given by $\mathcal{R}'F$. In suitable local coordinates

(y', y'') on Y such that $y'' = 0$ defines Y', the adjoint $\mathcal{I}'$ of $\mathcal{R}'$ is given by the map $(\mathcal{I}'f)(y', y'') = f(y')\delta(y'')$. Conditions such that $F^*\mathcal{I}'\psi'\mathcal{R}'F$ is pseudodifferential are given in [16], where ψ' is a suitable pseudodifferential cutoff. Reconstruction modulo a pseudodifferential operator is done in this case by first applying the map $\mathcal{I}'$ to the data, and then applying the previous procedure. Applying $\mathcal{I}'$ to the data simply means adding zeroes where there is no data in Y.

7. THE WAVE-EQUATION ANGLE TRANSFORM

We define the wave-equation angle transform A_{WE} by the following integral of the downward continued data, $\psi_2(z, 0)H(z, 0)\psi_Y d$,

$$(A_{\mathrm{WE}}d)(x, z, p) = \int_{\mathbb{R}^{n-1}} (\psi_2(z, 0)H(z, 0)\psi_Y d)(x - \tfrac{h}{2}, x + \tfrac{h}{2}, ph)\chi(x, z, h)\,\mathrm{d}h, \tag{35}$$

(cf. [5, 18]), where $h \mapsto \chi(x, z, h)$ is a compactly supported cutoff function the support of which contains $h = 0$. The first result in the following theorem implies that, when the correct background medium c_0 is given, A_{WE} maps data to a p-family of reconstructions of δc in the form of pseudodifferential operators acting on δc. Hence artifact-free images are obtained, unlike the images obtained with the generalized Radon transform derived operator $\psi_L L$ in (11)-(12).

Theorem 7.1. *Suppose Assumption 3 holds. Let C_0 be an upper bound for c_0. Assume that*

$$|p| < p_{\max} < \tfrac{1}{2}C_0^{-1}. \tag{36}$$

Then A_{WE} is a Fourier integral operator such that $A_{\mathrm{WE}}F$ is a smooth p-family of pseudodifferential operators in (x, z). Let C_1 be an upper bound for $\frac{\partial c_0^{-2}}{\partial x}$. If in addition the function $h \mapsto \chi(x, z, h)$ is supported in $B(0, R)$, where R depends on C_1 and ϵ, then the canonical relation of A_{WE} corresponds to an invertible map from a subset of $T^*\mathbb{R}^{2n-1}_{(s,r,t)}$ to a subset of $T^*\mathbb{R}^{2n-1}_{(x,z,p)}$ that has nonempty intersection with the set $\theta = 0$ (where θ is the p-covector).

Proof. The map $d \mapsto \psi_2(z, 0)H(z, 0)\psi_Y d$ is a Fourier integral operator with canonical relation given in (29).

The Schwarz kernel of the map $\psi_2(z, 0)H(z, 0)\psi_Y d \mapsto A_{\mathrm{WE}}d$ equals

$$\delta(x - \tfrac{s+r}{2})\,\delta(p(r - s) - t)\,\delta(z - z')\,\chi(x, z, r - s) \\ = (2\pi)^{-n-1} \int \mathrm{e}^{\mathrm{i}(\langle \xi, x - \frac{s+r}{2}\rangle + \tau(p(r-s)-t) + \zeta(z-z'))}\,\mathrm{d}\xi\,\mathrm{d}\tau\,\mathrm{d}\zeta.$$

It is a Fourier integral operator with canonical relation that is contained in $T^*\mathbb{R}^{2n-1}_{(x,z,p)}\backslash 0 \times T^*\mathbb{R}^{2n}_{(s,r,t,z)}\backslash 0$ and given by

$$\{(\tfrac{s+r}{2}, z, p, \xi, \zeta, (r - s)\tau; s, r, p(r - s), z, \tfrac{\xi}{2} + p\tau, \tfrac{\xi}{2} - p\tau, \tau, \zeta) \mid \\ (s, r, z, p, \xi, \zeta, \tau) \in \text{subset of } \mathbb{R}^{4n-1}\}. \tag{37}$$

This canonical relation can be parameterized by the coordinates of $T^*\mathbb{R}^{2n}_{(s,r,t,z)}\backslash 0$ except t, that is $(s, r, z, \sigma, \rho, \tau, \zeta)$. The projection of (37) on $T^*\mathbb{R}^{2n}_{(s,r,t,z)}\backslash 0$ is a hypersurface defined by

$$t = \left\langle \frac{\sigma - \rho}{2\tau}, (r - s) \right\rangle . \tag{38}$$

The canonical relation of the map $d \mapsto \psi_2(z, 0)H(z, 0)\psi_Y d$, considered as a function of (s, r, t, z), is parameterized by $(s_0, r_0, t_0, \sigma_0, \rho_0, \tau, z)$. It is given in (29), it is contained in $T^*\mathbb{R}^{2n}_{(s,r,t,z)}\backslash 0 \times T^*\mathbb{R}^{2n-1}_{(s_0,r_0,t_0)}\backslash 0$, and is time translation invariant (in t_0). The line in the projection of this canonical relation on $T^*\mathbb{R}^{2n}_{(s,r,t,z)}\backslash 0$ parameterized by t_0 for fixed $(s_0, r_0, \sigma_0, \rho_0, \tau)$ intersects the hypersurface (38) transversally. It follows that the composition of the canonical relations (29) and (37) is transversal. The composition is parameterized by $(s_0, r_0, \sigma_0, \rho_0, \tau, z)$. It follows that A_{WE} is a Fourier integral operator.

The composition $\psi_2(z, 0)H(z, 0)\psi_Y F$, that maps $\delta c = \delta c(x, z')$ to the downward continued data as a function of (s, r, t, z), is a Fourier integral operator with canonical relation

$$\begin{aligned} \{(&\boldsymbol{x}(z, z', x, \sigma', \tau), \boldsymbol{x}(z, z', x, \rho', \tau), \boldsymbol{t}(z, z', x, \sigma', \tau) + \boldsymbol{t}(z, z', x, \rho', \tau), z, \\ &\boldsymbol{\xi}(z, z', x, \sigma', \tau), \boldsymbol{\xi}(z, z', x, \rho', \tau), \tau, \zeta; x, z', \sigma' + \rho', \zeta') \mid \\ &(x, z', \sigma', \rho', \zeta', z) \in \text{ a subset of } \mathbb{R}^{3n}, \tau \text{ such that } \zeta' = \boldsymbol{\Gamma}(x, x, \sigma', \rho', \tau, z'), \\ &\zeta = \boldsymbol{\Gamma}(\boldsymbol{x}(z, z', x, \sigma', \tau), \boldsymbol{x}(z, z', x, \rho', \tau), \boldsymbol{\xi}(z, z', x, \sigma', \tau), \boldsymbol{\xi}(z, z', x, \rho', \tau), \tau, z)\}. \end{aligned} \tag{39}$$

The propagation of singularities upward by F and downward by $H(z, 0)$ is along the same DSR bicharacteristics.

We show that the composition $A_{\mathrm{WE}}F$ is a Fourier integral with canonical relation contained in

$$\{(x, z, p, \xi, \zeta, 0; x, z, \xi, \zeta) \mid (x, z, \xi, \zeta) \in T^*\mathbb{R}^n_{(x,z)}\backslash 0, p \in]-p_{\max}, p_{\max}[\}. \tag{40}$$

From that it follows that $A_{\mathrm{WE}}F$ is a p-family of pseudodifferential operators. The projection of (39) on $T^*\mathbb{R}^{2n}_{(s,r,t,z)}$ intersects the hypersurface (38) at $z' = z$, since then $r - s = 0$ and $t = 0$, leading to elements in (40). Since singularities propagate with speed less than C_0, if $(v_{\mathrm{s}}, v_{\mathrm{r}}, v_{\mathrm{t}}, v_{\mathrm{z}}, v_\sigma, v_\rho, 0, v_\zeta)$ is a tangent vector to the DSR bicharacteristic, then $\frac{v_{\mathrm{s}} - v_{\mathrm{r}}}{v_{\mathrm{t}}} \leq 2C_0$. Therefore, by (36), the composition of (39) with (37) is transversal and contains no elements outside (40).

The projection of the canonical relation of A_{WE} on the second component $T^*\mathbb{R}^{2n-1}_{(s,r,t)}\backslash 0$ is invertible if each DSR bicharacteristic with initial values $(s_0, r_0, t_0, \sigma_0, \rho_0, \tau)$, parameterized by z, intersects the hypersurface (38) at most once and transversally. Let $\boldsymbol{p}(z)$ denote $\frac{\sigma - \rho}{2\tau}$ along a certain DSR bicharacteristic and let $\boldsymbol{t}(z)$ denote the time and $\boldsymbol{h}(z)$ denote the value of $r - s$. The elements of the canonical relation of A_{WE} correspond to solutions of $\boldsymbol{t}(z) - \langle \boldsymbol{p}(z), \boldsymbol{h}(z) \rangle = 0$. To estimate the derivative of the left hand side we observe that

$$\frac{\partial \boldsymbol{t}}{\partial z} - \langle \boldsymbol{p}(z), \frac{\partial \boldsymbol{h}}{\partial z}(z) \rangle < -\epsilon_0$$

for some $\epsilon_0 > 0$ depending on ϵ and on $p_{\max}$. Since

$$\frac{\partial \boldsymbol{\xi}}{\partial z} = \frac{\partial b}{\partial x} = -\frac{\tau}{\sqrt{c_0^{-2} - \tau^{-2}\|\xi\|^2}} \frac{\partial c_0^{-2}}{\partial x},$$

there is a bound on $\frac{\partial \boldsymbol{p}}{\partial z}$ in terms of C_1 and on ϵ (to bound the square root from below). It follows that for some $\epsilon_1 < \epsilon_0$

$$\left| \langle \frac{\partial \boldsymbol{p}}{\partial z}(z), \boldsymbol{h}(z) \rangle \right| < \epsilon_1 < \epsilon_0$$

if $\|h\| < C_2 C_1^{-1}$ for some constant C_2 depending on ϵ and $p_{\max}$. This implies that the function $z \mapsto \boldsymbol{t}(z) - \langle \boldsymbol{p}(z), \boldsymbol{h}(z) \rangle$ is monotone. Hence the projection of the canonical relation of A_{WE} on $T^*\mathbb{R}^{2n-1}_{(s,r,t)} \backslash 0$ is invertible. It follows from the above reasoning that the projection on $T^*\mathbb{R}^{2n-1}_{(x,z,p)} \backslash 0$ is invertible as well. This establishes the last statement of the theorem. □

To conclude this section we determine, at the principal symbol level, the modification of (35) that leads to reconstruction of singularities of δc microlocally. For each p_i, $i = 1, \ldots, n-1$ such a reconstruction is analogous to the reconstruction in (34) of Theorem 6.1. The three cutoffs χ, ψ_Y, ψ_2 must of course be taken into account. The canonical relation of A_{WE} defines a map $(s, r, t, \sigma, \rho, \tau) \to (x, z, p, \xi, \zeta, \theta)$ (where θ is the p-covector), and there is also an associated value of $h = r - s$ through (37). By pull back one can map the symbols ψ_Y, $\psi_2(z, 0, s, r, \sigma, \rho, \tau)$ and $\chi(x, z, h)$ to symbols in the variables $(x, z, p, \xi, \zeta, \theta)$. We define $\Psi_{\mathrm{WE}} = \Psi_{\mathrm{WE}}(x, z, p, \xi, \zeta, \theta)$ as the product of these symbols. With this definition we have

Proposition 7.2. *Define* $\widetilde{A}_{\mathrm{WE}}$ *by*

$$(41)\quad (\widetilde{A}_{\mathrm{WE}} d)(x, z, p) = \int_{\mathbb{R}^{n-1}} \mathrm{d}h\, \chi(x, z, h) 2 c_0(x, z)^3$$
$$\times \big(\Xi(z) Q^*_{-,s}(z)^{-1} Q^*_{-,r}(z)^{-1} \psi_2(z, 0) H(z, 0) Q_{-,s}(0)^{-1} Q_{-,r}(0)^{-1} D_t^{-2} \psi_Y d\big)(x - \tfrac{h}{2}, x + \tfrac{h}{2}, ph).$$

Suppose Assumption 3 holds. Suppose that $\chi(x, z, 0) = 1$ *and* $h \mapsto \chi(x, z, h)$ *is supported in* $B(0, R)$ *(cf. Theorem 7.1), then* $\widetilde{A}_{\mathrm{WE}}$ *is an invertible Fourier integral operator. Let the symbol* Ψ_{WE} *be as defined above. The composition* $\widetilde{A}_{\mathrm{WE}} F$ *is a p-family of pseudodifferential operators with principal symbol* $\Psi_{\mathrm{WE}}(x, z, p, \xi, \zeta, 0)$*. The microlocal reconstruction hence follows from*

$$(42)\qquad \Psi_{\mathrm{WE}}(x, z, p, D_x, D_z, 0) \delta c = \widetilde{A}_{\mathrm{WE}} d.$$

Proof. It is sufficient to show that the map

$$(43)\qquad \delta c \mapsto \int \left(\Xi(z) H(z, 0) \int H(0, z') \mathcal{I}_2 \mathcal{I}_1 \delta c \, \mathrm{d}z' \right) (x - \tfrac{h}{2}, x + \tfrac{h}{2}, ph) \, \mathrm{d}h,$$

microlocally has principal symbol equal to 1. In this proof we will omit the cutoff functions that are part of the symbols; the calculations will be valid microlocally on the support of a cutoff.

Using an oscillatory integral representation of H similar to the one in the proof of Theorem 6.1, we find that the principal contribution to the kernel of this map, as a function of $(x,z,p;x',z')$, can be written as

$$(2\pi)^{-(2n-1)}\int \Xi(x-\tfrac{1}{2}h, x+\tfrac{1}{2}h, ph, -\frac{\partial S}{\partial s}, -\frac{\partial S}{\partial r}, -\frac{\partial S}{\partial t}, z)$$
$$\times\, \overline{A(z, x-\tfrac{1}{2}h, x+\tfrac{1}{2}h, ph, y_0, \eta_{0J})} A(z', x', x', 0, y_0, \eta_{0J})$$
$$\times\, \mathrm{e}^{\mathrm{i}[-S(z,x-\frac{1}{2}h,x+\frac{1}{2}h,ph,y_{0I},\eta_{0J})+S(z',x',x',0,y_{0I},\eta_{0J})]}\, \mathrm{d}y_{0I}\mathrm{d}\eta_{0J}\,\mathrm{d}h.$$

We expand the phase in a Taylor series around $(x',z',h)=(x,z,0)$ and identify the gradient at $(x,z,0)$,

$$-\frac{\partial S}{\partial x}(z,x,x,0,y_{0I},\eta_{0J}) = \sigma(z,x,x,0,y_{0I},\eta_{0J}) + \rho(z,x,x,0,y_{0I},\eta_{0J}),$$
$$-\frac{\partial S}{\partial z}(z,x,x,0,y_{0I},\eta_{0J}) = \zeta(z,x,x,0,y_{0I},\eta_{0J}),$$
$$-\frac{\partial S}{\partial h}(z,x,x,0,y_{0I},\eta_{0J})$$
$$= -\tfrac{1}{2}\sigma(z,x,x,0,y_{0I},\eta_{0J}) + \tfrac{1}{2}\rho(z,x,x,0,y_{0I},\eta_{0J}) + p\tau(z,x,x,0,y_{0I},\eta_{0J}).$$

Applying a change of variables, $(y_{0I},\eta_{0J}) \mapsto (\sigma,\rho,\zeta)$, the phase takes the form

$$\langle \sigma+\rho, x-x'\rangle + \zeta(z-z') + \langle \tfrac{1}{2}(\rho-\sigma), h\rangle.$$

The amplitude factor $\Xi\overline{A}A$ becomes equal to one by the calculations in the proof of Theorem 6.1. Upon changing integration variables, $\sigma = \frac{1}{2}\xi - \theta, \rho = \frac{1}{2}\xi + \theta$, the oscillatory integral takes the leading-order form (microlocally)

$$(2\pi)^{-(2n-1)}\int \mathrm{e}^{\mathrm{i}(\langle\xi,x-x'\rangle+\zeta(z-z')+\langle\theta,h\rangle)}\,\mathrm{d}\xi\mathrm{d}\zeta\mathrm{d}\theta\mathrm{d}h.$$

It follows by integrating the θ, h variables that (43) indeed has principal symbol equal to 1 microlocally. □

8. ANNIHILATORS

We address the inverse problem of determining the background medium c_0. This inverse problem can be addressed by making use of the redundancy in the data. If the data are in the range of the operator F (cf. (24)) we accept the model c_0 as a suitable background medium. In general, the model c_0 is not uniquely determined by the data. It was observed in [22] that if the data are in the range of F, they are annihilated by appropriate pseudodifferential operators microlocally. These operators, annihilators, are here derived from the angle transform introduced in the previous section. If the data are in the range of F, the angle transform generates a collection of identical images of δc, parameterized by p_i, $i=1,\ldots,n-1$. Hence, on transformed data $\widetilde{A}_{\mathrm{WE}}d$, the annihilators are given by $\frac{\partial}{\partial p_i}$, $i=1,\ldots,n-1$. It follows that the annihilators of the data are given by

$$\langle \widetilde{A}_{\mathrm{WE}}\rangle^{-1}\frac{\partial}{\partial p_i}\widetilde{A}_{\mathrm{WE}},$$

Where $\langle \widetilde{A}_{\rm WE}\rangle^{-1}$ is a regularized inverse for $\widetilde{A}_{\rm WE}$, that is an inverse microlocally for a subset of $T^*\mathbb{R}^{2n-1}_{(x,z,p)}$ where $\widetilde{A}_{\rm WE}$ is invertible.

We define the operator $\widetilde{K}^*$ that maps d to a function of (s, r, z) by

$$(44)\quad (\widetilde{K}^* d)(s, r, z) = 2c_0((s + r)/2, z)^3 \\ \times \big(\mathcal{R}_2 \Xi(z) Q^*_{-,s}(z)^{-1} Q^*_{-,r}(z)^{-1} \psi_2(z, 0) H(z, 0) Q_{-,s}(0)^{-1} Q_{-,r}(0)^{-1} D_t^{-2} \psi_Y d\big)(s, r, z).$$

Using (34), we observe that the operator $\widetilde{K}^*$ acting on $d = F\delta c$ yields

$$\widetilde{K}^* d = \Psi(s, r, z, D_s, D_r, D_z)\mathcal{I}_1 \delta c, \tag{45}$$

where Ψ is a pseudodifferential cutoff the principal symbol of which was already described in the remarks above Theorem 6.1). Applying the operator $\mathcal{M}$ given by the multiplication by $r - s$ to this equation yields

$$\mathcal{M}\widetilde{K}^* d = [\mathcal{M}, \Psi(s, r, z, D_s, D_r, D_z)]\, \mathcal{I}_1 \delta c, \tag{46}$$

i.e. there is only a lower order contribution, since $(r - s)\delta(r - s) = 0$ and hence $\mathcal{M}\mathcal{I}_1 = 0$ (cf. (22)).

For the subset in $T^*\mathbb{R}^{2n-1}_{(s,r,z)}$ where the operator $\Psi(s, r, z, D_s, D_r, D_z)$ is microlocally elliptic, the operator

$$\widetilde{\mathcal{M}} := \mathcal{M} - [\mathcal{M}, \Psi(s, r, z, D_s, D_r, D_z)]\Psi(s, r, z, D_s, D_r, D_z)^{-1} \tag{47}$$

is an annihilator of $\widetilde{K}^* d$ to all orders. We summarize this result in the following

Theorem 8.1. *A pseudodifferential annihilator of the data is given by*

$$W = K\widetilde{\mathcal{M}}\widetilde{K}^*. \tag{48}$$

Note that $W = W[c_0]$ depends on the background medium. The semi-norm $\|W[c_0]d\|$ detects whether c_0 was an acceptable choice or not. It can be viewed as the wave equation analog of the differential semblance functional [23]. By using this semi-norm the inverse problem of determining the background medium c_0 can be cast in an optimization problem.

Appendix A. First order pseudodifferential equations with order δ dissipative term

Let z be in an interval $Z =]z_0, \infty[$, $x \in \mathbb{R}^n$ and let

$$P(z, x, D_z, D_x) = \frac{1}{\mathrm{i}}\frac{\partial}{\partial z} - B(z, x, D_x)$$

where $B(z, x, D_x)$ is a standard pseudodifferential operator with real homogeneous principal symbol $b(z, x, \xi)$ of order 1. Let C be a standard pseudodifferential operator of order δ, $0 \le \delta < \frac{1}{2}$, with real homogeneous principal symbol $c(z, x, \xi) \ge 0$. In this appendix we consider differential/pseudodifferential equations of the form

$$(P(z, x, D_z, D_x) - \mathrm{i}C(z, x, D_x))u(z, x) = 0. \tag{A.1}$$

We consider the initial value problem with an initial condition at the hyperplane $z = z_0$

$$u(z_0, \cdot) = u_0, \tag{A.2}$$

where $u_0 \in \mathcal{E}'(\mathbb{R}^n)$. The equation is dissipative, microlocally where $C > 0$. Using energy estimates one can show that there is a well defined solution for $z > z_0$ [12, section 23.1]. We will denote the solution operator by $E : \mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(]z_0, \infty[\times\mathbb{R}^n)$. When $C = 0$ there is also a solution for $z < z_0$, and in that case it is well known that the solution operator is a Fourier integral operator [8, section 5.1] (note that $B(z, x, D_x)$ is not strictly a pseudodifferential operator in (z, x), but a z-family of pseudodifferential operators in x, but the same arguments can be used.) Here we will construct in a similar way a Fourier integral operator that solves the initial value problem for the case $C \geq 0$, $z > z_0$, and we compare this Fourier integral operator with the one for the case $C = 0$. Our objective is to show that (A.1) in principle yields a way to construct a regularized solution operator as described in section 4. Although problems of a similar nature have been treated in the literature, we have not found the result presented below.

For $C = 0$ the singularities (elements of the wave front set of the solution) propagate on the characteristic surface given by $\zeta = b(z, x, \xi)$, according to the Hamilton vector field of P, which reads

$$\frac{\partial}{\partial z} - \frac{\partial b}{\partial \xi}(z, x, \xi) \cdot \frac{\partial}{\partial x} + \frac{\partial b}{\partial x} \cdot \frac{\partial}{\partial \xi}.$$

It follows that ζ is invariant under the flow and that the solution curves, the bicharacteristics, can be parameterized by z. We denote the (x, ξ) components by $(\mathbf{x}(z, z_0; x_0, \xi_0)$, $\xi(z, z_0; x_0, \xi_0))$, and the curves by $\gamma(z, z_0; x_0, \xi_0) = (z, \mathbf{x}(z, z_0; x_0, \xi_0), \xi(z, z_0; x_0, \xi_0))$.

Consider the canonical relation

$$\begin{aligned} C_0 = \{(\mathbf{x}(z, z_0; x_0, \xi_0), z, \xi(z, z_0; x_0, \xi_0), b(z_0, x_0, \xi_0); x_0, \xi_0) \mid \\ (x_0, \xi_0) \in \mathbb{R}^n \times \mathbb{R}^n\backslash 0, z \in]z_0, \infty[\}. \end{aligned} \tag{A.3}$$

Let

$$I(z, z_0, x_0, \xi_0) = \int_{z_0}^{z} c(\gamma(z, z_0; x_0, \xi_0))\, \mathrm{d}z',$$

and let $\widetilde{I}(z, z_0, x, \xi) = I(z, z_0, \mathbf{x}(z_0, z; x, \xi), \xi(z_0, z; x, \xi))$. We will show that the map E is a Fourier integral operator, such that in the symbol or amplitude a factor $\exp(-I(z_0, z, x_0, \xi_0))$ occurs. For bicharacteristics that travel some finite distance in the region of phase space where $C > 0$, this leads to exponential decay of the symbol, that is hence in $S^{-\infty}$ for corresponding elements of the canonical relation. An additional consequence is that the amplitude will no longer satisfy standard symbol estimates, but will be in some class S^{μ}_{ρ}, where $\rho = 1 - \delta$, which is defined by the symbol estimates

$$\Big|\frac{\partial^\alpha}{\partial y^\alpha} \frac{\partial^\beta}{\partial \theta^\beta} a(y, \theta)\Big| \leq C_{\alpha,\beta}(1 + \|\theta\|)^{\mu+|\alpha|(1-\rho)-|\beta|\rho}. \tag{A.4}$$

(The case $\rho = 1$ is of course the standard case.) The result of this appendix, that we present now, shows that E is really a regularized solution operator of the type described in section 4, i.e. can be written as E_0 followed and/or preceded by appropriate pseudodifferential cutoff operators.

Theorem A.1. *Let P, C be as described and let $\rho = 1 - \delta$. The map E from the initial value u_0 to the solution u of (A.1) and (A.2) is a Fourier integral operator in*

$I_\rho^{-1/4}(]z_0,\infty[\times\mathbb{R}^n,\mathbb{R}^n;C_0)$. *If E_0 denotes the solution operator with $C=0$, then there is a pseudodifferential operator A with principal symbol 1 such that*

$$E=\exp(-\widetilde{I}(z,z_0,x,D_x))AE_0. \tag{A.5}$$

In the following we will denote by p_1 the unique homogeneous principal symbol of P, by $p=P(z,x,\zeta,\xi)-\frac{1}{2i}\sum_{j=1}^{n+1}\frac{\partial^2 P}{\partial(z,x)_j\partial(\zeta,\xi)_j}$ the invariantly defined principal symbol of P, and by $p_0=p-p_1$ the subprincipal symbol [8, chapter 4]. We wil use the notation $Z=]z_0,\infty[$ and $X=\mathbb{R}^n$.

Proof. To prove the result we follow the steps in [8, sections 4.3, 5.1], where for the proof of Theorem 4.3.2 we follow the reference to [7, proof of Theorem 5.3.1].

First we observe that Theorem 4.3.2 in [8] is still valid for the operator P. So if $E\in I_\rho^{m'}(Z\times X,X;C_0)$ and $e\in S_\rho^{m'+(2n+1)/4)}(C_0,\Omega_{\frac{1}{2}}\otimes L_{C_0})$ is a principal symbol of E, then PE is a Fourier integral operator in $I_\rho^{m'+1-\rho}(Z\times X,X;C_0)$ with principal symbol $\frac{1}{i}\mathcal{L}_{H_{p_1}}e+p_0e$; here, $\mathcal{L}_{H_p}$ denotes the Lie derivative along H_p lifted to the canonical relation C_0 via the projection on the first factor. The composition CE is a Fourier integral operator with principal symbol ce that is also in $I_\rho^{m'+1-\rho}(Z\times X,X;C_0)$. It follows that $(P-\mathrm{i}C)E$ is a Fourier integral operator in $I_\rho^{m'+1-\rho}(Z\times X,X;C_0)$ with principal symbol

$$\frac{1}{i}\mathcal{L}_{H_{p_1}}e+(-\mathrm{i}c+p_0)e. \tag{A.6}$$

From the definition (A.3) it follows that (z,x_0,ξ_0) are coordinates on C_0, and therefore we may use these as coordinates for the principal symbol, so we will use the notation $e(z,x_0,\xi_0)$. In these coordinates the Lie derivative $\mathcal{L}_{H_{p_1}}e$ of e is given by

$$(\mathcal{L}_{H_{p_1}}e)(z,x_0,\xi_0)=\frac{\partial e}{\partial z}(z,x_0,\xi_0). \tag{A.7}$$

We define an operator $E^{(0)}$ with principal symbol $e^{(0)}$ such that (A.6) vanishes, so that $B^{(1)}$ defined by

$$B^{(1)}=(P-\mathrm{i}C)E^{(0)} \tag{A.8}$$

is in $I_\rho^{m'+1-\rho-(2\rho-1)}(Z\times X,X;C_0)$, with principal symbol $b^{(1)}$. We will then search lower order contributions $E^{(j)}\in I_\rho^{m'-j(2\rho-1)}(Z\times X,X;C_0)$, $j=1,2,\ldots$, such that the principal symbol of $(P-\mathrm{i}C)E^{(j)}$, given by (A.6), is equal to $b^{(j)}$; hence, $B^{(j+1)}$ defined by

$$B^{(j+1)}=(P-\mathrm{i}C)E^{(j)}-B^{(j)} \tag{A.9}$$

is in $I_\rho^{m'+1-\rho-(j+1)(2\rho-1)}(Z\times X,X;C_0)$. Then, by adding the symbols of the $E^{(j)}$, the symbol of a parametrix is obtained.

To prove the first part of the theorem it hence remains to show that such $e^{(j)}$ can be found. By (A.6) and (A.7), equations (A.8) and (A.9) result in the following differential equation for for the $e^{(j)}$,

$$\left(\frac{\partial}{\partial z}+(c+\mathrm{i}p_0)(z,\mathbf{x}(z,z_0;x_0,\xi_0),\xi(z,z_0;x_0,\xi_0))\right)e^{(j)}(z,x_0,\xi_0)=b^{(j)}(z,x_0,\xi_0), \tag{A.10}$$

where $b^{(0)}(z, x_0, \xi_0) = 0$. The solution to this equation is given by

(A.11)
$$e^{(j)}(z, x_0, \xi_0) = e^{(j)}(z_0, x_0, \xi_0) \exp\left(- I(z_0, z, x_0, \xi_0) - \mathrm{i} \int_{z_0}^{z} p_0(\gamma(z' z_0; x_0, \xi_0))\, \mathrm{d}z' \right) + \int_{z_0}^{z} b^{(j)}(\tilde{z}, x_0, \xi_0) \exp\left(- I(\tilde{z}, z, x_0, \xi_0) - \mathrm{i} \int_{\tilde{z}}^{z} p_0(\gamma(z' z_0; x_0, \xi_0))\, \mathrm{d}z' \right) \mathrm{d}\tilde{z}.$$

For the multiple derivatives of an exponential there is the formula

(A.12)
$$\frac{\partial^\alpha}{\partial y^\alpha} \mathrm{e}^{f(y)} = \sum_{\{\alpha^{(1)},\dots,\alpha^{(k)}\} \text{ partition of } \alpha} c_{\{\alpha^{(1)},\dots,\alpha^{(k)}\}} \left(\prod_{j=1}^{k} \frac{\partial^{\alpha^{(j)}} f}{\partial y^{\alpha^{(j)}}} \right) \mathrm{e}^{f(y)},$$

for some integer constants $c_{\{\alpha^{(1)},\dots,\alpha^{(k)}\}}$. It follows that if $e^{(j)}|_{z=z_0} \in S_\rho^{m'+(2n+1)/4-j(2\rho-1)}$ and $b^{(j)} \in S_\rho^{m'+(2n+1)/4+1-\rho-j(2\rho-1)}$ then indeed $e^{(j)}$ obtained by (A.11) is a symbol in $S_\rho^{m'+(2n+1)/4-j(2\rho-1)}$. Hence $b^{(j+1)} \in S_\rho^{m'+(2n+1)/4+1-\rho-(j+1)(2\rho-1)}$, and by induction the $e^{(j)}$ are all in the right symbol space. Like in [8] the initial condition (A.2) now yields the values of $e^{(j)}|_{z=z_0}$, and the order of $-1/4$ of the Fourier integral operator.

For the second statement observe that $e^{(0)}(z, x_0, \xi_0) = \exp(-I(z, z_0, x_0, \xi_0)) e_0(z, x_0, \xi_0)$. Then $E^{(0)}$ has a decomposition as in (A.5). From (A.10) it follows that if

(A.13)
$$\exp(I(z, z_0, x_0, \xi_0)) b^{(j)} \in S_\rho^{m'+(2n+1)/4+1-\rho-j(2\rho-1)}$$

(the same symbol space as $b^{(j)}$ itself) then $\exp(I(z, z_0, x_0, \xi_0)) e^{(j)} \in S_\rho^{m'+(2n+1)/4-j(2\rho-1)}$. Since the principal symbol of E_0 is nonzero, we can write $E^{(j)}$ to highest order as the product of E_0 and $\exp(-\widetilde{I}(z, z_0, x, D_x))$ and a lower order pseudodifferential operator that contributes to A in (A.5). Then (A.13) is true with $j+1$ instead of j and the statement follows by induction. □

REFERENCES

[1] G. Beylkin, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform, J. Math. Phys., 26 (1985), pp. 99-108.

[2] S. Brandsberg-Dahl, M. V. de Hoop, and B. Ursin, Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers, Geophysics, 2002, In print.

[3] J. F. Claerbout, Imaging the Earth's Interior, Blackwell Scientific Publications, Oxford, 1985.

[4] R. W. Clayton, Common midpoint migration, Technical Report SEP-14, Stanford University, 1978.

[5] C. G. M. de Bruin, C. P. A. Wapenaar, and A. J. Berkhout. Angle-dependent reflectivity by means of prestack migration. *Geophysics*, 55(09):1223–1234, 1990.

[6] M. V. de Hoop, J. H. Le Rousseau, and R.-S. Wu, Generalization of the phase-screen approximation for the scattering of acoustic waves, Wave Motion, 31 (2000), pp. 43-70.

[7] J. J. Duistermaat and L. Hörmander. Fourier integral operators II. *Acta. Math.*, 128:183–269, 1972.

[8] J. J. Duistermaat, Fourier Integral Operators, Birkhäuser, Boston, MA, 1996.

[9] V. Guillemin, On some results of Gel′fand in integral geometry, in Pseudodifferential operators and applications (Notre Dame, Ind., 1984), Amer. Math. Soc., Providence, RI, 1985, pp. 149-155.

[10] S. Hansen, Solution of a hyperbolic inverse problem by linearization, Comm. Partial Differential Equations, 16 (1991), pp. 291–309.

[11] Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 1. Springer-Verlag, Berlin, 1983.

[12] Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 3. Springer-Verlag, Berlin, 1985.

[13] L. Hörmander, The Analysis of Linear Partial Differential Operators, volume 4, Springer-Verlag, Berlin, 1985.

[14] A. P. E. ten Kroode, D. J. Smit, and A. R. Verdel, A microlocal analysis of migration, Wave Motion, 28 (1998), pp. 149-172.

[15] A. Melin and J. Sjöstrand, Fourier integral operators with complex-valued phase functions, in Fourier integral operators and partial differential equations (Colloq. Internat., Univ. Nice, Nice, 1974), Lecture Notes in Math., Vol. 459, Springer-Verlag, Berlin, 1975, pp. 120-223.

[16] C. J. Nolan and W. W. Symes, Global solution of a linearized inverse problem for the wave equation, Comm. Partial Differential Equations, 22 (1997), pp. 919-952.

[17] A. M. Popovici, Prestack migration by split-step DSR, Geophysics, 61 (1996), pp. 1412-1416.

[18] M. Prucha, B. Biondi, and W. Symes, Angle-domain common image gathers by wave-equation migration, in 69th Ann. Internat. Mtg., Soc. Of Expl. Geophys., 1999, pp. 824-827.

[19] Rakesh, A linearized inverse problem for the wave equation, Comm. Partial Differential Equations, 13 (1988), pp. 573-601.

[20] C. C. Stolk, Microlocal analysis of a seismic linearized inverse problem, Wave Motion, 32 (2000), pp. 267-290.

[21] C. C. Stolk, Microlocal analysis of the scattering angle transform, Preprint, Dept. of Computational and Applied Mathematics, Rice University, http://www.caam.rice.edu/˜cstolk/angle.ps, July 2001.

[22] C. C. Stolk and M. V. de Hoop, Microlocal analysis of seismic inverse scattering in anisotropic elastic media, Comm. Pure Appl. Math., 55 (2002), pp. 261-301.

[23] W. W. Symes, A differential semblance algorithm for the inverse problem of reflection seismology, Comput. Math. Appl., 22 (1991), pp. 147-178.

[24] M. E. Taylor, Reflection of singularities of solutions to systems of differential equations, Comm. Pure Appl. Math., 28 (1975), pp. 457-478.

[25] M. E. Taylor, Pseudodifferential Operators and Nonlinear PDE, Birkhäuser, Boston, MA, 1991.

[26] François Treves. Introduction to Pseudodifferential and Fourier Integral Operators, volume 2. Plenum Press, New York, NY, 1980.

[27] Sheng Xu, Hervé Chauris, Gilles Lambaré, and Mark Noble. Common angle migration: A strategy for imaging complex media. *Geophysics*, 66(6):1877–1894, 2001.

CHRISTIAAN C. STOLK, THE MATHEMATICAL SCIENCES RESEARCH INSTITUTE, 1000 CENTENNIAL DRIVE #5070, BERKELEY CA 94720-5070 USA

E-mail address: cstolk@caam.rice.edu

MAARTEN V. DE HOOP, THE MATHEMATICAL SCIENCES RESEARCH INSTITUTE, 1000 CENTENNIAL DRIVE #5070, BERKELEY CA 94720-5070 USA

E-mail address: mdehoop@Mines.EDU