Modeling and inversion of seismic data in anisotropic elastic media

Christiaan C. Stolk* and Maarten V. de Hoop†

March 31, 2000

Abstract

Seismic data is modeled in the high frequency limit. We consider general anisotropic media, and our method is also valid in the case of multipathing (caustics). The data is modeled in two ways. First using the Kirchhoff approximation (where the medium is assumed to be piecewise smooth, and reflection and transmission occurs at the interface). Secondly the data is modeled using the Born approximation, in other words by a linearization in the medium parameters.

The main result is a characterisation of seismic data. We construct a Fourier integral operator and a "reflectivity function", which is a function of subsurface position and scattering angle and azimuth, such that the data is given by the invertible Fourier integral operator acting on the reflectivity function.

Using this new transformation of seismic data to subsurface position/angle coordinates we obtain the following results on the problem of reconstructing the medium coefficients. Given the medium above the interface in the Kirchhoff approximation one can reconstruct the position of the interface and the angular dependent reflection coefficients on the interface. We also obtain a criterium to determine whether the medium above the interface (the background medium in the Born approximation) is correctly chosen. These results are new in medium with caustics. In the Born approximation the singular medium perturbation can be reconstructed.

Acknowledgments We thank Prof. J.J. Duistermaat for many helpful discussions. The first author thanks Shell International Exploration and Production for their financial support.

*Mathematics Department, Utrecht University, Budapestlaan 6, 3584 CD Utrecht, The Netherlands, email: stolk@math.uu.nl

†Center for Wave Phenomena, Colorado School of Mines, Golden CO 80401, USA, email: mdehoop@dix.mines.edu

1 Introduction

In the seismic experiment one generates elastic waves in the earth using sources at the surface. The waves that return to the surface of the earth are observed (in fact sources and receivers are not always on the surface of the earth, this case is also considered). The problem is to reconstruct the elastic properties of the subsurface from the data thus obtained.

The subsurface is given by an open set $X \subset \mathbb{R}^n$. In practice $n = 2$ or 3, but we leave it unspecified. Sources and receivers are contained in the boundary ∂X of X. Their position is denoted by $\tilde{x}, \hat{x}$. Measurement of data takes places during a time interval $]0, T[$. The set of $(\hat{x}, \tilde{x}, t)$ for which data is taken is called the acquisition manifold Y'. We assume that the displacement of the medium is measured for pointsources at $\tilde{x}, t = 0$ and that data is taken for all the elastic components, both at the source and at the receiver. Thus we assume that (after preprocessing) the data is given by the Green's function $G_{il}(\hat{x}, \tilde{x}, t)$, for $(\hat{x}, \tilde{x}, t) \in Y'$.

We refer to the codimension of the set of $Y' \in \partial X \times \partial X \times]0, T[$ as the codimension of the acquisition manifold and we denote it by c. For example in marine data the receivers are along a line behind the source and we may have $n = 3, c = 1$, $\partial X = \{x \in \mathbb{R}^n \mid x_3 = 0\}$, $Y' = \{(\hat{x}, \tilde{x}, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times]0, T[\mid \hat{x}_3 = \tilde{x}_3 = \hat{x}_2 - \tilde{x}_2 = 0\}$. So the data is a function of $2n - 1 - c$ variables. From this data we want to determine a function of n variables, hence there is a redundancy in the data of dimension $n - 1 - c$.

We analyze the highfrequency content of the data. Highfrequency methods (in particular ray theory) are applied very often to seismic data, and turn out to be successful. We use the methods of microlocal analysis, see Hörmander [12], Duistermaat [10], Treves [24, 25].

The data is modeled for general elastic media, allowing for multipathing (leading to caustics). We model the data in two ways. In Section 3 we assume that the medium consists of different pieces with smooth interfaces between the different pieces. The medium parameters are assumed to be smooth on each piece, and smoothly extendible across each interface, but they vary discontinuously at the interface. We discuss how to model the high frequency part of the data using Fourier integral operators, which is new for seismic data. Data modeled in this way are called Kirchhoff data. In Section 4 we discuss the Born approximation. This is essentially a linearization, where the mediumparameters are written as the sum of a background medium and a perturbation that is assumed to be small. It is assumed that the background is smooth and that the perturbation contains the singularities of the medium.

The main result is the characterization of seismic data in Theorem 6.2. The data can be written as an invertible Fourier integral operator H_{MN} acting on a "reflectivity" function $r_{MN}(x, e)$, that is a function of subsurface position x and an additional variable e, essentially parametrizing the scattering angle and azimuth. In the Kirchhoff approximation the function $r_{MN}(x, e)$ equals to highest order $R_{MN}(x, e)\delta(z_n(x))$, where $R_{MN}(x, e)$ is the appropiately normalized reflection coefficient for the pair of modes (M, N) and $\delta(z_n(x))$ is the singular function of the interface. For the Born approximation something similar holds. The result holds microlocally away from points in the cotangent space $T^*Y' \backslash 0$ that violate Assumptions 1 to 5, discussed in the text.

As a consequence of Theorem 6.2 we obtain results about the reconstruction of the mediumparameters. Given the medium above the interface the function $r_{MN}(x,e)$ and hence the position of the interface and the reflection coefficients can be reconstructed by acting with the inverse H_{MN}^{-1} on the data, see Corollary 6.3. For the Born approximation a similar result holds, but an inverse is also obtained directly in Theorem 4.4.

When the data is redundant there is in addition a criterium to determine whether the medium above the interface (the background medium in the Born approximation) is correctly chosen. The position of the singularities of the function $r_{MN}(x,e)$, obtained by acting with H_{MN}^{-1} on the data, should not depend on e. There exist pseudodifferential operators $Q_{MN}(y', D_{y'})$ that, if the medium above the interface is correctly chosen, annihilate the data, see Corollary 7.1.

The exact choice of the variable e is unspecified. When multipathing occurs a suitable choice is the scattering angle, because in these coordinates the caustics are "unfolded". In that case the operator H_{MN} transforms the data to subsurface position and scattering angle coordinates, which is new. In other cases one can use for instance the offset (difference between source and receiver coordinates).

We discuss some of the literature on this subject. There have been many publications about highfrequency methods to invert seismic data in acoustics media. The reconstruction of the singular component of the medium coefficients in the Born approximation, without caustics has been done in the paper by Beylkin [2]. Bleistein [4] discusses the case of a smooth jump using Beylkin's results. It has been shown by Rakesh [19] that the modeling operator in the Born approximation is a Fourier integral operator. Hansen studied the inversion in an acoustic medium with multipathing for both the Born approximation and the case of a smooth jump. Ten Kroode, Smit and Verdel [23] also treat the case of seismic imaging in the presence of multipathing. They discuss in detail the assumptions (most importantly Assumption 5ii) below) that are made about the geometry of the rays involved in the scattering. Stolk [20] discusses the case when this assumption is violated. Nolan and Symes discuss the imaging with different acquisition geometries. The article by Symes [21] discusses the reconstruction of the background medium in the Born approximation.

The mathematical treatment of systems of equations, such as the elastic equations, in the highfrequency approximation has been described by Taylor [22]. This fundamental paper also discusses the interface problem. Beylkin and Burridge [3] discuss the imaging of seismic data in the Born approximation in isotropic elastic media, under a no caustics assumption. De Hoop and Bleistein [7] discuss the imaging in general anisotropic elastic media, using a Kirchhoff type approximation. The Born approximation in anisotropic elastic media allowing for multipathing is discussed by De Hoop and Brandsberg-Dahl [8].

We give an overview of the paper. In Section 2 we discuss the propagation of waves in smooth elastic media. First we discuss how in asymptotically the elastic system can be decoupled by conjugating with pseudodifferential operators (a technique that is common in mathematics, but not in the seismic literature). Then we discuss the construction of asymptotic solutions for the decoupled equations using Fourier integral operators.

In Section 3 we discuss the reflection and transmission of waves at a smooth interface. We explicitly construct Fourier integral operator solutions describing reflected and transmitted waves. These solutions where already discussed, but not explicitly constructed, by Taylor [22]. Thus we prove directly the validity of the Kirchhoff approximation, which is not obvious from e.g. De Hoop and Bleistein [7].

In Section 4 we discuss the modeling and inversion of seismic data in the Born approximation. This is important both in its own right and for the reconstruction problem if we model using a smooth jump. We give an efficient presentation for the case of general anisotropic media with general acquisition geometry. We discuss in detail the assumptions that are needed.

In Section 5 we essentially discuss the geometry of the wave front set of the data. Under the assumptions of Section 4 this set is contained in a coisotropic submanifold L of the cotangent space $T^*Y'\backslash 0$. We discuss the extension of symplectic coordinates on L to a neighborhood of L.

After the preparations of Sections 2 to 5 the derivation of our main result in Section 6 is relatively simple. We discuss a characterization of seismic data and some consequences, in particular the reconstruction of the position of the interface and the reflection coefficients given the medium above the interface.

Finally in Section 7 we discuss the reconstruction of the smoothly varying medium parameters above the interface (or of the background medium in the Born approximation).

2 The elastic wave equation with smooth coefficients

2.1 Decoupling the elastic equations

The elastic wave equation is given by

$$\left(\rho\,\delta_{il}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}c_{ijkl}\frac{\partial}{\partial x_k}\right)(\text{displacement})_l = (\text{vol. force density})_i. \tag{1}$$

Here $\rho(x)$ is the volume density of mass and $c_{ijkl}(x)$ is the elastic stiffness tensor, and $i, j, k, l = 1, \ldots, n$.

In order to diagonalize this system, thus decoupling the modes, it is convenient to remove the x-dependent coefficient ρ in front of the time derivative. Thus we introduce the equivalent system

$$P_{il}u_l = f_i, \tag{2}$$

where

$$u_l = \sqrt{\rho}(\text{displacement})_l, \qquad f_i = \frac{1}{\sqrt{\rho}}(\text{force density})_i, \tag{3}$$

and

$$P_{il} = \delta_{il}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}\frac{c_{ijkl}}{\rho}\frac{\partial}{\partial x_k} + \text{l.o.t.}\ . \tag{4}$$

Here we use that ρ is smooth and bounded away from zero. Both systems (1) and (2) are real, time reversal invariant, and satisfy reciprocity.

We describe how the system (2) can be decoupled by transforming it with appropriate pseudodifferential operators see Taylor [22], Dencker [9]. It turns out that microlocally, away from certain exceptional points in $T^*X\backslash 0$, there are a matrix valued pseudodifferential operator $Q(x, D)_{iM}$, $D = -i\frac{\partial}{\partial x}$, and scalar pseudodifferential operators $P_M(x, D, D_t)$ such that

$$Q(x, D)^{-1}_{Mi}\,P_{il}(x, D, D_t)\,Q_{lN}(x, D) = \text{diag}(P_M(x, D, D_t)\,;\,M = 1, \ldots, n)_{MN}. \tag{5}$$

Here the indices M, N denote the mode of propagation, they range from 1 to n. Let

$$u_M = Q(x, D)^{-1}_{Mi}u_i, \qquad f_M = Q(x, D)^{-1}_{Mi}f_i. \tag{6}$$

The system (2) is then equivalent to the n scalar equations

$$P_M(x, D, D_t)u_M = f_M. \tag{7}$$

The time derivative in P_{il} is already on diagonal form, hence we only have to diagonalize the spatial part

$$A_{il}(x, D) = -\frac{\partial}{\partial x_j}\frac{c_{ijkl}}{\rho}\frac{\partial}{\partial x_k} + \text{l.o.t.}\ .$$

So we have to find Q_{iM} and A_M such that (5) is valid with P_{il}, P_M replaced by A_{il}, A_M. The operator P_M is then given by

$$P_M(x, D, D_t) = \frac{\partial^2}{\partial t^2} + A_M(x, D).$$

The principal symbol $A_{il}^{\text{prin}}(x, \xi)$ is a positive symmetric matrix, so it can be diagonalized by an orthogonal matrix. On the level of principal symbols, composition of pseudodifferential operators is given by multiplication. Therefore, we let $Q_{iM}^{\text{prin}}(x, \xi)$ be this orthogonal matrix, and we let $A_M^{\text{prin}}(x, \xi)$ be the eigenvalues, so that

$$Q_{Mi}^{\text{prin}}(x, \xi)^{-1} A_{il}^{\text{prin}}(x, \xi) Q_{lN}^{\text{prin}}(x, \xi) = \text{diag}(A_M^{\text{prin}}(x, \xi)\,;\, M = 1, \dots, n)_{MN}. \tag{8}$$

The principal symbol $Q_{iM}^{\text{prin}}(x, \xi)$ is the matrix, that has as its columns the orthonormalized polarization vectors associated with the modes of propagation.

If the multiplicities of the eigenvalues are constant then $Q^{\text{prin}}(x, \xi)_{iM}$ depends smoothly on (x, ξ) and microlocally equation (8) carries over to an operator equation. Taylor [22] has shown that if this condition is satisfied then decoupling can be accomplished to all orders. We summarize this result in the following lemma.

Lemma 2.1 *Suppose the multiplicities of the eigenvalues of $A_{il}(x, \xi)$ are equal to one on some neighborhood. Then we can find pseudodifferential operators $Q_{iM}(x, D)$, $A_M(x, D)$ with principal symbol as described above such that microlocally (5) is valid.*

Remark 2.2 For generic elastic systems the case where the multiplicity of an eigenvalue is equal to two is investigated in Braam and Duistermaat [5]. They give a normal form for such systems and investigate the behavior of bicharacteristics and polarization spaces. In this case the system cannot be decoupled. On the other hand if the multiplicities are constant, but not equal to 1 such as in the isotropic elastic case, then the system can still be decoupled with the right hand side of (5) replaced by a blockdiagonal matrix, each block corresponding to a different eigenvalue.

The second order equations (7) clearly describe the decoupling of the original system in n elastic modes. In addition equations (7) inherit the symmetries of the original system. To start it is easy to see that they are time reversal invariant. The operators Q, A can be chosen such that $Q_{iM}(x, \xi) = -\overline{Q}_{iM}(x, -\xi), A_M(x, \xi) = \overline{A}_M(x, \xi)$. This means that Q_{iM}, A_M are real. We argue that they also satisfy reciprocity. For the causal Green's function $G_{ij}(x, x_0, t - t_0)$ reciprocity means that $G_{ij}(x, x_0, t - t_0) = G_{ji}(x_0, x, t - t_0)$. We show that such a relationship holds (modulo smoothing operators) for the Green's function $G_M(x, x_0, t - t_0)$. The transpose operator $Q(x, D)_{Mi}^t$ (obtained by interchanging x, x_0 and i, M in the distribution kernel $\mathcal{Q}_{iM}(x, x_0)$ of $Q_{iM}(x, D)$) is also a pseudodifferential operator, with principal symbol $Q(x, \xi)_{Mi}^t$. It follows from the fact that $A_{ij}^t = A_{ij}$ that we can choose Q orthogonal, i.e. such that $Q(x, D)_{iM} Q(x, D)_{Mj}^t = \delta_{ij}$. From the fact that

$$G_M(x, x_0, t - t_0) = Q(x, D)_{Mi}^{-1} G_{ij}(x, x_0, t - t_0) Q(x_0, D_{x_0})_{jN}$$

it follows that microlocally G_M is reciprocal, $G_M(x, x_0, t - t_0) = G_M(x_0, x, t - t_0)$, modulo smoothing operators.

The values $\tau = \pm\sqrt{A_M^{\mathrm{prin}}(x, \xi)}$ are precisely the solutions to the equation

$$\det P_{il}^{\mathrm{prin}}(x, \xi, \tau) = 0. \tag{9}$$

Because $P_{il}^{\mathrm{prin}}(x, \xi, \tau)$ is homogeneous in ξ, τ, one often uses the slowness $-\tau^{-1}\xi$ in calculations. The set of $-\tau^{-1}\xi$ such that (9) holds is called the slowness surface. It can easily be visualized and may for instance look like Figure 1.

2.2 The Green's function

To calculate the Green's function we use the first order system for u_M that is equivalent to (7). It is given by

$$\frac{\partial}{\partial t}\begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A_M(x, D) & 0 \end{pmatrix}\begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} + \begin{pmatrix} 0 \\ f_M \end{pmatrix}. \tag{10}$$

This system can be decoupled in a similar way as above. Let $B_M(x, D) = \sqrt{A_M(x, D)}$, which exists because $A_M(x, D)$ is positive definite. The principal symbol of $B_M(x, D)$ is given by $B_M^{\mathrm{prin}}(x, \xi) = \sqrt{A_M^{\mathrm{prin}}(x, \xi)}$. We find that (10) is equivalent to the following two first order equations

$$\left(\frac{\partial}{\partial t} \pm iB_M(x, D)\right) u_{M,\pm} = f_{M,\pm}, \tag{11}$$

where

$$\begin{aligned} u_{M,\pm} &= \tfrac{1}{2}u_M \pm \tfrac{1}{2}iB_M(x, D)^{-1}\frac{\partial u_M}{\partial t} \\ f_{M,\pm} &= \pm\tfrac{1}{2}iB_M(x, D)^{-1}f_M. \end{aligned} \tag{12}$$

We construct operators $G_{M,\pm}$ that solve the initial value problem for (11). The operators $G_{M,\pm}$ are Fourier integral operators. Their construction is well known, see e.g. Duistermaat [10], chapter 5. The singularities are propagated along the bicharacteristics, that are determined by Hamilton's equations from the principal symbol (factor i divided out) $\tau \pm B_M(x, \xi)$ of (11). These equations read

$$\begin{aligned} \frac{\partial x}{\partial \lambda} &= \pm\frac{\partial}{\partial \xi}B_M(x, \xi), & \frac{\partial t}{\partial \lambda} &= 1, \\ \frac{\partial \xi}{\partial \lambda} &= \mp\frac{\partial}{\partial x}B_M(x, \xi), & \frac{\partial \tau}{\partial \lambda} &= 0. \end{aligned} \tag{13}$$

The solution may be parameterized by t. We denote the solution with the $+$ sign and initial values x_0, ξ_0 by $(x_M(x_0, \xi_0, t), \xi_M(x_0, \xi_0, t))$. The solution with the $-$ sign is given by reversing the time direction, i.e. it is given by $(x_M(x_0, \xi_0, -t), \xi_M(x_0, \xi_0, -t))$.

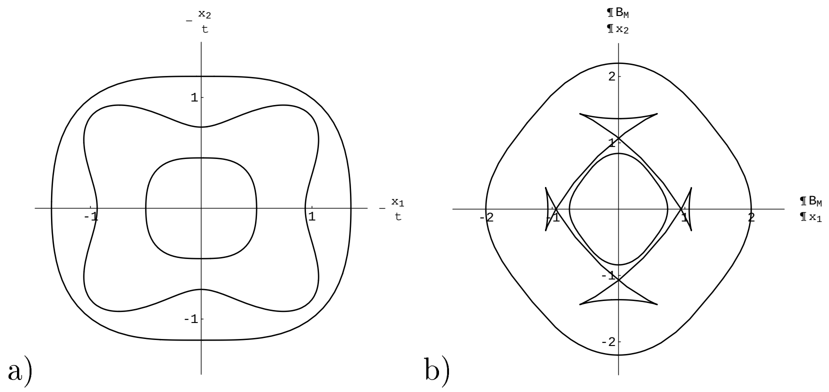

Figure 1: Slowness surface, and the set of corresponding velocities. Note the caustics that are due to the face that middle component of the slowness surface is not convex.

Observe that the group velocity (the velocity of the bicharacteristic) $\pm\frac{\partial B_M}{\partial\xi}$ is orthogonal to the slowness surface. If the slowness surface is not convex caustics may arise instantly from a point source. An example is given in Figure 1.

The canonical relation of the operator $G_{M,\pm}$ is given by

$$C_{M,\pm} = \{(x_M(x_0,\xi_0,\pm t), t, \xi_M(x_0,\xi_0,\pm t), \mp B_{M,\pm}(x_0,\xi_0); x_0,\xi_0)\}. \tag{14}$$

A convenient choice of phase function is described in Maslov and Fedoriuk [17]. His results state that one can always use a subset of the cotangent vector components as phase variables. Let us choose coordinates for $C_{M,+}$ of the form

$$(x_I, \xi_J, \tau, x_0)$$

where I, J is a partition of $\{1, \dots, n\}$. Now it follows from Theorem 4.21 in Maslov and Fedoriuk [17] that there is a function $S_{M,+}(x_I, \xi_J, \tau, x_0)$, such that locally $C_{M,+}$ is given by

$$\begin{aligned} x_J &= -\frac{\partial S_{M,+}}{\partial \xi_J}, & t &= -\frac{\partial S_{M,+}}{\partial \tau}, \\ \xi_I &= \frac{\partial S_{M,+}}{\partial x_I}, & \xi_0 &= -\frac{\partial S_{M,+}}{\partial x_0}. \end{aligned} \tag{15}$$

Here we take into account that $C_{M,+}$ is a canonical relation which introduces a minus sign for ξ_0. A nondegenerate phase function for $C_{M,+}$ is then defined by

$$\phi_{M,+}(x, t, x_0, \xi_J, \tau) = S_{M,+}(x_I, \xi_J, x_0, \tau) + \langle x_J, \xi_J\rangle + \tau t. \tag{16}$$

The canonical relation $C_{M,-}$ for is given by

$$C_{M,-} = \{(x, t, -\xi, -\tau; x_0, -\xi_0) \,|\, (x, t, \xi, \tau; x_0, \xi_0) \in C_{M,+}\}.$$

Thus a phase function for $C_{M,-}$ is $\phi_{M,-}(x, t, x_0, \xi_J, \tau) = -\phi_{M,+}(x, t, x_0, -\xi_J, -\tau)$. We may define the canonical relation for G_M as $C_M = C_{M,+} \cup C_{M,-}$ and a phase function $\phi_M = \phi_{M,-}$ if $\tau > 0$, $\phi_M = \phi_{M,+}$ is $\tau < 0$.

We have to assume that the decoupling is valid microlocally around the bicharacteristic. In that case Theorem 5.1.2 of Duistermaat [10] gives that the operator $G_{M,\pm}$ is microlocally a Fourier integral operator of order $-\frac{1}{4}$. Hence microlocally we have an expression for $G_{M,\pm}$ of the form

$$G_{M,\pm}(x, x_0, t) = (2\pi)^{-\frac{|J|+1}{2}-\frac{2n+1}{4}} \int A_{M,\pm}(x_I, x_0, \xi_J, \tau) \mathrm{e}^{i\phi_{M,\pm}(x,x_0,\xi_J,\tau)}\, \mathrm{d}\xi_J\, \mathrm{d}\tau. \tag{17}$$

The factors of (2π) in front of the integral are according to the convention of Treves [25], Hörmander [13].

The amplitude $A_{M,\pm}(x_I, x_0, \xi_J, \tau)$ satisfies a transport equation along the bicharacteristics. It is an element of $M_{C_M} \otimes \Omega^{1/2}(C_M)$, the tensor product of the Maslov bundle M_{C_M} and the halfdensities on the canonical relation C_M. The Maslov bundle gives a factor i^k, which we will not explicitly calculate. We will however give an expression for the absolute value of the amplitude, using the fact that energy is conserved to highest order.

For this purpose consider the Green's function with t, t_0 fixed. It can be viewed as an invertible FIO, mapping the displacement at t_0, $u|_{t_0} \in \mathcal{E}'(X)$ to the displacement at t, $u|_t \in \mathcal{D}'(X)$. We denote this FIO by $G_{M,\pm}(t - t_0)$. For this FIO on can find a Maslov type phase function using (x_I, ξ_J, t, x_0) to parameterize $C_{M,\pm}$. We will calculate the absolute value of the corresponding amplitude $A_{M,\pm}(x_I, \xi_J, t, x_0)$. To highest order the energy at time t is given by

$$\int |B_M(x, D) u_{M,\pm}(x, t)|^2 \, \mathrm{d}x.$$

This gives the relation

$$G_{M,\pm}(t - t_0)^* B_M(x, D)^* B_M(x, D) G_{M,\pm}(t - t_0) = B_{M,\pm}(x_0, D_0)^* B_{M,\pm}(x_0, D_0),$$

where $G_{M,\pm}(t - t_0)^*$ denotes the adjoint of $G_{M,\pm}(t - t_0)$. The left hand side is a product of invertible Fourier integral operators, so we can use the theory of section 8.6 in Treves [25]. We find that to highest order

$$\left|(2\pi)^{-\frac{1}{4}} A_{M,\pm}(x_I, \xi_J, t, x_0)\right|^2 = \left|\det \frac{\partial \xi_0}{\partial(x_I, \xi_J)}\right| \left|\frac{B_M(x_0, \xi_0)}{B_M(x, \xi)}\right|^2,$$

The value of $B_M(x, \xi)$ is conserved along the bicharacteristic. Also we may use that $\left|\det \frac{\partial \xi_0}{\partial(x_I, \xi_J)}\right| = \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_I, \xi_J, x_0, t)}\right|$. It follows that to highest order

$$|A_{M,\pm}(x_I, \xi_J, t, x_0)| = (2\pi)^{\frac{1}{4}} \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_I, \xi_J, x_0, t)}\right|^{\frac{1}{2}}. \tag{18}$$

Since the absolute value of the amplitude is a halfdensity on the canonical relation we can easily transform this to different variables.

We collect the results of this section, using equations (12), (18) to obtain the statement about the amplitude. We require that around some bicharacteristic from (x_0, ξ_0, t_0) to (x, ξ, t) the decoupling is valid, i.e. we have

Assumption 1 *On the bicharacteristic the multiplicity of the eigenvalue $A_M(x, \xi)$ in (8) is equal to one.*

Lemma 2.3 *Suppose that for some bicharacteristic given by $(x, t, \xi, \tau; x_0, \xi_0) \in C_M$ Assumption 1 is satisfied. Then microlocally we have*

$$u_M(x, t) = \int G_M(x, x_0, t - t_0) f_M(x_0, t_0) \, \mathrm{d}x_0 \, \mathrm{d}t_0, \tag{19}$$

where $G_M(x, x_0, t)$ is the kernel of a Fourier integral operator with canonical relation C and order $-1\frac{1}{4}$, mapping functions of x_0 to functions of x, t. It can be written as

$$G_M(x, x_0, t) = (2\pi)^{-\frac{|J|+1}{2} - \frac{2n+1}{4}} \int A_M(x_I, x_0, \xi_J, \tau) \mathrm{e}^{i\phi_M(x, x_0, \xi_J, \tau)} \, \mathrm{d}\xi_J \, \mathrm{d}\tau. \tag{20}$$

For the amplitude $A_M(x_I, x_0, \xi_J, \tau)$ we have to highest order

$$|A_M(x_I, x_0, \xi_J, \tau)| = (2\pi)^{\frac{1}{4}} \tfrac{1}{2} |\tau|^{-1} \left| \det \frac{\partial(x_0, \xi_0, t)}{\partial(x_I, \xi_J, x_0, \tau)} \right|^{\frac{1}{2}}. \tag{21}$$

3 Reflection at an interface

A popular way to model the subsurface is to assume that it consists of different layers that have different physical properties, in our case the elastic coefficients c_{ijkl} and the density ρ. In this section we will model the reflection of waves at a smooth interface between two regions with smoothly varying parameters.

The amplitude of the scattered waves is determined essentially by the reflection coefficients. It is well known how to calculate these for two constant coefficient media and a plane interface (see e.g. Aki and Richards [1], chapter 5). In the case of smoothly varying media they determine the scattering in the limit of high frequency, see Taylor [22] for a treatment of reflection and transmission of waves using microlocal analysis. For the acoustic case see also Hansen [11]

Mathematically the reflection and transmission of waves is described by an interface problem. Let ν be the normal to the interface. At the interface the displacement and the normal traction have to be continuous

$$\begin{aligned} P_{il}u_l &= f_i && \text{away from the interface} \\ u_i &= 0 && \text{for } t < 0 \\ \rho^{-1/2}u_i & && \text{is continuous at the interface} \\ \nu_j c_{ijkl}\frac{\partial}{\partial x_k}(\rho^{-1/2}u_l) & && \text{is continuous at the interface.} \end{aligned} \tag{22}$$

Here we have the factors ρ because of our definition (3). We assume the source vanishes on a neighborhood of the interface. That this is a well-posed problem can for instance be shown using energy estimates (see e.g. Lions and Magenes [16], section 3.8).

The solutions to the PDE with $f = 0$ follow from the theory discussed in Section 2. The singularities are propagated along the bicharacteristics, curves in $T^*(X \times \mathbb{R})\backslash 0$ given by

$$(x_M(x_0, \xi_0, \pm t), t, \xi_M(x_0, \xi_0, \pm t), \mp B_M(x_0, \xi_0)).$$

This is the bicharacteristic associated with the $M, \pm$ part of the solution, see Section 2. We define a bicharacteristic to be incoming if its direction is from inside the medium towards the interface in positive time. We define a bicharacteristic to be outgoing if its direction is from the interface inside the medium in positive time, see Figure 3.

Assume that the incoming bicharacteristic stays inside the medium from $t = 0$ until it hits the interface, then the solution along such a bicharacteristic is determined completely by the PDE and the initial condition. On the other hand the solution along the outgoing bicharacteristics is not determined by the PDE and the initial condition. We can put an arbitrary source at the interface. We will show that the solution along the outgoing bicharacteristics is determined by the interface conditions.

Let's consider the consequences of the interface condition. We have

$$WF(u_i|_{x_n=0}) = \{(x', t, \xi', \tau) \mid \text{there is } \xi_n \text{ with } (x', 0, t, \xi', \xi_n, \tau) \in WF(u_i)\}.$$

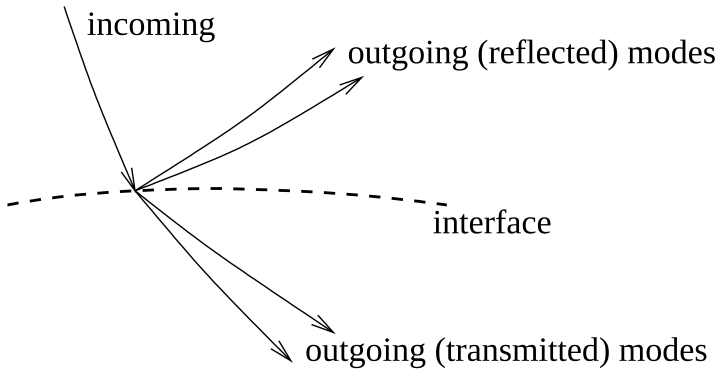

Figure 2: Incoming and outgoing rays

It follows that waves travelling along bicharacteristics that intersect the boundary at some point $x', x_n = 0, t$, with the same value of ξ', τ interact (reflect and transmit into each other). This is depicted in Figure 3

Depending on x' and the "tangential" slowness $-\tau^{-1}\xi'$ the number of interacting bichararacteristics may vary. For large values of $-\tau^{-1}\xi'$ there will be no incoming or outgoing modes, for small values there are n incoming and n outgoing modes. The situation where the vertical line in Figure 3 is tangent to the slowness surface corresponds to rays tangent to the interface. This is not treated here. Equation (9) gives that the incoming and outgoing modes correspond to the real solutions ξ_n of

$$\det P_{il}(x', 0, \xi', \xi_n, \tau) = 0.$$

This equation has $2n$ real and complex roots. The complex roots correspond to evanescent waves.

In the following theorem we show that if none of the rays involved is tangent there exists a pseudodifferential operator type relation between the amplitudes of the different modes at the interface and we calculate the principal symbol in the proof. Let $x \mapsto z(x) : \mathbb{R}^n \to \mathbb{R}^n$ be a coordinate transformation such that the interface is given by $z_n = 0$. The corresponding cotangent vector is denoted by ζ. We have the following result

Assumption 2 *There are no tangent rays at z', ζ', τ.*

Theorem 3.1 *Suppose Assumption 2 holds microlocally on some neighborhood in $T^*(Z' \times \mathbb{R})\backslash 0$. Let $u^{\text{in}}_{N(\nu)}$ be microlocal parts of a solution for the incoming modes, and suppose $G_{M(\mu)}$ refers to an outgoing mode (17). Microlocally, the singly reflected/transmitted part of the solution is given by*

$$u_{M(\mu)}(x,t) = \int_{z_n=0} G_{M(\mu)}(x, x(z), t - t_0) 2iD_{t_0} \left(R_{\mu\nu}(z, D_{z'}, D_{t_0}) u^{\text{in}}_{N(\nu)}(x(z), t_0)\right) \mathrm{d}z' \, \mathrm{d}t_0, \tag{23}$$

where $R_{\mu\nu}(z, D_{z'}, D_t)$ is a pseudodifferential operator of order 0.

In the proof we derive the explicit form of $R_{\mu\nu}(z, D_{z'}, D_t)$.

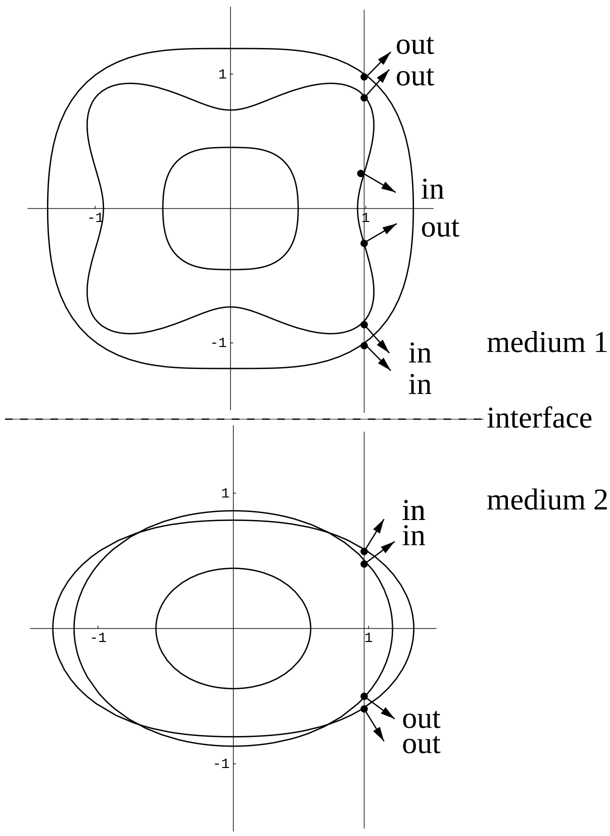

Figure 3: 2-dimensional section of 3-dimensional slowness surfaces at some point of the interface, for the medium on both sides of the interface. The slownesses of the modes that interact (i.e. reflect and transmit into each other) are the intersection points with a line that is parallel to the normal of the interface. The group velocity, which is normal to the slowness surface determines whether the mode is incoming or outgoing.

Proof For the moment we assume we have a reflector at $x_n = 0$, and smooth coefficients on either side. We show that at the interface

$$u_{M,\mathrm{out}}(x', 0, t) = R^0_{MN}(x', 0, D', D_t) u_{N,\mathrm{in}}. \tag{24}$$

We will use the notation $c_{jk;il} = c_{ijkl}$ and also $(c_{jk})_{il} = c_{ijkl}$. The partial differential equation reads in this notation

$$\left(\rho\delta_{il}\frac{\partial^2}{\partial t^2} - c_{jk;il}\frac{\partial^2}{\partial x_j \partial x_k}\right)(\rho^{-1/2}u_l) + \text{l.o.t.} \ = 0.$$

This equation can be rewritten in a first order system for the vector v_a of length $2n$ that contains both the displacement and the normal traction (normal to the surface x_n constant)

$$v_a = \begin{pmatrix} \rho^{-1/2}u_i \\ c_{nk;il}\frac{\partial(\rho^{-1/2}u_l)}{\partial x_k}. \end{pmatrix}. \tag{25}$$

The first order system is

$$\frac{\partial v_a}{\partial x_n} = iC_{ab}(x, D', D_t)v_b,$$

where C is a matrix differential operator given to highest order by

$$C_{ab}(x, D', D_t) = -i\begin{pmatrix} -\sum_{q=1}^{n-1}\sum_{j=1}^{n}(c_{nn})^{-1}_{ij}c_{nq;jl}\frac{\partial}{\partial x_q} & (c_{nn})^{-1}_{il} \\ -\sum_{p,q=1}^{n-1} b_{pq;il}\frac{\partial^2}{\partial x_p \partial x_q} + \delta_{il}\frac{\partial^2}{\partial t^2} & -\sum_{p=1}^{n-1}\frac{\partial}{\partial x_p}c_{pn;ij}(c_{nn})^{-1}_{jl} \end{pmatrix}_{ab}.$$

Here $b_{pq;il} = c_{pq;il} - \sum_{j,k=1}^{n} c_{pn;ij}(c_{nn})^{-1}_{jk}c_{nq;kl}$ (we indicated the summations explicitly because the summations over p, q are $1, \dots, n-1$, while for the j, k indices they are still $1, \dots, n$).

If the matrix $C_{ab}(x, \xi', \tau)$ has no degenerate eigenvalues then microlocally this first order system can be decoupled. This means there are scalar pseudodifferential operators $C_\mu(x, D', D_t)$ and a matrix pseudodifferential operator $L_{a\mu}(x, D', D_t)$ such that

$$C_{ab}(x, D', D_t) = L_{a\mu}(x, D', D_t)\ \mathrm{diag}(C_\mu(x, D', D_t))_{\mu\nu}\, L^{-1}_{\nu b}(x, D', D_t).$$

The index μ is $1, \dots, 2n$. The principal symbols $C^{\mathrm{prin}}_\mu(x, \xi', \tau)$ are the $2n$ solutions ξ_n to

$$\det P^{\mathrm{prin}}_{il}(x, (\xi', \xi_n), \tau) = 0, \tag{26}$$

while the principal symbol $L^{\mathrm{prin}}_{a\mu}$ (the columns appropiately normalized) is given by

$$L^{\mathrm{prin}}_{a\mu}(x, \xi', \tau) = \begin{pmatrix} Q_{iM(\mu)}(x, (\xi', C^{\mathrm{prin}}_\mu(x, \xi', \tau))) \\ c_{inkl}(-i(\xi', C^{\mathrm{prin}}_\mu(x, \xi', \tau))_k)Q_{lM(\mu)}(x, (\xi', C^{\mathrm{prin}}_\mu(x, \xi', \tau))) \end{pmatrix}_{aM}.$$

We define $v_\mu = L^{-1}_{\mu a} v_a$. (The index μ here is not the same as in (23), where it refers to the outgoing modes on both sides).

If the principal symbol of $C_\mu(x, \xi', \tau)$ is real the decoupled equation for mode μ is of hyperbolic type. It corresponds to an outgoing wave or to an incoming wave, depending on the direction of the corresponding ray. If the principal symbol of $C_\mu(x, \xi', \tau)$ is complex the decoupled operator for mode μ is of elliptic type. Depending on the sign of the imaginary part it corresponds to a mode that "blows up" going into the medium, a backward parabolic equation, or one that "dies out", a forward parabolic equation. The blow up mode has to be absent.

The matrix $L_{a\mu}$ is fixed up to normalization of its columns. For the elliptic modes ($\operatorname{Im} C^{\mathrm{prin}}_\mu(x, \xi', \tau) \neq 0$) the normalization is unimportant. For the hyperbolic modes the normalization can be such that the vector $v_\mu = L^{-1}_{\mu a} v_a$ agrees microlocally with the corresponding mode defined in section 2. To see see this assume v_μ refers to the same mode as u_M. In that case there is an invertible pseudodifferential operator $\psi(x, D, D_t)$ of order 0 such that $v_\mu = \psi u_M$. Now we can define $v_{\mu,\mathrm{new}} = \psi^{-1} v_{\mu,\mathrm{old}}$. Because ψ may depend on ξ_n, this factor cannot directly be absorbed in L. However, since $v_{\mu,\mathrm{old}}$ satisfies a first order hyperbolic equation the dependence on ξ_n can be eliminated and the factor ψ^{-1} can be absorbed in L.

For the purpose of this proof let the in-modes be the modes for which the amplitude is known, that is the incoming hyperbolic and the "blow up" elliptic modes. Denote by $L^{(1)}_{a\mu}, L^{(2)}_{a\mu}$ the matrix $L_{a\mu}$ on each side of the interface. We define the $2n \times 2n$ matrix L^{in} such that it contains the columns related to incoming modes of $L^{(1)}_{a\mu}, L^{(2)}_{a\mu}$ as follows

$$L^{\mathrm{in}}_{a\mu} = \begin{pmatrix} L^{(1),\mathrm{in}} & -L^{(2),\mathrm{in}} \end{pmatrix}_{a\mu},$$

and define $L^{\mathrm{out}}_{a\mu}$ similarly (so here μ is slightly different). The boundary condition now reads

$$L^{\mathrm{out}}_{a\mu} v^{\mathrm{out}}_\mu + L^{\mathrm{in}}_{a\mu} v^{\mathrm{in}}_\mu = 0,$$

so if we set $R^0_{\mu\nu} = -(L^{\mathrm{out}})^{-1}_{\mu a} L^{\mathrm{in}}_{a\nu}$, (for the question whether the inverse exists, see the remark after the proof) then the part referring to the hyperbolic modes give (24).

By (24) the $u_{M,\mathrm{out}}$ are determined at the interface, finding how they propagate into the medium is a (microlocal) initial value problem similar to the problem for $G_{M,\pm}$ above, where now the x_n variable plays the role of time. The solution is again a Fourier integral operator, with canonical relation given by the rays. it follows that we can use $\phi_{M,\pm}(x, x_0, t - t_0, \xi_J, \tau)$ as phase function (take care that $n \notin J$). The amplitude $A_{M,\pm}(x_I, \xi_J, \tau, x_0)$ satisfies the transport equation, however, the restriction of the FIO to the "initial surface" $x_n = 0$ so constructed is a pseudodifferential operator that is not necessarily the identity. Let's therefore try

$$\int G_{M,\pm}(x', x_n, t - t_0, x_0, 0) \psi(x, D', D_t) u_{M,\mathrm{out}}(x'_0, 0, t_0) \, \mathrm{d}x'_0 \, \mathrm{d}t_0, \tag{27}$$

where $\psi(x, D', D_t)$ is to be chosen such that the restriction of this formula to $x_n = 0$ is the identity. We can use again section 8.6 of Treves [25] to find that the principal symbol of this pseudodifferential operator should be

$$\psi(x, D', D_t) = \frac{\partial B_M}{\partial \xi}(x, \xi', B_M(\xi', \tau)) = \left| \frac{\partial x_M}{\partial t}(x, \xi', B_M(\xi', \tau), 0) \right|, \tag{28}$$

i.e. the velocity of the ray, the group velocity. After this we take into account that $G_M = \frac{1}{2} i G_{M,+} B_M(x, D)^{-1} - \frac{1}{2} i G_{M,-} B_M(x, D)$ and that $B_M^{\mathrm{prin}}(x, \xi) = \mp \tau$.

We have now obtained (23) for the case that $z = x$ (no coordinate transformation). We argue that (23) is also true for $z(x)$ is a general coordinate transformation. We start with the following equivalent of (24)

$$u_{M(\mu),\mathrm{out}}(x(z', 0), t) = R^{0,\mathrm{transf}}_{\mu\nu}(z', 0, D_{z'}, D_t) u_{N(\nu),\mathrm{in}}(x(z', 0), t). \tag{29}$$

This follows transforming the system (22) to z coordinates. The symbol of the (pseudo)differential operators transforms as $\psi^{\mathrm{transf}}(z, \zeta) = \psi(x(z), \left(\frac{\partial z}{\partial x}\right)^{\mathrm{t}} \zeta)$. Tracing the steps of the proof we find (29).

When the interface is at $z_n = 0$ we can obtain (27) in z coordinates instead of x coordinates. Transforming G_M, u_M back to x coordinates we find that for x away from the interface

$$u_M(x) = \int_{z_n=0} G_M(x, x(z), t - t_0) \left| \det \frac{\partial x}{\partial z} \right| \left| \frac{\partial z_{M,n}}{\partial t}(z, D_{z'}, D_t) \right| u_M(x(y), t_0) \, \mathrm{d}z' \, \mathrm{d}t_0.$$

Here $\left| \frac{\partial z_{M,n}}{\partial t}(z, D_{z'}, D_t) \right|$ is the transformed version of (28). Thus expression (23) follows. □

Remark 3.2 The reflection coefficients satisfy unitary relations, see Chapman [6], Kennett [15], the appendix to chapter 5. These follow essentially from conservation of energy. It follows that the matrix of reflection coefficients is well defined and in particular that the inverse $L^{\mathrm{out}}_{a\mu}$ exists. Chapman also gives a direct proof of the reciprocity relations for the reflection coefficients.

Remark 3.3 We have shown that the reflected/transmitted signal is given by a composition of Fourier integral operators acting on the source. In the case of multiple reflections or transmissions (for instance in a medium consisting of a number of smooth pieces with smooth interfaces) this is also the case. It follows that microlocally the solution operator describing the reflected solutions is itself a Fourier integral operator, where the canonical relation is given by the generalized bicharacteristics (i.e. the reflected and transmitted bicharacteristics), and the amplitude is essentially the product of the ray amplitudes and the reflection/transmission coefficients.

4 The Born approximation

We discuss the modeling and inversion of seismic data in the Born approximation, where the mediumparameters are written as the sum of smooth background and a singular perturbation. This is important in its own right, and it will also be a motivation for our approach to the model with smooth jumps described in the previous section.

The Born approximation has been discussed by a number of authors. In the acoustic case, allowing for multipathing (caustics), see Ten Kroode e.a. [23], Hansen [11]. For the acoustic problem with nonmaximal acquisition geometry, see Nolan and Symes [18]. For the elastic case with maximal acquisition geometry (and from a more applied point of view), see De Hoop and Brandsberg-Dahl [8]. We extend their results, and give an efficient, partly new presentation. Also we discuss in detail the different assumptions that are needed for the modeling and inversion.

4.1 Modeling

In the Born approximation one assumes that the total value of the medium parameters c_{ijkl}, ρ can be written as the sum of a smooth background constituent and a singular perturbation,

$$c_{ijkl} + \delta c_{ijkl}, \qquad \rho + \delta\rho$$

This leads to a perturbation of P

$$\delta P_{il} = \delta_{il} \frac{\delta\rho}{\rho} \frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j} \frac{\delta c_{ijkl}}{\rho} \frac{\partial}{\partial x_k}.$$

We denote the causal Green's operator for (2) by G_{il} and its distribution kernel by $G_{il}(x, x_0, t - t_0)$. The first order perturbation δG_{il} of G_{il} is derived by demanding that the first order term in $(P_{ij} + \delta P_{ij})(G_{jk} + \delta G_{jk})$ vanishes. This gives

$$\delta G_{il}(\hat{x}, \tilde{x}, t) = -\int_0^t \int_X G_{ij}(\hat{x}, x_0, t - t_0)\, \delta P_{jk}(x_0, D_{x_0}, D_t)\, G_{kl}(x_0, \tilde{x}, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0. \quad (30)$$

Because the background model is smooth the operator δG contains only once reflected data.

We use the decoupled equations (7). Omitting the factors $Q_{iM}(x, D), Q(x, D)^{-1}_{Nl}$ at the beginning and end of the product, we obtain an expression for the perturbation of the Green's function for the pair of modes (M, N),

$$\begin{aligned} \delta G_{MN}(\hat{x}, \tilde{x}, t) = &-\int_0^t \int_X G_M(\hat{x}, x_0, t - t_0) Q(x_0, D_{x_0})^{-1}_{Mi} \\ &\times \left(\delta_{il} \frac{\partial}{\partial t} \frac{\delta\rho}{\rho} \frac{\partial}{\partial t} - \frac{\partial}{\partial x_{0,j}} \frac{\delta c_{ijkl}}{\rho} \frac{\partial}{\partial x_{0,k}} \right) Q(x_0, D_{x_0})_{lN} G_N(x_0, \tilde{x}, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0. \end{aligned} \quad (31)$$

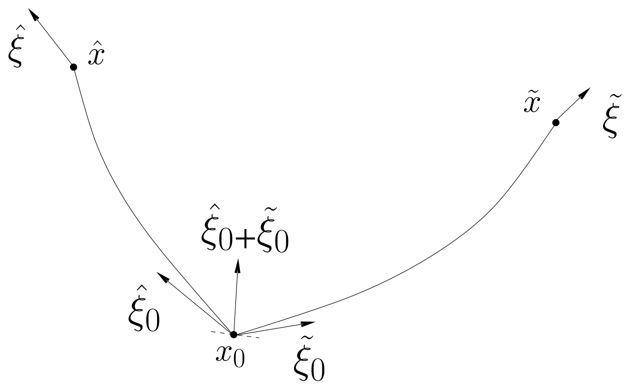

Figure 4: The scattering

Microlocally we can write G_M as in (20). For G_N we use in addition the reciprocity relation $G_N(x_0, \tilde{x}, t_0) = G_N(\tilde{x}, x_0, t_0)$. The product of operators $G_M Q(x_0, D_{x_0})^{-1}_{Mi} \frac{\partial}{\partial x_{0,j}}$ is a Fourier integral operator with the same phase as G_M, and amplitude that is to highest order the product $A_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) Q(x_0, \hat{\xi}_0)^{-1}_{Mi} i\hat{\xi}_{0,j}$, where $\hat{\xi}_0 = \xi_0(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$. In the resulting expression one of the two frequency variables $\hat{\tau}, \tilde{\tau}$ can now be eliminated using the integral over t_0. Asymptotically, by stationnary phase, $\int_0^t \mathrm{e}^{\mathrm{i}t_0(\hat{\tau}-\tilde{\tau})} \, \mathrm{d}t_0 = 2\pi\delta(\hat{\tau} - \tilde{\tau})$. The resulting formula for δG_{MN} is, modulo lower order terms in the amplitude,

$$\begin{aligned} \delta G_{MN}(\hat{x}, \tilde{x}, t) = (2\pi)^{-\frac{3n+1}{4} - \frac{\hat{J}+\tilde{J}+1}{2}} \int B_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \\ \times \left(w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta c_{ijkl}(x_0)}{\rho(x_0)} + w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta\rho(x_0)}{\rho(x_0)} \right) \\ \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x}, \tilde{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)} \, \mathrm{d}x_0 \, \mathrm{d}\hat{\xi}_{\hat{J}} \, \mathrm{d}\tilde{\xi}_{\tilde{J}} \, \mathrm{d}\tau. \end{aligned} \tag{32}$$

Here (see (16) for the construction of ϕ_M, ϕ_N)

$$\Phi_{MN}(\hat{x}, \tilde{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = \phi_M(\hat{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tau) + \phi_N(\tilde{x}, t, x_0, \tilde{\xi}_{\tilde{J}}, \tau) - \tau t. \tag{33}$$

The amplitude factors B_{MN} and $w_{MN;ijkl}$ are given by

$$\begin{aligned} B_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \omega) &= (2\pi)^{-\frac{n-1}{4}} A_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) \, A_N(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau) \\ w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) &= Q_{iM}(x_0, \hat{\xi}_0) Q_{lN}(x_0, \tilde{\xi}_0) \hat{\xi}_{0,j} \tilde{\xi}_{0,k}, \\ w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) &= -Q_{iM}(x_0, \hat{\xi}_0) Q_{iN}(x_0, \tilde{\xi}_0) \tau^2 \end{aligned} \tag{34}$$

where in the second and third equation $\hat{\xi}_0 = \xi_0(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$, $\tilde{\xi}_0 = \xi_0(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau)$. The scattering is depicted in Figure 4.1.

We investigate the map $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}) \mapsto \delta G_{MN}(\hat{x}, \tilde{x}, t)$. We use the notation C_{ϕ_M} for the canonical relation associated to ϕ_M.

Lemma 4.1 *Assume that if* $(\hat{x}, \hat{t}, \hat{\xi}, \tau; x_0, \hat{\xi}_0) \in C_{\phi_M}$, $(\tilde{x}, \tilde{t}, \tilde{\xi}, \tau; x_0, \tilde{\xi}_0) \in C_{\phi_N}\}$, *then* $\hat{\xi}_0 + \tilde{\xi}_0 \neq 0$. *Then the map* $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}) \mapsto \delta G_{MN}(\hat{x}, \tilde{x}, t)$ *given by (32) is a Fourier integral operator* $\mathcal{E}'(X) \to \mathcal{D}'(X \times X \times]0, T[)$. *The canonical relation is given by*

$$\{(\hat{x}, \tilde{x}, \hat{t} + \tilde{t}, \hat{\xi}, \tilde{\xi}, \tau; x_0, \hat{\xi}_0 + \tilde{\xi}_0) \,|\, (\hat{x}, \hat{t}, \hat{\xi}, \tau; x_0, \hat{\xi}_0) \in C_{\phi_M}, (\tilde{x}, \tilde{t}, \tilde{\xi}, \tau; x_0, \tilde{\xi}_0) \in C_{\phi_N}\}. \tag{35}$$

Proof We show that $\Phi_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, t, x, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)$ is a nondegenerate phase function. The derivatives with respect to the phase variables are given by

$$\begin{aligned}
\frac{\partial \Phi_{MN}}{\partial \tau} &= -\hat{t}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \omega) - \tilde{t}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \omega) + t \\
\frac{\partial \Phi_{MN}}{\partial \hat{\xi}_{\hat{J}}} &= -\hat{x}_{\hat{J}}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \omega) + \hat{x}_{\hat{J}} \\
\frac{\partial \Phi_{MN}}{\partial \tilde{\xi}_{\tilde{J}}} &= -\tilde{x}_{\tilde{J}}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \omega) + \tilde{x}_{\tilde{J}},
\end{aligned}$$

where $\hat{x}_{\hat{J}}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau), \tilde{x}_{\tilde{J}}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau)$ are as defined in (15). The derivatives of these expressions with respect to the variables $(\hat{x}_{\hat{J}}, \tilde{x}_{\tilde{J}}, t)$ are linearly independent, so Φ_{MN} is nondegenerate. From the expression (33) it follows that the canonical relation of this operator is given by (35). By the assumption it contains no elements with $\hat{\xi}_0 + \tilde{\xi}_0 = 0$, so it continuous $\mathcal{E}'(X) \to \mathcal{D}'(X \times X \times]0, T[)$. $\square$

We show that the condition is violated if and only if $M = N$ and there is a direct bicharacteristic from $\tilde{x}, \tilde{\xi}$ to $\hat{x}, -\hat{\xi}$. From the symmetry of the bicharacteristic under the transformation $\xi \to -\xi, t \to -t$ it follows that indeed in this case the condition is violated. On the other hand we have $B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau$. If $\hat{\xi}_0 = -\tilde{\xi}_0$ then we must have $M = N$, because $B_M(x_0, \hat{\xi}_0) = B_M(x_0, -\hat{\xi}_0)$ and the condition that the multiplicities of the eigenvalues in (8) are equal to one. If $M = N$ and $\hat{\xi}_0 = -\tilde{\xi}_0$ then we have a direct bicharacteristic.

The data is supposed to be given by $\delta G_{MN}(\hat{x}, \tilde{x}, t)$ for $(\hat{x}, \tilde{x}, t)$ in the acquisition manifold. To make this explicit, let $y \mapsto (\hat{x}(y), \tilde{x}(y), t(y))$ be a coordinate transformation, such that $y = (y', y'')$ and the acquisition manifold is given by $y'' = 0$. Assume that the dimension of y'' is $2 + c$, where c is the codimension of the geometry. We assume the data is given by

$$\delta G_{MN}(\hat{x}(y', 0), \tilde{x}(y', 0), t(y', 0)). \tag{36}$$

It follows that the map $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho})$ to the data may be seen as the compose of the map of Lemma 4.1 with the restriction operator to $y'' = 0$. The restriction operator that maps a function $f(y)$ to $f(y', 0)$ is a FIO with canonical relation given by $\Lambda_r = \{(y', \eta'; (y', y''), (\eta', \eta'')) \in T^*Y' \times T^*Y \,|\, y'' = 0\}$. The composition of the canonical relations $\Lambda_{0,MN}$ and Λ_r is well defined if the intersection of $\Lambda_r \times \Lambda_{0,MN}$ with $T^*Y' \backslash 0 \times \mathrm{diag}(T^*Y \backslash 0) \times T^*X \backslash 0$ is transversal. In this case we must have that the intersection of $\Lambda_{0,MN}$ with the manifold $y'' = 0$ is transversal.

Let us repeat our assumptions, and state the final result of this subsection.

Assumption 3 *There are no elements* $(y', 0, \eta', \eta'') \in T^*Y\backslash 0$ *such that there is a direct bicharacteristic from* $(\hat{x}(y', 0), \hat{\xi}(y', 0, \eta', \eta''))$ *to* $(\tilde{x}(y', 0), -\tilde{\xi}(y', 0, \eta', \eta''))$ *with time* $t(y', 0)$.

Assumption 4 *The intersection of* $\Lambda_{0,MN}$ *with the manifold* $y'' = 0$ *is transversal. In other words*

$$\frac{\partial y''}{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})} \text{ has maximal rank.} \tag{37}$$

Theorem 4.2 *If Assumptions 3, 4 are satisfied then the operator* $F_{MN;ijkl}$ *that maps the medium perturbation* $(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho})$ *to the data is microlocally a Fourier integral operator with canonical relation given by*

$$\begin{aligned}\Lambda_{MN} =& \{(y'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}), \eta'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}); x_0, \hat{\xi}_0 + \tilde{\xi}_0) \mid \\ & B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau, y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0\}\end{aligned} \tag{38}$$

The order equals $\frac{n-1+c}{4}$. *The amplitude is given to highest order (in coordinates* (y'_I, η'_J, x_0) *for* Λ_{MN}, *where* I, J *are a partition of* $\{1, \dots, 2m-1-c\}$*) by the product* $w_{MN;ijkl}(y'_I, \eta'_J, x_0) B_{MN}(y'_I, \eta'_J, x_0)$, *where*

$$|B_{MN}(y'_I, y'', \eta'_J, x_0)| = \frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n+1+c}{4}} \left|\frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-\frac{1}{2}} \left|\frac{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x_0, y'_I, y'', \eta'_J, \Delta\tau)}\right|^{\frac{1}{2}} \tag{39}$$

Proof The first statement is argued above. The order is given by

$$\mu + \frac{N}{2} - \frac{\dim X + \dim Y'}{4},$$

where μ is the degree of homogeneity of the amplitude and N is the number of phase variables. Now the factor $w_{MN;ijkl}, w_{MN;0}$ are homogeneous of order 2 the degree of homogeneity of the factor B_{MN} follows from (18). We have

$$\text{order } F_{MN;ijkl} = 2 + (-2 - \frac{|\hat{J}| + |\tilde{J}| + 2}{2} + n) + \frac{|\hat{J}| + |\tilde{J}| + 1}{2} - \frac{3n - 1 - c}{4}.$$

This gives the order.

We calculate the amplitude of the Fourier integral operator in Lemma 4.1. The factor $w_{MN;ijkl}$ is simply multiplicative. Suppose we parameterize (35) by $x_0, \hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \hat{\tau}, \tilde{\tau}$, where $\hat{\tau} = \tilde{\tau}$. Define $\tau = \frac{\hat{\tau}+\tilde{\tau}}{2}$, $\Delta\tau = \hat{\tau} - \tilde{\tau}$. Using (18), (34) we find that the amplitude $B_{MN}(x_0, \hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)$ satisfies

$$|B_{MN}(x_0, \hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)| = \frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n-1}{4}} \left|\frac{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x_0, \hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau, \Delta\tau)}\right|^{\frac{1}{2}}$$

The transformation to y coordinates in (36), and not as a halfdensity, gives a factor $|\frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}|^{-\frac{1}{2}}$, (for the Fourier integral operators it would be more natural to transform as a half density). The amplitude transforms as a halfdensity on the canonical relation, this gives a factor

$$\left|\frac{\partial(y'_I, y'', \eta'_J)}{\partial(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)}\right|^{\frac{1}{2}}.$$

The additional factor $(2\pi)^{-\frac{2+c}{4}}$ is because of the normalization. We find (39). $\square$

The canonical relation is naturally parameterized by $(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})$ such that $B_M(x_0, \hat{\xi}_0) - B_N(x_0, \tilde{\xi}_0) = 0, y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0$. There is also a natural density associated to this set, the quotient density. The Jacobian in (39) means that the amplitude factor $|B_{MN}(y'_I, y'', \eta'_J, x_0)|$ is given in fact by the associated halfdensity times $\frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n-3}{4}}\left|\frac{\partial(\hat{x},\tilde{x},t)}{\partial y}\right|^{-\frac{1}{2}}$.

If $c = 0$ and there are no tangent rays, i.e.

$$\operatorname{rank}\frac{\partial y''}{\partial(\hat{t}, \tilde{t})} = 2,$$

then a practical way to parameterize the canonical relation is by using the vectors $\hat{\alpha} = \frac{-\tau^{-1}\hat{\xi}}{\|\tau^{-1}\hat{\xi}\|}, \tilde{\alpha} = \frac{-\tau^{-1}\tilde{\xi}}{\|\tau^{-1}\tilde{\xi}\|} \in S^{n-1}$ and the frequency τ.

4.2 Inversion

Let us now consider the reconstruction of $\left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right)$ from the data. We define some new notation, let

$$g_\alpha = \left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right),$$

the forward operator in the Born approximation is denoted by $F_{MN;\alpha}$.

Supppose we want to invert data from one pair of modes (M, N) (the general case is discussed at the end of this section). The standard procedure to deal with the fact that the problem is overdetermined is to use the method of least squares. Define the normal operator $N_{MN;\alpha\beta}$ as the product of $F_{MN;\alpha}$ and its adjoint $F^*_{MN;\alpha}$

$$N_{MN;\alpha\beta} = F^*_{MN;\alpha}F_{MN;\beta} \tag{40}$$

(no summation over M, N). If $N_{MN;\alpha\beta}$ is invertible (as a matrix valued operator with indices $\alpha\beta$), then

$$F^{-1}_{MN;\alpha} = (N_{MN})^{-1}_{\alpha\beta}F^*_{MN;\beta} \tag{41}$$

(no summation over M, N) is a left inverse of $F_{MN;\alpha}$ that is optimal in the sense of least squares[1].

[1]Equation (40) is for case where one wants to minimize the L^2 norm $\|d_{MN} - F_{MN;\alpha}g_\alpha\|$. It can be easily adapted to the case where one wants to minimize a different norm. This introduces extra factors in the amplitude.

The properties of the compose (40) depend on Λ_{MN}. Let $\pi_{Y'}, \pi_X$ be the projection mappings of Λ_{MN} to $T^*Y'\backslash 0$, resp. $T^*X\backslash 0$. We will show that under the following assumption $N_{MN;\alpha\beta}$ is a pseudodifferential operator, so that the problem of inverting $N_{MN;\alpha\beta}$ reduces to a finite dimensional problem for each (x, ξ).

Assumption 5 *The projection $\pi_{Y'}$ of Λ_{MN} on $T^*Y'\backslash 0$ or on a open conical subset of $T^*Y'\backslash 0$, for the case we apply a microlocal cutoff to the data, is an embedding, i.e. it is*
i) immersive
ii) injective
iii) proper

This assumption implies that the image of $\pi_{Y'}$ is a submanifold of $T^*Y'\backslash 0$. Let us discuss these requirements, starting with the first. Using that Λ_{MN} is a canonical relation we have

Lemma 4.3 *The projection $\pi_{Y'}$ of Λ_{MN} on $T^*Y'\backslash 0$ is an immersion if and only if the projection π_X of Λ_{MN} on $T^*X\backslash 0$ is a submersion. In this case the image of $\pi_{Y'}$ is locally a coisotropic submanifold of $T^*Y'\backslash 0$.*

Proof This is a property of Lagrangian manifolds. It follows from Lemma 25.3.6 in Hörmander [13]. We give an independent proof.

The symplectic forms $\sigma_X, \sigma_{Y'}$ on $T^*X\backslash 0, T^*Y'\backslash 0$ can be viewed as 2-forms on Λ_{MN}. Because Λ_{MN} is a canonical relation $\sigma_{Y'} = \sigma_X$ on Λ_{MN}, and in particular $\operatorname{rank} \sigma_{Y'} = \operatorname{rank} \sigma_X$. Now consider π_X. Clearly $\operatorname{rank} \sigma_X = 2n$ if and only if π_X is submersive.

Consider $\pi_{Y'}$. If this projection is immersive then the image has dimension $n+m$ (in this proof $m = \dim Y' = 2n - 1 - c$), while $\dim T^*Y'\backslash 0 = 2m$. Then $\operatorname{rank} \sigma_{Y'}$ is at least $2n$, so it must be equal to $2n$. On the other hand, if $\operatorname{rank} \sigma_{Y'} = 2n$, then the tangent space of Λ_{MN} at that point is given by the span of a set vectors of the form

$$\{(v_1, w_1), \ldots, (v_{2n}, w_{2n}), (0, w_{2n+1}), \ldots, (0, w_{n+m})\}.$$

The $w_i, i \in \{1, \ldots, 2n\}$ must be linearly independent because $\operatorname{rank} \sigma_{Y'} = 2n$. For $w_i, w_j, i \leq 2n, j > 2n$ we have $\sigma_{Y'}(w_i, w_j) = 0$, so the w_j are linearly independent from the w_i. The $w_i, i > 2n$ must be linearly independent, because $(0, w_i)$ are basisvectors for the tangent space to Λ_{MN}. So if $\operatorname{rank} \sigma_{Y'} = 2n$ then $\pi_{Y'}$ is an immersion. Because $\operatorname{rank} \sigma_{Y'} = 2n$ in that case the image is locally a coisotropic submanifold. □

Thus if the first part of Assumption 5 is satisfied then we can use $(x, \xi) \in T^*X\backslash 0$ as coordinates on Λ_{MN}. In addition we need a coordinate on the set $(x, \xi) =$ constant, that we denote by e. The new parameterization of Λ_{MN} is

$$\Lambda_{MN} = \{(y'(x, \xi, e), \eta'(x, \xi, e); (x, \xi))\}. \tag{42}$$

The results do not depend on the precise definition of e. A natural choice is something like scattering angle and azimuth, (such that ξ, e together parameterize

$(\hat{\xi}_0, \tilde{\xi}_0)$ of equation (38), assuming that the variables $(\hat{t}, \tilde{t})$ can be solved from the constraints), because the scattering can always be parameterized in this way. Often people use part of the y' coordinates, such as the offset $\hat{x} - \tilde{x}$ if $X \subset \mathbb{R}^n_{x_n > 0}$. However, these cannot always be used, because of caustics.

We show that the first part of Assumption 5 implies that $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) + \frac{\partial B_M}{\partial \xi}(x, \tilde{\xi}_0) \neq 0$, in other words the group velocities at the scattering point do not add up to 0. We may parameterize Λ_{MN} by parameters $x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}$, where $\hat{\xi}_0, \tilde{\xi}_0$ are such that $B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau$. The projection π_X is given by $(x, \hat{\xi}_0 + \tilde{\xi}_0)$. Consider tangent vector to Λ_{MN} given by vectors $v_{\hat{\xi}_0}, v_{\tilde{\xi}_0}$. They must satisfy

$$\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) \cdot v_{\hat{\xi}_0} = \frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0) \cdot v_{\tilde{\xi}_0} = \pm v_\tau. \tag{43}$$

So if $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$, then (43) implies that $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) \cdot (v_{\hat{\xi}_0} + v_{\tilde{\xi}_0}) = 0$, so that the projection of Λ_{MN} on $T^*X\backslash 0$ is not submersive. If $c = 0$, and rank $\frac{\partial y''}{\partial(\hat{t}, \tilde{t})} = 2$ (no tangent rays), then (43) is the only condition on $(\hat{\xi}_0, \tilde{\xi}_0)$. In that case if $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) \neq -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$ then the projection is submersive. In other cases the set of $(\hat{\xi}_0, \tilde{\xi}_0)$ is in general a smaller subset of $T^*_x X\backslash 0 \times T^*_x\backslash 0$, which should also be taken into account.

Let us now discuss the second and third part of Assumption 5. The second part is a well known condition, see Ten Kroode e.a. [23], Hansen [11]. Essentially the condition is that there are no two different singularities in g_α mapped to the same position in $T^*Y\backslash 0$. For an analysis of the case where this condition is violated see Stolk [20].

The definition of proper is that the preimage of a compact set is a compact set. So assume we have a compact subset of $T^*Y'\backslash 0$, The elements of Λ_{MN} correspond to those points where the source and receiver rays intersect. This can be written as a set where some continuous function vanishes. Therefore this set is closed. It is also bounded, and hence it is compact. So the third part of the assumption is automatically satisfied.

When constructing the compose (40) there is a subtlety that we have to take into account, namely that the linearized forward operator is only *microlocally* a Fourier integral operator. To make it globally a Fourier integral operator we apply a pseudodifferential cutoff $\psi(y', D_{y'})$ with compact support. Due to the third part of Assumption 5 the forward operator is then a finite sum of local Fourier integral operators.

Theorem 4.4 *Let $\psi(y', D_{y'})$ be a pseudodifferential cutoff with conically compact support in $T^*Y'\backslash 0$, such that for the set*

$$\{(y', \eta'; x_0, \xi_0) \in \Lambda_{MN} \,|\, (y', \eta') \in \operatorname{supp}\psi\} \tag{44}$$

Assumptions 3, 4, 5 are satisfied. Then

$$F^*_{MN;\beta}\psi(y', D_{y'})^*\psi(y', D_{y'})F_{MN;\alpha} \tag{45}$$

is a pseudodifferential operator of order $n-1$. Its principal symbol is given by

$$N_{MN;\beta\alpha}(x,\xi) = \frac{1}{16}(2\pi)^{-n} \int |\psi(y'(x,\xi,e),\eta'(x,\xi,e))|^2 \tau^{-4} \overline{w_{MN;\beta}(x,\xi,e)} w_{MN;\alpha}(x,\xi,e) \times \left|\frac{\partial(\hat{x},\tilde{x},t)}{\partial y}\right|^{-1} \left|\frac{\partial(x,\hat{\xi}_0,\tilde{\xi}_0,\hat{t},\tilde{t})}{\partial(x,\xi,e,y'',\Delta\tau)}\right| \mathrm{d}e. \tag{46}$$

Proof We use the clean intersection calculus for Fourier integral operators (see e.g. Treves [25]) to show that (45) is a Fourier integral operator. The canonical relation of F^*_{MN} is given by

$$\Lambda^*_{MN} = \{(x,\xi;y',\eta') \,|\, (y',\eta';x,\xi) \in \Lambda_{MN}\}.$$

Let $L = \Lambda^*_{MN} \times \Lambda_{MN}$ and $M = T^*X\backslash 0 \times \mathrm{diag}(T^*Y'\backslash 0) \times T^*X\backslash 0$. We have to show that the intersection of $L \cap M$ is clean, i.e.

$$L \cap M \text{ is a manifold} \tag{47}$$

$$TL \cap TM = T(L \cap M). \tag{48}$$

It follows from Assumption 5 that $L \cap M$ is given by

$$L \cap M = \{(x,\xi,y',\eta',y',\eta',x,\xi) \,|\, (y',\eta';x,\xi) \in \Lambda_{MN}\}. \tag{49}$$

Because Λ_{MN} is a manifold this set satisfies (47). The property (48) follows from the assumption that the map π'_Y is immersivene. The excess is given by

$$\begin{aligned} e = {} & \dim(L\cap M) - (\dim L + \dim M - \dim T^*X\backslash 0 \times T^*Y'\backslash 0 \times T^*Y'\backslash 0 \times T^*X\backslash 0) \\ = {} & n - 1 - c. \end{aligned} \tag{50}$$

Taking into account that we apply the pseudodifferential cutoff $\psi(y', D_{y'})$ it follows that (45) is a Fourier integral operator. The canonical relation $\Lambda^*_{MN} \circ \Lambda_{MN}$ is contained in the diagonal of $T^*X\backslash 0 \times T^*X\backslash 0$, so it is a pseudodifferential operator. The order is given by $2\,\mathrm{order}\, F_{MN;\alpha} + \frac{e}{2} = n - 1$.

We write $\psi(y',\eta') = \sum_i \psi^{(i)}(y',\eta')$, where each $\psi^{(i)}$ is such that the distribution kernel of $\psi^{(i)}(y',D_{y'})F_{MN;\alpha}(y',x)$ can be written as

$$\psi^{(i)}(y',D_{y'})F_{MN;\alpha}(y',x) = (2\pi)^{-\frac{3n-1-c}{4}-\frac{|J|}{2}} \int \psi^{(i)}(y'_I,\eta'_J,x) \times B_{MN}(y'_I,\eta'_J,x) w_{MN;\alpha}(y'_I,\eta'_J,x) \mathrm{e}^{\mathrm{i}(S(y'_I,x,\eta'_J)+\langle \eta'_J, y'_J\rangle)} \,\mathrm{d}\eta'_J, \tag{51}$$

where $\psi^{(i)}(y'_I,\eta'_J,x) = \psi^{(i)}(y'_I, y'_J(y'_I,\eta'_J,x), \eta'_I(y'_I,\eta'_J,x), \eta'_J)$. The distribution kernel of the normal operator is given by a sum of terms

$$\int (\overline{\psi(y', D_{y'}) F_{MN;\beta}(y', x)})(\psi(y', D_{y'}) F_{MN;\alpha}(y', x_0))\,\mathrm{d}y'$$
$$= (2\pi)^{-\frac{3n-1-c}{2}-|J|} \int \overline{\psi^{(i)}(y'_I, \eta'_J, x)}\psi^{(i)}(y'_I, \eta'_{0,J}, x_0)$$
$$\times \overline{B_{MN}(y'_I, \eta'_J, x)} B_{MN}(y'_I, \eta'_{0,J}, x_0) \overline{w_{MN;\beta}(y'_I, \eta'_J, x)} w_{MN;\alpha}(y'_I, \eta'_{0,J}, x_0)$$
$$\times \mathrm{e}^{\mathrm{i}(S(y'_I, x_0, \eta'_{0,J}) - S(y'_I, x, \eta'_J) + \langle \eta'_{0,J}, y'_J\rangle - \langle \eta'_J, y'_J\rangle)}\,\mathrm{d}\eta'_{0,J}\,\mathrm{d}\eta'_J\,\mathrm{d}y'.$$

We now perform stationnary phase. One can integrate out the variables $y'_J, \eta'_{0,J}$. For the remaining variables we use that

$$S(y'_I, x_0, \eta'_J) - S(y'_I, x, \eta'_J) = \langle x - x_0, \xi(y'_I, \eta'_J, x_0)\rangle + O(|x - x_0|^2).$$

Thus we find (to highest order)

$$(2\pi)^{-\frac{3n-1-c}{2}} \int |\psi^{(i)}(y'_I, \eta'_J, x)|^2 |B_{MN}(y'_I, \eta'_J, x)|^2 \overline{w_{MN;\beta}(y'_I, \eta'_J, x)} w_{MN;\alpha}(y'_I, \eta'_J, x)$$
$$\times \mathrm{e}^{\mathrm{i}\langle x - x_0, \xi(y'_I, \eta'_J, x_0)\rangle}\,\mathrm{d}\eta'_J\,\mathrm{d}y'_I.$$

We now do the change of variables $(x, y'_I, \eta'_J) \to (x, \xi, e)$, and we use (39). In addition we can do the summation over i. We find

$$N_{MN;\beta\alpha}(x, x_0) = \frac{(2\pi)^{-2n}}{16} \int |\psi(y'(x,\xi,e), \eta'(x,\xi,e))|^2 \tau^{-4} \overline{w_{MN;\beta}(x,\xi,e)} w_{MN;\alpha}(x,\xi,e)$$
$$\times \left|\frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-1} \left|\frac{\partial(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x, \xi, e, y'', \Delta\tau)}\right| \mathrm{e}^{\mathrm{i}\langle x - x_0, \xi\rangle}\,\mathrm{d}\xi\,\mathrm{d}e. \tag{52}$$

It follows that the principal symbol of $N_{MN;\beta\alpha}$ is given by (46). □

So far we concentrated on inversion of data from one pair of modes (M, N). Often data d_{MN} will be available for some subset I of all possible pairs of modes. Define the normal operator for this case

$$N = \sum_{(M,N)\in I} F^*_{MN} F_{MN} = \sum_{(M,N)\in I} N_{MN}.$$

If all the N_{MN} are pseudodifferential operators then N is also a pseudodifferential operator. A left inverse is now given by

$$N^{-1} F^*,$$

where F^* is the vector containing the $F_{MN}, (M,N) \in I$.

5 Symplectic geometry of data

In the previous section we saw that the wavefront set of the modeled data can not be arbitrary. This is due to the redundancy in the data, in the Born approximation the singular part of the medium parameters is a function of n variables, while the data is a function of $2n-1-c$ variables. This redundancy is also important in the reconstruction of the background medium (or the medium above the interface in the case of a smooth jump). This will be explained below.

Consider again the canonical relation Λ_{MN}. Denote by F in this section the map $(x,\xi,e) \mapsto (y'(x,\xi,e), \eta'(x,\xi,e)) : T^*X\backslash 0 \times E \to T^*Y'\backslash 0$. This map conserves the symplectic form of $T^*X\backslash 0$. That is, if $w_{x_i} = \frac{\partial(y',\eta')}{\partial x_i}$, and similar for w_{ξ_i}, w_{e_i}, we have

$$\begin{aligned}
&\sigma_{Y'}(w_{x_i}, w_{x_j}) = \sigma_{Y'}(w_{\xi_i}, w_{\xi_j}) = 0 \\
&\sigma_{Y'}(w_{\xi_i}, w_{x_j}) = \delta_{ij} \\
&\sigma_{Y'}(w_{e_i}, w_{x_j}) = \sigma_{Y'}(w_{e_i}, w_{\xi_j}) = \sigma_{Y'}(w_{e_i}, w_{e_j}) = 0.
\end{aligned} \tag{53}$$

The (x,ξ,e) are "symplectic coordinates" on the projection of Λ_{MN} on $T^*Y'\backslash 0$, which is a subset of $T^*Y'\backslash 0$.

The image L of the map F is coisotropic. The sets $(x,\xi) =$ constant are the isotropic fibers of the fibration of Hörmander [14], Theorem 21.2.6, see also Theorem 21.2.4). Duistermaat [10] calls them characteristic strips (see Theorem 3.6.2). We have sketched the situation in Figure 5. The wavefront set of the data is a union of fibers.

Using the following result we can extend the coordinates (x,ξ,e) to symplectic coordinates on an open neighborhood of L.

Lemma 5.1 *Let L be an embedded coisotropic submanifold of $T^*Y'\backslash 0$, with coordinates (x,ξ,e) such that (53) holds. Denote $(y',\eta') = F(x,\xi,e)$. We can find a homogeneous canonical map G from an open part of $T^*(X\times E)\backslash 0$ to an open neighborhood of L in $T^*Y'\backslash 0$, such that $G(x,e,\xi,\epsilon=0) = F(x,\xi,e)$.*

Proof The e_i can be viewed as functions on L. We will first extend them to functions on the whole $T^*Y'\backslash 0$ such that the Poisson brackets $\{e_i, e_j\}$ satisfy

$$\{e_i, e_j\} = 0, \qquad 1 \le i,j \le m-n, \tag{54}$$

where $m = \dim Y' = 2n-c-1$. This can be done successively for $e_1,\dots,e_{m-n}$ by the method that we describe now, see Treves [25], chapter 7, the proof of Theorem 3.3, or Duistermaat [10], the proof of Theorem 3.5.6. Suppose we have extended $e_1,\dots,e_l$, we extend e_{l+1}. In order to satisfy (54) e_{l+1} has to be a solution u of

$$H_{e_i} u = 0, \qquad 1 \le i \le l,$$

with initial condition on some manifold transversal to the H_{e_i}. For any $(y',\eta') \in L$ the covectors $\mathrm{d}e_i$, $1 \le i \le l$ restricted to $T_{(y',\eta')}L$ are linearly independent, so the H_{e_i} are transversal to L and they are linearly independent modulo L. So we can give the

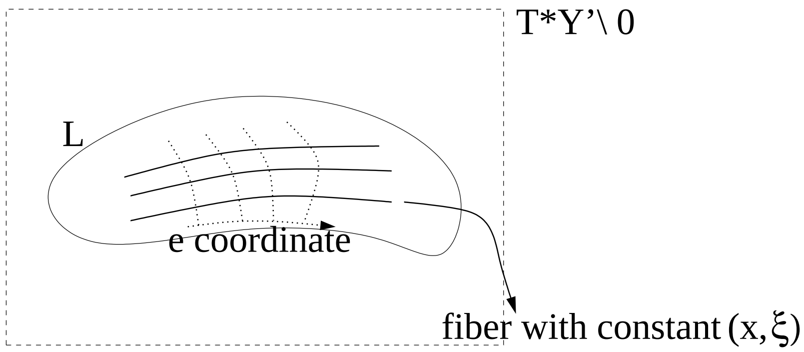

Figure 5: Symplectic structure of Λ_{MN}.

initial condition $u = e_{l+1}$ for u on L and even prescribe u on a larger manifold, which lead to nonuniqueness of the extensions e_i.

We now have $m - n$ commuting vectorfields H_{e_i} that are transversal to L and linearly independent on some open neighborhood of L. The Hamilton systems with parameters ϵ_i reads

$$\frac{\partial y'_j}{\partial \epsilon_i} = \frac{\partial e_i}{\partial \eta'_j}(y', \eta'), \qquad \frac{\partial \eta'_j}{\partial \epsilon_i} = -\frac{\partial e_i}{\partial y'_j}(y', \eta').$$

Let $G(x, e, \xi, \epsilon)$ be the solution of the Hamilton systems with initial value $(y', \eta') = F(x, \xi, e)$ with "flowout parameters" ϵ. This gives a diffeomorphic map of a neigborhood of the set $\epsilon = 0$ in $T^*(X \times E)\backslash 0$ to a neighborhood of L in $T^*Y'\backslash 0$. One can check from the Hamilton system that this map is homogeneous.

Remains to check the commutation relations. The relations (53) are valid for any λ, because the Hamilton flow conserves the symplectic form on $T^*Y'\backslash 0$. The commutation relations for $\frac{\partial(y', \eta')}{\partial \epsilon_i}$ follow, using that $\frac{\partial(y', \eta')}{\partial \epsilon_i} = H_{e_i}$. □

Let M_{MN} the canonical relation associated to the map G we just constructed, i.e. $M_{MN} = \{(G(x, e, \xi, \epsilon); x, e, \xi, \epsilon)\}$. We construct a Maslov type phase function for M_{MN} that is directly related to a phase function for Λ_{MN}.

Suppose (y'_I, η'_J, x) are suitable coordinates for Λ_{MN}. For ϵ small the constant-ϵ subset of M_{MN}, can be coordinatized by the same set of coordinates, thus we can use coordinates $(y'_I, \eta'_J, x, \epsilon)$ on M_{MN}. Now there is (see Theorem 4.21 in Maslov and Fedoriuk [17]) a function $S(y'_I, \eta'_J, x, \epsilon)$ such that M_{MN} is given by

$$e = \frac{\partial S}{\partial \epsilon}, \qquad \xi = -\frac{\partial S}{\partial x},$$
$$y'_J = -\frac{\partial S}{\partial \eta'_J}, \qquad \eta'_I = \frac{\partial S}{\partial u'_I}.$$

Thus a phase function for M_{MN} is given by

$$\Psi_{MN}(y', x, e, \eta'_J, \epsilon) = S(y'_I, \eta'_J, x, \epsilon) + \langle y'_J, \eta'_J \rangle - \langle e, \epsilon \rangle. \tag{55}$$

A Maslov type phase function for Λ_{MN} is given by $\Psi_{MN}(y', x, e, \eta'_J, 0)$.

6 Characterization of seismic data and the imaging reflection coefficients

In this section we give our main result, which is a characterization of seismic data, modeled with the Born approximation or using a model with "smooth jumps" as in Section 3. We also give a discussion of this result.

First we give an expression for the data modeled using the smooth jump approximation that is very similar to the expressions for the Born modeled data we obtained in Section 4. The smooth medium above the interface plays the role of the background medium in the Born approximation.

Recall the coordinates $x, \hat{\xi}_0, \tilde{\xi}_0$ that played a role in the Born approximation. A signal with mode N and covector $\tilde{\xi}_0$ can be reflected into mode M, covector $\hat{\xi}_0$ if the frequencies τ are equal and $\hat{\xi}_0 + \tilde{\xi}_0$ is normal to the interface. To highest order the pseudodifferential reflection "coefficient" $R_{\mu\nu}(z', \zeta', \tau)$ leads to a factor $R_{MN}(x, \hat{\xi}_0, \tilde{\xi}_0) = R_{\mu(M),\nu(N)}(z'(x), \zeta'(\hat{\xi}_0), \tau)$. The indices μ, ν are not necessary here. This factor can now be viewed as a function of coordinates (x, ξ, e) on Λ_{MN} (strictly speaking only defined for x in the interface, and ξ normal to the interface). To highest order it does not depend on $\|\xi\|$ and is it simply a function of (x, e). We obtain the following result, which is a generalization of the Kirchhoff approximation.

Theorem 6.1 *Suppose Assumptions 1, 2, 3, 4 are satisfied, microlocally for the relevant part of the data. Let $\Phi_{MN}(y', x, \eta'_J)$, $B_{MN}(y'_I, x, \eta'_J)$ be phase and amplitude as in Theorem 4.2, but now for the smooth medium above the interface. The data modeled with the smooth jump model is given microlocally by*

$$d_{MN}(y') = (2\pi)^{-\frac{|J|}{2} - \frac{3n-1-c}{4}} \int (B_{MN}(y'_I, x, \eta'_J) 2i\tau(\eta') R_{MN}(y'_I, x, \eta'_J) + \text{l.o.t.}) \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(y', x, \eta'_J)} \delta(z_n(x)) \, \mathrm{d}\eta'_J \, \mathrm{d}x. \tag{56}$$

i.e. by a Fourier integral operator with canonical relation Λ_{MN} and order $\frac{n-1+c}{4} - 1$ acting on the function $\delta(z_n(x))$.

Proof We write the distribution kernel of the reflected data in a similar form as (32). First recall the reciprocal expression for the Green's function.

$$G_N(x(z), \tilde{x}, t_0) = (2\pi)^{-\frac{|\tilde{J}|+1}{2} - \frac{2n+1}{4}} \int A_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) \mathrm{e}^{i\phi_N(\tilde{x}, x(z), t_0, \tilde{\xi}_{\tilde{J}}, \tau)} \, \mathrm{d}\tilde{\xi}_{\tilde{J}} \, \mathrm{d}\tau.$$

By using Theorem 3.1, and doing an integration over an t and a τ variable one finds that the Green's function for the reflected part is given by

$$\begin{aligned} G^{\text{refl}}_{MN}(\hat{x}, \tilde{x}, t) = {} & (2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2} - n} \\ & \times \int_{z_n=0} \Big(2i\tau A_M(\hat{x}_{\hat{I}}, x(z), \hat{\xi}_{\hat{J}}, \tau) A_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) R_{\mu(M)\nu(N)}(z, \zeta', \tau) + \text{l.o.t.}\Big) \\ & \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x}, \tilde{x}, t, x(z), \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)} \left|\det \frac{\partial x}{\partial z}\right| \, \mathrm{d}\hat{\xi}_{\hat{J}} \, \mathrm{d}\tilde{\xi}_{\tilde{J}} \, \mathrm{d}\tau \, \mathrm{d}z', \end{aligned} \tag{57}$$

where ζ' depends on $(x(z), \hat{\xi}_0)$ (the indices μ, ν for the reflection coefficients have been explained in Section 3). The integration $\int \mathrm{d}z'$ is now replaced by $\int \delta(z_n)\mathrm{d}z$. The latter can be transformed back to an integral over x. Thus we have the following expression

$$(2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2}-n} \int \Big(2i\tau A_M(\hat{x}_{\hat{I}}, x(z), \hat{\xi}_{\hat{J}}, \tau) A_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) R_{\mu(M)\nu(N)}(z, \zeta') + \text{l.o.t.}\Big)$$
$$\times\, \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x},\tilde{x},t,x,\hat{\xi}_{\hat{J}},\tilde{\xi}_{\tilde{J}},\tau)} \delta(z_n(x))\, \mathrm{d}\hat{\xi}_{\hat{J}}\, \mathrm{d}\tilde{\xi}_{\tilde{J}}\, \mathrm{d}\tau\, \mathrm{d}x. \tag{58}$$

This formula is very similar to (32), only the amplitude is different and $\frac{\delta c_{ijkl}(x)}{\rho(x)}, \frac{\delta\rho(x)}{\rho(x)}$ is replaced by the δ-function $\delta(z_n(x))$. The phase function Φ_{MN} now comes from the smooth medium above the reflector.

The data is modeled by $G^{\text{refl}}_{MN}(\hat{x}, \tilde{x}, t)$, with $(\hat{x}, \tilde{x}, t)$ in the acquisition manifold, see the text following Lemma 4.1. We follow the approach of Section 4, and do a coordinate transformation $(\hat{x}, \tilde{x}, t) \mapsto (y', y'')$, such that the acquisition manifold is given by $y'' = 0$. It follows that under Assumptions 3, 4 the data is the image of a Fourier integral operator acting on $\delta(z_n(x))$ and that it is given by (56). $\square$

We now show our main result, by applying the results of the previous section on the Kirchhoff modeling formula (56), and its equivalent in the Born approximation (32).

Theorem 6.2 *Suppose microlocally Assumptions 1, 2, 3, 4, 5 are satisfied. Let H_{MN} be the Fourier integral operator with as canonical relation the extended map $(x, \xi, e, \epsilon) \mapsto (y', \eta')$ constructed in Section 5, and with amplitude to highest order given by $(2\pi)^{\frac{n}{2}}(2i\tau)B_{MN}(y'_I, x, \eta'_J, \epsilon)$, such that $B_{MN}(\epsilon = 0)$ is as given in Theorem 4.2. Then the data in both Born and Kirchhoff approximation is given by H_{MN} acting on a function $r_{MN}(x, e)$. For the Kirchhoff approximation*

$$r_{MN}(x, e) = (\text{pseudo}(x, D_x, e))\delta(z_n(x)), \tag{59}$$

and to highest order $r_{MN}(x, e) = R_{MN}(x, e)\delta(z_n(x))$. For the Born approximation the function $r_{MN}(x, e)$ is given by a pseudodifferential operator acting on $\left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right)_\alpha$, with principal symbol $(2i\tau(x, \xi, e))^{-1} w_{MN;\alpha}(x, \xi, e)$, see (34).

Proof We do the proof for the Kirchhoff approximation using (56), for the Born approximation it is similar. Since Assumption 5 is satisfied, the projection $\pi_{Y'}$ of Λ_{MN} into $T^*Y'\backslash 0$ is an embedding, and the image is a coisotropic submanifold of T^*Y'. Therefore we can apply Lemma 5.1. Formula (55) gives that the factor $\mathrm{e}^{\mathrm{i}\Phi_{MN}}$ can be written as

$$\begin{aligned}\mathrm{e}^{\mathrm{i}(S_{MN}(y'_I,x,\eta'_J,0)+\langle y'_J,\eta'_J\rangle)} &= (2\pi)^{-(n-1-c)} \int \mathrm{e}^{\mathrm{i}(S_{MN}(y'_I,x,\eta'_J,\epsilon)+\langle y'_J,\eta'_J\rangle-\langle e,\epsilon\rangle)}\, \mathrm{d}\epsilon\, \mathrm{d}e \\ &= (2\pi)^{-(n-1-c)} \int \mathrm{e}^{\mathrm{i}\Psi_{MN}(y',x,e,\eta'_J,\epsilon)}\, \mathrm{d}\epsilon\, \mathrm{d}e.\end{aligned} \tag{60}$$

So the number of phase variables is increased by using stationnary phase. Let $B_{MN}(y'_I, x, \eta'_J, \epsilon)$ be as described. Then we obtain

$$
\begin{aligned}
d_{MN}(y') = {} & (2\pi)^{-\frac{|J|+n-1-c}{2}-\frac{2n-1-c}{2}} \\
& \times \int \left((2\pi)^{\frac{n}{2}} 2i\tau(\eta') B_{MN}(y'_I, x, \eta'_J, \epsilon) R_{MN}(x,e) + \text{l.o.t.} \right) \\
& \times \mathrm{e}^{\mathrm{i}\Psi_{MN}(y',x,e,\eta'_J,\epsilon)} \delta(z_n(x)) \, \mathrm{d}\eta'_J \, \mathrm{d}\epsilon \, \mathrm{d}x \, \mathrm{d}e. \qquad (61)
\end{aligned}
$$

In this formula the data is represented as a Fourier integral operator acting on a function of (x, e) given by $\delta(z_n(x))$. Multiplying by H_{MN}^{-1} gives a pseudodifferential operator of the form described acting on $\delta(z_n(x))$. Thus we obtain the result. □

Thus given the medium above the reflector (in Kirchhoff approximation) the function $r_{MN}(x, e)$ can be reconstructed by applying the Fourier integral H_{MN}^{-1} to the data. Hence we have the following result for Kirchhoff data.

Corollary 6.3 *Suppose that the medium above the reflector is given, and that it satisfies Assumptions 1, 2, 3, 4, 5. Then one can reconstruct position of the interface and angle dependent reflection coefficient $R_{\mu\nu}(x, e)$ on the interface.*

The operator H_{MN} transforms the data to (x, e) coordinates. If e is chosen as scattering angle and azimuth we have a transformation of the data to subsurface position and scattering angle/azimuth coordinates, which is new. The advantage of these coordinates is that multipathing is incorporated.

The motivation for Lemma 5.1 can now also be explained. Suppose there is high frequency data that is not from a given model. In the Kirchhoff case this may be because the medium above the interface is not correctly chosen, or because the data cannot be modeled at all by Kirchhoff modeling. To such data there is no natural value of the scattering angle/azimuth associated. So to transform it to (x, e) coordinates the value of e must be chosen. This is precisely the choice that we have in the proof of Lemma 5.1, where the function $e(y', \eta')$ on $T^*Y'\backslash 0$ is chosen.

It is known how to transform data to the (x, e) domain when e is chosen to be the offset. Assume $X = \mathbb{R}^n_{x_n > 0}, \partial X = \mathbb{R}^{n-1} = \{x \in \mathbb{R}^n \mid x_n = 0\}$, and the acquisition manifold is given by $\partial X \times \partial X \times]0, T[$. In the absence of certain degenerate ray geometries e can be the offset $e = \hat{x} - \tilde{x} \in \mathbb{R}^{n-1}$. This coordinate is automatically defined on all of $T^*Y'\backslash 0$, which gives automatically an extension as in Lemma 5.1. Define the midpoint coordinate $m = \frac{\hat{x}-\tilde{x}}{2}$. The operator H_{MN} is now for each fixed value of h an invertible Fourier integral operator mapping functions (m, t) to functions of $x \in X$.

7 Reconstructing the smooth part of the mediumparameters

The result of the previous section gives information on the problem of reconstructing the medium above the interface, or, in the Born approximation, of the background medium. Suppose there is a redundancy in the data, i.e. the dimension of the variable e, $n-1-c>0$. If the smooth mediumparameters above the interface are correct, then applying the operator H_{MN}^{-1} of Theorem 6.2 to the data results in a reflectivity function $r_{MN}(x,e)$, such that the position of the singularities does not depend on e. This can be used as a criterion to determine whether the medium above the interface or the background medium is correct. This technique is called velocity analysis, because in the acoustic case one determines in this way the local propagation speed of the acoustic waves. If the smooth medium parameters above and below the interface are correct, then also the amplitude of the singularities of $r_{MN}(x,e)$ should be proportional to the reflection coefficients. This could also be used in the determination of the medium above the interface. However, information about the position of the singularities (traveltimes) is often more reliable than information about the amplitudes.

This is well known in the case where e is given by the offset, $e=\hat{x}-\tilde{x}$. Thus we have generalized this method to using any coordinate e, in particular we can use the scattering angle, which depends only on the phase directions at the scattering point.

We mention the two most important criteria to measure how well the data "line up", see Symes [21] for a discussion. The first criterium is called "stacking power". Assume the reflection coefficient that has the same sign for each e, then the integral

$$\int \left| \int r_{MN}(x,e)\,\mathrm{d}e \right|^2 \mathrm{d}x$$

is maximal when the data line up.

The second way to measure how well the data lines up is essentially by taking the derivative with respect to e. If $r_{MN}(x,e)$ depends smoothly on e as in (59), then $\frac{\partial}{\partial e} r_{MN}(x,e)$ is one order less singular (for instance in a Sobolev space) than if it would not have this smooth dependence on e. One can now try to find the medium above the reflector by minimizing the semblance norm

$$\left\| \frac{\partial r_{MN}}{\partial e}(x,e) \right\|^2, \tag{62}$$

where a suitable norm should be chosen (we do not go into this). Taking also the factor in front of the δ function of r_{MN} into account, see (59), we obtain that to the highest two orders

$$\left(R_{MN}(x,e)\frac{\partial}{\partial e} - \frac{\partial R_{MN}}{\partial e}(x,e) \right) r_{MN}(x,e) = 0. \tag{63}$$

If $R_{MN}(x,e)$ is nonzero then the lower order terms can be chosen such that this equation is valid to all orders.

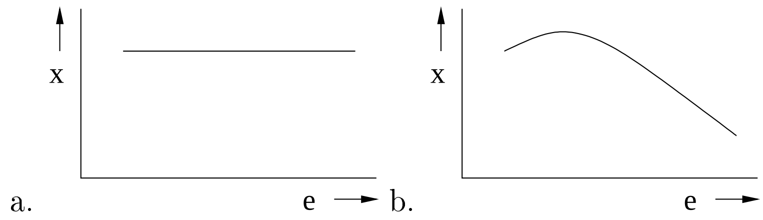

Figure 6: Examples of the singularities of $r_{MN}(x,e)$, when the medium above the reflector is correct (a.) or incorrect (b.)

Conjugating the differential operator of (63) with the invertible FIO H_{MN} we obtain pseudodifferential operator on $\mathcal{D}'(Y')$. Thus we obtain the following corollary of Theorem 6.2

Corollary 7.1 *Let the pseudodifferential operators $Q_{MN}(y', D_{y'})$ be given by*

$$Q_{MN}(y', D_{y'}) = H_{MN}\left(R_{MN}(x,e)\frac{\partial}{\partial e} - \frac{\partial R_{MN}}{\partial e}(x,e)\right)H_{MN}^{-1}.$$

Then for Kirchhoff data $d_{MN}(y')$ we have to the highest two orders

$$Q_{MN}(y', D_{y'})d_{MN}(y') = 0. \tag{64}$$

For values of e where $R_{MN}(x,e) \neq 0$ the operator $Q_{MN}(y', D_{y'})$ can be chosen such that (64) is valid to all orders.

Notation

n	dimension of space
x	position in medium
X	subset of $\mathbb{R}^n$ where the medium is
subscript i, j, k, l	indices space variables and elastic indices
δ_{ij}	Kronecker delta
t	time
ξ, τ, etc.	cotangent vectors corresponding to x, t,etc.
y'	coordinates on acquisition manifold
$\rho(x)$	mass density in the medium
$c_{ijkl}(x)$	elastic tensor
u_i	normalized displacement, see (3)
f_i	normalized elastic force density, see (3)
P_{il}	normalized elastic wave operator, see (4)
subscript M, N	indices over elastic mode
$P_M(x, D)$	wave operator for each mode (ψdo), see (5)
$Q_{iM}(x, D)$	pseudodifferential operator that diagonalizes P_{il} (contains the polarization vectors)
u_M, f_M	amplitude and force density of each mode
A_{il}	spatial part of the wave operator
$A_M(x, D)$	spatial part of decoupled wave operator (ψdo)
$B_M(x, D)$	square root of $A_M(x, D)$, see above (11)
$u_{M,\pm}, f_{M,\pm}$	amplitude and force density for first order equations
$x_M(x_0, \xi_0, t), \xi_M(x_0, \xi_0, t)$	bicharacteristic, see (13)
$G_{M,\pm}$	Green's function for first order decoupled equations
$C_{M,\pm}, C_M$	canonical relation of $G_{M,\pm}, G_M$, see (14)
$\phi_{M,\pm}, \phi_M$	phase function for $G_{M,\pm}, G_M$
$A_{M,\pm}(x_I, \dots)$	amplitude function for $G_{M,\pm}$
subscript I, J	partition of some set $\{1, \dots, k\}$ in two disjoint subsets
x_I	$\{x_i \mid i \in I\}$
superscript prin	to indicate that we have a principal symbol
subscript a	index for the $2n$ component vector of (25)
$R^0_{\mu\nu}$	reflection coefficients ΨDO, for amplitudes
$R_{\mu\nu}$	reflection coefficients ΨDO, see Theorem 3.1
v_a	displacement and traction (25)
v_μ	decoupled displacement and traction, see below (26)

$\delta c_{ijkl}, \delta\rho$	medium perturbation in Born approximation
$\delta G_{il}, \delta G_{MN}$	perturbation of Green's function
$d_{MN}(y')$	data, for a pair of modes (M, N)
g_α	$\left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right)$
$F_{MN\alpha}$	operator mapping $g_\alpha(x) \mapsto d_{MN}^{\mathrm{Born}}(y')$
Φ_{MN}	phase function for $F_{MN\alpha}$
$B_{MN}, w_{MN;ijkl}, w_{MN;0}$	amplitude factors for $F_{MN\alpha}$
Λ_{MN}	canonical relation of $F_{MN\alpha}$
$e \in E$	coordinate on Λ_{MN} together with (x, ξ), see (42)
$\pi_X, \pi_{Y'}$	projection from Λ_{MN} to $T^*X\backslash 0$, resp. $T^*Y'\backslash 0$
$N_{MN\alpha\beta}$	normal operator $F_{MN\alpha}^* F_{MN\beta}$
$\sigma_{Y'}$	symplectic form on $T^*Y'\backslash 0$
M_{MN}	canonical relation extending Λ_{MN}, see Section 5
Ψ_{MN}	phase function extending Φ_{MN}, see (55)
$r_{MN}(x, e)$	"reflectivity" function, see Theorem 6.2
H_{MN}	operator mapping $r_{MN}(x, e)$ to data, see Theorem 6.2
$Q_{MN}(y', D_{y'})$	pseudodifferential operator that annihilates data, see Corollary 7.1

We use the Einstein summation convention (summation over repeated indices), unless explicitly mentioned. We use the notation $Q(x, D)$ for a pseudodifferential operator with symbol $Q(x, \xi)$.

References

[1] Keiiti Aki and Paul G. Richards. *Quantitative seismology: theory and methods*, volume 1. Freeman, San Francisco, 1980.

[2] G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *Journal of Mathematical Physics*, 26:99–108, 1985.

[3] G. Beylkin and R. Burridge. Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion*, 12:15–52, 1990.

[4] Norman Bleistein. On the imaging of reflectors in the earth. *Geophysics*, 52:931–942, 1987.

[5] P.J. Braam and J.J. Duistermaat. Normal forms of real symmetric systems with multiplicity. *Indag. Mathem.*, 4:197–232, 1993.

[6] C.H. Chapman. Reflection/transmission coefficient reciprocities in anisotropic media. *Geophysical Journal International*, 116:498–501, 1994.

[7] Maarten V. de Hoop and Norman Bleistein. Generalized Radon transform inversions for reflectivity in anisotropic elastic media. *Inverse Problems*, 13:669–690, 1997.

[8] Maarten V. de Hoop and Sverre Brandsberg-Dahl. Maslov asymptotic extension of generalized Radon transform in anisotropic elastic media: A least squares approach. *Inverse Problems*, To appear, 1999.

[9] Nils Dencker. On the propagation of polarization sets for systems of real principal type. *Journal of Functional Analysis*, 46:351–372, 1982.

[10] J.J. Duistermaat. *Fourier Integral Operators.* Birkhäuser, Boston, 1996.

[11] S. Hansen. Solution of a hyperbolic inverse problem by linearization. *Communications in Partial Differential Equations*, 16:291–309, 1991.

[12] Lars Hörmander. *The Analysis of Linear Partial Differential Operators*, volume 1-4. Springer-Verlag, Berlin, 1983-1985.

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

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

[15] B.L.N. Kennett. *Seismic Wave Propagation in Stratified Media.* Cambridge University Press, Cambridge, 1983.

[16] J.L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications*, volume 1. Springer-Verlag, Berlin, 1972.

[17] V.P. Maslov and M.V. Fedoriuk. *Semi-Classical Approximation in Quantum Mechanics*. D. Reidel Publishing Company, 1981.

[18] C.J. Nolan and W.W. Symes. Global solution of a linearized inverse problem for the wave equation. *Communications in Partial Differential Equations*, 22(5-6):919–952, 1997.

[19] Rakesh. A linearized inverse problem for the wave equation. *Communications in Partial Differential Equations*, 13(5):573–601, 1988.

[20] Christiaan C. Stolk. Microlocal analysis of a seismic linearized inverse problem. Preprint 1095, Department of Mathematics, Utrecht University, March 1999.

[21] William W. Symes. A differential semblance algorithm for the inverse problem of reflection seismology. *Comput. Math. Appl.*, 22(4-5):147–178, 1991.

[22] Michael E. Taylor. Reflection of singularities of solutions to systems of differential equations. *Communications on Pure and Applied Mathematics*, 28:457–478, 1975.

[23] A. P. E. ten Kroode, D. J. Smit, and A. R. Verdel. A microlocal analysis of migration. *Wave Motion*, 28:149–172, 1998.

[24] François Treves. *Introduction to Pseudodifferential and Fourier Integral Operators*, volume 1. Plenum Press, New York, 1980.

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