CANADIAN APPLIED
MATHEMATICS QUARTERLY
Volume 10, Number 2, Summer 2002

SEISMIC WAVEFIELD 'CONTINUATION' IN THE SINGLE SCATTERING APPROXIMATION: A FRAMEWORK FOR DIP AND AZIMUTH MOVEOUT

Based on a presentation at the PIMS–MITACS Workshop on Inverse Problems and Imaging, University of British Columbia, June 9–10, 2001.

MAARTEN V. DE HOOP, ALISON E. MALCOLM
AND JÉRÔME H. LE ROUSSEAU

ABSTRACT. Seismic data are commonly modeled by a high-frequency single scattering approximation. In this paper we use methods from microlocal analysis and the theory of Fourier integral operators to study continuation of the seismic wavefield in this single scattering approximation. This amounts to a linearization in the medium coefficient about a smooth background. The discontinuities are contained in the medium perturbation. We use the smooth background to derive the continuation as the composition of imaging, modeling and restriction operators.

1 Introduction. In reflection seismology one places point sources and point receivers on the Earth's surface. The source generates acoustic waves in the subsurface, that are reflected where the medium properties vary discontinuously. The recorded reflections that can be observed in the data are used to reconstruct these discontinuities. In principle, the recordings are taken on an acquisition manifold, made up of all source and receiver positions and a time interval. In practice, however, certain subsets in the acquisition manifold are not covered. In this paper, we discuss how, and conditions when, data can be continued from any open subset of the acquisition manifold to a more complete acquisition manifold.

The data are commonly modeled by a high-frequency single scattering approximation. This amounts to a linearization in the medium coefficient about

M. V. d. H. thanks The Mathematical Sciences Research Institute for partial support, through NSF grant DMS-9810361. He also thanks Christiaan Stolk and Clifford Nolan for helpful discussions, and Gunther Uhlmann for providing a very stimulating environment during the Inverse Problems program in Fall 2001.

AMS subject classification: 35S30, 35L05, 86A22, 35R30, 86A15.

a smooth background. The discontinuities are contained in the medium perturbation [**2**]. Thus a linear operator, the modeling operator, depending on the background, that maps the perturbation to the data is obtained. The smooth background (C^∞) is associated with a computational medium rather than a physical one, the distributional ($\mathcal{E}'$) perturbation accounts for geological transitions and the medium's singularities across interfaces and faults. We will consider seismic wavefield continuation in the single scattering approximation and will use an image of the distributional perturbation as an intermediary. We will require some knowledge of the medium properties, viz. the smooth background, but not their discontinuities.

The framework of Fourier integral operators (FIOs) and their composition through the clean intersection calculus [**11**], [**19**], [**32**] yield the tools to carry out the following processes: seismic modeling (FIO), acquisition (restriction FIO), imaging (adjoint FIO), resolution (normal operator, the sum of a pseudodifferential operator and a nonlocal FIO) and inversion [**29**].

The wavefront set of the data is, under the so-called Bolker condition [**14**], a coisotropic submanifold of the acquisition cotangent bundle. It reveals a structure: that of characteristic strips. Restricting in the imaging FIO the seismic data to a common coordinate value on these strips, yields a generalized Radon transform (GRT [**2**], [**8**], [**9**]) that maps the reflection data into a seismic image. (Under certain conditions this GRT is an FIO [**28**].) Collecting these seismic images from the points on the characteristic strips corresponding to available data results in the set of so-called common-image-point gathers (CIGs). In the presence of caustics, a filter needs to be designed and applied prior to extracting a trace from each of the CIGs in the set, to form a model image of the singular component of the medium [**6**], [**30**].

From this image, we model seismic data that correspond to a different coordinate value on the characteristic strips. The result of this procedure is a composition of FIOs yielding seismic wavefield continuation, be it in the single scattering approximation. Relevant examples of seismic wavefield continuation are the 'transformation to zero offset' (TZO [**15**]) and the 'transformation to common (prescribed) azimuth' (TCA [**3**]). The distribution kernel of TZO is called dip moveout (DMO); the distribution kernel of TCA is called azimuth moveout (AMO).

In practice, DMO/AMO is applied to data sets using a constant coefficient model. This is done because, on the one hand, the traditional transforms were derived in constant media, and on the other hand, to make the algorithms which apply DMO/AMO to data simpler and more efficient. Here, we develop a framework for DMO/AMO in heterogeneous, smoothly varying, models allowing the formation caustics. We can thus assess the error in applying DMO/AMO in a simplified model if the 'true' model were to have (strong) variations. (In fact, by composition, the error generating operator can be obtained.)

We mention some of the applications of DMO/AMO:

(i) DMO/AMO effectively corresponds to 'partial stacking' of seismic data; such stacking generates reduced data sets viz. on acquisition (zero-offset, common-azimuth) submanifolds. It is important to note that rather than obtaining a reduced dataset from restricting the complete dataset, DMO/AMO will enhance the signal-to-noise ratio by using all available data in the reduction. Under certain conditions addressed in this paper, the reduced data set produces an image consistent with the complete data set. The advantage of using reduced data sets is computational efficiency.

(ii) DMO can be employed as a tool for 'velocity analysis', i.e. estimating the smooth background.

(iii) AMO can be employed to carry out approximate (based on a linearized scattering model) seismic data 'regularization'.

The basic idea of investigating the composition of imaging and modeling operators dates back in particular to the work of Goldin [**13**].

2 High-frequency Born modeling and imaging. We consider the scalar wave equation for acoustic waves in a constant density medium in $\mathbb{R}^n$. We introduce coordinates $x \in \mathbb{R}^n$. The scalar acoustic wave equation is given by

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

where $D_x = -\mathrm{i}\frac{\partial}{\partial x}$. The equation is considered in a time interval $]0, T[$.

If $c \in C^\infty$ the solution operator of (1) propagates singularities along bicharacteristics. These are the solutions of a Hamilton system with Hamiltonian given by the principal symbol of P,

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

The Hamilton system is given by

$$\frac{\partial(x,t)}{\partial\lambda} = \frac{\partial P}{\partial(\xi,\tau)}, \quad \frac{\partial(\xi,\tau)}{\partial\lambda} = -\frac{\partial P}{\partial(x,t)}. \tag{2}$$

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

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

and similarly for $\boldsymbol{t}, \boldsymbol{\xi}$; τ is invariant along the Hamilton flow. The evolution parameter λ is the time t.

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

$$u(x,t) = \int_0^t \int G(x, t - t_0, x_0) f(x_0, t_0)\, \mathrm{d}x_0\, \mathrm{d}t_0, \tag{3}$$

where G defines a Fourier integral operator (FIO) with canonical relation, Λ_G, that is essentially a union of bicharacteristics,

$$\Lambda_G = C_+ \cup C_-,$$

$$C_\pm = \Big\{ \Big(\boldsymbol{x}(x_0, \alpha, \tau, \pm t), t, \boldsymbol{\xi}(x_0, \alpha, \tau, \pm t), \mp\tau; x_0, -\underbrace{\mp\big(\tau/c(x_0)\big)\alpha}_{\xi_0} \Big) \Big\}.$$

Let

$$(x_I, x_0, \underbrace{\xi_J, \tau}_{\theta}) \quad \text{with} \quad I \cup J = \{1, \ldots, n\}, N := |J| + 1,$$

denote coordinates on $C_\pm$. A function S will locally describe C_+ according to

$$\begin{aligned} x_J &= -\frac{\partial}{\partial \xi_J} S, & t &= -\frac{\partial}{\partial \tau} S, \\ \xi_I &= \frac{\partial}{\partial x_I} S, & \xi_0 &= -\frac{\partial}{\partial x_0} S, \end{aligned} \tag{4}$$

and generates the non-degenerate phase function

$$\phi_+(x, x_0, t, \xi_J, \tau) = S(x_I, x_0, \xi_J, \tau) + \langle \xi_J, x_J \rangle + \tau t. \tag{5}$$

With the canonical relation Λ_G is thus associated the (non-degenerate) phase function ϕ defined by $\phi = \phi_-$ if $\tau > 0$, $\phi = \phi_+$ if $\tau < 0$. The kernel of the mentioned FIO can then be written as a sum of oscillatory integral (OI) contributions

$$G(x, t, x_0) = \sum_i \int_{\mathbb{R}^{N^{(i)}}} a^{(i)}(x, t, x_0, \theta) \exp[\mathrm{i}\phi^{(i)}(x, x_0, t, \theta)]\, \mathrm{d}\theta, \tag{6}$$

where the $a^{(i)}$ are suitable symbols, see [**11,** chapter 5].

We adopt the linearized scattering approximation, in which the linearization is in the coefficient c around a smooth background c_0, $c = c_0 + \delta c$. The perturbation δc may contain singularities. We assume that its support is contained in $X \subset \mathbb{R}^n$. The perturbation in G at (s, r, t) with $s, r \in X$ and $t \in]0, T[$ is given by (cf. (3))

(7)

$$\delta G(r, t, s) = \int_X \int_0^t G(r, t - t_0, x_0) 2c_0^{-3}(x_0)\delta c(x_0)\partial_{t_0}^2 G(x_0, t_0, s)\,\mathrm{d}t_0\,\mathrm{d}x_0.$$

The singular part of δG is obtained by substituting (6) into (7). This defines the data modeling map

$$F = F[c_0]\colon \delta c \longmapsto \mathcal{R}\delta G,$$

where $\mathcal{R}$ is the restriction to the acquisition manifold $Y \ni (s, r, t)$ with $Y = O_s \times O_r \times]0, T[$ subject to $O_s, O_r \subset \partial X$ open. Throughout the paper s and r denote source and receiver positions, respectively.

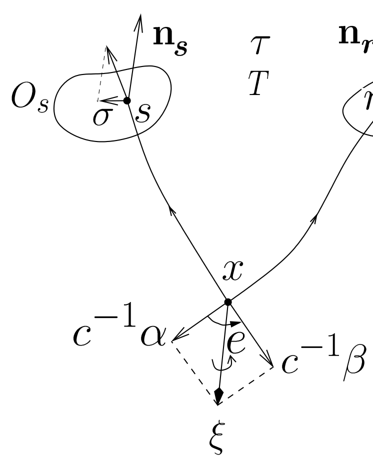

FIGURE 1: Source-receiver bicharacteristics; parameterization of Λ_F.

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

Theorem 2.1 ([**26**]). *With Assumption* 1 *the map* F *is a Fourier integral operator* $\mathcal{D}'(X) \to \mathcal{D}'(Y)$ *of order* $(n-1)/4$ *with canonical relation*

(8)

$$\begin{aligned}\Lambda_F = \Big\{ \Big(&\boldsymbol{s}(x_0,\beta), \boldsymbol{r}(x_0,\alpha), \overbrace{T(x_0,\alpha)+T(x_0,\beta)}^{T(x_0,\alpha,\beta)}, \boldsymbol{\sigma}(x_0,\beta), \boldsymbol{\rho}(x_0,\alpha), \tau; \\ &x_0, -\underbrace{\big(\tau/c(x_0)\big)(\alpha+\beta)}_{\boldsymbol{\xi}(x_0,\alpha,\beta,\tau)} \Big) \;\Big|\; (x_0,\alpha,\beta) \in K, \tau \in \mathbb{R}\setminus 0 \Big\} \\ &\subset T^*Y\setminus 0 \times T^*X \setminus 0,\end{aligned}$$

where $K \subset \mathbb{R}^n \times \{(\alpha,\beta) \in S^{n-1} \times S^{n-1} \mid \alpha+\beta \neq 0\}$. *Here,*

$$\underbrace{\boldsymbol{x}\big(x_0,\alpha,\tau,T(x_0,\alpha)\big)}_{=:\,\boldsymbol{r}(x_0,\alpha)} \in O_r, \quad \underbrace{\boldsymbol{x}\big(x_0,\beta,\tau,T(x_0,\beta)\big)}_{=:\,\boldsymbol{s}(x_0,\beta)} \in O_s,$$

which expresses that the time T *is locally solved from the equation describing the intersection of the rays with the acquisition manifold, while*

$$\boldsymbol{\rho}(x_0,\alpha) = (I - \mathbf{n}_{\boldsymbol{r}} \otimes \mathbf{n}_{\boldsymbol{r}}) \cdot \boldsymbol{\xi}\big(x_0,\alpha,\tau,T(x_0,\alpha)\big) \tag{9}$$

where $\mathbf{n}_{\boldsymbol{r}}$ *is the unit normal to* O_r *at* $\boldsymbol{r}(x_0,\alpha)$. *A similar expression holds for* $\boldsymbol{\sigma}(x_0,\beta)$.

The parameterization of Λ_F is illustrated in Figure 1. The cotangent vectors σ, ρ can be identified with acquisition 'slopes' p_s, p_r in the data, in accordance with $\boldsymbol{\sigma}(x_0,\beta) = -\tau p_s(x_0,\beta)$ and $\boldsymbol{\rho}(x_0,\alpha) = -\tau p_r(x_0,\alpha)$.

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

We assume also that ψ vanishes outside Y. To image the singularities of δc from the singularities in the data we consider the adjoint $F^*\psi$, which is a Fourier integral operator also.

Assumption 2 ([**14**]). The projection of the canonical relation (8) on $T^*Y \setminus 0$ is an embedding.

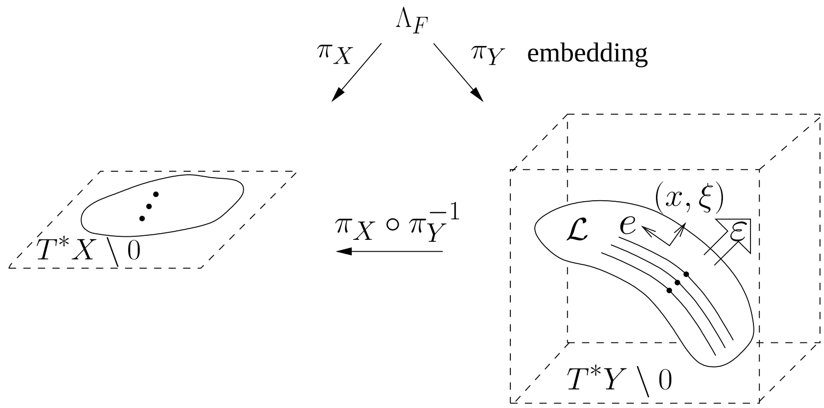

FIGURE 2: Canonical relation and characteristic strips [**29**]. (Their parameterization is illustrated in Figure 1.)

This assumption is also known as the Bolker condition.

Since (8) is a canonical relation that projects submersively on the subsurface variables (x, ξ), the projection of (8) on $T^*Y \setminus 0$ is immersive [**19,** Lemma 25.3.6 and (25.3.4)]. Therefore only the injectivity in the assumption needs to be verified [**22**]. In fact, it is precisely the injectivity condition that has been assumed in what seismologists call 'map migration'; see [**20**] for a recent summary. Figure 2 illustrates this schematically.

The following theorem describes the reconstruction of δc modulo a pseudodifferential operator with principal symbol that is nonzero at (x, ξ) whenever there is a point $(s, r, t, \sigma, \rho, \tau; x, \xi)$ in the canonical relation (8) with $(s, r, t, \sigma, \rho, \tau)$ in the support of ψ (i.e. whenever there is illumination).

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

For the purpose of wavefield continuation within the acquisition manifold Y, we parameterize Λ_F with acquisition coordinates s, r rather than β, α. To describe the kernel of the operator F as an OI on a neighborhood of the point on Λ_F parameterized by $(x_0, \alpha, \beta, \tau)$, the minimum number of phase variables is given by the corank of the projection

$$D\pi \colon T\Lambda_F \longrightarrow T(T^*Y \times T^*X)$$

at $(x_0, \alpha, \beta, \tau)$, which is here given by

$$\operatorname{corank} D\pi = 1 + \operatorname{corank} \frac{\partial \boldsymbol{s}}{\partial \beta}(x_0, \beta) + \operatorname{corank} \frac{\partial \boldsymbol{r}}{\partial \alpha}(x_0, \alpha).$$

This corank is > 1 when s or r is in a caustic point relative to x_0. Let

$$\Lambda'_F = \Lambda_F \setminus \big\{\text{closed neighborhood of } \{\lambda \in \Lambda_F \mid \operatorname{corank} D\pi > 1\}\big\}. \tag{10}$$

Λ'_F can be described by phase functions of the 'traveltime' form $\tau(t - T^{(m)})$ with the only phase variable being τ. Here, $T^{(m)}$ is the value of the time variable in (8). The index m labels the branches of the multi-valued traveltime function. Thus the set $\{T^{(m)}\}_{m\in M}$ describes the canonical relation (8) except for a neighborhood of the subset of the canonical relation where the mentioned projection is degenerate. Each $T^{(m)}$ can be viewed as a function defined on a subset $D^{(m)}$ of $X \times O_s \times O_r$. We define $F^{(m)}$ to be a contribution to F with phase function given by $\tau\big(t - T^{(m)}(x, s, r)\big)$, and symbol $A^{(m)}$ in a suitable class such that on the subset Λ'_F of the canonical relation F is given microlocally by $\sum_{m\in M} F^{(m)}$.

3 Generalized Radon transform. We can use $(x, \xi) \in T^*X \setminus 0$ as the first $2n$ local coordinates on the canonical relation (8) (cf. [**18,** Prop. C.3.3]). In addition, we need to parameterize the subsets (these are characteristic strips) of the canonical relation given by $(x, \xi) =$ constant; we denote such parameters by e. The canonical relation (8) was parameterized by (x, α, β, τ). We relate (x, ξ, e) by a coordinate transformation to (x, α, β, τ): A suitable choice when $\alpha \neq \beta$ is the scattering angles given by [**9**]

$$\boldsymbol{e}(x, \alpha, \beta) = \left(\arccos(\alpha\cdot\beta), \frac{-\alpha + \beta}{2\sin\big(\arccos(\alpha\cdot\beta)/2\big)}\right) \in \,]0, \pi[\, \times S^{n-2}. \tag{11}$$

On $D^{(m)}$ there is a map $(x, \alpha, \beta) \mapsto (x, s, r)$. We define $\boldsymbol{e}^{(m)} = \boldsymbol{e}^{(m)}(x, s, r)$ as the composition of $\boldsymbol{e}$ with the inverse of this map, see Figure 1.

In preparation for the generalized Radon transform (GRT) we define the 'angle' transform, $\check{L}$, via a restriction in F^* of the mapping $\boldsymbol{e}^{(m)}$ to a prescribed value e, i.e. the distribution kernel of each contribution ${F^{(m)}}^*$ is multiplied by $\delta\big(e - \boldsymbol{e}^{(m)}(x, s, r)\big)$ (which is justified by [**17,** Thm. 8.2.10]). Invoking the Fourier representation of this δ, the kernel of $\check{L}$ follows as

$$\check{L}(x, e, r, s, t) = \sum_{m\in M} (2\pi)^{-(n-1)} \int \overline{A^{(m)}(x, s, r, \tau)} \cdot \exp[\mathrm{i}\Phi^{(m)}(x, e, s, r, t, \varepsilon, \tau)]\, \mathrm{d}\tau\, \mathrm{d}\varepsilon, \tag{12}$$

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

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

In these expressions, ε is the cotangent vector corresponding to e, as in [**29**].

Let $\psi_L = \psi_L(D_s,D_r,D_t)$ be a pseudodifferential cutoff such that $\psi_L(\sigma,\rho,\tau) = 0$ on a closed conic neighborhood of $\tau = 0$ ($(\sigma,\rho) \neq (0,0)$). Then $\psi_L\check{L}$ is a Fourier integral operator [**28**] with canonical relation

(13)

$$\begin{aligned}\Lambda_{\check{L}} = \bigcup_{m\in M} \Big\{ \big(x, \boldsymbol{e}^{(m)}(x,s,r), \boldsymbol{\xi}^{(m)}(x,s,r,\tau,\varepsilon), \varepsilon; s, r,\\ T^{(m)}(x,s,r), \boldsymbol{\sigma}^{(m)}(x,s,r,\tau,\varepsilon), \boldsymbol{\rho}^{(m)}(x,s,r,\tau,\varepsilon), \tau\big)\\ \Big|\ (x,s,r)\in D^{(m)}, \varepsilon\in\mathbb{R}^{n-1}, \tau\in\mathbb{R}\setminus 0\Big\}\\ \subset T^*(X\times E)\setminus 0 \times T^*Y\setminus 0,\end{aligned}$$

where

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

with similar expressions for $\boldsymbol{\sigma}^{(m)}$ and $\boldsymbol{\rho}^{(m)}$ from $\partial_s\Phi^{(m)}$ and $\partial_r\Phi^{(m)}$.

With the choice (11) for $\boldsymbol{e}$, the following assumption is implied. However, for other choices of $\boldsymbol{e}$ it needs to be verified.

Assumption 3. Consider the mapping

$$\Xi\colon \Lambda_F \longrightarrow T^*X\setminus 0\times E, \quad \lambda(x,\alpha,\beta,\tau) \longmapsto (x,\xi,e),$$
$$\text{with } \xi = -\big(\tau/c(x)\big)(\alpha+\beta).$$

Composing this mapping with the inverse of the mentioned map $(x,\alpha,\beta)\mapsto(x,s,r)$, yields per branch m a mapping $\Xi^{(m)}$ from (x,s,r,τ) to an element of $T^*X\setminus 0\times E$. $\Xi^{(m)}$ is locally diffeomorphic, i.e.

$$\operatorname{rank} \frac{\partial(\boldsymbol{\xi}^{(m)}, \boldsymbol{e}^{(m)})}{\partial(s,r,\tau)}\bigg|_{\varepsilon=0} \text{ is maximal, at given } x \text{ and branch } m.$$

Let d be the Born modeled data in accordance with Theorem 2.1. To reveal any artifacts generated by $\check{L}$, i.e. singularities in $\check{L}d$ at positions not corresponding to an element of $\mathrm{WF}(\delta c)$, we consider the composition $\check{L}\,F$. With Assumptions 2 and 3 this composition is equal to the sum of a smooth e-family of pseudodifferential operators and, in general, a non-microlocal operator the wavefront set of which contains no elements with $\varepsilon = 0$ [**28,** Thm. 6.1]. This non-microlocal operator will be a concern in the development of single reflection wavefield continuation. The origin of contributions from $\varepsilon \neq 0$ is illustrated in Figure 3. A filter needs to be applied [**6**], [**30**] removing contributions from $|\varepsilon| \geq \varepsilon_0 > 0$: We define the GRT L as the FIO, $\mathcal{D}'(Y) \to \mathcal{D}'(X \times E)$, with canonical relation $\Lambda_L = U_{\check{L}}$ given as a *neighborhood* of $\Lambda_{\check{L}} \cap \{\varepsilon = 0\}$ in $\Lambda_{\check{L}} \subset T^*(X \times E) \setminus 0 \times T^*Y \setminus 0$.

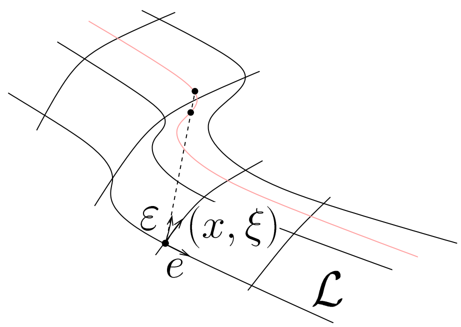

FIGURE 3: The origin of artifacts generated by the GRT. (Inside the $T^*Y \setminus 0$ box of Figure 2.) The dashed line is associated with the restriction to a fixed e.

The artifacts in the compose of canonical relations of $\check{L}$ with F can be evaluated through solving the system of equations

$$r = \boldsymbol{r}(x, \alpha), \tag{15}$$

$$s = \boldsymbol{s}(x, \beta), \tag{16}$$

$$T^{(m)}(z, s, r) = T(x, \alpha) + T(x, \beta), \tag{17}$$

$$\boldsymbol{\rho}^{(m)}(z, s, r, \tau, \varepsilon) = -\tau p_r(x, \alpha), \tag{18}$$

$$\boldsymbol{\sigma}^{(m)}(z, s, r, \tau, \varepsilon) = -\tau p_s(x, \beta). \tag{19}$$

(The frequency is preserved.) Equations (15)–(17) imply that the image point z must lie on the isochron determined by (x, s, r). Equations (18)–(19) enforce a match of slopes (apparent in the appropriate 'slant stacks') in the measurement process,

$$-\tau \partial_r T^{(m)}(z, s, r) + \langle \varepsilon, \partial_r \boldsymbol{e}^{(m)}(z, s, r) \rangle = -\tau p_r(x, \alpha), \tag{20}$$

$$-\tau \partial_s T^{(m)}(z, s, r) + \langle \varepsilon, \partial_s \boldsymbol{e}^{(m)}(z, s, r) \rangle = -\tau p_s(x, \beta). \tag{21}$$

For $\varepsilon \neq 0$ the take-off angles of the pairs of rays at (r, s) following from the right-hand sides of (15)–(19) may be distinct from those following from the left-hand sides. Equations (20)–(21) imply the matrix compatibility relation (upon eliminating ε/τ)

$$\begin{aligned} &[\partial_r \boldsymbol{e}^{(m)}(z, s, r)]^{-1}[p_r(x, \alpha) - \partial_r T^{(m)}(z, s, r)] \\ &\qquad = [\partial_s \boldsymbol{e}^{(m)}(z, s, r)]^{-1}[p_s(x, \beta) - \partial_s T^{(m)}(z, s, r)]. \end{aligned} \tag{22}$$

The geometrical composition equations determining the artifacts are solved as follows: For each $(x, \alpha, \beta) \in K$ solve the $(3n - 2)$ equations (15)–(17), (22) for the $(3n-2)$ unknowns (z, s, r). (From (20) we then obtain ε/τ, hence ε.)

The GRT reconstructs a distribution in $\mathcal{E}'(X \times E)$. We can extend the domain of the modeling operator F from $\mathcal{E}'(X)$ to $\mathcal{E}'(X \times E)$ in accordance with Theorem 7.1 [**29**]; the resulting operator is denoted by H. At $\epsilon = 0$, Λ_H reduces to Λ_F. Hence we can remodel, or what seismologists call 'de-migrate', the image Ld of data d.

4 Modeling restricted to an acquisition submanifold. Single reflection seismic wavefield continuation aims at generating from reflection data (through the canonical relation (8)) measured on an open subset of Y parameterized by an open subset of $T^*X \setminus 0 \times E$ denoted by the subscript i, reflection data on a larger open subset of Y parameterized by an open subset of $T^*X \setminus 0 \times E$ denoted by the subscript o. By abuse of notation we indicate the initial parameter subset by E_i and the final parameter subset by $E_\mathrm{o} \supset E_\mathrm{i}$. Such continuation, within the acquisition manifold Y, is accomplished through the composition of Fourier integral operators generating an intermediate image of δc. In the previous section, we analyzed a Fourier integral operator, the GRT, that generates δc from data on E_i. In this section we consider, once data are modeled from δc as in Theorem 2.1, the restriction to an acquisition submanifold. In the following sections, the restriction, modeling and GRT imaging operators will be composed to yield the continuation. In this composition, the coefficient function c_0 is used, but, naturally, δc does not appear. The continuation is illustrated in Figure 4.

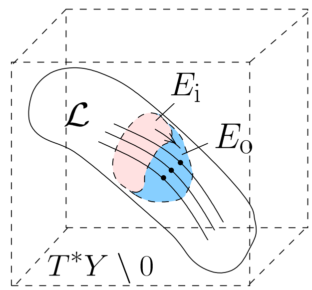

FIGURE 4: Continuation and characteristic strips. (Inside the $T^*Y \setminus 0$ box of Figure 2.)

A further restriction of the acquisition manifold Y to a submanifold $Y^c = \Sigma^c \times]0, T[$, with $\Sigma^c \xrightarrow{\imath} O_s \times O_r$ representing an embedded manifold of codimension $c \geq 0$, yields the following extension of Assumption 1. Let $(y'_1, \ldots, y'_{2n-2-c})$ denote a local coordinate system on Σ^c and let $(y'_1, \ldots, y'_{2n-2-c}, y''_{2n-1-c}, \ldots, y''_{2n-2})$ denote a local coordinate system on $O_s \times O_r$ such that Σ^c is given by $(y''_{2n-1-c}, \ldots, y''_{2n-2}) = (0, \ldots, 0)$ locally. (The coordinates on Y are completed by identifying y_{2n-1} with t:

$$(\underbrace{y'_1, \ldots, y'_{2n-2-c}}_{y'}, \underbrace{y''_{2n-1-c}, \ldots, y''_{2n-2}}_{y''}, \underbrace{y_{2n-1}}_{t}).\ )$$

Assumption 4. The projection

$$\Lambda_F \longrightarrow O_s \times O_r \setminus \Sigma^c, \quad (y', y'', t, \eta', \eta'', \tau; x, \xi) \longrightarrow y''$$

has full rank. In other words

$$\frac{\partial y''}{\partial (x, \alpha, \beta, \tau)} \text{ has maximal rank.}$$

Applying [**11,** Thm. 4.2.2] to the pair F and the restriction $\mathcal{R}^c$ from $O_s \times O_r \to \Sigma^c$ with Assumption 4 implies that $\mathcal{R}^c F$ is an FIO of order $(n+c-1)/4$

with canonical relation

$$\Lambda_F^c = \{(y', t, \eta', \tau; x, \xi) \mid \exists\, (y', y'', \eta', \eta'') \text{ such that } y'' = 0 \text{ and } (y, \eta; x, \xi) \in \Lambda_F\} \subset T^*Y^c \setminus 0 \times T^*X \setminus 0. \tag{23}$$

We will encounter two examples: Zero offset (ZO), where $c = n-1$ and $\Sigma^c := \Sigma_0 \subset \operatorname{diag}(\partial X)$ (subject to the $n-1$ constraints $r = s$ when $\arccos(\alpha\cdot\beta) = 0$ and $e_{\rm o}$ at x follows from (11)), and common azimuth (CA), where $c = 1$ and $\Sigma^c := \Sigma_A$ subject to one constraint typically of the form that the $(n-1)$-st coordinate in $r - s$ is set to zero, while $E_{\rm o} \ni e$ at x follows from the mapping $\boldsymbol{e}^{(m)}$. We set $Y_0 = \Sigma_0 \times\,]0, T[$ and $Y_A = \Sigma_A \times\,]0, T[$.

The restriction to acquisition submanifolds is placed in the context of inversion in [**25**].

5 Exploding reflector modeling. In this section we introduce a procedure to model zero-offset (ZO) data: data with coinciding sources and receivers. To ensure that the zero-offset experiment can be modeled by an FIO we invoke Assumption 4 with $\Sigma^c := \Sigma_0$. We denote its canonical relation by Λ_0.

For the zero-offset reduction to be 'image preserving', i.e. for the associated normal operator to be pseudodifferential, we mention:

Assumption 5. The projection

$$\pi_{Y_0} \colon \Lambda_0 \longrightarrow T^*Y_0 \setminus 0$$

is an embedding.

(In fact, Assumption 4 with $\Sigma^c = \Sigma_0$ implies that π_{Y_0} is an immersion.) This assumption is most easily verified by checking whether an element (y_0, η_0) in $T^*Y_0 \setminus 0$ uniquely determines an element $(x_0, \xi_0 = \partial_{x_0} T_0)$ in $T^*X \setminus 0$ smoothly; here, T_0 is the zero-offset traveltime. (In fact, Assumption 5 implies that the projection π_{Y_0} is a diffeomorphism, which coincides with Beylkin's condition [**2**].)

Remark 5.1. Assumptions 4 and 5 precisely allow the introduction of so-called map migration-demigration between the wavefront set of zero-offset data and the wavefront set of the singular medium perturbation.

In the absence of Assumption 5 we introduce the notion of the exploding reflector (ER) model in the following:

Lemma 5.2. *Let Φ_{ER} be the phase function given by $2S(x_I, x_0, \xi_J, \tau) + 2\langle \xi_J, x_J\rangle + \tau t$ (cf.* (5)*). Let A_{ER} be the symbol given by $[a(x,t,x_0,\xi_J,\tau)]^2$ (cf.* (6)*). A_{ER} and Φ_{ER} generate an oscillatory integral and define an FIO, $\delta G_0 \colon \mathcal{E}'(X) \to \mathcal{D}'\big(X \times]0,T[\ \big)$,*

$$\delta G_0(x,t,x_0) = \sum_i \int_X \int_{\mathbb{R}^{N(i)}} A^{(i)}_{\mathrm{ER}}(x,t,x_0,\theta)(-\tau^2) \\ \cdot \exp[\mathrm{i}\Phi^{(i)}_{\mathrm{ER}}(x,x_0,t,\theta)] 2c_0^{-3}(x_0)\delta c(x_0)\,\mathrm{d}\theta\,\mathrm{d}x_0.$$

Its canonical relation, $\Lambda_{0,\mathrm{ER}}$, is a scaled version of Λ_G obtained by replacing c in Hamilton system (2) *by $\frac{1}{2}c_0$.*

Proof. Φ_{ER} follows from the nondegenerate phase function ϕ associated with G upon replacing c_0 by $\frac{1}{2}c_0$, and is hence nondegenerate. The source f in (3) and (6) is replaced by $2c_0^{-3}(x_0)\delta c(x_0)$. □

Let $\mathcal{R}_r$ denote the restriction of $X \times]0,T[$ to $O_r \times]0,T[$. Let $x = (x', x'')$ denote local coordinates on X such that O_r is defined by $x'' = 0$.

Assumption 6. The intersection of $\Lambda_{0,\mathrm{ER}}$ with the manifold $Y_0 = O_r \times]0,T[$ is transversal. In other words

$$\frac{\partial x''}{\partial(x_0,\alpha_0,\tau_0)} \text{ has maximal rank.}$$

Corollary 5.3. *Subject to Assumption* 6*, the restriction $F_0 = \mathcal{R}_r \delta G_0$ is a local FIO, $F_0 = F_0[c_0] \colon \mathcal{E}'(X) \to \mathcal{D}'(Y_0)$, of order $(n-1)/2$. Its canonical relation is given by*

$$\Lambda_E = \Big\{ \Big(\boldsymbol{z}(x_0,\alpha_0), \overbrace{2T(x_0,\alpha_0)}^{T_0(x_0,\alpha_0)}, \boldsymbol{\zeta}(x_0,\alpha_0), \tau_0; x_0, -\underbrace{2\big(\tau/c(x_0)\big)\alpha_0}_{\boldsymbol{\xi}_0(x_0,\alpha_0,\tau_0)} \Big) \\ \Big|\ (x_0,\alpha_0) \in K_0, \tau_0 \in \mathbb{R}\setminus 0 \Big\} \\ \subset T^*Y_0 \setminus 0 \times T^*X \setminus 0, \tag{24}$$

where $K_0 \subset \mathbb{R}^n \times S^{n-1}$. In the notation of (8)*,*

$$\boldsymbol{z}(x_0,\alpha_0) = \boldsymbol{r}(x_0,\alpha_0), \quad \boldsymbol{\zeta}(x_0,\alpha_0) = 2\boldsymbol{\rho}(x_0,\alpha_0).$$

Proof. Let $\Lambda_{\mathcal{R}_r}$ denote the canonical relation of $\mathcal{R}_r$,

$$\Lambda_{\mathcal{R}_r} = \big\{ \big(x', t, \xi', \tau; (x', x'', t), (\xi', \xi'', \tau)\big) \\ \in T^*Y_0 \setminus 0 \times T^*(X \times]0, T[\,) \setminus 0 \;\big|\; x'' = 0 \big\}.$$

With Assumption 6 it follows that the intersection of $\Lambda_{\mathcal{R}_r} \times \Lambda_{0;\mathrm{ER}}$ with $T^*Y_0 \setminus 0 \times \mathrm{diag}\big(T^*(X \times]0, T[\,) \setminus 0\big) \times T^*X \setminus 0$ is transversal. Now apply [**11,** Thm. 4.2.2] to the pair δG_0 and the restriction $\mathcal{R}_r$. □

Note that canonical relation Λ_E is related to canonical relation $\Lambda_{\check{L}}$ by fixing the value of e in the latter in accordance with $\beta = \alpha$ (cf. (11)). Thus, in the framework of the ER model, the inverse problem is formally determined.

Remark 5.4. Subjecting the configuration to Assumption 5, the exploding reflector modeling, F_0, is, as far as the phase function is concerned, equivalent to restricting the multiple-offset modeling to zero offset, $\mathcal{R}_0 \delta G$, where $\mathcal{R}_0$ is the restriction of $X \times X \times]0, T[$ to Y_0.

6 Transformation to zero offset: Dip MoveOut. In applications, the data at zero offset is usually missing: Receivers cannot be placed on top of sources. Hence, as a first example, we analyze the continuation of multiple finite-offset seismic data to zero-offset seismic data. Dip MoveOut is the process following upon composing ER modeling with L, the imaging GRT for a neighborhood of a given value of e (conventionally for given value of offset $r - s$); the $\mathrm{sing\,supp}$ of the Lagrangian-distribution kernel of the resulting operator is what seismologists call the DMO 'impulse response'. The compose, $F_0 \check{L}$, is a well-defined operator $\mathcal{D}'(Y) \to \mathcal{D}'(Y_0)$. Its wavefront set is contained in the composition of the wavefront sets of F_0 and $\check{L}$ [**11,** Thm. 1.3.7], hence in the composition of canonical relations,

(25)
$$\Lambda_E \circ \Lambda'_L = \{(z, t_0, \zeta, \tau_0; s, r, t, \sigma, \rho, \tau) \mid \exists\, (x, \xi, \varepsilon) \text{ such that} \\ (z, t_0, \zeta, \tau_0; x, \xi) \in \Lambda_E \text{ and } (x, e, \xi, \varepsilon; s, r, t, \sigma, \rho, \tau) \in \Lambda_{\check{L}}\} \\ \subset T^*Y_0 \setminus 0 \times T^*Y \setminus 0.$$

with

$$\Lambda'_L = \{(x, \xi; s, r, t, \sigma, \rho, \tau) \mid \exists\, \varepsilon \text{ such that } (x, e, \xi, \varepsilon; s, r, t, \sigma, \rho, \tau) \in \Lambda_{\check{L}}\}.$$

Whether the compose is an FIO is yet to be investigated.

Using the parameterization of Λ_E in (24) and the parameterization of $\Lambda_{\breve{L}}$ in (13), the compose (25) can be evaluated through solving a system of equations, the first n being trivial fixing the scattering point $x_0 = x$, the second n equating the cotangent vectors

$$\underbrace{2\tau_0 \partial_x T(x_0, \alpha_0)}_{\boldsymbol{\xi}_0(x_0,\alpha_0,\tau_0)} = \underbrace{\tau \partial_x T^{(m)}(x_0, s, r) - \langle \varepsilon, \partial_x \boldsymbol{e}^{(m)}(x_0, s, r)\rangle}_{\boldsymbol{\xi}^{(m)}(x_0,s,r,\tau,\varepsilon)} . \tag{26}$$

Given x_0, these constitute n equations with the n unknowns (α_0, τ_0). Thus for each $(s, r, \tau, \varepsilon)$ we need to solve these equations.

Note that, given $e = \boldsymbol{e}^{(m)}(x_0, s, r)$, we can obtain r from s (cf. (11)). Thus we can parameterize the composition $\Lambda_E \circ \Lambda'_L$ by $(x_0, s, \tau, \varepsilon)$. We can interpret the computation of the composition as follows:

(i) Given (x_0, s) we compute r and then $T^{(m)}$;
(ii) then, given (τ, ε) we compute $\sigma = \boldsymbol{\sigma}^{(m)}$ and $\rho = \boldsymbol{\rho}^{(m)}$;
(iii) we solve (26) for (α_0, τ_0);
(iv) with these initial values, we solve the Hamiltonian flow (with (2) in the exploding reflector model) up to its intersection with the acquisition manifold Y_0, from which we deduce t_0 and z, as well as ζ.

Theorem 6.1. *With Assumptions* 2 *and* 3 *the composition $F_0 L$ yields a smooth family of FIOs parameterized by e. The compose is called Dip MoveOut. Its canonical relation is given by* (25)

$$\Lambda_D = \Lambda_E \circ \Lambda_L = \{(z, t_0, \zeta, \tau_0; s, r, t, \sigma, \rho, \tau)\}$$

parameterized by $(x_0, s, \tau, \varepsilon)$, where (s, r, t, σ, ρ) are given in (13) subject to the substitution $x = x_0$ and r is obtained from s through $\boldsymbol{e}^{(m)} = e$ which mapping is defined below equation (11), and (z, t_0, ζ) are given in Corollary 5.3 in which (α_0, τ_0) are obtained by solving (26).

Proof. First we extend the operator F_0 to act on distributions in $\mathcal{E}'(X \times E)$ by assuming that the action does not depend on $e \in E$. The calculus of FIOs gives sufficient conditions that the composition of two FIOs, here F_0 and L, is again an FIO. The essential condition is that the composition of canonical relations is transversal, i.e. that

$$\Lambda_E \times \Lambda_L \text{ and } T^*Y_0 \setminus 0 \times \operatorname{diag}\big(T^*(X \times E) \setminus 0\big) \times T^*Y \setminus 0$$

intersect transversally. We have

(27)

$$\begin{array}{ccccc} & \Lambda_E & & \Lambda_L & \\ \swarrow & & \searrow \quad \swarrow & & \searrow \\ T^*Y_0 \setminus 0 & & T^*X \setminus 0(\times E) & & T^*Y \setminus 0 \end{array}$$

where the inner two projections are submersions.

In a neighborhood of a point in $\Lambda_{\check{L}}$ given by (13), $\Lambda_{\check{L}}$ can be parameterized as in Λ'_F. Using this parameterization one finds that the composition of Λ_E and Λ_L is transversal if and only if the matrix

$$\frac{\partial}{\partial(s,r,\alpha_0,\tau,\varepsilon,\tau_0)}\big(\boldsymbol{\xi}_0(x_0,\alpha_0,\tau_0)-\boldsymbol{\xi}^{(m)}(x_0,s,r,\tau,\varepsilon)\big)$$

has maximal rank (cf. (26)). This follows, for example, just from the $\boldsymbol{\xi}_0$ contribution to this matrix. However, it follows also from the $\boldsymbol{\xi}^{(m)}$ contribution: Parameterizing Λ_L by (x,ξ,ε) and restricting Λ_L further to $\varepsilon=0$, results in a parameterization in terms of (x,ξ) (with the artifacts filtered out). Then $\boldsymbol{\xi}^{(m)}$ reduces to ξ and it follows that the composition of Λ_E and Λ_L is transversal if and only if

$$\operatorname{rank}\frac{\partial}{\partial(\xi,\alpha_0,\tau_0)}\big(\boldsymbol{\xi}_0(x_0,\alpha_0,\tau_0)-\xi\big) \text{ is maximal.}$$

This is indeed the case. $\square$

Remark 6.2. The Normal MoveOut is the relation obtained by the intersection

$$\Lambda_E \circ \big(U_L \cap \{\xi/\|\xi\| = (0,\dots,0,1)\}\big)$$

and defines a special case of the time function t_0 which is denoted by t_n. Such a relation accounts for δc with $\mathrm{WF}(\delta c) \subset X \times \{\xi \mid \xi/\|\xi\| = (0,\dots,0,1)\}$ only.

Using all the data (when available), integration over the $(n-1)$ dimensional e removes the artifacts under the Bolker condition, Assumption 2: We obtain the transformation to zero offset (TZO).

Corollary 6.3. *Let $\langle N^{-1}\rangle$ denote the regularized inverse of the normal operator in Theorem* 2.2. *With Assumptions* 1, 2 *and* 4 *(with $\Sigma^c = \Sigma_0$), the composition $F_0\langle N^{-1}\rangle F^* = \int \mathrm{d}e F_0\langle N^{-1}\rangle \check{L}$ is an FIO, $\mathcal{D}'(Y) \to \mathcal{D}'(Y_0)$. With Assumption* 5 *the reduced dataset is image preserving.*

The proof follows that of Theorem 2.2 closely (see [**29,** Thm. 4.5]).

Remark 6.4. The adjoint $(F_0 L)^*$ is by Theorem 6.1 also an FIO. This operator is called 'inverse' DMO.

7 Continuation, transformation to common azimuth.

Continuation. We analyze the 'continuation' of multiple finite-offset seismic data.

The compose $F\check{L}$ is a well-defined operator $\mathcal{D}'(Y) \to \mathcal{D}'(Y)$. Its wavefront set is contained in the composition of the wavefront sets of F and $\check{L}$ [**11,** Thm. 1.3.7], hence in the composition of canonical relations,

$$\begin{aligned}(28)\qquad \Lambda_F \circ \Lambda_L' = \{&(s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1) \mid \exists\,(x, \xi, \varepsilon) \\ &\text{such that } (s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; x, \xi) \in \Lambda_F \\ &\text{and } (x, e, \xi, \varepsilon; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1) \in \Lambda_L\} \\ \subset\, & T^*Y \setminus 0 \times T^*Y \setminus 0.\end{aligned}$$

Whether the compose is an FIO is yet to be investigated.

Using the parameterizations of Λ_F in (8) and $\Lambda_{\check{L}}$ in (13), the compose (28) can be evaluated through solving a system of equations, the first n being trivial fixing the scattering point $x_0 = x$, the second n equating the cotangent vectors

$$(29)\qquad \underbrace{\tau_2 \partial_x T(x_0, \alpha, \beta)}_{\boldsymbol{\xi}(x,\alpha,\beta,\tau_2)} = \underbrace{\tau_1 \partial_x T^{(m)}(x, s, r) - \langle \varepsilon, \partial_x \boldsymbol{e}^{(m)}(x, s, r)\rangle}_{\boldsymbol{\xi}^{(m)}(x,s,r,\tau_1,\varepsilon)}.$$

Given $\boldsymbol{e}(x, \alpha, \beta) = e$ ($n - 1$ constraints) these constitute n equations with the $2n - 1$ unknowns (α, β, τ_2). (On $D^{(m)}$ the constraints on e can be invoked on s, r instead, viz. via the inverse of the map $(x, \alpha, \beta) \mapsto (x, s, r)$ as before.)

Lemma 7.1. *With Assumptions* 2 *and* 3 *the composition FL yields a smooth family of FIOs parameterized by e. Their canonical relations are given by*

$$\Lambda_C = \Lambda_F \circ \Lambda_L = \{(s_2, r_2, t_2, \sigma_2, \rho_2, \tau_2; s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1)\}$$

parameterized by $(x_0, \alpha, s_1, \tau_1, \varepsilon)$, where upon substituting $x = x_0$ and once r_1 is obtained from s_1 through the value e of $\mathbf{e}^{(m)}$ (which mapping is defined below equation (11)*), $(s_1, r_1, t_1, \sigma_1, \rho_1)$ are given in* (13)*, and, given (α, ε), $(s_2, r_2, t_2, \sigma_2, \rho_2)$ are given in Theorem* 2.1 *in which (β, τ_2) are obtained by solving* (29)*.*

Proof. First we extend the operator F to act on distributions in $\mathcal{E}'(X \times E)$ which yields the operator H in Section 4. The calculus of FIOs gives sufficient conditions that the composition of two FIOs, here F and L, is again an FIO. The essential condition is that the composition of canonical relations is transversal, i.e. that $\mathcal{L} = \Lambda_F \times \Lambda_L$ and $\mathcal{M} = T^*Y \setminus 0 \times \operatorname{diag}\big(T^*(X \times E) \setminus 0\big) \times T^*Y \setminus 0$ intersect transversally. We have

(30)

$$\begin{tikzcd}[column sep=small]
 & \Lambda_{F(H)} \arrow[dl] \arrow[dr] & & \Lambda_L \arrow[dl] \arrow[dr] & \\
T^*Y \setminus 0 & & T^*X \setminus 0 \, (\times E) & & T^*Y \setminus 0
\end{tikzcd}$$

where the inner two projections are submersions.

On the other hand, in a neighborhood of a point in Λ_F given by (13), Λ_F can be parameterized as in Λ'_F. Using this parameterization one finds that the composition of Λ_F and Λ_L is transversal if and only if the matrix

$$\frac{\partial}{\partial(s, r, \alpha, \beta, \tau_2, \varepsilon, \tau_1)} \big(\boldsymbol{\xi}(x, \alpha, \beta, \tau_2) - \boldsymbol{\xi}^{(m)}(x, s, r, \tau_1, \varepsilon)\big)$$

has maximal rank (cf. (29)). This follows, for example, just from the $\boldsymbol{\xi}$ contribution in view of the submersivity of the projection $\pi_X : \Lambda_F \to T^*X \setminus 0$. However, it follows also from the $\boldsymbol{\xi}^{(m)}$ contribution: Parameterizing Λ_L by (x, ξ, ε) and restricting Λ_L further to $\varepsilon = 0$, results in a parameterization in terms of (x, ξ) (with the artifacts filtered out). Then $\boldsymbol{\xi}^{(m)}$ becomes ξ and it follows that the composition of Λ_F and Λ_L is transversal if and only if

$$\operatorname{rank} \frac{\partial}{\partial(\xi, \alpha, \beta, \tau_2)} \big(\boldsymbol{\xi}(x, \alpha, \beta, \tau_2) - \xi\big)t) \text{ is maximal.}$$

This is indeed the case. $\square$

Subjecting the operator F in the composition to the constraint that $\boldsymbol{e}$ (cf. (11)) attains a prescribed value, the parameter α in the lemma will be eliminated.

Remark 7.2. Following seismological convention, we have used the terminology wavefield continuation. In fact, this is continuation in the context of continuation theorems also. We consider the continuation of the wavefield in the acquisition manifold from one subset to a larger subset. This continuation is unique in the sense that $FLd = 0$ implies $F^*FLd = 0$ and, since

$F^*F = N$ is strictly elliptic and pseudodifferential, then $Ld = 0$ so that the image of δc vanishes. In the single scattering approximation this implies that $d = F\delta c = 0$, all modulo smoothing contributions.

Remark 7.3. The subject of data regularization is the transformation of measured reflection data, sampled in accordance with the actual acquisition, to data associated with a regular sampling of the acquisition manifold Y. In our approach the operator $\mathcal{R}^c F \int \mathrm{d}e \langle N^{-1}\rangle \check{L}$ replaces the forward interpolation operator in the usual regularization procedures.

Transformation to common azimuth: Azimuth MoveOut. Azimuth MoveOut is the process following composing $\mathcal{R}^1_A$ restricting Y to Y_A with modeling operator F with the imaging GRT L centered at a given value of e (conventionally for given value of offset $r - s$); the $\operatorname{sing\,supp}$ of the Lagrangian-distribution kernel of the resulting operator is what seismologists call the AMO 'impulse response'. The composition FL has been addressed in Lemma 7.1. The general restriction has been addressed in Section 4. Here we combine these results in the following:

Theorem 7.4. *With Assumptions* 2, 3 *and* 4 *with* $Y^c = Y_A$, *the composition* $\mathcal{R}^1_A FL$ *yields a smooth family of FIOs parameterized by* e. *The resulting operator is called Azimuth MoveOut.*

The following Bolker-like condition ensures that the restriction to common azimuth is 'image preserving'. Let Λ_A denote the canonical relation of $\mathcal{R}^1_A F$ in accordance with the analysis of Section 4,

Assumption 7. The projection

$$\pi_{Y_A} : \Lambda_A \longrightarrow T^*Y_A \setminus 0$$

is an embedding.

This assumption is most easily verified whether an element in $T^*Y_A \setminus 0$ uniquely determines an element in $T^*X \setminus 0$ smoothly given the medium c_0.

Using 'all' the data (when available), integration over the $(n-1)$ dimensional e removes the artifacts under the Bolker condition, Assumption 2: We obtain the transformation to common azimuth (TCA).

Corollary 7.5. *Let* $\langle N^{-1}\rangle$ *denote the regularized inverse of the normal operator in Theorem* 2.2. *With Assumptions* 1, 2 *and* 4 *(with* $\Sigma^c = \Sigma_A$*), the composition* $\mathcal{R}^1_A F\langle N^{-1}\rangle F^* = \int \mathrm{d}e \mathcal{R}^1_A F \langle N^{-1}\rangle \check{L}$ *is an FIO,* $\mathcal{D}'(Y) \to \mathcal{D}'(Y_A)$. *With Assumption* 7 *the reduced dataset is image preserving.*

The proof follows that of Theorem 2.2 closely (see [**29,** Thm. 4.5]).

8 Examples. We give a seismologists' perspective on Dip and Azimuth MoveOut. We illustrate their perspective in the constant coefficient c_0 case. This is the common case where the associated transformations are applied. In this paper, however, we have established the methodology to honor the heterogeneity in the subsurface.

The smooth background coefficient function c_0 is called the (seismic) velocity model and characterizes the speed at which waves travel through the medium. Invoking Cartesian coordinates, the acquisition manifold is obtained by setting the n-th coordinate of s and r to zero. Then O_s and O_r are open subsets of a plane hypersurface. In seismology, the midpoint in this hypersurface is defined as $y = \frac{1}{2}(s + r)$ and the offset is defined as $h = \frac{1}{2}(r - s)$. In some sense, the midpoint is associated with the direction of ξ while the offset is a particular choice for e. Here, we assume that c_0 is constant. We will illustrate both DMO and AMO, i.e. the singular supports of their respective kernels. In this section, we will highlight the transition from a parameterization including (y, h, t) to a parameterization including (s, e, t) where e relates to the scattering angles.

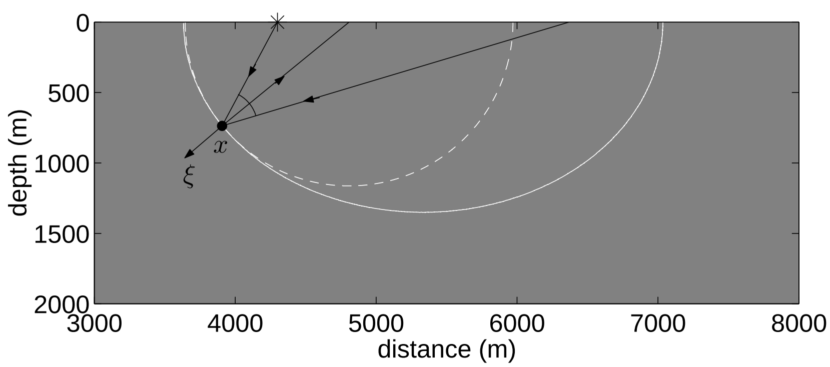

FIGURE 5: Constant velocity medium ($c_0 = 1.7$ km/s) the solid curve is the finite offset isochron, the dashed curve the zero-offset isochron, the black lines are the rays and the dot shows the location of the scattering point, which is the same as that marked by a dot in Figure 6.

Dip MoveOut. The "dip" in Dip MoveOut refers to the direction of the cotangent vector ξ in the canonical relations for modeling or imaging. Here, we illustrate DMO for $n = 2$. For the history of DMO, see [**1**], [**4**], [**10**], [**15**], [**21**], [**23**], [**24**], [**27**]. The relevant phase functions and canonical relations are derived in Appendix A, where also a parameteric representation of the impulse response is given.

We first illustrate the (transversal) composition of the canonical relations in a way familiar to seismologists. To this end, we view a canonical relation in X: For each (y, h) (or, equivalently, (s, r)) in the canonical relation Λ_F, an isochron is obtained by fixing the time t. We can, however, also view isochrons for each (s, e) instead, where e is given by (11).

The composition of canonical relations that determines the DMO canonical relation, implies the 'matching' in (x, ξ) of exploding reflector (F_0) isochrons with modeling (F) isochrons. Figure 5 illustrates this composition in the (y, h, t) parameterization; the finite-offset imaging operator maps data at (y, h, t) to the associated finite-offset isochron (white ellipse), indicated by the two arrows pointing towards the scattering point x. The exploding reflector modeling operator maps the image of the medium perturbation from the zero-offset isochron (white dashed circle) to the acquisition manifold, indicated by an arrow pointing away from the scattering point.

The analogous construction in the (s, e, t) parameterization is shown in Figure 6. Note that the shape of the finite e 'isochron' differs from the one of the finite h isochron, but that the shapes of exploding reflector model isochrons are the same.

Figure 7 shows the 'isochron' in the (s, e, t) parameterization for different values of e. All isochrons, except the exploding reflector one, have two points in common. One of these is the point at which the source ray travels for one time sample less than the full (fixed) time before the ray is scattered and returns to the acquisition surface; the other is the source point.

The impulse response of the DMO operator is the zero-offset traveltime t_0 and the distance d_0 from the source (s) to the exploding reflector source/receiver position (z) both as a function of the direction (θ_s) of the ray at the source (related to σ); all other parameters are fixed. In Figure 8 we plot these functions parameterically against one another. They are derived in Appendix A.

Azimuth MoveOut. The azimuth in Azimuth MoveOut [**3**] is the polar angle associated with the two-dimensional offset ($n = 3$) in the acquisition manifold. As the key parameter, we will employ the azimuthal angle in e rather than azimuth in h.

The composition of canonical relations that determines the AMO canonical

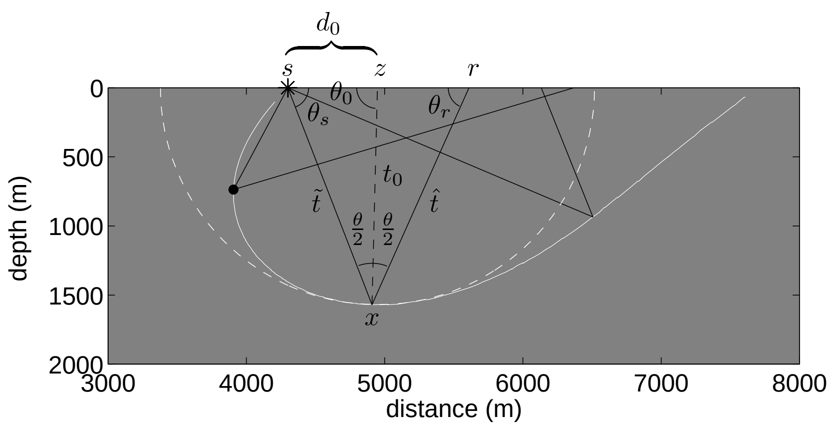

FIGURE 6: Notation for the derivation of the constant medium impulse response (cf. Appendix A). Black lines are rays, the dashed white curve is the zero-offset isochron, the solid white curve is the e 'isochron' and the black dot is the location of the scattering point for the rays shown in Figure 5.

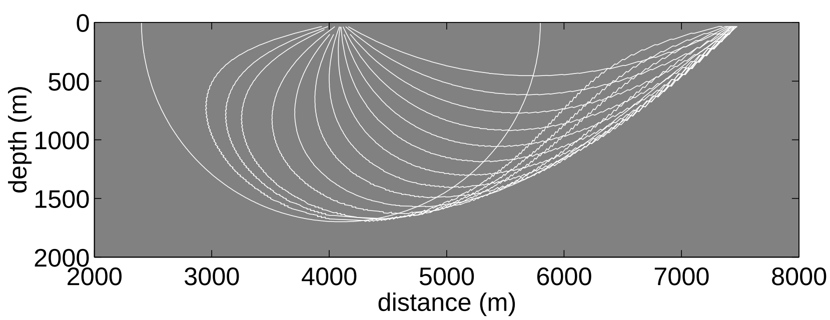

FIGURE 7: Constant velocity medium ($c_0 = 1.7$ km/s) the lines are the locations of the scattering points. Each line represents a different scattering angle, the circular line is $e = 0$, and the other lines are at increments of 0.1 radians from 0.1 radians (leftmost line) to 2.6 radians (shallowest line). All other parameters are the same as Figure 6.

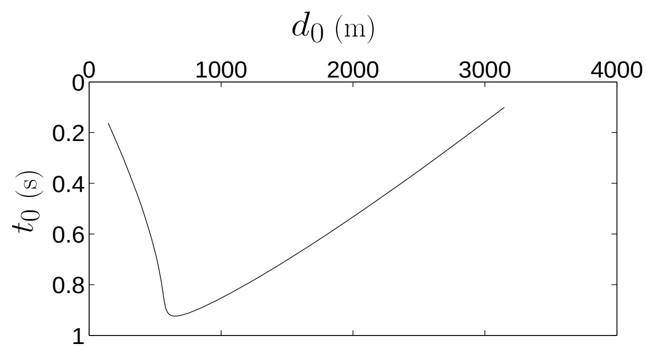

FIGURE 8: Constant velocity medium DMO impulse response, scattering angle 0.7 radians, $c_0 = 1.7$ km/s, $t = 2$ s.

relation, implies the 'matching' in (x, ξ) of two 'isochrons', one associated with the imaging operator L and one associated with the modeling operator F. The points at which these two 'isochrons' touch and share the same (co)tangent plane are the points which contribute to the canonical relation of the AMO operator.

The impulse response of the AMO operator is the traveltime t_2 as a function of the direction ((θ_s, ψ_s)) of the ray at the source (related to σ); all other parameters are fixed. In Figure 9 we plot this function. The expression is derived in Appendix B.

Appendix A. Dip MoveOut: $n = 3$ and constant coefficient. The aim of this appendix is two-fold:

(i) to show that the analysis presented in the main text encompasses the usual DMO analysis in the absence of caustics as practiced in seismology, and
(ii) to clarify the issue of number of phase variables needed in the OI representation of the DMO kernel for $n = 3$ in the constant coefficient case.

It is noted, that in the case of constant c, e can be chosen to be offset $h = \frac{1}{2}(r - s)$ in the acquisition manifold. We define an acquisition submanifold, Y', by prescribing the value of h. Throughout the analysis, in particular of the operator L, the manifold Y can be replaced by the submanifold Y' and the cotangent bundle $T^*Y \setminus 0$ by $T^*Y' \setminus 0$.

A.1 Modeling and imaging operators. In the case of a medium with constant velocity c, the generating function S in (4) is simply given by $-\tau T(x, x_0)$

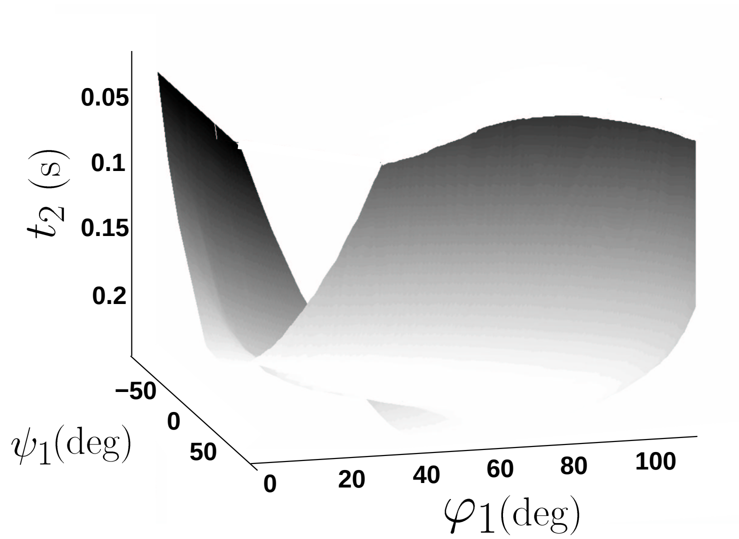

FIGURE 9: Surface plot of the AMO impulse response as a function of input ray directions.

with $T(x, x_0)$ the traveltime function along the ray connecting x with x_0, viz.

(A.1) $$T(x, x_0) = \frac{|x - x_0|}{c}.$$

Since, away from the point source, no caustics occur, the traveltime function is single valued and only one phase variable, namely τ, is required in the phase function. We will use Cartesian coordinates.

The Green's function, G, is given by the OI (cf. (6))

$$G(x, t, x_0) = \int \frac{1}{8\pi^2 |x - x_0|} \exp\big[\mathrm{i}\tau\big(t - T(x, x_0)\big)\big]\, \mathrm{d}\tau,$$

from which the modeling operator kernel of F is derived,

$$F(s_1, s_2, r_1, r_2, t, x_0) = \int \frac{-\tau^2}{16c^3\pi^3 |r - x_0|\, |s - x_0|} \cdot \exp\big[\mathrm{i}\tau\big(t - T(x_0, s_1, s_2, 0, r_1, r_2, 0)\big)\big]\, \mathrm{d}\tau,$$

in which

(A.2) $$T(x, s, r) = T(x, s) + T(x, r),$$

and the acquisition manifold, Y, is given by $(s_3, r_3) = (0, 0)$ and X is given by $(x_0)_3 > 0$. The canonical relation of F follows as

$$\begin{aligned}\Lambda_F = \big\{ \big(& \boldsymbol{s}_{1,2}(x_0,\beta) = (x_0)_{1,2} + cT(x_0,\beta)\beta_{1,2}, \boldsymbol{r}_{1,2}(x_0,\alpha) = (x_0)_{1,2} \\ & + cT(x_0,\alpha)\alpha_{1,2}, T(x_0,\beta) + T(x_0,\alpha), \boldsymbol{\sigma}_{1,2}(\beta,\tau) = \tau\beta_{1,2}/c, \\ & \boldsymbol{\rho}_{1,2}(\alpha,\tau) = \tau\alpha_{1,2}/c, \tau; x_0, \tau(\beta+\alpha)/c \big) \;\big|\; (x_0)_3 > 0, \\ & (\alpha,\beta) \in S^2 \times S^2, \alpha_3 > 0, \beta_3 > 0 \big\},\end{aligned}$$

where

$$\begin{aligned}T(x_0,\beta) &= (x_0)_3/(c\beta_3), \\ T(x_0,\alpha) &= (x_0)_3/(c\alpha_3).\end{aligned}$$

The kernel of the exploding reflector modeling operator, F_0, is given by

$$F_0(z_1,z_2,t_0,x_0) = \int \frac{-\tau_0^2}{16c^3\pi^3|z-x_0|^2} \exp\big[\mathrm{i}\tau_0\big(t - 2T(z_1,z_2,0,x_0)\big)\big]\,\mathrm{d}\tau_0,$$

with $z = (z_1, z_2, 0)$. The corresponding canonical relation follows as

$$\begin{aligned}\Lambda_E = \big\{ \big(& \boldsymbol{z}_{1,2}(x_0,\alpha) = (x_0)_{1,2} + (x_0)_3\alpha_{1,2}/\alpha_3, \\ & T_0(x_0,\alpha) = 2(x_0)_3/(\alpha_3 c), \boldsymbol{\zeta}_{1,2}(x_0,\alpha) = 2\tau_0\alpha_{1,2}/c, \tau_0; \\ & x_0, 2\tau_0\alpha/c \big) \;\big|\; (x_0)_3 > 0, \alpha \in S^2, \alpha_3 > 0 \big\}.\end{aligned}$$

In the absence of caustics, we have the freedom to follow a hybrid formulation, as we did, that encompasses replacing the exploding reflector modeling operator by an operator that takes the phase from the exploding reflector operator but the amplitude from the modeling operator F (see also [**7,** (38)]. This is justified, since in the absence of caustics we can trivialize the half-density bundle over Λ_{ER}.

A.2 The Dip MoveOut operator. We compose the exploding reflector modeling operator and the GRT to form the DMO operator. We introduce the midpoint(y)-offset(h) parameterization, i.e. $s = y - h$, $r = y + h$. In the constant coefficient case, in the absence of caustics, we can set $e = h$. Then the GRT, $L = L_U$, is replaced by the 'common-offset imaging' operator. The phase function associated with the common-offset imaging operator is simply given by $\Phi(y,t,x_0,\tau) = \tau(\frac{|x_0-y-h|}{c} + \frac{|x_0-y+h|}{c} - t)$. We choose our coordinates such that $h = (h_1, 0, 0)$.

The phase function associated with the exploding reflector modeling is $\Phi_{\mathrm{ER}}(z, t_0, x_0, \tau_0) = \tau_0(t_0 - 2\frac{|x_0 - z|}{c})$. The phase function of the DMO operator then becomes $\Psi(y, t, z, t_0, x_0, \tau, \tau_0) = \Phi(y, t, x_0, \tau) + \Phi_{\mathrm{ER}}(z, t_0, x_0, \tau_0)$. (Observe that (x_0, τ, τ_0) are the phase variables.)

Theorem A.1. *Ψ is a non-degenerate phase function. The composition of Λ_E and Λ_L is transversal.*

Proof. The partial derivatives of Ψ with respect to the phase variables are given by

$$\frac{\partial \Psi}{\partial (x_0)_i} = -\frac{2\tau_0}{c} \frac{((x_0)_i - z_i)}{|x_0 - z|} + \frac{\tau}{c}\left(\frac{((x_0)_i - y_i - h_i)}{|x_0 - y - h|} + \frac{((x_0)_i - y_i + h_i)}{|x_0 - y + h|} \right), \tag{A.3}$$

$$\frac{\partial \Psi}{\partial \tau} = -t + \frac{1}{c}\big(|x_0 - y - h| + |x_0 - y + h|\big), \tag{A.4}$$

$$\frac{\partial \Psi}{\partial \tau_0} = t_0 - \frac{2|x_0 - z|}{c}, \tag{A.5}$$

$i = 1, 2, 3$. The form of the differentials with respect to all the variables is:

		$\mathrm{d}(\partial_{(x_0)_i}\Psi)$ $(i = 1, 2, 3)$	$\mathrm{d}(\partial_\tau \Psi)$	$\mathrm{d}(\partial_{\tau_0}\Psi)$
(i)	y_j $(j = 1, 2)$	$*$	$*$	$*$
(ii)	t	$\boxed{0}$	$\boxed{-1}$	$\boxed{0}$
(iii)	z_j $(j = 1, 2)$	$-\frac{2\tau_0}{c}\Big(-\frac{\delta_{ij}}{\lvert x_0 - z\rvert} + \frac{((x_0)_i - z_i)((x_0)_j - z_j)}{\lvert x_0 - z\rvert^3}\Big)$	$*$	$*$
(iv)	t_0	$\boxed{0}$	$\boxed{0}$	$\boxed{1}$
(v)	$(x_0)_j$ $(j = 1, 2, 3)$	$*$	$*$	$*$
(vi)	τ	$*$	0	0
(vii)	τ_0	$-\frac{2}{c}\frac{((x_0)_i - z_i)}{\lvert x_0 - z\rvert}$	0	0

Because of the entries related to t and t_0 [rows (ii) and (iv)] the rank of the matrix is $2 + \operatorname{rank}\big(\mathrm{d}(\frac{\partial \Psi}{\partial x_0})\big)$. Now, $\frac{c}{2\tau_0} \times (iii)_j - \frac{c}{2}\frac{((x_0)_j - z_j)}{|x_0 - z|^2} \times (vii)$, for $j = 1, 2$, yields the form of rows (iii) and (vii):

$$\begin{array}{c|ccc} (iii)_1 & \frac{1}{|x_0 - z|} & 0 & 0 \\ (iii)_2 & 0 & \frac{1}{|x_0 - z|} & 0 \\ (vii) & * & * & -\frac{2}{c}\frac{((x_0)_3 - z_3)}{|x_0 - z|} \end{array}.$$

Since $|x_0 - z| > 0$, $(0 = z_3 < (x_0)_3)$, $\mathrm{rank}\big(\mathrm{d}(\frac{\partial\Psi}{\partial x_0})\big) = 3$. The rank of the differentials is therefore maximal; the phase is nondegenerate. It follows that the composition of the two canonical relations is transversal [**18,** Thm. 21.2.19]. □

Parameterization of the canonical relation. As already mentioned above in the case $e = h$ one can restrict the DMO operation to a constant offset one, replacing Y by Y'. Comparing to the main text, observe that one does not need the cotangent variable ε to parameterize the canonical relation: we make use of only s, x_0, τ here. In this case Λ_D follows as

$$\Lambda_D = \Big\{ \big(\boldsymbol{z}(s,x_0), \boldsymbol{t}_0(s,x_0), \boldsymbol{\zeta}(s,x_0,\tau), \boldsymbol{\tau}_0(s,x_0,\tau); \\ \boldsymbol{y}(s), \boldsymbol{t}(s,x_0), \boldsymbol{\eta}(s,x_0,\tau), \tau\big) \Big\}.$$

The midpoint y is $y = s + h$, the receiver location is $r = s + 2h$, and the traveltime is $t = (|x_0 - s| + |x_0 - r|)/c$. Using the phase function Ψ we immediately obtain that

$$\eta_i = \frac{\tau}{c}\left(\frac{s_i - (x_0)_i}{|x_0 - s|} + \frac{r_i - (x_0)_i}{|x_0 - r|} \right), \quad i = 1, 2.$$

We introduce ξ as

$$\xi = \partial_{x_0}\Psi = \frac{\tau}{c}\left(\frac{(x_0)_i - s_i}{|x_0 - s|} + \frac{(x_0)_i - r_i}{|x_0 - r|} \right), \quad i = 1, 2, 3.$$

Observe that, in a constant medium, we naturally have $\xi_i = -\eta_i$ for $i = 1, 2$. According to the main text, we define $\alpha_0 \in S^2$ as

$$(2\tau_0/c)\alpha_0 = \xi\ ,$$

which yields τ_0. Note that $\sin\theta_0 = (\alpha_0)_3 > 0$ (cf. Figure 6). The zero-offset travel time is then given by $t_0 = (x_0)_3/(c\sin\theta_0)$. The zero-offset source location, z, then follows as $z = ct_0\alpha_0 + x_0$.

The 'impulse response', $n = 2$. In the case $n = 2$, with a 'horizontal' acquisition manifold as in the previous subsection, we have $s = (s_1, 0)$, $r = (r_1, 0)$ for source and receiver locations. We parameterize the canonical relation of the DMO operator, Λ_D, with $(x_0, s, \tau, \varepsilon)$, where s is the source, x_0 is the scattering point (cf. Figure 6), τ the frequency, and ε the cotangent variable corresponding to e. Here, e is taken to be the scattering angle, θ,

$$\theta = \boldsymbol{\theta}(x_0, s, r) = \arccos\left(\frac{\langle x_0 - r, x_0 - s\rangle}{|x_0 - r||x_0 - s|} \right).$$

The canonical relation, Λ_D, will then be of the form

$$\Lambda_D = \big\{ \big(\boldsymbol{z}(x_0, s, \tau, \varepsilon), \boldsymbol{t}_0(x_0, s, \tau, \varepsilon), \boldsymbol{\zeta}(x_0, s, \tau, \varepsilon), \boldsymbol{\tau}_0(x_0, s, \tau, \varepsilon); \\ s, \boldsymbol{r}(x_0, s, \tau, \varepsilon), T(x_0, s, \tau, \varepsilon), \boldsymbol{\sigma}(x_0, s, \tau, \varepsilon), \boldsymbol{\rho}(x_0, s, \tau, \varepsilon), \tau\big)\big\}.$$

The zero-offset case corresponds to $\theta = 0$, which we exclude in the neighborhood E_{i} (cf. Figure 4). We make use of only one connected component of E_{i} and thus assume that $\theta > 0$ in E_{i}.

To determine the cotangent variables, σ, ρ, ζ and τ_0, we will make use of the derivatives

$$\partial_{(x_0)_i}\theta = -\frac{1}{\sin\theta}\left\{\left(\frac{\delta_{ij}}{|x_0 - r|} - \frac{(x_0)_j - r_j}{|x_0 - r|^3}\right)\left(\frac{(x_0)_j - s_j}{|x_0 - s|}\right) \right.\\ \left. + \left(\frac{\delta_{ij}}{|x_0 - s|} - \frac{(x_0)_j - s_j}{|x_0 - s|^3}\right)\left(\frac{(x_0)_j - r_j}{|x_0 - r|}\right)\right\},$$

$$\partial_{s_1}\theta = \frac{1}{\sin\theta}\left(\frac{\delta_{1j}}{|x_0 - s|} + \frac{(x_0)_j - s_j}{|x_0 - s|^3}\right)\left(\frac{(x_0)_j - r_j}{|x_0 - r|}\right),$$

and a similar expression for $\partial_{r_1}\theta$. From (x_0, s) we determine the direction of the ray at the source,

$$-\beta = \frac{s - x_0}{|s - x_0|}, \tag{A.6}$$

and the traveltime

$$\tilde{t} = \frac{|s - x_0|}{c}. \tag{A.7}$$

The angle θ_s is defined through

$$\beta = (\cos\theta_s, \sin\theta_s). \tag{A.8}$$

Using the relation $\theta_r = \theta_s + \theta$ (cf. Figure 6) we find the angle θ_r which defines the direction of the ray at the receiver

$$-\alpha = -(\cos\theta_r, \sin\theta_r). \tag{A.9}$$

The receiver ray traveltime then follows from

$$\sin\theta_r c\hat{t} = \frac{(x_0)_3}{c}. \tag{A.10}$$

Then the receiver position is found to be

$$r = -\hat{t}c\alpha + x_0. \tag{A.11}$$

The total traveltime is simply given by

$$T = \tilde{t} + \hat{t}. \tag{A.12}$$

The cotangent variables σ and ρ are then given by

$$\sigma = -\frac{\tau}{c}\cos\theta_s - \varepsilon\partial_{s_1}\theta(x_0, s, r), \tag{A.13}$$

$$\rho = -\frac{\tau}{c}\cos\theta_r - \varepsilon\partial_{r_1}\theta(x_0, s, r). \tag{A.14}$$

We determine (α_0, τ_0) from the equality (cf. (26))

$$\frac{\tau}{c}(\alpha + \beta) - \varepsilon\partial_{x_0}\theta(x_0, s, r) = \xi = 2\left(\frac{\tau_0}{c}\right)\alpha_0. \tag{A.15}$$

The zero-offset traveltime, t_0, thus follows as

$$t_0 = (x_0)_3/(c\sin\theta_0),$$

with $\sin\theta_0 = (\alpha_0)_3$. The zero-offset source position, z, is then given by

$$z = ct_0\alpha_0 + x_0,$$

while

$$\zeta = -\xi_1$$

(cf. (A.15)).

For Born modeled data the only contribution comes from $\varepsilon = 0$, in which case these formulae simplify to

$$\sigma = -\frac{\tau}{c}\cos\theta_s, \quad \rho = -\frac{\tau}{c}\cos\theta_r,$$
$$\alpha_0 = (\alpha + \beta)/(|\alpha + \beta|),$$
$$t_0 = (x_0)_3/(c\sin\theta_0) = T\sin\theta_r\sin\theta_s/[\sin\theta_0(\sin\theta_r + \sin\theta_s)].$$

The distance between the source location, s, and the zero-offset source location, z, is given by

$$d_0 = z - s = c\big(\tilde{t}\beta_1 - t_0(\alpha_0)_1\big) = cT \sin\theta_r \sin(\theta/2)/[(\sin\theta_s + \sin\theta_r)\sin\theta_0].$$

A.3 Parameterization of the canonical relation by a phase function. In the case $e = h$ discussed above, it is possible that the number of phase variables used in Ψ, here, (x_0, τ, τ_0), is unnecessarily large. Since the canonical relation and the stationary point set are locally diffeomorphic, we can investigate this question on the stationary point set, $S_\Psi = \{(y, t, z, t_0, x_0, \tau, \tau_0) \mid \partial_{x_0}\Psi = 0, \partial_\tau \Psi = 0, \partial_{\tau_0}\Psi = 0\}$.

Minimum number of phase variables. Let us first project S_Ψ onto the natural base coordinates (y, t, z, t_0). Let $\pi\colon (y, t, z, t_0, x_0, \tau, \tau_0) \mapsto (y, t, z, t_0)$, then

$$\begin{aligned}\operatorname{rank}(D\pi|_{S_\Psi}) &= \dim\{(y, t, z, t_0)\}\\ &\qquad + \operatorname{rank}(\partial_{x_0} f, \partial_\tau f, \partial_{\tau_0} f) - \dim\{(x_0, \tau, \tau_0)\},\end{aligned}$$

where $f = 0$ is the defining equation for S_Ψ, i.e. $f = (\partial_{x_0}\Psi, \partial_\tau\Psi, \partial_{\tau_0}\Psi)$. Then

$$\operatorname{corank}(D\pi|_{S_\Psi}) = \dim\{(x_0, \tau, \tau_0)\} - \operatorname{rank}(\partial_{x_0} f, \partial_\tau f, \partial_{\tau_0} f)$$

is the minimal number of phase variables required to characterize the canonical relation.

Corollary A.2. *The minimum number of phase variables that locally parameterizes $\Lambda_D = \Lambda_E \circ \Lambda_L$ is 2.*

Proof. The structure of the differentials of f with respect to (x_0, τ, τ_0) is:

	(a) $\mathrm{d}(\partial_{(x_0)_j}\Psi)$ $(j = 1, 2, 3)$	(b) $\mathrm{d}(\partial_\tau \Psi)$	(c) $\mathrm{d}(\partial_{\tau_0}\Psi)$
(i) $(x_0)_i$	$(a)_{(i)}$	$(b)_{(i)}$	$\frac{2}{c}\frac{\big((x_0)_i - z_i\big)}{\lvert x_0 - z\rvert}$
(ii) τ	$(a)_{(ii)}$	0	0
(iii) τ_0	$-\frac{2}{c}\frac{\big((x_0)_j - z_j\big)}{\lvert x_0 - z\rvert}$	0	0

with $i = 1, 2, 3$, and where

$$\begin{aligned}(a)_{(i)} = &-\frac{2\tau_0}{c}\left(\frac{\delta_{ij}}{|x_0 - z|} - \frac{\big((x_0)_i - z_i\big)\big((x_0)_j - z_j\big)}{|x_0 - z|^3}\right)\\ &+ \frac{\tau}{c}\left(\frac{\delta_{ij}}{|x_0 - y - h|} - \frac{\big((x_0)_i - y_i - h_i\big)\big((x_0)_j - y_j - h_j\big)}{|x_0 - y - h|^3}\right)\\ &+ \frac{\tau}{c}\left(\frac{\delta_{ij}}{|x_0 - y + h|} - \frac{\big((x_0)_i - y_i + h_i\big)\big((x_0)_j - y_j + h_j\big)}{|x_0 - y + h|^3}\right),\end{aligned}$$

$$(a)_{(ii)} = \frac{1}{c}\left(\frac{\big((x_0)_j - y_j - h_j\big)}{|x_0 - y - h|} + \frac{\big((x_0)_j - y_j + h_j\big)}{|x_0 - y + h|}\right), \text{ and}$$

$$(b)_{(i)} = \frac{1}{c}\left(\frac{\big((x_0)_i - y_i - h_i\big)}{|x_0 - y - h|} + \frac{\big((x_0)_i - y_i + h_i\big)}{|x_0 - y + h|}\right).$$

On the stationary point set, S_Ψ, in view of equation (A.3), rows (ii) and (iii) are linearly dependent and the same holds for columns (b) and (c). Equation (A.3) for $(x_0)_3$ ($(x_0)_3 \neq 0$) gives

$$-\frac{2\tau_0}{c}\frac{1}{|x_0 - z|} + \frac{\tau}{c}\left(\frac{1}{|x_0 - y - h|} + \frac{1}{|x_0 - y + h|}\right) = 0$$

and hence a simplification of the upper left 3×3 matrix. Therefore $\operatorname{rank}(\partial_{x_0} f, \partial_\tau f, \partial_{\tau_0} f)$ is that of:

$$\begin{array}{c|cc} & (a) & (b) \\ \hline (i) & \begin{array}{l} 2\tau_0 \frac{((x_0)_i - z_i)((x_0)_j - z_j)}{|x_0 - z|^3} \\ \quad -\tau \frac{((x_0)_i - y_i - h_i)((x_0)_j - y_j - h_j)}{|x_0 - y - h|^3} \\ \quad -\tau \frac{((x_0)_i - y_i + h_i)((x_0)_j - y_j + h_j)}{|x_0 - y + h|^3} \end{array} & \big((x_0)_i - z_i\big) \\ (ii) & \big((x_0)_j - z_j\big) & 0 \end{array}$$

with $i, j = 1, 2, 3$. By subtracting

$$\left[2\tau_0 \frac{\big((x_0)_j - z_j\big)}{|x_0 - z|^3} - \tau\left(\frac{\big((x_0)_j - y_j - h_j\big)}{|x_0 - y - h|^3} + \frac{\big((x_0)_j - y_j + h_j\big)}{|x_0 - y + h|^3}\right)\right] \times (b)$$

from $(a)_j$, $j = 1, 2, 3$, and using the fact that $h_2 = h_3 = 0$, $z_3 = y_3 = 0$, and $z_2 = y_2$ on S_Φ, the matrix further simplifies to

$$\begin{array}{c|cc} & (a) & (b) \\ \hline (i)_1 & -\tau\frac{(z_1-y_1-h_1)((x_0)_j-y_j-h_j)}{|x_0-y-h|^3} - \tau\frac{(z_1-y_1+h_1)((x_0)_j-y_j+h_j)}{|x_0-y+h|^3} & \big((x_0)_1 - z_1\big) \\ (i)_2 & 0 & \big((x_0)_2 - y_2\big) \\ (i)_3 & 0 & (x_0)_3 \\ & & \\ (ii) & \big((x_0)_j - z_j\big) & 0 \end{array}$$

Performing a similar operation with rows instead yields that the rank is that of $((x_0)_3 > 0)$

$$\begin{array}{c|ccc} & (a)_1 & (a)_3 & (b) \\ \hline (i)_1 & -\tau\frac{(z_1-y_1-h_1)^2}{|x_0-y-h|^3} - \tau\frac{(z_1-y_1+h_1)^2}{|x_0-y+h|^3} & 0 & \big((x_0)_1 - z_1\big) \\ (i)_3 & 0 & 0 & (x_0)_3 \\ & & & \\ (ii) & \big((x_0)_1 - z_1\big) & (x_0)_3 & 0 \end{array} \tag{A.16}$$

the determinant of which, $\tau(x_0)_3^2(\frac{(z_1-y_1-h_1)^2}{|x_0-y-h|^3} + \frac{(z_1-y_1+h_1)^2}{|x_0-y+h|^3})$, is not zero if $h_1 \neq 0$. We conclude that $\operatorname{rank}(\partial_{x_0} f, \partial_\tau f, \partial_{\tau_0} f) = 3$. $\square$

Choice of phase variables. A similar argument as that of the proof of Corollary A.2, shows that the corank of the projection on $\big(y_1, y_2, t, z_1, (x_0)_2, \tau\big)$ is 0. It makes use of the fact that $h_1 \neq 0$, $(x_0)_3 \neq 0$ as well as that $(x_0)_1 > y_1 \Leftrightarrow (x_0)_1 > z_1$ on S_Ψ. We can therefore use $(x_0)_2$ and τ as phase variables or $\alpha_2 = \big(z_2 - (x_0)_2\big)/|z - x_0|$ and τ. Figure A.1 illustrates the need for a second phase variable in addition to τ.

The amplitude. We apply the stationary phase formula to achieve the parameterization of the OI representation of the DMO kernel with only $(x_0)_2$ and τ as phase variables. We compute the Hessian of Ψ/τ with respect to $(x_0)_1$, $(x_0)_3$, τ_0, which does not vanish because of the previous remarks. Its evaluation at stationarity yields

$$\begin{aligned} H &= \left|\frac{\partial^2\Psi/\tau}{\partial^2\big((x_0)_1, (x_0)_3, \tau_0\big)}\right| \\ &= \frac{4}{\tau^2 c^3}\frac{(x_0)_3^2}{|x_0 - z|^2}\left(\frac{(z_1 - y_1 - h_1)^2}{|x_0 - y - h|^3} + \frac{(z_1 - y_1 + h_1)^2}{|x_0 - y + h|^3}\right). \end{aligned}$$

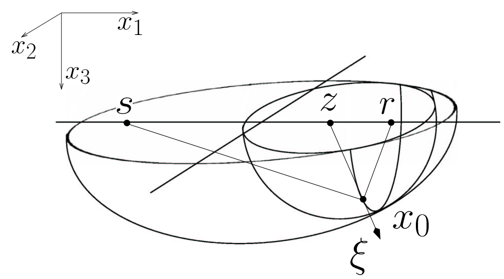

FIGURE A.1: Isochrones for DMO, $n = 3$ constant coefficient.

We can relate this Hessian geometrically to the curvatures of the finite-offset and zero-offset isochrons. In this form, the Hessian simplifies, using the law of sines, to

$$H = \frac{-2(x_0)_3^2}{\tau^3 c^2}(\kappa_O - \kappa_I)K\frac{1}{h_1}\left(\frac{1}{\sin(\theta_r)} + \frac{1}{\sin(\theta_s)}\right)\frac{1}{\sin(\theta/2)}\frac{1}{\sin^2(\theta_0)},$$

where the κ_O is the curvature of the zero-offset isochron, κ_I is the curvature of the constant-offset isochron, and $K = -\frac{2\tau\sin^2(\theta/2)\cos(\theta/2}{c|x-z|}$ as defined in [**5,** (7.6.10)–(7.6.13)].

The signature of the second order differential $\frac{\partial^2\Psi/\tau}{\partial^2((x_0)_1,(x_0)_3,\tau_0)}$ is constant. It is easy to compute it at a point x_0 half-way between source and receiver. The signature is then -1.

At stationarity the phase function simplifies to

$$\tau\left(\frac{|x_0 - y - h|}{c} + \frac{|x_0 - y + h|}{c} - t\right),$$

and the entire amplitude of the associated oscillatory integral representation becomes

$$\frac{\tau^{1/2}\tau_0^2}{32c^6(2\pi)^{3/2}|z - x_0|^2\,|y + h - x_0|\,|y - h - x_0|}\exp(-\mathrm{i}\pi/4)\frac{1}{|H|^{1/2}}.$$

Appendix B. Azimuth MoveOut: $n = 3$ and constant coefficient. In a constant velocity medium, it is possible to derive an expression for the impulse

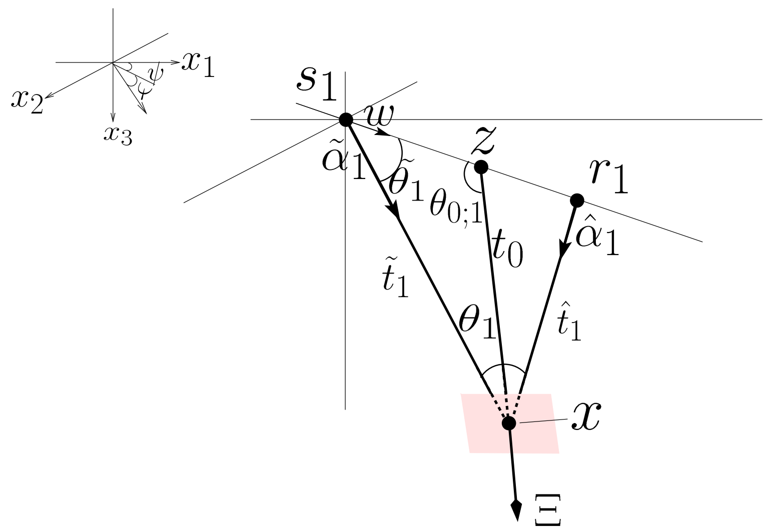

FIGURE B.1: DMO geometry and notation related to $(s_1, r_1, t_1, \sigma_1, \rho_1, \tau_1)$.

response in closed form. In [**3**] the impulse response of AMO was derived as the time t_2 as a function of translation in midpoint location, $\frac{1}{2}(r_2 + s_2) - \frac{1}{2}(r_1 + s_1)$ for given offsets $\frac{1}{2}(r_1 - s_1)$ and $\frac{1}{2}(r_2 - s_2)$. Here, we determine time t_2 as a function of ray direction at s_1 (associated with σ_1) for given e_1 and (subsurface) scattering angle from e_2 and (acquisition surface) azimuth direction, i.e. direction of $\frac{1}{2}(r_2 - s_2)$.

We derive the impulse response in two steps. First, we determine the three-dimensional DMO zero-offset traveltime (t_0) from (s_1, r_1, t_1), and then we determine the AMO time (t_2) by performing inverse DMO from the zero-offset ray to (s_2, r_2, t_2). Since the zero-offset ray will always be in the plane defined by the source and receiver rays, we need only compute the scattering angle in that plane, as a function of the initial ray angles at the source s_1. We then apply the DMO formula derived in Appendix A.

In Figure B.1 we introduce the unit vectors

$$\tilde{\alpha}_1 = \big(\cos(\varphi_1)\cos(\psi_1), \cos(\varphi_1)\sin(\psi_1), \sin(\varphi_1)\big), \tag{B.1}$$

$$\Xi = \big(\cos(\varphi_2)\cos(\psi_2), \cos(\varphi_2)\sin(\psi_2), \sin(\varphi_2)\big). \tag{B.2}$$

($\tilde{\alpha}_1$ determines σ_1 and Ξ determines ξ.) We observe that $w = \tilde{\alpha}_1 - \lambda\Xi$, while

w lies in the $x_3 = 0$ plane. We evaluate λ by setting $w_3 = 0$,

$$\lambda = \frac{\sin(\varphi_1)}{\sin(\varphi_2)}. \tag{B.3}$$

Then

$$w = \begin{pmatrix} \cos(\varphi_1)\cos(\psi_1) - \frac{\sin(\varphi_1)}{\sin(\varphi_2)}\cos(\varphi_2)\cos(\psi_2) \\ \cos(\varphi_1)\sin(\psi_1) - \frac{\sin(\varphi_1)}{\sin(\varphi_2)}\cos(\varphi_2)\sin(\psi_2) \\ 0 \end{pmatrix} \tag{B.4}$$

with

$$\|w\|^2 = \cos^2(\varphi_1) + \frac{\sin^2(\varphi_1)}{\tan^2(\varphi_2)} - \frac{\sin(2\varphi_1)}{\tan(\varphi_2)}\cos(\psi_1 - \psi_2). \tag{B.5}$$

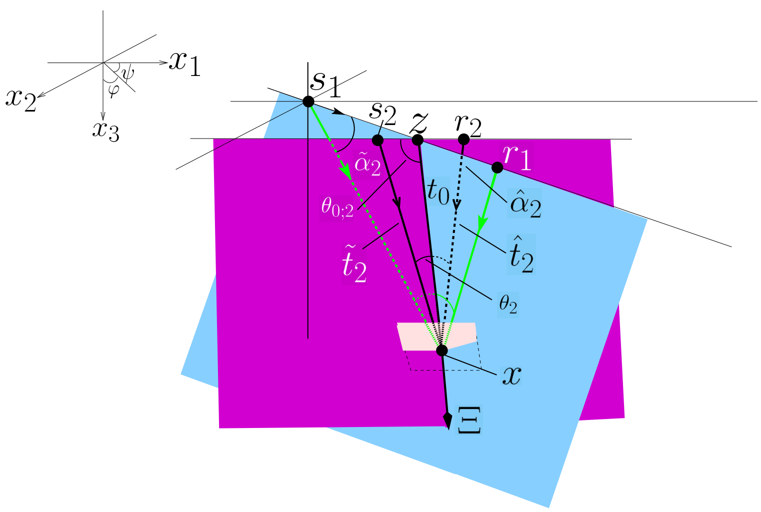

FIGURE B.2: Rotation to set azimuth showing the notation for the output (black) rays. The input rays are shown in gray. The plane which contains the output rays (dark gray) is the result of rotating the plane which contains the input rays (light gray) about the Ξ vector.

The angle $\tilde{\theta}_1$ is defined in Figure B.1 and is given by

$$\tilde{\theta}_1 = \operatorname{acos}\left(\frac{\tilde{\alpha}_1 \cdot w}{\|w\|}\right). \tag{B.6}$$

With $\tilde{\theta}_1$ we derive the zero-offset time t_0 using (A.2),

$$t_0 = \frac{t_1 \sin(\tilde{\theta}_1)\sin(\hat{\theta}_1)}{\big(\sin(\tilde{\theta}_1) + \sin(\hat{\theta}_1)\big)\sin(\theta_{0;1})}. \tag{B.7}$$

We now rotate this DMO ray geometry about the Ξ axis (Figure B.2)) to obtain the desired azimuthal orientation. We have chosen our coordinates such that this orientation coincides with the 1 axis, which implies that $w := (1,0,0)$. We determine the time t_2 by applying inverse DMO to the rotated geometry. Thus,

$$\theta_{0;2} = \operatorname{acos}\big(\Xi \cdot (1,0,0)\big) = \operatorname{acos}\big(\cos(\varphi_2)\cos(\psi_2)\big), \tag{B.8}$$

and it follows that

$$t_2 = \tilde{t}_2 + \hat{t}_2 = \frac{t_0\big(\sin(\tilde{\theta}_2) + \sin(\hat{\theta}_2)\big)\sin(\theta_{0;2})}{\sin(\tilde{\theta}_2)\sin(\hat{\theta}_2)}. \tag{B.9}$$

REFERENCES

1. C. Artley and D. Hale, *Dip-moveout processing for depth-variable velocity*, Geophysics **59** (1994), 610–622.
2. G. Beylkin, *Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform*, J. Math. Phys. (1) **26** (1985), 99–108.
3. B. Biondi, S. Fomel and N. Chemingui, *Azimuth moveout for 3-D prestack imaging*, Geophysics **63** (1998), 574–588.
4. J. L. Black, K. L. Schleicher and L. Zhang, *True-amplitude imaging and dip moveout*, Geophysics **58** (1993), 47–66.
5. N. Bleistein, J. K. Cohen and J. W. Stockwell, *Mathematics of multidimensional seismic imaging, migration, and inversion*, Springer-Verlag, New York, 2001.
6. S. Brandsberg-Dahl, M. V. de Hoop and B. Ursin, *Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers*, Geophysics **68** (2003), 232–254.
7. G. Cheng and S. Coen, *The relationship between Born inversion and migration for common-midpoint stacked data*, Geophysics **49** (1984), 2117–2131.
8. M. V. de Hoop, R. Burridge, C. Spencer and D. Miller, *Generalized Radon transform/amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media*, SPIE (San Diego, CA) **2301**, 1994.
9. M. V. de Hoop, C. Spencer and R. Burridge, *The resolving power of seismic amplitude data: An anisotropic inversion/migration approach*, Geophysics **64** (1999), 852–873.

10. S. G. Deregowski and F. Rocca, *Geometrical optics and wave theory of constant offset sections in layered media*, Geophys. Prospect. **29** (1981), 374–406.
11. J. J. Duistermaat, *Fourier Integral Operators*, Birkhäuser, Boston, 1996.
12. S. Fomel, *Amplitude preserving offset continuation in theory, Part 1: the offset continuation equation*, Stanford Exploration Project, preprint **SEP-84** (1995), 179–196.
13. S. Goldin, *Superposition and continuation of transformations used in seismic migration*, Russian Geology and Geophysics **35** (1994), 131–145.
14. V. Guillemin, *On some results of Gel'fand in integral geometry*, in *Pseudodifferential operators and applications* (Notre Dame, Ind., 1984), Amer. Math. Soc., Providence, RI, 1985, 149–155.
15. D. Hale (ed.), *DMO processing*, Geophysics Reprint Series **16**, Soc. Explor. Geophys., 1995.
16. S. Hansen, *Solution of a hyperbolic inverse problem by linearization*, Comm. Partial Differential Equations **16** (1991), 291–309.
17. L. Hörmander, *The Analysis of Linear Partial Differential Operators*, vol. **1**, Springer-Verlag, Berlin, 1983.
18. ———, *The Analysis of Linear Partial Differential Operators*, vol. **2**, Springer-Verlag, Berlin, 1985.
19. ———, *The Analysis of Linear Partial Differential Operators*, vol. **4**, Springer-Verlag, Berlin, 1985.
20. E. Iversen, H. Gjøystdal, and J. O. Hansen, *Prestack map migration as an engine for parameter estimation in TI media*, 70th Ann. Mtg. Soc. Explor. Geoph., Expanded Abstracts, 2000, 1004–1007.
21. H. Jakubowicz, *A simple efficient method of dip-moveout correction*, Geophys. Prospect. **38** (1990), 221–245.
22. A. P. E. ten Kroode, D. J. Smit and A. R. Verdel, *A microlocal analysis of migration*, Wave Motion **28** (1998), 149–172.
23. C. L. Liner, *Born theory of wave-equation dip moveout*, Geophysics **56** (1989), 182–189.
24. D. Miller and R. Burridge, *Multiparameter inversion, dip-moveout, and the generalized Radon transform*, in *Geophysical inversion* (J. B. Bednar, L. R. Lines, R. H. Stolt and A. B. Weglein, eds.), SIAM, 46–58, 1992.
25. C. J. Nolan and W. W. Symes, *Global solution of a linearized inverse problem for the wave equation*, Comm. Partial Differential Equations **22** (1997), 919–952.
26. Rakesh, *A linearized inverse problem for the wave equation*, Comm. Partial Differential Equations **13** (1988), 573–601.
27. V. Sorin and S. Ronen, *Ray-geometrical analysis of dip moveout amplitude distribution*, Geophysics **54** (1989), 1333–1335.
28. C. C. Stolk, *Microlocal analysis of the scattering angle transform*, preprint, Dept. of Computational and Applied Mathematics, Rice University, http://www.caam.rice.edu/~cstolk/angle.ps, July 2001.
29. C. C. Stolk and M. V. de Hoop, *Microlocal analysis of seismic inverse scattering in anisotropic elastic media*, Comm. Pure Appl. Math. (3) **55** (2002), 261-301.
30. ———, *Seismic inverse scattering in the 'wave-equation' approach*, MSRI #**2001–047**, 2001.
31. M. E. Taylor, *Reflection of singularities of solutions to systems of differential equations*, Comm. Pure Appl. Math. **28** (1975), 457–478.
32. F. Treves, *Introduction to Pseudodifferential and Fourier Integral Operators*, vol. **2**, Plenum Press, New York, 1980.

Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401-1887, USA
E-mail address: mdehoop@Mines.EDU

Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401-1887, USA
E-mail address: amalcolm@dix.Mines.EDU

Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401-1887, USA
Current address:
Laboratoire d'Analyse, Topologie et Probabilités, CNRS UMR6632, Centre de Mathématiques et Informatique, Université de Provence, 13453 Marseille cedex 13, France
E-mail address: jlerous@cmi.univ-mrs.fr