Multi-scale propagation and imaging with wave packets

Herwig Wendt, Maarten V. de Hoop and Fredrik Andersson, Purdue University*

SUMMARY

Wave propagation, downward continuation, and imaging can be expressed in terms of members of a certain class of Fourier integral operators (FIOs). We present an expansion and discretization of such operators following a multi-scale approach, making use of wave packets or "curvelets". The discretization leads to an algorithm exploiting compression of the "data", via approximation by sums of wave packets, to which the operator is applied. We demonstrate the accuracy of our approach in a couple of basic, numerical examples.

INTRODUCTION

Wave propagation, downward continuation, and imaging can be expressed in terms of members of a certain class of Fourier integral operators (FIOs). In the case of imaging, in the presence of caustics, an extension needs to be constructed (Stolk and de Hoop (2002, 2005, 2006)) to arrive at a description in this class. The action of an FIO F in this class on a function u is given by [*]

$$(Fu)(y) = \int a(y,\xi)\exp(\mathrm{i}S(y,\xi))\hat{u}(\xi)\mathrm{d}\xi, \qquad (1)$$

where the amplitude function $a(y,\xi)$ and the generating function $S(y,\xi)$ are determined by the ray geometry of the background medium, and where the latter describes the propagation of singularities by the operator according to de Hoop et al. (2009)

$$\left(\frac{\partial S}{\partial \xi},\xi\right) \rightarrow \left(y,\frac{\partial S}{\partial y}\right);$$

we denote this map by χ. The operator F has a sparse matrix representation with respect to the frame of curvelets Smith (1998). We will refer to curvelets (Candès et al. (2006) and references therein) by their original name "wave packets".

Recently, De Hoop *et al.* (de Hoop et al. (2009)) proposed an approximation of the action of operator F on a single wave packet, with error $\mathscr{O}(2^{-k/2})$ for a wave packet at frequency scale 2^k (cf. (5) below). The work presented here elaborates on this result and proposes a numerical procedure for evaluating it in practice: First, a numerical procedure for obtaining the tensor product representation of the complex exponential involved in the approximation is proposed. Then, we develop a strategy for obtaining a efficient parallel computational scheme for approximating the action of the operator F on input functions composed of wave packets. We show numerical results both for propagation, in particular the formation of caustics, and imaging.

[*]We denote the Fourier transform of a function by $\hat{\cdot}$, and by ξ the Fourier (frequency) variables.

OPERATOR EXPANSION AND APPROXIMATION

We briefly summarize a result in de Hoop et al. (2009). We let $\varphi_\gamma(x)$, $\gamma=(j,\nu,k)$, denote a wave packet with central position $x_j^{\nu,k}$ and orientation ν at scale k, that is central wave vector $2^k\nu$. An example of a discrete, almost symmetric wave packet, which have the necessary properties to provide the approximation with estimates given below, is shown in Fig. 1.

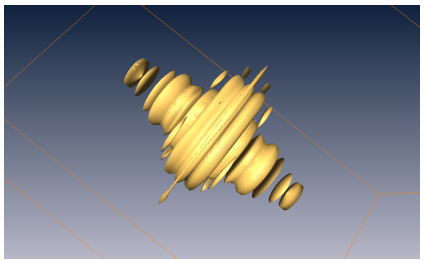

Figure 1: A discrete, almost symmetric wave packet (Duchkov et al. (2009)).

We have

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

where $\hat{\varphi}_\gamma(\xi)=\rho_k^{-1/2}\hat{\chi}_{\nu,k}(\xi)\exp[-\mathrm{i}\langle\xi,x_j\rangle]$ is the Fourier transform of φ_γ. The matrix representation, $[F]$, of F is defined as $[F]_{\gamma'\gamma}=\langle\varphi_{\gamma'},F\varphi_\gamma\rangle$. The strategy of de Hoop et al. (2009) for obtaining an approximation of $(F\varphi_\gamma)(y)$ is to replace S by a sufficient number of terms of its Taylor expansion in ξ on the frequency support of the wave packet φ_γ.

We introduce the "coordinate transform",

$$y \rightarrow T_{\nu,k}(y) = \frac{\partial S}{\partial \xi}(y,\nu), \qquad (3)$$

which describes the propagation of the wave packet φ_γ along a ray according to geometrical optics, without altering the support or shape of the wave packet. Refining the approximation by including the leading term of the second-order Taylor expansion of S on the support of $\hat{\varphi}_\gamma$, and neglecting higher order terms yields the result (de Hoop et al., 2009, Theorem 4.1): *With functions* $T_{\nu,k}(y)$ *defined by* (3), $\alpha^1_{r;\nu,k}(y)$ *and* $\hat{\alpha}^2_{r;\nu,k}(\xi)$ *defined by:*

$$a(y,\nu)\exp\left(\mathrm{i}\frac{1}{2}\frac{\xi_2^2}{\xi_1}\frac{\partial^2 S}{\partial\xi_2^2}(y,\nu)\right) \approx \sum_{r=1}^{R}\alpha^1_{r;\nu,k}(y)\hat{\alpha}^2_{r;\nu,k}(\xi), \qquad (4)$$

where $\xi=(\xi_1,\xi_2)$ *with* ξ_1 *representing the coordinate in the* ν *direction, one may express with* $R\sim k/\log(k)$*:*

$$(F\varphi_\gamma)(y) = \sum_{r=1}^{R}\alpha^1_{r;\nu,k}(y)\left(\alpha^2_{r;\nu,k}*\varphi_\gamma\right)(T_{\nu,k}(y)) + 2^{-k/2}f_\gamma, \qquad (5)$$

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

Thus $(F\varphi_\gamma)(y)$ can be approximated, to order $\mathcal{O}(2^{-k/2})$, by a sum over R modified wave packets $\tilde{\varphi}_{r;\gamma}(x) = (\alpha^2_{r;\nu,k} * \varphi_\gamma)(x)$ with amplitude corrections $\alpha^1_{r;\nu,k}(y)$, followed by a coordinate transform $T_{\nu,k}(y)$. The functions $T_{\nu,k}(y)$, $\alpha^1_{r;\nu,k}(y)$ and $\hat{\alpha}^2_{r;\nu,k}(\xi)$ do not depend on the position x_j of the wave packet φ_γ, but only on the scale k and orientation ν, and are hence the same for all wave packets with given orientation ν at scale k.

We develop an algorithm for the evaluation of (5) at discrete frequency and output points ξ_n and y_m, respectively[†]. We write

$$(F\varphi_\gamma)(y) = \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y) \sum_{\xi} e^{\mathrm{i}\langle T_{\nu,k}(y),\xi\rangle} \hat{\alpha}^2_{r;\nu,k}(\xi)\hat{\varphi}_\gamma(\xi) + 2^{-k/2} f_\gamma. \quad (6)$$

COMPUTATIONAL APPROACH

Tensor product representation

The functions $\alpha^1_{r;\nu,k}(y)$ and $\hat{\alpha}^2_{r;\nu,k}(\xi)$ in the separated representation (4) need to be found for each orientation ν and scale k. Obviously, the representation should have small separation rank R, and should be obtainable with low computational complexity. The amplitude term $a(y,\nu)$ on the left hand side of (4) does not depend on ξ, can hence always be trivially accounted for in the final separation.

The argument of the complex exponential in (4) is separated in y and ξ, and the left-hand side can hence be arranged as a matrix $[\Gamma]$, say, with row and column indices m and n pointing at the discrete sets y_m and ξ_n, respectively. Hence, a discrete tensor product representation as in (4) can always be obtained by truncating the singular value decomposition of $[\Gamma]$: $[\Gamma] = [B_1][B_2]$ with $[B_1] = [U]$, $[B_2] = [\Lambda][V]^*$, where the left and right singular vectors in $[U]$ and $[V]$ correspond with singular values collected in $[\Lambda]$ that lie above a certain threshold that determines the precision of the truncated decomposition. This generic approach is computationally expensive, since it requires the computation of the full SVD of $[\Gamma]$, while only a small number, R, of singular vector – singular value pairs will yield a significant contribution to the final expansion. Candès *et al.* (Candès et al. (2007)) proposed to exploit the structure (i.e. low rank) of a complex exponential matrix in a similar problem by an algorithm based on random subsets of its rows and columns, which in average reduces the size of the SVD problem. Their algorithm is based on work by Kapur *et al.* (Kapur and Long (1997)) [‡].

Here, we follow an approach that does not rely on a SVD and is tailored to directly match the specific form of (4). We exploit the fact that the leading-order term of the above mentioned Taylor expansion yielding (3) captures the highly oscillatory part of $\exp[\mathrm{i}S(y,\xi)]$ on the support of $\hat{\varphi}_\gamma$. We first construct a computationally inexpensive, but potentially high rank, expansion of the complex exponential using orthogonal polynomials, and then use a fast alternative least square (ALS) algorithm with low complexity (Beylkin and Mohlenkamp (2005)) to find a representation with $\sim k/\log(k)$ terms.

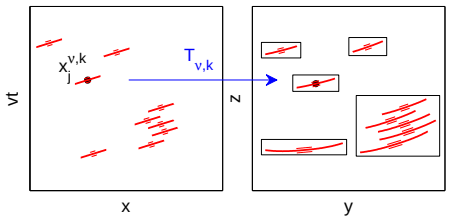

Figure 2: Image regions corresponding to data wave packets.

Coordinate transform

The sum over ξ in (6) corresponds to a non-equispaced (inverse) Fourier transform from the discrete set ξ to space points which are, for each k and ν, determined by the coordinate transform $T_{\nu,k}(y)$.

We start with a decomposition of u into wave packets,

$$u(x) = \sum_\gamma c_\gamma \varphi_\gamma(x) = \sum_\gamma \langle \varphi_\gamma, u\rangle \varphi_\gamma(x) = \sum_\xi \sum_{\nu,k} e^{\mathrm{i}\langle x,\xi\rangle} \hat{u}(\xi)\hat{\chi}_{\nu,k}(\xi)^2. \quad (7)$$

We obtain

$$(Fu)(y) = \sum_{\nu,k} \sum_{r=1}^{R} \alpha^1_{r;\nu,k}(y) \sum_\xi e^{\mathrm{i}\langle T_{\nu,k}(y),\xi\rangle} \hat{\alpha}^2_{r;\nu,k}(\xi)\hat{u}(\xi)\hat{\chi}_{\nu,k}(\xi)^2 + \mathcal{O}(2^{-k_{max}/2}), \quad (8)$$

where the non-equispaced Fourier transform ($\sum_\xi[\,\cdot\,]$) can be taken per box ν,k, since neither the functions $\alpha^{\{1,2\}}_{r;\nu,k}$ nor the coordinate transform $T_{\nu,k}(y)$ depend on the index j. Hence, from a computational point of view, approximating the action of F on the function $u(x)$ by wave packets induces the following differences with respect to (7) (i.e., the recomposition of the function $u(x)$ from its constituting wave packets):

1. An additional sum over the elements of the tensor product representation, which increases the computational complexity by a factor R. The sum over r can be taken either inside or outside of the sum $\sum_{\nu,k}$.
2. Due to the (ν,k)-dependence of the coordinate transform $T_{\nu,k}(y)$ and the amplitudes $\alpha^1_{r;\nu,k}(y)$ in (8), the non-equidistant Fourier transform $\sum_\xi[\,\cdot\,]$ can not be calculated over the entire ξ-domain at once, but needs to be calculated for each (ν,k) separately. The sum over ξ has to be calculated only for ξ points within the support of the wave packets at (ν,k), yet still needs to be evaluated for each y. This increases the computational complexity by a factor proportional to the number of (ν,k)-pairs, i.e., in 2D, approximately by $\sqrt{N}$.

[†] With slight abuse of notation, we will continue to write ξ and y for these discrete sets.

[‡] See e.g. Golub and Loan (1994) for an overview of deterministic methods for reducing the complexity of the SVD for matrices with certain specific structures

Propagator and GRT compression

A - Initialization fix numerical precision and define regular discrete grid on image domain y

B - Action per ν,k For each orientation ν with non-zero coefficient c_γ at scale k:

1. Fix image region y corresponding to data domain x of interest
2. Evaluate $\frac{\partial S}{\partial \xi}(y,\nu)$ and $\frac{\partial^2 S}{\partial \xi_2^2}(y,\nu)$
3. Obtain tensor product representation (4) with rank reduction
4. Evaluate $\sum_\xi [\,\cdot\,]$ in (8) for each tensor product term r
5. Evaluate $\sum_r$ in (8)

C - Action $(Fu)(y)$ Evaluate $\sum_{\nu,k}$ in (8)

Table 1: Outline of a parallel numerical scheme for evaluating (8).

Table 1 summarizes the main steps of our scheme for evaluating (8). All of the computationally costly operations are performed in step **B**, independently for each ν,k, and can hence be performed using massive parallelization. Step **A** merely consists in setting up a discrete target grid in the image domain by fixing a reference point y_0 and target resolution, such that the final image can be obtained as a simple sum of the contributions from different ν,k in step **C**.

Since wave packets with different orientation ν do not propagate to the same image region, in step **B**1, the image region corresponding to data wave packets at ν,k is determined in order to decrease the size of the y domain to be considered for each orientation. This can be performed either by an iterative search for y points for which $T_{\nu,k}(y)$ lies on the boundaries of the data domain x, or by inversion of $T_{\nu,k}(y)$, and by considering the decay properties of wave packets and the compression obtained in the image domain, as we illustrate in Fig. 2.

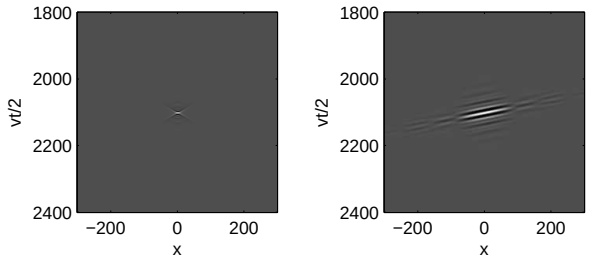

Figure 3: Common offset synthetic data: (a) bandlimited delta function, (b) a single constituent wave packet.

IMAGING – COMMON-OFFSET MIGRATION

We demonstrate the algorithm by imaging a bandlimited delta function at $(x=0, vt/2=2100)$ (units are meters). We imagine this function to be obtained by a superposition of wave packets, with orientations or slopes, p_u, within a certain range $[p_u-\Delta_p, p_u+\Delta_p]$. To simplify the computation of a and S, we use a constant velocity model Douma and de Hoop (2007). Let h denote the offset; we take $h/z \approx 0.05$. In Fig. 3 we show the synthetic data and one of its wave packet constituents. We have taken (in degrees from horizontal) $[p_u-\Delta_p, p_u+\Delta_p] = [63^\circ, 117^\circ]$. The data consists of 256×256 samples yielding $k_{max}=5$ available scales. The numerical precision for the tensor product expansion is set to $2\cdot 10^{-6}$ and, on average over scales and orientations, to $R=12$. In Fig. 4, we show images using a single wave packet at scale $k=3$ (cf. Fig 3(b)) as the data. The coordinate transform $T_{\nu,k}$ propagates the wave packet to the correct position in the image (Figs. 4(a) and 4(c)). With approximation (6), the wave packet is properly spreading, bending and shearing (Figs. 4(b) and 4(d)). In Fig. 5, we show images using the bandlimited delta function in Fig. 3(a) as the data. The image in Fig. 5(a) has gaps along the isochrone, resulting from the fact that it is obtained purely by propagation of wave packets along rays according to geometrical optics: The image of a singularity at large time / depth eventually breaks apart into its constituting wave packets; this was also observed

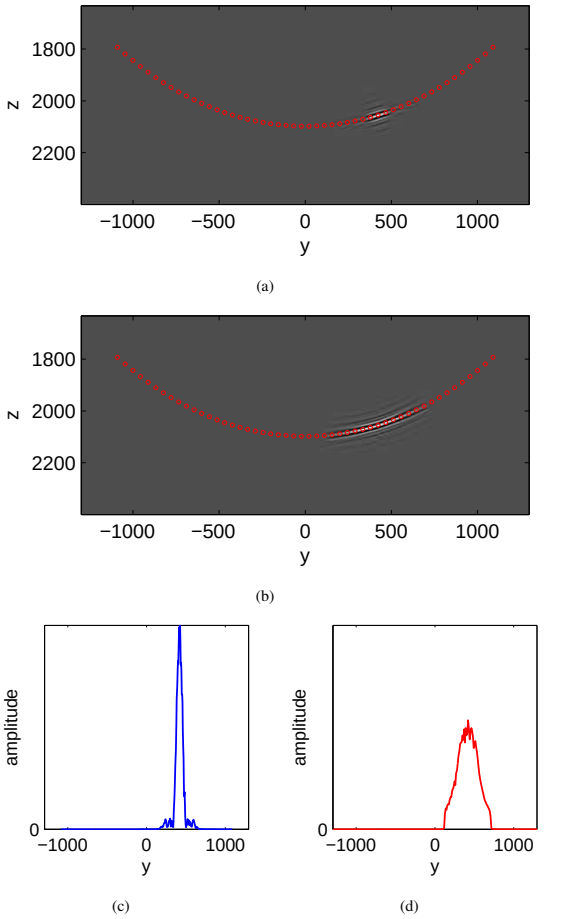

Figure 4: Images using a single wave packet at scale $k=3$ as the data (cf. Fig. 3(b)): (a) approximation with error $\mathcal{O}(2^0)$, (b) approximation (8), with error $\mathcal{O}(2^{-k/2})$, (c) amplitude profile of (a) measured along the associated isochron, (d) amplitude profile of (b) measured along the associated isochron.

in Douma and de Hoop (2007). In contrast, approximation (6) produces a satisfactory image (Fig. 5(b)): The contributing wave packets are spreading, bending and shearing such that they remain connected.

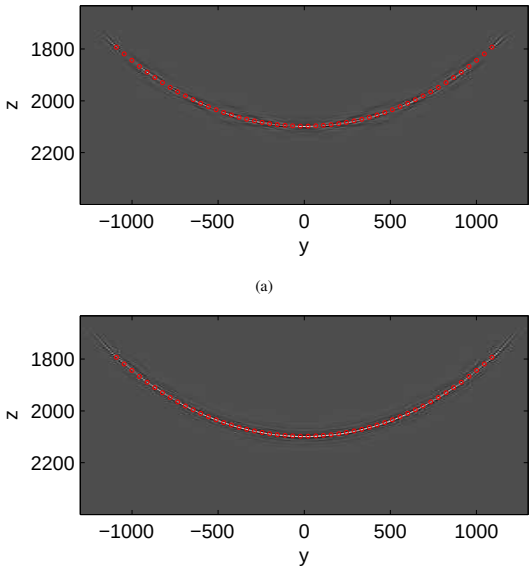

Figure 5: Images using the bandlimited delta function in Fig. 3(a) as the data: (a) approximation with error $\mathcal{O}(2^0)$, (b) approximation (8), with error $\mathcal{O}(2^{-k/2})$.

WAVE PROPAGATION – CAUSTICS

In the case of wave propagation, we write (y,t) for y and (ξ,τ) for ξ to make the evolution in time explicit. We let $T(y,x)$ denote the travel time. Then

$$S(y,t,\xi,\tau) = (\tau t - \tau T(y,x) + \langle \xi, x\rangle)|_{x=x(y,\xi)}, \qquad (9)$$

where $x(y,\xi)$ is the solution to $\frac{\partial T(x,y)}{\partial x} = \tau^{-1}\xi$. The representation is valid in points (y,ξ) where $\det \frac{\partial^2 T}{\partial x^2}|_{x=x(y,\xi)} \neq 0$.
In Fig. 6, we show the propagation of an elliptic wave front with multi-pathing in constant velocity medium from data approximated by a sum wave packets. The presence of caustics does not affect our method and approximation (6).

In the case of downward continuation, we encounter the so-called thin-slab double-square-root (DSR) propagators. The evolution parameter, here, is depth. We have $y = (s,r,t)$ and $\xi = (\xi_s,\xi_r,\omega)$. At depth z, the symbol of the single-square-root operator is given by

$$b(z,s,\xi_s,\omega) = \omega[c(z,s)^{-2} - \omega^{-2}|\xi_s|^2]^{1/2} + \text{l.o.t.}$$

and similarly for $b(z,r,\xi_r,\omega)$. The FIO becomes $F(z-\Delta,z)$, which acts on $u = u(z,.)$. We get $a(z,y,\xi) = 1$ and $S(z,y,\xi) = -\Delta[b(z,s,\xi_s,\omega) + b(z,r,\xi_r,\omega)]$.

Limited smoothness, full wave
We can construct also *full wave* solutions using wave packets

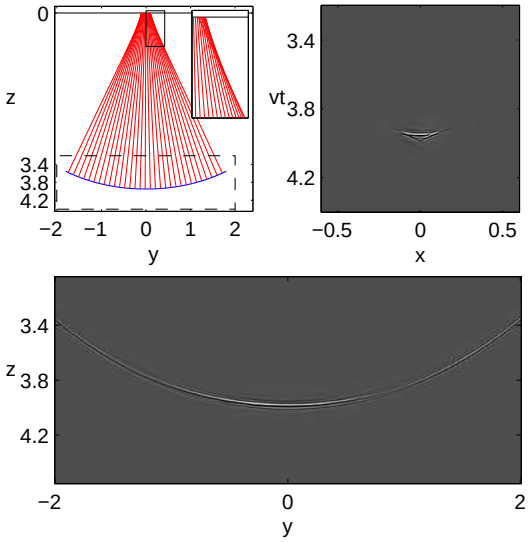

Figure 6: Elliptic wave front, corresponding rays and formation of caustics (top left), corresponding wave packet data (top right), and result of wave propagation (bottom).

Andersson et al. (2008). Let z denote the evolution parameter; the solution operator is then written as $F(z,z_0)$ if the initial values are specified at $z = z_0$. We assume, roughly, that the velocity model is in C^s (with Hölder regularity $s \geq 2$). In the full-wave solution, the propagation of singularities following ray theory is replaced by the following concentration property: For all $\alpha < s$

$$\sup_{\gamma'} \sum_{\gamma} 2^{2|k-k'|\alpha} \left(2^{\min(k,k')}\, \overline{d}(\gamma, \chi_{z,z_0}(\gamma'))^{2\alpha} |[F(z,z_0)]_{\gamma\gamma'}|^2 \leq C$$

where $\overline{d}(\gamma;\gamma') = 2^{-\min(k,k')} + d(x,\nu;x',\nu')$ if $\gamma = (x,\nu,k)$, $\gamma' = (x',\nu',k')$, in which

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

The matrix representations of the FIOs discussed above (following the asymptotic viewpoint), with $s \to \infty$, satisfy this estimate implying that they are sparse.

CONCLUSIONS

We introduced a multi-scale, numerical scheme for the application of operators representing wave propagation, downward continuation, and imaging in smoothly varying velocity models, following the analysis developed in de Hoop et al. (2009). The scheme exploits the compression of the data to which the operator is applied. In the case of imaging, the velocity-dependent "diffraction stack" is replaced by the velocity-independent wave packet transform, while a method of partial reconstruction emerges. In the case of downward continuation, we have obtained a wave packet or "curvelet" propagator reminiscent of the generalized screen propagator.

REFERENCES

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

Beylkin, G. and M. Mohlenkamp, 2005, Algorithms for numerical analysis in high dimensions: SIAM J. Sci. Comput., **26**, 2133–2159.

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

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

de Hoop, M., H. Smith, G. Uhlmann, and R. van der Hilst, 2009, Seismic imaging with the generalized radon transform: a curvelet transform perspective: Inverse Problems, **25**, 025005+.

Douma, H. and M. de Hoop, 2007, Leading-order seismic imaging using curvelets: Geophysics, **72**, 231–248.

Duchkov, A., F. Andersson, and M. de Hoop, 2009, Discrete almost symmetric wave packets and multi-scale geometric representation of (seismic) waves: preprint.

Golub, G. and C. V. Loan, 1994, Matrix computations (3rd ed.). Series in Mathematical Science: Johns Hopkins University Press.

Kapur, S. and D. Long, 1997, Ies3: A fast integral equation solver for efficient 3-dimensional extraction: Proc. IEEE/ACM Int. Conf. Computer Aided Design, 448–455.

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

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

Stolk, C. and M. V. de Hoop, 2005, Modeling of seismic data in the downward continuation approach: SIAM J. Appl. Math., **65**, 1388–1406.

———, 2006, Seismic inverse scattering in the downward continuation approach: Wave Motion, **43**, 579–598.