Preface

The electronic version of this report\textsuperscript{1} makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR] are promises by the author about the reproducibility of each figure result. Reproducibility is a way of organizing computational research that allows both the author and the reader of a publication to verify the reported results. Reproducibility facilitates the transfer of knowledge within SEP and between SEP and its sponsors.

\textbf{ER} denotes Easily Reproducible and are the results of processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, be easily available to all researchers (e.g. SEG-EAGE data sets), or be available in the SEP data library\textsuperscript{2}. We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Windows system and the software downloadable from our website (SEP makerules, SEPlib, and the SEP latex package), or other free software such as SU. Before the publication of the electronic document, someone other than the author tests the author’s claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

\textbf{CR} denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. SEP staff have only attempted to make sure that the makefile rules exist and the source codes referenced are provided. The primary reasons for the CR designation is that the processing requires 20 minutes or more, or commercial packages such as Matlab or Mathematica.

\textbf{M} denotes a figure that may be viewed as a movie in the web version of the report. A movie may be either ER or CR.

\textbf{NR} denotes Non-Reproducible figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.4 (using the Portland Group Fortran90 compiler), but the code should be portable to other architectures. Reader’s suggestions are welcome. For more information on reproducing SEP’s electronic documents, please visit <http://sepwww.stanford.edu/research/redoc/>.

\textsuperscript{1}http://sepwww.stanford.edu/private/docs/sep115
\textsuperscript{2}http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html/
# Table of Contents

## Imaging

*Shan G. and Biondi B.*, Imaging overturned waves by plane-wave migration in tilted coordinates ......................................................... 1  

*Shragge J. and Sava P.*, Adaptive phase-ray wavefield extrapolation ..................... 13  

*Shragge J. and Biondi B.*, Updating ray-coordinate systems with phase-rays .......... 29  

*Artman B., Shragge J., and Biondi B.*, Aliasing in prestack wavefield continuation migration 39  

*Rosales D. A. and Guitton A.*, Ocean-bottom hydrophone and geophone coupling ...... 57  

*Valenciano A. A., Tisserant T., and Biondi B.*, Surface boundary condition for one-way wave equation shot-profile migration ......................................................... 71  

*Vlad I. and Tisserant T.*, Improving the amplitude accuracy of downward continuation operators (Part 2) ......................................................... 81

## Multiples

*Brown M.*, Least-squares joint imaging of multiples and primaries applied to 3-D field data 89  

*Guitton A.*, Multidimensional multiple attenuation in complex geology: illustration on the Sigsbee2B dataset ......................................................... 109  

*Guitton A.*, Subtraction of surface-related multiples: adaptive subtraction vs. pattern recognition ......................................................... 127  

*Alvarez G., Biondi B., and Guitton A.*, Attenuation of diffracted multiples with an apex-shifted tangent-squared radon transform in image space .............. 139  

*Shan G. and Guitton A.*, Migration of surface-related multiples: tests on the Sigsbee2B dataset ......................................................... 153

## Velocity

*Biondi B. and Tisserant T.*, 3-D Angle-domain common-image gathers for migration velocity analysis ......................................................... 163  

*Tisserant T. and Biondi B.*, Residual move-out analysis with 3-D angle-domain common-image gathers ......................................................... 191  

*Sava P. and Biondi B.*, Sensitivity kernels for wave-equation migration velocity analysis 199
Sava P., Biondi B., and Etgen J., Diffraction-focusing migration velocity analysis with application to seismic and GPR data ...................................................... 213

Clapp R. G., Velocity uncertainty in tomography .................................................. 233

Guitton A., Claerbout J., and Lomask J., First-order lateral interval velocity estimates without picking .......................................................... 249

Vlad I., Velocity estimation for seismic data exhibiting focusing-effect AVO (Part 4) .... 265

Wolf K., Rosales D., Guitton A., and Claerbout J., Robust moveout without velocity picking 273

Inversion

Clapp M. L. and Clapp R. G., Regularized inversion for subsalt imaging: real data example 283

Clapp M. L., AVA effects of regularized least-squares inversion .............................. 295

Valenciano A. A. and Biondi B., Target-oriented shot-profile one-way wave equation inversion ................................................................. 305

Ji J., Conjugate-guided-gradient (CGG) method for robust inversion and its application to velocity-stack inversion .............................................. 315

Image processing

Liner C. and Clapp R. G., Dynamic programming and trace alignment: Part 2 .............. 337

Curry W., Regularizing Madagascar: PEFs from the data space? ......................... 347

Lomask J., Biondi B., and Shragge J., Improved image segmentation for tracking salt boundaries ................................................................. 357

Lomask J. and Guitton A., Analytical flattening with adjustable regularization ........... 367

Imaging of non-conventional data

Shragge J., Irving J., and Artman B., Shot-profile migration of GPR data ................. 383

Artman B., Draganov D., Biondi B., and Wapenaar K., Migration methods for passive seismic data ................................................................. 391

Wilson C. K., Shragge J., and Artman B., Teleseismic imaging with wave equation migration 399

Guitton A. and Vlad I., Imaging oceanic thermohaline structure with reflection seismology 411
Rock Physics

*Berryman J. G.*, Seismic waves in finely layered VTI media: Poroelasticity, Thomsen parameters, and fluid effects on shear waves ........................................ 417
*Berryman J. G.*, Poroelastic fluid effects on shear for rocks with soft anisotropy .......... 445
*Haines S.*, Electric fields created at the seismic source - a new electroseismic phenomenon 467

Computing

*Clapp R. G.*, Parallel datasets in SEPlib ....................................................... 479

SEP article published or in press, 2003-04 .................................................. 489
SEP phone directory .......................................................... 493
Research personnel .......................................................... 495
SEP sponsors for 2003-04 .......................................................... 503
Imaging overturned waves by plane-wave migration in tilted coordinates

Guojian Shan and Biondo Biondi

ABSTRACT
We image overturned waves by decomposing the source and receiver wavefields into plane-waves. For each plane-wave, we extrapolate the source and receiver wavefields in a tilted coordinate system and cross-correlate them to obtain Common Image Gathers (CIGs). The tilting angle for the coordinate system is determined by the propagation direction of the plane wave. In tilted coordinates, the propagation direction is close to the extrapolation direction, so we can image steeply dipping reflectors and overturned waves with the one-way wave equation. We can also obtain robust, dip-dependent angle-domain CIGs (ADCIGs) by the same method employed in reverse-time migration. These gathers provide moveout information and thus are very important for velocity analysis on steeply dipping reflectors. Since plane-wave migration needs no padding and the extrapolation is one-way wave equation based, our method is very efficient. We demonstrate our method on the Marmousi model by computing the Green’s function for a point source on the surface. We also apply our method to a North Sea real dataset with overturned events.

INTRODUCTION
Downward continuation migration (Claerbout, 1985) increases in popularity with the continuing development of computer power. It is accurate, and can handle lateral velocity change and multi-pathing naturally. It is used to image complex geological structure, especially sub-salt reflectors. However, the dip angle of reflectors is limited in downward continuation, which makes it difficult to image steeply dipping reflectors and overturned waves. Downward continuation is based on the one-way wave equation and extrapolates the wavefields in the downward direction, although in reality, waves propagate in all directions. When the propagation direction of waves is too far from the extrapolation direction, the accuracy of extrapolation is limited.

Kirchhoff methods can handle steeply dipping reflectors and overturned waves, but they are based on a high-frequency approximation and are less reliable for imaging complex geological structure, where multi-path events are present. Reverse-time migration (Whitmore, 1983; Baysal et al., 1983; Biondi and Shan, 2002), which is based on the two-way wave equation, can handle waves propagating in all directions, but it is still too expensive for today’s
Many methods have been developed to handle waves that propagate at large angles to the extrapolation direction. They solve the problem either by developing a more accurate extrapolation operator to image high-angle waves, such as Fourier finite difference (Ristow and Ruhl, 1994; Biondi, 2002) and generalized screen propagators (de Hoop, 1996; Huang and Wu, 1996), or by making the extrapolation direction closer to the propagation direction, such as beam migration (Hill, 2001; Gray et al., 2002; Albertin et al., 2001; Brandsberg-Dahl and Etgen, 2003), beamlet migration (Chen et al., 2002) and coordinate-transformation based methods (Zhang and McMechan, 1997; Etgen, 2002; Sava and Fomel, 2004). Our method is of the latter type. We decompose both source and receiver wavefields into plane-waves and run plane-wave migration (Rietveld, 1995; Duquet et al., 2001; Liu et al., 2002; Zhang et al., 2003) on each of them within a tilted coordinate system (Etgen, 2002).

Offset-domain CIGs for shot-profile migration and reverse-time migration are generated by cross-correlating the source and receiver wavefields with a horizontal shift (Rickett and Sava, 2002), and can be transformed into ADCIGs by slant stack (Sava and Fomel, 2003). However, CIGs obtained from downward continuation are contaminated by smearing noise at steeply dipping reflectors (Biondi and Shan, 2002). Reverse-time migration can provide both horizontal and vertical CIGs. Both are then merged into dip-dependent CIGs, which are robust and immune to the smearing noise (Biondi and Symes, 2003). By performing plane-wave migration in tilted coordinates, we can obtain similar dip-dependent ADCIGs. These CIGs are robust and are useful for velocity analysis in the presence of steeply dipping reflectors in the subsurface.

**TILTED CARTESIAN COORDINATES**

In Cartesian coordinates \((x, z)\), the one-way wave equation is

\[
\frac{\partial P}{\partial z} = \frac{i \omega}{v} \sqrt{1 + \frac{v^2}{w^2} \frac{\partial^2}{\partial x^2}} P. \tag{1}
\]

It can be obtained by factoring the two-way acoustic wave equation. If the coordinates are rotated by an angle \(\theta\), the two-way acoustic wave equation doesn’t change in the new coordinates \((x', z')\). Factoring the two-way wave equation in the tilted coordinates, we can obtain the same one-way wave equation. In the tilted coordinates, wavefields are now extrapolated in the \(z'\) direction, rather than in the downward direction as done in downward continuation. Also the source and receiver data are not on the line \(z' = 0\), but on \(x' \sin \theta - z' \cos \theta = 0\) (Figure 1a).

For a point source, the waves propagate in all directions, making it impossible to cover all propagation directions with only one Cartesian coordinate system. We decompose the point source into plane waves and extrapolate each plane wave in tilted coordinates, whose extrapolation direction \(z'\) is determined by the propagation direction of the plane wave. The propagation direction of the plane wave at the surface can be used as the extrapolation direction.
Figure 1: Tilted Cartesian coordinates: (a) sources and receivers are on the line $x' \sin \theta - z' \cos \theta = 0$ and the extrapolation direction is the $z'$ direction; (b) sources and receivers are usually very far from the reflection point for overturned waves and the opening angle is usually small.

$z'$ of the tilted coordinates. In this coordinate system, the source and receiver wavefields of overturned waves can usually be caught, since the opening angle between them is usually small and the propagation directions of the source and receiver wavefields are close to each other (Figure 1b).

Another advantage of using a plane-wave rather than a point source is that no padding is needed to image steeply dipping reflectors. To catch the energy from steeply dipping reflectors, the locations of sources and receivers are usually far from the reflector point (Figure 1b). Large extrapolation aperture is required to image these reflectors using a point-source method. In contrast, plane waves inherently have large apertures, which can easily cover the locations of source, receiver and reflector points. This greatly reduces the cost of imaging the steeply dipping reflectors.

ANGLE-DOMAIN CIGS IN TILTED CARTESIAN COORDINATES

In shot-profile migration, offset-domain CIGs can be obtained by cross-correlating the source and receiver wavefields with a horizontal shift (Rickett and Sava, 2002). The offset-domain CIGs can be transformed into ADCIGs by slant-stack (Sava and Fomel, 2003). In reverse-time migration, in addition to horizontal offset-domain CIGs, we have vertical offset-domain CIGs, which are obtained by cross-correlating the source and receiver wavefields with a vertical shift (Biondi and Shan, 2002). Both horizontal and vertical offset-domain CIGs can be transformed into ADCIGs and merged into dip-dependent CIGs, by transforming the horizontal and vertical
offsets into apparent geological offset as follows (Biondi and Symes, 2003):

\[
h_x = \frac{h_0}{\cos \alpha}, \quad (2)\\
h_z = \frac{h_0}{\sin \alpha}, \quad (3)
\]

where \(\alpha\) is the dip angle of the subsurface reflector, \(h_x\) is the horizontal offset, \(h_z\) is the vertical offset and \(h_0\) is the apparent geological offset.

Figure 2: The Marmousi model: (a) wavefields obtained by downward continuation with FFD; (b) wavefields obtained by plane-wave decomposition and extrapolation in tilted coordinates.

In the tilted coordinates, wavefields are extrapolated in the \(z'\) direction. The offset-domain CIGs are generated by cross-correlating the source and receiver wavefields with an \(x'\) direction shift. So the subsurface offset is in the \(x'\) direction. As with the apparent geological offset \(h_0\),
the $x'$ direction offset $h_{x'}$ can be transformed to horizontal and vertical offsets as follows:

$$h_x = \frac{h_{x'}}{\cos \theta},$$  \hspace{1cm} (4)

$$h_z = \frac{h_{x'}}{\sin \theta},$$  \hspace{1cm} (5)

where $\theta$ is the tilting angle in Figure 1. As for reverse-time migration, horizontal and vertical offset-domain CIGs can be transformed into ADCIGs and merged into dip-dependent ADCIGs. A simple way to merge them is with the following weights:

$$w_{h_x} = \cos^2 \alpha,$$  \hspace{1cm} (6)

$$w_{h_z} = \sin^2 \alpha,$$  \hspace{1cm} (7)

where $\alpha$ is the apparent dip angle of the reflector. Dip-dependent residual moveout (Shan and Biondi, 2003) can be used to analyze dip-dependent ADCIGs to provide useful moveout information for velocity analysis.

In tilted coordinates, the direction of subsurface offset is close to that of apparent geological offset, since the extrapolation direction of the wavefields is close to the propagation direction of the waves. Within a limited length of subsurface offsets, we can obtain much more accurate CIGs at steeply dipping reflectors than with standard downward continuation.

**EXAMPLES**

Our first example concerns waves propagating in the Marmousi model. Figure 2 presents (a) the wavefields obtained by downward continuation with Fourier finite difference (FFD) overlaid velocity model, and (b) the wavefields obtained by plane-wave decomposition and extrapolation of each plane wave with FFD in tilted coordinates overlaid velocity model. The source location is at $x = 5000$ m, and the traveltime is 3 s. In Figure 2a, although we extrapolate the wavefield with a wide-angle extrapolation operator, the wavefield is not accurate for large incident angles. In contrast, Figure 2b does not have this angle limitation, and the wavefield is extrapolated better for high-angle waves.

Our second example images the overturned waves of a 2-D line from the ELF L7D North Sea dataset. This dataset presents challenges caused by the interaction between the salt edge and a chalk layer. The migration velocity is a smooth version of a tomographic result, which was obtained by inverting the moveouts measured on ADCIGs computed from downward-continuation prestack migration (Clapp, 2001). The image from reverse-time migration (Biondi and Shan, 2002) shows the presence of overturned energy. Figure 3 shows (a) the image obtained by downward continuation, (b) the image obtained by reverse-time migration, (c) the image obtained by plane-wave migration in tilted coordinates and (d) the rays shot from the steeply dipping reflector. Note that, to save computational cost, we use large space-sampling in reverse-time migration, which make reflectors in Figure 3b not as continuous as the others. There are two steeply dipping reflectors at the edge of the salt, one of which is almost vertical. Rays from these two reflectors (Figure 3d) show that the waves are overturned.
Figure 3: Migration of North Sea real dataset: (a) the image from downward continuation; (b) the image from reverse-time migration; (c) the image from plane-wave migration in tilted coordinates; (d) the rays shot at $x = 5000\text{m}$, $z = 1900\text{m}$. guojian1-elf.image [CR]
Downward continuation (Figure 3a) doesn’t image these two steeply dipping reflectors, failing to handle the overturned energy. In contrast, both reverse-time prestack migration (Figure 3b) and plane-wave migration in tilted coordinates (Figure 3c) preserve the overturned energy and image the steeply dipping reflectors. The images of the steep reflectors by these two methods are comparable, but the cost of plane-wave migration in tilted coordinates is much less than that of reverse-time migration.

Figure 4 compares the horizontal CIGs at the surface location \( x = 4000m \) from plane-wave migration in tilted coordinates with those from reverse-time migration. Figure 4a and Figure 4c show offset and angle domain CIGs from reverse-time migration, respectively. Figure 4b and Figure 4d show offset and angle domain CIGs from plane-wave migration in tilted coordinates, respectively. The ADCIGs from plane-wave migration in tilted coordinates look similar to those from reverse-time migration.

Figure 5 compares the horizontal CIGs at the surface location \( x = 5000m \) from plane-wave migration in tilted coordinates with those from reverse-time migration. Panels are organized as in Figure 4. The vertical salt edge is located at surface location \( x = 5000m \), in the vertical interval \( 1500 < z < 2000m \). At the vertical salt edge, we show that the ADCIGs from reverse-time migration and from plane-wave migration in tilted coordinates are similarly smeared along the angle axis. This smearing noise makes horizontal CIGs a poor source of information for velocity analysis at these steeply dipping reflectors.

Figure 6 compares the vertical CIGs at depth \( z = 1900m \) from plane-wave migration in tilted coordinates with those from reverse-time migration. Panels are organized as in Figure
4 but in the vertical direction. The vertical ADCIGs from both methods are similar. In contrast to horizontal CIGs, vertical CIGs focus at the vertical salt edge and can provide useful information for velocity analysis at the salt edge.

Figure 7 shows the dip-dependent ADCIGs obtained by merging horizontal and vertical ADCIGs with the weights in equations (6) and (7). Comparing against horizontal and vertical ADCIGs, we find that dip-dependent ADCIGs are more robust. The smearing noise is avoided and both horizontal and vertical reflectors are well focused.

**CONCLUSION**

We image the overturned waves by decomposing the source and receiver wavefields into plane waves. Each plane wave is extrapolated in tilted Cartesian coordinates. Offset-domain CIGs are generated by cross-correlating the source and receiver wavefields, with a shift in the direction normal to the extrapolation direction. The offset-domain CIGs are decomposed into horizontal and vertical CIGs, which are merged into robust, dip-dependent ADCIGs.

We apply our method to a North Sea dataset and compare the image with the ones obtained from downward continuation and reverse-time migration. The image obtained by plane-wave migration in tilted coordinates is better than the image obtained from standard downward continuation. The steeply dipping salt edge is imaged by one-way plane-wave migration in tilted coordinates, but it is missing in the image obtained by downward continuation. The
Figure 6: Vertical CIGs at $z = 1850\text{m}$: (a) offset domain CIGs obtained by reverse-time migration; (b) offset domain CIGs obtained by plane-wave migration in tilted coordinates; (c) ADCIGs obtained by reverse-time migration; (d) ADCIGs obtained by plane-wave migration in tilted coordinates.

Plane-wave migration results are also comparable to those obtained by reverse-time migration and produces similar horizontal and vertical CIGs, while its computational cost is much less than reverse-time migration. The dip-dependent ADCIGs merged from horizontal and vertical CIGs are robust and provide useful moveout information for reflectors with a wide range of dips.

ACKNOWLEDGMENTS

We would like to thank TotalFinalElf for making the North Sea dataset available.

REFERENCES


Figure 7: Dip-dependent ADCIGs: Notice that both horizontal and vertical reflectors are focused very well.


A letter to the editor of the Stanford Daily

The Geophysics Department Student Body

Dear Editor,

William Curry, a regular contributor to the letters to the editor section and staunch critic of others’ moral shortcomings, has audaciously eschewed responsibility of the office newly placed in his charge. As one of the three geophysical representatives to the School of Earth Sciences graduate student advisory committee, the weighty and weekly responsibility of doughnut boy was bestowed on him Thursday last. With the hopes and needs the entire student body resting on the success of his mission, Mr. Curry did flagrantly snooze well past the scheduled bagel delivery hour, to the displeasure of the frenzied masses.

We call upon President Hennessy to immediately form a truth and reconciliation committee to establish the root cause of the Mr. Curry’s miserable failure in this his first responsibility in office. Should the committee find gross negligence or nefarious intent on his part, we ask Mr. Curry to recuse himself from office post-haste.

Sincerely,

The Geophysics Department Student Body

Stanford, May 4, 2004

2email: we.are.hungry.while.you.snooze@pangea.stanford.edu
Adaptive phase-ray wavefield extrapolation

Jeff Shragge and Paul Sava

ABSTRACT

Riemannian wavefield extrapolation (RWE) is a generalization of downward continuation to coordinate systems that closely conform to the orientation of extrapolated wavefields. If the coordinate system overturns, so does the computed wavefield, despite being extrapolated with a one-way solution to the acoustic wave-equation. This allows for accurate imaging of structures of arbitrarily steep dips with simple operators equivalent to standard extrapolators. An obvious question for RWE is which is an optimal coordinate system for a given velocity model. One option is to compute ray coordinates as a solution to the wide-band eikonal equation in a smoothed velocity model. However, this solution ignores the natural variability and frequency dependence of wavepaths in cases of complicated velocity models, for example under salt bodies. The solution advocated in this paper is a recursive bootstrap procedure where a frequency-dependent coordinate system is computed on-the-fly at every step from the gradient of the monochromatic wavefield phase of the preceding few steps, coupled with standard RWE.

INTRODUCTION

Wavefield extrapolation extends surface-recorded data to depth through application of a wave-equation operator. The choice of operator depends mainly on practical considerations (e.g. computer memory, total flop count); however, one persistent theoretical constraint is the degree of velocity model complexity. In laterally invariant media, closed-form Fourier-domain operators (single square root, SSR) can accurately extrapolate surface recorded wavefields up to 90° (Gazdag, 1978; Claerbout, 1985). However, such solutions are inapplicable in media characterized by lateral velocity variation, and approximate solutions to the SSR equation are employed. Consequently, the accuracy of the extrapolation operators degrades, particularly at high angles relative to the downward extrapolation axis, and more sophisticated procedures are required to ensure wavefield accuracy.

Improved wavefield extrapolation can be achieved in many ways. First, one may improve the high-angle accuracy of an operator while retaining a Cartesian computational grid. Examples of this include incorporating higher-order terms in the expansion of Fourier domain operators (e.g. Fourier finite-difference (Ristow and Ruhl, 1994), generalized screen propagator (de Hoop et al., 2000)), or using tilted Cartesian coordinate systems (Zhang and McMechan,
1997; Etgen, 2002; Shan and Biondi, 2004) that extend the accuracy of high-angle propagation. Second, seismic wavefields may be spatially partitioned into more manageable sections and then independently extrapolated in preferred directions. For example, decomposing data sections to form local beams for extrapolating along tubes of finite thickness (Hill, 2001; Albertin et al., 2001; Gray et al., 2002; Brandsberg-Dahl and Etgen, 2003).

A third option is to abandon the strictures of Cartesian coordinates altogether and represent the physics of one-way wavefield extrapolation in a generalized coordinate system that obeys the tenets of differential geometry (Guggenheimer, 1977). In particular, one could use a basis (or coordinate system) that conforms to where the wavefronts propagate (Sava and Fomel, 2004). In this reference frame low-angle operators remain applicable, and the extrapolation procedure is of high fidelity, even at arbitrarily large angles to depth axis. The strategy espoused in this paper is the latter: it is more prudent to adjust the coordinate system to conform better with the physics than to force the physics to work in Cartesian coordinates, or on an a priori spatial partition of the data or model space.

One judicious choice of non-Cartesian coordinate system is a basis derived from a suite of rays. In this approach, the natural wavefield extrapolation direction is travel-time along a ray, with orthogonal coordinates directed across the rayfront at a constant time step. However, unlike for Cartesian coordinates, the distance between adjacent rays may freely expand or contract according to the lateral variations in the velocity model. Thus, properly defining the coordinate metric requires additional parameters that account for the Jacobian-like coordinate spreading. Given a rayfield and the associated Jacobian parameters, the solution to the corresponding one-way acoustic wave-equation is generated in an ordinary fashion (Sava and Fomel, 2004). The ray-coordinate wavefield solution is then interpolated back to a Cartesian mesh through a simple mapping operation.

One of the practical difficulties of ray-coordinate-based wavefield extrapolation is developing a robust procedure for handling triplicating rayfields that naturally arise due to wavefield multipathing. In particular, we need to prevent numerical instabilities from arising when calculating coordinate Jacobian spreading and related parameters that require computing finite-difference derivatives at the ray-crossing locations. Sava and Fomel (2004) apply a regularization parameter that prevents division by zero. This procedure, though, can lead to anomalous extrapolation amplitudes, which motivates us to seek out new methods for calculating rayfields and circumventing the ubiquitous problem of ray-coordinate triplication.

Underlying ray-coordinate systems may be generated by assuming that the rayfield is frequency-independent, and computing the solution to the wide-band eikonal equation (Červený, 2001). However, this approximation can be inappropriate for complex geology where a stationary ray-coordinate system inadequately describes monochromatic wave propagation over a range of frequencies. One example is the significant frequency-dependence of rayfield illumination across salt-sediment interfaces characterized by large impedance contrasts and rugose topography. Hence, an important question is how does one expect to maintain a sufficient and consistent wavefield illumination when the underlying rayfield is itself strongly dependent on frequency? Hence, a frequency-dependent ray-coordinate systems should be an invaluable tool for enhancing imaging practice in complex media.
This paper presents a procedure for constructing a frequency-dependent ray-coordinate system in an adaptive manner directly from the wavefield. The key idea is that the rayfront vectors at any given step are directly calculable from the phase-gradient of previous wavefield solution steps. This naturally leads to a bootstrapping procedure where one alternates between calculating the coordinate system for the next step, and the corresponding wavefield solution at that step. The methodology is similar to the Riemannian wavefield extrapolation (RWE) technique presented by Sava and Fomel (2004), where rayfields are traced through a smoothed velocity model using a Huygens’ wavefront tracer (Sava and Fomel, 2001). The method differs, though, in that frequency-stationarity of the rayfield is not assumed, and the rayfield is instead calculated on-the-fly from the monochromatic wavefield. This method also differs from Shragge and Biondi (2003) in that an initial wavefield is not required as a precondition for solution. Also included in this report is a companion paper, (Shragge and Biondi, 2004), that discusses the strategy of using wavefield solutions precomputed on a background velocity model to train an updated ray-coordinate system using phase-rays.

We begin this paper with a review of phase-ray theory, frequency-dependent coordinate system generation, and ray-coordinate wavefield extrapolation. Then, we introduce the bootstrap procedure by which the ray-coordinate system and accompanying wavefield solutions are computed. Next, we show examples of wavefields extrapolated in adaptive phase-ray coordinates, and conclude with a discussion of the complications posed by triplicating coordinate systems. A more general formulation involving the oriented wave equation (Fomel, 2003) has the potential to address this problem in a robust theoretical framework, although such opportunity remains subject to future research.

THEORY

Phase-rayfields

A monochromatic acoustic wavefield, $\mathcal{U}$, at frequency, $\omega$, and spatial location, $\mathbf{x}$, may be represented by,

$$\mathcal{U}(\mathbf{x}, \omega) = A(\mathbf{x}, \omega) e^{i\phi(\mathbf{x}, \omega)},$$

(1)

where $A(\mathbf{x}, \omega)$ and $\phi(\mathbf{x}, w)$ are the amplitude and phase functions, respectively. For monochromatic waves propagating through isotropic media, the gradient of phase function, $\nabla \phi(\mathbf{x}, w)$, represents the instantaneous direction of energy transport and is a characteristic to the solution of the governing Helmholtz equation (Foreman, 1989). Analogous to the ray precept in broadband theory, this vector quantity defines the instantaneous direction and magnitude of one ray in a continuous ray manifold. However, to differentiate between the broadband and monochromatic ray representations, we term the latter quantity phase-rays (Shragge and Biondi, 2003). The governing differential equations for a phase-ray, $r_i$, are presented in the Exact-ray formulation of Foreman (1989). In Cartesian coordinates, the subscript $i$ on $r$ refers to the projection of the ray along the $x$ and $z$ axes - $r_x$ and $r_z$, respectively. The phase-ray
equations, in summation notation, are,

$$\frac{dr_i}{ds} = \frac{\partial \phi}{\partial x_i} \left[ \left( \frac{\partial \phi}{\partial x_k} \right) \left( \frac{\partial \phi}{\partial x_k} \right) \right]^{-\frac{1}{2}}, \quad (2)$$

where $\phi$ is the above phase function, $x_i$ is a coordinate of the underlying Cartesian grid, and the repeated index $k$ here (and throughout the paper) represents a summation over all coordinate indices. Scalar step magnitude, $ds$, is given by,

$$ds(x) = v(x) \, d\tau, \quad (3)$$

where $v(x)$ is the velocity in the neighborhood of ray, $r_i(x)$, and $d\tau$ is an element of time along the ray.

Calculating phase-rays thus requires isolating the gradient of the monochromatic phase function. An efficient procedure is to calculate the ratio of the wavefield gradient to the wavefield itself,

$$\frac{\nabla U}{U} = \frac{\nabla A}{A} + i\nabla \phi, \quad (4)$$

which eliminates the oscillatory nature of the wavefield. Taking the imaginary component of equation (4),

$$\nabla \phi = \Im \left( \frac{\nabla U}{U} \right), \quad (5)$$

yields the required phase gradient. The right hand side of equation (5) is calculable only when a wavefield solution is known. The solution for a ray, $r_i$, is computed through integrating the right hand sides of equations (2) using a one-sided, non-stiff integration method (e.g. Simpson’s 1/3 rule). Interestingly, ray solutions are uniquely determined given an initial starting position by reason that equations (2) form a decoupled system of differential equations of first-order. Accordingly, a phase-ray coordinate system is uniquely defined by specifying of a set of initial coordinate points and a frequency, $\omega$. Note that this specification makes the coordinate system frequency dependent. Additional information on the theory of phase-rays is discussed in both Shragge and Biondi (2003) and Foreman (1989).

Ray-coordinate wavefield extrapolation

Wavefield extrapolation in ray-coordinates requires casting the acoustic wave-equation not in the usual Cartesian representation, but rather in a system parameterized by phase-ray variables. In 2-D, these variables consist of $\tau$, the one-way travel time from a source/receiver point along the direction of a ray, and $\gamma$, the direction across the rayfront at a constant time step. A cartoon illustrating ray-coordinate geometry is presented in Figure 1.

Note that the dimensions of time and space coordinates $\tau$ and $\gamma$ are seconds and meters, respectively.
The 2-D acoustic wave-equation for wavefield, $U$, at frequency, $\omega$, in ray-coordinates is (Sava and Fomel, 2003),

$$\frac{1}{vJ} \left[ \frac{\partial}{\partial \tau} \left( \frac{J}{v} \frac{\partial U}{\partial \tau} \right) + \frac{\partial}{\partial \gamma} \left( \frac{v}{J} \frac{\partial U}{\partial \gamma} \right) \right] = -\frac{\omega^2}{v^2} U,$$

where $v$ is the velocity function, and $J$ is the ray-coordinate Jacobian or geometrical ray spreading factor given by,

$$J = \left[ \frac{\partial x_k}{\partial \gamma} \frac{\partial x_k}{\partial \gamma} \right]^{\frac{1}{2}}.$$

Importantly, parameter $J$ is solely a component of ray-coordinates and is independent of wavefield extrapolated on the coordinate system.

Analogous to wave-equation extrapolation in Cartesian coordinates, a dispersion relation must be specified that forms the basis for all derived ray-coordinate extrapolation operators. The relation being sought is the wavenumber along the ray direction, $k_\tau$. Following Sava and Fomel (2003), the partial derivative operators in equation (6) are expanded out to generate a second-order partial differential equation with non-zero cross derivatives. Fourier-domain wavenumbers are then substituted for the partial differential operators acting on wavefield, $U$, and the quadratic formula is applied to yield the expression for $k_\tau$,

$$k_\tau = \frac{iv}{2J} \frac{\partial}{\partial \tau} \left( \frac{J}{v} \right) \pm \left[ \omega^2 - \left[ \frac{v}{2J} \frac{\partial}{\partial \gamma} \left( \frac{J}{v} \right) \right]^2 + \frac{iv}{J} \frac{\partial}{\partial \gamma} \left( \frac{v}{J} \right) k_\gamma - \frac{v^2}{J^2} k_\gamma^2 \right]^{\frac{1}{2}}.$$

Note, again, that quantity $\frac{J}{v}$ depends solely on the coordinate system and is independent of the wavefield being propagated.
One relatively straightforward manner to apply wavenumber $k_x$ in an extrapolation scheme is to develop the ray-coordinate equivalent of Claerbout’s classic 15° equation (Claerbout, 1985). This involves a second-order Taylor series expansion of the radical in equation (8), and the identification of Fourier dual parameters $k_x$ and $k_y$ with their space domain derivative counterparts $-i \frac{\partial}{\partial \tau}$ and $-i \frac{\partial}{\partial \gamma}$. The ray-coordinate formula corresponding to the 15° equation is,

$$\frac{\partial U}{\partial \tau} \approx - \frac{v}{2J} \frac{\partial}{\partial \tau} \left( \frac{J}{v} \right) + \frac{i}{2\omega_0} \frac{\partial}{\partial \gamma} \left( \frac{v}{J} \right) \frac{\partial U}{\partial \gamma} + \frac{-i}{2\omega_0} \left[ \left( \frac{v}{2J\omega_0} \frac{\partial}{\partial \gamma} \left( \frac{v}{J} \right) \right)^2 - \frac{v^2}{J^2} \right] \frac{\partial^2 U}{\partial \gamma^2},$$

(9)

where $\omega_o$ may be considered as the effective (non-stationary) frequency,

$$\omega_o = \omega \left[ 1 - \left( \frac{v}{2\omega_0} \frac{\partial}{\partial \tau} \left( \frac{J}{v} \right) \right)^2 \right]^{\frac{1}{2}}.$$

(10)

Equation (9) may be solved in 2-D using fully implicit finite difference methods (e.g. Crank-Nicolson) and fast tridiagonal solvers. After wavefield solution, $U(\tau, \gamma, \omega)$, has been computed at all rayfield locations, the result is mapped to Cartesian coordinates using sinc-based interpolation operators in a neighborhood about each mapped point.

The chicken and the egg

The 2-D phase-ray extrapolation approach detailed above is analogous to the fabled ’chicken and egg’ conundrum: which to compute first? Stated explicitly, phase-rays must be calculated from a known wavefield solution; however, the wavefield is itself the quantity being computed. Because the wavefield is not known a priori, clearly a new strategy is required to resolve these disparate observations.

There are (at least) two possible ways to circumvent this issue. The first procedure involves using precomputed wavefields to train the phase-ray coordinate system by: i) establishing the longer wavelength rayfield structure by raytracing in a background velocity model using a broadband solver; ii) calculating an initial monochromatic wavefield solution using the background rayfield as the coordinate system; iii) computing an updated ray-coordinate system from the previous wavefield solution; and iv) calculating an updated wavefield on the improved phase-ray coordinate system. This procedure is addressed in companion paper Shragge and Biondi (2004) in this report.

A second approach is to use wavefields parameterized in phase-ray coordinates, rather than Cartesian, to dictate the direction of the next rayfront step. For judiciously chosen $\Delta \tau$ steps, both the rayfield direction and resulting wavefield evolves slowly, and ray directions differ by only small, incremental amounts in a neighborhood of $\tau$. Hence, the orientation of previous few ray steps provide a good estimate of the required ray direction at the present $\tau$ step. Thus, a phase-ray coordinate system may be computed on the fly using the magnitude of the wavefield phase gradient of the previous few steps and the local velocity function.
Using a wavefield parameterized in ray-coordinates to generate the underlying coordinate system requires that the governing phase-ray equations are transformed accordingly. Fortunately, the magnitude of a scalar field gradient remains invariant to coordinate transformation, and is related through a change of variables,

$$\frac{\partial \phi}{\partial x_l} = \frac{\partial \phi}{\partial y_m} \frac{\partial y_m}{\partial x_l},$$  \hspace{1cm} (11)

where \(x_l = [x, z]\) and \(y_m = [\tau, \gamma]\) are the Cartesian and ray coordinate basis, respectively, and \(l\) and \(m\) are dummy indices. This reparameterization leads to the ray-coordinate phase-ray equations,

$$\frac{dr_i}{ds} = \frac{\partial \phi}{\partial y_j} \frac{\partial y_j}{\partial x_i} \left[ \left( \frac{\partial \phi}{\partial y_m} \frac{\partial y_m}{\partial x_l} \right) \left( \frac{\partial \phi}{\partial y_m} \frac{\partial y_m}{\partial x_l} \right) \right]^{-\frac{1}{2}}. \hspace{1cm} (12)$$

Equations (11) can be written explicitly in Cartesian and ray variables,

$$\begin{bmatrix} \frac{\partial \phi}{\partial x} \\ \frac{\partial \phi}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial x} + \frac{\partial \phi}{\partial y} \frac{\partial y}{\partial x} \\ \frac{\partial \phi}{\partial \tau} \frac{\partial \tau}{\partial z} + \frac{\partial \phi}{\partial y} \frac{\partial y}{\partial z} \end{bmatrix}. \hspace{1cm} (13)$$

The partial derivatives between the two coordinate systems in equations (11) are directly related to traditional ray parameters. Cartesian derivatives of \(\tau\) are the horizontal and vertical plane-wave slownesses, while those with respect to \(\gamma\) are the local rotation angle to the Cartesian coordinate system (illustrated in Figure 1). Explicitly, these functions are,

$$\frac{\partial \tau}{\partial x} = \frac{\sin \theta}{v(x)}, \quad \frac{\partial \tau}{\partial z} = \frac{\cos \theta}{v(x)},$$

$$\frac{\partial \gamma}{\partial x} = \cos \theta, \quad \frac{\partial \gamma}{\partial z} = \sin \theta,$$

where parameter \(\theta\) is the angle formed between a ray and the \(z\)-axis (assumed to be positive downward).

Computing a weighted average direction from the previous \(M\) steps requires saving \(M + 1\) previous wavefield steps. The discrete version of equation (12) for ray step at index \(t\), \(\Delta r^t_i\), is,

$$\Delta r^t_i \approx v(r_i(x)) \Delta \tau \sum_{m=1}^{M} \beta_m \frac{\Delta r^{t-m}_i}{\Delta s^{t-m}},$$ \hspace{1cm} (15)

where \(\Delta s^t\) is the scalar step magnitude at step \(t\), and \(\beta_m\) are a set of weights subject to,

$$\sum_{m=1}^{M} \beta_m = 1. \hspace{1cm} (16)$$

Weights \(\beta_m\) may be chosen to yield a spline fit of at least second-order accuracy.

The use of previous wavefield solutions to compute solutions to equations (12) naturally gives rise to a broadened finite difference stencil. Figure 2 presents the finite difference stencil for the \(M=2\) case that has second-order accuracy in \(\gamma\) and \(\tau\).
Figure 2: Finite difference stencil for calculating wavefield solution at present step using the previous two steps ($M=2$). The solid square represents the location of the desired rayfield solution, and the in-filled circles connected by lines are the points contributing to the solution point.

Bootstrapping the chicken to get the egg

Having parameterized the phase-ray equations in ray-coordinates, and specified a method for updating rayfront directions, it is possible to detail the bootstrap method that forms the core of adaptive phase-ray extrapolation procedure. Figure 3 presents a flowchart representation of the bootstrap procedure. Rayfields must be computed prior to wavefield extrapolation.

Accordingly, the coordinate system is first initialized by assuming the first $M+1$ steps using an educated guess of where wavefront energy will propagate. Two examples are an expanding circular mesh for a point source (illustrated in Figure 1) or a tilted coordinate system for a dipping plane-wave source (illustrated in Figure 4). After coordinate system initialization, $M$ wavefield extrapolation steps are carried out to generate the required $M+1$ step wavefield.
Figure 4: The first 4 steps of an initial coordinate mesh appropriate for initializing a dipping plane-wave source. At coordinate locations above the ground surface the velocity model is assumed to be constant so that extrapolated energy enters the model as a monochromatic plane-wave (i.e. in both $\omega$ and $k_x$).

The bootstrap process is a loop around three separate calculations: i) ray step $\Delta r_i$ from the previous $M$ wavefield steps; ii) rayfield Jacobian spreading, $J$, and associated functions; and iii) wavefield $U$ at the current step. The final step involves interpolating the wavefield from the ray to the Cartesian coordinate basis, and is done independently after extrapolation.

**PHASE-RAY EXTRAPOLATION EXAMPLES**

In this section we illustrate the utility of adaptive phase-ray wavefield extrapolation using 2-D synthetic examples involving progressively more complex velocity models. Underlying ray-coordinate systems were calculated according to equation (12), and wavefields were extrapolated on the computed rayfields using the ray-coordinate 15° equation (equation (9)). All images were computed using uniform time steps with individual $\Delta \tau$ dependent on velocity model complexity. Orthogonal coordinate $\gamma$ was parameterized as either a punctual or a plane wave source. Point source images required a parameterization of $\gamma$ over shooting angle starting at radius, $R$. The initial 4-step coordinate mesh was created from a circularly expanding rayfront with an angular range bounded by $\gamma_{min}$ and $\gamma_{max}$. Initial wavefields consisted of constant amplitude lines separated in $\tau$ and distributed uniformly over coordinate $\gamma$. Plane-wave images required a parameterization of $\gamma$ over surface coordinate $x_o$. The initial 4-step coordinate system was a Cartesian mesh tilted at the dip angle of the plane-wave being extrapolated. Initial wavefields again consisted of constant amplitude lines separated in $\tau$ and distributed uniformly over surface position, $x_o$.

We designed the first extrapolation example to test the method on a smooth velocity function of sufficient contrast to enable overturning waves. The velocity model, shown in the left panel of Figure 5, consists of a broad Gaussian velocity anomaly that is 86% slower than the background velocity of 3000m/s. Superposed on the Gaussian anomaly are vertical and horizontal gradients of 0.1 and -0.05 s$^{-1}$, respectively. The first test involved using a point source located at 3000m. Rays were computed between $-60^\circ$ and $60^\circ$ assuming an initial radius of 200m. Phase-ray wavefield extrapolation was then carried out with a constant $\Delta \tau$ spacing of 0.0005s at a 5Hz frequency. The resulting ray-coordinate system, superposed on the left panel of Figure 5, is smooth and triplication-free. The middle panel presents the corresponding 5Hz
Figure 5: Point source wavefield extrapolation example. Left: Velocity model with Gaussian-shaped anomaly 86% slower than the background velocity of 3000m/s. Superposed over top are vertical and horizontal gradients of 0.1 and -0.05km/s per km, respectively, and the 5 Hz ray-coordinate system. Middle: The 5Hz monochromatic wavefield. Right: Broadband wavefield calculated using a 0.2-25Hz frequency band. [jeff1-Circle.ps] [CR]

The second example, shown in Figure 6, is the plane-wave equivalent to Figure 5. The panels in Figure 6 are similar to those presented in the previous figure. The initial coordinate system and wavefield tilt angle was -10°, and the spacing between individual rays was set at 20m. The left and middle panels show the 5Hz rayfield and the corresponding 5Hz wavefield, respectively. Again, the rayfield is observed to compress along coordinate γ as it nears the center of the Gaussian anomaly leading to increased amplitudes and shorter wavelengths. The right panel presents the broadband wavefield calculated in the 0.2-25Hz frequency band. Again, all wavefield frequencies were extrapolated on a stationary 5Hz ray coordinate system.

The next example demonstrates adaptive phase-ray wavefield extrapolation in a Gulf of Mexico salt model. The background velocity of the model, shown in the left panel of Figure 7,
is a typical Gulf of Mexico $v(z)$ velocity gradient. The superposed salt body is characterized by a higher wavespeed (4700m/s) and a somewhat rugose bottom of salt interface. A point source was modeled at surface position 12000m with a starting radius of 200m. The initial angular coverage was bounded by $\gamma_{\text{min}} = -22^\circ$ and $\gamma_{\text{max}} = 20^\circ$ because shooting beyond this angular range lead to ray-coordinate triplication. The superposed rayfield in the left panel demonstrates the effect of strong velocity contrasts and a rugose interface between the bottom of salt body and the enveloping sediments. At angles tending away from vertical (i.e. $\theta=0$), rays increasingly refract in accordance with Snell’s law, become horizontal, impinge on the salt-sediment interface, and eventually refract upward at fairly steep angles. The middle and right panels present the 5Hz monochromatic and 2-35Hz broadband wavefields, respectively. Again, the ray-coordinate system in the right panel was assumed to be stationary, and all frequencies were extrapolation on 5Hz rayfield.

**RAY-COORDINATE TRIPPLICATION**

The phase-ray extrapolation examples discussed so far have intentionally avoided triplicating wavefields. However, wavefield triplications are commonly observed in seismic data, especially in areas of complex geology where the phase-ray extrapolation technique shows most potential. In this formulation, the ray coordinate system is computed from previous wavefield steps; hence, the underlying basis will begin to triplicate immediately after the wavefield does. Thus, a contingency plan must exist to prevent numerical instabilities associated with coordinate triplication.

Wavefield triplication naturally occurs when a propagating wavefield is focused by lateral velocity variation acting as an optic lens. One canonical example is a Gaussian-shaped slow velocity anomaly, where continued wavefields exhibit a characteristic bow-tie signature beneath the anomaly. Numerical instabilities occur when calculating the ray coordinate system in the vicinity of the bow-tie because neighboring rays overlap while following their respective branches of the bow-tie. At the crossing point, the Jacobian in equation (8) is identically
zero leading to infinite values of wavenumber $k_f$ in equation (6). Infinite wavenumbers, of course, are not realizable in practice and are only a theoretical artifice of the wavefield being multivalued at that point. Accordingly, instabilities with ray coordinate triplication may be rectified through an appropriate accounting for the wavefield’s multivalued nature during numerical calculations.

One way to deal with multivalued functions is to treat the individual branches of the triplication bow-tie as independent wavefield components that should be held incommunicado. This idea, borrowed from ideas in the mathematical field of complex analysis (Cohn, 1967), is illustrated in Figure 8. Isolating triplication branches requires computing the locations of wavefield triplications from crossing ray segments in the rayfield. In 2-D, crossing ray segments may be identified by modeling the segments as infinite lines, computing their intersection point, and testing whether this location falls within the area bounded by the ray segments. Where this test reveals a crossing point (or branch point), the rayfield has triplicated and should be cut into individual branches. Jacobian coordinate and other related functions that require the computation of derivatives, can be then be calculated on their respective branches (i.e. on one of the three planes in Figure 8). For locations not on branch cuts, centered finite-difference stencils may be used; however, at branch-cut locations appropriate left- and right-sided derivatives are required. Importantly, the locations of branch cuts are kept for all subsequent computations. Finally, we acknowledge that this treatment of rayfield triplication is cursory and remains a topic of ongoing research. However, similarities between our proposed method for handling coordinate triplications and the standard branch cut technique of complex analysis should provide us with a powerful set of tools for further development.

The canonical example of a slow Gaussian-shaped velocity anomaly is presented in Figure 9. The velocity model used in this example is presented in the left panel, and consists of a slow Gaussian anomaly of maximum -50% perturbation of the 2000m/s background velocity. The 10Hz phase-ray coordinate system is also overlain. The middle panel shows the 10Hz wavefield. The hatched pattern in the lower center of the figure is created by the superposition of the phases of the two competing triplication branches (as discussed in Shragge and Biondi (2003)). The right panel presents the broadband result (0.1-30Hz) computed on a stationary 10Hz phase-ray coordinate system. In the lower part of the figure, the signature bow-ties of the
triplicating wavefield are evident. The slight undulations on the centered part of the bow-tie, though, should not be present. Our conjecture is that these are an artifact of coordinate system interpolation.

CONCLUSIONS

We demonstrate the utility of a ray-based Riemannian wavefield extrapolation method using an adaptive bootstrap approach to calculate the ray-coordinate system directly from the wavefield. This bootstrap procedure allows for frequency-dependent ray-coordinate systems to be computed on-the-fly from the wavefield phase gradient and the velocity model. Coupling this procedure with RWE leads to a general frequency-dependent extrapolation procedure capable of following a wavefront as it propagates, overturns, and even triplicates. We propose that in locations where the ray-coordinate system triplicates, the Jacobian spreading and other required functions may be calculated by cutting the triplication into its constituent branches and using appropriate one-sided finite difference stencils. The full proof of this conjecture, though, remains a topic of future research.

ACKNOWLEDGMENTS

We acknowledge Western Geco for permission to use the Gulf of Mexico velocity model. We thank Biondo Biondi and Sergey Fomel for helpful discussions.
REFERENCES


Claerbout, J., 1985, Imaging the earth’s interior: Stanford University.


Červený, V., 2001, Seismic ray theory: Cambridge Univ. Press.

Short Note

Updating ray-coordinate systems with phase-rays

Jeff Shragge and Biondo Biondi

INTRODUCTION

Riemannian wavefield extrapolation (RWE) casts the problem of wavefield extrapolation in a framework independent of a particular coordinate system (Sava and Fomel, 2003). A practical implementation of RWE, though, requires specifying the computation domain on which to perform extrapolation. One judicious choice is a rayfield where the natural extrapolation direction is stepping in time along an individual ray. Rays in simple media are characterized by smooth curves, and are regularly distributed and (usually) triplication-free. However, rays in more complex media often exhibit tortuous behavior, are irregularly distributed, and are full of triplications. These characteristics, especially triplication, are problematic for the calculation of the derivatives necessary for RWE. Accordingly, a procedure must exist to minimize these adverse effects.

One solution to the ray-coordinate triplication problem is to use velocity models sufficiently smooth such that computed rayfields are triplication-free. The true velocity model is then mapped to the calculated smooth rayfield and RWE is performed as usual. Often, through, the amount of smoothing required to meet this objective causes rays to deviate significant distances from their true ray paths. This spatial relocation leads to a degradation of ray-coordinate extrapolation operator fidelity, because of the non-conformal orientations of the wavefront and extrapolation axes. However, operator accuracy may be partially restored by updating the ray-coordinate system to a basis better aligned with true extrapolation direction. One way to accomplish this is to extrapolate an initial wavefield on an initial ray-coordinate system, and to use this result to calculate a corresponding phase-rayfield (Shragge and Biondi, 2003). This rayfield then forms an improved coordinate basis on which to extrapolate wavefields using RWE.

This short note describes a procedure for updating ray-coordinate systems using phase-rays and RWE. This procedure differs from that of companion paper Shragge and Sava (2004), which presents a recursive bootstrap procedure for calculating a ray-coordinate system on-the-fly using the wavefield phase gradient of the previous few extrapolation steps. This short note is comprised of an overview of the processing flow required to generate the updated phase-rayfield, and examples of the method’s application using the Sigsbee 2A velocity model.

1email: jef@sep.stanford.edu, biondo@sep.stanford.edu
METHODOLOGY

This section outlines our proposed strategy for updating ray-coordinate systems. This exposition of the required processing steps is intended to be didactic rather than technical. Accordingly, we review in this short note neither the calculation of phase-rays from monochromatic wavefields, nor Riemannian wavefield extrapolation. Instead, we refer the reader interested in the details of either procedure to the more detailed discussions in Shragge and Biondi (2003), Shragge and Sava (2004), and Sava and Fomel (2003).

Calculating an updated coordinate system using phase-ray tracing is a fairly simple procedure involving five processing steps. The goal of this processing is the formation of a coordinate system that better conforms to the orientation of propagating wavefronts. A flowchart of a typical processing sequence is presented in Figure 1. The first step involves developing a velocity model with adequate smoothness such that the corresponding rayfield is triplication-free. To do this, we apply stationary triangular smoothing operators to the true velocity model. The second step is the generation of an initial rayfield using the prepared velocity model. Point source rayfields are modeled by specifying a set of initial ray points located on an arc of some small radius about the shot point. Plane-wave rayfields are generated by seeding rays at all areal acquisition points, and shooting at a constant angle. Planar ray-coordinate systems at different shooting angles are used for extrapolating plane-waves of similar dip. Generally, steps 1 and 2 are repeated until a triplication-free ray-coordinate system is produced.

The third step is to generate an initial monochromatic wavefield solution by: i) performing RWE on the calculated smooth rayfield; and ii) interpolating the resulting wavefield to a Cartesian basis. This step requires mapping a velocity model to the initial ray-coordinate system. Importantly, we are free to choose a velocity function that is not identical to the true model. One strategy is to choose a velocity model that is somewhat rougher than that used to calculate an initial rayfield, but smoother than the exact model. A judicious model choice may yield a wavefield that triplicates, but it should not be too complex to inhibit phase-ray computation.

The fourth step involves using the initial wavefield’s phase gradient to calculate an updated ray-coordinate system. This procedure is identical to the calculation of phase-rays discussed in Shragge and Biondi (2003). Attention need be paid to locations where the wavefield magnitude is zero, as they lead to unstable phase-ray calculation. However, because the rayfield only
serves as a proxy for the true wavefront extrapolation direction, we may freely regularize it through application of triangular smoothing operators to generate a smooth ray-coordinate system.

The final step in the procedure comprises: i) a RWE of the broadband wavefield on the updated ray-coordinate system using the true velocity model; ii) and interpolating the resulting wavefield to a Cartesian basis. In principle, steps three and four may be repeated incorporating velocity models of increasing roughness that tend to true velocity model; however, we do not apply this extension here.

**EXAMPLE - SIGSBEE 2A**

This section presents an example of the ray-coordinate system updating procedure using the Sigsbee 2A velocity model. The velocity model, presented in the lower panel of Figure 2, consists of a typical Gulf of Mexico \( v(z) \) velocity gradient overlain by a rugose salt body characterized by significantly higher wavespeed.

The first step of the procedure is shown in the upper panel of Figure 2. The rugosity of the true salt velocity model leads to a plane-wave rayfield containing a substantial number of triplications. To avoid this problem, we iteratively smooth the velocity model with stationary triangular operators. The velocity model shown in the upper panel is the roughest model where rays did not cross. The result of the second step of the procedure, calculating an initial rayfield, is presented in the middle panel. Here, the rays are calculated with a Huygens’ wavefront tracing procedure (Sava and Fomel, 2001).

The third step of the procedure is to extrapolate a monochromatic wavefield on the initial rayfield generated in the step 2. This is done using a 5Hz wavefield and the RWE procedure employing the ray-coordinate 15° equation. The resulting wavefield, interpolated to Cartesian coordinates, is shown in the upper panel of Figure 3. The wavefield above and to the left of the salt body is well behaved; however, beneath the top of the salt boundary it exhibits a cross-hatched pattern characteristic of the superposition of different triplication branch phases. This phenomenon is more evident in the sub-salt regions where the lens-like focusing effects of the salt are visible.

The fourth step is to compute the phase-rayfield from the monochromatic wavefield produced in step 3. Results for the 5Hz phase-rayfield are shown in the lower panel of Figure 2. The horizontal extent of the computation grid is less than that of the wavefield in the upper panel, because phase-rays can only be calculated at locations where the initial wavefield was computed. The effects of the rougher velocity model are evident in the shorter wavelength features of the coordinate system. For example, coordinates in the salt body canyons are delayed relative to neighboring points inside the salt body, leading to the sharper corners of the ray-coordinate system below the salt body. Also, the kink in the rayfield in the lower right corner of the lower panel is caused by side boundary reflections masquerading as wavefield triplications. The fifth step of the procedure, calculating a monochromatic wavefield on the updated ray-coordinate system, is illustrated in the lower panel of Figure (3). The resulting
Figure 2: Velocity models and traced rayfields. Upper panel: smoothed Sigsbee 2A velocity model that forms the initial velocity model used for tracing an initial ray-coordinate system. Middle panel: rougher velocity model used for tracing the updated wavefield. The initial ray-coordinate system is overlain; Lower panel: updated 5Hz ray-coordinate system overlying the unsmoothed Sigsbee 2A velocity model. [jeff2-Ziggy.raycomparison] [CR]
Figure 3: Calculated monochromatic 5Hz wavefields. Upper panel: initial wavefield solution calculated on initial ray-coordinate system using the rougher velocity model; Middle panel: Sigsbee 2A velocity model; Lower panel: updated wavefield solution calculated on updated ray-coordinate system.
wavefield is full of triplications, and the illumination gaps (areas of low wavefield intensity) observed by Clapp (2003) are now plainly visible.

Broadband wavefields computed on the initial and updated ray-coordinate systems are presented in the upper panels of Figure 4. For interest, we include the velocity model mapped into both of these coordinate systems (lower panels). The horizontal lines represent surfaces of constant extrapolation time. The upper panels show the initial plane-wave after propagation from the surface to depth through the respective velocity models. Results are presented at three travel-time steps spaced out at one second intervals. If the coordinate system were perfect, extrapolated wavefronts would be collinear with the three straight lines. This would indicate that the wavefront remains perfectly conformal with the ray-coordinate system. (Diffracted energy, though, moves sub-parallel to the wavefront and therefore does not form a straight line).

Differences in wavefront curvature between the upper two panels illustrate that the angles formed between the wavefront and extrapolation axes have, indeed, changed. In particular, they are now greater on the left side than on the right, which is not surprising because our goal is creating a coordinate system more conformal with the wavefront orientation. (Recall that a decreased extrapolation angle corresponds to increased operator fidelity.) This improved alignment, though, does come at a cost: the flanks of the canyon between 17 and 20km in midpoint are now steeper than in the original velocity model. This means that extrapolation across the salt flanks is now less robust because the salt flank angles are too severe. A result of this problem is observable in the less realistic extrapolated wavefield in the salt canyon between 2-3s in time and 17 and 20km in midpoint.

From this example, we point out that the ability to create a ray-coordinate system actually leads to an extra degree of freedom in the extrapolation process. As a result, the practitioner must resolve the trade-off between how well a wavefront conforms with the ray-coordinate system and how steep the structural dips of the velocity model are mapped as a result. Thus, one must ask the question: is it better to account for the steepest structural dips in the velocity model, but with lower extrapolation accuracy? or is it better to handle only shallower structural dips but with greater accuracy? Our conjecture is that the answer lies somewhere between these two end members and strongly depends on the particular velocity model in question.

The wavefields of Figure 4, interpolated to a Cartesian basis, are presented in Figure 5. The upper and lower panels present the broadband extrapolation results computed using the initial and updated ray-coordinate systems, respectively. Snapshots of the wavefields at the first depth level is fairly similar; however, the second and third are markedly different. The wavefield in the lower panel at the fourth snapshot is more continuous across the breadth of the computational grid. However, beneath the salt canyon described above, the extrapolated wavefields in the upper panel seem to be more representative of the expected propagation results. This lack of global improvement is an example of the trade-off discussed above.
Figure 4: Calculated broadband wavefield solutions in ray-coordinates (2-20Hz). Upper left panel: wavefield computed on initial ray-coordinate system; Upper right panel: wavefield computed on updated ray-coordinate system; Lower left panel: velocity model mapped to the initial ray-coordinate system; Lower right panel: velocity model mapped to the updated ray-coordinate system.
Figure 5: Calculated broadband wavefield solutions in Cartesian coordinates (2-25Hz). Upper panel: wavefield calculated on initial ray-coordinate system; Middle panel: Sigsbee 2A velocity model; lower panel: wavefield calculated on a updated ray-coordinate system.
[jeff2-Ziggy.allcartsoln] [CR]
CONCLUSIONS

We present a procedure for calculating ray-coordinate system updates incorporating RWE and phase-ray tracing. This procedure can successfully generate a ray-coordinate systems that better conforms with the direction of wavefield propagation. This fact then leads to an improved wavefield extrapolation operator fidelity. However, coordinate system realignment is not always a global advantage. This is because, as a result, structural dips may be mapped more steeply in the updated ray-coordinate velocity models. Importantly, steeper structural dip angles in the velocity model also inhibit the accuracy of wavefield extrapolation. Accordingly, we suggest that a trial-and-error (or a more sophisticated) approach could be used to generate a ray-coordinate system for RWE that represents an optimal trade-off between improving operator fidelity and lowering structural dips in resulting velocity models.

ACKNOWLEDGMENTS

We acknowledge Western Geco for permission to use the Gulf of Mexico velocity model. We also thank Paul Sava and Antoine Guitton for helpful discussions.

REFERENCES


Aliasing in prestack wavefield continuation migration

Brad Artman, Jeff Shragge, and Biondo Biondi

ABSTRACT

With the widespread adoption of wavefield continuation methods for prestack migration, the concept of operator aliasing warrants revisiting. While zero-offset migration is unaffected by spatial aliasing due to the migration operator, this is not the case for prestack migration. This problem arises in any situation where sources and receivers are not collocated at every sampling point. Once anti-aliasing criteria have been calculated, aliased energy may be prevented from entering the image space by using a source function of appropriate band-width, or band-limiting the energy of the contributing data. As shot-profile migration is the most accurate and expensive imaging algorithm, data axes are commonly subsampled to save cost. We analyse the costs and benefits of implementing anti-aliasing measures to remediate unequal sampling intervals. While some bandwidth of the output image is lost in this process, it will attend to aliasing problems that will be most apparent in the shallow overburden and steeply dipping reflectors. Despite the loss in resolution, any proposed method still enjoys better bandwidth than source-receiver migration with the same data.

INTRODUCTION

The potential for aliasing phenomena to be introduced into Kirchhoff migration images from unaliased data is well documented. However, common wisdom holds that wavefield continuation migration does not introduce aliasing artifacts into the image. This is strictly true only for zero-offset migration. Thus, with the increasing use of prestack wave equation migration, it is important to establish when and how aliasing artifacts are introduced into the image through the wavefield-continuation migration process.

Migration of seismic data may give rise to aliasing problems in four distinct situations. Two of these aliasing situations we consider to be well understood and have effective solutions. These are: 1) improper discretization of the wavefield recorded at the surface can lead to aliasing of the features in the raw data (i.e. poorly sampled hyperbolae), and 2) not establishing an image space with twice finer sampling to accommodate the sum of spatial frequencies due to the multiplication of the wavefields in the imaging condition. Proper planning and layout of the acquisition mitigate the first problem, while Zhang et al. (2003) points out that interpolation of source and receiver wavefields by a factor of two eliminates the aliasing from multiplication.

1email: brad@sep, jeff@sep, biondo@sep
For the duration of this discussion we will assume both of these issues have been effectively controlled.

This paper presents an evaluation of two additional aliasing situations that have yet to be examined by the geophysical community. First, and the focus of this presentation, aliasing phenomenon can arise during the adoption of a wavefield coordinate system on which both source and receiver wavefields can be propagated and combined to create the image. Zero-offset migration, by definition, has common source and receiver locations. However, prestack migration requires choosing from two probably unequal source and receiver sampling intervals (doubly compounded for 3D surveys). This paper explores the ramifications of this problem and defines appropriate bandlimits. Finally, the extrapolation operator can introduce aliasing by moving energy to too high wavenumbers as it convolves the wavefields with the earth velocity model. During later propagation steps, that same energy could move back within appropriate limits again. We have not identified how to capitalize on this effect.

The discretized representation of seismic wavefields and wavefield continuation operators requires a strategy to eliminate contamination from aliasing. Fourier sampling theory allows for the development of the rigorous Nyquist limits for arbitrary sampling of data and image axes. Using these requirements to restrict the discretized wavefield continuation process, we present criteria for determining appropriate image space Nyquist limits for arbitrary sampling choices.

As an example, we show a simple numerical case where aliased energy is introduced into the image space during migration after subsampling the shot axis. We then present three ways in which operator aliasing problems may be resolved in shot-profile migration strategies, and discuss the implications of operator aliasing on source-receiver migration formulations. Finally, an Appendix is included to provide, a rigorous development of the appropriate energy wavenumber limits as a function of data axes sampling, the extension of that development to explain the equivalence of shot-profile migration and source-receiver migration, and the introduction of imaging condition aliasing in the source-receiver migration algorithm.

This analysis has several important ramifications. Shot axis subsampling is a common practice before migration of large data sets to save time or cost. Narrow azimuth acquisition strategies, common to marine surveys, have inherent trade-offs between strike and dip resolution that are sometimes difficult to quantify. Wide azimuth land surveys are also often constrained by unequal in-line and cross-line sampling. Migrations of ocean bottom cable data also suffer from this problem due to their acquisition idiosyncrasies (though the reciprocal of our example presented later). Thus for any situation, be it acquisition design or processing choices, where one is forced to migrate data without equal numbers of shots and receivers at the same locations in both surface directions, the aliasing criteria explained herein can easily be implemented in standard wavefield continuation migration programs to enhance the quality of the image.
AXES AND ALLIES

The standard seismic acquisition grid is presented in panel (a) of Figure 1. Ease of drafting and understanding conventionally has lead geophysicists to draw these axes $90^\circ$ to each other. However, there is no reason to do so. Further, imposing the intuition that these axes are mathematically orthogonal leads to difficulties in interpreting Fourier sampling criteria that we aim to investigate here. These axes inhabit the same physical space along a 2D seismic acquisition line. Plotting them orthogonally casts an inappropriate feeling of a second physical dimension. At the limit of this argument, we contend that it is much easier to de-couple completely the origins of these axis and plot them parallel to each other. Having performed a Fourier transform across space of both the source and receiver axes, we present them in Figure 2. Viewing these three distinct axes separately aids in the interpretation of this entire argument. Unfortunately, there is an historic tendency when analyzing the acquisition grid coordinates to include midpoint-offset, $m_h$, axes as diagonal axes to those shown in Figure 1. We will avoid the use of midpoint while casting this presentation largely in the terms of shot-profile migration, as well as explain later the development of our $x$-axis during the imaging condition. Further, when they are superposed, an incorrect stretching is implied. We will briefly consider the mid-point axis, in order to highlight the danger of this practice.

The mapping transformation of energy from one coordinate frame to the other has been historically defined as:

$$m = \frac{s + r}{2}$$  \hspace{1cm} (1)

$$h = \frac{s - r}{2}.$$  \hspace{1cm} (2)

The Jacobian of this system is thus,

$$J^2 = \left( \frac{\partial m}{\partial r} \frac{\partial m}{\partial s} \right) = \left( \frac{1}{2} - \frac{1}{2} \right).$$  \hspace{1cm} (3)
Figure 2: Lengths of the source, receiver and image axes in the Fourier domain. The circle
with an x indicates the multiplication of the data axes to produce the image surface location
axis during the imaging condition. Aliased replications from the source axis are represented
by the bold lines pointing toward the origin. (a) If the image spaced is sampled as finely as the
receiver axis, alias contamination will enter into the image. (b) Aliased contamination from
the subsampled source axes is avoided when the two fields are compared during the imaging
condition.

which makes the determinant $1/\sqrt{2}$. Drawing a midpoint axis along the $45^\circ$ line of the $sr$-axis
is confusing when someone attempts to find value of a particular midpoint location on the
plane. Some measurement of distance must be employed as a zero-offset location does not
lie on one of the original axes. The next problem we then face is which of the multitude of
distance measures we should select: $l_2$. The historic choice has been the square-root of the
sum of the squares: $l_2$. If we make that choice, we then apply the determinant presented above
to cancel the $\sqrt{2}$ factor associated with the norm that so naturally lends itself to pieces of
paper. We could have chosen any norm. Each of them would return a different number, and
none of them have any more or less value for locating a seismic image on the surface of the
earth.

The cross-correlation imaging condition with subsurface offset (Rickett and Sava, 2001)

$$I(x, h, z) = \sum_{s} \sum_{\omega} R(x = r - h, z, \omega) S^\ast(x = s + h, z, \omega),$$

(4)

combines the two independently propagated wavefields, Source and Receiver, and generates
the surface location axis $x$ with the relationship within the arguments of the two wavefields.
Note that these relations are equations for a line where both the source coordinate, $s$, and
receiver coordinate, $r$, coexist. Here we see the distinction between the surface location coor-
dinate $x$ and midpoint. The $x$-axis follows directly from the mathematics of migration. With
the multiplication (across space) associated with the correlation (along time) of equation (4),
the cartoons in Figure 2 should be recalled.

The Appendix develops in rigorous detail the wavenumber limits acceptable in the image
to eliminate completely alias contamination. The analysis of the problem centers around the
effects of the migration process on the data grid, without needing to consider the values of
the data on each grid node. We thus draw an analogy to the body of work available from crystallography, where structure can be analyzed mathematically without need to know what atom resides at any particular location. Thankfully, the regular Cartesian grid on which we normally acquire and process seismic data is a simple rectilinear crystal, though of several more dimensions than seen under a microscope.

The reference crystal we will consider will be the archetypal seismic grid where sources and receivers occupy all locations and share the same spacing increment. The suspicious or simply inquisitive reader can now turn to the appendix to work through the details of the following result. The maximum allowable wavenumbers, $B_{x_{\xi}}$, to avoid artifacts due to migration operator aliasing is

\[
B_{x_{\xi}} = \frac{1}{\text{lcf}(2a_1 \Delta r_\xi, 2a_2 \Delta s_\xi)} = \text{lcf}\left(\frac{N_{r_\xi}}{a_1}, \frac{N_{s_\xi}}{a_2}\right) = \min(B_{r_\xi}, B_{s_\xi})
\]

where $N_{r_\xi}$ and $N_{s_\xi}$ are the Nyquist frequencies defined by fundamental sampling intervals $\Delta r_\xi = \Delta s_\xi$, $a$’s are subsampling factors, and lcf stands for least common factor (which will change to min if the subsampling factors are integers. The subscript $\xi$ denotes the sampling associated with the model space and is included to maintain parallelism with the appendix.

We consider three approaches to remove the aliasing problems associated with the acquisition and subsampling situations mentioned above during shot-profile migration. First, wavenumbers from the source and receiver wavefields at each depth level are band-limited to prevent the entry of aliased duplications into the image during the imaging condition. This does not require eliminating these components from the propagating wavefields, as we can save appropriate portions of the wavefields in temporary buffers for imaging condition evaluation. Second, a band-limited source function, with a wavenumber spectrum limited to the cutoff frequencies imposed by the resampled shot axis, is propagated throughout the migration process. This effectively zeros energy in the aliased band during the convolution in the imaging condition. No additional computational overhead is required for the latter alternative, though anti-aliasing by band-limited imaging requires two additional Fourier transforms for a split-step Fourier migration strategy. It should be noted, however, that both of these approaches will remove energy across both $k_x$ and $k_h$ axes.

A third alternative is to restrict the wavenumbers of the subsurface offset axis $k_h$ during imaging. Casting the imaging condition in terms of it Fourier dual can allow similar mitigation options. Because $k_x - k_r = k_h$, we can select $(k_s, k_r)$ combinations during the imaging condition that do not exceed our prescribed bandwidth. The multiplication of the source and receiver wavefields shown above takes the form of a convolution in the Fourier domain, which can be utilized to insert our anti-aliasing criteria. Lastly, decimating the receiver wavefield to match the shot increment, will be discussed in more detail with reference to shot-geophone style migration.
FLAT EARTH SYNTHETIC

A synthetic data set was generated to test the first two methods of preventing aliasing. Shot-gathers were modeled over a 2000m/s earth model with one flat reflector at 1000m. Nominal shot, \( \Delta s \), and receiver, \( \Delta r \), spacing is 10m. Dominant frequency of the wavelet is 30 Hz. Migration experiments were performed using both shot-profile and source-receiver algorithms. The analysis and conclusions are valid for both shot-geophone and shot-profile migrations due to their mathematical equivalence (Biondi, 2003; Shan and Zhang, 2003). However, in practice there are subtle differences between the two that will be discussed later.

The data so generated has been imaged with a split-step Fourier wave equation continuation migration kernel. The same continuation operator has been implemented in both shot-profile and source-receiver formulations. The complex valued image at the depth of the reflector was extracted without the normal summation of the frequency axis to aid in interpretation. Fourier transforms of the surface location and subsurface offset axes were then applied. Both migration strategies were used to examine several cases of data completeness. For the shot-profile migration example, four cases were considered: 1) shots at every receiver location, 2) shots at every tenth receiver location, 3) shots at every tenth receiver location imaged with a band-limited source function, and 4) shots at every tenth receiver location imaged with the selective energy imaging condition. For the source-receiver example, three cases were considered: 1) migration of all CMP’s, 2) migration of CMP’s from shot placements every tenth receiver location, and 3) migration of subsampled data appropriately filled with zero-traces to regain the size of the original data set. Despite the large decimation of the data for these experiments, care was taken to assure that the data were still appropriately sampled after decimation to assure that the conclusions drawn from these results are not due to acquisition aliasing.

Shot-Profile Migration

The introduction of aliasing artifacts is demonstrated by comparing panels a and b in Figure 3. Both images were constructed with the shot-profile migration algorithm using a single trace source function. Each panel shows the Fourier transformed image surface location and subsurface offset axes, \((k_x, k_h, \omega)\). In this manner the spatial energy components of the image from a particular depth in the model can be viewed. These figures capture the model space at the depth of a perfectly imaged flat reflector. Migrating shot-profiles at every receiver location as in panel a, shows marked difference to the wavenumber components that make up the image produced by migrating shots at every tenth shot in panel b. As the \(k_x = 0\) aliased energy moves into the image from the replicated Fourier spectra, bands of aliased energy appear at multiples of \(10m^{-1}\) on the \(k_x\) axis.

Panel c, also produced using every tenth shot, illustrates the effectiveness of the first two methods in eliminating the aliased energy introduced by subsampling the shot axis. The image does not suffer from aliased energy due to the use of a spatially band-limited source function with some width in space rather than a single trace. Notice also the diminished \(k_h\) bandwidth as a result of restricting energy content of the migration.
Figure 3: Panels showing image wavenumbers from migrations of different selections of available shots from the data set. Left panel was created by migrating all available shots, while the center and right panels include only every tenth shot. Note the replication of the flat reflector every 10 m⁻¹ in the center panel. The right panel is the result of restricting the migrated energy with a fat source function. Notice that the length of the reflector energy along the kₓ = 0 axis has been limited from 10 m⁻¹ to 5 m⁻¹ indicating an imposed restriction on the number of offsets contributing to the image at any given dip. Selective energy imaging yields an identical result to the band-limited wavelet result.

Selective energy imaging, our second proposed method, uses band-limited versions of the source and receiver wavefields for the imaging condition. The resultant image is identical to the right panel and is accordingly not shown.

Source-Receiver Migration

Figure 4 shows analogous plots to those presented in the previous discussion. The same earth model and analysis strategy is employed for these experiments, though a source-receiver geometry for the data is employed requiring sorting the data to midpoint-offset coordinates. The same migration split-step Fourier migration kernel is used. While the panel a result employs a fully populated, regular data set, the others only used shots every 10 receiver locations. Panel b was simply sorted and migrated. Panel c used the same amount of live data traces as panel b, with zero traces replacing nine out of ten traces from the full data volume. Thus, the first and last experiments migrated the same size data cube, while second and last experiments contain the same amount of non-trivial data though they are an order of magnitude different in size.

Importantly, the second panel does not show the aliasing problems present in the second panel of Figure 3 despite the same level of shot decimation. In this experiment, the subsampling of the shot-axis is partially mapped into both of the two new coordinate axes before migration. The coordinate transform of equation (1) thus distributes the axis’ lower Nyquist
limit equally to the new coordinates prior to migration. Since the output image coordinates inherit this same sampling, the resampled data naturally adhere to the band-limiting criterion of equation (5).

The image location and subsurface offset variables of these migrations have the same meaning as those discussed in the shot-profile section previously (Biondi, 2003). There have been two important modifications however due to the initial coordinate transform. First, notice the range of wavenumbers included in the second image space is drastically limited from the panel to the left showing the migration of all available shots. The resorting has effectively band-limited the image space honoring the Nyquist requirement appropriate for the image given the shot axis subsampling. Thus, the algebraic combination of source and receiver coordinates in the numerator makes this an inherently band-limited propagation method. Second, the division by two of both new axes stretches their Fourier dual domains. Notice that the alias replications in the right panel of Figure 4 appear at a wavenumber of 20 rather than the 10 seen in the shot-profile migration example. Despite the fact that the same number and sampling interval for the shot axis was used in both experiments, the division associated with the coordinate mapping has decreased the sampling interval in the space domain and stretched energy along the Fourier domain. This has happened independently of the three modes of aliasing described above and needs undoing separately as well.

**Figure 4**: Source-receiver migration results. Left panel imaged with all data. Center panel imaged with every tenth shot and sorted to midpoint-offset coordinates. Right panel has zero traces inserted to fill out the decimated data set migrated in the center panel to the data set’s original size. [brad1-sg](CR)

**Discussion**

The analytical band-limit of equation (5) is the necessary criteria to appropriately delimit non-aliased wavenumbers for either wavefield continuation migration method. The amount of compression of a data axis in the Fourier domain by removing samples from the space domain...
dictates the areal extent of the source function or the wavenumber limit for the limited energy imaging strategy. This problem will manifest itself primarily with shot-profile migration strategies. This is due to the fact that two conflicting sampling schemes are simultaneously available. Choosing the finer sampling for migration, results in an aliased image. Choosing the coarser sampling for the migration throws away valuable information. We have shown two methods thus far to eliminate aliased contributions to the image based on equation (A-20). We feel that it is more appropriate to use selective energy imaging conditions rather than a fat source function or band-limited continuation steps while propagating on the finest available data grid.

In practice, lateral velocity variation will cause individual wavenumber energy constituents to move about the $f_k$-plane. At some continuation step, energy could move back and forth across the prescribed image space band-limit, and yet still be appropriate for that individual profile propagation. To allow for any beneficence from this movement, it would be unwise to either: a) propagate a band-limited source function, or b) to eliminate energy from the source and receiver wavefields during propagation steps (by either coarse resampling, or band-limiting the propagation wavefields). Therefore we recommend migrating individual shots on the fine grid, and accounting for operator aliasing in the imaging step at a modest cost increase. The output resolution will still be at least as good as a source-geophone migration.

**COMPLEX EARTH EXPERIMENTS**

Despite the success in removing alias artifacts with the above two methodologies, use of the bandlimited source function or the bandlimited imaging condition in areas of complex geology has one substantial limitation. Both strategies bandpass the wavefields at some point during the migration which effectively introduces focusing of the source illumination beam that propagates through the model. This is due to elimination of $k_x$ energy in the process. Figure 5 illustrates this effect using a shot modeled over the Marmousi data set. The left panel shows an image generated from the single shot with no restrictions imposed during migration. The right panel however used the spatially bandpassed source function. The introduction of low spatial frequencies into the initial condition of the source wavefield effectively changes the impulse into a short planewave. While this is appropriate to limit aliasing when all the shots are summed, the focusing of the beam directly down is inappropriate. To limit the dipspectrum of the geology appropriate to our sampling thereof, we have limited the directionality of the source wavefield to near vertical as well.

Instead, our last methodology attempts to include all propagation directions by choosing portions of the wavenumber spectrum of the receiver wavefield that are appropriately limited for each component of the source wavefield (or vice-versa). While the goal of the imaging condition is a cross-correlation of the two wavefields followed by extraction of the zero time-lag as shown in equation (4), the process in effect multiplies across the space axis. Further, to calculate offset, a spatial cross-correlation without summation is employed. The Fourier dual of these two implicit operations are convolution and multiplication respectively. Thus, by transforming our wavefields into the wavenumber domain, the imaging condition takes the
form of the convolution

\[ I(x; h)|_{z, \omega} = R(x - h)S^*(x + h)I(k_x; h)|_{z, \omega} = \sum_{\omega} \sum_{j} \tilde{R}(k_x = k_r - j, z, \omega) \tilde{S}^*(k_x = j, z, \omega)e^{-ijh}. \] (6)

Inspecting the above, we notice that the counting subscript \( j \) is actually the wavenumber offset axis with the exponential inverse Fourier transforming the axis during summation to give a single offset panel. Thus, by not summing this dimension, we build the Fourier transform of the offset axis.

Using this formulation, and the fact that \( k_r - k_x = k_h \), we can bandlimit the image space by only allowing offset wavenumber combinations where \( k_r - k_x \) is less than the prescribed bandlimit. Thus, while calculating \( k_x \) wavenumbers for the image space, only a limited and varying band from the offset axis is considered. In this manner, we can limit reflectors to different offset spectra depending on their structural dip. Figure (6) shows the result of this implementation. Several features are prominent. First, the thick, fast layer at \((3000m, 2500m)\) contains dipping energy that is not in the impulse response. This is a dip ringing due to implementing a hard cutoff in the Fourier domain when selecting wavenumbers for imaging. This noise should cancel during summation of many shots. Unfortunately, the convolutional imaging condition has Fourier domain periodicity problems that are well avoided by operating in the space domain. Further, there did not seem to be significant improvement over illumination angle as compared to the two previous methods proposed. Finally, given that our effective range of sub-surface offset is actually quite limited, the huge cost differential, \( O(n_x n_h) \) vs. \( O(n_x^2) \), makes the decision for us.
Figure 6: Single shot image from the band-limited convolutional imaging condition. The wave-number bandwidth was limited to 1/8 of the receiver Nyquist limit.

CONCLUSIONS

We conclude that operator aliasing artifacts are indeed introduced during prestack wavefield continuation migration when \( s \)- or \( r \)-axes are allowed to deviate from a regular grid. This problem is thus manifest both when we choose to subsample the number of shots available from the data or design a survey without shots at every receiver location (or in reciprocal cases like OBC surveys).

Source-receiver migration strategies conveniently side-step the bulk of the operator aliasing problem by effectively performing all propagation and imaging on the coarsest available grid. Thus during every wavefield continuation step, we are only propagating energy that lies below the Nyquist wavenumber for the final image space. However, this constraint can be much too stringent when compared to the sampling criteria set out, and presumably met, during the acquisition effort. This migration style will propagate energy along fewer traces than shot-profile methodologies, although it lacks some ability to control the quality of the image. The fact that the model output of source-receiver migration is on a twice finer grid than acquisition or the model produced with shot-profile migration does not exclude this migration strategy from the imaging condition aliasing described by Zhang et al. (2003). The division in the mapping of equation (1) is responsible for this, but does not perform the necessary
interpolation to avoid imaging condition aliasing. Performing the imaging condition in the Fourier domain does not entail additional cost for this split-step Fourier migration kernel, and conveniently eliminates the need to interpolate the wavefields by a factor of two to address the aliasing due to multiplication as the output of the convolution is inherently twice the size of the two inputs.

Given some knowledge of the dip content of the data, it is possible to extend the boundaries of the rigorous anti-aliasing criteria presented. Using both positive and negative one-sided band-limits can include high wavenumber energy that can improve the image in areas of steep dips and the shallow section of the model. Therefore, when challenged with imaging important steeply dipping targets, decisions concerning acquisition design or the level of decimation along different directions for migration can be made with a better understanding of the consequences to the final product. Finally, it has been noticed that the rigorous Nyquist limits are substantially too restrictive in practice with real data. In reality, the benefits of some level of anti-aliasing are realized when inspecting the dipping canyon features of the Marmousi data, though the limitation imposed during the imaging condition was much more relaxed than indicated by the theory and needed to be found through experimentation.

Further, if a subsample of a data set is imaged with a shot profile migration strategy, full bandwidth source and receiver wavefields, \( W_s(x, z, \omega) \) and \( W_r(x, z, \omega) \), could be saved for future migration efforts. An imaging condition with the appropriate band-limitation across the spatial axes can be applied with partial or complete sets of these migrated volumes. Thus incremental increases in image quality can be achieved while avoiding re-migration of data by augmenting a library of wavefields (if adequate storage capacity is available).

REFERENCES


Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, 36, no. 03, 467–481.


APPENDIX A

Surface seismic acquisition involves the acquisition of a set of discrete wavefield measurements using sources and receivers that populate a 2D recording surface. In wave equation migration, the wavefield continuation component of the experiment involves a downward extrapolation of the recorded wavefield from the acquisition surface to subsurface layers within an earth model. Finally, the imaging component requires the extraction of energy from the wavefield by the evaluation of an imaging condition (Claerbout, 1971). Central to these three procedures is the geometry on which the experiment and accompanying processing are based. From a practical processing viewpoint, the ideal data set would be defined on a uniformly-sampled 3D (for 2D surveys) or 5D (for 3D surveys) acquisition grid. This Appendix seeks to analyze the data grid through all of the afore-mentioned steps as if studying the structure of a crystal.

To address the issue of aliasing, Fourier sampling theory will be applied. This theory provides the necessary and sufficient conditions for preserving the information content of a continuous physical wavefield represented in a discrete manner. One important tenet of sampling theory is that the highest frequency recoverable from a regularly-sampled data set is independent of values at sample locations, but dependent on the interval between neighboring samples. Aliasing considerations for seismic wavefields are likewise dependent on the spacing between individual points in the lattice. In light of the above, we will dissociate the acquisition geometry of a seismic experiment from the values recorded at the acquisition points, and represent the former with a multi-dimensional Shah function. Using this representation, the effects on the lattice structure of the processes of downward continuation and imaging condition evaluation are readily examined.

Throughout this presentation, we will maintain the formulation of the seismic experiment in the shot–geophone coordinate system. Shan and Zhang (2003) points out that a correlation of the source and receiver wavefields is implied in this migration formulation. This observation introduces a convolution of source and receiver lattices. A lattice convolution, whether in a shot–geophone or shot–profile migration setting, gives rise to the phenomenon of image condition aliasing (Zhang et al., 2003). This phenomenon arises when two wavefields multiplicatively interfere to yield an aliased Moiré pattern with frequencies up to twice those of the original wavefields. To account for this phenomenon we will assume that the wavefield is already interpolated by a factor of two, and that the lattice upon which this discussion begins is not the acquisition grid, but is of twice finer spacing. Finally, although the theory as developed below is strictly for 2D seismic experiments, the extension to 3D is a trivial matter and is omitted for clarity.

Lattice basis of seismic wavefields

We choose to represent a seismic wavefield, $W$, by dissociating the underlying interpolated acquisition lattice (IAL), $\mathcal{L}$, from the discretely sampled values of the continuous wavefield, $f^W$. Lattice $\mathcal{L}$ is defined over experimental source and receiver coordinates, $r_\xi$ and $s_\xi$, and
time and depth coordinates, \( t_\xi \) and \( z_\xi \). Importantly, the underlying grid function is only a construct of the experimental process and is associated, accordingly, with model space coordinates \( \mathcal{L}(r_\xi, s_\xi, t_\xi, z_\xi) \). This stands in contrast to the continuous wavefield function defined by corresponding continuous variables, \( f^W(r, s, t, z) \). Written in this manner, it becomes natural to associate the act of observation with the mapping from physical (continuous) to experimental (discrete) variables,

\[
W(r_\xi, s_\xi, t_\xi, z_\xi) = \mathcal{L}(r_\xi, s_\xi, t_\xi, z_\xi) f^W(r, s, t, z) \delta(r_\xi - r) \delta(s_\xi - s) \delta(z_\xi - z) \delta(t_\xi - t). \tag{A-1}
\]

Note, that we define \( W \) as the entire experimental wavefield including traces from all source and receiver pairs. This volume is separable into many different subsets, but we will keep it in tact. One way to represent lattice \( \mathcal{L} \) is with a 4D infinite sum over delta functions (i.e. a 4D Shah function),

\[
\mathcal{L}(r_\xi, s_\xi, t_\xi, z_\xi) = \sum_{u_r, u_s, u_t, u_z = -\infty}^{\infty} \delta(r_\xi - a_r u_r \Delta r_\xi) \delta(s_\xi - a_s u_s \Delta s_\xi) \delta(t_\xi - a_t u_t \Delta t_\xi) \delta(z_\xi - a_z u_z \Delta z_\xi). \tag{A-2}
\]

In equation (A-2), variables \((a_r, a_s, a_t, a_z)\) are subsampling factors over the fundamental discretization intervals \((\Delta r_\xi, \Delta s_\xi, \Delta t_\xi, \Delta z_\xi)\), and \((u_r, u_s, u_t, u_z)\) are the associated summation indicies of the delta functions. It is assumed that for our ideal grid \( \Delta r_\xi = \Delta s_\xi \) and any departures from this equality may be represented through the subsampling factors. Throughout the development, unless specified otherwise, the summation from \(-\infty\) to \(\infty\) in equation (A-2) is assumed. Due to the fact that no experiment is ever carried out with infinite extent, padding the wavefields to infinity with zero traces maintains the rigor of this evaluation.

**Band-limited lattices**

In a seismic experiment, measurements are necessarily acquired at discrete sampling intervals. This fact requires placing a restriction on the frequency content representable on the 4-D lattice. One convenient manner to do this is to apply a 4-D Rect function in the frequency domain to cut off frequencies greater than predefined values \([e.g. (B_r, B_s, B_t, B_z)]\). To accomplish this, equation (A-2) is first Fourier transformed over all variables to yield,

\[
\mathcal{L}(k_{r\xi}, k_{s\xi}, \omega, k_{z\xi}) = \sum_{u_1, u_2, u_3, u_4} \delta(k_{r\xi} - \frac{u_1}{a_1 \Delta r_{\xi}}) \delta(k_{s\xi} - \frac{u_2}{a_2 \Delta s_{\xi}}) \delta(\omega - \frac{u_3}{a_3 \Delta t_{\xi}}) \delta(k_{z\xi} - \frac{u_4}{a_4 \Delta z_{\xi}}). \tag{A-3}
\]

A 4D Rect function, \( \Pi \), with arguments (in 1D)

\[
\Pi(B_{r\xi}) = \begin{cases} 
1 & \text{for } |r_{\xi}| < \frac{1}{2B_{r\xi}} \\
0 & \text{for } |r_{\xi}| \geq \frac{1}{2B_{r\xi}}
\end{cases},
\]

is then applied to the infinite Fourier domain lattice \( \mathcal{L} \) to yield a band-limited version \( \mathcal{L}_{FL} \),

\[
\mathcal{L}_{FL}(k_{r\xi}, k_{s\xi}, \omega, k_{z\xi}) = \mathcal{L}(k_{r\xi}, k_{s\xi}, \omega, k_{z\xi}) \Pi(B_{r\xi}, B_{s\xi}, B_{t\xi}, B_{z\xi}) \tag{A-4}
\]

\( \mathcal{L}_{FL} \),...
Applying an inverse Fourier transform over all 4 dimensions to lattice $\mathcal{L}_{FL}$ yields,

$$\mathcal{L}_{FL}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}) = \mathcal{L}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}) \ast r_{\xi} \ast s_{\xi} \ast t_{\xi} \ast z_{\xi} \mathcal{F}_{L}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}), \quad (A-5)$$

where

$$\mathcal{F}_{L}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}) = \frac{\text{sinc}(\frac{r_{\xi}}{B_{r_{\xi}}} \cdot \frac{s_{\xi}}{B_{s_{\xi}}} \cdot \frac{t_{\xi}}{B_{t_{\xi}}} \cdot \frac{z_{\xi}}{B_{z_{\xi}}})}{B_{r_{\xi}} B_{s_{\xi}} B_{t_{\xi}} B_{z_{\xi}}}, \quad (A-6)$$

and the subscripts on the convolution symbol, $\ast$, delimit the coordinate over which the convolution is applied. Thus the seismic wavefield may be represented by,

$$W(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}) = \mathcal{L}_{FL}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}) \delta(z_{\xi}) f^{W}(r_{\xi}, s_{\xi}, t_{\xi}, z_{\xi}). \quad (A-7)$$

It is important to emphasize that the lattice $\mathcal{L}_{FL}$ in equation (A-7) represents only the lattice structure on which the data wavefield is overlaid. For individual seismic experiments, the values at each location will vary, while the lattice structure remains invariant. Utilizing the crystal structure analogy again, any atom may inhabit a node in the lattice, but there must be one and only one atom present.

**Downward continuation with the DSR**

In this section we apply the double square root (DSR) operator to extrapolate the recorded wavefield of equation (A-7) through the depth coordinate of the lattice. After initially Fourier transforming the wavefield into $k_{r_{\xi}}, k_{s_{\xi}},$ and $\omega_{r_{\xi}},$ the first application of the DSR yields a new wavefield at depth step $z_{\xi} = a_{z} \Delta z_{\xi}$. Mathematically, this is represented by:

$$W(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) = \delta(z_{\xi} - a_{z} \Delta z_{\xi}) \ast z_{\xi} \left[ \text{DSR}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) W(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi} = 0) \right] \quad (A-8)$$

$$\quad = \delta(z_{\xi} - a_{z} \Delta z_{\xi}) \ast z_{\xi} \left[ \text{DSR}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) f^{W} \mathcal{L}_{FL}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) \delta(z_{\xi}) \right]$$

$$\quad = \left[ \text{DSR}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) f^{W} \right] \mathcal{L}_{FL}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi} + a_{z} \Delta z_{\xi}) \delta(z_{\xi} - a_{z} \Delta z_{\xi}).$$

The periodicity of the lattice over depth coordinate $z_{\xi}$ enables the Shah function index $u_{z}$ to be shifted by $a_{z} \Delta z_{\xi}$ such that equation (A-8) reads,

$$W(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) = \left[ \text{DSR}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) f^{W} \right] \mathcal{L}_{FL}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) \delta(z_{\xi} - a_{z} \Delta z_{\xi}) \quad (A-9)$$

By extension, any continuation step operating on a wavefield will take the same form. Applying an inverse Fourier transform over coordinates $k_{r_{\xi}}$ and $k_{s_{\xi}}$ yields,

$$W(r_{\xi}, s_{\xi}, \omega, z_{\xi} = a_{z} \Delta z_{\xi}) = H(r_{\xi}, s_{\xi}, \omega, z_{\xi}) \ast r_{\xi} \ast s_{\xi} \left[ \mathcal{L}_{FL}(r_{\xi}, s_{\xi}, \omega, z_{\xi}) \delta(z_{\xi} - a_{z} \Delta z_{\xi}) \right] \quad (A-10)$$

where, for convenience, $H$ is defined by,

$$H(r_{\xi}, s_{\xi}, \omega, z_{\xi}) = \mathcal{F}_{k_{r_{\xi}}, k_{s_{\xi}}}^{-1} \left[ \text{DSR}(k_{r_{\xi}}, k_{s_{\xi}}, \omega, z_{\xi}) f^{W} \right], \quad (A-11)$$

where $\mathcal{F}$ is the Fourier transform operator. It is important to note here that the convolution of lattice $\mathcal{L}_{FL}$ with filter $H$ does not change the location of the sample points. Rather, it operates only on the amplitudes at the predefined locations.
Evaluation of imaging condition

While we have maintained our coordinate system thus far in parallel with a shot-receiver migration strategy, we will now detail the evaluation of the imaging condition with the order of operations normally associated with shot-profile migration. In fact, the convolution of the source and receiver grids associated with the imaging condition (and gives rise to imaging condition aliasing) can be performed/implied before the migration which is common to resorted \( mh \)-coordinate migrations or at each depth step within the migration. As we have maintained the distinctiveness of the source and receiver grids to this point, the convolution of their axes must now be evaluated.

Calculating Common Image Gather (CIG) offsets involves the evaluation of an imaging condition at all acceptable values of \( r_\xi \) and \( s_\xi \). This gives rise to two new image space variables: the horizontal image coordinate for the earth model, \( x_\xi \), and the subsurface horizontal offset coordinate, \( h_\xi \). These variables have much similarity to the data space variables midpoint and offset. In a strictly \( v(z) \) medium, these axes overlay. However, in more complicated media, the midpoint variable is somewhat misleading. This is because the wavefield continuation extrapolates energy from a midpoint on the surface to a different midpoint as the wavefields are successively downward-continued. Thus, mixing these two ideas is inappropriate. Source and receiver coordinates and horizontal and subsurface offset coordinates are related through transforms \( r_\xi Dx \) and \( s_\xi Dx \). The new coordinate, \( h_\xi \), a derived parameter with a magnitude equal to an integer multiple of \( 1x \), is naturally represented as the product of an integer multiplication factor \( h \) and horizontal image space discretization interval \( \Delta x_\xi \) (i.e. \( h_\xi = h \Delta x_\xi \)) that will most commonly be unity.

Using these definitions, the image cube may be constructed by applying the general correlation imaging condition to wavefield \( W \),

\[
I(x_\xi, z_\xi, h_\xi) = \sum_\omega \left[ \delta(r_\xi - h_\xi) * \delta(s_\xi - h_\xi) \right] \delta(r_\xi - x_\xi) \delta(s_\xi - x_\xi) \tag{A-12}
\]

Note that this expression reduces to the familiar zero subsurface offset form when \( h_\xi = 0 \),

\[
I(x_\xi, z_\xi) = \sum_\omega W(r_\xi, s_\xi, \omega, z_\xi) \delta(r_\xi - x_\xi) \delta(s_\xi - x_\xi). \tag{A-13}
\]

The convolution arguments applied to wavefield \( W \), in equation (A-12), yield

\[
I(x_\xi, z_\xi, h_\xi) = \sum_\omega \left[ \delta(r_\xi - h \Delta x_\xi) * \delta(s_\xi + h \Delta x_\xi) \right] W(r_\xi = x_\xi, s_\xi = x_\xi, \omega, z_\xi) \delta(s_\xi + h \Delta x_\xi) \\
= \sum_\omega W(r_\xi - h \Delta x_\xi = x_\xi, s_\xi + h \Delta x_\xi = x_\xi, \omega, z_\xi). \tag{A-14}
\]

Before continuing with this development, it is useful here to stop and interpret the meaning of equations (A-12) and (A-14). The imaging condition itself builds the image-space coordinate axes \( x \) and \( h \) during the convolution expressed above. The arguments within the wavefield \( W \) of equation (A-14) are the equations of a line. This line, \( x_\xi \), defines the axis for surface location of the image, and is independent of any assumptions about surface midpoints during the experiment. This is one reason \( ^2 \) we have avoided using the midpoint variable, \( m \). These

\[ ^2 \text{Recall that the other is the in measuring along it.} \]
two coordinates indeed share many traits, though the midpoint concept is an arbitrary, while intuitive and convenient, coordinate transform. Surface location, $x_\xi$, is a rigorous development required by the imaging process.

Continuing our derivation, we now reintroduce the lattice in equation (A-10) to the imaging condition which yields,

$$
I(x_\xi, z_\xi, h_\xi) = \sum_\omega H(r_\xi, s_\xi, \omega, z_\xi) *_{r_\xi} *_{s_\xi} L_{FL}(x_\xi = r_\xi - h \Delta x_\xi, x_\xi = s_\xi + h \Delta x_\xi, \omega, z_\xi)
$$

(A-15)

However, $h \Delta x_\xi$ is an integer shift by $\Delta x_\xi$ and is defined only at known points on the lattice allowing the index of the Shah function to be shifted to yield,

$$
I(x_\xi, z_\xi, h_\xi) = L_{FL}(r_\xi = x_\xi, s_\xi = x_\xi, z_\xi) *_{r_\xi} *_{s_\xi} \sum_\omega L_{FL}(\omega) H(r_\xi, s_\xi, \omega, z_\xi)
$$

(A-16)

Expanding lattice $L_{FL}$ into its components $L$ and $FL$,}

$$
I(x_\xi, z_\xi, h_\xi) = L(r_\xi = x_\xi, s_\xi = x_\xi, z_\xi) *_{r_\xi} *_{s_\xi} F L(r_\xi = x_\xi, s_\xi = x_\xi, z_\xi) *_{r_\xi} *_{s_\xi} \sum_\omega L_{FL}(\omega) H(r_\xi, s_\xi, \omega) F L(\omega)
$$

and applying a Fourier transform over coordinates $r_\xi$, $s_\xi$, and $z_\xi$ yields,

$$
I(k_{x_\xi}, k_{z_\xi}, k_{h_\xi}) = \left[ L(k_{r_\xi} = k_{x_\xi}, k_{s_\xi} = k_{x_\xi}, k_{z_\xi}) \Pi(B_{r_\xi} k_{r_\xi} = B_{r_\xi} k_{x_\xi}, B_{s_\xi} k_{s_\xi} = B_{s_\xi} k_{x_\xi}, B_{z_\xi} k_{z_\xi}) \right] G,
$$

$$
G(k_{r_\xi} = k_{x_\xi}, k_{s_\xi} = k_{x_\xi}, \omega) = \mathcal{F}_{r_\xi,s_\xi} \left[ \sum_\omega L_{FL}(\omega) H(r_\xi, s_\xi, \omega) F L(\omega) \right].
$$

(A-17)

The Rect functions of coordinates $k_{r_\xi}$ and $k_{s_\xi}$ are collapsed back to a single Rect function in $k_{x_\xi}$, where the frequency limit is given by $\min(B_{r_\xi}, B_{s_\xi})$. The $\min$ function arises because the maximum grid-spacing along either shot or receiver axis alone dictates the aliasing criteria for the $k_x$-axis. This also allows for simplified calculations in the particular case. Generally, however, the bracketed expression in equation (A-17) is

$$
L(k_{r_\xi} = k_{x_\xi}, k_{s_\xi} = k_{x_\xi}, k_{z_\xi}) = \sum_{u_1, u_2 = -\min(B_{r_\xi}, B_{s_\xi})}^{\min(B_{r_\xi}, B_{s_\xi})} \sum_{u_4} \delta(k_{x_\xi} - \frac{u_1}{a_1 \Delta r_\xi}) \delta(k_{x_\xi} - \frac{u_2}{a_2 \Delta s_\xi}) \delta(k_{z_\xi} - \frac{u_4}{a_4 \Delta z_\xi}).
$$

(A-18)

The summations of the delta functions over $u_1$ and $u_2$ collapse to a single summation over the variable with the lowest common factor ($\text{lcf}$),

$$
L(k_{x_\xi}, k_{z_\xi}) = \sum_{u = -\min(B_{r_\xi}, B_{s_\xi})}^{\min(B_{r_\xi}, B_{s_\xi})} \sum_{u_4 = -d_1 \Delta z_\xi}^{d_1 \Delta z_\xi} \delta(k_{x_\xi} - \frac{u}{\text{lcf}(a_1 \Delta r_\xi, a_2 \Delta s_\xi)}) \delta(k_{z_\xi} - \frac{u_4}{a_4 \Delta z_\xi}).
$$

(A-19)
The horizontal image coordinate is being sampled at a spacing $\text{lcf}(a_1 \Delta r_\xi, a_2 \Delta s_\xi)$. Thus, for aliasing to be absent the following condition must hold,

$$B_{x_\xi} = \min(B_{r_\xi}, B_{s_\xi}) = \frac{1}{\text{lcf}(2a_1 \Delta r_\xi, 2a_2 \Delta s_\xi)} = \text{lcf}(\frac{N_{r_\xi}}{a_1}, \frac{N_{s_\xi}}{a_2})$$

(A-20)

where $N_{r_\xi}$ and $N_{s_\xi}$ are the Nyquist frequencies defined by fundamental sampling interval $\Delta r_\xi$ and $\Delta s_\xi$. Thus, the alias-free wavefield is given by the following geometry

$$\mathcal{L}(k_{x_\xi}, k_{z_\xi}) = \sum_{u=-\min(\frac{N_{r_\xi}}{a_1}, \frac{N_{s_\xi}}{a_2})}^{\min(\frac{N_{r_\xi}}{a_1}, \frac{N_{s_\xi}}{a_2})} \sum_{d_1 \Delta z_\xi} \delta(k_{x_\xi} - \frac{u}{\max(a_1 \Delta r_\xi, a_2 \Delta s_\xi)}) \delta(k_{z_\xi} - \frac{u_4}{a_4 \Delta z_\xi})$$

(A-21)

Notice that for the simplified case of zero-offset migration, the pre-supposed notion that there are no operator aliasing artifacts introduced can be shown conclusively within the presentation above. Without two different sampling intervals, be they source/receiver or orthogonal surface coordinates, there are no choices for the $\min$ operator in equation (A-20) nor the $\max$ operator of equation (A-21). Instead, the sole variable available, surface location $x$, dictates the sampling of the model space.

Notice that for the simplified case of zero-offset migration, the pre-supposed notion that there are no operator aliasing artifacts introduced can be shown conclusively within the presentation of the above results. Without two possibly different sampling intervals, for source and receiver grids, there are no choices for the min operator in equation (A-20) nor the max operator of equation (A-21). Instead, the sole variable available, surface location, dictates the sampling of the model space. This does not however release zero-offset migrations from the ramifications of image condition aliasing, as the implied correlation of the source wavefield associated with source-receiver migrations is still present.
Ocean-bottom hydrophone and geophone coupling

Daniel A. Rosales and Antoine Guitton

ABSTRACT

We compare two methods for combining hydrophone and geophone components for an ocean-bottom seismic experiment to eliminate the receiver ghosts associated with this type of seismic acquisition. One approach is in the time domain, the other in the frequency domain. Both approaches are compared with the 2D OBS data over the Mahogany field in the Gulf of Mexico. The receiver ghosts are eliminated more efficiently with the frequency domain method, because this method combines the data in two different steps: i) calibration, and ii) deghosting.

INTRODUCTION

Ocean-bottom cable acquisition results in a receiver ghost problem. An operational method to solve this problem is to use paired hydrophone and geophone detectors. Combining the hydrophone and geophone takes the advantage of the fact that the two types of detectors generate signals of the same polarity for the upcoming wavefield, and opposite polarity for the downgoing wavefield (Gal’perin, 1974; Barr and Sanders, 1989; Soubaras, 1996). The main challenge of this method is that the hydrophone and geophone must be properly calibrated to produce a deghosted output.

Barr and Sanders (1989) propose a technique in the time domain that calibrates the geophone measurement and eliminates the ghost reflection in one simple step. According to Soubaras (1996), however, the geophone calibration and the deghosting process must be done separately. He proposes a method in the frequency domain to separately calibrate the geophone measurement and eliminate the receiver ghost.

A 2D line over the Mahogany field in the Gulf of Mexico helps to test both of these approaches. First, we present a pre-processing technique over this 2D line. We present two methods of combining the hydrophone and geophone components and use the results to obtain preliminary estimates of the P velocity field for this dataset.

1email: daniel@sep.stanford.edu, antoine@sep.stanford.edu
PRE-PROCESSING

This section illustrates some of the problems with the Mahogany data set. Figure 1 shows the hydrophone (left) and the geophone (right) components of a common shot gather. Note the events with a predominantly linear moveout on the hydrophone component. These events represent an interface wave, one that travels in the first layer below the water bottom with a characteristic velocity of 1500m/s. Due to the high energy of these events and their dispersive characteristic it is not possible to observe and analyze the contribution of the far offset traces to the moveout of the reflections. Hence, it is important that we eliminate these events without destroying the main reflections. Because this noise has a characteristic linear moveout, a radial noise model serves to approximate and eliminate these events.

The pseudo-unitary implementation of the radial-trace transform (Brown and Claerbout, 2000) promises to be an efficient technique to suppress the noise in the hydrophone component, because the radial-trace transform lowers the apparent temporal frequency of these radial events.

After the radial noise suppression, we also performed a mute with the water velocity and a bandpass filter. Figure 2 shows the same gathers as in Figure 1 after the radial-trace noise suppression for the hydrophone component and the mute and bandpass filtering for both the hydrophone and geophone components. Now, it is possible to observe more clear reflections in the hydrophone component of the shot gather. Most of these events match with those observed in the geophone component of the shot gather.

HYDROPHONE AND GEOFONE COMBINATION

We discuss two methods to combine the pressure and velocity detectors at an identical location on a 2D OBC line. Both methods perform a calibration over the velocity detector or geophone. The goal of both of these methods is to eliminate the ghost reflection. The first method is in the time domain and combines the geophone calibration and the deghosting in one step. The second method is in the frequency domain and performs the geophone calibration and the deghosting in two steps.

Time domain methodology

The method that Barr and Sanders (1989) proposed to combine the hydrophone and geophone is simple and easy to implement. They simply add the hydrophone and the calibrated geophone in the shot domain. The calibrated geophone is computed with a constant factor equal to the ratio of the amplitudes of the hydrophone and geophone. We calculate a constant factor per trace, we average all of them, and finally apply the averaged constant factor to the entire shot gather.

This procedure not only calibrates the geophone but also eliminates the ghost reflection.
The final combined signal \( s(t) \) is given by the following:

\[
s(t) = h(t) + \frac{\rho v_p}{\cos \gamma'_p} \frac{kr(1+kr)}{kt(1-kr)} z(t),
\]

where \( h(t) \) and \( z(t) \) correspond to the hydrophone and geophone, respectively, \( \rho \) is the water density, \( v_p \) is the P-wave water velocity, \( \gamma'_p \) is the P-wave refraction angle in water, and \( kr, kt \) are the reflection coefficient and the refraction coefficient, respectively.

Figure 3 presents the physical model for this approach. Solving the boundary conditions for the elastic wave-equation at the water bottom (left panel on Figure 3) gives the amplitudes of the reverberations (right panel on Figure 3). This model explains that combining the hydrophone and the geophone components as in equation (1) results in a reverberation-free signal.

The right panel on Figure 3 also explains how to obtain the scale factor for equation (1). Comparing the amplitudes of the reverberations shows that the scale factor is just the absolute value of the ratio between the amplitudes of the hydrophone and the geophone. The final result, \( s(t) \), is a deghosted output.

Figure 4 shows the result of this approach over the common-shot gather from Figure 1.
Although it was possible to eliminate some of the multiples, the final result, $s(t)$, is not totally multiple-free.

**Frequency domain methodology**

Soubaras (1996) proposes to split into two procedures the calibration of the hydrophone and geophone and the ghost elimination. Figure 5 shows the physical model proposed by Soubaras (1996). The fields $U_0$, $D_0$, and $S_0$ represent the initial upgoing, downgoing and source wavefields, respectively. Similarly, the fields $U$, $D$, and $S$ are the upgoing, downgoing and source wavefields at the water-bottom level (receiver level).

**Calibration**

The pressure component ($P$) and the vertical component ($Z$) of the receiver gather are both in the frequency domain. The available data are the hydrophone component ($P$) and the non-
Figure 3: Physical model for the reverberations. It solves for the boundary conditions of the elastic wave field for the model on the left. On the right are shown the reverberations as a function of time for the hydrophone (h) and geophone (z) components. The first arrival corresponds to event $a_h$ and $a_z$ with an amplitude equal to $a_h = kt$, $a_z = \frac{kt}{\rho vp} \cos \gamma_p'$. The first reverberation, $b_h$ and $b_z$, has an amplitude of $b_h = -kt(1 + kr)$ and $b_z = \frac{kt(1 - kr)}{\rho vp} \cos \gamma_p'$. The second reverberation, $c_h$ and $c_z$, has an amplitude of $c_h = ktr(1 + kr)$ and $c_z = \frac{ktr(1 - kr)}{\rho vp} \cos \gamma_p'$. The third reverberation, $d_h$ and $d_z$, has an amplitude of $d_h = -ktr^2(1 + kr)$ and $d_z = \frac{ktr^2(1 - kr)}{\rho vp} \cos \gamma_p'$.

calibrated geophone component ($\tilde{Z} = \frac{Z}{C}$, $C$ is the calibration factor we need to compute):

\[
\begin{align*}
P &= U + D, \\
Z &= \frac{U - D}{\rho vp}.
\end{align*}
\]

The initial source wavefield is given as follows:

\[
S_0 = U_0 + D_0.
\]

The propagated upgoing and downgoing wavefields at the water-bottom surface are, respectively,

\[
\begin{align*}
U &= e^{-i\omega \Delta t} U_0, \\
D &= e^{i\omega \Delta t} D_0,
\end{align*}
\]

where $\Delta t = 2\Delta z/v$, $\Delta z$ is the water depth and $v$ is the water velocity. From equations (3) and (4) the propagated source at the water-bottom surface is as follows:

\[
S = D + e^{i\omega \Delta t} U.
\]
Figure 4: Hydrophone-geophone summation. From left to right: hydrophone component, geophone component, summation.

Figure 5: Physical model in study. From Soubaras (1996).
The calibration methodology assumes that the source energy should be zero after a time equal to the sum of the source-receiver propagation time and the source duration, which is a few hundred milliseconds. Combining equations (2) and (5) yields the following relation between the propagated source \( S \) and the hydrophone \( P \) and geophone \( Z \) components:

\[
S = P' - CZ',
\]

where:

\[
P' = \frac{1 + e^{i\omega \Delta t}}{2} P,
\]

\[
Z' = \frac{1 - e^{i\omega \Delta t}}{2} Z.
\]

The propagated source vanishes after a certain period of time if the hydrophone and geophone are calibrated. This corresponds to finding \( C \) such that the propagated source \( S \) has minimum energy after a period of time:

\[
\min_S ||S_{[a,b]}||^2.
\]

The solution for this simple least-squares problem is as follows:

\[
C = \frac{P'Z'}{Z'Z' + \epsilon^2},
\]

where \( \epsilon \) is a small constant to avoid dividing by zero.

The filter \( C \) [equation (8)] is for a single trace. To obtain a filter for the entire gather, we compute the filter \( C \) for each trace and average them.

Figure 6 shows the hydrophone component of the receiver gather (left), the geophone component of the receiver gather (center) and the calibrated geophone (left).

**Deghosting**

After the calibration, the deghosting is as simple as taking the average between the hydrophone and calibrated geophone components:

\[
U = \frac{P + Z}{2}.
\]

Figure 7 compares the receiver gather of both the geophone component and the combined signal; observe that most of the ghost reflections have been eliminated. This can also be seen in the CMP gather of both the geophone component and the combined signal (Figure 8), where the arrows point to some of the multiples that have been removed.
MIGRATION RESULTS

The previous two sections show that the separate procedure of calibration and data combination provides better results than just calibrating and combining the data in one step. However, to verify whether we have effectively eliminated some multiples, we perform a poststack migration on the data and compare the results before and after the combination.

Wolf et al. (2004) propose a methodology that calculates stacking velocities without picking through a robust median estimator manufactured from neighboring traces only. The methodology depends on the estimation of the local step out; therefore, its accuracy strongly depends on the estimation of the dip field. Finally, the local estimate of the RMS velocity is:

$$V_{RMS}^2 = \frac{x dx}{t dt}$$

The local step outs are estimated with Fomel (2000) method. Figure 9 shows the result of this methodology over three characteristic CMP gathers of the combined P-component Mahogany data set. From left to right, the figure shows the CMP gather, the dip field, and the RMS velocity function. The first CMP gather corresponds to the East part of the salt body, the second gather corresponds to the center part of the section, the third gather corresponds to the West part of the salt body. We estimate the velocity model for several CMPs, then perform linear interpolation and smoothing. Figure 10 shows the final slowness model.
Figure 11 presents the first 4000 m of the migrated seismic line. The top part shows the result of the combination and the bottom part shows the geophone component alone. In the same way, Figure 12 exhibits a close-up view of the migration result.

Several multiples have been attenuated; as indicated by the arrows in both Figures 11 and 12. This is an encouraging result; it reflects that our method produces a reasonable result. However, some multiples are still present in our final result. For example, notice the strong event that follows the water bottom reflection. Note that the water bottom can be considered flat, at a constant depth of approximately 118 m.

These multiples correspond to the source ghost. Further multiple-attenuation processes should be performed. For example, a wave-equation based multiple reduction technique can easily remove the source ghosts, since the water bottom is almost flat. This is a technique we have yet to test on this data set.

CONCLUSIONS

- Pseudo-unitary Radial-trace Transform can be used to successfully eliminate radial noise present in the hydrophone component of this 2D OBC line.
Figure 8: Geophone and deghosted gather, CMP gather.

- The time-domain methodology certainly removes some of the ghost energy. However, the frequency-wavenumber-domain method is more efficient in eliminating the ghost reflection, because it splits the geophone calibration and the receiver ghost elimination into two steps.

- We satisfactorily removed the receiver ghost on this data set, as is clearly shown in a comparison with the poststack migration result. However, further multiple-attenuation process is needed.

REFERENCES


Figure 9: An example of the initial velocity model. From left to right: A typical combined CMP gather, dip field, RMS velocity function. From top to bottom, a gather taken from: The East part of the salt body, the center part of the section, the West part of the salt body.
Figure 10: Initial P slowness model.

Figure 11: Comparison of the zero-offset section for the migration result of the combined signal (top) and the geophone component alone (bottom).
Figure 12: Detailed view of the migration result of Figure 11.


Short Note

Surface boundary condition for one-way wave equation shot-profile migration

Alejandro A. Valenciano, Thomas Tisserant, and Biondo Biondi

INTRODUCTION

The one-way approximation to the full wave equation has been widely used for imaging the earth’s interior (Claerbout, 1985). This approximation ignores back-scattering in the wavefield but usually works well with surface seismic data. In the shot-profile migration scheme, two one-way wave equations need to be solved. One downward extrapolates the source wavefield and the other downward extrapolates the receiver wavefield. By using an imaging condition, the subsurface image is formed (Claerbout, 1971).

Usually, the initial surface boundary condition for downward extrapolating the source wavefield is chosen to be an impulse convolved with a wavelet at the shot position. However, downward extrapolation with the one-way wave equation requires boundary conditions on the surface, \( z = 0 \), at all locations and all times. Since horizontal (or high-propagation-angle) waves, reflected waves, and overturned waves can all contribute to the surface wavefield, it appears that the full solution must already be known in order to supply these initial data (Nichols, 1994).

Zhang (1993) proposes two correction terms to the traditional one-way wave equations. One is a correction for the source and receiver wavefields extrapolation, while the other is a new boundary condition for the source wavefield. This new boundary condition is not only an impulse at the shot position but also includes contributions at different times and surface positions, depending on the surface velocity.

In this short note, we introduce an alternative surface boundary condition for the source wavefield downward extrapolation i.e, the Green function corresponding to the Helmholtz equation for a constant velocity medium at \( z = 0 \). It includes contributions at different times and surface positions depending on the surface velocity, similar to Zhang’s boundary condition, but differs in the amplitude.

We compare the wavefront amplitudes calculated using the new surface boundary condition, the traditional surface boundary condition, and Zhang’s surface boundary condition. The

1email: valenciano@sep.stanford.edu, thomas@sep.stanford.edu, biondo@sep.stanford.edu
comparisons are made using three synthetic velocity models: constant, vertically varying, and Marmousi. The results show that the new surface boundary condition produces more balanced wavefronts in the constant-velocity model and in the vertically varying velocity model (VOZ) and behaves similarly to Zhang’s in the Marmousi velocity model.

**SURFACE BOUNDARY CONDITIONS**

**Point source**

The traditional boundary condition for the source wavefield extrapolation with the one-way wave equation (Claerbout, 1985) can be written as follows:

\[ D(x, z = 0; \omega) = W(\omega) \delta(x - x_s), \]  

where \( D(x, z = 0; \omega) \) is the source wavefield at the surface, \( W(\omega) \) is the wavelet, and \( \delta(x - x_s) \) is an impulse at the shot position \( x_s \). Although kinematically correct, Claerbout’s scheme fails to provide the correct wavefield amplitudes.

**WKBJ approximation**

Zhang (1993) proposes a different surface boundary condition, together with a correction for the source and receiver wavefields extrapolation, aiming to improve the dynamic information of the one-way wave equation. Zhang et al. (2001) apply this surface boundary condition to obtain a true-amplitude shot profile migration result in the WKBJ sense (Shan and Biondi, 2003). Zhang’s surface boundary condition is stated as follows:

\[ p_D(x, z = 0; \omega) = \frac{1}{2} \Lambda^{-1} W(\omega) \delta(x - x_s), \]  

where

\[ \Lambda = ik_z = i\sqrt{\frac{\omega^2}{v^2} - k_x^2}, \]  

\( p_D(x, z = 0; \omega) \) is the new source wavefield that satisfies the corrected one-way wave equation (Zhang et al., 2001), \( \Lambda \) is the square-root operator, \( v \) is the medium velocity, \( k_z \) is the vertical wavenumber, and \( \mathbf{k} = (k_x, k_y) \) is the horizontal wavenumber vector.

This boundary condition is not only an impulse at the shot position (Figure 1a) but also includes a contribution at different times and surface positions depending on the surface velocity (Figure 1b). This appears to resolve the contradiction discussed by Nichols (1994), by creating a V-shaped curve as the surface boundary condition. It mimics a wavefield with a high angle of propagation at the surface, resulting in more homogeneous wavefronts.

The implementation of Zhang’s surface boundary condition has the disadvantage that the square root operator (\( \Lambda \)) is undefined for high values of the spatial wavenumber. That is why there is the need of establishing a cut off for the spatial wavenumber limiting the accuracy of the steep angles.
Another possible surface boundary condition could be the Green function of the Helmholtz equation for a constant velocity medium at $z = 0$, convolved with the wavelet:

$$D_g(x, z = 0; \omega) = W(\omega)G(x, z = 0; \omega),$$

(4)

where

$$G(x, z = 0; \omega) = \frac{1}{4\pi r} e^{i\frac{\pi}{r}},$$

(5)

is the Green function at the surface, $W(\omega)$ is the wavelet, and $r$ is the distance to the shot position.

The idea is illustrated in Figure 2, where a linear moveout depending on the surface velocity dictates the time-position of the wavefront. As time increases, the amplitude should decay as the inverse of the distance to the shot position (geometrical spreading). Figure 1c shows the new boundary condition.

As oppose to Zhang’s approximation the Green function is computed without the need of the square root operator. Thus, there is no need of cutting off the spatial wavenumbers. This allows to keep the accuracy of the steep angles.
RESULTS

Constant velocity medium

Figure 3 shows a comparison of three snapshots in the \((x, z)\) plane of the source wavefield computed for the three different surface boundary conditions in a constant-velocity medium, with phase-shift extrapolation. Ideally the wavefront amplitude should be homogeneous with propagation angle.

Figure 3: Snapshot of the source wavefield in a constant velocity medium using (a) the traditional surface boundary condition, (b) Zhang’s surface boundary condition, and (c) the new surface boundary condition.

Figure 4 shows a comparison of the amplitudes along the wavefront depending on the propagation angle. We can see that for the traditional surface boundary condition (dotted curve), the amplitude along the wavefront is not uniform—it is high for sub-vertical propagation, but it fades away at higher angles and completely disappears for horizontal propagation. The curve corresponding to Zhang’s surface boundary condition (dashed curve), although behaving better than the traditional one, fails to produce the correct amplitude at high propagation angles. On the other hand, the proposed surface boundary condition (solid curve) presents homogeneous amplitude at every angle.

Figure 4: Amplitudes along the wavefields in a constant velocity medium.
Vertically varying velocity (VOZ) medium

Figure 5 shows a comparison of three snapshots in the \((x, z)\) plane of the source wavefield computed for the three different surface boundary conditions in a VOZ medium. The extrapolation method used was phase-shift.

Figure 5: Snapshot of the source wavefield in a vertically varying velocity using (a) the traditional surface boundary condition (b) Zhang’s surface boundary condition (c) the new surface boundary condition.

Figure 6 shows a comparison of the amplitudes along the wavefront depending on the propagation angle. We can see that for the traditional surface boundary condition (dotted curve), the amplitude along the wavefront fails to follow the theoretical trend (Winbow, 1995) (dot-dash curve)–it is high for sub-vertical propagation, but it fades away at higher angle and completely disappears for horizontal propagation. The curve corresponding to Zhang’s surface boundary condition (dashed curve) and the curve corresponding to the proposed surface boundary condition (solid curve) match the theoretical curve up to 70°. After that angle, Zhang’s surface boundary condition gives a lower value than the theoretical while the proposed surface boundary condition gives a higher value than the theoretical. Overall, the curve corresponding to the proposed surface boundary condition agrees better with the theoretical curve.

Figure 6: Amplitudes along the wavefields in a VOZ medium.
Complex velocity medium (Marmousi)

The Marmousi synthetic data set (Bourgeois et al., 1991) was first released as a blind test for velocity estimation. It has become a popular testbed for migration algorithms. Its structural style is dominated by growth faults, which arise from salt creep and give rise to the complicated velocity structure in the upper part of the model.

Figure 7 shows a comparison of three snapshots in the \((x, z)\) plane of the source wavefield, computed using the three different surface boundary conditions in the Marmousi model. The extrapolation method used was split-step with 7 reference velocities (Stoffa et al., 1990). In the Marmousi model, the wavefront amplitude should not be homogeneous with the propagation angle. This is because of the complexity of the velocity model.

![Figure 7: Snapshot of the source wavefield in the Marmousi velocity model using (a) the traditional surface boundary condition, (b) Zhang’s surface boundary condition, and (c) the new surface boundary condition.](alejandro2-comp_wfld_Marm)

Figure 8 shows a comparison of the amplitudes along the wavefront depending on the polar angle (polar system center at the shot position). We can see that the amplitudes along the wavefront have a complex behavior. In general, for the traditional surface boundary condition (dotted curve), the wavefront shows lower amplitudes at higher angles than do the curves corresponding to Zhang’s surface boundary condition (dashed curve) and the proposed surface boundary condition (solid curve).

![Figure 8: Amplitudes along the wavefields in the Marmousi velocity model.](alejandro2-comp_max_Marm)
We migrated the Marmousi model data with a split-step algorithm using 7 reference velocities. Figures 9a and 10a show the image obtained using the traditional surface boundary condition, figures 9b and 10b show the image obtained using Zhang’s surface boundary condition, and figures 9c and 10c show the image obtained using the new surface boundary condition. Both figure 9b (10b) and figure 9c (10c) show steep dips more accurately than 9a (10a).

CONCLUSIONS

We presented a new surface boundary condition for downward continuation using the one-way wave equation which improves the amplitude distribution along the wavefronts. This new surface boundary condition is the Green function of the Helmholtz equation for a constant velocity medium at $z = 0$. The results, in both the constant velocity and VOZ models, show that the new surface boundary condition produces wavefront amplitudes that match better the theoretical curve than the traditional and Zhang’s.

Both Zhang’s surface boundary condition and the Green function surface boundary condition behave similarly in the Marmousi velocity model, resulting in better imaging of steep dips than those obtained with the traditional surface boundary condition.

More research needs to be done to study the AVA amplitude responses and phase differences due to the use of different surface boundary conditions.

ACKNOWLEDGMENT

We thank Tom Dickens for pointing to a reference for the theoretical Green function in the 2D VOZ medium and for important suggestions. We also thank Anatoly Baumstein and John Anderson for insightful discussions.

REFERENCES


Figure 9: Marmousi image using (a) the traditional surface boundary condition, (b) Zhang’s surface boundary condition, and (c) the new surface boundary condition.
Figure 10: Close-up of the Marmousi image using (a) the traditional surface boundary condition, (b) Zhang’s surface boundary condition, and (c) the new surface boundary condition.


Short Note

Improving the amplitude accuracy of downward continuation operators (Part 2)

Ioan Vlad and Thomas Tisserant

INTRODUCTION

One-way wavefield continuation methods correctly account for traveltimes, but the amplitude and phase of the images they produce can still be improved. Zhang et al. (2002; 2003a; 2003b) present theoretical formulations of amplitude-improving corrections for shot-profile migration. Vlad et al. (2003) show concrete ways of implementing Zhang’s theory for both finite-difference and mixed-domain extrapolators, with applications to constant-velocity, laterally smooth and heterogeneous cases.

The above-described corrections consist of two parts: one that is applied at the $z = 0$ boundary, and one that is applied in the propagation operator. The boundary condition correction depends only on the velocity at the surface, while the propagation correction takes the entire interval velocity model into account and is directly proportional to its vertical gradient, $v_z$, vanishing where this becomes zero. As a result, the effect of the boundary condition correction can be tested in isolation, using constant velocity models. Figure 1 in Vlad et al. (2003) shows that the Zhang boundary condition correction improves the phase and brings the amplitudes very close to the ones computed analytically, especially in the case of the mixed-domain implementation. Valenciano et al. (2004) show the effect of applying it to propagation through the Marmousi model. They also present a more intuitive interpretation of the boundary condition correction, and an alternative method for computing it.

Vlad et al. (2003) show in their Figure 2 that the propagation operator correction is equivalent to the WKBJ one in a $v(z)$ medium. They also apply it to the $v(x, z)$ model presented in Figure 3 of their paper. However, the results (their Figure 4) did not show a visible effect of the correction. In the first section that follows, we explain why that was the case for that very particular velocity model, and we quantitatively show that the propagation operator correction does have visible effects. In the last section of this paper, we discuss the application of the Zhang corrections to the linearized downward continuation operator used in wave equation migration velocity analysis.

1email: nick@sep.stanford.edu, thomas@sep.stanford.edu
APPLICATION TO PROPAGATION IN A MORE GENERAL $V(X, Z)$ MODEL

The velocity model used by Vlad et al. (2003) for examining the effects of the amplitude corrections was particularly unsuitable for the problem. It belonged actually to a limit case in which the amplitude corrections canceled themselves. We will show below why that was the case and we will analyze the effects of the corrections by picking amplitudes at each midpoint in a wavefield depth slice.

At each downward continuation step, the propagation amplitude correction applied to the wavefield is:

$$P_z' = P_z e^{-\frac{v_z}{2\pi k_z} \left[ \frac{1}{1 + \frac{v_z}{2\pi k_z}} \right]^2},$$  \hspace{1cm} (1)

where $v_z$ denotes the vertical gradient of the velocity. When variations of velocity in the $x$ direction exist, this exponential is strictly noncommutative with the downward continuation step. The noncommutativity however becomes weak in the particular case when the laterally varying velocity is symmetrical with respect to a horizontal plane. In the case of such symmetry, corrections of the same magnitude along midpoint, but of different sign (because of an opposing sign for $v_z$) cancel each other. The corrections in the lower half of the velocity model in Figure 3 of Vlad et al. (2003) were therefore erasing the effects of the ones performed in the upper half. As a result, Figure 4 of Vlad et al. (2003) was not showing any results of the correction.

Recognizing that such a symmetrical configuration is not very plausible geologically, we downward continued (with split-step) only through the upper half of the respective velocity model. This half is depicted in Figure 1. The top panels in Figure 2 show the wavefield at a depth of 1000m. The four panels represent all the combinations of applying ("+") or not ("-”) the boundary condition correction ("L") and/or the propagation correction ("G"). The lower half of the figure shows the maximum amplitudes picked for each midpoint. The effect of the propagation correction (Curves “-L+G” and “+L+G”) is visible now as increased focusing and reduced amplitude decay with offset. The boundary condition correction (Curves “+L-G” and “+L+G”) also has a strong effect; it especially diminishes the amplitude decay. We expect the two Zhang corrections to increase, when cummulated, power for large dips and for large incidence angles on the reflector.

APPLICATION TO LINEARIZED DOWNWARD CONTINUATION

Linearized downward continuation – purpose and description

Let us define a slowness perturbation $\Delta s = s - s_0$ as a difference between two slowness models, one of which ($s_0$) is named the “background slowness”. By undertaking several approximations, the most notable of which is Born, the mixed-domain downward continuation operator can be written as an explicit function of the slowness perturbation. This allows an explicit relation between the slowness perturbation and the wavefield. In conjunction with the imaging
Figure 1: **Top panel:** Upper half of the velocity model used by Vlad et al. (2003). Figure 2 shows the amplitudes produced by propagation through this velocity model. **Bottom panel:** Rays shot from (0,0), shown for an enhancement of the intuitive appreciation of the focusing during the propagation.

condition, this allows writing an explicit relation between the slowness perturbation and the image. This relation is the basis of Wave-Equation Migration Velocity Analysis (WEMVA), a complete flowchart of which is presented in Figure 3. This procedure finds the velocities in the following way: using the recorded data and the background slowness, it creates a background image. This image is then improved so that it is closer to the optimally focused one, then the two images are subtracted to create an image perturbation. This is transformed into a wavefield perturbation through an inverse imaging condition, then is upward continued, to create an adjoint scattered wavefield. This in turn is transformed into a slowness update by inverting the linearized downward continuation operator. This operator is linearized so that its inversion will be computationally cheap. A complete derivation is provided by Biondi and Sava (1999), with more explanations in Sava (2000).

If we denote the wavefield as $U$, the linearized downward continuation (complexified local Born-Fourier method), according to Appendix B in Vlad (2002), is given by:

$$U_{z=n\Delta z} = \left( \prod_{1}^{n} \mathcal{T} \right) U_{z=0} + \sum_{j=1}^{n} \left( \prod_{1}^{n-j} \mathcal{T} \right) \delta \Delta s_{z=j\Delta z} \left( \prod_{1}^{j} \mathcal{T} \right) U_{z=0}, \quad (2)$$

where $\mathcal{T}$ is the background wavefield downward continuation operator:

$$\mathcal{T} = e^{i \Delta z \sqrt{\omega^2 k_0^2 - (1-i\eta)^2 |k_m|^2}}, \quad (3)$$
Figure 2: “L” denotes the boundary condition correction, “G” denotes the propagation correction, a "+" shows that the respective correction was applied, and a “-” shows that it was not. **Top panels:** Wavefield generated by a shot at (0,0) in the velocity model from Figure 1, and recorded at the bottom of that velocity model. **Bottom panel:** Maximum amplitudes picked for each x location in the top panels.
Figure 3: WEMVA flowchart, provided for illustrating the use of linearized downward continuation.

\[ \delta = \frac{i \Delta z \omega^2 s_o}{\sqrt{\omega^2 s_o^2 - (1 - i \eta)^2 |k_m|^2}}. \] (4)

**The boundary condition correction**

We tested the amplitude correction at the boundary condition by downward propagating a shot and picking the amplitudes of the wavefield recorded at a certain depth. Since our particular purpose of using WEMVA is finding velocity anomalies that generate focusing-effect AVO (Vlad and Biondi, 2002), we used the same focusing-generating velocity model as that presented in the lower right panel of Figure 6 of Vlad (2002). For convenience, we present it in the upper panel of Figure 4, also tracing wavefronts through it for a better visualization of the kinematics of propagation. The time delays induced by the presence of the low-velocity slab have not been shown because they are under the common picking threshold – this is a typical focusing-effect AVO case.

To obtain the curves in the lower panel of Figure 4, for each of the three methods, we propagated a wavefield through the velocity model in the upper panel of the figure (with the low-velocity slab), and another wavefield through the constant-velocity background only.
each x location, we picked the maximum amplitudes of each wavefield at the depth of 6000m, and we divided the amplitudes obtained from the model with the slab by the amplitudes obtained from the constant-velocity background. A deviation from the value of 1 indicates the presence of the wavefield scattered by the slab. We performed this procedure using three different algorithms: (1) – Linearized downward continuation with no amplitude correction applied; (2) – Linearized downward continuation with boundary condition amplitude correction applied at the surface; (3) – Pseudospectral wave propagation (Biondi, 2002), for reference. The application of the boundary condition correction has brought the values closer to those of the reference curve. A possible shot-profile formulation of WEMVA would therefore benefit from the application of the boundary condition correction.

Figure 4: **Top panel:** Velocity model for testing the amplitude behavior of the linearized downward continuation operator. A shot has been generated at (0,0) and the wavefield is recorded at a depth of 6 km. The thin slab has a velocity of 1647 m/s, contrasting to the background of 1830 m/s. Wavefronts have been traced for better visualization of propagation kinematics. **Bottom panel:** Curve “-L” – Linearized downward continuation with no amplitude correction applied; Curve “+L” – Linearized downward continuation with boundary condition amplitude correction; Curve “Ref” – Pseudospectral wave propagation.
Improving amplitudes

The propagation operator correction

Things are different, however, for the amplitude correction which is applied during the propagation. Equation 28 of Vlad et al. (2003) states that the correction to be applied at each downward continuation step is

\[
U'_z = U_z e^{-\frac{v_z \Delta z}{2} \left( k_x \left( \frac{k_x}{\omega} \right)^2 \right) + \frac{v_z \Delta z v^2}{4} \left( \frac{v_z \Delta z}{4} - v \right) \left( \frac{k_x}{\omega} \right)^4}.
\]

(5)

Counting a possible linearization of the exponential in front and the presence of \( v_z \), this is an expression of degree 5 in velocity. Moreover, at each depth step this has to be multiplied with the first-order-in-slowness propagation step in Equation 2. The result would be a expression of at least degree 6. Discarding terms of order higher than 1 would result in losing an important amount accuracy in the propagation step itself, since a part of it will be multiplied with higher-order amplitude correction terms. We conclude that the propagation operator amplitude correction is not applicable to linearized downward continuation because of compounding linearization errors that will affect both the kinematics and the amplitudes.

CONCLUSIONS

We continue the work of Vlad et al. (2003) on implementing the Zhang amplitude corrections for wavefield extrapolation imaging. We apply the corrections to the case of propagation through a less particular \( v(x, z) \) velocity model than the one used in the previous paper. The corrections result in an increase in wavefield focusing. We apply with encouraging results the boundary condition correction to the linearized downward continuation used in wave-equation migration velocity analysis. The propagation operator correction, however, is not applicable to linearized downward continuation.

ACKNOWLEDGMENTS

We thank Guojian Shan for performing the pseudospectral wave modeling used in Figure 4.

REFERENCES


Sava, P., 2000, A tutorial on mixed-domain wave-equation migration and migration velocity analysis: SEP–105, 139–156.


Least-squares joint imaging of multiples and primaries applied to 3-D field data

Morgan Brown

ABSTRACT
In this paper I outline the extension to 3-D of the Least-squares Joint Imaging of Multiples and Primaries (LSJIMP) method for simultaneously separating multiples and primaries and combining their information. I apply LSJIMP to a 3-D field dataset and demonstrate that the method cleanly removes surface-related multiples from the data while preserving the prestack amplitude signature of the primaries. LSJIMP compares favorably to least-squares Radon demultiple, both in terms of computational performance and result quality.

INTRODUCTION
Multiple reflections often erect the highest barrier to the successful imaging and interpretation of marine seismic data. Despite their nuisance, however, multiples illuminate the prospect zone, and moreover, illuminate different angular ranges and reflection points than primaries. In theory and in practice, multiples provide subsurface information that primaries do not. To use the information in the multiples, we must map multiples and primaries to a domain where they are comparable, like a prestack image domain, and then combine them.

The semi-independent measurements provided by primaries and multiples overlap each other in one data record. In theory, averaging the multiple and primary images can improve signal fidelity and fill coverage gaps, but this encounters two problems in practice. First, unless we correct multiples for their different raypaths and additional reflections, the signal events are incomparable. Secondly, just as multiples represent noise on the primary image, primaries and higher order multiples represent noise on a first-order multiple image. Corresponding noise, or “crosstalk” events on the images are often kinematically consistent, so adding the images may actually degrade signal fidelity. Prestack separation of multiples and primaries is a prerequisite to image averaging.

I presented the LSJIMP method to simultaneously solve the separation and integration problems, as a global inversion (Brown, 2003b). The model space is a collection of images, with the energy from each mode partitioned into only one image. The forward model contains amplitude corrections which ensure that the signal events in the multiple and primary images are directly comparable, in terms of both kinematics and amplitudes. I presented an efficient
kinematic (Brown, 2003c) and amplitude (Brown, 2003a) modeling scheme which combines the efficiency necessary to realistically run on real 3-D data with the accuracy to model pegleg multiples in a moderately complex earth. Three model regularization operators exploit multiplicity between and within the images to discriminate between crosstalk and signal, combine the images, fill coverage gaps, and increase signal fidelity.

I previously applied LSJIMP to a difficult 2-D Gulf of Mexico data example (Brown, 2003b). In this paper, I apply LSJIMP to a 3-D data example, also from the Gulf of Mexico. CGG donated the data, which were acquired in Green Canyon. I demonstrate that LSJIMP can cleanly separate the multiples and primaries in the data, in spite of sparse data geometry that inhibits many other multiple suppression techniques. I show that LSJIMP compares favorably to least-squares Radon demultiple, both in terms of computational performance and result quality. Last, I show that the LSJIMP estimated primaries are more robust with respect to amplitude-versus-offset (AVO) analysis than the raw data.

**LSJIMP REVIEW**

LSJIMP models the recorded data as the sum of primary reflections and \( p \) orders of pegleg multiples from \( n_{\text{surf}} \) multiple generators. An \( i^{\text{th}} \)-order pegleg splits into \( i + 1 \) legs. Denoting the primaries \( d_0 \) and the \( k^{\text{th}} \) leg of the \( i^{\text{th}} \) order pegleg from the \( m^{\text{th}} \) multiple generator \( d_{i,k,m} \), the modeled data takes the following form:

\[
d_{\text{mod}} = d_0 + \sum_{i=1}^{p} \sum_{k=0}^{n_{\text{surf}}} \sum_{m=1}^{n_{\text{surf}}} d_{i,k,m}. \tag{1}
\]

If we have designed imaging operators that map primaries and multiples to comparable signal events (kinematics and angle-dependent amplitudes), we can write the \( d_{i,k,m} \) as linear functions of prestack images. We can similarly denote the modeling operators (adjoint to imaging) for primaries and peglegs as \( L_0 \) and \( L_{i,k,m} \), respectively, and the images of the primaries and peglegs as \( m_0 \) and \( m_{i,k,m} \), respectively. Rewriting equation (1), we have:

\[
d_{\text{mod}} = L_0 m_0 + \sum_{i=1}^{p} \sum_{k=0}^{n_{\text{surf}}} \sum_{m=1}^{n_{\text{surf}}} L_{i,k,m} m_{i,k,m} = L m \tag{2}
\]

The LSJIMP method optimizes the primary and multiple images, \( m \), by minimizing the \( \ell_2 \) norm of the difference between the recorded data, \( d \), and the modeled data, \( d_{\text{mod}} \):

\[
\min_m \| d - L m \|^2. \tag{3}
\]

Minimization (3) is under-determined for most choices of \( L_0 \) and \( L_{i,k,m} \), implying infinitely many solutions. Crosstalk leakage is a symptom of the problem. For instance, \( L_0 \) maps residual first-order multiple energy in \( m_0 \) to the position of a first-order multiple in data space. Minimization (3) alone cannot distinguish between crosstalk and signal. Without model regularization, the basic LSJIMP problem is intractable.
Previously (Brown, 2003b), I devised discriminants between crosstalk and signal, and used them to derive three model regularization operators which choose the set of primary and multiple images which are optimally free of crosstalk. Moreover, these operators exploit signal multiplicity—within and between images—to increase signal fidelity, fill coverage gaps, and combine multiple and primary information. These model regularization operators are the key to the LSJIMP method’s novelty.

To solve the regularized LSJIMP problem, we supplement minimization (3) with the three model regularization operators:

\[
\min_m \|L_m - d\|^2 + \epsilon_1^2 \|r_m^{[1]}\|^2 + \epsilon_2^2 \|r_m^{[2]}\|^2 + \epsilon_3^2 \|r_m^{[3]}\|^2. \tag{4}
\]

\(r_m^{[1]}\), \(r_m^{[2]}\), and \(r_m^{[3]}\) are the model residuals corresponding to differencing across images, differencing across offset, and crosstalk penalty weighting, respectively. Scalars \(\epsilon_1\), \(\epsilon_2\), and \(\epsilon_3\) balance the relative weight of the three model residuals with the data residual. I use the conjugate gradient method for minimization (4).

Previously (Brown, 2003c,a), I outlined an efficient prestack, true relative amplitude imaging scheme for pegleg multiples. We can rewrite \(L_{i,k,m}\) as the cascade of a differential geometric spreading correction for multiples \((G_{i,m})\), Snell Resampling to normalize a multiple’s AVO to its primary \((S_{i,m})\), HEMNO (Heterogeneous Earth Multiple NMO Operator) to kinematically image split peglegs \((N_{i,k,m})\), and finally, application of the multiple generator’s spatially-variant reflection coefficient \((R_{i,k,m})\):

\[
L_{i,k,m} = R_{i,k,m} N_{i,k,m} S_{i,m} G_{i,m}. \tag{5}
\]

The imaging operator for primaries, \(L_0\), is simply NMO. The imaged multiples and primaries on the \(m_{i,k,m}\) in equation (2) are directly comparable in terms of kinematics and amplitudes. We can exploit this important fact (with model regularization) to discriminate crosstalk from signal, and also to spread information between the images. A true relative amplitude imaging scheme like equation (5) is crucial to fully leverage the multiples in this joint imaging algorithm. Since we apply the operator and its adjoint many times in iterative optimization, computational efficiency is crucial. Because HEMNO images (like NMO) by vertical stretch, equation (5) is fast, memory-conserving, and robust to poorly sampled wavefields, which are the norm with 3-D acquisition. Brown and Guitton (2004) demonstrated that this modeling/imaging approach can accurately model multiples, even in a moderately complex earth.

### 3-D THEORY

In marine environments, three-dimensional reflection seismic data is normally acquired in a so-called “wide tow” streamer configuration, illustrated in Figure 1. Ironically, the crossline azimuth range of this data is less than with most land acquisition geometries, so geophysicists often call towed streamer data “narrow azimuth” data. Note that the crossline shot interval, \(\Delta s_y\), is chosen such that the outermost receiver line on one swath overlaps the innermost receiver line on the previous swath. Figure 2 illustrates that such an acquisition geometry
produces a regularly sampled crossline CMP axis, if cable feathering is absent. In one sense, this geometry boasts some degree of optimality, as it produces a well-sampled 3-D image at a minimum cost.

Figure 1: Wide tow marine geometry. The acquisition boat tows many (usually 4-12) receiver lines and steams in parallel sail lines.

Unfortunately, geometry shown in Figure 2 causes the 3-D extension of the SRME method of multiple prediction to fail spectacularly. SRME requires \( \Delta s_y \) to be relatively small—in practice, roughly the same as commonly chosen crossline receiver line spacing parameters (van Borstelen, 2003). 3-D field datasets commonly have a crossline shot interval of up to ten times the crossline receiver line spacing. Workarounds for the 3-D sampling problem include: ignoring crossline structure and using a 2-D prediction, massive (270,000 CPU hours) shot interpolation (Kleemeyer et al., 2003), sparse inversion of the crossline multiple contribution gathers (van Dedem and Verschuur, 2002; Hokstad and Sollie, 2003), and novel acquisition geometries (Paffenholz, 2003). Currently, none of these methods combines proven accuracy with computational/cost efficiency.

**LSJIMP AND WIDE TOW MARINE DATA**

For the particular geometry shown in Figures 1 and 2, notice that each crossline midpoint gather is occupied by one, and only one receiver line. Therefore, we can conceptualize a 3-D CMP gather as a 2-D CMP gather with nonzero crossline offset, and thus remove the crossline offset axis from five-dimensional CMP-sorted data, saving considerable memory and computational waste. This approach is similar in spirit to Biondi’s (1997) combination of azimuth moveout (AMO) (Biondi et al., 1998) and common-azimuth wave equation depth migration (Biondi and Palacharla, 1996) for the prestack imaging of primaries in data with narrow-azimuth geometries. Unfortunately, the AMO transformation is not generally valid for multiples.

HEMNO is well-suited to image multiples with the data geometry described above, primarily because HEMNO images pegleg multiples with a vertical time shift. Rather than correlating wavefields across possibly-undersampled axes like migration, HEMNO uses a measure-
Figure 2: Midpoint locations from two adjacent sail lines of the CGG Green Canyon 3-D dataset. The boat tows four streamers and fires two sources alternatively in a “flip-flop” configuration. Midpoints from shot “A” are labeled “a”, and so on. Two shot pairs from each sail line are shown. For a fixed ship speed, this geometry doubles crossline midpoint density, at the cost of reduced inline resolution, compared to a single-source configuration. Flip-flop shooting allows one airgun to be recharged while the other shoots, thereby allowing the ship to sail faster than would be possible with one gun. Cable feathering is evident, though not severe. [morgan1-cgg-midpoints] [ER]
ment of the data’s zero-offset time dip to account for structure-induced moveout variations. Because the crossline offset axis is removed, the computational cost increase of applying HEMNO to 3-D data versus 2-D data is only proportional to the number of crossline midpoints.

My particular LSJIMP implementation, outlined in two previous works (Brown, 2003c,a), uses HEMNO, an extension of the NMO equation for multiples, in conjunction with three amplitude normalization operators to produce a “true relative amplitude” image of pegleg multiples. One of these, the Snell Resampling operator, moves multiple energy across offset to make the multiple’s AVO response comparable to its primary. However, the use of Snell Resampling in the crossline direction on narrow azimuth data runs contrary to the stated assumption that we store only one crossline offset bin per 3-D CMP gather. Therefore, for the results shown in this paper, I do not apply crossline Snell Resampling. In practice, little useful angular information is anyway obtained in the crossline direction, since in most cases the data will have a maximum crossline offset of only a few hundred meters.

**CGG GREEN CANYON IV 3-D DATA**

In 2003, CGG donated a large portion of a 161-block speculative seismic survey it acquired in the Green Canyon region of the Gulf of Mexico. The data were acquired in the transition zone between the edge of the Continental Shelf and the Sigsbee escarpment which signals the edge of the abyssal plain. Geologically, the Green Canyon region is characterized by sedimentary “minibasins” interrupted by complex salt bodies (AAPG, 1998).

CGG’s 3-D data were acquired by a ship sailing east-to-west, in the strike direction relative to the dominant geologic dip. The subset of the data that I process in this paper contains fairly significant crossline dip (> 3°) in most places. Figure 3 shows a stacked section of the subset, which contains 192 midpoints inline and 14 midpoints crossline. The stacked section includes contributions from two adjacent sail lines, the geometries of which are illustrated in Figure 4.

The subset shown in Figure 3 is situated in a sedimentary minibasin, with strong reflections visible at a two-way traveltime of well over 5 seconds. Thanks to a strong velocity gradient and the sparse offset sampling, surface-related multiples are largely absent from the stacked section. Still, as we shall see, the multiples are fairly strong in the prestack data and inhibit prestack amplitude analysis. The section exhibits moderate reflector dip – an anticlinal structure in the inline direction and effectively constant crossline dip of several degrees.

The acquisition ship sailed quite fast, with a flip-flop source interval of 37.5 meters, and an interval of 75 meters between like sources. The fast ship speed leads to reduced resolution along the inline offset axis: for an 8100-meter cable with receiver group spacing of 25 meters, the nominal CMP fold is only 54, implying a nominal inline offset spacing of 150 meters. Figure 5 illustrates the sparse sampling of the inline offset axis. While the nominal inline offset bin size of 150 meters ensures that all bins will contain a live trace, such sparsity will greatly inhibit the estimation of reasonable stacking velocities and create “checkerboard” artifacts in the shallow portions of a stacked image.
Therefore, in my processing of this dataset, I use an offset bin spacing of 25 meters. While this fine sampling better honors the physics of the experiment, it leads to a fivefold increase in empty bins. Moreover, although I have cast LSJIMP primarily as a wavefield separation algorithm, recall that one major motivation of integrating multiples and primaries is to use the multiples as a constraint on the primaries in zones where we do not record data. Multiples sample reflectors more finely in reflection angle/offset than do primaries. Moreover, the regularization strategies discussed earlier provide the infrastructure to exploit the inherent multiplicity of signal within an image and between multiple and primary images. Although designed to separate signal and noise, these same strategies also prove adept at interpolating signal in missing traces.

Stacking velocities were computed by a conventional velocity scan, coupled with maximum amplitude autopicking and local weighted (stack power) mean smoothing. The residual weight, simply zero for missing traces, but one elsewhere, is particularly important to achieve a successful LSJIMP result.
Figure 4: Acquisition geometry for two sail lines contributing to subset shown in Figure 3. “S” symbols illustrate source positions. “M” and “F” symbols illustrate the medium- and far-offset receivers, respectively, on each of the boat’s four streamers. The nominal sail line spacing is 300 meters, although it varies considerably in this case from about 200 to 500 meters. The subset processed in this section contains shot records from about 40,000 to 50,000 meters inline location. Cable feathering in this zone is present, though not severe. The two sail lines overlap to some extent, which reduces the number of crossline midpoint locations from these two sail lines to 14 from 16 over the subset.

Figure 5: Illustration of sparsity of inline offset axis of CGG Green Canyon IV data. Grid represents nominal bin size of 150x12.5 meters. Black dots correspond to trace locations.
RESULTS

Figure 6 shows stacked sections from the multiple-infested zone of the CGG subset before and after application of LSJIMP. Figure 7 is in the same format, but shows a zoom of the multiple-infested region. CMP stacking strongly suppresses the multiples, but from the difference panel, notice that LSJIMP has nonetheless subtracted most of the remaining surface-related multiple energy, and has preserved the stronger primaries to a great extent. The timeslice on the 3-D cube transects the seabed pegleg from reflector R1; it shows up prominently on the raw data stack, as well as on the difference panel, but has been largely suppressed from the LSJIMP estimated primaries stack.

Figure 6: Stacked subset (192 midpoints inline, 14 midpoints crossline) of CGG 3-D data before and after LSJIMP. All panels windowed in time from 3.8 to 6.0 seconds and gained with $t^2$. Left: Raw data stack. Center: Stack of estimated primary image, $m_0$. Right: Stack of the subtracted multiples. Naming convention for pure first-order multiples: (reflecter)M, e.g., R1M. Naming convention for first-order pegleg multiples: (target)PL(multiple generator), e.g., R1PLWB.

Figures 8 and 9 illustrate LSJIMP’s performance on two individual CMP gathers extracted from different portions of the CGG 3-D data. It is in this domain where the strength of the LSJIMP method shines most. The raw data panels show strong surface-related multiples with an onset of around 4.3 seconds, and also fairly strong primary events under the curtain of multiples. The LSJIMP estimated primaries in panel (b) are effectively free of multiples, and moreover, since the data residual panel (f) barely contains any noticeable flat primary energy,
we have preserved the primary events. Also notice that the data residual contains little structured energy. This implies that the LSJIMP forward model accurately models the primaries and important multiples in the data. Unfortunately, much of this “unstructured” energy likely belongs to fairly weak pegleg multiples that simply appear incoherent with the data’s poor inline resolution. On Figure 9, notice that cable feathering has caused missing traces at far offsets. LSJIMP has used the data’s multiplicity and model constraints to reasonably extrapolate the missing traces.

**LSJIMP versus Radon Demultiple**

Radon demultiple remains the default multiple suppression technique in many situations, particularly in 3-D, where acquisition sparsity may inhibit other techniques. On CMP gathers, primaries and multiples normally have different apparent velocities, and a Radon transform which sums across offset using various curvature parameters will focus the two types of events in different parts of the transform panel. The most natural curvature parameter for CMP data is the velocity of the hyperbola defined by the NMO equation (Foster and Mosher, 1992). While the Hyperbolic Radon transform is a linear mapping, it is not time-invariant, and thus cannot be implemented efficiently as a Fourier domain operator. However, a multiple’s resid-
Figure 8: LSJIMP results on individual midpoint location (CMPx=100, CMPy=4). All panels decimated in offset by a factor of 6 and NMO’ed with stacking velocity, for display purposes. Panel (a): Raw data. Panel (b): LSJIMP estimated primaries. Panels (c) and (d): Estimated seabed and R1 pegleg multiples. Panel (e): Modeled data (sum of panels (b), (c), and (d)). Panel (f): Data residual (difference of panels (a) and (e)), with residual weight applied.
Figure 9: LSJIMP results on individual midpoint location (CMPx=100, CMPy=12). All panels decimated in offset by a factor of 6 and NMO’ed with stacking velocity, for display purposes. Panel (a): Raw data. Panel (b): LSJIMP estimated primaries. Panels (c) and (d): Estimated seabed and R1 pegleg multiples. Panel (e): Modeled data (sum of panels (b), (c), and (d)). Panel (f): Data residual (difference of panels (a) and (e)), with residual weight applied.
ual moveout after NMO is approximately parabolic (quadratic) with offset, so a time-invariant Parabolic Radon transform is much faster, though not as accurate (Hampson, 1986; Kabir and Marfurt, 1999).

To remove multiples, the multiple energy in the transform panel is muted, and the inverse Radon transform applied to produce multiple-free CMP data. If we define \( d \) as a raw CMP gather, \( H \) as the linear mapping between Radon transform space and data space, \( M \) as a mute operator that zeroes multiple energy in Radon transform space, and \( d_r \) as the estimated primaries, then we can express the Radon demultiple process in equation form:

\[
d_r = HM^T H^T d. \tag{6}
\]

Operator \( H \) is non-unitary (\( H^T H \neq I \)), so the amplitude of the estimated primaries will not match the recorded primaries. By casting Radon demultiple as a least-squares optimization problem, the Radon transform panel can be scaled such that \( d_r \) and \( d \) are directly comparable. We first optimize a Radon transform panel, \( p \), to minimize the data misfit:

\[
\min_p Q(p) = \|Hp - d\|^2. \tag{7}
\]

and then apply the mute operator and adjoint of \( H \) to produce the estimated primaries:

\[
d_r = HM^T p. \tag{8}
\]

The finite frequency content of the data, limited extent of the array, and the intrinsic unresolvability of velocity information at zero offset all contribute to the non-uniqueness of the least-squares Radon demultiple problem. At far offsets, events with many zero-offset traveltimes and different velocities are fit equally well by a single curvature parameter. Low-frequency data makes moveout discrimination between multiples and primaries more difficult. At near offsets, all the events are fit equally well by all curvature parameters. All these pitfalls lead to reduced resolution of events in the Radon domain. So-called “high resolution” least-squares Radon transform implementations partially overcome these problems by imposing sparsity constraints in either the hyperbolic or parabolic Radon domain (Thorson and Claerbout, 1985; Sacchi and Ulrych, 1995).

I implemented and tested least-squares Hyperbolic Radon demultiple (LSHRTD) on the CGG 3-D data subset. I performed 10 conjugate gradient iterations to produce an optimal \( p \), then applied a mute function which is zero for velocities less than 85% and greater than 115% of the known stacking velocity. The mute tapers linearly from 0.0 to 1.0 at 90% and 110% of the known stacking velocity, respectively. The computational cost of LSHRTD is very similar to the cost of applying LSJIMP.

Figure 10 compares the results of applying LSJIMP and LSHRTD on a single CMP gather from the CGG 3-D data (CMPx=100, CMPy=4). The LSHRTD results are quite good, as we expect, given the high velocity gradient and relatively simple moveout seen in this region of the data. Note some “smearing” of primaries in the LSHRTD result, as well as a generally higher level of energy removed from the data. Both effects would likely be lessened by a more conservative, tapered mute. LSJIMP is a more “surgical” separation technique, although
the model regularization operators also exploit moveout differences to separate multiples and primaries.

Figure 11, a stack of the LSHRTD estimated primaries, can be compared directly with the LSJIMP result, Figure 7. Like before, the multiples predominantly stack out, since the moveout separation is so significant. Still, a noticeable amount of multiple energy has been removed by LSHRTD, perhaps more than by LSJIMP. However, we immediately see some removed primary energy: for example, the strong primary near the bottom of the section.

**AVO Analysis Before and After LSJIMP**

Amplitude-versus-offset or AVO analysis is perhaps the most commonly utilized direct hydrocarbon indicator in exploration reflection seismology today. The stacked results before and after LSJIMP on the Green Canyon 3-D dataset showed that stacking greatly attenuates multiples. However, from the prestack data, we saw that the multiples are prominent, and would surely inhibit signal processing techniques, like AVO analysis, if left intact. In this section, I illustrate how LSJIMP improves the AVO analysis of the Green Canyon 3-D data.

Shuey (1985) showed that in a 1-D earth, the measured reflection strength of an event at the surface is approximately linear with the square of its incidence angle, at angles less than 30
degrees. In a 1-D earth, the NMO equation gives an approximate relationship between offset and incidence angle. Claerbout (1995) defines the “stepout”, $p$, as the spatial derivative of an event’s traveltime curve:

$$ p = \frac{dt}{dx} = \frac{\sin \theta}{V}, $$

(9)

where $\theta$ is the incidence angle and $V$ is the velocity at the surface; in the marine case, simply water velocity. In a 1-D earth, the traveltime curve of a primary reflection is approximately given by the NMO equation. Taking its derivative with respect to offset, then substituting into equation (9) gives the following expression for the sine of incidence angle as a function of offset:

$$ \sin \theta = \frac{V}{V_{\text{rms}}^{2}(\tau)} \sqrt{\frac{x}{\tau^{2} + \frac{x^{2}}{V_{\text{rms}}^{2}(\tau)}}}. $$

(10)

Reflection data as a function of offset may be mapped to $\sin^{2} \theta$ via equation (10), at which point the AVO “slope” and “intercept” parameters may be estimated, usually via a linear least-squares fit to the data after resampling from offset to $\sin^{2} \theta$.

Figure 12 illustrates the estimation of AVO slope and intercept parameters on a deep reflector in the Green Canyon 3-D data, before and after application of LSJIMP. The reflector, which is well under the multiples in the data, is denoted on the zero offset section with “O” symbols. The maximum amplitude in a small time window around the reflection were picked automatically, and make up the input data to the least-squares estimation.

We see that while the parameter estimates contain the same trends before and after LSJIMP, the LSJIMP result is more consistent and less “noisy” across midpoint. My implementation of LSJIMP works on a CMP-by-CMP basis, so the results shown in Figure 12 are not smoothed across midpoint. The similarity across midpoint is an expression of the true lithology – lithology which LSJIMP better reveals.

Figure 13 illustrates, as a function of midpoint, the small time windows taken around the deep reflector shown in Figure 12, before and after LSJIMP. The input data to an AVO parameter estimation are picked maximum amplitudes within the time window as a function of $\sin^{2} \theta$. Notice the significant increase in reflector clarity after LSJIMP. Also recall that the data residuals (e.g., in Figures 8 and 9) are quite small. Therefore, the cleaner reflection events after LSJIMP in Figure 13 are not only better for AVO analysis – they also fit the recorded data in a quantitative fashion. LSJIMP is not an ad hoc post-processing technique.

CONCLUSIONS

On the CGG Green Canyon 3-D dataset, LSJIMP again demonstrated an excellent ability to separate primaries and multiples. The data subset shown in the paper came from a sedimentary minibasin, which boasts a simple velocity profile, high signal-to-noise ratio, and fairly mild (but non-trivial) dips. The full dataset released by CGG contains much data recorded over
The data subset is particularly well-suited for Radon demultiple, with its large velocity gradient and gentle geology. I tested least-squares hyperbolic Radon demultiple (LSHRTD) and found that LSJIMP compares quite favorably, both in terms of computational efficiency, multiple separation, and amplitude preservation.

A quantitative study of prestack reflection amplitudes confirmed what was suspected; LSJIMP’s ability to remove multiples and random noise, as well as its ability to use multiples and other constraints to interpolate missing traces, improves prestack amplitude analysis.
Figure 13: Small time windows around the deep reflector shown in Figure 12, before and after LSJIMP. Individual panels along the vertical axis correspond to windows taken at different midpoint locations. Left: Data windows before LSJIMP. Right: Data windows after LSJIMP.

REFERENCES


Multidimensional multiple attenuation in complex geology: illustration on the Sigsbee2B dataset

Antoine Guitton

ABSTRACT
A pattern-based multiple attenuation technique is tested on the Sigsbee2B dataset. The results of multiple removal are analyzed in the image space after migration to better understand the impact of this process on the primaries. When an accurate model of the primaries (signal) and the multiples (noise) exist for the estimation of the noise and signal annihilation filters (i.e., non-stationary multidimensional prediction-error filters), this method removes the multiples extremely well while preserving the primaries. When the model of the multiples is provided by the Spitz approximation, which consists in convolving the data with the noise annihilation filters to get a signal model, good results are obtained if 3D filters are utilized.

INTRODUCTION
In the presence of complex geology where multipathing, illumination gaps, and coherent noise are present, the most advanced techniques need to be used for both preprocessing and imaging. For multiple attenuation, it has been established that methods that take the wavefield propagation into account (e.g., Verschuur et al. (1992); Weglein et al. (1997)) are the most successful in complex geology (Matson et al., 1999; Bishop et al., 2001; Paffenholz et al., 2002).

For multiple removal in complicated geology, the standard processing workflow is usually divided into a prediction step (i.e., modeling of the multiples) and a subtraction step. In the subtraction step, multiples are removed according to some assumptions made on the signal distribution (primaries) in the data. Assuming that the signal has minimum energy, the multiple model is often simply subtracted by adaptive subtraction with a $\ell^2$ norm. However, the least-squares assumption might not hold all the time (Spitz, 1999). For instance, Guitton and Verschuur (2004) show that when primaries are much stronger than the multiples, the $\ell^1$ norm should be used instead. In Guitton (2003b), I showed that a subtraction scheme based on the assumption that both primaries and multiples have different patterns leads to a successful separation. This technique approximates the multivariate spectrum (patterns) of the noise and signal with non-stationary multidimensional prediction-error filters (PEFs).

In this paper, I investigate the multiple attenuation technique with multidimensional PEFs with the Sigsbee2B dataset. This dataset is particularly challenging due to the complex geom-

1email: antoine@sep.stanford.edu
etry of the salt body. In the ideal case where an accurate noise (multiples) and signal model are known in advance, the PEF processing leads to an excellent attenuation of the multiples. If only a multiple model is known such as with the Delft approach, 3D filters should be used instead of 2D filters. This result is consistent with the conclusions in Guitton (2003b).

Often in the processing of multiples, the final results are displayed on common shot gathers, common offset sections or stacks. Because the end-product of the seismic processing workflow is always a migrated image, the outcome of a multiple attenuation technique should be analyzed in the image space (after migration) as often as possible. Therefore, I will concentrate most of my efforts in displaying multiple attenuation results in the image space with migrated images at zero-offset and angle domain common-image gathers (ADCIG).

In the next section, I derive the basic equations governing the multiple attenuation technique with multidimensional PEFs. Then, I present the results of multiple attenuation on the Sigsbee2B dataset with or without a known signal model and when 2D or 3D filters are used.

THEORY OF MULTIPLE ATTENUATION AND FILTER ESTIMATION

The key assumption of the proposed multiple attenuation technique is that the primaries and multiples have different multidimensional spectra that PEFs can approximate (Claerbout, 1992; Spitz, 1999). In this approach, the multiple attenuation is similar to Wiener filtering (Abma, 1995). One important approximation is that the multiples and primaries are uncorrelated so that the cross-spectrum between them is not needed. Multiple attenuation with multidimensional PEFs is performed in two steps. First the PEFs for the multiples and primaries are estimated. Then multiples are separated from the primaries. In the next section, I describe the multiple removal step first, assuming that the PEFs for both the primaries and multiples are known. Then I describe how PEFs are estimated and show how the signal PEFs can be computed when no signal model is given, leading to the Spitz approximation.

Multiple attenuation

First consider that any seismic data \( d \) are the sum of signal (primaries) and noise (multiples) as follows:

\[
d = s + n, \tag{1}
\]

where \( s \) is the signal we want to preserve and \( n \) the noise we wish to attenuate.

Now assuming that the multidimensional PEFs \( N \) and \( S \) are known for the noise and signal components, respectively (see the following section for a complete description of the PEFs estimation step), we have

\[
Nn \approx 0, \\
Ss \approx 0. \tag{2}
\]
by definition of the PEFs. Equations (1) and (2) can be combined to solve a constrained problem to separate signal from noise as follows:

\[
\begin{align*}
0 & \approx r_n = \mathbf{N} \mathbf{n}, \\
0 & \approx r_s = \mathbf{S} \mathbf{s}, \\
\text{subject to} & \quad \mathbf{d} = \mathbf{s} + \mathbf{n}.
\end{align*}
\] (3)

The noise \( \mathbf{n} \) can be eliminated in the last equation of the fitting goal (3) by convolving with \( \mathbf{N} \). Doing so, equation (3) becomes:

\[
\begin{align*}
0 & \approx r_n = \mathbf{N} \mathbf{s} - \mathbf{N} \mathbf{d}, \\
0 & \approx r_s = \mathbf{S} \mathbf{s}.
\end{align*}
\] (4)

Sometimes it is useful to add a masking operator on the noise and signal residuals \( r_n \) and \( r_s \) when performing the noise attenuation. It happens for example when we want to isolate and preserve parts of the data where no multiples are present. For instance, a mute zone can be taken into account very easily. Calling \( \mathbf{M} \) this masking operator, the fitting goals in equation (4) are weighted as follows:

\[
\begin{align*}
0 & \approx r_n = \mathbf{M} (\mathbf{N} \mathbf{s} - \mathbf{N} \mathbf{d}), \\
0 & \approx r_s = \mathbf{M} \mathbf{S} \mathbf{s}.
\end{align*}
\] (5)

Solving for \( \mathbf{s} \) in a least-squares sense leads to the objective function

\[
f(\mathbf{s}) = \|\mathbf{r}_n\|^2 + \epsilon^2 \|\mathbf{r}_s\|^2,
\] (6)

where \( \epsilon \) is a trade-off parameter that relates to the signal/noise ratio. The least-squares inverse for \( \mathbf{s} \) becomes

\[
\hat{\mathbf{s}} = (\mathbf{N}'\mathbf{MN} + \epsilon^2 \mathbf{S}'\mathbf{MS})^{-1} \mathbf{N}'\mathbf{MN} \mathbf{d},
\] (7)

where \( (') \) stands for the adjoint. Soubaras (1994) uses a very similar approach for random noise attenuation and more recently for coherent noise attenuation (Soubaras, 2001) with F-X PEFs. Because the size of the data space can be quite large, \( \mathbf{s} \) is estimated iteratively with a conjugate-gradient method. In the next section, I describe how both \( \mathbf{N} \) and \( \mathbf{S} \) are estimated.

**Filter estimation**

The PEFs are time-space domain non-stationary filters to cope with the variability of seismic data with time, offset and shot position. The basic equations for non-stationary PEFs estimation are (Guitton, 2003b):

\[
\begin{align*}
0 & \approx \mathbf{r}_y = \mathbf{Y} \mathbf{K} \mathbf{a} + \mathbf{y}, \\
0 & \approx \mathbf{r}_a = \mathbf{R} \mathbf{a},
\end{align*}
\] (8)

where \( \mathbf{Y} \) is a non-stationary combination matrix (Margrave, 1998), \( \mathbf{K} \) is a masking operator, \( \mathbf{a} \) a vector of the unknown PEFs coefficients, \( \mathbf{y} \) the data vector from which we want to estimate the PEFs and \( \mathbf{R} \) a regularization operator.
Often with seismic data, the amplitude varies across offset, midpoint and time. These amplitude variations can be troublesome with least-squares inversion because they tend to bias the final result (Claerbout, 1992). Therefore, it is important to make sure that these amplitude variations do not affect our processing. One solution is to apply a weight to the data like Automatic Gain Control (AGC) or a geometrical spreading correction. However, a better way is to incorporate this weight inside the inversion by scaling the residual (Guitton, 2003a). Introducing a weighting function \( W \) in the PEFs estimation leads to:

\[
0 \approx r_y = W(YK\alpha + y), \quad 0 \approx r_{\alpha} = R\alpha. \tag{9}
\]

As shown by Guitton (2003a), this weighting improves the signal/noise separation results and can incorporate a mute zone where no PEFs are to be estimated. This weight can be different for the noise and the signal PEFs. Estimating \( \alpha \) in a least-squares sense gives:

\[
f(\alpha) = \|r_y\|^2 + \epsilon^2\|r_{\alpha}\|^2, \tag{10}\]

which leads to the least-squares estimate of \( \alpha \)

\[
\hat{\alpha} = -(K'Y'W^2YK + \epsilon^2R'R)^{-1}K'Y'W^2y. \tag{11}
\]

Because many filter coefficients are estimated, \( \alpha \) is estimated iteratively with a conjugate-gradient method.

Now, prior to the signal estimation in equation (5), \( S \) and \( N \) need to be computed from a signal and noise model, respectively. The multiple model is often (but not necessarily) derived by auto-convolving the recorded data (Verschuur et al., 1992), thus obtaining a prestack model of the multiples later used to estimate a bank of non-stationary PEFs \( N \).

At this stage, a key assumption is that the relative amplitude of all order of multiples is preserved. In theory, an accurate surface-related multiple model can be derived if (1) the source wavelet is known, (2) the surface coverage is large enough, and (3) all the terms of the Taylor series that model different orders of multiples are incorporated (Verschuur et al., 1992). In practice, however, a single convolution is performed (first term of the Taylor series) which gives a multiple model with erroneous relative amplitude for high-order multiples (Wang and Levin, 1994; Guitton et al., 2001). In addition, the surface coverage might not be sufficient. This can generate wrong amplitudes for short offset traces and complex structures. Because PEFs estimate patterns, wrong relative amplitude can affect the noise estimation. However, as we shall see later, 3D filters seem to better cope with noise modeling inadequacies.

The signal PEFs are more difficult to estimate since the signal is usually unknown. However, Spitz (1999) shows that for uncorrelated signal and noise, the data PEFs \( D \) can be approximated with

\[
D = SN. \tag{12}
\]

As demonstrated by Claerbout and Fomel (2000), equation (12) is a good approximation for the data PEFs because PEFs are important where they are small. Both \( N \) and \( D \) can be estimated from the model of the multiples and the data (primaries plus multiples), respectively.
Estimation of the signal PEFs involves a deconvolution in equation (12) that can be unstable with non-stationary filters. To avoid the deconvolution step, the noise PEFs are convolved with the data:

\[ u = Nd. \]  

Then the PEFs \( U \) are estimated for \( u \) such that

\[ Uu = UNd \approx 0. \]  

From Spitz's approximation in equation (12), the following relationships hold:

\[ UN = D = SN, \]  

and \( U = S \). Therefore, by convolving the data with the noise PEFs, signal PEFs consistent with the Spitz approximation can be computed. Equation (12) insures that the PEFs \( S \) and \( N \) will not span similar components of the data space.

Thanks to the Helix (Mersereau and Dudgeon, 1974; Claerbout, 1998), the PEFs can have any dimension. In this paper, I use 2D and 3D filters and demonstrate that 3D filters lead to the best noise attenuation result. When 2D filters are used, the multiple attenuation is performed on one shot gather at a time. When 3D filters are used, the multiple attenuation is performed on one macro-gather at a time. A macro-gather is a cube made of fifty consecutive shots with all the offsets and time samples. When the multiple attenuation is done, the macro-gathers are reassembled to form the final result. There is an overlap of five shots between successive macro-gathers. In the next section, I show a prestack multiple attenuation example with the synthetic Sigsbee2B dataset.
Figure 2: Two constant offset sections (h=1125 ft) of the Sigsbee2B dataset with (a) and without (b) free surface condition. The multiples are very strong below 5 s.
ATTENUATION OF MULTIPLES WITH THE SIGSBEE2B DATASET

The Sigsbee2B dataset was designed to generate strong surface-related multiples. Figure 1 shows the true stratigraphic interval velocity model for this dataset. The data were created with a 2D acoustic finite difference modeler with constant density. Two datasets were generated: one with a free surface (Figure 2a) and one without a free surface (Figure 2b). These two datasets are such that a direct subtraction of the two leads to an almost true prestack model of the surface-related multiples. Because this multiple attenuation technique deals with the subtraction step only, the multiple model obtained by subtraction of the two datasets with and without free-surface conditions is used for the noise PEFs estimation for all my computations.

As stated in the introduction, I focus my analysis in the image space after migration. In complex geology, multiple attenuation results should always be assessed after migration; the effects of the multiple attenuation technique on the amplitudes of the primaries in ADCIGs (e.g., Figure 3), or on migrated images at zero offset (e.g., Figure 4) can be then directly assessed. For the Sigsbee2B dataset, a split-step double square-root (DSR) migration code with three reference velocities is used.

Ideally, because the data were created with finite differences, a finite differences migration code should be used to take the full complexity of the wave propagation into account. Figure 3d illustrates the limits of the migration algorithm by showing non-flat events below the salt. In contrast, Figure 3b displays flat gathers in the sedimentary section left of the salt body. An illumination effect is clearly visible below 20,000 ft, between 20 and 30 deg in Figure 3b. Figures 3a and 3c highlight the effects of the multiples on ADCIG by creating numerous artifacts and fictitious events.

It is interesting and somewhat surprising to see that in Figure 4a the multiples are very weak after migration below the salt compared to the constant offset sections in Figure 2a. In particular, the water bottom multiple seems to disappear. This is because the multiples are extremely distorted by the migration process in the vicinity of the complex salt structure. Compared with the migration of the primaries only in Figure 4b, the multiples in Figure 4a are masking a lot of primaries in the deepest part of the model and thus need to be removed.

In the next section, I demonstrate that in the ideal case where a model for both the primaries and multiples exist, the primaries can be recovered with almost no bias from the attenuation technique. Then, without any model for the primaries, I show that the Spitz approximation gives an excellent multiple attenuation result when 3D filters are used.

Estimating biases

A bias is a processing footprint left by the multiple attenuation technique, e.g., edge effects from the non-stationary PEFs. In an ideal but unrealistic case, a model for both the primaries and the multiples might be available. In this case, a bias is also any difference between the true primaries and the estimated primaries after attenuation of the multiples. In this section I demonstrate that the bias is minimum with the pattern-based approach.
Figure 3: ADCIGs for the migrated data at two different surface locations inside (i.e., 33 kft) and outside (i.e., 30 kft) the salt boundaries. (a) ADCIG at 30 kft for the data with multiples. (b) ADCIG at 30 kft for the data without multiples. (c) ADCIG at 33 kft for the data with multiples. (d) ADCIG at 33 kft for the data without multiples. Theoretically, with the appropriate migration algorithm, this last gather should be flat everywhere.
Figure 4: Migrated images at zero-offset for the data with (a) and without (b) free surface condition. Comparing with Figure 2, the multiples appear much weaker below the salt after migration. However, some reflectors near 22 kft are hidden in (a).
For the model of the primaries, the answer, i.e., the data modeled without the free surface condition is used. For the multiples, the difference between the modeled data with (Figure 2a) and without (Figure 2b) free surface is used. Because the noise and signal PEFs are estimated from almost the exact answer, only 2D filters are estimated. 3D filters can help if the primaries and multiples are correlated in time and offset but uncorrelated across shot position. With 2D filters, the attenuation is performed one shot gather at a time. Figure 5a displays the estimated primaries and Figure 5b the difference with the true primaries (Figure 2b). The bias introduced by the attenuation method is very small. 3D filters would have probably given better results where the difference between Figure 5a and 2b is the strongest (e.g., near 60 kft).

Looking now at the same estimated primaries after migration in Figure 6a, we see again that the attenuation gives an excellent result with almost no bias. Some energy is visible in the difference plot in Figure 6b between Figures 6a and 4b where no multiples are actually present. These artifacts come from the fact that the modeled data with and without free surface condition are not perfectly similar where the primaries are located.

From these results it appears that the quality of the multiple attenuation depends essentially on the filters. If the primaries and multiples are known, the primaries with almost no bias are recoverable. Therefore, we should always try to find the best models for the signal and the noise. In practice, a very accurate model of the multiples can be estimated with the auto-convolutional process of the Delft approach. For the primaries, the next section shows that the Spitz approximation gives a very good model if 3D filters are used.

Testing the Spitz approximation

Now I assume that only a model of the multiples is known. The Spitz approximation in equation (12) shows how the PEFs for the signal can be estimated. Because deconvolution with non-stationary filters is unstable, the noise PEFs is convolved with the data and the signal PEFs is then estimated from the convolution result.

The primaries are recovered with 2D and 3D filters. Figure 7 displays two constant offset sections after multiple attenuation with 2D and 3D PEFs. 3D PEFs give by far the best results and attenuate multiples very well. After migration, we see again in Figure 8 that the 3D PEFs better attenuate the multiples. A close-up in Figure 9 demonstrates in more detail how the two results with 2D or 3D filters differ below the salt. Events are more continuous and better preserved with 3D filters. Comparing with the true reflectors in Figure 9a, important primaries are attenuated with both 2D and 3D filters. This is particularly true where the signal is the weakest (between midpoint positions 35 and 50 kft).

This important observations could not have been made before migration in the prestack domain because the primaries are much weaker than the surface-related multiples below the salt. This illustrates the point that for complex geology, the quality of a multiple removal technique should be assessed in the image space as often as possible. For this analysis to be successful, accurate velocity and migration algorithm must be provided as well. Finally, the fact that some primaries are attenuated in Figure 9 should motivate us in devising improved strategies for building better noise and signal models.
Figure 5: Two constant offset panels at h=1125 m. for the estimated primaries (a) and the difference with the true primaries (b). The true primaries and multiples are used to estimate the PEFs.
Figure 6: (a) Migration result after multiple attenuation when the true primaries and multiples are used to estimate the PEFs. (b) Difference between (a) and Figure 6b. The estimated primaries are almost exact.
Figure 7: Two constant offset sections (h=1125 ft) after multiple attenuation with the Spitz approximation using (a) 2D and (b) 3D filters.
Figure 8: Two migration results of the estimated primaries with (a) 2D and (b) 3D filters.
Figure 9: Close-up of Figure 8 showing two migrated images when (b) 2D and (c) 3D filters are used. The true primaries are shown in (a).

DISCUSSION-CONCLUSION

I demonstrated with the Sigsbee2B dataset that multiples can be reliably attenuated without introducing artifacts and/or damaging primaries as long as very accurate models for both noise and signal are provided. Because it is often not possible to obtain such a model for the signal, the Spitz approximation is used. As such, the Spitz approximation yields a very good separation of primaries and multiples with 3D filters. However, analyzing this separation in the image space, we notice that some weak primaries are attenuated.

This illustrates the necessity to evaluate multiple removal techniques on migrated images as much as possible. Different improvements are possible to make the noise removal better; for instance, by changing the trade-off parameter in equation (7). In addition, estimated primaries with 3D filters could be used as a signal model for a new iteration of multiple removal. Another possibility is to migrate the multiple model and apply the PEFs in the image space directly. There the effects of multiple removal could be quantified in a more interactive way.

Therefore, in addition to the fact that the image space should be used as much as possible for multiple removal, for quality control and/or attenuating multiples (Sava and Guitton, 2003), one important lesson learned with this dataset is that finding an accurate model for primaries and multiples before noise removal is crucial. We can relate this to the need for imaging to find the right velocity model; to paraphrase Claerbout (1999), everything depends on it.
ACKNOWLEDGMENTS

I thank SMAART-JV for providing the Sigsbee2B dataset and Paul Sava for providing the migration codes.

REFERENCES


Subtraction of surface-related multiples: adaptive subtraction vs. pattern recognition

Antoine Guitton

ABSTRACT
Two multiple subtraction techniques are tested on a 2D synthetic dataset. The first technique is adaptive subtraction, where the signal is assumed to have minimum energy. The second technique is pattern recognition, where the signal and noise are assumed to have different multivariate spectra. Overall, the pattern based technique leads to a better subtraction of the multiples.

INTRODUCTION
In complex geology, multiples and primaries can be accurately separated by estimating a model of the multiples and subtracting it from the data (Verschuur et al., 1992). The subtraction of a given model is usually made by assuming that the signal (the primaries) has minimum energy. As exemplified by numerous authors (e.g., Spitz (1999); Kelamis et al. (2002); Lu (2003); Guo (2003); Guitton and Verschuur (2004)), the minimum energy assumption might not hold and other norms or subtraction techniques might be needed.

In Guitton (2003a), I demonstrated on a field data example that the pattern recognition technique presented in Guitton (2003b) gives the best multiple attenuation result for complex geology, thus providing an alternative to the popular adaptive subtraction method. In addition, I showed that 3D prediction error filters (estimated in the time, shot, offset domain) were more effective in characterizing the multivariate spectra of the noise and data than a series of 2D filters (time, offset) (Guitton, 2003a, 2004).

In this paper, I investigate both the adaptive subtraction technique with the \( \ell^2 \) norm and the pattern-based method for a 2D synthetic data example. This example is perfectly suited for the adaptive subtraction since all the recorded events are propagating in the inline direction. In this comparison, I show that the pattern-based method performs generally better than the adaptive subtraction technique. However, where the noise and signal have similar patterns, I demonstrate that the pattern-based technique damages some primary reflections. This problem is mainly caused by the Spitz approximation which only provides an accurate signal model where the noise and signal are uncorrelated.

In the next section I briefly present the two subtraction techniques and the parameters

\(^1\text{email: antoine@sep.stanford.edu}\)
used for the modeling of the multiples. Then I show the multiple attenuation results on a 2D synthetic model provided by BP.

**SUBTRACTION OF MULTIPLES**

Both adaptive subtraction and pattern recognition assume that a model is known. For this paper I use the method described by Verschuur et al. (1992) to compute the multiple model. In the sections below I describe how the two methods use the data and the multiple model to produce an estimate of only the signal.

**Adaptive subtraction**

Given a model of the multiple \( M \) and the data \( d \), a bank of non-stationary filters \( f \) is estimated such that

\[
g(f) = \| Mf - d \|^2 + \epsilon^2 \| Rf \|^2
\]

is minimum (Rickett et al., 2001). In equation (1), \( R \) is the Helix derivative (Claerbout, 1998) that smooths the filter coefficients across micro-patches (Crawley, 2000) and \( \epsilon \) a trade-off parameter between data fitting and smoothing. The filters have two dimensions. In the following results, the filters are \( 20 \times 3 \) and the patch size is \( 44 \times 20 \) (the first number corresponds to the time axis and the second number corresponds to the offset axis). Once the filters are estimated, the signal becomes

\[
\hat{s} = Mf - d.
\]

The subtraction is done one shot gather at a time.

**Pattern recognition**

As shown in (Guitton, 2003b), the estimated signal is given by:

\[
\hat{s} = (N'KN + \epsilon^2 S'KS)^{-1} N'KND,
\]

where \( N \) and \( S \) are the noise and signal prediction error filters (PEFs), respectively, \( \epsilon \) a trade-off parameter and \( K \) a masking operator. The noise PEFs are estimated from the noise model. The signal PEFs are estimated with the Spitz approximation (Guitton, 2004). As we shall see later, the Spitz approximation works very well when the noise and signal are uncorrelated. 3D filters are estimated since they lead to the best multiple attenuation results (Guitton, 2003a, 2004).

For the following results, the filters size is \( 15 \times 3 \times 3 \) (the last number corresponds to the shot axis) and the patch size is \( 16 \times 8 \times 5 \). These numbers are identical for both noise and signal filters. With 3D filters, because of memory limitations, we cannot estimate the signal for a complete 2D line on one computer only. Therefore, we segment the 2D line into macro-patches of 50 successive shots. There is an overlap of 5 shots between adjacent macro-patches. Each macro-patch is processed on one node before being merged into the final file.
Modeling of the multiples

In the shot domain, for one frequency, the multiple model is given by the spatial non-stationary convolution of shot gathers (Verschuur et al., 1992; Dragoset and Jericevic, 1998). The synthetic model has an offset spacing of 12.5 m and a shot separation of 50 m. To make the multiple prediction work, the offset axis is sampled down to 50 m. Figure 1 shows one constant offset section from -15,000 m to +15,000 m.

This section of the dataset is particularly interesting because of the diffractions visible throughout. Because no velocity model or sedimentary section is available, a possible interpretation of these diffractions is the presence of salt bodies with a rugose top (similar to what we see with the Sigsbee2B dataset). The multiple model is shown in Figure 2 for the same offset. DT points to diffraction tails where the model is not properly rendering the multiples in the data. Besides these few imperfections in the model, the model looks very faithful to the actual multiples. The pattern recognition technique and the adaptive subtraction are compared in the following section.
Figure 2: Constant offset section (h=550 m) of the estimated multiples. The multiples are accurately modeled except for the diffracted multiples, shown as DT, for which the limited range of offsets and number of shots hamper any attempt at modeling the diffraction tails.

SIGNAL/NOISE SEPARATION RESULTS

The goal of this section is to compare the two subtraction techniques on the synthetic dataset. This dataset was primarily designed to conduct blind tests for velocity estimation methods. Consequently, no structural information is known.

The result of adaptive subtraction is shown for one offset section in Figure 3 and the result of pattern-based subtraction is shown in Figure 4. The adaptive subtraction is doing a decent job everywhere. However, some multiples are still visible. For example, '1' in Figure 3 points to a location where multiples overlap with primaries and are not attenuated. In contrast, the pattern-based subtraction technique (i.e., Figure 4) seems to do a better job attenuating these events. The same is true for arrows '2' and '5'. The diffracted multiples (arrows '3' and '4') are also better attenuated with the pattern-based technique.

Because no velocity analysis was conducted with this dataset, no stacks are presented.
Alternatively, close-ups of constant offset sections are shown to illustrate strengths and weaknesses of the two different approaches. Figure 5 shows a comparison between the input data, the multiple model, the estimated primaries with the adaptive subtraction and the estimated primaries with the pattern recognition technique. The offset is 700 m. As shown by the arrows, the pattern-based method performs generally better. The same conclusions hold in Figure 6. Note in Figure 6b aliasing artifacts due to the coarse sampling of the offset axis for the multiple prediction (van Dedem, 2002).

Sometimes, it can be rather difficult to see if multiples are removed or not by simply looking at 2D planes. Figure 7c shows one event at '2' that seems to be a primary. However, by looking at the shot gathers (not shown here), it appears that this event is a multiple that the pattern-based approach was able to attenuate.

One shortcoming of the pattern-recognition technique is that it relies on the Spitz approximation to provide a signal model if nothing else is available. By construction, the signal and noise filters will span different components of the dataspace. Therefore, the estimated primaries and multiples are uncorrelated. This simple fact proves that with the Spitz approx-
imation, higher dimension filters are preferred because primaries and multiples have fewer chances to look similar.

Figure 8 shows an example where primaries are damaged by the pattern-based method. For instance in Figure 8a, we see at ’2’ a primary that is attenuated by the PEFs (Figure 8d) but well preserved by the adaptive subtraction (Figure 8c). Here the primaries and multiples (Figure 8b) exhibit similar patterns. Using the Spitz approximation, event ’2’ is identified as noise and removed as such. For event ’3’, it is quite difficult to say if multiples are better removed in Figure 8d or if primaries are better preserved in Figure 8c. Looking at the corresponding shot gathers did not help to make a decision because the multiples are very strong. Event ’4’ looks clearly better with the adaptive subtraction and ’1’ and ’5’ are well recovered with the pattern-based approach.
Figure 5: Constant offset sections (h=700 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction and (d), estimated primaries with the pattern-based approach. Arrows point to locations where the pattern-based approach attenuates multiples significantly better than the adaptive subtraction.
Figure 6: Constant offset sections (h=4550 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction and (d), estimated primaries with the pattern-based approach. Arrow A points to aliasing effects due to the offset sampling of the shot gathers. The pattern-based approach attenuates the multiples better than the adaptive subtraction in '1' and '2'. [antoine2-compwin6] [CR,M]
Figure 7: Constant offset sections (h=3300 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction and (d), estimated primaries with the pattern-based approach. '1' points to a primary that the pattern-recognition preserves very well. '2' points to an event that is attenuated with the pattern-based approach but not with the adaptive subtraction in (c). Though not shown here, a close inspection of the corresponding shot gathers suggests that '2' is actually a multiple.
Figure 8: Constant offset sections (h=5050 m) for (a) the input data, (b) the multiple model, (c) the estimated primaries with adaptive subtraction and (d), estimated primaries with the pattern-based approach. ’1’ and ’5’ show events better preserved with the pattern-based method. ’2’ and ’1’ are better recovered with the adaptive subtraction. ’4’ seems to point to a primary that the adaptive subtraction is able to save. ’3’ is undecided.
CONCLUSION

Two methods for subtracting surface-related multiples were presented. One method removes the multiples by adaptive subtraction, assuming that the primaries have minimum energy. One method removes the multiples with a pattern-recognition technique, assuming that the primaries and multiples have different multivariate spectra (patterns). Tests on a 2D synthetic dataset show that the pattern-based technique tends to separate primaries and multiples better than adaptive subtraction. In cases where the primaries and multiples are correlated, however, pattern-recognition can damage primaries. This effect is amplified by the Spitz approximation which prevents the noise and signal PEFs from spanning similar areas of the data space.

ACKNOWLEDGMENTS

I thank BP for providing the synthetic dataset.

REFERENCES


Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: SEP–104.


Attenuation of diffracted multiples with an apex-shifted tangent-squared radon transform in image space

Gabriel Alvarez, Biondo Biondi, and Antoine Guitton

ABSTRACT

We propose to attenuate diffracted multiples with an apex-shifted tangent-squared Radon transform in angle domain common image gathers (ADCIG). Usually, where diffracted multiples are a problem, the wavefield propagation is complex and the moveout of primaries and multiples in data space is irregular. In our method, the complexity of the wavefield is handled by the migration provided reasonably accurate migration velocities are used. As a result, the moveout of the multiples is well behaved in the ADCIGs. For 2D data, our apex-shifted tangent-squared Radon transform maps the 2D image space into a 3D model space cube whose dimensions are depth, curvature and apex-shift distance. Well-corrected primaries map to or near the zero curvature plane and specularly-reflected multiples map to or near the zero apex-shift plane. Diffracted multiples map elsewhere in the cube according to their curvature and apex-shift distance. Thus, specularly reflected as well as diffracted multiples can be attenuated simultaneously. We illustrate our approach with a segment of a 2D seismic line over a large salt body in the Gulf of Mexico. We show that ignoring the apex shift compromises the attenuation of the diffracted multiples, whereas our approach attenuates both the specularly-reflected and the diffracted multiples without compromising the primaries.

INTRODUCTION

Surface-related multiple elimination (SRME) uses the recorded seismic data to predict and iteratively subtract the multiple series (Verschuur et al., 1992). 2D SRME can deal with all kinds of 2D multiples, provided enough data are recorded given the offset limitations of the survey line. Diffracted multiples from scatterers with a cross-line component cannot be predicted by 2D SRME but in principle can be predicted by 3D SRME as long as the acquisition is dense enough in both in-line and cross-line directions. With standard marine streamer acquisition, the sampling in the cross-line direction is too coarse and diffracted multiples need to be removed by other methods (Hargreaves et al., 2003) or the data need to be interpolated and extrapolated to a dense, large aperture grid (van Dedem and Verschuur, 1998; Nekut, 1998; Biersteker, 2001). In general, multiples may not have their moveout apex at zero offset on a CMP gather. Peg-leg multiples “split” into independent events when reflectors dip. These events look similar to diffracted multiples and may similarly hamper standard Radon demul-

1email: gabriel@sep.stanford.edu, biondo@sep.stanford.edu, antoine@sep.stanford.edu
multiple and velocity analysis. Hargreaves et al. (2003) proposed a shifted hyperbola approach to attenuate split or diffracted multiples in CMP gathers. This approach, however, relies on the moveout of the multiples to be well approximated by hyperbolas in data space, which is problematic in complex media. A similar apex-shifted Radon transform was proposed by Trad (2002) for data interpolation.

In most situations in which diffracted multiples are a serious problem, the wave propagation is rather complex, for example for multiples diffracted off the edge of salt bodies. Thus, the moveout of primaries and multiples tend to be very complex, making the application of data-space moveout-based methods for the removal of multiples difficult. In ADCIGs, however, since the complexity of the wavefield has already been taken into account by prestack migration (to the extent that the presence of the multiples allows an accurate enough estimation of the migration velocity field), the residual moveout of multiples is generally smoother and better behaved (Sava and Guitton, 2003).

In this paper we focus on attenuating diffracted multiples in ADCIGs by redefining the tangent-squared Radon transform of Biondi and Symes (2003) to add an extra dimension to account for the shift in the apexes of the moveout curves of the diffracted multiples. We show with a 2D seismic line from the Gulf of Mexico that our approach is effective in attenuating both, the specularly-reflected and the diffracted multiples. In contrast, ignoring the apex shift results in poor attenuation of the diffracted multiples.

The real impact of our method for attenuating diffracted multiples is likely to be in 3D rather than in 2D, though the results that we show in this paper are limited to 2D. Biondi and Tisserant (2003) have presented a method for computing 3D ADCIGs from full 3D prestack migration. These 3D ADCIGs are functions of both the aperture angle and the reflection azimuth. Simple ray tracing modeling shows that out-of-plane multiples map into events with shifted apexes (like the 2D diffracted multiples) and different reflection azimuth than the primaries. Attenuation of these multiples from 3D ADCIGs can be accomplished with a methodology similar to one we present in this paper.

**DIFRACTED MULTIPLES ON ADCIGS**

Figure 1 shows a stack over aperture-angle of a wave-equation migrated 2D line from the Gulf of Mexico over a large salt body. The presence of the salt creates a host of multiples that obscure any genuine subsalt reflection. Most multiples are surface-related peg-legs with a leg related to the water bottom, shallow reflectors or the top of salt. Below the edges of the salt we also encounter diffracted multiples (e.g., CMP position 6000 m below 4000 m depth in Figure 1). Figure 2 shows four ADCIGs obtained with wave-equation migration as described by Sava and Fomel (2003). The top two ADCIGs correspond to lateral positions directly below the left edge of the salt body (CMP positions 6744 m and 22056 m in Figure 1). Notice how the apexes of the diffracted multiples are shifted away from the zero aperture angle (e.g., the “seagull”-looking event at about 4600 m in panel (a)). For comparison, the bottom panels in Figure 2 show two ADCIGs that do not have diffracted multiples. Figure 2c corresponds to an ADCIG below the sedimentary section (CMP 3040 m in Figure 1) and Figure 2d to and
Figure 1: Angle stack of migrated ADCIGs of 2D seismic line in the Gulf of Mexico. Notice that multiples below the salt obscure any primary reflections.

ADCIG below the salt body (CMP position 12000 m in Figure 1). In these ADCIGs all the multiples are specularly-reflected and thus have their apexes at zero aperture angle. Notice also that although the data is marine, the ADCIGs show positive and negative aperture angles. We used reciprocity to simulate negative offsets and interpolation to compute the two shortest-offset traces not present in the original data. The offset gathers were then converted to angle gathers. The purpose of having both positive and negative aperture angles is to see more clearly the position of the apexes of the diffracted multiples.

**APEX-SHIFTED RADON TRANSFORM**

In order to account for the apex-shift of the diffracted multiples \((h)\), we define the forward and adjoint Radon transforms as a modified version of the “tangent-squared” Radon transform introduced by Biondi and Symes (2003). We define the transformation from data space (ADCIGs) to model space (Radon-transformed domain) as:

\[
m(h, q, z') = \sum_{\gamma} d(\gamma, z = z' + q \tan^2(\gamma - h)),
\]

and from model space to data space as

\[
d(\gamma, z) = \sum_{q} \sum_{h} m(h, q, z' = z - q \tan^2(\gamma - h)),
\]
Figure 2: Angle domain common image gathers. (a) under the left edge of the salt, CMP at 6744 m; (b) under the right edge of the salt, CMP at 22056 m; (c) below the sedimentary section, CMP at 3040; (d) below the salt body, CMP at 12000 m.
where \( z \) is depth in the data space, \( \gamma \) is the aperture angle, \( z' \) is the depth in the model space, \( q \) is the moveout curvature and \( h \) is the lateral apex shift. In this way, we transform the two-dimensional data space of ADCIGs, \( d(z, \gamma) \), into a three-dimensional model space, \( m(z', q, h) \).

In the ideal case, primaries would be perfectly horizontal in the ADCIGs and would thus map in the model space to the zero-curvature \((q = 0)\) plane, \( i.e. \), a plane of dimensions depth and apex-shift distance \((h, z')\). Specularly-reflected multiples would map to the zero apex-shift distance \((h = 0)\) plane, \( i.e. \), a plane of dimensions depth and curvature \((q, z')\). Diffracted multiples would map elsewhere in the cube depending on their curvature and apex-shift distance.

**SPARSITY CONSTRAINT**

As a linear transformation, the apex-shifted Radon transform can be represented simply as

\[
d = Lm,
\]

where \( d \) is the image in the angle domain, \( m \) is the image in the Radon domain and \( L \) is the forward apex-shifted Radon transform operator. To find the model \( m \) that best fits the data in a least-squares sense, we minimize the objective function:

\[
f(m) = \|Lm - d\|^2 + \epsilon^2 b^2 \sum_{i=1}^{n} \ln \left( 1 + \frac{m_i^2}{b^2} \right),
\]

where the second term is a Cauchy regularization that enforces sparseness in the model space. Here \( n \) is the size of the model space, and \( \epsilon \) and \( b \) are two constants chosen a-priori: \( \epsilon \) which controls the amount of sparseness in the model space and \( b \) related to the minimum value below which everything in the Radon domain should be zeroed (Sava and Guitton, 2003). The least-squares inverse of \( m \) is

\[
\hat{m} = \left[ L' L + \epsilon^2 \text{diag} \left( \frac{1}{1 + \frac{m_i^2}{b^2}} \right) \right]^{-1} L' d,
\]

where \( \text{diag} \) defines a diagonal operator. Because the model space can be large, we estimate \( m \) iteratively. Notice that the objective function in Equation (2) is non-linear because the model appears in the definition of the regularization term. Therefore, we use a limited-memory quasi-Newton method (Guitton and Symes, 2003) to find the minimum of \( f(m) \).

**A LOOK AT THE 3D RADON DOMAIN**

In this section we will illustrate the mapping of the different events between the image space \((z, \gamma)\) and the 3D Radon space \((z', q, h)\) using the ADCIG at CMP 6744 m (Figure 2a). This ADCIG has no discernible primaries below the salt, but nicely shows the apex-shifted moveout of the diffracted multiples.
Figure 3a shows the $h = 0$ plane from the 3D volume. This plane corresponds to the zero apex-shift and is therefore similar to the standard 2D Radon transform. Primaries are mapped near the $q = 0$ line and specularly-reflected multiples are mapped to other $q$ values. Diffracted multiples, since their moveout apex is not zero, are not mapped to this plane and so do not obscure the mapping of the specularly-reflected multiples. For comparison, Figure 3b shows the 2D Radon transform, plotted at the same clip value as Figure 3a. Notice how the diffracted-multiple energy is mapped as background noise, especially at the largest positive and negative $q$ values. Notice also that the primary energy is higher than in Figure 3a since with the 3D transform the primary energy is mapped not only to the $h = 0$ plane but to other $h$ planes as well. This is further illustrated on Figure 4 which shows the zero-curvature ($q = 0$) plane for the same ADCIG. The nearly flat primaries have zero curvature for all values of the apex-shift distance $h$, so in this plane they appear as horizontal lines. Neither the specularly-reflected nor the diffracted multiples map to this plane.

Figure 3: Radon transforms of the ADCIG in Figure 2b. (a): 2D transform. (b): $h = 0$ plane of the apex-shifted 3D transform.

ATTENUATION OF DIFFRACTED AND SPECULARLY-REFLECTED MULTIPLES

With ideal data, attenuating both specularly-reflected and diffracted multiples could, in principle, be accomplished simply by zeroing out (with a suitable taper) all the $q$-planes except the one corresponding to $q = 0$ in the model cube $m(z', q, h)$ and taking the inverse apex-shifted Radon transform. In practice, however, the primaries may not be well-corrected and primary energy may map to a few other $q$-planes. Energy from the multiples may also map to those planes and so we have the usual trade-off of primary preservation versus multiple attenuation. The advantage now is that the diffracted multiples are well focused to their corresponding
Figure 4: Zero curvature plane from 3D Radon transform of ADCIG in Figure 2b. Flat primaries are mapped to the zero curvature line for all $h$ values.

$h$-planes and can therefore be easily attenuated. Rather than suppressing the multiples in the model domain, we chose to suppress the primaries and inverse transform the multiples to the data space. The primaries were then recovered by subtracting the multiples from the data.

Figure 5 shows a close-up comparison of the primaries extracted with the standard 2D transform (Sava and Guitton, 2003) and with the apex-shifted Radon transform for the two ADCIGs at the top in Figure 2. The standard transform (Figures 5a and 5c) was effective in attenuating the specularly-reflected multiples, but failed at attenuating the diffracted multiples (below 4000 m), which are left as residual multiple energy in the primary data. Again, this is a consequence of the apex shift of these multiples. There appears not to be any subsalt primary in Figures 5a and 5b and only one clearly visible subsalt primary in Figures 5c and 5d (just above 4400 m). This primary was well preserved with both transformations.

Figure 6 shows a similar comparison for the extracted multiples. Notice how the diffracted multiples were correctly identified and extracted by the 3D Radon transform, in particular in Figure 6b. In contrast, the standard 2D transform misrepresent the diffracted multiples as though they are specularly-reflected multiples as seen in Figure 6a. We can take advantage of the 3D model representation to separate the diffracted multiples from the specularly-reflected ones. This is shown in Figure 7. The diffracted multiples are clearly seen in Figure 7c.

In order to assess the effect of better attenuating the diffracted multiples on the angle stack of the ADCIGs we processed a total of 310 ADCIGs corresponding to horizontal positions 3000 m to 11000 m in Figure 1. Figure 8 shows a close-up view of the stack of the primaries extracted with the 2D Radon transform, the stack of the primaries extracted with the 3D Radon transform and their difference. Notice that the diffracted multiple energy below the edge of the salt (5000 m to 7000 m) that appears as highly dipping events with the 2D transform, has been attenuated with the 3D transform. This is shown in detail in Figure 8c. It is very difficult to identify any primary reflections below the edge of the salt, so it is hard to assess if the
Figure 5: Comparison of primaries extracted with the 2D Radon transform (a) and (c) and with the apex-shifted Radon transform (b) and (d). Notice that some of the diffracted multiples remain in the result with the 2D transform.
Figure 6: Comparison of multiples extracted with the 2D Radon transform (a) and (c) and with the apex-shifted Radon transform (b) and (d).
primaries have been equally preserved with both methods. It is known, however, that for this dataset, there are no multiples above a depth of about 3600 m, between CMP positions 3000 m to 5000 m. The fact that the difference panel appears nearly white in that zone shows that the attenuation of the diffracted multiples did not affect the primaries. Of course, this is only true for those primaries that were correctly imaged, so that their moveout in the ADCIGs was nearly flat. Weak subsalt primaries may not have been well-imaged due to inaccuracies in the migration velocity field and may therefore have been attenuated with both the 2D and the 3D Radon transform.

For the sake of completeness, Figure 9 shows the extracted multiples with the 2D and the 3D Radon transform and their difference. Again, the main difference is largely in the diffracted multiples.

**DISCUSSION**

The results shown in the previous section demonstrate that with the 3D Radon transform it is possible to attenuate, although not completely remove, the diffracted multiples. It should be noted, however, that in this seismic section it is very difficult to find a legitimate primary reflection below the salt and in particular below the edge of the salt, where the contamination by the diffracted multiples is stronger. It is somewhat disappointing that the attenuation of the diffracted multiples didn’t help in uncovering any meaningful primary reflections in this case. We expect the situation to be different with other datasets.
Figure 8: Comparison of angle stacks for primaries.

(a): Angle stack of Primitives with 2D RT

(b): Angle stack of Primitives with 3D RT

(c): Difference between angle stacks of primaries
Figure 9: Comparison of angle stacks for multiples.
We mentioned before that working in image space (ADCIGs in this case) is convenient because the migration takes care of the complexity of the wavefield propagation, but attenuating the multiples after migration does not come without a price. The estimation of the migration velocities may be more difficult and less accurate because of the presence of the multiples. There is therefore an inherent trade-off when choosing to work in image space. Good migration velocities for weak subsalt primaries may be particularly difficult to estimate in the presence of the multiples. On the other hand, the parabolic or hyperbolic assumption for the moveout of the multiples in data space may not be appropriate at all in complex media. An alternative could be to do a standard Radon demultiple before prestack migration to facilitate the choice of the migration velocities and 3D Radon demultiple on the ADCIGs to attenuate residual multiples, in particular diffracted multiples.

We should also emphasize that adding the extra dimension to deal with the diffracted multiples does not in itself resolve the usual problem that non-flat primaries may map to the multiple region and therefore we have to trade primary preservation for multiple attenuation. We saw this limitation in this case, which forced us to let some residual multiple energy leak into the extracted primaries. Obviously, the flatter the primaries in the ADCIGs, the better our chances to reduce the residual multiple energy.

**CONCLUSION**

The combination of choosing the image space in the form of ADCIGs and the apex-shifted tangent-squared transformation from \((z, \gamma)\) to \((z', q, h)\) has proven to be effective in attenuating specularly-reflected and diffracted multiples in 2D marine data. The residual moveout of both multiples in ADCIGs is well-behaved and the extra dimension provided by the apex-shift allows the attenuation of the multiples without compromising the integrity of the primaries.

**ACKNOWLEDGMENTS**

We thank the sponsors of the Stanford Exploration Project for their support to carry out this study and WesternGeco for providing the dataset.

**REFERENCES**


Migration of surface-related multiples: tests on the Sigsbee2B dataset

Guojian Shan and Antoine Guitton

ABSTRACT
We present a theory to generate pseudo-primary shot gathers from multiple and primary reflections by performing a surface-consistent cross-correlation. The estimated pseudo-primaries exhibit the same kinematics as the original dataset with few transformation artifacts. We demonstrate that pseudo-primaries can accurately estimate missing traces as long as the gaps are within the acquisition spread. Pseudo-primaries can also help to extrapolate the data outside the acquisition spread. The image obtained by migrating the pseudo-primary gathers shows that multiple migration can provide valuable information under complex geology.

INTRODUCTION
When only primary reflections are considered for imaging, multiple reflections are usually attenuated as a preprocessing step (Verschuur et al., 1992; Weglein et al., 1997; Guitton, 2003; Sava and Guitton, 2003). However, multiples contain subsurface reflectivity information, and can be treated as signal. For instance, Brown (2004) shows how a joint inversion of both primaries and multiples can provide more knowledge of the earth’s properties.

Multiples can be imaged by Kirchhoff (Reiter et al., 1991) or crosscorrelogram migration (Sheng, 2001), by transforming the traveltimes of multiple reflections to those of primary reflections. Multiples can be also imaged by shot-profile migration, considering the primary reflections as areal shot records and the multiple reflections as receiver wavefields (Berkhout and Verschuur, 1994; Guitton, 2002). Instead of being transformed into primaries implicitly, multiples also can be explicitly mapped into primaries by cross-correlation (Shan, 2003; Berkhout and Verschuur, 2003) or deconvolution (Shan, 2003). We call “pseudo-primary” any events resulting from the cross-correlation of multiples with the original dataset (primaries+multiples). The pseudo-primaries are similar to the original data and can be imaged by source-receiver migration (Shan, 2003).

In this paper, we estimate pseudo-primary shot gathers by cross-correlating primary and multiple reflections. These pseudo-primary shot gathers are then migrated with shot-profile migration. The images obtained from the pseudo-primaries are then compared to images of the primaries alone.

1email: shan@sep.stanford.edu, antoine@sep.stanford.edu
We applied our method to the Sigsbee2B synthetic dataset. This dataset is challenging because of the complex geometry of the salt body (Figure 1). Two versions of the Sigsbee2B dataset were generated: one with surface-related multiples (FS) and one without (NFS). The multiples are known and can be obtained by subtracting the two datasets (i.e. FS-NFS). But with field data, surface-related multiples need to be separated prior to the migration using the method of Verschuur et al. (1992). From this dataset, we show that multiples can (1) fill acquisition holes, (2) extrapolate data beyond the acquisition spread, and (3) provide an image of the subsurface under complex geology.

![Figure 1: Stratigraphic interval velocity model of the Sigsbee2B dataset](image)

**PSEUDO-PRIMARY FROM MULTIPLES**

We can generate pseudo-primary shot gathers $W$ by computing:

$$W(x_p, x_m, \omega) = \sum_{x_s} M(x_s, x_m, \omega) \overline{P}(x_s, x_p, \omega),$$

where $\omega$ is the frequency, $x_s$ is the shot location, $\overline{P}(x_s, x_p, \omega)$ is the complex conjugate of the original trace recorded at the surface location $x_p$, and $M(x_s, x_m, \omega)$ is the multiple reflection data recorded at the surface location $x_m$. In equation (1), note that first-order multiples in $M$
are transformed into primaries (cross-correlation with primaries in $P$) and zero-lag components (cross-correlation with first-order multiples in $P$) in the pseudo-primaries $W$. Similarly, second-order multiples in $M$ are transformed into primaries (cross-correlation with first-order multiple in $P$), first-order multiples (cross-correlation with the primaries in $P$) and zero-lag components (cross-correlation with second-order multiples in $P$) in the pseudo-primaries $W$, and so on.

In contrast to primaries in the original dataset, pseudo-primaries can illuminate different areas at different angles. They contain subsurface information that primaries do not. Because of recording geometries, it is often difficult to obtain near-offset data, as well as far-offset data. Pseudo-primaries created from multiples can help fill these acquisition holes.

Figure 2: Near-offset recovery example: In (a), no trace is recorded at $x_m$ for the shot at $x_p$. In (b), a primary reflection is recorded at $x_p$ and a multiple reflection is recorded at $x_m$ for the shot at $x_s$. With the pseudo-primaries, we recover the trace with a source at $x_p$ and receiver at $x_m$. [NR]

Figure 3: Far-offset recovery example: In (a), no trace is recorded at $x_m$, since it is outside the acquisition spread for the shot at $x_p$. In (b), both $x_p$ and $x_m$ are within the acquisition spread for the shot at $x_s$. A primary reflection is recorded at $x_p$ and a multiple reflection is recorded at $x_m$. With the pseudo-primaries, we recover the trace with a source at $x_p$ and receiver at $x_m$. [NR]
Figure 2 illustrates how missing near-offset traces are recovered in the pseudo-primary dataset. In Figure 2a, near offsets are missing when the source is at $x_p$. In Figure 2b, the primary reflection is recorded at $x_p$ and the multiple reflection is recorded at $x_m$ for the shot at $x_s$. By cross-correlating the traces at $x_p$ and $x_m$ in Figure 2b, we obtain the pseudo-primary trace, whose shot and receiver locations are $x_p$ and $x_m$, respectively. So the near-offset missing trace is recovered in the pseudo-primaries.

Figure 3 illustrates how missing far-offsets are recovered by the pseudo-primary dataset in some special cases. Using reciprocity, we can obtain the negative-offset data by mirroring sources and receivers. In Figure 3a, the primary reflection with a source at $x_p$ and receiver at $x_m$ is outside the acquisition spread. In Figure 3b, a primary reflection is recorded at $x_p$ and a multiple reflection is recorded at $x_m$ for the shot at $x_s$. Both $x_p$ and $x_m$ are within the acquisition spread. By cross-correlating the traces at $x_p$ and $x_m$ in Figure 3b, we obtain a trace of the pseudo-primaries, whose shot and receiver locations are $x_p$ and $x_m$, respectively. So the far-offset missing trace is recovered with pseudo-primaries. Note that to recover the far-offset trace, as is illustrated in Figure 3, we need a steeply dipping reflector.

We now illustrate our technique on the Sigsbee2B dataset. Figure 4 shows four shot gathers with a source at 50,000 ft: (a) the original dataset (primaries + multiples) with near and far offsets removed, (b) the surface-related multiples with near and far offsets removed, (c) the original dataset (primaries + multiples) with full offsets, and (d) the pseudo-primary dataset. We mirrored the sources and receivers to get negative offsets. To demonstrate that pseudo-primaries can interpolate data inside the acquisition holes, we removed the offsets that are less than 2,000 ft and greater than 20,000 ft from the original dataset and the surface-related multiples, which are illustrated in Figure 4a and 4b. Figure 4d shows the pseudo-primary shot gather obtained by cross-correlating the original dataset with the multiple dataset without near and far offsets. Comparing the original shot gather in Figure 4c (no mute) with the pseudo-primary shot gather in Figure 4d, we conclude that the pseudo-primary shot gather is very similar to the original shot gather. Note that some artifacts and noisy events appear in the pseudo-primary gather. The noise arises during the cross-correlation of unpaired events at different surface locations. It might be better handled by the method of Berkhout and Verschuur (2003). The near-offset data are recovered very well in the pseudo-primary gather. For far offsets, notice that the event in the oval of Figure 4c, which is removed in Figure 4a, is recovered in Figure 4d. We think that this far-offset event is recovered in the pseudo-primary gather due to the presence of two large canyons with steeply dipping walls on the salt body (Figure 1), similar to the structure illustrated in Figure 3. Note that more shots would help to recover more events at far offsets in Figure 4d.

Figure 5 compares the original zero-offset dataset with the pseudo-primary zero-offset dataset. Same as Figure 4, the pseudo-primary dataset is generated by cross-correlating the original dataset with the multiple dataset without near and far offsets. These two zero-offset datasets have similar structures.
Figure 4: Comparison of shot gathers at 50,000 ft for (a) the original dataset (primaries + multiples) with near and far offsets removed, (b) surface-related multiples with near and far offsets removed, (c) the original dataset (primaries + multiples) with full offsets, and (d) the pseudo-primary, generated by cross-correlating the original dataset with the multiple dataset without near and far offsets.
Figure 5: Comparison of zero-offset datasets for (a) the original data (primaries+multiples) and (b) the pseudo-primary, generated by cross-correlating the original dataset with the multiple dataset without near and far offsets.
MIGRATION OF THE PSEUDO-PRIMARY DATASET

As illustrated in Figure 4, shot gathers of the pseudo-primaries are similar to those of the original dataset. We now migrate the pseudo-primary dataset by shot-profile, downward-continuation migration. We use Fourier finite difference (Ristow and Ruhl, 1994) as our wavefield extrapolation operator for the migration. Figure 6 compares the image from the primary-only dataset (NFS), the image from the original dataset (FS), and the image from the pseudo-primaries, which consists of primaries (from first-order multiples) and multiples (from higher-order multiples). Figure 6a shows the migration result of primaries, Figure 6b shows the migration result of the original dataset and Figure 6c shows the migration result of pseudo-primaries. The pseudo-primary image is noisier than the images from the primaries only and the original dataset. This is caused by the noise from the cross-correlation and first-order multiples in the pseudo-primary gathers. In the image obtained by migrating the pseudo-primaries (Figure 6c), the salt body is clearly imaged, and reflectors below the salt are also well imaged. However the reflectors below 20,000 ft are contaminated by first-order multiples in the pseudo-primaries, which are similar to those in the original-dataset image (Figure 6b).

CONCLUSIONS

In this paper, we show how multiples can be transformed into primaries (i.e., pseudo-primaries) by being cross-correlated with the original dataset (primaries+multiples). Pseudo-primaries can recover missing near-offset data as well as some far-offset events that are not recorded. The comparison between images from the primary dataset and from the pseudo-primary dataset on the Sigsbee2B model demonstrates that multiple migration can be used to image complex geological structures.

ACKNOWLEDGMENTS

We would like to thank SMAART-JV for providing the Sigsbee2B dataset.

REFERENCES


Figure 6: Comparison of migration results for (a) the primaries only, (b) the original dataset (primaries+multiples) and (c) the pseudo-primary dataset.


3-D Angle-domain common-image gathers for migration velocity analysis

Biondo Biondi and Thomas Tisserant

ABSTRACT

Angle-Domain Common Image Gathers (ADCIGs) are an essential tool for Migration Velocity Analysis (MVA). We present a method for computing ADCIGs in 3-D from the results of wavefield-continuation migration. The proposed methodology can be applied before or after the imaging step in a migration procedure. When computed before imaging, 3-D ADCIGs are functions of the offset ray parameters \((p_{xh}, p_{yh})\); we derive the geometric relationship that links the offset ray parameters to the aperture angle \(\gamma\) and the reflection azimuth \(\phi\). When computed after imaging, 3-D ADCIGs are directly produced as functions of \(\gamma\) and \(\phi\).

The mapping of the offset ray parameters \((p_{xh}, p_{yh})\) into the angles \((\gamma, \phi)\) depends on both the local dips and the local interval velocity; therefore, the transformation of ADCIGs computed before imaging into ADCIGs that are function of the actual angles is difficult in complex structure. In contrast, the computation of ADCIGs after imaging is efficient and accurate even in presence of complex structure and a heterogeneous velocity function. On the other hand, the estimation of the offset ray parameters \((p_{xh}, p_{yh})\) is less sensitive to velocity errors than the estimation of the angles \((\gamma, \phi)\). When ADCIGs that are functions of the offset ray parameters \((p_{xh}, p_{yh})\) are adequate for the application of interest (e.g. ray-based tomography), the computation of ADCIGs before imaging might be preferable.

Errors in the migration velocity cause the image point in the angle domain to shift along the normal to the apparent geological dip. By assuming stationary rays (i.e. small velocity errors), we derive a quantitative relationship between this normal shift and the traveltime perturbation caused by velocity errors. This relationship can be directly used in a MVA procedure to invert depth errors measured from ADCIGs into migration velocity updates. In this paper, we use it to derive an approximate 3-D Residual Moveout (RMO) function for measuring inconsistencies between the migrated images at different \(\gamma\) and \(\phi\). We tested the accuracy of our kinematic analysis on a 3-D synthetic data set with steeply dipping reflectors and a vertically varying propagation velocity. The tests confirm the accuracy of our analysis and illustrate the limitations of the straight-rays approximation underlying our derivation of the 3-D RMO function.

\[1\text{email: biondo@sep.stanford.edu}\]
INTRODUCTION

Wavefield-continuation migration methods have the potential of producing high-quality images even when complex overburden severely distorts the wavefield. However, as for all migration methods, the quality of the final image is strongly dependent on the accuracy of the velocity model. In complex area the velocity model is usually estimated in an iterative process called Migration Velocity Analysis (MVA). At each iteration of an MVA process the velocity is updated based on the information extracted from the current migrated image, and, in particular, from the Common Image Gathers (CIGs). The computation of accurate CIG is thus crucial for any MVA method. Most of the current MVA methods for wavefield-continuation migration employ Angle-Domain CIGs (ADCIGs) (Biondi and Sava, 1999; Clapp and Biondi, 2000; Mosher et al., 2001; Liu et al., 2001).

The computation of ADCIGs with wavefield-continuation migration is based on a plane-waves decomposition of the wavefield either before imaging (de Bruin et al., 1990; Mosher et al., 1997; Prucha et al., 1999; Xie and Wu, 2002), or after imaging (Sava and Fomel, 2003; Rickett and Sava, 2002; Biondi and Shan, 2002). The methods previously presented in the literature are limited to the computation of 2-D ADCIGs that are functions of the aperture angle $\gamma$ only. These methods have been applied to 3-D marine data by assuming zero-azimuth reflections; this assumption is approximately correct for marine data that have been acquired with a narrow-azimuth acquisition geometry.

In this paper we present the computation and the geometric interpretation of full 3-D ADCIGs that decompose the image not only according to the aperture angle $\gamma$, but also to the reflection azimuth $\phi$. We extend to 3-D both methods for computing ADCIGs: before and after imaging. Our analysis shows that the 2-D equation that relates the in-line offset ray parameter $p_{xh}$ to the aperture angle $\gamma$ is also valid in 3-D for zero-azimuth reflections, when used in combination with common-azimuth migration (Biondi and Palacharla, 1996). This result supports the previous use of 2-D ADCIGs computed before imaging for MVA of 3-D marine data (Biondi and Vaillant, 2000; Clapp, 2001; Mosher et al., 2001; Liu et al., 2001). In contrast, the 2-D transformation to angle domain performed after imaging (Sava and Fomel, 2003) is not correct in 3-D, not even in the case of zero-azimuth reflections; its use in presence of cross-line dips leads to the overestimation of the reflection-aperture angle $\gamma$. In either case (before or after imaging), when the azimuth of the reflections is not oriented along the in-line direction, we must use the full 3-D methodology to obtain accurate ADCIGs.

The geometrical understanding we developed when generalizing 2-D ADCIGs to 3-D ADCIGs enables us to generalize the analysis of the kinematics of 2-D ADCIGs in the presence of migration-velocity errors that was presented by Biondi and Symes (2003). However their purely ray-theoretical analysis of 2-D ADCIGs cannot be directly extended because in 3-D the source and receiver rays are not guaranteed to be coplanar. Fortunately, a plane-wave interpretation of 3-D ADCIGs overcomes these difficulties because the two plane waves corresponding to the source and receiver rays define a plane of coplanarity that passes through the angle-domain image point. Once this plane is defined, the 2-D kinematic analysis carries over to 3-D and we can define a quantitative relationship between the reflectors’ movement along the apparent geological dip and the traveltime perturbations caused by velocity errors.
This relationship can be directly used in a tomographic inversion of 3-D ADCIGs; we use it to define a 3-D RMO function for measuring kinematic errors from 3-D ADCIGs.

ADCIGs have been introduced also for integral migration methods (e.g. Kirchhoff and Generalized Radon Transform); Xu et al. (2001) defined and applied them in 2-D, Brandsberg-Dahl et al. (2003) defined them in 3-D and applied them in 2.5-D, and Brandsberg-Dahl et al. (1999) applied 2-D ADCIGs to MVA. ADCIGs computed by Kirchhoff-like migration share many properties with ADCIGs computed by wavefield-continuation migration. However, in complex media, the two types of ADCIG have subtle kinematic differences, as clearly demonstrated by Stolk and Symes (2004). A thorough comparison of these two types of ADCIG would be of theoretical and practical value, but we consider it beyond the scope of this paper.

ANGLE-DOMAIN COMMON IMAGE GATHERS BEFORE IMAGING

This paper presents two methods for computing 3-D ADCIGs in conjunction with wavefield-continuation migration. Both methods rely on a decomposition of the wavefield into plane-wave components; this decomposition is performed by slant stacks (Schultz and Claerbout, 1978) along the subsurface offset axes. The two methods differ in the stage of the migration process during which the plane-wave decomposition is performed (before imaging or after imaging), but in both cases the plane-wave decomposition links the ray-theoretical and the plane-wave interpretation of ADCIGs that we use in this paper. We equate the propagation direction of plane waves with the propagation direction of specular rays, and thus we implicitly make a stationary-phase approximation that assumes that the seismic events are not dispersive; that is, that wavefronts are locally planar and coherent for all frequencies. This assumption is not strictly necessary because our analysis is valid for each frequency component, but nonetheless we believe that the advantages in simplicity are worthwhile the apparent loss of generality.

We start by presenting the method based on an angle decomposition before imaging (de Bruin et al., 1990; Prucha et al., 1999) because it has the most direct physical interpretation. The physical interpretation of the ADCIGs computed after imaging (Sava and Fomel, 2003; Rickett and Sava, 2002; Biondi and Shan, 2002) follows directly from the interpretation of the ADCIGs computed before imaging.

A 3-D prestack data set can be decomposed according to the offset ray parameters \((p_{xh}, p_{yh})\) by applying slant stacks along the offset axes. These slant stacks produce plane-wave components called offset plane waves. If the velocity function is horizontally invariant, each offset plane wave can be accurately propagated and imaged independently from the others. Each prestack partial image produced by this procedure corresponds to one offset ray parameter, and thus ADCIG functions of the offset ray parameters \([\text{ADCIG}(p_{xh}, p_{yh})]\) can be extracted from the whole prestack image cube. The offset ray parameters measured at the surface are functions of the surface take-off angles of the source and receiver rays, and are obviously linked with the reflection opening angle \(\gamma\) and the reflection azimuth \(\phi\).

Ottolini and Claerbout (1984) presented a procedure for obtaining angle-dependent images
based on the slant-stack of 2-D CMP gathers at the surface and on the propagation of offset plane waves. Mosher et al. (1997) generalized the 2-D procedure to 3-D marine data and derived a more accurate downward-continuation method for the offset plane waves. However, there is a limit on the accuracy of a migration based on an offset plane-wave decomposition at the surface. When the velocity function is laterally varying, the source and receiver wavefields are refracted differently, according to the local velocities. Consequently, the independent propagation of each offset plane wave is inaccurate when strong lateral velocity variations occur. In these situations, the idea of decomposing the wavefield into offset plane waves is still useful, but it must be generalized; the prestack wavefield must be slant-stacked recurrently at each depth level, instead of only once at the surface.

When source-receiver migration (i.e. survey sinking) is applied, the whole prestack wavefield is obtained at every depth level as the result of recursively downward continuing the recorded data using the Double Square Root (DSR) operator (Claerbout, 1985). Therefore, once we have introduced the concept of decomposing the prestack wavefield at depth, the derivation of the procedure for computing ADCIGs before imaging using source-receiver downward continuation is straightforward (Prucha et al., 1999).

In conventional source-receiver migration the image is then obtained by extracting the values of the wavefield at zero time \( t = 0 \). The migration process can be schematically represented as follows:

\[
P (\omega, m, h; z = 0) \xrightarrow{\text{DSR}} P (\omega, m, h; z) \quad (1)
\]

\[
P (\omega, m, h; z) \xrightarrow{\text{Imaging}} P (t = 0, m, h; z), \quad (2)
\]

where \( P (\omega, m, h; z) \) is the wavefield at the temporal frequency \( \omega \), the midpoint \( m \), the offset \( h \), and at depth \( z \).

To compute ADCIGs, we need to introduce a slant-stack transformation between downward continuation [step (1)] and imaging [step (2)]. The results of slant stacks are functions of the offset ray parameter \( p_h \) and the intercept time \( \tau \). The image is then obtained by extracting the values of the wavefield at zero intercept time; i.e. \( \tau = 0 \). The migration process used to produce ADCIGs can be schematically represented as follows:

\[
P (\omega, m, h; z = 0) \xrightarrow{\text{DSR}} P (\omega, m, h; z) \quad (3)
\]

\[
P (\omega, m, h; z) \xrightarrow{\text{Slant stack}} P (\tau, m, p_h; z) \quad (4)
\]

\[
P (\tau, m, p_h; z) \xrightarrow{\text{Imaging}} P (\tau = 0, m, p_h; z). \quad (5)
\]

Figure 1 illustrates the subsequent stages of the process represented in (3–5). The synthetic data set imaged to create this figure was modeled assuming two reflectors (one dipping and one flat) immersed in a constant-velocity medium. The downward-continuation process [step (1) or step (3)] focuses the wavefield toward zero offset; if the continuation velocity is correct, a migrated image can be obtained by extracting the value of the wavefield at zero offset and zero time. Figure 1a shows the zero-offset section of the downward-continued wavefield at the depth of the flat reflectors; that is, at \( z = 700 \) meters. As expected, the flat reflector is well imaged at zero time. Figure 1b shows the wavefield at the same depth as in Figure 1a, but
Figure 1: Illustration of the method for computing ADCIGs before imaging. Orthogonal slices of the prestack wavefield after downward continuation with the correct velocity to the depth of 700 meters; that is, the depth of the flat reflector: zero-offset section (panel a), and the common-midpoint gather at 1,410 meters (panel b). Panel c) shows the complete (i.e. for all depths) ADCIG at 1,410 meters. 

as a function of offset and at fixed midpoint ($x_m = 1,410$ meters). The flat reflector is well focused at zero offset and zero time, whereas the energy corresponding to the dipping reflector has already been defocused and thus it appears as a “time-reversed” hyperbola at negative times. The wavefield shown in Figure 1b is decomposed into plane waves by performing slant stacks for each midpoint location [step (4)], and the results of these slant stacks are imaged at zero time [step (5)]. The downward-continuation, slant stacks, and imaging steps are repeated for all depth levels. Figure 1c shows the ADCIG at $x_m = 1,410$ meters, for all depths. Both reflectors are imaged as flat events in the ADCIGs, because at each depth, slant stacks transform band-limited impulses located at zero offset (like the one shown in Figure 1b) into flat events.

Computing ADCIGs before imaging is more difficult when using shot-profile migration than when using source-receiver migration, because the prestack wavefield at depth is not easily available as a function of the subsurface offset. In theory, it could be computed by appropriate correlations (or deconvolutions) of the source and receiver wavefields (de Bruin, 1992). This procedure generates ADCIGs equivalent to those obtained by source-receiver migration, as it can be demonstrated by a generalization of the equivalence principle for the prestack images demonstrated by Wapenaar and Berkhout (1987) and Biondi (2003). The computation of the ADCIGs before imaging for shot-profile migration is computationally demanding, because it requires either decomposing the wavefield for each shot independently and accumulating the prestack images for all shots (computationally inefficient), or, accumulating the contributions of each shot to the whole prestack wavefield before decomposing it (storage and I/O inefficient). Therefore, ADCIGs are rarely computed before imaging when using shot-profile migration. In the next section we will see how computing ADCIGs after imaging addresses this problem.
Figure 2: A schematic of the geometry of an ADCIG gather in 2-D. Depending on the context, the angles can be either the angles formed by the propagation direction of the rays, or those formed by the propagation direction of the associated plane waves. The arrows indicate positive angles; that is, $\beta_s$, $\beta_r$, and $\alpha_x$ are negative and $\gamma$ is positive. This sign convention is consistent with upward propagating rays (plane waves).

**Geometric interpretation of 3-D ADCIG** \((p_{xh}, p_{yh})\)

Regardless of whether they have been computed using source-receiver migration or shot-proile migration, ADCIGs computed before imaging have the same geometrical interpretation. Strictly speaking, they are functions of the offset ray parameters \((p_{xh}, p_{yh})\) and not of the reflection opening angle $\gamma$ and the reflection azimuth $\phi$. However, the offset ray parameters \((p_{xh}, p_{yh})\) are easily related to $\gamma$ and $\phi$ by simple trigonometric relationships. We now derive these relationships, starting from the simpler 2-D case. The understanding gained from the 2-D analysis will help the analysis of the more complex 3-D case.

The schematic in Figure 2 defines the angles relevant to 2-D ADCIGs: the reflection opening angle $\gamma$ and the geological dip angle $\alpha_x$, which is defined as the angle formed by the vertical direction and the normal to the apparent geological dip $\mathbf{n}$, both oriented in the upward direction. The box in Figure 2 signifies that for the purpose of our discussion the geometric relationships are local around the imaging point $\mathbf{i}$.

The opening angle and the dip angle are related to the source and receiver rays (plane
waves) propagation angles \((\beta_s, \beta_r)\) by the following simple relationships:

\[
\gamma = \frac{\beta_r - \beta_s}{2}, \quad \text{and} \quad \alpha_s = \frac{\beta_s + \beta_r}{2}.
\] (6)

The source ray parameter \(p_{xs}\) and the receiver ray parameter \(p_{xs}\) are related to the midpoint ray parameters \(p_{xm}\) and offset ray parameter \(p_{xh}\) by the following relationships:

\[
p_{xm} = p_{xs} + p_{xh}, \quad \text{and} \quad p_{xh} = p_{xs} - p_{xh}.
\] (7)

In 2-D the source and receiver ray parameters \((p_{xs}, p_{xs})\) are functions of the source and receiver ray propagation angles \((\beta_s, \beta_r)\) and of the velocity at the reflector \(v(z, x)\) as follows:

\[
p_{xs} = \frac{\sin \beta_s}{v(z, x)} \quad \text{and} \quad p_{xs} = \frac{\sin \beta_r}{v(z, x)}.
\] (8)

Substituting the relationships in (8) into equation (7), then using the angular relationship expressed in equation (6) and applying fundamental trigonometric identities, we obtain the desired relationship:

\[
p_{xh} = \frac{\sin \beta_r - \sin \beta_s}{v(z, x)} = \frac{2\sin \left(\frac{\beta_r - \beta_s}{2}\right) \cos \left(\frac{\beta_r + \beta_s}{2}\right)}{v(z, x)} = \frac{2\sin \gamma \cos \alpha_s}{v(z, x)},
\] (9)

which directly links the offset ray parameter \(p_{xh}\) to the reflection opening angle \(\gamma\).

The relationship in equation (9) is also valid in 3-D when the reflection azimuth \(\phi\) is equal to zero. But in the general 3-D case we need to add another relationship linking the cross-line offset ray parameter \(p_{yh}\) to \(p_{xh}\) and the angles \(\gamma\) and \(\phi\) at the reflection point. To derive this relationship, we impose the constraint that the source and the receiver rays must become coplanar before meeting at the reflection point. We will thus refer to this relationship as the coplanarity condition.

The coplanarity of the two rays also simplifies the 3-D geometry of ADCIGs and enables us to generalize the 2-D geometric interpretation in a straightforward manner. The schematic in Figure 3 is the 3-D generalization of the 2-D schematic shown in Figure 2. The plane that is shared by the two rays in 3-D corresponds to the vertical plane in 2-D. As for Figure 2, all the geometric relationships represented in Figure 3 are local in a neighborhood of the image point \(\bar{I}\). This locality implies that the source and receiver rays (plane waves) need to satisfy the coplanarity condition only at the image point, not during propagation in the overburden.

We define the reflection azimuth angle \(\phi\) as the angle formed by the in-line axis \(x\) with the line defined by the intersection of the plane of coplanarity with a constant depth plane. After rotation by \(\phi\), the horizontal coordinates \(x\) and \(y\) become \(x'\) and \(y'\), respectively. Once rotated by \(\phi\), the plane of coplanarity is tilted with respect to the vertical by the cross-line dip angle \(\alpha_y\). All the angles in Figure 2 that are formed between the ray (plane wave) directions and the true vertical axis \(z\) \((\beta_s', \beta_r', \text{and} \alpha_y)\) are now relative to the tilted vertical axis \(z'\); these angles have the same meaning on the tilted plane as they have on the vertical plane in 2-D. Furthermore, the horizontal ray parameters \((p_{xs}, p_{xg})\) are not affected by the tilt, since they
Figure 3: A schematic of the geometry of an ADCIG gather in 3-D. The geometry is analogous to the 2-D case illustrated in Figure 2. In contrast with the schematic of Figure 2, the plane of coplanarity is not vertical but it is tilted by $\alpha_{y'}$ and rotated by $\phi$ with respect to the horizontal coordinates. Notice also that the angles $\delta_{y'}$ and $\delta_{r'}$ formed by the rays (plane waves) propagation directions with the “true” vertical axis are different from the angles $\beta_{y'}$ and $\beta_{r'}$ formed by the rays with the tilted vertical axis.

This relationship was derived by Biondi and Palacharla (1996) to define common-azimuth migration, and it is demonstrated in the Appendix of their paper.
So far we have interpreted the coplanarity condition in terms of source and receiver rays, but it has a similar interpretation in terms of plane waves. The phase vectors of two arbitrary plane waves are coplanar by construction, since they meet at the origin. The tilt angle $\alpha_y$ and the azimuth $\phi$ of the plane of coplanarity for the incident and the reflected plane waves are implicitly determined by the coplanarity condition expressed in equation (10). The cross-correlation of the two plane waves does not define an imaging point but an imaging line that is orthogonal to the plane of coplanarity. The tilted plane shown in Figure 3 is an arbitrary plane parallel to the plane of coplanarity and the imaging point $\bar{I}$ is the intersection of the imaging line with this plane. The source and receiver rays are defined as the lines parallel to the phase vectors of the corresponding plane waves and passing through the imaging point $\bar{I}$.

Using the relationships among the in-line ray parameters expressed in equations (7) (and their analogues along the cross-line direction $y$), and equations (8), and after applying trigonometric identities, we can rewrite equation (10) as a function of the local velocity $v(z', x', y')$ and the angles $\alpha_x, \alpha_y, \gamma$ and $\phi$ as follows:

$$p_{y'h} = \frac{\sin \alpha_y (\cos \beta_x - \cos \beta_y \cos \phi \tan \beta_x)}{v(z', x', y')(\cos \beta_x + \cos \beta_y \cos \phi \tan \beta_x)} = \frac{-\sin \alpha_y \tan \gamma \tan \alpha_y}{v(z', x', y')}.$$  (11)

The in-line offset ray parameter from equation (9) can be rotated by the azimuth $\phi$ to give the following expression:

$$p_{x'h} = \frac{2 \sin \gamma \cos \alpha_y}{v(z', x', y')}.$$  (12)

which together with equation (11), can be used to map the image from the offset ray parameter domain $(p_{x'h}, p_{y'h})$ into the angle domain $(\gamma', \phi)$.

This mapping of the ray parameters into angles depends on both the local velocity and the geological dips. When the spatial velocity variations are significant and the geological structure is complex, unraveling this dependence might be a challenge because of the well known difficulties in reliably measuring rapidly changing local dips. For some applications this dependency does not cause difficulties. For example, in a ray-based MVA (Clapp and Biondi, 2000), there is no need to perform the mapping from ray parameters to angles, since ray parameters are adequate initial conditions for ray tracing. For other applications, the reflection angles are needed, and thus the mapping must be performed. When this mapping is difficult, it might be convenient to use the ADCIGs computed after imaging, which are presented in the next section. This kind of ADCIG circumvents the need to estimate local dips by transforming the image directly into angles.

ANGLE-DOMAIN COMMON IMAGE GATHERS AFTER IMAGING – ADCIG($\gamma', \phi$)

The computation of ADCIGs before imaging has a straightforward physical interpretation, and thus it was the first methodology to be developed for both and shot-profile migration (de Bruin et al., 1990) and source-receiver migration (Prucha et al., 1999). However, as we have discussed in the previous section, the computation of ADCIGs before imaging with shot-profile
migration can be either computationally or I/O demanding. Sava and Fomel (2003) recognized that for source-receiver migration, the plane-wave decomposition can be performed after applying the imaging condition, as well as before. Rickett and Sava (2002) applied this insight to compute 2-D ADCIGs after shot-prole downward continuation. Their method is based on the generalization of the conventional imaging condition for shot-profile migration that we summarize in the next paragraph. When using this imaging condition we compute a prestack image function of the subsurface half offset $h$, as well as of depth $z$ and the midpoint $m$.

If $P^s_z(\omega, x, y; s_i)$ and $P^g_z(\omega, x, y; s_i)$ are respectively the source wavefield and the recorded wavefield downward-continued to depth $z$ for the $i$-th source location $s_i$, as functions of the subsurface horizontal coordinates $(x, y)$, the image is formed by first cross-correlating the two wavefields along the time axis (multiplication by the complex conjugate in the frequency domain) and then evaluating the correlation at zero time (summation over frequencies) as following:

$$I(m, h; z) = \sum_i \sum_\omega P^g_z(\omega, x + x_h, y + y_h; s_i) P^s_z(\omega, x - x_h, y - y_h; s_i).$$  (13)

The physical interpretation of the subsurface offsets from equation (13) is not immediate. However, it can be demonstrated (Wapenaar and Berkhout, 1987; Biondi, 2003) that the prestack image $I(m, h; z)$ obtained by equation (13) is equivalent to the image obtained by source-receiver migration applying the procedure outlined in equations (1) and (2). The subsurface offsets can thus be equated to the data offset of the recorded data datumed at depth.

The computation of the ADCIGs after imaging is based on a plane-wave decomposition of the prestack image $I(m, h; z)$ – which is obtained by either source-receiver migration or shot-profile migration – by applying slant-stacks along the offset axes, similarly to the computation before imaging that we discussed in the previous section. The only difference between the two methods (before and after imaging) is that the dips along the offset axes are affected by the transformation from time to depth that is implicit in the imaging step. Therefore, the offset-dip parameters are linked to the reflection opening angle $\gamma$ and azimuth $\phi$ differently than in the previous case. We will now derive and discuss the analytical relationships between reflection angles and offset dips after imaging. We will start with the simpler 2-D case (Sava and Fomel, 2003), and then address the general 3-D case (Biondi et al., 2003).

The application of the imaging condition transforms a wavefield propagating in time into an image cube that is a function of depth. The transformation from time to depth depends on the local dips in the wavefield and the local propagation velocity. In the frequency-wavenumber domain this transformation is represented by the DSR operator, which in 2-D can be expressed as a function of the angles $\beta_s$ and $\beta_r$ as follows:

$$k_z = -\frac{\omega}{v(z, x)}(\cos \beta_r + \cos \beta_s).$$  (14)

Recalling the relationship between the in-line offset wavenumber $k_{x_h}$ and the in-line offset ray parameter $p_{x_h}$:

$$k_{x_h} = p_{x_h} \omega,$$  (15)
and substituting both equation (14) and equation (9) in equation (15), we obtain the following relationship:

$$k_{x_b} = -p_{s_b} \frac{k_z v(z,x)}{\cos \beta_x + \cos \beta_s} = -\frac{2 \sin \gamma \cos \alpha_x}{v(z,x)} \frac{k_z v(z,x)}{2 \cos \gamma \cos \alpha_x} = -k_z \tan \gamma.$$  \hfill (16)

This relationship directly links the dips in the depth-offset domain ($k_{x_b}/k_z$) to the aperture angle ($\tan \gamma$).

Figure 4 illustrates the computation of 2-D ADCIGs after imaging with a synthetic example. It is analogous to Figure 1, which illustrates the computation of 2-D ADCIGs before imaging. Figure 4a shows a vertical section of the prestack image taken at constant midpoint. We will refer to this kind of section as an Offset Domain Common Image Gather (ODCIG). This ODCIG was obtained with the correct velocity, and the energy is well focused at zero offset for both the dipping and the flat reflectors. Slant stacks transform the impulses at zero offset into flat events in the angle domain (Figure 4b).

To generalize the 2-D offset-to-angle transformation to 3-D, we use the coplanarity condition and the geometrical model shown in Figure 3. In this context, the 2-D DSR expressed in equation (14) describes upward-propagating waves on the tilted plane shown in Figure 3, instead of the vertical plane shown in Figure 2. Consequently, the vertical wavenumber in equation (14) is now the vertical wavenumber along the tilted vertical axis $k_z'$, and not the vertical wavenumber along the “true” vertical direction. This distinction is irrelevant in the case of ADCIGs computed before imaging because the computation is performed at each depth level independently, but it is required in this case because the plane-wave decomposition is performed in the depth domain.

If $\delta_s'$ and $\delta_r'$ are the angles that the source and receiver rays (plane waves) form with the “true” vertical direction (as indicated in Figure 3), simple trigonometry relates these an-
gles to $\beta_{y'}$ and $\beta_{y''}$ through the tilt angle $\alpha_{y'}$ as follows: $\cos \delta_{y'} = \cos \beta_{y'} \cos \alpha_{y'}$, and $\cos \delta_{y''} = \cos \beta_{y''} \cos \alpha_{y''}$. The vertical wavenumber $k_{z'}$ is thus related to the vertical wavenumber $k_z$ as follows: $k_{z'} = k_z \cos \alpha_{y'} = \sqrt{k_z^2 + k_{y'}^2}$. Substituting this relationship into equation (16) leads to its 3-D equivalent:

$$k_{x'_{h}} = -k_{z'} \tan \gamma = -\sqrt{k_z^2 + k_{y'}^2} \tan \gamma.$$  \hspace{1cm} (17)

This 3-D expression is not independent of the geological dips as was its 2-D equivalent [equation (16)]. In the presence of significant cross-line dips, it is thus important to use the correct 3-D expression in place of the approximate 2-D one. Figures 5–6 illustrate with a data example the effects of the cross-line dip correction in equation (17).

When deriving the after-imaging equivalent of equation (11), we need to take into account that $\alpha_{y'}$ is also measured along the tilted axis and thus that $\tan \alpha_{y'} = -k_{y'}/k_{z'}$. Using equation (11) and equation (17) we derive the following relationship, which expresses $k_{y'_{h}}$ as a function of the other wavenumbers in the image:

$$k_{y'_{h}} = -k_{y} \tan \gamma \tan \alpha_{y'} = -\frac{k_{y'_{h}} k_{x'_{h}}}{k_{z'}^2} = -\frac{k_{y'_{h}} k_{x'_{h}}}{k_{z'}^2 + k_{y'}^2}.$$ \hspace{1cm} (18)

The combination of equation (17) and equation (18) enables the computation of 3-D ADCIGs after imaging; using these two relationships, we can map the prestack image from the offset domain $(x_{h}, y_{h})$ to the angle domain $(\gamma, \phi)$. Notice that the reflection azimuth $\phi$ is implicitly included in these equations since, all the wavenumbers, with the exception of $k_z$, are dependent on $\phi$. Therefore, during computations the wavenumbers need to be properly transformed by rotating the horizontal axes by the azimuth angle $\phi$.

In contrast with the transformation described by equations (9) and (11), the transformation described by equations (17-18) is independent of the local velocity $v(z, x, y)$. Therefore there is no need to estimate the geological dips locally, but only globally. This estimate can be performed accurately and efficiently in the wavenumber domain. However, it is important to notice that the absence of the local velocity $v(z, x, y)$ from the expressions used to compute ADCIGs after imaging does not make the result independent from the local migration velocity. This dependence is indirect through the vertical wavenumber $k_z$; or, in other words, the estimates of $\gamma$ and $\phi$ depend on the apparent vertical wavelength of the imaged reflectors. Therefore, the advantages derived from the absence of $v(z, x, y)$ in equations (17-18) are purely computational. The sensitivity of the estimates of $\gamma$ and $\phi$ on the accuracy of the local velocity $v(z, x, y)$ is the same, regardless of whether $\gamma$ and $\phi$ are estimated indirectly through $p_{x_{h}}$ and $p_{y_{h}}$ by using equations (9) and (11), or directly by using equations (17-18). This concept is well illustrated by the example shown in the last section (Figures 11–12).

There are two alternative computational algorithms to numerically perform the transformation described by equations (17-18); they differ in whether the computations are performed with the offset axes in the space or wavenumber domain. In either case, it is computationally advantageous to perform the computation in the wavenumber domain $(k_z, k_x, k_y)$ for the physical coordinates, because of the dependence of the mapping on the geological dips. If the ADCIGs are computed for many values of $\gamma$ and $\phi$, it is less expensive to perform the
computation in the offset-wavenumber domain by a 3-D generalization of the 2-D radial-trace transform used by Ottolini and Claerbout (1984) and Sava and Fomel (2003). However, this approach can generate artifacts, because the subsurface-offset axes are usually short, and the Fourier transforms have circular boundary conditions. The computation of the slant stack by integration in the offset-space domain avoids these artifacts and can be preferable when high-quality gathers are needed.

Examples of 3-D ADCIGs

This section presents several examples of 3-D ADCIGs. The first example is taken from the SEG-EAGE salt data set (Aminzadeh et al., 1996). It illustrates the error introduced when the approximate 2-D relationship [equation (16)] is used instead of the correct 3-D one [equation (17)] to estimate the aperture angle in presence of significant geological dips.

Figure 5 shows a depth slice (z=580 meters) of the migrated image of the SEG-EAGE salt data set. The crosshair is centered at a horizontal location where the top of the salt dips at approximately 50 degrees in the cross-line direction and is flat along the in-line direction. Figure 6 shows the ADCIGs computed at the location marked by the crosshair. Figure 6a was computed using the 2-D relationship [equation (16)], whereas Figure 6b was computed using the 3-D relationship [equation (17)]. The aperture angle is overestimated in the gather on the left (apparent maximum aperture is about 60 degrees), and correctly estimated in the gather on the right (apparent maximum aperture is about 48 degrees). This error is consistent with the factor $\cos \gamma_0$ that is neglected in the 2-D case. Notice that the bottom of the salt reflection ($z \approx 2,100$ meters) is unchanged, because it is flat.

In 3-D, ADCIGs are five-dimensional objects, and thus it can be challenging to gain an intuitive understanding of their behavior. The next set of figures shows 2-D slices of the 5-D space generated by computing 3-D ADCIGs from the migrated cube obtained from a simple synthetic data set. The data set contains 5 dipping planes, dipping at 0°, 15°, 30°, 45° and 60° toward increasing $x$ and $y$. The azimuth of the planes is 45 degrees with respect to the direction of the acquisition. The velocity is $V(z) = (1,500 + .5 z)$ m/s, which roughly corresponds to typical gradients found in the Gulf of Mexico. The acquisition geometry had a single azimuth oriented along the $x$ axis, and the maximum source-receiver offset was 3,000 meters. Figure 7 shows the geometry of the reflectors. The data were imaged with a full source-receiver 3-D prestack migration.

Because of the velocity gradient and the oblique azimuthal orientation, the azimuths of the reflections are not equal to the azimuth of the acquisition ($\phi = 0$). The reflection azimuths are within the range of $0° \leq \phi \leq 15°$ and depend on the reflector dip and on the aperture angle $\gamma$. The steeper the reflector dip and the wider the aperture angle are, the larger the azimuth rotation is at the reflection point.

All the following figures show slices of the ADCIGs at one fixed horizontal location with $x = y = 450$ meters; that is, they show slices through the 3-D image cube described as ADCIG$(z, \gamma, \phi)$. The most familiar of these slices display the image as a function of the depth ($z$) and the aperture angle ($\gamma$). Figure 8 shows two of these ADCIGs, for two different
Figure 5: Depth slice (z=580 meters) of the migrated image of the SEG-EAGE salt data set. The crosshair is centered at a horizontal location where the top of the salt dips at approximately 50 degrees in the cross-line direction.

Figure 6: ADCIG computed using the approximate 2-D relationship (panel a), and ADCIG computed using the correct 3-D relationship (panel b). The aperture angle $\gamma$ of the top of the salt reflection ($z \simeq 600$ meters) is overestimated in panel (a), whereas the bottom of the salt reflection ($z \simeq 2,100$ meters) is the same in the two panels.
reflection azimuths: $\phi = 10$ degrees (panel a) and $\phi = 15$ degrees (panel b). The reflections from the deepest – and steepest – reflector ($z \simeq 1,430$ meters) are well focused within the range delimited by these two azimuths. In contrast, the reflections from the other two reflectors (dipping at 30 and 45 degrees), are not well focused at these azimuths, and thus they frown downward even if the migration velocity is correct.

Figure 9 shows a slice taken at the constant depth of $z = 1,430$ meters; this depth corresponds to the deepest reflector. The reflection amplitudes are thus shown as functions of both the aperture angle ($\gamma$) and the reflection azimuth ($\phi$). Because of the poor azimuthal resolution close to normal incidence, the azimuthal range is wide for small $\gamma$; it narrows around $\phi = 15$ degrees as $\gamma$ increases.

The increasing azimuthal resolution with aperture angle is clearly demonstrated in Figure 10. The three panels in Figure 10 display the image as a function of depth ($z$) and reflection azimuth ($\phi$), and at constant aperture angle. The aperture angles are: a) $\gamma = 4$ degrees, b) $\gamma = 20$ degrees, and c) $\gamma = 30$ degrees. The curvature of the reflectors as a function of the azimuth increases with increasing aperture angle, indicating that the azimuthal resolution increases as the aperture angle widens. In other words, the common-azimuth data “illuminates” all the reflection azimuths for narrow aperture angles, but “illuminates” only a narrow range of reflection azimuths at wide aperture angles.

Figure 7: Reflectors’ geometry for the synthetic data set used to illustrated 3-D ADCIGs. The reflectors are slanted planes, dipping at 0°, 15°, 30°, 45° and 60° toward increasing $x$ and $y$; they are oriented with an azimuth of 45° with respect to the in-line direction. [NR]

ANGLE-DOMAIN COMMON IMAGE GATHERS AND VELOCITY

Migration Velocity Analysis (MVA) is one of the most important applications of ADCIGs, and the main focus of this paper. MVA from ADCIGs computed using wavefield-continuation migration is conceptually analogous to MVA from conventional offset-domain CIGs obtained by Kirchhoff migration. The ADCIGs are analyzed for departure from flatness, usually by applying some kind of Residual Moveout (RMO) analysis. The RMO analysis provides measurements of depth errors that are first transformed into travelt ime perturbations and finally inverted into velocity perturbations by a ray-based tomographic inversion. This process is repeated several times in order to converge to a satisfactory result.
Figure 8: ADCIGs as functions of depth ($z$) and aperture angle ($\gamma$) for two different reflection azimuths and at constant horizontal location ($x = y = 450$ meters): $\phi = 10^\circ$ (a), and $\phi = 15^\circ$ (b).

Figure 9: ADCIG as a function of aperture angle ($\gamma$) and reflection azimuth ($\phi$) at constant depth ($z = 1,430$ meters) and horizontal location ($x = y = 450$ meters).

Figure 10: ADCIGs as functions of depth ($z$) and reflection azimuth ($\phi$) for three different aperture angles and at constant horizontal location ($x = y = 450$ meters): $\gamma = 4^\circ$ (a), $\gamma = 20^\circ$ (b), and $\gamma = 30^\circ$ (c).
In this section we analyze the kinematic properties of 3-D ADCIGs when the migration velocity is not correct. First, we qualitatively analyze the sensitivity of the estimation of the angles \((\gamma, \phi)\) to velocity errors. Second, we develop a quantitative relationship between depth errors in ADCIGs and traveltime perturbations caused by velocity errors. Third, we present a 3-D RMO function for measuring the depth errors in ADCIGs. Finally, we compare the imaging-point shifts predicted by our results with the actual images produced from a synthetic data set migrated with the wrong velocity.

Sensitivity of the estimation of \((\gamma, \phi)\) to velocity errors

We start with a simple 2-D qualitative analysis of the effects of velocity errors on both kinds of ADCIGs that we presented; that is, ADCIGs computed before imaging and ADCIGs computed after imaging. This first step illustrates the sensitivity of the estimation of the angles \((\gamma, \phi)\) to velocity errors. Figure 11 shows the effects of velocity errors on ADCIGs computed before imaging. It shows the same panels as Figure 1 [i.e. wavefield at zero-offset (a), wavefield at fixed midpoint (b), and ADCIG (c)], but when the migration velocity is 10% higher than the correct one. At the correct depth of the flat reflector \((z = 700\) meters), the energy for both reflectors has started defocusing and it forms time-reversed hyperbolas. Most of the energy for the flat reflector has not reached the zero-time line yet, causing the flat reflector to be imaged deeper than the correct depth. In the ADCIG both events frown downward, and thus they indicate too high of a migration velocity. The residual moveout caused by velocity errors in ADCIGs is thus qualitatively similar to the moveout observed in conventional surface-offset CIGs computed by Kirchhoff migration. If the velocity function is too low the reflections will smile upward; if the velocity function is too high the reflections will frown downward.

Figure 12 shows the effects of velocity errors on ADCIGs computed after imaging. It shows the same panels as Figure 4 [i.e. ODCIG (a), and ADCIG (b)], but when the migration velocity is 10% higher than the correct one. The energy is defocused in the ODCIGs and the events have a hyperbolic moveout (Figure 12a); the apexes of the hyperbolas are deeper than the reflectors’ true depths. Slant stacks transform the hyperbolas into frowns (Figure 12b) that are similar to the frowns shown in Figure 11c. However, there is a subtle but important difference between the two cases. The \(p_{sb}\) range does not change between the ADCIG shown in Figure 1c and the one shown in Figure 11c, whereas the \(\gamma\) range increases when the velocity is too high (Figure 12b). This increase is due to the increase in the apparent vertical wavelength in the image, which correspondingly causes a decrease of \(k_z\) in equation (16). Or, from a different viewpoint, if we were to use equation (9) to map \(p_{sb}\) into \(\gamma\), the mapping would be affected by both the increase in \(v(z,x)\) and the increase in apparent geological dip \(\alpha_x\) — and the corresponding decrease in \(\cos\alpha_x\). This simple example illustrates the fact that the estimates of \(\gamma\) and \(\phi\) are similarly sensitive to the accuracy of the local velocity \(v(z,x,y)\), regardless of whether ADCIGs are computed before or after imaging.

Kinematic properties of 3-D ADCIGs

Biondi and Symes (2003) analyze in detail the kinematic properties of 2-D ADCIGs when
Figure 11: Illustration of the sensitivity to velocity errors of the ADCIGs computed before imaging. The panels are the same as in Figure 1; that is, they display orthogonal slices of the prestack wavefield after downward continuation to the depth of 700 meters, but the velocity used for downward continuation was 10% higher than the correct one. Panel a) displays the zero-offset section, panel b) displays the common-midpoint gather at 1,410 meters, and panel c) shows the complete (i.e. for all depths) ADCIG at 1,410 meters.

Figure 12: Illustration of the sensitivity of ADCIGs computed after imaging to velocity errors. The panels are the same as in Figure 4, but after migration with a velocity 10% higher than the correct one. Panel a) displays the ODCIG and panel b) the ADCIG(γ). Notice that the range of γ increases from Figure 4b to Figure 12b.
the migration velocity is not correct. Their analysis can be extended to 3-D by following the same geometrical considerations we used in the previous sections to generalize ADCIGs from 2-D to 3-D (Figure 3). However, there is an important conceptual difference between the 2-D case and the 3-D case. In 2-D we can assume that the source and receiver rays cross even when the data were migrated with the wrong velocity; below the imaging point in case of too low migration velocity and above the imaging point in the opposite case. In 3-D, this assumption is easily violated because two rays are not always coplanar. Fortunately, the plane-wave interpretation is still valid and the plane of coplanarity of Figure 13 is determined by the coplanarity condition expressed in equation (10) applied to the phase vectors of the incident and reflected plane waves, or, alternatively by the coplanarity condition expressed in equation (18) applied to the offset dips present in the prestack image. The source and receiver plane waves constructively interfere along an imaging line; the angle-domain image point I_y is defined as the intersection of this imaging line with the tilted plane of coplanarity. Once this plane is defined, the geometrical relations between the objects (rays and imaging points) that lie on the vertical plane in 2-D (Biondi and Symes, 2003), directly apply to 3-D on the tilted plane of coplanarity, as it is schematically represented in Figure 13. In particular, the relative position of the angle-domain image point I_y and of the crossing point between the source and receiver ray I_i are defined on the tilted plane. The 2-D results translate into the following 3-D results:

1. The transformation to the angle domain shifts the image point along the tilted vertical direction z' from the offset-domain image point I_{x_0} to the angle-domain image point I_y.

2. I_y lies on the normal to the apparent geological dip passing through the crossing point of the source and receiver rays (I_i). I_y is located at the crossing point of the lines passing through S_0 and R_0 and orthogonal to the source ray and receiver ray, respectively.

3. From the previous geometric results, by invoking Fermat’s principle and applying simple trigonometry, we can also easily derive a relationship between the total normal shift \Delta n_{tot} = (I_y - I_i) and the total traveltime perturbation caused by velocity errors as follows:

   \[ \Delta n_{tot} = \frac{\Delta t}{2S \cos \gamma}, \]  

   where \( S \) is the background slowness around the image point and \( \Delta t \) is defined as the difference between the perturbed traveltime and the background traveltime.

These results are based on the assumption that the velocity is locally smooth in a neighborhood of the imaging point. Furthermore, the relationship expressed in equation (19) depends on the assumption of stationary raypaths, since we need to invoke Fermat’s principle. We can assume stationary raypaths if the velocity perturbations are small. Biondi and Symes (2003) discuss these assumptions in details.

The relationship expressed in equation (19) forms the basis of using ADCIGs for MVA. Similar relationships have been extensively used in MVA methods based on constant-offset
Kirchhoff migration (Etgen, 1990; Stork, 1992; Meng et al., 1999a,b). Equation (19) can be directly used to transform measurements of depth errors in ADCIGs into traveltime perturbations that can be inverted by a ray-based tomographic inversion; or it can be the foundation for the derivation of accurate RMO functions to be applied to measure depth errors in ADCIGs.

**Residual moveout in ADCIGs**

The inconsistencies between the migrated images at different aperture angles are the primary source of information for velocity updating during Migration Velocity Analysis (MVA). An effective and robust method for measuring inconsistencies between images is to compute first semblance scans as a function of one “residual moveout” (RMO) parameter, and then pick the maxima of the semblance scan. This procedure is most effective when the residual moveout function used for computing the semblance scans closely approximates the true moveouts in the images.

Appendix A presents the derivation of a RMO functions for scanning 3-D ADCIGs. The starting point of our derivation is the 2-D RMO function presented by Biondi and Symes.
(2003). Its generalization is based on the 3-D geometrical understanding developed in the previous sections, and it leads to the following expression:

$$\Delta n_{\text{RMO}} = \frac{\rho - 1}{\cos \alpha} \frac{\sin^2 \gamma}{(1 - \sin^2 \alpha \cos^2 (\eta - \phi) - \sin^2 \gamma)} z_0 n,$$

where $\rho$ is the ratio between the true velocity and the migration velocity, $z_0$ is the migrated depth of the normal incidence event, $\alpha$ is the geological dip, and $\eta$ is the azimuth angle of the normal to the geological dip (see Figure 13).

Equation (20) is the general expression of the RMO function in 3-D. For flat reflectors (i.e. $\alpha = 0$) equation (20) becomes the much simpler RMO function (Biondi and Symes, 2003) that follows:

$$\Delta n_{\text{RMO}} = (\rho - 1) \tan^2 \gamma z_0 n. \hspace{1cm} (21)$$

It is also easy to verify that when the reflection azimuth is aligned with the geological dip azimuth (i.e. “dip reflection” with $\eta = \phi$), equation (20) simplifies to the following 2-D RMO function:

$$\Delta n_{\text{2D-RMO}} = \frac{\rho - 1}{\cos \alpha_{\psi}} \frac{\sin^2 \gamma}{(\cos^2 \alpha_{\psi} - \sin^2 \gamma)} z_0 n, \hspace{1cm} (22)$$

with $\alpha_{\psi} = \alpha$. Similarly, when the reflection azimuth is orthogonal to the geological dip (i.e. “strike reflection” with $\eta - \phi = \pm 90^\circ$), equation (20) simplifies into the following expression:

$$\Delta n_{\text{RMO}} = \frac{\rho - 1}{\cos \alpha} \tan^2 \gamma z_0 n. \hspace{1cm} (23)$$

Equation (23) is the RMO function for flat events, [equation (21)], except for the scaling factor of $1/\cos \alpha$ (with $\alpha = \alpha_{\psi}$) that takes into account the inclination of the tilted plane in Figure 13.

The special cases of dip-reflections and strike-reflections depend on the orientation of the geological dip relative to the reflection azimuth and not on the data azimuth. They are thus subsurface (wavefields at depth) properties and not surface (recorded data) properties; that is, they depend on the velocity model in the overburden as well as on the relative orientation of the acquisition direction with respect to the geological dips.

**Examples of 3-D ADCIGs in presence of velocity errors**

To verify the accuracy of our kinematic analysis of 3-D ADCIGs in presence of migration velocity errors, we compare the predictions of our results [equation (19) and equation (20)] with the actual reflector shifts in the angle-domain image cube obtained from a simple synthetic data set. The data set is the same as the one used in the previous section (5 dipping planes immersed in a vertically layered velocity function). We have migrated the data with a velocity function 3% slower than the correct one; that is, the slowness function is scaled by 1.03 at every depth level.
As we have discussed in the previous sections, the azimuth of the reflections $\phi$ increases as the aperture angle $\gamma$ increases. Therefore, equation (19) and the RMO function define curved lines that live in a 5-D space. To verify their accuracy in predicting actual reflector movements we display the image cube for the following three pairs of $\gamma$ and $\phi$: $(\gamma = 0^\circ, \phi = 10^\circ)$ in Figure 14, $(\gamma = 20^\circ, \phi = 11^\circ)$ in Figure 15, and $(\gamma = 30^\circ, \phi = 13^\circ)$ in Figure 16. These are the reflections’ angles for the steepest (60 degrees dip) reflector when the velocity is correct, consistently with our assumption of unperturbed rays. All these figures display orthogonal sections cut through the image cube at fixed $\gamma$ and $\phi$ and function of the physical coordinates $z, x$, and $y$. In panels a) the cross-hairs are centered at the coordinates predicted using equation (19); whereas in panels b) the cross-hairs are centered at the coordinates predicted by the 3-D RMO function expressed in equation (20). The numerical values of these coordinates are shown by the numbers at the edges of the black lines that define the cross-hairs.

The predictions corresponding to equation (19) were computed starting from a numerical raytracing of the reflected events using the correct velocity. The traveltime perturbation $\Delta t$ was then computed by integrating the slowness perturbations along the unperturbed rays. Given $\Delta t$, $\Delta n_{tot}$ was computed using equation (19). If the background velocity were constant (straight rays) the procedure outlined above would result in image-point shifts exactly equal to the image-point shifts computed using equation (20). In our example the background velocity increases with depth and therefore the two predictions differ for large aperture angles.

The cross-hairs in the panels on the left accurately track the reflector movements in the corresponding image cubes. These panels demonstrate that equation (19) accurately predicts the actual movements of the reflector across the whole range of aperture angles, even if the reflector analyzed is steeply dipping at 60 degrees.

As expected, the predictions computed by applying equation (20) (panels b) are not as accurate as the predictions computed using equation (19) (panels b). Consistently to the intended use of RMO functions, the shift at normal incidence ($\gamma = 0^\circ$) was picked from the image, whereas the shifts at $\gamma = 20^\circ$ and at $\gamma = 30^\circ$ were computed using equation (20). The predicted shifts track fairly well the reflector movements in the image up to an aperture angle of 20 degrees (Figure 15b), but the straight-rays approximation underlying our derivation of the 3-D RMO function breaks down when either the source or the receiver rays are close to overturning. As it is easy to verify by simple trigonometry, when $\gamma = 30^\circ$ and $\phi = 13^\circ$ (Figure 16b) the take-off angle at the reflection point is about 81 degrees for one of the two rays. A straight ray propagating at such an oblique angle is a poor approximation of the true ray that starts with the same take-off angle at the reflection point. The true ray quickly becomes more vertical as it travels toward the surface, and consequently it is much shorter than the straight ray. This break down of the assumptions at the basis of the derivation of the 3-D RMO function, causes the substantial overprediction of the reflector movements observed in Figure 16b.
Figure 14: Orthogonal sections cut through the migrated image cube obtained with a velocity function too low by 3%. The image cube is at fixed $\gamma = 0^\circ$ and $\phi = 10^\circ$. The cross-hairs are centered at the coordinates predicted by using equation (19) in panel a), and by using equation (20) in panel b).

Figure 15: Orthogonal sections cut through the migrated image cube obtained with a velocity function too low by 3%. The image cube is at fixed $\gamma = 20^\circ$ and $\phi = 11^\circ$. The cross-hairs are centered at the coordinates predicted by using equation (19) in panel a), and by using equation (20) in panel b).
**CONCLUSIONS**

We present a new methodology for computing 3-D Angle Domain Common Image Gathers (ADCIGs) from the results of wavefield-continuation migration. The proposed methodology is general and can be applied to the results obtained with any wavefield-continuation migration method, including: downward-continuation shot-profile migration, reverse-time shot-profile migration, source-receiver migration, and plane wave migration. Furthermore, our method can be used to compute ADCIGs before imaging and after imaging. These two kinds of ADCIG are related, though we show that they have a different sensitivity to velocity errors.

We successfully applied the proposed methods to the computation of 3-D ADCIGs from a simple synthetic data set. 3-D ADCIGs represent reflections as a function of the reflection azimuth $\phi$, in addition to the aperture angle $\gamma$. Whereas reflection azimuth is obviously related to the data azimuth, our synthetic-data example demonstrates that they may differ when velocity variations in the overburden distort the wavefield. In the general case, both $\gamma$ and $\phi$ are needed to characterize the events and to use ADCIGs for either amplitude analysis or Migration Velocity Analysis (MVA).

To enable the application of 3-D ADCIGs to MVA, we generalize the kinematic analysis of 2-D ADCIGs in presence of migration-velocity errors (Biondi and Symes, 2003). Our analysis yields a quantitative relationship between depth errors measurable in 3-D ADCIGs and traveltine perturbations caused by velocity errors. This relationship can be directly used in a
ray-based MVA procedure. In this paper, we use it for deriving an approximate 3-D Residual Moveout (RMO) function that predicts reflectors’ movement when migration velocity is inaccurate. Our derivation of the 3-D RMO function assumes straight rays.

We confirm the accuracy of our kinematic analysis by comparing the image-point shifts predicted by its results with the actual reflector’s shifts in the angle-domain image cube obtained from a 3-D synthetic data set with steeply dipping reflectors and a vertically varying propagation velocity. The proposed 3-D RMO function is accurate for a wide range of dips and aperture angles. However, as expected, when rays are close to overturn the straight-rays approximation is inadequate, and the 3-D RMO function overpredicts the actual reflector’s movements.

ACKNOWLEDGMENTS

We would like to thank Sam Gray for useful discussions on ADCIGs and for pointing out to us that a plane-wave interpretation of 3-D ADCIGs would overcome the difficulties encountered by a ray-theoretical one.

REFERENCES


Biondi, B., and Palacharla, G., 1996, 3-D prestack migration of common-azimuth data: Geophysics, 61, no. 6, 1822–1832.


**APPENDIX A**

In this Appendix we derive the expression for the residual moveout (RMO) function to be applied to 3-D ADCIGs computed by wavefield continuation.

We start from the 2-D expression for RMO, as derived by Biondi and Symes (2003):

\[
\Delta n_{2D-RMO} = \frac{\rho - 1}{\cos \alpha_y} \frac{\sin^2 \gamma}{\left(\cos^2 \alpha_y - \sin^2 \gamma\right)} z_0 \mathbf{n}, \tag{A-1}
\]

In 3-D, this relationship is valid on the tilted plane shown Figure 3, where the migrated depth of the normal incidence event \(z_0\) is now the depth along the tilted plane \(z_0'\). Therefore, the migrated depth \(z_0\) needs to be scaled by a factor of \(1/\cos \alpha_y\), where \(\alpha_y\) is the inclination of the tilted plane. In 3-D equation (A-1) becomes:

\[
\Delta n_{RMO} = \frac{\rho - 1}{\cos \alpha_y} \frac{\sin^2 \gamma}{\left(\cos^2 \alpha_x - \sin^2 \gamma\right)} z_0' \mathbf{n} = \frac{\rho - 1}{\cos \alpha_x \cos \alpha_y} \frac{\sin^2 \gamma}{\left(\cos^2 \alpha_x - \sin^2 \gamma\right)} z_0 \mathbf{n}. \tag{A-2}
\]
When $\alpha_x'$ and $\alpha_y'$ are available, the RMO function could be directly evaluated using the expression in equation (A-2). However, in several situations it is more useful to express the RMO function explicitly as a function of the geological dip $\alpha$, the azimuth angle of the normal to the geological dip $\eta$, and the azimuth angle of the reflected event $\phi$ (see Figure 13). To derive the desired expression, we use the following two trigonometric relationships among the angles:

$$\cos \alpha = \cos \alpha_x' \cos \alpha_y', \quad \text{and} \quad \sin \alpha_x' = \sin \alpha \cos (\eta - \phi). \quad (A-3)$$

Substituting equations (A-3) into equation (A-2), we obtain the following final result:

$$\Delta \mathbf{n}_{\text{RMO}} = \frac{\rho - 1}{\cos \alpha} \frac{\sin^2 \gamma}{\left(1 - \sin^2 \alpha \cos^2 (\eta - \phi) - \sin^2 \gamma\right)^2} \mathbf{n}. \quad (A-4)$$
Residual move-out analysis with 3-D angle-domain common-image gathers

Thomas Tisserant and Biondo Biondi

ABSTRACT

We describe a method to update the velocity model from the residual move-out information contained in 3-D angle-domain common-image gathers. The 3-D angle-domain common-image gathers computed after wave-equation migration are functions of the aperture angle and the reflection azimuth angle. We perform a velocity error analysis by semblance using 2-D and 3-D residual move-out functions. Both functions enable us to update the velocity model. The 3-D function leads to a better estimation of the velocity error in the case of 3-D events.

INTRODUCTION

Migration velocity analysis combines migration and velocity analysis in an iterative process. After each migration, if the velocity model is not correct, the events in the common-image gathers (CIGs) are not flat. Migration velocity analysis uses the residual move-out (RMO) in the gathers to update the velocity model.

Wavefield-continuation migration offset-domain common-image gathers (ODCIGs) generated are inconvenient for extracting velocity information. Conversely, wavefield-continuation migration angle-domain common-image gathers (ADCIGs) are convenient because their energy is spread over the gather, similar to ODCIGs in Kirchhoff migration. In addition, ADCIGs avoid ambiguities of the reflector positions in Kirchhoff ODCIGs when multipathing occurs (Prucha et al., 1999).

The ADCIGs can be formed during the imaging step of the wavefield-continuation migration (de Bruin et al., 1990; Prucha et al., 1999). In this case, the ADCIGs are expressed in terms of the offset ray-parameter instead of the aperture angle. Sava and Fomel (2003) propose a method applied after imaging to obtain ADCIGs as functions of the direct aperture angle. Pointing out that the method is valid only for reflectors orthogonal to the acquisition direction (dip-reflection), and consequently ignores the azimuth angle of the reflection, Biondi and Tisserant (2004) generalize it to 3-D. The resulting ADCIGs are not only function of aperture angle, but also function of reflection azimuth angle.

As well as providing additional information to perform velocity analysis, the knowledge of
the reflection azimuth can also be used to correctly backproject rays in reflection tomography, or to perform amplitude-versus-angle analysis along the reflection azimuth at the image point rather than along the acquisition azimuth.

Biondi and Symes (2003) analyze the kinematics of the ADCIGs in the presence of velocity error and derive an RMO function. Biondi and Tisserant (2004) extend this study to 3-D. This paper describes the estimation of the velocity error from 3-D ADCGIs in order to update the velocity model. We first review the additional information contained by the 3-D ADCIGs. We then quantify the velocity error from the ADCIGs with a 2-D RMO function. The last part describes how accounting for the 3-D wave propagation improves the estimation of the velocity error. The method is tested on the synthetic model displayed in Figure 1. The model consists of 5 dipping planes in a linearly increasing velocity $v(z) = 1500 + 0.5z$ m/s. The planes have a 45° azimuth. The dataset is migrated with a full prestack wavefield-continuation migration and a velocity that is 3% too low.

3-D ANGLE-DOMAIN COMMON-IMAGE GATHERS

Our velocity-error estimation method uses ADCIGs computed after imaging through an offset-to-angle transformation (Sava and Fomel, 2003). Although valid in 2-D, Biondi et al. (2003) point out that the transformation assumes the azimuth of the reflectors to be orthogonal to the azimuth of the acquisition (dip-reflections). The transformation also ignores the reflection azimuth. The reflection azimuth is equal to the acquisition azimuth when the rays do not bend (in constant velocity), or in case of dip-reflections. Biondi and Tisserant (2004) extend the method to 3-D. The transformation maps the migrated inline and cross-line offsets into the aperture angle ($\gamma$) and reflection azimuth angle ($\phi$). Hence, after being transformed to the angle domain, each CIG is a $(z,\gamma,\phi)$ cube.
following the characteristics of the fourth reflectors on the model displayed in Figure 1, Figure 2 simulates a reflection from a reflector with a 45° dip and a 45° azimuth in a vertically increasing velocity. It illustrates how different the reflection azimuth can be from the acquisition azimuth in a 3-D configuration. The gray lines symbolize the survey azimuth at constant depth. The black lines symbolize the azimuth between the source and receiver rays at each depth. At the reflection position, the angle between the black and green lines is the reflection azimuth.

Figure 3 shows slices of the $(z, \gamma, \phi)$ cube at constant $\phi$ for two CIGs from the model. The top row is the CIG taken at $x=700m$ and $y=700m$, where the steepest reflector is not visible. The bottom row is the CIG taken at $x=425m$ and $y=425m$, where only the two steepest reflectors are visible. For the remainder of this paper, we will refer to these two CIGs as CIG$_{700}$ and CIG$_{425}$. On both rows, the first panel is the stack over the $\phi$ dimension of the cube. The other panels are slices of the cube at constant $\phi$. As $\phi$ increases, the gathers frown downward, then smile upward, reach a maximum, and frown down again. Note that, although the velocity is incorrect, the gathers are flat for some values of $\phi$. Each event is stationary at one particular value of $\phi$ when the smile is at its maximum. For CIG$_{700}$, $\phi$ is around 0° for 0° dip and 15° dip reflectors, around 3° for the 30° dip reflector, and a bit higher than 6° for the 45° dip reflector. For CIG$_{425}$, $\phi$ is around 5° for 45° dip, and around 16° for the 60° dip reflector. Notice that the values of $\phi$ from one CIG to another are comparable for reflectors with the same dip. Since the gathers are stationary for the correct $\phi$ value, the stack carries some information about the reflection azimuth.

This analysis can be done whether the migration velocity is correct or not. If the velocity is incorrect, however, the measured values are apparent reflection azimuth angles. The difference between the true reflection azimuth and the apparent one is due to the incorrect ray-bending when the velocity is incorrect.

We next describe how to estimate the velocity error from the ADCIGs with a RMO function.
ESTIMATING THE VELOCITY ERROR WITH A 2-D RMO FUNCTION

Migrating with an incorrect velocity results in gathers frowning downward if the velocity is too high, and gathers smiling upward if the velocity is too slow. Biondi et al. (2003) analyze the kinematics of ADCIGs. Assuming stationary-raypath, or straight rays, the authors show that the apparent image point in the angle domain is located on the normal to the apparent reflector, and provide a 2-D RMO function to quantify the amount of the shift of the image point:

$$\Delta n_{RMO} = \frac{1 - \rho}{\cos \alpha} \frac{\sin^2 \gamma}{(\cos^2 \alpha - \sin^2 \gamma) z_0},$$  \hspace{1cm} (1)$$

where $\rho$ is the ratio between the slowness used for the migration and the true slowness, i.e., a measure of the velocity error, $\alpha$ the dip of the reflector, $\gamma$ the aperture angle, and $z_0$ the depth of the reflector. In the case of a flat reflector ($\alpha = 0$), equation (1) reduces to

$$\Delta n_{RMO} = (\rho - 1) \tan^2 \gamma z_0.$$  \hspace{1cm} (2)$$

The parameter we want to estimate is $\rho$. The RMO is measured along the normal to the reflector but can be applied along the vertical after scaling with the cosine of the dip.

Both functions are tested on two CIGs of the model displayed in Figure 1. Figure 4 illustrates a velocity error analysis by semblance. The top row is CIG$_{700}$, and the bottom row...
Figure 4: Velocity error analysis by semblance using a 2-D RMO function on CIG\textsubscript{700} on the top row, and on CIG\textsubscript{425} on the bottom row. a) Stack over $\phi$ of CIG (425,425). b) Semblance analysis of the first panel with a zero-dip assumption. c) Semblance analysis of the first panel with the dip information. Both semblance panels are squared to increase resolution.

The dips for CIG\textsubscript{700} range from 0° to 45°. The dips for CIG\textsubscript{700} range from 45° to 60°. Figure 4a is the stack over $\phi$ for the two CIGs. The upward smiles indicate that the migration velocity is too low. The velocity error being 3%, the peak of semblance will ideally be centered at $\rho = 1.03$.

Figure 4b is the velocity error spectrum of the left panels in the absence of dip information and using $\rho = 0$ (equation (2)). As expected, the peak of semblance is correctly positioned for the flat reflector at depth 960 m. However, the other peaks of semblance are mispositioned for the steeper reflectors, where the zero-dip assumption is not valid. For the steepest reflector, the peak of semblance is out of the semblance panel.

Figure 4c is the velocity error spectrum of the left panel, using the RMO function with the knowledge of the dip (equation (1)). The dip does not have to be known \textit{a priori}; it can be extracted from the migrated cube between the migration step and the velocity error estimation step. The peaks of semblance are better focused but are still not accurate enough, especially for the steepest reflector, on which we read $\rho = 1$, signifying that the velocity is correct. To correct this, we need to take into account 3-D events.

The 2-D RMO functions can be used to extract the velocity error from the gathers of the migrated cube. The simple RMO function works for flat reflectors. When the dip information is included, the RMO function corrects the error generated by steeply dipping reflectors. However, the peaks are not correctly positioned because of 3-D events that are not appropriately processed by equation (1).

**IMPROVING THE ESTIMATION WITH A 3-D RMO FUNCTION**

Biondi and Tisserant (2004) extend the 2-D RMO function (equation (1)) to a 3-D function
Figure 5: Velocity error analysis by semblance using a 3-D RMO function on CIG$_{700}$ on the top row, and on CIG$_{425}$ on the bottom row. a) CIG after stack over $\phi$. b) Semblance analysis of the left panel using the 3-D RMO function at $\phi = 0$. c) $\phi$ semblance analysis. d) Semblance analysis of the first panel using the 3-D RMO function with a $\phi(z)$ law. All semblance panels are squared to increase resolution.

that includes the azimuth of the reflector ($\eta$), and the azimuth of the reflection ($\phi$):

$$
\Delta n_{RMO} = \frac{\rho - 1}{\cos \alpha} \frac{\sin^2 \gamma \sqrt{1 - \sin^2 \alpha \cos^2 (\eta - \phi)}}{\cos^2 (\eta - \phi) - \sin^2 \gamma} z_0. \tag{3}
$$

Equation (1) is a special case of equation (3), where the azimuth of the acquisition and the reflection are the same ($\eta = \phi$): the 2-D RMO function implicitly assumes dip-reflections just as the 2-D offset-to-angle transformation does. In the case of a flat reflector ($\alpha = 0$), the RMO functions are the same in 2-D and 3-D, because horizontal planes have no defined azimuth and do not generate 3-D reflection.

Figure 5 synthesizes all the velocity analysis done with equation (3) on two CIGs taken from the model. CIG$_{700}$ (top) and CIG$_{425}$ (bottom) stacked over $\phi$ are represented in Figure 5a.

Figure 5b is the semblance analysis of the stack using the 3-D RMO function (equation (3)), with $\eta = 45^\circ$ (the true azimuth of the reflectors) and assuming $\phi = 0^\circ$ in the absence of $\phi$ information. For the first CIG, the comparison with the semblance panel obtained with the 2-D RMO function (Figure 4c) shows a more accurate resolution of the velocity error when the reflector azimuth is provided.

We perform a semblance analysis to determine the reflection azimuth. The procedure consists in scanning all the slices at constant $\phi$ of the $(z, \gamma, \phi)$ cube using equation (3), with the velocity error obtained from a previous analysis performed on the stack with the $\phi = 0$ assumption (Figure 5b). Then, only the trace corresponding to the measured velocity error is taken from each semblance panel at constant $\phi$. All the traces displayed side by side form the $\phi$ semblance panel displayed in Figure 6. The peaks of semblance give the reflection azimuth for each event. The value of $\phi$ ranges from $0^\circ$ to $7^\circ$ for the first CIG, and from $6^\circ$ to $16^\circ$ for the second CIG. They are consistent with the $\phi$ values read on Figure 3. Once again, the reflection
azimuths are comparable at similar reflector dips.

Figure 5c adds the reflection azimuth information picked in Figure 6 to the velocity error analysis. This extra information does not improve the resolution of the velocity error for the first CIG, where the range of $\phi$ is limited and satisfies the $\phi = 0$ assumption because of the limited dips. Note that the peaks of semblance are not centered perfectly on $\rho = 1.03$ but are closer to $\rho = 1.025$. This underestimated velocity error is due to the straight rays assumption that over-corrects the move-out.

For the second CIG, where the $\phi = 0$ assumption is not valid, we expect some changes. Paradoxically, the estimation seems less accurate for the $60^\circ$ reflector when the value of $\phi$ is used ($\rho = 1.02$) instead of setting $\phi$ to 0 ($\rho = 1.03$). The reason is when the $\phi = 0$ approximation is not valid, the RMO function over-estimates the velocity error. Simultaneously, the RMO function under-estimates the velocity error because of the straight rays approximation. Hence the errors cancel out and make the estimation apparently more accurate with the $\phi = 0$ assumption than with the correct value of $\phi$.

The 3-D analysis of the RMO in the ADCIGs improves the resolution of the velocity error, even if the knowledge of the reflection azimuth does not bring visible improvement. The precise estimation of the reflection azimuth still holds for tomographic or amplitude-versus-angle studies.

**CONCLUSION**

We have presented a method to estimate the error in the migration velocity. It is based on the analysis of the residual move-out in 3-D angle-domain common-image gathers. The 3-D ADCIGs are functions of the aperture angle and the reflection azimuth angle. We tested a 2-D
and a 3-D RMO function that led to an estimation of the velocity error and made it possible to update the velocity model. Because of 3-D events, the 2-D estimation of the velocity error is not accurate enough. The 3-D approach accounted for 3-D events and improved the resolution of the error. All the parameters required to estimate precisely the RMO function (the dips of the reflectors, the azimuth of the reflectors, and, to a minor extent, the reflection azimuth) can be estimated from the migrated data. The method can be automated and can serve as a basis for migration velocity analysis.

REFERENCES


Sensitivity kernels for
wave-equation migration velocity analysis

Paul Sava and Biondo Biondi

ABSTRACT
The success of migration velocity analysis methods is strongly dependent on the characteristics of the linearized tomographic operator that is inverted to estimate velocity updates. To study the properties of wave-equation migration velocity analysis, we analyze its sensitivity kernels. Sensitivity kernels describe the dependence of data space elements to small changes of model space elements. We show that the sensitivity kernels of wave-equation MVA depend on the frequency content of the recorded data and on the background velocity model. Sensitivity kernels computed assuming the presence of a salt body in the background velocity show that these kernels are drastically different from idealized “fat rays”. Consequently sensitivity kernels cannot be approximated by artificial fattening of geometrical rays. Furthermore, our examples illustrate the potential of finite-frequency MVA as well as the frequency-dependent nature of illumination for subsalt regions.

INTRODUCTION
Depth imaging of complex structures depends on the quality of the velocity model. However, conventional Migration Velocity Analysis (MVA) procedures often fail when the wavefield is severely distorted by lateral velocity variations and thus complex multipathing occurs. Biondi and Sava (1999) introduce a method of migration velocity analysis using wave-equation techniques (WEMVA), which aims to improve the quality of migrated images, mainly by correcting moveout inaccuracies of specular energy. WEMVA finds a slowness perturbation which corresponds to an image perturbation. It is thus similar to ray-based migration tomography (Al-Yahya, 1989; Stork, 1992; Etgen, 1993), where the slowness perturbation is derived from depth errors, and to wave-equation inversion (Tarantola, 1986) or tomography (Woodward, 1992; Pratt, 1999; Dahlen et al., 2000) where the slowness perturbation is derived from measured wavefield perturbations.

WEMVA has the potential of improving velocity estimation when complex wave propagation makes conventional ray-based MVA methods less reliable. Imaging under rugged salt bodies is an important case where WEMVA has the potential of making a difference in the imaging results. In this paper, we analyze the characteristics of the tomographic operator inverted in WEMVA to update the velocity model, and contrast these characteristics with the

1email: paul@sep.stanford.edu, biondo@sep.stanford.edu
well-known characteristics of ray-based tomographic operators.

One way of characterizing integral operators, e.g. tomography operators, is through sensitivity kernels, which describe the sensitivity of a component of a member of the data space to a change of a component of a member of the model space. In this paper, we formally introduce the sensitivity kernels for wave-equation migration velocity analysis and show 2D and 3D examples.

The analysis of WEMVA sensitivity kernels provide an intuition on WEMVA’s potential for overcoming limitations of ray-based MVA. Some of these limitations are intrinsic, other are practical. An important practical difficulty encountered when using rays to estimate velocity below salt bodies with rough boundaries is the instability of ray tracing. Rough salt topographies create poorly illuminated areas, or even shadow zones, in the subsalt region. The spatial distribution of these poorly illuminated areas is very sensitive to the velocity function. Therefore, it is often extremely difficult to trace the rays that connect a given point in the poorly illuminated areas with a given point at the surface (two-point ray-tracing). Wavefield-extrapolation methods are robust with respect to shadow zones and they always provide wavepaths (i.e. sensitivity kernels) usable for velocity inversion.

Ray-tracing has intrinsic limitations when modeling wave-propagation through salt bodies with complex geometry, because of the asymptotic assumption on which it is based. This intrinsic limitation prevent ray-tracing from modeling the frequency-dependency of full-bandwidth wave propagation. The comparison of sensitivity kernels computed assuming different frequency bandwidths illustrates clearly the drawbacks of the asymptotic assumptions. Top-salt rugosity causes the WEMVA sensitivity kernels to be strongly dependent on the bandwidth. Furthermore, in these conditions, sensitivity kernels are drastically different from simple “fat rays”. Therefore, they cannot be approximated by kernels computed by a bandwidth-dependent fattening of geometric rays (Lomax, 1994).

We compute the sensitivity kernels for perturbations in the phase as well as perturbations in the amplitude. We observe that there is a 90 degree phase shift between these two types of kernels. It is interesting to notice that the 3-D kernels for phase perturbations are hollow in the middle, exactly where the geometric rays would be. This result is consistent with the observations first made by Woodward (1990); then extensively discussed in the global seismology community (Marquering et al., 1999; Dahlen et al., 2000), and further analyzed at SEP by Rickett (2000).

**THEORY**

**Fréchet derivative integral kernels**

Consider a (nonlinear) function \( g \) mapping one element of the functional model space \( m \) to one element of the functional data space \( d \):

\[
d = g(m) .
\]
The tangent linear application to $g$ at point $m = m_0$ is a linear operator $G_0$ defined by the expansion

$$g(m_0 + \delta m) = g(m_0) + G_0 \delta m + \ldots,$$

(2)

where $\delta m$ is a small perturbation in the model space. The tangent linear application $G_0$ is also known under the name of Fréchet derivative of $g$ at point $m_0$ (Tarantola, 1987).

Equation (2) can be written formally as

$$\delta d = G_0 \delta m,$$

(3)

where $\delta m$ is a perturbation in the model space, and $\delta d$ is a perturbation in the image space. If we denote by $\delta d^i$ the $i^{th}$ component of $\delta d$, and by $\delta m(x)$ an infinitesimal element of $\delta m$ at location $x$, we can write

$$\delta d^i = \int_{\mathcal{V}} G^i_0(x) \, \delta m(x) \, dv(x).$$

(4)

$G^i_0$ is, by definition, the integral kernel of the Fréchet derivative $G_0$, $\mathcal{V}$ is the volume under investigation, $dv$ is a volume element of $\mathcal{V}$ and $x$ is the integration variable over $\mathcal{V}$. The sensitivity kernel, a.k.a. Fréchet derivative kernel, $G^i_0$ expresses the sensitivity of $\delta d^i$ to a perturbation of $\delta m(x)$ for an arbitrary location $x$ in the volume $\mathcal{V}$.

Sensitivity kernels occur in every inverse problem and have different meanings depending of the physical quantities involved:

- For wideband traveltime tomography (Bishop et al., 1985; Kosloff et al., 1996; Stork, 1992), $\delta d$ is represented by traveltime differences between recorded and computed traveltimes in a reference medium. The sensitivity kernels are infinitely-thin rays computed by ray tracing in the background medium.

- For finite-frequency traveltime tomography (Marquering et al., 1999; Dahlen et al., 2000; Hung et al., 2000; Rickett, 2000), $\delta d$ is represented by time shifts measured by crosscorelation between the recorded wavefield and a wavefield computed in a reference medium. The sensitivity kernels are represented by hollow fat rays (a.k.a. “banana-doughnuts”) which depend on the background medium.

- For wave-equation tomography (Woodward, 1992; Pratt, 1999), $\delta d$ is represented by perturbations between the recorded wavefield and the computed wavefield in a reference medium. The sensitivity kernels are represented by fat rays with similar forms for either the Born or Rytov approximation.

- For wave-equation migration velocity analysis (Biondi and Sava, 1999; Sava and Fomel, 2002; Sava and Biondi, 2004a,b), $\delta d$ is represented by image perturbations. The sensitivity kernels are discussed in the following sections.
Wave-equation migration velocity analysis

Wave-equation migration velocity analysis (WEMVA) is based on a linear relation established between perturbations of the slowness model $\delta s$ and perturbations of migrated images $\delta r$. $\delta d$ and $\delta r$ correspond, respectively, to $\delta m$ and $\delta d$ in equation (3).

Formally, we can write

$$\delta r = L_0 \delta s,$$

where $L_0$ is the linear first-order Born wave-equation MVA operator. The operator $L_0$ incorporates all first-order scattering and extrapolation effects for media of arbitrary complexity. The major difference between WEMVA and wave-equation tomography is that $\delta d$ is formulated in the image space for the former as opposed to the data space for the later. Thus, with WEMVA we are able to exploit the power of residual migration in perturbing migrated images – a goal which is much harder to achieve in the space of the recorded data.

By construction, the linear operator $L_0$ depends on the wavefield computed by extrapolation of the surface data using the background slowness, which corresponds to $m_0$ in equation (2). Thus, the operator $L_0$ depends directly on the type of recorded data and its frequency content, and it also depends on the background slowness model. Thus, the main elements that control the shape of the sensitivity kernels are

- the frequency content of the background wavefield,
- the type of source from which we generate the background wavefield (e.g. point source, plane wave), and
- the type of perturbation introduced in the image space, which for this problem corresponds to the data space.

In our examples, we define two types of image perturbations: a purely kinematic type $\delta r_k$, implemented simply as a derivative of the image with respect to depth, which can be implemented as a multiplication in the depth wavenumber domain (Sava and Biondi, 2004a,b) as follows:

$$\delta r_k = -i k_z r ,$$

and a purely dynamic type $\delta r_a$, implemented by scaling the reference image $r$ with an arbitrary number:

$$\delta r_a = \epsilon r .$$

In both cases, the perturbations are limited to a small portion of the image. The main difference between $\delta r_k$ and $\delta r_a$ is given by the $90^\circ$ phase-shift between the two image perturbations.
EXAMPLES

In our first example, we consider a background wavefield emerging from a fixed point source on the surface, but investigate the sensitivity kernels for various points in the image. Each panel in Figure 1 depicts a superposition of three elements: the velocity model, the band-limited wavefield corresponding to a point source on the surface, and the sensitivity kernel corresponding to a point in the subsurface.

A fundamental problem with ray-based MVA is that rays are poor approximations of the actual wavepaths when a band-limited seismic wave propagates through a rugose top of the salt. Figure 1 illustrates this issue quite clearly. It shows three sensitivity kernels for frequencies of $1 - 26$ Hz. The top panel in Figure 1 shows a wavepath that could be reasonably approximated using the method introduced by Lomax (1994) to trace fat rays using asymptotic methods. In contrast, the wavepaths shown in both the middle and bottom panels cannot be well approximated using Lomax’ method. The amplitude and shapes of these wavepaths are much more complex than a simple fattening of a geometrical ray could ever describe. The bottom panel illustrates the worst situation for ray-based tomography because the rugosity of the top of the salt has the same scale as the spatial wavelength of the seismic wave.

The fundamental reason why the true wavepaths cannot be approximated using fattened geometrical rays is that they are frequency dependent. Figure 2 illustrates this dependency by depicting the wavepath shown in the bottom panel of Figure 1 as a function of the temporal bandwidth: $1 - 5$ Hz (top), $1 - 16$ Hz (middle), and $1 - 64$ Hz (bottom). The width of the wavepath decreases as the frequency bandwidth increases, and the focusing and defocussing of the energy varies with the frequency bandwidth.

In the next example (Figures 3 and 4) we compare the shapes of sensitivity kernels when we change the type of source for the background wavefield, its frequency content and the method used to generate an image perturbation in the subsurface. As for the preceding example, we show the results as a superposition of the velocity model, the background wavefield and the sensitivity kernel from a fixed point in the subsurface.

Figure 3 shows the sensitivity kernels for a point source on the surface, and Figure 4 shows the sensitivity kernels for a plane-wave propagating vertically at the surface. In both pictures, the left column corresponds to kinematic image perturbations of equation (6), and the right column corresponds to amplitude image perturbations of equation (7) obtained by scaling of the background image by an arbitrary number. From top to bottom, we show sensitivity kernels of increasing frequency range: $1 - 4$ Hz, $1 - 8$ Hz, $1 - 16$ Hz and $1 - 32$ Hz. Once again, we can see the large frequency dependence of the sensitivity kernels. The area of sensitivity reduces with increased frequency which is a clear indication that a frequency dependent migration velocity analysis method like WEMVA can better handle subsalt environments with patchy illumination and that illumination itself is a frequency dependent phenomenon which needs to be addressed in this way.

Finally, we show wave-equation MVA sensitivity kernels for a 3D velocity model Figure 5 corresponding to a salt environment. We consider the case of a point source on the surface and data with a frequency range of $1 - 16$ Hz. Figure 6 shows the sensitivity kernel for the kinematic
Figure 1: Kinematic sensitivity kernels for frequencies between 1 and 26 Hz for various locations in the image and a point on the surface. Each panel is an overlay of three elements: the slowness model, the wavefield corresponding to a point source on the surface at $x = 16$ km, and wave paths (sensitivity kernels) from a point in the subsurface to the source.
Figure 2: Frequency dependence of kinematic sensitivity kernels between a location in the image and a point on the surface. Each panel is an overlay of three elements: the slowness model, the wavefield corresponding to a point source on the surface at $x = 16$ km, and wave paths (sensitivity kernels) from a point in the subsurface to the source. The different wave paths correspond to frequency bands of $1 - 5$ Hz (top), $1 - 16$ Hz (middle) and $1 - 64$ Hz (bottom). The larger the frequency band, the narrower the wave path. The end member for an infinitely wide frequency band corresponds to an infinitely thin geometrical ray.
Figure 3: The dependence of sensitivity kernels to frequency and image perturbation. From top to bottom, the frequency range is $1 - 4$ Hz, $1 - 8$ Hz, $1 - 16$ Hz and $1 - 32$ Hz. The left column corresponds to kinematic image perturbations, and the right column corresponds to dynamic image perturbations. The wavefield is produced from a point source. [paul1-fat2d.Tray2a] [CR,M]
Figure 4: The dependence of sensitivity kernels to frequency and image perturbation. From top to bottom, the frequency range is $1 - 4$ Hz, $1 - 8$ Hz, $1 - 16$ Hz and $1 - 32$ Hz. The left column corresponds to kinematic image perturbations, and the right column corresponds to dynamic image perturbations. The wavefield is produced by a horizontal incident plane-wave.
image perturbation, while Figure 7 for an amplitude image perturbation. In both cases, the shapes of the kernels are complicated, which is an expression of the multipathing occurring as waves propagate through rough salt bodies. The horizontal slice indicates multiple paths linking the source point on the surface with the image perturbation in the subsurface.

One noticeable characteristic is that the sensitivity kernels constructed from amplitude image perturbations show the largest sensitivity in the center of the kernel, as opposed the kinematic kernels which show the largest sensitivity away from the central path. This phenomenon was discussed by Dahlen et al. (2000) in the context of finite-frequency traveltime tomography. We illustrate it for WEMVA in Figure 8 with two horizontal slices in the sensitivity kernels shown in Figures 6 and 7.
Figure 6: 3D sensitivity kernels for wave-equation MVA. The frequency range is 1 – 16 Hz. The kernels are complicated by the multipathing occurring as waves propagate through the rough salt body. The image perturbation corresponds to a kinematic shift.
Figure 7: 3D sensitivity kernels for wave-equation MVA. The frequency range is 1 – 16 Hz. The kernels are complicated by the multipathing occurring as waves propagate through the rough salt body. The image perturbation corresponds to an amplitude scaling. [paul1-fat3.fq3] [CR]
CONCLUSIONS

We construct 2D and 3D sensitivity kernels for wave-equation MVA and study their dependence with respect to the data source, the frequency content, and the type of image perturbation. Sensitivity kernels illustrate the complications of wave propagation in the cases of rough salt bodies and the frequency-dependent illumination subsalt. With this type of analysis, we also illustrate some of the reasons why traveltime tomography is less than ideal for subsalt velocity analysis.

ACKNOWLEDGMENTS

We acknowledge BP and ExxonMobil for the 3D velocity model.

REFERENCES


Diffraction-focusing migration velocity analysis with application to seismic and GPR data

Paul Sava, Biondo Biondi, and John Etgen

ABSTRACT

We propose a method for estimating interval velocity using the kinematic information in diffractions. We extract velocity information from migrated diffracted events by analyzing their residual focusing in physical space (depth and midpoint) using prestack residual migration. The results of this residual-focusing analysis are fed to a linearized inversion procedure that produces interval velocity updates. Our inversion procedure employs a wavefield-continuation operator linking perturbations of interval velocities to perturbations of migrated images, based on the principles of Wave Equation Migration Velocity Analysis (WEMVA) introduced in recent years. We measure the accuracy of the migration velocity by using a diffraction-focusing criterion, instead of the criterion of flatness of migrated common-image gathers that is commonly employed in Migration Velocity Analysis (MVA). This new criterion enables us to extract velocity information from events that would be challenging to use with conventional MVA methods, and thus makes our method a powerful complement to conventional MVA methods. We demonstrate our method with synthetic and real Ground-Penetrating Radar data.

INTRODUCTION

Migration velocity analysis (MVA) using diffracted events is not a new concept. Harlan (1986) addresses this problem and proposes methods to isolate diffraction events around faults, quantifies focusing using statistical tools, and introduces MVA techniques applicable to simple geology, e.g. constant velocity or $v(z)$. Similarly, de Vries and Berkhout (1984) use the concept of minimum entropy to evaluate diffraction focusing and apply this methodology to MVA, again for the case of simple geology. Soellner and Yang (2002) estimate interval velocities from focusing of diffractions simulated using data-derived parameters.

Sava and Biondi (2004a,b) introduce a method of migration velocity analysis using wave-equation techniques (WEMVA), which aims to improve the quality of migrated images, mainly by correcting moveout inaccuracies of specular energy. WEMVA finds a slowness perturbation which corresponds to an image perturbation. Thus, it is similar to ray-based migration tomography (Al-Yahya, 1989; Stork, 1992; Etgen, 1993), where the slowness perturbation is derived from depth errors, and to wave-equation tomography (Tarantola, 1986; Woodward, 1982).
1992; Pratt, 1999; Dahlen et al., 2000) where the slowness perturbation is derived from measured wavefield perturbations.

The moveout information given by the specular energy is not the only information contained by an image migrated with the incorrect slowness. Non-specular diffracted energy is present in the image and clearly indicates slowness inaccuracies. Traveltime-based MVA methods cannot easily deal with the diffraction energy, and are mostly concerned with moveout analysis. In contrast, a difference between an inaccurate image and a perfectly focused target image contains both specular and non-specular energy; therefore WEMVA is naturally able to derive velocity updates based on both these types of information. Our proposed method can benefit, and thus be used in conjunction with, methods to isolate diffracted energy from the seismic data, such as the one proposed by Khaidukov et al. (2004).

In this paper, we use WEMVA to estimate slowness updates based on focusing of diffracted energy using residual migration. One possible application of this technique in seismic imaging concerns areas with abundant, clearly identifiable diffractions. Examples include highly fractured reservoirs, carbonate reservoirs, rough salt bodies and reservoirs with complicated stratigraphic features. Another application is related to imaging of zero-offset Ground-Penetrating Radar (GPR) data, where moveout analysis is simply not an option.

Of particular interest is the case of salt bodies. Diffractions can help estimate more accurate velocities at the top of the salt, particularly in the cases of rough salt bodies. Moreover, diffraction energy may be the most sensitive velocity information we have from under the salt, since most of the reflected energy we record at the surface has only a narrow range of angles of incidence at the reflector, rendering the analysis of moveout ambiguous.

We begin with a summary of the wave-equation MVA methodology, specialized to diffraction focusing, followed by synthetic and real-data examples from seismic and GPR applications.

**WEMVA THEORY**

In this section, we summarize the main elements of wave-equation migration velocity analysis, by closely following the theory presented in Sava and Biondi (2004a,b). The reader familiar with those details can safely skip to the next section.

We begin with a quick discussion of wavefield scattering in the context of wavefield extrapolation and then define the objective function of our method and the linearized image perturbations which enable us to overcome the limitations of the first-order Born approximation.
Wavefield scattering

Imaging by wavefield extrapolation (WE) is based on recursive continuation of the wavefields \( \mathcal{U} \) from a given depth level to the next, by means of an extrapolation operator \( \mathcal{E} \):

\[
\mathcal{U}_{z+\Delta z} = \mathcal{E}_z \left[ \mathcal{U}_z \right].
\]  

(1)

Here and hereafter, we use the following notation conventions: \( A[x] \) stands for the linear operator \( A \) applied to \( x \), and \( f(x) \) stands for function \( f \) of argument \( x \).

At any depth \( z \), the wavefield (\( \tilde{\mathcal{U}} \)), extrapolated through the background medium characterized by the background velocity (\( \tilde{s} \)), interacts with medium perturbations (\( \Delta s \)) and creates wavefield perturbations (\( \Delta \mathcal{V} \)):

\[
\Delta \mathcal{V}_z = \mathcal{S}_z \left( \tilde{\mathcal{U}}_z \right) \left[ \Delta s_z \right].
\]  

(2)

\( \mathcal{S} \) is a scattering operator relating slowness perturbations to wavefield perturbations. The total wavefield perturbation at depth \( z + \Delta z \) is the sum of the perturbation accumulated up to depth \( z \) from all depths above (\( \Delta \mathcal{U}_z \)), plus the scattered wavefield from depth (\( \Delta \mathcal{V}_z \)) extrapolated one depth step (\( \Delta z \)):

\[
\Delta \mathcal{U}_{z+\Delta z} = \mathcal{E}_z \left[ \Delta \mathcal{U}_z \right] + \mathcal{E}_z \left[ \mathcal{S}_z \left( \tilde{\mathcal{U}}_z \right) \left[ \Delta s_z \right] \right].
\]  

(3)

We can use the recursive equation (3) to compute a wavefield perturbation, given a pre-computed background wavefield and a slowness perturbation. In a more compact notation, we can write equation (3) as follows:

\[
\Delta \mathcal{U} = (1 - \mathcal{E})^{-1} \mathcal{E} \mathcal{S} \Delta s ,
\]  

(4)

where \( \Delta \mathcal{U} \) and \( \Delta s \) stand respectively for the wavefield and slowness perturbations at all depth levels, and \( E, S \) and \( I \) are respectively the wavefield extrapolation operator, the scattering operator and the identity operator. In our current implementation, \( \mathcal{S} \) refers to a first-order Born scattering operator.

From the wavefield perturbation (\( \Delta \mathcal{U} \)), we can compute an image perturbation (\( \Delta \mathcal{R} \)) by applying an imaging condition, \( \Delta \mathcal{R} = I \Delta \mathcal{U} \). For example, the imaging operator, (I) can be a simple summation over frequencies. If we accumulate all scattering, extrapolation and imaging into a single operator \( \mathcal{L} = I (1 - \mathcal{E})^{-1} \mathcal{E} \mathcal{S} \), we can write a simple linear expression relating an image perturbation (\( \Delta \mathcal{R} \)) to a slowness perturbation (\( \Delta s \)):

\[
\Delta \mathcal{R} = \mathcal{L} \Delta s .
\]  

(5)

For wave-equation migration velocity analysis, we use equation (5) to estimate a perturbation of the slowness model from a perturbation of the migrated image by minimizing the objective function

\[
J (\Delta s) = ||W (\Delta \mathcal{R} - \mathcal{L} \Delta s)||^2 ,
\]  

(6)
where $W$ is a weighting operator related to the inverse of the data covariance, indicating the reliability of the data residuals. Since, in most practical cases, the inversion problem is not well conditioned, we need to add constraints on the slowness model via a regularization operator. In these situations, we use the modified objective function

$$J(\Delta s) = ||W(\Delta R - L \Delta s)||^2 + \epsilon^2 ||A \Delta s||^2 .$$

(7)

$A$ can be a regularization operator which penalizes rough features of the model, and $\epsilon$ is a scalar parameter which balances the relative importance of the data residual, $W(\Delta R - L \Delta s)$, and the model residual, $(A \Delta s)$.

An essential element of our velocity analysis method is the image perturbation, $\Delta R$. For the purposes of the optimization problem in equation (7), this object is known and has to be precomputed, together with the background wavefield used by the operator $L$. In the next section, we discuss how $\Delta R$ is estimated in practice.

**Image perturbations**

A simple way to define the image perturbation ($\Delta R$) is to take the image obtained with the background slowness and improve it by applying an image enhancement operator. There are many techniques that can be used to obtain an enhanced image.

- One possibility is to flatten events in angle-domain common-image gathers (ADCIG) by using residual moveout.
- Another possibility is to use residual migration to flatten ADCIGs and, at the same time, focus diffractions which can be observed in common-offset sections.

In principle, both focusing in space (along the midpoint axis) and focusing in offset are velocity indicators, and they should be used together to achieve the highest accuracy. In this paper, however, we emphasize migration velocity analysis using only focusing of diffractions along the spatial axes.

In our current implementation, we use prestack Stolt residual migration (Stolt, 1996; Sava, 2003) as the image enhancement operator ($K$). This residual migration operator applied to the background image creates new images ($R$), functions of a scalar parameter ($\rho$), which represents the ratio of a new slowness model relative to the background one:

$$R = K(\rho) [R_b] .$$

(8)

We can now take the image perturbation to be the difference between the improved image ($R$) and the background image ($R_b$):

$$\Delta R = R - R_b .$$

(9)

The main challenge with this method of constructing image perturbations for WEMVA is that the two images, $R$ and $R_b$, can get out of phase, such that they risk violating the requirements
of the first-order Born approximation (Sava and Biondi, 2004a). For example, we might end up subtracting unfocused from focused diffractions at different locations in the image.

We address this challenge by using *linearized image perturbations*. We run residual migration for a large number of parameters \( \rho \) and pick at every location the value where the image is best focused. Then we estimate at every point the gradient of the image relative to the \( \rho \) parameter and construct the image perturbations using the following relation:

\[
\Delta R \approx K\left[ \rho = 1 \right] [R_b] \Delta \rho ,
\]

(10)

where, by definition, \( \Delta \rho = 1 - \rho \).

The main benefit of constructing image perturbations with equation (10) is that we avoid the danger of subtracting images that are out of phase. In fact, we do not subtract images at all, but we simply construct the image perturbation that corresponds to a particular map of residual migration parameters (\( \rho \)). In this way, we honor the information from residual migration, but we are safe relative to the limits of the first-order Born approximation.

Figures 1 and 2 illustrate the migration velocity analysis methodology using residual migration and linearized image perturbations. Figure 1 shows three simple models with diffractors and reflectors with a constant velocity \( v = 2000 \text{ m/s} \). We use these three models to illustrate different situations: an isolated diffractor at location \( x = 2000 \text{ m} \) and depth \( z = 900 \text{ m} \), (Figure 1, left), the same diffractor flanked by other diffractors at \( z = 1100 \text{ m} \) (Figure 1, middle), and finally the same diffractor next to a short reflector at \( z = 1100 \text{ m} \) (Figure 1, right).

We migrate each synthetic datum with an incorrect velocity, \( v = 1800 \text{ m/s} \), and then run residual migration with various velocity ratios from \( \rho = 0.7 \) to \( \rho = 1.1 \). From top to bottom, each row corresponds to a different velocity ratio as follows: 0.7, 0.8, 0.9, 1.0, 1.1. For all residual migration examples, we have eliminated the vertical shift induced by the different velocities, such that only the diffraction component of residual migration is left. Thus, we can better compare focusing of various events without being distracted by their vertical movement.

The images at \( \rho = 0.9 \) are the best focused images. Since both the backgrounds and the perturbations are constant, the images focus at a single ratio parameter. The ratio difference between the original images at \( \rho = 1.0 \) and the best focused images at \( \rho = 0.9 \) is \( \Delta \rho = 0.1 \). In general, the images focus at different ratios at different locations; therefore \( \Delta \rho \) is a spatially variable function.

Using the background images and the measured \( \Delta \rho \), we compute the linearized image perturbations (Figure 2, top), and using the WEMVA operator we compute the corresponding slowness perturbations after 15 linear iterations (Figure 2, bottom). The image perturbations closely resemble the background image (Figure 1, fourth row from top), with a \( \pi/2 \) phase shift and appropriate scaling with the measured \( \Delta \rho \).

For all models in Figures 1 and 2, we measure focusing on a single event (the main diffractor at \( x = 2000 \text{ m} \) and depth \( z = 900 \text{ m} \)), but assign the computed \( \Delta \rho \) to other elements of the image in the vicinity of this diffractor. The rationale for doing so is that we can assume that all elements at a particular location are influenced by roughly the same part of the model.
Figure 1: Residual migration applied to three simple synthetic models (from left to right). From top to bottom, the images correspond to the ratios $\rho = 0.7, 0.8, 0.9, 1.0, 1.1$. The middle row corresponds to the correct velocity, when all diffractors are focused.
Therefore, not only is a priori separation of the diffractors from the reflectors not required, but the additional elements present in the image perturbation add robustness to the inversion.

Figure 2: Migration velocity analysis for the three simple synthetic models in Figure 1. The top row depicts image perturbations, and the bottom row depicts slowness perturbations obtained after 15 linear iterations. [paul2-DIFLAmva][CR]

EXAMPLES

The first example concerns a synthetic dataset obtained by acoustic finite-difference modeling over a salt body. Although, in this example, we use our technique to constrain the top of the salt, we would like to emphasize that we can use the same technique in any situation where diffractions are available. For example, in sub-salt regions where angular coverage is small, uncollapsed diffractions carry substantial information which is disregarded in typical MVA methodologies.

The second example is a real dataset of single-channel, Ground-Penetrating Radar (GPR) data. Many GPR datasets are single-channel and no method has thus far been developed to estimate a reasonable interval velocity models in the presence of lateral velocity variations. Typically, the velocity estimated by Dix inversion at sparse locations along the survey line is smoothly extrapolated, although this is not even close to optimal from an imaging point of view.

Delineation of rough salt bodies

Figure 3 shows the zero-offset data we use for velocity analysis to delineate the top of the rough salt body. The section contains a large number of diffractors, whose focusing allows us to constrain the overburden velocity model.

Figure 4(a) depicts the starting velocity model, and Figure 4(b) depicts the initial image obtained by zero-offset migration. The starting velocity is a typical Gulf of Mexico \( v(z) \) function hanging from the water bottom. Uncollapsed diffractions are visible at the top of the salt, indicating that the velocity in the overburden is not accurate. Such defocusing also hampers our ability to pick accurately the top of the salt and, therefore, degrades imaging at depth.
As we did for the preceding synthetic example, we run residual migration on the background image (Figure 4). Figure 5 shows this image after residual migration with various velocity ratios (Sava, 2003). From top to bottom, the ratios are: 1.04, 1.00, 0.96, 0.92, 0.88. At $\rho = 1.00$ we recover the initial image. Different parts of the image come into focus at different values of the velocity ratio.

Figure 6(a) shows the picked velocity ratios at various locations in the image. The white background corresponds to picked $\Delta \rho = 0$, and the gray shades correspond to $\Delta \rho$ between 0 and 0.08. Figure 6(b) shows a map of the weights ($W$) associated to each picked value. The white background corresponds to $W = 0$, indicating low confidence in the picked values, and the dark regions correspond to $W = 1$, indicating high confidence in the picked values. In this example, we disregard regions where we did not pick any diffractions. All other regions receive an arbitrary ratio value ($\rho = 1.0$), but also a low weight such that they do not contribute to the inversion. Exceptions include the water bottom, for which we assign a high weight of the picked ratio $\rho = 1.0$, and a few other reflectors for which we did not have any diffraction focusing information.

Figure 7(a) shows the slowness perturbation obtained after 20 iterations of zero-offset
Figure 4: Zero-offset migrated image for the synthetic data in Figure 3: velocity model (a), and migrated image (b). Migration using the initial $v(z)$ velocity model. [paul2-BPATimg1] [CR.M]
Figure 5: Residual migration applied to the image migrated with the initial velocity model, Figure 4. From top to bottom, the images correspond to the ratios $\rho = 1.04, 1.00, 0.96, 0.92, 0.88$. [paul2-BPAITsrm] [CR,M]
Figure 6: Residual migration picks (a) and the associated confidence weights (b).

paul2-BPAITpck [CR,M]
inversion from the image perturbation in Figure 7(b). The image perturbation is non-zero only in the regions where we had diffractions we could pick, as indicated by Figure 6. The smooth slowness perturbation is further constrained by the regularization operator we use, which is a simple Laplacian penalizing the rough portions of the model.

Figure 8(a) shows the updated slowness model and Figure 8(b) shows the zero-offset migrated image corresponding to the updated model. Most of the diffractions at the top of the salt have been collapsed, and the rough top of the salt can be easily picked. The diffractions corresponding to the salt bodies at $x = 2000 - 4000$ ft, $z = 3500$ ft are not fully collapsed, indicating that another nonlinear iteration involving residual migration and picking might be necessary.

Finally, Figure 9 shows prestack migrated images using the initial velocity model (a) and the one updated using zero-offset focusing (b). The top panels depict stacks, and the bottom panels depict angle-domain common-image gathers (ADCIG) (Sava and Fomel, 2003). The ADCIGs show substantial bending after migration with the initial velocity, but they are mostly flat after migration with the updated velocity, although none of the moveout information has been used for velocity update. Figure 10 shows two ADCIGs at $x = -2350$ ft from the images obtained with the initial velocity model (a) and the updated velocity model (b). The ADCIG in panel (a) corresponds to a notch in the top of the salt and is complicated to use for velocity analysis. However, after migration with the updated velocity model, panel (b), the ADCIG is much simpler, and the small residual moveouts can be picked for velocity updates.

A comparison of Figure 8(b) with Figure 9(b) shows a potential limitation of our technique in the presence of prismatic waves (Biondi, 2003). Both images are obtained with the same velocity, the first one with zero-offset data and the second one with prestack data. The imaging artifacts visible at the bottom of the deep canyons at the top of the salt in Figure 8(b) are created by prismatic waves that are not properly imaged from zero-offset data. Prismatic waves are better (though not perfectly) handled by full prestack migration, and thus the artifacts are not visible in the prestack-migrated image shown in Figure 9(b). Since these artifacts resemble uncollapsed diffractions, they may mislead the analysis of the residual migrated images and be interpreted as symptoms of velocity inaccuracies.

**Imaging of GPR data**

Our next example concerns a zero-offset GPR dataset over a lava flow region. In this situation, diffraction focusing is the only option available for migration velocity analysis. The data depicted in Figure 11 show many diffractions spread over the entire dataset. A few obvious ones are at $x = 22$ ft, $t = 27$ ns, at $x = 28$ ft, $t = 22$ ns, and at $x = 35$ ft, $t = 23$ ns.

We follow the same procedure for migration velocity analysis as the one described for the preceding example. Figure 12(a) shows the initial image obtained by migration with a constant velocity of 0.2 ft/ns, and Figure 12(b) shows the final image obtained after velocity update. We can notice that the image has been vertically compressed, since the velocity update indicated a faster velocity, and most of the diffractions have been collapsed.
Figure 7: Slowness perturbation (a), derived from an image perturbation (b) derived from the background image in Figure 4 and the velocity ratio picks in Figure 6. [CR,M]
Figure 8: Zero-offset migrated image for the synthetic data in Figure 3: velocity model (a), and migrated image (b). Migration using the updated velocity. [paul2-BPAITimg2] [CR,M]
Figure 9: Prestack migrated images using the initial velocity model (a) and the updated velocity model (b). The top panels depict image stacks and the bottom panels depict angle-domain common image gathers.

Figure 10: Angle-domain common image gather obtained after migration with the initial velocity model (a) and the updated velocity model (b).
Figure 11: Zero-offset GPR data used for focusing migration velocity analysis.
Figure 12: Zero-offset migrated images for the data in Figure 11 using the initial velocity (a) and the updated velocity (b).
Figures 13 and 14 are detailed views of the initial and final images and slownesses at various locations of interest. Figure 13 shows collapsed diffractions in the left part of the image. We can also observe features with better continuity in the updated image than in the original image, for example at $x = 20 - 24$ ft and $z = 2$ ft in Figure 13(a,c). Likewise, Figure 14 shows a better focused image than in the original, for example at $x = 34$ ft and $z = 1.8$ ft in Figure 14(a,c).

![Figure 13: Detail of the images depicted in Figure 12. Migration with the initial velocity (a), updated slowness model (b) and migration with the updated slowness (c). The window corresponds to $x = 20 - 24$ ft and $z = 2$ ft.](image1)

![Figure 14: Detail of the images depicted in Figure 12. Migration with the initial velocity (a), updated slowness model (b) and migration with the updated slowness (c). The window corresponds to $x = 34$ ft and $z = 1.8$ ft.](image2)
CONCLUSIONS

Diffracted events contain useful velocity information that is overlooked by conventional MVA methods, which use flatness of common image gathers as the only criterion for the accuracy of migration velocity. In this paper, we demonstrate that accurate interval-velocity updates can be estimated by inverting the results of a residual-focusing analysis of migrated diffracted events. To convert residual-focusing measurements into interval-velocity updates, we employ the WEMVA methodology (Biondi and Sava, 1999; Sava and Fomel, 2002; Sava and Biondi, 2004a,b). Our WEMVA methodology is ideally suited for this task because it is capable of inverting image perturbations directly, without requiring an estimate of the reflector geometry. In contrast, ray-based MVA methods require the reflector geometry to be provided by interpreting the migrated image. However, since the interpretation of partially-focused diffracted events is an extremely difficult task, ray-based methods are never employed for diffraction-focusing velocity analysis.

Our seismic-data example demonstrates how the proposed method can exploit the velocity information contained in the event generated by a rugose salt-sediment interface. This kind of event is present in many salt-related data sets, and the ability of using the diffracted energy to further constrain the velocity model might significantly improve the final imaging results.

The GPR-data example demonstrates the significant potential of our method for improving the imaging of GPR data. We demonstrate that the interval-velocity model obtained by extracting velocity information from the diffracted events improves the reflector continuity in the migrated image and facilitates geological interpretation of the images. Since a large number of GPR data sets are limited to zero-offset data, the possibility of using diffractions to define the lateral variations in interval velocity can substantially widen the range of applications of GPR methods.

ACKNOWLEDGMENTS

We would also like to acknowledge Frederic Billette (BP) for providing the synthetic salt-dome dataset and Antoine Guitton (SEP) for suppressing multiples in this dataset. Kevin Williams of the Smithsonian Institution in Washington, D.C. provided the GPR dataset.

REFERENCES


Velocity uncertainty in tomography

Robert G. Clapp

ABSTRACT

The conventional method for migration velocity analysis scans over one (or possibly multiple) move-out parameter(s) along the offset or angle axis. The appropriate move-out parameter is then selected based on what produces the best coherence. The coherence as a function of the scanning parameter provides important information on how confident we are of our given move-out measure. In this paper, the move-out parameter selection is set up as an inverse problem and multiple, equi-probable move-out fields are generated. These various realization of move-out are then used to update the velocity model and migrated image. Early results are promising.

INTRODUCTION

Risk assessment is a key component to any business decision. Geostatistics has recognized this need and has introduced methods such as simulation to attempt to assess uncertainty in their estimates of earth properties (Isaaks and Srivastava, 1989). The problem is that the geostatistical methods are generally concerned with local, rather than global, solutions to problems and therefore can not be easily applied to the global inversion problems that are common in geophysics.

In previous works (Clapp, 2000, 2001b,c), I showed how we can modify standard geophysical inverse techniques by adding random noise into the model styling goal to obtain multiple realizations. In Clapp (2002a) and Chen and Clapp (2002), these multiple realizations were used to produce a series of equi-probable velocity models. The velocity models were used in a series of migrations, and the effect on Amplitude vs. Angle (AVA) was analyzed. In Clapp (2003a), I showed how the data fitting goal can be modified to account for data uncertainty. I demonstrated its applicability to a simple Super Dix (Clapp et al., 1998) problem.

In this paper I extend the data uncertainty concept to a ray based tomography problem (Stork and Clayton, 1991; Clapp, 2001b). I show how we can produce multiple, realistic velocity models. On a complex synthetic I produce a range of equi-probable velocity models, for migration to produce kinetically different images.

\[email: \text{bob@sep.stanford.edu}\]

233
MULTIPLE REALIZATIONS REVIEW

Inverse problems obtain an estimate of a model \( m \), given some data \( d \) and an operator \( L \) relating the two. We can write our estimate of the model as minimizing the objective function in a least-squares sense,

\[
f(m) = \| d - Lm \|^2. \tag{1}
\]

We can think of this same minimization in terms of fitting goals as

\[
0 \approx r = d - Lm, \tag{2}
\]

where \( r \) is a residual vector.

Bayesian theory tells us (Tarantola, 1987) that convergence rate and the final quality of the model is improved the closer \( r \) is to being Independent Identically Distributed (IID). If we include the inverse noise covariance \( \mathbf{N} \) in our inversion our data residual becomes IID,

\[
0 \approx r = \mathbf{N}(d - Lm). \tag{3}
\]

A regularized inversion problem can be thought of as a more complicated version of (3) with an expanded data vector and an additional covariance operator,

\[
0 \approx r_d = \mathbf{N}_{\text{noise}}(d - Lm) \quad 0 \approx r_m = \epsilon \mathbf{N}_{\text{model}}(0 - \mathbf{I}m). \tag{4}
\]

In this new formulation, the first expression is the “data fitting goal” and the second is the “model styling goal”, \( r_d \) is the residual from the data fitting goal, \( r_m \) is the residual from the model styling goal, \( \mathbf{N}_{\text{noise}} \) is the inverse noise covariance, \( \mathbf{N}_{\text{model}} \) is the inverse model covariance, \( \mathbf{I} \) is the identity matrix, and \( \epsilon \) is a scalar that balances the fitting goals against each other. Normally we think of \( \mathbf{N}_{\text{model}} \) as the regularization operator \( A \). Simple linear algebra leads to a more standard set of fitting goals:

\[
0 \approx r_d = \mathbf{N}_{\text{noise}}(d - Lm) \quad 0 \approx r_m = \epsilon A m. \tag{5}
\]

A problem with this approach is that we never know the true inverse noise or model covariance and therefore are only capable of applying approximate forms of these matrices.

One way to approximate \( \mathbf{N}_{\text{noise}} \) is to think of it as a chain of two operators (Clapp, 2003a). One operator will describe the two-point covariance of the matrix \( \mathbf{N}_{\text{n,c}} \) and one operator that describes the variance \( \mathbf{N}_{\text{n,v}} \). We have a number of options in designing \( \mathbf{N}_{\text{n,c}} \). We can use a Laplacian or some type of symmetric operator, a stationary Prediction Error Filter (Claerbout, 1999), a steering filter (Clapp, 2001b), or a non-stationary PEF (NSPEF) (Crawley, 2000). What we use is based on our \textit{a priori} knowledge of our noise. The variance component \( \mathbf{N}_{\text{n,v}} \) can be thought of simply in terms on how reliable we consider a given component of our data.
A simple example of this is the Super Dix (Clapp and Biondi, 1999) problem where $N_{n,v}$ is constructed from stack power.

Another problem with fitting goals (5) is that we produce a single answer, with no information on the variability of different components in the model space. Our single answer is the minimum energy solution. In (Clapp, 2001b) I showed how, for the interpolation problem, we can produce a range of equi-probable solutions by replacing $r_m$ with a random noise vector $n$. The resulting models all had a more realistic texture than the minimum energy solution because the regularization operator did not fully describe the inverse model covariance.

In (Clapp, 2003a) I showed, how by replacing $r_d$ with a random noise vector, we could produce a range of equi-probably interval velocity estimates for the Super Dix problem. The Super Dix example was a 1-D problem. The $N_{n,c}$ was a derivative operator. The $N_{n,v}$ operator was constructed based on the semblance scan. The variance was based on how quickly the semblance fell off from the peak value used to construct the data. The resulting models provided a fuller description of the potential interval velocity models but the 1-D nature of the problem limited its usefulness. A more interesting implementation of the methodology is the migration velocity analysis problem.

TOMOGRAPHY REVIEW

Tomography for migration velocity analysis is a non-linear problem that we linearize around an initial slowness model. In this discussion, I will be talking about the specific case of ray based tomography but most of the discussion is valid for other tomographic operators. We can linearize the problem around an initial slowness model and obtain a linear relation $T$ between the change in travel times $\Delta t$ and change in slowness $\Delta s$:

$$\Delta t \approx T \Delta s. \tag{6}$$

Certain components of the velocity model are better determined than others. As a result we need to either choose an intelligent model parameterization (Cox et al., 2003) or regularize the model. I prefer the regularizing approach so that I can easily incorporate a priori information. I impose a regularization operator that tends to smooth the velocity according to reflector dip (Clapp, 2001b). My resulting fitting goals then become

$$\Delta t \approx T \Delta s \quad 0 \approx A(s_0 + \Delta s). \tag{7}$$

where $A$ is a steering filter. In this paper I will do downward continuation based migration (Stoffa et al., 1990; Ristow and Ruhl, 1994) and then constructing angle gathers (Prucha et al., 1999; Sava and Fomel, 2000). Biondi and Symes (2003) and Biondi and Tisserant (2004) showed that the travel time error is related to the reflector dip $\alpha$, local slowness $s_0$, depth $z_0$, and the velocity model scaled by $\gamma$ through

$$\Delta t_{RMO} = \frac{1 - \rho}{\cos \alpha} \frac{\sin^2 \gamma}{(\cos^2 \alpha - \sin^2 \gamma) s_0 z_0}. \tag{8}$$
One of the best ways to scan over $\gamma$ is by doing Stolt Residual Migration (SRM) (Sava, 1999b,a). In Clapp (2002b) I showed how we can obtain better ray-based tomography results by doing SRM scanning versus simple vertical move-out analysis.

**THEORY**

My goal in this paper is to produce multiple, equi-probable velocity models. A straightforward implementation of the approach described above, just accounting for data uncertainty, would lead to the fitting goals,

$$
0 \approx n = N_{n,v}N_{n,c}(\Delta t - T\Delta s)
$$

$$
0 \approx r_m = \epsilon A(s_o + \Delta s). \tag{9}
$$

Implementing these fitting goals is problematic because of the space that $\Delta t$ reside in. Unlike the Super Dix problem of Clapp (2003a), our data space is not regular. Normally our $\Delta t$ values lie along a series of reflectors or a semi-random set of points (Clapp, 2001a). In either case, constructing $N_{n,c}$ is problematic. When selecting points along reflectors, we are limited to a covariance description only along the reflectors with no easy way to describe continuity between reflectors. If we choose the random point methodology, we are limited to simply minimizing differences of nearby points, an unsatisfying option.

One solution to this problem is to introduce a mapping operator that maps $\Delta t$ from an irregular space to a regular space. This solution holds promise, but the interaction between $N_{n,v}$ and $N_{n,c}$ becomes confusing.

Another option is to move the data variability problem. As mentioned earlier, our data isn’t actually travel time differences, but $\gamma$ values calculated by doing SRM. Normally we choose the $\gamma$ value corresponding to the maximum semblance at a given location or some smooth version of the maximum (Clapp, 2003b). The selecting of the $\gamma$ value is really where our data uncertainty problem lies. The selection problem has some convenient and some not so convenient properties. On the positive side, we are working with a regular grid and we know that we want some consistency along reflectors. As a result, a steering filter becomes a very obvious choice for our covariance description. On the negative side, the selection problem shares all of the non-linear aspects of the semblance problem (Toldi, 1985).

To get around these issues I decides to borrow something from both the geostatistics world and the geophysics world. Instead of thinking of the problem in terms of selecting the best $\gamma$ value, I am going to think of the problem in terms of selecting a value within a distribution. I am going to construct my distributions in a similar manner to Rothman (1985). Rothman (1985) was trying to solve the non-linear residual statics problem using simulated annealing. He built a distribution based on stack power values from static-shift traces based on the surface locations of the sources and receivers. In this case, my distribution is going to be constructed based on the semblance values at given $\gamma$ values.

I do not want the rough solution that Rothman (1985) was looking for, instead I am looking
for a smooth solution. If I set up the inverse problem

\[ \mathbf{n} \approx \mathbf{c} \]
\[ \mathbf{0} \approx \epsilon \mathbf{A} \mathbf{c}, \]

(10)

I will get a vector \( \mathbf{c} \) that contains random numbers that have been colored according to the spectrum of \( \mathbf{A} \). I now have a field that has the spectrum I want but the vector \( \mathbf{c} \) will tend to have a normal distribution with zero mean. At this stage, I am going to borrow something from the geostatistical community. When dealing with problems where the variable does not have a normal distribution, they use what they refer to as a normal-score transform (Isaaks and Srivastava, 1989; Deutsch and A.G., 1992). They build an operator \( \mathbf{G} \) that relates a value \( a \) in an arbitrary distribution to a value \( c \) in standard normal distribution, \( a = \mathbf{G}c \), based on the cumulative distribution function (cdf) of both distributions. I am going to apply the same trick. I will solve for all values of the variable \( \mathbf{c} \) simultaneously and then apply \( \mathbf{G} \) to convert these values to \( \gamma \) values.

From theory to practice

I found that squaring the semblance values \( \text{semb}(\gamma) \) produced a better result than the semblance values themselves. My resulting cdf \( a_{\text{cdf}}(\gamma) \) function was then

\[ a_{\text{cdf}}(\gamma) = \frac{\sum_{\gamma_0}^{\gamma_1} \text{semb}^2(\gamma)}{\sum_{\gamma_0}^{\gamma_1} \text{semb}^2(\gamma)} \]  

(11)

In addition, I need to account for the fact that I am not scanning over all possible \( \gamma \) values. I introduced a constant parameter \( g_{\max} \) that scaled my \( \mathbf{c} \) values so that one standard deviation of \( \mathbf{c} \) might correspond to 2-5 standard deviations \( \mathbf{a} \).

Finally, we are going to a precondition the model (Fomel et al., 1997) with the inverse of \( \mathbf{A} \) and solve the problems in terms of the variable \( \mathbf{p} = \mathbf{A} \mathbf{m} \),

\[ \mathbf{n} \approx \mathbf{A}^{-1} \mathbf{p} \]
\[ \mathbf{0} \approx \epsilon \mathbf{p}. \]

(12)

Limitations

The method described above isn’t quite what we would like to have. Ideally we would rather phrase the problem as

\[ \mathbf{n} \approx \mathbf{G} \mathbf{A}^{-1} \mathbf{p} \]
\[ \mathbf{0} \approx \epsilon \mathbf{p}, \]

(13)

where \( \mathbf{G} \) transforms our residual from the arbitrary distribution indicated by our semblance scan to a standard normal distribution. In this way we would be smoothing our semblance values rather than residual space shaping as indicated by fitting goals (12). The problem is that \( \mathbf{G} \) can not be described in terms of a linear operator. If \( \mathbf{G} \) was a stationary function we also wouldn’t have a problem.
Figure 1: A 2-D synthetic. A realistic reservoir bounded by a faulted anticline and a basement rock. The left panel shows the velocity, the panel is the result of migrating with the correct velocity.

EXAMPLE

To test the methodology, I started with a fairly complex 2-D synthetic. The synthetic, shown in Figure 1, contains a realistic reservoir bounded by a faulted anticline and a basement rock. The velocity generally follows structure but with some low spatial frequency anomalies that range up to 5% of the background velocity. The synthetic reservoir is based on a real North Sea reservoir. The velocity and density structures vary significantly within the reservoir. The overburden was designed to break most conventional model characterization schemes. A layer-based approach would have difficulty with the anomalies and a gridded approach would find the sharp contrasts introduced by the faulting troublesome.

As an initial velocity estimate, I smoothed significantly the correct model. Figure 2 shows the initial velocity and the resulting migration. The basement reflectors are no longer flat and the anticline structure is more compressed in shape. The common reflection point gathers, five of which are shown in Figure 3, show significant move-out.

I then performed SRM semblance analysis on the dataset using $\gamma$ ranging from .7 to 1.3. Figure 4 shows the semblance scans for the same five locations as in Figure 3. I used the initial migration image to construct a steering filter for the colored random number generation.
Figure 2: The left panel shows the initial velocity model. The right panel shows the resulting of migrating with this velocity model. [bob3-initial] [CR,M]

Figure 5 shows nine realizations of applying fitting goals (12). Note how the generated random numbers follow the image structure shown in Figure 2, but vary dramatically from realization to realization. I converted these random numbers to $\gamma$ values using cdfs generated at every location using equation (11). Thirty realizations of $\gamma$ values overlay the semblance scans in Figure 4. The various realizations are smooth as function of depth, which is reasonable. The amount of variance also seems reasonable. Note how when we have a sharp semblance maximum we see almost no variation, while when our blob is more spread out, the selected gamma values are more diverse. Note the fourth from the left gather at 1.8 seconds. Two different maxima are present in the semblance gather and we see that both are picked by various realizations. Figure 6 shows nine gamma maps as a function of space. The $\gamma$ values are white when $\gamma = 1$ and become more red has $\gamma$ becomes lower, and more blue as $\gamma$ increases. Note how we see a spatial consistency within a panel but differences between the panels. The top-right portion of the various realizations is especially interesting. We see $\gamma$ values greater than one, less than one, and some mix of both are selected by the various realizations.

The $\gamma$ value were then used to update the velocity model using a preconditioned version of fitting goals (7). At this stage we are so far away from the correct answer that I limited the tomography to selecting back projection shallower than 1.7 km. Figure 7 shows the selected back projection points overlaid upon the migrated model. Figure 8 shows nine realizations of the tomography problem using the $\gamma$ values shown in Figure 6. Note the variance in the
velocity structure from realization to realization. It is especially noticeable at the fault through the anticline and to the left the fault. We see velocity variations of 1\,km/s or greater between the various realizations. This isn’t surprising given the variance present in our semblance scans (Figure 4).

As a final step, I migrated the data with the various velocity models. Figure 9 shows the nine images corresponding to the nine velocity models shown in Figure 8. Note the differences in the various images. The top-right image shows a much flatter basement reflector than any of the other realizations. The variation in positioning of the fault reflector is also significant between the various realizations.

**CONCLUSION**

The effect of velocity uncertainty on reflector positioning is assessed. Random numbers are colored using geologic dip information. These random numbers are used to select moveout descriptors which are in turn used in a reflection tomography problem. Initial results show that the methodology works as expected, but the variation when starting from such an
Figure 4: Five different SM semblance scans at the CRP locations shown in Figure 4. Overlaying the semblance scans are 30 different \( \gamma \) realizations. Note the fourth from the left gather at 1.8 seconds. Two different maxims are present in the semblance gather and we see that both are picked by various realizations.

Erroneous initial velocity model causes large differences in the various image realizations. More meaningful results can be obtained when the estimated velocity is closer to the true velocity.

**ACKNOWLEDGMENTS**

I would like to thank Satomi Suzuki and Jef Caers of the Stanford Center for Reservoir Forecasting (SCRF) for providing the reservoir description used in this paper.

**REFERENCES**


Figure 5: Nine realizations of random numbers colored by applying fitting goals (12). Note how the random numbers follow geologic dip but vary significantly from image to image.
Figure 6: Nine realizations of $\gamma$ values using the random numbers shown in Figure 5. Note how we see a spatial consistency within a panel but differences between the panels.
Figure 7: The points used to update the velocity model using ray-based tomography overlaying the initial migrated image.


Clapp, R., 2001a, Ray-based tomography with limited picking: SEP–110, 103–112.


Figure 8: Nine realizations of tomography using the $\gamma$ values shown in Figure 6. Note the variation in the velocity structure especially to the left of the fault cutting through the anticline.
Figure 9: Nine realizations corresponding to the nine velocity models shown in Figure 8. Note the differences in the various images. The top-right image shows a much flatter basement reflector than any of the other realizations. The variation positioning of the fault reflector is also significant between the various realizations.


First-order lateral interval velocity estimates without picking

Antoine Guitton, Jon Claerbout, and Jesse Lomask

ABSTRACT
A new method for estimating interval velocities without picking is proposed. The first step applies a normal move-out correction with a $v(z)$ stacking velocity to common mid-point gathers. The second step estimates local stepouts at every offset and time for each gather. Local stepouts across offset are then integrated to obtain local time shifts. The integration is done in the Fourier domain for increased speed. Finally the interval velocity is estimated in the $\tau$ space by fitting the time shifts with a tomographic inversion procedure based on a straight rays geometry. This approach is tested on a Gulf of Mexico dataset with flat geology where recovery of lateral velocity variations across faults is demonstrated.

INTRODUCTION

Interval velocity estimation requires picking in many circumstances. For example, one might pick parameters indicating the flatness of a common depth point (CDP) gather for an initial velocity model (Al-Yahya, 1989; Etgen, 1990) or how well an image focuses after residual migration (Biondi and Sava, 1999). Stereotomography (Billette et al., 2003) requires picking slopes and traveltimes for the velocity estimation process. Other imaging techniques such as common-focus-point migration involve semblance analysis for updating focusing operators (Berkhout and Verschuur, 2001). Moreover, for most of these methods, specifically those based on tomography, several reflectors need to be selected (picked) for the velocity inversion (Clapp, 2001).

Very few velocity estimation techniques do not require any picking. For instance, Toldi (1989) derives a relationship between interval and stacking slowness perturbations that is valid for flat geology with constant velocity background. Closest to our approach, Symes and Carazzone (1991) directly invert time shifts between adjacent traces to estimate interval velocities.

It is our belief that picking is inherently flawed and should be replaced by more robust techniques requiring as little human interpretation as possible. The major shortcomings of human intervention are unrepeatability and subjectivity. Results of velocity analysis invariably differ from one person to another based on the tools used to perform the picking or on the experience of the interpreter. Our conjecture is as follows:

---

1email: antoine@sep.stanford.edu, jon@sep.stanford.edu, lomask@sep.stanford.edu
It is difficult to pick arrivals or events at different spatial locations reliably. However, it is easy to estimate local stepouts between adjacent traces. Event picking should always be replaced by dip estimation.

Therefore, we propose a fully automated interval velocity estimation technique based on (1) dip estimation, (2) dip integration and (3), tomographic inversion. The ultimate goal of this work is to be able to provide a robust technique that affords a first order estimate of interval velocities.

The initial velocity, or starting guess, is a $v(z)$ model. From this simple model we apply a NMO correction to the CMP gathers. In general, NMO is unable to completely flatten CMP gathers because of laterally varying velocity. Flat gathers are then obtained by estimating a trace-by-trace local stepout from the NMO corrected gathers. Local stepouts are then integrated to form absolute time shifts at every time, offset, and midpoint location. If data are noise-free (bad traces, random noise), estimated time shifts flatten gathers regardless of the subsurface complexity.

Time shifts are then used to perform a tomographic inversion in $(x, \tau)$ space (Clapp and Biondi, 2000), where $x$ is the mid-point position and $\tau$ the zero-offset travel time. Straight rays are assumed between the subsurface location and the source/receiver positions; however, this assumption is evidently violated for any realistic geological setting. We none-the-less show that this simplistic model leads to a reasonable velocity update.

Once the interval velocity is updated, more iterations of tomography are usually required. Because of the assumptions made in the tomographic inversion, we stop at the first velocity update. One way to check whether the estimated velocity perturbations flatten the gathers or not is by applying the forward modeling operator to the estimated velocity perturbations; the modeled time shifts can then be applied to the NMO corrected gathers. From this approach we are able to obtain updated interval velocities and flat CMP gathers. In the next two sections, the time-shifts estimation step and the tomographic inversion are described.

**ESTIMATION OF TIME SHIFTS**

Estimating time shifts is a two step procedure where local stepouts are first estimated and then integrated. The goal of dip estimation is to find a local stepout, $p_h$, that destroys the local plane wave such that,

$$0 \approx \frac{\partial u}{\partial h} + p_h \frac{\partial u}{\partial \tau},$$

where $u$ is the wavefield at time $\tau$, midpoint $x$ and offset $h$. For all gathers, we evaluate the slope $p_h$ with a method based on high-order plane-wave destructor filters (Fomel, 2002). This technique has the advantage of being accurate for steep dips. The estimation of $p_h$ is based on a non-linear algorithm (Gauss-Newton method) that includes a regularization term. This regularization smooths dips across offset and CMP location. One problem with the current technique is that the dip estimation algorithm cannot properly handle conflicting dips. This
Velocity estimation without picking

can be troublesome when, for instance, multiples are present in the data. To solve this problem, Fomel (2002) and Brown (2002) show how two dips can be estimated. Then if multiple reflections are present, the stepouts corresponding to the primaries are kept and those of the multiples are rejected.

Dips estimates lead to a vector of local stepouts $p_h$ that are integrated one CMP at a time to obtain time shifts. Dips are smoothed along both spatial directions but not in time $\tau$. To enforce smoothness in the $\tau$ direction Lomask and Guitton (2004) introduce a time component $p_\tau$ to $p$. The relationship between the local time shift vector, $t_s = t_s(\tau, x, h)$ and the local dip vector $p = (p_h, p_\tau)^T$ ($()^T$ being the transpose) at a constant $x$ is:

$$p = \begin{pmatrix} p_h \\ p_\tau \end{pmatrix} = \begin{pmatrix} \partial_h t_s \\ \partial_\tau t_s \end{pmatrix}, \tag{2}$$

where $\partial_\tau$ and $\partial_h$ are the partial derivative in $\tau$, $h$ respectively. In practice, we choose $p_\tau = 1$ and control the amount of smoothness by introducing a trade-off parameter $\epsilon$ as follows (Lomask and Guitton, 2004):

$$\begin{pmatrix} \partial_h t_s \\ \epsilon \partial_\tau t_s \end{pmatrix} = \begin{pmatrix} p_h \\ \epsilon p_\tau \end{pmatrix}. \tag{3}$$

Using equation (3) we wish to minimize the length of a vector $r_{ts}$ that measures the difference between $p_e$ and $\nabla_\tau t_s$ as follows:

$$0 \approx r_{ts} = \nabla_\epsilon t_s - p_e, \tag{4}$$

where $\nabla_\tau = (\partial_h, \epsilon \partial_\tau)^T$, and $p_e = (p_h, \epsilon p_\tau)^T$. We then minimize the following objective function:

$$f(t_s) = ||r_{ts}||^2, \tag{5}$$

where $||.||$ is the $L_2$ norm. By increasing $\epsilon$, the estimated time shifts $\hat{t}_s$ become smoother in time. We solve equation (5) analytically in the Fourier domain (Lomask, 2003), which speeds up the estimation of $t_s$:

$$\hat{t}_s = \text{FFT}_{2D}^{-1} \left[ \frac{\text{FFT}_{2D} \left[ \nabla_\epsilon p_e \right]}{-Z_h^{-1} - \epsilon Z_\tau^{-1} + 2 + 2\epsilon - Z_h - \epsilon Z_\tau} \right] \tag{6}$$

where $Z_h = e^{i\omega \Delta h}$ and $Z_\tau = e^{i\omega \Delta \tau}$. The dip integration yields the desired time shifts plus a constant, i.e., a DC frequency component. The zero frequency component is removed by subtracting the near offset panel from the other offsets. Therefore, the time shifts are a measure of the moveout errors relative to the near-offset panel. At the end of the dip integration process, we end up with a map of time shifts, $t_s(\tau, x, h)$. These time shifts can be used to flatten the CMP gathers without any velocity analysis. Our goal, however, is to find an interval velocity function consistent with the estimated time shifts using a tomographic inversion procedure.
From the time shifts, we estimate interval velocities in the $\tau$ domain. As pointed out by Alkhalifah (2003) and Clapp (2001), $\tau$ tomography is more robust than depth tomography to reflector position and velocity errors. However, going from depth to vertical travel time introduces new variables. As described by Biondi et al. (1997) and Alkhalifah (2003), the transformation from depth coordinates ($x, z$) into vertical-traveltime coordinates ($\tilde{x}, \tau$) is governed by the relationships:

$$
\tau(x, z) = \int_0^z \frac{2}{v(x, z')} dz',
$$

$$
\tilde{x}(x, z) = x.
$$

Therefore, we have the following relationships between the differential quantities ($dx, dz$) and ($d\tilde{x}, d\tau$):

$$
dz = \frac{v(x, z)}{2} d\tau - \frac{v(x, z)\sigma}{2} d\tilde{x},
$$

$$
dx = d\tilde{x},
$$

where $v(x, z)$ is the focusing velocity proportional to the mapping velocity (Clapp, 2001) and

$$
\sigma = \int_0^z \frac{\partial}{\partial x} \left( \frac{2}{v(x, z')} \right) dz'.
$$

In this paper, it is assumed that $x = \tilde{x}$ and $\sigma = 0$ because the initial slowness field is horizontally invariant.

The data space for this inverse problem is a cube of time-shifts at every time, offset and midpoint location. This differs from most tomographic techniques where a few reflectors are usually selected and picked for the inversion. The number of the model space unknowns (the velocity update) is the product of the number of gathers and the number of time samples. A velocity perturbation is computed for each pixel in the model space. For a CMP location $x$ at time $\tau$ and offset $h$, a total time shift $t_s$ is estimated. The forward problem relating velocity perturbation and time shift is derived from Fermat’s principle:

$$
t_s(\tau, x, h) = \int_0^\tau (\Delta s^-(r) + \Delta s^+(r)) dr,
$$

with

$$
dr = \frac{d(dt)}{dS},
$$

where $\Delta s^-(r)$ and $\Delta s^+(r)$ are the slowness perturbations along the down- and up-going rays respectively (from $x - h/2$ to $x$ and $x + h/2$ to $x$), $S$ is the focusing slowness, and $dt$ the time increment along the ray (Clapp and Biondi, 2000). Note that slownesses are actually estimated and not velocities, as it is usually done in tomography. To simplify the problem, we assume that the up- and down-going rays are straight lines in the $(\tau, x)$ space.
Equation (10) is a linear relationship between the time shifts and the slowness perturbations allowing us to write,

\[ d = Lm, \]  

(12)

where \( d \) are the estimated time shifts, \( L \) is the tomographic operator in equation (10) and \( m \) is a field of slowness perturbations. Our goal is to find \( m \) such that,

\[ 0 \approx r_d = Lm - d. \]  

(13)

The addition of a regularization operator to enforce smoothness in the horizontal direction gives:

\[ 0 \approx r_d = Lm - d, \]

\[ 0 \approx \epsilon r_m = \epsilon \nabla_x m, \]  

(14)

where \( \nabla_x \) is the horizontal gradient. Next, \( m \) is estimated in a least-squares sense by minimizing objective function,

\[ f(m) = \|r_d\|^2 + \epsilon^2 \|r_m\|^2. \]  

(15)

In practice, \( m \) is estimated with a conjugate-gradient method and \( \epsilon \) is estimated by trial and error. Because tomography is inherently non-linear, more iterations are needed to converge toward a satisfying velocity model. However, the assumptions made in this paper do not allow us to iterate without using more sophisticated imaging operators or ray tracing tools. We now test our method on a 2D Gulf of Mexico dataset.

**A GULF OF MEXICO 2D FIELD DATA EXAMPLE**

Figure 1 displays a near-offset section of a 2D dataset. The geology is relatively simple with mostly flat layers and few normal faults. A first 1D interval slowness model is estimated by assuming a \( v(z) = v_0 + \alpha z \) function that leads to \( v(\tau) = v_0 e^{\alpha \tau^2} \) in \((x, \tau)\) space. With this dataset \( v_0 = 1.6 km/s \) and \( \alpha = 0.5 s^{-1} \). We then transform the velocity into an interval slowness function shown Figure 2.

Figure 3 shows every twenty-fifth CMP gathers after NMO correction. Note that these gathers are not perfectly flat and that the noise level is quite high, especially in the deepest part of the section. In addition, there are both missing and bad traces at different offsets. We expect that the estimation of stepouts is robust enough to the noise level present in the gathers to give reasonable dips.

Local stepouts and time shifts are estimated from the CMP gathers. Figure 4 displays the estimated time shifts for the five selected CMP gathers. It is interesting to notice that the time shifts increase with offset. The time shifts are also relatively smooth in the \( \tau \) direction thanks to the dip regularization in equation (3). The smoothing in both offset and midpoint directions during the stepouts estimation allows us to have time shifts where traces were originally
Figure 1: Near-offset section of the 2D field dataset from the Gulf of Mexico. Some normal faults are visible.}

Figure 2: Initial slowness function.
Figure 3: Five CMP gathers every 1.6 km after NMO correction with the RMS velocity derived from the interval slowness in Figure 2. Some residual curvature is apparent throughout the section.

missing (e.g., gathers four and five in Figure 3). The fact that the estimated time shifts change with midpoint for a fixed time and offset prove that lateral velocity variations exist. These time shifts can be checked by applying a moveout correction to the input gathers in Figure 3 according to the shift values in Figure 4. Figure 5 shows the same gathers after moveout correction. These gathers are now flat and demonstrate that the estimated time shifts after integration of the local stepouts are correct.

The velocity perturbations are then estimated from the time shifts with the $\tau$ tomography. Figure 6 shows estimated slowness perturbations and Figure 7 displays the updated slowness field. We used 40 iterations and set $\epsilon = 1$ in equation (15) to obtain this result. Lateral velocity variations are visible throughout. In Figure 8, four fault locations interpreted from the seismic are superimposed (Figure 15). These faults locations seem to be aligned with velocity variations in Figure 7. In particular, it is pleasing to see the change of velocities across the different faults.

To check whether the method converged, modeled time shifts are estimated from the slowness perturbations in Figure 6 by applying the forward operator in equation (10). The re-modeled time shifts are shown in Figure 9. Comparing Figures 4 and 9, it appears that the re-modeled time shifts are smoother. Yet, applying these time shifts to the NMO corrected data in Figure 3 yield flat gathers (Figure 10). The difference between Figures 4 and 9 is
Figure 4: Estimated time shifts for five CMP gathers. The maximum time shift is around 0.05 s. Note that the first trace is set to zero.

Figure 5: Flattened CMP gathers after applying the time shifts in Figure 4. Comparing with Figure 3, most of the events are now flat.
that the re-modeled time shifts are constrained by the physics of the tomographic inversion, thus giving well-behaved amplitude variations. In Figure 4, however, the time shifts take any value according to estimated dips. The forward operator of the tomographic inversion can be interpreted as a velocity-consistent, time shifts estimator.

![Figure 6: Slowness perturbations estimated from the tomography in the \( \tau \) space.](antoine3-delta.dslow)

Finally, the flattened gathers in Figure 10 are stacked (Figure 14). We compare this result with the stacked section of the data with a 1D slowness model shown in Figure 11. This slowness function flattens the CMP gathers well for every midpoint position (Figure 12). We show the stacked section of the input data with this 1D slowness function in Figure 13. The reflectors are stronger and better defined in Figure 14 wherever the signal level is strong. In the lower part of the section, however, some continuous events in Figure 13 are attenuated in Figure 14. This effect is due to the difficulty to estimate meaningful stepouts when the noise level is too high. Again, four interpreted faults are shown on the stacked section in Figure 15.

**DISCUSSION-CONCLUSION**

We present a method for estimating interval velocities without picking. In this approach, we estimate time shifts from NMO corrected gathers by first computing local stepouts and then integrating them across offset. These time shifts are then fitted with a tomographic inversion in \((x, \tau)\) space. This approach seems to be robust to a reasonable noise level in the data. In addition, only a few parameters need to be chosen.
Figure 7: Updated slowness field by adding Figures 2 and 6.

Figure 8: Same as Figure 7 with four interpreted faults from the seismic. Note the changes of velocity across the faults.
This technique presents some limitations that require improvements. For instance the local stepouts need to be single valued with no conflicting dips in the data. This limitation can be overcome by estimating a few dips and retaining the ones of interest only. Second, an initial 1D velocity model is needed as a starting guess. This approximation, along with the simplified geometry of the rays, prevent us from recovering lateral velocity variations for complex geology, e.g., salt environment. This limitation can be overcome by incorporating the work of Clapp and Biondi (2000).

In spite of these approximations, this new velocity estimation technique is able to recover lateral velocity variations without picking for a 2D Gulf of Mexico dataset. In addition, the updated velocity field seems to match the geological environment: we can see velocity changes across faults at various locations. We finally show that the estimated velocity perturbations yield a map of time shifts that can be used to flatten CMP gathers.

In theory, more iterations of velocity updating should be performed. One problem with more updates is the need for more sophisticated time or depth imaging algorithms. In addition, more updates would mean improving on the tomographic inversion by allowing any type of ray geometry and background slowness field. These changes go beyond the scope of this paper. We believe, however, that all these sources of improvements should be investigated further to provide a robust and picking-free interval velocity estimation tool.
Figure 10: Flattened data with the remodeled time shifts in Figure 9. The gathers are flat showing that the slowness perturbations in Figure 6 fit the estimated time shifts in Figure 4 very well.

Figure 11: A 1D stacking slowness function.
Figure 12: CMP gathers after NMO with the slowness function in Figure 11. The gathers are almost flat.

Figure 13: Stacked section of the data with the our best picked 1D slowness function in Figure 11.
Figure 14: Stacked section of the moveout corrected data from the remodeled time shifts in Figure 9.

Figure 15: Stacked section of the moveout corrected data from the remodeled time shifts in Figure 9 with four picked faults.
REFERENCES


Velocity estimation for seismic data exhibiting focusing-effect AVO (Part 4)

Ioan Vlad

ABSTRACT
Focusing-effect AVO (FEAVO) can be eliminated by migrating with a velocity model obtained by Wave-Equation Migration Velocity Analysis (WEMVA). For this specific problem, WEMVA requires the extraction of an image perturbation that contains all and only the FEAVO effects. Such an image perturbation can be generated by using a “discriminate-focus-filter-mask” strategy. A simplified version of the first step of this approach was implemented with acceptable results. A simple, effective and cheap FEAVO detector was also conceived and implemented.

INTRODUCTION
Focusing-effect AVO (FEAVO) is the focusing of seismic wavefield amplitudes through velocity lenses too small to generate fully developed triplications. The amplitude effects are large enough to thwart proper AVO analysis in the affected area (Kjartansson, 1979). The effects can be eliminated by migrating with a velocity model containing the respective lenses (Bevc, 1994). However, the traveltime effects are too small to allow classical velocity analysis approaches such as Dix inversion or traveltome tomography to succeed. New approaches are needed to deal with such small velocity anomalies.

Biondi and Sava (1999) introduced Wave-Equation Migration Velocity Analysis (WEMVA), which finds the unknown part of the true velocity field by optimizing image quality in the angle domain after prestack depth migration. Vlad et al. (2003) have shown on a synthetic dataset that WEMVA resolves the FEAVO-causing velocity anomalies. The solution was verified by showing that the amplitude anomalies disappear after migration with the updated velocity model.

Error can be introduced in WEMVA by two classes of factors. The first one consists of the approximations undertaken by the mathematical and numerical procedures. The second one encompasses the limitations of the ability to transform a given image into an optimal one. In a synthetic dataset, in which the optimal image is known, this second source of error is completely eliminated. This allowed Vlad et al. (2003) to evaluate solely the effect of the first class of error-generating factors (Born approximation, linearized wavefield continuation,

1email: nick@sep.stanford.edu
etc.), and they found that it is too small to impede WEMVA from resolving the small FEAVO-causing velocity lenses. This paper, on the other hand, deals with eliminating the second class of error-generating factors by obtaining the optimal image in a way that is as error-free as possible.

**EXTRACTING THE IMAGE PERTURBATION**

WEMVA proceeds in three steps. The first one consists of prestack depth-migrating the data with a starting velocity model, then transforming the image to the angle domain. The second step starts by improving the image so that it is closer to the result of a migration with the correct velocity. The improvement may be performed by straightening the gathers or by smoothing the amplitudes along already flat gathers. The improved image is subtracted from the original one to create an image perturbation. The third step is inverting the image perturbation into a slowness perturbation.

The WEMVA flow presented above is actually the one for the procedure named Target Image Fitting (TIF) WEMVA. In a more radical inversion-theory approach, called Differential Semblance Optimization (DSO) WEMVA, steps 2 and 3 are combined into a single step, with the image improvement being performed by a weight operator during the inversion. In theory, DSO WEMVA will not exhibit the same Born-related problems as TIF WEMVA (Sava and Symes, 2002). However, it assumes that the quality of the image-improvement operator that is embedded as a weight in the inversion can be trusted. I am currently just trying to develop such an image-improvement operator specific to the FEAVO problem, and I need to be able to isolate its effects, so for the time being I will use TIF WEMVA on velocity anomalies inside the Born approximation.

In the context of the FEAVO problem, the image perturbation is the difference between the FEAVO-free image and the original one. I therefore need to eliminate the FEAVO effects from the image. To do so, I have to discriminate between AVO due to focusing and AVO due to actual changes in reflectivity with the incidence angle. I will proceed by first obtaining an estimate of the AVO due to the properties of the reflector.

The variation of amplitudes due to changes in reflectivity with incidence angles lower than $30^\circ$ has been modeled by Shuey (1985) as follows:

$$R(\theta) = I + G \sin^2(\theta),$$

where $I$ and $G$ are scalars depending only on the physical properties of the materials above and below the reflecting interface. If the amplitudes are picked at a single midpoint-depth location with no FEAVO or illumination problems, and they are plotted as a function of the squared sine of the incidence angle, the values will arrange close to a line with intercept $I$ and gradient $G$. The presence of FEAVO causes the linear dependence to break, as exemplified in Figure 1. The presence or absence of a linear trend in the $(\sin^2, \text{amplitude})$ space is therefore a FEAVO-discriminating criterion, which can be used to eliminate FEAVO. I will have to account also for the superposition of FEAVO and reflector-caused AVO. At each midpoint-depth location, I
Figure 1: **Top panel:** Midpoint-angle depth slice from the prestack migrated synthetic dataset shown in Figure 1 of Vlad et al. (2003). **Bottom panels:** Amplitudes at midpoints marked by vertical thin lines in the upper panel.
will compute the most plausible reflector-caused AVO as a linear trend in the \((\sin^2, \text{amplitude})\) space, and I will obtain FEAVO by subtracting this trend from the image.

The challenge is now fitting to the data at each midpoint-depth location a linear trend in the \((\sin^2, \text{amplitude})\) space. The trend should be as close as possible to that of the reflector-caused AVO, even when focusing effects are present. I will consider only the angles up to 30°. After that, the accuracy of the angle-gather amplitudes decreases, unless the offsets are extremely densely sampled, which would increase the costs of the migrations in WEMVA too much. Also, after 30°, the two-term Shuey approximation of reflector-caused AVO stops working. I have experimented with his three-term formula for larger angles. The extra degree of freedom allowed for a curve that fitted everything, including the FEAVO effects at angles lower than 30°. I have therefore decided to use only the two-term formula on angles under 30°.

The best way to fit the linear trend would be to formulate the fit as a 2-D inverse problem, penalizing nonphysical variations of AVO along the midpoint or along the reflector. While I intend to do so in the future, I experimented for the time being with a 1-D approach mapped in Figure 2. At each (midpoint, depth) location I fitted by least-squares a straight line through

![Figure 2: 1-D FEAVO extraction flowchart.](nick1-extract1D_flow)

the \([\text{reflectivity, } \sin^2(\text{angle})]\) pairs. I then subtracted the linear trend and applied a low-cut filter along angle. A depth slice from the results of the FEAVO extraction, together with the corresponding true FEAVO anomalies, is shown in Figure 3. The extracted anomalies are closer in morphology to the true FEAVO in the optimal image perturbation. Also, the extraction worked up to 30°, in contrast with the previous image-processing-based technique proposed by Vlad et al. (2003), which only worked up to 21°.

A 2-D extraction procedure in which the intercept and gradient would vary more smoothly
Figure 3: **Top panel:** Depth slice from the optimal angle-domain image perturbation in the top panel of Figure 6 of Vlad et al. (2003)). **Bottom panel:** FEAVO anomalies extracted by applying the workflow in Figure 2 to the image built by migrating with a constant velocity the dataset in Figure 1 of Vlad et al. (2003).

along the midpoint would make it possible to eliminate in the future the low-pass filter step, which is responsible for the reverse-polarity “borders” around the extracted anomalies. However, the results of the 1-D extracting procedure may still provide some valuable insight into the effectiveness of FEAVO extracting procedure, given that the “unfair” advantage of the lack of noise is compensated by the simplicity of the 1-D extraction.

The slowness update produced by inverting the extracted image perturbation is displayed in Figure 4. The slowness update produced this way puts anomalies with the correct polarities in the correct places. The shape of the anomalies could nevertheless be improved with the use of a 2-D FEAVO extraction procedure.

The method presented in this paper discriminates FEAVO on the basis of a local property—nonlinear variation with the squared sine of the angle. FEAVO anomalies, however, are correlated in tent-like shapes throughout the prestack volume, as shown in Figures 3, 4 and 9 in Vlad (2002). That paper proposed a “focus-filter-spread” approach to discriminate FEAVO from uncorrelated amplitude effects. The shapes of the anomaly surfaces in the prestack volume can be precomputed as functions of the known velocity. Summation along them would highlight the anomalies; only the “bright stars” would be kept. Then, the de-noised Radon-space image would be spread along the surfaces, to keep only the FEAVO anomalies in the image perturbation. Vlad et al. (2003) identified a major difficulty with this approach: the alternating polarities in the wavelet would lead to summed values canceling each other. The
obvious solution—taking the unsigned values—would, however, lead to the inability to discriminate between positive and negative velocity anomalies, and to the inversion process never converging if the data contains FEAVO-generating velocity anomalies of more than one sign.

The conundrum of FEAVO extraction based on its global properties can however be broken by using a procedure that I present below. Instead of a “focus-filter-spread” approach, I will use a “discriminate-focus-filter-mask” one. I will first process the prestack volume through a FEAVO discriminator based on the local properties of the anomalies. Then, as in the “focus-filter-spread” approach, I will square or take the modulus of the values in the prestack volume, sum along precomputed velocity-dependent surfaces, filter to keep only the “bright stars” in the Radon domain, then spread back. But instead of feeding the result of spreading into the inversion, I will use it as a mask that I will apply to the result of local property-based FEAVO discrimination. That will filter out all areas with amplitude variations that are nonlinear, but are not FEAVO (because they do not correlate vertically and laterally in the known tent-shape). Since the synthetic dataset used in this paper does not have non-FEAVO amplitude departures from the Shuey (1985) model, it represents a good benchmark for the accuracy of the inversion, even in the absence of the de-noising “focus-filter-mask” step.

**AUTOMATIC FEAVO DETECTION**

The automatic FEAVO-discriminating procedure outlined in the previous section offers a ready solution to an older problem: FEAVO detection. Until now, the only way to find whether the amplitudes at the reflector were corrupted by focusing effects was the method used by Kjartansson (1979). First, the interpreter would have to become suspicious that the AVO in his region of interest exhibited variations too strong and abrupt to be due to the rock properties expected at that interface. The dataset would be sent back to processing, which would sum the absolute values of the samples in the 2-D prestack line along time, and examine the midpoint-offset image looking for the specific “V” shapes of the anomalies. Since it is not to be expected that such a labor-intensive, custom procedure was undertaken routinely, compromised AVO measurements were the likely result.
FEAVO, however, can be detected and flagged automatically during processing in a much simpler way. Figure 2 depicts the removal of a linear trend from FEAVO-affected data. If the same workflow is applied to data that is not affected by FEAVO, the residual after subtraction of the trend will be very close to zero everywhere. High values in the variance of the residuals will therefore flag the presence of FEAVO. Figure 5 shows the result of automatic FEAVO detection, which is no longer a 5-D cube, but a 3-D one, reducing the volume of data to be examined by orders of magnitude through elimination of the offset axes. The presence of FEAVO anomalies is highly visible. I need not even do a good job of finding the values of the intercept and gradient. All that matters for detection is that the amplitude values in focusing-affected areas do not vary linearly with the squared sine of offset, and therefore no straight line whatsoever will approximate them well. The procedure is also quite cheap computationally.

![Figure 5](image)

Figure 5: FEAVO anomalies flagged in the midpoint-depth space by the automatic detection procedure. The stars denote the location of the lenses causing the focusing.

I could build an “anomaly locator” by summing Radon-transform-style the output of the FEAVO detector inside the bounds of precomputed, velocity-dependent, FEAVO-effect paths. That, however, would not give the magnitude of the velocity anomalies, so it would be of little use in eliminating the amplitude effects in the prestack volume.

**CONCLUSIONS**

I have identified a “discriminate-focus-filter-mask” strategy for constructing an image perturbation for WEMVA’s application to the FEAVO problem. I implemented, with acceptable results, a simplification of the first step of this approach—FEAVO discrimination based on local characteristics. I built and implemented an effective process for performing FEAVO detection in a much simpler and cheaper manner than the current practice.
ACKNOWLEDGMENTS

I thank Paul Sava for meaningful discussions and for software assistance.

REFERENCES


Robust moveout without velocity picking

Kevin Wolf, Daniel Rosales, Antoine Guitton, and Jon Claerbout

ABSTRACT

At every point in a CMP gather, a local estimate of RMS velocity is:

\[ V_{RMS}^2 = \frac{x \, dx}{t \, dt}, \]

where \( dt/dx \) is the local stepout. We form a median stack of these local velocity estimates to obtain stable estimates of RMS velocity without the conventional need to form many hyperbolic stacks.

INTRODUCTION

Velocity estimation is still a fundamental problem in seismic exploration. There are numerous methods to estimate the stacking velocity model based on velocity spectra (Taner and Koehler, 1985; Lumley, 1992). However, all these methods depend on the ability to pick the velocity from a series of coherency panels. These methods of velocity estimation are sensitive to noise levels in the data.

A way to make velocity estimation robust in the presence of noise is to use median stacks within CMPs. A problem with CMP stacks is that data from a large range of offsets are merged despite intrinsic variations in gain, frequency, NMO stretch, array response, and AVO. We estimate \( V_{RMS} \) with a robust median estimator of slowness squared terms; each term manufactured from neighboring traces only. Our goal is to develop a robust code that will reasonably move out all 40 of the worldwide Yilmaz-Cumro shot profiles (Yilmaz, 1987) without need for individualized parameter choices.

METHOD

A simple way to represent a wave traveling at constant velocity with slowness \( s \) is as an expanding circle:

\[ t^2 = \tau^2 + x^2 s^2, \]  

\[ 1 \text{email: kxwolf@pangea.stanford.edu,daniel@sep.stanford.edu,antoine@sep.stanford.edu,jon@sep.stanford.edu} \]
where $t$ is traveltime, $\tau$ is traveltime depth and $x$ is offset (Claerbout, 1995). Differentiating with respect to $x$ at constant traveltime depth $\tau$ we obtain:

$$s^2 = \frac{t\, dt}{x\, dx},$$

where $dt/dx$ is Snell’s parameter $p$. Snell’s parameter is related to the apparent horizontal velocity. It can also be regarded as a measure of the local stepout (dip) at any given time and offset along a hyperbola. Therefore, from equation (2) it can be seen that multiplying the local dip of hyperbolas in the $(x, t)$ plane by the ratio of their time and space coordinates yields an estimate of slowness squared. This estimate of slowness is independent of where it is measured along the event in the $(x, t)$ plane; consequently, a NMO correction of the slowness squared section will result in horizontal lines of constant slowness.

In order to obtain a dip estimate for the events in the plane we employ the method of Fomel (2002). This technique estimates local stepouts with plane wave destructor filters. Only one dip is estimated at every time and offset position, which makes this method sensitive to crossing events or coherent noise. A solution to this problem is to estimate multiple dips at every location and to select those of interest (Fomel, 2002). Once the dips have been estimated, the slowness can be computed in a straightforward manner by multiplying each dip estimate by $t/x$. A map of local slowness (squared) is then obtained that need only be converted to a velocity profile for a given CMP gather.

To achieve this goal, a NMO correction with an approximate velocity trend can be applied to roughly flatten the hyperbolas. Finally, a median stack over the $x$ coordinate provides a reasonable estimate of $s^2$ as a function of $\tau$. To ensure local bad dip estimates do not skew the results of the method, data points corresponding to atypical slowness values are disregarded, and the final result is smoothed in time. Once the estimate of slowness squared is obtained, we convert it to $V_{RMS}$.

**TEST CASES**

To assess the usefulness of the proposed method we applied it to a synthetic data set and to several shots from the set of 40 worldwide Yilmaz-Cumro shot profiles.

**Synthetic Example**

The proposed methodology was first tested on a simple synthetic example in order to check the validity of the approach. Figure 1a shows the synthetic example, which is an idealized case with neither crossing events nor aliasing, allowing robust dip estimation to occur.

The variable brightness of the estimated local dip in Figure 1b represents the calculated value of dip, and shows that the estimate is robust for this simple synthetic example. Figure 1c, calculated $s^2$ after NMO correction with the synthetic velocity profile, shows the expected trend with slowness values decreasing slightly at later times. There are some anomalous values
of $s^2$ at small offsets due to the minimal dip of the reflectors at those locations, but they will be removed by the median stacking procedure.

The estimated RMS velocity function of the synthetic data is shown in Figure 2. The solid line represents the velocity function used to create the synthetic data. The two dashed lines show the estimated velocity, one estimate from the original $s^2$ panel, and the other from the $s^2$ panel that has been NMO corrected (Figure 1c). The two results are similar, suggesting that in order to obtain a robust estimate of slowness the velocity used to flatten the hyperbolas of $s^2$ before the median stack does not need to be very accurate. As long as data do not have extremely large offsets, a rough estimate of velocity for NMO correction should flatten the hyperbolas of $s^2$ adequately, allowing the median stacking routine to obtain a reasonable estimate of $s^2$.

The results of using the two estimated velocity functions for NMO correction are shown in Figure 3. Figure 3b uses the $s^2$ estimate without NMO correction, and Figure 3c uses the $s^2$ velocity estimate with NMO correction. In both cases the estimated velocity function does a good job of flattening the hyperbolas in the synthetic data. The results are encouraging and suggest that the method will work even if the velocity for NMO correction of the plane of $s^2$ is inaccurate.
Figure 2: RMS velocity profile used for the synthetic model, and estimated RMS velocity profiles with and without the NMO correction applied to the estimate of $s^2$. [kevin1-synvelocity] [ER]

Figure 3: (a) Synthetic data, (b) NMO correction using velocity estimated without NMO correction, and (c) NMO correction using velocity estimated with NMO correction. No picking was required to flatten this gather. [kevin1-synresults] [ER]
Field Data Examples

Once the method was shown to work on an idealized synthetic case, we tested it on some real shot gathers to see how robust the method is when working with real data and the problems inherent with them. Theoretically, CMP gathers should be used for this analysis, however we have decided to use the Yilmaz-Cumro shot gathers to test our method. The 40 shots in the dataset provide varying data quality and numerous challenges, which will thoroughly test the robustness of our method.

Figure 4a shows shot 14 from the Yilmaz-Cumro shot gather dataset. The shot is a fairly clean record consisting of many hyperbolic events that should give a good velocity estimate. Examining Figure 4b, which shows the dip estimate, there are areas at the top of the profile that have negative dips. These result from the low velocity direct arrivals visible in the data; the dip estimator picks the aliased energy for its dip estimate. In the same area in Figure 4c the negative dip estimates result in a negative estimate for the value of $s^2$. Although this is an undesirable result, these values will be accounted for and disregarded by the median stacking routine. The estimate of $s^2$ has had an NMO correction applied to it with an arbitrary velocity of 2.5 km/s. Although most of the hyperbolas are not well flattened, synthetic example results show that this has little effect on the output from the median stack.

Figure 4: (a) Shot gather 14 with AGC, (b) estimated local dip, and (c) estimated $s^2$ after NMO correction of 2.5 km/s.
The estimated RMS velocity profile created by median stacking Figure 4c is shown in Figure 5. At early times the estimate varies due to a lack of information from large offsets required to help constrain the velocity. For this reason the beginning of the estimate is erased and replaced by extending a reasonable value back to time zero. The estimate also fluctuates at later times in the record, because the data become noisier, causing dip estimates to become unstable.

The result of applying the NMO correction with the estimated velocity function is shown in Figure 6b. Data are flattened in most areas, especially away from the beginning of the section where there is little offset information, and before the data becomes noisier at later times. Examining the strong reflector at approximately 5.2 seconds in Figure 6a, which is in the noisier part of the record and clearly non-hyperbolic, Figure 6b shows that enough velocity information is present in order to effectively flatten the event. This result is encouraging.

The method was also tested on shot 27 from the Yilmaz-Cumro dataset, shown in Figure 7a. The shot has numerous hyperbolas at early times, but noise levels hide hyperbolic events at later times. The dip estimate (Figure 7b) and $s^2$ estimate (Figure 7c) each show negative values associated with the direct arrival, and the noisier part of the record but, again, these are handled by the median stacking routine.

The estimated RMS velocity (Figure 8) was also extended back to zero time in order to remove bad values where there is little far offset information. The estimate seems reasonable until around 4.5s; where it starts to fluctuate. This is where the noise level in the data increases, and the poor estimate in this location occurs. The results of the NMO correction with this velocity estimate is shown in Figure 9b. Again the data has been reasonably flattened, particularly at times where the velocity estimate was well constrained.
CONCLUSIONS

The proposed velocity estimation method performs well in the test cases presented. Estimates are robust for areas of reasonable data quality. In these areas enough information is available for the median stacking routine to eliminate poor data points, and gently smoothing the estimate in time gives a good estimate of velocity. In poor data quality areas, or at early times with little offset information, estimates are less reliable and unstable. Although estimates are not always exact in these areas, they do provide a decent starting estimate of velocity without velocity spectra analysis or manual picking.

REFERENCES


Figure 7: (a) Shot gather 27 with AGC, (b) estimated local dip, and (c) estimated slowness^2 after NMO correction with velocity of 2.5 km/s. [ER]

Figure 8: Estimated RMS velocity profile for the data in Figure 7. [ER]
Figure 9: (a) Shot gather 27 with AGC, (b) results of the NMO corection on shot 27 using the estimated RMS velocity in Figure 8.


Regularized inversion for subsalt imaging: real data example

Marie L. Clapp and Robert G. Clapp

ABSTRACT

Imaging the subsurface where seismic illumination is poor is a difficult exercise. Conventional imaging techniques such as migration are insufficient. Better results can be obtained from regularized least-squares inversion methods that use migration operators in a conjugate-gradient minimization. We demonstrate this regularized inversion using downward continuation migration and regularization along offset ray parameters (reflection angles) on a real 2-D seismic line. The result is cleaner than the migration result and has filled in some amplitude information where poor illumination caused gaps. We discuss a regularized inversion that uses common azimuth migration and the same type of regularization to image a real 3-D subsurface around a salt body.

INTRODUCTION

Properly imaging the subsurface in areas that are structurally complex is a daunting task. The migration algorithms typically used for imaging are unable to provide satisfactory images where shadow zones are common, particularly around salt bodies (Muerdter et al., 1996; Prucha et al., 1998). Since salt can make a good hydrocarbon trap, these areas are where we would really like to obtain good images.

There have been many improved migration algorithms that ameliorate the effects of the complex subsurface. Several authors have demonstrated that wave equation migration methods can provide better images than Kirchhoff migration methods (Geoltrain and Brac, 1993; O’Brien and Etgen, 1998). Additionally, some artifacts commonly seen in complex areas are caused by seismic energy that arrives at the receivers at the same time, but follow different paths through and reflect at different points in the subsurface (ten Kroode et al., 1994). These artifacts can be reduced by creating images with angle-domain common image gathers (AD-CIGs). Methods that produce ADCIGs through Kirchhoff techniques (Xu et al., 2001) may partially reduce artifacts caused by multipathing, but still have difficulties (Stolk and Symes, 2002). Wave equation methods to create ADCIGs (Prucha et al., 1999; Mosher and Foster, 2000) handle multipathing better (Stolk and De Hoop, 2001; Stolk and Symes, 2004). However, regardless of how a migration algorithm is formulated, migration is generally insufficient to image poorly illuminated areas (Prucha et al., 2001).

To improve our seismic imaging in areas of poor illumination, we can use migration as

1email: marie @ sep.Stanford.edu, bob @ sep.Stanford.edu
an imaging operator in a least-squares inversion scheme (Nemeth et al., 1999; Duquet and Marfurt, 1999; Ronen and Liner, 2000; Prucha and Biondi, 2002b; Kuehl and Sacchi, 2001). In areas with poor illumination, the inversion problem is ill-conditioned; therefore, it is wise to regularize the inversion (Tikhonov and Arsenin, 1977). The regularization operator can be designed to exploit knowledge we have about the expected amplitude behavior and dip orientation of events in the image (Prucha and Biondi, 2002a).

In this paper, we will begin by reviewing a scheme for iterative regularized inversion. We will implement an inversion scheme that regularizes amplitudes along offset ray parameters (reflection angles) on a real 2-D seismic line from the Gulf of Mexico. We will also discuss how our regularized inversion scheme that can be applied to the real 3-D dataset from which we extracted the 2-D line.

**REVIEW OF REGULARIZED INVERSION**

The vast size of the seismic imaging problems makes performing a direct inversion impossible with today’s computer power, even if we are only dealing with a 2-D seismic line. Fortunately, we can closely approximate a direct inverse with iterative techniques. In particular, we can approximate a least-squares inversion with the conjugate-gradient minimization of this objective function:

\[
Q(m) = \| Lm - d \|^2
\]

where \( L \) is a linear modeling operator, \( d \) is the data, and \( m \) is the model. This minimization can be expressed more concisely as a fitting goal:

\[
0 \approx Lm - d. \tag{2}
\]

However, for the seismic imaging problem, this inversion can have a large null space, due in part to poor illumination. Any noise that exists within that null space can grow with each iteration until the problem becomes unstable. Fortunately, we can stabilize this problem with regularization (Tikhonov and Arsenin, 1977). The regularization adds a second fitting goal that we are minimizing at the same time:

\[
0 \approx Lm - d \tag{3}
0 \approx \epsilon A m.
\]

The first expression is the “data fitting goal,” meaning that it is responsible for making a model that is consistent with the data. The second expression is the “model styling goal,” meaning that it allows us to impose some idea of what the model should look like using the regularization operator \( A \). The strength of the regularization is controlled by the regularization parameter \( \epsilon \).

Unfortunately, the inversion process described by fitting goals (3) can take many iterations to produce a satisfactory result. We can reduce the necessary number of iterations by making
the problem a preconditioned one. We use the preconditioning transformation \( m = A^{-1}p \) (Fomel et al., 1997; Fomel and Claerbout, 2003) to give us these fitting goals:

\[
0 \approx LA^{-1}p - d \\
0 \approx \epsilon p.
\]

\( A^{-1} \) is obtained by mapping the multi-dimensional regularization operator \( A \) to helical space and applying polynomial division (Claerbout, 1998). This process is called Regularized Inversion with model Preconditioning (RIP).

**The migration operator**

The migration operator and its adjoint (\( L \)) that are used in this inversion scheme are linear operators. For the 2-D case, we choose to use the downward continuation migration operator introduced by Prucha et al. (1999). This 2-D downward continuation migration operator takes an input of seismic data with the dimensions of common midpoint (CMPX), offset \( (h_x) \), and frequency \( (\omega) \). Its output is a model (image) with the dimensions of depth (\( z \)), common reflection point (CRPX), and offset ray parameter \( (p_{hx}) \), which is related to the reflection angle for a given subsurface point. This downward continuation migration operator can be formulated as a 3-D process by adding the crossline common midpoint (CMPY) and crossline offset \( (h_y) \) to the input, but that would be a very computationally expensive process. Fortunately, to reduce costs in 3-D, we could also use a Common Azimuth Migration (CAM) operator (Biondi and Palacharla, 1996). For this, we add the CMPY dimension, but not the crossline offset.

**The regularization operator**

The regularization operator (\( A \)) should be designed based on the expected model covariance (Tarantola, 1986). Since we are particularly concerned with the effects of poor illumination, we need to design \( A \) to compensate for these effects. In this case, since our downward continuation operator or CAM operator will create a model that is a cube of depth, CRPX, CRPY (for the CAM operator), and \( p_{hx} \), we can expect illumination problems to appear as gaps in events in the CRP plane(s) and along the \( p_{hx} \) axis. Prucha et al. (2001) demonstrated the use of steering filters (Clapp et al., 1997) as a regularization operator to compensate for sudden changes in amplitude along events with an expected dip. Accomplishing this in the CRP planes requires some interpretation of the result of migration, but along the \( p_{hx} \) we expect the events to be flat and horizontal as long as the correct velocities have been used for imaging. In this paper, to keep \( A \) simple, we will just be applying the steering filters horizontally along the \( p_{hx} \) axis. This regularization scheme will minimize changes in amplitude along the \( p_{hx} \) axis, penalizing large amplitude changes the most.
RESULTS

To demonstrate our 2-D Regularized Inversion with model Preconditioning (RIP), we choose to extract a 2-D line from a real 3-D Gulf of Mexico dataset provided to us by BP and ExxonMobil. A portion of the 3-D velocity model for this dataset can be seen in Figure 1. The velocity model is believed to be accurate, which is important given our choice of regularization operator (Clapp, 2003).

![Figure 1: Subset of the BP Gulf of Mexico velocity model.](marie2-bpvel)

We first performed downward continuation migration on this 2-D line. The migration results can be seen in Figures 2 and 4. In these figures, the left part shows a common offset ray parameter section, taken from $p_{hx} = .153$ and $p_{hx} = .317$ respectively. The right part shows a common image gather taken from $CRPX = 20.5$ and $CRPX = 22.225$ respectively.

In Figure 2, note the clear shadow zones visible in the common offset ray parameter section. The poor illumination that causes these shadow zones is manifested in the common image gather as gaps in the events. These gaps are what we hope to fill with our regularization operator.

Figure 3 is the result of 10 iterations of RIP. The common $p_{hx}$ section and common image gather correspond to those of the migration result in Figure 2. The inversion process has cleaned up many of the artifacts seen in the migration result. More importantly, the common image gathers show that we are filling the gaps in the events. Our regularization operator is successfully compensating for the illumination problems.

The common $p_{hx}$ section shown in the migration result in Figure 4 and the RIP result in Figure 5 is interesting. At the CRPX location from which we have extracted the common image gather ($CRPX = 22.225$), we note that the result after 10 iterations of RIP (Figure 5 is beginning to show real events beneath the salt that are not visible in the migration result (Figure 4. The common image gather shows that we have extended the events in $p_{hx}$, essentially allowing the inversion image to “recapture” energy that left the survey area. Additionally, as we saw in Figure 3, the inversion result is cleaner than the migration result.
Figure 2: Result of downward continuation migration of 2-D line. Left part is a common offset ray parameter section, right part is a common image gather taken from $CRPX = 20.5$.

It is also interesting to stack the migration and RIP results for comparison. Figure 6 shows the stacked migration result and Figure 7 shows the stacked RIP result after 10 iterations. Each of these figures has been zoomed in to concentrate on the poorly illuminated areas under the salt. The RIP stack is slightly higher in frequency content, due to artifacts in the migration result that stack into lower frequencies. More importantly, the RIP stack has improved the imaging of events within the shadow zones. The events extend farther into the poorly illuminated areas, are more continuous, and have more consistent amplitudes.

CONCLUSIONS

The Regularized Inversion with model Preconditioning (RIP) process that we have previously used only on synthetic data has proven to be effective on real data as well. The regularization scheme used for RIP in this paper helps to compensate for poor illumination by penalizing
Figure 3: Result of 10 iterations of RIP. Left part is a common offset ray parameter section, right part is a common image gather taken from $CRPX = 20.5$. Large amplitude changes along the offset ray parameter axis. It also helps to clean up artifacts that plague migration results.

FUTURE WORK

As discussed in this paper, we wish to extend RIP to work on 3-D data. We plan to do so by switching from our downward continuation migration to a Common Azimuth Migration (CAM). The regularization operator will still act along the inline offset ray parameter axis.
Figure 4: Result of downward continuation migration of 2-D line. Left part is a common offset ray parameter section, right part is a common image gather taken from $CRPX = 22.225$. marie2-bpcubes.mig2d

ACKNOWLEDGMENTS

We would like to thank BP and ExxonMobil for providing SEP with the 3-D GoM dataset used in this paper.

REFERENCES


Figure 5: Result of 10 iterations of RIP. Left part is a common offset ray parameter section, right part is a common image gather taken from $CRPX = 22.225$. 


Figure 6: Stack of the downward continuation migration result, zoomed in under the salt body.


Figure 7: Stack of the RIP result after 10 iterations, zoomed in under the salt body.


AVA effects of regularized least-squares inversion

Marie L. Clapp

ABSTRACT
As we search for hydrocarbons in areas where the earth’s subsurface is too complex to accurately image with migration algorithms, we find ourselves turning to imaging techniques such as least-squares. My version of least-squares inversion imaging uses a downward continuation operator to produce offset ray parameter gathers (equivalent to angle gathers) and a regularization operator that is a derivative along the angle axis. The regularization operator stabilizes the inversion and helps to fill in illumination gaps. Essentially, I assume that any large, sudden changes in amplitude along the reflection angle axis are caused by poor illumination. This methodology is effective for reducing artifacts and helping to compensate for poor illumination. However, there is still the question of how this regularization will affect any real amplitude variation with angle (AVA) that should be seen in the model. In this paper, I address the question of how the derivative type regularization operator affects expected AVA in a simple model with no illumination problems. I experiment with various numbers of iterations and various strengths of regularization. Overall, I find that this operator, as implemented in this paper, can have a minor effect on the true AVA due to edge effects. However, it does not affect all possible AVA information, so I remain hopeful that the derivative-type regularization operator can be modified to allow us to trust AVA information from models produced by my regularized least-squares inversion.

INTRODUCTION

Our ongoing quest for hydrocarbons requires that we improve our ability to image the earth’s subsurface. This is particularly true in areas around salt bodies, which can be good hydrocarbon traps but cause poor seismic illumination in the surrounding subsurface. Conventional imaging techniques such as migration cannot provide an adequate picture of these poorly illuminated areas (Muerdter et al., 1996; Prucha et al., 1998). In such areas, random noise and processing artifacts can easily obscure the small amount of signal that exists. A common type of artifact seen in these areas is caused by multipathing. Many authors have reduced these artifacts by generating images through Kirchhoff-type migration that create angle domain common image gathers (Xu et al., 2001). The artifacts are even better handled by downward continuation migration (Prucha et al., 1999a; Stolk and Symes, 2004).

However, reducing multipathing artifacts does not significantly improve the image where

1email: marie @ sep.Stanford.edu
illumination is poor. If we wish to try to gain information on rock properties through amplitude analysis (de Bruin et al., 1990), we will have to deal with both these artifacts and the poor illumination. To improve the image and potentially recover accurate information on rock properties, we must move beyond migration.

Although migration is not sufficient to image the subsurface in areas with poor illumination, we can use migration as an imaging operator in a least-squares inversion scheme (Nemeth et al., 1999; Duquet and Marfurt, 1999; Ronen and Liner, 2000). In areas with poor illumination, the inversion problem is ill-conditioned, therefore it is wise to regularize the inversion scheme (Tikhonov and Arsenin, 1977). The regularization operator can be designed to exploit knowledge we have about the expected amplitude behavior and dip orientation of events in the image (Prucha and Biondi, 2002).

When using regularized inversion for imaging, the choice of regularization operator is critical. An intelligent and fairly safe choice is to penalize large amplitude changes as the reflection angle varies for a given point in the subsurface (Kuehl and Sacchi, 2001; Prucha and Biondi, 2002). I refer to this as “geophysical” regularization. This process will help to reduce artifacts and improve the image, but its impact on possible amplitude variation with angle (AVA) analysis must be considered. Kuehl and Sacchi (2002) found that a similar regularization scheme used to compensate for incomplete data rather than the problem of poor illumination could still provide accurate AVA information.

In this paper, I examine the effects of geophysically regularized inversion on a simple synthetic dataset with known AVA. I begin by reviewing the theory of Regularized Inversion with model Preconditioning (RIP). Then I explain the regularization operator used for geophysical RIP. I apply this RIP algorithm to the simple synthetic, using varying numbers of iterations and different strengths of regularization to evaluate the impact of the regularization on the AVA.

**REVIEW OF REGULARIZED INVERSION**

Iterative least-squares inversion can be expressed simply as the conjugate-gradient minimization of this objective function:

\[
min \{ Q(\mathbf{m}) = ||\mathbf{Lm} - \mathbf{d}||^2 \}
\]

where \( \mathbf{L} \) is a linear modeling operator, \( \mathbf{d} \) is the data, and \( \mathbf{m} \) is the model. In this paper, \( \mathbf{L} \) is the adjoint of the downward continuation migration operator explained by Prucha et al. (1999b). This migration algorithm produces a model with the axes depth \( (z) \), common reflection point (CRP), and offset ray parameter \( (p_h) \). Offset ray parameter is related to the reflection angle \( \theta \) (where \( \theta \) is half of the opening angle between incident and reflected rays) by:

\[
p_h = \frac{2 \sin \theta \cos \phi}{V(z, CRP)}.
\]

where \( \phi \) is the local dip and \( V(z, CRP) \) is the local velocity at the reflection point.
The minimization can be expressed more concisely as a fitting goal:

$$0 \approx Lm - d. \quad (3)$$

However, many issues exist, including poor illumination, that make this iterative inversion likely to have a large null space. Any noise that exists within that null space will not be constrained during the inversion and can grow with each iteration until the problem becomes unstable. Fortunately, we can stabilize this problem with regularization (Tikhonov and Arsenin, 1977). The regularization adds a second fitting goal that we are minimizing at the same time as we minimize the first one:

$$0 \approx Lm - d$$

$$0 \approx \epsilon Am. \quad (4)$$

The first expression in (4) is the “data fitting goal,” meaning that it is responsible for making a model that is consistent with the data. The second expression is the “model styling goal,” meaning that it allows us to impose some idea of what the model should look like using the regularization operator $A$. The strength of the regularization is controlled by the regularization parameter $\epsilon$.

Unfortunately, the inversion process described by fitting goals (4) can take many iterations to produce a satisfactory result. We can reduce the necessary number of iterations by making the problem a preconditioned one. We use the preconditioning transformation $m = A^{-1}p$ (Fomel et al., 1997; Fomel and Claerbout, 2003) to give us these fitting goals:

$$0 \approx LA^{-1}p - d$$

$$0 \approx \epsilon p. \quad (5)$$

$A^{-1}$ is obtained by mapping the multi-dimensional regularization operator $A$ to helical space and applying polynomial division (Claerbout, 1998). This makes our imaging method a Regularized Inversion with model Preconditioning (RIP).

THE REGULARIZATION OPERATOR

The regularization operator $A$ is designed to shape the model to conform to some expectations of its characteristics, particularly its covariance (Tarantola, 1986; Prucha et al., 2000). When using downward continuation migration as the linear imaging operator in the least-squares inversion problem, it is known that poor illumination will cause the resulting model to have gaps in events along the offset ray parameter (or the equivalent reflection angle) axis (Prucha et al., 2000). When attempting to image the subsurface where we know illumination is poor, it is reasonable to design $A$ to regularize amplitudes along the offset ray parameter ($p_h$) axis.

To create a simple regularization operator, I begin with the assumption that the velocity model used by the imaging operator is correct and therefore there is no moveout along the $p_h$ axis. Therefore, $A$ can act horizontally along the $p_h$ axis, penalizing sudden changes in amplitude. This is called “geophysical” regularization. Essentially, $A$ is a derivative operator.
It will act most forcefully on very large, sudden amplitude changes along the $p_h$ axis. This should compensate for the sudden holes caused by poor illumination without changing AVA trends that generally vary smoothly (Shuey, 1985).

Previous experiments using this geophysical regularization where illumination is poor have shown that it helps to “heal” the gaps along the $p_h$ axis (Prucha et al., 2000). While this is encouraging, the question of how much the geophysical regularization affects any genuine amplitude variation with offset ray parameter or angle (AVA) remains open. To answer this question, I have designed a simple test with known AVA and no illumination problems.

**AMPLITUDE EFFECTS**

The synthetic model I used to test the AVA effects of geophysical regularization is a simple subsurface with four flat layers (Figure 1). The velocity within each layer is constant and I assume a constant density throughout the model. I generated synthetic finite difference data from this velocity model. The expected AVA at each of the three interfaces can be calculated from the velocities of the adjacent layers.

![Figure 1: Simple velocity model.](image)

**Effects of migration operator**

I first investigated the effects of my downward continuation migration operator on the AVA. I migrated the synthetic data using the correct velocity model and extracted my calculated amplitudes along each of the three interfaces. These results can be seen in Figure 2. In each of the panels, the solid line represents the theoretical AVA values and the dots show the AVA values obtained after migration. The horizontal axis is shown in offset ray parameter ($p_h$), which is not an intuitively obvious physical unit. Based on the relationship in equation (2), the range of opening angles for the shallowest interface (left panel of Figure 2) is $0^\circ - 44^\circ$, the range for the second interface (center panel of Figure 2) is $0^\circ - 46^\circ$, and the range for the deepest interface (right panel of Figure 2) is $0^\circ - 68^\circ$. Overall, the results for each of the three interfaces are good. The calculated values for the second and third interfaces are not as
consistent as the shallowest interface, partly due to the presence of more migration artifacts at depth. In the case of the deepest interface, it is also an effect of the survey geometry: we cannot expect the deep event to have reliable amplitude information at large $p_h$ (larger than an opening angle of $36^\circ$) because the common midpoint and offset ranges of the seismic data are limited, so energy reflecting at large angles from the deep event are lost. However, the overall trends of all of the calculated AVA values are fairly accurate.

![Figure 2](marie1-ava.comp.mig) [CR]

**Effects of RIP**

Having obtained satisfactory AVA results with my downward continuation migration operator, I moved on to test my geophysically Regularized Inversion with model Preconditioning (RIP). For my first test, I set $\epsilon$ to zero. Although this essentially sets the model styling goal of fitting goals (5) to zero, I am preconditioning the problem so I am still regularizing the inversion through the $A^{-1}$ in the data fitting goal. Using this formulation, I ran tests using 5, 10 and 15 iterations. The results after 5 iterations can be seen in Figure 3. The results after 10 iterations are shown in Figure 4 and the results after 15 iterations are shown in Figure 5.

For all three exercises, we notice that the shallowest interface (left panels) has developed an artificial increasing trend. This is partly due to an edge effect at the small $p_h$, and partially due to the effects of the regularization. It is more pronounced for this interface because the true AVA trend is expected to be almost flat. The edge effect at small $p_h$ is present for the other two interfaces as well.

Looking at the results for the second interface (center panels of Figures 3 through 5), it appears that our results after regularized inversion are more accurate than the result we saw from the migration (center panel of Figure 2). For the second interface, other than the edge effect at small $p_h$, the AVA trend is quite accurate after 5, 10 and 15 iterations. The best result for the second interface appears to be that after 10 iterations. The result after 15 iterations is deteriorating slightly as the inversion is trying harder to accommodate artifacts that exist in the data.

The results for the deepest interface (right panels of Figures 3 through 5) are also better
than the result from the migration (right panel of Figure 2). They have the same problems at large $p_h$ due to survey geometry, and have the edge effect at small $p_h$ seen for the other interfaces, but the AVA trend between these two extremes is close to the expected trend. Due to the known problem at large $p_h$, I have actually chosen to turn the regularization operator off halfway along the $p_h$ axis to keep the inversion from spending all of its effort trying to correct the sudden decrease in amplitude. Once again, it seems that the AVA trend in the result after 10 iterations is the best.

Figure 3: Amplitude variation with $p_h$ for geophysical regularization with 5 iterations and $\epsilon = 0$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion. \[marie1-ava.comp.5it0eps\]

To examine the effect of a stronger regularization, I next set $\epsilon$ to .002. This means that the data fitting goal still influences the resulting model more than the regularization, but now the regularization is acting through the model styling goal as well as the data fitting goal. This value for $\epsilon$ was selected based on previous trial-and-error experiments in which poor illumination was a problem (Clapp, 2003). The results after 5 iterations of this test can be seen in Figure 6. The results after 10 and 15 iterations are in Figures 7 and 8, respectively. Since there are no sudden, large amplitude changes along the $p_h$ axis, this stronger regularization should not affect the results much more than the inversion results with $\epsilon = 0$. As expected, these results are similar to the results with $\epsilon = 0$. We see the same unfortunate edge effect at small $p_h$ for all of the interfaces in all three experiments (5, 10 and 15 iterations). We also see the effects of the survey geometry on the deepest interface (right panels of Figures 6-8), where I have once again elected to turn off the regularization operator. Overall, the increased strength of regularization has not affected the AVA trends of any of the interfaces any more than the previous inversion experiment.

CONCLUSIONS

The geophysically regularized least-squares inversion described in this paper is designed to minimize large, sudden changes in amplitudes along the offset ray parameter axis (or equivalent reflection angle axis). This behavior is based on the assumption that such changes are due to poor illumination, and true AVA trends should be relatively smooth in comparison. I have demonstrated that the derivative-type operator used for the regularization in the experiments in
Figure 4: Amplitude variation with $p_h$ for geophysical regularization with 10 iterations and $\epsilon = 0$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion.

Figure 5: Amplitude variation with $p_h$ for geophysical regularization with 15 iterations and $\epsilon = 0$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion.

this paper do leave the true AVA trends intact, with some error due to edge effects. This result is encouraging as it indicates that with some further work to reduce the edge effects, I can still extract reliable AVA information from images produced by my geophysically regularized inversion.

REFERENCES


Figure 6: Amplitude variation with $p_h$ for geophysical regularization with 5 iterations and $\epsilon = 0.002$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion.

Figure 7: Amplitude variation with $p_h$ for geophysical regularization with 10 iterations and $\epsilon = 0.002$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion.

Figure 8: Amplitude variation with $p_h$ for geophysical regularization with 15 iterations and $\epsilon = 0.002$. Left panel shows the results for the shallowest interface as seen in Figure 1, center panel for the second interface, and right panel for the deepest interface. The solid line shows the theoretical value and the dots show the values obtained after inversion.
Duquet, B., and Marfurt, K. J., 1999, Filtering coherent noise during prestack depth migration: 
Geophysics, 64, no. 4, 1054–1066.

Fomel, S., and Claerbout, J., 2003, Multidimensional recursive filter preconditioning in geo-
physical estimation problems: Multidimensional recursive filter preconditioning in geo-

Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter 
preconditioning: SEP–95, 15–25.

Kuehl, H., and Sacchi, M., 2001, Generalized least-squares dsr migration using a common 
Abstracts, 1025–1028.


Muerdter, D. R., Lindsay, R. O., and Ratcliff, D. W., 1996, Imaging under the edges of salt 

Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflec-
tion data: Geophysics, 64, no. 1, 208–221.


Prucha, M. L., Clapp, R. G., and Biondi, B. L., 1998, Imaging under the edges of salt bodies: 
Analysis of an Elf North Sea dataset: SEP–97, 35–44.

Prucha, M., Biondi, B., and Symes, W., 1999a, Angle-domain common image gathers by 
Abstracts, 824–827.

Prucha, M. L., Biondi, B. L., and Symes, W. W., 1999b, Angle-domain common image gathers 
by wave-equation migration: SEP–100, 101–112.

Prucha, M. L., Clapp, R. G., and Biondi, B., 2000, Seismic image regularization in the reflec-

Ronen, S., and Liner, C. L., 2000, Least-squares DMO and migration: Geophysics, 65, no. 5, 
1364–1371.

Shuey, R. T., 1985, A simplification of the zoeppritz-equations: A simplification of the 

Stolk, C. C., and Symes, W. W., 2004, Kinematic artifacts in prestack depth migration: Geo-
phys., 69, 562–575.


Target-oriented shot-profile one-way wave equation inversion

Alejandro A. Valenciano and Biondo Biondi

ABSTRACT

Least-squares shot-profile inversion could improve image amplitudes while remaining consistent with the data. This could be done efficiently by approximating the scalar two-way wave-equation operator by two one-way wave-equation operators: one for the source wavefield and the other for the receiver wavefield. In addition, a target-oriented scheme could help reduce the computation time. Instead of recursively computing the Green functions at each depth step for each inversion iteration, the Green functions from the surface to a target and from the target to the surface could be calculated and stored in the first iteration. In the subsequent iterations, the Green functions are retrieved from the disk.

INTRODUCTION

The uneven illumination of reflectors can lead to misinterpretation of their amplitudes after migration. This can be due to the seismic experiment acquisition geometry (incomplete data), or to energy focusing or defocusing caused by obstacles in the wave path. In the worst case, this problem can create shadow zones at reservoir depths (Muerdter et al., 1996; Prucha et al., 1998).

Some attempts to solve this subsurface imaging problem have used the power of geophysical inverse theory (Tarantola, 1987). It compensates for the experimental deficiencies (acquisition geometry, obstacles, etc.) while being consistent with the acquired data. Usually the inversion is implemented by using iterative least-squares algorithms (Nemeth et al., 1999; Duquet and Marfurt, 1999; Ronen and Liner, 2000; Prucha et al., 2000; Kuehl and Sacchi, 2001). The main inconvenient of this approach is that it is computationally expensive since it iteratively apply the seismic modeling and migration operators to build the whole image.

Since reflection amplitudes are more important at reservoir depths, we propose to apply a target-oriented shot-profile one-way wave equation inversion strategy. Instead of recursively computing the Green functions at each depth step and at each inversion iteration, they could be computed from the surface to the target and from the target to the surface during the first iteration, stored, and reused in subsequent iterations. In that way, the computational cost could be kept reasonable.

In this paper, we first discuss how a target-oriented least-squares inversion could help to

1email: valencia@sep.stanford.edu, biondo@sep.stanford.edu

305
improve the image amplitudes while keeping computational cost reasonable. We also review target-oriented modeling by shot-profile extrapolation with the one-way wave equation.

LINEAR LEAST SQUARES INVERSION

Tarantola (1987) formalizes the geophysical inverse problem from a Bayesian point of view. He gives a theoretical background to compensate for the reflection experiment’s deficiencies (acquisition geometry, obstacles, etc.), while being consistent with the acquired data.

Under this scheme, data and model (subsurface image) are assumed to have Gaussian distribution (a priori data and model probability density are Gaussian). The forward operator is assumed to be linear (or weakly non-linear). The resulting subsurface image is the mean of a posterior Gaussian probability density in the model space. Its expression for a general linear (forward modeling) operator $L$ is

$$\hat{m} = \left(L' C_D^{-1} L + C_M^{-1}\right)^{-1} \left(L' C_D^{-1} d_{obs} + C_M^{-1} m_{prior}\right),$$

(1)

where $\hat{m}$ is the mean of a posterior Gaussian probability density, $L'$ is the transpose of the linear forward operator, $C_D$ is the data covariance, $C_M$ is the model covariance, $d_{obs}$ is the measured data and $m_{prior}$ is the prior model.

Equation (1) can be solved by using gradient-based methods like steepest descent. This iterative algorithm can be written in the form $m_{n+1} = m_n + \delta m_n$ (Tarantola, 1987), where the term $\delta m_n$ depends on the gradient of the function to be optimized, a metric in the model space, and an ad-hoc constant.

Target-oriented least squares inversion strategy

In the case of wave equation migration or inversion, the operator $L$ is expensive to apply. Thus, iteratively applying this operator and its transpose is sometimes prohibitive. The computational cost is proportional to the number of depth steps the wavefields need to be propagated (Audebert, 1994), and the number of iterations, among other factors.

Since reflection amplitudes are more important in the neighborhood of the reservoir, it makes sense to apply a target-oriented strategy to reduce the number of depth steps. A way to achieve this objective is to write the modeling operator $L$ in a target-oriented fashion. Instead of recursively computing the Green functions at each depth step and at each inversion iteration, we can compute them from the surface to the target and from the target to the surface during the first iteration, store them, and reuse them in subsequent iterations. By storing the Green functions to the target and from the target, we add a new problem, since we then require approximately twice the disk space for data storage; however, we save computing time.

In the next sections, we write a target-oriented, one-way approximation to the scalar wave equation operator.
FORWARD OPERATOR

Ideally, the elastic wave equation should be used to model the propagation of seismic waves in the earth. But in exploration reflection seismology, except in very specific cases, only the vertical component of the propagated wavefield is recorded at the surface. This recorded wavefield is often assumed to be a scalar (converted modes projected in the vertical wavefield component are treated as noise). Thus a scalar wave equation is a good approximation of the physics involved in an exploration reflection seismology experiment.

Additionally, modeling or migrating with the scalar wave equation is computationally expensive. That is why it is rarely used for migration (Claerbout, 1985; Stolt and Benson, 1986). Instead, two one-way wave equations are used to mimic the scalar wave equation solution (Claerbout, 1971; Stolt and Benson, 1986). This approximation is reasonable in computational time and accuracy.

In the next three subsections we review the three steps involved in shot-profile modeling with the one-way wave equation: first, downward continuation of the source wavefield, second, upward continuation of the receiver wavefield to the surface, and third, truncation of the receiver wavefield at zero depth.

Source wavefield downward extrapolation

The first step of shot-profile modeling with the one-way wave equation consists of downward continuation of the source wavefield (for each shot, at each frequency), which is done by using the following equation:

\[ p^+(z_{i+j}) = w^+(z_{i+j}, z_i) \ p^+(z_i), \]  

initialized by the wavefield at the surface, as follows:

\[ p^+(z_0) = f, \]  

where \( p^+(z_i) \) is the source wavefield at depth \( z_i \), \( p^+(z_0) \) is the source wavefield at zero depth, \( w^+(z_{i+j}, z_i) \) is the downward continuation operator (Green function) that downward propagates the source wavefield from depth \( z_i \) to \( z_{i+j} \), and \( f \) is the source wavefield at the surface.

The Green functions can be computed recursively at each depth step (from \( z_i \) to \( z_{i+1} \)) (Figure 1), or precomputed and stored to allow jumps (from \( z_i \) to \( z_{i+j} \)) (Figure 2). The equations (2) and (3) can be written also in matrix form (see APPENDIX A).

Including all the frequencies and all the shot positions in the data, it follows from equation (A-3) that the source wavefield \((P^+)\) can be computed as a function of the source signature \((f)\), as follows:

\[ P^+ = (I - W^+)^{-1} f. \]  

(4)
Defining the the multi-frequency, multi-shot, downward propagation operator as $\mathbf{B}^+ = (\mathbf{I} - \mathbf{W}^+)^{-1}$, equation (4) can be written as

$$\mathbf{P}^+ = \mathbf{B}^+ \mathbf{f}.$$  (5)

**Receiver wavefield extrapolation**

The second step of shot-profile modeling consists of upward continuation of the receiver wavefield to the surface, which is done by using the following equation:

$$p^-(z_i) = w^-(z_i, z_{i+j}) p^-(z_{i+j}) + p^+(z_i) r(z_i) + p^-_r(z_i)$$  (6)

where

$$p^-_r(z_i) = \left\{ \begin{array}{ll} \sum_{m=1}^{j-1} w^-(z_i, z_{i+m}) p^+(z_{i+m}) r(z_{i+m}) & \text{if } j \neq 1, \\
0 & \text{otherwise,} \end{array} \right.$$  (7)

with the final condition

$$r(z_0) = 0,$$  (8)

where $p^-(z_i)$ is the receiver wavefield at depth $z_i$, $p^+(z_i)$ is the source wavefield at depth $z_i$, $w^-(z_i, z_{i+j})$ is the upward continuation operator (Green function) from depth $z_{i+j}$ to $z_i$, $r(z_i)$ is the reflectivity at depth $z_i$, and $p^-_r(z_i)$ is the contribution by the source wavefield and the reflectivity when jumps bigger than a depth step are used.
The Green functions can be computed recursively at each depth step (from \( z_{i+1} \) to \( z_i \)) (Figure 3), or precomputed and stored to allow jumps (from \( z_{i+j} \) to \( z_i \)) (Figure 4). In addition, the third term in equation (6) must also be stored. Equation (6) can be written also in matrix form (see APPENDIX B).

\[
\begin{align*}
p^-(z_0) & \quad \vdots \quad z_0 \\
p^-(z_{i-1}) + p^+(z_i) r(z_i) & \quad z_{i-1} \\
p^-(z_i) w^-(z_{i+1}, z_i) & \quad z_i \\
p^-(z_{i+1}) & \quad z_{i+1} \\
p^-(z_N) & \quad z_N
\end{align*}
\]

\[
\begin{align*}
p^-(z_0) & \quad \vdots \quad z_0 \\
p^-(z_{i-1}) + p^+(z_i) r(z_i) + p^-(z_i) & \quad z_{i-1} \\
p^-(z_i) w^-(z_{i+j}, z_i) & \quad z_{i+j} \\
p^-(z_{i+j}) & \quad z_{i+j} \\
p^-(z_N) & \quad z_N
\end{align*}
\]

Including all the frequencies and all the shot positions in the data, it follows from equation (B-4) that the receiver wavefield \( (P^-) \) can be computed as follows:

\[
P^- = [I - W^-]^{-1} P^+ \Sigma_{oas}^T \mathbf{r}.
\]  

(9)

Defining the the multi-frequency, multi-shot, upward propagation operator as \( B^- = (I - W^-)^{-1} \), equation (4) can be written as

\[
P^- = B^- P^+ \Sigma_{oas}^T \mathbf{r}.
\]  

(10)

**Linear forward operator \((L)\)**

Since in the conventional surface seismic experiment, geophones are located only at the surface, the data is the receiver wavefield at \( z = 0 \). This can be represented in equation (9) as a truncation operation (transpose of zero padding \( Z \)):

\[
d = Z' P^- = Z' B^- P^+ \Sigma_{oas}^T \mathbf{r}.
\]  

(11)

Then the expression for linear shot-profile modeling one-way wave equation operator is

\[
L = Z' B^- P^+ \Sigma_{oas}^T \mathbf{r},
\]  

(12)

a chain of linear operators.
SUMMARY

Target-oriented shot-profile least-squares inversion could help improve amplitudes while keeping computational cost reasonable. We propose a methodology that combines first, an approximation of the scalar two-way wave equation operator by two one-way wave equation operators, and second, precomputed Green functions to the target. This methodology will allow to perform the shot-profile least-squares inversion without having to recursively compute the Green functions for all the depth steps at each iteration.

REFERENCES


APPENDIX A

Source wavefield downward extrapolation

The recursion in equations (2) and (3) can be also written in matrix form as

\[(I - W^+) P^+ = F, \quad (A-1)\]

\[
\begin{bmatrix}
1 & 0 & 0 & \ldots & 0 & 0 \\
-w^+(z_i, z_0) & 1 & 0 & \ldots & 0 & 0 \\
0 & -w^+(z_{i+j}, z_i) & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -w^+(z_{N_z}, z_{N_z-1}) & 1
\end{bmatrix}
\begin{bmatrix}
p^+(z_0) \\
p^+(z_i) \\
p^+(z_{i+j}) \\
\vdots \\
p^+(z_{N_z})
\end{bmatrix}
= \begin{bmatrix}
f \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix},
\]

where

- \( W^+ \) is a lower bidiagonal matrix containing the downward continuation operator for all depth levels,
- \( P^+ \) is a column vector containing the source wavefield at all depth levels, and
- \( F \) is a column vector containing the source signature.

Equation (A-1) represents the downward continuation recursion written for a given frequency. We can write a similar relationship for each of the frequencies in the data, and group them all in a matrix relationship:

\[(I - W^+) P^+ = \mathcal{F}, \quad (A-2)\]

\[
\begin{bmatrix}
I - W^+(\omega_1) & 0 & \ldots & 0 \\
0 & I - W^+(\omega_2) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & I - W^+(\omega_{N_\omega})
\end{bmatrix}
\begin{bmatrix}
p^+(\omega_1) \\
p^+(\omega_2) \\
\vdots \\
p^+(\omega_{N_\omega})
\end{bmatrix}
= \begin{bmatrix}
F(\omega_1) \\
F(\omega_2) \\
\vdots \\
F(\omega_{N_\omega})
\end{bmatrix},
\]

where

- \((I - W^+)\) is a lower bidiagonal matrix containing the downward continuation operators for all the frequencies in the data,
- \( P^+ \) is a column vector containing the wavefield data for all the frequencies, and
- \( \mathcal{F} \) is a column vector containing the source signature for all the frequencies.
Equation (A-2) represents the downward continuation recursion written for a given shot position. We can write a similar relationship for each of the shot positions in the data, and group them all in a matrix relationship:

\[(I - W^+) P^+ = f, \quad \text{(A-3)}\]

\[
\begin{bmatrix}
1 - W^+(s_1) & 0 & \cdots & 0 \\
0 & 1 - W^+(s_2) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 - W^+(s_{N_s}) \\
\end{bmatrix}
\begin{bmatrix}
P^+(s_1) \\
P^+(s_2) \\
\vdots \\
P^+(s_{N_s}) \\
\end{bmatrix}
= 
\begin{bmatrix}
F(s_1) \\
F(s_2) \\
\vdots \\
F(s_{N_s}) \\
\end{bmatrix},
\]

where

- \((I - W^+)\) is a lower bidiagonal matrix containing the downward continuation operators for all the shots in the data,
- \(P^+\) is a column vector containing the wavefield data for all the shots, and
- \(f\) is a column vector containing the source signature for all the shots.

**APPENDIX B**

**Receiver wavefield extrapolation**

The recursion in equation (6) can be also written in matrix form as

\[(I - W^-) P^- = P^+ r, \quad \text{(B-1)}\]

\[
\begin{bmatrix}
1 & -w^-(z_0,z_i) & 0 & \cdots & 0 & 0 \\
0 & 1 & -w^-(z_1,z_{i+j}) & \cdots & 0 & 0 \\
0 & 0 & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & \cdots & 1 & -w^-(z_{N_z-1},z_{N_z}) \\
\end{bmatrix}
\begin{bmatrix}
p^-(z_0) \\
p^-(z_1) \\
\vdots \\
p^-(z_{i+j}) \\
p^-(z_{N_z}) \\
\end{bmatrix} = 
\begin{bmatrix}
p^+(z_0) \\
p^+(z_1) \\
\vdots \\
p^+(z_{i+j}) \\
p^+(z_{N_z}) \\
\end{bmatrix}
\begin{bmatrix}
r(z_0) \\
r(z_1) \\
\vdots \\
r(z_{i+j}) \\
r(z_{N_z}) \\
\end{bmatrix} + 
\begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
0 \\
\end{bmatrix},
\]

where

- \(W^-\) is a upper bidiagonal matrix containing the upward continuation operator for all depth levels,
• $P^-$ is a column vector containing the receiver wavefield at all depth levels,
• $P^+$ is a diagonal square matrix containing the source wavefield at all depth levels, and
• $r$ is the reflectivity at all depth levels.

Equation (B-1) represents the upward continuation recursion written for a given frequency. We can write a similar relationship for each of the frequencies in the data, and group them all in a matrix relationship:

$$(I - W^-) P^- = P^+ \Sigma^t \omega r,$$

where

• $(I - W^-)$ is a upper bidiagonal matrix containing the upward continuation operators for all the frequencies in the data,

$$(I - W^-) = \begin{bmatrix}
  I - W^-(\omega_1) & 0 & \ldots & 0 \\
  0 & I - W^-(\omega_2) & \ldots & 0 \\
  \ldots & \ldots & \ldots & \ldots \\
  0 & 0 & \ldots & I - W^-(\omega_{N_\omega})
\end{bmatrix},$$

• $P^-$ is a column vector containing the receiver wavefield for all the frequencies,

$$P^- = \begin{bmatrix}
  P^-(\omega_1) \\
  P^-(\omega_2) \\
  \vdots \\
  P^-(\omega_{N_\omega})
\end{bmatrix},$$

• $P^+$ is a diagonal square matrix containing the source wavefield for all the frequencies,

$$P^+ = \begin{bmatrix}
  P^+(\omega_1) & 0 & \ldots & 0 \\
  0 & P^+(\omega_2) & \ldots & 0 \\
  \ldots & \ldots & \ldots & \ldots \\
  0 & 0 & \ldots & P^+(\omega_{N_\omega})
\end{bmatrix},$$

• and $\Sigma^t_\omega$ is the sum over frequency matrix with dimensions $N_Z \times (N_Z \times N_\omega)$ (the transpose of the spreading over frequencies),

$$\Sigma^t_\omega = \begin{bmatrix}
  1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & \ldots & 1 & 0 & \ldots & 0 \\
  0 & 1 & \ldots & 0 & 0 & 1 & \ldots & 0 & \ldots & 0 & 1 & \ldots & 0 \\
  \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
  0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 1 & \ldots & 0 & 0 & \ldots & 1
\end{bmatrix}.$$
• \((I - W^-)\) is a upper bidiagonal matrix containing the upward continuation operators for all the shots in the data,
\[
(I - W^-) = \begin{bmatrix}
1 - W^-(s_1) & 0 & \ldots & 0 \\
0 & 1 - W^-(s_2) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 - W^-(s_N)
\end{bmatrix},
\]

• \(P^-\) is a column vector containing the receiver wavefield for all the shots,
\[
P^- = \begin{bmatrix}
P^-(s_1) \\
P^-(s_2) \\
\vdots \\
P^-(s_N)
\end{bmatrix},
\]

• \(P^+\) is a diagonal square matrix containing the source wavefield for all the shots,
\[
P^+ = \begin{bmatrix}
P^+(s_1) & 0 & \ldots & 0 \\
0 & P^+(s_2) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & P^+(s_N)
\end{bmatrix},
\]

• \(\Sigma_{\omega s}\) is the sum over the frequency matrix with dimensions \(N_Z \times (N_Z \times N_\omega \times N_s)\) (the transpose of the spreading over frequencies),
\[
\Sigma_{\omega s} = \begin{bmatrix}
1 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 & \ldots & 1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & 0 & 1 & \ldots & 0 & \ldots & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1 & 0 & 0 & \ldots & 1 & \ldots & 0 & 0 & \ldots & 1
\end{bmatrix}.
\]
Conjugate-guided-gradient (CGG) method for robust inversion and its application to velocity-stack inversion

Jun Ji

ABSTRACT

This paper proposes a modified conjugate-gradient (CG) method, called the conjugate-guided-gradient (CGG) method, as an alternative iterative inversion method that is robust and easily manageable. The CG method for solving least-squares (LS) (i.e. $L^2$-norm minimization) problems can be modified to solve for a different norm or different minimization criteria by guiding the gradient vector appropriately. The guiding can be achieved by iteratively weighting either the residual vector or the gradient vector during iteration steps. Weighting the residual vector can guide the solution to the minimum $L^p$-norm solution, and weighting the gradient vector can guide the solution to one constrained by a priori information imposed in the model space. In both cases, the minimum solutions are found in a least-squares sense along the gradient direction guided by the weights. Therefore, the solution found by the CGG method can be interpreted as the LS solution located in the guided gradient direction. I applied the CGG method to the velocity stack inversion, and the results suggest that the CGG method gives a far more robust model estimation than the standard $L^2$-norm solution, with results comparable to, or better than, an $L^1$-norm IRLS (Iteratively Reweighted LS) solution.

INTRODUCTION

The inverse problem has received considerable attention in various geophysical applications. One of the most popular inverse solutions is the least-squares (LS) solution. The LS solution is a member of a family of generalized $L^p$-norm solutions that are deduced from a maximum-likelihood formulation. This formulation allows the design of various statistical inversion solutions. Among the various $L^p$-norm solutions, the $L^1$-norm solution is more robust than the $L^2$-norm solution, being less sensitive to spiky, high-amplitude noise (Claerbout and Muir, 1973; Taylor et al., 1979; Scales and Gersztenkorn, 1987; Scales et al., 1988). However, the implementation of the algorithm to find $L^1$-norm solutions is not a trivial task. Iterative inversion algorithms called IRLS (Iteratively Reweighted LS) (Gersztenkorn et al., 1986; Scales et al., 1988) are a good choice for solving $L^p$-norm minimization problems for $1 \leq p \leq 2$. If the number of unknown model values is very large, LS problems are often solved by iterative solvers like the popular conjugate-gradient (CG) method. IRLS approaches for nonlinear inversion can be adapted to solve linear inverse problems by modifying the CG method (Darche,
This paper introduces a way to modify the CG method so that it can not only handle the general $L^p$-norm problem but also impose *a priori* constraints on the solution space. This method is called conjugate-guided-gradient (CGG), and is achieved by guiding the gradient vector during the iteration step. In the first section, I review the conventional CG method for solving LS problems and show how the IRLS approach differs from the standard LS approach. Next, I explain the CGG method and contrast it with both LS and IRLS. Finally I test the proposed CGG method on velocity-stack inversions with both noisy synthetic data and real data. I compare the results of the CGG method with conventional LS and $L_1$-norm IRLS results.

**CG METHOD FOR LS AND IRLS INVERSION**

Most inversion problems start by formulating the forward problem, which describes the forward operator, $L$, that transforms the model vector $m$ to the data vector $d$:

$$ d = Lm. \quad (1) $$

In general, the measured data $d$ may be inexact, and the forward operator $L$ may be ill-conditioned. In that case, instead of solving the above equation directly, different approaches are used to find an optimum solution $m$ for a given data $d$. The most popular method is finding a solution that minimizes the misfit between the data $d$ and the modeled data $Lm$. The misfit, or the residual vector, $r$, is described as follows:

$$ r = Lm - d. \quad (2) $$

In least-squares inversion, the solution $m$ is the one that minimizes the squares of the residual vector as follows:

$$ \min_m (r \cdot r) = \min_m (Lm - d)^T (Lm - d). \quad (3) $$

Iterative solvers for the LS problem search the solution space for a better solution in each iteration step, along the gradient direction (in steepest-descent algorithms), or on the plane made by the current gradient vector and the previous descent-step vector (in conjugate-gradient algorithms). Following Claerbout (1992), a conjugate-gradient algorithm for the LS solution can be summarized as follows:

\[
\begin{align*}
\text{iterate} & \quad r \leftarrow Lm - d \\
\Delta m & \leftarrow L^T r \\
\Delta r & \leftarrow L \Delta m \\
(m,r) & \leftarrow \text{cgstep}(m,r,\Delta m,\Delta r)
\end{align*}
\]

where the subroutine cgstep() remembers the previous iteration descent vector, $\Delta s = m_i - m_{i-1}$, where $i$ is the iteration step, and determines the step size by minimizing the quadrature...
function composed from $\Delta r$ (the conjugate gradient) and $\Delta s$ (the previous iteration descent vector), as follows (Claerbout, 1992):

$$Q(\alpha, \beta) = (r - \alpha \Delta r - \beta \Delta s)^T (r - \alpha \Delta r - \beta \Delta s).$$

Notice that the gradient vector ($\Delta m$) in the CG method for LS solution is the gradient of the squared residual and is determined by taking the derivative of the squared residual (i.e. the $L^2$-norm of the residual, $r \cdot r$, with respect to the model $m^T$):

$$\Delta m = \frac{\partial}{\partial m^T} (Lm - d)^T (Lm - d) = L^T r. \quad (4)$$

**Iteratively Reweighted Least Squares (IRLS)**

Instead of $L^2$-norm solutions obtained by the conventional LS solution, $L^p$-norm minimization solutions, with $1 \leq p \leq 2$, are often tried. Iterative inversion algorithms called IRLS (Iteratively Reweighted Least Squares) algorithms have been developed to solve these problems, which lie between the least-absolute-values problem and the classical least-squares problem. The main advantage of IRLS is to provide an easy way to compute the approximate $L^1$-norm solution. $L^1$-norm solutions are known to be more robust than $L^2$-norm solutions, being less sensitive to spiky, high-amplitude noise (Claerbout and Muir, 1973; Taylor et al., 1979; Scales and Gersztenkorn, 1987; Scales et al., 1988). The problem solved by IRLS is a minimization of the weighted residual in the least-squares sense:

$$r = W(Lm - d). \quad (5)$$

The gradient for the weighted residual in the least-squares sense becomes

$$L^T W r = \frac{\partial}{\partial m^T} (Lm - d)^T W^T W (Lm - d). \quad (6)$$

The particular choice for $W$ is the one that results in minimizing the $L^p$ norm of the residual. Choosing the $i^{th}$ diagonal element of $W$ to be a function of the $i^{th}$ component of the residual vector as follows:

$$\text{diag}(W) = |r|^{(p-2)/2}, \quad (7)$$

the norm of the weighted residual is then

$$r^T W^T W r = r^T |r|^{(p-2)} r = |r|^p. \quad (8)$$

Therefore, this can be thought of as a method that estimates the gradient in the $L^p$-norm of the residual. This method is valid for norms where $1 \leq p \leq 2$. When the $L^1$-norm is desired, the weighting is as follows:

$$\text{diag}(W) = |r|^{-1/2}. \quad (9)$$

This will reduce the contribution of large residuals and improve the fit to the data that is already well-estimated. Thus, the $L^1$-norm-based minimization is robust, less sensitive to noise bursts
in the data. Huber proposed a hybrid $L^1/L^2$-norm (Huber, 1973) that treats the small residuals in an $L^2$-norm sense and the large residuals in an $L^1$-norm sense. This approach deals with both bursty and Gaussian-type noise, and can be realized by weighting as follows:

$$\text{diag}(W) = \begin{cases} |r|^{-1/2}, & |r| > \epsilon \\ 1, & |r| \leq \epsilon \end{cases}$$

where $\epsilon$ is a value that is used as a threshold between $L^1$ and $L^2$-norms. IRLS can be easily incorporated in CG algorithms by including a weight $W$ such that the operator $L$ has a post-multiplier $W$ and the the adjoint operator $L^T$ has a premultiplier $W^T$ (Claerbout, 2004). Even though we do not know the real $L^p$-norm residual vector at the beginning of the iteration, we can approximate the residual with a residual of the previous iteration step, and it will converge to a residual that is very close to the $L^p$-norm residual as the iteration step continues. This can be summarized as follows:

$$r \leftarrow Lm - d$$

iterate {

$$W \leftarrow \text{diag}[f(r)]$$

$$\Delta m \leftarrow L^T W^T r$$

$$\Delta r \leftarrow W L \Delta m$$

$$(m, r) \leftarrow \text{cgstep}(m, r, \Delta m, \Delta r)$$

}.

CONJUGATE GUIDED GRADIENT(CGG) METHOD

Within the CG method, the IRLS algorithm can be considered as the LS method, but with its operator, $L$, modified by the weight, $W$. The only change that distinguishes the IRLS algorithm from the LS one is the substitution of $LW$ and $W^T L^T$ for $L$ and $L^T$, respectively. Instead of modifying the operator, we can choose a way to guide the minimizing search to find the minimum $L^2$-norm in a specific model subspace so as to obtain a solution that meets a user’s specific criteria. The specific model subspace could be guided by a specific $L^p$-norm’s gradient or constrained by an a priori model. Such guiding of the model vector can be realized by weighting the residual vector or gradient vector in the CG algorithm.

CGG with iteratively reweighted residual

If we apply the same weight $W$ we used in the IRLS, but do not change the operator from $L$ to $WL$, the weight affects only the gradient direction. This corresponds to guiding the gradient direction with a weighted residual, and the resultant weighted gradient will be the same gradient as we used in the IRLS method. This algorithm can be implemented as follows:

$$r \leftarrow Lm - d$$

iterate {


\[ W \leftarrow \text{diag}[f(r)] \]
\[ \Delta m \leftarrow L^T W^T r \]
\[ \Delta r \leftarrow L \Delta m \]
\[ (m, r) \leftarrow \text{cgstep}(m, r, \Delta m, \Delta r) \]
\]

Notice that the above algorithm is different from the original CG algorithm only at the step of gradient computation; the modification of the gradient is performed by changing the residual before the gradient is computed from it. By choosing the weight as a function of the residual of the previous iteration step, as we did in the IRLS, we can guide the gradient to the gradient of the \( L^p \)-norm. Thus the result obtained by weighting the residual could be interpreted as an LS solution located along the direction of the \( L^p \)-norm gradient, according to the weight applied. If, during the iteration, any intermediate solution is found at the minimum \( L^2 \)-norm location in the model space, it will be the final solution of the algorithm, and it is the same as the solution of the conventional LS problem. However, the minimum \( L^2 \)-norm location is unlikely to fall along the gradient of the different \( L^p \)-norm determined by the applied weight. Therefore, it is more likely that the solution will be close to the minimum \( L^p \)-norm location determined by the applied weight.

**CGG with iteratively reweighted gradient**

Another way to modify the gradient direction is to modify the gradient vector after the gradient is computed from a given residual. Since the gradient vector is in the model space, any modification of the gradient vector imposes some constraint in the model space. If we know some characteristics of the solution which can be expressed in terms of weighting in the solution space, we can use that \textit{a priori} knowledge to redirect the gradient vector by applying a weight to it. This algorithm can be implemented as follows:

\[ r \leftarrow L m - d \]
iterate {
\[ W \leftarrow \text{diag}[f(m)] \]
\[ \Delta m \leftarrow W L^T r \]
\[ \Delta r \leftarrow L \Delta m \]
\[ (m, r) \leftarrow \text{cgstep}(m, r, \Delta m, \Delta r) \]
}

Even though weighting the gradient has different meaning from weighting the residual, the analysis is similar in both cases. As we redefined the contribution of each residual element by weighting it with the absolute value of itself to some power: we can do the same with each model element in the solution,

\[ W = |m|^p, \quad (9) \]

where \( p \) is a real number that depends on the problem we wish to. When we have a finite model space we are applying a uniform weight to the finite model space and zero weight to
the outlying space. If the operator used in the inversion is close to unitary, the solution obtained
after the first iteration already closely approximates the real solution. Therefore, weighting the
gradient with some power of the absolute value of the previous iteration means that we down-
weight the importance of small model values and improve the fit to the data by emphasizing
model components that already have large values.

CGG with iteratively reweighted residual and gradient

In the previous two subsections, we examined the meaning of weighting the residual and
the gradient vector, respectively. Since applying the weighting in both residual space and
model space is nothing but changing the direction of the descent for the solution search, the
weighting is not limited either to residual or to model space. We can weight both the residual
and the gradient,

\[
\begin{align*}
\mathbf{r} & \leftarrow \mathbf{Lm} - \mathbf{d} \\
\text{iterate} \{ & \\
\mathbf{W}_r & \leftarrow \text{diag}[f(r)] \\
\mathbf{W}_m & \leftarrow \text{diag}[f(m)] \\
\Delta \mathbf{m} & \leftarrow \mathbf{W}_m \mathbf{L}^T \mathbf{W}_r \mathbf{r} \\
\Delta \mathbf{r} & \leftarrow \mathbf{L} \Delta \mathbf{m} \\
(\mathbf{m}, \mathbf{r}) & \leftarrow \text{cgstep}(\mathbf{m}, \mathbf{r}, \Delta \mathbf{m}, \Delta \mathbf{r}) \\
\} .
\end{align*}
\]

Again, the above CGG algorithm is different from the conventional CG method only in the
step of gradient computation. Whether we modify the gradient in the residual sense or in the
model sense, it changes only the gradient direction, or the direction in which the solution is
sought. Therefore the CGG algorithm always converges to a solution.

APPLICATION OF THE CGG METHOD IN VELOCITY-STACK INVERSION

In this section, the CGG method is tested on a velocity-stack inversion, which is useful not
only for velocity analysis but also for various data processing. The applications of the velocity-
stack inversion include separating multiples from the signal in the velocity domain (Lumley et
al., 1995; Kostov and Nichols, 1995), multiple-attenuation techniques using parabolic Radon
transforms (Kabir and Marfurt, 1999; Herrmann et al., 2000), missing-trace interpolation
in the CMP domain (Ji, 1994), and so on. In these applications, the velocity-stack panels
obtained by inversion are usually required to be as spiky and sparse as possible. Then the
hyperbolic events represented by the isolated peaks in the velocity-stack panel are more easily
distinguished from the rest of the noise. The conventional velocity stack is performed by
summing or estimating semblance (Taner and Koehler, 1969) along the various hyperbolas in
a CMP gather, resulting in a velocity-stack panel. Ideally a hyperbola in a CMP gather should
be mapped onto a point in a velocity-stack panel. Summation along a hyperbola, or hyperbolic
Radon transform (HRT), does not give such resolution. To obtain a velocity-stack panel with better resolution, Thorson and Claerbout (1985) formulated it as an inverse problem in which the velocity domain is the unknown space. If we find an operator $H$ that transforms a point in a model space (velocity-stack panel), $m$, into a hyperbola in data space (CMP gather), $d$,

$$d = Hm,$$  \hspace{1cm} (10)

and also find its adjoint operator $H^T$, we can pose the velocity stack problem as an inverse problem. Inverse theory helps us to find a velocity-stack panel which synthesizes a given CMP gather via the operator $H$. The usual process is to implement the inverse as the minimization of a least-squares problem and calculate the solution by solving the normal equation:

$$H^T Hm = H^T d.$$  \hspace{1cm} (11)

Since the number of equations and unknowns may be large, an iterative least-squares solver such as CG is usually preferred to solving the normal equation directly. The least-squares solution has some attributes that may be undesirable. If the model space is overdetermined and has bursty noise in data, the least-squares solutions usually will be spread over all the possible solutions. Other methods may be more useful if we desire a parsimonious representation. To obtain a more robust solution, Nichols (1994) used the IRLS method for $L^1$-norm minimization, and Guitton and Symes (2003) used the L-BFGS method for Huber-norm minimization. Another possibility is the CGG method proposed in the preceding section. In the next subsections the results of the CGG method for the velocity-stack inversion are compared with the results of conventional LS and $L^1$-norm IRLS.

**Examples on synthetic data**

To examine the performance of the proposed CGG method, a synthetic CMP data set with various types of noise is used. Figure 1 shows the synthetic data with three types of noise — Gaussian noise in the background, bursty noise, and a very noisy trace. Figure 1 (right) is the same data as Figure 1 (left), but displayed in wiggle format to clearly show the bursty noise that was not discernable because of the clipping in the raster-format display. The amplitudes of the three bursty spikes are eight times the maximum amplitude of the hyperbolas. Figure 2(a) shows the modeled data from the velocity-stack panel obtained using the conventional CG algorithm for LS solution. The inversion was obtained with 30 iterations and the same number of iteration was used for all the other examples including the real data cases. We can clearly see the limit of $L^2$-norm minimization. The noise with Gaussian statistics is removed quite well, but some spurious events are generated around the bursty noise spikes and noisy trace. Figure 2(b) shows the modeled data from the velocity-stack panel obtained using the IRLS algorithm in an $L^1$-norm sense. In Figure 2(b) we can see the robustness of $L^1$-norm minimization. The bursty noise is reduced significantly, but the removal of the background noise seems to be worse than the result of $L^2$-norm minimization. Figure 2(c) shows the modeled data from the velocity-stack panel obtained using the CGG method, with the iteratively reweighted residual in an $L^1$-norm sense. The result is comparable to the IRLS inversion (Figure 2(b)). This tells us that guiding the gradient vector toward the $L^1$-norm gradient gives a solution similar to the $L^1$-norm solution with the IRLS method. Figure 2(d) shows the modeled data from
the velocity-stack panel obtained using the CGG method with iteratively reweighted gradient as follows:

$$\text{diag}(W_i) = |m_{i-1}|^{1.5},$$

that is the diagonal element of the i-th iteration weighting matrix is the absolute value of the model vector from the previous iteration, raised to the power 1.5. The result shows that the Gaussian background noise and the very noisy trace are better removed than with any of the $L^1$-norm approaches, either in CGG or IRLS. However, some spurious events around the bursty spikes still exist. This is because of the high amplitude of the noise. Since the weight is dependent on the amplitude of the model vector, and because the high amplitude in the CMP gather is also mapped to a high amplitude in the velocity-stack panel, the bursty noise with high amplitude would have higher weighting than the noise with low amplitude. However, this kind of artifact would easily be removed if the bursty noise had an amplitude similar to the rest of signal, which is the case when AGC (automatic gain control) is applied to the data. Figure 3 shows the modeled data obtained using the CGG algorithm with both residual and gradient weightings. In this case, the result looks like modeled data without any noise, because the bursty noise is reduced with residual weighting ($L^1$-norm criteria), and background noise is removed with gradient weighting. The velocity stacks obtained from the various inversions are shown in Figure 4. From left to right, the velocity-stack panels correspond to the results from LS inversion, IRLS inversion, CGG with residual weighting, CGG with gradient weighting, and CGG with both residual and gradient weighting. From these velocity-stack panels, we can deduce why different inversion methods were successful with the different noise styles. If we want an application that distinguishes signals in the model space, we can see that gradient weighting is the preferred method, because it gives a more parsimonious representation than the others.
Figure 2: Remodeled data from the result of the LS inversion (a), from the result of the IRLS inversion with an $L^1$-norm minimization (b), from the CGG method with iteratively reweighted residuals (c) in an $L^1$-norm sense, and from the result of the CGG method with iteratively reweighted gradients (d). [jun1-fig2] [CR]

Figure 3: Remodeled data from the CGG method, with iteratively reweighted residual and gradient together. [jun1-fig3] [CR]
Examples on real data

I tested the proposed CGG method on two real data sets that contain various types of noise. The data are shot gathers from land surveys. The trajectories of the events from both data sets look "hyperbolic" enough to be tested with hyperbolic inversion. Figure 5 shows one of the real data sets tested. The noise in the data is mainly the strong ground roll, the amplitude anomalies early at near-offset and late at 0.8km offset, and the time shift around offsets 1.6 km and 2.0 km. Figure 6(a) shows the modeled data obtained using the IRLS algorithm in an $L^1$-norm sense. Figure 6(b) shows the modeled data obtained using the CGG algorithm with residual weighting in an $L^1$-norm sense. In both panels in Figure 6, the noise is greatly reduced, especially the ground roll, since the limited-size model space does not include the ground-roll velocity. However, we can see that the IRLS method reduce the signal at far offset too much and the signal is preserved better by CGG with residual weighting (Figure 6(b)) than with IRLS inversion (Figure 6(a)). Figure 7(a) shows the modeled data from the result obtained using the CGG algorithm with gradient weighting. The following weighting factor was used:

$$\text{diag}(W_i) = |m_{i-1}|^{1.5},$$
that is the diagonal element of the i-th iteration weighting matrix is the absolute value of the (i-1)-th iteration model vector raised to the power 1.5. Comparing this with Figure 6(b) shows that weighting the gradient directly could result in better noise reduction at far offset than CGG with residual weighting. Figure 7(b) shows the modeled data obtained using the CGG algorithm with both residual and gradient weightings. In this case, the result is very similar to that of gradient-only weighting (Figure 7(a)), and it tells us that most of the noise could removed by weighting only the gradient. The velocity stacks obtained from the various inversions are shown in Figure 8. As we saw in the synthetic data example, the velocity stacks obtained by the CGG method that includes gradient weighting, Figure 8 (c) and (d), show more parsimonious representation than the others. Figure 9 shows another real data set used for testing. The noise in the data can be characterized by anomalous shifts and noisy amplitude at near offset around 1.5 sec, several junk traces around middle offset, and widespread noise. Figure 10(a) shows the modeled data obtained using the IRLS algorithm in an $L_1$-norm sense. Figure 10(b) shows the modeled data obtained using the CGG algorithm with residual weighting in an $L_1$-norm sense. In both results, most of the noise is greatly reduced. However, we can see that some of the noise characterized by anomalous amplitude is better reduced by CGG with residual weighting (Figure 10(b)) than by IRLS inversion (Figure 10(a)) and the IRLS inversion shows overly reduced amplitude at far offset signal, again. We can see that for the same iteration numbers, residual-only weighting is more effective than IRLS style residual weighting. Figure 11(a) shows the modeled data obtained using the CGG algorithm with gradient weighting. The following weighting factor was used:

$$\text{diag}(W_i) = |m_{i-1}|^2,$$

that is the diagonal element of the i-th iteration weighting matrix is the square of the (i-1)-th iteration model vector. Comparing this with Figure 10(b) shows the interesting result that weighting the gradient directly could result in noise reduction similar to that of residual weighting. Figure 11(b) shows the modeled data obtained using the CGG algorithm with both residual and gradient weightings. In this case, the result differs little from that of gradient-only weighting (Figure 11(a)), and most of the noise could removed by weighting only the gradient. The velocity stacks obtained from the various inversions are shown in Figure 12. As we saw in the synthetic data example, the velocity stacks obtained by the CGG method with gradient weighting show more parsimonious representation than the others.

**CONCLUSIONS**

The proposed CGG (Conjugate Guided Gradient) inversion method is a modified CG (Conjugate Gradient) inversion method, which guides the gradient vector during the iteration and allows the user to impose various constraints for residual and model. The guiding is implemented by weighting the residual vector and the gradient vector, either separately or together. Weighting the residual vector with the residual itself corresponds to guiding the solution search toward the $L_p$-norm minimization; weighting the gradient vector with the model itself corresponds to guiding the solution search toward a priori information imposed. Testing the CGG algorithm for the velocity-stack inversion of noisy synthetic and real data demonstrates that
Figure 5: The real data used for the inversion. Notice the strong ground roll, the amplitude anomalies early at near-offset and late at 0.8km offset, and the time shift around offsets 1.6 km and 2.0 km.
Figure 6: Remodeled data from the IRLS inversion with $L^1$-norm minimization (a) and from the result of the CGG method with iteratively reweighted residual in an $L^1$-norm sense (b).
Figure 7: Remodeled data from the CGG method with iteratively reweighted gradient (c) and from the CGG method with iteratively reweighted residual and gradient together (d).
Figure 8: Velocity-stack panels obtained by various inversions. From left to right: IRLS, CGG with residual weighting, CGG with gradient weighting, and CGG with both residual and gradient weighting. [jun1-fig4v][CR]
Figure 9: The real data used for the inversion. Notice the amplitude anomalies and time shift at near-offset traces, unrealistic junk traces at 1.8 km and 2.7 km, and widespread noise.
Figure 10: Remodeled data from the IRLS inversion with $L^1$-norm minimization (a), and remodeled data from the CGG method with iteratively reweighted residual in the $L^1$-norm sense (b).
Figure 11: Remodeled data from the CGG method with iteratively reweighted gradient (c), and from the CGG method with iteratively reweighted residual and gradient together (d).
Figure 12: Velocity-stack panels obtained by various inversions. From left to right: IRLS, CGG with residual weighting, CGG with gradient weighting, and CGG with both residual and gradient weighting. [jun1 Fig5v] [CR]
residual weighting appears to be comparable to or better than IRLS for the $L^1$-norm solution. Gradient weighting produces a more spiky velocity spectrum than any of the $L^p$-norm solutions, which are preferable for velocity picking. Therefore, I think the CGG method is a possible alternative to the more traditional IRLS method for robust inversion of seismic data.

REFERENCES


Short Note

Dynamic programming and trace alignment: Part 2

Chris Liner and Robert G. Clapp

INTRODUCTION

Dynamic programming is an effective tool for finding a solution for certain types of relatively small, non-linear problems. In biology, dynamic programming is used for pairwise alignment of amino acid sequences (Needleman and Wunsch, 1970). In electrical engineering, it is used for error correction in wireless communication and speech recognition (Hosom et al., 1999) among many other things. We can also find examples of its use in geophysics. Kruse (1988) used dynamic programming for signal correlation and trace interpolation. He calculates an error function based on the difference in instantaneous frequency between all points along two signals. Dynamic programming is then used to find the error path with the least energy. Zhang (1991) used it for a starting solution when doing event picking.

In this paper we will continue the work started in Liner and Clapp (2002). We use dynamic programming for trace alignment. The Needleman-Wunsch algorithm (Needleman and Wunsch, 1970) proves effective when we expect significant stretch between compared traces, while alternate dynamic programming approaches can be more effective when we expect smoother variations and desire continuity between alignments.

In this paper we begin by discussing the basic tenets of dynamic programming and the specific features of both algorithms. We show how to extend the algorithms to seismic trace alignment and apply them on several different types of alignment problems.

DYNAMIC PROGRAMMING

Dynamic programming offers a way to solve certain classes of non-linear problems. It is useful for problems that can be thought of as making one decision after another. To understand how it works, it is easiest to start with our final decision. Our goal is to maximize our ‘score’ $S$ of a series of decision $1...n$. Each decision has several potential outcomes $1...m$. Our final score is going to be based on some score we have calculated for all of our possible options ‘states’ at $j = n - 1$, and the best score we can get from moving from all of the possible states at $n - 1$
to all possible states at $n$. We can write the score as

$$S(i, j) = \max_{k=1}^{m}[S(k, j-1) + v(i, j, k)],$$

where $i$ is the given state and $v(i, j, k)$ is the value obtained from moving from state $k$ to state $i$ at decision $j$. The best series of decisions is then found by going backwards, taking the state at each decision $i$ that corresponded to the highest score.

In this most general form, the algorithm is quite expensive. The cost is on order $n \times m^2$. For a large number of states the problem quickly becomes impractical. The easiest way to reduce the cost is to limit the number of states that we must search when moving from one decision to another. The next two subsections describe ways to limit the search that are practical for certain classes of problems.

**Needleman-Wunsch algorithm**

The Needleman-Wunsch (NW) algorithm (Needleman and Wunsch, 1970) is a method that was developed for amino acid sequence alignment in proteins. This was the first of many important alignment techniques which now find application in the Human Genome Project.

Human DNA consists of some 30,000 genes which in turn composed of 20 amino acids represented by letters of a reduced alphabet (ADCEFGHILKMNPQRSTVWY). The total genome is composed of about 3 billion letters, or 100,000 per gene. Finding where a particular string of amino acids fits on a protein is an optimization problem that aims to find the optimal alignment of two character strings with respect to a defined set of rules and parameter values for comparing different alignments.

In terms of dynamic programming, we have two strings of length $n$ and $m$ that we would like to align. The key observation of the algorithm, and what makes it a computationally attractive process, is that the strings can be compressed and stretched but must remain in order. As a result, when comparing the two strings we have only to evaluate three possible states: a match, a compression, or a dilation.

To implement the NW algorithm we start by constructing a `similarity' matrix $\sigma$ of size $n$ by $m$. In the genome example, the matrix is going to have binary values: either the amino acid matches (1) or it doesn’t (0). We then construct our score matrix. The optimality of the path is defined by our similarity measurement $\sigma$. At each point in the scoring matrix, we evaluate whether the optimal path is from a point below ($E$), a point to the right ($F$), or a point along the diagonal ($G$). Our preferred solution is to move along the diagonal (we want to give preference to finding as much similarity between the two strings as possible) so we penalize the other two options. The scoring matrix ($V$) will then be recursively built by selecting the maximum of $E$, $F$, and $G$. We can write this in terms of the following equations,

$$V(i, j) = \max\left[G(i, j), F(i, j), E(i, j)\right]$$

$$G(i, j) = \sigma(x_i, y_j) + V(i + 1, j + 1)$$

$$F(i, j) = -(p + q) + \max\left[V(i + 1, j), F(i + 1, j) + p\right]$$

$$E(i, j) = -(p + q) + \max\left[V(i, j + 1), E(i, j + 1) + p\right].$$
The parameters \((p, q)\), often referred to as ‘gap’ penalties, are an important aspect of the algorithm. They allow the user to control how discontinuous a path to allow through the similarity matrix.

The NW algorithm has some nice properties. The cost is relatively low, \(n \ast m \ast 3\). In addition, it can handle very discontinuous traces. This can also be a disadvantage, sometimes we would like to force a greater level of continuity than the gap penalties allow.

**Extension to seismic trace alignment**

There are two obvious problems with a straightforward implementation of the NW algorithm. Both are related to the fact that our data is continuous, rather than discrete. We do not have the limited alphabet of the DNA example, so our similarity matrix can’t simply be matching of values. Therefore, a more appropriate measure of similarity in the seismic case is a short-window correlation. We capture this idea in a similarity function as

\[
\sigma(i, j) = \frac{\sum_{a=-n}^{n} x_{i+a} y_{j+a}}{\sum_{a=-n}^{n} x_{i+a} \sum_{a=-n}^{n} y_{j+a}}
\]

(6)

or a semblance measure

\[
\sigma(i, j) = \frac{\sum_{a=-n}^{n} (x_{i+a} y_{j+a})^2}{(\sum_{a=-n}^{n} x_{i+a})^2 (\sum_{a=-n}^{n} y_{j+a})^2}^{1/2}
\]

(7)

where \(x\) and \(y\) are again the two traces, and \(n\) is the correlation length.

The second problem is how to use our alignment data. We are not dealing with discrete points but smooth functions. We can estimate a prediction error filter (Claerbout, 1998) or a time variant, non-stationary prediction error filter upon the original data to describe the wavelet. We have adopted this approach, allowing us to estimate our aligned model \(m\) from our unaligned data \(d\) by minimizing

\[
\min_m = ||m - d||^2 + \epsilon^2 ||Am||^2
\]

(8)

where \(A\) is filtering with the estimated prediction error filter and \(\epsilon\) is a scalar controlling how much weight to give the aligning. This controls the relative importance of the alignment versus fitting the wavelet.

**NW EXAMPLES**

To illustrate the application and results of NW seismic alignment, we present three examples. The first is alignment of events in synthetic P-wave and S-wave zero-offset sections. The second is flattening of events in a common image gather (CIG) from a marine 2D seismic
Figure 1: The left panel (a) shows a synthetic 3-layer P-wave section. The right panel (b) shows the S-wave section for the same earth model. Both simulations represent zero-offset data.

In the final case we attempt to align a real P- and S-wave section. For the first example, two synthetic zero-offset seismic sections (Figure 1) were generated by 2D acoustic Kirchhoff modeling. The data is moderately complex, containing low and steep dips, diffractions, and a buried focus. The subsurface reflector geometry is identical in each case. The S-wave section (Figure 1b) was modeled using the P-wave velocities and a VS-VP ratio that varied both laterally and with depth. This means a spatially-variant, nonlinear stretch is needed to align the sections. The alignment algorithm works here in a pairwise fashion on each trace pair that has the same trace number in each section. Every alignment is calculated independently, so there is no error propagation in this case.

The alignment result (Figure 2) represents the P-wave section after application of nonlinear stretch coefficients. We could just as easily have aligned the S-wave section to the P-wave data. The gap penalties for this example are \((p,q) = (0.0,0.1)\) and \(\epsilon = 10\). Comparison of the S-wave section (Figure 1) with the alignment section demonstrates the accuracy and sensitivity of this method. Considering the complexity of the input data, we feel the alignment is of excellent quality. For a second example, we chose a common reflection point gather from a 2D marine data set (left panel of Figure 3). This collection of traces forms an common image gather in the angle domain (Prucha et al., 1999; Sava and Fomel, 2000) after phase-shift plus-
Figure 2: The left panel (a) shows the P-wave section after trace alignment based on the S-wave section from Figure 1a. The right panel (b) displays the S-wave section again for comparison.

interpolation (PSPI) migration (Gazdag and Sguazzero, 1985). Note that we see some residual moveout in the CIG. The number of events, low-slope residual moveout, and lateral variations make this a significant alignment problem. The goal here is to align events in the gather with the first (zero-angle) trace. Constructing the aligned model in this case is far more complex than the procedure necessary for independent pairwise trace alignments. The effect of aligning trace two to trace one affects the alignment of traces three and two, etc. In our view, there are three potential ways around this connectivity. The first is to do the alignments sequentially. Align trace n-1 to trace n, then align both n and n-1 to n-2, and repeat the procedure until all traces are aligned. This procedure can be effective but is prone to error propagation. The second is to recognize that all possible pairs of traces can be aligned to each other and setup an inverse problem that takes all possible alignments into account. This would probably lead to the best solution but how best to relate the different combinations isn’t obvious and the computational cost is significantly increased (we now have to compute and account for $\frac{n(n-1)}{2}$ comparisons). We chose a third approach to multi-trace alignment involving a linear operator that performs sequential shifting. The idea is to begin at an initial trace (for example, the near offset) and to work outward toward the final trace (far offset). As each trace pair is aligned, these alignment coefficients (ACs) are applied to all remaining traces. In this way, an intermediate trace receives many partial alignments, and one last alignment from its nearest
neighbor which itself has been modified several times. The algorithm can be summarized as follows:

```c
for trace i from 1 to (n-1) {
    for trace j from (i+1) to n {
        modify trace j using ACs between traces i and (i+1)
    }
}
```

This method evolved from extensive testing of alternative approaches, each of which introduced serious error propagation problems.

In this case, single application of the algorithm was not sufficient. In order to obtain a reasonable solution, a larger gap penalty \((p,q) = (0.1,0.5)\) was found to be necessary, along with \(\epsilon = 100\) as in the CMP test. The resulting section still had residual moveout. As a result we reversed the output of the first alignment along the time axis and used this as input to a second run of the alignment procedure. Total compute time on a workstation-class machine was 25 seconds. The right panel of Figure 3 shows the result after the two-pass alignment. The flattening is not perfect, but the events are much better aligned than before while maintaining the same amplitude versus angle (AVA) characteristics.

For the final example, we attempted to align a PP and PS stacked section from the Alba field in the North Sea. Figure 4 shows the PP and PS section before alignment. This problem poses significant additional problems compared to the synthetic used in the first example. We are dealing with much higher frequency data than the first example and have many more events. The left panel of Figure 5 shows the output of the alignment procedure. Note the jumps in the sections, and the gaps at certain CMP locations. The first problem is an inherent property of the algorithm. The penalty we can impose for jumps is limited. The second problem is because we are performing a series of independent alignments. The result of the previous alignment has no affect on the current alignment. As a result we get the inconsistencies seen in the image.

**BACK TO DYNAMIC PROGRAMMING**

There are several ways we can try to improve the result seen in the left panel of Figure 5. We can think of smoothing the paths through the score matrix. We can also make some reasonable *a priori* assumptions. We can add to our similarity matrix. We might put large weights in the matrix by saying that time 0 should be the same in the PP and PS section or by visually matching a single reflector. These approaches improve our result, but still lead to something less than ideal.

A better solution is to return to a more generic form of dynamic programming. We can improve our 1-D estimates by remembering that what we are really doing is estimating a \(\frac{v_p}{v_s}\) ratio. This ratio has to be a continuous function of time, as a result it is not reasonable to have large gaps in our trace alignment. First, if we consider the traces to be samples in a PP gather,
Figure 3: A CIG gather after PSPI migration and conversion to angle. Note that the gather still has significant residual moveout. The CIG gather shown in the right panel after multi-trace alignment. Note how almost all residual moveout has been removed while preserving relative amplitude variations.

we can assume that the alignment function will be single valued. As a result, we can make our series of decisions one PP time element at a time. Changing our decision definition takes away the nice feature of the NW algorithm that we only have to consider three states at each decision. On the other hand, we can limit ourselves to considering the subset of states that correspond to a reasonable $\frac{v_p}{v_s}$ path. We know that each PP sample must correspond to the previous PP sample $+x$, where $x$ is a very small number. As a result, we again need to only consider two to four states at each decision. The right panel of Figure 5 shows the result of applying this dynamic programming concept, along with the constraints discussed above, to the Alba alignment problem. Note how the solution is better than the result shown in the left panel of Figure 5, but still could be improved.

CONCLUSION

We have shown how dynamic programming can be applied to a series of seismic trace alignment problems. For many cases, a straight implementation of an algorithm developed for protein sequences is desirable. For problems with relatively high coherent noise, and when extending from 1-D to multiple dimensions, more complex dynamic programming algorithms
Figure 4: The left panel shows a PP section from the Alba field in the North Sea. The right panel shows a PS section from the same line. [bob1-alba.in][ER]

can be useful.

With further work, this approach may supply a general tool for nonlinear alignment of seismic traces for use in processing and interpretation.

**ACKNOWLEDGMENTS**

We would like to thank ChevronTexaco for providing the data and Daniel Rosales for the preliminary processing of the data used in this paper.

**REFERENCES**

Claerbout, J., 1998, Geophysical Estimation by Example: Environmental soundings image enhancement:

Figure 5: The PP section from the Alba field in the North Sea after alignment. Note the jumps in both time and CMP.


Regularizing Madagascar: PEFs from the data space?

William Curry

ABSTRACT
The Madagascar seasat dataset presents a problem where data are collected along crossing tracks. These tracks are not straight, and appear to be irregular in the model space. Previous methods assumed that the data were regularly sampled in the model space coordinate system. I warp the model space to look more like the data space, so that prediction-error filters can be estimated in the more regularly-sampled data space. I try two different approaches to this problem on the warped space, one where the data is single-valued and fixed, while the other is multivalued and allowed to vary. The former method works very well, while the second one works well.

INTRODUCTION
The Madagascar seasat sea level dataset is a collection of two passes (ascending and descending) of the GEOSAT satellite over a region of the Southwest Indian Ridge in the Indian Ocean. There is a densely-acquired region of the dataset in the south, which ranges from 40 to 70 degrees (E) longitude and 30 to 40 degrees (S) latitude, while the latitude of the sparsely-acquired data ranges from 20 to 40 degrees (S) latitude.

The satellite tracks are much like feathered marine geophone cables, sail lines, or shot lines in a 3D seismic survey. Any method that hopes to succeed on 3D seismic data should be able to deal with this toy problem.

There are several issues to account for when dealing with this data. The crossing tracks of the dataset result in single locations with multivalued data, which can be substantially different due to tidal variations, weather patterns, and currents. In addition, there is spiky noise throughout the dataset, which is infamous for its effect on least-squares based methods (Clarenboult and Muir, 1973; Guitton and Symes, 2003). Finally, the only tracks in the northern half of the survey area have a very wide spacing, so an interpolation problem also arises.

Previous work on this dataset at SEP (Ecker and Berlioux, 1995; Lomask, 1998, 2002) has mainly dealt with the systematic errors present in the dense dataset (Ecker and Berlioux, 1995), or with ways in which to use information in the dense portion of the data to regularize the missing bins in the northern, sparse portion of the data (Lomask, 1998, 2002).

Some previous methods of dealing with sparse data rely on creating proxy data, either by

1email: bill@sep.stanford.edu
an initial guess of a prediction-error filter (PEF) (Claerbout, 1999) or by using coarser scales of the data (Curry, 2002). I propose a method where we solve the interpolation problem in the data space coordinate system. The model space is warped to fit the data location and two different approaches are applied. In the first approach te data are preprocessed so that they are single valued. In the second, two sets of co-located tracks are used to create a map.

**BACKGROUND**

The Madagascar regularization problem has been approached using the following fitting goals (Lomask, 2002):

\[
W \frac{d}{dt}[Lm - d] \approx 0 \\
\epsilon Am \approx 0.
\]  

(1)

In these fitting goals: \(W\) corresponds to a weight for ends of tracks and spikes in the data, \(\frac{d}{dt}\) is a derivative along each track used to eliminate low frequency variations along each track, \(L\) is a linear interpolation operator that moves from values on a regular grid to the data points, \(m\) is the desired gridded model, \(d\) are the data points along the tracks, \(A\) is a regularization operator, and \(\epsilon\) is a trade-off parameter between the two fitting goals.

The regularization operator (\(A\)) typically is a Laplacian, a prediction-error filter (PEF), or a non-stationary PEF (Crawley, 2000). When using a PEF, it first must be estimated on some training data, using a least-squares fitting goal,

\[
W(DKf + d) \approx 0.
\]

(2)

in which \(W\) is a weight to exclude equations with missing data, \(D\) is convolution with the data, \(K\) constrains the first filter coefficient to 1, \(f\) is the unknown filter, and \(d\) is a copy of the data.

In order to set a benchmark for how effective a prediction-error filter can be as a regularization operator, a PEF is estimated on the densely-sampled portion of the Madagascar dataset and is then used to interpolate the sparse tracks in the same area. The data are interpolated by using the following fitting goals:

\[
K_{\text{data}}m \approx m_k \\
\epsilon Am \approx 0.
\]

(3)

Here \(K_{\text{data}}\) is a mask for known data, \(m_k\) are the data, \(m\) is the model, and \(A\) is the regularization operator. In this case, the input data is the output of the fitting goals in equation (1), with \(\epsilon\) set to zero. Results for using a Laplacian, a PEF estimated on well-sampled data, and a non-stationary PEF estimated on dense data as the regularization operator \(A\) are all shown in Figure 1. As we can see, as the complexity of the regularization operator increases, the interpolated result improves. This is because the Laplacian assumes isotropic behavior in the data while the PEF and non-stationary PEF are based on statistical information in the data. The PEF assumes statistical stationarity in the data while the non-stationary recognizies and
Now that the best-case scenarios are out of the way, we can see what we can accomplish without cheating and using the dense data. By using only the sparse tracks, we are not able to capture nearly as much information about the model as we have in the previous case. The results for interpolating with a Laplacian as well as stationary and non-stationary PEFs estimated solely on the sparse tracks with the multi-scale method (Curry and Brown, 2001; Curry, 2002, 2003) are shown in Figure 2.

In this case, there are regions of the data where the Laplacian gives the best result, and regions where the non-stationary PEF gives the best result. One example of this is the spreading ridge. The non-stationary PEF is able to interpolate some fine features of the ridge that Laplacian interpolation is incapable of. However, in other regions, the non-stationary PEF performs poorly, and the far simpler Laplacian interpolation gives a more reasonable result. This is because the PEFs estimated using the multi-scale approach do not always properly characterize the data, as local information is destroyed in the multi-scale averaging process.

Some of the problems encountered by the non-stationary PEF are due to the spatial distribution of the data. The multi-scale approach used in Figure 2 blindly rescales the data without regard to their location. By simply making the grid cells larger, very coarse bins must be used before the data become contiguous, and many data points along the tracks fall into these bins. We are clearly in need of a method that will take into account that these data are collected along curved tracks that are at an angle.

**PEFS IN THE DATA SPACE**

The known data points in the model space are distributed along curved crossing tracks, making it very difficult to estimate a PEF in this space. However, in the data space of the fitting goals in equation (1) of the previous section, the data are sampled in a regular space: a series of regularly-sampled tracks. By transforming the coordinates of the data and model, we can keep the data space regularly-sampled while warping the model space so that a PEF estimated in the data space can be used in the model space.

The existing data are collected in a series of ascending and descending crossing tracks. The goal is to transform the model space so that it more closely resembles the pattern in which the data was collected, which is in two sets of crossing tracks. This can be done by first stretching the longitudinal axis so that the tracks are orthogonal, subtracting the mean values of the coordinates of the data, rotating the data so that the tracks are aligned North-South and East-West, applying two parabolic shifts along a rotated axis, and then finally breaking the data up into ascending and descending tracks and gridding the data. This process is shown in Figure 3.

The method used to warp the model space is not especially important. It would perhaps be better to use specific coordinates of the data space, but this warping accomplishes the same
Figure 1: When we know the answer, from top to bottom: (a) The sparse tracks on the lower half of the data set; (b) Those tracks interpolated with Laplacian regularization; (c) The sparse tracks interpolated with a PEF estimated on the co-located dense tracks; (d) the same as (c), but using a non-stationary PEF; (e) the dense tracks from the same area.
Figure 2: When we don’t know the answer, from top to bottom: (a) The sparse tracks on the lower half of the data set; (b) Those tracks interpolated with Laplacian regularization; (c) The sparse tracks interpolated with a PEF estimated on the same sparse tracks; (d) the same as (c), but using a non-stationary PEF; (e) the dense tracks from the same area.
Figure 3: The steps used in warping the coordinate system. (a): Original model; (b): Stretching and centering; (c): Rotation; (d): Parabolic shifts. [bill1-warping][ER]
end goal of having straightened orthogonal tracks. We can use the same warped grid for our model space, allowing us to use a similar methodology that we used for the regularly-sampled data problem. The two sets of orthogonal tracks with missing data is very similar to a problem shown by Claerbout (1999), where a pair of 1D PEFs are used to fill in missing data. This can be expressed as

\[
K_{\text{data}}m \approx m_k \\
\epsilon A_a m \approx 0 \\
\epsilon A_d m \approx 0,
\]  

where \( K, m \) and \( m_k \) are the same as above, and \( A_a \) and \( A_d \) are 1D PEFs that are estimated on the ascending and descending tracks in the data space, respectively. This approach assumes that the known data is single-valued, and fixes it so that it does not change. The two sets of tracks are reduced to a single map with the fitting goals in equation (1), with \( \epsilon \) set to 0. This becomes the data, and the obtained model is shown in Figure 4. The results of this approach

![Figure 4](image)

**Figure 4**: Results in a warped space using the fitting goals in equation (4). Above: model obtained after 500 iterations. Below: residual. [bill1-warpfill1] [ER,M]

are good. The general trend of the data is interpolated, and the spreading ridge is clearly
present in the model, which is not the case with Laplacian interpolation. Since this method only deals with the merged tracks, another method which uses the two track sets separately is tested next.

In this second method, the following fitting goals are solved:

\[ A_a(m - d_a) \approx 0 \]
\[ A_d(m - d_d) \approx 0, \]

where \( A_a \) and \( A_d \) are again 1D PEFs that are estimated in the data space, which is now two panels containing separately gridded ascending and descending tracks, represented by \( d_a \) and \( d_d \). The results for this method are shown in Figure 5.

The residual space is now twice the size as that for the fitting goals in equation (4), since the ascending and descending tracks are now dealt with separately. This approach appears to also be successful, although some effects of the PEF are present in the data, making the result look like a roughened result. The spreading ridge can be easily delineated. The residual space is interesting to look at, as the influence of the two sets of tracks can be seen in the different portions of the residual. The imprint of the data has disappeared from one half of the residual and has appeared in the other.

**CONCLUSIONS AND FUTURE WORK**

By estimating a pair of 1D PEFs in a warped data space, the sparse, curved tracks of the Madagascar dataset were successfully interpolated. Other methods that require a starting guess or rescaled proxy data were much less successful, as they were estimated in a domain where the data are not optimally distributed.

The tracks in the Madagascar dataset have several similarities with the tracks seen in 3D seismic data, such as sail lines, receiver cables, and cut lines for sources. When estimating PEFs on this data, the choice of doing it in the model space and the data space needs to be investigated.

To obtain a better result for this data, the next obvious step is to use non-stationary PEFs, as the character of the Madagascar data changes greatly with position.

**ACKNOWLEDGMENTS**

I would like to thank Jon Claerbout, Jesse Lomask and Antoine Guitton for several fruitful discussions.

**REFERENCES**

Figure 5: Results in a warped space using the fitting goals in equation (5). (a): model obtained after 150 iterations. (b),(c): two parts of the residual space. The data space contains the two sets of tracks, and is twice the size of the model space. [bill1-warpfill2a] [ER,M]


Curry, W., 2002, Non-stationary, multi-scale prediction-error filters and irregularly sampled data: SEP–111, 327–337.


Improved image segmentation for tracking salt boundaries

Jesse Lomask, Biondo Biondi and Jeff Shragge

ABSTRACT
Normalized cut image segmentation can be used to track salt boundaries within an iterative velocity analysis scheme. To overcome the formidable computational expense and storage requirements of normalized cut image segmentation, three cost saving approaches are proposed. First, pixels are sampled from windows centered at powers of 2. This greatly increases the sparseness of the weight matrix. Second, initial solutions are provided to subsequent segmentations for multiple segmentation passes in iterative velocity analysis. Third, an iterative multi-scale approach would be necessary for tracking of the bright salt events in large 3D cubes.

INTRODUCTION
An accurate velocity model is needed to image beneath a salt body. This velocity model is typically created by manually picking the top of the salt. Improvements are then made to the velocity model and the data is remigrated causing the salt boundary to move and refocus. This process may be repeated several times and the manual picking of the salt boundary can be time consuming. Using image segmentation to pick the boundary could make this process easier. The interpreter is then only required to check the results and make some minor modifications.

Hale and Emanuel (2003, 2002) apply the normalized cut image segmentation method developed by Shi and Malik (2000) to paint a 3D coherency-based reservoir model. Our approach for tracking salt boundaries is similar (Lomask, 2003). This image segmentation technique creates a matrix containing weights relating each pixel to every other pixel in a local neighborhood. The matrix is then used to cut the image where the normalized sum of weights cut is minimized. We have modified the weight calculation to be dependent on the negative absolute value of the complex trace (instantaneous amplitude) of the seismic. This makes the weights very weak at salt boundaries, causing the segmentation algorithm to cut along the boundary.

In this paper, we give a very general review of the normalized cut segmentation technique. We then describe how we modified it for application to salt dome seismic data. We test this technique on synthetic seismic sections to illustrate its efficacy with discontinuous salt boundaries. Approaches to make this method more robust and cost effective are also presented.

1email: lomask@sep.stanford.edu, biondo@sep.stanford.edu, jeff@sep.stanford.edu
SEGMENTATION METHODOLOGY

The normalized segmentation method described by Shi and Malik (2000) is designed to look for clusters of pixels with similar intensity. To do this, a weight matrix is created that relates each pixel to every other pixel within a local neighborhood. The strongest weights are given to pixels of similar intensity and close proximity. The method then seeks to partition the image into two groups, \( A \) and \( B \), by minimizing the normalized cut:

\[
N_{\text{cut}} = \frac{\text{cut}}{\text{total}_A} + \frac{\text{cut}}{\text{total}_B},
\]

where \( \text{cut} \) is the sum of the weights cut by the partition. \( \text{total}_A \) is the sum of all weights in Group \( A \), and \( \text{total}_B \) is the sum of all weights in Group \( B \). Normalizing the cut by the sum of all the weights in each group prevents the partition from selecting overly small groups of nodes.

The minimum of \( N_{\text{cut}} \) can be found by solving the generalized eigensystem:

\[
(D - W)y = \lambda Dy,
\]

created from weight matrix \( W \) and a diagonal matrix \( D \), with each value on the diagonal being the sum of each column of \( W \). The eigenvector \( y \) with the second smallest eigenvalue \( \lambda \) is used to partition the image by taking all values greater than zero to be in one group, and its complement to be in the other.

APPLICATION TO SEISMIC DATA

To apply this segmentation method to seismic data, the weight calculation needs to be modified. Rather than looking for clusters of pixels with similar intensity, we are now looking for groups of pixels on each side of the bright amplitude salt boundary. Therefore, we want the weights connecting pixels on either side of the salt boundary to be low and the weights connecting pixels on the same side of the salt boundary to be relatively high. We get the best results when the weight for every pair of pixels is made weak (set to zero) when the following two criteria are met, otherwise the weight is made strong (set to one). The first criteria is that the minimum negative instantaneous amplitude along the shortest path between pixels is below a threshold. The threshold in our examples thus far is 85 percent of the maximum negative amplitude. The second criteria is the minimum is less than the two pixels themselves. The second criteria insures that the minimum is actually located between the two pixels. This is illustrated in Figure 1.

The top panel of Figure 2 shows the negative of the instantaneous amplitude of a synthetic salt boundary. This is the input for the weight calculation. The resulting weight matrix is extremely large: \((m \times n)^2\) where \( m \) and \( n \) are the dimensions of the image. Because this matrix is so large, checking the quality of weights is problematic. We can however view all of the weights for particular pairs of pixels. For instance, the bottom of Figure 2 is the weight...
Figure 1: A cartoon illustrating some sample weights on the negative instantaneous amplitude. Several pixels are shown in squares and their corresponding weights in circles. The salt boundary is defined as areas where the negative instantaneous amplitude is below a threshold.

Figure 2: Top is negative absolute value of the complex trace of a synthetic salt boundary. Bottom is the weight connecting each pixel with the pixel directly above it. It is zero where two pixels are separated by the salt boundary.
between each pixel and the adjoining pixel directly above it. Notice that it follows the peak of the amplitude.

We found that the best results are obtained when the maximum search distance is greater than 30 pixels. The maximum search distance is the maximum distance between pixels, beyond which the weights are set to zero. Using large search distances causes two problems. First, the matrix becomes more dense and requires more storage space. Second, the weights of the pixels at greater distances far outnumber those at close distances causing a bias in the weight matrix. Shi and Malik (2000) used a decaying distance weight to reduce this bias. We found that if we only use windows that are centered at powers of 2, then the matrix is still sufficiently sparse while still benefiting from greater search distances. However, this obviously causes many distances to be not sampled at all. Figure 3 shows the $\log_2$ sampling of one pixel up to a distance of 16 samples. To correct this, we plan to randomly sample from within the maximum search distance so that the number of non-zero points is approximately the same for all distances.

**Iterative Velocity Analysis**

This method can be used to pick salt boundaries that are discontinuous. An updated velocity model can then be created and used to remigrate the data.

Additionally, if the salt boundary changes only slightly, then the eigenvector solution to equation (2) from the previous segmentation can be used as an initial solution to the current segmentation.

The top panel of Figure 4 is a synthetic 2D section that has been migrated with a preliminary velocity model. The results of applying the segmentation method can be seen in the bottom. In general it does a good job picking the salt boundary, however at cmp locations -750 and -1950 it has some difficulty.

The top panel of Figure 5 contains the eigenvector with the second smallest eigenvalue from equation (2) that was partitioned to get the result in Figure 4. Below is a contour plot of this eigenvector. Notice that the spreading contours correspond to areas where the picking had some difficulty.
Figure 4: Segmentation can be an effective picker even in the presence of noise and discontinuities. Top is a synthetic salt boundary. Bottom is the resulting partition from the segmentation method.
Figure 5: Areas of uncertain picking can be found by inspection of the eigenvector to be partitioned. Top is the eigenvector used to find the boundary in the lower part of Figure 4. Bottom is a contour plot of the eigenvector. Notice the areas of uncertainty where the contours are spreading. [jesse1-p1.eig] [ER]
Figure 6: Top is a synthetic salt boundary after the velocity has been adjusted and remigrated. Bottom is the resulting partition from the segmentation method.
Figure 7: Top is the eigenvector used to find the boundary in the lower part of Figure 6. Bottom is a contour plot of the eigenvector. Notice the areas of uncertainty where the contours are spreading.
The top of Figure 6 is the same synthetic data used in Figure 4 except migrated with an updated velocity. Again, notice that the partitioning result in the bottom successfully picks the salt boundary except in a couple of places. The corresponding eigenvector is presented in Figure 7. Notice its similarity to the eigenvector from Figure 5. Initializing subsequent segmentation processes with previous eigenvectors should speed convergence.

CONCLUSIONS AND FUTURE WORK

Our modified segmentation method successfully tracked the salt boundaries in our test cases. These test cases present challenges including non-continuous salt boundaries with noise. An interpreter would still have to slightly adjust the tracking result in a few places.

Because of this method’s excessive storage requirements, a multi-scale approach will be necessary. The size of the weight matrix in equation (2) is \((m \times n)^2\) where \(m\) and \(n\) are the dimensions of the image. Even using sparse matrices in 2 dimensions, this matrix can be prohibitively large. In 3D, things get much worse as the matrix is now \((m \times n \times o)^2\), where \(o\) is the 3rd dimension. The segmentation method applied to a coarsely sub-sampled input cube will still select the salt boundary as long as it is the brightest amplitude in the cube. Smaller, more finely sampled cubes can then be segmented along the boundary.

Application of this method to 3D datasets is going to be a computational challenge that will require us to take advantage of maximizing the sparseness of the weight matrix, using accurate starting solutions, and taking a multi-scale approach.

ACKNOWLEDGMENTS

We would like to thank Frederick Billette and BP for making available to SEP the synthetic data used for our tests. We would also like to thank Antoine Guitton and Paul Sava for suggestions and data.

REFERENCES


Analytical flattening with adjustable regularization

Jesse Lomask and Antoine Guittion

ABSTRACT
We add an adjustable regularization parameter to the analytical flattening method which integrates dips in the Fourier domain. The regularization penalizes roughness in depth in the integration result which, in turn, insures that the flattening result is monotonic and continuous. This preserves the data which is necessary for multiple flattening passes or for undoing the flattening result. Because we perform the integration in the Fourier domain, this method is still highly efficient. 2D field gathers and a stacked section are provided as examples. This can easily be extended to 3D, allowing adjustable regularization along shot gathers.

INTRODUCTION
When flattening seismic data, it is important that the relative position of adjacent data points be preserved. In other words, it is important that the process be continuous and monotonic. If the flattening process is not monotonic, then points can be swapped creating artifacts in the data. This is important if the flattened data is to be unflattened or if multiple iterations are to be performed as these artifacts can worsen with each iteration. Simply smoothing the dip in depth prior to integrating should achieve a monotonic result, however, it is more desirable to penalize roughness in depth in a least squares sense while the dip is being integrated. In short, we want to apply an adjustable model styling goal to insure smoothness of the dip integration result.

In Lomask and Claerbout (2002); Lomask (2003) dips are integrated (summed) by a Fourier method with regularization. However, this regularization was not adjustable. Although it did insure that the integrated dip result was smooth, in many cases the depth regularization goal over-whelmed the dip integration goal, causing the result to be too smooth.

In this paper, we present a modification to the analytical flattening method that uses an adjustable weighting parameter for regularization in depth. This method is applied to several 2D field gathers provided by WesternGeco and compared to 2D field gathers that are flattened without regularization. A 2D section created by applying NMO prior to stacking is compared to the same 2D section created by flattening prior to stacking. We show that applying the analytical flattening method with regularization can preserve the integrity of the data even after multiple passes of flattening.

1email: lomask@sep.stanford.edu, antoine@sep.stanford.edu
METHODOLOGY

In Lomask and Claerbout (2002), we found that we could integrate local dip information \( p_x \) into total time shifts \( t_x(t_x) \) quickly in the Fourier domain with:

\[
 t_x \approx \text{FFT}^{-1}_1 \left[ \text{FFT}_1 \left[ \frac{\nabla' p_x}{Z_x^{-1} + 2 - Z_x} \right] \right],
\]

where \( Z_x = e^{i w \Delta x} \).

We also found that if we initialized the dips in the \( y \) direction \( p_y \) to zero, then this equation would apply some kind of regularization in the \( y \) direction:

\[
 t \approx \text{FFT}^{-1}_2 \left[ \frac{\nabla' p}{Z_x^{-1} - Z_y^{-1} + 4 - Z_x - Z_y} \right],
\]

where \( t = t(x, y) \), \( p = (p_x, p_y) \), \( Z_x = e^{i w \Delta x} \) and \( Z_y = e^{i w \Delta y} \). This would cause the integration to be smooth in the \( y \) direction. However, we were not able to control how smooth it would be.

Here we will add an adjustable regularization parameter \( \epsilon \) to equation (2). We begin with the fitting goal:

\[
 \nabla t = p.
\]

We can minimize the difference between the estimated slope and the theoretical slope with:

\[
 0 \approx \nabla t - p.
\]

Next, we write the quadratic form to be minimized as:

\[
 Q(t) = (\nabla t - p)'(\nabla t - p).
\]

Because the gradient is \( \nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}) \), we can write:

\[
 Q(t) = \left[ \frac{\partial t}{\partial x} - p_x \right]' \left[ \frac{\partial t}{\partial x} - p_x \right] + \left[ \frac{\partial t}{\partial y} - p_y \right]' \left[ \frac{\partial t}{\partial y} - p_y \right].
\]

This can be rewritten as:

\[
 Q(t) = \left( \frac{\partial t}{\partial x} - p_x \right)^2 + \epsilon^2 \left( \frac{\partial t}{\partial y} - p_y \right)^2.
\]

The second term in equation (7) is the regularization term and only needs a scalar parameter \( \epsilon \) to adjust its weight relative to the first term. Now we have:

\[
 Q(t) = \left( \frac{\partial t}{\partial x} - p_x \right)^2 + \epsilon^2 \left( \frac{\partial t}{\partial y} - p_y \right)^2.
\]
Working backwards we see that it is now necessary to define a gradient operator that has an 
epsilon weight applied to one direction as:

\[ \nabla_\epsilon = \left( \frac{\partial}{\partial x}, \epsilon \frac{\partial}{\partial y} \right). \]  

(9)

It is also necessary to apply the scalar to the dip in the y direction as:

\[ p_\epsilon = (p_x, \epsilon p_y). \]  

(10)

Lastly, the y components of the z-transform in the denominator of equation (2) also need 
to be scaled. The final analytical solution with an adjustable regularization parameter is:

\[ t \approx \text{FFT}_{2D}^{-1} \left[ \frac{\text{FFT}_{2D} \left[ \nabla_\epsilon p_\epsilon \right]}{-Z_x^{-1} - \epsilon Z_y^{-1} + 2 + 2\epsilon - Z_x^z - \epsilon Z_y^z} \right], \]  

(11)

where \( Z_x = e^{i\omega \Delta x} \) and \( Z_y = e^{i\omega \Delta y} \).

**EXAMPLES**

We use the adjustable regularized flattening method on field CMP gathers from the Gulf of 
Mexico Mississippi Canyon provided by WesternGeco. Methods for using flattening for ve-
locity analysis are currently being developed ((Guitton et al., 2004; Wolf et al., 2004)).

**CMP gathers**

Figure 1a shows a raw CMP gather. Figure 1b shows the same data after constant velocity 
(2000 ms\(^{-1}\)) NMO has been applied. This reduces the maximum dip and therefore makes the 
dip estimation more robust. Figure 1c shows the result of one iteration of flattening with no 
regularization applied (\( \epsilon = 0 \)). This means that each slice of dip is integrated independently. 
The estimated dip can be seen in Figure 2a. Fluctuations in dip can cause horizontal striations 
in the time shifts as in Figure 2b. These sudden changes in time shift can cause points to 
be swapped in depth or time. This can create the artifacts seen in Figure 1c. These artifacts 
cause errors in the dip estimation which cause the problem to grow with subsequent iterations. 
The estimated dip for the third iteration can be seen in Figure 1c. These dips are very small 
compared to the dips in 1a because the figures are clipped differently. However, notice the 
severity of the striations in the time-shifts after the third iteration shown in Figure 2d. 

Figures 3 and 4 show the same data as Figures 1 and 2 except using regularization (\( \epsilon = 2 \)). 
Notice that the integrated dip shown in Figure 4b is much smoother than the result shown in 
Figure 2b.

After 3 iterations of flattening, Figure 3d, the gather is flat and the data is still preserved. 
If necessary, this data can be easily unflattened because the data integrity is intact.
Figure 1: (a) Raw CMP gather. (b) CMP gather with constant velocity applied, 2000 m/s. (c) Flattened version of (b) with no regularization, one iteration. (d) Flattened version of (b) with no regularization, three iterations. Notice the artifacts in (c) and (d) caused by data points being shifted without maintaining continuity and monotonicity.
Figure 2: (a) Estimated dip from (b) in Figure 1. (b) Integration result of the dip in (a). (c) Estimated dip for third iteration. (d) Integration result of the dip in (c).
Figure 3: (a) Raw CMP gather. (b) CMP gather with constant velocity applied. (c) Flattened version of (b) with regularization ($\epsilon=2$), one iteration. (d) Flattened version of (b) with regularization ($\epsilon=2$), three iterations. Notice the significant improvement of the flattening result compared to Figure 1.
Figure 4: (a) Estimated dip from (b) in Figure 3. (b) Integration result of the dip in (a). (c) Estimated dip for third iteration. (d) Integration result of the dip in (c).
Figure 5: 9 sample gathers from the same data used to create a 2D stack section. Each gather is 2600 meters apart.
2D stacked section

Nine sample CMP gathers are shown in Figure 5. Radon de-multiple pre-processing has been applied but significant multiple energy can still be seen in the bottom half of the gathers. For comparison to the flattening method, conventional processing is first applied. Using the simple velocity field shown in Figure 6, NMO is applied to the gathers and the results are shown in Figure 7. This velocity is obviously not exact but many primaries are almost flat. The data is stretched badly and could benefit from a stretch mute.

To reduce dip estimation errors, we first apply NMO with a constant velocity of 2000 m/s. Applying NMO with this simple velocity does not require knowing any velocity information.
Figure 7: Same gathers as Figure 5 with NMO applied using the velocity function in Figure 6. This velocity is obviously incorrect in several areas but it does flatten many of the primaries and it does not flatten the multiples.
at all. Next, the gathers are flattened. The nine flattened gathers are shown in Figure 10. These gathers don’t have the same stretching errors as in Figure 7 but events with the strongest amplitude are flattened regardless of whether they are primaries, multiples, or headwaves. Evidence of this can be seen in the lower half where the flattening method has flattened the undesirable multiple energy. At the top of the gathers, the flattening method flattens the gathers significantly better than the NMO in Figure 7.

![Figure 8: Same gathers as Figure 5 with NMO applied using a constant velocity function, 2000 m/s⁻¹.](image)

The resulting stacked sections with conventional processing and flattening applied can be seen in Figures 9 and 11, respectively. The flattening method does a better job imaging the diffractions seen in the ellipse at the top of the section. However, in the conventional processing image, the primaries are imaged better in the lower left ellipse. In the ellipse on
the lower right, the flattening does a much better job imaging the bright reflections.

Figure 9: Stacked 2D section using the velocity in Figure 6.

In the stacked section after flattening was applied in Figure 11, there is significant jitter between CMP’s. This is because each CMP was flattened independently. We can easily correct this by estimating dip from gather to gather and solve the analytical flattening method in 3D. An adjustable parameter that controls the weight of the flattening between gathers can be added as well.
Figure 10: Flattened version of the gathers in Figure 8.
Figure 11: Stacked 2D section using flattened gathers.
CONCLUSIONS

The analytical flattening method with adjustable regularization flattens data without creating discontinuities. Consequently, the data integrity is preserved allowing for multiple flattening iterations or for undoing the flattening.

There are other ways to achieve a continuous, monotonic flattening result. Smoothing the input dip in depth can achieve similar results. Alternatively, the method that we use to apply the time-shifts can, in principle, be modified so that it is constrained to preserve the relative data order.

The regularized analytical method can be used to find a smooth solution but it cannot fill in missing data. One of the main benefits of regularization is the ability to fill missing data. This requires a weight matrix or mask to discern known data from unknown data. Unfortunately, this weight matrix is singular and the analytical solution would require us to know its inverse. Therefore, in the case of regions of unknown dip, it is necessary to abandon the analytical method and solve the flattening problem in the time domain.

ACKNOWLEDGMENT

We would like to thank WesternGeco for providing the field data used in this paper.

REFERENCES


Shot-profile migration of GPR data

Jeff Shragge, James Irving, and Brad Artman

ABSTRACT
Multi-offset ground-penetrating radar data possess a number of important advantages over constant-offset data, and are becoming increasingly popular within the GPR community. With the availability of these data comes the opportunity to experiment with state-of-the-art seismic imaging techniques. Here, we consider the application of shot-profile migration, a prestack scalar wave-equation imaging method, to 2-D multi-offset GPR data. With this method, source and receiver wavefields of individual shot records are propagated separately and combined at depth with application of an imaging condition. Receiver wavefields are comprised of recorded traces, and source wavefields are modeled from point sources at the transmitter locations. The complete migrated image is the sum of all overlapping shot-record migrations.

INTRODUCTION
At present, the majority of ground penetrating radar (GPR) work involves the collection, processing, and interpretation of constant-offset data. However, the increased availability of multi-channel GPR systems is making multi-offset data collection increasingly popular. Advantages of working with multi-offset GPR data include the improved estimation of subsurface velocities and imaging of reflectors, and also the ability to perform amplitude versus offset (AVO) or angle (AVA) analysis (Fisher et al., 1992a; Greaves et al., 1996; Baker, 1998). These advantages allow for better estimation of sedimentary facies and subsurface properties.

Because of the many similarities between the GPR and seismic methods, numerous exploration seismic imaging techniques have been directly transferred to the radar community. One family of imaging methods commonly used in seismic exploration, especially in areas of complex geology, is based on wave-equation wavefield extrapolation. Methods in this family employ a one-way scalar wave equation to extrapolate a wavefield recorded at the surface through a subsurface velocity model. In theory, this procedure generates the wavefield that would have been recorded had the instruments been located at a surface deeper within the earth. Images of geologic structure are then constructed through the evaluation of a physical imaging condition at each subsurface model point.

Here, we consider the application of shot-profile migration, a prestack imaging algorithm belonging to the family of scalar wave-equation methods, to 2-D multi-offset GPR data. This

1email: jeff@sep.stanford.edu, jdirving@pangea.stanford.edu, brad@sep.stanford.edu
method is similar to the zero-offset, survey-sinking migration method proposed by Sena et al. (2003), which involves using split-step Fourier operators to extrapolate GPR wavefields. However, the algorithm presented here is designed for application to multi-offset data in a shot-profile configuration, and allows for the easy formation of angle-dependent images that are suitable for migration velocity analysis (MVA) and AVA studies.

**THEORY**

**Scalar Wave Equation for GPR Imaging**

Modeling a GPR experiment is a complicated process that, for complete accuracy, requires taking into account such effects as antenna radiation patterns and the vectorial nature of electromagnetic (EM) wave propagation and scattering (van der Kruk et al., 2003). It is well known, however, that seismic processing techniques based on a scalar wave equation can often be applied very successfully to GPR data (Fisher et al., 1992b). This latter point is not mere coincidence. In many situations, isotropic scattering and scalar wave propagation effectively model the kinematics of a GPR experiment. Here, in applying an imaging algorithm based on a scalar wave equation to GPR data, we argue that radar propagation kinematics are well represented. In addition, we suggest that with further development, effects such as antenna radiation patterns and realistic scattering could be accounted for in the source and receiver wavefields and imaging condition.

Considering a situation where Maxwell’s equations can be represented by a 2-D scalar wave equation involves making two approximations. First, we implicitly assume that the subsurface geology and sources are strictly 2-D. This results in the decoupled transverse electric (TE) and transverse magnetic (TM) propagation modes (Jackson, 1975). Choosing the TE-mode, we next assume that heterogeneities within the earth are small such that the gradients of EM constitutive parameters can be neglected (Sena et al., 2003). The result is a scalar wave equation for transverse electric field, \( E \), which in the frequency (\( \omega \)) domain is given by,

\[
\frac{\partial^2}{\partial z^2} E + \frac{\partial^2}{\partial x^2} E + \omega^2 s^2 E = 0,
\]

where the slowness of wave propagation (i.e. inverse of velocity), \( s \), is dependent on the medium’s dielectric permittivity, \( \epsilon \), and conductivity, \( \sigma \), through,

\[
s = \sqrt{\mu(\epsilon - \frac{i\sigma}{\omega})} \approx \sqrt{\mu\epsilon}.
\]

The magnetic permeability of the medium, \( \mu \), is roughly constant for most material likely to be encountered in a routine radar application. Hence, in low conductivity media (i.e. \( \sigma \ll \omega \)), the slowness of wavefield propagation is directly proportional to the dielectric permittivity.
Shot-profile wavefield continuation

Assuming applicability of the scalar wave equation given by equation (1), the first step in GPR shot-profile migration is the extrapolation of surface-recorded data to depth. This is done by applying a one-way wave-equation operator to an arbitrary wavefield, as a function of space and frequency, to yield the wavefield at a deeper level:

\[ W(z + \Delta z, x, \omega) = W(z, x, \omega) e^{\pm k_z \Delta z}. \]  

Here, the positive or negative exponent corresponds to causal or acausal propagation, respectively, and \( \Delta z \) is the size of the downward continuation step. The vertical wavenumber, \( k_z \), is calculated from the scalar wave-equation dispersion relation,

\[ k_z = \sqrt{s^2 \omega^2 - k_x^2}, \]  

where \( k_x \) is the horizontal Fourier wavenumber component of the data wavefield.

Surface-recorded wavefields are extrapolated to all depths within the model through successive applications of equation (3) using a vertical wave-number given by equation (4). Although equation (4) is strictly valid only for vertically stratified media, techniques exist to extend it to laterally varying media. We employ a split-step Fourier approach (Stoffa et al., 1990) that involves approximating \( k_z \) in equation (4) using a Taylor series expansion about a reference slowness, \( s_0 \):

\[ k_z \approx \omega (s - s_0) + \sqrt{s_0^2 \omega^2 - k_x^2}. \]  

The first, mixed-domain term in equation (5) acts as a local correction to the second term that handles the bulk of the propagation. Increased accuracy can be achieved by summing the results of multiple reference velocity steps in order to minimize the quantity \( s(x) - s_0 \).

Shot-profile migration directly mimics the data collection process by migrating individual shot records. Receiver wavefields are comprised of individual shot profiles and are propagated acausally. Source wavefields have the same geometry, but are initially zero except for an appropriate source function at the transmitter location, and are propagated causally.

The imaging condition

The second step of shot-profile migration is forming a subsurface image through extraction of appropriate information from the independently extrapolated source and receiver wavefields. Claerbout’s imaging principle (Claerbout, 1971) asserts that energy in the receiver wavefield, \( R \), that is spatially collocated with energy in the source wavefield, \( S \), at time \( t = 0 \) originates from a reflector at that model point. Mathematically, this is accomplished through the extraction of the zero-lag of the cross-correlation of the two wavefields. In practice, this translates to a summation over frequency after the multiplication of the wavefields (Claerbout, 1985),

\[ I(z, x) = \sum_s I_s(z, x) = \sum_s \sum_{\omega} S_s(z, x, \omega) R_s^*(z, x, \omega), \]
Here, $I(z,x)$ represents the image point as a function of horizontal distance and depth, * represents complex conjugation, and the subscript $s$ refers to individual shot-profile image results.

Additional subsurface reflectivity constraints are obtained by extending (6) to include subsurface offset (Rickett and Sava, 2002). Offset domain common image gathers (ODCIGs) are created by multiplication of the source and receiver wavefields after a lateral shift of $h$:

$$I(z,x,h) = \sum_s I_s(z,x,h) = \sum_s \sum_\omega S_s(z,x+h,\omega) R^*_s(z,x-h,\omega).$$  \hspace{1cm} (7)

### Angle domain common image gathers

ODCIGs can be transformed into an angle domain representation (ADCIGs) that describe reflectivity as a function of incidence angle at the reflector, $\gamma$ (Sava and Fomel, 2003). ADCIGs are generated after migration with the relation,

$$\tan \gamma = -\frac{k_h}{k_z},$$  \hspace{1cm} (8)

which uses the Fourier transformed wavenumbers associated with offset and depth.

Angle gathers may be used to examine the accuracy of the imaging velocity model at all subsurface locations. A reflector focused with the correct velocity model should be at a consistent depth for all illumination angles. Deviations in reflector position due to incorrect velocity models result in angle-dependent "smiles" and "frowns", which indicate over- and under-migration, respectively. Corresponding velocity model errors can be estimated from angular moveout through the process of migration velocity analysis (MVA). More importantly, after migration velocity inaccuracy has been minimized, angle-dependent amplitude variations should be attributed to AVA response. However, for GPR images, this will not be the case unless the proper radiation patterns and scattering physics have been taken into account.

### APPLICATION TO FIELD DATA

We applied the shot-profile migration methodology outlined above to a 2-D multi-offset GPR data set collected near Langley, British Columbia, Canada. The geology of the field site consists of a sand and gravel glacial outwash deposit underlain by a conductive marine clay, the depth of which varies from near surface to approximately 10 m across the profile. The data were collected along a road using a PulseEkko 100 GPR system with 100 MHz antennas oriented perpendicular to the survey line. A shot-profile survey configuration was used with 30 receiver offsets ranging from 0.5 m to 15 m, with a step size of 0.5 m. The transmitter spacing along the line was also 0.5 m, with a total of 200 shot gathers making up the approximately 100 m long profile.

Pre-processing of the Langley data included residual median filtering to remove the low frequency inductive response upon which the radar reflection signal was superimposed, and
correction for drift in the zero time of the GPR instrument, most likely due to temperature fluctuations. Because the road surface was very flat, no topographic correction was needed. A time-varying exponential gain was also applied to the data to compensate for energy losses incurred by the GPR pulse during propagation.

Figure (1) shows the unmigrated, near-offset (0.5 m) time section extracted from the Langley data set. This is what would be recorded in a typical, constant-offset GPR survey. Notice that the section is quite complicated, with numerous diffraction hyperbolae present and very few laterally continuous reflectors. The dipping boundary between the sand/gravel and clay layers can be seen as the region where the radar signal rapidly attenuates. However, due to the numerous diffraction hyperbolae and conflicting dips present in the section, the exact location of this boundary is difficult to determine.

In order to perform shot-profile migration, a subsurface velocity model was required. To obtain this model, the radar data were sorted into common-midpoint (CMP) gathers, and semblance analysis was performed. After picking the maximum points on the semblance scans, the resulting root-mean-square (RMS) velocity model was interpolated and converted into a map of interval velocity. The resulting model that was used for the migration is approximately \( v(z) \), with the boundary between the vadose (high-velocity) and saturated (low-velocity) zones at approximately 4.5 m.

Figure (2) shows the Langley multi-offset data after shot-profile migration. As outlined in the previous section, each shot-gather was migrated through independent extrapolation of the source and receiver wavefields, and application of the imaging condition. The migrated
Figure 2: Corresponding depth image after shot-profile migration, and a number of representative ADCIGs.
shot gathers were then summed to produce the image shown. Also displayed in Figure (2) are a number of representative ADCIGs from the migrated data. Finite length acquisition leads to a decrease in available incidence angles for reflectors at increasing depth. For this reason, the ADCIGs become less coherent at large angle deeper in the section.

Comparing Figures (1) and (2), notice that the original time section has been significantly improved after shot-profile migration of the multi-offset data. The migrated image is much cleaner, as diffraction hyperbolas have been largely collapsed and conflicting dips due to the interference of diffraction tails have been eliminated. Many laterally continuous reflections which are absent on the time section can now be seen. Most notable is the water table, which is now very visible at approximately 4.5 m depth. Also, the boundary between the sand/gravel and clay layers is easily identified on the migrated section. On-lap of reflectors in the sand and gravel layer with this boundary can also be seen. Finally, notice that the ADCIG panels shown are very flat. This indicates that we have used an appropriate velocity model to image the data. Any remaining curvature in these gathers could be used to further refine this velocity model.

**DISCUSSION AND CONCLUSIONS**

The seismic community has identified a number of situations where prestack wave-equation migration has advantages over Kirchhoff techniques. The most important of these are the handling of wavefield triplications and earth models with complex structure or large velocity contrasts. Application of the shot-profile migration strategy to GPR shows encouraging results at our field location, and warrants further research. Further, the flexibility of the method allows for incorporation of more accurate source and receiver wavefield modeling, to include effects such as radiation patterns. Lastly, more advanced imaging conditions could be used to incorporate realistic EM scattering physics into the problem.

**ACKNOWLEDGMENTS**

We would like to thank Dr. Rosemary Knight for the use of the Stanford Environmental Geophysics Research Group’s radar system. We also thank Biondo Biondi and Paul Sava for helpful discussions.

**REFERENCES**


Claerbout, J., 1971, Toward a unified theory of reflector mapping: Geophysics, 36, 467–481.


Migration methods for passive seismic data

Brad Artman, Deyan Draganov, Biondo Biondi, and Kees Wapenaar

ABSTRACT

Passive seismic imaging is based on the fact that by cross-correlating the transmission responses of a medium, one can reconstruct its reflection response. Here, we show a method to directly migrate the transmission responses measured at the surface, based on the shot-profile migration. We also show that the results from direct migration of passive data and from migration of the simulated reflection response are identical. At the end, we also show comparisons between the behavior of the results from the migration process and the simulated reflection responses.

INTRODUCTION

In Wapenaar et al. (2004b), general relations are shown between the reflection and the transmission response of a 3-D inhomogeneous medium. The applications of these relations are in synthesis of transmission coda from reflection data, multiple elimination, seismic interferometry and acoustic daylight imaging. The relation used in the acoustic daylight imaging permits us to synthesize the reflection response of a medium by cross-correlating its transmission responses. Here, we show a method to directly migrate passive white-noise seismic data without the need to first perform the cross-correlations. We also compare the results of simulating the reflection response with the results from migration of the passive data.

DIRECT MIGRATION OF WHITE-NOISE DATA

To simulate the reflection response from the transmission responses measured at the surface in the presence of white-noise sources in the subsurface we can use the formula

\[
R^+ (x_A, x_B, \omega) + \left\{ R^+ (x_A, x_B, \omega) \right\}^* = \delta \left( x_{H,A} - x_{H,B} \right) - T^{-\omega}_{obs} (x_A, \omega) \left\{ T^{-\omega}_{obs} (x_B, \omega) \right\}^*. \tag{1}
\]

Above, \( R^+ (x_A, x_B, \omega) \) is the reflection response measured at point \( x_A \) at the surface \( (\partial \mathcal{D}_0) \) in the presence of an impulsive source at \( x_B \), while \( T^{-\omega}_{obs} (x_A, \omega) \) is the transmission response measured at the surface point \( x_A \) in the presence of noise sources in the subsurface.

---

1email: brad@sep, D.S.Draganov@CiTG.TUDelft.NL, biondo@sep.C.P.A.Wapenaar@CiTG.TUDelft.NL
extrapolation of $R^+(x_A, x_B, \omega)$ to common surface locations $\xi_A, \xi_B$ at an arbitrarily greater depth is described by
\[
R^+ (\xi_A, \xi_B, \omega) = \int_{\partial D_0} \int_{\partial D_0} \left\{ W^+ (\xi_A, x_A, \omega) \right\}^* R^+ (x_A, x_B, \omega) W^- (x_B, \xi_B, \omega) \] \tag{2}
\]
where $R^+ (\xi_A, \xi_B, \omega)$ is the reflection response extrapolated from the surface to some subsurface level and $W^+ (\xi_A, x_A, \omega)$ and $W^- (x_B, \xi_B, \omega)$ are forward-extrapolation operators. If we substitute equation (1) into equation (2) we obtain
\[
R^+ (\xi_A, \xi_B, \omega) = \int_{\partial D_0} \left\{ W^+ (\xi_A, x_A, \omega) \right\}^* T^-_{obs} (x_A, \omega) dx_A + \text{anti-causal terms} \tag{3}
\]
In the above relation, we used the fact that the reflection coefficient of the free-surface is $r = -1$ and the reciprocity relation of the forward-extrapolation operator $W^- (x_B, \xi_B, \omega) = W^+ (\xi_B, x_B, \omega)$. Equation (3) shows that by inverse-extrapolating the transmission response $T^-_{obs} (x_A, \omega)$ at all $x_A$ at the surface to a certain subsurface level, and forward-extrapolating the downward-reflected transmission response $r T^-_{obs} (x_B, \omega)$ at all $x_B$ to the same subsurface level, followed by cross-correlation of the resultant wave fields, we obtain the downward extrapolated reflection response. If we subsequently apply the imaging condition, we can image the subsurface at that level (Armtan et al., 2004)). If we compare this process with shot-profile migration (Claerbout, 1971) we can see that they are identical. This means that based on shot-profile migration we can directly migrate passive white-noise data without the need to first simulate the reflection shot gathers. As Figure 1 shows, we may thus use two paths for obtaining a migrated image from passive data. Following the first path, we first cross-correlate the transmission responses recorded at the surface to simulate reflection shot gathers, then we extrapolate the simulated shot gathers and apply the imaging condition (this process was also proposed by Schuster (2001) and named Interferometric Imaging). The other way is to directly migrate the passive data - first we extrapolate the transmission responses recorded at the surface to some subsurface level, then we cross-correlate them and apply the imaging condition. The left panel of Figure 2 shows a double syncline model used to generate transmission responses of white-noise sources in the subsurface. These transmissions were afterwards migrated using both migration methods described above. The results were identical (Figure 2).

SIMULATED REFLECTION VERSUS MIGRATION

In this section, we present some numerical results showing the behavior of the simulated reflection shot gathers and the migrated image when changing some source and receiver parameters. The migrations were performed with the exact velocity model. No multiple elimination
Figure 1: Two paths can be followed to obtain a migrated image from passive data.

\[
T_{\text{obs}}^-(x_A, \omega) \otimes \{T_{\text{obs}}^-(x_B, \omega)\}^* = \{W^+ (\xi_A, x_A, \omega)\}^* \{W^+ (\xi_B, x_B, \omega)\}^* = R^+ (x_A, x_B, \omega)
\]

\[
\{W^+ (\xi_A, x_A, \omega)\}^* \downarrow \{W^+ (\xi_B, x_B, \omega)\}^* \downarrow \{R^+ (\xi_A, \xi_B, \omega)\}
\]

Figure 2: Left: Double syncline model with white-noise sources regularly distributed at depth level \(x_3 = 800\) every 25m between \(x_1 = 1200\) and \(x_1 = 6800\) m. The transmission recordings were 66 minutes. Receivers at the surface are distributed between \(x_1 = 1200\) and \(x_1 = 6800\) m every 20 m. Right: The migrated image from the direct migration and the migration of the simulated reflection shot panels are identical.
schemes were applied. The quality of the end result depends strongly on the number of the present subsurface sources. The left panels from Figures 3 through 6 show simulated reflection shot gathers for a decreasing number of subsurface noise sources with a simulated surface shot position at $x_1 = 4000$ m. The right panels from Figures 3 through 6 show the results from direct migration of the same noise recordings. The noise recordings were 6 minutes long. We see that the simulated reflection responses very quickly decrease in quality (compared with the directly modelled reflection response in Figure 11), while the migration process delivers much better results. Fewer source locations cause illumination problems in the migration result. The

Figure 3: Left: Simulated reflection shot panel from 6 minutes long noise recordings with 113 regularly distributed subsurface sources with a simulated shot position at $x_1 = 4000$ m. Right: Result from direct migration of 6 minutes long noise recordings with 113 regularly distributed subsurface sources. [brad2-dx_50m] [NR]

left panels of Figures 7 through 10 show the change in quality of the simulated reflection shot gather (with simulated surface shot position at $x_1 = 4000$ m) when decreasing the recording time length of the receivers. The right panels of Figures 7 through 10 show the migrated image for the same noise recordings. We can see that while for short recording times the quality of the simulated reflection is strongly degraded, the result from migration is still good, only the signal-to-noise ratio has decreased. Note that multiple events that are hardly visible in the simulated reflection shot gathers at short recording times are still clearly present in the migrated image.

CONCLUSIONS

We showed the theoretical justification for directly migrating passive seismic noise recordings. This method is based on the shot-profile migration procedure, but is applied to transmission data. The results from the direct migration of passive data and from the migration of simulated reflection shot gathers are identical. Depending on the objective (to have intermediate results or not) one or the other can be used. The numerical examples showed that while with decreasing number of subsurface noise sources and shorter noise recordings drastically reduced the quality of the simulated reflection response, the migration process still delivers good results.
Lastly, correlation of each trace with every other trace produces \( N \) shot gathers each with \( N \) traces from a survey with \( N \) receivers. If correlations are performed in the frequency domain, \( N^2 \) traces must be inverse Fourier transformed after multiplication to produce the shot-gathers. This requirement is especially onerous when the record length is at least minutes long. Then, another Fourier transform of \( N^2 \) traces, though now only seconds long, must be performed to make a \( f-k \) based migration. Finally, the increased I/O associated with migrating \( N \) shots in a shot-profile migration are substantial compared with the single shot, even with many more frequencies, for direct migration utilizing an algorithm parallelized over frequency.

Figure 4: Left: Simulated reflection shot panel from 6 minutes long noise recordings with 57 regularly distributed subsurface sources with a simulated shot position at \( x_1 = 4000 \) m. Right: Result from direct migration of 6 minutes long noise recordings with 57 regularly distributed subsurface sources.

Figure 5: Left: Simulated reflection shot panel from 6 minutes long noise recordings with 11 regularly distributed subsurface sources with a simulated shot position at \( x_1 = 4000 \) m. Right: Result from direct migration of 6 minutes long noise recordings with 11 regularly distributed subsurface sources.
Figure 6: Left: Simulated reflection shot panel from 6 minutes long noise recordings with 6 regularly distributed subsurface sources with a simulated shot position at $x_1 = 4000$ m. (b) Result from direct migration of 6 minutes long noise recordings with 6 regularly distributed subsurface sources. [brad2-dx_1000m][NR]

ACKNOWLEDGMENTS

This research is performed as part of research projects financed by the Dutch Science Foundation STW (number DTN4915), the Dutch Research Centre for Integrated Solid Earth Sciences - ISES and the Stanford Exploration Project.

Figure 7: Left: Simulated reflection shot panel from 33 minutes long noise recordings with a simulated shot position at $x_1 = 4000$ m. Right: Result from direct migration of 33 minutes long noise recordings. [brad2-t_33m][NR]

REFERENCES

Figure 8: Left: Simulated reflection shot panel from 6 minutes long noise recordings with a simulated shot position at $x_1 = 4000$ m. Right: Result from direct migration of 6 minutes long noise recordings. [brad2-t_6m] [NR]

Figure 9: Left: Simulated reflection shot panel from 1 minute long noise recordings with a simulated shot position at $x_1 = 4000$ m. Right: Result from direct migration of 1 minute long noise recordings. [brad2-t_1m] [NR]

Figure 10: Left: Simulated reflection shot panel from 20 seconds long noise recordings with a simulated shot position at $x_1 = 4000$ m. Right: Result from direct migration of 20 seconds long noise recordings. [brad2-t_20s] [NR]
Figure 11: Directly modelled reflection response for the double syncline model from figure 2 (a) with a shot at $x = (4000,0) \text{ m}$. [NR]

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, 36, no. 03, 467–481.


Teleseismic imaging with wave equation migration

Charles K. Wilson, Jeff Shragge, and Brad Artman

ABSTRACT
Images of the lithosphere from three-component seismic arrays recording wavefields generated by teleseismic earthquakes (30°-90° epicentral distance) reveal important aspects of lithospheric structure. Unfortunately, interpretation of these images remains ambiguous due to improper reflector mapping and amplitude restoration. We introduce the shot-profile representation of wave equation migration as a novel way to cast the teleseismic imaging problem. We show results from two synthetic datasets using both forward- and backscattered modes to demonstrate the strength of this imaging scheme. To demonstrate the techniques effectiveness with real data, we present preliminary images from the forward-scattered P-to-S mode from data collected across the Proterozoic Cheyenne suture near Laramie, Wyoming.

INTRODUCTION
Deployments of regionally extensive seismic arrays in the last decade have provided earthquake data that afford images of the lithosphere at scales hitherto unattainable (Dueker and Sheehan, 1998; Rondenay et al., 2001; Poppeliers and Pavlis, 2003; Wilson et al., 2003). Attesting to their importance, these images now play a crucial role in improving our understanding of past and present processes of tectonic evolution. The next decade, owing to the EARTHSCOPE2 program and the associated US-ARRAY, will see a dramatic increase in the volume of three-component broad-band seismic data available for imaging the North American lithosphere. However, to fully exploit this opportunity current lithospheric imaging practice must be improved.

Most teleseismic imaging experiments use forward-scattered P-to-S converted waves isolated using receiver function analysis (Phinney, 1964; Langston, 1977). The reasons for concentrating only on this phase are mainly historical. Before the use of receiver arrays became commonplace in teleseismic imaging, only data from sparsely distributed three-component stations were available. Seismologists interested in structure beneath a seismic station were forced to use scattered phases that could be identified and isolated from a single set of three-component seismograms. One easily identified phase is horizontally polarized shear waves converted from lithospheric discontinuities by nearly vertically incident teleseismic P waves; However, identifying backscattered phases (e.g. free surface multiples) requires the use of

1email: wilsonck@sep.stanford, brad@sep.stanford.edu, jeff@sep.stanford.edu
2http://www.earthscope.org
moveout variations instead of polarization. For a single station, there is no moveout of scattered phases without several teleseismic sources from a variety of distances thereby making use of other scattered phases nearly impossible.

Contemporaneously, the introduction of spatially extensive three-component seismic arrays, earthquake seismologists began adapting imaging practices of exploration seismology to produce images of the crust and mantle with teleseismic waves. Initial work in teleseismic imaging used simple processing flows involving linear moveout corrections combined with common conversion point stacking (Dueker and Sheehan, 1998). Later efforts attempted to adopt Kirchoff imaging schemes to produce images of the Cascadia subduction zone (Rondenay et al., 2001) and an Archean continental suture in Southwestern Wyoming (Sheehan et al., 2000). Rondenay et al. (2001) and more recent work by Aprea et al. (2002) clearly demonstrated the utility of backscattered phases in lithospheric imaging and fomenting a new direction for future teleseismic imaging research.

Figure 1: Ray paths of forward and backscattered phases generated by an impinging teleseismic P wave front. Solid lines mark P ray paths with S rays marked by dashed lines. Uppercase letters denote downgoing phases with lowercase letters symbolizing upgoing phases. The idealized vertical and radial seismograms indicate the approximate arrival and amplitude relationships for some of the arrivals in the ray diagram above. The diagram demonstrates the natural separation of P and S phases from near vertically incident waves onto the vertical and horizontal components.

There are clear advantages of wave equation imaging over Kirchoff in exploration seismology. Most likely these advantages remain valid even in the modified geometry of teleseismic sources and receivers (Shragge and Artman, 2003) where lateral velocity heterogeneity is often encountered (e.g. in the presence of partial melt or along large offset strike-slip faults). The shot-profile formulation of wave-equation migration provides a framework in which to cast the teleseismic imaging experiment that allows the incorporation of other scattering modes besides forward-scattered P-to-S. In this paper, we demonstrate its utility for several scattered modes produced by a teleseismic P wavefield impinging on two simple (yet realistic) synthetic crustal models. We also show images from migrated forward-scattered P-to-S converted phases from a dense profile across the Cheyenne suture near Laramie, Wyoming and compare
them to images obtained using conventional receiver function stacking techniques (Dueker and Yuan, 2004).

**SOURCE-RECEIVER GEOMETRY OF TELESEISMIC IMAGING**

By definition, teleseismic earthquake sources originate at a great distance from the receiver array and have ray turning points in the middle-to-lower mantle (>400 km depth). As a result, source wavefields are nearly vertically incident, planar wavefronts arriving beneath the receiver array. Thus, the recorded wavefield is a combination of the source along with the P and S scattered modes generated by the incoming plane wave. Figure 1 shows a schematic example of the teleseismic wavefield and the modes we use for imaging. The synthetic teleseismic wavefield shown in Figure 2 illustrates the combined source and receiver wavefield recorded by the surface array. The first arrival is the source wavelet followed by the forward-scattered P-to-S conversion seen near 13 seconds on the radial and the backscattered P-to-P phase on the vertical near 20 seconds. As demonstrated by the synthetic seismograms, the plane-wave source creates secondary sources in the form of free surface reflections. These reflections retain a significant amount of energy and produce subsurface reflections often several times larger in amplitude than their forward-scattered counterparts. In theory, free-surface multiples provide additional information that may be used in a more traditional up-

![Figure 2: Synthetic seismograms calculated for a crustal model with an 8 km Moho offset.](image-url)
going/downgoing wavefield framework. However, without a means to discriminate between forward and backscattered phases propagating as both P and S waves we may expect a significant amount of contamination between modes. Component rotation from [V,N,E] into the [P,Sv,Sh] system using the free-surface transfer matrix (Kennett, 1991) allows us to identify phases based on polarization direction. Cross-contamination by modes not traveling nearly parallel to the predicted source wave propagation direction still occurs.

WHY SHOT-PROFILE MIGRATION?

The primary advantage of the shot-profile formulation of wave equation migration for teleseismic imaging comes from the ability to separate the source and receiver wavefields for independent propagation during propagation. Shot-profile migration separates the total wavefield into source and receiver wavefields that are each propagated separately (by Fourier-domain multiplication) with the appropriate single square root (SSR) operators (Claerbout, 1995).

\[
S_s(\omega, k_x, z + \Delta z) = S_s(\omega, k_x, z)e^{-ik_c\Delta z}
\]
\[
R_s(\omega, k_x, z + \Delta z) = R_s(\omega, k_x, z)e^{ik_c\Delta z}.
\]  

Here \(S_s\) is the source wavefield for a recorded teleseismic event \(s\) and \(R_s\) represents the receiver wavefield. The negative sign on the single square root operator exponent applied to the source wavefield denotes an initially downgoing wavefield. To represent source and receiver wavefields traveling in the same (upgoing) direction we apply a positive exponent to the single square root operator of each wavefield. The strength of this method comes from the implicit assignment of source and receiver wavefield mode (P or S waves) and propagation direction based on the choice of velocity model and operator exponent, respectively. Following the passive seismic methodology developed previously by Claerbout (1968), we attempt to exploit this direct observation of the source wavefield thereby minimizing the amount of prior assumptions and preprocessing (Artman and Shragge, 2003). Single components of our recorded wavefield may be treated as a source representation with the particular component chosen according to the scattering mode of interest. Table 1 shows the sign and velocity model used to propagate both the source and receiver wavefields for several possible scattering modes.

After propagation of the wavefields, image computation results from a frequency summation of the deconvolution at each depth.

\[
I(x, z) = \sum_\omega \frac{R_s(x, z, \omega)S_s^*(x, z, \omega)}{S_s(x, z, \omega)S_s^*(x, z, \omega) + \varepsilon^2}
\]

After computing individual images from each scattering mode for individual earthquakes, we create composite images for each scattering mode by summing images from all recorded earthquakes. We produce a final image by summing all images produced by all scattering modes.
Table 1: Table listing of the scattering modes, propagation velocity, and propagaton direction for each source and receiver wavefield for different scattering modes.

SYNTHETIC TESTING

In this section we show the results of migration of modes 2 through 4 from table 1. The synthetic data consists of 16 plane waves at angles of incidence between $-30^\circ$ and $30^\circ$ propagated through two crustal models: one containing an 8 km offset in the Moho near 30 km depth and another containing a dipping Moho (Figure 3). Figure 2 and Figure 4 show vertical and radial seismograms from the offset and dipping Moho models, respectively. Notice the clear diffracted arrival generated by the Moho offset in Figure 2 that are not present in Figure 4. Also, interesting is the focusing that occurs at the ramp due to the apparent increase of incidence angle due to the structural dip. Contamination from side reflections later in time prohibits our ability to use some later arriving backscattered phases.

Figure 3: Velocity models used to generate finite-difference synthetic seismograms.
Figure 4: Synthetic seismograms calculated for a crustal model with a dipping Moho.

We use a receiver spacing of 0.5 km to represent a densely sampled teleseismic experiment. In practice, this sampling is impractical but is necessary here to concentrate on the limitations of the procedure without adding complexity due to wavefield aliasing. Stacked images from individual modes and all modes are shown in the next sections.

Figure 5 shows the images created with the forward-scattered P-to-S phase for both synthetic models. Both structures are clearly visible although they appear blurred. The low frequency image results from the fact that the wavefields propagate in the same direction and therefore remain correlated over more imaging depths. We show images created with backscattered P-to-P for both synthetic models in Figure 6. Both structures appear better resolved. The high amplitude shallow structure results from erroneously migrated cross-contamination from the forward-scattered P-to-S phase. This indicates cross-contamination may render interpretation of shallow structures difficult. Figure 7 created using backscattered P-to-S conversion also appears heavily contaminated by other phases. This is most apparent in the Moho offset synthetic model (top panel) where the Moho appears to stretch completely across the image. Combining all images reduces the contamination from other phases and sharpens the real crustal features. However, a significant shallow artifact still remains in both images.

APPLICATION TO DATA RECORDED BY THE LARAMIE BROAD-BAND ARRAY

The Laramie broad-band array which consisted of 30 CMG-40T three-component seismometers spaced between one and two kilometers apart, recorded continuously for 8 months in 2000-2001. The focus of the project was to image the Proterozoic Cheyenne Suture represent-
Figure 5: Migrated image of all synthetic events computed using forward-scattered P-to-S waves for the Moho offset model (top) and the model with the dipping Moho (bottom).

Figure 6: Migrated image of all synthetic events computed using backscattered P-to-P waves for the Moho offset model (top) and the model with the dipping Moho (bottom).
Figure 7: Migrated image of all synthetic events computed using backscattered P-to-S waves for the Moho offset model (top) and the dipping Moho model (bottom).

Figure 8: Migrated image of all synthetic events and all modes shown previously for the Moho offset model (top) and the dipping Moho model (bottom).
ing a nearly 2 billion years old continental collision. The array trending northwest-southeast, straddled the northern edge of the suture. Dueker and Yuan (2004) produced an image us-

Figure 9: The X’s represent the stations associated with the Laramie array. Note large thickness of the Laramie basin shown by the contours. The sampling area of the teleseismic dataset at 40 km depth is shown as the gray ellipsoidal area. The Cheyenne belt suture (CB) surface expression is shows as the thick white line (dotted where inferred). Major NE-trending shear zones are marked as thick dark lines and denoted as: FM, Farwell Mountain; SG, Skin Creek Gulch; IC, Illinois Creek. Surrounding mountains that expose the basement are labeled as: SM, Sierra Madre; MB, Medicine Bow; LM, Laramie Mts. Image modified from Dueker and Yuan, 2004.

ing receiver function analysis where the vertical component is treated as the source function and deconvolved from the radial component. Application of a linear moveout operator to the processed seismograms followed by common conversion point stacking produced the image shown in Figure 10. Important features of this image include the mid-crustal reflector near 20 km depth and the Moho which is interpreted to be near 40 km depth across the line. A second feature dips down below the Moho to the north to a depth of 60 km. Dueker and Yuan (2004) interpret this feature to be a sliver of crust thrust underneath the Wyoming craton during collision.

We have created a migrated image of the forward-scattered P-to-S phase using the same data for comparison to the receiver function image. The results are consistent in many parts of the image including the feature near 20 km depth and the dipping feature between 40 and 60 km depth. However the feature interpreted to be the Moho in the north is absent from our image casting doubt on the previous interpretation. The migrated image appears to be very noisy and low frequency. A different source representation, perhaps as a band limited plane wave, and the incorporation of backscattered phases would improve the image.

**CONCLUSIONS**

We have demonstrated the potential of the wave equation based shot-profile migration developed by Artman and Shragge (2003) for imaging with the teleseismic wavefield using two
Figure 10: Common conversion point stacked image using receiver function analysis on data recorded by the Laramie Broad-Band Array.

Figure 11: Migrated image of all events computed using forward-scattered P-to-S converted waves recorded by the Laramie Broad-Band Array.
synthetic datasets. In addition, preliminary images produced with the forward-scattered P-to-S mode correlate well with previous methodology. The images may be enhanced with better source wavefield modeling and wavefield interpolation prior to processing. Other improvements could be made by imaging each scattering mode after multiple and/or primary suppression or simultaneously inverting the entire wavefield to solve for a single earth model defined by all scattering modes (e.g. simultaneous multiple and primary migration).

ACKNOWLEDGMENTS

The first author would like to thank everyone in SEP for their patience and assistance while becoming accustomed to the SEP environment. We also thank Hy Yuan and Ken Dueker for use of their P wave velocity model and also the map and images from the Laramie experiment.

REFERENCES


Short Note

Imaging oceanic thermohaline structure with reflection seismology

Antoine Guitton and Ioan Vlad

INTRODUCTION

Temperature and salinity contrasts between volumes of seawater can generate reflections that are recorded before the water-bottom arrival on marine seismic data. While explorationists commonly mute them without observing them, seawater reflections are of interest to researchers of ocean dynamics.

Yilmaz (2001), at page 1809, maintains that such reflections are due to density contrasts. Holbrook et al. (2003) use a towed submersible to directly measure the salinity and temperature of water while performing a seismic survey. They conclude that thermohaline anomalies result chiefly in acoustic velocity deviations of less than 20 m/s from the average. Velocities are found by calibrating reflectivity values with local direct measurements, in the context of the aquatic medium, which lacks illumination and focusing problems.

Images of the thermohaline fine structure of the seas can be a useful tool for oceanographers. We describe a processing flow designed to extract seawater reflections from under the shot noise using prediction-error filters (PEFs). We also show an image of thermohaline reflection, as well as evidence that velocity varies with depth and midpoint in the ocean. We propose performing wave-equation migration velocity analysis (WEMVA) to find the corresponding velocity anomalies in the seawater.

PREPROCESSING

Figure 1 shows the input data from the Gulf of Mexico with a strong source component that masks the water column reflections (Figure 1a). Our goal is to unravel these reflections by performing signal/noise separation. We opted for a pattern-based signal/noise separation technique with stationary prediction-error filters (Guitton, 2003b).

This technique separates the signal and the noise by assuming that both components have different multivariate spectra that PEFs approximate. Therefore, we needed one PEF for the

---

1email: antoine@sep.stanford.edu, nick@sep.stanford.edu
noise and one PEF for the signal. We designed 2D PEFs on common-shot gathers. The noise level is so overwhelming that designing a model for it was relatively easy. We selected nine shot gathers with no water column reflections. This choice was made by inspecting all shot gathers after applying a high-pass filter to them to remove most of the source energy. We then stacked the selected gathers to increase the noise coherency and attenuate any remaining signal.

From this noise model we estimated a stationary $25 \times 5$ PEF. The signal PEF was obtained using the Spitz approximation (Guitton, 2004). The size of the signal PEF is $5 \times 2$. Having had estimated the noise and signal PEFs, we proceeded to perform noise removal. Figure 2 displays the estimated signal. A reflection is clearly visible in Figure 2a. Figures 1 and 2 have the same clip for direct comparison. In Figure 2b, the thermohaline structure of the water column is revealed. We followed with a 10-60 Hz bandpass, a gain with the first power of time, and a mute. Since several near offsets needed for prestack wavefield-continuation migration were missing (not acquired or lost to PEFs boundaries), we used an offset continuation tool for regularly sampled data (Vlad and Biondi, 2001). An image of the denoised data prestack migrated with a constant velocity of 1520 m/s is shown in Figure 3. The reflections are clearly visible throughout the section. The coherency of these events decreases for depths greater than 600m.

Figure 1: Input data showing the source noise in (a) a shot gather ($x=2500$ m) and (b) a constant-offset section ($h=130$ m). The water bottom is visible below 1.6 s. R points to a water reflection mostly hidden by the noise. [nick2-data][ER]
Figure 2: Estimated signal for (a) a shot gather (x=2500 m) and (b) a constant offset section (h=130 m). The water reflections are now clearly visible above 1.6 s. The four bad traces in (a) come from the boundary conditions for the PEFs.

Figure 3: Migration of the denoised data with a constant velocity of 1520 m/s.
FUTURE WORK: FINDING THE VELOCITY

The left panel in Figure 2 shows that even after successful elimination of coherent non-thermohaline events, large amounts of random noise are left in the data. To be able to obtain RMS velocities in a time-efficient manner, with a high-amplitudes autopicker, we obtain the semblance of each individual CMP gather. The prior knowledge of the limits of velocity variations in water allows for a small-range, high-resolution transformation. The result is seen in Figure 4. The presence of velocity variations with depth and midpoint is apparent. The absolute value of these variations does not surpass 30 m/s.

Velocity departures from the background can be seen in the left panel of Figure 5 as residual curvatures in angle-domain common image gathers (ADCIGs). This figure contains only the angles between 10° and 30°. Because of the small depth of the thermohaline reflections and of the missing near-offset information, information from incidence angles smaller than 10° was obtained only from the offset-continuation fill. We discarded it, since it was characterized completely by our estimate of constant velocity. The angles larger than 30° were unusable because of a loss of bandwidth during the transformation to ADCIGs. The small curvatures place the velocity anomalies well within the limits of the Born approximation. This means that WEMVA would be a suitable tool for resolving them.

WEMVA is an iterative inversion scheme that attempts to optimize the focusing of the migrated image (Biondi and Sava, 1999). Specifically, the result of wavefield-continuation migration is transformed to ADCIGs, the gathers are flattened, the difference from the unflattened image is taken to obtain an image perturbation, which is finally inverted into a velocity update. We plan to perform this procedure in the future, using moveout shifts computed by dip field integration (Guitton, 2003a) to flatten the gathers for the image perturbation.

CONCLUSIONS

Variations in the temperature and salinity of seawater generate velocity anomalies of the order of a few m/s that cause recordable reflections of acoustic waves in the seismic exploration frequency range. Imaging these bodies of seawater can provide useful insight into ocean dynamics. A specific processing flow is designed to enhance the visibility of these reflections, by eliminating shot noise with prediction error filters. The presence of velocity variations with depth and midpoint is shown in semblance scans and angle-domain common image gathers. Wave-equation migration velocity analysis is proposed as a possible tool for resolving the thermohaline acoustic velocity structure of seawater.

REFERENCES


Figure 4: Semblance scan volume denoised with the procedure described in Vlad (2003). In the right panel, the vertical line delimits the 1500 m/s point on the slowness axis.

Figure 5: ADCIGs between 10° and 30° after migration with a constant velocity of 1520 m/s.


Seismic waves in finely layered VTI media: Poroelasticity, Thomsen parameters, and fluid effects on shear waves

James G. Berryman

ABSTRACT

Layered earth models are well justified by experience, and provide a simple means of studying fairly general behavior of the elastic and poroelastic characteristics of seismic waves in the earth. Thomsen’s anisotropy parameters for weak elastic and poroelastic anisotropy are now commonly used in exploration, and can be conveniently expressed in terms of the layer averages of Backus. Since our main interest is usually in the fluids underground, it would be helpful to have a set of general equations relating the Thomsen parameters as directly as possible to the fluid properties. This end can be achieved in a rather straightforward fashion for these layered earth models, and the present paper develops and then discusses these relations. Furthermore, it is found that, although there are five effective shear moduli for any layered VTI medium, one and only one effective shear modulus for the layered system contains all the dependence of pore fluids on the elastic or poroelastic constants that can be observed in vertically polarized shear waves in VTI media. The effects of the pore fluids on this effective shear modulus can be substantial. An increase of shear wave speed on the order of 10% is shown to be possible when circumstances are favorable, which occurs when the medium behaves in an undrained fashion, and the shear modulus fluctuations are large (resulting in strong anisotropy). These effects are expected to be seen at higher frequencies such as sonic and ultrasonic waves for well-logging or laboratory experiments, or at seismic wave frequencies for low permeability regions of reservoirs, prior to hydrofracing. Results presented are strictly for velocity analysis.

INTRODUCTION

Gassmann’s fluid substitution formulas for bulk and shear moduli (Gassmann, 1951) were originally derived for the quasi-static mechanical behavior of fluid saturated rocks. It has been shown recently (Berryman and Wang, 2001) that deviations from Gassmann’s results at higher frequencies, especially for shear modes, can be understood when the rock is heterogeneous on the microscale, and in particular when the rock heterogeneity anywhere is locally anisotropic. On the other hand, a well-known way of generating anisotropy in the earth is through fine layering. Then, Backus’ averaging (Backus, 1962) of the mechanical behavior of
the layered isotropic media at the microscopic level produces anisotropic mechanical behavior at the macroscopic level. For our present purposes, the Backus averaging concept can also be applied to fluid-saturated porous media, and thereby permits us to study how deviations from Gassmann’s predictions could arise in an analytical and rather elementary fashion. We study layers of isotropic elastic/poroelastic materials because this is a simple, explicitly calculable model that nevertheless produces surprising results on the overall poroelastic shear modulus behavior. [If we considered instead layers of anisotropic poroelastic materials, the effects we want to study here concerning fluid-shear interactions would arrive before we begin, because they are often automatically present in anisotropic poroelastic materials as was shown earlier by Gassmann (1951) and others (Schoenberg and Douma, 1988; Sayers, 2002). So we could not show what we have set out to show here concerning the fluid effects by considering such inherently anisotropic models.] By studying both closed-pore and open-pore boundary conditions between layers within the chosen model, we learn in great detail just how violations of Gassmann’s predictions can arise in undrained versus drained conditions, or for high versus low frequency waves.

We review some standard results concerning layered VTI media in the first two sections. Then, we discuss singular value composition of the elastic (or poroelastic) stiffness matrix in order to introduce the interpretation of one shear modulus (out of the five shear moduli present) that has been shown recently (Berryman, 2004) to contain all the important behavior related to pore fluid influence on the shear deformation response. These results are then incorporated into our analysis of the Thomsen parameters (originally derived for weak anisotropy, but used here for arbitrary levels of anisotropy). For purposes of analysis, expressions are derived for the quasi-P- and quasi-SV-wave speeds and these results are then discussed from this new point of view. Numerical examples show that the approximate analysis presented is completely consistent with the full theory for layered media. Our conclusions are summarized in the final section of the paper.

NOTATION AND SOME PRIOR RESULTS

Notation for VTI media

We begin by introducing some notation needed in the remainder of the paper. For transversely isotropic media with vertical symmetry axis, the relationship between components of stress $\sigma_{kl}$ and strain $e_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}) = \frac{1}{2} \left( \frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right)$ (where $u_j$ is the $j$th component of the displacement vector) is given by

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{31} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} a & b & f \\ b & a & f \\ f & f & c \end{pmatrix} \begin{pmatrix} 2l \\ 2l \\ 2m \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{22} \\ e_{33} \\ e_{23} \\ e_{31} \\ e_{12} \end{pmatrix}, \quad (1)$$
where \( a = b + 2m \) (e.g., Musgrave, 1970; Auld, 1973), with \( i, j, k, l \) each ranging from 1 to 3 in Cartesian coordinates. The matrix describes isotropic media in the special case when \( a = c = \lambda + 2\mu, \ b = f = \lambda, \) and \( l = m = \mu. \)

The Thomsen (1986) parameters \( \epsilon, \delta, \) and \( \gamma \) are related to these stiffnesses by

\[
\epsilon \equiv \frac{a - c}{2c},
\]

\[
\delta \equiv \frac{(f + l)^2 - (c - l)^2}{2c(c - l)},
\]

\[
\gamma \equiv \frac{m - l}{2l}.
\]

Certain interpretations are allowed for these parameters when they are small enough. For P-wave propagation in the earth near the vertical, the important anisotropy parameter is \( \epsilon. \) For SV-wave propagation near the vertical, the combination \( (c/l)(\epsilon - \delta) \) plays essentially the same role as \( \delta \) does for P-waves. For SH-waves, the pertinent anisotropy parameter is \( \gamma. \) All three of the Thomsen parameters vanish for an isotropic medium, and the interpretations mentioned are valid for weakly anisotropic media such that all these parameters are relatively small \((< 1)\). However, the definitions are also useful outside the range of these constraints, and we will use the same definitions (and also continue to call them the “Thomsen parameters”) even when the smallness condition is violated; there is no fundamental problem doing this as long as it is recognized that the interpretations already mentioned in this paragraph are not necessarily valid any more when the parameters are large. This generalization of the Thomsen parameters will however require us to be careful in our subsequent usage of the parameters, as they cannot always be assumed to be small here as is usual in other treatments. Unless explicitly stated otherwise, the parameters \( \epsilon, \gamma, \) and \( \delta \) are not small quantities in this paper.

It is also useful to note for later reference that

\[
a = c(1 + 2\epsilon), \quad m = l(1 + 2\gamma), \quad \text{and} \quad f \simeq c(1 + \delta) - 2l,
\]

where smallness of \( \delta \) was in fact assumed in the third expression. In TI media, \( c \) and \( l \) are directly related to the velocities normal to the layering. Then, \( \epsilon, \gamma, \) and \( \delta \) measure the deviations from these normal velocities at other angles. We present the relevant details of the phase velocity analysis later in the paper.

**Gassmann results for isotropic poroelastic media**

To understand the significance of the results to follow, we briefly review a well-known result due to Gassmann (1951) [also see Berryman (1999b) for a tutorial]. Gassmann’s equation relates the bulk modulus \( K^* \) of a saturated, undrained isotropic porous medium to the bulk modulus \( K_{dr} \) of the same medium in the drained case:

\[
K^* = K_{dr}/(1 - \alpha B),
\]
where the parameters $\alpha$ and $B$ [respectively, the Biot-Willis parameter (Biot and Willis, 1957) and Skempton’s pore-pressure buildup coefficient (Skempton, 1954)] depend on the porous medium and fluid compliances. For the shear moduli of drained ($\mu_{dr}$) and saturated ($\mu^*$) media, Gassmann’s quasi-static theory gives

$$\mu^* = \mu_{dr}. \quad (7)$$

We want to emphasize once more that (7) is a result of the theory, not an assumption. The derivation of (6) and (7) shows that both results are elementary (and coupled) consequences of the theory. Furthermore, the two equations (6) and (7) taken together show that, for isotropic microhomogeneous media, the entire fluid effect on the overall elastic behavior is all contained in the parameter $\lambda^* = K^* - \frac{2}{3} \mu^*$, where $\lambda$ and $\mu$ are the well-known Lamé parameters. This result is crucial for understanding the significance of our later results to oil and gas exploration.

**Backus averaging**

Backus (1962) presented an elegant method of producing the effective constants for a thinly layered medium composed of either isotropic or anisotropic elastic layers. This method applies either to spatially periodic layering or to random layering, by which we mean either that the material constants change in a nonperiodic (unpredictable) manner from layer to layer or that the layer thicknesses might also be random. For simplicity, we will assume that the physical properties of the individual layers are constant and isotropic. [For applications to porous earth materials, we implicitly make the typical assumptions of spatial stationarity within these layers as well as scale separation — i.e., the sizes of the pores are much smaller than the thickness of the individual layers in which they reside.] The key idea presented by Backus is that these equations can be rearranged into a form where rapidly varying (in depth) coefficients multiply slowly varying stresses or strains.

The derivation has been given many places including Schoenberg and Muir (1989) and Berryman (1999a). Another illuminating derivation has been given recently by Milton (2002). We will not repeat any of the derivations here. The final results will be expressed in terms of averages $\langle Q \rangle$, where the brackets $\langle \cdot \rangle$ surrounding a variable $Q(z)$ indicate the volume average (or, equivalently, the linear average with depth in the vertically layered medium under consideration) of the quantity $Q$. It follows that the anisotropy coefficients in equation (1) are then related to the layer parameters by the following well-known expressions:

$$c = \left( \frac{1}{\lambda + 2\mu} \right)^{-1}, \quad (8)$$

$$f = c \left( \frac{\lambda}{\lambda + 2\mu} \right), \quad (9)$$

$$l = \left( \frac{1}{\mu} \right)^{-1}. \quad (10)$$
Fluid effects on shear waves

\[ m = \langle \mu \rangle, \]  \hspace{1cm} (11)  

\[ a = \frac{f^2}{c} + 4m - 4\left( \frac{\mu^2}{\lambda + 2\mu} \right), \]  \hspace{1cm} (12)  

and

\[ b = a - 2m. \]  \hspace{1cm} (13)  

When the layering is fully periodic, these results may be attributed to Bruggeman (1937) and Postma (1955), while for more general layered media including random media they should be attributed to Backus (1962). The constraints on the Lamé parameters \( \lambda \) and \( \mu \) for each individual layer are \( 0 \leq \mu \leq \infty \) and \( -\frac{2}{3} \leq \lambda \leq \infty \). Although, for physically stable materials, shear modulus \( \mu \) and bulk modulus \( K = \lambda + \frac{2}{3}\mu \) must both be nonnegative, these relations mean that \( \lambda \) (and also Poisson’s ratio \( \nu \)) may be negative (but nevertheless bounded below, since \( \nu \geq -1 \), and \( \lambda \geq -2\mu/3 \)). Large fluctuations in \( \lambda \) for different layers are therefore entirely possible, in principle, but may or may not be an issue for any given region of the earth.

Large fluctuations in \( \mu \) are also possible, and the Backus averaging technique is fully capable of handling all such fluctuations properly. But, if these fluctuations are too large, then the weak anisotropy assumption of Thomsen’s original work (Thomsen, 1986) will be violated and some care must be taken when writing approximate equations. We do not at any point assume weak anisotropy in this paper [except equations (5) and (28)], since the shear behavior we are trying to study will be shown to depend on the presence of strong anisotropy in this sense. We will also find it useful to develop alternatives to some of Thomsen’s formulas in order to deal with the strong anisotropy that arises in our analysis.

One very important fact known about the Backus averaging equations (Backus, 1962) is that they reduce to isotropic results with \( a = c \), \( b = f \), and \( l = m \), if the shear modulus is a constant (\( = \mu \)) — regardless of the behavior of \( \lambda \). This fact is also very important for applications involving partial and/or patchy saturation (Mavko et al., 1998; Johnson, 2001). Furthermore, this fact is closely related to the well-known bulk modulus formula of Hill (1963) for isotropic composites having uniform shear modulus, and also to the Hashin-Shtrikman bounds (Hashin and Shtrikman, 1961), which can be used to provide an elementary proof of Hill’s equation. Nevertheless, this limit will not be of much interest to us here except as a boundary condition on the results obtained. Furthermore, one of the main purposes of the paper is to show how deviations from these limiting and rather restrictive results affect the predictions of the referenced work on partial and patchy saturation.

**THOMSEN PARAMETERS \( \epsilon \) AND \( \delta \)**

**Thomsen’s \( \epsilon \)**

An important anisotropy parameter for quasi-SV-waves (which is our main interest in this paper) is Thomsen’s parameter \( \epsilon \), defined in equation (2). Formula (12) for \( a \) may be rewritten
as

\[
a = \left( \frac{(\lambda + 2\mu)^2 - \lambda^2}{\lambda + 2\mu} \right) + c \left( \frac{\lambda}{\lambda + 2\mu} \right)^2,
\]

which can be rearranged into the convenient and illuminating form

\[
a = (\lambda + 2\mu) - c \left[ \left( \frac{\lambda^2}{\lambda + 2\mu} \right) - \left( \frac{\lambda}{\lambda + 2\mu} \right)^2 \right].
\]

This formula is very instructive because the term in square brackets is in Cauchy-Schwartz form \[|q_1^2| |q_2^2| \geq (q_1 q_2)^2], so this factor is nonnegative. Furthermore, the magnitude of this term depends mainly on the fluctuations in the \(\lambda\) Lamé parameter, and is largely independent of \(\mu\), since \(\mu\) appears only in the weighting factor \(1/(\lambda + 2\mu)\). Clearly, if \(\lambda = \text{constant}\), then this bracketed factor vanishes identically, regardless of the behavior of \(\mu\). Large fluctuations in \(\lambda\) will tend to make this term large. If in addition we consider Thomsen’s parameter \(\epsilon\) written in a similar fashion as

\[
2\epsilon = \left[ (\lambda + 2\mu) \left( \frac{1}{\lambda + 2\mu} \right) - 1 \right] - \left[ \left( \frac{\lambda^2}{\lambda + 2\mu} \right) - \left( \frac{\lambda}{\lambda + 2\mu} \right)^2 \right],
\]

we find that the term enclosed in the first bracket on the right hand side is again in Cauchy-Schwartz form showing that it always makes a positive contribution unless \(\lambda + 2\mu = \text{constant}\), in which case it vanishes. Similarly, the term enclosed in the second set of brackets is always non-negative, but the minus preceding the second bracket causes this contribution to make a negative contribution to \(2\epsilon\) unless \(\lambda = \text{constant}\), in which case it vanishes. So, in general the sign of \(\epsilon\) is indeterminate. The Thomsen parameter \(\epsilon\) may have either a positive or a negative sign for a TI medium composed of arbitrary thin isotropic layers. Thomsen (2002) states that \(\epsilon > 0\) if \(K\) and \(\mu\) are positively correlated. But (16) shows that such correlations only produce \(\epsilon > 0\) with certainty if they are also supplemented by the stronger condition that \(\lambda \simeq \text{const}\) [in fact, \(\lambda \simeq \text{const}\) implies that there is a positive correlation between \(K\) and \(\mu\), but the reverse does not necessarily hold unless we also assume that the fluctuations in \(\mu\) are quite small — an assumption that we do not make here].

Fluctuations of \(\lambda\) in the earth have important implications for oil and gas exploration. As we recalled in our earlier discussion, Gassmann’s well-known results (Gassmann, 1951) show that, when isotropic porous elastic media are saturated with any fluid, the fluid has no mechanical effect on the shear modulus \(\mu\), but — when these results apply — it can have a significant effect on the bulk modulus \(K = \lambda + \frac{2}{3}\mu\), and therefore on \(\lambda\). Thus, observed (high spatial frequency) variations in layer shear modulus \(\mu\) should have no direct information about fluid content, while such variations observed in layer Lamé parameter \(\lambda\), especially if they are large variations, may contain important clues about variations in fluid content. So the observed structure of \(\epsilon\) in (16) strongly suggests that small positive and all negative values of \(\epsilon\) may be important indicators of significant fluctuations in fluid content (Berryman et al., 1999).
**Thomsen’s δ**

Thomsen’s parameter $\delta$ defined by Eq. (3) is pertinent for near vertical quasi-$P$-waves and can also be rewritten as

$$\delta = -\frac{(c + f)(c - f - 2l)}{2c(c - l)}. \quad (17)$$

This parameter is considerably more difficult to analyze than either $\gamma$ or $\epsilon$ for various reasons, some of which we will enumerate shortly. Thomsen (2002) provides some insight into the behavior of $\delta$ by noting that its sign depends only on the variations of the ratio $V_s/V_p$. This can be seen to be true from its definition by noting that

$$c - f - 2l = 2cl\left[\left(\frac{1}{\mu}\right)\left(\frac{\mu}{\lambda + 2\mu}\right) - \left(\frac{1}{\lambda + 2\mu}\right)\right] = -2cl\left(\frac{1}{\mu}\cdot \Delta \left(\frac{V_s^2}{V_p^2}\right)\right). \quad (18)$$

where

$$\Delta \left(\frac{V_s^2}{V_p^2}\right) \equiv \frac{V_s^2}{V_p^2} - \left(\frac{V_s^2}{V_p^2}\right). \quad (19)$$

Because of a controversy surrounding the sign of $\delta$ for finely layered media (e.g., Levin, 1988; Thomsen, 1988; Anno, 1997), Berryman et al. (1999) performed a series of Monte Carlo simulations with the purpose of establishing the existence or nonexistence of layered models having positive $\delta$. Those simulation results should be interpreted neither as modeling of natural sedimentation processes nor as an attempt to reconstruct any petrophysical relationships. The main goal was to develop a general picture of the distribution of the sign of $\delta$ using many choices of constituent material properties. This analysis established a similarity in the circumstances between the occurrence of positive $\delta$ and the occurrence of small positive $\epsilon$ (i.e., both occur when Lamé $\lambda$ is fluctuating greatly from layer to layer). The positive values of $\delta$ are in fact most highly correlated with the smaller positive values of $\epsilon$. We should also keep in mind the fact that $\epsilon - \delta \geq 0$ is always true for models with isotropic layers (Postma, 1955; Berryman, 1979) and this fact also plays a role in these comparisons, determining the unoccupied upper left hand corner of a $\delta$ vs. $\epsilon$ plot.

**SINGULAR VALUE DECOMPOSITION FOR STIFFNESS MATRIX**

The singular value decomposition (SVD), or equivalently the eigenvalue decomposition, of the real symmetric stiffness matrix appearing in (1) is relatively easy to perform. We can immediately write down four eigenvectors:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad (20)$$
and their corresponding eigenvalues, respectively $2l$, $2l$, $2m$, and $a - b = 2m$. All four correspond to shear modes of the system. The two remaining eigenvectors must be orthogonal to all four of these and therefore both must have the general form

$$\begin{pmatrix} 1 \\ 1 \\ X \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  

(21)

Applying (21) to the stiffness matrix in (1) shows that the corresponding eigenvalue is

$$\chi = a + b + fX,$$

(22)

where the remaining condition that determines both $X$ and $\chi$ is

$$\chi X = 2f + cX.$$

(23)

After substitution for $\chi$, we obtain a quadratic equation having the solutions

$$X_\pm = \frac{1}{2} \left( -\left[ \frac{a+b-c}{f} \right] \pm \sqrt{8 + \left[ \frac{a+b-c}{f} \right]^2} \right).$$

(24)

The ranges of values for $X_\pm$ are $0 \leq X_\pm \leq \infty$ and, since $X_- = -2/X_+$, $-\infty \leq X_- \leq 0$. The interpretation of the solutions $X_\pm$ is simple for the isotropic limit where $X_+ = 1$ and $X_- = -2$, corresponding respectively to pure compression and pure shear modes. So, except for special angles of propagation, these two modes always have mixed character, indicating that pure compression cannot be excited in the system, and must always be coupled to shear. Some types of pure shear modes can still be excited even in the nonisotropic cases, because the other four eigenvectors in (20) are unaffected by this coupling, and they are all pure shear modes. Pure compressional and shear modes are obtained as linear combinations of these two mixed modes according to

$$r \begin{pmatrix} 1 \\ 1 \\ X_+ \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \\ X_- \\ 0 \\ 0 \\ 0 \end{pmatrix} = (1+r) \begin{pmatrix} 1 \\ 1 \\ -2 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

(25)

with $r = -2(X_+ - 1)/[X_+(X_+ + 2)]$ for pure shear, and

$$\begin{pmatrix} 1 \\ 1 \\ X_+ \\ 0 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 1 \\ X_- \\ 0 \\ 0 \\ 0 \end{pmatrix} = (1+s) \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

(26)
with \( s = X_+(X_+-1)/(X_++2) \) for pure compression.

To understand the behavior of \( X_+ \) in terms of the layer property fluctuations or, alternatively, in terms of the Thomsen parameters, it is first helpful to note that the pertinent functional
\[
F(x) = \frac{1}{2} \left[ -x + \sqrt{8 + x^2} \right]
\]
is easily shown to be a monotonic function of its argument \( x \). So it is sufficient to study the behavior of the argument \( x = (a+b-c)/f \).

**Exact results for isotropic layers**

Combining results from Eqs. (8)–(12), we find after some work on rearranging the terms that

\[
\frac{a+b-c}{f} = \left( \frac{\lambda}{\lambda+2\mu} \right)^{-1} \left[ \left( \frac{\lambda}{\lambda+2\mu} \right) + 6 \left( \frac{m-\mu}{\lambda+2\mu} \right) - 8 \left\{ \left( \frac{\mu^2}{\lambda+2\mu} \right) \left( \frac{1}{\lambda+2\mu} \right) - \left( \frac{\mu}{\lambda+2\mu} \right)^2 \right\} \right],
\]

where the correction involving \( m-\mu \) in the numerator is the difference of the shear modulus from the layer-averaged shear modulus \( m \), and will be the dominant correction when fluctuations in \( \mu \) are small. The fact that \( (m-\mu)/\mu = \langle \mu \rangle (1/\mu) - 1 \geq 0 \), suggests that this dominant correction to unity (since the leading term is exactly unity) for this expression will be positive if \( \lambda \) and \( \mu \) are positively correlated throughout all the layers, but the correction could be negative in cases where there is a strong negative correlation between \( \lambda \) and \( \mu \). On the other hand, the term in curly brackets in (27) is again in Cauchy-Schwartz form (i.e., \( \langle q^2 \rangle \langle Q^2 \rangle - \langle q \rangle \langle Q \rangle^2 \geq 0 \), and therefore is always non-negative. But, since it is multiplied by \(-1\), the contribution to this expression is non-positive. This term is also quadratic in the deviations of \( \mu \) from its layer average, and thus is of higher order than the term explicitly involving \( m-\mu \). So, if the fluctuations in shear modulus are very large throughout the layered medium, the quadratic terms can dominate — in which case the overall result could be less than unity. Numerical examples developed by applying a code of V. Grechka [used previously in a similar context by Berryman et al. (1999)] confirm these analytical results.

Our main conclusion is that the shear modulus fluctuations giving rise to the anisotropy due to stacks of thin isotropic layers are (as expected) the main source of deviations of (27) from unity. But now we can say more, since positive deviations of this parameter from unity are generally associated with smaller magnitude fluctuations of the layer shear modulus, whereas negative deviations from unity must be due to large magnitude fluctuations in these shear moduli.

**Approximate results if Thomsen parameters have small values**

Using the definitions of the Thomsen parameters, we can also rewrite the terms appearing in (27) in order to make connection with this related point of view. Recalling (5) and the fact that \( b = a - 2m \), we have

\[
\frac{a+b-c}{f} \approx 1 + \frac{3}{c-2l} (c\delta + 4l\gamma) + \frac{4}{c-2l} [c(\epsilon - \delta) - 4l\gamma],
\]

(28)
with some higher order corrections involving powers of $\delta$ and products of $\delta$ with $\epsilon$ and $\gamma$ that we neglected in this equation. We have added and subtracted equally some terms proportional to $\delta$, and others proportional to $\gamma$, in order to emphasize the similarities between the form (28) and that found previously in (27). In particular, the difference $\epsilon - \delta$ is known (Postma, 1955; Berryman, 1979) to be non-negative and its deviations from zero depend on fluctuations in $\mu$ from layer to layer, behavior similar to that of the final term in (27). Since the formula (28) is only approximate and its interpretation requires the use of various other results we derive subsequently for other purposes, for now we will delay further discussion of this to a point later in the paper. [See the discussion of Eq. (62).]

**DISPERSION RELATIONS FOR SEISMIC WAVES**

The equations of motion and their solutions for seismic waves in anisotropic media are well known, and have been derived in many places including Berryman (1979) and Thomsen (1986). The dispersion relations for phase velocities are

\[ \rho \omega_{+}^2 = \frac{1}{2} \left\{ (a + l)k_1^2 + (c + l)k_3^2 \pm \sqrt{[(a - l)k_1^2 - (c - l)k_3^2]^2 + 4(f + l)^2k_1^2k_3^2} \right\}, \quad (29) \]

for quasi-compressional ($+$) waves and quasi-SV ($-$) waves (i.e., vertically polarized quasi-shear waves, by which we mean the plane normal to the cross-product of the polarization vector and the propagation vector is vertical) and

\[ \rho \omega_{-}^2 = mk_1^2 + lk_3^2, \quad (30) \]

for horizontally polarized shear waves. In these equations, $\rho$ is the overall density (including fluids when present), $\omega$ is the angular frequency, $k_1$ and $k_3$ are horizontal and vertical wavenumbers (respectively), and the phase velocities are determined simply by $V = \omega/k$ with $k = \sqrt{k_1^2 + k_3^2}$. Elastically, the SH wave depends only on the two parameters $l$ and $m$, which are not dependent in any way on layer Lamé parameter $\lambda$ and, therefore, will play no role in the poroelastic analysis. The densities of any fluids present affect all three wave speeds equally, and cannot therefore contribute to shear wave bi-refringence by itself. Thus, we can safely ignore SH except when we want to check for shear wave splitting — in which case the SH results will be most useful as a baseline for such comparisons.

The dispersion relations for quasi-P- and quasi-SV-waves can be rewritten in a number of instructive ways. One of these that we will choose for reasons that will become apparent shortly is

\[ \rho \omega_{\pm}^2 = \frac{1}{2} \left\{ (a + l)k_1^2 + (c + l)k_3^2 \right\} \pm \sqrt{[(a + l)k_1^2 + (c + l)k_3^2]^2 - 4[(ak_1^2 + ck_3^2)lk^2 + (a - l)(c - l) - (f + l)^2k_1^2k_3^2]} \]. \quad (31) \]

Written this way, it is obvious that the following two relations hold:

\[ \rho \omega_{+}^2 + \rho \omega_{-}^2 = (a + l)k_1^2 + (c + l)k_3^2, \quad (32) \]
and
\[ \rho \omega_+^2 \cdot \rho \omega_-^2 = (ak_1^2 + ck_3^2)k^2 + [(a-l)(c-l) - (f+l)^2]k_1^2k_3^2, \] (33)
either of which could have been obtained directly from (29) without the intermediate step of (31).

We are motivated to write the equations in this way in order to try to avoid evaluating the square root in (29) directly. Rather, we would like to arrive at a natural approximation that is quite accurate, but does not involve the square root operation. The desire to do this is not new (Thomsen, 1986), but our goal is different since we must necessarily treat strong anisotropy in this paper. From a general understanding of the problem, it is clear that a reasonable way of making use of (32) is to make the identifications

\[ \rho \omega_+^2 \equiv ak_1^2 + ck_3^2 - \Delta, \] (34)
and
\[ \rho \omega_-^2 \equiv lk^2 + \Delta, \] (35)
with \( \Delta \) still to be determined. Then, substituting these expressions into (33), we find that
\[ (ak_1^2 + ck_3^2 - lk^2 - \Delta)\Delta = [(a-l)(c-l) - (f+l)^2]k_1^2k_3^2 \] (36)
Solving (36) for \( \Delta \) would just give the original results back again. So the point of (36) is not to solve it exactly, but rather to use it as the basis of an approximation scheme. If \( \Delta \) is small, then we can presumably neglect it inside the parenthesis on the left hand side of (36) — or we could just keep a small number of terms in an expansion.

The leading term, and the only one we will consider here (but see the Appendix for further discussion), is
\[ \Delta = \frac{[(a-l)(c-l) - (f+l)^2]k_1^2k_3^2}{(a-l)k_1^2 + (c-l)k_3^2 - \Delta} \approx \frac{[(a-l)(c-l) - (f+l)^2]}{(a-l)/k_1^2 + (c-l)/k_3^2}. \] (37)
The numerator of this expression is known to be a positive quantity for layers of isotropic materials (Postma, 1955; Berryman, 1979). Furthermore, it can be rewritten (without approximation) in terms of Thomsen’s parameters as
\[ [(a-l)(c-l) - (f+l)^2] = 2c(e-l)(e - \delta). \] (38)
Using the first of the identities noted earlier in (5), we can also rewrite the first elasticity factor in the denominator as \( a-l = (c-l)[1 + 2c\epsilon/(c-l)] \). Combining these results in the limit of \( k_1^2 \rightarrow 0 \) (for relatively small horizontal offset), we find that
\[ \rho \omega_+^2 \simeq ck^2 + 2c\delta k_1^2, \] (39)
and
\[ \rho \omega_-^2 \simeq lk^2 + 2c(e - \delta)k_1^2, \] (40)
with $\Delta \simeq 2c(\epsilon - \delta)k^2$ for very small angles from the vertical. These two equations may be recognized simply as small angle approximations to the weak-anisotropy equations of Thomsen (1986). However, the main thrust of this paper (as we will soon see) requires strong anisotropy and therefore also requires improved approximations, which can be obtained to any desired order with only a little more effort by using (36) instead of the first approximation derived here in (37). Note that Eqs. (39) and (40) were derived without assumptions about the smallness of $\epsilon$ or $\delta$.

Although the approximations being discussed in this section are of some practical interest in their own right, their elaboration at this point would lead us away from the main theme of the paper. So, to avoid further digression here from the issue of fluid effects on shear modulus, we collect our remaining results concerning these dispersion relation approximations in the Appendix.

**INTERPRETATION OF P AND SV COEFFICIENTS FOR LAYERED MEDIA**

**General analysis for VTI media**

The correction terms, *i.e.*, those contained in the factor $\Delta$ in (35) for quasi-SV waves in anisotropic media, are proportional to the factor

$$A \equiv (a - l)(c - l) - (f + l)^2 = 2c(c - l)(\epsilon - \delta),$$

(41)

which is sometimes called the *anellipticity parameter*. Similarly, we will call $\Delta$ the *anellipticity correction*. For the case of strong anisotropy that we are considering here, the presence of $A/(c - l)$ in (40) just introduces ellipticity into the move out, but the higher order corrections that we neglected can introduce deviations from ellipticity — hence anellipticity.

Clearly, from (40) for quasi-SV-waves [and in layered media at this order of approximation], the anellipticity parameter holds all the information about the presence or absence of fluids that is not already contained in the density factor $\rho$. So it will be worth our time to study this factor in more detail. First note that, after rearrangement, we have the general identity

$$A = (f + l)(a + c - 2f - 4l) + (a - f - 2l)(c - f - 2l),$$

(42)

which is true for all transversely isotropic media.

In some earlier work (Berryman, 2003), the author has shown that it is convenient to introduce two special-purpose effective shear moduli $\mu_1^*$ and $\mu_3^*$ associated with $a$ and $c$, namely,

$$\mu_1^* \equiv a - m - f \quad \text{and} \quad 2\mu_3^* \equiv c - f.$$

(43)

Furthermore, it was shown that the combination defined by

$$G_{eff} = (\mu_1^* + 2\mu_3^*)/3$$

(44)
plays a particular role in the theory, as it is only this effective shear modulus for the anisotropic system that can also contain information about fluid content. It turns out that (42) can be rewritten in terms of this effective shear modulus if we first introduce two more parameters:

\[ K = f + l + \left( \frac{1}{a - f - 2l} + \frac{1}{c - f - 2l} \right)^{-1} \]  

(45)

and

\[ \mathcal{G} = \left[ 3G_{\text{eff}} + m - 4l \right] / 3 \]  

(46)

Then, (42) can be simply rewritten as

\[ A = 3K \mathcal{G} \]  

(47)

This result is analogous to, but distinct from, a product formula relating the effective shear modulus \( G_{\text{eff}} \) and the bulk modulus \( K_D \) to the eigenvalues of the elastic matrix according to

\[ \chi_+ \chi_- = 6K G_{\text{eff}}. \]  

(49)

Eq. (49) can be motivated by noting that, in the isotropic limit, the eigenvalues are \( 3K \) and \( 2\mu \).

[Side notes concerning layered materials: In the isotropic limit, when \( \mu \to \text{constant} \), we have \( K \to f + 2\mu / 3 \), while \( \mathcal{K} \to f + \mu \). So these two parameters are not the same, but they do have strong similarities in their behavior. In contrast, \( G_{\text{eff}} \to \mu \), while \( \mathcal{G} \to 0 \) in the same limit. It is also possible to show for layered materials that in general \( l \leq \mathcal{K} - f \leq m \), with the lower limit being optimum, i.e., attainable.]

Also, since Thomsen’s \( \delta \) plays an important role in (39), it is helpful to note that (17) can also be rewritten as

\[ c\delta = -(c - f - 2l) \left[ 1 - \frac{c - f - 2l}{2(c - l)} \right], \]  

(50)

which shows that, at least for weakly anisotropic media (in which case the deviation from unity inside the brackets is neglected), \( c\delta \) is very nearly a direct measure of the quantity \( c - f - 2l \).

**Analysis for isotropic layers**

The analysis presented in the previous subsection is general for all VTI elastic media. But we can say more by assuming now that the anisotropy arises due to layers of isotropic elastic (or possibly poroelastic) media. Then, using (8)-(12), we have the following relations

\[ f + 2l = c \left( \frac{\lambda + 2l}{\lambda + 2\mu} \right), \]  

(51)
\[ \begin{align*}
c - f - 2l &= 2c \left( \frac{\mu - l}{\lambda + 2\mu} \right), \\
an - f - 2l &= 2c \left\{ \left( \frac{2m - \mu - l}{\lambda + 2\mu} \right) - 2 \left( \frac{\mu^2}{\lambda + 2\mu} \right) \right\}. 
\end{align*} \] (52) and

Eq. (51) is an easy consequence of the Backus averaging formulas. Then, (52) shows that \( c \) differs from \( f + 2l \) only by a term that measures the difference in the weighted average of \( \mu \) and \( l \). Eq. (53) shows that \( a \) differs from \( f + 2l \) in a more complicated fashion that depends on the difference in the weighted average of \( (2m - l) \) and \( \mu \), as well as a term that is higher order in the fluctuations of the layer \( \mu \) values. Combining these results, we have

\[ G_{\text{eff}} = m - \frac{4c}{3} \left[ \frac{\mu^2}{\lambda + 2\mu} \left( \frac{1}{\lambda + 2\mu} \right)^2 \right], \] (54)

showing that all the interesting behavior (including strong \( \mu \) fluctuations in the layers together with \( \lambda \) dependence) is collected in \( G_{\text{eff}} \). Since the product of \( (52) \) and \( (53) \) is clearly of higher order in the fluctuations of the layer shear moduli, it is not hard to see that, to leading order when these fluctuation effects are small,

\[ \mathcal{A} \simeq (c - l)(3G_{\text{eff}} + m - 4l). \] (55)

To give a quick estimate, note that if all the layers have the same value of Poisson’s ratio, then the ratio \( r = \lambda / \mu \) is constant. Then, it is easy to show that \( G_{\text{eff}} = m - 4(m - l)/3(2 + r) \). Since \(-2/3 \leq r \leq \infty\), the effective shear modulus for this class of models lies in the range \( I \leq G_{\text{eff}} \leq m \). From this fact, we can conclude that the important coefficient in \( (40) \) is given to a good approximation by

\[ 2c(\epsilon - \delta) \simeq 3G_{\text{eff}} + m - 4l, \] (56)

and ranges from \( 2I \gamma \) to \( 8I \gamma \).

To study the fluid effects, the drained Lamé parameter \( \lambda \) in each layer should be replaced under undrained conditions by

\[ \lambda^* = K^* - 2\mu/3, \] (57)

where \( K^* \) was defined by \( (6) \). Then, for small fluctuations in \( \mu \), Eq. (56) shows that the leading order terms due to these shear modulus variations contributing to \( \epsilon - \delta \) actually do not depend on the fluids at all (since \( m - l \) does not depend on them). With no fluid in the pores, there is a contribution to the shear wave speed for SV in layered media, just due to the fluctuations in the shear moduli. One part of the contribution is always independent of any fluids that might be present, but the magnitude of this contribution (which is always positive) is small whenever the difference \( m - l \) is also small. If \( m - l \) is large, then the magnitude of the additional increase due to liquids in the pores can be very substantial as we will see in the following examples.
So the effects of liquids on $G_{\text{eff}}$ will generally be weak when the fluctuations in $\mu$ are weak, and strong when they are strong.

Furthermore, when the product $\alpha B \neq 0$, we first choose to define

$$ratio_{\alpha B} = \frac{m - G_{\text{eff}}}{m - l}.$$ (58)

so that, for all possible layered models, we have $0 \leq ratio_{\alpha B} \leq 1$. Then, we consider plotting the quantity $1 - ratio_{\alpha B}/ratio_0$ versus $\gamma$ (which we treat as a simple quantitative measure of the fluctuations in the layer shear moduli). To generate a class of 900 models for each of three choices of $\alpha$ (treated as a single constant for all layers in each individual model) in order to illustrate the behavior of these quantities, I made use of a code of V. Grechka [used previously in a joint publication (Berryman et al., 1999)]. This code chooses layer parameters randomly from within the following (arguable, but generally reasonable) range of values: $1.5 \leq V_p \leq 5.0$ km/s, $0.1 \leq V_s/V_p \leq 0.8$, and $1.8 \leq \rho \leq 2.8 \times 10^3$ kg/m$^3$. The results are displayed in Figure 1 for $\alpha = 0.5$, 0.8, and 0.9. We find empirically that (for $B = 1$) the values never exceed $\alpha$ for any set of choices for the layer model parameters. This apparent fact (as determined by these computer experiments) does not appear to be easy to prove from the general formula. But one simple though nontrivial calculation we can do is based again on an assumption that the bulk moduli in the layers are always proportional to the shear modulus, so $K = s\mu$, for some fixed value of the proportionality factor $s > 0$. Then, for a given model, we find that

$$1 - \frac{ratio_{\alpha B}}{ratio_0} = \frac{\alpha B}{1 + 4(1 - \alpha B)/3s} \leq \alpha B,$$ (59)

in agreement with the empirical result from the synthetic data shown in Figure 1.

To check the corresponding result for P-waves, we need to estimate $\delta$. Making use of (50), we have

$$c\delta = -2c\left\{\frac{\mu - l}{\lambda + 2\mu}\right\}\left[1 - l^{-1}\left\{\frac{\lambda + \mu}{\mu(\lambda + 2\mu)}\right\}^{-1}\left\{\frac{\mu - l}{\lambda + 2\mu}\right\}\right].$$ (60)

Working to the same order as we did for the final expression in (56), we can neglect the second term in the square brackets of (60). What remains shows that pore fluids would have an effect on this result. The result is

$$c^*\delta^* \approx -2c^*\left\{\frac{\mu - l}{\lambda^* + 2\mu}\right\}.$$ (61)

If desired, a similar replacement can also be made for $G_{\text{eff}}$ in (44) using the fact that $2(\mu^*_3 - l) = c - f - 2l$. Eq. (61) shows that, since $c^*$ and $\delta^*$ both depend on the $\lambda^*$'s (although in opposite ways, since one increases while the other decreases as $\lambda^*$ increases), the product of these factors will have some dependence on fluids. The degree to which fluctuations in $\lambda^*$ and $\mu$ are correlated, or anticorrelated, as they vary from layer to layer will also affect these results in predictable ways.
Blue dots are for $\alpha = 0.9$, red for $\alpha = 0.8$, and green for $\alpha = 0.5$. Note, that in each case, all the points for a particular choice of $\alpha$ are bounded above precisely by the value of $\alpha$. (A general proof of this empirical observation is currently lacking.) Scatter plot illustrating how $G_{\text{eff}}$ varies over a physically sensible range of layered isotropic media (see text for details) with 2700 distinct models and $B = 1$ [see Eq. (58) in the text for the definition of $\text{ratio}_\alpha$].
Interpretation of the results

Now we have derived all the results needed to interpret Eq. (28) and show how it is related to (27). First, we note some of the main terms missing from (28) are those due to approximations made to $\delta$ and the denominators of (27), which have been approximated as $f \approx c - 2l$ instead of $f \approx c(1 + \delta) - 2l$. Then, from (56), it is easy to see that the final term in (28) vanishes to lowest order, and that the remainder is given exactly by the shear modulus fluctuation terms in brackets in (53) — in complete agreement with the final terms of (27). Then, from (60), it follows that the leading contribution to the factor $c\delta + 4l\gamma$ is

$$c\delta + 4l\gamma \approx 2c \left( \frac{m - \mu}{\lambda + 2\mu} \right),$$

in complete agreement with the second term on the right hand side of (27).

In the case of very strong fluctuations in the layer shear moduli, then (53) and (60) both show that pore fluids effects are magnified due to the fluctuations in layer shear moduli and, therefore, contribute more to the anisotropy correction factors $2c^*\delta^* - \delta^*$ and $2c^*\delta^*$ for undrained porous media. So these effects will be more easily observed in seismic, sonic, or ultrasonic data under these circumstances. When these effects are present, the vertically polarized quasi-shear mode will show the highest magnitude effect, the horizontally polarized shear mode will show no effect, and the quasi-compressional mode will show an effect of intermediate magnitude. It is known that these effects, when present, are always strongest at $45^\circ$, and are diminished when the angle of propagation is either $0^\circ$ or $90^\circ$ relative to the layering direction. We will test these analytical predictions with numerical examples in the next section.

To summarize our main result here: The most significant contributions of the liquid dependence to shear waves comes into the wave dispersion formulas through coefficient $a$ (or equivalently $\epsilon$). Equations (53) and (54) show that

$$a = 2f - c + m + 3G_{eff}.$$

For small fluctuations in $\mu$, coefficients $a$ and $c$ have comparable magnitude dependence on the fluid effects, but of opposite sign. For large fluctuations, the effects on $a$ are much larger (quadratic) than those on $c$ (linear). Propagation at normal incidence will never show much effect due to the liquids, while propagation at angles closer to $45^\circ$ can show large enhancements in both quasi-P and quasi-SV waves (when shear fluctuations are large), but still no effect on SH waves.

**COMPUTED EXAMPLES**

From previous work (Berryman, 2003), we know that large fluctuations in the layer shear moduli are required before significant deviations from Gassmann’s quasi-static constant result, thereby showing that the shear modulus dependence on fluid properties can become noticeable. To generate a model that demonstrates these results, again I made use of the same code of V.
Grechka as described when presenting Figure 1. But this time I arbitrarily picked just one of the models that seemed to be most interesting for the present purposes. The parameters of this model are displayed in Table 1. The results for the various elastic coefficients and Thomsen parameters are displayed in Table 2. The results of the calculations for \( V_p \) and \( V_s \) are shown in Figures 1 and 2.

The model calculations were simplified in one way: the value of the Biot-Willis parameter was chosen to be a uniform value of \( \alpha = 0.8 \) in all layers. We could have actually computed a value of \( \alpha \) from the other layer parameters, but to do so would require another assumption about the porosity values in each layer. Doing this seemed an exercise of little value because we are just trying to show in a simple way that the formulas given here really do produce the types of results predicted analytically, and also to get a feeling for the magnitude of the effects. Furthermore, if \( \alpha \) is a constant, then it is only the product \( \alpha B \) that matters. Whatever choice of constant \( \alpha \leq 1 \) is made, it mainly determines the maximum value of the product \( \alpha B \) for \( B \) in the range [0, 1]. So, for a parameter study, it is only important not to choose too a small value of \( \alpha \), which is why the choice \( \alpha = 0.8 \) was made. This means that the maximum amplification of the bulk modulus due to fluid effects can be as high as a factor of 5 \( [= 1/(1 - \alpha)] \) for the present examples.

<table>
<thead>
<tr>
<th>Constituent</th>
<th>( K ) (GPa)</th>
<th>( \mu ) (GPa)</th>
<th>( z ) (m/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.4541</td>
<td>0.0965</td>
<td>0.477</td>
</tr>
<tr>
<td>2</td>
<td>14.7926</td>
<td>4.0290</td>
<td>0.276</td>
</tr>
<tr>
<td>3</td>
<td>43.5854</td>
<td>8.7785</td>
<td>0.247</td>
</tr>
</tbody>
</table>

Table 1. Layer parameters for the three materials in a simple layered medium used to produce the examples in Figures 2 and 3. For this model, \( \gamma = 7.882 \) (indicating strong anisotropy).
We took the porosity to be $\phi = 0.2$, and the overall density to be $\rho = (1 - \phi)\rho_s + \phi S\rho_l$, where $\rho_s = 2650.0$ kg/m$^3$, $S$ is liquid saturation ($0 \leq S \leq 1$), and $\rho_l = 1000.0$ kg/m$^3$. Then, three cases were considered: (1) Gas saturation $S = 0$ and $B = 0$, which is also the drained case, assuming that the effect of the saturating gas on the moduli is negligible. (2) Partial liquid saturation $S = 0.95$ and $B = \frac{1}{2}$ [which is intended to model a case of partial liquid saturation], intermediate between the other two cases. For smaller values of liquid saturation, the effect of the liquid might not be noticeable, since the gas-liquid mixture when homogeneously mixed will act much like the pure gas in compression, although the density effect will still be present. When the liquid fills most of the pore-space, and the gas occupies less than about 3% of the entire volume of the rock, the gas starts to become disconnected, and we expect the effect of the liquid to start becoming noticeable, and therefore we choose $B = \frac{1}{2}$ to be representative of this case. And, finally, (3) full liquid saturation $S = 1$ and $B = 1$, which is also the fully undrained case. We assume for the purposes of this example that a fully saturating liquid has the maximum possible stiffening effect on the locally microhomogeneous, isotropic, poroelastic medium.

The results shown in Figures 2 and 3 are in complete qualitative and quantitative agreement with the analytical predictions described, as expected.

**DISCUSSION AND CONCLUSIONS**

The primary question we address in this paper is this: Does the effective shear wave speed of a long-wavelength quasi-SV wave in a finely layered VTI material depend on the fluid in the porous layers, even though Gassmann’s results (Gassmann, 1951) say that — without doubt — the shear modulus in each individual layer is mechanically independent of the fluid? Perhaps surprisingly, we show that the answer to the question is positive. The quasi-SV wave always does depend on the fluid mechanics, unless the shear modulus of all the layers is exactly a uniform constant. Furthermore, the magnitude of this effect is largest when the layer shear modulus fluctuations are large.
Figure 2: Compressional wave speed $V_p$ as a function of angle $\theta$ from the vertical. Two curves shown correspond to choices of Skempton’s coefficient $B = 0$ for the drained case (dashed line) and $B = 1$ for the undrained case (solid line). The case $B = \frac{1}{2}$ (dot-dash line) is used to model partial saturation conditions as described in the text. The Biot-Willis parameter was chosen to be $\alpha = 0.8$, constant in all layers.
Figure 3: Vertically polarized shear wave speed $V_s$ as a function of angle $\theta$ from the vertical. Two curves shown correspond to choices of Skempton’s coefficient $B = 0$ for the drained case (dashed line) and $B = 1$ for the undrained case (solid). The case $B = \frac{1}{2}$ (dot-dash line) is used to model partial saturation conditions as described in the text. The Biot-Willis parameter was chosen to be $\alpha = 0.8$, constant in all layers. [NR]
In addition, our analysis leads us to consider some different ways of expanding the formulas for the dispersion relationships for the quasi-P and quasi-SV modes. These secondary results may also have some practical benefits and are illustrated in the Appendix.

Although there are five effective shear moduli for any layered VTI medium, the main result of the paper is that there is just one effective shear modulus for the layered system that contains all the dependence of elastic or poroelastic constants on pore fluids — all that can be observed in vertically polarized shear waves in VTI media. The relevant modulus $G_{eff}$ is related to uniaxial shear strain and the relevant axis of symmetry is the vertical one, normal to the bedding planes. The pore-fluid effects on this effective shear modulus can be substantial when the medium behaves in an undrained fashion, as might be expected at higher frequencies such as sonic and ultrasonic for well-logging or laboratory experiments, or at seismic frequencies for lower permeability regions of reservoirs. These predictions are clearly illustrated by the example in Figure 2.

The stiffness coefficients $a$, $b$, $c$, and $f$, all contain contributions from fluid effects for undrained layers. However, only stiffness $a$ and Thomsen parameter $\epsilon$ contain terms quadratic in layer shear modulus fluctuations, and these contributions are the ones creating the most significant effects on shear waves for strong anisotropy.

ACKNOWLEDGMENTS

I thank P. A. Berge and V. Y. Grechka for their insight and collaboration during our earlier related studies in this research area. I thank S. R. Pride and L. Thomsen for constructive criticism that helped improve the final form of the paper.

REFERENCES


Berryman, J. G., 1999a, Transversely isotropic elasticity and poroelasticity arising from thin isotropic layers: in *Theoretical and Computational Acoustics ’97, Proceedings of the Third*


**APPENDIX A**

**APPROXIMATE PHASE VELOCITIES**

Probably the most common way to write the linearized equations for the phase velocities in VTI media (Thomsen, 1986; 2002; Rüger, 2002) is

\[
V_p(\theta) \simeq V_{p0} \left[ 1 + \delta \sin^2 \theta \cos^2 \theta + \epsilon \sin^4 \theta \right], \quad (A-1)
\]

\[
V_s(\theta) \simeq V_{s0} \left[ 1 + \frac{V_{p0}^2}{V_{s0}^2} (\epsilon - \delta) \sin^2 \theta \cos^2 \theta \right], \quad (A-2)
\]

and

\[
V_{sh}(\theta) \simeq V_{s0} \left[ 1 + \gamma \sin^2 \theta \right]. \quad (A-3)
\]

The approximations in all three cases are made based on assumed smallness of the parameters $\delta$, $\epsilon$, $\gamma$, which in fact may or may not hold for any particular layered medium. However, the work in this paper demands better approximations than these, because the assumptions of weak anisotropy are always violated in the cases of most interest, *i.e.*, when the SV-wave velocity actually does depend in a significant way on fluid content. Thomsen (2002) (in his Appendix III) quotes another form of the dispersion relation for $V_p(\theta)$ that is more useful for our purposes (modified here to correct an obvious typo in the leading term):

\[
V_p^2(\theta) \simeq V_{p0}^2 + 2V_{nmo}^2 (\delta \sin^2 \theta + \eta \sin^4 \theta), \quad (A-4)
\]
where \( V_{p0}^2 = c/\rho, \) \( V_{nmo} \simeq V_{p0}(1 + \delta), \) and \( \eta = (\epsilon - \delta)/(1 + 2\delta) \) is the combination of parameters introduced by Alkhalifah and Tsvankin (1996). Although (A-4) is still an approximation, it is much closer in form to the dispersion relation quoted here in (34). The form (A-4) still assumes smallness of the anisotropy parameters, but the usual square root approximation has not been made yet, so the correspondence with (34) is easier to scan. If we neglect higher order contributions to \( \Delta \) and thereby make the approximation that
\[
\Delta \simeq 2k^2c(\epsilon - \delta)\sin^2\theta\cos^2\theta, \tag{A-5}
\]
then (34) becomes
\[
V_p^2(\theta) \simeq V_{p0}^2 + 2V_{p0}^2\delta\sin^2\theta + 2V_{p0}^2(1 + 2\delta)\eta\sin^4\theta. \tag{A-6}
\]
If in addition we also make the small anisotropy approximations \( V_{p0}^2\delta \simeq V_{nmo}^2\delta \) and \( V_{p0}^2(1 + 2\delta) \simeq V_{nmo}^2 \) in (A-6), then the result recovers (A-4).

Our main goal in this Appendix is to make a direct comparison between the exact formulas (34) and (35), the approximate formulas resulting from (34) and (35) when \( \Delta \) is replaced by its first approximation (37), and either standard equations (A-1) and (A-2) or approximation (A-4) and some yet to be determined companion equation for quasi-SV waves. The easiest and most consistent way to arrive at an appropriate approximate form for \( V_{sv}^2 \) is to use the exact relations (34) and (35) to determine what effective value of \( \Delta_{eff} \) has been used in (A-4) and then use it again in (35). We find
\[
\Delta_{eff}(\theta) \simeq 2k^2\rho V_{p0}^2(\epsilon - \delta)\sin^2\theta\cos^2\theta - 2k^2\rho V_{p0}^2\delta^2(2 + \delta)\sin^2\theta. \tag{A-7}
\]
However, this formula has the undesirable characteristic that it does not vanish as it should for \( \theta = 90^\circ \). The offending terms are second order in \( \delta \) and therefore are usually neglected for weak anisotropy. But the weak anisotropy assumptions implicit in (A-4) are not valid in the present context, so this is nevertheless a problem for us here. Making proper allowance for this, we can arrive at a corrected \( \Delta \) that has the desired behavior and still agrees with the prior results under weak anisotropy conditions:
\[
\Delta_{corr}(\theta) \simeq 2k^2\rho V_{nmo}^2\eta\sin^2\theta\cos^2\theta, \tag{A-8}
\]
and, therefore, that a good choice for \( V_{sv}^2 \) to the same level of approximation is
\[
V_{sv}^2(\theta) = V_{so}^2 + \Delta_{corr}(\theta)/k^2\rho. \tag{A-9}
\]
But these modifications have led us back to the approximation (34) and therefore provide nothing new. So instead of comparisons to (A-4) and (A-9), we will choose to make our comparisons to (A-1) and (A-2). In particular, these two equations amount to using
\[
\Delta_{eff} = 2k^2\rho V_{p0}^2(\epsilon - \delta)\sin^2\theta\cos^2\theta, \tag{A-10}
\]
extcept for some higher order corrections in \( \Delta_{eff} \) which would always be small if the anisotropy were really always weak. Although Eqs. (A-8) and (A-10) are apparently the same, this fact is a bit misleading since \( \Delta \) and its corrections arise in the final results in different ways.
Figure A-1: Compressional wave velocities as computed exactly from the dispersion relation (34), by (34) using approximation (37) for $\Delta$, and by the linear approximation (64). This layered model is the same as in Figures 2 and 3 for the case $B = 1$. [NR]

because of differing square root approximations and different assumptions about the presence or absence of strong anisotropy.

Numerical comparisons of these three sets of results for quasi-P and quasi-SV waves are summarized in Figures 4 and 5 for one strong anisotropy example. The comparison is obviously not a fair one for the weak anisotropy equations since they are being used beyond their acknowledged (and expected) range of validity. The main point of the exercise is to see that the approximations made here give reasonable approximations to the exact results for the full range of possible incidence angles for strong anisotropy conditions, while the standard results do not fair as well. All the methods agree quite well for compressional waves in this model. The evaluation of (35) using (37) to approximate $\Delta$ gives a clear improvement over (A-2) for the quasi-SV wave velocity in a range of intermediate angles. Overall, the weak anisotropy formulas (A-1) and (A-2) give better results for strong anisotropy in this case than might have been expected.
Figure A-2: Shear wave velocities as computed exactly from the dispersion relation (35), by (35) using approximation (37) for $\Delta$, and by the linear approximation (66). This layered model is the same as in Figures 2 and 3 for the case $B = 1$. [jim1-shearvels] [NR]
Poroelastic fluid effects on shear for rocks with soft anisotropy

James G. Berryman

ABSTRACT
A general analysis of poroelasticity for vertical transverse isotropy (VTI) shows that four eigenvectors are pure shear modes with no coupling to the pore-fluid mechanics. The remaining two eigenvectors are linear combinations of pure compression and uniaxial shear, both of which are coupled to the fluid mechanics. After reducing the problem to a $2 \times 2$ system, the analysis shows in a relatively elementary fashion how a poroelastic system with isotropic solid elastic frame but with anisotropy introduced through the poroelastic coefficients interacts with the mechanics of the pore fluid and produces shear dependence on fluid properties in the overall poroelastic system. The analysis shows, for example, that this effect is always present (though sometimes small in magnitude) in the systems studied, and can be quite large (up to a definite maximum increase of 20 per cent) in some rocks — including Spirit River sandstone and Schuler-Cotton Valley sandstone.

INTRODUCTION
An important paper by Gassmann (1951) concerns the effects of fluids on the mechanical properties of porous rock. His main result is the well-known fluid-substitution formula (that now bears his name) for the bulk modulus in undrained, isotropic poroelastic media. He also postulated that the effective undrained shear modulus would (in contrast to the bulk modulus) be independent of the mechanical properties of the fluid when the medium is isotropic. That the independence of shear modulus from fluid effects is guaranteed for isotropic media at very low or quasistatic frequencies was shown recently by Berryman (1999) to be tightly coupled to the original bulk modulus result of Gassmann; each result implies the other in isotropic media. It has gone mostly without discussion in the literature that Gassmann (1951) also derived general results for anisotropic porous rocks in the same 1951 paper. It is not hard to see that these results imply that, contrary to the isotropic case, some of the overall undrained shear moduli in fact may depend on fluid properties in anisotropic media, thus mimicking the bulk modulus behavior. However, Gassmann’s paper does not remark at all on this difference in behavior between isotropic and anisotropic porous rocks. Brown and Korringa (1975) also address the same class of problems, including both isotropic and anisotropic cases, but again they do not remark on the shear modulus results in either case. Norris (1993) studies partial saturation in isotropic layered materials in the low-frequency regime ($\lesssim 100$ Hz) and takes as a
fundamental postulate that Gassmann’s results hold for the low frequency shear modulus, but it seems that some justification should be provided for such an assumption, and furthermore some indication of its range of validity established.

On the other hand, Hudson (1981), in his early work on cracked solids, explicitly demonstrates differences between fluid-saturated and dry cracks and relates his work to that of Walsh (1969) and O’Connell and Budiansky (1974), but does not make any connection to the work of either Gassmann (1951), or Brown and Korringa (1975). Mukerji and Mavko (1994) show numerical results based on work of Gassmann (1951), Brown and Korringa (1975) and Hudson (1981) demonstrating the fluid dependence of shear in anisotropic rock, but again they do not remark on these results at all. Mavko and Jizba (1991) use a simple reciprocity argument to establish a direct, but approximate, connection between undrained shear response and undrained compressional response in rocks containing cracks. Berryman and Wang (2001) show that deviations from Gassmann’s results sufficient to produce shear modulus dependence on fluid mechanical properties require the presence of some anisotropy on the microscale, thereby explicitly violating the microhomogeneous and microisotropy conditions implicit in Gassmann’s original derivation. Berryman et al. (2002a) go further and make use of differential effective medium analysis to show explicitly how the undrained, overall isotropic shear modulus can depend on fluid trapped in penny-shaped cracks. Meanwhile, laboratory results for wave propagation [see Berryman et al. (2002b)] show conclusively that the shear modulus does indeed depend on fluid mechanical properties for low-porosity, low-permeability rocks, and high-frequency laboratory experiments ($f > 500$ kHz).

One thing lacking from all the preceding work is a simple example showing how the presence of anisotropy influences the shear modulus, and specifically when and how the shear modulus becomes fluid dependent. Our main purpose in the present work is therefore to demonstrate, in a set of rather simple examples, how the overall shear behavior becomes coupled to fluid compressional properties at high frequencies in anisotropic media — even though overall shear modulus is always independent of the fluid properties in microhomogeneous isotropic media at sufficiently low frequencies, whether drained or undrained. Two other distinct but related analyses addressing this topic have been presented recently by the author (Berryman, 2004b,c). Both of these prior papers have made explicit use of layered media, composed of isotropic poroelastic materials, together with exact results for such media based on Backus averaging (Backus, 1962). In contrast, the present analysis does not make use of such a specific model, and is therefore believed to be about as simple as possible, while still achieving the level of understanding desired for this rather subtle technical issue. One important simplification we make here in order to separate the part that is due to poroelastic effects, from the part that would be present in any elastic (i.e., possibly zero permeability porous medium) is to model each material as if the elastic part is entirely isotropic, while the poroelastic effects [i.e., the Biot-Willis coefficients (Biot and Willis, 1957) for the anisotropic overall material] provide the only sources of anisotropy in the system. Thus, we specifically distinguish two possible sources of anisotropy, the elastic or “hard” anisotropy that is assumed not to be present here, and the poroelastic or “soft” anisotropy that is the source of the effects we want to study in this paper.

Our analysis for general transversely isotropic media is presented in the next three sec-
tions. In particular, the section on eigenvectors also introduces the effective undrained shear modulus relevant to our general discussion. Examples are then presented for glass, granite, and sandstone. The paper’s results and conclusions are summarized in the final section. Some mathematical details are collected in the Appendix.

**FLUID-SATURATED POROELASTIC ROCKS**

In contrast to traditional elastic analysis, the presence in rock of a saturating pore fluid introduces the possibility of an additional control field and an additional type of strain variable. The pressure $p_f$ in the fluid is a new field parameter that can be controlled. Allowing sufficient time for global pressure equilibration permits us to consider $p_f$ to be a constant throughout the percolating (connected) pore fluid, while restricting the analysis to quasistatic processes. (But ultimately we are not interested in such quasi-static processes in this paper, as we are trying to reconcile laboratory wave data with the theory.) The change $\zeta$ in the amount of fluid mass contained in the pores [see Biot (1962) or Berryman and Thigpen (1985)] is a new type of strain variable, measuring how much of the original fluid in the pores is squeezed out during the compression of the pore volume while including the effects of compression or expansion of the pore fluid itself due to changes in $p_f$. It is most convenient to write the resulting equations in terms of compliances rather than stiffnesses, so the basic equation to be considered takes the following form for isotropic media:

\[
\begin{pmatrix}
e_{11} \\
e_{22} \\
e_{33} \\
-\zeta
\end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} & s_{12} & -\beta \\
s_{12} & s_{11} & s_{12} & -\beta \\
s_{12} & s_{12} & s_{11} & -\beta \\
-\beta & -\beta & -\beta & \gamma
\end{pmatrix} \begin{pmatrix} \sigma_{11} \\
\sigma_{22} \\
\sigma_{33} \\
-p_f
\end{pmatrix}.
\] (1)

The constants appearing in the matrix on the right hand side will be defined in the following two paragraphs. It is important to write the equations this way rather than using the inverse relation in terms of the stiffnesses, because the compliances $s_{ij}$ appearing in (1) are simply related to the drained elastic constants $\lambda_{dr}$ and $G_{dr}$ in the same way they are related in normal elasticity, whereas the individual stiffnesses obtained by inverting the equation in (1) must contain coupling terms through the parameters $\beta$ and $\gamma$ that depend on the pore and fluid compliances. Thus, we find that

\[
s_{11} = \frac{1}{E_{dr}} = \frac{\lambda_{dr} + G_{dr}}{G_{dr}(3\lambda_{dr} + 2G_{dr})}
\] (2)

and

\[
s_{12} = \frac{\nu_{dr}}{E_{dr}},
\] (3)

where the drained Young’s modulus $E_{dr}$ is defined by the second equality of (2) and the drained Poisson’s ratio is determined by

\[
\nu_{dr} = \frac{\lambda_{dr}}{2(3\lambda_{dr} + 2G_{dr})}.
\] (4)
When the external stress is hydrostatic so \( \sigma = \sigma_{11} = \sigma_{22} = \sigma_{33} \), equation (1) telescopes down to

\[
\begin{pmatrix}
e & -\zeta \\
-\zeta & 1/K_{dr} & -\alpha/K_{dr} \\
-\alpha/K_{dr} & \alpha/BK_{dr} & -p_f
\end{pmatrix}
\]

where \( e = e_{11} + e_{22} + e_{33} \), \( K_{dr} = \lambda_{dr} + \frac{2}{3}G_{dr} \) is the drained bulk modulus, \( \alpha = 1 - K_{dr}/K_m \) is the Biot-Willis parameter (Biot and Willis, 1957) with \( K_m \) being the bulk modulus of the solid minerals present, and Skempton’s pore-pressure buildup parameter \( B \) (Skempton, 1954) is given by

\[
B = \frac{1}{1 + K_p(1/K_f - 1/K_m)}
\]

New parameters appearing in (6) are the bulk modulus of the pore fluid \( K_f \) and the pore modulus \( K_p^{-1} = \alpha/\phi K_{dr} \) where \( \phi \) is the porosity. The expressions for \( \alpha \) and \( B \) can be generalized slightly by supposing that the solid frame is composed of more than one constituent, in which case the \( K_m \) appearing in the definition of \( \alpha \) is replaced by \( K_s \) and the \( K_m \) appearing explicitly in (6) is replaced by \( K_\phi \) (see Brown and Korringa, 1975; Rice and Cleary, 1976; Berryman and Wang, 1995). This is an important additional complication (Berge and Berryman, 1995), but — for the sake of desired simplicity — we will not pursue the matter further here.

Comparing (1) and (5), we find that

\[
\beta = \frac{\alpha}{3K_{dr}}
\]

and

\[
\gamma = \frac{\alpha}{BK_{dr}}.
\]

As we develop the ideas to be presented here, we will need to treat Eqs. (1)–(6) as if they are true locally, but perhaps not globally. In particular, if we assume overall drained conditions, then \( p_f = \) a constant everywhere. But, if we assume locally undrained conditions, then \( p_f \approx \) a constant in local patches, but these local constant values may differ from patch to patch. This way of thinking about the system is intended to mimic the behavior expected when a high frequency wave propagates through a system having highly variable (or just uniformly very low) fluid permeability everywhere.

**RELATIONS FOR ANISOTROPY IN POROELASTIC MATERIALS**

Gassmann (1951), Brown and Korringa (1975), and others have considered the problem of obtaining effective constants for anisotropic poroelastic materials when the pore fluid is confined within the pores. The confinement condition amounts to a constraint that the increment of fluid content \( \zeta = 0 \), while the external loading \( \sigma \) is changed and the pore-fluid pressure \( p_f \) is allowed to respond as necessary and thus equilibrate.
To provide an elementary derivation of the Gassmann equation for anisotropic materials, we consider the anisotropic generalization of (1)

$$
\begin{pmatrix}
\epsilon_{11} \\
\epsilon_{22} \\
\epsilon_{33}
\end{pmatrix} =
\begin{pmatrix}
s_{11} & s_{12} & s_{13} - \beta_1 \\
s_{12} & s_{22} & s_{23} - \beta_2 \\
s_{13} & s_{23} & s_{33} - \beta_3
\end{pmatrix}
\begin{pmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33}
\end{pmatrix} + \begin{pmatrix}
\gamma \\
-\beta_1 \\
-\beta_2 \\
-\beta_3
\end{pmatrix}
\cdot
\begin{pmatrix}
p_f
\end{pmatrix}.
$$

(9)

Three shear contributions have been immediately excluded from consideration since they can easily be shown not to interact mechanically with the fluid effects. This form is not completely general in that it includes orthorhombic, cubic, hexagonal, and all isotropic systems, but excludes triclinic, monoclinic, trigonal, and some tetragonal systems that would have some nonzero off-diagonal terms in the full elastic matrix. Also, we have assumed that the material axes are aligned with the spatial axes. But this latter assumption is not significant for the derivation that follows. Such an assumption is important when properties of laminated materials having arbitrary orientation relative to the spatial axes need to be considered, but we do not treat this more general problem here.

If the fluid is confined (or undrained on the time scales of interest), then $\zeta \equiv 0$ in (9) and $p_f$ becomes a linear function of $\sigma_{11}, \sigma_{22}, \sigma_{33}$. Eliminating $p_f$ from the resulting equations, we obtain the general expression for the strain dependence on external stress under such confined conditions:

$$
\begin{pmatrix}
\epsilon_{11} \\
\epsilon_{22} \\
\epsilon_{33}
\end{pmatrix} =
\begin{pmatrix}
s_{11} & s_{12} & s_{13} \\
s_{12} & s_{22} & s_{23} \\
s_{13} & s_{23} & s_{33}
\end{pmatrix} - \gamma^{-1} \begin{pmatrix}
\beta_1 \\
\beta_2 \\
\beta_3
\end{pmatrix}
\begin{pmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33}
\end{pmatrix}
\equiv
\begin{pmatrix}
s_{11}^{*} & s_{12}^{*} & s_{13}^{*} \\
s_{12}^{*} & s_{22}^{*} & s_{23}^{*} \\
s_{13}^{*} & s_{23}^{*} & s_{33}^{*}
\end{pmatrix}
\begin{pmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33}
\end{pmatrix}.
$$

(10)

The $s_{ij}$’s are fluid-drained constants, while the $s_{ij}^{*}$’s are the fluid-undrained (or fluid-confined) constants. The fundamental result (10) was obtained earlier by both Gassmann (1951) and Brown and Korringa (1975), and may be written simply as

$$
s_{ij}^{*} = s_{ij} - \frac{\beta_i \beta_j}{\gamma}, \quad \text{for} \quad i, j = 1, 2, 3.
$$

(11)

This expression is just the anisotropic generalization of the well-known Gassmann equation for isotropic, microhomogeneous porous media.

**EIGENVECTORS FOR TRANSVERSE ISOTROPY**

The $3 \times 3$ system (10) can be analyzed fairly easily, and in particular the eigenfunctions and eigenvalues of this system can be obtained in general. However, such general results do not provide much physical insight into the problem we are trying to study, so instead of proceeding in this direction we will now restrict attention to transversely isotropic materials. This case
is relevant to many layered earth materials and also industrial systems, and it is convenient because we can immediately eliminate one of the eigenvectors from further consideration. Three mutually orthogonal (but unnormalized) vectors of interest are:

$$\begin{align*}
v_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \\
v_2 &= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \\
v_3 &= \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.
\end{align*}$$  \hspace{1cm} (12)

Treating these vectors as stresses, the first corresponds to a simple hydrostatic stress, the second to a planar shear stress, and the third to a pure shear stress applied uniaxially along the z-axis (which would also be the symmetry axis for a layered system, but we are not treating such layered systems here). Transverse isotropy of the system under consideration requires:

$$s_{11} = s_{22}, \ s_{13} = s_{23}, \text{ and for the poroelastic problem } \beta_1 = \beta_2. \text{ Thus, it is immediately apparent that the planar shear stress } v_2 \text{ is an eigenvector of the system, and furthermore it results in no contribution from the pore fluid. Therefore, this vector will be of no further interest here, and the system can thereby be reduced to } 2 \times 2.

### Compliance formulation

If we define the effective compliance matrix for the system as $S^*$ having the matrix elements given in (11), then the bulk modulus for this system is defined in terms of $v_1$ by

$$\frac{1}{K_u} = v_1^T S^* v_1 = \frac{1}{K_{dr}} - \gamma^{-1} (2\beta_1 + \beta_3)^2,$$  \hspace{1cm} (13)

where the $T$ superscript indicates the transpose, and $1/K_{dr} \equiv \sum_{i,j=1}^{3} s_{ij}$. This is the result usually quoted as Gassmann’s equation for the bulk modulus of the undrained (or confined) anisotropic (VTI) system. Also, note that in general

$$\sum_{i=1}^{3} \beta_i = 2\beta_1 + \beta_3 = \alpha/K_{dr}.$$

\hspace{1cm} (14)

Thus, even though $v_1$ is not an eigenvector of this system, it nevertheless plays a fundamental role in the mechanics. Furthermore, this role is quite well-understood. What is perhaps not so well-understood then, especially for poroelastic systems, is the role of $v_3$. Understanding this role will become our main focus for the remainder of this discussion.

The true eigenvectors of the $2 \times 2$ subproblem of interest (i.e., in the space orthogonal to the four pure shear eigenvectors already discussed) are necessarily linear combinations of $v_1$ and $v_3$. We can construct the relevant contracted operator for the $2 \times 2$ subsystem by considering:

$$\begin{pmatrix} v_1^T \\ v_3^T \end{pmatrix} S^* \begin{pmatrix} v_1 \\ v_3 \end{pmatrix} \equiv \begin{pmatrix} 9A_{11}^* \\ 18A_{13}^* \end{pmatrix},$$

\hspace{1cm} (15)

(in all cases the * superscripts indicate that the pore-fluid effects are included) and the reduced matrix

$$\Sigma^* = A_{11}^* v_1 v_1^T + A_{13}^* (v_1 v_3^T + v_3 v_1^T) + A_{33}^* v_3 v_3^T,$$  \hspace{1cm} (16)
where

\[
A_{11}^* = \frac{2(s_{11}^* + s_{12}^* + 2s_{13}^*) + s_{33}^*}{9},
\]
\[
A_{13}^* = \frac{(s_{11}^* + s_{12}^* - s_{13}^* - s_{33}^*)}{9},
\]
\[
A_{33}^* = \frac{(s_{11}^* + s_{12}^* - 4s_{13}^* + 2s_{33}^*)}{18}.
\]

Providing some understanding of these connections and the implications for shear modulus dependence on fluid content is one of our goals.

First we remark that \(A_{11}^* = 1/9K_u\), where \(K_u\) is again the undrained (or Gassmann) bulk modulus for the system in (13). Therefore, \(A_{11}^*\) is proportional to the undrained bulk compliance of this system. The other two matrix elements cannot be given such simple interpretations in general. To simplify the analysis we note that, at least for purposes of modeling, anisotropy of the compliances \(s_{ij}\) and the poroelastic coefficients \(\beta_i\) can be treated independently. Anisotropy displayed in the \(s_{ij}\)’s corresponds mostly to the anisotropy in the solid elastic components of the system, while anisotropy in the \(\beta_i\)’s corresponds mostly to anisotropy in the shapes and spatial distribution of the porosity. We will therefore distinguish these contributions by calling anisotropy appearing in the \(s_{ij}\)’s the “hard anisotropy,” and the anisotropy in the \(\beta_i\)’s will in contrast be called the “soft anisotropy.”

Now, it is clear (also see the discussion in the Appendix for more details) that the eigenvectors having unit magnitude \(f(\theta)\) for this problem (i.e., for the reduced operator \(\Sigma^*\)) necessarily take the form

\[
f(\theta) = \overline{v}_1 \cos \theta + \overline{v}_3 \sin \theta,
\]

where \(\overline{v}_1 = v_1/\sqrt{3}\) and \(\overline{v}_3 = v_3/\sqrt{6}\) are the normalized eigenvectors. Two solutions for the rotation angle are: \(\theta_-\) and \(\theta_+ = \theta_- + \frac{\pi}{3}\), guaranteeing that the two solutions (the eigenvectors) are orthogonal. It is easily seen that the eigenvalues are given by

\[
\Lambda_{\pm}^* = 3 \left[ A_{33}^* + A_{11}^*/2 \pm \sqrt{(A_{33}^* - A_{11}^*/2)^2 + 2(A_{13}^*)^2} \right]
\]

and the rotation angles are determined by

\[
\tan \theta_{\pm}^* = \frac{\Lambda_{\pm}^* / 3 - A_{11}^*}{\sqrt{2}A_{13}^*} = \frac{A_{33}^* - A_{11}^*/2 \pm \sqrt{(A_{33}^* - A_{11}^*/2)^2 + 2(A_{13}^*)^2}}{\sqrt{2}A_{13}^*}.
\]

One part of the rotation angle is due to the drained (fluid free) “hard anisotropic” nature of the rock frame material. We will call this part \(\overline{\theta}\). The remainder is due to the presence of the fluid in the pores, and we will call this part \(\delta \equiv \theta^* - \overline{\theta}\) for the “soft anisotropy.” Using a standard formula for tangents, we have

\[
\delta \theta_{\pm} = \tan^{-1} \left[ \frac{\tan \theta_{\pm}^* - \tan \overline{\theta}_{\pm}}{1 + \tan \theta_{\pm}^* \tan \overline{\theta}_{\pm}} \right].
\]

Furthermore, definite formulas for \(\overline{\theta}_{\pm}\) are found from (20) by taking \(\gamma \to \infty\) (corresponding to air saturation of the pores).
Since
\[ \tan\theta_+ \tan\theta_- = -1, \tag{22} \]
it is sufficient to consider just one of the signs in front of the radical in (20). The most convenient choice for analytical purposes turns out to be the minus sign (which corresponds to the eigenvector with the larger component of pure compression). Furthermore, it is also clear from the form of (20) that often the behavior of most interest to us here occurs for cases when \( A_{13}^* \neq 0 \).

In the limit of a nearly isotropic solid frame (so the “hard anisotropy” vanishes and thus we will also call this the “quasi-isotropic” limit), it is not hard to see that
\[ A_{33}^* \simeq \frac{1}{12G_{dr}} - \frac{(\beta_1 - \beta_3)^2}{9\gamma}, \tag{23} \]
where \( G_{dr} \) is the drained shear modulus of the quasi-isotropic solid frame. Similarly, the remaining coefficient
\[ A_{13}^* \simeq -\frac{(\beta_1 - \beta_3)(2\beta_1 + \beta_3)}{9\gamma}, \tag{24} \]
since all the solid contributions approximately cancel in this limit.

To clarify the situation further, we will enumerate three cases:

**Case I.** \( A_{33}^* - A_{11}^*/2 \neq 0, A_{13}^* = 0 \).

Whenever \( A_{33}^* - A_{11}^*/2 \neq 0 \) and \( A_{13}^* \to 0 \), we find easily that \( \theta^* \to 0 \), while \( \theta^*_+ \to \pi/2 \). In this case, \( v_1 \) and \( v_3 \) are themselves the eigenvectors, while the eigenvalues are proportional to \( A_{11}^* \) and \( A_{33}^* \). In the quasi-isotropic limit, \( A_{13}^* \) can vanish only if \( \beta_1 - \beta_3 = 0 \), in which case \( A_{33}^* \) also does not depend on fluid properties. For media differing significantly from the quasi-isotropic limit, \( A_{13}^* \) could vanish for some physically interesting situations, but the resulting physical constraints are too special (and complicated) for us to consider them further here.

**Case II.** \( A_{33}^* - A_{11}^*/2 = 0, A_{13}^* \neq 0 \).

For this case, \( \tan\theta^*_\pm = \pm 1 \), so \( \theta^*_\pm = \pm \pi/4 \). The two eigenvectors are \( v_1/\sqrt{6} \pm v_3/\sqrt{12} \), with no dependence on the fluid properties. However, the eigenvalues continue to be functions of the fluid properties. This seems to be a rather special case, but again considering the quasi-isotropic limit, we find that \( A_{33}^* - A_{11}^*/2 \simeq v/2E + [(2\beta_1 + \beta_3)^2 - 2(\beta_1 - \beta_3)^2]/18\gamma \), where \( v \) is Poisson’s ratio and \( E \) is Young’s modulus. For this combination of the parameters to vanish for special values does not appear to violate any of the well-known constraints (such as positivity, etc.) on these parameters. For example, if \( \beta_1 = 0 \), the term depending on the fluid properties clearly makes a negative contribution, which might be large enough to cancel the contribution from the solid. But, for now, this case seems rather artificial, so we will not consider it further here.
Case III. \( A_{33}^* - A_{11}^*/2 \neq 0, A_{13}^* \neq 0 \).

This case is the most general one of the three, and the one we will study at greater length in the remainder of this discussion.

We want to understand how the introduction of liquid into the pore space affects the shear modulus. We also want to know how the anisotropy influences, \( i.e. \), aids or hinders, the impact of the liquid on the shear behavior. To achieve this understanding, it should be sufficient to consider the case when \((A_{13}^*)^2 \ll (A_{33}^* - A_{11}^*/2)^2\), assuming as we do that both factors are nonzero. Then, expanding the square root in (19), we have

\[
\Lambda_+^* = 6A_{33}^* + \Delta \quad \text{and} \quad \Lambda_-^* = 3A_{11}^* - \Delta,
\]

where \( \Delta \) is defined consistently by either of the two preceding expressions or by \(2\Delta \equiv \Lambda_+^* - \Lambda_-^* + 3A_{11} - 6A_{33}\) and is also given approximately for cases of interest here by

\[
\Delta \simeq \frac{3(A_{13}^*)^2}{A_{33}^* - A_{11}^*/2}.
\]

In the quasi-isotropic soft anisotropy limit under consideration, we find

\[
\Delta \simeq \frac{2(\beta_1 - \beta_3)^2(2\beta_1 + \beta_3)^2/27\gamma^2}{\nu/E + [(2\beta_1 + \beta_3)^2 - 2(\beta_1 - \beta_3)^2]/9\gamma}.
\]

All of the mechanical effects of the liquid that contribute to this formula appear in the factor \( \gamma \). The order at which \( \gamma \) appears depends on the relative importance of the two terms in the denominator of this expression. If the second term ever dominates, then one factor of \( \gamma \) cancels, and therefore \( \Delta \sim O(\gamma^{-1}) \), and furthermore \( \Delta \sim 2(\beta_1 - \beta_3)^2/3\gamma \) if \( |\beta_1 - \beta_3| << |2\beta_1 + \beta_3| \). If instead what seems to be the more likely situation holds and the first term in the denominator dominates, then \( \Delta \sim O(\gamma^{-2}) \). So in either of these cases, as long as \( \beta_1 - \beta_3 \neq 0 \) (which is the condition for soft anisotropy), we always have contributions to \( \Delta \) from liquid mechanical effects. There do not appear to be any combinations of the parameters for which the fluid effects disappear whenever the material is in the class of anisotropic solids considered here.

**Stiffness formulation**

The dual to the problem just studied replaces compliances everywhere with stiffnesses, and then proceeds as before. Equations (15)–(18) are replaced by

\[
\begin{pmatrix}
v_1^T \\
v_3^T
\end{pmatrix}
C^*(v_1 \ v_3) \equiv
\begin{pmatrix}
9B_{11}^* & 18B_{13}^* \\
18B_{13}^* & 36B_{33}^*
\end{pmatrix}
\]

(in all cases the * superscripts indicate that the pore-fluid effects are included) and the reduced matrix

\[
(\Sigma^*)^{-1} = B_{11}^*v_1v_1^T + B_{13}^*(v_1v_3^T + v_3v_1^T) + B_{33}^*v_3v_3^T,
\]
where
\[
\begin{align*}
B_{11}^* &= \frac{[2(c_{11}^* + c_{12}^* + 2c_{13}^*) + c_{33}^*]}{9}, \\
B_{13}^* &= \frac{(c_{11}^* + c_{12}^* - c_{13}^* - c_{33}^*)}{9}, \\
B_{33}^* &= \frac{(c_{11}^* + c_{12}^* - 4c_{13}^* + 2c_{33}^*)}{18}.
\end{align*}
\] (30)

It is a straightforward exercise to check that the two reduced problems are in fact inverses of each other. We will not repeat this analysis here, as it is wholly repetitive of what has gone before. The main difference in the details is that the expressions for the \(B\)'s in terms of the \(\alpha\)'s are rather more complicated than those for the compliance version, which is also why we chose to display the compliance formulation instead.

**Effective and undrained shear moduli** \(G_{\text{eff}}\) and \(G_u\)

Four shear moduli are easily and unambiguously defined for the anisotropic system under study. Furthermore, since we are treating only soft anisotropy, all of these moduli are the same, i.e., \(G_i = G_{dr}\) for \(i = 1, \ldots, 4\). These are all related to the four shear eigenvectors of the systems, and they do not couple to the pore-fluid mechanics. But, the eigenvectors in the reduced \(2 \times 2\) system studied here are usually mixed in character, being quasi-compressional or quasi-shear modes. It is therefore somewhat problematic to find a proper definition for a fifth shear modulus. The author has analyzed this problem previously (Berryman, 2004b), and concluded that a sensible (though approximate) definition can be made using \(G_5 = G_{\text{eff}}\). There are several different ways of arriving at the same result, but for the present analysis the most useful of these is to express \(G_{\text{eff}}\) in terms of the product \(\Lambda_+ \Lambda_-\) (the eigenvalue product, which is also the determinant of the \(2 \times 2\) compliance system). The result, which will be quoted here without further discussion [see Berryman (2004b) for details], is

\[
\frac{1}{3K_u} \cdot \frac{1}{2G_{\text{eff}}} \equiv \Lambda_+ \Lambda_- = 18 \left[ A_{11}^* A_{33}^* - (A_{13}^*)^2 \right],
\] (31)

which we take as the definition of \(G_{\text{eff}}\) here. And, since \(A_{11}^* = 1/9K_u\), we have

\[
\frac{1}{G_{\text{eff}}} = 12 \left[ A_{33}^* - (A_{13}^*)^2 / A_{11}^* \right].
\] (32)

To obtain an isotropic average overall undrained shear modulus, we next take the arithmetic mean of these five shear compliances:

\[
\frac{1}{G_u} = \frac{1}{5} \sum_{i=1}^{5} \frac{1}{G_i}.
\] (33)

Combining these definitions and results gives:

\[
\frac{1}{G_u} - \frac{1}{G_{dr}} = - \frac{4}{15} \frac{(\beta_i^* - \beta_j^*)^2}{1 - \alpha B} \frac{\alpha B}{K_{dr}} = \frac{4}{15} \frac{(\beta_i^* - \beta_j^*)^2}{1 - \alpha B} \left[ \frac{1}{K_u} - \frac{1}{K_{dr}} \right],
\] (34)
where the $\beta'$s are defined by $\beta'_i = \beta_i K_{dr}/\alpha$. The final equality is presented to emphasize the similarity of the present results to those of both Mavko and Jizba (1991) and Berryman et al. (2002b). Setting $\beta'_1 = 0$, $\beta'_3 = 1$, $B = 1$, and $\alpha \approx 0$ recovers the form of Mavko and Jizba (1991) for the case of a very dilute system of flat cracks.

**Table.** Elastic and poroelastic parameters of the three rock samples considered in the text. Bulk and shear moduli of the grains $K_m$ and $G_m$, bulk and shear moduli of the drained porous frame $K_{dr}$ and $G_{dr}$, the effective and undrained shear moduli $G_{eff}$ and $G_u$, and the Biot-Willis parameter $\alpha = 1 - K_{dr}/K_m$. The porosity is $\phi$.

<table>
<thead>
<tr>
<th>Elastic/Poroelastic Parameters</th>
<th>Sierra White Granite</th>
<th>Schuler-Cotton Valley Sandstone</th>
<th>Spirit River Sandstone</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_m$ (GPa)</td>
<td>31.7</td>
<td>36.7</td>
<td>69.0</td>
</tr>
<tr>
<td>$G_u$ (GPa)</td>
<td>28.3</td>
<td>17.7</td>
<td>12.4</td>
</tr>
<tr>
<td>$G_{dr}$ (GPa)</td>
<td>26.4</td>
<td>15.7</td>
<td>11.33</td>
</tr>
<tr>
<td>$G_{eff}$ (GPa)</td>
<td>39.8</td>
<td>35.8</td>
<td>20.11</td>
</tr>
<tr>
<td>$K_m$ (GPa)</td>
<td>57.7</td>
<td>41.8</td>
<td>30.0</td>
</tr>
<tr>
<td>$K_{dr}$ (GPa)</td>
<td>38.3</td>
<td>13.1</td>
<td>7.04</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.336</td>
<td>0.687</td>
<td>0.765</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.008</td>
<td>0.033</td>
<td>0.052</td>
</tr>
</tbody>
</table>

**EXAMPLES AND DISCUSSION**

It is clear from (25) that fluid effects in $\Delta$ cannot increase the overall compliance eigenvalues simultaneously for both the quasi-bulk and the quasi-shear modes. Rather, if one increases, the other must decrease. Furthermore, it is certainly always true that the presence of pore liquid either has no mechanical effect or else strengthens (i.e., stiffens) the porous medium in compression. But this effect on the bulk modulus has been at least partially accounted for in $A^{*}_{11} = 1/K^{*}$ through the original contribution derived by Gassmann (1951). So presumably the contribution of $\Delta$ to compliance cannot be so large as to negate completely the liquid effects on the undrained bulk modulus.

**Examples**

To clarify the situation, we show some examples in Figures 1–8. The details of the analysis that produces these figures are summarized in the Appendix. The main point is that, for the compliance version of the analysis, the contours of constant energy are ellipses when the vector $f$ in (18) is interpreted as a stress. Analogously, when the vector is treated as a strain, the contours of constant energy are ellipses for the dual (or stiffness) formulation. If we choose to think of these figures as diagrams in the complex plane, then we note that — while circles and lines transform to circles and lines when transforming back and forth between these two planes — the shapes of ellipses are not preserved (except, of course, in the special case —
Figure 1: For a glassy porous material having bulk modulus $K_{dr} = 18.52$ GPa and shear modulus $G_{dr} = 13.89$ GPa, the locus of points $z = Re^{i\theta}$ — see equation (36) — having constant energy $U = 900$ GPa, when the linear combination of pure compression and pure uniaxial shear is interpreted as strain field applied to the stiffness matrix (solid black line). The plot is in the complex $z$-plane, with the inverse of the corresponding expression for the compliance energy superposed for comparison (dashed blue line). Red circles at the two points of intersection correspond to the two eigenvectors of the system of equations. The ellipse (solid black line) in this plane corresponds to the more complex curve in Figure 2.

which is precisely that of isotropy – when the ellipses degenerate to circles). Eigenvectors are determined by the directions in which the points of contact of these two curves lie (indicated by red circles).

Figures 1 and 2 present an example based on a glassy material. Typical values for the bulk and shear moduli of glass were used: $K_m = 46.3$ GPa and $G_m = 30.5$ GPa, respectively. The value of the Biot-Willis coefficient was arbitrarily chosen as $\alpha = 0.6$, so $K_{dr} = 18.52$ GPa. Taking Poisson’s ratio as $\nu_{dr} = 0.2$, we have $G_{dr} = 13.89$ GPa. Skempton’s coefficient was chosen for simplicity to be $B = 1$ in this and all the other examples as well. (This choice is extreme because it implies that $K_u = K_m$. But, since our interest here is in analysis of the undrained shear modulus, the study of this limit is particularly useful to us.) The most anisotropic choices of $\beta_1$ and $\beta_3$ were used that would not produce absurd (negative) values of the diagonal coefficients for either $s_{ij}^*$ or $c_{ij}^*$, and that also would not produce $G_u > G_m$. [The bound determined by (32) and (33) is a type of upper bound – actually the Voigt average. Values of this bound that might exceed $G_m$ need not be considered.] For glass, these values were found to be $\beta_1 = 0.15\alpha/K_{dr}$ and $\beta_3 = 0.70\alpha/K_{dr}$. The value of the energy used for normalization was $U = 900.0$ GPa. Computed values for the effective and undrained shear moduli were $G_{eff} = 25.43$ GPa and $G_u = 15.28$ GPa.
Figure 2: Same parameters as Figure 1, but the linear combination of pure compression and pure uniaxial shear is interpreted as a stress field and is applied to the compliance matrix (dashed blue line). The plot is again in the complex $\mathbb{z}$-plane, with the inverse of the corresponding expression for the stiffness energy superposed for comparison (solid black line). Red circles at the two points of intersection correspond to the two eigenvectors of the system of equations. The ellipse (dashed blue line here) corresponds to the more complex curve in Figure 1.
For the remaining three sets of examples, the values used for the moduli of the samples are taken from results contained in Berryman (2004a), wherein it was shown how certain laboratory data could be fit using an elastic differential effective medium scheme. These results are summarized in the TABLE.

<table>
<thead>
<tr>
<th>Normalized Compressional Strain</th>
<th>Normalized Uniaxial Shear Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>0.1</td>
<td>0.4</td>
</tr>
</tbody>
</table>

**SIERRA WHITE GRANITE**

Figure 3: Same as Figure 1 for Sierra White Granite using the parameters from the TABLE.

Figures 3 and 4 present results for Sierra White granite. Laboratory data on this material were presented by Murphy (1982). The values chosen for $\beta_1$ and $\beta_3$ were $\beta_1 = 0.05\alpha/K_{dr}$ and $\beta_3 = 0.90\alpha/K_{dr}$. The value of the energy used for normalization was $U \simeq 900.0$ GPa. Computed values for the effective and undrained shear moduli were $G_{eff} = 39.8$ GPa and $G_u = 28.3$ GPa.

Figures 5 and 6 present results for Schuler-Cotton Valley sandstone. Laboratory data on this material were also presented by Murphy (1982). The values chosen for $\beta_1$ and $\beta_3$ were $\beta_1 = 0.20\alpha/K_{dr}$ and $\beta_3 = 0.60\alpha/K_{dr}$. The value of the energy used for normalization was $U \simeq 900.0$ GPa. Computed values for the effective and undrained shear moduli were $G_{eff} = 35.8$ GPa and $G_u = 17.7$ GPa.

Figures 7 and 8 present results for Spirit River sandstone. Laboratory data on this material were presented by Knight and Nolen-Hoeksema (1990). The values chosen for $\beta_1$ and $\beta_3$ were $\beta_1 = 0.25\alpha/K_{dr}$ and $\beta_3 = 0.50\alpha/K_{dr}$. The value of the energy used for normalization was $U \simeq 900.0$ GPa. Computed values for the effective and undrained shear moduli were $G_{eff} = 20.11$ GPa and $G_u = 12.41$ GPa.
Figure 4: Same as Figure 2 for Sierra White Granite using the parameters from the TABLE.
DISCUSSION

We can compare the results obtained with results obtained for the same rocks using differential effective medium theory to fit data. The two characteristics that will interest us here are: (1) comparisons between the values chosen in our examples for the anisotropic β′s and the best fitting crack aspect ratios found in Berryman (2004a), and (2) comparisons between the magnitudes of changes in the overall shear moduli from their drained to undrained values.

The preferred crack aspect ratios found for Sierra White granite, Schuler-Cotton Valley sandstone, and Spirit River sandstone in Berryman (2004a) were respectively, 0.005, 0.015, and 0.0125. Here we found that (β′₁,β′₂) for the same samples were, respectively, (0.05,0.90), (0.20,0.60), and (0.25,0.50). Clearly, these values are at least weakly correlated with those of the aspect ratios for the same samples, but no stronger conclusions can be reached at the present time concerning these values.

Similarly, the comparisons of the changes in shear modulus magnitude from drained to undrained also show a weak correlation. The increases in shear moduli observed in the measured laboratory data for Sierra White granite, Schuler-Cotton Valley sandstone, and Spirit River sandstone are, respectively, about 10%, 10%, and 20%. As seen in the TABLE, the magnitude of the changes predicted here is essentially about 10% in all three of these cases. Thus, agreement is good both qualitatively and semi-quantitatively in all cases. We conclude that the theory presented here is correctly predicting the magnitudes of these shear modulus enhancements due to pore-fluid effects.
Figure 6: Same as Figure 2 for Schuler-Cotton Valley Sandstone using the parameters from the TABLE.
SUMMARY AND CONCLUSIONS

The preceding discussion shows how overall shear modulus dependence on pore-fluid mechanics arises in simple anisotropic (the specific example used was transversely isotropic) media. The results demonstrate in an entirely elementary fashion how compression-to-shear coupling enters the analysis for anisotropic materials, and furthermore how this coupling leads to overall shear dependence on mechanics of fluids in the pore system.

These effects need not always be large. However, the effect can be very substantial (on the order of a 10% to 20% increase in the overall shear modulus) in cracked or fractured materials, when these pores are liquid-filled. The anisotropy and liquid stiffening effects then both come strongly into play in the results we see, such as those illustrated in Figures 1–8. In particular, if $\beta_1 \approx \beta_3$, then soft anisotropy does not make a significant contribution. But, if either $\beta_1 << \beta_3$ or $\beta_1 >> \beta_3$, then the contribution can be significant. For example, we might expect these cases to be relevant for systems with either vertical or horizontal fractures.

REFERENCES


Berge, P. A., and Berryman, J. G., 1995, Realizability of negative pore compressibility in
Figure 8: Same as Figure 2 for Spirit River Sandstone using the parameters from the TABLE.


**APPENDIX A**

The equation of an ellipse centered at the origin whose semi-major and semi-minor axes are of lengths $a$ and $b$ and whose angle of rotation with respect to the $x$-axis in the $(x, y)$-plane is $\psi$ is given by

$$(x \cos \psi + y \sin \psi)^2/a^2 + (-x \sin \psi + y \cos \psi)^2/b^2 = 1. \tag{A-1}$$

For comparison, when $x = r \cos \theta$, $y = r \sin \theta$, and a stress of magnitude $r = \sqrt{x^2 + y^2}$ is applied to a poroelastic system, the energy stored in the anisotropic media of interest here [using (16) and (18)] is given by

$$r^2 \vec{f}(\theta) \Sigma^* \vec{f}(\theta) \equiv U(r, \theta) = 3r^2 \left[ A_{11} \cos^2 \theta + 2 \sqrt{2} A_{13} \cos \theta \sin \theta + 2 A_{33} \sin^2 \theta \right] = R^2 U(r_0, \theta). \tag{A-2}$$

In the second equation $R \equiv r/r_0$, and $r_0$ in an arbitrary number (say unity) having the dimensions of stress (i.e., dimensions of Pa). It is not hard to see that, when $U(r, \theta) = \text{const}$, the two equations (A-1) and (A-2) have the same functional form and, therefore, that contours of constant energy in the complex ($z = x + iy$) plane are ellipses. Furthermore, we can solve for the parameters of the ellipse by setting $U = 1$ (in arbitrary units for now) in (A-2) and then factoring $r^2$ out of both equations. We find that

$$3A_{11} = \frac{\cos^2 \psi}{a^2} + \frac{\sin^2 \psi}{b^2},$$

$$6\sqrt{2} A_{13} = \sin 2\psi \left( \frac{1}{a^2} - \frac{1}{b^2} \right), \tag{A-3}$$

$$6A_{33} = \frac{\sin^2 \psi}{a^2} + \frac{\cos^2 \psi}{b^2}.$$
These three equations can be inverted for the parameters of the ellipse, giving:

\[
\frac{1}{a^2} = \frac{3A_{11}\cos^2\psi - 6A_{33}\sin^2\psi}{\cos 2\psi},
\]

\[
\frac{1}{b^2} = -\frac{3A_{11}\sin^2\psi - 6A_{33}\cos^2\psi}{\cos 2\psi},
\]

\[
\tan 2\psi = \frac{2\sqrt{2}A_{13}}{A_{11} - 2A_{33}}.
\]

Although contours of constant energy are of some interest, it is probably more useful to our intuition for the poroelastic application to think instead about contours associated with applied stresses and strains of unit magnitude, \textit{i.e.}, for \( r = 1 \) (in appropriate units) and \( \theta \) varying from 0 to \( \pi \) [again see definition (18)]. We then have the important function \( U(1, \theta) \). [Note that, when \( \theta \) varies instead between \( \pi \) and \( 2\pi \), we just get a copy of the behavior for \( \theta \) between 0 and \( \pi \). The only difference is that the stress and strain vectors have an overall minus sign relative to those on the other half-circle. For a linear system, such an overall phase factor of unit magnitude is irrelevant to the mechanics of the problem.] Then, if we set \( U(r, \theta) = const = R^2U(r_0, \theta) \) and plot \( z = Re^{i\theta} \) in the complex plane, we will have a plot of the ellipse of interest with \( R \) determined analytically by

\[
R = \sqrt{U(r, \theta)/U(r_0, \theta)} = \sqrt{const/U(r_0, \theta)}.
\]

We call \( R \) the magnitude of the normalized stress (\textit{i.e.}, normalized with respect to \( r_0 \)).

The analysis just outlined can then be repeated for the stiffness matrix and applied strain vectors. The mathematics is completely analogous to the case already discussed, so we will not repeat it here. Since strain is already a dimensionless quantity, the factor that plays the same role as \( r_0 \) above can in this case be chosen to be unity if desired, as the main purpose of the factor \( r_0 \) above was to keep track of the dimensions of the stress components.
Electric fields created at the seismic source - a new electroseismic phenomenon

Seth Haines

ABSTRACT
Electroseismic data may show two different forms of source-related energy. The first is the electroseismic direct field which is produced at any impact source. It is approximately an electric dipole created by the asymmetrical pressure field of the impact. The second field is the electric field created by the impact of a mass on a metal hammer plate. The impact moves the hammer plate within the earth’s magnetic field, and creates an electric field described by Lorentz’s equation. Both show no moveout, and the amplitude pattern of a dipole. The direct field may be differentiated from the Lorentz field, however, by its reversed polarity on opposite sides of the shot point.

INTRODUCTION
Electroseismic phenomena have been shown to produce two forms of energy: the interface response and the coseismic field (Butler et al., 1996; Garambois and Dietrichz, 2001; Haines and Guitton, 2002). As explained in detail by Pride and Haartsen (1996), these phenomena depend on the pressure-induced flow of pore fluid relative to the grain matrix as a P-wave passes through a fluid-saturated porous medium. The pore fluid carries with it a small amount of electric charge due to the electric double layer that exists at the grain/fluid boundary. This mechanism produces an electric charge separation within the P-wave. The associated electric field is termed the “coseismic field”. When the P-wave encounters an interface in mechanical or elastic properties, the charge distribution is disrupted and made asymmetrical; this results in what can be approximated as an oscillating disk of electric dipoles at the first Fresnel zone. The corresponding electric field is termed the “interface response” and can be observed at the Earth’s surface or other remote location. It shows reversed polarity on opposite sides of the shot point, virtually zero moveout ($V_{em} >> V_p$), and the amplitude pattern of a dipole.

A third type of electroseismic energy is predicted by Equation 144 of (Pride and Haartsen, 1996):

$$E_x = T_{E,sem}^F \left( \frac{i}{\omega s_{em} x} - \frac{1}{(\omega s_{em} x)^2} \right) \frac{e^{i\omega s_{em} x}}{4\pi x} 6 \cos \theta \sin \theta F$$  

1email: shaines@pangea.stanford.edu
where

\[ T_{E,se}^F = \frac{i \omega \mu_0 L}{G(s_{em}^2 - s_z^2)} \]  

(with original typographic error corrected). It was first reported in the literature by Haines et al. (2004). We refer to this form of electroseismic energy as the "direct field" because it is analogous to the seismic direct wave. The direct field is the electric field of a vertical electric dipole created at the location of an impact source (Figure 1). The impact source (a sledgehammer strike, for instance) creates an asymmetric fluid-pressure distribution (enhanced pressure beneath the source and decreased pressure above) which results in a similarly asymmetrical charge distribution. This charge distribution has a strong vertical dipole component, so the measured field shows reversed polarity on opposite sides of the shot point and the amplitude pattern of a dipole. The dipole is localized at the shot point and begins at the time of the source impulse and continues until the earth has relaxed to its original state.

Electroseismic data may also demonstrate the existence of a similar-looking, but entirely unrelated, electric field. This is the field of a conductor moving with velocity \( v \) within the Earth’s magnetic field \( \mathbf{B} \), described by Lorentz’s equation:

\[ \mathbf{E} = \mathbf{v} \times \mathbf{B}. \]  

If a metal hammer plate is in good electrical contact with the soil then the electric field resulting from the hammer impact may be observed at nearby electrode dipoles during the time that the hammer plate is moving, typically not more than 10 ms. The field can often be discriminated from electroseismic fields by the fact that it shows non-reversed polarity on the two sides of the shot point.

In this contribution we provide a detailed analysis of these observations and further examples of them from recently-acquired field data.

**QUALITATIVE OBSERVATIONS**

In order to develop our understanding of the source-related electric fields that may be observed in electroseismic data, we begin with the simplest possible data collection scenario, and add complexity one step at a time. In this way we can better identify the impact of each individual element of electroseismic data collection.
Figure 2: Electroseismic shot gathers collected at the vineyard field site testing various source options. a) stack of manual triggers. Note absence of any coherent arrivals. b) Wooden fence post on plastic hammer plate. Note flat direct field arrivals and dipping coseismic energy. c) Metal sledgehammer on plastic hammer plate. Gather appears very similar to that in b). d) Metal sledge on aluminum hammer plate. Note addition of flat Lorentz field energy in the upper \( \sim 0.01 \) s.

The series of gathers shown in Figure 2 was collected at the vineyard field site described in detail by Haines and Guitton (2002). The site is a small meadow at a vineyard in St. Helena, CA. The soil is fairly homogeneous and clay-rich, and extends to a depth of at least 3m. Although the homogeneity has been disrupted by the construction of two sand-filled trenches, the data in Figure 2 were collected away from the trenches such that they should have no impact on the displayed data. The data were collected with a source point in the center of an array of 24 electrode pairs at a spacing of 0.7m. The distance across each pair of electrodes (the dipole width) is 1.05m. All gathers are the result of stacking individual impacts of the source (frequently a sledgehammer) on a metal or plastic hammer plate. Generally 25 or 30 impacts were recorded separately and then those that do not show any strong electrical noise in the time window of interest are stacked to produce the gathers shown. Final gathers are generally the result of stacking between 10 and 25 individual impacts.
The simplest possible data collection example is carried out by manually triggering the recording seismograph. Figure 2a shows data collected by hitting the trigger switch against a stationary object. Thus the data represent electrical background noise and the lack of any coherent energy demonstrates that the trigger mechanism produces no electrical noise. Next we add a level of complexity by putting seismic energy into the ground, but with no moving metal objects. Figure 2b shows data collected using the impact of a wooden source (a fence post) on a plastic hammer plate. We now see the expected dipping coseismic energy (with seismic moveout). We also observe flat (no moveout) energy in the upper ~17 ms of the record. This energy appears to show the amplitude pattern of a dipole and reversed polarity on opposite sides of the shot point. If the site geology included any shallow interfaces, we might conclude that this flat energy was the electroseismic interface response. However, it does not, so we interpret this energy as the electroseismic direct field. We will re-visit this interpretation in the next section. We add another level of complexity by using a metal sledgehammer on the plastic hammer plate (Figure 2c) and observe that the result is very similar to that of Figure...
2b. Thus we can conclude that the moving metal hammer head does not create a noticeable electric field. We move one step further by employing a metal hammer plate (an aluminum cylinder ~0.2m long and ~0.15m in diameter, positioned with its axis horizontal). We now observe (Figure 2d) an additional form of flat energy in the upper ~10 ms of the record. It shows no moveout, and an amplitude pattern suggestive of a dipole. But unlike the interface response and the direct field, this energy shows the same polarity on the two sides of the shot point. Thus we conclude that this energy is due to a horizontal electric dipole oriented along the electrode transect line. The Lorentz field (Equation 3) offers the most likely explanation for the observed energy. The motion (v) of the conductive hammer plate in the Earth’s magnetic field (B) produces an electric field E. We further examine this field later in this contribution.

In order to gain more certainty in our interpretations, we examine data from a separate field site. The data in Figure 3 were collected at the Thompson tree farm in the Santa Cruz mountains of California. The site is remote from cultural noise (both electrical and seismic) and has a subsurface geology that we consider to be free of any distinct interfaces in the upper few meters. Figure 3a shows a gather collected using a wooden source on a plastic hammer plate with dipping coseismic energy and (faint) flat direct field energy clearly visible. The gather in Figure 3b looks very similar, and was collected using a metal sledgehammer on the plastic hammer plate. Figure 3c shows data collected with a metal hammer on the aluminum hammer plate, and it shows the strong Lorentz electric field as well as the dipping coseismic energy. The gather in Figure 3d was also collected with the hammer on the aluminum hammer plate, but in this case the plate was insulated from the earth by a thin layer of wool material. The Lorentz field is not observed, demonstrating the need for electrical contact between the metal hammer plate and the earth for observation of this field. Thus a metal hammer plate may be used for electroseismic work if it is insulated from the earth.

A MORE DETAILED ANALYSIS

Electroseismic direct field

In order to characterize both the direct field and the Lorentz field, we conducted a series of experiments involving the recording of various sources with electrode pairs deployed in a circular geometry. The electrode array was spaced evenly around a circle of radius 4.3m in a homogeneous part of the vineyard meadow, with 12 electrode pairs oriented radially and 12 oriented tangentially. Thus tangential and radial pairs of electrodes were co-located at 30 degree intervals around the circle. The source point is in the center of the circle for all shot gathers. Figure 4a shows the radial traces of a shot gather collected with a sledgehammer striking the plastic plate. Based on arrival times from the data shown in Figure 2, we can interpret the strong coherent arrival at 0.02 seconds as the coseismic energy, and the weaker arrival at 0.01 seconds as the direct field. These arrivals do not appear in the tangential part of the same shot gather (Figure 4b), as is to be expected for a vertical dipole (the direct field) and radially propagating seismic energy (the coseismic arrival). Further confirmation of our interpretation of the 0.02 second arrival as coseismic energy is provided by the corresponding radial horizontal geophone data shown in Figure 4c, where we see that the first seismic arrival
closely matches the interpreted coseismic arrival in Figure 4b. The lack of any energy at \sim 0.01 seconds in Figure 4c supports our interpretation of the 0.01 second energy in Figure 4a as the direct field, or at least as an electroseismic arrival. As is to be expected, the tangential geophone data (Figure 4d) do not show coherent arrivals.

Figure 4: Stacked shot gathers recorded by the circular electrode array described in the text. a) Radial component of electric field created by the impact of a sledgehammer on a plastic hammer plate. Note coherent direct field arrival at \sim 0.01 sec and coseismic arrival at \sim 0.02 sec. b) Tangential component of the same shot gather (hammer on plastic plate) with no coherent arrivals. c) Radial component of horizontal geophone data, with first seismic arrival at \sim 0.02 sec, corresponding with coseismic arrival in a). d) Tangential component of geophone data, with no clear coherent arrivals. [shaines1-circle1]

The fact that the direct field energy in Figure 4a shows (approximately) constant amplitudes around the circle is consistent with the interpretation that this energy is due to a vertical dipole. The deviations from constant amplitude are likely caused by imperfect electrode coupling. We can constrain the size and location of this dipole by considering the amplitude pattern of the in-line shot record shown in Figure 2c. Figure 5a shows the same data, but with a lower-frequency bandpass filter (so as to better represent the full direct field, which can be
clearly seen as a single strong arrival at ~0.005 seconds). The amplitude of the maximum of this arrival is plotted as dots in Figure 5b. Using the equation for the amplitude pattern of a dipole

\[ V(x,z) = \frac{qd}{4\pi\epsilon_0} \frac{z}{(x^2 + y^2)^{3/2}}, \]  

(where \( q \) is the charge and \( d \) is the separation between poles) we model amplitudes corresponding with a disk of dipoles with radius 0.8m and located at a distance of 0.8m from the electrode receiver line, plotted as a solid line in Figure 5b. The fit of this model to the data broadly indicates that the direct field is produced within a volume of earth of radius ~0.8m. The absolute magnitude of the modeled dipole is entirely arbitrary; it is simply scaled to match the real data. The numerous variables that contribute to the real magnitude are too complex to permit exact modeling.

Figure 5: Direct field amplitudes. (a) Shot gather from Figure 2c, shown with bandpass filter 20 to 800 Hz. (b) Amplitude pattern for the direct field arrival at ~0.005 seconds of gather in (a) plotted as dots, and modeled amplitudes as a solid line. Modeled amplitudes correspond with a disk of electric dipoles of radius 0.8m centered at 0.8m away from the electrode array.

Electric field of the metal hammer plate

One important observation about the Lorentz field is that its polarity reverses between sequential hammer strikes, such that approximately half of the raw hammer gathers show one polarity, and the other half show the opposite polarity. The stacks shown in Figures 2d and 3c are made from shot gathers selected on the basis of the polarity of the Lorentz field. The other
gathers would produce a stack with a Lorentz field arrival with opposite polarity. A stack of all of the shot gathers would show very little Lorentz energy as it tends to stack out.

Data collected by the circular electrode array using the sledgehammer on the aluminum hammer plate are shown in Figure 6. Figure 6a and b are the radial and tangential parts of a partial stack of hammer strikes, and 6c and d are the radial and tangential parts of a stack of the other hammer strikes. These two sets of impacts were selected from the individual hammer strike gathers based on the presence and polarity of the events that appear at a time of ~0.001 seconds at certain radial positions (90-180° and 240-330°) in the radial component and 90° out of phase (0-90° and 180-270°) in the tangential component. Note that the polarity of these arrivals is reversed between the two stacks (Figure 6a and b, versus Figure 6c and d).

The radial pattern of these arrivals suggests that they are due to a horizontal electric dipole oriented a few tens of degrees west of north, with the orientation of the dipole reversed between the two sets of gathers. Because the field occurs only for hammer impacts on a metal plate, we assume that it is caused by the metal plate, and that it is the Lorentz field (Equation 3). Because the orientation of the dipole reverses phase between sequential hammer impacts, we must assume that it is caused by a component of $\mathbf{v} \times \mathbf{B}$ that can reverse from one strike to the next. The earth’s magnetic field $\mathbf{B}$ is essentially constant (oriented toward magnetic north, and inclined at an angle of ~60° from horizontal), so we must look to $\mathbf{v}$ for this reversal. Although the dominant component of $\mathbf{v}$ is vertical, there is also a small horizontal component due to the imperfect impact of the hammer on the rounded top of the aluminum block. For the case of the in-line data (Figure 2d and Figure 3c), the aluminum cylinder is oriented along the electrode receiver line, and thus the hammer strikes will tend to cause horizontal motion perpendicular to the line. If we take the cross product of this velocity with the vertical component of $\mathbf{B}$, we get a horizontal electric field $\mathbf{E}$ oriented along the electrode line, just as we observe. The orientation of the horizontal component of $\mathbf{v}$ will vary from strike to strike, but will generally be perpendicular to the electrode transect line, in one of two primary polarities. We conclude that the observed electric field is due to the horizontal component of the hammer plate velocity crossed with the vertical component of the earth’s magnetic field.

Next we extract the amplitudes of the observed arrivals and compare them with modeled amplitudes. Figure 7a shows the amplitude of the Lorentz field event shown in Figure 6a and b, while Figure 7b shows the amplitude of the Lorentz in Figure 6c and d. The radial component is plotted as a solid line and the tangential component as a dashed line. The amplitude in Figure 7a corresponds with the third of the three phases of the Lorentz field arrival in Figure 6a (0.0042 to 0.0065 seconds) while the amplitude in Figure 7b was extracted from the second of the three main phases of the Lorentz event in Figure 6b (0.0025 to 0.0045 seconds), thus the two amplitude patterns are in-phase while the two displayed Lorentz events are 180° out-of-phase. We use Equation (4) to model a horizontal dipole at the source point, and find that a best fit is achieved with a dipole oriented ~50° west of north. This alignment corresponds with the alignment of the hammer plate and the person swinging the hammer, not with magnetic north, confirming our interpretation that the horizontal $\mathbf{v}$ of the hammer plate and the vertical component of the earth’s field $\mathbf{B}$ are responsible for the Lorentz field. The horizontal component of $\mathbf{B}$ does not seem to play a role in the creation of this field.
Figure 6: Electroseismic data collected with a sledgehammer on the aluminum hammer plate and recorded by the circular electrode array. Lorentz energy can be seen at \( \sim 0.001 \) s in each gather. 

a) Radial component of stack of selected hammer strikes. Note Lorentz energy at \( \sim 0.001 \) s at certain radial positions. 

b) Tangential component of the same stacked gathers, with faint Lorentz energy at \( \sim 0.001 \) s at positions orthogonal to the energy in a). 

c) Radial component of a stack of other shot gathers, with similar energy at \( \sim 0.001 \) s but with reversed polarity relative to a). 

d) Tangential component of stack of the same gathers as c), again showing Lorentz energy at positions orthogonal to the energy in c).
Figure 7: Comparison between real and modeled Lorentz energy as a function of radial position. a) Amplitude (solid=radial, dashed=tangential) patterns extracted from the stacks in Figure 6a and b. b) Amplitudes extracted from stacks in Figure 6c and d. c) Amplitude pattern modeled for a dipole oriented 50° west of north.

We can gain further knowledge about the Lorentz field by extracting amplitudes from the in-line data (Figure 2d). Figure 8a and c show stacks of two different sets of impacts collected with the metal hammer plate processed with a broad bandpass filter (20 to 800 Hz); the data in Figure 8a is the same stack as in Figure 2d. Both of these data plots show a strong flat arrival at about 0.002 seconds which we interpret as the Lorentz field, followed by another flat event with reversed polarity on opposite sides of the shot point. This second arrival is the direct field. Amplitudes extracted from these stacks for the Lorentz field are show in Figure 8b and d as dots. Modeled amplitudes for a horizontal electric dipole matching the hammer plate (charge separation of 0.2m between ends of the dipole, lateral offset of 0.25m from the receiver line) are plotted as solid lines. Only the magnitude and polarity of the modeled dipole is varied between the two plots. The central two traces show polarity opposite that of the rest of the Lorentz field because they are located along the horizontal dipole and so are measuring the field off of its main axis, where the field is opposite to the direction of the dipole.

DISCUSSION

We have identified and described two different forms of electrical energy that may be created by the standard hammer-on-metal-plate seismic source option. The motion of the metal plate itself creates an electric field, which we refer to as the Lorentz field, due to the physics implied by Equation (3). The second effect is electrokinetic in origin, and is termed the “direct field”. It is the field of the electrical charge separation caused by the asymmetrical pressure gradient created at an impact source (such as a sledgehammer).
Figure 8: In-line data showing the Lorentz field at $\sim 0.001$ s, the direct field at $\sim 0.005$ s, and modeled and real amplitudes. a) Stacked data shown in Figure 2d, but with lower-frequency bandpass filter (20 to 800 Hz). b) Real amplitudes (dots) of Lorentz field arrival of data in a), and modeled amplitudes (solid line) corresponding with the metal hammer plate acting as an electric dipole oriented along the electrode array. c) Stack of other selected shot gathers (those with Lorentz field arrival opposite in polarity to stack in part a). d) Real (dots) and modeled (solid line) amplitudes for metal plate as a horizontal dipole as predicted by the Lorentz equation.
The Lorentz field is unlikely to prove useful, and so is a noise source to be avoided. Fortunately this can easily be achieved with the use of a non-metal hammer plate, or insulation between a metal plate and the earth.

The direct field could potentially be used to measure physical properties of the region where it exists (immediately around the source). Though not terribly interesting at the surface, measurement of the direct field in a down-hole setting could prove useful.

ACKNOWLEDGMENTS

We are very grateful to Jim and Carolyn Pride for the use of their land (the vineyard site) and for their ongoing generosity. We are also very grateful to George and Anita Thompson for the use of their property, and to George for his continued support and general guidance. Steve Pride has played an integral role throughout this project, and made important contributions to this paper. Biondo Biondi, Simon Klemperer, Jerry Harris, and Art Thompson have all made major contributions to this research project as a whole, and have provided valuable guidance throughout. The data presented in this paper were made possible by the sweat of those who swung the hammer: Mike Beman, Morgan Brown and Phil Resor. We also greatly appreciate the efforts of the many others who have swung the hammer as part of other data collection efforts. Financial support has been provided by The American Chemical Society’s Petroleum Research Fund, an Achievement Rewards for College Scientists fellowship, the Stanford Exploration Project, the American Association of Petroleum Geologists Grants-in-Aid program, the Geological Society of America’s graduate student fellowship program, the Colorado Geological Survey, and the Stanford Earth Sciences McGee Fund.

REFERENCES


Short Note

Parallel datasets in SEPlib

Robert G. Clapp

Cluster computing, with relatively inexpensive computational units, looks to be the most cost effective high performance computing option for at least the next five years. Unfortunately, cluster computing is not the most convenient processing environment for the type of problems routinely performed in seismic exploration.

Many of the tasks we want to perform require both large memory and massive disk space. These requirements make grid based computing (Foster and Kesselman, 1998) generally impractical. Storing and accessing these large datasets is also problematic. The Parallel Virtual File System (PVFS) (Ross et al., 2004) is an attempt to create a virtual file system across several machines. It is still relatively new, and doesn’t allow good user control of data locality (e.g. you often know you want to store frequency ‘x’ on node one because that is where you plan to process it).

When writing an application for a cluster you often face difficult additional challenges. Problems specific to coding in SEPlib, but which are indicative of more wide-spread problems include: each process may not be able to read a shared a history file; each process needs a unique tag descriptor when writing out a temporary file; all shared binary files need to be read locally and then distributed (through some mechanism such as MPI (Snir et al., 1995)); and writing effective check-pointing in a parallel code can be extremely cumbersome. These problems significantly increase the complexity or writing and debugging applications.

In this paper I discuss an extension of the basic SEP data description to help solve the above problems. It extends the definition of a SEPlib dataset to include one that is stored on multiple machines. It allows the same code base to be used for both serial and parallel applications. All that changes is which libraries you link with. Parallel datasets can be easily created and accessed, greatly simplifying coding and debugging in a cluster environment. In this paper I describe how the library works, and provide several examples of its use.

DESIGN

The extension of SEPlib to handle parallel datasets has many unintended similarities to the parallel data capability of HDF (NCSA, 2004). A parallel dataset is a composition of several datasets with a master file describing the relationship between the various parts. All programming with the parallel dataset library is done through the superset library (Clapp, 2003). To

1email: bob@sep.stanford.edu
handle parallel datasets with minimal changes to coding style requires a more abstract definition for IO than simple SEPlib or SEP3d allows. The library has both C and Fortran90 interfaces which allows almost any SEPlib code to be changed quickly. In this section I will describe how to interact how the library interacts with the various sub-datasets.

Files

Similar to SEP3D (Biondi et al., 1996), the master file is still the standard SEP history file. The difference is that the history file will not have a pointer to the dataset (in=) and it will have an additional pointer to a text file (Distribution Format File (DFF)) describing the distribution pattern (dff=).

The distribution format file is made up of several parameters:

- **nsect**  The number of sections the dataset is broken up into.
- **axis**  The axis along which the data is distributed.
- **naxis_convert**  The length of the axis that the data is distributed along.
- **pattern**  The distribution pattern along the distributed axis. Right now two distribution patterns are defined, BLOCK and SPREAD. In the first option sequential elements along the first axis will belong to the same section. In the second option a given section will contain elements $i\text{sect} + k*n\text{sect}$ where $i\text{sect}$ is the section number, $n\text{sect}$ is the number of sections, and $k$ is the number of axis elements the section owns. One obvious future extension to the library is to handle more arbitrary distribution patterns.
- **tag_secti**  The tag associated with section number $i$.
- **mach_secti**  The machine in which section $i$ was written.

The final set of files are the various tag_sect. Each tag_sect is a valid SEP3D dataset.

Initialization

There are three mechanisms to initialize a new distributed dataset. A distributed dataset can be created by a call to the library. The call defines which axis to distribute, how many sections to create, and what distribution pattern to use. The library then determines the number of threads it is running on and then distributes the sections to the various threads in a round-robin fashion. Normally you would think of having a single section for each thread, but this is not required, and in certain situations might not be desirable.

A distributed dataset can be read from a file. When initializing a distributed dataset from a tag, data locality tries to be preserved. The various sections are assigned to threads on the same machine as much as possible without sacrificing load balance. For example, if a dataset
has been broken up into six sections, three on machine A and three on machine B, and you then start a job on machine A, B, and C. A and B will each be assigned two sections that were written locally, while C be assigned one section from both the original A and B group. Whenever a dataset is initialized from a tag, the tag is only read by the master node. Its contents are then sent to all slave processes. This keeps every process from having to have access, and to simultaneously read, a single file. When scaling to a large number of processors, this ability can be useful.

The final method to create a distributed dataset is from another distributed dataset. If you are running on multiple threads and you initialize a new dataset from an already distributed dataset, the new dataset will inherit the original dataset’s distribution pattern.

I/O

The library operates in three different modes when reading and writing a dataset. When reading a non-distributed dataset in a parallel environment, the dataset will be by default read by the master thread and sent to all slave threads. You can also define the read as being exclusively local (imagine a local temporary file). When reading a distributed dataset, the read will just return the portion of the given read request that it locally owns. This requires you always to check the amount of data that is returned. In the next section I will provide an example.

When writing a distributed dataset, the library will by default assume that you are writing only the portion of the given window that it actually owns. When writing a non-distributed dataset you can again specify that a write is exclusively local, but otherwise only the master thread will write data, and it is the programmer’s job to make sure they have collected the dataset.

Distributing and collecting

Distributing a dataset is rather trivial. You simply have the master thread read the dataset locally and then use the distribute dataset routine to partition the dataset to the various threads. Collecting operates in a similar manner. The various threads read the data locally and then use a collect routine to combine the various subsections. The library also offers the ability to compress dataset along an axis. An example of this would be migration, where each thread will have its own image and the master thread wants the sum of the various portions.

Parameter handling

One of the biggest problems with running on cheap hardware is the failure rate is high. Writing migration code that is able to figure out where it died and continue is challenging. When the migration is part of a larger inversion problem the tasking becomes even more difficult. One of the goals of the library is to make check-pointing easier. Any thread can write a status parameter to a distributed tag. This status parameter is written to its local sections rather than
the global tag, so clobbering of the text file isn’t an issue. Restarting becomes a much simpler
matter. You can request the status parameter from each section with a single call. Figuring out
what portion of job has finished, and what portion is remaining becomes a trivial matter.

**Global parameters**

The library offers two options that are available to any distributed dataset. The first option, 
`master_data`, allows you to switch back and forth between a master-slave and master among
equals programming paradigm. If `out.master_data=1`, the default, the master thread will own
section(s) of the distributed tag `out`, otherwise it will not. Running in a true master-slave mode
can be useful when your data is for a file server which you do not want to do computations on.

The second option is to guarantee that each section of dataset will have a unique name.
By default all sections are given a name based on the output tag name (when accessible²),
otherwise a random, unique name is given. A unique name can be guaranteed by specifying `out.unique=1`, where `out` is the tag, on the command line. A related option is function call, `sep3d_make_local_name` which will create a tag on a local machine with an unique name.

**Extendability**

The library was written with the assumption that current parallel programming paradigms are
likely to change. All of the MPI specific code is limited to two files. Adding support for Parallel
Virtual Machine (PVM) or the next ‘latest and greatest’ parallel programming library simply
involves defining how to transfer data between processes and how to synchronize their flow.

**WRITING APPLICATIONS**

The best way to learn how to write code is to look at examples. Rather than just make list of
function calls, I will take portions from several different codes.

**Serial**

The first example is from the SEPlib program `Nmo3d`. In SEPlib there is a module that helps
initialize the parallel environment. Our only `ifdef` will be in whether to include the MPI
version of the parallel environment or a dummy serial version. We also need to initialize our
parallel environment. The optional arguments signify the input and output to the program. If
provided to the `sep_par_start` routine, `intag` will be set to "in" for a serial run and "intag" for a parallel run. The variable `outtag` will be set to "out" or "outtag" and will be initialized
with a copy of the `intag` history file.

²There is no standard way in unix to find the name of a file when it isn’t written the local directory
#ifdef SEP_MPI
  use sep_par_mpi
#else
  use sep_par
#endif
  call sep_par_start(intag=intag,outtag=outtag)

Initializing the output dataset is unaffected. If we are running in a parallel environment and the input is sectioned, the output tag will be sectioned. Otherwise a standard SEPlib file will be created.

    call init_sep3d(in,out,"OUTPUT",ctag=outtag) !Initialize SEP

It is often useful to know the thread number for the current process. The below code section bases its decision on how often to print progress on whether it is the master or slave process.

    if(sep_thread_num()==0) then
      call from_param("pct_master",pct_print,2.)
    else
      call from_param("pct_slave",pct_print,10.)
    end if

For NMO our data can be gridded in an arbitrary style. We simply need to loop through all of the data. The init_loop_calc function figures out looping parameters based on the dimensionality of the data in%ndims, the dimensions of the data in%n, and the maximum amount of data we want to read ntr*in%n(1)*in%n(2). The identifier "MAIN" just gives a name for the current loop. The do_sep_loop will increment windowing parameters fwind and nwind until we have looped through entire dataset, at which point the return value will change.

    if(0/=init_loop_calc(in%ndims,in%n,"MAIN",ntr*in%n(1)*in%n(2))) &
      call seperr("trouble initing loop")
    do while(0==do_sep_loop("MAIN",nwind,fwind))

To read the data (whether it is a regular or irregular cube) we first request the number of traces that fall in our current window. In this case, the library will return nh, the number of traces that the current thread locally owns in the given window.

    call sep3d_grab_headers(intag,in,nh,fwind=fwind(2:),nwind=nwind(2:))

If the thread owns any of data in the window, read it.

    if (nh.ne.0) then
      if (.not. sep3d_read_data(intag,in,input(:,:nh))) then
        call erexit("trouble reading data")
      end if
For a regular dataset, which elements of the current window are owned by the current thread is important. We can request to know the coordinates of the traces. The returned 2-D array is of dimensions \((ndim - 1) \times nh\), where \(ndim\) is the dimensions of the data and \(nh\) is the number of traces in the current block.

```fortran
if(.not. sep3d_grab_coords(in,coords)) &
call sepperr("trouble grabbing coords")
```

Writing is again automatic. We will automatically write just the local portion of our data.

```fortran
if(.not. sep3d_write_data(outtag,out,output,nwind=nwind,fwind=fwind)) then
call sepperr("trouble writing out ")
end if
```

The only remaining step we need to do is to make sure that the total number of traces is correctly calculated and make any calls needed to end the current parallel environment. The library keeps track of how many traces it has written to a given file. The `sep3d_update_ntraces` call tells the master thread how many traces each slave thread has written. The total number is then written to the output history (and possibly header format file).

```fortran
if(.not. sep3d_update_ntraces(out)) call sepperr("trouble updating ntraces")
call sep3d_rite_num_traces(outtag,out)
call sep_par_stop()
```

**Distributing a dataset**

The second example is taken from the SEPlib program `Transf`. It does a Fourier transform of a dataset and then distributes the data along the frequency axis. Again, only the portions relevant (and different from the previous example) will be shown.

In this case we want to know how many threads we are running on.

```fortran
nmpi=sep_num_thread()
```

By default, when converting to frequency `inv=.false.`, we are going to create a dataset with \(nmpi\) sections, but we will allow the user to override this option. We are going to section the dataset along the last axis \((ndimc)\) which in this case is the frequency axis.

```fortran
nsect=1; if(.not. inv) nsect=nmpi
call from_param("nsect",nsect,nsect)
if(nsect>1) then
  if(.not. inv) then
    if(.not. sep3d_section_tag(output,ndimc,"BLOCK",nsect)) &
      call sepperr("trouble sectioning tag")
  endif
endif
```
Our input and our output could involve a sectioned dataset. Given our current window parameters (nc,fc,jc for output and nr,fr, and jr for our input), we need to know the size of the data with which we will be working. The \texttt{sep\_local\_buffer\_size} gives the dimensions (number of samples by number of traces) of the local buffer size given the current window for a given dataset.

\begin{verbatim}
if(.not. sep3d\_local\_buffer\_size(output,n_c,nc,fc,jc) .or. &
   .not. sep3d\_local\_buffer\_size(input,n_r,nr,fr,jr)) then
   call seperr("trouble getting local buffer sizes")
end if
\end{verbatim}

Once we have transformed the data, we need to pass it to the appropriate thread. We first describe the window that we want to pass \texttt{sep\_3d\_set\_window} and then distribute the data.

\begin{verbatim}
call sep3d\_set\_window(space,nwind=nc,fwind=fc)
   if(.not. sep3d\_transfer\_data(input,space,blockc\_in,blockc\_out)) &
      call seperr("trouble transfering data")
\end{verbatim}

Everything else that is different from how you would conventionally program is the same as the first example.

\section*{Restarting and combining}

The following examples are taken from the SEPlib WEI library (Sava and Clapp, 2002). They demonstrate how to read a dataset, recognize which elements of the dataset are locally owned, and write status parameter and how to combine a dataset.

This first code portion is executed every time a frequency has been fully migrated. The frequency number currently being processed \texttt{i\_w} are looped through. A variable named \texttt{ifreq\_doneX}, where ‘X’ is the frequency number is created. Then that frequency is set to to 1 by the \texttt{sep\_3d\_set\_sect\_param}. In this case the image sep file (the \texttt{sep\_3d} structure \texttt{rsep}) is potentially a distributed dataset along an artificial axis corresponding to the number of processes that are computing the image. The \texttt{sep\_3d\_set\_sect\_param} call will automatically write this parameter to its local portion of the image.

\begin{verbatim}
do i=1,size(i\_w)
   par="ifreq\_done"
   call parcat(par,i\_w(i))
   if(adj) then
      if(.not. sep3d\_set\_sect\_param(rsep,par,1)) then
         write(0,*) "trouble writing sect\_params"
         return
      end if
end if
\end{verbatim}
The second example is from the portion of the code that checks for a restart request. Here, all of the threads are looped over. For each frequency, we check all of the sections of the image tag to see if any of the threads have processed this frequency.

```fortran
  do ifreq=1,wsep%n(6) !loop over frequencies
    par="ifreq_done"
    call parcat(par,ifreq)
    iw=-1;
    if(.not. sep3d_grab_sect_param(sep,par,iw)) then
      write(0,*), "trouble grabbing section parameters"
      return
    end if
    if(any(iw==1)) then
      Oftentimes we need to make a decision whether to proceed based on whether we own a specific element of the axis we spread along. This code fragment is marking whether or not it owns a given frequency.

  do i=1,wsep%n(6)
    iw_own(i)=sep3d_own(dsep,i)
  end do

  end

In the case of migration, we need to combine all the local images into a global image. This code fragment adds all of the local data (bigc) from the distributed image (big_sep), into the standard SEP dataset (small_sep) buffer bigc2.

  if(.not. sep3d_compress_data(big_sep,small_sep,bigc,bigc2)) then
    write(0,*), "trouble combining data ", fwind
    return
  end if

FUTURE WORK

There are several additional features that could be added to the current libraries’ capabilities. As mentioned earlier an arbitrary description about how the data is distributed along an axis would be useful. This would allow for better load balancing, redistributing of tasks on a restart, and for operations that perform some type of windowing.

The most obvious place for work is to add recognition of distributed datasets to more SEPlib programs. Currently only the generic SEPlib utilities Window3d and In3d are aware of distributed datasets. Adding awareness to program like Headermath would be helpful. More application programs could/should be converted. Currently Transf, Nmo3d, Phase, CAM, Rtm2D, and Fdmod are distributed dataset aware. Almost all of SEP’s current application programs could benefit from conversion.
CONCLUSIONS

Developing code for a parallel environment requires significant, sometimes non-obvious additional coding overhead. By expanding the concept of a SEPlib dataset to include a dataset that is spread over several machines, much of the difficulty associated with cluster computing can be avoided by the programmer. The included examples show how with relatively minor changes in coding style a parallel code can be created virtually free.

REFERENCES


SEP ARTICLES PUBLISHED OR IN PRESS


Guitton, A., 2004a, Amplitude and kinematic corrections of migrated images for non-unitary imaging operators: accepted for publication in Geophysics.


Sava, P., Biondi, B., and Etgen, J., 2004b, Diffraction-focusing migration velocity analysis with application to seismic and GPR data: submitted for publication in Geophysics.


SEP PHONE DIRECTORY

<table>
<thead>
<tr>
<th>Name</th>
<th>Phone</th>
<th>Login Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alvarez, Gabriel</td>
<td>724-0461</td>
<td>gabriel</td>
</tr>
<tr>
<td>Artman, Brad</td>
<td>723-6007</td>
<td>brad</td>
</tr>
<tr>
<td>Berryman, James</td>
<td>723-4417</td>
<td>berryman</td>
</tr>
<tr>
<td>Biondi, Biondo</td>
<td>723-1319</td>
<td>biondo</td>
</tr>
<tr>
<td>Brown, Morgan</td>
<td>723-6006</td>
<td>morgan</td>
</tr>
<tr>
<td>Claerbout, Jon</td>
<td>723-3717</td>
<td>jon</td>
</tr>
<tr>
<td>Clapp, Marie</td>
<td>725-1334</td>
<td>marie</td>
</tr>
<tr>
<td>Clapp, Robert</td>
<td>723-4417</td>
<td>bob</td>
</tr>
<tr>
<td>Curry, William</td>
<td>723-5911</td>
<td>bill</td>
</tr>
<tr>
<td>Guitton, Antoine</td>
<td>723-6007</td>
<td>antoine</td>
</tr>
<tr>
<td>Lau, Diane</td>
<td>723-1703</td>
<td>diane</td>
</tr>
<tr>
<td>Lomask, Jesse</td>
<td>724-0461</td>
<td>lomask</td>
</tr>
<tr>
<td>Rosales, Daniel</td>
<td>725-1334</td>
<td>daniel</td>
</tr>
<tr>
<td>Sava, Paul</td>
<td>723-0463</td>
<td>paul</td>
</tr>
<tr>
<td>Shan, Guojian</td>
<td>723-0463</td>
<td>shan</td>
</tr>
<tr>
<td>Shragge, Jeff</td>
<td>724-0461</td>
<td>jeff</td>
</tr>
<tr>
<td>Tisserant, Thomas</td>
<td>723-1250</td>
<td>thomas</td>
</tr>
<tr>
<td>Valenciano, Alejandro</td>
<td>723-6006</td>
<td>valencia</td>
</tr>
<tr>
<td>Vlad, Ioan</td>
<td>723-1250</td>
<td>nick</td>
</tr>
<tr>
<td>Wilson, Charlie</td>
<td>725-2234</td>
<td>wilsonck</td>
</tr>
</tbody>
</table>

SEP fax number:  (650) 723-0683

E-MAIL

Our Internet address is “sep.stanford.edu”; i.e., send Jon electronic mail with the address “jon@sep.stanford.edu”.

WORLD-WIDE WEB SERVER INFORMATION

Sponsors who have provided us with their domain names are not prompted for a password when they access from work. If you are a sponsor, and would like to access our restricted area away from work, visit our website and attempt to download the material. You will then fill out a form, and we will send the username/password to your e-mail address at a sponsor company.
### STEERING COMMITTEE MEMBERS, 2003-2004

<table>
<thead>
<tr>
<th>Name</th>
<th>Company</th>
<th>Telephone</th>
<th>E-Mail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raymond Abma</td>
<td>BP America</td>
<td>(281) 366-4604</td>
<td><a href="mailto:abmar1@bp.com">abmar1@bp.com</a></td>
</tr>
<tr>
<td>Francois Audebert</td>
<td>CGG Americas</td>
<td>(281) 646-2524</td>
<td><a href="mailto:faudebert@cgg.com">faudebert@cgg.com</a></td>
</tr>
<tr>
<td>Dimitri Bevc</td>
<td>3DGeo</td>
<td>(650) 969-3886</td>
<td><a href="mailto:dimitri@3dgeo.com">dimitri@3dgeo.com</a></td>
</tr>
<tr>
<td>Biondo Biondi</td>
<td>SEP</td>
<td>(650) 723-1319</td>
<td><a href="mailto:biondo@sep.stanford.edu">biondo@sep.stanford.edu</a></td>
</tr>
<tr>
<td>Luis Canales</td>
<td>WesternGeco</td>
<td>(713) 806-5271</td>
<td><a href="mailto:lcanales@houston.westerngeco.slb.com">lcanales@houston.westerngeco.slb.com</a></td>
</tr>
<tr>
<td>Jon Claerbout</td>
<td>SEP</td>
<td>(650) 723-3717</td>
<td><a href="mailto:jon@sep.stanford.edu">jon@sep.stanford.edu</a></td>
</tr>
<tr>
<td>Richard Cook</td>
<td>Shell</td>
<td>(713) 245-7195</td>
<td><a href="mailto:richard.cook@shell.com">richard.cook@shell.com</a></td>
</tr>
<tr>
<td>Helmut Jakubowicz</td>
<td>Veritas DGC</td>
<td>(44) 1293 443219</td>
<td><a href="mailto:helmut_jakubowicz@veritasdgc.com">helmut_jakubowicz@veritasdgc.com</a></td>
</tr>
<tr>
<td>Stewart Levin</td>
<td>Landmark Graphics</td>
<td>(303) 779-8080</td>
<td><a href="mailto:salevin@lgc.com">salevin@lgc.com</a></td>
</tr>
<tr>
<td>Simon Spitz</td>
<td>CGG</td>
<td>(281) 646-2502</td>
<td><a href="mailto:sspitz@cgg.com">sspitz@cgg.com</a></td>
</tr>
</tbody>
</table>


**Research Personnel**

**Gabriel Alvarez** received a B.S. degree in Physics from the National University of Colombia in 1985 and an M.Sc. in Geophysics from the Colorado School of Mines in 1995, where he was a member of the Center for Wave Phenomena. From 1989 to 2000, he worked at the Instituto Colombiano del Petroleo (ICP), the Research and Development Division of Ecopetrol, the National Oil Company of Colombia. Joined SEP in 2000, and is currently working towards a Ph.D. in geophysics at Stanford University.

**Brad Artman** received his B.Sc. in Geophysical Engineering from the Colorado School of Mines in December 1996. He worked at Shell Deepwater Development Company in New Orleans in petrophysical and geophysical capacities until joining SEP in the fall of 2000 to work toward a Ph.D. Brad is a member of SEG and SPWLA.

**James G. Berryman** received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on seismic waves in rocks and sediments – at AT&T Bell Laboratories (1978-81) and at Lawrence Livermore National Laboratory (1981- ), where he is currently a physicist in the Earth Sciences Division. Continuing research interests include acoustic, seismic, and electrical methods of geophysical imaging and waves in porous media containing fluids. He is a member of ASA, AGU, APS, and SEG.
**Biondo L. Biondi** graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987, he worked as a Research Geophysicist for TOTAL, Compagnie Francaise des Petroles in Paris. After his Ph.D. at Stanford, Biondo worked for three years with Thinking Machines Co. on the applications of massively parallel computers to seismic processing. After leaving Thinking Machines, Biondo started 3DGeo Development, a software and service company devoted to high-end seismic imaging. Biondo is now Associate Professor (Research) of Geophysics and leads SEP efforts in 3-D imaging. He is a member of SEG and EAGE.

**Morgan Brown** received a B.A. in Computational and Applied Mathematics from Rice University in 1997 and is currently completing a Ph.D. at SEP. Morgan worked as a research intern with Western Geophysical in 1997 and Landmark Graphics in 2000. He is a member of SEG and EAGE.

Robert Clapp received his B.Sc. (Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.

Marie Clapp, formerly Marie Prucha, received her B.Sc. in Geophysical Engineering from Colorado School of Mines in May 1997. She joined SEP in September 1997 and received her M.S. in June 1999. She married one of her fellow SEPers in 2001 and finally changed her last name in the summer of 2002. She is currently edging towards a Ph.D. in geophysics. She is a member of SEG.

William Curry has been with SEP since 2000. He graduated in 2000 from the University of Alberta with a B.Sc. (Hons.) in Geophysics, and received his M.S. from SEP in 2002. From 1998-1999, he worked as a seismic interpreter with Jet Energy Corporation in Calgary, Canada, and during the summer of 2002 with Marathon Oil in Houston. Bill is a student member of the AGU, CGU, CSEG, EAGE, IEEE, SEG, and SIAM.
John Etgen works for BP in Houston. He spends his time taking good ideas from SEP (and a few other places), coding them up, and processing huge surveys with them. He finally learned to love the 45 degree equation just last year.

Antoine Guitton received a M.Sc. in geophysics from Universite de Strasbourg, France in 1996 and from Stanford University in 2000. Received a "Diplome d’ingenieur de L’Ecole de Physique du Globe de Strasbourg" in 1996. He received the Best Student Paper Award from the SEG in 1999. Assistant research geophysicist at the Institut Francais du Petrole (Paris-1996/97) working on well seismic imaging. Assistant research geophysicist at CGG Houston (1997-98) working on multiples attenuation. He joined SEP in September 1998. Current research topics are nonlinear inversion and noise attenuation. He is a member of the SEG.

Seth Haines received a B.A. (Hons) in Geology and Physics (double major) from Middlebury College in 1997, and began graduate work in the Stanford Geophysics Department in 1998 with a project involving seismic reflection and refraction profiling of the Tibetan crust. He received his M.Sc. in 2001, and is currently working on applying the electroseismic method to geophysical exploration problems as he continues toward a Ph.D. He is a member of AGU, GSA, and EEGS.
Jun Ji received a B.S. in Engineering from the Seoul National University in 1987, an M.S. in Engineering from the Graduate School of Seoul National University in 1989, and a Ph.D. in Geophysics from Stanford in 1995. During 1996, he worked as a Research Geophysicist for 3DGeo Development Inc. Since 1997, he has been a professor with Software System at Hansung University in Korea. During 2003, he is visiting Stanford for his sabbatical year. He is a member of the SEG and KSEG.

Jesse Lomask graduated in 1993 with a B.S. in Geology from Temple University in Philadelphia. He then worked as a field engineer for Anadrill-Schlumberger in the Gulf of Mexico for two years. In 1998, he completed a Masters in Exploration and Development at Stanford which included a summer internship at Mobil Exploration and Producing in Houston. He then worked as a staff geophysicist for Occidental Oil and Gas at Elk Hills, California, where he interpreted recently acquired 3D seismic survey. In the fall of 2001, he returned to Stanford to join SEP.

Daniel A. Rosales received his B.S. in Geophysics from Universidad Simon Bolivar, Venezuela, in 1997. From 1996 to 1998, he worked at PDVSA-INTEVEP S.A. in seismic modeling, 2-D prestack depth migration, and velocity estimation by wave-field downward continuation. In 1998, he joined Simon Bolivar University as an instructor in geophysics. He obtained his M.Sc in geophysics from Stanford University in June 2001, and is currently working towards his Ph.D. in geophysics with the Stanford Exploration Project. He is a member of SEG and AAPG.
Paul Sava graduated in June 1995 from the University of Bucharest, with an Engineering Degree in Geophysics. Between 1995 and 1997, he was employed by Schlumberger GeoQuest. He joined SEP in 1997, received his M.Sc. in 1998, and continues his work toward a Ph.D. in Geophysics. His main research interest is in seismic imaging using wave-equation techniques. He is a member of SEG, EAGE, SIAM and AGU.

Guojian Shan received his B.Sc. in Mathematics School of Peking University in July, 1998. From 1998 to 2001, he studied in Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences (CAS), and received his M.S. in Applied Mathematics in July, 2001. He joined SEP in 2001 and is currently working towards a Ph.D. in geophysics. He is a member of the SEG.

Jeff Shragge graduated in 1998 with a BScH in Honours Physics from Queen’s University in Kingston, Canada. After completing a MSc degree in 2001 in teleseismic imaging at the UBC in Vancouver, Canada, he spent 2001 and 2002 working for a small geophysical consulting company based out of Woburn, MA. He joined SEP in 2002, and is working towards a Ph.D. in Geophysics. His main research interest is migration and wavefield inversion. He is a member of SEG and AGU.

Thomas Tisserant received a “Diplôme d’ingénieur de l’Ecole de Physique du Globe de Strasbourg” in 2002. He joined SEP in September 2002, and is working towards a M.S. in Geophysics. He is a member of SEG.
**Alejandro A. Valenciano** received a B.Sc. degree in Physics from Havana University (Cuba) in 1994, and a M.Sc. in Physics from Simon Bolivar University (Venezuela) in 1998. He worked in the Earth Science Department of PDVSA-INTEVEP from 1995 to 2001. He joined SEP to work towards a Ph.D in geophysics in the Fall of 2001.

**Ioan Vlad** graduated in June 2000 with an Engineer Diploma (5-year degree) in Geophysics from the University of Bucharest, with a thesis on gravity and geodynamical modeling of the lithosphere. He joined SEP in 2000 and is currently working towards a Ph.D. in geophysics at Stanford. He is a member of SEG.
Charlie Wilson received his B.Sc. degree in geology from the University of Arizona in 1998 and a Ph.D. in Geophysics from the University of Colorado, Boulder in 2003. His thesis was on imaging the continental lithosphere using teleseismic earthquakes to solve tectonic problems. He joined the Stanford Geophysics Department as a Post-Doc in December 2003 and is currently working on several projects related to tectonics and earthquake imaging. He is a member of AGU.

Kevin Wolf received a B.Sc. (Hons.) in Earth Sciences Geophysics option from the University of Waterloo in June 2001 and is currently working towards a Ph.D. in geophysics at Stanford. Kevin worked as a research geophysicist at Imperial Oil Resources on a co-op work term in 1999 and with the Geological Survey of Canada in 2003. He is a member of SEG.
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2003-2004

Amerada Hess Corporation
One Allen Center
500 Dallas St.
Houston, TX 77002
USA
tel: (713) 609-5829
fax: (713) 609-5666
contact: Scott A. Morton
e-mail: morton@hess.com

ChevronTexaco E&P Technology Company
935 Gravier St., Room 1668
New Orleans, LA 70112
USA
tel: (504) 592-6153
fax: (504) 592-6742
contact: Ming Zhao
e-mail: ming.zhao@chevronexaco.com

Aramco Services Company
Saudi Aramco
Geophysics Technology
P.O. Box 750, EXPECL Bldg., X-1370
Dhahran 31311
Saudi Arabia
tel: 966 (3) 874 7262
fax: 966 (3) 874 1020
contact: Mohammed N. Alfaraj
e-mail: mohammed.faraj@aramco.com

ConocoPhillips
P.O. Box 2197
Lobo 3000
Houston, TX 77252
USA
tel: (281) 293-5696
fax: (281)
contact: Robert M. Habiger
e-mail: Rob.M.Habiger@ConocoPhillips.com

BP
200 Westlake Park Blvd., #1018
Houston, TX 77079
USA
tel: (281) 366-3611
fax: (281) 366-5856
contact: John T. Etgen
e-mail: john.etgen@bp.com

Ecopetrol-ICP
Laboratorio de Geofisica
A.A. 4185 Bucaramanga
Colombia
tel: (57) 76 445420
fax: (57) 76 445444
contact: Alfredo Tada
e-mail: atada@ecopetrol.com.co

CGG Americas, Inc.
16430 Park Ten Place
Houston, TX 77079
USA
tel: (281) 646-2525
fax: (281) 646-2620
contact: Francois S. Audebert
e-mail: faudebert@cgg.com

ENI - E & P Division
Dept. RIGE
via Unione Europea 3
20097 S. Donato Milanese
Italy
tel: 39 (02) 520 55308
fax: 39 (02) 520 45694
contact: Vittorio De Tomasi
e-mail: vittorio.detomasi@agip.it
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2003-2004

ExxonMobil Upstream Research Company
Seismic Processing Research
URC-GW3-834A
P.O. Box 2189, 3319 Mercer St.
Houston, TX 77081-2189
USA
tel: (713) 431-6011
colorado@exxonmobil.com

Norsk Hydro ASA
PB 7190
Sandsliveien 90
Bergen, N-5020
Norway
tel: (47) 5599 6861
fax: (47) 5599 6970
contact: Arnnn Sagehaug
e-mail: Arnnn.Sagehaug@nho.hydro.com

ExxonMobil Upstream Research Company
Seismic Processing Research
URC-GW3-834A
P.O. Box 2189, 3319 Mercer St.
Houston, TX 77081-2189
USA
tel: (713) 431-6011
colorado@exxonmobil.com

Norsk Hydro ASA
PB 7190
Sandsliveien 90
Bergen, N-5020
Norway
tel: (47) 5599 6861
fax: (47) 5599 6970
contact: Arnnn Sagehaug
e-mail: Arnnn.Sagehaug@nho.hydro.com

Fairfield Industries, Inc.
14100 Southwest Fwy.
Suite 100
Sugar Land, TX 77478
USA
tel: (281) 275-7597
colorado@exxonmobil.com

Paradigm Geophysical Corporation
Two Memorial Plaza
820 Gessner, Suite 400
Houston, TX 77024
USA
tel: (713) 393-4979
colorado@exxonmobil.com

Fairfield Industries, Inc.
14100 Southwest Fwy.
Suite 100
Sugar Land, TX 77478
USA
tel: (281) 275-7597
colorado@exxonmobil.com

Paradigm Geophysical Corporation
Two Memorial Plaza
820 Gessner, Suite 400
Houston, TX 77024
USA
tel: (713) 393-4979
colorado@exxonmobil.com

GX Technology Corporation
5847 San Felipe
Suite 3500
Houston, TX 77057-3010
USA
tel: (713) 789-7250
colorado@exxonmobil.com

Petrobras
Avenida Chile 65, s/1302
Rio de Janeiro
20031-912 RJ
Brazil
tel: 55 (21) 2534-2706
colorado@exxonmobil.com

GX Technology Corporation
5847 San Felipe
Suite 3500
Houston, TX 77057-3010
USA
tel: (713) 789-7250
colorado@exxonmobil.com

Petrobras
Avenida Chile 65, s/1302
Rio de Janeiro
20031-912 RJ
Brazil
tel: 55 (21) 2534-2706
colorado@exxonmobil.com

Landmark Graphics Corporation
1805 Shea Center Dr.
Suite 400
Highlands Ranch, CO 80129-2258
USA
tel: (303) 779-8080
colorado@exxonmobil.com

PGS Seismic Processing Technology
10550 Richmond Ave.
Suite 250
Houston, TX 77042
USA
tel: (713) 735-6322
colorado@exxonmobil.com

Landmark Graphics Corporation
1805 Shea Center Dr.
Suite 400
Highlands Ranch, CO 80129-2258
USA
tel: (303) 779-8080
colorado@exxonmobil.com

PGS Seismic Processing Technology
10550 Richmond Ave.
Suite 250
Houston, TX 77042
USA
tel: (713) 735-6322
colorado@exxonmobil.com
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2003-2004

Shell International E&P Inc. Deepwater USA, UNOCAL
3737 Bellaire Blvd. 14141 Southwest Fwy.
P.O. Box 481 Sugar Land, TX 77478
Houston, TX 77001-0481 USA
USA
tel: (713) 245-7195 fax: (281) 287-5360
contact: Richard W. Cook email: phil.schultz@unocal.com
email: richard.cook@shell.com
tel: (281) 287-7481 fax: (281) 287-5360
contact: Philip S. Schultz
Statoil Veritas DGC Ltd.
Arkitekt Ebbellsvei 10 Crompton Way
Rotvoll Manor Royal Estate
Trondheim, N-7005 Crawley, West Sussex RH10 2QR
Norway England
tel: (47) 73584917 tel: 44 (1293) 443219
fax: (47) 73584665 fax: 44 (1293) 443010
contact: Constantin Gerea contact: Helmut Jakubowicz
eemail: geco@statoil.com email: Helmut_Jakubowicz@veritasdgc.com
3DGeo Development Inc. Weinman GeoScience
4633 Old Ironsides Dr. 17103 Preston Road
Suite 401 Suite 200 N.
Santa Clara, CA 95054 Dallas, TX 75248
USA USA
tel: (408) 450 7840, x102 tel: (972) 818-2550
fax: (408) 450 7809 fax: (972) 818-2553
contact: Dimitri Bevc contact: Barry L. Weinman
eemail: dimitri@3dgeo.com email: barryw@weinmangeoscience.com
TOTAL E&P USA, Inc. WesternGeco
800 Gessner Houston Technology Center
Suite 700 10001 Richmond Ave.
Houston, TX 77024 Houston, TX 77042-4299
USA USA
tel: (713) 647-3940 tel: (713) 806-5271
fax: (713) 647-3875 fax: (713) 689-5757
contact: Henri Calandra contact: Luis L. Canales
eemail: henri.calandra@total.com email: LCanales@houston.westerngeco.slb.com