Seismic imaging with the generalized Radon transform: A curvelet transform perspective‡

M V de Hoop[1], H Smith[2], G Uhlmann[2] and R D van der Hilst[3]

[1] Center for Computational and Applied Mathematics, and Geo-Mathematical Imaging Group, Purdue University, West Lafayette IN 47907, USA
[2] Department of Mathematics, University of Washington, Seattle WA 98195-4350, USA
[3] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge MA 02139, USA

E-mail: mdehoop@purdue.edu

Abstract. A key challenge in the seismic imaging of reflectors using surface reflection data is the subsurface illumination produced by a given data set and for a given complexity of the background model (of wavespeeds). The imaging is described here by the generalized Radon transform. To address the illumination challenge and enable (accurate) local parameter estimation, we develop a method for partial reconstruction. We make use of the curvelet transform, the structure of the associated matrix representation of the generalized Radon transform, which needs to be extended in the presence of caustics, and phase-linearization. We pair an image target with partial waveform reflection data, and develop a way to solve the matrix normal equations that connect their curvelet coefficients via diagonal approximation. Moreover, we develop an approximation, reminiscent of Gaussian beams, for the computation of the generalized Radon transform matrix elements only making use of multiplications and convolutions, given the underlying ray geometry; this leads to computational efficiency. Throughout, we exploit the (wavenumber) multi-scale features of the dyadic parabolic decomposition underlying the curvelet transform and establish approximations that are accurate for sufficiently fine scales. The analysis we develop here has its roots in and represents a unified framework for (double) beamforming and beam-stack imaging, parsimonious pre-stack Kirchhoff migration, pre-stack plane-wave (Kirchhoff) migration, and delayed-shot pre-stack migration.

PACS numbers: 43.40.Ph, 43.60.Fg, 91.30.Ab

Submitted to: *Inverse Problems*

‡ This research was supported in part under NSF CMG grant DMS-0724808. MVdH was also funded in part by START-grant Y237-N13 of the Austrian Science Fund. HS was supported by NSF grant DMS-0654415. GU was also funded in part by a Walker Family Endowed Professorship.

1. Introduction

1.1. Seismic Imaging with Arrays – Beyond Current Capabilities

Much research in modern, quantitative seismology is motivated – on the one hand – by the need to understand subsurface structures and processes on a wide range of length scales, and – on the other hand – by the availability of ever growing volumes of high fidelity digital data from modern seismograph networks and access to increasingly powerful computational facilities.

Passive-source seismic tomography, a class of imaging techniques (derived from the geodesic X-ray transform and) adopted from medical applications in the late 1960's, has been used to map the smooth variations in the propagation speed of seismic P and S waves below the earth's surface (see, e.g., Romanowicz [1], for a review and pertinent references). To image singularities in the earth's medium properties one needs to resort to scattered waves or phases. Exploration seismologists have developed and long used a range of imaging and inverse scattering techniques with scattered waves, generated by active sources, to delineate and characterize subsurface reservoirs of fossil fuels (e.g., Yilmaz [2]). A large class of these imaging and inverse scattering techniques can be formulated and analyzed in terms of a Generalized Radon Transform (GRT [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]) and its extension [12] using techniques from microlocal analysis.

In this paper we utilize the matrix representation of GRT operators in the curvelet frame [13, 14, 15, 16, 50] to address issues of approximation of the imaging operator and its inverse. Under appropriate hypotheses, the GRT has sparse representation in the curvelet frame [50, 51, 54]. We strengthen this result by constructing simple approximations to the action of the GRT on a curvelet, with error term of size $2^{-k/2}$ on curvelets at frequency scale 2^k. The approximation is expressed as the composition of a quadratic change of variables and Fourier multiplication by a quadratic oscillatory function; the coefficients of the quadratic forms are determined by the ray-geometry of the underlying background medium. For pseudodifferential operators the approximating matrix is diagonal [57], allowing a simple construction of an approximate inverse for the normal operator associated to the GRT. The approximate inverse can then be used to exactly invert the imaging operator in the high-frequency region.

The action of the GRT on a source curvelet is highly localized in phase space near the image curvelet produced by ray-tracing [50]. This permits the localization of the inversion scheme (partial reconstruction). In particular, the singular perturbations to the background medium can be determined in a region of interest by the image of a collection of source curvelets which illuminates that region.

The analysis we develop here has its roots in (double) beamforming, double beam migration [18, 19] and beam-stack imaging [20, 21], which pose less stringent

requirements on data coverage than the GRT. Seismic data can be sparsely represented by curvelet-like functions [22]. Therefore, the results presented here shed new light on the concept of parsimonious pre-stack Kirchhoff migration [23]. Our approach also retains aspects of pre-stack plane-wave (Kirchhoff) migration [24, 25], offset plane-wave migration [26, 27], and delayed-shot pre-stack migration [28]. For example, synthesizing "incident" plane waves from point sources has its counterpart in the curvelet transform of the data.

Recently, while using tomographic models as a background, passive-source seismic imaging and inverse scattering techniques have been developed for the exploration of Earth's deep interior. For the imaging of crustal structure and subduction processes, see Bostock *et al.* [29] and Rondenay *et al.* [30] – here, the incident, teleseismic, waves are assumed to be "plane" waves. Wang *et al.* [31] present an inverse scattering approach based upon the GRT to image selected neighborhoods of Earth's core-mantle boundary (CMB) using broadband wavefields including the main "topside" reflections off the CMB and its precursors and coda (generated by scattering off interfaces above the CMB). Through joint interpretation with data from mineral physics this method enabled the estimation of temperatures at and near the CMB [32]. Enabling GRT-like transforms of "underside" reflections, Cao *et al.* [33] used SS precursors (see figure 1) to produce high resolution images of the upper mantle transition zone discontinuities. Mantle discontinuities near the CMB and in the transition zone are associated with phase transformations.

There exists a rich literature on the use of regional (dense) seismic arrays to detect and locate the origin of scattered energy in the seismic wavefield. Recent reviews of such array processing techniques are given by, for instance, Rost and Thomas [34] and Rondenay *et al.* [35]. In general, these techniques involve some type of beamforming [36]; that is, they assume (or aim to detect) the wave vector (or the horizontal slowness – related to the angle of incidence and back azimuth) of the incoming waves, and use this information to separate the coherent from the incoherent parts of the recorded signal. Implicitly, these methods aim to detect the wavefront set of the scattered wavefield [37]; this detection can then be used in migration. In beam-stack imaging [21] a region of the crust is subdivided into sub-areas. For each sub-area to be scanned, the seismograms from an event suite are incoherently stacked after beam-correcting each trace, computing new beams for each crustal sub-area, and migrating the results by applying appropriate time offsets, in the spirit of time migration or geophysical diffraction tomography. Deuss *et al.* [38] use an imaging approach through waveform stacking, in particular, of SS precursors: After selecting a bin of scattering (or image) points, which implies a selection of source-receiver pairs, the authors correct for the moveout (observed reference arrival times) of SS in the seismic records, and then stack the records at different slownesses (dependent on the bin) for given (array specific) times relative to the SS arrival time. (This stacking can be viewed as

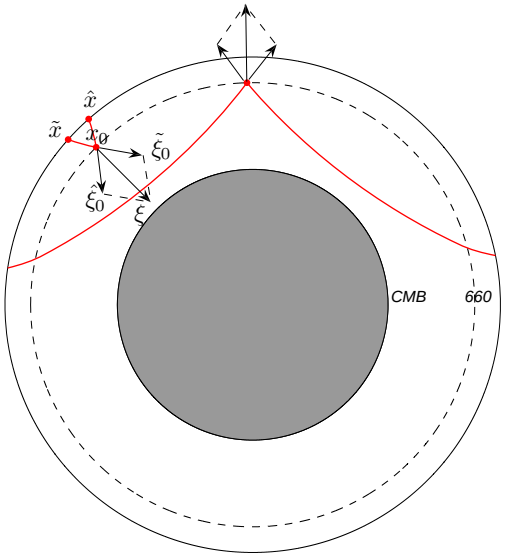

Figure 1. Scattered rays (broken geodesics) for imaging discontinuities (here the "660" corresponding with a phase transition at 660 km depth) in Earth's mantle. (CMB stands for core-mantle boundary.)

beamforming.) For this family of imaging techniques, see also Flanagan and Shearer [39].

Migration methods have been applied to regional data sets with a weighting factor which depends on the incident angles of the rays. To this end, the migration operators have been limited to the Fresnel volume of the reflected ray paths [40] to reduce artifacts caused by truncated wavefield observations. In this context, the migration operator has been further subjected to slowness-backazimuth weighting with the aid of Gaussian window functions [41]. The desired artifact reduction is implied by the rigorous partial reconstruction proposed and developed in this paper.

The outline of the paper is as follows. In Section 2 we summarize the extension of the generalized Radon transform viewed as a Fourier integral operator and bring its kernel in a particular oscillatory integral form. In Section 3, we introduce the relevant matrix classes. We review the (co-)frame of curvelets and the underlying dyadic parabolic decomposition in Appendix A. We then prove a result pertaining to the diagonal approximation of pseudodifferential operators (Lemma 3.1) and the computation of their inverses on the range of the curvelet transform restricted to sufficiently fine scales. To this end, we introduce the symbol class $S^0_{\frac{1}{2},rad}$ and the notion of a "curvelet-like function". In Section 4 we prove results (Theorems 4.1-4.3) pertaining to matrix approximations to the generalized Radon transform. The approximations are characterized by multiplications

and convolutions, the consequence of an underlying separation of variables in phase space of the relevant symbols. These lead to fast algorithms, and we speak of imaging "in the curvelet domain". The results of this section also apply, for example, to the Fourier integral operator representing the parametrix of the wave equation with smooth coefficients. In Section 5 we introduce a method of partial reconstruction incorporating "illumination correction" and prove the necessary estimates (Lemma 5.1). The results of this section can be directly extended to other imaging schemes as long as the canonical relation describing the propagation of singularities by the scheme is locally the graph of an invertible canonical transformation. In Section 6 we discuss how the results presented in Sections 4 and 5 can be implemented while replacing techniques currently in practice in seismic array processing.

1.2. Modelling, Scattering Operator

The propagation and scattering of seismic waves is governed by the elastic wave equation, which is written in the form

$$P_{il}u_l = f_i, \tag{1}$$

where

$$u_l = \sqrt{\rho(x)}(\text{displacement})_l, \qquad f_i = \frac{1}{\sqrt{\rho(x)}}(\text{volume force density})_i, \tag{2}$$

and

$$P_{il} = \delta_{il}\frac{\partial^2}{\partial t^2} + A_{il} + \text{l.o.t.}\ ,\ A_{il} = -\frac{\partial}{\partial x_j}\frac{c_{ijkl}(x)}{\rho(x)}\frac{\partial}{\partial x_k}, \tag{3}$$

where l.o.t. stands for "lower-order terms", $x \in \mathbb{R}^n$ and $i,j,k,l \in \{1,\dots,n\}$; ρ is the density of mass while c_{ijkl} denotes the stiffnesss tensor. The system of partial differential equations (1) is assumed to be of principal type. This excludes, essentially, characteristics with changing multiplicities. Isotropic media satisfy this assumption. The system supports different wave types (also called modes), one "compressional" and $n-1$ "shear". We label the modes by $M, N, \dots$

For waves in mode M, singularities are propagated along bicharacteristics, which are determined by Hamilton's equations with Hamiltonian B_M; that is

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}\lambda} &= \frac{\partial}{\partial \xi}B_M(x,\xi) \quad , \qquad \frac{\mathrm{d}t}{\mathrm{d}\lambda} = 1, \\ \frac{\mathrm{d}\xi}{\mathrm{d}\lambda} &= -\frac{\partial}{\partial x}B_M(x,\xi) \quad , \qquad \frac{\mathrm{d}\tau}{\mathrm{d}\lambda} = 0. \end{aligned} \tag{4}$$

The $B_M(x,\xi)$ follow from the diagonalization of the principal symbol matrix of $A_{il}(x,\xi)$, namely as the (distinct) square roots of its eigenvalues. Clearly, the solution of (4) may be parameterized by t (that is, $\lambda = t$). We denote the solution of (4) with initial values (x_0,ξ_0) at $t=0$ by $(x_M(x_0,\xi_0,t), \xi_M(x_0,\xi_0,t))$.

To introduce the scattering of waves, the total value of the medium parameters ρ, c_{ijkl} is written as the sum of a smooth background component, $\rho(x), c_{ijkl}(x)$, and a singular perturbation, $\delta\rho(x), \delta c_{ijkl}(x)$, namely $\rho(x)+\delta\rho(x)$, $c_{ijkl}(x)+\delta c_{ijkl}(x)$. This decomposition induces a perturbation of P_{il} (cf. (3)),

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

The scattered field, δu_l, in the single scattering approximation, satisfies

$$P_{il}\delta u_l = -\delta P_{il} u_l.$$

Data are measurements of the scattered wave field, δu. When no confusion is possible, we denote data by u, however. We assume point sources (consistent with the far field approximation) and point receivers. Then the scattered wave field is expressible in terms of the Green's function perturbations, $\delta G_{MN}(\widehat{x}, \widetilde{x}, t)$, with incident modes of propagation N generated at $\widetilde{x}$ and scattered modes of propagation M observed at $\widehat{x}$ as a function of time. Here, $(\widehat{x}, \widetilde{x}, t)$ are contained in some acquisition manifold. This is made explicit by introducing the coordinate transformation, $y \mapsto (\widehat{x}(y), \widetilde{x}(y), t(y))$, such that $y = (y', y'')$ and the acquisition manifold, Y say, is given by $y'' = 0$. We assume that the dimension of y'' is $2 + c$, where c is the codimension of the acquisition geometry. In this framework, the data are modeled by

$$\left(\frac{\delta\rho(x)}{\rho(x)}, \frac{\delta c_{ijkl}(x)}{\rho(x)}\right) \mapsto \delta G_{MN}(\widehat{x}(y', 0), \widetilde{x}(y', 0), t(y', 0)). \tag{5}$$

When no confusion is possible, we use the notation $\delta G_{MN}(y')$.

We denote scattering points by x_0; $x_0 \in X \subset \mathbb{R}^n$, reflecting that $\operatorname{supp}\delta\rho \subset X$ and $\operatorname{supp}\delta c \subset X$. The bicharacteristics connecting the scattering point to a receiver (in mode M) or a source (in mode N) can be written as solutions of (4),

$$\widehat{x} = x_M(x_0, \widehat{\xi}_0, \widehat{t}\,)\ ,\ \widetilde{x} = x_N(x_0, \widetilde{\xi}_0, \widetilde{t}\,)\ ,$$
$$\widehat{\xi} = \xi_M(x_0, \widehat{\xi}_0, \widehat{t}\,)\ ,\ \widetilde{\xi} = \xi_N(x_0, \widetilde{\xi}_0, \widetilde{t}\,)\ ,$$

with appropriately chosen "initial" $\widehat{\xi}_0$ and $\widetilde{\xi}_0$, respectively. Then $t = \widehat{t} + \widetilde{t}$ represents the "two-way" reflection time. The frequency τ satisfies $\tau = -B_M(x_0, \widehat{\xi}_0)$. We obtain $(y(x_0, \widehat{\xi}_0, \widetilde{\xi}_0, \widehat{t}, \widetilde{t}), \eta(x_0, \widehat{\xi}_0, \widetilde{\xi}_0, \widehat{t}, \widetilde{t}))$ by transforming $(\widehat{x}, \widetilde{x}, \widehat{t} + \widetilde{t}, \widehat{\xi}, \widetilde{\xi}, \tau)$ to (y, η) coordinates. We then invoke the following assumptions that concern scattering over π and rays grazing the acquisition manifold:

Assumption 1. *There are no elements $(y', 0, \eta', \eta'')$ with $(y', \eta') \in T^*Y\backslash 0$ such that there is a direct bicharacteristic from $(\widehat{x}(y', 0), \widehat{\xi}(y', 0, \eta', \eta''))$ to $(\widetilde{x}(y', 0), -\widetilde{\xi}(y', 0, \eta', \eta''))$ with arrival time $t(y', 0)$.*

Assumption 2. *The matrix*

$$\frac{\partial y''}{\partial(x_0, \widehat{\xi}_0, \widetilde{\xi}_0, \widehat{t}, \widetilde{t}\,)} \text{ has maximal rank.} \tag{6}$$

With Assumptions 1 and 2, equation (5) defines a Fourier integral operator of order $\frac{n-1+c}{4}$ and canonical relation, that governs the propagation of singularities, given by

$$\begin{aligned}\Lambda_{MN} = \{(y'(x_0,\widehat{\xi}_0,\widetilde{\xi}_0,\widehat{t},\widetilde{t}),\eta'(x_0,\widehat{\xi}_0,\widetilde{\xi}_0,\widehat{t},\widetilde{t});x_0,\widehat{\xi}_0+\widetilde{\xi}_0)\,| \\ B_M(x_0,\widehat{\xi}_0)=B_N(x_0,\widetilde{\xi}_0)=-\tau, y''(x_0,\widehat{\xi}_0,\widetilde{\xi}_0,\widehat{t},\widetilde{t})=0\} \\ \subset T^*Y\backslash 0\times T^*X\backslash 0\end{aligned} \tag{7}$$

[4, 9, 12]. The condition $y''(x_0,\widehat{\xi}_0,\widetilde{\xi}_0,\widehat{t},\widetilde{t}) = 0$ determines the traveltimes $\widehat{t}$ for given $(x_0,\widehat{\xi}_0)$ and $\widetilde{t}$ for given $(x_0,\widetilde{\xi}_0)$. The canonical relation admits coordinates, (y'_I, x_0, η'_J), where $I \cup J$ is a partition of $\{1,\dots,2n-1-c\}$, and has an associated phase function, $\Phi_{MN} = \Phi_{MN}(y', x_0, \eta'_J)$. While establishing a connection with double beamforming, we will also use the notation $x^s = \widetilde{x}(y',0)$, $x^r = \widehat{x}(y',0)$; when no confusion is possible, we use the simplified notation $y' = (x^s, x^r, t)$.

We refer to the operator above as the scattering operator. Its principal symbol can be explicitly computed in terms of solutions of the transport equation [12]. In the further analysis we suppress the subscripts $_{MN}$, and drop the prime and write y for y' and η for η'.

2. Generalized Radon Transform

Through an extension, the scattering operator becomes, microlocally, an invertible Fourier integral operator, the canonical relation of which is a graph. The inverse operator acts on seismic reflection data and describes inverse scattering by the generalized Radon transform.

2.1. Extension

Subject to the restriction to the acquisition manifold Y, the data are a function of $2n-1-c$ variables, while the singular part of the medium parameters is a function of n variables. Here, we discuss the extension of the scattering operator to act on distributions of $2n-1-c$ variables, equal to the number of degrees of freedom in the data acquisition. We recall the commonly invoked

Assumption 3. *(Guillemin [42]) The projection π_Y of Λ on $T^*Y\backslash 0$ is an embedding.*

This assumption is known as the Bolker condition. It admits the presence of caustics. Because Λ is a canonical relation that projects submersively on the subsurface variables (x,ξ) (using that the matrix operator P_{il} is of principal type), the projection of (7) on $T^*Y\backslash 0$ is immersive [43, Lemma 25.3.6 and (25.3.4)]. Indeed, only the injectivity part of the Bolker condition needs to be verified. The image $\mathcal{L}$ of π_Y is locally a coisotropic submanifold of $T^*Y\backslash 0$.

Since the projection π_X of Λ on $T^*X\backslash 0$ is submersive, we can choose (x,ξ) as the first $2n$ local coordinates on Λ; the remaining $\dim Y - n = n-1-c$ coordinates are denoted by $e \in E$, E being a manifold itself. Moreover, $\nu = \|\xi\|^{-1}\xi$ is identified as the

seismic *migration dip*. The sets $\mathcal{X} \ni (x,\xi) = \text{const.}$ are the isotropic fibers of the fibration of Hörmander [44], Theorem 21.2.6; see also Theorem 21.2.4. The wavefront set of the data is contained in $\mathcal{L}$ and is a union of such fibers. The map $\pi_X \pi_Y^{-1} : \mathcal{L} \to \mathcal{X}$ is a canonical isotropic fibration, which can be associated with seismic *map migration* [45].

With Assumption 3 being satisfied, we define Ω as the map (on Λ),

$$\Omega : (x,\xi,e) \mapsto (y(x,\xi,e), \eta(x,\xi,e)) : T^*X\backslash 0 \times E \to T^*Y\backslash 0 ;$$

this map conserves the symplectic form of $T^*X\backslash 0$. The (x,ξ,e) are "symplectic" coordinates on the projection $\mathcal{L}$ of Λ on $T^*Y\backslash 0$. In the following lemma, these coordinates are extended to symplectic coordinates on an open neighborhood of $\mathcal{L}$, which is a manifestation of Darboux's theorem stating that T^*Y can be covered with symplectic local charts.

Lemma 2.1. *Let $\mathcal{L}$ be an embedded coisotropic submanifold of $T^*Y\backslash 0$, with symplectic coordinates (x,ξ,e). Denote $\mathcal{L} \ni (y,\eta) = \Omega(x,\xi,e)$. We can find a homogeneous canonical map G from an open part of $T^*(X \times E)\backslash 0$ to an open neighborhood of $\mathcal{L}$ in $T^*Y\backslash 0$, such that $G(x,e,\xi,\varepsilon = 0) = \Omega(x,\xi,e)$.*

Let M be the canonical relation defined as the graph of map G in this lemma, i.e.

$$M = \{(G(x,e,\xi,\varepsilon); x,e,\xi,\varepsilon)\} \subset T^*Y\backslash 0 \times T^*(X \times E)\backslash 0 .$$

One can then construct a Maslov-type phase function for M that is directly related to a phase function for Λ. Suppose (y_I, x, η_J) are suitable coordinates for Λ. For $|\varepsilon|$ small, the constant-ε subset of M allows the same set of coordinates, thus we can use coordinates $(y_I, \eta_J, x, \varepsilon)$ on M. Now there is (see Theorem 4.21 in Maslov and Fedoriuk [46]) a function $S(y_I, x, \eta_J, \varepsilon)$, called the generating function, such that M is given by

$$\begin{aligned} y_J &= \frac{\partial S}{\partial \eta_J} , \qquad \eta_I = -\frac{\partial S}{\partial y_I} , \\ \xi &= \frac{\partial S}{\partial x} , \qquad e = -\frac{\partial S}{\partial \varepsilon} . \end{aligned} \tag{8}$$

A phase function for M is hence given by

$$\Psi(y,x,e,\eta_J,\varepsilon) = S(y_I,x,\eta_J,\varepsilon) - \langle \eta_J, y_J \rangle + \langle \varepsilon, e \rangle . \tag{9}$$

A phase function for Λ is then recovered by

$$\Psi(y, x, \tfrac{\partial S}{\partial \varepsilon}|_{\varepsilon=0}, \eta_J, 0) = \Phi(y, x_0, \eta_J) .$$

We then obtain a mapping from a reflectivity function (illustrated in figure 2) to reflection data that extends the mapping from contrast to data (cf. (5)). We recall

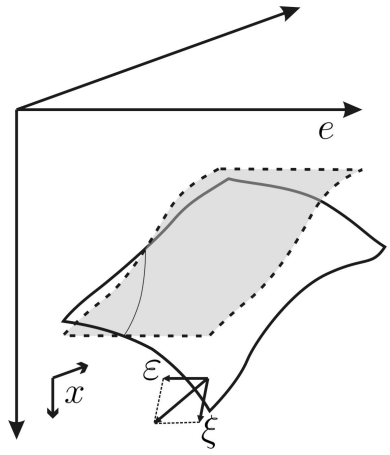

Figure 2. Wavefront set of an extended image, $r = r(x, e)$. The gray surface (singular support) corresponds with $\varepsilon = 0$ and maps into the range of the scattering operator before extension. The transparent surface exemplifies the extension to ε values away from zero.

Theorem 2.2. *[12] Suppose microlocally that Assumptions 1 (no scattering over π), 2 (transversality), and 3 (Bolker condition) are satisfied. Let F be the Fourier integral operator,*

$$F : \mathcal{E}'(X \times E) \to \mathcal{D}'(Y)\,,$$

with canonical relation given by the graph of the extended map $G : (x, \xi, e, \varepsilon) \mapsto (y, \eta)$ constructed in Lemma 2.1. Then the data can be modeled by F acting on a distribution $r(x, e)$ of the form

$$r(x, e) = \mathsf{R}(x, D_x, e)\,\mathsf{c}(x), \tag{10}$$

where R *stands for a smooth e-family of pseudodifferential operators and* $\mathsf{c} \in \mathcal{E}'(X)$ *with* $\mathsf{c} = \left(\frac{\delta c_{ijkl}}{\rho}, \frac{\delta\rho}{\rho}\right)$.

The operator F is microlocally invertible. By composing with an elliptic pseudodifferential operator we can assume without loss of generality that F is a zeroth-order Fourier integral operator associated to a (local) canonical graph. We recall that for Fourier integral operators the canonical relations of which are locally the graphs of canonical transformations, we have the property that their orders equal their Sobolev orders [44, Cor. 24.3.2]. The Bolker condition pertains to the background model. With maximal acquisition geometry it is generically satisfied. A further understanding - in terms of geometry of characteristics - how the Bolker condition can be violated can be found in Stolk [47], and for the case of common-source acquisition geometry in Nolan and Symes [8].

Remark. The operator F extends the procedure applied in [31, 32] to image, with the adjoint F^*, Earth's lowermost mantle, in particular the so-called D'' layer, using core reflected ScS "phases", their precursors and their coda, to the generic case admitting the formation of caustics. The e dependence in $r(x,e)$ can be exploited in a formulation of inference of singularities in the presence of (coherent) "noise" [48].

2.2. Oscillatory Integral Representation

If we have a canonical transformation from a neighborhood of $(x_0, e_0, \xi_0, \varepsilon_0) \in T^*(X \times E)\backslash 0$ to a neighborhood of $(y_0, \eta_0) \in T^*Y\backslash 0$, then one can choose local coordinates (y, ξ, ε) on a neighborhood of $(y_0, \eta_0, x_0, e_0, \xi_0, \varepsilon_0)$ on M [43, Prop. 25.3.3], that is, $M : (y, \eta, x, e, \xi, \varepsilon) \to (y, \xi, \varepsilon)$ is a local diffeomorphism. We denote the associated generating function by $\widetilde{S} = \widetilde{S}(y, \xi, \varepsilon)$ and obtain the phase function

$$\phi(x, e, y, \eta) = \widetilde{S}(y, \xi, \varepsilon) - \langle \xi, x\rangle - \langle \varepsilon, e\rangle \tag{11}$$

(cf. (9)). In fact, on M locally we can regard η and (x, e) as functions of (y, ξ, ε); then we can take $\widetilde{S}(y, \xi, \varepsilon) = \langle \eta(y, \xi, \varepsilon), (x(y, \xi, \varepsilon), e(y, \xi, \varepsilon))\rangle$ [44, Thm. 21.2.18].

We introduce the shorthand notation, $x := (x, e)$, $\xi := (\xi, \varepsilon)$, resetting $n := 2n - 1$, and $S(y, \xi) := \widetilde{S}(y, \xi, \varepsilon)$ and $\Sigma : (x, \xi) \to (y, \eta) = (\Sigma_1(x, \xi), \Sigma_2(x, \xi))$ corresponding with $G(x, e, \xi, \varepsilon)$, cf. Lemma 2.1. We identify $v(x)$ with $r(x, e)$, and we get, since F is a Fourier integral operator,

$$(Fv)(y) = \int A(y, x)v(x)\,\mathrm{d}x. \tag{12}$$

The kernel admits an oscillatory integral representation

$$A(y, x) = \int a(y, \xi)\ \exp[\mathrm{i}\phi(y, x, \xi)]\,\mathrm{d}\xi, \tag{13}$$

with non-degenerate phase function

$$\phi(y, x, \xi) = S(y, \xi) - \langle \xi, x\rangle \tag{14}$$

and amplitude $a = a(y, \xi)$, a standard symbol of order zero, with principal part homogeneous in ξ of order 0. With the above form of the phase function, it follows immediately that operator F propagates singularities according to the map,

$$\left(\frac{\partial S}{\partial \xi}, \xi\right) \to \left(y, \frac{\partial S}{\partial y}\right), \tag{15}$$

which can be identified as Σ. Substituting (14) into (12)-(13) yields the representation

$$(Fv)(y) = \int a(y, \xi)\ \exp[\mathrm{i}\, S(y, \xi)]\,\widehat{v}(\xi)\,\mathrm{d}\xi, \tag{16}$$

in which S satisfies the homogeneity property $S(y, c\xi) = cS(y, \xi)$ for $c > 0$; $\widehat{v}$ denotes the Fourier transform of v, and $\mathrm{d}\xi$ denotes $(2\pi)^{-n}$ times Lebesgue measure.

We remark that the above representation is valid microlocally. In Section 4 we study the action of operators of the form (16) to curvelets. The results for the global Fourier integral operator F are obtained by taking a superposition of the above representations using an appropriate microlocal partition of the unity in phase space.

3. Matrix Classes and Operator Approximations

3.1. "Curvelets", Matrix Classes and Operators

The (co)frame of curvelets, φ_γ, ψ_γ, is defined in (A.3). We introduce the notation C for the curvelet transform (analysis): $v_\gamma = (\mathrm{C}v)_\gamma$ (cf. (A.4)), and also define $\mathrm{C}^{-1}\{c_\gamma\} = \sum_\gamma c_\gamma \varphi_\gamma$ for the inverse transform (synthesis). We observe that $\mathrm{C}^{-1}\mathrm{C} = I$ on $L^2(\mathbb{R}^n)$, and that $\mathrm{C}\mathrm{C}^{-1} \equiv \Pi$ is a (not necessarily orthogonal) projection operator of ℓ^2_γ onto the range of the analysis operator C. It holds that $\Pi^2 = \Pi$, but Π is generally not self-adjoint unless $\psi_\gamma = \varphi_\gamma$. Observe that, as a matrix on ℓ^2_γ,

$$\Pi_{\gamma'\gamma} = \langle \psi_{\gamma'}, \varphi_\gamma \rangle \, .$$

If $A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$, then the matrix $[A] = \mathrm{C}A\mathrm{C}^{-1}$ preserves the range of C, since $\mathrm{C}^{-1}\Pi = \mathrm{C}^{-1}$, and $\Pi\mathrm{C} = \mathrm{C}$. In particular, $[A]\Pi = \Pi[A] = [A]$. Here, and when convenient, we identify operators on ℓ^2_γ with matrices.

Let d denote the pseudodistance on $S^*(X)$ introduced in [49, Definition 2.1]

$$d(x,\nu;x',\nu') = |\langle \nu, x - x' \rangle| + |\langle \nu', x - x' \rangle| \\ + \min\{\|x - x'\|, \|x - x'\|^2\} + \|\nu - \nu'\|^2.$$

If $\gamma = (x,\nu,k)$ and $\gamma' = (x',\nu',k')$, let

$$\overline{d}(\gamma;\gamma') = 2^{-\min(k,k')} + d(x,\nu;x',\nu'). \tag{17}$$

The weight function $\mu_\delta(\gamma,\gamma')$ introduced in [50] is given by

$$\mu_\delta(\gamma,\gamma') = (1 + |k' - k|^2)^{-1} 2^{-(\frac{1}{2}n+\delta)|k'-k|} 2^{-(n+\delta)\min(k',k)} \overline{d}(\gamma,\gamma')^{-(n+\delta)}.$$

We summarize [50, Definitions 2.6-2.8]. If χ is a mapping on $S^*(\mathbb{R}^n)$, the matrix M with elements $M_{\gamma'\gamma}$ belongs to the class $\mathcal{M}^r_\delta(\chi)$, if there is a constant $C(\delta)$ such that

$$|M_{\gamma'\gamma}| \le C(\delta)\, 2^{kr} \mu_\delta(\gamma', \chi(\gamma)) \qquad (2^{kr} \approx \|\xi\|^r); \tag{18}$$

here, $\chi(\gamma) = (\chi(x_j,\nu),k)$. Furthermore, $\mathcal{M}^r(\chi) = \cap_{\delta>0}\mathcal{M}^r_\delta(\chi)$. If χ is the projection of a homogeneous canonical transformation, then by [49, Lemma 2.2] the map χ preserves the distance d up to a bounded constant; that is $\overline{d}(\chi^{-1}(\gamma),\gamma') \approx \overline{d}(\gamma,\chi(\gamma'))$. Hence, the transpose operation takes matrices in $\mathcal{M}^r(\chi)$ to $\mathcal{M}^r(\chi^{-1})$. We note that the projection map $\Pi = \mathrm{C}\mathrm{C}^{-1}$ belongs to $\mathcal{M}^0(I)$, see [50, Lemma 2.9].

It is also useful to introduce norms on the class of matrices determined by distance-weighted ℓ^2_γ norms on columns and rows. Precisely, for $\alpha \geq 0$ and a given χ,

$$\|M\|^2_{2;\alpha} = \sup_\gamma \sum_{\gamma'} 2^{2|k-k'|\alpha} 2^{2\min(k,k')\alpha} \overline{d}(\gamma';\chi(\gamma))^{2\alpha} \, |M_{\gamma'\gamma}|^2 + \sup_{\gamma'} \sum_{\gamma} 2^{2|k-k'|\alpha} 2^{2\min(k,k')\alpha} \overline{d}(\gamma';\chi(\gamma))^{2\alpha} \, |M_{\gamma'\gamma}|^2 \,. \quad (19)$$

We remark that any matrix bounded on ℓ^2_γ must have finite $(2;0)$ norm, since this corresponds to rows and columns being square summable. Additionally, it follows immediately that

$$\|M\|_{2;\alpha+n} < \infty \quad \Rightarrow \quad M \in \mathcal{M}^0_\alpha(\chi) \,. \quad (20)$$

Inclusion in the other direction follows from the proof of [50, Lemma 2.4]

$$M \in \mathcal{M}^0_\alpha(\chi) \quad \Rightarrow \quad \|M\|_{2;\alpha} < \infty. \quad (21)$$

The technique of $(2;\alpha)$ bounds has been designed for propagation and scattering problems in rough background metrics (density normalized stiffness), but the $\mathcal{M}^r_\delta$ conditions lead more directly to desired mapping properties.

3.2. Pseudodifferential Operators and Diagonal Approximation

Pseudodifferential operators, of order r, with appropriate symbols are the most important example of operators with matrices of class $\mathcal{M}^r(I)$.

Let

$$Av(x) \equiv a(x,D)v(x) = \int \exp[\mathrm{i}\langle x,\xi\rangle] \, a(x,\xi)\widehat{u}(\xi) \, \mathrm{d}\xi,$$

where the symbol satisfies, for all j, α, β,

$$|\langle \xi, \partial_\xi\rangle^j \partial^\alpha_\xi \partial^\beta_x a(x,\xi)| \leq C_{j,\alpha,\beta}(1+\|\xi\|)^{-\frac{1}{2}|\alpha|+\frac{1}{2}|\beta|} \,. \quad (22)$$

We denote the class of symbols satisfying these estimates as $S^0_{\frac{1}{2},rad}$. Thus, $a \in S^0_{\frac{1}{2},rad}$ precisely when $\langle \xi, \partial_\xi\rangle^j a \in S^0_{\frac{1}{2},\frac{1}{2}}$ for all j. More generally, $a \in S^r_{\frac{1}{2},rad}$ precisely when $\langle \xi, \partial_\xi\rangle^j a \in S^r_{\frac{1}{2},\frac{1}{2}}$ for all j. Let A be a pseudodifferential operator with symbol in $S^r_{\frac{1}{2},rad}$. A stationary phase analysis then shows that $A\varphi_\gamma = 2^{kr} f_\gamma$, where

$$\widehat{f_\gamma}(\xi) = \rho_k^{-1/2} \widehat{g}_{\nu,k}(\xi) \, \exp[-\mathrm{i}\langle x_j, \xi\rangle], \quad (23)$$

in which $\widehat{g}_{\nu,k}$ satisfies the estimates

$$|\langle \nu, \partial_\xi\rangle^j \partial^\alpha_\xi \widehat{g}_{\nu,k}| \leq C_{j,\alpha,N} 2^{-k(j+\frac{1}{2}|\alpha|)} (1 + 2^{-k}|\langle \nu,\xi\rangle| + 2^{-k/2}\|\xi - B_{\nu,k}\|)^{-N}$$

for all N, where $\|\xi - B_{\nu,k}\|$ denotes the distance of ξ to the rectangle $B_{\nu,k}$ supporting $\widehat{\chi}_{\nu,k}$. Such an f_γ will be called a "curvelet-like function" centered at γ, cf. (A.3). In particular,

$$|\langle \psi_{\gamma'}, f_\gamma\rangle| \leq C(\delta) \, \mu_\delta(\gamma',\gamma)$$

for all $\delta > 0$, so that $\langle \psi_{\gamma'}, f_\gamma \rangle \in \mathcal{M}^0(I)$.

If the principal symbol of A is *homogeneous of order* 0, $a_0(x,\xi) = a_0(x, \xi/\|\xi\|)$, we have the following diagonalization result, which is a simple variation of the phase-linearization of Seeger-Sogge-Stein [52]

Lemma 3.1. *Suppose that A is a pseudodifferential operator with homogeneous principle symbol $a_0(x,\xi)$ of order 0. Then*

$$A\varphi_\gamma = a_0(x_j,\nu)\,\varphi_\gamma + 2^{-k/2} f_\gamma\,, \tag{24}$$

where f_γ is a curvelet-like function centered at γ.

Proof. The precise assumption we need is that the symbol of A equals a_0 plus a symbol of class $S^{-\frac{1}{2}}_{\frac{1}{2},rad}$. The terms of order $-\frac{1}{2}$ can be absorbed into f_γ, while

$$a_0(x,D)\varphi_\gamma(x) = \rho_k^{-1/2} \int \exp[\mathrm{i}\langle x - x_j, \xi\rangle] a_0(x,\xi)\widehat{\chi}_{\nu,k}(\xi)\,\mathrm{d}\xi.$$

For convenience we assume that $\nu = (1,0,\dots,0)$ lies on the ξ_1 axis. By homogeneity, $a_0(x,\xi) = a_0(x,1,\xi''/\xi_1)$, where $\xi'' = (\xi_2,\dots,\xi_n)$. We take the first-order Taylor expansion on a cone about the ξ_1 axis, that is,

$$a_0(x,1,\xi''/\xi_1) - a_0(x_j,\nu) = b_1(x,\xi)\cdot(x - x_j) + b_2(x,\xi)\cdot\xi''/\xi_1\,,$$

where b_1 and b_2 are smooth homogeneous symbols. The term with ξ''/ξ_1 is bounded by $2^{-k/2}$ on the support of $\widehat{\chi}_{\nu,k}$, and preserves the derivative bounds (23) on $\widehat{\chi}_{\nu,k}$ with a gain of $2^{-k/2}$. The term $b_1\cdot(x-x_j)$ leads to a contribution

$$\rho_k^{-1/2} \int \exp[\mathrm{i}\langle x - x_j, \xi\rangle]\, D_\xi(b_1(x,\xi)\widehat{\chi}_{\nu,k}(\xi))\,\mathrm{d}\xi,$$

which also yields a curvelet-like function of order $-\frac{1}{2}$. □

In (24) we write $r_\gamma = 2^{-k/2} f_\gamma$. Taking inner products with $\psi_{\gamma'}$ yields

$$[A]_{\gamma'\gamma} = a_0(x_j,\nu)\,\Pi_{\gamma'\gamma} + \langle \psi_{\gamma'}, r_\gamma\rangle\,. \tag{25}$$

If A is *elliptic*, we have uniform upper and lower bounds on the symbol $a_0(x,\xi)$, that is $C^{-1} \le |a_0(x,\xi)| \le C$ for some positive constant C. By (25) we then have

$$a_0(x_j,\nu)^{-1}[A]_{\gamma'\gamma} - \Pi_{\gamma'\gamma} \in \mathcal{M}^{-\frac{1}{2}}(I)\,. \tag{26}$$

Also, by (25),

$$|\, a_0(x_j,\nu) - \langle \psi_\gamma, \varphi_\gamma\rangle^{-1}[A]_{\gamma\gamma}\,| \le C\,2^{-k/2}\,.$$

It follows that (26) holds with $a_0(x_j,\nu)$ replaced by the normalized diagonal

$$D_\gamma = \Pi_{\gamma\gamma}^{-1}[A]_{\gamma\gamma}\,,$$

after modifying $[A]_{\gamma\gamma}$, if necessary, by terms of size $2^{-k/2}$, to allow for the possibility that the diagonal elements of $[A]$ may vanish for small k.

We remark that (26) also holds with $a_0(x_j, \nu)$ replaced by $a_0(x'_j, \nu')$. (The latter appears from applying the procedure of diagonal approximation to the *adjoint* of A.) This follows by (25) and the fact that

$$|a_0(x_j, \nu) - a_0(x'_j, \nu')| \le C\,(\,|x_j - x'_j| + |\nu - \nu'|\,) \le C\,d(x_j, \nu; x'_j, \nu')^{1/2}\,,$$

hence the commutator $(a_0(x'_j, \nu') - a_0(x_j, \nu))\Pi_{\gamma'\gamma}$ belongs to $\mathcal{M}^{-\frac{1}{2}}(I)$. As above, it then follows that

$$D^{-1}_{\gamma'}[A]_{\gamma'\gamma} = \Pi_{\gamma'\gamma} + R_{\gamma'\gamma}\,, \qquad R \in \mathcal{M}^{-\frac{1}{2}}(I)\,. \tag{27}$$

While A need not be invertible, (27) implies that one can invert $[A]$ on the range of C restricted to k sufficiently large. Precisely, let $\mathit{\Gamma}_0$ be a collection of indices γ. We denote by $\mathbf{1}^{\mathit{\Gamma}_0}$ the multiplication operator (diagonal) on ℓ^2_γ that truncates a sequence to $\mathit{\Gamma}_0$. Then $\Pi^{\mathit{\Gamma}_0} = \Pi\ \mathbf{1}^{\mathit{\Gamma}_0}$ is an approximate projection into the range of C, with rapidly decreasing coefficients away from Γ_0. In practice, it is desirous to take $\mathbf{1}^{\mathit{\Gamma}_0}$, at each fixed scale k, to be a smooth truncation to a neighborhood of $\mathit{\Gamma}_0$, such that $|\,\mathbf{1}^{\mathit{\Gamma}_0}_\gamma - \mathbf{1}^{\mathit{\Gamma}_0}_{\gamma'}| \le C\,\overline{d}(\gamma, \gamma')^{1/2}$. In this case,

$$(\,\mathbf{1}^{\mathit{\Gamma}_0}_\gamma - \mathbf{1}^{\mathit{\Gamma}_0}_{\gamma'})\Pi_{\gamma\gamma'} \in \mathcal{M}^{-\frac{1}{2}}\,, \tag{28}$$

so that $\mathbf{1}^{\mathit{\Gamma}_0}$ preserves the range of C at any fixed scale k up to an operator of norm $2^{-k/2}$, hence the difference between $\Pi^{\mathit{\Gamma}_0}$ and $\mathbf{1}^{\mathit{\Gamma}_0}$ is small on the range of C for large k.

If we multiply (27) on the right by $\mathbf{1}^{\mathit{\Gamma}_0}_\gamma$, and use that $R = R\,\Pi$, then

$$D^{-1}[A]\,\mathbf{1}^{\mathit{\Gamma}_0} = \Pi\,\mathbf{1}^{\mathit{\Gamma}_0} + R\,\mathbf{1}^{\mathit{\Gamma}_0} = (I + R_0)\Pi^{\mathit{\Gamma}_0},$$

where R_0 is the matrix R restricted to the scales k occuring in $\mathit{\Gamma}_0$. Hence, if $\mathit{\Gamma}_0$ is supported by k sufficiently large, then $I + R_0$ can be inverted, and

$$(I + R_0)^{-1}D^{-1}[A]\,\mathbf{1}^{\mathit{\Gamma}_0} = \Pi^{\mathit{\Gamma}_0}\,,$$

using a Neumann expansion. To leading order the inverse is *diagonal*. We will exploit this result in Section 5, while solving the normal equations derived from the composition F^*F, yielding "illumination correction" and partial reconstruction of the reflectivity function.

4. Generalized Radon Transform Matrix Approximation

We consider the action of the generalized Radon transform operator F on a single curvelet, that is $v = \varphi_\gamma$ in (16),

$$(F\varphi_\gamma)(y) = \rho_k^{-1/2} \int a(y, \xi)\widehat{\chi}_{\nu,k}(\xi)\ \exp[\mathrm{i}\,(S(y, \xi) - \langle \xi, x_j\rangle)]\,\mathrm{d}\xi. \tag{29}$$

With the outcome, we can associate a "kernel"

$$A_{\nu,k}(y, x_j) = (F\varphi_\gamma)(y)\,. \tag{30}$$

The infinite generalized Radon transform matrix is given by

$$[F]_{\gamma'\gamma} := \int \overline{\psi_{\gamma'}(y)}(F\varphi_\gamma)(y)\,\mathrm{d}y = \int \overline{\psi_{\gamma'}(y)}A_{\nu,k}(y, x_j)\,\mathrm{d}y\,. \tag{31}$$

We then have $F = \mathrm{C}^{-1}[F]\mathrm{C}$.

We seek an approximation of $F\varphi_\gamma$ via expansions of the generating function $S(y,\xi)$ and the symbol $a(y,\xi)$ near the microlocal support of φ_γ. The first-order Taylor expansion of $S(y,\xi)$ along the ν axis, following [52], yields

$$S(y,\xi) - \langle \xi, x_j\rangle = \left\langle \xi, \frac{\partial S}{\partial \xi}(y,\nu) - x_j \right\rangle + h_2(y,\xi)\,, \tag{32}$$

where the error term $h_2(y,\xi)$ satisfies the estimates (22) on the ξ-support of $\widehat{\chi}_{\nu,k}$. Consequently, $\exp[\mathrm{i}h_2(y,\xi)]$ is a symbol of class $S^0_{\frac{1}{2},rad}$ if ξ is localized to the rectangle $B_{\nu,k}$ supporting $\widehat{\chi}_{\nu,k}$.

We introduce the coordinate transformation (note that ν depends on k)

$$y \to T_{\nu,k}(y) = \frac{\partial S}{\partial \xi}(y,\nu)\,.$$

If $b_{\nu,k}(x,\xi)$ is the order 0 symbol

$$b_{\nu,k}(x,\xi) = (a(y,\xi)\exp[\mathrm{i}h_2(y,\xi)]\,)|_{y=T^{-1}_{\nu,k}(x)}\,,$$

then

$$(F\varphi_\gamma)(y) = [b_{\nu,k}(x,D)\varphi_\gamma]_{x=T_{\nu,k}(y)}.$$

This decomposition expresses the generalized Radon transform operator as a (ν,k) dependent pseudodifferential operator followed by a change of coordinates, also depending on the pair (ν,k). This decomposition can be used to show that the matrix $[F]$ belongs to $\mathcal{M}^0(\chi)$, where χ is the projection of the homogeneous canonical transformation Σ (cf. (15)) to the co-sphere bundle. (See also Theorem 4.3 below.)

We use an expansion of the symbol and phase of the oscillatory integral representation to obtain an approximation for the generalized Radon transform matrix elements up to error of size $2^{-k/2}$; more precisely, the matrix errors will be of class $\mathcal{M}^{-\frac{1}{2}}(\chi)$. The principal part $a_0(y,\xi)$ of symbol $a(y,\xi)$ is homogeneous of order 0. Following Lemma 3.1, we may replace $a_0(y,\xi)$ by either $a_0(y,\nu)$ or $a_0(y_j,\nu)$, where

$$x_j = \frac{\partial S}{\partial \xi}(y_j,\nu) = T_{\nu,k}(y_j)\,,$$

with the effect of modifying the generalized Radon transform matrix by a matrix of class $\mathcal{M}^{-\frac{1}{2}}(\chi)$.

The symbol $h_2(y,\xi)$ is homogeneous of order 1 and of class $S^0_{\frac{1}{2},rad}$ on the support of $\widehat{\chi}_{\nu,k}$, whence we need account for the second-order terms in its Taylor expansion to obtain an approximation within order $-\frac{1}{2}$. The relevant approximation is to Taylor expand in ξ in directions perpendicular to ν, preserving homogeneity of order 1 in the radial direction; this is dictated by the non-isotropic geometry of the second-dyadic (or dyadic parabolic) decomposition.

For convenience of notation, we consider the case that ν lies on the ξ_1 axis. Then (compare (32))

$$S(y,\xi_1,\xi'') = \xi_1 S(y,1,\xi''/\xi_1) = \xi\cdot\frac{\partial S}{\partial\xi}(y,\nu) + \frac{1}{2}\,\frac{\xi''^2}{\xi_1}\cdot\frac{\partial^2 S}{\partial\xi''^2}(y,\nu) + h_3(y,\xi),$$

where $h_3(y,\xi)\in S^{-\frac{1}{2}}_{\frac{1}{2},rad}$ if ξ is restricted to the support of $\widehat{\chi}_{\nu,k}$. Replacing the symbol $\exp[\mathrm{i}h_3(y,\xi)]$ by 1 changes the matrix by terms of class $\mathcal{M}^{-\frac{1}{2}}(\chi)$, as in the proof of Lemma 3.1. Consequently, up to errors of order $-\frac{1}{2}$, one can replace the symbol $a(y,\xi)\exp[\mathrm{i}h_2(y,\xi)]$ on $B_{\nu,k}$ by

$$a(y,\nu)\exp[\mathrm{i}\,\tfrac{1}{2}\,\xi_1^{-1}\xi''^2\cdot\partial^2_{\xi''}S(y,\nu)]\;\mathbf{1}_{B_{\nu,k}}(\xi)$$

with $\mathbf{1}_{B_{\nu,k}}$ a smooth cutoff to the rectangle $B_{\nu,k}$ supporting $\widehat{\chi}_{\nu,k}$.

The exponent separates the variables y and ξ, and is bounded by a constant, independent of (ν,k). Approximating the complex exponential for bounded (by C) arguments by a polynomial function leads to a tensor-product representation of the symbol:

$$a(y,\nu)\exp[\mathrm{i}\,\tfrac{1}{2}\,\xi_1^{-1}\xi''^2\cdot\partial^2_{\xi''}S(y,\nu)]\approx\sum_{s=1}^{N}\alpha^1_{s;\nu,k}(y)\,\widehat{\alpha}^2_{s;\nu,k}(\xi).$$

To obtain an error of size $2^{-k/2}$ requires $C^N/N!\le 2^{-k/2}$, or $N\sim k/\log k$:

Theorem 4.1. *With $N\sim k/\log k$, one may express*

$$(F\varphi_\gamma)(y)=\sum_{s=1}^{N}\alpha^1_{s;\nu,k}(y)\,(\alpha^2_{s;\nu,k}*\varphi_\gamma)\circ T_{\nu,k}(y)+2^{-k/2}f_\gamma, \tag{33}$$

where f_γ is a curvelet-like function centered at $\chi(\gamma)$.

An alternative approximation starts with replacing $a(y,\xi)$ or $a(y,\nu)$ by $a(y_j,\nu)$ with $y_j=T^{-1}_{\nu,k}(x_j)$ (and $\gamma=(x_j,\nu,k)$). Similarly, up to an error of order $-\frac{1}{2}$, one may replace $\xi_1^{-1}\xi''^2\cdot\partial^2_{\xi''}S(y,\nu)$ by $\xi_1^{-1}\xi''^2\cdot\partial^2_{\xi''}S(y_j,\nu)$. Consequently, replacing $b_{\nu,k}(x,\xi)$ by the *x-independent* symbol

$$b_\gamma(\xi)=a(y_j,\nu)\exp[\mathrm{i}\,\tfrac{1}{2}\,\xi_1^{-1}\xi''^2\cdot\partial^2_{\xi''}S(y_j,\nu)]\,\mathbf{1}_{B_{\nu,k}}(\xi)=\widehat{\alpha}_\gamma(\xi),$$

modifies the generalized Radon transform matrix by terms in $\mathcal{M}^{-\frac{1}{2}}(\chi)$. Precisely,

Theorem 4.2. *One may express*

$$(F\varphi_\gamma)(y) = (\alpha_\gamma * \varphi_\gamma) \circ T_{\nu,k}(y) + 2^{-k/2} f_\gamma \tag{34}$$

where f_γ is a curvelet-like function centered at $\chi(\gamma)$.

This is a generalization of the geometrical, zeroth-order approximation of the common-offset realization – valid in the absence of caustics – of the generalized Radon transform considered in [53]. In Theorem 4.2, as well as Theorem 4.3 below, the terms in the approximation are determined by the geometry of the underlying canonical transformation (ray-tracing). An approach based upon uniform approximation of the symbol and zeroth-order phase terms by functions which separate space and frequency variables, as in Theorem 4.1, was studied in greater depth in [55].

The change of variables $T_{\nu,k}$ can also be suitably approximated by a local expansion of the generating function about (y_j, ν). This requires an approximation of the phase $\langle \xi, \frac{\partial S}{\partial \xi}(y,\nu) - x_j \rangle$ up to an error of size $2^{-k/2}$ (cf. (32)), which is accomplished by taking the second-order expansion in y about y_j. Precisely, we write

$$\begin{aligned} \frac{\partial S}{\partial \xi}(y,\nu) - x_j = \frac{\partial^2 S}{\partial \xi \partial y}(y_j,\nu)\cdot(y-y_j) + \frac{1}{2}\frac{\partial^3 S}{\partial \xi \partial y^2}(y_j,\nu)\cdot(y-y_j)^2 \\ + h_3(y,\nu), \end{aligned} \tag{35}$$

where $h_3(y,\nu)$ vanishes to third order at $y = y_j$, and hence $\xi \cdot h_3(y,\nu)$ leads to terms of order $2^{-k/2}$ as in Lemma 3.1. The first two terms on the right hand side of (35) are exactly the quadratic expansion of $T_{\nu,k}$ about $y = y_j$.

In the expression $\xi \cdot \frac{\partial^3 S}{\partial \xi \partial y^2}(y_j,\nu)\cdot(y-y_j)^2$ the terms in ξ perpendicular to ν are of size $2^{k/2}$ as opposed to 2^k for the component of ξ parallel to ν, hence lead to terms of size $2^{-k/2}$. This allows one to replace the third-order derivative term by the quadratic expression

$$\begin{aligned} \frac{1}{2}\left[\nu \cdot \frac{\partial^3 S}{\partial \xi \partial y^2}(y_j,\nu)\cdot(y-y_j)^2\right]\nu \\ = \frac{1}{2}\left[\frac{\partial^2 S}{\partial y^2}(y_j,\nu)\cdot(y-y_j)^2\right]\nu = Q_\gamma \cdot (y-y_j)^2 \end{aligned} \tag{36}$$

with $y_j = T_{\nu,k}^{-1}(x_j)$ (and $\gamma = (x_j,\nu,k)$) as before:

Theorem 4.3. *One may express*

$$(F\varphi_\gamma)(y) = (\alpha_\gamma * \varphi_{\nu,k}) \circ [DT_\gamma \cdot (y-y_j) + Q_\gamma \cdot (y-y_j)^2] + 2^{-k/2} f_\gamma, \tag{37}$$

where f_γ is a curvelet-like function centered at $\chi(\gamma)$.

Here, the affine map $DT_\gamma = \frac{\partial T_{\nu,k}}{\partial y}(y_j) = \frac{\partial^2 S}{\partial \xi \partial y}(y_j,\nu)$ can be decomposed into a rigid motion and a shear. The shear factor acts in a bounded manner on the curvelet, in that it preserves position and direction; see also [51, 53].

The contribution $Q_\gamma \cdot (y - y_j)^2$ captures the curvature of the underlying canonical transformation applied to the infinitesimal plane wave attached to φ_γ. As with the shear term it acts in a bounded manner on a curvelet, and can be neglected in a zeroth-order approximation. This is the case in [50], where rigid approximations to $T_{\nu,k}$ were taken. Both shear and curvature terms must be accounted for to obtain an approximation up to errors of size $2^{-k/2}$.

The expansion in Theorem 4.3 is analogous to the Gaussian beam expansion for isotropic wave packets evolving under the wave equation, that is, if F were the forward parametrix of the wave equation. A Gaussian beam is frequency localized to a ball of diameter $2^{k/2}$ in ξ, and in the Gaussian beam expansion one considers quadratic expansions in ξ about the center ξ_0 of the packet. For curvelets, the support is of dimension 2^k in radial directions, and the approximations to the phase must preserve homogeneity in the radial variable.

Remark. The matrix $[F^*]$, essentially, provides the means to perform generalized Radon transform imaging entirely in the curvelet domain (that is, "after double beamforming"). In this context, "beam stack migration" can be understood as "scanning" the magnitude of $\langle F\,\delta_{x_0}, u\rangle = \sum_\gamma \langle \delta_{x_0}, F^*\varphi_\gamma\rangle\, u_\gamma = \sum_\gamma F^*\varphi_\gamma(x_0)\, u_\gamma$ as a function of x_0.

5. Partial Reconstruction

In applications, the image will admit a sparse decomposition into curvelets. Suppose the goal is to reconstruct the image contribution composed of a small set of curvelets (a "target"). The aim is to reconstruct this contribution by the available acquisition of data with the least "artifacts" (hence curvelets).

Let v denote a model of reflectivity, as before, and w its image, interrelated through $w = F^*Fv$. We write

$$N = F^*F,$$

so that $[N] = [F^*]\,[F]$. The operator N is a pseudodifferential operator with polyhomogeneous symbol of order 0; in particular N has homogeneous principal symbol of order 0, and the results of Section 3.2 apply to N.

We describe a target region by the set of indices $\mathit{\Gamma}_0$. Our resolution-illumination analysis is thus focused on the product $[N]\,\Pi^{\mathit{\Gamma}_0}$. The acquisition of data is accounted for by $\Pi^{\mathcal{S}} = \Pi\,\mathbf{1}^{\mathcal{S}}$, where $\mathcal{S}$ stands for the (finite) set of curvelets that can be observed given the acquisition geometry. The resolution is thus described by the operator, and matrix,

$$\widetilde{N} = F^*\mathrm{C}^{-1}\,\mathbf{1}^{\mathcal{S}}\mathrm{C}\,F\,, \qquad [\widetilde{N}] = [F^*]\;\mathbf{1}^{\mathcal{S}}[F] = [F^*]\,\Pi^{\mathcal{S}}[F]\,, \tag{38}$$

and the normal equation to be solved, yielding the partial reconstruction, is given by $[\widetilde{N}]\,\mathrm{C}v = [F^*]\,\Pi^{\mathcal{S}}\mathrm{C}u$ where $\Pi^{\mathcal{S}}\mathrm{C}u$ represents the observed data. The set $\mathcal{S}$ is assumed

to contain a suitable neighborhood of $\chi(\Gamma_0)$, in that $\overline{d}(\gamma, \chi(\gamma_0)) \gg 2^{-k}$ for $\gamma \in \mathcal{S}^c$ and $\gamma_0 \in \Gamma_0$ at scale k. (Otherwise, Γ_0, or $\mathcal{S}$, need to be adjusted.) The matrix $[\widetilde{N}]$ then approximates the matrix $[N]$ near Γ_0 in the following sense:

Lemma 5.1. *Let*

$$\Delta_{\Gamma_0} = \inf_{\gamma \in \mathcal{S}^c, \gamma_0 \in \Gamma_0} 2^{|k_0 - k|} 2^{\min(k_0,k)} \overline{d}(\gamma; \chi(\gamma_0)).$$

Then for all α, and m arbitrarily large, there exists a constant $C_{\alpha,m}$ such that

$$\|([N] - [\widetilde{N}])\Pi^{\Gamma_0}\|_{2;\alpha} \leq C_{\alpha,m} \Delta_{\Gamma_0}^{-m}.$$

Proof. Since $[F^*]\Pi = [F^*]$ and $[F]\Pi = [F]$, the matrix $[N]\Pi^{\Gamma_0} - [\widetilde{N}]\Pi^{\Gamma_0}$ takes the form

$$\sum_{\gamma''} [F^*]_{\gamma\gamma''} \, \mathbf{1}^{\mathcal{S}^c}_{\gamma''} [F]_{\gamma''\gamma'} \, \mathbf{1}^{\Gamma_0}_{\gamma'} .$$

The sum is dominated by

$$C_{\delta,m} \sum_{\gamma''} \mu_\delta(\chi(\gamma), \gamma'') \; \mathbf{1}^{\mathcal{S}^c}_{\gamma''} \, \mu_{\delta+m}(\gamma'', \chi(\gamma')) \; \mathbf{1}^{\Gamma_0}_{\gamma'} .$$

We use the bound $\mu_{\delta+m}(\gamma'', \chi(\gamma')) \leq \Delta_{\Gamma_0}^{-m} \mu_\delta(\gamma'', \chi(\gamma'))$ and [50, Lemma 2.5], together with invariance of the distance under χ, to bound the sum by $C_{\delta,m} \Delta_{\Gamma_0}^{-m} \mu_\delta(\gamma, \gamma')$. The result follows, since $\|\mu_\delta(.,.)\|_{2,\alpha} \lesssim 1$ if $\delta \geq \alpha$. □

Finally, we explore the invertibility of $[\widetilde{N}]$ on the range of Π^{Γ_0}. To this end, we introduce an intermediate index set Γ_1 with $\Gamma_0 \subset \Gamma_1 \subset \chi^{-1}(\mathcal{S})$, for which $\Delta_{\Gamma_1} \approx \Delta_{\Gamma_0}$, and with

$$\|\Pi^{\Gamma_1}\Pi^{\Gamma_0} - \Pi^{\Gamma_0}\|_{2;\alpha} \lesssim \Delta_{\Gamma_0}^{-m} \tag{39}$$

for m arbitrarily large. For γ in a set containing Γ_1, then $|[\widetilde{N}]_{\gamma\gamma} - [N]_{\gamma\gamma}| \ll 1$. We introduce the inverse diagonal,

$$\widetilde{D}_\gamma^{-1} = \Pi_{\gamma\gamma} [\widetilde{N}]_{\gamma\gamma}^{-1} \tag{40}$$

for γ near Γ_1, and smoothly truncate $\widetilde{D}_\gamma^{-1}$ to 0 away from Γ_1. Then

$$\widetilde{D}^{-1} [\widetilde{N}] \, \Pi^{\Gamma_1} = \Pi^{\Gamma_1} + R \, ,$$

where $\|R\|_{2,\alpha} \ll 1$ if Δ_{Γ_0} is sufficiently large, depending on the given α.

If Π^{Γ_1} were a true projection then we would have $R = R \, \Pi^{\Gamma_1}$, and applying $(I+R)^{-1}$ would yield the desired inverse of $[\widetilde{N}]$ on the range of Π^{Γ_1}. In the case of the approximate projections Π^{Γ_0}, Π^{Γ_1}, one can obtain an approximate inverse against Π^{Γ_0}. We write

$$(I+R)^{-1} \widetilde{D}^{-1} [\widetilde{N}] \, \Pi^{\Gamma_1} \Pi^{\Gamma_0} = \Pi^{\Gamma_0} + (I+R)^{-1} (\Pi^{\Gamma_1} \Pi^{\Gamma_0} - \Pi^{\Gamma_0}) \, .$$

By (39) this yields

$$(I+R)^{-1}\widetilde{D}^{-1}[\widetilde{N}]\,\Pi^{\Gamma_0} = \Pi^{\Gamma_0} + \widetilde{R}\,,$$

where $\|\widetilde{R}\|_{2,\alpha} \ll 1$, provided Δ_{Γ_0} is sufficiently large, depending on the given α. Thus, by applying $(I+R)^{-1}\widetilde{D}^{-1}$ to $[F^*]\,\Pi^{\mathcal{S}}\mathrm{C}u$, we obtain the desired, approximate, partial reconstruction of the reflectivity function, where C has replaced the notion of double beamforming, and $[F]$ and $[F^*]$ can now be replaced by their approximations developed in the previous section.

Remark. In practical applications, R and $\widetilde{R}$ are neglected. In general, with limited illumination, the diagonal elements $[\widetilde{N}]_{\gamma\gamma}$ have to be estimated numerically through "demigration" followed by "remigration" against Π^{Γ_0}. In the case of full illumination, the diagonal elements can be directly approximated using (25). For an optimization approach to solving the normal equation, in this context, see Symes [56] and Herrmann *et al.* [57].

Remark. The image of a single data curvelet is naturally given by $w = F^*\varphi_\gamma = \sum_{\gamma'}[F^*]_{\gamma'\gamma}\varphi_{\gamma'}$ whence $w_{\gamma'} = [F^*]_{\gamma'\gamma}$. From the fact that the matrix $[F^*]$ belongs to $\mathcal{M}^0(\chi^{-1})$, it is immediate that for α arbitrarily large (cf. (19))

$$\sum_{\gamma} 2^{2|k-k'|\alpha}2^{2\min(k,k')\alpha}\overline{d}(\gamma';\chi^{-1}(\gamma))^{2\alpha}\,|[F^*]_{\gamma'\gamma}|^2 \le C$$

illustrating that the curvelet decomposition of the data eliminates the "isochrone smear" associated with imaging individual data samples.

6. Discussion

The results presented in this paper essentially provide a novel approach to imaging, based on the generalized Radon transform, replacing the notions of "plane-wave migration" and "beam-stack imaging" by matrix approximations using curvelets on the one hand, and addressing the problem of partial reconstruction on the other hand. However, the results presented in Section 3.2 apply to general, elliptic, pseudodifferential operators, while the results presented in Section 4 pertain to all Fourier integral operators (of order zero) the canonical relation of which is (locally) a canonical graph.

Application of the results presented in this paper, in the context of seismic array processing, consists of data decomposition into curvelets, imaging, and reconstruction. We briefly discuss each step.

Decomposition. The curvelet transform applied to the data replaces the notion of double beamforming [36], a tool used in seismic array processing [34]. Double beamforming can be introduced as follows. Let g be a real, even Schwartz function in $\mathbb{R}^n$ with $\|g\|_{L^2} = (2\pi)^{-n/2}$; suppose that $\widehat{g}$ is supported in the unit ball. For $\lambda \ge 1$, define

$$g_\lambda(x;x_0,\xi_0) = \lambda^{-n/4}\exp(\mathrm{i}\langle\eta, x-x_0\rangle)\,g(\lambda^{1/2}(x-x_0)). \tag{41}$$

The FBI transform [58] of u is then given by

$$U(x_0,\xi_0) = T_\lambda u(x_0,\xi_0) = \int u(x)\,\overline{g_\lambda(x;x_0,\xi_0)}\,\mathrm{d}x. \tag{42}$$

The adjoint, T_λ^*, of T_λ follows as

$$T_\lambda^* U(x) = \int U(x_0,\xi_0)\,g_\lambda(x;x_0,\xi_0)\,\mathrm{d}x_0\mathrm{d}\xi_0. \tag{43}$$

We have the property $T_\lambda^* T_\lambda = I$.

For application to the data (see Subsection 1.2), we identify $y = (x^s, x^r, t)$ (replacing x) and $\eta_0 = \tau\,(\pi^s, \pi^r, 1)$ (replacing ξ_0). In the context and terminology of double beamforming, $\pi^{s,r}$ represent the slowness vectors. The remaining normal components of the associated covectors can be obtained by using that their norms equal the reciprocal wavespeeds at $x_0^{s,r}$. In the above, x_0^r coincides typically with the geometrical center of the receiver array. It has been noted that the applicability of beamforming requires relatively narrow-aperture arrays. To use earthquakes for source-array processing, normalizations have to be applied acounting for source mechanisms and depths. Normalizations can be implemented as deconvolutions in time for each event. Furthermore, x_0^s in the above would be the location of a master event.

Double beamforming is obtained as (changing order of integration §)

$$(\mathrm{B}u)(y_0,\pi^s,\pi^r) = \frac{1}{2\pi}\int (T_{\lambda=1}u)(y_0,\tau\,(\pi^s,\pi^r,1))\,\mathrm{d}\tau. \tag{44}$$

(Typically, one subjects the data, u, to rotations to separate the polarizations and identify a seismic phase prior to applying double beamforming.) Double beamforming aims to estimate the variables (t_0,π^s,π^r), that is, (y_0,π_0^s,π_0^r) for events (singularities) in the data ([18]).

In our procedure, T_λ is replaced by the curvelet transform, C, while the integration over τ is no longer carried out. One can think of setting $\lambda = 2^k$ and identifying in (42) x_0 with x_j and ξ_0 with $2^k\nu$, and $U(x_0,\xi_0)$ with u_γ. In [59] discrete, almost symmetric wave packets and a higher-dimensional curvelet transform, acting on unevenly sampled data, have been developed. To find a sparse decomposition, one can follow an ℓ^1 optimization procedure [60]; we arrive at the index set $\mathcal{S}$.

Imaging. Imaging structure in the earth with seismic array data has been carried out through first-order Taylor expansions in x^r (and x^s) of the traveltime function (appearing in the canonical relation Λ) about x_0^r (and x_0^s, with corresponding traveltime t_0) relative to a prescibed scattering point (x_0 in (7)) defining slowness vectors; in the process, the amplitude (or semblance) of the outcome using these traveltime derived slowness vectors is displayed on a grid of scattering points.

§ If g were replaced by a Gaussian function, $T_{\lambda=1}$ would be identified with the Gabor transform.

Here, we apply Theorem 4.2, or 4.1, to generate an image from the decomposed data, which requires dynamical ray tracing computations. The image is decomposed into curvelets yielding coefficients pertaining to the index set $\mathit{\Gamma}_1$.

Reconstruction. For each index γ in the set $\mathit{\Gamma}_1$, the diagonal entry $[\widetilde{N}]_{\gamma\gamma}$ is computed. This is most straightforwardly done by methods of so-called demigration-remigration, that is, following (38). One can apply the right-hand sides of (38) to the image obtained in the preceding step, and estimate a diagonal matrix acting on the curvelet-transformed image that recovers the result. One then evaluates the inverse diagonal matrix (cf. (40)) and applies the result to the outcome of the imaging step to obtain the reconstruction (solution to the normal equations) on the index set $\mathit{\Gamma}_0$.

Acknowledgment

The authors would like to thank A. Deuss for stimulating discussions on beamforming.

Appendix A. Dyadic Parabolic Decomposition and "Curvelets"

We introduce boxes (along the ξ_1-axis, that is, $\xi' = \xi_1$)

$$B_k = \left[\xi'_k - \frac{L'_k}{2}, \xi'_k + \frac{L'_k}{2}\right] \times \left[-\frac{L''_k}{2}, \frac{L''_k}{2}\right]^{n-1},$$

where the centers ξ'_k, as well as the side lengths L'_k and L''_k, satisfy the parabolic scaling condition

$$\xi'_k \sim 2^k, \quad L'_k \sim 2^k, \quad L''_k \sim 2^{k/2}, \quad \text{as } k \to \infty.$$

Next, for each $k \geq 1$, let ν vary over a set of approximately $2^{k(n-1)/2}$ uniformly distributed unit vectors. (We can index ν by $\ell = 0, \ldots, N_k - 1$, $N_k \approx \lfloor 2^{k(n-1)/2} \rfloor$: $\nu = \nu(\ell)$ while we adhere to the convention that $\nu(0) = e_1$ aligns with the ξ_1-axis.) Let $\Theta_{\nu,k}$ denote a choice of rotation matrix which maps ν to e_1, and

$$B_{\nu,k} = \Theta_{\nu,k}^{-1} B_k.$$

In the (co-)frame construction, we have two sequences of smooth functions, $\widehat{\chi}_{\nu,k}$ and $\widehat{\beta}_{\nu,k}$, on $\mathbb{R}^n$, each supported in $B_{\nu,k}$, so that they form a co-partition of unity

$$\widehat{\chi}_0(\xi)\widehat{\beta}_0(\xi) + \sum_{k\geq 1}\sum_{\nu} \widehat{\chi}_{\nu,k}(\xi)\widehat{\beta}_{\nu,k}(\xi) = 1, \tag{A.1}$$

and satisfy the estimates

$$|\langle \nu, \partial_\xi\rangle^j \, \partial_\xi^\alpha \widehat{\chi}_{\nu,k}(\xi)| + |\langle \nu, \partial_\xi\rangle^j \, \partial_\xi^\alpha \widehat{\beta}_{\nu,k}(\xi)| \leq C_{j,\alpha}\, 2^{-k(j+|\alpha|/2)}.$$

We then form

$$\widehat{\psi}_{\nu,k}(\xi) = \rho_k^{-1/2}\widehat{\beta}_{\nu,k}(\xi)\,, \quad \widehat{\varphi}_{\nu,k}(\xi) = \rho_k^{-1/2}\widehat{\chi}_{\nu,k}(\xi), \tag{A.2}$$

with ρ_k the volume of B_k. These functions satisfy the estimates

$$\left.\begin{matrix}|\varphi_{\nu,k}(x)| \\ |\psi_{\nu,k}(x)|\end{matrix}\right\} \leq C_N 2^{k(n+1)/4}\,(\,2^k|\langle\nu,x\rangle| + 2^{k/2}\|x\|\,)^{-N}$$

for all N. To obtain a (co-)frame, one introduces the integer lattice: $X_j := (j_1,\ldots,j_n)$, the dilation matrix

$$D_k = \frac{1}{2\pi}\begin{pmatrix} L'_k & 0_{1\times n-1} \\ 0_{n-1\times 1} & L''_k I_{n-1}\end{pmatrix}, \qquad \det\, D_k = (2\pi)^{-n}\rho_k,$$

and points $x_j = \Theta_{\nu,k}^{-1} D_k^{-1} X_j$. The frame elements ($k \geq 1$) are then defined in the Fourier domain as

$$\widehat{\varphi}_\gamma(\xi) = \rho_k^{-1/2}\widehat{\chi}_{\nu,k}(\xi)\,\exp[-\mathrm{i}\langle x_j,\xi\rangle], \quad \gamma = (x_j,\nu,k), \tag{A.3}$$

and similarly for $\widehat{\psi}_\gamma(\xi)$. We obtain the transform pair

$$v_\gamma = \int v(x)\overline{\psi_\gamma(x)}\,\mathrm{d}x, \qquad v(x) = \sum_\gamma v_\gamma\varphi_\gamma(x) \tag{A.4}$$

with the property that $\sum_{\gamma':\ k'=k,\ \nu'=\nu} v_{\gamma'}\widehat{\varphi}_{\gamma'}(\xi) = \widehat{v}(\xi)\widehat{\beta}_{\nu,k}(\xi)\widehat{\chi}_{\nu,k}(\xi)$, for each ν, k.

Remark. If we write $\widehat{v}_{\nu,k}(\xi) = \rho_k^{1/2}\widehat{v}(\xi)\widehat{\beta}_{\nu,k}(\xi)$, the curvelet transform pair attains the form of a quadrature applied to the convolution,

$$v(x) = \sum_{\nu,k} v_{\nu,k} * \varphi_{\nu,k}(x)\,. \tag{A.5}$$

This observation can be exploited to obtain sparse approximations, of v, by sums of wavepackets [22].

References

[1] B. Romanowicz. Global mantle tomography: Progress status in the past 10 years. *An. Rev. Earth Planet. Sci.*, 31:303–328, 2003.

[2] Ö. Yilmaz. *Seismic data processing*. Society of Exploration Geophysicists, Tulsa, 1987.

[3] G. Beylkin. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform. *J. Math. Phys.*, 26:99–108, 1985.

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

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

[6] M.V. De Hoop, R. Burridge, C. Spencer, and D. Miller. Generalized Radon Transform/amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media. In *Mathematical Methods in Geophysical Imaging II*, pages 15–27. Proceedings of SPIE – Volume 2301, 1994.

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

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

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

[10] M.V. De Hoop, C. Spencer, and R. Burridge. The resolving power of seismic amplitude data: An anisotropic inversion/migration approach. *Geoph.*, 64:852–873, 1999.

[11] M.V. De Hoop and S. Brandsberg-Dahl. Maslov asymptotic extension of generalized Radon transform inversion in anisotropic elastic media: A least-squares approach. *Inverse Problems*, 16:519–562, 2000.

[12] C.C. Stolk and M.V. De Hoop. Microlocal analysis of seismic inverse scattering in anisotropic, elastic media. *Comm. Pure Appl. Math.*, 55:261–301, 2002.

[13] E.J. Candès and D. Donoho. New tight frames of curvelets abd optimal representations of objects with piecewise-C^2 singularities. *Comm. Pure Appl. Math.*, 57:219–266, 2004.

[14] E.J. Candès and D. Donoho. Continuous curvelet transform: I. Resolution of the wavefront set. *Applied and Computational Harmonic Analysis*, 19:162–197, 2005.

[15] E.J. Candès and D. Donoho. Continuous curvelet transform: II. Discretization and frames. *Applied and Computational Harmonic Analysis*, 19:198–222, 2005.

[16] E.J. Candès, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *SIAM Multiscale Model. Simul.*, 5-3:861–899, 2006.

[17] A.P. Kiselev and M.V. Perel. Highly localized solutions of the wave equation. *J. Math. Phys.*, 41:1934–1955, 2000.

[18] F. Scherbaum, F. Krüger, and M. Weber. Double beam imaging: Mapping lower mantle heterogeneities using combinations of source and receiver arrays. *J. Geophys. Res.*, 102:507–522, 1997.

[19] F. Krüger, M. Baumann, F. Scherbaum, and M. Weber. Mid mantle scatterers near the Mariana slab detected with a double array method. *Geophys. Res. Lett.*, 28:667–670, 2001.

[20] D. Davies, E.J. Kelly, and J.R. Filson. Vespa process for analysis of seismic signals. *Nature Phys. Sci.*, 232:8–13, 1971.

[21] M.A.H. Hedlin, J.B. Minster, and J.A. Orcutt. Beam-stack imaging using a small aperture array. *Geophys. Res. Lett.*, 18:1771–1774, 1991.

[22] F. Andersson, M. Carlsson, and M.V. De Hoop. Dyadic parabolic decomposition and approximation of functions by sums of wave packets. *Applied and Computational Harmonic Analysis*, submitted, 2008.

[23] B. Hua and G.A. McMechan. Parsimonious 2D prestack kirchhoff depth migration. *Geophysics*, 68:1043–1051, 2003.

[24] F. Akbar, M. Sen, and P. Stoffa. Pre-stack plane-wave kirchhoff migration in laterally varying media. *Geophysics*, 61:1068–1079, 1996.

[25] F. Liu, D.W. Hanson, N.D. Whitmore, R.S. Day, and R.H. Stolt. Toward a unified analysis for source plane-wave migration. *Geophysics*, 71:S129, 2006.

[26] C.C. Mosher, D.J. Foster, and S. Hassanzadeh. Seismic imaging with offset plane waves. In *Mathematical Methods in Geophysical Imaging IV*, pages 52–63. Proceedings of SPIE – Volume 2822, 1996.

[27] D.J. Foster, C.C. Mosher, and S. Jin. Offset plane wave migration. In *Expanded Abstracts*, pages B–14. Eur. Assoc. Expl. Geophys., 2002.

[28] Y. Zhang, J. Sun, C. Notfors, S. Gray, L. Chernis, and J. Young. Delayed-shot-3D prestack depth migration. In *Expanded Abstracts*, page 1027. Society of Exploration Geophysicists, 2003.

[29] M.G. Bostock, S. Rondenay, and J. Shragge. Multi-parameter two-dimensional inversion of scattered teleseismic body waves, 1, Theory for oblique incidence. *J. Geophys. Res.*, 106:30,771–30,782, 2001.

[30] S. Rondenay, M.G. Bostock, and J. Shragge. Multi-parameter two-dimensional inversion of scattered teleseismic body waves, 3, Application to the Cascadia 1993 data set. *J. Geophys. Res.*, 106:30,795–30,808, 2001.

[31] P. Wang, M.V. De Hoop, R.D. Van der Hilst, P. Ma, and L. Tenorio. Imaging of structure at and near the core mantle boundary using a generalized Radon transform: 1. Construction of image gathers. *J. Geophys. Res.*, 111, B12304:doi:10.1029/2005JB004241, 2006.

[32] R.D. Van der Hilst, M.V. De Hoop, P. Wang, S.-H. Shim, P. Ma, and L. Tenorio. Seismo-stratigraphy and thermal structure of Earth's core-mantle boundary region. *Science*, 315:1813–1817, 2007.

[33] Q. Cao, P. Wang, M.V. De Hoop, R.D. Van der Hilst, and R. Lamm. High-resolution imaging of upper mantle discontinuities with SS precursors. *Physics of Earth and Planetary Interiors*, submitted, 2008.

[34] S. Rost and T. Thomas. Array seismology: Methods and applications. *Rev. of Geophys.*, 40:Art. No. 1008, 2002.

[35] S. Rondenay, M.G. Bostock, and K.M. Fischer. Multichannel inversion of scattered teleseismic body waves: Practical considerations and applicability. In A. Levander and G. Nolet, editors, *Seismic Data Analysis and Imaging With Global and Local Arrays*, volume 157, pages 187–204. Am. Geophys. Un., Geophysical Monograph Series, 2005.

[36] J. Capon. Investigation of long-period noise at the large aperture seismic array. *J. Geophys. Res.*, 74:3182–3194, 1969.

[37] L. Riabinkin. Fundamentals of resolving power of controlled directional reception (CDR) of seismic waves. In *Slant Stack Processing*, pages 36–60. Society of Exploration Geophysicists (Tulsa), 1991.

[38] A. Deuss, S.A.T. Redfern, K. Chambers, and J.H. Woodhouse. The nature of the 660-kilometer discontinuity in Earth's mantle from global seismic observations of PP precursors. *Science*, 311:198–201, 2006.

[39] M.P. Flanagan and P.M. Shearer. Global mapping of topography on transition zone velocity discontinuities by stacking SS precursors. *J. Geophys. Res.*, 103:2673–2692, 1998.

[40] S. Lüth, S. Buske, R. Giese, and A. Goertz. Fresnel volume migration of multicomponent data. *Geophysics*, 70:S121–S129, 2005.

[41] T. Kito, A. Rietbrock, and C. Thomas. Slowness-backazimuth weighted migration: A new array approach to a high-resolution image. *Geophys. J. Int.*, 169:1201–1209, 2007.

[42] V. Guillemin. On some results of Gel'fand in integral geometry. In *Pseudodifferential Operators and Applications*, pages 149–155, Providence, RI, 1985. AMS.

[43] L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume IV. Springer-Verlag, Berlin, 1985.

[44] L. Hörmander. *The Analysis of Linear Partial Differential Operators*, volume III. Springer-Verlag, Berlin, 1985.

[45] H. Douma and M.V. De Hoop. Explicit expressions for pre-stack map time-migration in isotropic and VTI media and the applicability of map depth-migration in heterogeneous anisotropic media. *Geophysics*, 71:S13–S28, 2006.

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

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

[48] P. Ma, P. Wang, L. Tenorio, M.V. De Hoop, and R.D. Van der Hilst. Imaging of structure at and near the core mantle boundary using a generalized Radon transform: 2. Statistical inference of singularities. *J. Geophys. Res.*, 112, B08303:doi:10.1029/2006JB004513, 2007.

[49] H.F. Smith. A Hardy space for Fourier integral operators. *Jour. Geom. Anal.*, 8:629–654, 1998.

[50] H.F. Smith. A parametrix construction for wave equations with $C^{1,1}$ coefficients. *Ann. Inst. Fourier, Grenoble*, 48:797–835, 1998.

[51] F. Andersson, M.V. De Hoop, H.F. Smith, and G. Uhlmann. A multi-scale approach to hyperbolic evolution equations with limited smoothness. *Comm. Partial Differential Equations*, 33:988–1017, 2008.

[52] A. Seeger, C.D. Sogge, and E.M. Stein. Regularity properties of Fourier integral operators. *Annals Math.*, 133:231–251, 1991.

[53] H. Douma and M.V. De Hoop. Leading-order seismic imaging using curvelets. *Geophysics*, 72:S231–S248, 2007.

[54] E.J. Candès and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Comm. Pure Appl. Math.*, 58:1472-1528, 2005.

[55] E.J. Candès, L. Demanet, and L. Ying. Fast computation of Fourier integral operators. *SIAM J. Sci. Comput.*, 29:2464–2493, 2007.

[56] W.W. Symes. Optimal scaling for reverse-time migration. *Technical Report 06-18*, Department of Computational and Applied Mathematics, Rice University, 2006.

[57] F.J. Herrmann, P.P. Moghaddam, and C.C. Stolk. Sparsity- and continuity-promoting seismic image recovery with curvelet frames. *Applied and Computational Harmonic Analysis*, in print, 2007.

[58] J. Bros and D. Iagolnitzer. Support essentiel et structure analytique des distributions. *Séminaire Goulaouic-Lions-Schwartz*, exp. no. 19, 1975-1976.

[59] A. Duchkov, S. Dong, F. Andersson, and M.V. De Hoop. Discrete, almost symmetric wave packets and higher-dimensional "curvelet" transform: A parallel algorithm. *preprint*, 2008.

[60] I. Daubechies, M. Defrise, and C. de Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.*, LVII:1413–1457, 2004.