Microlocal Analysis of Seismic Inverse Scattering in Anisotropic Elastic Media

CHRISTIAAN C. STOLK
Utrecht University
AND
MAARTEN V. DE HOOP
Colorado School of Mines

Abstract

Seismic data is modeled in the high-frequency approximation, using the techniques of microlocal analysis. We consider general, anisotropic elastic media. Our methods are designed to allow for the formation of caustics. The data is modeled in two ways. First, we give a microlocal treatment of the Kirchhoff approximation, where the medium is assumed to be piecewise smooth, and reflection and transmission occur at interfaces. Second, we give a refined view on the Born approximation based upon a linearization of the scattering process in the medium parameters around a smooth background medium. The joint formulation of Born and Kirchhoff scattering allows us to take into account general scatterers as well as the nonlinear dependence of reflection coefficients on the medium parameters. The latter allows the treatment of scattering up to grazing angles.

The outcome of the analysis is a characterization of the singular part of seismic data. We obtain a set of pseudodifferential operators that annihilate the data. In the process we construct a Fourier integral operator and a reflectivity function such that the data can be represented by this operator acting on the reflectivity function. In our construction this Fourier integral operator becomes invertible. We give the conditions for invertibility for general acquisition geometry. The result is also of interest for inverse scattering in acoustic media.

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. Subsurface position is denoted by x. Sources and receivers are contained in the boundary ∂X of X. Their position is denoted by $\tilde{x}$ and $\hat{x}$,

Communications on Pure and Applied Mathematics, Vol. LV, 0261–0301 (2002)

respectively. Measurement of data takes place 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 coordinates y' on Y' are given. We also assume that the displacement of the waves is measured for point sources at $\tilde{x}, t = 0$ with all its components, both at the source and at the receiver. Thus we assume that (after preprocessing) the data matches 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 may lie along a line behind the source, in which case we have $n = 3, c = 1, \partial X = \{x \in \mathbb{R}^n : x_3 = 0\}, Y' = \{(\hat{x}, \tilde{x}, t) \in \mathbb{R}^3 \times \mathbb{R}^3 \times]0, T[: \hat{x}_3 = \tilde{x}_3 = \hat{x}_2 - \tilde{x}_2 = 0\}$, so the data is a function of $2n - 1 - c$ variables. From this data we try to determine a function of n variables; hence there is a redundancy in the data of dimension $n - 1 - c$.

Our approach follows the work of Beylkin [2, 3] and other authors (see the references below), applying microlocal analysis to the seismic inverse problem. Microlocal analysis and the theory of Fourier integral operators are described in the books by Hörmander [16, 17, 18], Duistermaat [10], and Treves [30, 31]. Beylkin [3] considered the seismic inverse scattering problem in acoustic media with constant density. He modeled the data using the Born approximation, where the scattering is linearized in the medium coefficients. The medium perturbation $\delta c(x)$ acts as a distribution of scatterers superimposed on a smooth background medium $c(x)$. Given the background medium $c(x)$, an operator was given to reconstruct $\delta c(x)$ microlocally from an n-dimensional subset of the data (from data that is a function of n variables). This was done under certain conditions on the rays. In particular, the situation where the wavefronts form caustics was excluded.

When the data is redundant in the sense that the available data is a function of more than n variables, then the data can be seen as a family of n-dimensional datasets, where each n-dimensional subset in the family has a fixed value of some coordinate, which we refer to as e (in terms of acquisition variables, this would be the collection of offset coordinates $\hat{x} - \tilde{x}$). The result of the inversion, let us call this the reflectivity $r(x, e)$, should not depend on e. This is the criterion that must be used to determine the background medium from the data; see, for example, Symes [28]. In summary, under the assumptions made by Beylkin [3], there is microlocally an invertible map that maps seismic data to a reflectivity function $r(x, e)$, of which the singular part should not depend on e.

In this paper we will generalize the results of Beylkin and Symes in two directions. First, we remove the restriction that there cannot be caustics. Such a transformation from data to a reflectivity function $r(x, e)$ was previously not defined for data in the neighborhood of a caustic from the scattering point, even in acoustic media. Second, we consider general elastic media instead of acoustic media. We decouple the system into scalar equations (see Taylor [29], Ivrii [19], Dencker [9]) and apply Fourier integral operator techniques to take into account the situation where the wavefronts form caustics. We give results for all orders.

The fact that we consider general elastic media makes the result technically more complicated and may make it more difficult to see some of the essential ideas that can also be applied to the acoustic case. On the other hand, there are several good reasons why the results are particularly useful in elastic media. For instance, caustics form "instantaneously" in elastic media; they may even occur in constant-coefficient media. Also, for elastic media the dependence of reflection coefficients on the scattering angle is more complicated, and it is more important to use this information in the inversion of seismic data.

The data is modeled 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 extendable 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, following the approach of Taylor [29]. In this way we construct a generalization of the Kirchhoff approximation. In Section 4 we discuss the Born approximation. This is essentially a linearization where the medium parameters 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 7.1. The multimodal data can be written as an invertible Fourier integral operator H_{MN} acting on a "reflectivity" distribution $r_{MN}(x, e)$, that is, a function of subsurface position x and the additional variable e, essentially parameterizing the scattering angle and azimuth. The position of the singularities of $r_{MN}(x, e)$ does not depend on e. In the Kirchhoff approximation for elastic media the function $r_{MN}(x, e)$ equals to highest order $R_{MN}(x, e)\| \frac{\partial z_n}{\partial x} \| \delta(z_n(x))$, where $R_{MN}(x, e)$ is the appropriately normalized reflection coefficient for the pair of elastic modes (M, N), and $\| \frac{\partial z_n}{\partial x} \| \delta(z_n(x))$ is the singular function of the interface. For the Born approximation, $r_{MN}(x, e)$ is given by pseudodifferential operators that take into account the radiation patterns acting on the medium perturbation.

The result is new even for acoustic media. In that case the coordinate e can be chosen as scattering angle and azimuth. For acoustic media such a map has been proposed in the geophysical literature to highest order only and when the acquisition point is not at a caustic from the scattering point; see Xu, Chauris, Lambaré, and Noble [32].

The new step in the proof that is needed to deal with the presence of caustics is given in Section 5. The coordinate e is a priori only defined on the co-isotropic subset $\mathcal{L} \subset T^*Y' \setminus 0$ that contains the wavefront set of the data (the subset L_ϕ in Ten Kroode, Smit, and Verdel [21]). To construct an invertible Fourier integral operator from data to the function $r_{MN}(x, e)$, the coordinate e has to be defined on an open part of $T^*Y' \setminus 0$. This is done in Lemma 5.1, where we construct an extension of the coordinate function e from $\mathcal{L}$ to an open neighborhood of $\mathcal{L}$ in $T^*Y' \setminus 0$.

The extension is not unique. In Xu et al. [32] the choice of the extension is made implicitly by using the fact that, given the scattering point x, there is a locally diffeomorphic map from the source and receiver coordinates $(\hat{x}, \tilde{x}) \in \partial X \times \partial X$ to the dip and the scattering angle/azimuth (in the notation of Section 5 given by $(\xi/\|\xi\|, e)$), assuming that the acquisition points are not at a caustic from x.

The results hold microlocally away from points in the cotangent space $T^*Y' \setminus 0$ that violate our assumptions 1 to 5, introduced in the main text. The assumptions exclude certain degenerate ray (bicharacteristic) geometries. For example, assumptions 1, 2, and 3 exclude rays that go through a singularity of the slowness surface, rays tangent to an interface, and direct rays from source to receiver, respectively. In general, the set of $(y', \eta') \in T^*Y' \setminus 0$ where the assumptions are violated has lower dimension than the dimension of $T^*Y' \setminus 0$. The data associated to such (y', η') can be muted using a pseudodifferential cutoff.

As a consequence of Theorem 7.1, we obtain results about the reconstruction of the medium parameters. 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 7.3. For the Born approximation a similar result holds, but an inverse is also obtained directly in Theorem 4.5.

When the data is redundant (c sufficiently small) there is in addition a criterion 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 $W_{MN}(y', D_{y'})$ that, if the medium above the interface is correctly chosen, annihilate the data; see Corollary 7.4. This allows one to do differential semblance optimization [28] in elastic media with caustics.

We discuss some of the literature on this subject. There have been many publications about high-frequency methods to invert seismic data in acoustic media. The reconstruction of the singular component of the medium coefficients in the Born approximation, without caustics, has been done in the papers by Beylkin [2, 3]. Bleistein [5] discusses the case of a smooth jump using Beylkin's results. It has been shown by Rakesh [26] that the modeling operator in the Born approximation is a Fourier integral operator. Hansen [13] studied the inversion in an acoustic medium with multipathing for both the Born approximation and the case of a smooth jump. Ten Kroode et al. [21] also treat the case of seismic imaging in the presence of multipathing. They discuss in more detail the assumptions (most importantly Assumption 5(ii) below) that are made about the geometry of the rays underlying the scattering. Stolk [27] discusses the case when Assumption 5(ii) is violated. Nolan and Symes [25] discuss the imaging with different acquisition geometries. The article by Symes [28] 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 high-frequency approximation has been given by Taylor [29]. This fundamental paper also discusses the interface problem. Beylkin and Burridge [4] discuss the imaging of seismic data in the Born approximation in isotropic elastic media under a no-caustics assumption. De Hoop and Bleistein [14] 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 [15].

An overview of the paper follows. In Section 2 we discuss the propagation of waves in smooth elastic media. First, we discuss how asymptotically the elastic system can be decoupled by conjugating with appropriately chosen 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 operators describing reflected and transmitted waves. These solutions were already discussed but not explicitly constructed by Taylor [29]. Thus we prove directly the validity of the Kirchhoff approximation, which is not obvious from, for example, De Hoop and Bleistein [14]. 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 a comprehensive 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 characterize the geometry of the wave front set of the data. Under the assumptions of Section 4 this set is contained in a co-isotropic submanifold $\mathcal{L}$ of the cotangent space $T^*Y' \setminus 0$. We discuss the extension of symplectic coordinates on $\mathcal{L}$ to a neighborhood of $\mathcal{L}$ in $T^*Y' \setminus 0$.

In Section 6 we establish microlocally a correspondence between the Kirchhoff approximation and the Born approximation. After the preparations of Sections 2 through 6, the derivation of our main result in Section 7 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, we construct pseudodifferential operators that annihilate the high-frequency part of the data. In principle, these can be used for the reconstruction of the smoothly varying medium parameters above the interface (or of the background medium in the Born approximation).

2 Propagation of Elastic Waves in Smoothly Varying Media

2.1 Decoupling the Modes

The elastic wave equation is given by

$$\left(\rho\delta_{il}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}c_{ijkl}\frac{\partial}{\partial x_k}\right)(\text{displacement})_l = (\text{vol. force density})_i\,. \tag{2.1}$$

Here $\rho(x)$ is the volume density of mass and $c_{ijkl}(x)$ is the elastic stiffness tensor, and $i, j, k, l = 1, 2, \ldots, n$.

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

$$P_{il}u_l = f_i\,, \tag{2.2}$$

where

$$u_l = \sqrt{\rho}(\text{displacement})_l\,, \qquad f_i = \frac{1}{\sqrt{\rho}}(\text{vol. force density})_i\,, \tag{2.3}$$

and

$$P_{il} = \delta_{il}\frac{\partial^2}{\partial t^2} - \frac{\partial}{\partial x_j}\frac{c_{ijkl}}{\rho}\frac{\partial}{\partial x_k} + \text{lower-order terms} \tag{2.4}$$

is the partial differential operator. Here we use that ρ is smooth and bounded away from zero. Both systems (2.1) and (2.2) are real and time-reversal invariant, and satisfy reciprocity.

We describe how system (2.2) can be decoupled by transforming it with appropriate pseudodifferential operators; see Taylor [29], Ivrii [19], and Dencker [9]. The goal is to transform the operator P_{il} by conjugation with a matrix-valued pseudodifferential operator $Q(x, D)_{iM}$, $D = D_x = -\mathrm{i}\frac{\partial}{\partial x}$, to an operator that is of diagonal form modulo a regularizing part,

$$Q(x, D)^{-1}_{Mi}\, P_{il}(x, D, D_t)\, Q(x, D)_{lN} = \operatorname{diag}(P_M(x, D, D_t))_{MN}\,, \tag{2.5}$$

$M, N \in \{1, 2, \ldots, n\}$, $D_t = -\mathrm{i}\frac{\partial}{\partial t}$. Here the indices M and N denote the mode of propagation. In fact, for the construction of Fourier integral operator solutions as in the scalar case, it is sufficient to transform the partial differential operator to block-diagonal form, where each of the blocks $P_M(x, D, D_t)$ has scalar principal part (proportional to the identity matrix). In this case we will use the indices M and N to denote the block, and we will omit indices for the components within each block. Let

$$u_M = Q(x, D)^{-1}_{Mi}u_i\,, \qquad f_M = Q(x, D)^{-1}_{Mi}\, f_i\,. \tag{2.6}$$

The system (2.2) is then equivalent to the uncoupled equations

$$P_M(x, D, D_t)u_M = f_M\,. \tag{2.7}$$

The time derivative in P_{il} is already in diagonal form; hence, we only have to diagonalize its 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 (2.5) is valid with P_{il} and P_M replaced by A_{il} and A_M. The operator P_M is then given by

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

Given the properties of stiffness, the principal symbol $A_{il}^{\text{prin}}(x, \xi)$ of $A_{il}(x, D)$ is a positive symmetric matrix, so it can be diagonalized by an orthogonal matrix. On the level of principal symbols, composition of pseudodifferential operators reduces to 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 of $A_{il}^{\text{prin}}(x, \xi)$ so that

$$Q_{Mi}^{\text{prin}}(x, \xi)^{-1} A_{il}^{\text{prin}}(x, \xi) Q_{lN}^{\text{prin}}(x, \xi) = \text{diag}(A_M^{\text{prin}}(x, \xi))_{MN} . \tag{2.8}$$

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 $A_M^{\text{prin}}(x, \xi)$ are constant, then the principal symbol $Q_{iM}^{\text{prin}}(x, \xi)$ depends smoothly on (x, ξ), and microlocally equation (2.8) carries over to an operator equation. Taylor [29] has shown that if this condition is satisfied, then decoupling can be accomplished to all orders, where each block corresponds to a different eigenvalue. In fact, he proved the following slightly more general result.

LEMMA 2.1 (Taylor) *Suppose the pseudodifferential operator $Q_{iM}(x, D)$ of order 0 is such that*

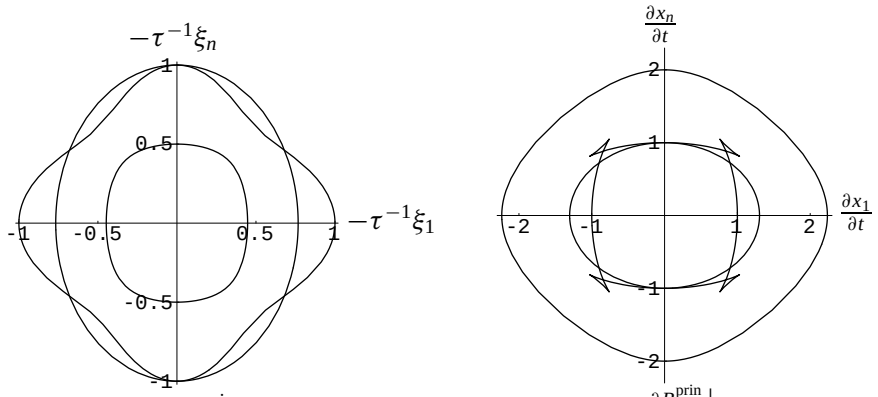
$$Q(x, D)_{Mi}^{-1} A(x, D)_{il} Q(x, D)_{lN} = \begin{pmatrix} A_{(1)}(x, D) & 0 \\ 0 & A_{(2)}(x, D) \end{pmatrix}_{MN} + a(x, D)_{MN} ,$$

where $A_{(1)}(x, \xi)$ and $A_{(2)}(x, \xi)$ are homogeneous of order 2, and $a(x, \xi)_{MN}$ is polyhomogeneous of order 1. Suppose the spectra of $A_{(1)}(x, \xi)$ and $A_{(2)}(x, \xi)$ are disjoint on a conic neighborhood of some $(x_0, \xi_0) \in T^*X \setminus 0$. Then by modifying Q with lower-order terms, the system can be transformed such that

$$a(x, D)_{MN} = \begin{pmatrix} a_{(1)}(x, D) & 0 \\ 0 & a_{(2)}(x, D) \end{pmatrix}_{MN} + \text{smoothing remainder}$$

microlocally around (x_0, ξ_0).

This implies that if the multiplicity of a particular eigenvalue $A_M^{\text{prin}}(x, \xi)$ is constant, then the system can be transformed such that the part related to this eigenvalue decouples from the rest of the system modulo a smoothing remainder. In this work we will assume that at least some of the modes decouple (microlocally). This is stated as Assumption 1 below. At that point we will also discuss whether this assumption is satisfied in relevant cases.



(a) Cotangent: $B_M^{\text{prin}}(x, \tau^{-1}\xi) = 1$ (b) Tangent: $\frac{\partial B_M^{\text{prin}}}{\partial \xi}\Big|_{B_M^{\text{prin}}(x,\tau^{-1}\xi)=1}$

FIGURE 2.1. (a) Section of a slowness surface (the characteristic surface) for a transversely isotropic medium in $n = 3$ dimensions. (b) Set of velocities associated to the slowness surface in (a). Note the caustics that occur due to the fact that one of the sheets is not convex.

We give an alternative characterization of the quantities

$$A_M^{\text{prin}}(x, \xi) \quad \text{and} \quad Q_{iM}^{\text{prin}}(x, \xi)\,.$$

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

$$\det P_{il}^{\text{prin}}(x, \xi, \tau) = 0\,. \tag{2.9}$$

The multiplicity of $A_M^{\text{prin}}(x, \xi)$ is equal to the multiplicity of the corresponding root of (2.9). The columns of $Q_{iM}^{\text{prin}}(x, \xi)$ satisfy

$$Q_{iM}^{\text{prin}} \in \ker P_{il}^{\text{prin}}(x, \xi, \sqrt{A_M^{\text{prin}}(x, \xi)})\,.$$

Since $P_{il}^{\text{prin}}(x, \xi, \tau)$ is homogeneous in (ξ, τ), one may choose to use the slowness vector $-\tau^{-1}\xi$ instead of the cotangent or wave vector ξ in calculations. The set of $-\tau^{-1}\xi$ such that (2.9) holds is called the slowness surface, which can be easily visualized. A section of the slowness surface for the case of a transversely isotropic medium in three dimensions is given in Figure 2.1(a). Note that the slowness surface need not be convex. The multiplicity of the eigenvalues changes at the points (directions) where the different sheets intersect.

The second-order equations (2.7) clearly describe the decoupling of the original system into different elastic modes. In addition, equations (2.7) inherit the symmetries of the original system. It is easy to see that they are time-reversal invariant. The operators $Q_{iM}(x, D)$ and $A_M(x, D)$ can be chosen in such a way that $Q_{iM}(x, \xi) = -\overline{Q_{iM}(x, -\xi)}$ and $A_M(x, \xi) = \overline{A_M(x, \xi)}$. Then Q_{iM} and A_M are real. We argue that equations (2.7) 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 also holds (modulo smoothing operators) for the Green's function $G_M(x, x_0, t - t_0)$ associated with (2.7). The transpose operator $Q(x, D)^{\mathsf{T}}_{Mi}$ (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 $Q^{\text{prin}}(x, \xi)^{\mathsf{T}}_{Mi}$. As noted before for the principal symbol, it follows from the fact that $A^{\mathsf{T}}_{ij} = A_{ij}$ that we can choose Q orthogonal, i.e., such that $Q(x, D)_{iM} Q(x, D)^{\mathsf{T}}_{Mj} = \delta_{ij}$. From the fact that

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

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

Remark 2.2. We already observed that if an eigenvalue $A^{\text{prin}}_M(x, \xi)$ has constant multiplicity $m_M > 1$, then u_M is an m-dimensional vector and (2.7) is a $m_M \times m_M$ system, with scalar principal symbol. For such a system a microlocal solution can be constructed in the same way as for scalar systems; see the next subsection. In this case all kinematic quantities, such as bicharacteristics, phase functions, and canonical relations, depend only on M. Other quantities such as u_M and $Q_{iM}(x, D)$ will have multiple components. The Green's function G_M and its amplitude $\mathcal{A}_M$, introduced in (2.20), are then $m_M \times m_M$ matrices. To simplify notation we do not take this into account explicitly. However, the reader may check that the results of this work can be generalized to this case.

2.2 The Green's Function

To evaluate the Green's function we use the first-order system for u_M that is equivalent to (2.7). It is given by

$$\frac{\partial}{\partial t} \begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A_M(x, D) & 0 \end{pmatrix} \begin{pmatrix} u_M \\ \frac{\partial u_M}{\partial t} \end{pmatrix} + \begin{pmatrix} 0 \\ f_M \end{pmatrix}. \tag{2.10}$$

This system can be decoupled also. Let $B_M(x, D) = \sqrt{A_M(x, D)}$, which is a pseudodifferential operator of order 1 that exists because $A_M(x, D)$ is positive definite. The principal symbol of $B_M(x, D)$ is given by $B^{\text{prin}}_M(x, \xi) = \sqrt{A^{\text{prin}}_M(x, \xi)}$. We then find that (2.10) is equivalent to the following two first-order equations:

$$\left(\frac{\partial}{\partial t} \pm \mathrm{i} B_M(x, D) \right) u_{M,\pm} = f_{M,\pm} \tag{2.11}$$

upon transforming

$$u_{M,\pm} = \frac{1}{2} u_M \pm \frac{1}{2} \mathrm{i} B_M(x, D)^{-1} \frac{\partial u_M}{\partial t} ,$$

$$f_{M,\pm} = \pm \frac{1}{2} \mathrm{i} B_M(x, D)^{-1} f_M . \tag{2.12}$$

We construct operators $G_{M,\pm}$ with distribution kernel $G_{M,\pm}(x, x_0, t)$ that solve the initial value problem for (2.11). Then using Duhamel's principle we find that

$$u_{M,\pm}(x,t) = \int_0^t G_{M,\pm}(x, x_0, t-t_0)\, f_{M,\pm}(x_0, t_0)\mathrm{d}x_0\, \mathrm{d}t_0\,.$$

It follows from (2.12) that the Green's function for the second-order decoupled equation is then given by

$$G_M(x, x_0, t) = \left(\frac{1}{2}\mathrm{i}G_{M,+}(x, x_0, t) - \frac{1}{2}\mathrm{i}G_{M,-}(x, x_0, t)\right) B_M(x_0, D_{x_0})^{-1}\,. \tag{2.13}$$

The operators $G_{M,\pm}$ are Fourier integral operators. Their construction is well known; see, for example, Duistermaat [10, chap. 5]. The singularities are propagated along the bicharacteristics that are determined by Hamilton's equations generated by the principal symbol (factor i divided out) $\tau \pm B_M^{\mathrm{prin}}(x,\xi)$ of (2.11). These equations read

$$\begin{aligned} \frac{\partial x}{\partial \lambda} &= \pm\frac{\partial}{\partial \xi} B_M^{\mathrm{prin}}(x,\xi)\,, & \frac{\partial t}{\partial \lambda} &= 1\,,\\ \frac{\partial \xi}{\partial \lambda} &= \mp\frac{\partial}{\partial x} B_M^{\mathrm{prin}}(x,\xi)\,, & \frac{\partial \tau}{\partial \lambda} &= 0\,. \end{aligned} \tag{2.14}$$

The solution may be parameterized by t. We denote the solution of (2.14) with the $+$ sign and initial values x_0 and ξ_0 by $(x_M(x_0,\xi_0,t), \xi_M(x_0,\xi_0,t))$. The solution with the $-$ sign is found by reversing the time direction; in other words, it is given by $(x_M(x_0,\xi_0,-t), \xi_M(x_0,\xi_0,-t))$. Observe that the group velocity (the velocity $\frac{\partial x}{\partial t}$ of the bicharacteristic) is orthogonal to the slowness surface. Where the slowness surface fails to be convex, caustics may arise instantly from a point source. An example is shown in Figure 2.1(b).

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

$$C_{M,\pm} = \left\{(x_M(x_0,\xi_0,\pm t), t, \xi_M(x_0,\xi_0,\pm t), \mp B_{M,\pm}(x_0,\xi_0); x_0, \xi_0)\right\}. \tag{2.15}$$

A convenient choice of phase function is described in Maslov and Fedoriuk [23]. They 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

$$(x_I, x_0, \xi_J, \tau)\,, \tag{2.16}$$

where $I \cup J$ is a partition of $\{1, 2, \ldots, n\}$. It follows from theorem 4.21 in Maslov and Fedoriuk [23] that there is a function $S_{M,+}(x_I, x_0, \xi_J, \tau)$ such that locally $C_{M,+}$ is given by

$$\begin{aligned} x_J &= -\frac{\partial S_{M,+}}{\partial \xi_J}\,, & t &= -\frac{\partial S_{M,+}}{\partial \tau}\,,\\ \xi_I &= \frac{\partial S_{M,+}}{\partial x_I}\,, & \xi_0 &= -\frac{\partial S_{M,+}}{\partial x_0}\,. \end{aligned} \tag{2.17}$$

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

$$\phi_{M,+}(x, x_0, t, \xi_J, \tau) = S_{M,+}(x_I, x_0, \xi_J, \tau) + \langle \xi_J, x_J \rangle + \tau t \,. \tag{2.18}$$

On the other hand, the canonical relation $C_{M,-}$ is given by

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

Thus a phase function for $C_{M,-}$ is

$$\phi_{M,-}(x, x_0, t, \xi_J, \tau) = -\phi_{M,+}(x, x_0, t, -\xi_J, -\tau) \,.$$

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 $\tau > 0$, and $\phi_M = \phi_{M,+}$ if $\tau < 0$.

We have to assume that the decoupling is valid microlocally around the bicharacteristic. In that case theorem 5.1.2 of Duistermaat [10] implies 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

$$G_{M,\pm}(x, x_0, t) = (2\pi)^{-\frac{|J|+1}{2} - \frac{2n+1}{4}} \int \mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau) \mathrm{e}^{\mathrm{i}\phi_{M,\pm}(x, x_0, t, \xi_J, \tau)} \, \mathrm{d}\xi_J \, \mathrm{d}\tau \,. \tag{2.19}$$

The factors of (2π) in front of the integral are according to the convention of Treves [31] and Hörmander [18].

The amplitude $\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)$ satisfies a transport equation along the bicharacteristics $(x_M(x_0, \xi_0, \pm t), \xi_M(x_0, \xi_0, \pm t))$. Properties of amplitudes are described, for instance, in Treves [31, sec. 8.4]. The amplitude is an element of $M_{C_M} \otimes \Omega^{1/2}(C_M)$, the tensor product of the Keller-Maslov bundle M_{C_M}, and the half-densities on the canonical relation C_M. If the subprincipal part of $A_M(x, D)$ is a matrix, then the amplitude is also a matrix; see Remark 2.2. The Keller-Maslov bundle gives a factor i^k where k is an index, which we will absorb in the amplitude. So the amplitude should be seen as a function on the canonical relation $C_{M,\pm}$, coordinatized by (x_I, x_0, ξ_J, τ); see (2.16). It is possible to choose a Maslov phase function with a different set of phase variables, for instance, $\xi_{\tilde{J}}$ (and not τ), where $\tilde{I} \cup \tilde{J}$ is a partition of $\{1, 2, \ldots, n\}$ and $C_{M,\pm}$ is parameterized by $(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$. In that case the transformed amplitude $\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$ contains a Jacobian factor to the power $\frac{1}{2}$, that is,

$$|\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})| = |\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)| \left| \frac{\partial(x_I, x_0, \xi_J, \tau)}{\partial(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})} \right|^{1/2} , \tag{2.20}$$

where in the Jacobian both sets of variables are coordinates on $C_{M,\pm}$.

We will calculate the left-hand side of (2.20). For this purpose, consider the Green's function $G_{M,\pm}(x, x_0, t - t_0)$ with t and $t_0 = 0$ fixed. It can be viewed as an invertible Fourier integral operator, mapping the displacement at $t = 0$, $u|_{t=0} \in$

$\mathcal{E}'(X)$, to the displacement at t, $u|_t \in \mathcal{D}'(X)$, with phase $\tilde{\phi}_{M,\pm}(x, x_0, t, \xi_{\tilde{J}})$ and amplitude $\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})$. To the highest order, the energy at time t is given by

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

Conservation of this quantity is reflected by the relation

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

where the left-hand side denotes a composition of Fourier integral operators and * denotes the adjoint. Since the left-hand side is a product of invertible Fourier integral operators, we can use the theory of section 8.6 in Treves [31]. We find that to highest order

$$\left|(2\pi)^{-1/4} \tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})\right|^2 = \left|\det \frac{\partial \xi_0}{\partial(x_{\tilde{I}}, \xi_{\tilde{J}})}\right| \left|\frac{B_M(x_0, \xi_0)}{B_M(x, \xi)}\right|^2 .$$

The value of $B_M(x, \xi)$ equals the frequency τ and is conserved along the bicharacteristic. Recall that (x_0, ξ_0, t) are valid coordinates for $C_{M,\pm}$ (cf. (2.15)). The Jacobian $|\frac{\partial(x_0, \xi_0, t)}{\partial(x_I, x_0, t, \xi_J)}|$ is equal to the factor $|\det \frac{\partial \xi_0}{\partial(x_I, \xi_J)}|$. It follows that to highest order

$$|\tilde{\mathcal{A}}_{M,\pm}(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})| = (2\pi)^{1/4} \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_{\tilde{I}}, x_0, t, \xi_{\tilde{J}})}\right|^{1/2} . \tag{2.21}$$

From (2.20) it now follows that

$$|\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)| = (2\pi)^{1/4} \left|\det \frac{\partial(x_0, \xi_0, t)}{\partial(x_I, x_0, \xi_J, \tau)}\right|^{1/2} . \tag{2.22}$$

We give our result about the Green's function for (2.7), collecting the results of this section, and using equations (2.12) and (2.22) to obtain a statement about the amplitude. We will assume that microlocally around the relevant bicharacteristics the decoupling is valid. Let $\mathrm{Char}(P_M)$ be the characteristic set of P_M given by $\{(x, t, \xi, \tau) : P(x, \xi, \tau) = 0\}$. The Green's function is such that precisely the singularities of f_M at $\mathrm{Char}(P_M)$ propagate; see Hörmander [17, theorem 23.2.9]. Thus we have the following:

Assumption 1. On a neighborhood of the bicharacteristic, the multiplicity of the eigenvalue $A_M^{\mathrm{prin}}(x, \xi)$ in (2.8) is constant.

Let $\mathrm{WF}(f)$ denote the wave front section of f [16, sec. 8.1].

LEMMA 2.3 *Suppose that for the bicharacteristics through* $\mathrm{WF}(f_M) \cap \mathrm{Char}(P_M)$ *Assumption* 1 *is satisfied. Then* u_M *is given microlocally, away from* $\mathrm{WF}(f_M)$, *by*

$$u_M(x, t) = \int G_M(x, x_0, t - t_0) \, f_M(x_0, t_0) \mathrm{d}x_0 \, \mathrm{d}t_0 \tag{2.23}$$

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

$$G_M(x, x_0, t) = (2\pi)^{-\frac{|J|+1}{2}-\frac{2n+1}{4}} \int \mathcal{A}_M(x_I, x_0, \xi_J, \tau) e^{i\phi_M(x,x_0,t,\xi_J,\tau)}\, d\xi_J\, d\tau\,. \tag{2.24}$$

For the amplitude $\mathcal{A}_M(x_I, x_0, \xi_J, \tau)$ we have to highest order

$$|\mathcal{A}_M(x_I, x_0, \xi_J, \tau)| = (2\pi)^{1/4}\frac{1}{2}|\tau|^{-1}\left|\det\frac{\partial(x_0, \xi_0, t)}{\partial(x_I, x_0, \xi_J, \tau)}\right|^{1/2}. \tag{2.25}$$

The implications of Assumption 1 for elastic media depend on which class of media one is interested in. By a class of media we mean a set of media parameterized by a number of (position dependent) parameters. From a physical point of view, one may be interested in media where the elastic tensor is characterized by certain symmetry properties.

Isotropic media are characterized by the mass density ρ and the Lamé parameters λ and μ. The matrix $A_{il}^{\text{prin}}(x, \xi)$ has two eigenvalues, $A_1^{\text{prin}}(x, \xi) = \frac{\lambda+2\mu}{\rho}\|\xi\|^2$ with polarization vector proportional to ξ (referred to as the P-mode), and $A_2^{\text{prin}}(x, \xi) = \frac{\mu}{\rho}\|\xi\|^2$ with polarization space normal to ξ (the two S-modes). Thus this system can be decoupled.

If the matrix $A_{il}^{\text{prin}}(x, \xi)$ of an isotropic medium is perturbed by a small amount, then one eigenvalue of the perturbed matrix will be close to the P-eigenvalue of the isotropic medium, and two eigenvalues will be close to the S-eigenvalue. The two eigenvalues close to the S-eigenvalue of the isotropic medium will not coincide in general but may coincide for certain values of (x, ξ). So in elastic media sufficiently close to an isotropic medium, there will still be a quasi-P mode that decouples from the other modes, but the two quasi-S modes will in general not decouple.

The elastic system for generic elastic media has been investigated by Braam and Duistermaat [6]. The set of singular points is generically of codimension 3 (thus 1 lower than one would expect naively) and is of conical form in the neighborhood of the singular point. They give a normal form for such systems and investigate the behavior of its associated bicharacteristics and polarization spaces. In this case the system cannot be decoupled. However, in a generic elastic medium there cannot be an open set of bicharacteristics that pass through a singular point, since the singular points form a set of codimension 3. In this sense the set of bicharacteristics that have to be excluded is small.

In case the elastic tensor has symmetries it is determined by less than 21 coefficients. The characteristic sets of such media are analyzed by Musgrave [24]. In this case the singularities can be of different types. For instance, in some classes of media, such as transversely isotropic media, the determinant factors into smooth

factors. In that case the multiplicities of the eigenvalues $A_M^{\text{prin}}(x,\xi)$ can vary on a larger (codimension 2) subset of $T^*X \setminus 0$. Since the bicharacteristics are curves on a codimension 1 surface, Assumption 1 can be violated on an open set of bicharacteristics.

3 Reflection at an Interface: Microlocal Analysis of the "Kirchhoff" Approximation

A particular 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 such layers with smoothly varying medium parameters.

The amplitude of the scattered waves is determined essentially by the reflection coefficients and, implicitly, by the curvature of the interface. It is well known how to calculate these for two constant-coefficient media and a plane interface (see, e.g., Aki and Richards [1, chap. 5]). In the case of smoothly varying media, they determine the scattering in the high-frequency limit; see Taylor [29] for a treatment of reflection and transmission of waves using microlocal analysis. For the acoustic case, see also Hansen [13].

Mathematically, the reflection and transmission of waves is formulated as a boundary value problem. The displacement u_l must satisfy the partial differential equation and initial conditions. In addition, the displacement and the normal traction have to be continuous at the interface. Denote by ν the normal to the interface. The following equations must hold:

$$\begin{aligned} P_{il}u_l &= f_i \quad \text{away from the interface,} \\ u_l &= 0 \quad \text{for } t < 0\,, \end{aligned} \tag{3.1}$$

while

$$\begin{aligned} \rho^{-1/2}u_l &\quad \text{is continuous at the interface,} \\ \nu_j c_{ijkl}\frac{\partial}{\partial x_k}(\rho^{-1/2}u_l) &\quad \text{is continuous at the interface.} \end{aligned} \tag{3.2}$$

Here, we have the factors ρ because of our normalization (2.3). We assume that the source vanishes on a neighborhood of the interface. That this is a well-posed problem can be shown using energy estimates (see, e.g., Lions and Magenes [22, sec. 3.8].

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

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

This is the bicharacteristic associated with the $M,\pm$ constituent of the solution; see Section 2. We define a bicharacteristic to be incoming if its direction is from inside

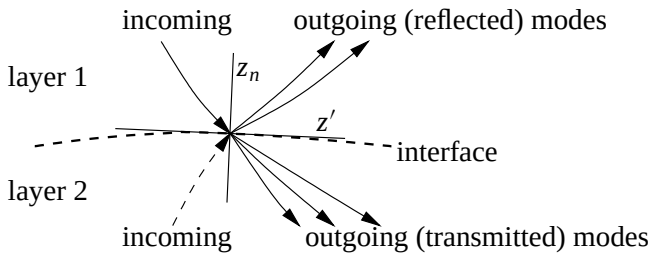


FIGURE 3.1. Incoming and outgoing rays.

a layer towards the interface for increasing time. We define a bicharacteristic to be outgoing if its direction is away from the interface into a layer for increasing time; see Figure 3.1.

Assume that the incoming bicharacteristic stays inside a layer from $t = 0$ until it hits the interface; then the solution along such a bicharacteristic is determined completely by the partial differential equation and the initial condition. On the other hand, the solution along the outgoing bicharacteristics is not determined by the partial differential equation and the initial condition. We will show that the solution along the outgoing bicharacteristics is determined by the partial differential equation and the interface conditions in (3.2).

Let us consider the consequences of the interface conditions. Assume for the moment that the interface is located at $x_n = 0$. We denote $x' = (x_1, x_2, \ldots, x_{n-1})$, $x = (x', x_n)$, and similarly for ξ. The wavefront set of the restriction of u_l to $x_n = 0$ satisfies

$$\mathrm{WF}(u_l|_{x_n=0}) = \left\{(x', t, \xi', \tau) : \text{there is } \xi_n \text{ such that } (x', 0, t, \xi', \xi_n, \tau) \in \mathrm{WF}(u_l)\right\}.$$

It follows that a solution traveling along a bicharacteristic that intersects the interface at some point $(x', 0, t)$ interacts with any other such solution as long as the associated values for ξ' and τ in their wavefront sets coincide (Snell's law). This is depicted in Figure 3.2.

Depending on the interface coordinate x' and the interface tangential part of the slowness $-\tau^{-1}\xi'$, the number of interacting bicharacteristics may vary. For large values of $-\tau^{-1}\xi'$ there will be neither incoming nor outgoing modes; for small values there are n incoming and n outgoing modes. The situation where the vertical line in Figure 3.2 is tangent to the slowness surface corresponds to rays tangent to the interface. Such rays are associated with head waves and are not treated in our analysis. Equation (2.9) implies that the incoming and the outgoing modes correspond to the real solutions ξ_n of

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

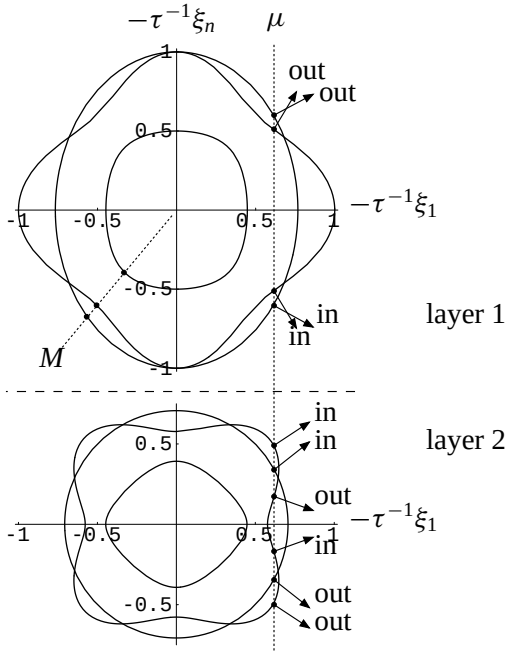


FIGURE 3.2. Two-dimensional section of an $n = 3$-dimensional slowness surface at some point of the interface for the medium on each side of the interface. The slownesses of the modes that interact (i.e., reflect and transmit into each other) are the intersection points with a line that is parallel to the normal of the interface. The group velocity, which is normal to the slowness surface, determines whether the mode is incoming or outgoing.

This equation has $2n$ real or complex-conjugated roots. The complex roots correspond to "evanescent" wave constituents. To number the roots we use an index μ.

In the following theorem we show that if none of the rays involved is tangent, there exists a pseudodifferential operator-type relation between the different modes restricted to the surface $x_n = 0$; we calculate its 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, denoted by ζ, satisfies $\zeta_i(\xi) = (\frac{\partial z}{\partial x})^{-1\mathsf{T}}_{ij}\xi_j$.

Assumption 2. There are no rays tangent to the interface $z_n = 0$ microlocally at (z', t, ζ', τ).

THEOREM 3.1 *Suppose the roots* τ *of* (2.9) *have constant multiplicity and Assumption* 2 *is valid microlocally on some neighborhood in* $T^*(Z' \times \mathbb{R}) \setminus 0$. *Let* $u^{\mathrm{in}}_{N(\nu)}$ *be microlocal constituents of a solution describing the "incoming" modes, and suppose* $G_{M(\mu)}$ *refers to an "outgoing" Green's function* (2.19). *Microlocally,*

the single reflected/transmitted constituent of the solution is given by

$$u_{M(\mu)}(x,t) = \int_{z_n=0} G_{M(\mu)}(x, x(z), t-t_0) \times 2\mathrm{i}D_{t_0}\big(R_{\mu\nu}(z, D_{z'}, D_{t_0})u^{\mathrm{in}}_{N(\nu)}(x(z), t_0)\big)\left|\det\frac{\partial x}{\partial z}\right|\left\|\frac{\partial z_n}{\partial x}\right\| \mathrm{d}z'\,\mathrm{d}t_0 \tag{3.3}$$

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 $R^{\mathrm{prin}}_{\mu\nu}(z, \zeta', \tau)$; see Remark 3.2 below. The integral $|\det\frac{\partial x}{\partial z}|\|\frac{\partial z_n}{\partial x}\|\mathrm{d}z'$ is the Euclidean surface integral over the surface $z_n = 0$.

PROOF: For the moment we assume $z(x) = x$, i.e., that we have a reflector at $x_n = 0$ and smooth coefficients on either side. We show that at the interface there is a relation of the type

$$u^{\mathrm{out}}_{M(\mu)}(x', 0, t) = R^0_{\mu\nu}(x', 0, D', D_t)u^{\mathrm{in}}_{N(\nu)}\,. \tag{3.4}$$

We will use the notation $c_{jk;il} = c_{ijkl}$ and also $(c_{jk})_{il} = c_{ijkl}$. The partial differential equation (2.1) 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 as a first-order system in the variable x_n for the vector V_a of length $2n$ that contains both the displacement and the normal traction (normal to the surface $x_n = \text{const}$)

$$V_a = \begin{pmatrix} \rho^{-1/2}u_i \\ c_{nk;il}\frac{\partial(\rho^{-1/2}u_l)}{\partial x_k}\end{pmatrix}, \qquad i = 1, 2, \ldots, n\,, \tag{3.5}$$

in preparation for the boundary value problem (3.1)–(3.2). Here, a is an index in $\{1, 2, \ldots, 2n\}$. The first-order system then is

$$\frac{\partial V_a}{\partial x_n} = \mathrm{i}C_{ab}(x, D', D_t)V_b\,,$$

where $C_{ab}(x, D', D_t)$ is a matrix partial differential operator given to highest order by

$$-\mathrm{i}\begin{pmatrix} -\sum_{q=1}^{n-1}\sum_{j=1}^{n}(c_{nn})^{-1}_{ij}c_{nq;jl}\frac{\partial}{\partial x_q} & (c_{nn})^{-1}_{il} \\ -\sum_{p,q=1}^{n-1}b_{pq;il}\frac{\partial^2}{\partial x_p\partial x_q} + \rho\delta_{il}\frac{\partial^2}{\partial t^2} & -\sum_{p=1}^{n-1}\frac{\partial}{\partial x_p}c_{pn;ij}(c_{nn})^{-1}_{jl}\end{pmatrix}_{ab}.$$

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 and q are $1, 2, \ldots, n-1$, while still $j \in \{1, 2, \ldots, n\}$).

The next step is to decouple this first-order system microlocally, as in Section 2.1. This means that we want to find 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 principal symbols $C^{\mathrm{prin}}_\mu(x, \xi', \tau)$ are the solutions for ξ_n of

$$\det P^{\mathrm{prin}}_{il}(x, (\xi', \xi_n), \tau) = 0\,. \tag{3.6}$$

In fact, it suffices for the transformed operator (the matrix $\mathrm{diag}(C_\mu(x, D', D_t))_{\mu\nu}$) to be block-diagonal with a block for each different real root of (3.6), a block with eigenvalues with positive imaginary part, and a block with eigenvalues with negative imaginary part. This has also been discussed by Taylor [29]. Under the assumptions of the lemma, this situation can be obtained, since when varying ξ' and τ, the multiplicity of a real eigenvalue only changes when the multiplicity of the corresponding root of (2.9) changes or when two real eigenvalues become complex.

The number of complex eigenvalues with positive or negative imaginary part changes only when two real eigenvalues become complex or vice versa. The latter case occurs only when there are tangent rays and hence is excluded. The $2n \times 2n$ principal symbol $L^{\mathrm{prin}}_{a\mu}(x, \xi', \tau)$ (the columns appropriately normalized) is given by

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

(The polarization vector $Q_{iM}(x, \xi)$ can also be defined for complex ξ). We define $V_\mu = L(x, D', D_t)^{-1}_{\mu a} V_a$. (The index mapping $\mu \mapsto M(\mu)$ assigns the appropriate mode to the normal component of the wave vector).

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 grows in the n-direction, a backward parabolic equation, or one that decays, a forward parabolic equation. The growing mode has to be absent; see, for instance, Hörmander [18, sect. 20.1].

The matrix $L_{a\mu}$ is fixed up to normalization of its columns. For the elliptic modes ($\mathrm{Im}\, C^{\mathrm{prin}}_\mu(x, \xi', \tau) \neq 0$) the normalization is unimportant. For the hyperbolic modes the normalization can be such that the vector $V_\mu = L(x, D', D_t)^{-1}_{\mu a} V_a$ agrees microlocally with the corresponding mode $u_{M,\pm}$ defined in Section 2. To see this, assume V_μ refers to the same mode as $u_{M,\pm}$. In that case there is an invertible pseudodifferential operator $\psi(x, D, D_t)$ of order 0 such that $V_\mu = \psi u_{M,\pm}$. Now we can define $V_{\mu,\mathrm{new}} = \psi^{-1} V_{\mu,\mathrm{old}}$. Because ψ may depend on ξ_n, this factor

cannot be directly 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.

In this proof let the in-modes be the modes for which the amplitude is known, that is, the incoming hyperbolic and the growing elliptic modes. Denote by $L^{(1)}_{a\mu}$ and $L^{(2)}_{a\mu}$ the matrix $L_{a\mu}$ on either side of the interface. We define the $2n \times 2n$ matrix L^{in} such that it contains the columns related to incoming modes out of $L^{(1)}_{a\mu}$ and $L^{(2)}_{a\mu}$, i.e.,

$$L^{\mathrm{in}}_{a\mu} = \left(L^{(1),\mathrm{in}} \quad -L^{(2),\mathrm{in}}\right)_{a\mu},$$

and define $L^{\mathrm{out}}_{a\mu}$ similarly (so, here, μ is slightly different). The interface conditions (3.2) now read

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

If we set $R^0_{\mu\nu} = -(L^{\mathrm{out}})^{-1}_{\mu a} L^{\mathrm{in}}_{a\nu}$ (for the question of whether the inverse exists, see the remark after the proof), then the part referring to the hyperbolic modes gives (3.4).

By (3.4) the u^{out}_M are determined at the interface; finding how they propagate away from the interface 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 generated by the bicharacteristics. It follows that we can use $\phi_{M,\pm}(x, t - t_0, x_0, \xi_J, \tau)$ as a phase function (taking care that $n \notin J$). The amplitude $\mathcal{A}_{M,\pm}(x_I, x_0, \xi_J, \tau)$ satisfies the transport equation as before. However, the restriction of the Fourier integral operator to the "initial surface" $x_n = 0$ so constructed is a pseudodifferential operator that is not necessarily the identity. Let us assume

$$u^{\mathrm{out}}_M(x,t) = \int G_{M,\pm}(x, (x'_0, 0), t - t_0)\psi(x, D_{x'_0}, D_{t_0}) u^{\mathrm{out}}_M(x'_0, 0, t_0)\mathrm{d}x'_0\,\mathrm{d}t_0\,, \tag{3.7}$$

where $\psi(x, D', D_t)$ is to be found such that the restriction of this representation to $x_n = 0$ is the identity. The $\pm$ sign is chosen such that $G_{M,\pm}$ is the outgoing mode. We can use again section 8.6 of Treves [31] to find that the principal symbol of this pseudodifferential operator should be

$$\psi(x, \xi', \tau) = \left|\frac{\partial B_M}{\partial \xi_n}(x, \xi', C^{\mathrm{prin}}_{\mu}(\xi', \tau))\right| = \left|\frac{\partial x_{M,n}}{\partial t}(x, \xi', C^{\mathrm{prin}}_{\mu}(\xi', \tau), 0)\right|; \tag{3.8}$$

i.e., the normal component of the velocity of the ray, the group velocity. We now replace $G_{M,\pm}$ by (the relevant part of) G_M, using that $G_M = \frac{1}{2}\mathrm{i}G_{M,+}B_M(x, D)^{-1} - \frac{1}{2}\mathrm{i}G_{M,-}B_M(x, D)^{-1}$. Taking this into account, and the fact that $B^{\mathrm{prin}}_M(x, \xi) = \mp\tau$, we have now obtained (3.3) for the case that $z = x$ (no coordinate transformation).

We argue that (3.3) is also true when $z(x)$ is a general coordinate transformation. This follows from transforming the equations (3.1) and (3.2) to z-coordinates.

To highest order, the symbol of (pseudo)differential operators transforms as

$$\psi^{\text{transf}}(z,\zeta,\tau) = \psi(x(z), \left(\frac{\partial z}{\partial x}\right)^{\mathsf{T}} \zeta, \tau)\,.$$

Tracing the steps of the proof, we find the following equivalent of (3.4)

$$u^{\text{out}}_{M(\mu)}(x(z',0),t) = R^0_{\mu\nu}(z',0,D_{z'},D_t)u^{\text{in}}_{N(\nu)}(x(z',0),t)\,. \tag{3.9}$$

When the interface is at $z_n = 0$, we can obtain (3.7) in z-coordinates instead of x-coordinates. Transforming G_M and u_M back to x-coordinates, we find that for x away from the interface,

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

Here $|\frac{\partial z_{M,n}}{\partial t}(z,D_{z'},D_t)|$ is the transformed version of (3.8). Thus expression (3.3) follows, with

$$R_{\mu\nu}(z,\zeta',\tau) = \left|\frac{\partial z_{M,n}}{\partial t}(z,\zeta',\tau)\right| \left\|\frac{\partial z_n}{\partial x}\right\|^{-1} R^0_{\mu\nu}(z,\zeta',\tau)\,.$$

□

Remark 3.2. The principal symbol $R^{0,\text{prin}}_{\mu\nu}(z,\zeta',\tau)$ that occurs in the proof is simply the reflection coefficient for the amplitudes. The principal symbol $R^{\text{prin}}_{\mu\nu}(z,\zeta',\tau)$ is obtained by multiplying $R^{0,\text{prin}}_{\mu\nu}$ with the normal component of the velocity of the ray, given (for $z(x) = x$) by (3.8). The reflection coefficients satisfy unitary relations; see Chapman [8] and Kennett [20] (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 of $L^{\text{out}}_{a\mu}$ exists. Chapman [8] also gives a direct proof of the reciprocity relations for the reflection coefficients.

Remark 3.3. We have shown that the reflected/transmitted wave 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 domains separated by smooth interfaces), this is also the case (cf. Frazer and Sen [11]). 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. The integration over z' accounts for the effects associated with the interface's curvature.

4 The Born Approximation

We discuss the modeling and inversion of seismic data in the Born approximation. The medium parameters are written as the sum of a 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 with an allowance for multipathing (caustics), see Hansen [13] and Ten Kroode et al. [21]. For the acoustic problem with nonmaximal acquisition geometry, see Nolan and Symes [25]. For the elastic case with maximal acquisition geometry (and from a more applied point of view), see De Hoop and Brandsberg-Dahl [15]. We extend their results and give an efficient, novel presentation. Also, we discuss in detail the different assumptions that are needed for the modeling and inversion of seismic data.

4.1 Modeling: Perturbation of the Green's Function

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

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

This decomposition induces a perturbation of P_{il} (cf. (2.4)),

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

We denote the causal Green's operator associated with (2.2) 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 $(P_{ij} + \delta P_{ij})(G_{jk} + \delta G_{jk})$ vanish. This results in the following expression for $\delta G_{il}(\hat{x}, \tilde{x}, t)$:

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

Here, $\tilde{x}$ denotes a source location, $\hat{x}$ a receiver location, and x_0 a scattering point. Because the background model is smooth, the operator δG_{il} contains only the single scattered field.

We use the decoupled equations (2.7). Omitting the factors $Q_{iM}(\hat{x}, D_{\hat{x}})$, $Q(\tilde{x}, D_{\tilde{x}})^{-1}_{Nl}$ at the beginning and end of the product, we obtain an expression for the

perturbation of the Green's function $\delta G_{MN}(\hat{x}, \tilde{x}, t)$ for the pair of modes M (scattered) and N (incident)

$$\delta G_{MN}(\hat{x}, \tilde{x}, t) = -\int_0^t \int_X G_M(\hat{x}, x_0, t - t_0) Q(x_0, D_{x_0})^{-1}_{Mi} \left(\delta_{il} \frac{\delta\rho}{\rho} \frac{\partial^2}{\partial t_0^2} - \frac{\partial}{\partial x_{0,j}} \frac{\delta c_{ijkl}}{\rho} \frac{\partial}{\partial x_{0,k}} \right) Q(x_0, D_{x_0})_{lN} G_N(x_0, \tilde{x}, t_0) \mathrm{d}x_0 \, \mathrm{d}t_0 .$$

Microlocally we can write G_M as in (2.24), with appropriate substitutions for its arguments. 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 to highest order equals the product $\mathcal{A}_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) Q(x_0, \hat{\xi}_0)^{-1}_{Mi} \mathrm{i}\hat{\xi}_{0,j}$, where $\hat{\xi}_0 = \xi_0(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$. Assuming that the medium perturbation vanishes around $\hat{x}$ and $\tilde{x}$, a cutoff is introduced for t_0 near 0 and t. In the resulting expression, one of the two frequency variables $\hat{\tau}$ and $\tilde{\tau}$ can now be eliminated using the integral over t_0 (see, for instance, Duistermaat [10, sect. 2.3]). In this case the result can be readily obtained by noting that the integral over t_0 can be extended to the whole of $\mathbb{R}$ (the phase is not stationary for t_0 outside $[0, t]$), and then using that $\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}t_0(\hat{\tau}-\tilde{\tau})} \mathrm{d}t_0 = 2\pi\delta(\hat{\tau} - \tilde{\tau})$. The resulting formula for δG_{MN} is, modulo lower-order terms in the amplitude,

$$\delta G_{MN}(\hat{x}, \tilde{x}, t) = (2\pi)^{-\frac{3n+1}{4} - \frac{|\hat{J}|+|\tilde{J}|+1}{2}} \int \mathcal{B}_{MN}(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, x_0, \tau)$$

$$\times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x}, \tilde{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)} \left(w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta c_{ijkl}(x_0)}{\rho(x_0)} + w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) \frac{\delta\rho(x_0)}{\rho(x_0)} \right) \mathrm{d}x_0 \, \mathrm{d}\hat{\xi}_{\hat{J}} \, \mathrm{d}\tilde{\xi}_{\tilde{J}} \, \mathrm{d}\tau . \tag{4.1}$$

Here (see (2.18) for the construction of ϕ_M, ϕ_N),

$$\Phi_{MN}(\hat{x}, \tilde{x}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = \phi_M(\hat{x}, x_0, t, \hat{\xi}_{\hat{J}}, \tau) + \phi_N(\tilde{x}, x_0, t, \tilde{\xi}_{\tilde{J}}, \tau) - \tau t . \tag{4.2}$$

The amplitude factors $\mathcal{B}_{MN}$ are given by

$$\mathcal{B}_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = (2\pi)^{-\frac{n-1}{4}} \mathcal{A}_M(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau) \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}, \tau) . \tag{4.3}$$

We will refer to the factors $w_{MN;ijkl}$ and $w_{MN;0}$ as the radiation patterns. They are given by

$$w_{MN;ijkl}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = Q_{iM}(x_0, \hat{\xi}_0) Q_{lN}(x_0, \tilde{\xi}_0) \hat{\xi}_{0,j} \tilde{\xi}_{0,k} , \tag{4.4}$$

$$w_{MN;0}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau) = -Q_{iM}(x_0, \hat{\xi}_0) Q_{iN}(x_0, \tilde{\xi}_0) \tau^2 , \tag{4.5}$$

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

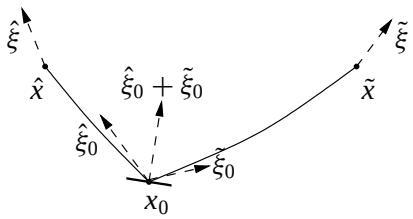


FIGURE 4.1. The scattering cotangent vectors.

We investigate the map $(\delta c_{ijkl}/\rho, \delta\rho/\rho) \mapsto \delta G_{MN}(\hat{x}, \tilde{x}, t)$ induced by (4.1). We use the notation C_{ϕ_M} to indicate the subset of the global canonical relation C_M that is associated to a phase function ϕ_M (cf. (2.15)).

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

$$\begin{aligned}\Lambda_{0,MN} = \big\{ &(\hat{x}, \tilde{x}, \hat{t} + \tilde{t}, \hat{\xi}, \tilde{\xi}, \tau; x_0, \hat{\xi}_0 + \tilde{\xi}_0) : \\ &(\hat{x}, \hat{t}, \hat{\xi}, \tau; x_0, \hat{\xi}_0) \in C_{\phi_M}, (\tilde{x}, \tilde{t}, \tilde{\xi}, \tau; x_0, \tilde{\xi}_0) \in C_{\phi_N} \big\}.\end{aligned} \tag{4.6}$$

PROOF: We show that $\Phi_{MN}(\hat{x}_{\hat{I}}, \tilde{x}_{\tilde{I}}, t, x_0, \hat{\xi}_{\hat{J}}, \tilde{\xi}_{\tilde{J}}, \tau)$ is a nondegenerate phase function. The derivatives with respect to the phase variables are given by

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

where $\hat{x}_{\hat{J}}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tau)$ and $\tilde{x}_{\tilde{J}}(\tilde{x}_{\tilde{I}}, x_0, \tilde{\xi}_{\tilde{J}}\tau)$ are as defined in (2.17), for the receiver side and the source side, respectively. 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 expression (4.2) it follows that the canonical relation of this operator is given by (4.6). By the assumption it contains no elements with $\hat{\xi}_0 + \tilde{\xi}_0 = 0$, so it is continuous as a map $\mathcal{E}'(X) \to \mathcal{D}'(X \times X \times]0, T[)$. □

We show that the condition in Lemma 4.1 is violated if and only if $M = N$ and there exists 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 eigenvalues in (2.8) are different for different modes. If $M = N$ and $\hat{\xi}_0 = -\tilde{\xi}_0$, then we have the mentioned direct bicharacteristic.

4.2 Restriction: Acquisition

The data are assumed to be representable 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 (the 2 enforces "remote sensing"). Then the data are given by the map

$$y' \mapsto \delta G_{MN}(\hat{x}(y', 0), \tilde{x}(y', 0), t(y', 0)) \,. \tag{4.7}$$

It follows that the map $(\delta c_{ijkl}/\rho, \delta\rho/\rho)$ to the data may be seen as the composition 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 Fourier integral operator 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 $T^*Y' \setminus 0 \times \mathrm{diag}(T^*Y \setminus 0) \times T^*X \setminus 0$ 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 $(y', 0, \eta', \eta'') \in T^*Y \setminus 0$ 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 arrival time $t(y', 0)$.

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

$$\frac{\partial y''}{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})} \quad \text{has maximal rank.} \tag{4.8}$$

In the following theorem we parameterize (4.6) by $(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})$ using the parameterization of C_{ϕ_M} given by (2.15). Thus we let $\tau = \mp B_M(x_0, \hat{\xi}_0)$ and

$$\hat{x} = x_M(x_0, \hat{\xi}_0, \pm\hat{t}) \,, \qquad \tilde{x} = x_N(x_0, \tilde{\xi}_0, \pm\tilde{t}) \,,$$
$$\hat{\xi} = \xi_M(x_0, \hat{\xi}_0, \pm\hat{t}) \,, \qquad \tilde{\xi} = \xi_N(x_0, \tilde{\xi}_0, \pm\tilde{t}) \,.$$

We suppose that $(y'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}), \eta'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}))$ is obtained by transforming $(\hat{x}, \tilde{x}, \hat{t} + \tilde{t}, \hat{\xi}, \tilde{\xi}, \tau)$ to (y, η)-coordinates.

THEOREM 4.2 *If Assumptions* 3 *and* 4 *are satisfied, then the operator* $F_{MN;ijkl}$ (*respectively,* $F_{MN;0}$) *that maps the medium perturbation* $\delta c_{ijkl}/\rho$ (*respectively,* $\delta\rho/\rho$)

to the data as a function of y' (4.7) *is microlocally a Fourier integral operator with canonical relation given by*

$$\Lambda_{MN} = \big\{(y'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}), \eta'(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}); x_0, \hat{\xi}_0 + \tilde{\xi}_0) : \\ B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau, y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0\big\}. \tag{4.9}$$

The order equals $(n-1+c)/4$. *The amplitude is given to highest order* (*in coordinates* (y'_I, η'_J, x_0) *for* Λ_{MN}, *where* I, J *is a partition of* $\{1, 2, \ldots, 2n-1-c\}$) *by the product* $\mathcal{B}_{MN}(y'_I, \eta'_J, x_0) w_{MN;ijkl}(y'_I, \eta'_J, x_0)$ (*respectively,* $\mathcal{B}_{MN}(y'_I, \eta'_J, x_0)$ $w_{MN;0}(y'_I, \eta'_J, x_0)$), *where*

$$|\mathcal{B}_{MN}(y'_I, \eta'_J, x_0)| = \\ \frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n+1+c}{4}} \left|\det \frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-\frac{1}{2}} \left|\det \frac{\partial(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x_0, y'_I, y'', \eta'_J, \Delta\tau)}\right|^{\frac{1}{2}}_{\Delta\tau=0\,.y''=0}. \tag{4.10}$$

Here we define $\Delta\tau = \hat{\tau} - \tilde{\tau}$ *so that the first constraint in* (4.9) *reads* $\Delta\tau = 0$. *The map* $(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) \mapsto (x_0, y'_I, y'', \eta'_J, \Delta\tau)$ *is bijective.*

PROOF: The first statement has been argued above. The order of the operator is given by

$$\chi + \frac{K}{2} - \frac{\dim X + \dim Y'}{4},$$

where χ is the degree of homogeneity of the amplitude and K is the number of phase variables. The factors $\{w_{MN;ijkl}, w_{MN;0}\}$ are homogeneous of order 2 in the ξ- and τ-variables; the degree of homogeneity of the factor $\mathcal{B}_{MN}$ follows from (2.22). We find

$$\begin{aligned}\text{order } F_{MN;ijkl} &= 2 + \left(-2 - \frac{|\hat{J}| + |\tilde{J}| + 2}{2} + n\right) \\ &\quad + \frac{|\hat{J}| + |\tilde{J}| + 1}{2} - \frac{3n-1-c}{4} \\ &= \frac{n-1+c}{4}.\end{aligned}$$

We now calculate the amplitude of the Fourier integral operator in Lemma 4.1. The factor $w_{MN;ijkl}$ is simply multiplicative. Suppose we choose coordinates on $\Lambda_{0,MN}$ to be $(\hat{x}_{\hat{I}}, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \hat{\tau}, \tilde{\tau}, x_0)$, with ultimately $\hat{\tau} = \tilde{\tau}$. Define $\tau = \frac{\hat{\tau}+\tilde{\tau}}{2}$, $\Delta\tau = \hat{\tau} - \tilde{\tau}$. Using (2.25) and (4.3), we find that the amplitude $\mathcal{B}_{MN}(\hat{x}_{\hat{I}}, x_0, \hat{\xi}_{\hat{J}}, \tilde{x}_{\tilde{I}}, \tilde{\xi}_{\tilde{J}}, \tau)$ is given by

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

The transformation from $(\hat{x}, \tilde{x}, t)$- to y-coordinates in the Fourier integral (4.7) induces an additional factor $|\det \frac{\partial(\hat{x},\tilde{x},t)}{\partial y}|^{-1/2}$ (note that for the Fourier integral operators, it would be more natural to transform as a half-density). The amplitude transforms as a half-density on the canonical relation, and we obtain the factor

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

The additional factor $(2\pi)^{-(2+c)/4}$ arises from the normalization. We find (4.10). □

Natural coordinates for the canonical relation are given 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 a natural density directly associated with this set, the quotient density. The Jacobian in (4.10) reveals that the amplitude factor $|\mathcal{B}_{MN}(y'_I, \eta'_J, x_0)|$ is in fact given by the associated half-density times

$$\frac{1}{4}\tau^{-2}(2\pi)^{-\frac{n+1+c}{4}}\left|\frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-1/2}.$$

If $c = 0$ and there are no rays tangent to the acquisition manifold, that is,

$$\operatorname{rank} \frac{\partial y''}{\partial(\hat{t}, \tilde{t})} = 2\,, \tag{4.11}$$

then a convenient way to parameterize the canonical relation is by using the phase directions $\hat{\alpha} = \hat{\xi}_0/\|\hat{\xi}_0\|, \tilde{\alpha} = \tilde{\xi}_0/\|\tilde{\xi}_0\| \in \mathbb{S}^{n-1}$ and the frequency τ.

4.3 Inversion

Let us now consider the reconstruction of $((\delta c_{ijkl})/\rho, (\delta\rho)/\rho)$ from the data. We simplify the notation and collect the medium perturbations into

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

The forward operator $(F_{MN;ijkl}, F_{MN;0})$ in the Born approximation is represented by $F_{MN;\alpha}$.

Let us consider data from a single 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 this inverse 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 $F^*_{MN;\alpha}$,

$$N_{MN;\alpha\beta} = F^*_{MN;\alpha}F_{MN;\beta} \tag{4.12}$$

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

$$F^{-1}_{MN;\alpha} = (N_{MN})^{-1}_{\alpha\beta}F^*_{MN;\beta} \tag{4.13}$$

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

The properties of the composition (4.12) depend on those of Λ_{MN}. Let $\pi_{Y'}$ and π_X be the projection mappings of Λ_{MN} to $T^*Y' \setminus 0$ and $T^*X \setminus 0$, respectively. 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 $(x, \xi) \in \pi_X(\Lambda_{MN})$.

Assumption 5. The projection $\pi_{Y'}$ of Λ_{MN} on $T^*Y' \setminus 0$ is an embedding; i.e., it is

(i) immersive,
(ii) injective, and
(iii) proper.

This assumption appears in Guillemin [12], and is sometimes referred to as the Bolker condition. It implies that the image of $\pi_{Y'}$ is a submanifold, $\mathcal{L}$ say, of $T^*Y' \setminus 0$. Let us discuss these requirements, starting with the first. Using that Λ_{MN} is a canonical relation, we have the following:

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

PROOF: This is a property of Lagrangian manifolds. It follows from lemma 25.3.6 in Hörmander [18]. We give an independent proof.

The symplectic forms σ_X and $\sigma_{Y'}$ on $T^*X \setminus 0$ and $T^*Y' \setminus 0$, respectively, 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 $n + m$, assuming $\dim T^*Y' \setminus 0 = 2m$ (in this proof $m = \dim Y' = 2n - 1 - c$). 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

$$\{(v_1, w_1), (v_2, w_2), \ldots, (v_{2n}, w_{2n}), (0, w_{2n+1}), (0, w_{2n+2}), \ldots, (0, w_{n+m})\}.$$

The w_i, $i \in \{1, 2, \ldots, 2n\}$, must be linearly independent because $\operatorname{rank} \sigma_{Y'} = 2n$. For w_i, $i \leq 2n$, and w_j, $j > 2n$, we have $\sigma_{Y'}(w_i, w_j) = 0$, so the w_j are linearly independent of the w_i. The w_i, $i > 2n$, must be linearly independent, because $(0, w_i)$ are basis vectors 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 co-isotropic submanifold. □

[1] Equation (4.12) is for the case where one minimizes the difference with the data δG_{MN} in L^2 norm $\|\delta G_{MN} - F_{MN;\alpha} g_\alpha\|$. It can easily be adapted to the case where one minimizes a Sobolev norm of different order or a weighted L^2 norm. This would introduce extra factors in the amplitude.

As a consequence, if part (i) of Assumption 5 is satisfied, then we can use $(x, \xi) \in T^*X \setminus 0$ as (local) coordinates on Λ_{MN}. In addition, we need to parameterize the subsets of the canonical relation given by $(x, \xi) = \text{const}$; we denote such local coordinates by e. The new parameterization of Λ_{MN} is (identifying x_0 with x)

$$\Lambda_{MN} = \{(y'(x, \xi, e), \eta'(x, \xi, e); (x, \xi))\}\,. \tag{4.14}$$

The results do not depend on the precise definition of e. As noted before, if the variables $(\hat{t}, \tilde{t})$ can be solved from the second constraint in (4.9) (cf. equation (4.11)), then Λ_{MN} can be parameterized using $(x, \hat{\alpha}, \tilde{\alpha}, \tau)$, where $(\hat{\alpha}, \tilde{\alpha})$ are phase directions. In that case (x, ξ, e) should be related by a coordinate transformation to $(x, \hat{\alpha}, \tilde{\alpha}, \tau)$. In acoustic media (where $\|\hat{\xi}_0\| = \|\tilde{\xi}_0\|$) a suitable choice is the pair scattering angle/azimuth given by

$$\left(\arccos(\hat{\alpha} \cdot \tilde{\alpha}),\ \frac{-\hat{\alpha} + \tilde{\alpha}}{2\sin(\arccos(\hat{\alpha} \cdot \tilde{\alpha})/2)}\right) \in]0, \pi[\times \mathbb{S}^{n-2}$$

(cf. Burridge and Beylkin [7]). The azimuth, the second component, defines together with ξ the plane spanned by $(\hat{\alpha}, \tilde{\alpha})$. It is not very difficult to show that in elastic media the scattering angle (the first component) can be used as coordinate when the slowness sheets are convex, but not always when one of the slowness sheets fails to be convex.

Remark 4.4. We show that the first part of Assumption 5 implies that $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) + \frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0) \neq 0$; in other words, the group velocities at the scattering point do not add up to 0. We have seen in Theorem 4.2 that Λ_{MN} may be parameterized by $(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})$, where $(\hat{\xi}_0, \tilde{\xi}_0)$ are such that

$$B_M(x_0, \hat{\xi}_0) = B_N(x_0, \tilde{\xi}_0) = \pm\tau$$

(and we have the additional constraint $y''(x_0, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t}) = 0$). The projection π_X is given by $(x, \hat{\xi}_0 + \tilde{\xi}_0)$. Consider tangent vectors to Λ_{MN} given by vectors $v_{\hat{\xi}_0}$ and $v_{\tilde{\xi}_0}$. They must satisfy

$$v_{\hat{\xi}_0} \cdot \frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = v_{\tilde{\xi}_0} \cdot \frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0) = \pm v_\tau\,. \tag{4.15}$$

So if $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$, then (4.15) implies that $(v_{\hat{\xi}_0} + v_{\tilde{\xi}_0}) \cdot \frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) = 0$, so that the projection of Λ_{MN} on $T^*X \setminus 0$ is not submersive. If $c = 0$ and $\operatorname{rank} \partial y''/\partial(\hat{t}, \tilde{t}) = 2$ (no tangent rays), then the constraint $y'' = 0$ may be used to solve for the parameters $\hat{t}$ and $\tilde{t}$, and (4.15) is the only condition on $(\hat{\xi}_0, \tilde{\xi}_0)$. In that case $\frac{\partial B_M}{\partial \xi}(x, \hat{\xi}_0) \neq -\frac{\partial B_N}{\partial \xi}(x, \tilde{\xi}_0)$ implies that the projection is submersive. In other cases the set of $(\hat{\xi}_0, \tilde{\xi}_0)$ is in general a smaller subset of $T^*_x X \setminus 0 \times T^*_x X \setminus 0$.

Let us now discuss the second and third parts of Assumption 5. The second part is a well-known condition; see Hansen [13] and Ten Kroode et al. [21]. Essentially the condition is that there are no two different singularities in g_α mapped to the

same position in $T^*Y' \setminus 0$. For an analysis of the case where this condition is violated, see Stolk [27].

The definition of *proper* is that the preimage of a compact set is a compact set. So assume we have a compact subset of $T^*Y' \setminus 0$. The elements of Λ_{MN} correspond to those "points" where the source and receiver rays intersect. The set of these points can be written as a set on which 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 composition (4.12) 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 4.5 *Let $\psi(y', D_{y'})$ be a pseudodifferential cutoff with conically compact support in $T^*Y' \setminus 0$ such that for the set*

$$\{(y', \eta'; x_0, \xi_0) \in \Lambda_{MN} : (y', \eta') \in \operatorname{supp} \psi\}. \tag{4.16}$$

Assumptions 3, 4, *and* 5 *are satisfied. Then*

$$F^*_{MN;\beta}\psi(y', D_{y'})^*\psi(y', D_{y'})F_{MN;\alpha} \tag{4.17}$$

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

$$\begin{aligned} N_{MN;\beta\alpha}(x,\xi) = \frac{1}{16}(2\pi)^{-n} \int |\psi(y'(x,\xi,e), \eta'(x,\xi,e))|^2 \tau^{-4} \\ \times \overline{w_{MN;\beta}(x,\xi,e)} w_{MN;\alpha}(x,\xi,e) \left|\det \frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y}\right|^{-1} \\ \times \left|\det \frac{\partial(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x, \xi, e, y'', \Delta\tau)}\right|_{\Delta\tau=0, y''=0} \mathrm{d}e\,, \end{aligned} \tag{4.18}$$

where $\tau = \tau(x, \xi, e)$.

PROOF: We use the clean intersection calculus for Fourier integral operators (see, e.g., Treves [31]) to show that (4.17) is a Fourier integral operator. The canonical relation of $F^*_{MN;\alpha}$ is given by

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

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

$$L \cap M \text{ is a manifold,} \tag{4.19}$$

$$TL \cap TM = T(L \cap M). \tag{4.20}$$

It follows from Assumption 5(ii) that $L \cap M$ must be given by

$$L \cap M = \{(x, \xi, y', \eta', y', \eta', x, \xi) : (y', \eta'; x, \xi) \in \Lambda_{MN}\}. \tag{4.21}$$

Because Λ_{MN} is a manifold, this set satisfies (4.19). The property (4.20) follows from the assumption that the map $\pi_{Y'}$ is immersive. The excess is given by

$$\begin{aligned} E &= \dim T^*X \setminus 0 \times T^*Y' \setminus 0 \times T^*Y' \setminus 0 \times T^*X \setminus 0 + \dim(L \cap M) \\ &\quad - \dim L - \dim M = n - 1 - c\,. \end{aligned} \tag{4.22}$$

Taking into account the pseudodifferential cutoff $\psi(y', D_{y'})$, it follows that (4.17) is a Fourier integral operator. The canonical relation

$$\Lambda_{MN}^* \circ \Lambda_{MN} \quad \text{of } F_{MN;\beta}^* \psi^* \psi F_{MN;\alpha}$$

is contained in the diagonal of $T^*X \setminus 0 \times T^*X \setminus 0$, so it is a pseudodifferential operator. The order is given by $E/2 + 2\,\text{order}(F_{MN;\alpha}) = n - 1$ (note that c drops out).

We write $\psi(y', D_{y'})^* \psi(y', D_{y'}) = \sum_i \chi^{(i)}(y', D_{y'})$, where the symbols $\chi^{(i)}(y', \eta')$ have small enough support, so that the distribution kernel of $\chi^{(i)}(y', D_{y'}) F_{MN;\alpha}$ can be written as the oscillatory integral

$$\begin{aligned} \chi^{(i)}(y', D_{y'}) \mathcal{F}_{MN;\alpha}(y', x) &= (2\pi)^{-\frac{3n-1-c}{4} - \frac{|J|}{2}} \int \chi^{(i)}(y'_I, \eta'_J, x) \\ &\quad \times \mathcal{B}_{MN}(y'_I, \eta'_J, x) w_{MN;\alpha}(y'_I, \eta'_J, x) \mathrm{e}^{\mathrm{i}(S_{MN}^{(i)}(y'_I, x, \eta'_J) + \langle \eta'_J, y'_J \rangle)}\, \mathrm{d}\eta'_J\,, \end{aligned} \tag{4.23}$$

where $\psi^{(i)}(y'_I, \eta'_J, x) = \psi^{(i)}(y'_I, y'_J(y'_I, \eta'_J, x), \eta'_I(y'_I, \eta'_J, x), \eta'_J)$. We can write $\Phi_{MN}^{(i)}(y', x, \eta'_J) = S_{MN}^{(i)}(y'_I, x, \eta'_J) + \langle \eta'_J, y'_J \rangle$ (cf. (2.18) and (4.2)). We do not indicate the dependence of J on i explicitly. The distribution kernel of the normal operator is then given by a sum of terms

$$\begin{aligned} &\int (\overline{\psi(y', D_{y'}) \mathcal{F}_{MN;\beta}(y', x)})(\psi(y', D_{y'}) \mathcal{F}_{MN;\alpha}(y', x_0))\mathrm{d}y' \\ &\quad = (2\pi)^{-\frac{3n-1-c}{2} - |J|} \sum_i \int \chi^{(i)}(y'_I, \eta'_{0,J}, x_0) \\ &\qquad \times \overline{\mathcal{B}_{MN}(y'_I, \eta'_J, x)} \mathcal{B}_{MN}(y'_I, \eta'_{0,J}, x_0) \overline{w_{MN;\beta}(y'_I, \eta'_J, x)} w_{MN;\alpha}(y'_I, \eta'_{0,J}, x_0) \\ &\qquad \times \mathrm{e}^{\mathrm{i}(S_{MN}^{(i)}(y'_I, x_0, \eta'_{0,J}) - S_{MN}^{(i)}(y'_I, x, \eta'_J) + \langle \eta'_{0,J}, y'_J \rangle - \langle \eta'_J, y'_J \rangle)}\, \mathrm{d}\eta'_{0,J}\, \mathrm{d}\eta'_J\, \mathrm{d}y'\,. \end{aligned}$$

We now apply the method of stationary phase and integrate out the variables y'_J and $\eta'_{0,J}$. For the remaining variables we use

$$S_{MN}^{(i)}(y'_I, x_0, \eta'_J) - S_{MN}^{(i)}(y'_I, x, \eta'_J) = \langle x - x_0, \xi(y'_I, \eta'_J, x_0) \rangle + O(|x - x_0|^2)\,.$$

Thus we find (to highest order)

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

We now change variables $(x, y'_I, \eta'_J) \to (x, \xi, e)$ and use (4.10). We sum over i and arrive at

$$\begin{aligned} \mathcal{N}_{MN;\beta\alpha}(x, x_0) = \frac{(2\pi)^{-2n}}{16} \int & |\psi(y'(x, \xi, e), \eta'(x, \xi, e))|^2 \tau^{-4} \\ & \times \overline{w_{MN;\beta}(x, \xi, e)} w_{MN;\alpha}(x, \xi, e) \left| \det \frac{\partial(\hat{x}, \tilde{x}, t)}{\partial y} \right|^{-1} \\ & \times \left| \det \frac{\partial(x, \hat{\xi}_0, \tilde{\xi}_0, \hat{t}, \tilde{t})}{\partial(x, \xi, e, y'', \Delta\tau)} \right|_{\Delta\tau=0, y''=0} e^{i\langle x - x_0, \xi\rangle} \, d\xi \, de . \end{aligned} \tag{4.24}$$

It follows that the principal symbol of $N_{MN;\beta\alpha}$ is given by (4.18). □

So far we have focused on inversion of data from one pair of modes (M, N). Often data will be available for some subset S of all possible pairs of modes. Define the normal operator for this case as

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

If all the $N_{MN;\alpha\beta}$ are pseudodifferential operators, then $N_{\alpha\beta}$ is also a pseudodifferential operator. A left inverse is now given by

$$N^{-1}_{\alpha\beta} F^*_\beta ,$$

where F^*_β is the vector of Fourier integral operators containing $F^*_{MN;\beta}$, $(M, N) \in S$.

5 Symplectic Geometry of the Data

In the previous section we saw that the wavefront set of the modeled data cannot 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 employed in the parameter reconstruction and is important in the reconstruction of the background medium (or the medium above the interface in the case of a smooth jump) as well. This will be explained below.

Consider again the canonical relation Λ_{MN}. Suppose Assumption 5 is satisfied. In this section we will denote by Ω the map

$$\Omega : (x, \xi, e) \mapsto (y'(x, \xi, e), \eta'(x, \xi, e)) : T^*X \setminus 0 \times E \to T^*Y' \setminus 0$$

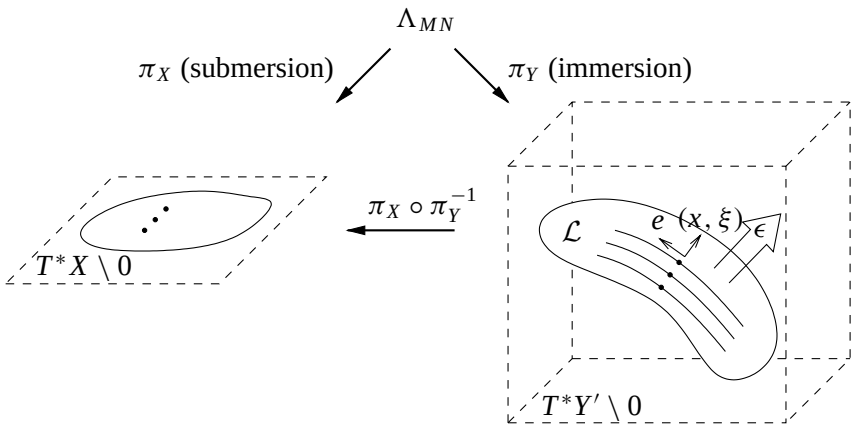


FIGURE 5.1. Visualization of the symplectic structure of Λ_{MN} (cone structure omitted).

introduced above (4.14). This map conserves the symplectic form of $T^*X \setminus 0$. That is, if $w_{x_i} = \partial(y', \eta')/\partial x_i$ and similarly for w_{ξ_i} and w_{e_i}, we have

$$\begin{aligned} &\sigma_{Y'}(w_{x_i}, w_{x_j}) = \sigma_{Y'}(w_{\xi_i}, w_{\xi_j}) = 0\,, \\ &\sigma_{Y'}(w_{\xi_i}, w_{x_j}) = \delta_{ij}\,, \\ &\sigma_{Y'}(w_{e_i}, w_{x_j}) = \sigma_{Y'}(w_{e_i}, w_{\xi_j}) = \sigma_{Y'}(w_{e_i}, w_{e_j}) = 0\,. \end{aligned} \tag{5.1}$$

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

The image $\mathcal{L}$ of the map Ω is co-isotropic, as noted in Lemma 4.3. The sets $(x, \xi) = \text{const}$ are the isotropic fibers of the fibration of Hörmander [17] described in 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 5.1. The wavefront set of the data is contained in $\mathcal{L}$ and 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 $\mathcal{L}$.

LEMMA 5.1 *Let $\mathcal{L}$ be an embedded co-isotropic submanifold of $T^*Y' \setminus 0$ with coordinates (x, ξ, e) such that* (5.1) *holds. Denote $\mathcal{L} \ni (y', \eta') = \Omega(x, \xi, e)$. We can find a homogeneous canonical map G from an open part of $T^*(X \times E) \setminus 0$ to an open neighborhood of $\mathcal{L}$ in $T^*Y' \setminus 0$ such that $G(x, e, \xi, \epsilon = 0) = \Omega(x, \xi, e)$.*

PROOF: The e_i can be viewed as (coordinate) functions on $\mathcal{L}$. We will first extend them to functions on the whole $T^*Y' \setminus 0$ such that the Poisson brackets $\{e_i, e_j\}$ satisfy

$$\{e_i, e_j\} = 0\,, \qquad 1 \le i, j \le m - n\,, \tag{5.2}$$

where $m = \dim Y' = 2n - c - 1$. This can be done successively for $e_1, e_2, \ldots, e_{m-n}$ by the method that we describe now; see the proof of theorem 3.3 in Treves [31,

chap. 7] or the proof of theorem 3.5.6 in Duistermaat [10]. Suppose we have extended $e_1, e_2, \dots, e_l$; we extend e_{l+1}. In order to satisfy (5.2) e_{l+1} has to be a solution u of

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

where H_{e_i} is the Hamilton field associated with the function e_i, with initial condition on some manifold transversal to the H_{e_i}. For any $(y', \eta') \in \mathcal{L}$ the covectors $\mathrm{d}e_i$, $1 \le i \le l$, restricted to $T_{(y',\eta')}\mathcal{L}$ are linearly independent, so the H_{e_i} are transversal to $\mathcal{L}$, and they are linearly independent modulo $\mathcal{L}$. So we can give the initial condition $u|_{\mathcal{L}} = e_{l+1}$ and even prescribe u on a larger manifold, which leads to nonuniqueness of the extensions e_i.

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

$$\frac{\partial y'_j}{\partial \epsilon_i} = \frac{\partial e_i}{\partial \eta'_j}(y', \eta')\,, \qquad \frac{\partial \eta'_j}{\partial \epsilon_i} = -\frac{\partial e_i}{\partial y'_j}(y', \eta')\,, \qquad 1 \le i, j \le m - n\,.$$

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

It remains to check the commutation relations. The relations (5.1) are valid for any ϵ, because the Hamilton flow conserves the symplectic form on $T^*Y' \setminus 0$. The commutation relations for $\partial(y', \eta')/\partial\epsilon_i$ follow, using that $\partial(y', \eta')/\partial\epsilon_i = H_{e_i}$. □

Let M_{MN} be 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 now 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} ($\epsilon = 0$). 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 a function $S_{MN}(y'_I, x, \eta'_J, \epsilon)$ (see theorem 4.21 in Maslov and Fedoriuk [23]) such that M_{MN} is given by

$$y'_J = -\frac{\partial S_{MN}}{\partial \eta'_J}\,, \quad \eta'_I = \frac{\partial S_{MN}}{\partial y'_I}\,, \quad \xi = -\frac{\partial S_{MN}}{\partial x}\,, \quad e = \frac{\partial S_{MN}}{\partial \epsilon}\,.$$

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

$$\Psi_{MN}(y', x, e, \eta'_J, \epsilon) = S_{MN}(y'_I, x, \eta'_J, \epsilon) + \langle \eta'_J, y'_J \rangle - \langle \epsilon, e \rangle\,. \tag{5.3}$$

A Maslov-type phase function for Λ_{MN} then follows as

$$\begin{aligned}\Phi_{MN}(y', x, \eta'_J) &= \Psi_{MN}\left(y', x, \left.\frac{\partial S_{MN}}{\partial \epsilon}\right|_{\epsilon=0}, \eta'_J, 0\right) \\ &= S_{MN}(y'_I, \eta'_J, x, 0) + \langle \eta'_J, y'_J \rangle\,.\end{aligned}$$

6 Modeling: Joint Formulation

In this section we match the expression for the data modeled using the smooth jump (Kirchhoff) approximation 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.

From Theorem 3.1 it follows that reflection of an incident N-mode with covector $\tilde{\xi}_0$ into a scattered M-mode with covector $\hat{\xi}_0$ can take place if the frequencies are equal and $\hat{\xi}_0 + \tilde{\xi}_0$ is normal to the interface. In other words, $\hat{\xi}_0 + \tilde{\xi}_0$ must be in the wavefront set of the singular function of the interface, $\delta(z_n(x))$. Given $\tilde{\xi}_0$ and $\hat{\xi}_0$, one can identify $\mu(M)$ and $\nu(N)$ and define (at least to highest order) the reflection coefficient as a function of $(x, \hat{\xi}_0, \tilde{\xi}_0)$, $R^{\text{prin}}_{MN}(x, \hat{\xi}_0, \tilde{\xi}_0) = R^{\text{prin}}_{\mu(M),\nu(N)}(z'(x), \zeta'(\tilde{\xi}_0), \tau)$. This factor can now be viewed as a function of coordinates (y'_I, x, η'_J) or of coordinates (x, ξ, e) on Λ_{MN} (strictly speaking, defined only for x in the interface and ξ normal to the interface). To highest order it does not depend on $\|\xi\|$ and is simply a function of (x, e).

We obtain the following result, which is a generalization of the Kirchhoff approximation. The normalization factor $\|\frac{\partial z_n}{\partial x}\|$ of the δ-function is such that integral $\int \|\frac{\partial z_n}{\partial x}\| \delta(z_n(x)) \mathrm{d}x$ is an integral over the surface $z_n = 0$ with Euclidean surface measure in x-coordinates.

THEOREM 6.1 *Suppose Assumptions* 1, 2, 3, *and* 4 *are satisfied, microlocally for the relevant part of the data. Let* $\Phi_{MN}(y', x, \eta'_J)$ *and* $\mathcal{B}_{MN}(y'_I, x, \eta'_J)$ *be phase and amplitude, respectively, 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*

$$G^{\text{refl}}_{MN}(y') = (2\pi)^{-\frac{|J|}{2} - \frac{3n-1-c}{4}} \times \int \left(\mathcal{B}_{MN}(y'_I, x, \eta'_J) 2\mathrm{i}\tau(\eta') R_{MN}(y'_I, x, \eta'_J) + \textit{lower-order terms}\right) \times \mathrm{e}^{\mathrm{i}\Phi_{MN}(y', x, \eta'_J)} \left\| \frac{\partial z_n}{\partial x} \right\| \delta(z_n(x)) \mathrm{d}\eta'_J \, \mathrm{d}x \, ; \tag{6.1}$$

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

PROOF: We write the distribution kernel of the reflected data (3.3) in a form similar to (4.1). First, recall the reciprocal expression for the Green's function (2.24),

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

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

$$G^{\text{refl}}_{MN}(\hat{x},\tilde{x},t) = (2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2}-n} \int_{z_n=0} \big(2\mathrm{i}\tau \mathcal{A}_M(\hat{x}_{\hat{I}}, x(z), \hat{\xi}_{\hat{J}}, \tau)$$

$$\times \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x(z), \tilde{\xi}_{\tilde{J}}, \tau) R_{\mu(M)\nu(N)}(z,\zeta',\tau) + \text{lower-order terms}\big)$$

$$\times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x},\tilde{x},t,x(z),\hat{\xi}_{\hat{J}},\tilde{\xi}_{\tilde{J}},\tau)} \left|\det \frac{\partial x}{\partial z}\right| \left\| \frac{\partial z_n}{\partial x} \right\| \mathrm{d}\hat{\xi}_{\hat{J}}\, \mathrm{d}\tilde{\xi}_{\tilde{J}}\, \mathrm{d}\tau\, \mathrm{d}z' ,$$

where ζ' depends on $(x(z), \tilde{\xi}_0)$ (the indices μ and ν for the reflection coefficients are 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 obtain

$$(2\pi)^{-\frac{|\hat{J}|+|\tilde{J}|+1}{2}-n} \int \big(2\mathrm{i}\tau \mathcal{A}_M(\hat{x}_{\hat{I}}, x, \hat{\xi}_{\hat{J}}, \tau) \mathcal{A}_N(\tilde{x}_{\tilde{I}}, x, \tilde{\xi}_{\tilde{J}}, \tau)$$

$$\times R_{\mu(M)\nu(N)}(z(x), \zeta'(\tilde{\xi}_J, x), \tau) + \text{lower-order terms}\big)$$

$$\times \mathrm{e}^{\mathrm{i}\Phi_{MN}(\hat{x},\tilde{x},t,x,\hat{\xi}_{\hat{J}},\tilde{\xi}_{\tilde{J}},\tau)} \left\| \frac{\partial z_n}{\partial x} \right\| \delta(z_n(x))\, \mathrm{d}\hat{\xi}_{\hat{J}}\, \mathrm{d}\tilde{\xi}_{\tilde{J}}\, \mathrm{d}\tau\, \mathrm{d}x .$$

This formula is very similar to (4.1), only the amplitude is different and

$$\frac{\delta c_{ijkl}(x)}{\rho(x)} \quad \text{and} \quad \frac{\delta\rho(x)}{\rho(x)}$$

are replaced by the δ-function $\|\partial z_n/\partial x\|\delta(z_n(x))$. Also, the factors $w_{MN;ijkl}$ and $w_{MN;0}$ depend only on the background medium, while $R_{\mu(M)\nu(N)}$ depends on the total medium. 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, as explained 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 and 4, the data is the image of a Fourier integral operator acting on $\|\partial z_n/\partial x\|\delta(z_n(x))$ and that it is given by (6.1). □

7 Inverse Scattering Revisited

In this section we present the main results of the paper. We first construct a Fourier integral operator and a reflectivity function, which is a function of subsurface position and the additional coordinate e. The data is modeled by letting the Fourier integral operator act on the reflectivity. The construction is such that this Fourier integral operator is invertible. We discuss its inverse. Finally, a set of pseudodifferential operators is constructed that annihilates the data if the smooth part of the medium above the reflector is correctly chosen.

7.1 Invertible Transformation into Subsurface Coordinates

We now construct the reflectivity function and the operator that maps it to seismic data. This is done by applying the results of Section 5 to the Kirchhoff modeling formula (6.1) and its equivalent in the Born approximation (4.1).

THEOREM 7.1 *Suppose microlocally Assumptions* 1, 2, 3, 4, *and* 5 *are satisfied. Let H_{MN} be the Fourier integral operator with canonical relation given by the extended map $(x,\xi,e,\epsilon)\mapsto(y',\eta')$ constructed in Section* 5 *and with amplitude to highest order given by $(2\pi)^{\frac{n}{2}}(2i\tau)\mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)$ such that $\mathcal{B}_{MN}(\epsilon=0)$ is as given in Theorem* 4.2. *Then the data, in both Born and Kirchhoff approximations, is given by H_{MN} acting on a distribution $r_{MN}(x,e)$ of the form*

$$r_{MN}(x,e)=(\text{pseudo})(x,D_x,e)(\text{distribution})(x)\,. \tag{7.1}$$

For the Kirchhoff approximation this distribution equals $\|\partial z_n/\partial x\|\delta(z_n(x))$, while the principal symbol of the pseudodifferential operator equals $R_{MN}(x,e)$, so to highest order $r_{MN}(x,e)=R_{MN}(x,e)\|\partial z_n/\partial x\|\delta(z_n(x))$. For the Born approximation the function $r_{MN}(x,e)$ is given by a pseudodifferential operator acting on $(\frac{\delta c_{ijkl}}{\rho},\frac{\delta\rho}{\rho})_\alpha$, with principal symbol $(2i\tau(x,\xi,e))^{-1}w_{MN;\alpha}(x,\xi,e)$; see (4.3).

PROOF: We do the proof for the Kirchhoff approximation using (6.1); for the Born approximation the proof is similar. Since Assumption 5 is satisfied, the projection $\pi_{Y'}$ of Λ_{MN} into $T^*Y'\setminus 0$ is an embedding, and the image is a co-isotropic submanifold of $T^*Y'\setminus 0$. Therefore we can apply Lemma 5.1. Formula (5.3) implies that the phase factor $e^{i\Phi_{MN}}$ can be written in the form

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

we define

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

Thus the number of phase variables is increased by making use of a stationary-phase argument.

Let $\mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)$ be as described. Then we obtain

$$\begin{aligned}G^{\text{refl}}_{MN}(y')&=(2\pi)^{-\frac{|J|+n-1-c}{2}-\frac{2n-1-c}{2}}\\&\times\int\big((2\pi)^{\frac{n}{2}}2i\tau(\eta')\mathcal{B}_{MN}(y'_I,x,\eta'_J,\epsilon)R_{MN}(x,e)+\text{lower-order terms}\big)\\&\qquad\times e^{i\Psi_{MN}(y',x,e,\eta'_J,\epsilon)}\left\|\frac{\partial z_n}{\partial x}\right\|\delta(z_n(x))\,d\eta'_J\,d\epsilon\,dx\,de\,.\end{aligned}\tag{7.2}$$

In this formula, the data is represented as a Fourier integral operator acting on $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$ considered as a function of (x,e). Multiplying by H^{-1}_{MN} gives a

pseudodifferential operator of the form described acting on $\|\frac{\partial z_n}{\partial x}\|\delta(z_n(x))$. Thus we obtain the result. □

Remark 7.2. It is implicitly assumed that $\mathcal{B}(y_I', x, \eta_J', \epsilon)$ vanishes outside a conic (η', ϵ) neighborhood of $\{(y_I', x, \eta_J') \in \operatorname{supp} \mathcal{B}_{MN}(y_I', x, \eta_J')\} \times \{\epsilon = 0\}$ (cf. (4.10)). The operator H_{MN} is defined microlocally on a neighborhood of $\mathcal{L}$, and its inverse on a neighborhood of $\{\epsilon = 0\}$.

7.2 The Inversion Operator

The operator H_{MN} is invertible. A choice of phase function and amplitude for its inverse is given by (see Treves [31, chap. 8])

$$-\Psi_{MN}(y', x, e, \eta_J', \epsilon), \qquad \mathcal{B}_{MN}(y_I', x, \eta_J', \epsilon)^{-1} \left| \det \frac{\partial(y', \eta')}{\partial(y_I', x, \eta_J', \epsilon)} \right|,$$

respectively. Thus microlocally an explicit expression for $r_{MN}(x, e)$ in terms of the data is given by

$$r_{MN}(x, e) = \int \mathcal{B}_{MN}(y_I', x, \eta_J', \epsilon)^{-1} \left| \det \frac{\partial(y', \eta')}{\partial(y_I', x, \eta_J', \epsilon)} \right| \times \mathrm{e}^{-\mathrm{i}\Psi_{MN}(y', x, e, \eta_J', \epsilon)} d_{MN}(y') \mathrm{d}\eta_J' \, \mathrm{d}\epsilon \, \mathrm{d}y' . \tag{7.3}$$

Since the function $r_{MN}(x, e)$ is to highest order equal to the product of the reflection coefficient and the singular function of the reflector surface, this reconstruction of the function $r_{MN}(x, e)$ leads to the following result for Kirchhoff data:

COROLLARY 7.3 *Suppose that the medium in between the acquisition manifold and the reflector is known and satisfies Assumptions* 1, 2, 3, 4, *and* 5. *Then one can reconstruct the position of the interface and the angle-dependent reflection coefficient* $R_{\mu\nu}(x, e)$ *on the interface.*

The precise reconstruction assumes mode-decoupled and single-scattered data. To ensure that the data can be mode decoupled requires maximal acquisition geometry $c = 0$. Of course, the reflection coefficient $R_{\mu\nu}(x, e)$ can only be reconstructed for values of e that are associated with rays connected to sources and receivers in the acquisition manifold. In general, a finite set of values of e is required to reconstruct the medium jump.

The motivation for Lemma 5.1 can be explained in case e is chosen to be the scattering angle/azimuth. 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 $T^*Y' \setminus 0$ is chosen.

7.3 Annihilators of the Data

The result of the previous subsections gives information on the problem of reconstructing the smooth background medium (or, in the Kirchhoff approximation, the smooth medium parameters above/in between the interfaces). If $n-1-c>0$, there is a redundancy in the data through the variable e. If the smooth medium parameters (above the interface) are correct, then applying the operator H_{MN}^{-1} of Theorem 7.1 to the data results in a reflectivity function $r_{MN}(x,e)$ such that the position of the singularities does not depend on e. The fact that the inverted data should "line up" in the variable e can be used as a criterion to assess the accuracy of the background medium.

One way to measure how well the data line up is by taking the derivative with respect to e. If $r_{MN}(x,e)$ depends smoothly on e as in (7.1), then $\frac{\partial}{\partial e} r_{MN}(x,e)$ is one order less singular than if it does not have this smooth dependence on e (for instance, a δ-function versus its derivative in the Kirchhoff case). Taking also the factor in front of the δ-function of r_{MN} into account (see (7.1)), we obtain that to the highest two orders

$$\left(R_{MN}(x,e)\frac{\partial}{\partial e}-\frac{\partial R_{MN}^{\mathrm{prin}}}{\partial e}(x,e)\right)r_{MN}(x,e)=0\,. \tag{7.4}$$

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

Conjugating the differential operator of (7.4) with the invertible Fourier integral operator H_{MN}, we obtain a pseudodifferential operator on $\mathcal{D}'(Y')$. Thus we obtain the following corollary of Theorem 7.1:

COROLLARY 7.4 *Let the pseudodifferential operators $W_{MN}(y', D_{y'})$ be given by*

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

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

$$W_{MN}(y', D_{y'})d_{MN}(y')=0\,. \tag{7.5}$$

For values of e where $R_{MN}(x,e)\neq 0$, the operator $W_{MN}(y', D_{y'})$ can be chosen such that (7.5) *is valid to all orders.*

In principle the operators $W_{MN}(y', D_{y'})$ can be used to obtain a quantitative criterion of how well the data line up. Symes [28] discusses such criteria for acoustic media using the offset coordinate.

Appendix: Notation

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 $Q(x,\xi)$, $\mathcal{Q}(x,x_0)$ for its distribution kernel, and $Q^{\mathrm{prin}}(x,\xi)$ for its principal symbol.

Index of Notation

General

δ_{ij}	Kronecker delta
n	p. 261
x	p. 261
$X \subset \mathbb{R}^n$	p. 261
t	p. 262
$Y', y' \in Y'$	p. 262
$z = (z', z_n)$	p. 276
$e \in E$	p. 288
$\xi, \eta, \zeta, \tau, \epsilon$	cotangent vectors for x, y, z, t, e
$\pi_X, \pi_{Y'}$	p. 287

Subscripts

i, j, k, l	pp. 261, 265
M, N	p. 266
I, J	p. 270
x_I	p. 270
a	p. 277
μ, ν	p. 276

Pseudodifferential Operators

P_{il}	p. 266
A_{il}	p. 266
$Q_{iM}(x, D)$	p. 266
$P_M(x, D)$	p. 266
$A_M(x, D)$	p. 267
$B_M(x, D)$	p. 269
$R^0_{\mu\nu}(z, D_{z'}, D_t)$	p. 277
$R_{\mu\nu}(z, D_{z'}, D_t)$	p. 277
$N_{MN;\alpha\beta}(x, D)$	p. 286
$W_{MN}(y', D_{y'})$	p. 298

Field Quantities

$\rho(x)$	p. 265
$c_{ijkl}(x)$	p. 265
$\delta c_{ijkl}(x), \delta\rho(x)$	p. 281
$g_\alpha(x)$	p. 286
$u_i(x, t)$	p. 266
$f_i(x, t)$	p. 266
$u_M(x, t),\ f_M(x, t)$	p. 266
$u_{M,\pm}(x, t),\ f_{M,\pm}(x, t)$	p. 269
$V_a(x, t)$	p. 277
$V_\mu(x, t)$	p. 278
$d_{MN}(y')$	p. 297
$r_{MN}(x, e)$	p. 296

FIOs and Related

$x_M(x_0, \xi_0, t), \xi_M(x_0, \xi_0, t)$	p. 270
$C_{M,\pm}, C_M$	pp. 270, 271
$\phi_{M,\pm}, \phi_M$	p. 271
$\mathcal{A}_{M,\pm}(x_I, \dots), \mathcal{A}_M(\dots)$	p. 271
G_M	p. 269
$G_{M,\pm}$	p. 270
$\delta G_{il}, \delta G_{MN}$	p. 282
$F_{MN;ijkl}, F_{MN;0}$	p. 284
$F_{MN;\alpha}$	p. 286
Φ_{MN}	p. 282
$\mathcal{B}_{MN}, w_{MN;ijkl}, w_{MN;0}$	p. 282
$\Lambda_{0,MN}$	p. 283
Λ_{MN}	p. 285
$\mathcal{L}$	p. 287
M_{MN}	p. 293
Ψ_{MN}	p. 293
H_{MN}	p. 296

Acknowledgments. We thank Prof. J. J. Duistermaat for many helpful discussions. The first author thanks Shell International Exploration and Production for its financial support. This work was supported in part by the members of the Consortium Project at the Center for Wave Phenomena, Colorado School of Mines.

Bibliography

[1] Aki, K.; Richards, P. G. *Quantitative seismology: theory and methods*. Vol. 1. W. H. Freeman, San Francisco, 1980.

[2] Beylkin, G. The inversion problem and applications of the generalized Radon transform. *Comm. Pure Appl. Math.* **37** (1984), no. 5, 579–599.

[3] Beylkin, G. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform. *J. Math. Phys.* **26** (1985), no. 1, 99–108.

[4] Beylkin, G.; Burridge, R. Linearized inverse scattering problems in acoustics and elasticity. *Wave Motion* **12** (1990), no. 1, 15–52.

[5] Bleistein, N. On the imaging of reflectors in the earth. *Geophysics* **52** (1987), no. 7, 931–942.

[6] Braam, P. J.; Duistermaat, J. J. Normal forms of real symmetric systems with multiplicity. *Indag. Mathem. (N.S.)* **4** (1993), no. 4, 407–421.

[7] Burridge, R.; Beylkin, G. On double integrals over spheres. *Inverse Problems* **4** (1988), no. 1, 1–10.

[8] Chapman, C. H. Reflection/transmission coefficient reciprocities in anisotropic media. *Geophys. J. Internat.* **116** (1994), 498–501.

[9] Dencker, N. On the propagation of polarization sets for systems of real principal type. *J. Funct. Anal.* **46** (1982), no. 3, 351–372.

[10] Duistermaat, J. J. *Fourier integral operators*. Birkhäuser, Boston, 1996.

[11] Frazer, N. L.; Sen, M. K. Kirchhoff-Helmholtz reflection seismograms in a laterally inhomogeneous multi-layered medium. I. Theory. *Geophys. J. R. Astr. Soc.* **80** (1985), 121–147.

[12] Guillemin, V. On some results of Gel′fand in integral geometry. *Pseudodifferential operators and applications (Notre Dame, Ind., 1984)*, 149–155. Proceedings of Symposia in Pure Mathematics, 43. American Mathematical Society, Providence, R.I., 1985.

[13] Hansen, S. Solution of a hyperbolic inverse problem by linearization. *Comm. Partial Differential Equations* **16** (1991), no. 2-3, 291–309.

[14] Hoop, M. V. de; Bleistein, N. Generalized Radon transform inversions for reflectivity in anisotropic elastic media. *Inverse Problems* **13** (1997), no. 3, 669–690.

[15] Hoop, M. V. de; Brandsberg-Dahl, S. Maslov asymptotic extension of generalized Radon transform inversion in anisotropic elastic media: a least-squares approach. *Inverse Problems* **16** (2000), no. 3, 519–562.

[16] Hörmander, L. *The analysis of linear partial differential operators*. Vol. 1. Grundlehren der mathematischen Wissenschaften, 256-257. Springer, Berlin, 1983.

[17] Hörmander, L. *The analysis of linear partial differential operators*. Vol. 3. Grundlehren der mathematischen Wissenschaften, 274. Springer, Berlin, 1985.

[18] Hörmander, L. *The analysis of linear partial differential operators*, Vol. 4. Grundlehren der mathematischen Wissenschaften, 275. Springer, Berlin, 1985.

[19] Ivriĭ, V. Ja. Wave fronts of solutions of symmetric pseudodifferential systems. *Sibirsk. Mat. Zh.* **20** (1979), no. 3, 557–578, 692; translation in *Siberian Math. J.* **20** (1979), no. 3, 390–405.

[20] Kennett, B. L. N. *Seismic wave propagation in stratified media*. Cambridge University Press, Cambridge, 1983.

[21] Kroode, A. P. E. ten; Smit, D. J.; Verdel, A. R. A microlocal analysis of migration. *Wave Motion* **28** (1998), 149–172.

[22] Lions, J. L.; Magenes, E. *Non-homogeneous boundary value problems and applications*. Vol. 1. Grundlehren der mathematischen Wissenschaften, 181. Springer, New York-Heidelberg, 1972.

[23] Maslov, V. P.; Fedoriuk, M. V. *Semiclassical approximation in quantum mechanics*. Mathematical Physics and Applied Mathematics, 7. Contemporary Mathematics, 5. D. Reidel, Dordrecht–Boston, 1981.

[24] Musgrave, M. J. P. *Crystal acoustics*. Holden-Day, San Francisco, 1970.

[25] Nolan, C. J.; Symes, W. W. Global solution of a linearized inverse problem for the wave equation. *Comm. Partial Differential Equations* **22** (1997), no. 5-6, 919–952.

[26] Rakesh. A linearised inverse problem for the wave equation. *Comm. Partial Differential Equations* **13** (1988), no. 5, 573–601.

[27] Stolk, C. C. Microlocal analysis of a seismic linearized inverse problem. *Wave Motion* **32** (2000), 267–290.

[28] Symes, W. W. A differential semblance algorithm for the inverse problem of reflection seismology. *Comput. Math. Appl.* **22** (1991), no. 4-5, 147–178.

[29] Taylor, M. E. Reflection of singularities of solutions to systems of differential equations. *Comm. Pure Appl. Math.* **28** (1975), no. 4, 457–478.

[30] Treves, F. *Introduction to pseudodifferential and Fourier integral operators*. Vol. 1. University Series in Mathematics. Plenum, New York, 1980.

[31] Treves, F. *Introduction to pseudodifferential and Fourier integral operators*. Vol. 2. University Series in Mathematics. Plenum, New York, 1980.

[32] Xu, S.; Chauris, H.; Lambaré, G.; Noble, M. Common angle migration: a strategy for imaging complex media. Preprint. École des Mines de Paris, 1999.

CHRISTIAAN C. STOLK
Rice University
Department of Computational
and Applied Mathematics
6100 Main Street - MS 134
Houston, TX 77005-1892
E-mail: cstolk@caam.rice.edu

MAARTEN V. DE HOOP
Colorado School of Mines
Center for Wave Phenomena
Golden, CO 80401
E-mail: mdehoop@mines.edu

Received December 2000.