Preface

The electronic version of this report 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.

**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. 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.

**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.

**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.

**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/>.

---

SEP–114 — TABLE OF CONTENTS

Imaging

Sava P. and Fomel S., Seismic imaging using Riemannian wavefield extrapolation ........ 1
Shragge J. and Biondi B., Wavefield extrapolation in phase-ray coordinates .............. 31
Biondi B., Amplitude balancing of 3-D angle-domain common-image gathers ............ 45
Clapp M. L., Velocity sensitivity of subsalt imaging through regularized inversion ........ 57
Shan G. and Biondi B., WKBJ and amplitude preserving one-way wave equation ......... 67
Tisserant T. and Biondi B., Kinematics of 3-D angle-domain common-image gathers for migration velocity analysis ..................................................... 71

Velocity Analysis

Sava P. and Biondi B., Wave-equation MVA: Born, Rytov and beyond ....................... 83
Shan G. and Biondi B., Dip-dependent residual moveout ....................................... 95
Vlad I., Biondi B., and Sava P., Velocity estimation for seismic data exhibiting focusing-effect AVO (Part 3) ................................................................. 101
Chen W., Clapp R. G., and Biondi B., Reflection tomography with depth control ......... 111
Vlad I., Frequency-dependent velocity analysis? .................................................. 119

Novel inversion methods

Artman B. and Sacchi M., Basis pursuit for geophysical inversion .......................... 137
Valenciano A. A. and Brown M., Deblurring using edge-preserving regularization ...... 151

Irregular data

Clapp R. G., AMO regularization: Effective approximate inverses for amplitude preservation ................................................................. 159
Curry W., More fitting equations for PEF estimation on sparse data ....................... 171
Brown M., Dip estimation from irregularly-sampled seismic data .......................... 177
Curry W., Prediction-error filter estimation on irregular traces ......................... 185

Image Segmentation

Lomask J. and Biondi B., Image segmentation for tracking salt boundaries .............. 193
Ji J., Automatic discontinuity extraction for 3-D seismic images ...................... 201

Rock Physics

Berryman J. G., Elastic and poroelastic analysis of Thomsen parameters for seismic waves in finely layered VTI media ................................................... 215
Berryman J. G., Scale-up in poroelastic systems and applications to reservoirs .......... 237
Pride S. R., Berryman J. G., and Harris J. M., Seismic attenuation due to wave-induced flow ................................................................. 249

SEP article published or in press, 2002-03 ................................................. 283
SEP phone directory ................................................................. 287
Research personnel ................................................................. 289
SEP sponsors for 2002-03 .......................................................... 295
Seismic imaging using Riemannian wavefield extrapolation

Paul Sava and Sergey Fomel

ABSTRACT

Riemannian spaces are described by non-orthogonal curvilinear coordinates. We generalize one-way wavefield extrapolation to semi-orthogonal Riemannian coordinate systems, which include, but are not limited to, ray coordinate systems. We obtain one-way wavefield extrapolation methods which are not dip-limited, and which can even be used to image overturning waves. Ray coordinate systems can be initiated either from point sources, or from plane waves incident at various angles. Since wavefield propagation happens mostly along the extrapolation direction, we can use cheap finite-difference or mixed-domain extrapolators to achieve high angle accuracy. The main applications of our method include imaging of steeply dipping or overturning reflections.

INTRODUCTION

Imaging complex geology is one of the main challenges of today’s seismic processing. Of the many seismic imaging methods available, downward continuation (Claerbout, 1985) has proven to be accurate, robust, and capable of handling models with large and sharp velocity variations. In addition, such methods naturally handle the multipathing which occurs in complex geology and provide a band-limited solution to the seismic imaging problem. Furthermore, as computational power increases, such methods are gradually moving into the mainstream of seismic processing.

However, migration by downward continuation imposes strong limitations on the dip of reflectors that can be imaged since, by design, it favors energy which is propagating mainly in the downward direction. Upward propagating energy, e.g., overturning waves, can be imaged in principle using downward continuation methods (Hale et al., 1992), although the procedure is difficult, particularly for prestack data. In contrast, Kirchhoff-type methods based on ray-traced traveltimes can image steep dips and handle overturning waves, although those methods are far less reliable in complex environments given their high-frequency asymptotic nature.

The steep-dip limitation of downward continuation techniques has been addressed in several ways:

- A first option is to increase the angular accuracy of the extrapolation operator, for example by employing methods from the Fourier finite-difference (FFD) family (Ristow

1email: paul@sep.stanford.edu, sergey.fomel@beg.utexas.edu
and Ruhl, 1994; Biondi, 2002), or the Generalized Screen Propagator (GSP) family (de Hoop, 1996; Huang and Wu, 1996). The enhancements brought about by these methods come at a price, since they increase the cost of extrapolation without guaranteeing unconditional stability.

- A second option is to perform the wavefield extrapolation in tilted coordinate systems (Etgen, 2002), or by designing sources which favor illumination of particular regions of the image (Rietveld and Berkhout, 1994; Chen et al., 2002). We can thus increase angular accuracy, although those methods favor a subset of the model (a salt flank, for example), while potentially decreasing the accuracy in other regions. In complex models it is also not obvious what is an optimal tilt angle for the extrapolation grid.

- A third possibility is hybridization of wavefield and ray-based techniques, either in the form of Gaussian beams (Hill, 1990, 2001; Gray et al., 2002), coherent states (Albertin et al., 2001, 2002), or beam-waves (Brandsberg-Dahl and Etgen, 2003). Such techniques are quite powerful, since they couple wavefield methods with multipathing and band-limited properties, with ray methods, which deliver arbitrary directions of propagation, even overturning. The main strength of those techniques, is also their main weakness, i.e. they are formulated in terms of decoupled beams. Beams may leave shadow zones in various parts of the model, which hamper their imaging abilities. Furthermore, beams have limited size, which in turn limits the extent of the diffractions created by sharp features in the model to that of any particular beam, no matter how accurate the extrapolator within each beam is. In addition, the narrow extrapolation domain poses serious beam superposition problems, such as beam boundary effects.

The main idea of our paper is to couple the beams together and extrapolate within all of them at once. We, therefore, cannot talk about beams anymore, but instead we need to talk about continuously changing coordinate systems. We extend downward continuation in a regular Cartesian space to wavefield extrapolation in distorted coordinates, known as Riemannian spaces, thus the name of our method. We formulate the theory in arbitrary 3-D semi-orthogonal Riemannian spaces, e.g., ray coordinates, although those coordinate systems do not necessarily need to have a physical meaning as long as they fulfill the semi-orthogonality condition. Examples of such coordinate systems include, but are not limited to, fans of rays emerging from a source point, or bundles of rays initiated by plane waves of arbitrary initial dips at the source. A special case of our method is represented by the polar/spherical coordinate system (Nichols and Palacharla, 1994; Nichols, 1994, 1996).

Our method can be seen as a finite-difference solution to the acoustic wave-equation in ray coordinates. In this respect, it is closely related to Huygens wavefront tracing (Sava and Fomel, 2001), which represents a finite-difference solution to the eikonal equation in ray coordinates.

Another idea related to our method is that of wave-equation in ray-centered coordinates (Yedlin, 1981; Červený, 2001). However, our method is different since the ray-centered coordinate system is orthogonal in 3-D and defined around an individual ray. In contrast, our method is formulated in ray coordinates which are parameterizations of the wavefields at the source, and which are non-orthogonal in 3-D and defined globally for an entire family of rays.
The upside of our method is that the coordinate system may follow the waves, and can even overturn, such that we can use one-way extrapolators to image diving waves (Figure 1). We can also use extrapolators with small angle accuracy (e.g. 15°), since, in principle, we are never too far from the actual direction in which waves propagate. We are also not confined to the extent of any individual beam, therefore we can track diffractions for their entire spatial extent.

Figure 1: Ray coordinate systems are superior to tilted coordinate systems for imaging overturning waves using one-way wavefield extrapolators. Overturning reflected energy may become evanescent in tilted coordinate systems (a), but stays non-evanescent in ray coordinate systems (b).

ACOUSTIC WAVE-EQUATION IN 3-D RIEMANNIAN SPACES

The Laplacian operator of a scalar function $U$ in an arbitrary Riemannian space with coordinates $\{\xi_1, \xi_2, \xi_3\}$ has the form

$$\Delta U = \sum_{i=1}^{3} \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial \xi_i} \left( \sum_{j=1}^{3} g^{ij} \sqrt{|g|} \frac{\partial U}{\partial \xi_j} \right),$$  \hspace{1cm} (1)

where $g^{ij}$ is a component of the associated metric tensor, and $|g|$ is its determinant (Synge and Schild, 1978).

The expression simplifies if one of the coordinates (e.g. the coordinate of one-way wave extrapolation) is orthogonal to the other coordinates. Let $\xi_1 = \xi$, $\xi_2 = \eta$, and $\xi_3 = \zeta$, with $\zeta$ orthogonal to both $\xi$ and $\eta$. Then the metric tensor has the matrix

$$[g_{ij}] = \begin{bmatrix} E & F & 0 \\ F & G & 0 \\ 0 & 0 & \alpha^2 \end{bmatrix},$$  \hspace{1cm} (2)

where $E$, $F$, $G$, and $\alpha$ are differential forms that can be found from mapping Cartesian coordinates $x$ to the general coordinates $\{\xi, \eta, \zeta\}$, as follows:

$$E = x_\xi \cdot x_\xi,$$
\[ F = \mathbf{x}_\xi \cdot \mathbf{x}_\eta, \]
\[ G = \mathbf{x}_\eta \cdot \mathbf{x}_\eta, \]
\[ \alpha^2 = \mathbf{x}_\xi \cdot \mathbf{x}_\xi. \]  

The associated metric tensor has the matrix
\[
\begin{bmatrix}
  +G/J^2 & -F/J^2 & 0 \\
  -F/J^2 & +E/J^2 & 0 \\
  0 & 0 & 1/\alpha^2
\end{bmatrix},
\]
where \( J^2 = EG - F^2 \). The metric determinant takes the form
\[ |g| = \alpha^2 J^2. \]

Substituting equations (4) and (5) into (1), we can modify the Helmholtz wave equation
\[ \Delta \mathcal{U} = -\frac{\omega^2}{v^2(\mathbf{x})} \mathcal{U} \]
for propagating waves in a 3-D Riemannian space:
\[
\frac{1}{\alpha J} \left[ \frac{\partial}{\partial \xi} \left( \frac{J}{\alpha} \frac{\partial \mathcal{U}}{\partial \xi} \right) + \frac{\partial}{\partial \eta} \left( \frac{G}{J} \frac{\partial \mathcal{U}}{\partial \eta} - F \frac{\alpha}{J} \frac{\partial \mathcal{U}}{\partial \eta} \right) + \frac{\partial}{\partial \zeta} \left( \frac{E}{J} \frac{\partial \mathcal{U}}{\partial \zeta} - F \frac{\alpha}{J} \frac{\partial \mathcal{U}}{\partial \zeta} \right) \right] = -\frac{\omega^2}{v^2} \mathcal{U}. \]

In equation (6), \( \omega \) is temporal frequency, \( v[\mathbf{x}(\xi, \eta, \zeta)] \) is the wave propagation velocity, and \( \mathcal{U} \) represents a propagating wave.

For the special case of two dimensional spaces \((F = 0 \text{ and } G = 1)\), the Helmholtz wave equation reduces to the simpler form:
\[
\frac{1}{\alpha J} \left[ \frac{\partial}{\partial \zeta} \left( \frac{J}{\alpha} \frac{\partial \mathcal{U}}{\partial \zeta} \right) + \frac{\partial}{\partial \eta} \left( \alpha \frac{\partial \mathcal{U}}{\partial \eta} \right) + \frac{\partial}{\partial \xi} \left( \frac{\alpha}{J} \frac{\partial \mathcal{U}}{\partial \xi} + \frac{\alpha}{J} \frac{\partial \mathcal{U}}{\partial \xi} \right) \right] = -\frac{\omega^2}{v^2} \mathcal{U},
\]
which corresponds to a curvilinear orthogonal coordinate system.

Particular examples of coordinate systems for one-way wave propagation are:

**Cartesian (propagation in depth):** \( x_1 = \xi, x_2 = \eta, x_3 = \zeta, \)
\[ E = G = \alpha = J = 1, \]
\[ F = 0. \]

**Cylindrical (propagation in radius):** \( x_1 = \zeta \cos \xi, x_2 = \zeta \sin \xi, x_3 = \eta, \)
\[ E = J = \zeta^2, \]
\[ G = \alpha = 1, \]
\[ F = 0. \]
Spherical (propagation in radius): 

\[ x_1 = \zeta \sin \xi \cos \eta, \ x_2 = \zeta \sin \xi \sin \eta, \ x_3 = \zeta \cos \xi, \]

\[ E = \zeta^2, \]
\[ G = \zeta^2 \sin^2 \xi, \]
\[ \alpha = 1, \]
\[ J = \zeta^2 \sin \xi, \]
\[ F = 0. \]

Ray family (propagation along rays): \( \xi \) and \( \eta \) represent parameters defining a particular ray in the family (i.e. the ray take-off angles), \( J \) is the geometrical spreading factor, related to the cross-sectional area of the ray tube (Červený, 2001). The coefficients \( E, F, G, \) and \( J \) are easily computed by finite-difference approximations with the Huygens wavefront tracing technique (Sava and Fomel, 2001). If the propagation parameter \( \zeta \) is taken to be time along the ray, then \( \alpha \) equals the propagation velocity \( v \).

**ONE-WAY WAVE-EQUATION IN 3-D RIEMANNIAN SPACES**

Equation (6) can be used to describe two-way propagation of acoustic waves in a semi-orthogonal Riemannian space. For one-way wavefield extrapolation, we need to modify the acoustic wave equation (6) by selecting a single direction of propagation.

In order to simplify the computations, we introduce the following notation:

\[ c_{\xi \xi} = \frac{1}{\alpha J \frac{\partial}{\partial \zeta}} \left( \frac{J}{\alpha} \right), \]
\[ c_{\eta \eta} = \frac{1}{\alpha J} \left[ \frac{\partial}{\partial \eta} \left( \frac{G}{J} \right) \right] - \frac{\partial}{\partial \eta} \left( \frac{F}{J} \right), \]
\[ c_{\xi \eta} = \frac{1}{\alpha J} \left[ \frac{\partial}{\partial \xi} \left( \frac{E}{J} \right) \right] - \frac{\partial}{\partial \xi} \left( \frac{G}{J} \right), \]
\[ c_{\xi} = \frac{1}{\alpha J} \left[ \frac{\partial}{\partial \xi} \left( \frac{E}{J} \right) \right] - \frac{\partial}{\partial \xi} \left( \frac{G}{J} \right), \]
\[ c_{\eta} = \frac{1}{\alpha J} \left[ \frac{\partial}{\partial \eta} \left( \frac{E}{J} \right) \right] - \frac{\partial}{\partial \eta} \left( \frac{F}{J} \right). \] (8)

All quantities in equations (8) can be computed by finite-differences for any choice of a Riemannian coordinate system which fulfills the orthogonality condition indicated earlier. In particular, we can use ray coordinates to compute those coefficients. With these notations, the acoustic wave-equation can be written as:

\[ c_{\xi \xi} \frac{\partial^2 U}{\partial \zeta^2} + c_{\xi \xi} \frac{\partial^2 U}{\partial \xi^2} + c_{\eta \eta} \frac{\partial^2 U}{\partial \eta^2} + c_{\zeta} \frac{\partial U}{\partial \zeta} + c_{\xi} \frac{\partial U}{\partial \xi} + c_{\eta} \frac{\partial U}{\partial \eta} + c_{\xi \eta} \frac{\partial^2 U}{\partial \zeta \partial \eta} = -\frac{\omega^2}{v^2} U. \] (9)
For the particular case of Cartesian coordinates \((c_\xi = c_\eta = c_\zeta = 0, c_{\xi\xi} = c_{\eta\eta} = c_{\zeta\zeta} = 1, c_{\xi\eta} = 0)\), the Helmholtz equation (9) takes the familiar form
\[
\frac{\partial^2 U}{\partial \xi^2} + \frac{\partial^2 U}{\partial \eta^2} + \frac{\partial^2 U}{\partial \zeta^2} = -\frac{\omega^2}{v^2} U .
\] (10)

From equation (9), we can directly deduce the modified form of the dispersion relation for the wave-equation in a semi-orthogonal 3-D Riemannian space:
\[
-c_\xi k_\xi^2 - c_{\xi\xi} k_\xi^2 - c_{\eta\eta} k_\eta^2 + i c_\xi k_\xi + i c_\eta k_\eta - c_{\xi\eta} k_\xi k_\eta = -\omega^2 s^2 .
\] (11)

For one-way wavefield extrapolation, we need to solve the second order equation (11) for the wavenumber of the extrapolation direction \(k_\xi\), and select the solution with the appropriate sign to extrapolate waves in the desired direction:
\[
k_\xi = \pm \sqrt{(os)^2 - k_\xi^2 - k_\eta^2} .
\] (12)

The solution with the positive sign in equation (12) corresponds to propagation in the positive direction of the extrapolation axis \(\zeta\).

For the particular case of Cartesian coordinates \((c_\xi = c_\eta = c_\zeta = 0, c_{\xi\xi} = c_{\eta\eta} = c_{\zeta\zeta} = 1, c_{\xi\eta} = 0)\), the one-way wavefield extrapolation equation takes the familiar form
\[
k_\xi = \pm \sqrt{(os)^2 - k_\xi^2} .
\] (13)

**MIXED-DOMAIN SOLUTIONS TO THE ONE-WAY WAVE-EQUATION**

We can use equation (12) to construct a numerical solution to the one-way wave equation in the mixed \(\omega - k, \omega - x\) domain. The extrapolation wavenumber described in equation (12) is, in general, a function which depends on several quantities
\[
k_\xi = k_\xi (s, c_j) ,
\] (14)
where \(s(\xi, \eta, \nu)\) is the space-variable slowness, and \(c_j (\xi, \eta, \zeta) = \{c_\xi, c_\eta, c_\zeta, c_{\xi\xi}, c_{\eta\eta}, c_{\zeta\zeta}, c_{\xi\eta}\}\) are coefficients which are computed numerically from the definition of the coordinate system, as indicated by equations (8). For any given coordinate system, \(c_j\) can be regarded as known.

Next, we write the extrapolation wavenumber \(k_\xi\) as a first-order Taylor expansion relative to a reference medium:
\[
k_\xi = k_{\xi 0} + \left. \frac{\partial k_\xi}{\partial s} \right|_{s_0, c_{j0}} (s - s_0) + \sum_{c_j} \left. \frac{\partial k_\xi}{\partial c_j} \right|_{s_0, c_{j0}} (c_j - c_{j0}) ,
\] (15)
where \(s(\xi, \eta, \nu)\) and \(c_j (\xi, \eta, \nu)\) represent the spatially variable slowness and coordinate system parameters, and \(s_0\) and \(c_{j0}\) are the constant reference values in every extrapolation “slab” (Sava, 2000).
As usual, the first part of equation (15), corresponding to the extrapolation wavenumber in the reference medium \( k_{\xi 0} \), is implemented in the Fourier domain, while the second part, corresponding to the spatially variable medium coefficients, is implemented in the space domain.

If we make the further simplifying assumptions that \( k_{\xi} \approx 0 \) and \( k_{\eta} \approx 0 \), we can write

\[
k_{\xi} = k_{\xi 0} + \frac{\partial k_{\xi}}{\partial c_{\xi \xi}} (s - s_0) + \frac{\partial k_{\xi}}{\partial c_{\xi \eta}} (c_{\xi \xi} - c_{\xi \eta 0}) + \frac{\partial k_{\xi}}{\partial c_{\eta}} (c_{\xi} - c_{\xi 0}) ,
\]

where

\[
\frac{\partial k_{\xi}}{\partial s} \bigg|_0 = \frac{2\omega (\omega s_0)}{\sqrt{4c_{\xi \xi 0} (\omega s_0)^2 - c_{\xi 0}^2}} \\
\frac{\partial k_{\xi}}{\partial c_{\xi \xi}} \bigg|_0 = -\frac{i c_{\xi 0}}{2c_{\xi \xi 0}^2} + \frac{c_{\xi 0}^2 - 2c_{\xi \xi 0} (\omega s_0)^2}{2c_{\xi \xi 0}^2 \sqrt{4c_{\xi \xi 0} (\omega s_0)^2 - c_{\xi 0}^2}} \\
\frac{\partial k_{\xi}}{\partial c_{\xi \eta}} \bigg|_0 = \frac{i}{2c_{\xi \xi 0}} - \frac{c_{\xi 0}}{2c_{\xi \xi 0}^2 \sqrt{4c_{\xi \xi 0} (\omega s_0)^2 - c_{\xi 0}^2}} .
\]

Equation (16) is motivated by a wavefront normal propagation approximation. By “0”, we denote the reference medium \((s_0, c_{\xi 0}, c_{\eta 0})\). We could also use many reference media, followed by interpolation, similarly to the technique of Gazdag and Sguazzero (1984).

For the particular case of Cartesian coordinates \((c_\xi = 0, c_{\xi \xi} = 1)\), equation (16) reduces to

\[
k_{\xi} = k_{\xi 0} + \omega (s - s_0) ,
\]

which corresponds to the popular Split-Step Fourier (SSF) extrapolation method (Stoffa et al., 1990).

**FINITE-DIFFERENCE SOLUTIONS TO THE ONE-WAY WAVE EQUATION**

Alternative solutions to the one-way wave-equation are represented by pure finite-difference methods in the \(\omega - x\) domain, which can be implemented either as implicit (Claerbout, 1985), or as explicit methods (Hale, 1991). For the same stencil size, the implicit methods are more accurate and robust than explicit methods, although harder to implement, particularly in 3-D. However, explicit methods of comparable accuracy can be designed using larger stencils.

For the implicit methods, various approximations to the square root in equation (12) lead to approximate equations of various orders of accuracy. In the Cartesian space, those methods are known by their respective angular accuracy as the 15° equation, 45° equation and so on. Although the meanings of 15°, 45° are undefined in ray coordinates where the extrapolation axis is time, we can still write approximations for the numerical finite-difference solutions using analogous approximations.
If we introduce the notation

\[ k_o^2 = \frac{(\cos)^2}{c_{\zeta \zeta}} - \left( \frac{c_\xi}{2c_{\zeta \zeta}} \right)^2 \]  

we can simplify the one-way wave equation (12) as

\[ k_\xi = i \frac{c_\xi}{2c_{\zeta \zeta}} + \sqrt{k_o^2 - \left( \frac{c_{\xi \xi}}{c_{\zeta \zeta}} k_\xi^2 - i \frac{c_\xi}{c_{\zeta \zeta}} k_\eta \right) - \left[ \frac{c_{\eta \eta} k_\eta^2 - i \frac{c_\eta}{c_{\zeta \zeta}} k_\eta}{c_{\zeta \zeta}} \right] - \frac{c_{\xi \eta} k_\xi k_\eta}{c_{\zeta \zeta}} . \]  

(19)

\[ k_\eta = \frac{c_\eta}{2c_{\zeta \zeta}} + \sqrt{k_o^2 - \left( \frac{c_{\eta \eta}}{c_{\zeta \zeta}} k_\eta^2 - i \frac{c_\eta}{c_{\zeta \zeta}} k_\xi \right) - \left[ \frac{c_{\xi \xi} k_\xi^2 - i \frac{c_\xi}{c_{\zeta \zeta}} k_\xi}{c_{\zeta \zeta}} \right] - \frac{c_{\xi \eta} k_\xi k_\eta}{c_{\zeta \zeta}} . \]  

(20)

**15° wave-equation in a 3-D Riemannian space**

A simple way of deriving the 15° equation is by a second order Taylor series expansion of the extrapolation wavenumber \( k_\xi \) relative to \( k_\xi \) and \( k_\eta \):

\[ k_\xi \approx k_\xi (k_\xi = 0, k_\eta = 0) + \frac{\partial k_\xi}{\partial k_\xi} k_\xi + \frac{\partial k_\xi}{\partial k_\eta} k_\eta + \frac{1}{2} \frac{\partial^2 k_\xi}{\partial k_\xi^2} k_\xi^2 + \frac{1}{2} \frac{\partial^2 k_\xi}{\partial k_\eta^2} k_\eta^2 + \frac{1}{2} \frac{\partial^2 k_\xi}{\partial k_\xi \partial k_\eta} k_\xi k_\eta . \]  

(21)

If we plug equation (20) into equation (21), we obtain an equivalent form for the 15° equation in a semi-orthogonal 3-D Riemannian space:

\[ k_\xi \approx i \frac{c_\xi}{2c_{\zeta \zeta}} + k_o + \frac{ic_\xi}{2c_{\zeta \zeta} k_o} k_\xi + \frac{1}{2} k_o \left[ \left( \frac{c_\xi}{2c_{\zeta \zeta} k_o} \right)^2 - \frac{c_{\xi \xi}}{c_{\zeta \zeta}} \right] k_\xi^2 + \frac{ic_\eta}{2c_{\zeta \zeta} k_o} k_\eta + \frac{1}{2} k_o \left[ \left( \frac{c_\eta}{2c_{\zeta \zeta} k_o} \right)^2 - \frac{c_{\eta \eta}}{c_{\zeta \zeta}} \right] k_\eta^2 + \frac{1}{2} k_o \left[ \frac{c_\xi c_\eta}{2c_{\zeta \zeta} k_o^2 - \frac{c_{\xi \eta}}{c_{\zeta \zeta}}} \right] k_\xi k_\eta . \]  

(22)

For the particular case of Cartesian coordinates \( (c_\xi = c_\eta = c_\xi = 0, c_{\xi \xi} = c_{\eta \eta} = c_{\zeta \zeta} = 1, c_{\xi \eta} = 0, k_o = \omega s) \),

\[ k_\xi \approx \omega s - \frac{1}{2 \omega s} \left( k_\xi^2 + k_\eta^2 \right) , \]  

(23)

which is the usual form of the 15° equation.

**EXAMPLES**

We illustrate our method with several synthetic examples of various degrees of complexity. In all examples, we use extrapolation in 2-D orthogonal Riemannian spaces (ray coordinates), and compare the results with extrapolation in Cartesian coordinates. We present images obtained by migration of synthetic datasets represented by events equally spaced in time. When
we use point sources, those images are representations of Green’s functions. In all examples, \((x, z)\) are the Cartesian space coordinates, \((\tau, x_0)\) are the ray coordinates for plane wave sources, and \((\tau, \gamma)\) are ray coordinates for point sources. \(x_0\) stands for surface coordinate, \(\gamma\) for shooting angle, and \(\tau\) for one-way traveltime.

Our first example is designed to illustrate our method in a fairly simple model. We use a 2-D model with horizontal and vertical gradients \(v(x, z) = 250 + 0.2x + 0.15z\) m/s which gives waves propagating from a point source a pronounced tendency to overturn (Figure 2). The model also contains a diffracon located around \(x = 3800\) m and \(z = 3000\) m.

We use ray tracing to create an orthogonal ray coordinate system corresponding to a point source on the surface at \(x = 6000\) m. Figure 2(a) shows the velocity model and the rays in the original Cartesian coordinate system \((x, z)\). Figure 2(c) shows the velocity model mapped into the ray coordinate system \((\tau, \gamma)\). The diffracon is mapped to \(\tau = 2.4\) s and \(\gamma = -18^\circ\) measured from the vertical. The synthetic data we use is represented by impulses at the source location at every 0.25 s. In ray coordinates, this source is represented by a plane-wave evenly distributed over all shooting angles \(\gamma\). Ideally, an image obtained by migrating such a dataset is a representation of the acoustic wavefield produced by a source which pulsates periodically.

Figure 2(b) shows the image obtained by downward continuation in Cartesian coordinates using the standard 15° equation. Figure 2(d) shows the image obtained by wavefield extrapolation in ray coordinates using the equivalent 15° equation. The overlays in panels (b) and (d) are wavefronts at every 0.25 s and rays shot at every 20° to facilitate one-to-one comparisons between the images in ray and Cartesian coordinates.

Figure 3 is a direct comparison of the results obtained by extrapolation in the two coordinate systems. The image created by extrapolation in Cartesian coordinates (a) is mapped to ray coordinates (b). The image created by extrapolation in ray coordinates (c) is mapped to Cartesian coordinates (d). Since we use the same velocity for ray tracing and for wavefield extrapolation, we expect the wavefields and the overlain wavefronts to be in agreement. The most obvious mismatch occurs in regions where the 15° equation fails to extrapolate correctly at steep dips (around \(\gamma = (-20, -50)^\circ\)). This is not surprising since, as its name indicates, this equation is only accurate up to 15°. However, this limitation is eliminated in ray coordinates, because the coordinate system brings the extrapolator in a reasonable position and at a good angle, although the extrapolator uses an equation of a similar order of accuracy.

Another interesting observation in Figures 3 (a) and (d) concerns the diffracon we introduced in the velocity model. When we extrapolate in Cartesian coordinates, the diffraction is only accurate to a small angle relative to the extrapolation direction (vertical). In contrast, the diffraction develops relative to the propagation direction when computed in ray coordinates, thus being more accurate after mapping to Cartesian coordinates.

We can also observe that the diffracons created by the anomaly in the velocity model are not at all limited in the ray coordinates domain. In a beam-type approach, such diffraction would not develop beyond the extent of any particular beam which interacts with it. Neighboring beams would be completely insensitive to the presence of the velocity anomaly.

The second example concerns a smooth velocity with a negative Gaussian anomaly which
Figure 2: Simple linear gradient model: Cartesian coordinates (a,b) and ray coordinates (c,d). Velocity model with an overlay of the ray coordinate system initiated by a point source at the surface (a); image obtained by downward continuation in Cartesian coordinates with the $15^\circ$ equation (b); velocity model with an overlay of the ray coordinate system (c); image obtained by wavefield extrapolation in ray coordinates with the $15^\circ$ equation (d).
Figure 3: Simple linear gradient model: the image obtained by downward continuation in Cartesian coordinates with the 15° equation (a); the image in panel (a) interpolated to ray coordinates (b); image obtained by extrapolation in ray coordinates with the 15° equation (c); the image in panel (c) interpolated to Cartesian coordinates (d).
creates a triplication of the ray coordinate system (Figure 4). Everything other than the velocity model is identical to its counterpart in the preceding example. Similarly to Figure 2, panels (a) and (b) correspond to Cartesian coordinates, and panels (c) and (d) correspond to ray coordinates. Using regularization of the ray coordinates parameters, we are able to extrapolate through the triplication. The small discrepancy between the wavefields and the corresponding wavefronts indicate that our method of ray tracing is not perfectly accurate in the triplicating region, and the wavefield extrapolation is correcting for the kinematic differences. Figure 5 is similar to Figure 3. The “butterfly” in panel (b) is another indication that the ray coordinate system is triplicating, and different shooting directions pick up the same energy from the wavefield extrapolated in Cartesian coordinates (a). None of this happens when we extrapolate in ray coordinates (c) and interpolate to Cartesian coordinates (d).

Our next example concerns the more complicated Marmousi model. Figure 6 shows the velocity models mapped into the two different domains, and the wavefields obtained by extrapolation in each one of them. We create the ray coordinate system by ray tracing in a smooth version of the model, and extrapolate in the rough version. The source is located on the surface at \( x = 5000 \) m.

In this example, the wavefields triplicate in both domains (Figure 7). Since we are using a 15° equation, extrapolation in Cartesian coordinates is only accurate for the small incidence angles, as can be seen in panels (a) and (b). In contrast, extrapolating in ray coordinates (c) does not have the same angle limitation, which can also be seen after mapping back to Cartesian coordinates (d).

Figure 8 is a close-up comparison of the wavefields obtained by extrapolation with different methods in different domains. Panel (a) is a window of the velocity model for reference. Panels (b) and (c) are obtained by extrapolation in ray coordinates using the 15° and split-step equations, respectively. Panels (d), (e) and (f) are obtained by downward continuation in Cartesian coordinates using the 45°, 15° and split-step equations, respectively. The ray coordinates extrapolation results are similar to the Cartesian coordinates results in the regions where the wavefields propagate mostly vertically, but are much better in the regions where the wavefields propagate almost horizontally.

Next, we present another example in the Sigsbee 2A model. We consider two types of ray coordinates: one initiated by a plane wave at the surface (Figure 9), and one initiated by a point source at the surface at \( x = 16000 \) m. (Figure 10). Similarly to the preceding examples, we observe complicated wavefield propagation, with many triplications, of the extrapolated wavefields.

Figure 11 is a close-up comparison of the wavefields obtained by extrapolation with different methods in different domains. Panel (a) is a window of the velocity model for reference. Panel (b) is the wavefield obtained by finite-difference modeling using the two-way acoustic wave equation. This panel contains not only waves propagating forward, but also reflections which are not going to be modeled in the one-way extrapolation results. Panel (c) is obtained by downward continuation in Cartesian coordinates using the 45° equation, and panel (d) is obtained by extrapolation in ray coordinates using the 15° equation. We can observe good match of the forward propagating wavefields in (b) and (d), in contrast to the poor match with
Figure 4: Gaussian anomaly model: Cartesian coordinates (a,b) and ray coordinates (c,d). velocity model with an overlay of the ray coordinate system initiated by a point source at the surface (a); image obtained by downward continuation in Cartesian coordinates with the 15° equation (b); velocity model with an overlay of the ray coordinate system (c); image obtained by wavefield extrapolation in ray coordinates with the 15° equation (d).
Figure 5: Gaussian anomaly model: the image obtained by downward continuation in Cartesian coordinates with the 15° equation (a); the image in panel (a) interpolated to ray coordinates (b); image obtained by extrapolation in ray coordinates with the 15° equation (c); the image in panel (c) interpolated to Cartesian coordinates (d).
Figure 6: Marmousi model: Cartesian coordinates (a,b) and ray coordinates (c,d). Velocity model with an overlay of the ray coordinate system initiated by a point source at the surface (a); image obtained by downward continuation in Cartesian coordinates with the 15° equation (b); velocity model with an overlay of the ray coordinate system (c); image obtained by wavefield extrapolation in ray coordinates with the 15° equation (d).
Figure 7: Marmousi model: the image obtained by downward continuation in Cartesian coordinates with the 15° equation (a); the image in panel (a) interpolated to ray coordinates (b); image obtained by extrapolation in ray coordinates with the 15° equation (c); the image in panel (c) interpolated to Cartesian coordinates (d). [paul2-RCma1.f15.ps] [CR]
Figure 8: Marmousi model: Velocity model (a); image obtained by wavefield extrapolation in ray coordinates using the 15° equation (b) and the split-step equation (c); image obtained using downward continuation in Cartesian coordinates with the 45° equation (d), the 15° equation (e) and the split-step equation (f). 

[Figure images]
Figure 9: Sigsbee 2A model: Cartesian coordinates (a,b) and ray coordinates (c,d). velocity model with an overlay of the ray coordinate system initiated by a plane source at the surface (a); image obtained by downward continuation in Cartesian coordinates with the 15° equation (b); velocity model with an overlay of the ray coordinate system (c); image obtained by wavefield extrapolation in ray coordinates with the 15° equation (d).
Figure 10: Sigsbee 2A model: Cartesian coordinates (a,b) and ray coordinates (c,d). velocity model with an overlay of the ray coordinate system initiated by a point source at the surface (a); image obtained by downward continuation in Cartesian coordinates with the 15° equation (b); velocity model with an overlay of the ray coordinate system (c); image obtained by wavefield extrapolation in ray coordinates with the 15° equation (d).
Figure 11: Sigsbee 2A model: Velocity (a); finite-difference solution to the two-way acoustic wave equation for a point source at $x = 16000$ m (b); image obtained by downward continuation in Cartesian coordinates with the 45° equation (c); image obtained by wavefield extrapolation in ray coordinates with the 15° equation (d). [paul2-RCzg2.zom.ps] [CR,M]
Finally, we present an example of zero-offset migration of overturning reflections using Riemannian wavefield extrapolation. Figure 12 depicts the velocity model (a), the recorded data (b), and the migrated image (c). The overlay is a sketch of the ray coordinate system used for migration. The first event in the data corresponds to the overturning reflection from the ball and is imaged correctly, and the later events are multiple reverberations inside the ball which are not imaged with our method.

Figure 12: Imaging of overturning reflections using Riemannian wavefield extrapolation. Velocity model (a), zero-offset data (b), and migrated image (c).

We can simplify our Riemannian wavefield extrapolation method by dropping the first order terms in equation (12). According to the theory of second-order hyperbolic equations, these terms affect only the amplitude of the propagating waves. To preserve the kinematics, it is sufficient to keep only the second order terms of equation (12):

$$k_\xi = \pm \sqrt{\frac{(\omega s)^2}{c_{\xi\xi}} - \frac{c_{\xi\eta} k_\eta^2}{c_{\xi\zeta}} - \frac{c_{\eta\eta} k_\eta^2}{c_{\zeta\zeta}} - \frac{c_{\xi\eta} k_\xi k_\eta}{c_{\zeta\zeta}}}.$$ (24)

Figure 13 illustrates the difference between wavefield extrapolation using equation (12) (panel b) and wavefield extrapolation using equation (24) (panel c). Kinematically, the two images are equivalent and the main changes are related to amplitudes.
DISCUSSION

We now present a few issues that we think are important for our Riemannian wavefield extrapolation method. In some cases, we discuss ideas which we have not treated yet, while in other cases we speculate on directions of future research.

3-D Riemannian extrapolation: All our examples of Riemannian extrapolation in 2-D using finite-differences use implicit methods, since they are more stable and there is no reason not to use them. In 3-D, however, implicit solutions to the one-way wave-equation become much more difficult, even in Cartesian coordinates (Fomel and Claerbout, 1997; Rickett et al., 1998). This problem seems even more complicated for Riemannian wavefield extrapolation, since the extrapolation equation also contains mixed $k_x k_y$ terms. However, we speculate that for wavefield extrapolation in ray coordinates, this problem is not as difficult as it seems at first sight. The reason is that energy propagates roughly in the forward direction of the coordinate system and, therefore, we do not need extrapolators accurate at high dips. Thus, we could either use explicit methods with reasonably small stencils, or we could use low order mixed-domain methods from the split-step family which are easy to implement even in 3-D.

Prestack data: Our examples of Riemannian wavefield extrapolation are based on equation (12) which corresponds to the single-square root (SSR) equation of standard Cartesian wavefield extrapolation. Riemannian wavefield extrapolation can be extended to prestack data either for shot-profile, plane-wave or S-G migration by appropriate definitions of the underlying ray coordinate system. Figure 14 is a schematic representation of shot-profile migration in ray coordinates, where both source and receivers are extrapolated in the same ray coordinate system appropriate for overturning waves. However, the sources and receivers do not necessarily have to be migrated in the same coordinate system. We could extrapolate both sides differently and apply the imaging condition after interpolation to the Cartesian grid.

Time wave-equation migration: Our Riemannian wavefield extrapolation allows the output
Figure 14: Shot-profile migration sketch. Sources (a) and receivers (b) are extrapolated in the same ray coordinate system which is appropriate for overturning waves.

image to be presented either in one-way traveltime, which is the extrapolation coordinate, or in depth, after interpolation to Cartesian coordinates. A ray coordinate system initiated by a plane wave propagating vertically is related to what has been known in the literature as a “τ coordinate system” (Biondi et al., 1997; Alkhalifah et al., 2001). However, our τ ray coordinate system is different, because it allows energy to move laterally, in contrast to the vertical traveltime coordinate system which does not allow such movement. Thus, another application of Riemannian wavefield extrapolation is time wave-equation migration (Figure 9), which has interesting properties, e.g., for migration velocity analysis (Clapp, 2001). Furthermore, wave-equation MVA (Biondi and Sava, 1999; Sava and Fomel, 2002) could also be reformulated as a function of one-way propagation time.

Regularization at caustics: The coordinate system coefficients for Riemannian wavefield extrapolation given by equations (8) have singularities at caustics, e.g., when the geometrical spreading term $J$, defining a cross-sectional area of a ray tube, goes to zero. In our current examples, we have used simple numerical regularization, by adding a small non-zero quantity to the denominators to avoid division by zero. This strategy worked reasonably well for our current examples.

Adaptive grid: A potential difficulty of our method is represented by the uneven sampling of the wavefronts caused by focusing and defocusing of the rays defining the coordinate system. One solution to this problem is to use adaptive gridding, by increasing or decreasing sampling along the wavefronts, similarly to the techniques employed by the wavefront construction method (Vinje et al., 1993; Qian and Symes, 2002). Furthermore, each frequency in the data can be extrapolated on its own grid, sparser at lower frequencies and denser at higher frequency, thus reducing cost and increasing accuracy.

Interpolation: The images created with wavefield extrapolation in ray coordinates require interpolation to a Cartesian coordinate system. This is a shared difficulty of all methods that do not operate on a Cartesian grid. In our examples, we have successfully used a
simple sinc-type interpolation method. In principle, we could use better interpolation methods using prediction-error filters at higher cost, although we have not seen the need for this in our current examples.

Coordinate system construction: The ray coordinate systems do not need to be created using the same velocity model as the one used for extrapolation. We can use a smooth velocity model to create the coordinate system by ray tracing, and then interpolate the unsmoothed velocity, similarly to the method used by Brandsberg-Dahl and Etgen (2003). Such a strategy opens up the possibility of defining coordinate systems using arbitrary velocity models which favor selected parts of the image. For example, we could use for ray tracing a velocity model optimized to reduce, in a least-squares sense, the angle between the extrapolation grid and the dips in the image. An alternative method of creating ray coordinate systems is discussed by Shragge and Biondi (2003).

Amplitude preservation: Amplitude-preserving imaging using one-way wavefield extrapolation operators is difficult. Recent research has advanced our knowledge on this subject (Zhang et al., 2001; Sava et al., 2001; Shan and Biondi, 2003), but the goal of true-amplitude wave-equation migration is still unachieved. The biggest practical difficulty is associated with amplitude preservation at high scattering angles relative to the extrapolation direction. Since we are normally using low angle operators relative to the wave propagation direction, we speculate that Riemannian extrapolation can also improve the amplitude characteristics of wave-equation migration.

CONCLUSIONS

We extend one-way wavefield extrapolation to Riemannian spaces which are, by definition, described by non-orthogonal curvilinear coordinate systems. We choose semi-orthogonal Riemannian coordinates which include, but are not limited to, ray coordinate systems.

We define an acoustic wave-equation for semi-orthogonal Riemannian coordinates, from which we derive a one-way wavefield extrapolation equation. We use ray coordinates initiated either from a point source, or from an incident plane wave at the surface. Many other types of coordinates are acceptable, as long as they fulfill the semi-orthogonal condition of our acoustic wave equation.

Since wavefield propagation happens mostly along the extrapolation direction, we can use cheap 15° finite-difference or mixed-domain extrapolators to achieve high angle accuracy. If the ray coordinate system overturns, our method can be used to image overturning waves with one-way wavefield extrapolation.

A special case of extrapolation corresponds to coordinates initiated by a plane wave at the surface propagating initially in the vertical direction. Since our extrapolation is done as a function of one-way traveltimes, this case resembles imaging in vertical traveltimes, although it is more physically correct, since it allows lateral movement of energy, which is not the case for vertical τ imaging.
Two main applications of our method are imaging of steeply dipping or overturning reflections.

REFERENCES


Nichols, D. E., 1996, Maximum energy traveltimes calculated in the seismic frequency band: Geophysics, 61, no. 01, 253–263.


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


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


APPENDIX A

WKBJ ASYMPTOTIC SOLUTIONS TO THE ACOUSTIC WAVE-EQUATION IN 3-D RIEMANNIAN SPACES

Neglecting wave propagation in the directions orthogonal to \( \zeta \), one can reduce the wave equation (6) to the form of the ordinary differential equation

\[
\frac{1}{\alpha J} \frac{d}{d \zeta} \left( \frac{J}{\alpha} \frac{d U}{d \zeta} \right) + \frac{\omega^2}{v^2} U = \frac{1}{\alpha J} \frac{d (J/\alpha)}{d \zeta} \frac{d U}{d \zeta} + \frac{1}{\alpha^2} \frac{d^2 U}{d \zeta^2} + \frac{\omega^2}{v^2} U \approx 0 . \tag{A-1}
\]

The high-frequency (WKBJ) asymptotics for the solution of equation (A-1) can be obtained by using a trial solution \( U = A e^{i \omega \phi} \), substituting it in equation (A-1) and evaluating terms with the same order of \( \omega \). The highest asymptotic order yields an equation for the phase function \( \phi \):

\[
- \frac{1}{\alpha^2} \left( \frac{d \phi}{d \zeta} \right)^2 + \frac{\omega^2}{v^2} = 0 . \tag{A-2}
\]

The next asymptotic order produces an equation for the amplitude function \( A \):

\[
\frac{1}{\alpha J} \frac{d (J/\alpha)}{d \zeta} \frac{d \phi}{d \zeta} A + \frac{2}{\alpha^2} \frac{d \phi}{d \zeta} \frac{d A}{d \zeta} + \frac{1}{\alpha^2} \frac{d^2 \phi}{d \zeta^2} A = 0 . \tag{A-3}
\]

Rearranging equation (A-2) to the form

\[
\frac{d \phi}{d \zeta} = \pm \frac{\alpha}{v} \tag{A-4}
\]

and equation (A-3) to the form

\[
\frac{d (\log A)}{d \zeta} = -\frac{1}{2} \left[ \frac{d (\log (J/\alpha))}{d \zeta} + \frac{d (\log (\alpha/v))}{d \zeta} \right], \tag{A-5}
\]

we can solve them explicitly to obtain the WKBJ approximation for the wave traveling preferentially in the \( \zeta \) direction:

\[
U_1 \approx U_0 \left( \frac{v_1 J_0}{v_0 J_1} \right)^{1/2} \exp \left[ \pm i \omega \int_{\xi_0}^{\xi_1} \frac{\alpha}{v} d\zeta \right] \tag{A-6}
\]

In the case of the ray coordinate system, equation (A-6) corresponds to the Green’s function approximation commonly employed in Kirchhoff imaging.

Accounting for the wave propagation in the directions different from \( \zeta \) and constructing the solution numerically by finite differences allows us to account for the finite-bandwidth wave propagation effects.
APPENDIX B

2-D POINT-SOURCE RAY COORDINATES

For the case of 2-D point-source ray coordinates the acoustic wave equation (6) takes the form

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

where, by definition,

\[ \alpha = \sqrt{\left( \frac{\partial z}{\partial \tau} \right)^2 + \left( \frac{\partial x}{\partial \tau} \right)^2} = v, \]
\[ J = \sqrt{\left( \frac{\partial z}{\partial \gamma} \right)^2 + \left( \frac{\partial x}{\partial \gamma} \right)^2}. \]  
(B-2)

The extrapolation axis is \( \tau \) (one-way traveltime from the source) and \( \gamma \) is the shooting angle at the source.

We can expand the parentheses in equation (B-1)

\[ \frac{1}{v^2} \frac{\partial^2 U}{\partial \tau^2} + \frac{1}{vJ} \frac{\partial (J/v)}{\partial \tau} \frac{\partial U}{\partial \tau} + \frac{1}{vJ} \frac{\partial (v/J)}{\partial \gamma} \frac{\partial U}{\partial \gamma} + \frac{1}{J^2} \frac{\partial^2 U}{\partial \gamma^2} = - \frac{\omega^2}{v^2(\tau, \gamma)} U \]  
(B-3)

and make the notations

\[ c_{\tau \tau} = \frac{1}{\alpha^2}, \]
\[ c_{\tau} = \frac{1}{\alpha J} \frac{\partial}{\partial \tau} \left( \frac{J}{\alpha} \right), \]
\[ c_\gamma = \frac{1}{\alpha J} \frac{\partial}{\partial \gamma} \left( \frac{\alpha}{J} \right), \]
\[ c_{\gamma \gamma} = \frac{1}{J^2}, \]  
(B-4)

from which the acoustic wave equation for 2-D point-source ray coordinates becomes:

\[ c_{\tau \tau} \frac{\partial^2 U}{\partial \tau^2} + c_{\tau} \frac{\partial U}{\partial \tau} + c_\gamma \frac{\partial U}{\partial \gamma} + c_{\gamma \gamma} \frac{\partial^2 U}{\partial \gamma^2} = - \frac{\omega^2}{v^2} U. \]  
(B-5)

The 2-D dispersion relation is

\[ -c_{\tau \tau} k_\tau^2 + i c_{\tau} k_\tau + ic_\gamma k_\gamma - c_{\gamma \gamma} k_\gamma^2 = -\omega^2 s^2, \]  
(B-6)

from which we can obtain the one-way wave equation for 2-D point-source ray coordinates:

\[ k_\tau = i \frac{c_\gamma}{2c_{\tau \tau}} \pm \sqrt{\left( \frac{\omega s}{c_{\tau \tau}} \right)^2 - \left( \frac{c_\gamma}{2c_{\tau \tau}} \right)^2 + i \frac{c_{\tau}}{c_{\tau \tau}} k_\gamma - \frac{c_{\gamma \gamma}}{c_{\tau \tau}} k_\gamma^2}. \]  
(B-7)
This appendix details the computations associated with the finite-difference solution to the 15° equation in a 2-D orthogonal Riemannian space. The 3-D wave equation (22) takes in two dimensions the simpler form:

\[ k \xi \approx i \frac{c \xi}{2c \xi \xi} + k_o + \frac{i c \xi}{2c \xi \xi k_o} k_\xi + \frac{1}{2} k_o \left[ \left( \frac{c \xi}{2c \xi \xi k_o} \right)^2 - \frac{c \xi \xi}{c \xi \xi} \right] k_\xi^2. \]  

(C-1)

If we substitute the Fourier-domain wavenumbers by their equivalent space-domain partial derivatives, we obtain

\[ \frac{\partial U}{\partial k_\xi} \approx -\frac{c \xi}{2c \xi \xi} + i k_o + \frac{i c \xi}{2c \xi \xi k_o} \frac{\partial U}{\partial k_\xi} - i \frac{1}{2} k_o \left[ \left( \frac{c \xi}{2c \xi \xi k_o} \right)^2 - \frac{c \xi \xi}{c \xi \xi} \right] \frac{\partial^2 U}{\partial k_\xi^2}. \]  

(C-2)

A finite-difference implementation of equation (C-2) involving the Crank-Nicolson method is

\[ \frac{U_\xi^{t+1} - U_\xi^t}{\Delta \xi} \approx \frac{i c \xi}{2c \xi \xi k_o} \left( U_\xi^{t+1} - U_\xi^t \right) + \left( U_\xi^{t+1} - U_\xi^t \right) - \frac{1}{2} k_o \left[ \left( \frac{c \xi}{2c \xi \xi k_o} \right)^2 - \frac{c \xi \xi}{c \xi \xi} \right] \frac{\Delta \xi^2}{2 \Delta \xi^2}, \]  

(C-3)

If we make the notations

\[ \mu = \frac{i c \xi}{2c \xi \xi k_o} \frac{\Delta \xi}{4 \Delta \xi}, \]  

\[ \nu = -\frac{i}{2} k_o \left[ \left( \frac{c \xi}{2c \xi \xi k_o} \right)^2 - \frac{c \xi \xi}{c \xi \xi} \right] \frac{\Delta \xi}{2 \Delta \xi^2}, \]  

(C-4)

we can write equation (C-3) as

\[ U_\xi^{t+1} - U_\xi^t \approx \mu \left( U_\xi^{t+1} - U_\xi^t \right) + \nu \left( U_\xi^{t+1} - 2U_\xi^t + U_\xi^{t+1} \right) - \nu \left( U_\xi^{t+1} - 2U_\xi^t + U_\xi^{t+1} \right), \]  

(C-5)

or, if we isolate the terms corresponding to the two extrapolation levels as:

\[ U_\xi^{t+1} - \mu \left( U_\xi^{t+1} - U_\xi^{t-1} \right) - \nu \left( U_\xi^{t+1} - 2U_\xi^t + U_\xi^{t+1} \right) = \]  

\[ U_\xi^t + \mu \left( U_\xi^{t-1} - U_\xi^t \right), \]  

(C-6)

After grouping the terms, we obtain

\[ -(\nu - \mu) U_\xi^{t-1} + (1 + 2 \nu) U_\xi^{t+1} - (\nu + \mu) U_\xi^{t+1} = (\nu - \mu) U_\xi^{t-1} + (1 - 2 \nu) U_\xi^{t} + (\nu + \mu) U_\xi^{t+1}, \]  

(C-7)

which is a finite-difference representation of the 15° solvable using fast tridiagonal solvers.
Wavefield extrapolation in phase-ray coordinates

Jeff Shragge and Biondo Biondi

ABSTRACT

A ray theoretic formulation is developed that allows rays to be traced directly from existing solutions to the Helmholtz equation. These rays, termed phase-rays, are defined by the direction normal to surfaces of constant wavefield phase. Phase-rays exhibit a number of attractive characteristics, including triplication-free ray-fields, an ability to shoot rays forward or backward, and an ability to shoot infill rays for ensuring adequate ray density. Because of these traits, we use phase-rays as a coordinate basis on which to extrapolate wavefields using the generalized coordinate system approach. Examples of wavefields successfully extrapolated in phase-ray coordinates are presented, and the merits and drawbacks of this approach, relative to conventionally traced ray coordinates, are discussed.

INTRODUCTION

Ray theory is routinely applied to generate characteristics to solutions of the Helmholtz equation. The usual ray theoretic approach introduces an ansatz representation of the wavefield solution into the Helmholtz equation to yield coupled eikonal and transport equations. Computation of eikonal equation solutions is usually facilitated by the introduction of a high-frequency approximation that removes the amplitude dependence from the full eikonal equation. Consequently, conventionally computed rays are independent of frequency and often triplicate due to their broad-band nature.

In contrast, whenever independent solutions to the Helmholtz equation exist, full ray theoretic formulae may be used to trace rays (Foreman, 1989). These rays are termed phase-rays herein owing to a ray direction that is always orthogonal to surfaces of constant phase. One of their beneficial traits is that, unlike conventionally traced rays, computed ray-fields are caustic-free. A second advantage is that ray position is the only required initial condition for ray-tracing. However, one obstacle seems to restrict the use of the phase-ray formulation in seismic imaging problems: wavefield solutions to the Helmholtz equation must be known in advance of ray-field computation.

One situation where phase-rays (and conventional rays) are of use is in wavefield extrapolation in generalized (i.e. non-Cartesian) coordinates (Sava and Fomel, 2003). A natural set of coordinates for wavefield extrapolation is the general ray family represented by a ray

1email: jeff@sep.stanford.edu, biondo@sep.stanford.edu
direction and shooting angle(s). Wavefield extrapolation in ray coordinates thus uses a ray-
field as the coordinate system on which to extrapolate wavefields. This approach transforms
the physics of one-way wave propagation from the usual Cartesian grid so that it is valid in
a ray coordinate system. Wavefield extrapolation operators are then applied more accurately
because extrapolation can occur at lower angles to the ray direction than usually occurs with
Cartesian coordinates. The result is then mapped from ray coordinates to Cartesian. Although
ray domain extrapolated wavefields are generally more accurate, issues still remain when using
conventionally-traced rays; in particular, how to robustly deal with infinite amplitudes at
locations of coordinate system triplication.

Motivated by this issue, this paper examines the use of phase-rays as a coordinate system
for wavefield extrapolation. The main advantage of phase-ray coordinates is that they are
caustic-free, and thereby avoid complications arising from triplicating coordinates. The paper
begins with a general discussion of ray theory and an approach for calculating phase-rays
from solutions to the Helmholtz equation. Phase-ray examples from a salt body model are
then presented, and are followed by results illustrating wavefield extrapolation in phase-ray
coordinates for a Gaussian velocity perturbation model. Finally, a method is proposed for
propagating broadband wavefields using frequency-dependent phase-ray coordinates.

THEORY

The theory outlined in this section closely follows that of Foreman’s exact ray theory (Fore-
man, 1989), but is summarized here for completeness. Ray theory may be used to compute
the characteristics to the time-independent, homogeneous Helmholtz equation,
\[ \nabla^2 \Psi + k^2 \Psi = 0, \]  

where \( \Psi \) is the desired wavefield solution, and \( k \) is the wavenumber. In most ray theoretic
developments, the wavefield is represented by an ansatz solution, \( \Psi = A e^{i\phi} \), where \( A(r) \) and
\( \phi(r) \) are the amplitude and phase functions, respectively. Substituting this representation into
the Helmholtz equation yields the usual eikonal and transport equations,
\[ K^2 = K \cdot K = \nabla \phi \cdot \nabla \phi = k^2 + \frac{\nabla^2 A}{A}, \]  
\[ 2\nabla A \cdot \nabla \phi + A \nabla^2 \phi = 0, \]

where \( K \) is the phase gradient vector (see figure 1). Solutions to equations (2) are the ray
paths and the amplitudes along these ray paths, respectively. In isotropic media the gradient
of the phase function, \( K = \nabla \phi \), is orthogonal to surfaces of constant phase and represents
the instantaneous direction of propagation. Explicitly, this may be written,
\[ K = \nabla \phi = |K| \frac{d}{ds} r, \]  

where \( ds \) is an element of length, and \( dr \) is a vector element of the ray-path. A ray-path
equation is developed by taking the gradient of the phase gradient magnitude (i.e. \( \nabla K \)),
\[ K^2 = K \cdot K, \]
Figure 1: Schematic of phase-front gradient quantities. $\mathbf{K}$ is the phase gradient vector, $|\mathbf{K}|ds$ is the gradient magnitude along step $ds$, and $|\mathbf{K}|dx$ and $|\mathbf{K}|dz$ are the projections of $|\mathbf{K}|ds$ along the $x$ and $z$ coordinates, respectively.

\[ 2\mathbf{K} \nabla \mathbf{K} = 2(\mathbf{K} \cdot \nabla)\mathbf{K} + 2\mathbf{K} \times (\nabla \times \mathbf{K}), \]
\[ \nabla \mathbf{K} = \left[ \frac{\mathbf{K}}{|\mathbf{K}|} \cdot \nabla \right] \mathbf{K}, \quad (4) \]
\[ = \left( \frac{d}{ds} \mathbf{r} \cdot \nabla \right) \mathbf{K} = \frac{d}{ds} \mathbf{K}, \]

where $\nabla \times \mathbf{K} = \nabla \times \nabla \phi = 0$ has been employed. Using equation (3) this may be written explicitly,
\[ \nabla \mathbf{K} = \frac{d}{ds} \left( \mathbf{K} \frac{d}{ds} \mathbf{r} \right) \quad (5) \]

Coupling between the eikonal and transport equations is evident through the dependence of the eikonal equation on amplitude function, $A$. In many cases a high-frequency approximation (i.e. $\frac{\nabla^2 A}{A} \approx 0$) is used to decouple these equations. Use of this approximation yields the usual form of the ray path equation,
\[ \frac{d}{ds} \left( \frac{1}{c} \frac{d}{ds} \mathbf{r} \right) = \nabla \left( \frac{1}{c} \right), \quad (6) \]

where $c$ is the velocity. Use of this approximation also eliminates the frequency dependence of ray trajectories (see Appendix A). One manner of reintroducing frequency-dependent ray trajectories is discussed in a latter section.

**Phase-ray formulation**

When a solution, $\Psi$, to the Helmholtz equation is known, obtaining a ray trajectory using equation (5) is relatively trivial. The expression for the wavefield gradient, $\nabla \Psi = \nabla (A e^{i\phi})$, divided by the wavefield, $\Psi$, is,
\[ \frac{\nabla \Psi}{\Psi} = \frac{\nabla A}{A} + i\mathbf{K}. \quad (7) \]

An expression for the wavefield gradient vector, $\mathbf{K}$, is obtained by retaining the imaginary component of equation (7) and using the expression for $\mathbf{K}$ in equation (3),
\[ K \frac{d}{ds} \mathbf{r} = \text{Im} \left( \frac{\nabla \Psi}{\Psi} \right) \quad (8) \]
Equation (8) may be rewritten explicitly as a system of two decoupled ordinary differential equations,

\[
\frac{d}{ds} \begin{bmatrix} x \\ z \end{bmatrix} = \frac{1}{\sqrt{\left(\frac{d\Psi}{dx}\right)^2 + \left(\frac{d\Psi}{dz}\right)^2}} \text{Im} \left( \frac{1}{\Psi} \left[ \frac{\partial}{\partial x} \left( \frac{\partial}{\partial z} \right) \Psi \right] \right). \tag{9}
\]

The solution for ray-path, \( r \), is computed through an initial evaluation of the right hand side of equations (9), and an iterative forward step by a constant interval using the precomputed quantity to determine the proper apportioning of the step along each coordinate. Note that because these differential equations are first-order, only one initial condition (position) is required for ray computation, and rays may be started from any location in the wavefield solution. Finally, although the two-dimensional formulation is presented here, the extension to three dimensions is trivial.

PHASE-RAY EXAMPLES

Examples of traced phase-rays are presented in this section using a salt body velocity field as a didactic model. The background velocity of the model, shown in figure 2(e), is a typical Gulf of Mexico \( v(z) \) velocity gradient. The superposed salt body is characterized by higher wave speeds (4700 m/s) and a fairly rugose bottom of salt interface.

Figure 2 presents five phase-rays computed from four different monochromatic wavefields. The wavefields were generated for a shot point located at 11700 m using a split-step Fourier operator (Stoffa et al., 1990) in a Cartesian coordinate system. Each ray begins at the same point in all panels. The rays to the extreme left and right in each panel show little variability in their spatial location; however, the three remaining rays are attracted to regions of greater wavefield amplitude and their spatial locations vary with a range up to 2000 m. Accordingly, because rays originate at the same spot, observed phase-ray movement is caused by changes in wavefield solution and indicates frequency-dependent behavior.

Phase-rays computed according to equations (9) may be traced in reverse, from observation to source point, by using a negative step interval. Figure 3 illustrates this situation with the same model as in figure 2. Initial ray locations are points at regular intervals on a semicircular arc of radius 5000 m. Calculated phase-rays do not overlap and the ray-field is caustic-free. Phase-ray density, though, is frequency-dependent, with significant coverage gaps of variable size appearing in all four panels. This suggests that an additional condition is required to ensure that, when needed, ray density is more uniform. One solution is to shoot a new ray between two successive rays wherever intra-ray spacing is too large. These three traits provide the main impetus for using phase-ray coordinates as a generalized coordinate system for wavefield extrapolation (Sava and Fomel, 2003).
Figure 2: Five phase-rays traced through monochromatic wavefields. Wavefields were generated by a split-step Fourier operator in Cartesian coordinates for a shot point at 11700 m using the velocity model illustrated in e). a) 5 Hz wavefield; b) 10 Hz wavefield; c) 15 Hz wavefield; d) 20 Hz wavefield; and e) smoothed salt body model velocity field.
Figure 3: 200 phase-rays traced backwards through monochromatic wavefields for velocity model shown in e). Initial ray locations are points at regular intervals on a semicircular arc of radius 5000 m. a) 5 Hz wavefield; b) 10 Hz wavefield; c) 15 Hz wavefield; d) 20 Hz wavefield; and e) smoothed salt body model velocity field.
WAVEFIELD EXTRAPOLATION IN PHASE-RAY COORDINATES

This section examines the use of phase-rays as a coordinate system for wavefield extrapolation. The velocity model examined here (shown in figure 4a) is characterized by a slow, Gaussian-shaped velocity anomaly (-600 m/s) in a medium of otherwise constant velocity (1200 m/s). This model was chosen to ensure that extrapolated wavefields triplicate, as illustrated by figure 4b. This wavefield was generated for a shot point located at 8000 m using a split-step Fourier operator in a Cartesian coordinate system. Figure 4c shows phase-rays traced through the wavefield of figure 4b. Phase-rays in the upper portions of the model have fairly smooth coverage. In areas of wavefield triplication, though, significant coverage gaps are noticeable.

The phase-rays shown in figure 4c were subsequently used as a coordinate system for the generalized coordinate wavefield extrapolation approach (Sava and Fomel, 2003). Figure 5a shows the velocity model of figure 4a in phase-ray coordinates. Figure 5b presents the results of wavefield extrapolation in phase-ray coordinates using a split-step Fourier operator and the velocity model shown in figure 5a. Figure 5c shows this result mapped to Cartesian coordinates. The wavefield presented in figure 5d was computed in Cartesian coordinates using a split-step Fourier operator. In areas where wavefields are present in significant amplitude, the phase-ray and Cartesian results are similar. Areas of low wavefield amplitude beneath the Gaussian velocity anomaly (e.g. [z,x]=[4200 m,6800 m]), though, are markedly different. This difference is related to the inability to map the results from ray coordinates to Cartesian in areas of minimal or non-existent ray coverage.

This experiment highlights a consequence of using phase-ray coordinates for wavefield extrapolation. Monochromatic wavefield triplication is generally identified by interference patterns created by converging wavefield components. (See, for example, the checkerboard pattern beneath the Gaussian velocity anomaly.) Because phase-ray direction is dependent on the total wavefield gradient, it is similarly dependent on the gradients of each converging wavefield component. The gradient vector, being unable to unwrap individual convergent phases, chooses a weighted average of individual gradients. Accordingly, phase-rays are usually steered in the direction of the convergent component with the largest individual gradient magnitude, but they will never triplicate since the weighted gradient is uniquely defined at each wavefield point. This fact suggests that phase-ray coordinates represent a trade off between introducing inaccuracy associated with triplicating coordinates and inaccuracy of wavefield extrapolation at greater angles to the phase-ray direction.

Figure 6 presents a comparison between wavefields extrapolated in phase-ray and conventional ray coordinates (Sava and Fomel, 2003). Figures 6a and 6b present wavefields extrapolated in phase-ray coordinates and after interpolation into Cartesian coordinates, respectively. Figures 6c and 6d present similar results, but with conventionally traced rays. Of the two ray coordinate systems tested, the phase-ray coordinate extrapolated wavefield (figure 6b) better resembles the wavefield calculated in Cartesian coordinates (figure 5c). However, the sampling of phase-ray extrapolated wavefields, and their Cartesian maps, must be greatly improved before a definitive comparison is possible.
Figure 4: Gaussian-shaped velocity anomaly model. a) Gaussian-shaped anomaly of -600 m/s maximum velocity perturbation superposed on a constant 1200 m/s velocity field; b) 5 Hz wavefield computed for a shot located at 8000 m using a single-velocity split step Fourier operator in Cartesian coordinates; and c) phase-rays traced through the wavefield of b). [jeff1-modelrays] [ER]
Figure 5: A comparison of wavefield extrapolation results computed in phase-ray and Cartesian coordinates. a) Velocity model of figure 4a mapped to phase-ray coordinates; b) wavefield extrapolated in phase-ray coordinates using a split-step Fourier method; c) the map of wavefield in b) to Cartesian coordinates; d) wavefield solved in Cartesian coordinates using a split-step Fourier operator. [jeff1-Reg.ER1.ssf.ps] [ER]
Figure 6: A comparison of wavefield extrapolation results computed in phase-ray and conventional ray coordinates. a) wavefield extrapolated in phase-ray coordinates; b) wavefield of a) interpolated into Cartesian coordinates; c) wavefield extrapolated in conventional ray coordinates (Sava and Fomel, 2003); and d) wavefield of c) interpolated in Cartesian coordinates.
Toward broadband propagation

The goal of using phase-rays as a coordinate system for wavefield extrapolation is to enable a more accurate propagation of broadband wavefields through the subsurface. However, this requires extrapolating wavefields at many frequencies. One interesting observation is that phase-rays exhibit frequency-dependent behavior. Thus, one strategy for wavefield extrapolation is to use phase-ray coordinates that adapt to frequency-dependent illumination.

The proposed approach for frequency-dependent wavefield extrapolation is as follows: i) an initial wavefield extrapolation in Cartesian or polar coordinate at the lowest frequency; ii) phase-ray tracing using the current wavefield solution; iii) extrapolating the next wavefield using the traced phase-rays as the new coordinate system; iv) a mapping of the latest wavefield result from phase-ray to Cartesian coordinates; and v) repeating steps ii), iii) and iv) until wavefields at all frequencies are calculated. The broadband wavefield is then obtained by a summation of wavefields over all extrapolated frequencies. Alternatively, because frequency-dependent wavefield illumination usually varies slowly over individual frequency steps, phase-rays could be traced only periodically to save computational cost.

CONCLUSIONS

This paper has introduced a method for tracing phase-rays through monochromatic wavefield solutions of the Helmholtz equation. The resulting phase-rays exhibit a number of attractive characteristics, including: i) a triplication-free ray-field; ii) an ability to shoot rays forward or backward from areas of strong or weak wavefield amplitude alike; and iii) an ability to easily infill rays to ensure adequate phase-ray density. Phase-rays may then be successfully used as a coordinate system on which to extrapolate wavefields. These coordinates avoid coordinate system triplication that can debilitate wavefields extrapolated using conventional ray-field coordinates.

The phase-ray formulation, though, cannot unwrap individual triplicating phases, and chooses a weighted average between interfering phases. Because of this fact phase-ray coordinates represent a trade off between introducing inaccuracy associated with triplicating coordinates and inaccuracy of wavefield extrapolation at greater angles to the ray direction. However, before a critical comparison of the relative merits and drawbacks of phase-ray and conventional ray coordinate extrapolated wavefields is attempted, phase-ray extrapolated wavefields need to be better sampled so that their Cartesian maps are more comparable.

ACKNOWLEDGMENTS

The first author would like to thank Paul Sava for many helpful discussions and the use of his computer programs.
REFERENCES


APPENDIX A

FREQUENCY DEPENDENCE OF THE RAY PATH EQUATION

The frequency-dependent eikonal equation is,

\[ K^2 = k^2 + \frac{\nabla^2 A}{A} = \nabla \phi \cdot \nabla \phi, \quad (A-1) \]

where \( A \) and \( \phi \) are the amplitude and phase functions, respectively. Defining parameter \( \gamma = \frac{\nabla^2 A}{A} \) and using \( k = \omega s \), where \( \omega \) is angular frequency and \( s \) is slowness, one may use the definition of \( K \) in equation (A-1) to write the following as the ray path equation (see equation (4))

\[ \nabla K = \frac{\omega s \nabla s + \nabla \gamma}{\sqrt{\omega^2 s^2 + \gamma}} = \frac{d}{ds} \left( K \frac{d}{ds} r \right). \quad (A-2) \]

To examine how equation (A-2) varies as a function of frequency, a frequency derivative is applied to yield,

\[ \frac{\partial}{\partial \omega} (\nabla K) = \frac{\gamma \omega s \nabla s - \frac{\partial}{\partial \omega} \nabla \gamma}{\omega^2 s^2 + \gamma} \left( \frac{\omega s^2 + \gamma}{\omega^2 s^2 + \gamma} \right)^{\frac{3}{2}}. \quad (A-3) \]

High frequency approximation

The ray theoretic ansatz assumes that \( K \approx k = \omega s \), which involves setting the contribution of \( \gamma \) and its derivatives to zero. Thus, the frequency dependence of the ray path equation is,

\[ \frac{\partial}{\partial \omega} (\nabla K) = 0, \quad (A-4) \]

which requires that the ray paths are stationary with respect to frequency.
Amplitude balancing of 3-D angle-domain common-image gathers

Biondo Biondi

ABSTRACT
The azimuthal resolution of 3-D Angle Domain Common Image Gathers (ADCIGs) strongly varies with the reflection aperture angle. This dependence may cause severe distortions in the image when 3-D ADCIGs are averaged over azimuths. To correct for these distortions, I derive an effective weighting method based on the jacobian of the transformation to angle domain. The proposed method avoids the underweighting of the reflections close to normal incidence by properly taking into account the “folding” of the azimuth axis. A simple scheme to limit the range of the azimuthal averaging as a function of the opening angle further attenuates noise in the image. A synthetic example illustrates the practical application of the proposed methodology.

INTRODUCTION
Tisserant and Biondi (2003) presented a method to create Angle Domain Common Image Gathers (ADCIGs) in 3-D. In 3-D ADCIGs the image is decomposed at each physical location \((x, y, z)\) depending on the aperture angle \(\gamma\) and the azimuth \(\phi\) of the reflections. Given the limited azimuthal range of many common acquisition geometries (e.g. marine streamer data), it is often useful to average the ADCIGs over azimuths and to limit the azimuthal average to a subrange of the possible azimuths. These procedures can attenuate coherent noise that was recorded in the data (e.g. multiples) and/or caused by computational shortcuts (Biondi, 2003). However, because of the variable resolution of the angle decomposition in the azimuthal direction, the averaging over azimuths may cause distortions in the amplitudes and the phases of the final image.

In this report I address the problems of 1) balancing the amplitudes across aperture angles while performing the stack over azimuths, 2) determining an “optimal” azimuthal subrange as a function of the aperture angle. Both of these problems are related to the strong dependence of the azimuthal resolution with aperture angle. The azimuthal resolution decreases as the aperture angle gets closer to normal incidence; at the limit, all azimuths are equally illuminated at normal incidence. To preserve the relative amplitudes between the whole range of aperture angles, we need to introduce a proper weighting factor when stacking the ADCIGs over azimuths. If no normalization factor is applied during the summation, the reflections with

\(^1\)email: biondo@sep.stanford.edu
angles close to normal incidence would totally overshadow the reflections with wider aperture angles. This normalization factor must be obviously based on the jacobian of the transformation into angle domain, but a straight application of the jacobian would underweight the reflections close to normal incidence. To avoid such a problem, I define a weighting method that takes into account the “wrapping” of the azimuth axis close to normal incidence.

**AMPLITUDE CORRECTIONS OF THE MAPPING INTO ANGLE DOMAIN**

Tisserant and Biondi (2003) showed that in 3-D the transformation to angle domain is accomplished by mapping the image-domain offset wavenumber plane \((k_{xh}, k_{yh})\) into the plane defined by the reflection aperture angle \(\gamma\) and the reflection azimuth \(\phi\). This mapping is performed at fixed depth wavenumber \(k_z\) and midpoint wavenumbers \(k_{xm}\) and \(k_{ym}\).

Appendix A provides the analytical form of this transformation. The analytical form is fairly intricate and not easy to interpret because the reflection azimuth \(\phi\) enters only indirectly as a parameter for rotating the midpoint and offset wavenumbers. Figure 1 shows a graphical representation of the mapping to angle domain, and thus it illustrates the problem that I am addressing in this paper more intuitively than the formulas. The figure shows a uniformly sampled Cartesian grid in the \((\gamma, \phi)\) domain mapped into the \((k_{xh}, k_{yh})\) plane. Each dot corresponds to one value of \(\gamma\) and \(\phi\), for fixed \(k_{zm}\), \(k_{xm}\) and \(k_{ym}\). The ranges for \(\gamma\) and \(\phi\) are: \(-80^\circ \leq \gamma \leq 80^\circ \) and \(-30^\circ \leq \phi \leq 30^\circ\). The dots are densely clustered close to the origin (corresponding to small values of \(\gamma\)), and become sparse away from the origin (corresponding to large values of \(\gamma\)). If this strong variability in the mapping density were not taken into account when a summation is performed in the angle domain, the summation would result in a strongly distorted image.

In this paper I consider the effects of averaging the image over reflection azimuth because it is the most challenging and interesting situation, as we will see briefly. However, similar considerations are needed when summing over aperture angles (for example when computing the “stacked” image after the application of a residual moveout correction to improve reflectors’ coherency).
The effects of ignoring the variability in the mapping density are demonstrated in Figure 2. This figure shows two ADCIGs obtained by imaging a synthetic data set with a full source-receiver 3-D prestack migration. The data set contains 5 dipping planes, from zero dip to 60 degrees dip. The azimuth of the planes is 45 degrees with respect to the direction of the acquisition. The velocity was $V(z) = 1.5 + .5z$ km/s, which corresponds to the upper limit among the typical gradients found in the Gulf of Mexico. The acquisition geometry had one single azimuth and the source-receiver offset range was -1.6–3 km. More detailed description of this data set can be found in (Vaillant and Biondi, 2000; Biondi, 2001, 2003).

Figure 2a shows an ADCIG computed at fixed reflection azimuth $\phi$ of 16 degrees. This azimuth corresponds – only approximately, because the reflection azimuth changes with the aperture angle – to the reflection azimuth for the deepest reflector. Correspondingly, the deepest reflector has a flat moveout along the aperture angles, but the shallower ones are frowning downward. Figure 2b shows the result of averaging over azimuths all the ADCIGS within the range of $-60^\circ \leq \phi \leq 60^\circ$. In this case the moveouts are flat for all the reflection angles larger than 10 degrees, but the amplitude of the image is strongly attenuated for all these angles. This distortion of the image amplitudes is caused by the variable density of the mapping from the $k_{xh}, k_{yh}$ plane into the $(\gamma, \phi)$ plane illustrated in Figure 1.

The solution to this problem seems straightforward. We can correct the amplitudes by applying the jacobian of the transformation from the $k_{xh}, k_{yh}$ plane into the $(\gamma, \phi)$ plane (Appendix A presents the formulas to evaluate the jacobian). However, while this correction yields a much improved result, it is not sufficient. Figure 3a shows the effect of including the jacobian while transforming the image into angle domain. The amplitudes of the ADCIG are now distorted in the opposite direction of the previous result (Figure 2b). Now the wide aperture angles have a good amplitude response, but the narrow angles have been too strongly attenuated and the $\gamma = 0$ trace has been zeroed. This behavior is simply explained by the fact that the jacobian is zero at $\gamma = 0$. This singularity of the jacobian is graphically represented in Figure 1 by the fact that all the dots corresponding to $\gamma = 0$ fall into the origin of the plane, where the dot density becomes effectively infinite.

A simple solution to this problem is suggested when we examine the inverse mapping, as it is graphically illustrated in the sketch in Figure 4. In this case an integration segment
along the line defined by constant $\gamma$ and close to the origin of the $(k_{xh}, k_{yh})$ plane would expand into a long segment extending well beyond the usual range of $-180^\circ \leq \phi \leq 180^\circ$. However, because of the periodicity of the mapping, the segment is actually folded into the $-180^\circ \leq \phi \leq 180^\circ$, and its effective length is limited to 360 degrees ($2\pi$). Taking into account of this fact, we can correct the jacobian weighting and recover the image amplitudes close to normal incidence. The details of the correction and its derivation are presented in Appendix A.

The effects of taking into account the folding of the azimuth axis in the mapping are demonstrated in Figure 3b. Now the angles close to near incidence have been properly imaged. Figure 5 shows windows of the ADCIGs shown in Figure 3 with narrower aperture angle ranges. The comparison of Figure 5a and Figure 5b demonstrates the improvements achieved by taking into account the folding of the azimuth axis.

Figures 6–7 illustrate the effect of the amplitude correction from another point of view. They show depth slices taken at the depth of 1,140 meters (corresponding to the reflector dipping at 45 degrees) before the stacking over azimuths. The reflection amplitudes are thus shown as function of both the aperture angle $\gamma$ and the azimuth $\phi$. Because of the poor azimuthal resolution close to normal incidence, the azimuthal range is wide for small $\gamma$; it narrows as $\gamma$ increases. Comparing Figure 6 with Figure 7 it is evident that the amplitude correction boost up the relative amplitudes of the image at large $\gamma$.

Relative phase shift across aperture angles ($\gamma$)

A close examination of both Figure 3 and Figure 5 reveals that the smaller aperture angles ($\gamma \leq 8$) have a slight phase shift with respect to the rest of the aperture angles. A “frowning” artifact that is related to this phase shift is also visible in the gathers. This artifact is attenuated by limiting the azimuthal range of the summation, as described in the next section. Comparison of Figures 3–5 with Figure 2a also suggests that the phase shift is caused by the stacking over azimuths, since it is absent in the gather shown in Figure 2a.

The phase shift is indeed related to the summation over azimuth and it is easily explained by the analysis of the image as a function of the reflection azimuth $\phi$ and at constant aperture angle $\gamma$. The three panels in Figure 8 show such sections for three different aperture angles: a) $\gamma = 4^\circ$, b) $\gamma = 20^\circ$, and c) $\gamma = 35^\circ$. The curvature of the reflectors as a function of the azimuth is different in Figure 8a from both Figure 8b and Figure 8c. These differences in the curvature of the reflector cause the relative phase shift of the stacked gathers shown in Figures 3–5. In other words, the phase shift at large aperture angles ($\gamma > 8$) is caused by the interference of the flanks of the hyperbolic curves shown in Figures 3–5 with the correct summation of the flat spots of the same hyperbolic curves.

Fortunately, both the phase shift and the “frowning” artifact are related to the illumination of the reflectors by the data (common-azimuth acquisition geometry) and they are not caused by the methodology employed to image the data. As the azimuthal range of the data increases, the flat spots at the top of the hyperbolic curves shown in Figures 3–5 should widen. Consequently, when stacking over azimuths the influence of the flanks should decrease relative to
Figure 3: ADCIGs for a synthetic data set. Left: Image obtained when the simple jacobian weighting is applied before averaging over azimuths. Right: Image obtained when the jacobian weighting takes into account the folding of the azimuth axis.

Figure 4: Graphical representation of the stretching involved in the mapping from the \((k_{xh}, k_{yh})\) plane into the \((\gamma, \phi)\) plane.

Figure 5: Zoom into the ADCIGs shown in Figure 3 to examine the differences between the two panels close to normal incidence.
Figure 6: Depth slice taken at the depth of 1,140 meters (corresponding to the reflector dipping at 45 degrees) before the stacking over azimuths. Notice that the azimuthal resolution is strongly dependent on the aperture angle.

Figure 7: Depth slice taken at the same depth as the slice shown in Figure 6 (z=1,1140 meters) after application of the proposed angular dependent weighting. Notice that the amplitudes close to normal incidence have been attenuated, but not zeroed, and the ones at large aperture angle have been boosted up.
the influence of the flat spot, and the phase shift should disappear. It would be interesting to confirm this hypothesis with a real data example from a marine data set.

**Determining the azimuthal range**

It is often useful to limit the azimuthal range of the image. Given the strong dependence of the azimuthal resolution with aperture angle, it is natural to make the bounds of the azimuthal range functions of the aperture angle $\gamma$. Any smooth function of $\gamma$ is probably adequate to define the azimuthal boundaries. I decided to apply a trigonometric function to transition between the limits at normal incidence and the limits at 90 degrees. If $\gamma_{\text{max}}$ is the maximum aperture angle, $\phi_{\text{min}}^0$ and $\phi_{\text{max}}^0$ are respectively the minimum and maximum azimuth angles for $\gamma = 0$ degrees, and $\phi_{\text{min}}^{90}$ and $\phi_{\text{max}}^{90}$ are respectively the minimum and maximum azimuth angles for $\gamma = 90$ degrees, then I set the azimuthal limits $\phi_{\text{min}}^\gamma, \phi_{\text{max}}^\gamma$, by applying the following expressions:

\[
\begin{align*}
\phi_{\text{min}}^\gamma &= \phi_{\text{min}}^{90} + (\phi_{\text{min}}^0 - \phi_{\text{min}}^{90}) \left[ \sin \left( \frac{90 - \gamma}{90} \right) \right]^p, \\
\phi_{\text{max}}^\gamma &= \phi_{\text{max}}^{90} + (\phi_{\text{max}}^0 - \phi_{\text{max}}^{90}) \left[ \sin \left( \frac{90 - \gamma}{90} \right) \right]^p,
\end{align*}
\]

where $p$ is a free parameter that determines the shape of the azimuthal window.

Figure 9 is a generalization of Figure 1. The red (darker in gray scale) dots are the same as in Figure 1, and represent the mapping into angle domain with $\phi_{\text{min}}^0 = \phi_{\text{min}}^{90} = -30$ degrees, and $\phi_{\text{max}}^0 = \phi_{\text{max}}^{90} = 30$ degrees. The green (lighter in gray scale) is the mapping when the azimuthal range is broader close to normal incidence, that is with $\phi_{\text{min}}^0 = -90$ degrees, and $\phi_{\text{max}}^0 = 90$ degrees, and $p = 3$. Notice that the integration domain represented by the green dots (lighter in gray scale) does not shrink close to the origin, as the original integration domain does.

Figure 10 shows the effect of the variable azimuthal range on the synthetic data set. Figure 10b shows the same ADCIG as in Figure 3b. The azimuthal range was constant over $\gamma$; that is $\phi_{\text{min}}^0 = \phi_{\text{min}}^{90} = -60$ degrees, and $\phi_{\text{max}}^0 = \phi_{\text{max}}^{90} = 60$ degrees. Figure 10a shows the ADCIG extracted at the same location as the one in Figure 10b, but obtained with variable...
azimuthal range. The parameters were $\phi_{\min}^0 = -60$ degrees, $\phi_{\min}^{90} = -5$ degrees, $\phi_{\max}^{90} = 25$ degrees, $\phi_{\max}^0 = 60$ degrees, and $p = 3$. The reduction in the azimuthal range attenuates the numerical noise in the image. In particular, it attenuates the “frowning” artifacts that, as I discussed in the previous section, are related to the narrow azimuthal coverage of the data.

Figure 11 shows a depth slice extracted from the image at the same depth as the slices shown in Figures 6–7, but with the variable azimuthal range defined by the parameters listed above. The comparison of Figure 11 with Figure 7 demonstrates that the window defined using the relationships (1) and (2) preserves the coherent energy of the event, while removing noise.

Figure 9: Graphical representation of the effects of the variable azimuthal range on the mapping from the offset wavenumber $(k_{xh}, k_{yh})$ plane into the $(\gamma, \phi)$ plane. The green (lighter in gray scale) dots correspond to the mapping with variable azimuthal range.

Figure 10: ADCIGs for a synthetic data set. Left: Image obtained after application of both the angular dependent weighting and the variable azimuthal range. Right: Image obtained after application of the angular dependent weighting.

CONCLUSIONS

To avoid distortion in the image, 3-D ADCIGs must be properly weighted before averaging over azimuths. These weights can be derived from the jacobian of the transformation to angle domain, after taking into account the folding of the azimuth axis. The weighting method I presented was successful to correct the image obtained from a synthetic data set.
To attenuate both coherent noise and numerical artifacts, I also defined a simple scheme for limiting the azimuthal range of the image. This variable azimuthal range proved to be useful in reducing the artifacts caused by the limited azimuthal coverage of the synthetic data. Further analysis of the artifacts caused by the azimuthal coverage of real acquisition geometries might be required to make the computation of 3-D ADCIGs a robust tool for velocity analysis and amplitude analysis.

REFERENCES


This appendix derives the expressions for the weights to be applied to the ADCIGs before averaging over azimuths. These weights are based on the jacobian of the transformation into angle domain. The first step is therefore to find the expressions for evaluating this jacobian.

The starting point for computing the jacobian is the transformation into angle domain. Tisserant and Biondi (2003) showed that 3-D ADCIGs can be computed according to the following mappings:

\begin{align}
  k'_{xh} &= -\tan \gamma \sqrt{k'^2_{xm} + k'^2_z}, \\
  k'_{yh} &= -\frac{k'_{ym} k'_m k'_{zh}}{k'^2_z + k'^2_{ym}},
\end{align}

where the primes on the wavenumber indicate the rotation of the coordinate axis by $\phi$ according to the following relationships:

\begin{align}
  k'_{xm} &= \cos \phi k_{xm} - \sin \phi k_{ym}, \\
  k'_{ym} &= \sin \phi k_{xm} + \cos \phi k_{ym},
\end{align}

and similarly

\begin{align}
  k'_{xh} &= \cos \phi k_{xh} - \sin \phi k_{yh}, \\
  k'_{yh} &= \sin \phi k_{xh} + \cos \phi k_{yh}.
\end{align}

We need to compute the partial derivatives of the offset wavenumbers at constant aperture angle $\gamma$. Therefore, we start from rewriting the coplanarity condition in equation (A-2) in terms of reflections angles in the rotated coordinate system: the aperture angle $\gamma'$, the in-line dip angle $\alpha'_{x}$, and the cross-line dip angle $\alpha'_{y}$. The following relationships link the wavenumber in the image domain to these angles

\begin{align}
  \tan \alpha'_{x} &= \frac{k'_{xm}}{k'_z}, \\
  \tan \alpha'_{y} &= \frac{k'_{ym}}{k'_z},
\end{align}

and

\begin{align}
  \tan \gamma' &= -\cos \alpha'_{y} \frac{k'_{xh}}{k'_z}.
\end{align}

Then equation (A-2) becomes:

\begin{align}
  k'_{yh} &= k'_z \tan \gamma' \tan \alpha'_{x} \sin \alpha'_{y},
\end{align}

and equation (A-1) becomes:

\begin{align}
  k'_{xh} &= -k'_z \frac{\tan \gamma'}{\cos \alpha'_{y}}.
\end{align}
Evaluation of $\partial k'_{x_h}/\partial \phi$ and $\partial k'_{y_h}/\partial \phi$

Differentiating equation (A-12) and after some algebraic manipulation, we obtain the following:

$$\frac{\partial k'_{x_h}}{\partial \phi} = -k_z \tan \gamma \frac{\partial \left( \frac{1}{\cos \alpha'_{y}} \right)}{\partial \phi} = -k_z \tan \gamma \frac{\partial \left( \frac{1}{\cos \alpha'_{y}} \right)}{\partial k'_{y_m}} \frac{\partial k'_{y_m}}{\partial \phi} = -k_z \tan \gamma \frac{\sin \alpha'_{y} \partial k'_{y_m}}{k_z \partial \phi} = -\tan \gamma \sin \alpha'_{y} \frac{\partial k'_{y_m}}{\partial \phi}. \quad \text{(A-13)}$$

Differentiating equation (A-11) we can write the following:

$$\frac{\partial k'_{y_h}}{\partial \phi} = k_z \tan \gamma \left[ \frac{\partial \tan \alpha'_{x}}{\partial \phi} \sin \alpha'_{y} + \frac{\partial \sin \alpha'_{y}}{\partial \phi} \tan \alpha'_{x} \right]. \quad \text{(A-14)}$$

To evaluate equation (A-14) we need $\partial \tan \alpha'_{x}/\partial \phi$ and $\partial \sin \alpha'_{y}/\partial \phi$; that is,

$$\frac{\partial \tan \alpha'_{x}}{\partial \phi} = \frac{\partial \tan \alpha'_{y} \partial k'_{x_m}}{\partial \phi} = \frac{1}{k_z} \frac{\partial k'_{x_m}}{\partial \phi}, \quad \text{(A-15)}$$

and

$$\frac{\partial \sin \alpha'_{y}}{\partial \phi} = \frac{\partial \sin \alpha'_{y} \partial k'_{y_m}}{\partial \phi} = \frac{1}{k_z} \cos \alpha'_{y} \frac{\partial k'_{y_m}}{\partial \phi}. \quad \text{(A-16)}$$

Substituting equation (A-15) and equation (A-16) into equation (A-14), we finally obtain:

$$\frac{\partial k'_{y_h}}{\partial \phi} = k_z \tan \gamma \left[ \frac{\sin \alpha'_{y} \partial k'_{x_m}}{k_z \partial \phi} + \frac{\tan \alpha'_{x} \cos ^3 \alpha'_{y} \partial k'_{y_m}}{k_z \partial \phi} \right] = \tan \gamma \left[ \sin \alpha'_{y} \frac{\partial k'_{x_m}}{\partial \phi} + \tan \alpha'_{x} \cos \alpha'_{y} \frac{\partial k'_{y_m}}{\partial \phi} \right]. \quad \text{(A-17)}$$

Evaluation of $\partial k_{x_h}/\partial \phi$ and $\partial k_{y_h}/\partial \phi$

Equations (A-13) and (A-17) express the partial derivatives of the offset wavenumbers in the rotated coordinates. We need to evaluate the partial derivatives of the offset wavenumbers in the data coordinates. Therefore, we need to differentiate the expressions defining the inverse rotation; that is, differentiate with respect to $\phi$ the following expressions:

$$k_{x_h} = \cos \phi k'_{x_h} + \sin \phi k'_{y_h}, \quad \text{(A-18)}$$
$$k_{y_h} = -\sin \phi k'_{x_h} + \cos \phi k'_{y_h}. \quad \text{(A-19)}$$
The partial derivatives are then given by the following expressions:

\[
\frac{\partial k_x}{\partial \phi} = \cos \phi \frac{\partial k'_x}{\partial \phi} + \sin \phi \frac{\partial k'_y}{\partial \phi} - \sin \phi k'_x + \cos \phi k'_y, \quad (A-20)
\]

\[
\frac{\partial k_y}{\partial \phi} = -\sin \phi \frac{\partial k'_x}{\partial \phi} + \cos \phi \frac{\partial k'_y}{\partial \phi} - \cos \phi k'_x - \sin \phi k'_y. \quad (A-21)
\]

**Computation of scaling factor** \(W_{\gamma,\phi}\)

The last step is the computation of the scaling factor \(W_{\gamma,\phi}\) from the partial derivatives \(\frac{\partial k_x}{\partial \phi}\) and \(\frac{\partial k_y}{\partial \phi}\).

The unit vector \(u\) tangent to the integration line at constant \(\gamma\) is given by

\[
u = \left(\frac{\partial k_x}{\partial \phi} \frac{\partial k_y}{\partial \phi}, \frac{\partial k_y}{\partial \phi} - \frac{\partial k_x}{\partial \phi}, \frac{\partial k_x}{\partial \phi} \frac{\partial k_y}{\partial \phi}, \frac{\partial k_y}{\partial \phi} - \frac{\partial k_x}{\partial \phi}\right).
\] (A-22)

The mapping along the azimuth axis of a vector of the same direction as \(u\) and length proportional to the sampling in the offset wavenumber domain \((\Delta k_x, \Delta k_y)\) is a segment of length

\[
\delta \phi = \frac{2 \sqrt{\Delta k_x^2 + \Delta k_y^2}}{\frac{\partial k_x}{\partial \phi}^2 + \frac{\partial k_y}{\partial \phi}^2}.
\] (A-23)

If \(\Delta \phi\) is the azimuthal range of the stacking at \(\gamma = 0\), the scaling factor is set as in the following:

\[
W_{\gamma,\phi} = \begin{cases} 
1 & \text{if } \delta \phi \geq \Delta \phi \\
\Delta \phi & \text{if } \delta \phi < \Delta \phi.
\end{cases}
\] (A-24)
Velocity sensitivity of subsalt imaging through regularized inversion

Marie L. Clapp

ABSTRACT
The effects of inaccurate velocity models on migration are well known. Accurate velocity models are most difficult to obtain in complex areas where iterative inversion can provide a better image than migration. This paper investigates the velocity sensitivity of a regularized inversion scheme that explicitly assumes that the correct velocity is being used. This inversion uses a regularization operator that assumes that there is no moveout along the offset ray parameter axis. Experiments performed with various incorrect velocity models indicate that this assumption is valid for velocity models that can be reasonably produced by common velocity analysis techniques. Velocity models that are very inaccurate cause the inversion process to reject attempts by the regularization to produce an image that is inconsistent with the data.

INTRODUCTION

The difficulties of imaging below salt edges are compounded by the difficulty of generating an accurate velocity model in these areas. The majority of imaging techniques require an accurate velocity model in order to produce a well focused result. The artifacts seen in a migrated image caused by errors in the velocity model are well known (Claerbout, 1985). In fact, migration can be used as a tool for developing a velocity model (Biondi and Sava, 1999). However, in areas such as those around salt, where it is most difficult to obtain a good velocity model, poor illumination makes it impossible for migration alone to provide a satisfactory image.

In complex areas, imaging can be improved by using a migration operator in an iterative inversion scheme. Unfortunately, if the velocity model is inaccurate the artifacts that are seen in the migration result will affect the inversion result as well. Even more ominously, some iterative inversion techniques make the assumption of correct velocity a critical part of their theory. In particular, imaging using the regularized inversion described by Prucha et al. (2000) and Kuehl and Sacchi (2001) assumes that the correct velocity is being used to justify the choice of regularization operator. The regularization operators they use assume that there is no moveout along the offset ray parameter axis. In this paper, I will examine the sensitivity of this assumption for a variation of Prucha et al. (2000)’s implementation.

I will first explain the manner in which I carry out regularized inversion and the regular-
ization operator used. Then I will perform migration and regularized inversion on a synthetic dataset using the correct velocity model and two velocity models that have been perturbed in different ways. I will compare these results to make a statement on the validity of the zero-moveout assumption.

**THEORY**

My inversion scheme is based on the downward continuation migration explained by Prucha et al. (1999a). To summarize, this migration is carried out by downward continuing the wavefield in frequency space, slant stacking at each depth, and extracting the image at zero time. The result is an image in depth \( z \), common midpoint (CMP), and offset ray parameter \( p_h \) space.

This migration operator is used in a Tikhonov regularized (Tikhonov and Arsenin, 1977) conjugate-gradient least-squares minimization:

\[
\min \{ Q(m) = ||W(Lm - d)||^2 + \epsilon^2 ||Am||^2 \}. \tag{1}
\]

This inversion procedure can be expressed as fitting goals as follows:

\[
\begin{align*}
0 &\approx W(Lm - d) \\
0 &\approx \epsilon Am.
\end{align*} \tag{2}
\]

The first equation is the “data fitting goal,” meaning that it is responsible for making a model that is consistent with the data. The second equation 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 model styling goal also helps to prevent a divergent result.

In the data fitting goal, \( d \) is the input data and \( m \) is the image obtained through inversion. \( L \) is a linear operator, in this case it is the adjoint of the angle-domain wave-equation migration scheme summarized above and explained thoroughly by Prucha et al. (1999b). In the model styling goal, \( A \) is, as has already been mentioned, a regularization operator. \( W \) is a weighting operator. \( \epsilon \) controls the strength of the model styling.

Unfortunately, the inversion process described by fitting goals (2) can take many iterations to produce a satisfactory result. I can reduce the necessary number of iterations by making the problem a preconditioned one. I use the preconditioning transformation \( m = A^{-1}p \) (Fomel et al., 1997; Fomel and Claerbout, 2002) to give us these fitting goals:

\[
\begin{align*}
0 &\approx W(LA^{-1}p - d) \\
0 &\approx \epsilon p.
\end{align*} \tag{3}
\]

\( A^{-1} \) is obtained by mapping the multi-dimensional regularization operator \( A \) to helical space and applying polynomial division (Claerbout, 1998).

The question now is what the regularization operator \( A \) is. I built my regularization operator based on the same assumptions as Prucha et al. (2000). First, I assume that the correct
velocity is being used in the inversion, therefore there should be no moveout along the offset ray parameter \((p_h)\) axis. Second, I assume that the amplitudes of individual events should vary smoothly and any drastic changes in amplitude are caused by illumination problems, which are what we wish to overcome. These assumptions allow me to say that \(\mathbf{A}\) needs to act to minimize amplitude differences horizontally along the \(p_h\) axis. Rather than using the derivative operator used by Kuehl and Sacchi (2001) or the steering filter used by Prucha et al. (2000), I have created a symmetrical filter by cascading two steering filters that are mirror images of each other.

**RESULTS**

I applied the downward continuation migration and the preconditioned inversion scheme to a synthetic dataset provided to us by SMAART JV, using different velocity models. The correct velocity model can be seen in Figure 1. The result of migration using this model is Figure 2. In the CRP-depth panel, note the sudden decrease in amplitude of the reflectors as they pass beneath the salt edge, particularly within the oval. There are also strong artifacts in the shadow zone beneath the salt (inside the oval) which make it difficult to pick out any true events. In the \(p_h\)-depth panel, note the “holes” in the events at the mid-range of ray parameters (inside the oval). These holes are caused by the poor illumination under the salt edge. The steeply dipping events in the \(p_h\)-depth panel are artifacts caused by aliasing along the offset axis in Fourier space.

The result of 3 iterations of conjugate-gradient preconditioned least-squares inversion using the correct velocity model can be seen in Figure 3. Note that the artifacts have been largely cleaned up. It is now possible to reliably pick out events beneath the salt (see inside the oval). In the CRP-depth panel, the amplitude of the events is maintained farther beneath the salt (particularly within the oval). The holes in the \(p_h\)-depth panel (inside the oval) are being filled in.

To test the sensitivity of the preconditioned inversion, the first incorrect velocity model I tested simply increased the correct velocities by 5%. As expected, the migration result using

![Figure 1: The correct velocity model.](.marie1-vel.corvel)
Figure 2: The result of downward continuation migration using the correct velocity model. Note the low amplitudes of events as they pass beneath the salt in the CRP-depth panel and the artifacts obscuring events beneath the salt (indicated by ovals). In the offset ray parameter-depth panel note the holes in the events at the mid-range of offset ray parameters (particularly within the oval).

Recall that the preconditioning operator acts horizontally along the offset ray parameter axis. It is this sensitivity that we are interested in observing in the result of 3 iterations of preconditioned inversion using the high velocity model (Figure 5). Note that once again the preconditioned inversion has cleaned up many of the artifacts. In the CRP-depth panel, the events extend farther under the salt, in a similar way to the inversion result using the correct velocity (Figure 3). The more interesting result is the \( p_h \)-depth panel. The inversion is still successfully filling in the holes along the events at the mid-range of offset ray parameters. At large \( p_h \), where the moveout is more pronounced, the preconditioning has made some attempt to change the dips to be more horizontal, but the moveout is still visible. This means that this result is most likely not safe to use for velocity analysis, but this preconditioned
Figure 3: The result of 3 iterations of preconditioned inversion using the correct velocity model. Note the more consistent amplitudes of events as they pass beneath the salt in the CRP-depth panel and the lack of artifacts obscuring events beneath the salt (inside ovals). In the offset ray parameter-depth panel note the filling in of the holes in the events at the mid-range of offset ray parameters (inside ovals). The inversion technique was never intended as a velocity tool. Overall, this result indicates that this technique can produce a better image than migration alone, even when the velocity model is incorrect by up to 5%.

A more extreme velocity model I tested was a severely smoothed one (Figure 6). This model has been smoothed so much that the canyon in the top of the salt has disappeared. As expected, the migration result from this model isn’t very good (Figure 7). The depth positioning of events is fairly good away from the salt, but becomes poor near the salt. The salt top and bottom are very poorly imaged. The events in the $p_h$-depth panel appear to be mostly random. Once again, the ovals indicate the same absolute regions as the ovals in Figures 2 and 3.

The result of 3 iterations of preconditioned inversion using this smoothed velocity model can be seen in Figure 8. Although many of the artifacts have been cleaned up, overall the image is not any better than the migration result. The events in the $p_h$-depth panel are more
Figure 4: The result of downward continuation migration using a velocity model 5% higher than the correct model. The events are all positioned deeper than they should be and there is moveout along the offset ray parameter axis. The ovals still indicate the loss of amplitudes under the salt edge and the poor imaging beneath the salt in the CRP-depth panel and holes in the events in the $p_h$ panel. [marie1-mig.hivel] [CR]

CONCLUSIONS

The results of preconditioned inversion with incorrect velocity models are encouraging. As long as the velocity model is not too inaccurate, the preconditioning operator behaves as it would for the correct velocity model and produces a better image than migration alone. In the case of a highly inaccurate model, the inversion itself prevents us from producing an image that would conflict with the known data. Overall, as long as the velocity model is reasonably close to correct, the assumption of zero moveout made by the preconditioning operator is acceptable.
Figure 5: The result of 3 iterations of preconditioned inversion using the 5% too high velocity model. Despite the use of the incorrect velocity model, the image is quite comparable to the result using the correct velocity (Figure 3). In the CRP-depth panel, the events extend farther under the salt and events under the salt can be seen (inside the ovals). In the $p_h$-depth panel, the holes in the events are filled in (inside the oval). [marie1-geop.hivel][CR]

ACKNOWLEDGMENTS

I would like to thank SMAART JV for the synthetic dataset used in my experiments.

REFERENCES


Figure 6: The smoothed velocity model. Note that the canyon in the top of the salt has disappeared.


Figure 7: The result of downward continuation migration using a severely smoothed velocity model. The events in the CRP-depth panel are properly imaged away from the salt but are mispositioned near the salt. The offset ray parameter-depth panel is completely uninformative. The ovals indicate the same absolute regions as the ovals in Figures 2 and 3.
Figure 8: The result of 3 iterations of preconditioned inversion using the smoothed velocity model. The result is cleaner than the migration result, but not more believable. The ovals indicate the same absolute regions as the ovals in Figures 2 and 3.
Short Note

WKBJ and amplitude preserving one-way wave equation

Guojian Shan and Biondo Biondi

INTRODUCTION

The standard one-way wave equation (Claerbout, 1971) produces the correct phase of the wavefield, but it is not equivalent to the acoustic wave equation in terms of amplitude. Zhang (1993) suggests that to improve the dynamics information of one-way wave equation, an amplitude correction term should be included into the standard one-way wave equation. Zhang et al. (2001) apply the one-way wave equation with the amplitude correction to shot-profile migration and shows that it can provide the same amplitude as an image produced by the true amplitude Kirchoff migration (Hanitzsch, 1997) by asymptotic analysis. Vlad et al. (2003) compare the amplitude of one-way wave equation with amplitude correction with acoustic wave equation and WKBJ amplitude correction, and concludes that it has the same effect as the WKBJ correction (Clayton and Stolt, 1981). In this short note, we will demonstrate in theory that the one-way wave equation with the amplitude correction term is equivalent to the first order approximation of WKBJ amplitude correction.

AMPLITUDE PRESERVING ONE-WAY WAVE EQUATION

The standard one-way wave equation is

$$\frac{\partial P}{\partial z} = i \frac{\omega}{v} \sqrt{1 - \frac{v^2 k_z^2}{\omega^2}} P,$$

(1)

where $k_z$ is the wave number in the $x$ direction, and $v$ is the velocity. From the dispersion relation, we have

$$k_z = \frac{\omega}{v} \sqrt{1 - \frac{v^2 k_x^2}{\omega^2}},$$

(2)

where $k_z$ is the wavenumber in the $z$ direction. Equation (1) mimics the phase behavior of the acoustic wave equation:

$$\frac{1}{v^2} \frac{\partial^2}{\partial t^2} P - \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) P = 0,$$

(3)

email: shan@sep.stanford.edu, biondo@sep.stanford.edu
although the amplitude is not correct. Zhang (1993) suggests that to maintain the dynamics information of equation (3), an extra amplitude correction term needs to be included into the standard one-way wave equation. The new one-way wave equation is

$$\frac{\partial P}{\partial z} = \left[ i k_z - \frac{1}{2v} \frac{\partial v}{\partial z} \left( 1 + \frac{(vk_x)^2}{\omega^2 - (vk_z)^2} \right) \right] P. \quad (4)$$

We call it the amplitude preserving one-way wave equation. By equation (4), the wavefield can be extrapolated as the following:

$$P(z + \Delta z) = e^{ik_z \Delta z - \frac{1}{2v} \frac{\partial v}{\partial z} \left( 1 + \frac{(vk_x)^2}{\omega^2 - (vk_z)^2} \right) \Delta z} P(z) \quad (5)$$

$$= e^{-\frac{1}{2v} \frac{\partial v}{\partial z} \left( 1 + \frac{(vk_x)^2}{\omega^2 - (vk_z)^2} \right) \Delta z} e^{ik_z \Delta z} P(z). \quad (6)$$

In equation (6), first term defines the amplitude information of the wavefield, and second term $e^{ik_z \Delta z}$ defines the phase information of the wavefield.

**WKBJ AND ITS FIRST ORDER APPROXIMATION**

In this section, we demonstrate that the first order approximation of WKBJ is the same as the amplitude preserving one-way wave equation for $v = v(z)$.

The one-way wave equation with the WKBJ amplitude correction is

$$P(z + \Delta z) = \sqrt{\frac{k_z(z + \Delta z)}{k_z(z)}} e^{ik_z \Delta z} P(z). \quad (7)$$

The WKBJ amplitude correction term $\sqrt{\frac{k_z(z + \Delta z)}{k_z(z)}}$ can be rewritten as

$$\sqrt{\frac{k_z(z + \Delta z)}{k_z(z)}} = e^{\frac{1}{2} \ln \frac{k_z(z + \Delta z)}{k_z(z)}}. \quad (8)$$

Then $k_z(z + \Delta z)$ in equation (8) can be linearized to $k_z(z) + \frac{\partial k_z}{\partial z} \Delta z$, so we have

$$\frac{1}{2} \ln \frac{k_z(z + \Delta z)}{k_z(z)} \approx \frac{1}{2} \ln(1 + \frac{1}{k_z} \frac{\partial k_z}{\partial z} \Delta z). \quad (9)$$

From

$$\ln(1 + x) = x - x^2 + \cdots,$$

and

$$\ln(1 + x) \approx x,$$

we have

$$\frac{1}{2} \ln(1 + \frac{1}{k_z} \frac{\partial k_z}{\partial z} \Delta z) \approx \frac{1}{2} \frac{1}{k_z} \frac{\partial k_z}{\partial z} \Delta z. \quad (10)$$
Because
\[
\frac{k_z}{z} \frac{\partial k_z}{\partial z} = \frac{\partial}{\partial z} \ln k_z,
\]
and from the dispersion relation \( k_z = \sqrt{\frac{\omega^2}{v^2} - k_x^2} \), we have
\[
\frac{1}{2} \frac{\partial k_z}{\partial z} \Delta z = \frac{1}{2} \frac{\partial}{\partial z} \ln \left( \frac{\omega^2}{v^2} - k_x^2 \right) \Delta z = \frac{1}{4} \frac{\partial}{\partial z} \ln \left( \frac{\omega^2}{v^2} - k_x^2 \right) \Delta z,
\]
and
\[
\frac{1}{4} \frac{\partial}{\partial z} \ln \left( \frac{\omega^2}{v^2} - k_x^2 \right) = \frac{1}{4} \frac{\omega^2}{v^2} \frac{\partial}{\partial z} \ln \left( \frac{\omega^2}{v^2} - k_x^2 \right) = -\frac{1}{2} \frac{\partial v}{\partial z} \frac{\omega^2}{v^2} - (k_x v)^2.
\]
From equation (8) to equation (13), we have
\[
\sqrt{\frac{k_z(z + \Delta z)}{k_z(z)}} \approx e^{-\frac{1}{2} \frac{\partial v}{\partial z} \frac{\omega^2}{v^2} - (k_x v)^2} \Delta z.
\]
So equation (7) can be rewritten as
\[
P(z + \Delta z) = \sqrt{\frac{k_z(z + \Delta z)}{k_z(z)}} e^{ik_z \Delta z} P(z)
\]
\[
\approx e^{ik_z \Delta z - \frac{1}{2} \frac{\partial v}{\partial z} \frac{\omega^2}{v^2} - (k_x v)^2} \Delta z P(z).
\]
Comparing the amplitude preserving one-way wave equation (equation (4)) with first order approximation of WKBJ (equation (16)), we find they are same. So we demonstrate theoretically that the amplitude preserving one-way wave equation is equivalent to the first order approximation of WKBJ.

**CONCLUSION**

With the extra amplitude correction term, the amplitude preserving one-way wave equation improves the amplitude of the wavefield. The extra amplitude correction in the equation is mathematically equivalent to the first order approximation of WKBJ amplitude correction. But unlike the WKBJ correction, the amplitude preserving one-way wave equation can be used for a laterally varying velocity. The accuracy of this correction in the presence of lateral velocity variation has not been analyze yet.

**REFERENCES**

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


Kinematics of 3-D angle-domain common-image gathers for migration velocity analysis

Thomas Tisserant and Biondo Biondi

ABSTRACT
We extend the study of the kinematic properties of 2-D offset- and angle-domain common-image gathers to the general 3-D problem. We examine how the use of an incorrect migration velocity affects the behavior of the image in the offset and angle domains. We show that the general 3-D case with and without the correct migration velocity can be cast into a 2-D formulation, making it possible to apply existing theory from 2-D. We illustrate both ray-tracing and plane-wave approaches to the problem and verify our theoretical results with a synthetic model.

INTRODUCTION
Migration Velocity Analysis (MVA) and Amplitude Versus Angle (AVA) analysis are the two main applications of Angle-Domain Common-Image Gathers (ADCIGs). ADCIGs can be computed in the Fourier domain from offset-domain images generated by wavefield-continuation methods (Sava and Fomel, 2000). In previous work (Tisserant and Biondi, 2003), we presented an extension of the method to make it valid for 3-D geometries. Biondi and Symes (2003) analyze the kinematics of ADCIGs when an incorrect migration velocity has been used during the wavefield-continuation step. This paper aims to extend their analysis to 3-D.

We consider two approaches: one based on rays, the other based on plane-waves. In the ray approach, we split our analysis between the use of an correct and an incorrect migration velocity. When the migration velocity is correct, the source and receiver rays focus at the same correct image point where they are coplanar. This property of the rays makes it simple to transform the 3-D geometry into 2-D for which the theory is available (Biondi and Symes, 2003). We then analyze the problem when an incorrect migration velocity is used. Incorrect migration velocity yields non focused ray resulting in an apparent image point, and generally not coplanar rays. We introduce an apparent propagation plane in which the 2-D theory can be applied. We will then discuss a plane-wave approach of the problem. Finally a synthetic example is used to verify our theoretical results.

1 email: thomas@sep.stanford.edu, biondo@sep.stanford.edu
GEOMETRY

Extension from 2-D to 3-D requires the introduction of new angles in the geometric definition of the problem. However, we will see that a 3-D problem can be formulated as a 2-D problem whose analysis is done by Biondi and Symes (2003).

In 2-D, two angles define the geometry of the problem: the dip of the interface, $\alpha$, and the aperture of the reflection, $\gamma$. In 3-D, the reflector is not only defined by its dip $\alpha$, but also by its azimuth $\phi$. Further, in the full 3-D geometry, the rays are not necessarily coplanar during downward-continuation. As a consequence, the reflection may have a different azimuth than the azimuth of the survey. We call $\beta$ the reflection azimuth. When the correct migration velocity is used, the two rays focus correctly in one point. Near this image point the two rays define defines one plane. The most general configuration considers an incorrect migration velocity. Because the velocity is incorrect, the rays do not focus in one point. Instead, they stop as two different points with the same depth when using downward continuation. The distance between the two points is the offset at constant $z$. The middle of the segment is the image point in the offset domain domain at constant $z$ (zODCIG). In such configuration, the rays are not necessarily coplanar. We introduce a new angle, $\xi$, accounting for the non-coplanarity of the rays. To make the link with the 2-D case, we seek an apparent propagation plane containing all the information about the actual geometry. In all cases, the image point in the angle domain moves along the normal to the apparent interface by an amount dependent on the migration velocity used, and on the aperture angle. The image point in the angle domain is obtained by transforming of the image point in the offset-domain. We use the post-migration transformation in the Fourier domain introduced in Sava and Fomel (2000) for the 2-D geometry, and its 3-D extensions by Tisserant and Biondi (2003) for the 3-D full prestack migration. Our approach in the next two sections is based on a ray construction. Later, we will present in the third section another approach based on plane-waves.

CORRECT MIGRATION VELOCITY

Discussion

The 3-D geometry implies that the rays can now have different azimuth and can propagate out of the vertical plane. If the migration velocity is correct, the two rays focus at the same point at zero subsurface offset (Figure 1). By assuming the velocity is constant around the image point, all the elements (rays, normal, image points) are contained in one plane: the plane of the propagation. By an appropriate change of coordinates, this 3-D problem with correct migration velocity can be locally transformed in a 2-D problem, and the 2-D theory analyzed by Biondi and Symes (2003) be applied. The offset-to-angle transformation must be adapted though.
Figure 1: The velocity function is \( V(z) = 1.5 + .5z \) km/s. The target has a fixed position, azimuth (\( \phi = 45 \)) and dip (\( \alpha = 60 \)).

Offset-to-angle transformation

In 2-D, the offset-to-angle transformation is done with the relation

\[
\tan \gamma = -\frac{k_h}{k_z},
\]

where \( \gamma \) is the aperture angle of the reflection, \( k_h \) is the offset wavenumber associated with the subsurface horizontal offset, and \( k_z \) is the vertical wavenumber. Tisserant and Biondi (2003) presented a 3-D generalization of Equation 1:

\[
\tan \gamma' = -\frac{k_{h_y}'}{\sqrt{k_z^2 + k_{m_y}^2}},
\]

\[
k_{h_y}' = -\frac{k_{m_y}' k_{m_x}' k_{h_x}'}{k_z^2 + k_{m_y}^2}
\]

where \( k_{m} \) and \( k_{h} \) are the midpoint and offset vector wavenumber, respectively, and where the reflection azimuth, \( \beta \), is introduced through

\[
k_{m_x}' = \cos \beta k_{m_x} - \sin \beta k_{m_y}
\]

(4)

\[
k_{m_y}' = \cos \beta k_{m_x} + \sin \beta k_{m_y}
\]

(5)

\[
k_{h_x}' = \cos \beta k_{h_x} - \sin \beta k_{h_y}
\]

(6)

\[
k_{h_y}' = \cos \beta k_{h_x} + \sin \beta k_{h_y}
\]

(7)

The offset-to-angle transforms a \((k_z, k_{m_x}, k_{m_y}, k_{h_x}, k_{h_y})\) five dimensions cube into another 5-D one \((k_z, k_{m_x}, k_{m_y}, \gamma', \delta)\). Figure 2 is the measured aperture-azimuth distribution for the configuration displayed in Figure 1 obtained with ray-tracing. We set the lower boundary in \( \gamma \) because of an increased incertitude in the estimation of \( \beta \) as \( \gamma \) gets close to 0. The upper boundary in \( \gamma \) is reached when one of the two rays begins to overturn.

We now present a more complex 3-D extension: the one addressing the 3-D full prestack migration with a wrong migration velocity.
Figure 2: $\gamma - \beta$ distribution corresponding to Figure 1

Figure 3: Ray-tracing with the incorrect velocity function $V(z) = 1.41 + .47z$ km/s. Because the velocity is too slow, the rays end at too shallow of a depth.

**INCORRECT MIGRATION VELOCITY**

**Discussion**

Figure 3 presents the geometry of the problem throughout an example. It has been built considering the same events as in Figure 1 but by using an incorrect velocity. Because the velocity is too low, the rays stop at too shallow of a depth, where they are not coplanar. The gap between the source and the receiver endpoints is the horizontal offset ($z_{ODCIG}$, $z$ for constant $z$). The middle of the horizontal offset is the image point in the $z_{ODCIG}$ domain. One can measure on Figure 3 the azimuth of the source-receiver end points segment. We call it $\beta'$. $\beta'$ can be decomposed into two components: one, $\beta$, is the azimuth due to ray-bending, the other, $\xi$, accounts for the non-coplanarity due to the use of an incorrect migration velocity. If the migration velocity is too slow, the rays stop too early, yielding to an underestimated reflection azimuth. Conversely, if the migration velocity is too high, the rays stop too late, yielding to a overestimated reflection azimuth. When the velocity is exact, the rays are coplanar and the rotation of the reflection azimuth is only due to ray-bending and dips.
ODCIGs and ADCIGs properties

The velocity is assumed to be constant in a small volume around the image points. That is why the rays are straight lines in the next few pictures.

Biondi and Symes (2003) noted that, in 2-D, image points in the offset domains (at constant \(x\), \(x_{\text{ODCIG}}\), and constant \(z\), \(z_{\text{ODCIG}}\)) lie on an apparent interface. In 2-D the apparent interface is a line, whereas it is a plane in 3-D. Not only the image points belong to the interface, but they are also all collinear (Figure 4). From this property we can define a new plane that includes the normal and the common line of the image points in the offset domains. This plane has special properties since it contains all the image points: it is the link to the 2-D case. Let us further study this plane.

Figure 4: The endpoints and image points of all the offset-domain gathers are displayed (dotted lines). All the image points are defined in the interface and are collinear.

If the velocity is incorrect, the rays do not focus at the actual image point but at an apparent image point located in the middle of the end points of the two rays. The apparent image point is the image point in the offset domain. Further, in the case of full prestack migration, the rays are not coplanar, so there is no propagation plane. Physics requires one though. In the absence of a propagation plane, we define an apparent propagation plane. To find it, we start from the actual rays and find a rotation that makes their image coplanar. The rotation is done around the normal at the interface (Figure 5). Note that the new rays are parallel to the original ones but have different end points. The plane in which the rays are coplanar after rotation is the apparent propagation plane. The geometric location of the image points in the offset-domain is at the intersection between the propagation plane and the interface. The rotated rays thus define the same plane-wave as the original rays. The angle of the rotation, \(\xi\), is equal to the azimuth defined in the interface of both source and receiver rays. The reflection azimuth, \(\beta\), is equal to the azimuth of the apparent propagation plane.

We now analyze how the image in the angle domain is obtained from the image in the offset domain.
**Offset-to-angle transformation.**

The apparent propagation plane plays another interesting role during the offset-to-angle transformation: despite the non-coplanarity of the rays, the energy moves only in the apparent propagation plane. Figure 6 illustrates this property. Two offset-domain image points are displayed. Points in the offset domain to the image point in the angle domain along a direction orthogonal to the direction of the offset. All the possible directions define a plane. The two orthogonal planes are displayed on Figure 6. The intersection of the two planes with the normal gives the position of the image point in the angle domain. More specifically, the energy moves at the intersection between the orthogonal plane and the apparent propagation plane. To sum up, when an incorrect migration velocity is used, the rays are not coplanar. There is however a preferential plane, the apparent propagation plane, which allows one to cast the 3-D problem into a 2-D one.

**PLANE-WAVE APPROACH**

In the previous sections, we have talked about rays or ray-tracing. This approach has the advantage of being intuitive. The physics, however is governed by plane-waves. In this section we present a plane-wave approach of the same problem.

**Why plane-waves**

One reason to use plane-waves instead of rays is to avoid the asymptotic approximation introduced when using rays. Unlike rays, each plane-wave can be treated independently at each
Figure 6: Geometric construction of the image point in the angle domain from the image points in the offset domain.

Figure 7: Geometric construction of the image point in the angle domain from the source and receiver rays end points.
frequency. The use of plane-waves simply ignores the problem of non-coplanar rays. Indeed, the original and the rotated rays described earlier share the same plane-wave. The reflection azimuth is the azimuth of the apparent propagation plane which is directly deduced from the angles of the two plane-waves.

Construction

Given one image point in the angle domain, and the four angles defining the two plane-waves (two azimuths and two dips), the two plane-waves are positioned such that the image point is contained in both. The directions of the two plane-waves provide the azimuth and the dip of the reflector. The image point is contained in the intersection of the two plane waves. To find the plane of the reflection we define the apparent propagation plane as being orthogonal to the two plane-waves and passing through the image point. Once this plane defined, knowledge of the rays location and coplanarity is inconsequential.

SYNTHETIC EXAMPLE

We verify our theoretical results on a model with five slanted planes (Vaillant and Biondi, 2000). The velocity function is \( v(z) = 1.5 + .5z \) km/s. The velocity gradient allows us to highlight the rotation of the reflection azimuth due to ray-bending. The model consists of five planes with 0°, 15°, 30°, 45°, 60° dip and a 45° azimuth. The dataset is migrated with full prestack migration. We migrate first with the correct velocity and then with a velocity that is 6% too slow. The migrated model is then transformed from offset to angle for different reflection azimuths. Hence the image has five dimensions: in-line and cross-line midpoints, depth, aperture and reflection azimuth. We select one midpoint and obtain a cube of a midpoint gather sorted in aperture angle and reflection azimuth angle.
Behavior of the reflection azimuth

Figure 9 shows one depth slice in the \((z, \gamma, \beta)\) cube at \(z = 1450\) m. We observe that the reflection azimuth changes slowly with the aperture angle, except at small angles where the uncertainty is high. This slow variation is consistent with Figure 2 since the reflection azimuth changes by \(2^\circ\) while the aperture changes by \(25^\circ\). One can also notice that the maximum aperture angle obtained with ray-tracing, \(35^\circ\), is consistent with the aperture angle interval where the energy is high. The reason may be that rays begin to overturn when the aperture angle is higher than \(35^\circ\). Indeed, the upper limit in ray-tracing is due to one of the ray beginning to overturn.

![Figure 9: Depth slice in the \((z, \gamma, \beta)\) cube.](image)

\[\text{ADCIGs movement with wrong velocity}\]

We use the ray-tracing to predict the localization of the energy in the synthetic dataset migrated with the wrong velocity. The ray-tracing is illustrated by Figure 3. Starting from the true position of the image point (the diamond on the Figure), we seek the position of the image point (stars on the figure) when an incorrect migration velocity is used. The reflector has a fixed azimuth and dip. We first model the events recorded at the surface at one particular aperture and reflection azimuth. Once the true events are known, the source and the receiver rays are shot in a media with an incorrect velocity this time. If the velocity is too slow, then the rays stop at two distinct end points. Knowing the position of the end points and the ray parameters, the position of the image point in the angle domain for an incorrect migration velocity is determined through a geometric construction similar to the one in 6. It is also possible to use the normal shift described in the 2-D case since we have shown that the 3-D problem can be recast as 2-D one. We test this procedure on an image point on the 60° dip reflector and whose location is \((400, 400, 1300)\) in image. We choose a source-receiver pair such as the aperture angle is \(32^\circ\). Again, our goal is to find where the true image point has moved because of the use of an incorrect migration velocity. We observe on Figure 10 that the coordinates of the apparent image point in the angle domain computed by ray-tracing do match those of the image point in the model migrated with the wrong velocity.
CONCLUSIONS

We have presented a 3-D study of the kinematic of angle-domain common-image gather. Knowing the theory in 2-D, we showed how a 3-D configuration can be casted in a 2-D formulation. The transition is straightforward when a correct velocity is used since the plane of the propagation is the plane in which the problem becomes 2-D. In the less obvious problem where an incorrect migration velocity is used, we have defined an apparent propagation plane that fulfills the same objective. We have presented both a ray-based and a plane-wave approach. The kinematic observations made with ray-tracing have been verified on a synthetic dataset.

REFERENCES


Wave-equation MVA: Born, Rytov and beyond

Paul Sava and Biondo Biondi

ABSTRACT

The linearized wave-equation MVA operator can be used for velocity analysis using both Born and Rytov approximations. The distinction arises from the method used to compute the image perturbations. Both approximations suffer from limitations that limit their practicality: the Born approximation is usable only for small anomalies, while the Rytov approximation requires phase unwrapping. Differential image perturbations can be used for arbitrarily large slowness anomalies and do not require phase unwrapping, but their accuracy decreases with increasing deviation from the background image. For simple cases, the differential image perturbation method is equivalent with phase-unwrapped Rytov.

INTRODUCTION

Migration velocity analysis based on downward continuation methods, also known as wave-equation migration velocity analysis (WEMVA), is a technique designed as a companion to wave-equation migration (Biondi and Sava, 1999; Sava and Fomel, 2002). The main idea of WEMVA is to use downward continuation operators for migration velocity analysis (MVA), as well as for migration. This is in contrast with other techniques which use downward continuation for migration, but traveltime-based techniques for migration velocity updating (Clapp, 2001; Liu et al., 2001; Mosher et al., 2001).

WEMVA is an optimization problem where the objective function is defined in the image space. As for other MVA methods, it tries to maximize the quality of the migrated image instead of trying to match the recorded data (Sava and Symes, 2002). In this respect, our method is related to Differential Semblance Optimization (Symes and Carazzone, 1991) and Multiple Migration Fitting (Chavent and Jacowitz, 1995). However, with respect to these two methods, our method has the advantage of exploiting the power of residual prestack migration to speed up the convergence, and it also gives us the ability to guide the inversion by geologic interpretation.

WEMVA benefits from the same advantages over traveltime estimation methods as wave-equation migration benefits over Kirchhoff migration. The most important among them are the accurate handling of complex wavefields which are characterized by multipathing, and the band-limited nature of the imaging process, which can handle sharp velocity variations much better than traveltime-based methods. Complex geology, therefore, is where WEMVA

1email: paul@sep.stanford.edu, biondo@sep.stanford.edu
is expected to provide the largest benefits.

Figures 1 and 2 illustrate the complications encountered under salt. From top to bottom, we show the wavefield corresponding to a point-source at the surface, the background slowness model, and a “fat ray” (Woodward, 1992) from the source to a point in the subsurfaces. Both examples show multipathing between source and receiver which is naturally taken into account by WEMVA, but which cannot be handled by simple traveltime tomography. Also, the slowness model is not smooth as required by methods using ray tracing.

WEMVA is based on a linearization of the downward-continuation operator using the Born approximation. This approximation leads to severe limitations on the magnitude and size of the anomalies that can be handled. Therefore, it cannot operate successfully in the regions of highest complexity. Other methods of linearization are possible (Sava and Fomel, 2002), but none allow for arbitrarily large anomalies.

In our early tests (Biondi and Sava, 1999), we construct the image perturbation using
Prestack Stolt Residual Migration (PSRM) (Sava, 2003). In summary, this residual migration method provides updated images for new velocity maps that correspond to a fixed ratio ($\rho$) of the new velocity with respect to the original velocity map. Residual migration is run for various ratio parameters, and finally we pick the best image by selecting the flattest gathers at every point.

The main disadvantage of building the image perturbation using PSRM is that, for large velocity ratio parameters ($\rho$), the background and improved images can get more than $\pi/4$ out of phase. Therefore, the image perturbation computed by the Born forward operator and the one computed by residual migration are fundamentally different, and can have contradictory behaviors when using the linearized WEMVA operator for inversion.

Alternative methods can be used to create image perturbations for WEMVA, directly in compliance with the Born approximation (Sava and Biondi, 2003). Those methods are not limited by the same restrictions as the Born methods, although their accuracy decreases with increasing deviation from the reference images. We refer to image perturbations created with
this method as differential image perturbations.

In this paper we investigate various methods that can be used to compute image perturbations. Our goal is two fold: firstly, we show that we can do Rytov WEMVA basically using the same backprojection operator as the one involved in the Born approximation, but with a different definition of the image perturbation; secondly, we further investigate the differential approach to computing image perturbations and show that this method is similar to a phase-unwrapped Rytov, although it does not require any explicit phase unwrapping. Thus, the differential method is more practical and robust than a Rytov method, although it is less accurate and may require more non-linear iterations for convergence.

We begin with a review of wave-equation MVA with emphasis on the methods used to compute image perturbations. We then describe the Rytov approach to WEMVA and explore the relationship between the three different methodologies. Finally, we present a couple of synthetic examples for each of the image perturbation methods mentioned earlier and discuss the relationships between them.

**WEMVA THEORY**

The wavefield constructed by downward continuation from the surface to depth \( z \), \( \mathcal{U}(z) \) is

\[
\mathcal{U} = D e^{j \sum_z \Phi_z} ,
\]

where \( D \) is the data at the surface and \( \Phi_z \) is the complex phase shift at one depth level. We can write the phase \( \Phi_z \) at every depth level as a Taylor expansion around a reference medium of slowness \( s_o \)

\[
\Phi_z = \Phi_{z_o} + \left. \frac{d \Phi_z}{d s} \right|_{s=s_o} \Delta s \\
= \Phi_{z_o} + \Delta \Phi_z .
\]

If we plug equation (2) in equation (1) we can write the following expression for the wavefield \( \mathcal{U} \):

\[
\mathcal{U} = \mathcal{U}_o e^{j \sum_z \Delta \Phi_z} ,
\]

where \( \mathcal{U}_o \) corresponds to the background slowness \( s_o \), and \( \mathcal{U} \) corresponds to an arbitrary spatially varying slowness \( s = s_o + \Delta s \).

We can define a wavefield perturbation at depth \( z \) by the expression

\[
\Delta \mathcal{U} = \mathcal{U} - \mathcal{U}_o \\
= \mathcal{U}_o \left[ e^{j \sum_z \Delta \Phi_z} - 1 \right] \\
\]

or, if we use the notation \( \Delta \Phi = \sum_z \Delta \Phi_z \)

\[
\Delta \mathcal{U} = \mathcal{U}_o \left[ e^{j \Delta \Phi} - 1 \right] .
\]
In general, we can compute a wavefield perturbation $\Delta U(\omega, z)$ by applying a non-linear operator $L$ which depends on the background wavefield $U_o(\omega, z)$ to a slowness perturbation $\Delta s(z)$, according to equation (7):

$$\Delta U = L(U_o)[\Delta s] . \quad (8)$$

**Linearization**

The simplest linearization of equation (7) is done by the Born approximation, which involves an approximation of the exponential function by a linear function $e^{i\phi} = 1 + i\phi$. With this approximation, we obtain

$$\Delta U \approx U_o i \Delta \Phi . \quad (9)$$

We can, therefore, compute a linear wavefield perturbation $\Delta U(\omega, z)$ using a Born WEMVA operator:

$$\Delta U = B(U_o)[\Delta s] , \quad (10)$$

from which we can compute an image perturbation by summation over frequency:

$$\Delta R = \sum_\omega \Delta U . \quad (11)$$

For wave-equation MVA, we are interested in applying an inverse WEMVA operator to a given image perturbation. Therefore, the main challenge of the linearized WEMVA is to estimate correctly $\Delta R$, i.e. an image perturbation corresponding to the accumulated phase differences given by all slowness anomalies above each image point.

Given an image perturbation $\Delta R$, we can compute a wavefield perturbation $\Delta U$ by the adjoint of the imaging operator, from which we can compute a slowness perturbation based on the background wavefield $U_o$:

$$\Delta s = B^*(U_o)[\Delta U] . \quad (12)$$

**Born image perturbation**

The simplest way of computing image/wavefield perturbations is by simple subtraction of the wavefields for the background image $U_o$ from the wavefield of a better image $U$:

$$\Delta U_b = U - U_o . \quad (13)$$

Equation (13) is only valid for small perturbations of the wavefields ($\Delta U_b << 1$). In practice, this requirement means that the cumulative phase difference between the two different wavefields is small at all frequencies.

If this condition is satisfied, we can compute a slowness perturbation which corresponds to the Born approximation:

$$\Delta s_b = B^*(U_o)[\Delta U_b] . \quad (14)$$
In practice, the small perturbation requirement is hard to meet, since small slowness differences amount to large cumulative phase differences. Thus, with the wavefield perturbation definition in equation (13), we can only handle small slowness perturbations.

**Rytov image perturbation**

An alternative to the wavefield perturbation definition in equation (13) is given by the Rytov approximation. If we can estimate the accumulated phase differences between the two wavefields at every depth level

\[ \Delta \Phi_r = \Phi - \Phi_o , \]  

we can compute another wavefield perturbation using the relation:

\[ \Delta U_r = U_o i \Delta \Phi_r \]  

which is directly derived from equation (9). With this definition of the wavefield perturbation, we can compute another slowness perturbation which corresponds to the Rytov approximation

\[ \Delta s_r = B^* (U_o) [\Delta U_r] \]  

using the same backprojection operator.

**Differential image perturbation**

Various methods can be used to improve images created with inaccurate, reference velocity models. Residual migration (Stolt, 1996; Al-Yahya, 1989; Etgen, 1990) is one such option, although we could use other methods like residual moveout or image continuation.

If image enhancement is done with a Stolt-type residual migration operator $S$ (Stolt, 1996; Sava, 2000, 2003), we can write a relation for an improved image $R$ derived from a reference image $R_o$

\[ R = S(\rho) [R_o] , \]  

where $\rho$ is a spatially varying scalar parameter indicating the magnitude of residual migration at every image point. We can compute a linearized image perturbation by a simple first-order expansion relative to the parameter $\rho$

\[ \Delta R_\rho = \left. \frac{dS}{d\rho} \right|_{\rho=\rho_o} [R_o] \Delta \rho , \]  

from which we can compute a wavefield perturbation $\Delta U_\rho$ using the adjoint of the imaging operator.

The operator $\left. \frac{dS}{d\rho} \right|_{\rho=\rho_o}$ can be computed analytically, since it only depends on the background image, while $\Delta \rho$ can be picked at every location from a suite of images computed using different values of $\rho$ (Sava and Biondi, 2003). Similar formulations are possible for
Figure 3: Fat rays for an image perturbation defined using the Born equation (21). The slowness anomaly is gradually increasing from (a) to (c). Only the smallest anomaly is correctly handled by the Born image perturbation.

other kinds of operators (e.g., normal residual moveout), and are not restricted to residual migration, in general, or to Stolt residual migration, in particular.

With this definition of the wavefield perturbation, we can compute another slowness perturbation:

$$\Delta s_a = B^* (U_o) [\Delta U_a] .$$  (20)

**DISCUSSION**

We have shown in the preceding section several methods that can be used to create image perturbations to be inverted for slowness perturbations:

$$\Delta R_b = \sum_o (U - U_o)$$  (21)

$$\Delta R_r = i \sum_o (U_o (\Phi - \Phi_o))$$  (22)

$$\Delta R_a = \frac{dS}{d\rho} \bigg|_{\rho=\rho_0} [R_o] \Delta \rho .$$  (23)

Figure 3 shows the slowness backprojection for image perturbations computed using the definition in equation (21). From top to bottom, the panels correspond to increasing magnitudes of the slowness anomalies. In panel (a), the accumulated differences between the background and correct images is small, such that the Born approximation holds and the backprojection creates simple “fat rays” (Woodward, 1992). However, as the slowness anomaly increases, the fat rays are distorted by sign changes (b), and/or by the characteristic ellipsoidal side-lobes (c).

Figure 4 shows the slowness backprojection for image perturbations computed using the definition in equation (22). From top to bottom, the panels correspond to increasing magnitudes of the slowness anomalies. In panels (a) and (b), the accumulated phase differences
between the background and correct images are small and do not wrap. Backprojection by WEMVA also creates simple undistorted fat rays. At large magnitudes, however, the phases become large enough to wrap, and backprojection from image perturbations defined by equation (22) fails (c).

Both equation (21) and equation (22) employ the same operator for backprojection. The difference is in the method we use to define the wavefield perturbation. For equation (21) we use the difference between the complete wavefields, with the constraint of small total wavefield difference. For equation (22) we use the difference between the cumulative phases, which does not impose a constraint on the actual size of the wavefields. Thus, using equation (22), we could in principle handle arbitrarily large slowness perturbations.

However, the phases in equation (22) need to be unwrapped to obtain a meaningful wavefield differences. In complex environments, wavefields are can be quite complicated, and it is not at all trivial to estimate and unwrap their phases. Therefore, even if we could in theory use equation (22) for arbitrarily large perturbations, in practice we are constrained by our ability to unwrap the phases of complicated wavefields. Figure 5 shows the fat rays corresponding to the different magnitudes of slowness anomalies when the phases have been unwrapped.

The more practical alternative we can use to create image perturbations using equation (23) is illustrated in Figure 6. In this case, the fat rays are not distorted at any magnitude of slowness anomaly, behavior which is similar to that of the unwrapped Rytov.

The explanation for this behavior lies in the definition in equation (23). The image perturbation is created by estimating the gradient of the residual migration change on the background image, followed by scaling with the appropriate $\Delta \rho$ picked from a suite of images obtained with different values of the parameter $\rho$.

In essence, we are using the information provided by the background image to infer the direction and magnitude of the image change. There is no limitation to how far we can go from the background image similar to the limitations of the Born and Rytov definitions. However, since we are employing a first order linearization, the accuracy of the differential image

---

**Figure 4:** Fat rays for an image perturbation defined using the Rytov equation (22). The slowness anomaly is gradually increasing from (a) to (c). The phases are not unwrapped, thus the largest anomaly is not described correctly. [paul1-RYTOV3b.xbsryt](#)
Figure 5: Fat rays for an image perturbation defined using the Rytov equation (22). The slowness anomaly is gradually increasing from (a) to (c). The phases are unwrapped, thus all anomalies are described correctly.

Figure 6: Fat rays for an image perturbation defined using the differential equation (23). The slowness anomaly is gradually increasing from (a) to (c). All anomalies are described correctly.

perturbation decreases with increasing $\Delta \rho$.

Figure 7 is a schematic illustration of the transformation implied by equation (23). We can create enhanced images either by nonlinear residual migration, or by a first-order linearization around the background image. In principle, the accuracy of this approximation decreases with increasing $\Delta \rho$. Therefore, in practice we cannot go arbitrarily far from any given background image and we need to run several non-linear iterations involving slowness inversion, re-migration and re-linearization.

Figure 8 is a summary of fat rays computed using the methods described in the preceding section. The magnitude of the slowness anomaly increases from left to right. From bottom to top we show the fat rays for the Born definition, the Rytov definition without phase unwrapping, the Rytov definition with phase unwrapping, and the differential definition.

The four panels on the left are identical, since all methods work as well for small anomalies. In the middle four panels, the fat ray obtained with the Born definition starts to break,
Figure 7: A sketch of the approximations done when computing image perturbations with equation (23). In this plot, each multi-dimensional image is schematically depicted by a point. We compute a linear approximation of an image corresponding to a spatially varying $\rho$ from the gradient information computed on the background image and the $\Delta \rho$ picked from a suite of images. The accuracy of the linear approximation decreases with increasing $\Delta \rho$. 

while the Rytov (with and without phase unwrapping) and differential approaches work well. Finally, the panels on the right correspond to the highest anomaly, when only the Rytov with phase unwrapping and differential methods work.

CONCLUSIONS

We analyze various options for computing image perturbations for wave-equation migration velocity analysis. Our three choices are Born (amplitude difference), Rytov (phase difference) and differential image perturbations derived analytically from residual migration operators.

We show that we can use the linearized WEMVA operator to invert all three types of image perturbations. We find that Rytov WEMVA is possible if we unwrap phases, but fails without.

The differential image perturbation can be used for arbitrarily large slowness anomalies, but its accuracy decreases with deviation from the background image. For simple cases, the differential image perturbation method is equivalent with a phase-unwrapped Rytov method.

REFERENCES


Figure 8: Summary plot for WEMVA using image perturbations constructed with different methods and with anomalies of increasing magnitude. From bottom to top we show the fat rays for the Born definition, the Rytov definition without phase unwrapping, the Rytov definition with phase unwrapping, and the differential definition. The magnitude of the slowness anomaly increases from left to right.
Dip-dependent residual moveout

Guojian Shan and Biondo Biondi

ABSTRACT
Residual moveout is an effective tool for interval velocity analysis for depth migration. However, the conventional residual moveout method assumes that the reflectors in the subsurface are all flat, and thus estimates curvature parameters for velocity analysis that are inaccurate for steeply dipping reflectors. In this paper, we develop a dip-dependent residual moveout method, which is performed in the Fourier domain and handles dip effects.

INTRODUCTION

Prestack depth migration produces superior images of complex structures and has become practical with the continued improvement of computer performance. However, the method is sensitive to errors in the velocity model and requires a good velocity model from velocity analysis.

Velocity analysis, whether ray-based (Clapp, 2001) or wave-equation-based (Sava and Biondi, 2003), requires estimation of the curvature parameters of common image gathers (CIGs). Residual moveout and residual migration (Sava, 2003) are two leading methods for estimating the curvature parameters. Residual migration, the more accurate of the two methods, constrains only global velocity changes. However, residual moveout, although assuming stationary rays, estimates the curvature parameters of CIGs affected by local velocity changes.

For 2-D problems, Biondi and Symes (2003) show that the image point should be on a line that passes through the crossing point between source and receiver rays and is normal to the reflector. When the migration velocity is not the true velocity, the migrated image points move in the direction normal to the reflector at the reflection point. However, conventional residual moveout methods assume that the image point moves only vertically. For steeply dipping reflectors, the normal direction deviates from vertical, approaching horizontal for nearly vertical reflectors. Thus, the vertical moveout assumption in the conventional residual moveout methods may lead to significant errors in the estimation of the curvature parameters for velocity analysis.

In this paper, we suggest a new way to perform residual moveout, which we call dip-dependent residual moveout. The key idea is to perform phase shifts in the Fourier domain instead of image-point shifts in the space domain. In the Fourier domain, dip information

1email: shan@sep.stanford.edu, biondo@sep.stanford.edu
can be obtained easily, so that the new method can account for dip effects and result in more accurate curvature parameters.

**ANGLE-DOMAIN CIGS**

Most current migration velocity analysis methods (Sava and Biondi, 2003; Clapp, 2001) are based on the curvature information from Angle-Domain CIGs (ADCIGs), which are created from the migration cube (Sava and Fomel, 2003). When the migration velocity is the true velocity, the ADCIG at a reflection point is a flat line. When the migration velocity is inaccurate, the curvature parameters estimated from ADCIGs can be back-projected and inverted for velocity updates. Biondi and Symes (2003) demonstrates that in an ADCIG cube, the image point lies on the line normal to the apparent reflector dip, passing through the point where the source ray intersects the receiver ray. Figure 1 shows the ADCIGs at reflection points of different dip angles. Under the stationary-raypath assumption, the shift of the image point along the normal direction is (Biondi and Symes, 2003)

\[
\Delta n = \frac{1 - \rho}{1 - \rho(1 - \cos \alpha)} \frac{\sin^2 \gamma}{\cos^2 \alpha - \sin^2 \gamma} z_0 \mathbf{n},
\]

(1)

where \( \rho \) is the constant scaling factor of the slowness, \( \alpha \) is dip angle of the reflector, \( \gamma \) is the opening angle, \( \mathbf{n} \) is the normal direction vector of the reflection point, and \( z_0 \) is the depth at the reflection point. For flat reflectors, the shift (1) reduces to

\[
\Delta n = (1 - \rho) \tan^2 \gamma z_0 \mathbf{n}.
\]

(2)

Equations (1) and (2) can be used to estimate the curvature parameters for velocity analysis. However, to estimate the curvature parameters caused by local velocity perturbation, we don’t consider the effect of depth \( z_0 \).

**RESIDUAL MOVEOUT IN THE FOURIER DOMAIN**

Conventional residual moveout is usually performed in the space domain. It assumes that the reflectors are flat, and thus image points only need to move vertically to flatten the ADCIGs. However, in dip-dependent residual moveout, the image points are moved along the direction normal to the reflector to flatten the ADCIGs.

It is difficult to calculate an accurate dip map for an image in the space domain, but the dip map can be obtained easily in the Fourier domain by the following relation:

\[
\tan \alpha = \frac{k_x}{k_z},
\]

(3)

where \( \alpha \) is the dip angle, and \( k_x \) and \( k_z \) are the wave numbers of \( x \) and \( z \) direction, respectively. In the local versions of equation (1) and (2), the shift along the direction normal to the
Figure 1: ADCIGs with different dip angles. $\mathbf{n}$ is the normal direction vector, and $\gamma$ is the opening angle.

reflectors depends only on the dip angle $\alpha$ and the opening angle $\gamma$. In the Fourier domain, the dip angle $\alpha$ is calculated by equation (3), and the opening angle $\gamma$ by

$$\tan \gamma = \frac{k_h}{k_z},$$

where $k_h$ is the wave number of offset $h$ (Sava and Fomel, 2003). Let $\Delta n$ be the shift along the direction normal to the reflectors. From the geometric relation, the shift along the normal direction is equivalent to a horizontal shift $\Delta x = \sin \alpha \Delta n$ followed by a vertical shift $\Delta z = \cos \alpha \Delta n$. Figure 2 shows, for an ADCIG at a reflection point with a dip angle of $\alpha$, the relation between normal direction shift $\Delta n$, horizontal shift $\Delta x$ and vertical shift $\Delta z$. A shift in the space domain is equivalent to a phase shift in the Fourier domain. Let $I(x, z, h)$ be the image cube obtained by migration, and $I(k_x, k_z, k_h)$ is its Fourier transformation. Then in the space domain, the shift $\Delta n$ along the direction normal to the reflector is equivalent to a phaseshift of $k_x$ followed by a phaseshift of $k_z$:

$$I^{\text{RMO}}(k_x, k_z, k_h) = I(k_x, k_z, k_h) \cdot e^{i k_x \Delta x} \cdot e^{i k_z \Delta z}$$

in the Fourier domain.

**SYNTHETIC DATA EXAMPLE**

In this section, we present a 2-D synthetic example to verify the theory. The reflector of the synthetic model has a spherical shape with a radius of 500 m. Modeling and migration have been done in Biondi and Symes (2003). Figure 3 shows the images obtained with the
true (left panel) and a too low ($\rho = 1.04$) (right panel) velocity models. Figure 4 shows the conventional residual moveout and dip-dependent residual moveout with a same velocity scale parameter. The vertical axes of these three panels are the direction normal to the reflector and the horizontal axes are the opening angle. The ADCIG is at $x = 3100\text{m}$, $z = 720\text{m}$, where the dip angle is about 45 degree. Panel (a) is the ADCIG without residual moveout. Panel (b) is the ADCIG after conventional residual moveout. Panel (c) is the ADCIG after dip-dependent residual moveout. In Figure 4, the energy out of 40 degree is the noise due to the illumination. Figure 5 shows the ADCIGs at $x = 2900\text{m}$, $z = 1000\text{m}$, where the dip angle is about 60 degree. The energy out of 20 degree is the noise due to the illumination. Comparing to 45 degree case, we find that the 60 degree case has a smaller aperture and is more affected by the noise. Since the dip-dependent residual residual moveout moves image points in the direction normal to the reflector (vertical axes in Figure 4 and Figure 5 ), rather than the vertical direction ($z$ direction) as in the conventional residual moveout processing, the former method flattens the ADCIGs better.

CONCLUSION

We develop a dip-dependent residual moveout formulation and test it on a synthetic dataset with steep dips. We perform the new residual moveout method as a phase shift in the Fourier domain, and obtain more accurate curvature parameters than with traditional spatial residual moveout. Although a 2-D example is tested in this paper, the method can easily be extended to 3-D.
Figure 3: Migration results: Left panel is the image migrated with the correct velocity. Right panel is the image migrated with a too slow velocity.

Figure 4: Compares between conventional and dip-dependent residual moveout: (a) ADCIG at $x = 3100\, \text{m}, \, z = 720\, \text{m}$ without residual moveout; (b) ADCIG at $x = 3100\, \text{m}, \, z = 720\, \text{m}$ after conventional residual moveout; (c) ADCIG at $x = 3100\, \text{m}, \, z = 720\, \text{m}$ after dip-dependent residual moveout. The vertical axes $n$ is the direction normal to the reflector. The dip angle at $x = 3100\, \text{m}, \, z = 720\, \text{m}$ is about 45 degree.
Figure 5: Compares between conventional and dip-dependent residual moveout: (a) ADCIG at $x = 2900m$, $z = 1000m$ without residual moveout; (b) ADCIG at $x = 2900m$, $z = 1000m$ after conventional residual moveout; (c) ADCIG at $x = 2900m$, $z = 1000m$ after dip-dependent residual moveout. The vertical axes $n$ is the direction normal to the reflector. The dip angle at $x = 2900m$, $z = 1000m$ is about 60 degree.

REFERENCES


Velocity estimation for seismic data exhibiting focusing-effect AVO (Part 3)

Ioan Vlad, Biondo Biondi, and Paul Sava

ABSTRACT

Given a satisfactory image perturbation, Target Image Fitting (TIF) Wave Equation Migration Velocity Analysis (WEMVA) successfully produces a velocity model that completely eliminates focusing-effect AVO anomalies through prestack depth migration. However, TIF WEMVA ceases to converge to the desired result when the Born approximation is not fulfilled because the starting guess is too far from the correct velocity. Fortunately, there are several possible remedies to this problem. Extracting the image perturbation proves somewhat more complicated than expected, but new possible solutions based on differential semblance have been identified.

INTRODUCTION

Focusing-effect AVO (FEAVO) consists of focusing seismic wavefield amplitudes through velocity lenses too small to generate proper triplications. The amplitude effects are large enough to thwart proper AVO analysis in the affected area (Kjartansson, 1979). Migration with a velocity model containing the lenses that cause FEAVO eliminates the effect (Bevc, 1994). However, the traveltime effects are too small to allow classical velocity analysis approaches such as Dix inversion or traveltime tomography to succeed. New approaches are needed to deal with such small velocity anomalies.

In previous work (Vlad and Biondi, 2002; Vlad, 2002) Wave-Equation Migration Velocity Analysis (WEMVA; introduced by Biondi and Sava (1999)) was shown to resolve FEAVO-causing velocity anomalies, by optimizing image quality in the angle domain after prestack depth migration. The first step of the proof consisted in showing that FEAVO anomalies are recognizable in a field dataset migrated and transformed to the angle domain. WEMVA iterates by migrating with the current velocity model, extracting an image perturbation, and converting it into a velocity model update by inverting a linearized downward continuation operator. Vlad (2002) then showed that the inversion operator did not distort FEAVO anomalies into becoming unrecognizable, despite the linearizing approximations.

While the previous studies only demonstrated the possible suitability of WEMVA as a tool for the given problem, this paper shows the results of successfully running WEMVA on a FEAVO-affected synthetic dataset. In particular, it shows that the velocity lenses are

1email: nick@sep.stanford.edu, biondo@sep.stanford.edu, paul@sep.stanford.edu
found, and that the image migrated with the updated velocity model no longer exhibits FEAVO effects. This paper also explores the limitations of the type of WEMVA used and different ways of extracting the image perturbation to be fed into the inversion.

FEAVO BEFORE MIGRATION

We have generated two 2-D synthetic datasets, using two velocity models. The purpose of the first one is to show that FEAVO anomalies are indeed eliminated when migrating with a velocity model obtained by WEMVA. The purpose of the second one is to test what happens when the velocity anomalies are outside the range within which TIF WEMVA is accurate. We will denote the two datasets by the names “WEMVA works” and “WEMVA breaks”.

The velocity models are depicted in the same color scale in the upper panels of Figures 1 (“WEMVA works”) and 2 (“WEMVA breaks”). The background velocity is 2000 m/s for both. The peak anomalies, from left to right, as departures from the background, are (in m/s): -153, -188, +231 for the “WEMVA works” velocity model and -586, -766, +519 for the “WEMVA breaks” velocity model.

The “WEMVA works” velocity anomalies cause visible amplitude focusing, but minimal traveltime anomalies (middle panel in Figure 1). The intense “WEMVA breaks” velocity lenses, on the other hand, cause visible triplications and departures from hyperbolicity (middle panel in Figure 2). The specific signature of FEAVO before migration can be examined in both cases by squaring each of the values in the data cube, then summing along the time axis. The results of this operation is shown in the lower panels of Figures 1 and 2 respectively. The axes of the image are midpoint and offset, and specific “V” shapes are visible for each of the three velocity anomalies.

It is interesting to compare these FEAVO effects with those present in the only field dataset currently available to us that exhibits FEAVO anomalies. The dataset in question is the Grand Isle dataset analyzed by Kjartansson (1979) – the original dataset on which FEAVO was defined. The FEAVO effects in it are comparable with those in the lower panels of Figures 1 and 2 because the background velocities are approximately similar. Given that the intensity of the effects is linked to the magnitude of the velocity anomalies, we can empirically estimate that the deviation of the Grand Isle velocity anomalies from the background is around 200-300 m/s. Since the “thickness” of the V’s is linked to the dimensions of the velocity lenses, we can also infer that they are approximately 20-30m in diameter.

WEMVA ELIMINATES SYNTHETIC FEAVO

As described by Sava and Symes (2002), the image perturbation to be fed into WEMVA can be created by Target Image Fitting (TIF), or by Differential Semblance Optimization (DSO, Symes and Carazzone (1991), Chauris and Noble (1998)). We chose the TIF approach because the purpose of this test was to determine conclusively whether WEMVA-produced velocity models can eliminate realistic FEAVO effects independent of image perturbation extraction.
Figure 1: “WEMVA works” case: velocity model, dataset and FEAVO effects. The velocity model has the same color scale as the one in the upper panel of Figure 2.
Figure 2: “WEMVA breaks” case: velocity model, dataset and FEAVO effects
Since the dataset is synthetic and the correct velocity model known, we were able to generate an optimal image perturbation by subtracting the image migrated with the current velocity from the one migrated with the correct velocity.

We used the “WEMVA works” dataset. The starting estimate for the first WEMVA iteration was the 2000 m/s background velocity. FEAVO effects in a slice from the angle-domain image migrated with the background velocity are shown in the upper panel of Figure 4.

We ran 10 conjugate-gradient solver iterations, then updated the velocity model. The resulting model is shown in the upper panel of Figure 3. Its peak anomalies as departures from the background, in m/s, from left to right, are: -74, -84, +99. We migrated with this velocity model. FEAVO effects in a slice from this image are shown in the middle panel of Figure 4. They are weaker now, but still visible.

We performed another WEMVA inversion loop, with 10 solver iterations, starting from the updated velocity model. The resulting velocity model is shown in the bottom panel of Figure 3. Its peak anomalies as departures from the background, in m/s, from left to right, are: -105, -123, +149. The angle-domain image obtained by migrating with this new velocity model are shown in the lower panel of Figure 4. The FEAVO effects are no longer recognizable. When WEMVA’s assumptions, discussed in the next section, are satisfied and the image perturbation can be extracted in a satisfactory manner, the inversion process converges and the resulting velocity field is accurate enough to reliably eliminate FEAVO effects.

Figure 3: **Upper panel**: velocity model updated after one WEMVA iteration. **Lower panel**: velocity model updated after a second WEMVA iteration. Both panels are represented in the same color scale.
Figure 4: Angle-domain slices at a depth of 155m. Upper panel: from the image obtained with the constant background velocity; Middle panel: from the image obtained with the velocity in the upper panel of Figure 3; Lower panel: from the image obtained with the velocity in the lower panel of Figure 3.

WEMVA LIMITATIONS

The most important limitation of WEMVA is the range of validity of its assumptions. In the case of the Born approximation in TIF WEMVA, phase differences caused by the true velocity anomalies in the image must be smaller than $\pi/4$. The corresponding velocity anomaly magnitudes vary with the range of frequencies present in the dataset and with the spatial extent of the anomalies.

We performed one WEMVA pass with 10 conjugate-gradient solver iterations on the “WEMVA breaks” dataset. In this instance the phase differences caused by the velocity anomalies are too large to be satisfactorily approximated in a Born manner. The velocity update thus obtained is presented in Figure 5. It does not converge toward the true anomalies like the corresponding update in the upper panel of Figure 3. Since this occurs when using the ideal image perturbation, inaccuracies in extracting the perturbation would only amplify the trend.

There are several possible avenues for avoiding either the Born approximation limitations or the Born approximation altogether: 1. obtaining a very good velocity starting guess with other velocity analysis methods; 2. creating the image perturbation in connection with residual migration (Sava and Biondi, 2001); 3. employing linearizations other than Born (Sava and Fomel, 2002); 4. inverting for the lower frequencies first (Pratt, 1999); 5. using DSO WEMVA
Figure 5: Velocity model updated after one WEMVA iteration comprising 10 solver iterations on the “WEMVA breaks” dataset. 

(Sava and Symes, 2002) where the image perturbation is extracted directly from the image using a specific operator during the inversion, without the need to find a corrected image.

**EXTRACTING IMAGE PERTURBATIONS**

The upper panel of Figure 6 shows the optimal image perturbation, found by subtracting the image obtained by prestack depth migration with the constant (2000 m/s) background velocity from that obtained with the correct velocity. While the absolute values of the amplitudes in the FEAVO anomalies are stronger than the background as expected, the polarities change every time the seismic wavelet alternates sign. Vlad (2002) proposed an extraction approach that would involve sweeping the image space with summations along the analytically computable shapes of the FEAVO effects. To do that, we would have to use absolute or squared values to prevent the polarity alternation from cancelling the summed values. However, discarding the information stored in the sign of the anomaly forfeits the ability to distinguish between positive and negative velocity anomalies.

There are several ways of circumventing this problem. They can be simple and cheap, but case-specific (a priori knowledge whether the lenses are faster or slower than the background). They can be more general, but also more complex and expensive (do WEMVA iterations assuming a single sign until all anomalies of that sign are eliminated, allowing anomalies of the other sign to increase; then when only increases in the image perturbations are noticed, switch the sign and start again). Even with the polarity problem, more consideration will be given in the future to the focus-filter-spread approach as envisaged by Vlad (2002) because of its potential power of eliminating the non-FEAVO noise and of exploiting the entirety of the FEAVO morphological characteristics.

The polarity conundrum mentioned above can be eliminated by identifying anomalies by their local, rather than global, characteristics. The middle panel of Figure 6 shows the image perturbation for the TIF WEMVA obtained by eliminating the DC component and by performing several other small adjustments (correcting with a cosine of the angle factor for amplitude decay, windowing away irrelevant energy at higher angles, etc). The lower panel in Figure 6
presents the velocity update obtained after one WEMVA iteration. Additional iterations did not change the result too much. The results may have improved had we used a weighting operator to mask the unresolved anomalies at higher angles (so they would not be fitted when inverting).

CONCLUSIONS

Given a satisfactory image perturbation, TIF WEMVA successfully produces a velocity model that completely eliminates FEAVO anomalies through prestack depth migration. However, TIF WEMVA ceases to converge to the desired result when the Born approximation is not fulfilled because of a velocity starting model that is too far from the correct model. Fortunately, there are several possible fixes to this problem. Extracting the image perturbation proves somewhat more complicated than expected, but we have identified and are studying possible new solutions to the problem.

REFERENCES


Figure 6: **Top:** ideal image perturbation for TIF WEMVA obtained by subtracting the image produced with the correct velocity and the image produced with the background velocity; **Middle:** image perturbation for TIF WEMVA obtained by processing the background velocity image in a DSO manner; **Bottom:** Velocity model obtained after one WEMVA iteration taking as input the image perturbation shown in the middle panel. [nick1-f6] [CR]
Short Note

Reflection tomography with depth control

Weitian Chen, Robert G. Clapp, and Biondo Biondi

INTRODUCTION

Reflection tomography (Stork and Clayton, 1991; Stork, 1992; Clapp, 2001) is one of the most effective and widely used velocity estimation methods. However, reflection tomography has velocity-depth ambiguity problem (we do not know how much a traveltime error is due to a velocity error and how much is due to a reflector misposition) because of insufficient source-receiver offset and lateral velocity changes (Bickel, 1990; Lines, 1993; Ross, 1994; Tieman, 1994).

From borehole data, we can obtain the correct reflection positions around the borehole. The normal shift between the correct reflection positions and the apparent reflection positions can be linearly mapped to the traveltime perturbation along the normal ray (van Trier, 1990). However, from borehole data, we can only obtain the correct position for only a few reflection points along the borehole. In this paper, we assume all the reflection points within a local area around the borehole have the same normal shift. The normal ray traveltime perturbation for all these reflection points are then backpropagated simultaneously with the reflection traveltime perturbation. We applied this scheme on a synthetic model and obtained a better inversion result than using reflection tomography without this control. We further discuss how to improve this method for more complex datasets.

BASIC PRINCIPLES OF REFLECTION TOMOGRAPHY

For reflection data, there are two things that can cause traveltime perturbation: slowness perturbation $\Delta s$ and reflector movement $\Delta r$. Figure 1 demonstrates the basic geometry for the reflection tomography problem. Here, $l_n$ is the normal ray, $l_o$ is the offset ray with aperture angle $\theta$, and $\Delta r$ is the normal shift between exact reflector and apparent reflector.

According to Fermat’s principle, the traveltime perturbation caused by slowness perturbation, $\Delta t_o$, can be mapped approximately to slowness perturbation by the following linear
relationship:

$$\Delta t_0 \approx \int_{l_o} \Delta sdl_o.$$  \hspace{1cm} (1)

According to van Trier (1990), the reflector movement $\Delta r$ can be assumed equal to the residual zero-offset migration of the reflector. Consequently, $\Delta r$ can be mapped to the slowness perturbation along the normal (zero-offset) ray, which can be expressed by the following equation

$$\Delta r \approx -\frac{1}{s_0} \int_{l_n} \Delta sdl_n,$$  \hspace{1cm} (2)

where $s_0$ is the local slowness at the reflection point. According to Fermat’s principle, the reflector movement $\Delta r$ causes $-2\Delta r \cos \theta$ change in ray length. As a result, the traveltime perturbation caused by reflector movement is

$$\Delta t_n \approx 2s_0 \Delta r \cos \theta \approx -2\cos \theta \int_{l_n} \Delta sdl_n.$$  \hspace{1cm} (3)

Figure 1: Geometry for reflection wave propagation. $l_o$ is the offset ray. $l_n$ is the normal ray. $\theta$ is the aperture angle of the offset ray. $\Delta r$ is the normal shift between apparent reflector and correct reflector.

By summing $\Delta t_o$ and $\Delta t_n$, we can obtain the total traveltime perturbation:

$$\Delta t = \Delta t_o + \Delta t_n \approx \int_{l_o} \Delta sdl_o - 2\cos \theta \int_{l_n} \Delta sdl_n.$$  \hspace{1cm} (4)

Equation (4) provides a linear relationship between reflection traveltime perturbation $\Delta t$ and slowness perturbation $\Delta s$ which can be used for backpropagation.

For migration velocity analysis, reflection traveltime perturbation, $\Delta t$, can be effectively obtained from angle-domain common-image-gathers (ADCIG) (Clapp, 2001). Figure 2 is a sketch of ADCIG. Here, $\Delta r$ is the normal shift between correct reflection position and apparent reflection position; $\Delta r_2$ is the residual moveout; and $\Delta r_1$ is the total normal shift. According to Biondi and Symes (2003), the traveltime perturbation $\Delta t_o$ can be calculated from total normal shift by following equation:

$$\Delta t_o \approx 2s_0 \cos \theta \Delta r_1.$$  \hspace{1cm} (5)
Combining equation (5) and (3), we can obtain reflection traveltime perturbation from residual moveout by following equation:

\[ \Delta t \approx 2s_0 \cos \theta \Delta r_2. \]  
(6)

As we can see, \( \Delta t_0 \) and \( \Delta t_n \) can provide independent data information for velocity inversion. However, from reflection data, we can not obtain them separately since the reflection data alone can not provide the exact reflector position. Instead, we can only obtain \( \Delta t \) which is the sum of \( \Delta t_0 \) and \( \Delta t_n \) for reflection tomography.

Figure 2: Illustration of calculating \( \Delta t \) for reflection tomography from angle-domain CIGs

**DEPTH CONTROLLED REFLECTION TOMOGRAPHY**

Borehole seismic data can provide the exact position of the subsurface reflectors at the borehole. Therefore, from borehole data, we can obtain \( \Delta r \) at these locations. \( \Delta r \) can be backpropagated to slowness perturbation using equation (2). In order to integrate the depth control to reflection tomography effectively, we transfer the reflector movement \( \Delta r \) to traveltime perturbation along normal ray, \( \Delta t_n \). The traveltime perturbation \( \Delta t_n \) then is backpropagated to slowness perturbation using following equation:

\[ \Delta t_n \approx - \int_{l_n} \Delta s dl_n \]  
(7)

Combining equation (4) and (7), we can obtain a depth controlled reflection tomography (DCRT) scheme:

\[ \Delta t \approx T_r \Delta s \]  
(8)

\[ \Delta t_n \approx T_n \Delta s \]  
(9)

\[ 0 \approx \epsilon A \Delta s \]  
(10)

Here, fitting goal (8) and (9) correspond to equation (4) and (7), respectively. Fitting goal 10 is the model styling goal. \( A \) is a regularization operator. We use a Laplacian as regularization operator for the following application.
RESULTS

We apply our DCRT scheme to a synthetic anticline model. Figure 3a and b show the correct velocity model and initial velocity model, respectively. There are seven reflectors for this synthetic model, which are overlayed on Figure 3b. A well is assumed at surface location $x = 10km$. Figure 4 shows the migration result using initial velocity model. The overlayed points are those reflection points we choose for backpropagation. For each reflector, we can only obtain the exact position for the reflection points where the well and the reflector cross. We make the assumption that all the reflection points within a local area around those reflection points have same normal shift. In Figure 5a, we show the assumed normal shift for all the reflection points we choose for adding depth control, which was used for DCRT in this application. As a comparison, in Figure 5b, we show the exact normal shift for those reflection points. After multiplying local slowness, we can obtain the corresponding traveltime perturbation along the normal ray. Figure 6a and b shows the assumed and exact normal ray traveltime perturbation, respectively.

Figure 3: Synthetic anticline velocity model, a), and initial velocity model, b)

Figure 4: The migration result using initial velocity model. Overlayed are the reflection points chosen for adding depth points to reflection tomography.
Figure 5: a) The approximated normal shift and b) the exact normal shift, for all the reflection points chosen for adding depth control.

Figure 6: a) The approximated normal ray traveltime perturbation and b) the exact normal ray traveltime perturbation, for all the reflection points chosen for adding depth control.
The three panels in Figure 7, from left to right, show the inversion result of reflection tomography, DCRT, and their difference. Notice the obvious difference around the borehole between DCRT result and regular reflection tomography result. Figure 8 and Figure 9 shows the migration and angle-domain common-image-gathers (ADCIGs) using velocity obtained by regular reflection tomography and DCRT, respectively. The surface positions for 5 ADCIGs, from left to right, are 9.6, 9.8, 10, 10.2, 10.4 km, respectively. Notice the improvement of the image and the reduced residual moveout around the borehole after using the DCRT method.

Figure 7: a) Reflection tomography result, b) DCRT result, and c) the difference between a) and b)

Figure 8: a) Migration result and b) ADCGIs, using velocity from reflection tomography. The surface location for 5 ADCIGs, from left to right, are 9.6, 9.8, 10, 10.2, 10.4 km, respectively.
DISCUSSION

In this shortnote, we presented a method to add depth control to reflection tomography. We transfer the exact reflector movement, which can be obtained from borehole data, to the traveltime perturbation along the normal ray. The traveltime perturbation along normal ray provides another data fitting goal for reflection tomography. By simultaneously backpropagating normal ray traveltime perturbation and reflection traveltime perturbation, we can improve the inversion result.

In the completed work, we obtained the normal shift between the correct reflection point and apparent reflection point, $\Delta r$, then transfer it to the traveltime perturbation along normal ray for backpropagation. Notice in Figure 2, by summing $\Delta r$ and residual moveout $\Delta r_2$, we can obtain the total normal shift $\Delta r_1$, which can be transfered to traveltime perturbation along the offset ray according to equation (5). Therefore, instead of using equation (7) for backpropagation, we can backpropagate the traveltime along the offset ray by using the following linear relationship:

$$\Delta t_o = \int_{l_o} \Delta s dl_o.$$  \hspace{1cm} (11)

An obvious advantage of backpropagating along the offset ray is that we can obtain better ray coverage. Backpropagating along normal ray can only obtain velocity along normal ray direction, whereas, when backpropagating along offset ray, we can obtain a much wider ray coverage with varying aperture angle of offset ray.

In the completed work, we did not apply any weighting between fitting goal (8) and (9). With an appropriate weighting scheme, the DCRT should improve the inversion result.

Another way to improve DCRT result is to use a spatially-varying Lagrange multiplier $\epsilon$. 
We assume all the reflection points within a local area have the same normal shift. Such an approximation is more reliable for the reflection points near the borehole, and less reliable for those points away from the borehole. In order to take this into account during inversion, we can apply spatially-varying $\epsilon$ to the model styling goal (10). We can apply small $\epsilon$ for the area close to the well to emphasize the data fitting, whereas big $\epsilon$ for the area away from the well to emphasize the model styling.

ACKNOWLEDGMENT

The first author would like to thank Stanford Graduate Fellowships for generously providing financial aids.

REFERENCES


Short Note

Frequency-dependent velocity analysis?

Ioan Vlad¹

**INTRODUCTION**

There are two reasons for exploring imaging and velocity analysis with a frequency-dependent velocity model. First, significant dispersion may occur naturally in the volume of rock investigated by the seismic experiment. Second, the goal of Wave Equation Migration Velocity Analysis (WEMVA, Biondi and Sava (1999)) is image optimization, which some successful approaches (Pratt, 1999) perform by inverting one frequency at a time, from lower to higher frequencies. The velocity model that optimizes the image may vary with frequency within the physical error bar of the velocity analysis solution. In what follows, I will examine only whether dispersion from natural causes is strong enough to be considered in WEMVA.

Frequency dispersion of elastic waves is a frequent occurrence that is more noticeable in surface waves than in body waves (Pedersen et al., 2003). Some authors (Wang, 2001) highlight its importance and the infrequency with which it is explored by rock physics. Other articles (Marion et al., 1994) analyze the conditions in which dispersion appears. Techniques for modeling dispersion (Robinson, 1994), or for migration with a frequency and attenuation-dependent velocity model (Mittet et al., 1995) are currently available.

Dispersive phenomena in the dataset on which FEAVO was originally defined² are depicted in Figure 8 of Vlad and Biondi (2002). I further investigate quantitatively whether dispersion plays a large enough role to warrant performing the velocity analysis separately for each frequency, and migrating with a \( v(x, z, \omega) \) velocity model. The expense and coding overhead associated with using a frequency-dependent velocity model in wave-equation migration would be negligible, because wave-equation imaging and migration velocity analysis are parallelized over frequencies.

**MEASURING DISPERSION**

A simple and efficient way to measure the dispersion of plane waves in a shot gather has been described by McMechan and Yedlin (1980; 1981). The method transforms a shot gather from

¹email: nick@sep.stanford.edu
²Grand Isle (LA), shallow water, courtesy of Dr. Ralph Shuey, Gulf Science and Technology Company.
the original (t,x) coordinate system to a (velocity, frequency) coordinate system, on which velocity can be picked as a function of frequency for each dispersion mode present in the data. The events to be analyzed this way must be linear. Surface waves fulfill this condition. An example of applying this method to a shot gather with ground roll stronger than the reflections is presented in Figure 1. The method consists of two steps: first the shot gather is slant-stacked, then it is Fourier-transformed along the time axis to obtain the power spectrum.

To investigate whether reflected events are dispersive, I transform hyperbolas (from a CMP gather this time, not from a shot gather) into straight lines using a \( T^2 - X^2 \) stretch. Linear event slopes will correspond to RMS velocities. I then apply the previously described algorithm for transforming the gather into an easy-to-pick (velocity, frequency) panel.

In practice, however, transforming to \( T^2 - X^2 \) coordinates then slant stacking is equivalent to summing along hyperbolic paths – a bare-bones Radon transform with no amplitude corrections. This cuts the cost in half and is the way I implemented it. Similar hyperbolic Radon transforms were used in velocity analysis, but they have been replaced in common practice by semblance panels. To obtain a semblance panel, instead of summing along hyperbolas, a statistical measure of coherence is computed along them. Feeding such a semblance panel into the dispersion analysis flow did not show improvement. The mathematics of the McMechan and Yedlin (1980) method seem to be geared towards actual summations along velocity-dependent paths. Deconvolution (which usually improves velocity analysis) did not result in visible changes either, which means that predictive deconvolution with the same filter for all traces in the dataset did not affect dispersion.

The results (bandpassed to eliminate DC components) are presented in Figure 2. Dispersion curves are much less visible than in the case of surface waves, but this may be because the number of reflections in the left panel of Figure 2 is considerably greater than the number of distinct linear events in the left panel of Figure 1. This may lead to mutually destructive interference in the dispersion analysis panel. The inferable trend does not warrant a variation in velocity of more than 100m/s across the reflection seismic frequency spectrum. These conclusions from a single Mississippi Delta 2-D line cannot be generalized to all seismic reflection data. However, I did not find any literature or anecdotal evidence pointing to variations in body wave velocity of more than 100 m/s across the useful reflection seismic frequency range.

**CONCLUSIONS**

For the analyzed dataset, the velocity of seismic reflection events varies less than 100m/s with frequency across the usual exploration spectrum. There is no information available indicating greater variation for other datasets. Future wave-equation migration velocity analysis tools may allow for still greater variation, since the computational expense is not higher, and the inversion is more stable if the output of velocity analysis for low frequencies is taken as a starting model for higher frequencies. The dispersion in reflected waves does not warrant discarding proven velocity analysis and imaging software tools, but is worth considering when trying to build significantly more accurate ones.
Figure 1: Example of applying the McMechan and Yedlin (1980) dispersion analysis method. **Left:** input shot gather. **Right:** output dispersion panel. Curved events represent various dispersion modes for surface waves. Other events correspond to different arrivals in the data (direct arrival, reflections, etc.).
Figure 2: Application of the McMechan and Yedlin (1980) dispersion analysis method to a CMP gather from which the surface waves have been eliminated by muting and f-k filtering. **Left:** virtual input. **Right:** output of the method. The frequency range is different from Figure 1 because the data has been processed to highlight reflections and eliminate surface waves.
REFERENCES


Marion, D., Mukerji, T., and Mavko, G., 1994, Scale effects on velocity dispersion: From ray to effective medium theories in stratified media: Geophysics, 59, no. 10, 1613–1619.


Interval velocity estimation using edge-preserving regularization

Alejandro A. Valenciano, Morgan Brown, Mauricio D. Sacchi, and Antoine Guitton

ABSTRACT

We test two edge-preserving forms of model regularization in a least-squares implementation of Dix formula that leads to interval velocities with sharp edges in the \((\tau, x)\) plane. This characteristic in the interval velocity may be desirable in geologic environments with abrupt changes in velocity, like: carbonate layers, salt bodies, and strong faulting.

INTRODUCTION

Interval velocity estimation is a central problem in reflection seismology (Claerbout, 1999). Without an estimate of seismic velocity, we would be unable to transform prestack seismic data into an interpretable image. Advanced velocity estimation techniques (Clapp, 2001; Biondi and Sava, 1999) have been developed to estimate interval velocity in complex geological environments, though the cost of these methods is often considerable.

In the early stages of prospect evaluation, an inexpensive interval velocity estimate is often desired. The Dix equation (Dix, 1952) analytically inverts root-mean-square (RMS) velocity for interval velocity as a function of time. In addition to many physical shortcomings (assumption of a stratified \(v(z)\) earth), Dix inversion suffers from numerical problems that lead to poor velocity estimates. Dix inversion is unstable when RMS velocities vary rapidly, and may produce interval velocities with unreasonably large and rapid variations. For this reason, the problem is often cast as a least-squares problem, with is regularized in time with a differential operator to penalize such rapid variations and to produce a smooth result (Clapp et al., 1998).

While temporal smoothness may often be justified from a geologic stand point \(v(z)\), in some cases velocity can change abruptly (e.g., carbonate layers, salt bodies, strong faulting). In these cases we desire a regularization technique that gives smooth velocity in most places, but which preserves sharp “geologic” interval velocity contrasts when they are present, without requiring pre-defined edges to be supplied.

In this paper we present two automatic edge-preserving regularization methodologies for the least-squares implementation of Dix formula. The first uses IRLS to effectively change the norm of the problem to permit a “spiky” or “sparse” model residual, which leads to a “blocky” velocity model. The second uses an isotropic edge detector, the gradient magnitude,
in a nonlinear scheme to compute a measure of the edges of the model. This edge measure is then used as a model residual weight, along the lines of Clapp et al. (1998) and Lizarralde and Swift (1999).

**DIX EQUATION AS A LEAST-SQUARES PROBLEM**

The Dix Equation states the nonlinear relationship between RMS velocity and interval velocity. Harlan (1999) linearizes the relationship and solves the problem using a Gauss-Newton nonlinear iteration. However, the problem is linear in the square of the interval velocity. This “linear” problem was solved by Clapp et al. (1998). They apply a preconditioned least squares optimization to “invert” Dix equation, with spatial smoothness constraints.

Let us rewrite for completeness the Dix equation as a least-squares fitting goal:

\[ W(Cu - d) \approx 0 \]  \hspace{1cm} (1)

where \( u \) is the unknown model, a vector of squared interval velocities. \( d \) is the known data, a vector of squared RMS velocities multiplied by time. \( C \) is the causal integration operator. \( W \) is a data residual weighting function, which is proportional to our confidence in the RMS velocities.

Fitting goal (1) is notoriously unstable to high frequency variations in RMS velocity, and moreover, it is under-determined in the sense that only strong reflections really qualify as “data”. Therefore, Clapp et al. (1998) supplement the system with a regularization term which penalizes “wiggliness”. In our case we use first order derivatives, but as we will see later, other rougheners can be used:

\[ W(Cu - d) \approx 0 \]  \hspace{1cm} (2)

\[ \epsilon_t D_t u \approx 0 \]  \hspace{1cm} (3)

\[ \epsilon_x D_x u \approx 0 \]  \hspace{1cm} (4)

where \( D_t \) and \( D_x \) are first-order finite differences derivatives in time and midpoint, respectively, and the scalars \( \epsilon_t \) and \( \epsilon_x \) balance the relative importance of the two model residuals.

In hard rock environments like carbonates, velocities tend to be homogeneous for intervals, with abrupt discontinuities at changes in lithology. There, the desire for a blocky interval velocity model is well-justified.

**“BLOCKY” MODELS**

In the following subsections we introduce two schemes to weight the model residuals in equations (3) and (4) to preserve sharper edges in the estimated \( u \).
Edge-preserving regularization with the Cauchy norm

Imagine that the model residuals in equations (3) and (4) consisted of spikes separated by relatively large distances. Then the interval velocity $u$ would be piecewise smooth with jumps at the spike locations, which is what we desire. However in solving (2)–(4) we use the least-squares criterion – minimization of the $\ell_2$ norm of the residual. Any spikes in the residual will be attenuated. To do this, the solver smooths the velocity across the spike location.

It is known that the $\ell_1$ norm is less sensitive to spikes in the residual (Claerbout and Muir, 1973; Darche, 1989; Nichols, 1994). $\ell_1$ norm minimization makes the assumption that the residuals have an exponential distribution, a “long-tailed” distribution relative to the Gaussian. Here, we compute nonlinear model residual weights which force a Cauchy distribution, another long-tailed distribution which approximates an exponential distribution (Youzwishen, 2001).

We perform the following non linear iterations: starting with $Q_\tau^0 = Q_x^0 = I$, at the $k^{th}$ iteration the algorithm solves

\begin{align}
W(Cu^k - d) &\approx 0 \\
\epsilon_\tau Q_\tau^{k-1} D_\tau u^k &\approx 0 \\
\epsilon_x Q_x^{k-1} D_x u^k &\approx 0
\end{align}

where

\begin{align}
Q_\tau^{k-1} &= \frac{1}{1 + \left( \frac{D_\tau u^{k-1}}{\alpha_\tau} \right)^2}, \\
Q_x^{k-1} &= \frac{1}{1 + \left( \frac{D_x u^{k-1}}{\alpha_x} \right)^2},
\end{align}

and $u^k$ is the result of the $k^{th}$ nonlinear iteration, $Q_\tau^{k-1}$ and $Q_x^{k-1}$ are the $(k - 1)^{th}$ diagonal weighting operators, $D_\tau$ and $D_x$ are the first order derivatives in time and midpoint, $I$ is the identity matrix, the scalars $\alpha_\tau$ and $\alpha_x$ are the trade-off parameters controlling the discontinuities in the solution, and the scalars $\epsilon_\tau$ and $\epsilon_x$ balance the relative importance of the two model residuals.

Edge-preserving regularization with the gradient magnitude

In the previous section we changed the norm of the minimization problem to prevent the roughener from smoothing over the edges of the model. In this subsection we shift from a statistical to a more mechanical approach to attain the same goal.

To preserve the edges of the model Clapp et al. (1998) propose adding a weight to zero the model residual at the edges. Lizarralde and Swift (1999) implement a similar approach for
the inversion of VSP data for interval velocity. This approach requires human intervention for reflector picking. We want to design a weight which de-weights edges in the model residual, but which is estimated automatically.

The 2-D gradient magnitude is a good isotropic edge-detection operator that can be used to calculate the diagonal weights. As we show in the deblurring problem (Valenciano et al., 2003), using the gradient magnitude we can iteratively obtain sharp edges.

We perform the following non-linear iterations: starting with $Q_{\|\nabla\|}^0 = I$, at the $k^{th}$ iteration the algorithm solves

$$W(Cu^k - d) \approx 0$$
$$\epsilon Q_{\|\nabla\|}^{k-1} \nabla^2 u^k \approx 0$$

(8)

where

$$Q_{\|\nabla\|}^{k-1} = \frac{1}{1 + \|\nabla u^{k-1}\|_\alpha}$$

(9)

and $u^k$ is the result of the $k^{th}$ nonlinear iteration, $Q_{\|\nabla\|}^{k-1}$ is the $(k - 1)^{th}$ diagonal weight operator, $\|\nabla\|$ is the gradient magnitude, $\nabla^2$ is the Laplacian operator, $I$ is the identity matrix, the scalar $\alpha$ is the trade-off parameter controlling the discontinuities in the solution, and the scalar $\epsilon$ balances the relative importance of model and data residuals.

**REAL DATA RESULTS**

We used 125 CMP’s from a 2-D prestack dataset acquired in the Gulf of Mexico. This data is suitable for using Dix equation, since the main reflectors are flat. The area is heavily faulted which may imply strong lateral velocity variations with sharp edges to preserve.

First, we performed velocity analysis on each CMP, and then use an auto-picker to pick the maximum stacking power that corresponds to the best RMS velocity at each CMP location. The value of the stacking power at the auto-picked RMS velocity was used as a quality measure of the data, and used as the data residual weight (W) in equations (2), (5), and (8). Clapp (2003) shows an alternative way to calculate the data residual weights based on a multiple realization of the RMS velocity. His approach looks promising since the RMS velocity average is used as the data and the RMS velocity variance is used as the data residual weight. Figure 1 shows the auto-picked RMS velocity and a stack of the CMP’s.

Figures 2, 3, and 4 show the interval velocities resulting from solving the inverse problems stated in equations (2), (5), and (8) respectively. We also show in figure 5 a graph comparing the three methods and the RMS velocity used as input data at two CMP locations.

The resulting interval velocity models show what the regularization was designed to do. In figure 2 the resulting interval velocity is smooth in time and space. Figure 3 shows sharp edged rectangular shapes all over the image, looking reasonable in the faults but in general geological unappealing. Figure 4 shows sharp objects with more geological appeal.
Figure 1: A) Auto-picked RMS velocity of one CMP from a 2-D prestack dataset, B) raw RMS velocity, C) the CMP in question, D) and the stacked data using the raw RMS velocity.
Figure 2: Interval velocity computed by 2-D inversion of the RMS velocity (equation (2)).

Figure 3: Interval velocity computed by 2-D inversion of the RMS velocity using Cauchy norm (equation (5)).
The preferential shapes can also be seen in the diagonal weight operator. Figures 6 and 7 show $Q_r^N$ and $Q_x^N$, the last nonlinear iteration diagonal weight operator in equation (7). Notice the two preferential directions in what the edges are preserved. Figure 8 shows $Q_{||\mathbf{v}||}^N$, the last nonlinear iteration diagonal weight operator in equation (9). Notice the isotropic behavior of the diagonal weight calculated using the gradient magnitude operator.

**CONCLUSIONS**

Dix formula can be implemented in a nonlinear least-squares inversion scheme to attain interval velocities with sharp edges in the $(\tau,x)$ plane. In this paper we present two automatic edge-preserving regularization methodologies to achieve this goal.

The first uses IRLS to effectively change the norm of the problem to permit a “spiky” or “sparse” model residual, which leads to a “blocky” velocity model. The second uses an isotropic edge detector, the gradient magnitude, in a nonlinear scheme to compute a measure of the edges of the model. This edge measure is then used as a model residual weight.

Both methods give the expected results when applied in a 2-D real data set acquired in the Gulf of Mexico. Even though, the gradient magnitude method shows sharp objects with more geological appeal than the “blocky” method.
Figure 5: Comparison of the results of solving the inverse problems stated in equations (2) Smooth, (5) Cauchy norm, (8) Gradient Magnitude, and the RMS velocity at the midpoint positions 8.04 and 12.194 km.
Figure 6: $Q^N_r$ is the last nonlinear iteration diagonal weight operator in equation (6).

Figure 7: $Q^N_x$ is the last nonlinear iteration diagonal weight operator in equation (7).
Figure 8: $Q_{||Y||}^N$ is the last nonlinear iteration diagonal weight operator in equation (9).

REFERENCES


Basis pursuit for geophysical inversion

Brad Artman and Mauricio Sacchi

ABSTRACT
Accepting an inversion principle, it is possible to design an algorithm to meet any requirements or constraints. Given the context of representing a signal with an arbitrary overcomplete dictionary of waveforms within the signal, one can design an inversion algorithm that will focus energy into a small number of model space coefficients. With this principle in mind, an analogy to linear programming is developed that results in an algorithm with: the properties of an $l^1$ norm, a small and stable parameter space, definable convergence, and impressive denoising capabilities. Linear programming methods solve a nonlinear problem in an interior/barrier loop method similar to iteratively reweighted least squares (IRLS) algorithms, and are much slower than a least squares solution obtained with a conjugate gradient method. Velocity scanning with the hyperbolic radon transform is implemented as a test case for the methodology.

INTRODUCTION
Chen et al. (1999) introduces the idea of “basis pursuit” (BP) as a principle to aid in precision analysis of complex signals. The central tenet of the presentation lies in the assumption that a minimal number of constituent members of a model space dictionary are responsible for the signal being analyzed. Therefore, the distribution of energy across the coefficients associated with dictionary atoms (be they sinusoids, wavelets, chirps, velocities, etc.) during the analysis of a signal should be uneven and sparse. This idea of a sparse model space is counter to a smoothly distributed $l^2$ norm inversion, and thus a different algorithm needs development to satisfy these requirements.

The development of this inversion principle into an algorithm can take any number of forms. Guitton and Symes (2003) choose the Huber norm to effect an $l^1$-like measure of the inversion error. We will cast the problem through the primal-dual Linear Programming (LP) structure resulting in a methodology wherein the concept of convergence is central to the algorithm. This fact has two important consequences. Firstly, the precision of the output model space is one of (the very few) input parameters. Secondly, the parameter space is insensitive to manipulation as compared to $\epsilon$ in regularized least squares problems or the cutoff value needed for Huber norm approaches.

Conventionally, LP methods deal almost exclusively in a small world of conveniently short

1email: brad@sep.stanford.edu, sacchi@phys.ualberta.ca
time signals such as bursts of speech. Application of these methods to geophysical problems of much larger size may prove prohibitive. While at its best the complexity of this method can be comparable to IRLS, in practice the method is usually several times slower to produce optimal solutions.

As an example of the method, the hyperbolic radon transforms are used as analysis operators of seismic and synthetic seismic data. An exploration of the method comparable to Guitton and Symes (1999) will be used to highlight the strengths and weaknesses of the method compared to conventional least squares and Huber norm inversion for velocity from seismic data.

**DEVELOPMENT**

The development of the BP principle begins with the assumption that the dictionary of indexed elementary waveform atoms, $\phi_i$, used to analyze the signal is overcomplete. This means that a signal, $s$, is decomposed to the sum of all the atoms in the dictionary scaled by a scalar energy coefficient, $\alpha_i$ where the last index number is much larger than the length of the input signal with $n$ points. Thus,

$$s = \sum_i \alpha_i \phi_i \quad (1)$$

has many possible configurations where the waveform atoms are linearly independent, though not necessarily orthogonal. This differs from conventional Fourier analysis where a signal of length $n$ is decomposed into a list of $n$ independent, orthogonal bases. There, the representation of the signal through the waveform dictionary is unique if the dictionary is simply complete. Many such dictionaries exist including wavelets, cosines, chirps, etc. An overcomplete dictionary can be devised through the combination of multiple dictionaries, or oversampling a single choice. Using the Fourier example again, a four-fold overcomplete dictionary would be defined as one where the indexed frequencies of sinusoidal waveforms sweep through the standard frequency range definition, but at a four-fold finer sampling interval.

Given an overcomplete dictionary for analysis, BP simply states the goal of choosing the one representation of the signal whose coefficients, $\alpha_i$, have the smallest $l^1$ norm. The pursuit then is searching through the dictionary for a minimum number of atoms that will act as bases and represent the signal precisely. Mathematically, this takes the form

$$\min ||\alpha||_1 \quad \text{subject to} \quad \Phi \alpha = s \quad (2)$$

where $\Phi$ is the decomposition operator associated with the indexed waveforms $\phi_i$ above. This formulation, demanding an $l^1$ minimization, has been investigated through IRLS by many other authors in the geophysical literature (Nichols, 1994; Darche, 1989; Taylor et al., 1979). Endemic to this approach, however, are issues related to defining convergence and the difficulties in choosing appropriate values from the parameter space. Huber norm tactics also share

---

2 Donoho and Huo (1999) prove for one dimensional time series, with a “highly sparse” model-space representation, that the proposed decomposition is unique. Also included are definitions of concepts such as “highly sparse”, and “sufficiently sparse”.

---
similar issues. BP however has adopted the infrastructure of primal-dual Linear Programming to solve the system.

Linear programming (LP) has a large literature and several techniques by which to solve problems of the form

$$\min_{x} c^T x \quad \text{subject to} \quad Ax = b, \; x \geq 0.$$  \hfill (3)

This optimization problem minimizes an objective cost function, $c^T x$, subject to two conditions: some data and an operator, and a positivity requirement. The underlying proposition of Chen et al. (1999) is the equivalence of the equations (2) and (3). If the analysis operator $\Phi$ is used as $A$, the cost function $c$ is set to ones, and use the signal $s$ as the forcing vector $b$, the BP principle falls neatly into the LP format.

Having recognized the format, one needs only choose among the available methods to solve an LP problem. BP chooses an interior point (IP) method as opposed to the simplex method of Claerbout and Muir (1973) to solve the LP. IP introduces a two loop structure to the solver: a barrier loop that calculates fabrication and direction variables, and an interior solver loop that uses a conjugate gradient minimization of the system derived from the current status of the solution and the derived directions. Heuristically, we begin with non-zero coefficients across the dictionary, and iteratively modify them, while maintaining the feasibility constraints, until energy begins to coalesce into the significant atoms of the sparse model space. Upon reaching this concentration, the method pushes small energy values to zero and completes the basis pursuit decomposition. For this reason, models where the answer has a sparse representation converge much more rapidly. If the model space is not sparse, or cannot be sparse, this is the wrong tool.

A particular benefit of adopting this structure to solve the decomposition problem lies in the primal-dual nature of interior point (as opposed to simplex) LP methods. For any minimization problem, the primal, there is a equally valid maximization problem, its dual, that can be simultaneously evaluated during the iterative solving process. It is known that at convergence, the primal and dual variables are the same. Thus, the separation between the primal and dual model space can be evaluated to provide a rigorous test for convergence. The two loop structure then takes the form of setting up a least squares problem that solves for the update to the dual variable, then evaluating the new solutions from this answer and testing the duality gap. When the difference is less than the tolerance proscribed by the user, the algorithm terminates.

The interior loop that uses a CG solver solves a system that looks like

$$\min \| \left( \begin{array}{c} D^{1/2}A^T \\ \delta I \end{array} \right) \Delta y - \left( \begin{array}{c} D^{1/2}(t - X^{-1}v) \\ r/\delta \end{array} \right) \|_2$$ \hfill (4)

where matrix $D$, and vectors $t, v, r$ are calculated quantities from the IP method, $X = \text{diag}(x)$, $\Delta y$ is the barrier step update for the dual variable, and $\delta$ is an input parameter that controls the regularization of the search direction for the IPLP solver. $A^T$ is the original operator from equation (3), so we can see that during the inner loop we will be evaluating the operator and its adjoint many times. If there is not either a sparse matrix representation or a fast implicit algorithm for this operator, this step makes using the method prohibitive. Importantly, because
the interior loop solves for model updates, it is not very important to produce precise results, and thus any relatively accurate solution will help the iterative improvement of the next barrier step.

During this brief explanation of the concept, only two user defined parameters have been mentioned. The first is the desired tolerance associated with the duality gap, and the second is $\delta$. In a realistic implementation there are, of course, a few more parameters. They are largely functions of these two mentioned, automatically updated in the IPLP algorithm, or constants.

$\delta$ has interesting properties to discuss. Saunders and Tomlin (1996) describes in detail the mathematics associated with this approach. The regularization of the search direction introduced with $\delta$ actually makes this one of a class of LP problems called perturbed, and introduces a minor change to equation (3). Leaving those details to the reader, I will address the choice of the parameter. Thankfully, the BP will converge if $\delta$ is anywhere between 1 and $4 \times 10^{-4}$. If the regularization parameter is small, the BP converges slowly, but to a more precise solution that honors the data exactly. As $\delta \to 1$, the central equations solved in the inner loop become highly damped resulting in three affects: 1) speedy convergence, 2) effective denoising, and 3) less rigor attached to the “subject to” condition in equation (3). The penalty is that the result lacks some sparsity and precision compared to the result with a small value.

IMPLEMENTATION

The condition of using an overcomplete dictionary will hopefully be satisfied by choosing the number of model space variables to be roughly two times the number of data points. While Chen et al. (1999) uses a minimum of four times oversampling for impressive super resolution results with the Fourier transform, their examples are normally of the order $n = 1 \times 10^3$. Larger dictionaries result in too slow processing for testing purposes.

Due to the fact that the LP infrastructure imposes a positivity constraint on the solution, we are forced to solve for a model space twice as large as we would choose with both negativity and positivity constraints, and then combine the two. We will use the hyperbolic radon transform (HRT) as the analysis operator, $\Phi$ in equation (2) or $A$ in equation (3). This operator is normally approximately 1% full and, therefore, a tractable operator to use for this method. As such, the programming is implemented with a sparse matrix approach rather than operators as it is traditional at SEP.

EXPERIMENTS

Synthetic problems

We use the same data panels as Guitton and Symes (1999) to compare the results of BP to Huber norm results as they are both $l^1$ minimization strategies. Synthetic tests are designed

---

Guitton and Symes (1999) include least squares inversion results with a CG method for each example as well.
evaluate the method’s performance to invert the HRT under circumstances of missing data, a slow plane wave superimposed on hyperbolic events, and spiky data space noise. Two field CMP’s are also analyzed and compared to the Huber norm result. The dissimilarities in axes origin and formatting of the plots are unimportant. The spikes are at the correct values, and the important thing to note in the plots is the distribution of energy around the model space. The HRT operator used for this implementation has no AVO qualities, although the synthetics were modeled with a wavelet and amplitude variation.

Figure 1 shows the results of the BP method when addressing the problem of missing data. We can see that the predicted data looks as accurate as the Huber norm result. The velocity model space, however, shows considerable difference. Notice the resolution increase over the same range of velocities and the lack of appreciable chatter away from basis atoms. With this figure, and those to come dealing with the synthetic examples, the predicted data looses the wavelet character and the amplitude seems to diminish with depth.

Figure 2 shows the results of the BP method when a slow plane wave is superimposed on the data. The overcomplete dictionary now shows significantly less chatter about the velocity panel, and very distinguishable differences in the predicted data panel are emerging on the right side of the CMP where the events cross. Combination operators, linear and hyperbolic hybrid operators (Trad et al., 2001), may be ideal for this situation, but have not been tried exhaustively yet.

Figure 3 shows the results of the BP method when randomly distributed spikes contaminate the data. BP had significant trouble resolving this model. Unlike the Huber norm implementations of Guitton and Symes (1999), the method has no capacity to utilize the properties of the \( l^1 \) norm in the data space, and so cannot handle the large spikes. Manually limiting the number of outer loops to seven was the only way to avoid instability. However, this point is easy to find as the duality gap begins increasing and the CG solver fails repeatedly to attain the input tolerance. Regardless, the predicted data looks pretty bad, and while the model space is sparse, the atoms that do have energy are inappropriate.

Real data

Two field data CMP’s are also analyzed with the BP algorithm. With an order of magnitude increase in size, as well as much energy in the data that leads to a more full model space, convergence does not seem as well behaved for real data. For the CMP with bad traces in Figure 4, we needed only 10 minutes of CPU time. Normally, this computation requires user intervention to stop the process as it looked to become unstable. The multiple ridden data of Figure 7, however, required about 40 minutes to compute. The model space used in both examples was only approximately 2.5-fold overcomplete, and this fact may contribute to the problems experienced. Interestingly, with the regularization parameter \( \delta = 1 \), the algorithm has a drastic denoising effect as well.

Figure 4 compares the predicted data from CG least squares inversion, the Huber norm inversion, and the BP inversion. The noise reduction of the near traces is remarkable and deserves further research. A very powerful linear noise train bounds the data to the right,
Figure 1: Missing data. Left column is velocity model space. Right column is data space. Row 1 is input velocity and modeled data. Row 2 is Huber norm inversion and modeled data. Row 3 and 4 are BP inversion and predicted data results. Row 3 model space has approximately the same number of model variable as data points. Row 4 has four times the number of model variables. [brad2-miss] [NR]
Figure 2: Slow plane wave superposition. Same format as explained in the caption of Figure 1.
Figure 3: Randomly spiked data. Same format as explained in the caption of Figure 1.
which we hypothesize is the result of the near offset noise in the raw data. Figure 6 contains four powerful noisy traces between 2200 - 2700 m/s. Also noticeable is the tendency for the forward model to bifurcate real events into a correct and a fast event such as at 1.25 seconds. Replacing the high amplitude ringing trace with zeros did not fix the problem.

Figure 4: Modeled data after inversion compared to original a CMP that suffers from bad traces and substantial near offset noise.

Figure 7 compares the predicted data from CG least squares inversion, the Huber norm inversion, and the BP inversion. The BP solver had great difficulty with the multiples infested CMP. The garbage in the low velocity range above 1.4 seconds is troublesome. This may contribute to the problems analyzing this data, as I may not have made the model space large enough to achieve the necessary overcompleteness, or the linear events are not well described by the hyperbolic dictionary. This type of data is a good candidate to try the amalgamated linear/hyperbolic radon transform of Trad et al. (2001).
Figure 5: Velocity panel comparison. The different output of the different programs makes direct comparison impossible. The left panels scan to much higher velocity than was necessary.
Figure 6: Presentation of the envelope of the velocity scan provides a better look at the location of the focus of energy. The several vertical noise traces probably lead to the poor quality of the predicted data (right panel, Figure 4). Disappointingly, some events have bifurcated.

CONCLUSION

In the case of the synthetic examples using an overcomplete dictionary, the velocity panels are truly sparse, and the convergence of the result is achieved within 20 outer loop iterations. Interestingly, with these sparse model examples, the number of CG iterations required drastically diminishes through the process. The first step usually requires around 230 iterations, and then immediately drops by at least an order of magnitude for the next several loops. After around seven outer loops, it bumps back up to between 50 and several hundred CG iterations until it achieves convergence. If the number of CG iterations is limited, especially during the first few outer loops, the method does not recover within the limits of the authors’ patience. Seven minutes was the longest run of the overcomplete decompositions, while the merely complete dictionaries require only about one minute. In all cases, the sparsity of the velocity space is remarkable, and thus warrants further research into the use of this tool.

Chen et al. (1999) mandates the use of an overcomplete dictionary, and Donoho and Huo (1999) proves the uniqueness of the solution only for overcomplete dictionaries. While it is true that super resolution, for which an impressive Fourier decomposition method is presented by Chen et al. (1999), can only be achieved with the overcomplete dictionary, the use of such with these experiments doubles the computational cost and provides only marginally better results than a model space approximately the same size as the data space. It may be possible that the extension of the problem to handle positive and negative results may naturally provide sufficient overcompleteness during the solving process, but this is undetermined. It could also be that this is a contributing factor to the difficulties handling the real (larger) data sets, especially the multiples example.

The real data examples showed positive, though as yet inconclusive, quality results with
Figure 7: Modeled data after inversion compared to original a CMP that suffers from internal multiples and strong ground roll.
this method. Usable, though not optimal, results can be achieved within a user terminated
dozen loops of BP, though if allowed to run longer, a better product may result. The bifur-
cation/event manufacturing in Figure 4 is unacceptable and needs serious attention in further
inquiry to the usefulness of the technique. The multiple data of Figure 7 gives a reasonable
result, though not sufficiently better, to warrant the extra cost when compared to the CG ver-
sion. These data exhibit a tendency to require several hundred CG iterations during the first
few outer loops, and then drop to single digits for a dozen iterations before becoming unstable,
after which the process is terminated.

A reason for the difficulty the algorithm has in convergence is its lack of understanding of
the bandlimited nature of the data. This quality of the data makes the BP inversion unstable
as it spends too much effort trying to solve for a sparse model of spikes that is inappropriate.
A frequency domain Radon transform may well perform better with this thought in mind, as it
will not carry the infinite frequency assumption through the modeling operator. Alternatively,
a second bandpass operator could be chained with an operator similar to the one used in
this example. In this manner, the composite operation could produce more stable and less
demanding results from the IPLP algorithm.

A tangent concept that this work introduces evolves from the idea of the waveform dictio-
naries used in any type of inversion. Rather than accepting the frequency, wavelet, or chirp
dictionaries from mathematical context, it may be possible to compile “seismic waveform”
dictionaries that have characteristics more directly suitable to the structures and features regu-
larly exhibited in seismic data. This could include pinch-outs, lapping configurations, and/or
variations of simple hyperbolas. These could be useful for many other situations with different
algorithms, and would not be restricted to this particular inversion implementation.
REFERENCES


Short Note

Deblurring using edge-preserving regularization

Alejandro A. Valenciano and Morgan Brown

INTRODUCTION

In this short note, we test various edge-preserving regularization schemes in the context of deblurring a text image with random noise. The blurred text image was created by Nagy and O’Leary (2003a) as a test case. Even if the blurring filter is known exactly, as it is in this case, sharp features are nearly in the nullspace of the filter which we must “invert”, or deblur. Those eigenvalues of the filter matrix corresponding to edges may be well below the noise level, and thus difficult or impossible to resolve.

We know that letters should be homogeneous for intervals (piecewise constant), thus it makes sense to impose model smoothness using a regularization operator. But letters also have abrupt discontinuities, thus using a regularization operator that imposes model smoothness considerably degrades our ability to discern the letter edges.

We present three separate strategies which allow the model regularization operator to preserve edges in the deblurring process. All are implemented as regularization operators in an unconstrained least-squares deblurring problem.

The data

The data $g$ (figure 1B) (Nagy and O’Leary, 2003a) were generated by taking the original image $f$ (figure 1A), posted in the answer paper (Nagy and O’Leary, 2003b), multiplying it by a non-stationary convolution matrix $K$ and then adding random noise. The noise prevents us from completely recovering the initial image.

REGULARIZATION SCHEMES

In Nagy and O’Leary (2003a), regularization is used to make the least-squares deblurring problem less sensitive to the noise. Figures 2 and 3 show the result of using the identity operator for regularization. This is comparable to the results presented in Nagy and O’Leary

\[email: valencia@sep.stanford.edu, morgan@sep.stanford.edu]
We go a step further by using regularization to impose *a priori* information on the solution of the problem. We exploit the fact that letters should be homogeneous for intervals (piecewise constant) with abrupt discontinuities between them.

A way of forcing a function to be piecewise constant is forcing its first order derivatives to be sparse. Thus, using a first order derivative regularization operator and forcing the model residuals to follow a Cauchy distribution should make the letters “blocky” and preserve the letter edges (Youzwishen, 2001). Obtaining model residuals following a Cauchy distribution can be achieve posing the inverse problem as Iterative Reweighed Least Squares (IRLS) (Darche, 1989).

The Cauchy norm first order derivative edge-preserving regularization fitting goal was
Figure 3: Comparison between Figures 2A and 2B; A) Slice $y = 229$ and B) Slice $x = 229$.

set following the nonlinear iterations: starting with $Q_x^0 = Q_y^0 = I$, at the $k^{th}$ iteration the algorithm solves

$$
Kf^k - g \approx 0 \\
\epsilon Q_x^{k-1} D_x f^k \approx 0 \\
\epsilon Q_y^{k-1} D_y f^k \approx 0
$$

(1)

where

$$
Q_x^{k-1} = \frac{1}{1 + \left(\frac{D_x f^{k-1}}{\alpha} \right)^2}^{1/2},
$$

$$
Q_y^{k-1} = \frac{1}{1 + \left(\frac{D_y f^{k-1}}{\alpha} \right)^2}^{1/2}. 
$$

(2)

$K$ is a non-stationary convolution matrix, $f^k$ is the result of the $k^{th}$ nonlinear iteration, $Q_x^{k-1}$ and $Q_y^{k-1}$ are the $(k-1)^{th}$ diagonal weighting operators, and $D_x$ and $D_y$ are the first order derivative operator in the $x$ and $y$ directions, $I$ is the identity matrix, the scalar $\alpha$ is the trade-off parameter controlling the discontinuities in the solution, and the scalar $\epsilon$ balances the relative importance of the data and model residuals.

We were successful in obtaining what we designed the algorithm to produce. The result is blocky in the $x$ and $y$ directions (Figures 4 and 5). However, the derivatives in the $x$ and $y$ directions do not produce an isotropic result. We know that letters often have round shapes. Thus, the problem could benefit from using a more isotropic operator to calculate the diagonal weights.
Figure 4: A) Original image, B) Deblurred image using LS with the first order derivative edge-preserving regularization.

Figure 5: Comparison between Figures 4A and 4B; A) Slice $y = 229$ and B) Slice $x = 229$. 
Gradient magnitude and Laplacian regularization

In the previous section we changed the norm of the minimization problem to prevent the roughener from smoothing over the edges of the model. In this subsection we shift from a statistical to a more mechanical approach to attain the same goal.

The gradient magnitude (||\nabla||) is an isotropic edge-detection operator that can be used to calculate the diagonal weights. Unfortunately, it is a nonlinear operator thus couldn’t be used for regularization. Instead we used the Laplacian, which is a regularization operator used in several SEP applications (Claerbout and Fomel, 2001).

The gradient magnitude edge-preserving regularization fitting goal was set following the nonlinear iterations: starting with \( Q_{||\nabla||}^0 = I \), at the \( k^{th} \) iteration the algorithm solves

\[
Kf^k - g \approx 0 \\
\epsilon Q_{||\nabla||}^{k-1} \nabla^2 f^k \approx 0
\]

where

\[
Q_{||\nabla||}^{k-1} = \frac{1}{1 + \frac{||\nabla f^k||}{\alpha}}.
\]

\( K \) is a non-stationary convolution matrix, \( f^k \) is the result of the \( k^{th} \) nonlinear iteration, \( Q_{||\nabla||}^{k-1} \) is the \((k-1)^{th}\) diagonal weighting operators, \( I \) is the identity matrix, \( ||\nabla|| \) is the gradient magnitude, \( \nabla^2 \) is the Laplacian operator, the scalar \( \alpha \) is the trade-off parameter controlling the discontinuities in the solution, and the scalar \( \epsilon \) balances the relative importance of the data and model residuals.

Figures 6 and 7 show a considerable improvement over Figures 4 and 5. They are noise free but keep the round features of the original image. However, since we are not imposing blockiness on the model but rather on the derivative of the model (using the Laplacian as the regularization operator), the edges are not as sharp as the previous case.
Gradient magnitude and smooth regularization

Until now we have used a roughener as the regularization operator to take into account the prior knowledge that letters are piecewise constant. Turning it off at the edges and on away from the edges. To do the opposite also makes sense, using a smoothing operator at the edges and turning it off away from the edges doesn’t make the model smoother away from the edges but increases the sharpness of the model at the edges.

The gradient magnitude with triangular smoothing edge-preserving regularization fitting goal was set following the nonlinear iterations: starting with $W_0 = I$, at the $k^{th}$ iteration the algorithm solves

$$K f^k - g \approx 0$$
$$\epsilon W^{k-1} \Delta f^k \approx 0$$

where

$$W^{k-1} = ||\nabla f||^{k-1}.$$  

$K$ is a non-stationary convolution matrix, $f^k$ is the result of the $k^{th}$ nonlinear iteration, $W^{k-1}$ is the $(k - 1)^{th}$ diagonal weighting operators, $I$ is the identity matrix, $\Delta$ is triangular smoother operator, and the scalar $\epsilon$ balances the relative importance of the data and model residuals.

Figures 8 and 9 show what we were expecting, i.e. the edges look sharp, but away from the edges the noise dominates. However, we can see in figure 8 some details that we are not visible in Figures 6, 4, and 2. Looking that a combination of the styling goals in equations (3) (smoothing away from the edges) and (5) (roughening at the edges) should give a good result.
Figure 8: A) Original image, B) Deblurred image using LS with the edge-preserving regularization gradient magnitude with triangular smoothing.

Figure 9: Comparison between Figures 8A and 8B; A) Slice $y = 229$ and B) Slice $x = 229$. 

**Vision is the art of seeing what is invisible to others.**
CONCLUSIONS

Introducing an edge-preserving regularization helps to take into account prior knowledge about letter statistics into the least-squares deblurring problem. The proposed gradient magnitude and Laplacian regularization was the better option for getting rid of the noise and preserving the round sharp features present in the original model.

ACKNOWLEDGMENTS

The authors thank Guojian Shan for providing the Kronecker product module. We also thank Bob Clapp for important discussions.

REFERENCES


Short Note

AMO regularization: Effective approximate inverses for amplitude preservation

Robert G. Clapp

INTRODUCTION

Amplitude preservation in imaging is becoming increasingly important. The irregularity of seismic data, particularly 3-D data, in both the model domain (in terms of subsurface position and reflection angle) and the data domain (in terms of midpoint, offset, and time) can have deleterious effects on amplitude behavior. There have been several general approaches to correct for this irregularity. The imaging problem is fundamentally an inverse problem, relating some model \( m \) to some data \( d \) through a linear operator \( L \), which in this case is the adjoint of the migration operator. Ronen and Liner (2000); Duquet and Marfurt (1999); Prucha et al. (2000) cast the problem as such and then try to solve it with an iterative solver. These approaches have shown promise but are in many cases prohibitively expensive.

The problem is further complicated in that many migration algorithms assume the data is lying on regular mesh (downward continuation and finite difference schemes for example). Biondi and Vlad (2001) dealt with the problem of mapping the irregular data to a regular mesh for downward continuation migration. They set up an inverse problem relating the irregular input data to a regular model space. They regularized the problem by enforcing consistency between the various \((\text{time},\text{cmp}_x,\text{cmp}_y)\) cubes. The consistency took two forms. In the first a simple difference between two adjacent inline offset cubes was minimized. In the second the difference was taken after transforming the cubes to the same offset through Azimuth Moveout (AMO) (Biondi et al., 1998). For efficiency the model was preconditioned with the inverse of the regularization operator (Fomel et al., 1997). Instead of solving the least squares inverse problem, the Hessian is approximated by a diagonal operator computed from a reference model (Claerbout and Nichols, 1994; Rickett, 2001; Clapp, 2003).

In this paper I examine and extend the work in Biondi and Vlad (2001). I show that approximating the inverse matrix with a simple diagonal operator is not sufficient. The resulting regular dataset has artificial amplitude anomalies. I replace the simple derivative operators with a filter that smooths along not only offset\(_x\), but also offset\(_y\). I conclude by discussing

\[1\text{email: bob@sep.stanford.edu}\]
how the problem can be effectively parallelized and the computational and storage challenges of various estimation schemes.

**REVIEW**

Most downward continuation methods require that the data lie on a regular mesh. To map the irregular recorded seismic data onto the regular mesh is a far from trivial exercise. A common approach in industry is to think of the problems in the same way we approach Kirchhoff migration, namely to loop over data space and spread into our regular model space. The spreading operation is governed by something like AMO (Biondi et al., 1998), which maps data from one offset vector to another. If we think of the AMO operator $T$ as mapping from the regular model space $m$ to the regular data space $d$, our estimation procedure becomes,

$$m = T'd.$$  

This formulation suffers from all of the usual problems associated with applying an adjoint operation. We are spraying into a regular mesh, but the coverage is not regular. Areas with higher concentration of data traces will tend to map to artificially higher amplitudes in the model space. We can do some division by hit count to help minimize this effect but will still see some artifacts that come from approximating the inverse with an adjoint.

We can think of turning (1) into an inversion problem but, in addition to the high cost associated with the AMO operation $m$ we face the same stability issues that setting up the migration problem as an inverse problem encounters. The null space of the imaging operator tends to put high frequency noise in the model space when cast as inverse problem.

Fomel (2001) suggested thinking of the problem more as a missing data problem. We can write the missing data problem in terms of the fitting goals

$$d \approx Lm$$

$$0 \approx \epsilon Am,$$  

where $L$ is a simple interpolation operator (nearest-neighbor, linear, etc) and the real work is done by the regularization operator $A$ which describes the relationship between the irregular data and the regular sampled model.

We can speed up the convergence of (2) by preconditioning the model with $m = A^{-1}p = Bp$. Our new fitting goals become,

$$d \approx LBp$$

$$0 \approx \epsilon p.$$  

Biondi and Vlad (2001) suggested following the approach of Claerbout and Nichols (1994) and Rickett (2001). Instead of solving the inverse problem, they suggest filtering the adjoint
solution with a diagonal operator. We obtain our filtering operator by first noting the least squares inverse of the interpolation problem,

$$m = B(B'L'LB + \epsilon^2)^{-1}B'L'd.$$  \hspace{1cm} (4)

We can think of equation (4) as filtering the adjoint solution with the matrix $W^{-1}$ where,

$$W = B'L'LB + \epsilon^2.$$  \hspace{1cm} (5)

The weighting matrix is $(np \times np)$ where $np$ is the size of our preconditioned model space. This matrix will be generally diagonally dominant. We can think of estimating a diagonal filtering operator $W_{\text{diag}}$ by using a reference model (in preconditioned model space) $p$ and applying

$$W_{\text{diag}} = \frac{\text{diag}[(B'L'LB + \epsilon^2)p_{\text{ref}}]}{\text{diag}(p_{\text{ref}})}.$$  \hspace{1cm} (6)

We can then get an estimate of our model through

$$m = BW^{-1}B'L'd.$$  \hspace{1cm} (7)

Regularization

In the formulation above our model quality is now greatly determined by our choice of regularization operator. Biondi and Vlad (2001) implemented two different approaches. The first was simply applying a derivative filter, $1 - \rho$ along the offset axis. The $\rho$ controls the length of the smoother. For symmetry we can cascade a left derivative $D_l$ followed by a right derivative $D_r$ for the combined regularization operator $A = D_rD_l$.

Biondi and Vlad’s (2001) second choice is more interesting. Vlad and Biondi (2001) describe a very fast implementation of AMO (based on the DMO formulation in the logstretch domain of Zhou et al. (1996)) on a regularly sampled mesh. They suggested instead of minimizing the difference between two offset cubes, to minimize the difference between the two cubes continued to the same offset through AMO. If we now imagine the filter operating on $(t,cmp_x,cmp_y)$ cubes, the right derivative operation becomes $I - \rho T_{hi+1i}$ where $T_{hi+1i}$ is the AMO transformation of the $(t,cmp_x,cmp_y)$ at offset $i + 1$ to offset $i$.

COMPLICATIONS

The procedure described in the last section has several problems when applied to a large 3-D dataset. In this section I cover some of these; issues, solutions, and compromises that need to be made for a practical implementation.

Approximate inverses

The approximation of equation (7) at first gives a visually appealing result. But if we take a closer look, we see that it isn’t as close to the true inverse as we might hope. To see how, let’s
look at a simpler problem. The circles in Figure 1 show a series of irregular data points subsampled from the solid line curve. The dotted lines show the solution when applying equation (7). The dashed line represents the solution using fitting goals (3). Note how the approximate solution approach has the correct low frequency shape but varies significantly from the full inverse solution. With better choices of $p$, or additional diagonals (Guitton, 2003), a better solution might be possible, but the potential is limited. The full matrix can not be adequately described by the limited description we are allowing.

![Figure 1](image)

Figure 1: The solid line represents the input signal. The circles represent our data points. The dashed line shows the solution with fitting goals (3). The dotted lines show the solution using the approximate method in equation (7).

We can see the same effect in our regularization problem. Figure 2 shows the regularization result of a small portion of a 3-D land dataset. Figure 3 shows the result of doing five steps of conjugate gradient solving the fitting goals in (3). Note how we have a smoother, more believable, amplitude behavior as a function of midpoint. Unfortunately, the effect of the approximate solution translates directly to an effect on the amplitudes in our migration. Figure 4 is the result of migrating a single line from the 3-D dataset regularized by the method described in Biondi and Vlad (2001). Note the stripes of high and low amplitude indicated by ‘A’ and ‘B’. Wavefront healing helps minimize the effect at depth but there is still a noticeable effect on the amplitude.

Solving the full inverse introduces its own problems. First, we have now significantly increased the cost. The approximate solution (7) required calling both $L$ and $B$ three times. Even using a minimal number of iteration (3-5) increases the cost by a factor of two or three. In addition, we have significantly increased our disk space requirement. If we set up the inverse problem shown in fitting goals (3), we now must store six copies of our model space. We are quickly approaching the point of impracticality even for a small dataset.
Figure 2: A portion of the regularized data, \((t, cmp_x, cmp_y)\) cube, estimated using the approximate solution in equation (7). Note the dimming and brightening due to the irregular sampling of the input data.
Figure 3: A portion of the regularized data, \((t, cmp_x, cmp_y)\) cube, estimated by five conjugate gradient iterations using (3). Note how the improved amplitude continuity.
Figure 4: Migration of one line of dataset regularized with the approximation solution (7). Note the difference in amplitude at ‘A’ and ‘B’ caused by the approximate solution.

Dimensionality

A 3-D reflection dataset resides in a five dimensional space. Typically we describe this space in terms of \((t, s, g)\) or \((t, h, cmp)\) where \(t\) is time and \(s, g, cmp, h\) are the source, receiver, midpoint, and offset vectors. In Biondi and Vlad (2001) the dimensionality of the dataset was decreased by one by describing offset by a scalar rather than a vector. This is far from an ideal solution, especially in the case of a land dataset and/or data over complicated geology. In both cases the earth being sampled at different azimuths can vary significantly. By stacking we are making an implicit assumption that there isn’t any variation (or at least significant variation) due to azimuth. This can affect both our amplitudes and our ability to accurately estimate the model velocity (Clapp and Biondi, 1995).

If we use a five dimensional model space we must modify our estimation procedure. The solution is to perform individual estimations at different \(h_y\) or azimuths. Solving independent problems is not generally a workable solution. First we aren’t imposing any smoothness over \(h_y\) or azimuth, something that we know should physically exist. Our estimation procedure is likely to produce an answer far from smooth over the added axes. Figure 5 and Figure 6 show fold maps for a portion of a 3-D land and marine datasets. The left panel shows the \(cmp_x, cmp_y, h_x\) cube, the latter the \(cmp_x, cmp_y, h_y\) cube at the same midpoint location. Note how in both cases the coverage varies significantly as a function of \(h_y\). Take the marine case for
Figure 5: The fold from a portion of a 3-D land dataset. The left panel is a subset at a constant $h_y$. Three panels from the subset are shown. The right panel is a subset at a constant $h_x$. Both panels show the same $cmp_x, cmp_y$ location.  

example. Standard acquisitions techniques would lead to $cmp_y$ (cross-line direction) locations to be banded along different $h_y$ locations (caused by the multiple towed cables) and few large $h_x$ at small $h_y$ (due to cable feathering).

Data size

The estimation problem we have set up requires a model space larger than we will typically use in migration. As mentioned above, traditional implementation of AMO (Biondi et al., 1998) works like Kirchhoff migration. We define our model space (as sparse or as dense as we wish) and there sum in nearby traces with appropriate weights. The AMO procedure can be used as a fairly intelligent partial stack. By implementing the AMO as a regularization operator we are asking $L$ to map the trace from the irregular data space to the the regular space that our model exists on. If we have too coarse of a sampling in our model space we end up mapping numerous data points to each model point. If we think about data’s behavior as a function of offset (fairly variable even after NMO) the danger of making too large of bins becomes apparent.

The problem is that our full model space is enormous. A small to mid-size dataset might
Figure 6: The fold from a portion of a 3-D marine dataset. The left panel is a subset at a constant $h_y$. Three panels from the subset are shown. The right panel is a subset at a constant $h_x$. Both panels show the same $cmp_x, cmp_y$ locations. [bob1-fold.elf] [ER,M]

have 1500 time samples, 1000 $cmp_x$, 1000 $cmp_y$, 128$h_x$, and require 20 $h_y$. That amounts 15 TBs, exceeding the entire storage capacity of SEP. Even a small portion of the dataset (500 $cmp_x$, 200 $cmp_y$) will still consume 1.5 TB.

**Parallelization and Regularization**

The standard Beowulf cluster, and SEP’s cluster, consists of many single or dual processor nodes with communication between nodes having fairly large latency. As a result we want limited communication and as coarse a grain scheme as possible. The AMO operation (Vlad and Biondi, 2001) consists of:

$S_{log}$ log stretch of time axis,

$F_t$ FFT of the stretched time axis,

$F_{xy}$ FFT the $cmp_x$ and $cmp_y$ axes,

$C$ complex multiplication
\[\mathbf{F}_{xy}^{-1} = \mathbf{F}_{xy}'\text{ FFT the } cmp_x\text{ and } cmp_y\text{ axes,}\]

\[\mathbf{F}_t^{-1} = \mathbf{F}_t'\text{ FFT of the stretched time axis, and}\]

\[\mathbf{S}_{log}^{-1}\text{ inverse log stretch of time axis.}\]

Our regularization operator is not limited to the AMO operator. Our choice of regularization also affects our ability to parallelize the problem. Ideally we would like to have a regularization operator that assessed continuity over all axes (a non-stationary Prediction Error Filter for example), but that would either eliminate our ability to parallelize the problem or require massive communication between the nodes (to pass the boundary areas over the axes we parallelized over).

If we sacrifice regularizing along the time axis the problem becomes more manageable. We can redefine our data as

\[\mathbf{d}_{\text{new}} = \mathbf{F}_t \mathbf{S}_{log} \mathbf{d},\]  

along the first axis.

We can now split the data along the time axis and regularize along any of the remaining axes. For this paper I chose to only regularize along the offset axes. A 4-D prediction error filter would be preferable, but would require simultaneously infilling and estimating the filter or some ad-hoc scheme that is beyond the scope of this paper.

Regularizing only over offset also allows additional cost savings. We can pull out the \(\mathbf{F}_{xy}\) operator outside our filtering operation. For the cascaded derivative operation used in Biondi and Vlad (2001) we save the cost of six \(\mathbf{F}_{xy}\) per iteration step (approximately 67% reduction in cost).

Our expanded model space (\(h_y\) axis) requires a new regularization scheme. Two obvious choices come to mind. The first is to cascade derivative regularization along the \(h_y\) axis. Our new regularization operator becomes

\[\mathbf{A} = \mathbf{F}_{xy} \mathbf{D}_{hy,l} \mathbf{D}_{hx,r},\]  

where \(\mathbf{D}_{ha,b}\) is taking the derivative (after transforming the \((t, cmp_x, cmp_y)\) cube) in the \(b\) direction along the \(a\) axis. The other approach is to use some arbitrary filter, such as a factored Laplacian (Claerbout, 1999). I tested both methodologies. The first approach does not have a completely symmetric impulse response (Fomel, 2001), but proved to converge in fewer iterations.

On even a small problem the current formulation is still problematic on SEP’s current architecture. Having to store six copies of the model exceed our node’s disk capacity even after splitting the data along the time axes. The final simplification is instead of solving a single global inversion problem to solve for each frequency independently. This final simplification makes the problem manageable, but at a price. We are doing a low number of conjugate gradient iterations, therefore our solution step size (and direction after the first iteration) is going to be different for the global and the individual local problems.
CONCLUSIONS

Data regularization is an important problem when using migration methods that rely on the data being on a regular mesh. Traditional methods that apply the adjoint of a continuation operation such as AMO can lead to poor amplitude information in the regularized (and later migrated) cube. By setting up the regularization problem as inverse problem the amplitudes in the regular model space are significantly improved. The problem can be made computationally acceptable by intelligent parallelization and regularization choices.

REFERENCES


Short Note

More fitting equations for PEF estimation on sparse data

William Curry

INTRODUCTION

Prediction-error filters (PEFs) can be used to interpolate data (Spitz, 1991; Claerbout, 1999; Crawley, 2000). In order to generate a PEF, a least-squares problem is solved where regularly-sampled data is convolved with an unknown PEF. When estimating a PEF with missing data, the equations that contain missing data can be eliminated from the inversion. When dealing with sparse data, however, all equations may contain missing data, so a PEF cannot be estimated.

This problem has been addressed by regridding the data to multiple different scales, and then introducing those coarser copies of the data to the problem to gain more fitting equations (Curry and Brown, 2001; Curry, 2002). When creating coarser copies of the data, there is freedom in choosing the method used to regrid the data, as well as the choice of the coarser grid itself.

Existing methods use multiple different grid sizes in order to extract more information from the sparse data. More fitting equations can be generated by not only varying the size of the grid, but also by varying the positioning of that grid, i.e., the location of the origin of the grid. As the grid becomes coarser, the number of possible grid positions increases, and as the dimensionality of the problem increases, that number increases further.

By varying the grid location, a PEF can be more accurately determined on sparse data. 2D and 3D examples are shown for non-stationary PEF estimation, with a noticeable improvement in the interpolated result.

BACKGROUND

A PEF can be estimated by solving the minimization problem where known data ($d$) is convolved ($D$) with an unknown PEF ($f$), so that

$$W(Df + d) \approx 0,$$

(1)

\[email: bill@sep.stanford.edu]
where $W$ is a weight for missing data and $K$ constrains the first PEF coefficient to be 1.

When all of the equations contain missing data, $W$ is 0 everywhere, and the problem cannot be solved. In this case, rescaled copies of the data can be substituted for the original data in equation (1), resulting in

$$W \left( \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix} Kf + \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \right) \approx 0. \quad (2)$$

In this case, $d_i$ represents the various different rescaled copies of the data, $D_i$ is convolution with that rescaled data, and $W$ is now a weight for all scales of data. The data is rescaled by taking the original finely gridded data, and transforming it to a series of points located at the center of cells with data. Adjoint linear interpolation is then performed to move the data points onto the new grid.

For the case of a non-stationary PEF, the equations remain largely the same, except that the PEF varies with position. When convolving different sizes of data with a non-stationary PEF, the PEF must be sub-sampled so that the spatial dimensions of the non-stationary PEF and the data match. This is accomplished by the introduction of a sub-sampling operator $P_i$, so

$$W \left( \begin{bmatrix} D_0 \\ D_1 \\ D_2 \\ \vdots \\ D_n \end{bmatrix} P_i I + \begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \right) \approx 0. \quad (3)$$

In addition, since the model space has now increased substantially, a second fitting goal must be added,

$$\epsilon Af \approx 0, \quad (4)$$

that ensures that the PEF will vary smoothly over space. In equation (4), $A$ is a regularization operator (in this paper, a spatial Laplacian), and $\epsilon$ is a scale factor.

**GRID SHIFTING**

Instead of simply taking multiple scales of data ($d_i$), multiple different grid locations can be used for each scale, as there is freedom in how they are chosen. For the case of a grid twice as coarse as the original fine grid, there would be 4 possible grid locations for a 2D case, shown in Figure 1. This number is inversely proportional to the size of the bins, and increases exponentially with dimension. However, in the case of irregular traces, there is no point in shifting the grid along the time axis, as the number of equations would be unaffected. In the case of a non-stationary PEF, extra bookkeeping is required, as the different grid locations will correspond to different regions of the non-stationary PEF. This does make using these extra versions of the data more expensive, as the weight $W$ and sub-sampler $P$ are not the same at one scale and need to be recomputed.
Figure 1: Shifting grids, clockwise from lower left: original fine-scale data; data points extracted from that fine grid; and four different grid orientations for a single coarser scale. Dots signify data, colored bins contain data, and white bins are empty.

EXAMPLES

The CASC93 dataset (Nabelek and Zandt, 1993) is a passive seismic experiment where arrivals from distant earthquakes were recorded by an irregular 2D array of seismometers in central Oregon. As we can see in Figure 2, the interpolated result with the PEF estimated from the shifted data looks better than the original multiscale result. In this case 7 different scales were used for (b), while for (c) 10 additional versions of the data were used. The difference is most pronounced in the area with coarser sampling, since the different grid positions would return more unique information than in the densely-sampled area.

Next, a 3D synthetic case is examined. The “quarter-dome” model (Claerbout, 1993) has been sub-sampled randomly so that only 10 percent of traces remain. It was then interpolated with both a standard multiscale PEF as well as a multiscale PEF estimated with different grid positions. The results are shown in Figure 3. The number of versions of the data varies even more in this case, where 5 different scales were used in the multiscale estimation, while 39 additional versions of the data were used when grid position was varied. The results are improved, although nearly 8 times as many copies of data were used. However most of those versions of the data were at the coarser scales.
Figure 2: CASC93 data set. (a) original data; Interpolated with multiscale PEF with (b) no shifting and (c) with shifting; (d) difference between (b) and (c).
Figure 3: The quarter-dome model: (a) original data; (b) interpolated with multiscale PEF; (c) interpolated with multiscale PEF with grid position; (d) difference between (b) and (c).
CONCLUSIONS

By altering the position of a grid as well as the coarseness of a grid, additional information can be introduced to the problem of PEF estimation. This information can be used for both stationary and non-stationary PEFs in any number of dimensions. As the dimensionality of a problem increases, the number of possible grid positions increases exponentially, whereas the number of scales of data is unaffected by the dimensionality of the problem. The benefits of this method increase with dimension, as shown in the differences between the 2D and 3D examples. The method also shows greater effectiveness around larger gaps in the data. However, the cost of regridding the data many times as well as the associated bookkeeping is problematic.

This method can be used to glean more information from the data using existing techniques. The question of how to best regrid the data remains, as does the question of whether it is best to manipulate the filter or the data when attempting to estimate a PEF on sparse data. Overall, this method provides one more means of manipulating the data to better constrain the PEF, and shows improvement over existing methods.

REFERENCES


Curry, W., 2002, Non-stationary, multi-scale prediction-error filters and irregularly sampled data: SEP–111, 327–337.


INTRODUCTION

The geometry of reflection seismic experiments rarely conforms to an idealized, or nominal geometry. This is especially true in regions with unavoidable surface obstructions like rivers, towns, or existing offshore oil platforms. Such deviations from nominal geometry render many computer seismic processing applications ineffective. It is therefore of considerable importance to 1) accurately map irregularly-sampled seismic data onto a regular grid with sufficient spatial resolution, and 2) to interpolate any missing trace locations with reasonable values.

Any interpolation method fills missing traces using a prior estimate of the missing data’s spatial correlation; obtaining this correlation is often the main challenge. When the data geometry is regular, many existing methods can resample the data onto a finer grid. Sinc interpolation (e.g.: (Bracewell, 1986)) is optimal for the regular resampling of band-limited data. Crawley (2000) estimates a nonstationary prediction-error filter (PEF) on regularly sampled, spatially-aliased seismic data and uses inverse interpolation (Claerbout, 1999) to fill missing traces on a finer grid. Fomel (2001) solves the same problem, but substitutes “plane-wave-destructor” filters, derived from estimated reflector dip, for a PEF.

Unfortunately, irregular data geometry renders most conventional methods, including those mentioned previously, inapplicable. To reliably estimate autoregressive filters, like the PEF, all points in the filter stencil must, but generally do not, fall on known data locations. A multi-scale autoregression technique (Curry and Brown, 2001; Curry, 2002) has yielded some success by estimating the PEF simultaneously from data subsampled to a series of different resolutions. Biondi and Vlad (2001) uses azimuth moveout to transform known data to arbitrary azimuth/offset bins to constrain missing data. Other techniques (e.g: Liu and Sacchi (2001) and Zwartjes and Hindriks (2001)) solve an inverse interpolation problem in the Discrete Fourier Transform domain by regularizing the unknown coefficients. Nonetheless, an industry-standard technique does not yet exist.

In this short note, I present a different way to obtain the correlation between irregularly-sampled traces in three dimensions. Given a pair of traces, as shown in Figure 1, reflector dip can be measured along an arbitrary azimuth, by (for instance) Claerbout’s “puck” method

---

1email: morgan@sep.stanford.edu
(Claerbout, 1992). If the dip between a master trace and two or more neighbors is measured along two distinct azimuths, we can solve a simple least-squares problem for the dip in the \(x\)- and \(y\)-directions at the master trace’s location. Both geology and survey geometry often dictate that the reflector dip should vary smoothly in space. Therefore it is both natural and intuitive to extend the estimated reflector dip to missing data locations.

I test my dip estimation scheme on a sparsely and irregularly sampled 3-D synthetic model. I use the estimated dip to compute “steering filters” (Clapp et al., 1997) which regularize an inverse interpolation problem. This choice of regularization leads to a far better result than the spatial gradient operator, which corresponds to the assumption of zero dip between traces.

\[
t_2 - t_1 = p_x(x_2 - x_1) + p_y(y_2 - y_1)
\]

However, in actuality we do not know \(p_x\) and \(p_y\), but using Claerbout’s puck method, we can measure \(t_2 - t_1\). Implemented on a computer, the puck method computes the time shift (in

**IRREGULAR-GEOMETRY DIP ESTIMATION METHODOLOGY**

Figure 1 illustrates a simplified conceptual model of seismic data. Given two traces, we assume that a seismic event passes through each trace location at times \(t_1\) and \(t_2\), and that the event takes the form of a plane in the neighborhood of the two traces.

Imagine that we wanted to measure \(t_2 - t_1\), given the local dip of the plane, \(p_x\) and \(p_y\). The total time shift is simply the sum of the time shifts along the \(x\) and \(y\) planes, going first from \(x_1\) to \(x_2\) and then from \(y_1\) to \(y_2\). We can write an equation directly:

\[
t_2 - t_1 = p_x(x_2 - x_1) + p_y(y_2 - y_1)
\]
pixels) between two traces that optimally aligns a seismic event on the two traces as a function of time. In other words, the method measures

$$p_{21} = \frac{t_2 - t_1}{\Delta t},$$

(2)

where $\Delta t$ is the time sampling of the traces. We can now rewrite equation accordingly:

$$p_{21} = p_x(x_2 - x_1) + p_y(y_2 - y_1)$$

(3)

Equation (3) describes the linear relationship between the computer dip measured between traces 2 and 1, $p_{21}$, and the local 3-D reflector dip. We require two equations to obtain a unique estimate of the parameters, but the noise and incoherency inherent to real data make it desirable to use more than two traces and exploit the statistical “smoothing” of least-squares estimation.

If trace 1 is the “master” trace and traces 2 through $n$ its neighbors, let us define the forward modeling operator $A$

$$A = \begin{bmatrix}
\frac{x_2-x_1}{\Delta t} & \frac{y_2-y_1}{\Delta t} \\
\frac{x_3-x_1}{\Delta t} & \frac{y_3-y_1}{\Delta t} \\
\vdots & \vdots \\
\frac{x_n-x_1}{\Delta t} & \frac{y_n-y_1}{\Delta t}
\end{bmatrix}$$

(4)

and the data vector $d$

$$d = \begin{bmatrix}
p_{21} \\
p_{31} \\
\vdots \\
p_{n1}
\end{bmatrix}.$$  

(5)

The estimated $p_x$ and $p_y$ at trace location 1 are then the solution to the normal equations:

$$\begin{bmatrix} p_x \\ p_y \end{bmatrix} = (A^T A)^{-1} A^T d$$

(6)

Inversion of the $2 \times 2$ matrix $A^T A$ is trivial. The result of the process is a pair of dip measurements ($p_x$ and $p_y$) at each trace location. These dip measurements must be interpolated to fill the entire model grid. In this paper, I use an expanding-window smoothing algorithm to accomplish the task. If reflectors are continuous, their dips should be somewhat smooth in space, so to some extent, spatial smoothing is justified.

**REVIEW OF INVERSE INTERPOLATION**

In least-squares fitting goals, the regularized inverse interpolation problem can be stated as follows:

$$Lm \approx d$$

(7)

$$\epsilon_x A_x m \approx 0$$

(8)

$$\epsilon_y A_y m \approx 0$$

(9)
Operator \( \mathbf{L} \) maps traces in a gridded model \( \mathbf{m} \) to the earth’s continuous surface. Operators \( \mathbf{A}_x \) and \( \mathbf{A}_y \) are “steering filters” (Clapp et al., 1997) in the \( x \) and \( y \) direction, respectively. The steering filters are initialized with a space-variable dip function and decorrelate events which have that dip, and tend to steer the estimated model along the dip direction. In this fashion we impose our prior model covariance estimate on the missing traces. Scalars \( \epsilon_x \) and \( \epsilon_y \) balance the two model residuals [equations (8) and (9)] with the data residual [equation (7)].

**TESTS**

I test my irregular-geometry dip estimation method on a decimated subset of the “quarter dome” model (Claerbout, 1999), shown in Figure 2. Only 100 out of 1024 input traces are assumed known, for a decimation rate of over 90 percent. The dip estimated by my method will then be used to interpolate the missing trace locations by solving the system (7)-(9).

Figure 2: Left: 32x32-trace subset of the quarter dome model. The subset was selected to be less sensitive to spatial aliasing than the steeper-dipping portions of the model. Right: Missing data test, with 100 known traces and 924 missing traces.

panels of Figure 3 show the quarter dome model’s “known” dip, which was computed using a variant of Fomel’s (2002) dip estimation method. The right panels show the dip estimated from the irregularly-sampled traces shown in Figure 2, using the method described herein. To smooth the dip estimates between master trace locations, I use an expanding-window smoothing program.

While the dip estimates are decidedly imperfect, they nonetheless do contain the general trends seen in the known dip fields. Particularly note the unconformity deep in the section. We see that in the more steeply-dipping parts of the section, my method tends to underestimate the reflector dip. Since the decimation is severe, spatial aliasing may arise, even if the dips are not severe, because my method measures dip directly between two arbitrary traces. Claerbout’s
puck method is known to be sensitive to spatial aliasing. Fomel (2002) discusses ways to de-sensitize dip estimation with respect to spatial aliasing. Figure 4 compares the result of using: a) the known dip, b) the dip estimated from the irregular data, and c) zero dip in solving equations (7)-(9) for an infilled model. We see that in spite of the imperfections of the dip estimated from the irregular data, that it definitely leads to a better inverse interpolation result than the zero dip result, which just smooths laterally. All results were computed using 20 iterations of a linear conjugate gradient solver.

**CONCLUSIONS**

In this short note I presented a methodology to estimate 3-D reflector dip from seismic traces with arbitrary geometry. While the resultant dip estimates are imperfect, I show that they nonetheless represent a significant improvement over a zero dip assumption in the desired application: inverse interpolation.

I foresee this method as having some value as a first-order starting guess for nonlinear implementations of Claerbout’s “two-stage” inverse interpolation method (Claerbout, 1999), as tested by Curry and Brown (2001). Significant improvements are likely possible, particularly in the method for estimating the initial reflector dip between two arbitrary traces.
Figure 4: Left to right: Known result; Decimated data; Data filled with known-dip steering filters; Data filled with estimated-dip steering filters; Data filled with spatial gradient only (zero dip).

ACKNOWLEDGMENT

Discussions with Bill Curry initially motivated me to take on this project.

REFERENCES


Curry, W., 2002, Non-stationary, multi-scale prediction-error filters and irregularly sampled data: SEP–111, 327–337.


ABSTRACT

Regularly-sampled data is normally required to estimate a prediction-error filter (PEF). I show how a small PEF can be estimated when the data is irregularly sampled in one dimension. I then estimate a PEF on an irregularly-sampled 2D examples, and discuss extension to 3D.

INTRODUCTION

Data interpolation has been performed with the use of multidimensional prediction-error filters (PEFs) (Spitz, 1991; Claerbout, 1999). However, methods to generate a prediction-error filter require regularly-sampled data. When the data is not regularly sampled, the data is first transformed to a regular grid (or multiple grids), and then the PEF is estimated on those copies of the data.

In the case of interlaced data (where traces are sampled at regular intervals) the PEF is stretched (Claerbout, 1999; Crawley, 2000) such that the coefficients fall on the interlaced traces during convolution. This method can be used to estimate a PEF only in those circumstances, and relies on the assumption of scale invariance, where a stretched filter and an unstretched filter behave in a similar fashion.

In the case of a line of irregularly-sampled traces, a small PEF can be determined by dynamically stretching the filter so that it fits each trace pair. This method does not require the distance between traces to be cleanly divisible by the shortest distance, nor does it even require the data to be gridded in all dimensions. It also does away with many of the parameter choices needed by other PEF estimation methods for irregular data. The method is first tested on a simple plane-wave model, with promising results.

BACKGROUND

A prediction-error filter is estimated by solving a minimization problem where known data is convolved (Y) with an unknown filter (a), so

$$0 \approx r = Ya, \quad (1)$$

email: bill@sep.stanford.edu
where the first coefficient of \( a \) is constrained to be unity. This can be written as

\[
0 \approx r = YK a + y, \tag{2}
\]

where \( K \) is a mask for the first PEF coefficient, and \( y \) is simply a copy of the data. Writing out the matrices for a PEF with 3 coefficients and for 7 data samples looks like

\[
0 \approx r = \begin{bmatrix}
y_2 & y_1 & y_0 \\
y_3 & y_2 & y_1 \\
y_4 & y_3 & y_2 \\
y_5 & y_4 & y_3 \\
y_6 & y_5 & y_4
\end{bmatrix}
\begin{bmatrix}
0 & \cdot & \cdot \\
\cdot & 1 & \cdot \\
\cdot & \cdot & 1
\end{bmatrix}
\begin{bmatrix}
1 \\
a_1 \\
a_2
\end{bmatrix} + \begin{bmatrix}
y_2 \\
y_3 \\
y_4 \\
y_5 \\
y_6
\end{bmatrix}. \tag{3}
\]

The previous equations have all been for a 1D case. By use of the helical coordinate (Claerbout, 1998), these equations can easily be extended to higher dimensions.

In the case of missing data, a diagonal weight (\( W \)) can be introduced that is 0 when a missing data point is in the equation, and 1 where all data are present. This weight can also be used to eliminate edge effects caused by helical convolution.

When data are interlaced, a PEF can be estimated by spacing filter coefficients during convolution, so that they fall on known data. An example of this filter spacing is shown in Figure 2. The problem with this method is that the data must be regularly sampled in all dimensions.

Figure 1: Examples of PEFs on interlaced data. White bins are empty, gray have data. Left: A PEF cannot be estimated due to too much missing data. Right: The spaced PEF can be estimated on interlaced data.

---

**IRREGULAR TRACES**

Seismic data are recorded as a series of traces that are regularly-sampled in time, but are often spatially distributed in an irregular manner. Since the estimation of a PEF requires a regularly-sampled input, the data are placed on a regular grid. If this grid is too fine to allow a PEF to be estimated, a multi-scale method can be used to introduce more fitting equations (Curry and Brown, 2001; Curry, 2002).

Instead of relying on methods to transform the data onto a regular grid, we can examine if it is possible to manipulate the filter so that it can fit the irregular data. In the case of interlaced data, the filter can be spaced to fit the data. If the data are from a 2D grouping of irregular
Irregular PEF estimation

traces, we can take a small 2-column PEF and dynamically stretch it so that as the filter is convolved with the data, the PEF will always fall on known data.

When looking at 2-column PEF estimation, we can take equation (2) and include a lag matrix \( S \), which matches the appropriate data point to filter coefficient. For a simple \( 2 \times 5 \) PEF, \( S \) will look like:

\[
S_{\text{orig}} = \begin{bmatrix}
1 & . & . & . & . \\
. & 1 & . & . & . \\
. & . & 1 & . & . \\
. & . & . & 1 & . \\
. & . & . & . & 1 \\
\end{bmatrix}
\]

The gap in the diagonal matrix is the length of the \( n_1 \) axis (time) minus the filter length in the \( n_1 \) dimension. When stretching the data for interlaced traces, the matrix changes, so that

\[
S_{\text{inter}} = \begin{bmatrix}
1 & . & . & . & . \\
. & . & . & . & . \\
. & 1 & . & . & . \\
. & . & 1 & . & . \\
. & . & . & 1 & . \\
. & . & . & . & 1 \\
\end{bmatrix}
\]

In this case, every second data point is selected by this matrix, and the gap now corresponds to 2 times the length of the \( n_1 \) axis, since a row is skipped. Now for a dynamically stretched PEF, the matrix will look somewhat different, as the PEF is not stretched by an integer amount, so if a filter coefficient falls between two data points (as shown in Figure 2) a linear combination of the two data points will be multiplied with the PEF coefficient. For a single trace pair, \( S \)
will now be

$$S_{\text{irreg}} = \begin{bmatrix}
1 & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0.5 & \ldots & \ldots & \ldots & \ldots & \ldots \\
0.5 & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
\ldots & \ldots & \ldots & \ldots & 1 & \ldots \\
\ldots & \ldots & \ldots & \ldots & 0.5 & \ldots \\
\ldots & \ldots & \ldots & \ldots & 0.5 & \ldots \\
\ldots & \ldots & \ldots & \ldots & 1 & \ldots \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots
\end{bmatrix}$$

if the traces were considered to be 1.5 bins apart. For each trace pair, a separate set of equations where $Y_i$ corresponds to convolution with the trace pair, $S_i$ is a filter lag matrix that is a function of the distance between the two traces, $y_i$ is a copy of the second trace of the pair, and $K$ and $f$ remain unchanged. For a three-trace set, the fitting goal would be

$$0 \approx r = \begin{bmatrix}
Y_1 & \ldots & \ldots \\
\ldots & Y_2 & \ldots \\
\ldots & \ldots & Y_3 \\
\end{bmatrix} \begin{bmatrix}
S_1 & \ldots & \ldots \\
\ldots & S_2 & \ldots \\
\ldots & \ldots & S_3 \\
\end{bmatrix} \begin{bmatrix}
K \\
K \\
K \\
\end{bmatrix} f + \begin{bmatrix}
y_1 \\
y_2 \\
y_3 \\
\end{bmatrix}. \quad (7)$$

In methods where the data are transformed to a regular grid (Curry and Brown, 2001), many parameters (such as scales) need to be chosen. With this method of manipulating the filter to
fit the data, the only parameter that needs to be set is the aspect ratio of the PEF, meaning what distance would correspond to a single bin in time.

There are several assumptions being made when estimating a PEF in this manner, most of which are shared with Crawley’s (1998) approach. First of all, this method assumes that the filter is scale-invariant, meaning that the stretched filter would behave the same as if it were not stretched, which is valid when dealing with plane waves (Claerbout, 1999). Next, there is an assumption of stationarity. The filter can be made to be non-stationary, but the stretching of the filter assumes that the area that the stretched filter covers is stationary. Finally, since using a linear combination of two data points as a proxy for a data point located between those two is tantamount to linear interpolation, there is an assumption that the data are not aliased in the well-sampled (time) axis.

There are also several limitations to this approach. First of all, in order to dynamically stretch a PEF to fit data along one dimension, all other dimensions must be regularly-sampled and contain no aliasing. This limits the estimation to 2D with seismic data, since the only dimension with the required regularity is the time axis. This problem can be addressed by simultaneously estimating numerous 2D PEFs for other dimensions. Another issue that arises is multiple dips. A non-stationary filter can deal with different dips in different areas, but is not effective if those dips are co-located. A pair of PEFs could be simultaneously estimated, which might work in estimating both dips.

TEST CASE

Figure 3 shows a fully-sampled plane wave previously used as a test case for PEF estimation (Brown et al., 2000). Nine traces (approximately 15 percent of traces) were removed from this model and were used to estimate a PEF. The only inputs to the algorithm are the nine traces and the distances between the traces. The results are promising, as the dip is easily recovered. The inverse impulse response of the PEF, shown in Figure 3c, matches the dip of the original fully-sampled data. Figure 3d contains a re-sampled version of Figure 3b, which was then interpolated with the PEF estimated on the irregularly-sampled traces. The flat areas in that image are due to the re-sampling and not the PEF.

CONCLUSIONS AND FUTURE WORK

By dynamically stretching filter coefficients, a small PEF can be estimated on irregularly-sampled traces. This method works for a 2D test case with a simple plane-wave model. This method is computationally inexpensive, and only requires a single extra parameter in addition to those required for PEF estimation on a regular grid.

Certain limitations need to be addressed with this method: First, extension to three dimensions. One possible approach to this problem is to use a series of PEFs operating over different axes. Another problem that arises is what to do in the presence of multiple dips. Again, using a series of PEFs might address this issue, as well as the use of non-stationary filters.
Figure 3: Plane wave test case: (a) is the original fully-sampled data; (b) is the sub-sampled version; (c) is the inverse impulse response of the PEF; and (d) is an interpolated copy of coarsened data.
ACKNOWLEDGMENT

I would like to thank Morgan Brown for several useful discussions about this topic.

REFERENCES


Curry, W., 2002, Non-stationary, multi-scale prediction-error filters and irregularly sampled data: SEP–111, 327–337.

ABSTRACT
Image segmentation can be used to track salt boundaries when the salt boundary amplitude is greater than any other local reflections. We apply a modified version of the normalized cut image segmentation method to partition seismic images along salt boundaries. In principle our method should work even when the boundaries are not continuous, and conventional horizon tracking algorithms may fail. Our implementation of this method calculates a weight connecting each pixel in the image to each pixel in a local neighborhood. The magnitude of the weight is inversely proportional to the size of the maximum instantaneous amplitude along the shortest path between the two pixels. This method is demonstrated to be effective on two simple models and 2D seismic section.

INTRODUCTION
Salt boundaries are often the brightest, most prominent reflections in a seismic image. For many processing applications, this boundary needs to be tracked. However, this reflection can be discontinuous, making it difficult for traditional amplitude based auto-pickers to track. Here, we present a tracking method that can handle these issues.

Hale and Emanuel (2003, 2002) applied the normalized cut image segmentation method developed by Shi and Malik (2000) to paint a coherency based reservoir model. Our approach is very similar. 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 inversely proportional to the 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 overview 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 simple models illustrating its efficacy with discontinuous salt boundaries. We also apply this method to a 2D field section where the amplitude is inconsistent and challenging to pick. Methods for dealing with noisy data are also presented.

1email: lomask@sep.stanford.edu, biondo@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, it first creates a weight matrix relating each pixel to every other pixel within a local neighborhood. The strongest weights will be between pixels of similar intensity and close proximity to each other. The method then seeks to partition the image into two groups, $A$ and $B$, by minimizing the normalized cut:

$$N_{\text{cut}} = \frac{\text{cut}}{\text{total}_A} + \frac{\text{cut}}{\text{total}_B}$$  \hspace{1cm} (1)

where $\text{cut}$ is the sum of the weights cut by the partition, $\text{total}_A$ is the total number of weights in Group $A$, and $\text{total}_B$ is the total number of weights in Group $B$. Normalizing the cut by the total number of 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:

$$(D - W)\mathbf{y} = \lambda D\mathbf{y}$$  \hspace{1cm} (2)

created from the weight matrix ($W$) and a diagonal matrix ($D$), with each value on the diagonal being the sum of each column of $W$. The eigenvector ($\mathbf{y}_2$) with the second smallest eigenvalue 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

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. Taking the negative of the maximum amplitude along the shortest path between two nodes as the weight would ensure that the weights connecting pixels on either side of the salt boundary will be low. However, weights on the same side would be alternating from low to high as they go from peak to trough on the seismic data. This could make the grouping more uncertain. To correct this problem, we take the negative of the maximum of the absolute value of the complex trace (instantaneous amplitude) along the shortest path between two nodes.

If two pixels happen to be adjacent to each other, then it doesn’t make sense to take the minimum of the two as the weight. This would make the weight connecting pixels within the boundary itself extremely weak, causing the segmentation algorithm to fail. Our solution is to set the weight equal to unity for adjacent pixels. This puts emphasis on the relative difference between weights calculated from pixels that are not adjacent to one another.

Alternatively, for pixels that are adjacent to one another, we can make the weight a function of similarity in intensity, similar to the standard implementation of the segmentation method.
For instance, the weight connecting two adjacent pixels, $\text{Pixel}_1$ and $\text{Pixel}_2$ can be calculated as:

$$\text{Weight(1,2)} = 1 - \sqrt{(\text{Amplitude(Pixel}_1) - \text{Amplitude(Pixel}_2))^2}. \quad (3)$$

This causes weights parallel to the reflections to be stronger, making it more likely to partition along reflections. Thus far, both of these approaches give similar results.

**TEST CASES**

We applied this segmentation technique to two simple models of salt boundaries and a 2D field seismic section. The first model tests a simple continuous salt boundary. The second model introduces a discontinuity into the salt boundary that would cause conventional auto-trackers to fail. Finally, the 2D seismic section tests this method on real data with a salt boundary that would be difficult for conventional trackers to follow.

**Simple Amplitude Boundary**

![Figure 1](image1.png)

Figure 1: Left is a simple model of a high amplitude salt boundary. Right is the partitioned output of the segmentation method.

On the left in Figure 1 is a simple model of a salt boundary. The vertical stripe in the middle is the high amplitude salt reflection. The results of the segmentation method can be seen on the right. The pixels have been partitioned into two groups, one on either side of the boundary.

On the left in Figure 2 is the same simple model as Figure 1 yet with a discontinuity. Notice that the output result on the right partitions the image across the discontinuity. Although only
a simple test case, this demonstrates that this segmentation method can successfully partition data where the amplitude is discontinuous.

The resulting $y_2$ eigenvector for the discontinuous model is shown on the left in Figure 3. It is this eigenvector that is used to partition the data. The splitting point is at zero. All values greater than zero will be in one group and all values less than zero in the other group. However, a practical measure recommended by Shi and Malik (2000) is to calculate the normalized cut at several splitting points across the eigenvector and take the minimum. On the right in Figure 3 is a contour plot of the $y_2$ eigenvector. Each contour can be thought of as the partition for a different splitting point. Notice that the contours are spread out in the area of the discontinuity. Here the algorithm is unsure of where to track the salt boundary and basically opts for the shortest distance.

**Field Data**

We applied this normalized cut method to the Unocal Gulf of Mexico field data shown in Figure 4. The salt boundary is at approximately 3200 ms. The instantaneous amplitude is shown is Figure 5. Notice that it isn’t too obvious where to pick the salt boundary near Cmp 13500.

The resulting $y_2$ eigenvector is shown in Figure 6 and a contour plot is shown in Figure 7. Notice that the contours are spread in areas where the amplitude of the salt boundary is low and the tracking is therefore less certain.

Lastly, Figure 8 shows the instantaneous amplitude with the partition with the minimum
Figure 3: Left the $y_2$ eigenvector of Figure 2. Right is a contour plot of the $y_2$ eigenvector.

Figure 4: A 2D seismic section from Unocal. The major reflection at 3200 ms is the salt boundary.
Figure 5: This is the instantaneous amplitude of the seismic section in Figure 4. [jesse1-unocal.amp] [ER]

Figure 6: $y_2$ eigenvector calculated from the data in Figure 5. [jesse1-out_y] [ER]
Figure 7: A contour plot of the $y_2$ eigenvector in Figure 6. Notice the areas of uncertainty where the contours are spreading.

normalized cut overlain. It tracks the salt boundary even across the challenging area around Cmp 13500.

CONCLUSIONS AND FUTURE WORK

Our modified segmentation method successfully tracked the salt boundaries in all of our test cases. These test cases presented challenges that include a salt boundary that was not continuous and a field data example that had variable amplitude.

Application of this method to 3D datasets is going to be a computational challenge. The normalized cut method has already been applied to paint 3D reservoir models based on coherency data (Hale and Emanuel, 2003). Unfortunately, 3D seismic cubes can have many more data points. However, for our application of tracking the salt boundary, much of the data can be windowed away in the beginning, hopefully making it possible.

ACKNOWLEDGMENTS

We would like to thank Phil Schultz and Unocal for making available to SEP the field data used for our tests.
Figure 8: The instantaneous amplitude with the minimum normalized cut overlain.

REFERENCES


Automatic discontinuity extraction for 3-D seismic images

Jun Ji

ABSTRACT
A goal of seismic processing is to verify the spatial characteristics of the subsurface. Achieving this goal often requires human analysis to interpret geologically meaningful discontinuities from the seismic image. This interpretation is a challenging task even for an experienced interpreter if the image is 3D. This paper introduces an automatic method to extract seismic event discontinuities from a 3D seismic image. The proposed method consists of three steps. The first step is to evaluate the coherency of seismic events from the seismic image. The second step is to express the regions where discontinuities may exist in a binary image form. The third step is to locate the discontinuity surfaces by thinning the region found in the second step.

INTRODUCTION

In the oil and gas industry, subsurface imaging with the reflection seismic method has been performed for many decades and has been accepted as a successful approach. The procedure of subsurface imaging from raw seismic data consists of many complicated sub-processes and is often organized differently according to its various purposes. In many cases, however, human interpretation of event discontinuity that indicates faults, unconformities, and buried channels is required during processing, or at the end of the procedure. Such interpretations are usually accomplished by an experienced interpreter who has in-depth knowledge of geology. Even for an experienced expert, the interpretation task is challenging if the image is a 3D volume.

In order to help with such tasks, a process called “coherency evaluation” has been developed (Bahorich and Farmer, 1995; Marfurt et al., 1998). The coherency evaluation calculates continuity of seismic events from a 3D seismic image cube and produces a so-called “coherency cube” that shows the distribution of the event continuity in a 3D volume. The output of coherency evaluation processing helps the interpreters to locate geologically meaningful discontinuities more easily. However, interpreting and mapping 3D discontinuity surfaces by hand is still a challenging job because of the difficulty in visualizing a 3D image and the complexity of surfaces in a 3D sense.

As a supplementary tool for mapping discontinuities, I introduce an automatic discontinuity extraction method. The proposed method starts with a coherency evaluation for a 3D seismic image and produces a discontinuity map that locates the event’s discontinuities in the

1email: jun@sep.stanford.edu
form of arbitrarily shaped 3D surfaces. The output of the proposed method could be a good starting point for an interpreter to narrow down the various discontinuities into more meaningful geological features such as faults, unconformities, and buried channels.

In the following section I will shortly review the coherency evaluation method to clarify the meaning of a discontinuity in a seismic image. The subsequent two sections are devoted to explaining the following procedures. In each step, the algorithm is explained and also demonstrated with a real seismic image. The testing image used is from the Boonsville natural gas field located in north-central Texas and was acquired by the DOE and the Gas Research Institute as part of a secondary gas recovery technique development. According to Hardage (Hardage, 1996), the Boonsville seismic image is a full-fold time migrated section with $110m \times 110m$ in bin size and covers $5.5km^2$ area. The data consists of 97 lines along the crossline and each line has 113 traces. Figure 1 shows three plane sections of the image volume which are orthogonal to each other.

![Figure 1: Three selected sections of Boonsville 3D seismic image (time slice at $t = 250ms$, the 37th crossline section, and the 40th inline section.](jun1-seis.png) [NR]

**COHERENCY EVALUATION FOR SEISMIC EVENTS**

Coherency evaluation has been widely used to estimate the continuity of seismic events for a given seismic image. There are several approaches, including the cross-correlation method (Bahorich and Farmer, 1995) and the semblance method (Marfurt et al., 1998). In this paper, I used the semblance method which is an improvement on the the early cross-correlation
method. The following is a short review of the method. First of all, as you can see in Figure 2, I define an elliptic plane that contains J traces around the trace where the coherency is calculated. Then the semblance \( s \) at the center of the ellipse for every dip direction is defined as follows:

\[
s(t, x, y, p, q) = \frac{\sum_{k=-K}^{K} \left( \sum_{j=1}^{J} u(t_j, x, y_j) \right)^2 + \left( \sum_{j=1}^{J} u^H(t_j, x, y_j) \right)^2}{J \sum_{k=-K}^{K} \left( \sum_{j=1}^{J} \left[ u(t_j, x, y_j) \right]^2 + \left[ u^H(t_j, x, y_j) \right]^2 \right)}
\]

(1)

with

\[
t_{j,k} = t + k \Delta t - px_j - qy_j
\]

where \( u \) is the image data, the superscript \( H \) represent the Hilbert transform, \( p \) and \( q \) are the dips of elliptic plane along \( x \) and \( y \) axes direction, respectively. In order to suppress a possible high coherency value around the zero-crossing location, the average semblance is calculated along a time window that ranges from upper \( K \) to lower \( K \). As evident in equation (1), the coherency analysis requires average semblances for various dips (Figure 3). The method of choosing the dips is important not only for the computational cost, but also to obtain an even distribution of dips. I used a “Chinese checker” tessellation (Marfurt et al., 1998) to find a finite number of discrete angle combinations. Then the coherency value at each location was determined by choosing a maximum value among the semblance values for various dips as follows:

\[
coh(t, x, y) = \text{Max}_{p,q} s(t, x, y, p, q).
\]

(2)

This coherency evaluation is performed for the Boonsville image and the result is shown in Figure 4. In Figure 4, the coherency values are shown in grey scale so that the dark represent low coherency and the bright represent high coherency. A comparison of Figure 4 to Figure 1, shows that the coherency cube more clearly reveals the discontinuities in the seismic image. It shows not only the discontinuities that were obvious, but also the ones that are hard to recognize in the seismic image.
Figure 3: Calculation of coherency over an elliptical analysis window with apparent dips (Marfurt et al., 1998).

Figure 4: Selected sections of coherency cube of the Boonsville image.
LOCATING REGIONS FOR POTENTIAL DISCONTINUITY LOCATIONS

The coherency evaluation explained in the previous section will produce a coherency cube for a 3D seismic image. In the coherency cube, the discontinuity information is expressed as a distribution of the event continuity with numerical values ranging from 1 to 0 corresponding to the semblance of events at each location. A conventional interpretation process takes another step to map the discontinuity surfaces with the help of various visualization tools.

It is obvious that the discontinuity surfaces will be located somewhere in the region where the coherency value is lower than its neighbor. The shape of the discontinuity surface will be similar to the shape of the region that has lower coherency values than others. Therefore, a rough 3D shape of the discontinuity surfaces can be obtained in a binary image form obtained by thresholding the coherency cube as follows:

\[ B(t, x, y) = \begin{cases} 
0 & \text{if } coh(t, x, z) \leq coh_{th} \\ 
1 & \text{otherwise} 
\end{cases} \]  

(3)

where the thresholding value, \( coh_{th} \), is determined empirically with a trial and error approach. The image needs to be histogram-equalized before thresholding, which requires the image to be quantized. I quantized the coherency value by rounding it to the nearest hundredth value, then applied histogram equalization to make sure each coherency value is distributed evenly so that the thresholding value change effects for the binary image shape appropriately. Figures 5 and 6 show the histogram of quantized coherency values and its result after histogram equalization, respectively. The effect of the quantization and the equalization on the image can be seen in Figures 7 and 8. By increasing the contrast, the unclear low coherency region in Figure 7 becomes clear in Figure 8.

![Histogram of the quantized coherency cube.](jun1-histogram)

The effects on the binary image shape with respect to different thresholding values are shown in Figures 9 through 11. Figure 9, 10 and 11 are images obtained by thresholding Figure 8 with 0.2, 0.3, and 0.4, respectively. From those figures, we can see that increasing the thresholding value makes the binary image shape get thicker, as expected.
Figure 6: Histogram of the quantized coherency cube after histogram equalization.

Figure 7: The coherency cube after quantization applied with 0.01 quantization level.
Figure 8: The coherency cube after the histogram equalization applied. [jun1-coh-heq] [NR]

Figure 9: Binary image obtained by selecting points whose values are less than 0.2 from the histogram equalized coherency cube. [jun1-binary20] [NR]
Figure 10: Binary image obtained by selecting points whose values are less than 0.3 from the histogram equalized coherency cube. [jun1-binary-30] [NR]

Figure 11: Binary image obtained by selecting points whose values are less than 0.4 from the histogram equalized coherency cube. [jun1-binary-40] [NR]
DISCONTINUITY SURFACE EXTRACTION

The binary image obtained in the previous section represents the shape of the region where the discontinuity surfaces may exist. Therefore, the discontinuity location could be found by thinning the 3D binary image. There are various thinning algorithms developed in the image processing field and they have been used for various applications in topological analysis. I used a 2D thinning algorithm and adapted it into 3D image skeletonization.

The conventional 2D thinning algorithm can be explained as follows: for a given pixel location whose value is 1, its value will be changed to 0 if the neighboring pixels have values like one of the patterns shown in equation (4). This operation continues until no pixels change. It will leave thin lines a pixel in width. The structure elements used in 2D thinning are given as follows:

\[
L_1 = \begin{bmatrix}
0 & 0 & 0 \\
* & 1 & *
\end{bmatrix}, \quad L_2 = \begin{bmatrix}
* & 0 & 0 \\
1 & 1 & 0
\end{bmatrix}, \ldots
\]  

(4)

where the central location “1” corresponds to the pixel for operation and “*” corresponds to unused pixels. The rest of the 6 elements can be found by rotating the above element appropriately (Sonka et al., 1999). Direct extending this 2D algorithm into a 3D algorithm, will result in a 3D thinning algorithm that produces 3D lines (Ma and Sonka, 1996). Such line searching does not meet our goal of finding 3D surfaces by squeezing a 3D region. Therefore, I iteratively applied the 2D thinning algorithms for each orthogonal direction, resulting in the desired 3D thinning algorithm.

First, the 2D thinning algorithm is applied once for each \((x, y)\), \((y, z)\), and \((x, z)\) plane separately, and the pixel is considered to be removed only if the removal happens in at least two of the three planes. The thinning continues until no pixels change. Finally, this results in 3D surfaces with a width of one pixel, representing the locations of discontinuities. This process can be summarized in a pseudo code as follows:

```plaintext
old_data(x,y,z) = binary_data(x,y,z)
condition = 1
do while (condition .eq. 1)
{
    temp1(x,y,z) = 2D thinning along every (x,y) plane on old_data(x,y,z)
    temp2(x,y,z) = 2D thinning along every (y,z) plane on old_data(x,y,z)
    temp3(x,y,z) = 2D thinning along every (x,z) plane on old_data(x,y,z)
    new_data = [(temp1.and.temp2).or.(temp2.and.temp3).or.(temp1.and.temp3)]
    if(new_data .eq. old_data) { condition = 0 }
    else
        { old_data(x,y,z) = new_data(x,y,z) }
}
```

The above 3D thinning algorithm is applied to the three differently thresholded images shown in Figure 9, Figure 10, and Figure 11. The results are shown at Figure 12, Figure 13, and Figure 14, respectively. These figures show that as the thresholding value increases, more
detailed shapes of the discontinuities can be found without affecting the major structure of the discontinuity surfaces. This reveals the robustness of the method in choosing the threshold values. By overlapping the discontinuity map onto the original coherency cube, we can see that the locations of discontinuities are accurately positioned, as expected (Figure 15).

![Discontinuity surfaces](image)

Figure 12: Discontinuity surfaces found by thinning the binary image with thresholding value of 0.2.

CONCLUSIONS

I describe an automatic algorithm for discontinuity extraction from a 3D seismic image cube. The proposed algorithm consists of three steps. The first step is the coherency computation which results in a coherency cube that gives event semblance at each point. Then next step is to represent the potential discontinuity locations in binary image form. This is accomplished by thresholding the histogram-equalized coherency cube. The final step is finding location of the discontinuity by thinning the binary image obtained in the previous step. For thinning in the 3D sense, a 2D thinning algorithm is consecutively applied to produce an arbitrary shaped 3D surface.

Testing the algorithm on a real seismic image demonstrated that it can successfully find discontinuity surfaces. The extracted discontinuity surfaces could be interpreted as fault, unconformity, or buried channel. However, in order to make the proposed algorithm a more useful tool that produces geologically meaningful surfaces, further research needs to be followed. Two areas of research are: (1) The development of topological relationships between
Figure 13: Discontinuity surfaces found by thinning the binary image with thresholding value of 0.3.

Figure 14: Discontinuity surfaces found by thinning the binary image with thresholding value of 0.4.
Figure 15: Discontinuities with thresholding value of 0.3 are overlapped with coherency cube to examine the accuracy of the location. To visualize the location of the discontinuities, low values in the coherency cube are expressed in bright grey and high values are expressed in dark grey. jun1-thincoh30 [NR]
surfaces to determine whether they are connected or not, and (2) The inclusion of smoothness criteria, as geology may often dictate that some discontinuity surfaces are expected to be smooth.

**REFERENCES**


Elastic and poroelastic analysis of Thomsen parameters for seismic waves in finely layered VTI media

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. 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 (as much as a factor of 5 in the examples presented here) when the medium behaves in an undrained fashion, as might be expected 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. The results presented are strictly for velocity analysis, not for amplitude or attenuation.

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 it is possible to understand deviations from Gassmann’s results at higher frequencies when the rock is heterogeneous, 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 consider both closed-pore and open-pore boundary conditions between layers within this model in order to study in detail 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, 2003) 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 for weak 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 recalling 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 \( e_{ij} \) is given by

\[
\begin{pmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{33} \\
\sigma_{23} \\
\sigma_{31} \\
\sigma_{12}
\end{pmatrix} = \begin{pmatrix}
\alpha & \beta & \gamma \\
\beta & \alpha & \gamma \\
\gamma & \gamma & \gamma \\
\gamma & \gamma & \gamma \\
\gamma & \gamma & \gamma \\
\gamma & \gamma & \gamma
\end{pmatrix} \begin{pmatrix}
e_{11} \\
e_{22} \\
e_{33} \\
e_{23} \\
e_{31} \\
e_{12}
\end{pmatrix},
\]

where \( \alpha = b + 2m \) (e.g., Musgrave, 1970; Auld, 1973), with \( i, j, k, l \) obviously each ranging from 1 to 3 in Cartesian coordinates. The matrix describes isotropic media in the special case when \( a = c = \lambda + 2\mu, \beta = f = \lambda, \) and \( \lambda = m = \mu. \)

The Thomsen (1986) parameters \( \epsilon, \delta, \) and \( \gamma \) are related to these stiffnesses by

\[
\epsilon \equiv \frac{a - c}{2c},
\]

\[
\delta \equiv \frac{(f + l)^2 - (c - l)^2}{2c(c - l)},
\]

\[
\gamma \equiv \frac{2f + 2l}{c}.
\]
For P-wave propagation in the earth near the vertical, the important anisotropy parameter is \( \delta \). For SV-wave propagation near the vertical, the combination \((\epsilon - \delta)\) 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.

It is also useful to note for later reference that

\[
a = c(1 + 2\epsilon), \quad m = l(1 + 2\gamma), \quad \text{and} \quad f \simeq c(1 + \delta) - 2l. \tag{5}
\]

In TI media, \( c \) and \( l \) are 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:

\[
K^* = K_{dr}/(1 - \alpha B), \tag{6}
\]

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 the drained \( (\mu_{dr}) \) and saturated \( (\mu^*) \) media, Gassmann’s quasi-static theory gives

\[
\mu^* = \mu_{dr}. \tag{7}
\]

We want to emphasize once more that (7) is a result of the theory, not an assumption. It follows immediately from (6) for any isotropic poroelastic medium. Furthermore, the two equations (6) and (7) taken together show that, for isotropic microhomogeneous media, the fluid effect 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 isotropic. The key idea presented by Backus is that these equations can be rearranged into a form where rapidly varying coefficients multiply slowly varying stresses or strains.

The derivation has been given many places including Schoenberg and Muir (1989) and Berryman (1999a). One illuminating derivation given recently by Milton (2002) will be followed here, with the main difference being that we assume the layering direction is $z$ or 3. We break the equation down into $3 \times 3$ pieces so that

$$
\begin{pmatrix}
\sigma_{11} \\
\sigma_{22} \\
\sigma_{12}
\end{pmatrix} = A_{11} \begin{pmatrix}
e_{11} \\
e_{22} \\
e_{12}
\end{pmatrix} + A_{13} \begin{pmatrix}
e_{33} \\
e_{23} \\
e_{31}
\end{pmatrix},
$$

(8)

and

$$
\begin{pmatrix}
\sigma_{33} \\
\sigma_{23} \\
\sigma_{31}
\end{pmatrix} = A_{31} \begin{pmatrix}
e_{11} \\
e_{22} \\
e_{12}
\end{pmatrix} + A_{33} \begin{pmatrix}
e_{33} \\
e_{23} \\
e_{31}
\end{pmatrix},
$$

(9)

where the $3 \times 3$ matrices are

$$
A_{11} = \begin{pmatrix}
\lambda + 2\mu & \lambda & 0 \\
\lambda & \lambda + 2\mu & 0 \\
0 & 0 & 2\mu
\end{pmatrix}, \quad A_{13} = A_{31}^T = \begin{pmatrix}
\lambda & 0 & 0 \\
0 & \lambda & 0 \\
0 & 0 & 2\mu
\end{pmatrix}, \quad A_{33} = \begin{pmatrix}
\lambda + 2\mu & 0 & 0 \\
0 & 2\mu & 0 \\
0 & 0 & 2\mu
\end{pmatrix}.
$$

(10)

Noting that the variables $\sigma_{11}$, $\sigma_{22}$, $\sigma_{12}$, $e_{33}$, $e_{23}$, and $e_{31}$ are fast variables in the layers, and all the remaining variables are slow (actually constant), it is advantageous to rearrange these equations so the slow variables multiply the elastic parameter matrices and are all on one side of the equations, while the fast variables are all alone on the other side of the equations. Then, it is trivial to perform the layer averages, since they depend only on the (assumed known) values of the elastic parameters in the layers and are multiplied by constants. Having done this, we can then transform back into the standard forms of (8) and (9) with the stresses and strains now reinterpreted as the overall values, and find the following relationships (where the star indicates the effective property of the layered system):

$$
A_{33}^* = \left\langle A_{33}^{-1} \right\rangle^{-1},
$$

(11)

$$
A_{13}^* = (A_{31}^*)^T = \left\langle A_{13}A_{33}^{-1} \right\rangle A_{33}^*,
$$

(12)

and

$$
A_{11}^* = \left\langle A_{11} \right\rangle + A_{13}^* \left( A_{33}^* \right)^{-1} A_{31}^* - \left\langle A_{13}A_{33}^{-1}A_{31} \right\rangle.
$$

(13)

The brackets $\langle x \rangle$ indicate the volume (or equivalently the one-dimensional layer) average of the quantity $x$ in the simple layered medium under consideration. It follows that the anisotropy
coefficients in equation (1) are then related to the layer parameters by the following expressions:

\[ c = \left( \frac{1}{\lambda + 2\mu} \right)^{-1}, \]  

\[ f = c \left( \frac{\lambda}{\lambda + 2\mu} \right), \]  

\[ l = \left( \frac{1}{\mu} \right)^{-1}, \]  

\[ m = (\mu), \]  

\[ a = \frac{f^2}{c} + 4m - 4 \left( \frac{\mu^2}{\lambda + 2\mu} \right), \]  

and

\[ b = a - 2m. \]  

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 \( -\frac{2}{3}\mu \leq \lambda \leq \infty \). Although, for physically stable materials, \( \mu \) and the bulk modulus \( K = \lambda + \frac{2}{3}\mu \) must both be nonnegative, \( \lambda \) (and also Poisson’s ratio) may be negative. 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.

One very important fact that is known about these 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, it 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).

**THOMSEN PARAMETERS \( \epsilon \) AND \( \delta \)**

**Thomsen’s \( \epsilon \)**

An important anisotropy parameter for quasi-SV-waves is Thomsen’s parameter \( \epsilon \), defined in equation (2). Formula (18) for \( a \) may be rewritten as

\[ a = \left( \frac{(\lambda + 2\mu)^2 - \lambda^2}{\lambda + 2\mu} \right) + c \left( \frac{\lambda}{\lambda + 2\mu} \right)^2, \]  

(20)
which can be rearranged into the convenient and illuminating form
\[
a = (\lambda + 2\mu) - c \left[ \frac{\lambda^2}{\lambda + 2\mu} \left( \frac{1}{\lambda + 2\mu} \right) - \frac{\lambda}{\lambda + 2\mu} \right]^2. \tag{21}
\]
This formula is very instructive because the term in square brackets is in Cauchy-Schwartz form \([(\alpha^2)\beta^2] \geq (\alpha\beta)^2\), so this factor is nonnegative. Furthermore, the magnitude of this term depends mainly on the fluctuations in the \(\lambda\) Lamé parameter, 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 \(\varepsilon\) written in a similar fashion as
\[
2\varepsilon = \left[ (\lambda + 2\mu) \left( \frac{1}{\lambda + 2\mu} \right) - 1 \right] - \left[ \left( \frac{\lambda^2}{\lambda + 2\mu} \right) \left( \frac{1}{\lambda + 2\mu} \right) - \frac{\lambda}{\lambda + 2\mu} \right]^2, \tag{22}
\]
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 \(2\varepsilon\) unless \(\lambda = \text{constant}\), in which case it vanishes. So, the sign of \(\varepsilon\) is indeterminate. The Thomsen parameter \(\varepsilon\) may have either a positive or a negative sign for a TI medium composed of arbitrary thin isotropic layers.

Helbig and Schoenberg (1987) discuss an interesting case where large fluctuations in \(\mu\) combined with large fluctuations in \(\lambda\), including \(\lambda < 0\) for one component, lead to wavefronts with anomalous polarizations in layered TI media. Schoenberg (1994) also discusses several shale examples having large fluctuations in both \(\lambda\) and \(\mu\).

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 variations in layer \(\mu\) should have no direct information about fluid content, while observed variations in layer \(\lambda\), especially if they are large variations, may contain important clues about variations in fluid content. So the observed structure of \(\varepsilon\) in (22) strongly suggests that small positive and all negative values of \(\varepsilon\) may be important indicators of significant fluctuations in fluid content.

From (21), we can infer in general that
\[
a \leq (\lambda + 2\mu), \tag{23}
\]
so we have an upper bound on \(\varepsilon\) in finely layered media stating that
\[
2\varepsilon \leq \left( (\lambda + 2\mu) \left( \frac{1}{\lambda + 2\mu} \right) - 1 \right) = \left( (\rho v_p^2) (\rho^{-1} v_p^{-2}) - 1 \right), \tag{24}
\]
where \(\rho\) is the density.
Thomsen’s $\delta$

Thomsen’s parameter $\delta$ defined by Eq. (3) is pertinent for near vertical quasi-$P$-waves and can also be rewritten as

$$
\delta = \frac{(c + f)(c - f - 2l)}{2c(c - l)}.
$$

This parameter is considerably more difficult to analyze than either $\gamma$ or $\epsilon$ for various reasons, some of which we will enumerate shortly.

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.

The analysis of Berryman et al. (1999) 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 layered models 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

The singular value decomposition (SVD), or equivalently the eigenvalue decomposition in the case of a real symmetric matrix, for (1) is relatively easy to perform. We can immediately write down four eigenvectors:

$$
\begin{pmatrix}
0 \\
0 \\
0 \\
1
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
1
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}, \begin{pmatrix}
1 \\
1 \\
0 \\
0
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}, \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}.
$$

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

\[
\begin{pmatrix}
1 \\
1 \\
X \\
0 \\
0
\end{pmatrix},
\] (27)

with the corresponding eigenvalue

\[
\chi = a + b + fX.
\] (28)

The remaining condition that determines both \(X\) and \(\chi\) is

\[
\chi X = 2f + cX,
\] (29)

which, after substitution for \(\chi\), leads to a quadratic equation having the solutions

\[
X_\pm = \frac{1}{2} \left( -\left[ \frac{a + b - c}{f} \right] \pm \sqrt{8 + \left[ \frac{a + b - c}{f} \right]^2} \right).
\] (30)

The ranges of values for \(X_\pm\) are \(0 \leq X_+ \leq \infty\) 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. For all other cases, these two modes 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 (26) 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

\[
\alpha \begin{pmatrix}
1 \\
1 \\
X_+ \\
0 \\
0
\end{pmatrix} + \beta \begin{pmatrix}
1 \\
1 \\
X_- \\
0 \\
0
\end{pmatrix} = (1 + \alpha) \begin{pmatrix}
1 \\
1 \\
0 \\
0 \\
0
\end{pmatrix},
\] (31)

with \(\alpha = -2(X_+ - 1)/[X_+(X_+ + 2)]\) for pure shear, and

\[
\alpha \begin{pmatrix}
1 \\
1 \\
X_+ \\
0 \\
0
\end{pmatrix} + \beta \begin{pmatrix}
1 \\
1 \\
X_- \\
0 \\
0
\end{pmatrix} = (1 + \beta) \begin{pmatrix}
1 \\
1 \\
0 \\
0 \\
0
\end{pmatrix},
\] (32)
and with $\beta = X_+(X_+ - 1)/(X_+ + 2)$ 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 in terms of layer elasticity parameters**

Combining results from Eqs. (18)–(15), we find after some work on rearranging the terms that

$$
\frac{a + b - c}{f} = \left(\frac{\lambda}{\lambda + 2\mu}\right)^{-1} \left[\left(\frac{\lambda}{\lambda + 2\mu}\right) + 6\left(\frac{m - \mu}{\lambda + 2\mu}\right)
- 8 \left(\frac{\mu^2}{\lambda + 2\mu}\right)\left(\frac{1}{\lambda + 2\mu}\right) - \left(\frac{\mu}{\lambda + 2\mu}\right)^2\right],
$$

(33)

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 (33) is again in Cauchy-Schwartz form (i.e., $\langle a^2 \rangle \langle b^2 \rangle - \langle ab \rangle^2 \geq 0$), 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 layering are (as expected) the main source of deviations of (33) 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 for small values of Thomsen parameters**

Using the definitions of the Thomsen parameters, we can also rewrite the terms appearing in (33) in order to make connection with this related point of view. Recalling (5) and the fact that $b = a - 2m$, we have

$$
\frac{a + b - c}{f} \approx 1 + \frac{3}{c - 2l}(c\delta + 4l\gamma) + \frac{4}{c - 2l}\left[c(\epsilon - \delta) - 4l\gamma\right],
$$

(34)
with some higher order corrections involving powers of $\delta$ and products of $\delta$ with $\epsilon$ and $\gamma$ that we have neglected here. 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 (34) and that found previously in (33). 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 (33). Since the formula (34) is only approximate and its interpretation requires the use of various other results we derive later for other purposes, we will for now delay further discussion of this point to the end of the paper.

**DISPERSION RELATIONS FOR SEISMIC WAVES**

The general behavior of seismic waves in anisotropic media is well known, and the equations are derived in many places including Berryman (1979) and Thomsen (1986). The results are

$$\rho \omega^2_{\pm} = \frac{1}{2} \left\{ (a + l)k_1^2 + (c + l)k_3^2 \pm \sqrt{[(a - l)k_1^2 - (c - l)k_3^2]^2 + 4(f + l)^2k_1^2k_3^2} \right\},$$

(35)

for compressional (+) and vertically polarized shear (−) waves and

$$\rho \omega^2_s = mk_1^2 + l k_3^2,$$

(36)

for horizontally polarized shear waves, where $\rho$ is the overall density, $\omega$ is the angular frequency, $k_1$ and $k_3$ are the horizontal and vertical wavenumbers (respectively), and the velocities are given simply by $v = \omega / k$ with $k = \sqrt{k_1^2 + k_3^2}$. The SH wave depends only on elastic parameters $l$ and $m$, which are not dependent in any way on layer $\lambda$ and therefore will play no role in the poroelastic analysis. Thus, we can safely ignore SH except when we want to check for shear wave splitting (bi-refringence) – in which case the SH results will be useful for the 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

$$\rho \omega^2_{\pm} = \frac{1}{2} \left\{ (a + l)k_1^2 + (c + l)k_3^2 \pm \right.$$  
$$\sqrt{[(a + l)k_1^2 + (c + l)k_3^2]^2 - 4(a k_1^2 + c k_3^2)l k^2 + [(a - l)(c - l) - (f + l)^2k_1^2k_3^2]} \right\},$$

(37)

Written this way, it is then obvious that the following two relations hold:

$$\rho \omega^2_+ + \rho \omega^2_- = (a + l)k_1^2 + (c + l)k_3^2,$$

(38)

and

$$\rho \omega^2_+ \cdot \rho \omega^2_- = (a k_1^2 + c k_3^2)l k^2 + [(a - l)(c - l) - (f + l)^2k_1^2k_3^2],$$

(39)
either of which could have been obtained directly from (35) without the intermediate step of (37).

We are motivated to write the equations in this way in order to try to avoid evaluating the square root in (35) directly. Rather, we would like to arrive at a natural approximation that is quite accurate, but does not involve the square root operation. From a general understanding of the problem, it is clear that a reasonable way of making use of (38) is to make the identifications

\[ \rho \omega^2_+ = ak_1^2 + ck_3^2 - \Delta, \]  

and

\[ \rho \omega^2_- = lk^2 + \Delta, \]  

with \( \Delta \) still to be determined. Then, substituting these expressions into (39), we find that

\[ (ak_1^2 + ck_3^2 - lk^2 - \Delta)\Delta = [(a - l)(c - l) - (f + l)^2]k_1^2 k_3^2 \]  

Solving (42) for \( \Delta \) would just give the original results back again. So the point of (42) 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 (42), or we could just keep a small number of terms in an expansion.

The leading term, and the only one we will consider here, is

\[ \Delta = \frac{[(a - l)(c - l) - (f + l)^2]k_1^2 k_3^2}{(a - l)k_1^2 + (c - l)k_3^2 - \Delta} \approx \frac{[(a - l)(c - l) - (f + l)^2]}{(a - l) / k_1^2 + (c - l) / k_3^2}. \]

The numerator of this expression is known to be a positive quantity for layered materials (Postma, 1955; Berryman, 1979). Furthermore, it can be rewritten in terms of Thomsen’s parameters as

\[ [(a - l)(c - l) - (f + l)^2] = 2c(c - l)(\epsilon - \delta). \]  

Using the first of the identities noted earlier in (5), we can also rewrite the first elasticity factor in the denominator as \( a - l = (c - l)[1 + 2ce/(c - l)] \). Combining these results in the limit of \( k_1^2 \to 0 \) (for relatively small horizontal offset), we find that

\[ \rho \omega^2_+ \simeq ck^2 + 2c\delta k_1^2, \]  

and

\[ \rho \omega^2_- \simeq lk^2 + 2c(\epsilon - \delta)k_1^2, \]  

with \( \Delta \simeq 2c(\epsilon - \delta)k_1^2 \). Improved approximations to any desired order can be obtained with only a little more effort by using (42) or (43) instead of the first approximation used here. But (45) and (46) are satisfactory for our present purposes.
INTERPRETATION OF P AND SV COEFFICIENTS FOR LAYERED MEDIA

General analysis for VTI media

The correction terms for SV waves in weakly anisotropic media are proportional to the factor

$$A = (a - l)(c - l) - (f + l)^2 = 2c(e - l)(e - \delta),$$

which is sometimes called the *anellipticity parameter*. For the case of weak anisotropy that we are considering here, the presence of this term in (46) just introduces ellipticity into the move out, but the higher order corrections that we neglected can introduce deviations from ellipticity, hence anellipticity.

Clearly, from (46) for quasi-SV-waves [and in layered media at this order of approximation], the anellipticity parameter holds all the information about 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

$$A = (f + l)(a + c - 2f - 4l) + (a - f - 2l)(c - f - 2l),$$

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 effective shear moduli $\mu_1^*$ and $\mu_3^*$ associated with $a$ and $c$, namely,

$$\mu_1^* = a - m - f \quad \text{and} \quad 2\mu_3^* = c - f.$$  

Furthermore, it was shown that the combination defined by

$$G_{\text{eff}} = (\mu_1^* + 2\mu_3^*)/3$$

plays a special role in the theory, as it is the only effective shear modulus for the anisotropic system that may also contain information about fluid content. It turns out that (48) can be rewritten in terms of this effective shear modulus if we first introduce two more parameters:

$$\mathcal{K} = f + l + \left[ \frac{1}{a - f - 2l} + \frac{1}{c - f - 2l} \right]^{-1}$$

and

$$\mathcal{g} = \left[ 3G_{\text{eff}} + m - 4l \right]/3.$$ 

Then, (48) can be simply rewritten as

$$A = 3\mathcal{K}\mathcal{g}.$$ 

This result is analogous to, but distinct from, a product formula relating the effective shear modulus $G_{\text{eff}}$ and the bulk modulus

$$K = f + \left[ \frac{1}{a - m - f} + \frac{1}{c - f} \right]^{-1}$$
to the eigenvalues of the elastic matrix according to

\[ \chi_+ \chi_- = 6K G_{\text{eff}}. \]  

(55)

In the isotropic limit for layered materials, when \( \mu \to \text{constant} \), we have \( K \to f + 2\mu/3 \), while \( K \to f + \mu \). 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 \( G \to 0 \) in the same limit. It is also possible to show for layered materials that in general \( l \leq K - f \leq m \), with the lower limit being optimum.

Also, since Thomsen’s \( \delta \) plays an important role in (45), it is helpful to note that (25) can also be rewritten as

\[ c \delta = -(c - f - 2l) \left[ 1 - \frac{c - f - 2l}{2(c - l)} \right], \]  

(56)

which shows that, at least for weakly anisotropic media, \( c \delta \) is very nearly a direct measure of the quantity \( c - f - 2l \).

**Analysis for layered media**

The analysis presented so far is general for all VTI elastic media. But we can say more by assuming now that the anisotropy arises due to layering. Then, for example, we have the following relations

\[ f + 2l = c \left( \frac{\lambda + 2l}{\lambda + 2\mu} \right), \]  

(57)

\[ c - f - 2l = 2c \left( \frac{\mu - l}{\lambda + 2\mu} \right), \]  

(58)

and

\[ a - f - 2l = 2c \left\{ \left( \frac{2m - \mu - l}{\lambda + 2\mu} \right) - 2 \left[ \left( \frac{\mu^2}{\lambda + 2\mu} \right) \left( \frac{1}{\lambda + 2\mu} \right) - \left( \frac{\mu}{\lambda + 2\mu} \right)^2 \right] \right\}. \]  

(59)

Eq. (57) is an easy consequence of the Backus averaging formulas. Then, (58) shows that \( c \) differs from \( f + 2l \) only by a term that measures the difference in the weighted average of \( \mu \) and \( l \). Eq. (59) 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

\[ G_{\text{eff}} = m - \frac{4c}{3} \left[ \left( \frac{\mu^2}{\lambda + 2\mu} \right) \left( \frac{1}{\lambda + 2\mu} \right) - \left( \frac{\mu}{\lambda + 2\mu} \right)^2 \right], \]  

(60)

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 (58) and (59) 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,

\[ \mathcal{A} \simeq (c - l)(3G_{\text{eff}} + m - 4l) \]  

(61)

from which we can conclude that the important coefficient in (46) is given to a good approxi-
mation by

\[ 2c(\epsilon - \delta) \simeq 3G_{\text{eff}} + m - 4l \sim 4(m - l) = 8l\gamma, \]

(62)

where the final expression is a statement about the limiting behavior when either the \( \mu \) fluc-
tuations are very small, or when strong undrained behavior is present together with large \( \mu \)
fluctuations.

To study the fluid effects, the drained Lamé parameter \( \lambda \) in each layer should be replaced
under undrained conditions by

\[ \lambda^* = K^* - 2\mu/3, \]  

(63)

where \( K^* \) was defined by (6). Then, for small fluctuations in \( \mu \), Eq. (62) 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). There is an enhancement
in the shear wave speed for SV in layered media, just due to the changes in the shear moduli,
and independent of any fluids that might be present in that case, but the magnitude of this
enhancement is small because the difference \( m - l \) is also small. When \( m - l \) is large, then the
magnitude of the enhancement due to liquids in the pores can be very substantial as we will
see in the following examples. So the effects of liquid on \( G_{\text{eff}} \) will generally be weak when
the fluctuations in \( \mu \) are weak, and strong when they are strong.

To check the corresponding result for P-waves, we need to estimate \( \delta \). Making use of (56),
we also have

\[ c\delta = -2c \left\{ \frac{\mu - l}{\lambda + 2\mu} \right\} \left[ 1 - l^{-1} \left\{ \frac{\lambda + \mu}{\mu(\lambda + 2\mu)} \right\}^{-1} \left\{ \frac{\mu - l}{\lambda + 2\mu} \right\} \right]. \]  

(64)

Working to the same order as we did for the final expression in (62), we can neglect the second
term in the square brackets of (64). What remains shows that pore fluids would have an effect
on this result. The result is

\[ c^*\delta^* \simeq -2c^* \left\{ \frac{\mu - l}{\lambda^* + 2\mu} \right\}, \]  

(65)

and a similar replacement should also be made for \( G_{\text{eff}} \) in (60). Eq. (65) 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.
Now we have derived all the results needed to interpret (34) and show how it is related to (33). First, we note the some of the main terms missing from (34) are those due to approximations made to $\delta$ and the denominators of (33), which have been approximated as $f \simeq c - 2l$ instead of $f \simeq c(1 + \delta) - 2l$. Then, from (62), it is easy to see that the final term in (34) vanishes to lowest order, and that the remainder is given exactly by the shear modulus fluctuation terms in brackets in (59) — in complete agreement with the final terms of (33). Then, from (64), it follows that the leading contribution to the factor $c\delta + 4l\gamma$ is

$$c\delta + 4l\gamma \simeq 2c\left(\frac{m - \mu}{\lambda + 2\mu}\right),$$

in complete agreement with the second term in (33).

In the case of very strong fluctuations in the layer shear moduli, then (59) and (64) both show that effects of the pore fluids can be more strongly felt in the anisotropy correction factors $2c^2(\varepsilon^s - \delta^s)$ and $2c^2\delta^s$ for undrained porous media, and therefore more easily observed in seismic, sonic, or ultrasonic data. 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: All the liquid dependence in the shear moduli comes into the wave dispersion formulas through coefficient $a$ (or equivalently $e$). Equations (59) and (60) 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, I made use of a code of V. Grechka [used previously in a joint publication (Berryman et al., 1999)] and then I arbitrarily picked 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 $V_{sv}$ 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 $5 \times [1/(1 - \alpha)]$ for the present examples.

Table 1. Layer parameters for the three materials in a simple layered medium used to produce the examples in Figures 1 and 2.

<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>

We took the porosity to be $\phi = 0.2$, and the overall density to be $\rho = (1 - \phi)\rho_i + \phi S\rho_l$, where $\rho_i = 2650.0$ 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 is still 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 the liquid to start becoming more 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, poroelastic medium.
TABLE 2. The VTI elastic coefficients and Thomsen parameters for the materials (see Table 1) used in the computed examples of Figures 1 and 2.

<table>
<thead>
<tr>
<th>Elastic Parameters and Density</th>
<th>Case</th>
<th>Case</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B = 0$</td>
<td>$B = \frac{1}{2}$</td>
<td>$B = 1$</td>
</tr>
<tr>
<td>$a$ (GPa)</td>
<td>33.8345</td>
<td>50.3523</td>
<td>132.7003</td>
</tr>
<tr>
<td>$c$ (GPa)</td>
<td>33.1948</td>
<td>50.4715</td>
<td>134.2036</td>
</tr>
<tr>
<td>$f$ (GPa)</td>
<td>22.2062</td>
<td>38.5857</td>
<td>120.7006</td>
</tr>
<tr>
<td>$l$ (GPa)</td>
<td>4.0138</td>
<td>4.0138</td>
<td>4.0138</td>
</tr>
<tr>
<td>$m$ (GPa)</td>
<td>6.7777</td>
<td>6.7777</td>
<td>6.7777</td>
</tr>
<tr>
<td>$G_{eff}$ (GPa)</td>
<td>5.2797</td>
<td>5.8841</td>
<td>6.2417</td>
</tr>
<tr>
<td>$\delta$</td>
<td>-0.0847</td>
<td>-0.0733</td>
<td>-0.0399</td>
</tr>
<tr>
<td>$\epsilon - \delta$</td>
<td>0.0943</td>
<td>0.0745</td>
<td>0.0343</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.3443</td>
<td>0.3443</td>
<td>0.3443</td>
</tr>
<tr>
<td>$\rho$ (kg/m$^3$)</td>
<td>2120.0</td>
<td>2310.0</td>
<td>2320.0</td>
</tr>
</tbody>
</table>

The results shown in the Figures are in complete qualitative and quantitative agreement with the analytical predictions described, as expected.

CONCLUSIONS

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 pore fluids on the elastic or poroelastic constants — 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.
Figure 1: 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.

ACKNOWLEDGMENTS

I thank P. A. Berge and V. Y. Grechka for their insight and collaboration during our earlier related studies in this research area. Work performed under the auspices of the U. S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48 and supported specifically by the Geosciences Research Program of the DOE Office of Energy Research within the Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences.
Figure 2: Vertically polarized shear wave speed $V_{sv}$ 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.

REFERENCES


Scale-up in poroelastic systems and applications to reservoirs

James G. Berryman

ABSTRACT
A fundamental problem of heterogeneous systems is that the macroscale behavior is not necessarily well-described by equations familiar to us at the meso- or microscale. In relatively simple cases like electrical conduction and elasticity, it is true that the equations describing macroscale behavior take the same form as those at the microscale. But in more complex systems, these simple results do not hold. Consider fluid flow in porous media where the microscale behavior is well-described by Navier-Stokes’ equations for liquid in the pores while the macroscale behavior instead obeys Darcy’s equation. Rigorous methods for establishing the form of such equations for macroscale behavior include multiscale homogenization methods and also the volume averaging method. In addition, it has been shown that Biot’s equations of poroelasticity follow in a scale-up of the microscale equations of elasticity coupled to Navier-Stokes. Laboratory measurements have shown that Biot’s equations indeed hold for simple systems but heterogeneous systems can have quite different behavior. So the question arises whether there is yet another level of scale-up needed to arrive at equations valid for the reservoir scale? And if so, do these equations take the form of Biot’s equations or some other form? We will discuss these issues and show that the double-porosity equations play a special role in the scale-up to equations describing reservoir behavior, for fluid pumping, geomechanics, as well as seismic wave propagation.

INTRODUCTION

Earth materials composing either aquifers or oil and gas reservoirs are generally heterogeneous, porous, and often fractured or cracked. Distinguishing water, oil, and gas using seismic signatures is a key issue in seismic exploration and reservoir monitoring. Traditional approaches to seismic monitoring have often used Biot’s theory of poroelasticity (Biot, 1941, 1956a,b, 1962; Gassmann, 1951). Many of the predictions of this theory, including the existence of the slow compressional wave, have been confirmed by both laboratory and field experiments (Plona, 1980; Berryman 1980a; Johnson et al., 1982; Chin et al., 1985; Winkler, 1985; Pride and Morgan, 1991; Thompson and Gist, 1993; Pride, 1994). Nevertheless, this theory always has been limited by an explicit assumption that the porosity itself is homogeneous. Although this assumption is often applied to acoustic or ultrasonic studies of many core samples in the laboratory setting, heterogeneity of porosity still exists in the form of both pores

1email: berryman@sep.stanford.edu
and cracks. One approach to dealing with this source of heterogeneity is to construct a model that is locally homogeneous (i.e., a finite element). This approach may be adequate for some applications, and is certainly amenable to study with large computers. However, such methods necessarily avoid the question of how we are to deal with heterogeneity on the local scale (i.e., much smaller than the block size or wavelength in the cases being studied). Double porosity models have been introduced as a means of dealing with these problems. Rather than trying to deal with all the heterogeneity at once, we choose to consider a model intended to capture two main features of importance. Just two types of porosity are often key at the reservoir scale: (1) Matrix porosity occupies a finite and substantial fraction of the volume of the reservoir. This porosity is often called the storage porosity since it stores the fluids of interest. (2) Fracture or crack porosity may occupy very little volume overall, but nevertheless has two very big effects on reservoir behavior. First the fractures/cracks drastically weaken the rock mechanically, so that a change in a very low effective stress level may introduce nonlinear geomechanical responses. The second effect is that fractures/cracks introduce a fast pathway for the fluid to escape from the reservoir. This effect is obviously key to reservoir analysis and the economics of fluid withdrawal.

Many attempts have been made to incorporate fractures into rock models, and especially models that try to account for compressional wave attenuation in rocks containing fluids. But these models have often been viscoelastic rather than poroelastic (Budiansky and O’Connell, 1976; O’Connell and Budiansky, 1977). Berryman and Wang (1995) showed how to make a rigorous extension of Biot’s poroelasticity to include fractures/cracks by making a generalization to double-porosity/dual-permeability media modeling. That work concentrated on geomechanics and fluid flow aspects of the problem in order to deal with the interactions between fluid withdrawal and elastic closure of fractures during reservoir drawdown. The resulting equations were later applied to the reservoir consolidation problem by Lewallen and Wang (1998). Berryman and Wang (2000) then showed how the double-porosity approach could be applied to wave propagation problems, thereby generalizing Biot’s work on waves to allow for heterogeneous porosities and permeabilities.

The present paper addresses the question of scale-up in heterogeneous reservoirs. If Biot’s equations of poroelasticity are the correct equations at the mesoscale, then what are the correct equations at the macroscale? We show that Biot’s equations are not the correct equations at the macroscale when there is significant heterogeneity in fluid permeability. However, the double-porosity dual-permeability approach appears to permit consistent modeling of such reservoirs and also shows that no further up-scaling is required beyond the double-porosity stage.

**EQUATIONS OF BIOT’S SINGLE-POROSITY POROELASTICITY**

For long-wavelength disturbances ($\lambda >> h$, where $h$ is a typical pore size) propagating through a single-porosity porous medium, we define average values of the (local) displacements in the solid and also in the saturating fluid. The average displacement vector for the solid frame is $u$, while that for the pore fluid is $u_f$. The average displacement of the fluid relative to the frame is $\mathbf{w} = \phi (\mathbf{u} - \mathbf{u}_f)$. For small strains, the frame dilatation is $e$, while the increment of fluid content
is defined by
\[ \zeta = -\nabla \cdot \mathbf{w} = \phi (e - e_f). \] (1)

With time dependence of the form \( \exp(-i \omega t) \), the coupled wave equations that follow in the presence of dissipation are
\[ -\omega^2 (\rho \mathbf{u} + \rho_f \mathbf{w}) = H \nabla e - C \nabla \zeta + \mu_d \left( \nabla^2 \mathbf{u} - \nabla e \right), \]
\[ -\omega^2 (\rho_f \mathbf{u} + q \mathbf{w}) = C \nabla e - M \nabla \zeta = -\nabla p_f, \] (2)
where \( \mu_d \) is the drained shear modulus, \( H, C, \) and \( M \) are bulk moduli,
\[ \rho = \phi \rho_f + (1 - \phi) \rho_m, \] (3)
and
\[ q = \rho_f \left[ \alpha/\phi + i F(\xi) \eta/\kappa \omega \right]. \] (4)

The kinematic viscosity of the liquid is \( \eta \); the permeability of the porous frame is \( \kappa \); the dynamic viscosity factor is given approximately [or see Johnson et al. (1987) for more discussion], for our choice of sign for the frequency dependence, by
\[ F(\xi) = \frac{1}{\eta} [\xi T(\xi)/[1 + 2T(\xi)/i \xi]], \] (5)
where
\[ T(\xi) = \frac{\text{ber}'(\xi) - i \text{bei}'(\xi)}{\text{ber}'(\xi) - i \text{bei}'(\xi)} \] (6)
and
\[ \xi \equiv (\omega/\omega_0)^{\frac{1}{2}} = (\omega \alpha \kappa/\eta \phi)^{\frac{1}{2}} = (\omega h^2/\eta)^{\frac{1}{2}}. \] (7)

The functions \( \text{ber}(\xi) \) and \( \text{bei}(\xi) \) are the real and imaginary parts of the Kelvin function. The dynamic parameter \( h \) is a characteristic length generally associated with and comparable in magnitude to the steady-flow hydraulic radius. The tortuosity \( \alpha \geq 1 \) is a pure number related to the frame inertia which has been measured (Johnson et al., 1982) and has also been estimated theoretically (Berryman, 1980a; 1983a).

The coefficients \( H, C, \) and \( M \) are given by (Gassmann, 1951; Geertsma, 1957; Biot and Willis, 1957; Geertsma and Schmidt, 1961; Stoll, 1974)
\[ H = K_d + \frac{4}{3} \mu_d + (1 - K_d/K_m)^2 M, \] (8)
\[ C = (1 - K_d/K_m) M, \] (9)
where
\[ M = 1/[(1 - \phi - K_d/K_m)/K_m + \phi/K_f]. \] (10)
The constants are drained bulk and shear moduli $K_d$ and $\mu_d$, mineral bulk modulus $K_m$, and fluid bulk modulus $K_f$. Korringa (1981) showed equations (8)-(10) to be correct as long as the porous material may be considered homogeneous on the microscopic scale as well as the macroscopic scale. Also, see a recent tutorial on Gassmann’s equations (Gassmann, 1951) by Berryman (1999).

To decouple the wave equations (2) into Helmholtz equations for the three modes of propagation, we note that the displacements $u$ and $w$ can be decomposed as

$$u = \nabla \Upsilon + \nabla \times \vec{\beta}, \quad w = \nabla \psi + \nabla \times \vec{\chi},$$

(11)

where $\Upsilon$, $\psi$ are scalar potentials and $\vec{\beta}$, $\vec{\chi}$ are vector potentials. Substituting (11) into (2), we find (2) is satisfied if two pairs of equations are satisfied:

$$(\nabla^2 + k_s^2)\vec{\beta} = 0, \quad \vec{\chi} = -\rho_f \vec{\beta} / q$$

and

$$(\nabla^2 + k_{\pm}^2)A_{\pm} = 0.$$  

(13)

The wavenumbers in (12) and (13) are defined by

$$k_s^2 = \omega^2(\rho - \rho_f^2/q)/\delta \mu$$

and

$$k_{\pm}^2 = \frac{1}{2} \left[ b + f \mp \left( (b - f)^2 + 4cd \right)^{\frac{1}{2}} \right],$$

(15)

where

$$b = \omega^2(\rho M - \rho_f C)/\Delta, \quad c = \omega^2(\rho_f M - q C)/\Delta,$$

$$d = \omega^2(\rho_f H - \rho C)/\Delta, \quad f = \omega^2(q H - \rho_f C)/\Delta,$$

(16)

with

$$\Delta = HM - C^2.$$  

(17)

The linear combination of scalar potentials has been chosen to be

$$A_{\pm} = \Gamma_{\pm}\Upsilon + \psi,$$

(18)

where

$$\Gamma_{\pm} = d/(k_{\pm}^2 - b) = (k_{\pm}^2 - f)/c.$$  

(19)

With the identification (19), the decoupling is complete.
Figure 1: Thin layering of isotropic materials produces an effective transversely isotropic medium at low frequencies of propagation. Overall permeability $\kappa_{\text{eff}}$ normal to the layering depends most strongly on the most impermeable layers since $1/\kappa_{\text{eff}} = \int_0^L \kappa^{-1}(z)dz/L$, being the harmonic mean. In contrast, the seismic attenuation (in the usual band from 1–100 Hz) ordinarily depends most strongly on the ones that are most permeable, since $1/Q(z) \propto \kappa(z)$. The character of this relationship between attenuation and permeability changes significantly at higher frequencies as described in the text.

<table>
<thead>
<tr>
<th>Layer</th>
<th>$1/Q_n \propto \kappa_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_1$</td>
<td>$1/Q_1 \propto \kappa_1$</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>$1/Q_2 \propto \kappa_2$</td>
</tr>
<tr>
<td>$\kappa_3$</td>
<td>$1/Q_3 \propto \kappa_3$</td>
</tr>
<tr>
<td>$\kappa_4$</td>
<td>$1/Q_4 \propto \kappa_4$</td>
</tr>
<tr>
<td>$\kappa_5$</td>
<td>$1/Q_5 \propto \kappa_5$</td>
</tr>
</tbody>
</table>
| ... | ...
| $\kappa(z)$ | $1/Q(z) \propto \kappa(z)$ |
| ... | ...
| $\kappa_{n-1}$ | $1/Q_{n-1} \propto \kappa_{n-1}$ |
| $\kappa_n$ | $1/Q_n \propto \kappa_n$ |
LOW FREQUENCY ASYMPTOTICS FOR SINGLE-POROSITY

We will first demonstrate the dichotomy of interest by showing what Biot’s theory predicts if it is applied to heterogeneous reservoirs. The main issues with up-scaling in poroelasticity occur for the low frequency asymptotics, and so we limit discussion to this regime here. For low frequencies, all the wavelengths are long, thereby covering large regions of the heterogeneous medium, and so up-scaling is an issue that must always be addressed in this limit.

Compressional and Shear Waves

Compressional and shear waves have almost the same asymptotic behavior at low frequencies, but the analysis for shear waves is much shorter, so we will present only the shear wave analysis here.

The wavenumber \( k_s \) for shear wave propagation is determined by (14), and when \( \omega \to 0 \) we have \( q \to i \rho_f \eta / \kappa \omega \), so

\[
\frac{2}{\mu_d} k_s^2 = \frac{\omega^2 \rho}{\mu_d} \left[ 1 + i \frac{\rho_f \kappa \omega}{\rho \eta} \right].
\]  

Thus, when the loss tangent is a small number, we find the shear wave quality factor is

\[
1/Q_s \approx \frac{\rho_f \kappa \omega}{\eta \rho}.
\]  

Total attenuation along the path of a shear wave is then determined by the integral

\[
\int \frac{\rho_f \kappa \omega^2}{2 \eta (\rho M)^{1/2}} \, d \ell
\]

along the path of the wave. We assume for the sake of argument that the fluid is the same throughout the reservoir. So all fluid factors as well as frequency are constant. The solid material parameters \( \mu_d \) and \( \rho_m \) and also the porosity \( \phi \) (which is hidden in \( \rho \)) may vary in the reservoir, but these variations will be treated here as negligible compared the variations in the permeability \( \kappa \). Thus, we find that the total attenuation along a path of length \( L = \int d \ell \) is approximately proportional to \( \int \kappa \, d \ell \). The average attenuation per unit length of the travel path is therefore proportional to \( \int \kappa \, d \ell / L \), which is just the mean of the permeability along the wave’s path. This result is also true for the compressional waves, but the other multiplicative factors are a bit more complicated in that case.

Slow Waves

In contrast, the slow compressional wave can have two very different types of behavior at low frequency depending on the magnitude of the permeability. The wavenumber \( k_- \) for slow wave propagation is determined by (15). To simplify this equation, we note that it is an excellent approximation to take

\[
k_-^2 \approx b + f = \frac{\omega^2}{\Delta} \left[ q H - 2 \rho_f C + \rho M \right].
\]
So, at low frequencies, $k_+^2$ is proportional to $q$, whereas $k_-^2$ was inversely proportional to $q$. Then, for small frequencies but large values of the permeability, $q \to \rho_f[\alpha/\phi + i \eta/\kappa \omega]$. Substituting this into (22), we find that

$$k_-^2 = \frac{\omega^2}{\Delta} \left[ \alpha \rho_f H/\phi - 2 \rho_f C + \rho M + i \eta \rho_f H/\kappa \omega \right].$$

(23)

So as $\omega \to 0$ for large $\kappa$, there will be an intermediate frequency regime in which the slow wave has a well-defined quality factor

$$1/Q_\omega \simeq \eta \rho_f H/\kappa \omega (\alpha \rho_f H/\phi - 2 \rho_f C + \rho M),$$

(24)

which for strong frame materials reduces to

$$1/Q_\omega \simeq \eta \phi/\alpha \kappa \omega.$$

(25)

Except for some factors of density, porosity, and tortuosity, this expression is essentially the inverse of the corresponding expression for $1/Q_s$. Obviously both factors cannot be small simultaneously except for a very limited range of frequencies, which is determined by the factor $\alpha \rho_f/\phi \rho_f$. Although the tortuosity $\alpha \geq 1$ in general it can have a wide range of values, for granular media it is typical to find $\alpha \simeq 2$ or 3. In addition, $\alpha$ is also scale invariant, i.e., it does not depend on the size of the particles composing the granular medium. So, the presence of $\alpha$ multiplying $\kappa$ in (25) does not change the fact that the slow-wave attenuation is strongly influenced by fluctuations in the permeability $\kappa$. Being proportional to the square of the typical particle sizes, the permeability is itself not scale invariant. There is nevertheless a fairly small range of frequencies in which the approximation in (25) is valid, say from about 20 kHz to a few MHz for $\kappa$’s on the order of 1 D ($\simeq 10^{-12} \text{ m}^2$). This is the range where a propagating slow wave might be expected to be seen, and in fact has been observed in laboratory experiments (Plona, 1980).

For still smaller permeabilities or smaller frequencies or both, the leading approximation for the slow wave dispersion is instead given by

$$k_-^2 \simeq i \frac{\omega \eta \rho_f H}{\kappa \Delta}.$$

(26)

This type of dispersion relation corresponds to a purely diffusive process having a diffusion coefficient $\mathcal{D} \simeq M \kappa / \eta \rho_f$. This result follows directly from the second equation in (2) when the porous frame is sufficiently rigid.

We reach the same conclusion about how fluctuating permeability affects the propagation or diffusion of increments of fluid content (i.e., masses of excess fluid particles) in both of these cases. For the wave propagation situation of (25), we clearly have, by simple analogy to the arguments given already, that the average attenuation per unit length along the wave’s path is proportional to $\int \kappa^{-1} d\ell / L$. Similarly, in the limit of the diffusion process described by (26), then for a planar excitation diffusing through such a system in a direction perpendicular to the bedding planes, or for regions of isotropic random fluctuations in permeability, we again expect the overall effective diffusion rate to depend on the same average quantity: $\int \kappa^{-1} d\ell / L$. 
Thus, measurements of slow waves or of fluid increment diffusion on the macroscale will measure an effective permeability that is largely controlled by the smallest permeability present in the system. Clearly, this is exactly the opposite dependence we found for the dependence of the shear wave and also for the fast compressional wave, and must cause difficulties for up-scaling in Biot’s theory, where only one permeability parameter is available for the fitting of data.

Discussion

These observations show that there is a significant problem with up-scaling Biot’s theory, i.e., that the resulting system of equations is no longer of the same form as Biot’s theory. This is certainly no failing of Biot’s theory, but rather a failing of any attempted application of Biot’s theory directly to the up-scaled macro-system. Biot’s theory predicts correctly that compressional and shear wave attenuation both depend on the integral of the permeability $\kappa$ along the path of each wave. But the permeability itself along the same path averages as the inverse of the permeability (harmonic mean). Thus, the overall permeability depends most strongly on the smallest permeabilities present in the system, while the wave attenuation depends most strongly on the largest permeabilities in the system (Berryman, 1988). When we try to up-scale under these circumstances, we have an inherent problem due to the fact that Biot’s theory contains only one permeability; yet, for heterogeneous systems, there are two very distinct measures of permeability (the mean and the harmonic mean) that play significant roles.

SUMMARY OF DOUBLE-POROSITY WAVE PROPAGATION ANALYSIS

Berryman and Wang (2000) provide a formulation, as well as some specific examples of the predictions, of a double-porosity dual-permeability model for wave propagation in heterogeneous poroelastic media. The analysis is fairly tedious and we do not have space to present details here. The main conclusion of the double-porosity analysis is that the presence of the two porosities and permeabilities leads to new modes of propagation. In particular, bulk compressional and shear waves very similar to those in Biot’s single-porosity formulation are found, and now there are also two slow compressional waves. As the choices of parameters are varied, there are many types of interactions among these waves that are possible, but — in the simplest cases — each of the two slow waves acts individually like the one described here for single-porosity poroelasticity in the preceding section.

Two Slow Waves

We assume that the two permeabilities in the double-porosity model differ greatly in magnitude so that $\kappa_1 >> \kappa_2$ and that the corresponding porosities satisfy $0 < \phi_1 << \phi_2$. Thus, the first porosity type is transport-like and the second is storage-like. The analysis of the preceding section of the present paper would suggest that the smaller of the two permeabilities would
Figure 2: The elements of a double porosity model are: porous rock matrix intersected by fractures. Three types of macroscopic pressure are pertinent in such a model: external confining pressure $p_c$, internal pressure of the matrix pore fluid $p_f^{(1)}$, and internal pressure of the fracture pore fluid $p_f^{(2)}$.

result in a diffusive mode at all frequencies and the larger of the two would result in a propagating slow wave at high frequencies while then degenerating into another diffusive mode at low frequencies. This behavior is exactly what was found in the numerical examples presented by Berryman and Wang (2000).

**Shear and Compressional Waves**

Shear waves were not studied explicitly by Berryman and Wang (2000), but Eq. (5) of that paper can be used for that purpose simply by applying the *curl* operator to all three of the equations in the set. When this is done, the result is that the first equation describes the actual shear mode, while the other two equations provide constraints on the relative motion of the
pore fluid in each type of porosity versus the displacement of the solid frame. In particular, the shear components of the differences in fluid and solid displacements can be uniquely related by complex factors (that are known explicitly) to the displacement of the solid alone. Furthermore, as in the case for single-porosity poroelasticity, all of the interesting behavior of the shear mode — at least for isotropic media — comes from the inertial terms. The form of the resulting dispersion relation at low frequencies is identical to (20) with the replacement

\[ \kappa \rightarrow \kappa_1 + \kappa_2 \approx \kappa_1, \]  

(27)

since we assume here that \( \kappa_1 >> \kappa_2 \). A similar result follows for the compressional wave. Thus, as for single-porosity, the attenuation of the shear and compressional waves is dominated by the largest permeability present in the system. However, this leads to no contradiction in the double-porosity formulation. Thus, the problem inherent in up-scaling with single-porosity poroelasticity is resolved in an intellectually satisfying way in the double-porosity approach.

CONCLUSIONS

It is well-known that fluid flow in porous media is well-described at the microscale by Navier-Stokes’ equations for fluids in the pores but at the macroscale the behavior instead obeys Darcy’s equation. Rigorous methods for establishing the form of such equations for macroscale behavior include multiscale homogenization methods and also the volume averaging method. In particular, it has been shown that Biot’s equations of single-porosity poroelasticity follow in a scale-up of the microscale equations of elasticity coupled to Navier-Stokes (Burridge and Keller, 1981).

We have found that the equations of single-porosity poroelasticity are not the correct equations at the macroscale when there is significant heterogeneity in fluid permeability. However, the double-porosity dual-permeability approach appears to permit consistent modeling of such reservoirs and also shows that no further up-scaling is required beyond the double-porosity stage in many circumstances. Recent extensions of these ideas by Pride and Berryman (2003a,b) and Pride et al. (2003) confirm these conclusions.

REFERENCES


Seismic attenuation due to wave-induced flow

Steven R. Pride, James G. Berryman, and Jerry M. Harris

ABSTRACT
Analytical expressions for three P-wave attenuation mechanisms in sedimentary rocks are given a unified theoretical framework. Two of the models concern wave-induced flow due to heterogeneity in the elastic moduli at "mesoscopic" scales (scales greater than grain sizes but smaller than wavelengths). In the first model, the heterogeneity is due to lithological variations (e.g., mixtures of sands and shales) with a single fluid saturating all the pores. In the second model, a single uniform lithology is saturated in mesoscopic "patches" by two immiscible fluids (e.g., air and water). In the third model, the heterogeneity is at "microscopic" grain scales (broken grain contacts and/or micro-cracks in the grains) and the associated fluid response corresponds to "squirt flow." The model of squirt flow derived here reduces to proper limits as any of the fluid bulk modulus, crack porosity, and/or frequency is reduced to zero. It is shown that squirt flow is incapable of explaining the measured level of loss ($10^{-2} < Q^{-1} < 10^{-1}$) within the seismic band of frequencies (1 to $10^4$ Hz); however, either of the two mesoscopic scale models easily produce enough attenuation to explain the field data.

INTRODUCTION
Intrinsic seismic attenuation is often quantified by the inverse quality factor $Q^{-1}$ of sedimentary rock within the seismic band of frequencies, which we loosely define as 1 to $10^4$ Hz. For transmission experiments (earthquake recordings, VSP, cross-well tomography, sonic logs), the total measured attenuation can be decomposed as $Q_{\text{total}}^{-1} = Q_{\text{scat}}^{-1} + Q^{-1}$ where both contributions to the total attenuation are necessarily positive. The inverse quality factor $Q^{-1}$ for the intrinsic attenuation represents the fraction of wave energy irreversibly lost to heat during a wave period as normalized by the strain energy. For reflection seismic prospecting, there are other wave energy losses due to reflection and transmission effects at interfaces. These effects are neglected here for reflection seismic by exclusion. For the types of transmission experiments that we do consider here, we justify neglecting them since we are basically studying the losses within any single block of material, and not treating effects at interfaces between blocks since these can be handled independently with well-known methods.

Crosswell experiments in horizontally-stratified sediments produce negligible amounts of scattering loss so that essentially all loss is attributable to intrinsic attenuation. Quan and Harris (1997) use tomography to invert for the amplitudes of crosswell P-wave first arrivals to obtain

---

1email: SRPride@lbl.gov, berryman@pangea.stanford.edu, harris@pangea.stanford.edu
the $Q^{-1}$ for the layers of a stratified sequence of shaly sandstones and limestones (depths ranging from 500–900 m). The center frequency of their measurements is roughly 1750 Hz and they find that $10^{-2} < Q^{-1} < 10^{-1}$ for all the layers in the sequence. Sams et al. (1997) also measure the intrinsic loss in a stratified sequence of water-saturated sandstones, siltstones and limestones (depths ranging from 50–250 m) using VSP (30–280 Hz), crosswell (200–2300 Hz), sonic logs (8–24 kHz), and ultrasonic laboratory (500–900 kHz) measurements. Sams et al. (1997) calculate (with some inevitable uncertainty) that in the VSP experiments, $Q_{1}^{-1} / Q_{\text{scat}}^{-1} \approx 4$, while in the sonic experiments, $Q_{1}^{-1} / Q_{\text{scat}}^{-1} \approx 19$; i.e., for this sequence of sediments, the intrinsic loss dominates the scattering loss at all frequencies. Sams et al. (1997) also find $10^{-2} < Q^{-1} < 10^{-1}$ across the seismic band.

We demonstrate here that wave-induced fluid flow generates enough heat to explain these measured levels of intrinsic attenuation. Other attenuation mechanisms need not be considered, although they may in fact be present but contribute much smaller fractions of the overall observed attenuation. The induced flow occurs at many different spatial scales that can broadly be categorized as “macroscopic”, “mesoscopic”, and “microscopic.”

The macroscopic flow is the wavelength-scale equilibration occurring between the peaks and troughs of a P-wave. This mechanism was first treated by Biot (1956a) and is often simply called “Biot loss.” However, the flow at such macro scales drastically underestimates the measured loss in the seismic band (by as much as 5 orders of magnitude). Mavko and Nur (1979) therefore proposed a microscopic mechanism due to microcracks in the grains and/or broken grain contacts. When a seismic wave squeezes a rock having such grain-scale damage, the cracks respond with a greater fluid pressure than the main porespace resulting in a flow from crack to pore that Mavko and Nur (1979) named “squirt flow.” Dvorkin et al. (1995) have given a squirt-flow model applicable to liquid-saturated rocks. Although squirt flow seems entirely capable of explaining much of the measured attenuation in the laboratory at ultrasonic frequencies and may also turn out to be important for propagation in ocean sediments at ultrasonic frequencies (Williams et al., 2002) as well, we show here that this mechanism cannot explain the attenuation in the seismic band.

Thus, a third mechanism based on mesoscopic-scale heterogeneity seems required to explain seismic attenuation. Mesoscopic length scales are those larger than grain sizes but smaller than wavelengths. Heterogeneity across these scales may be due to lithological variations or to patches of different immiscible fluids. When a compressional wave squeezes a material containing mesoscopic heterogeneity, the effect is similar to squirt with the more compliant portions of the material responding with a greater fluid pressure than the stiffer portions. There is a subsequent flow of fluid capable of generating significant amounts of loss in the seismic band.

Prior models of such mesoscopic loss have focused on flow between the layers of a stratified material due to P-waves propagating normal to the layering [e.g., White et al. (1975), Norris (1993), Gurevich and Lopatnikov (1995), and Gelinksy and Shapiro (1997)]. The present study seeks to model the flow for arbitrary mesoscopic geometry, albeit under the restriction that only two porous phases are mixed together in each averaging volume.

In the next section, we review a recent model of Pride and Berryman (2003a,b) treating
the mesoscopic loss created by lithological patches having, for example, different degrees of consolidation. This so-called “double-porosity” model provides the theoretical framework that will be used throughout. Then, we derive a new patchy-saturation variant of the model and, in following section, a new squirt-flow variant. The results are then discussed in the concluding section. The main point of the paper is to derive models for patchy-saturation and squirt using the same notation and approach as the double-porosity theory; in so doing, we aim to draw conclusions about the nature of attenuation in the seismic band of frequencies.

**REVIEW OF THE DOUBLE-POROSITY THEORY**

In this theory, the mesoscopic heterogeneity is modeled as being a mixture of two porous phases saturated by a single fluid. Porous phase 1 is defined as being the stiffer lower-permeability phase and phase 2 the more compressible higher-permeability phase.

**Local Governing Equations**

Each porous phase is locally modeled as a porous continuum and obeys the laws of poroelasticity [e.g., Biot (1962)]

\[ \nabla \cdot \tau_i^D - \nabla p_{ci} = \rho \ddot{u}_i + \rho_f \dot{Q}_i \]  
\[ Q_i = \frac{k_i}{\eta} (\nabla p_{fi} + \rho_f \dot{u}_i) \]  
\[ \left[ \nabla \cdot \dot{u}_i \right] = \frac{1}{K_i^d} \left[ \frac{1}{-\alpha_i / B_i} \right] \left[ \ddot{p}_{ci} \right] \]  
\[ \tau_i^D = G_i \left( \nabla u_i + \nabla u_i^T - \frac{2}{3} \nabla \cdot u_i I \right) \]

where the index \( i \) represents the two phases \((i = 1, 2)\). The response fields in these equations are themselves local volume averages taken over a scale larger than the grain sizes but smaller than the mesoscopic extent of either phase. The local fields are: \( u_i \), the average displacement of the framework of grains; \( Q_i \), the Darcy filtration velocity; \( p_{fi} \), the fluid pressure; \( p_{ci} \), the confining pressure (total average pressure); and \( \tau_i^D \), the deviatoric (or shear) stress tensor. In the linear theory of interest here, the overdots on these fields denote a partial time derivative. In the local Darcy law (2), \( \eta \) is the fluid viscosity and the permeability \( k_i \) is a linear time-convolution operator whose Fourier transform \( k_i(\omega) \) is called the “dynamic permeability” and can be modeled using the theory of Johnson *et al.* (1987) (see the Appendix).

In the local compressibility law (3), \( K_i^d \) is the drained bulk modulus of phase \( i \) (confining pressure change divided by sample dilatation under conditions where the fluid pressure does not change), \( B_i \) is Skempton’s (1954) coefficient of phase \( i \) (fluid pressure change divided by confining pressure change for a sealed sample), and \( \alpha_i \) is the Biot and Willis (1957) coefficient of phase \( i \) defined as

\[ \alpha_i = (1 - K_i^d / K_i^u) / B_i, \]
where $K_i^u$ is the undrained bulk modulus (confining pressure change divided by sample dilatation for a sealed sample). In the present work, no restrictions to single-mineral isotropic grains will be made. Finally, in the deviatoric constitutive law (4), $G_i$ is the shear modulus of the framework of grains. At the local level, all these poroelastic constants are taken to be real constants. In the appendix we give the Gassmann (1951) uid-substitution relations that allow $B_i$ and $\alpha_i$ to be expressed in terms of the porosity $\phi_i$, the fluid and solid bulk moduli $K_f$ and $K_s$, and the drained modulus $K_i^d$.

**Double-Porosity Governing Equations**

In the double-porosity model, the goal is to determine the average uid response in each of the porous phases in addition to the average displacement of the solid grains (Berryman and Wang, 1995; 2000). The averages are taken over regions large enough to significantly represent both porous phases, but smaller than wavelengths. Assuming an $e^{-i\omega t}$ time dependence, Pride and Berryman (2003a) have found the volume averaged local laws (1)–(4) in order to obtain the macroscopic “double-porosity” governing equations in the form

$$\nabla \cdot \tau^D - \nabla P_c = -i\omega(\rho v + \rho_f q_1 + \rho_f q_2),$$

$$\left[ \begin{array}{c} q_1 \\ q_2 \end{array} \right] = -\frac{1}{\eta} \left[ \begin{array}{cc} \kappa_{11} & \kappa_{12} \\ \kappa_{12} & \kappa_{22} \end{array} \right] \left[ \begin{array}{c} \nabla \bar{P}_{f1} - i\omega\rho_f v \\ \nabla \bar{P}_{f2} - i\omega\rho_f v \end{array} \right],$$

$$\left[ \begin{array}{c} \nabla \cdot v \\ \nabla \cdot q_1 \\ \nabla \cdot q_2 \end{array} \right] = i\omega \left[ \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{array} \right] \left[ \begin{array}{c} P_c \\ \bar{P}_{f1} \\ \bar{P}_{f2} \end{array} \right] + i\omega \left[ \begin{array}{c} 0 \\ \zeta_{\text{int}} \\ -\zeta_{\text{int}} \end{array} \right],$$

$$-i\omega\zeta_{\text{int}} = \gamma(\omega)(\bar{P}_{f1} - \bar{P}_{f2}),$$

$$-i\omega\bar{P}_{f}^D = [G(\omega) - i\omega g(\omega)] \left[ \nabla v + (\nabla v)^T - \frac{2}{3}\nabla \cdot v I \right].$$

The macroscopic fields are: $v$, the average particle velocity of the solid grains throughout an averaging volume of the composite; $q_i$, the average Darcy flux across phase $i$; $P_c$, the average total pressure in the averaging volume; $\tau^D$, the average deviatoric stress tensor; $\bar{P}_{fi}$, the average fluid pressure within phase $i$; and $-i\omega\zeta_{\text{int}}$, the average rate at which uid volume is being transferred from phase 1 into phase 2 as normalized by the total volume of the averaging region. The dimensionless increment $\zeta_{\text{int}}$ represents the “mesoscopic ow.”

Equation (7) is the generalized Darcy law allowing for uid cross-coupling between the phases [c.f., Pride and Berryman (2003b)], equation (8) is the generalized compressibility law where $\nabla \cdot q_i$ corresponds to fluid that has been depleted from phase $i$ due to transfer across the external surface of an averaging volume, and equation (9) is the transport law for internal mesoscopic flow (fluid transfer between the two porous phases).

The coefficients in these equations have been modeled by Pride and Berryman (2003a,b). Before presenting these results, the nature of the waves implicitly contained in these laws is briefly commented upon. If plane-wave solutions for $v$, $q_1$ and $q_2$ are introduced, there is found to be a single transverse wave, and three longitudinal responses: a fast wave and
two slow waves (Berryman and Wang, 2000). The fast wave is the usual P-wave identified on seismograms, while the two slow waves correspond to fluid-pressure diffusion in phases 1 and 2. The only problem with analyzing the fast compressional wave in this manner is that the characteristic equation for the longitudinal slowness $s$ is cubic in $s^2$ and, therefore, analytically inconvenient.

**Reduction to an Effective Biot Theory**

The approach that we take instead is to first reduce these double-porosity laws (6)–(10) to an effective single-porosity Biot theory having complex frequency-dependent coefficients. The easiest way to do this is to assume that phase 2 is entirely embedded in phase 1 so that the average flux $q_2$ into and out of the averaging volume across the external surface of phase 2 is zero. By placing $\nabla \cdot q_2 = 0$ into the compressibility laws (8), the fluid pressure $p_{f2}$ can be entirely eliminated from the theory. In this case the double-porosity laws reduce to effective single-porosity poroelasticity governed by laws of the form (3) but with effective poroelastic moduli given by

$$\frac{1}{K_D} = a_{11} - \frac{a_{13}^2}{a_{33} - \gamma/i \omega},$$  \hfill (11)

$$B = \frac{-a_{12}(a_{33} - \gamma/i \omega) + a_{13}(a_{23} + \gamma/i \omega)}{(a_{22} - \gamma/i \omega)(a_{33} - \gamma/i \omega) - (a_{23} + \gamma/i \omega)^2},$$  \hfill (12)

$$\frac{1}{K_U} = \frac{1}{K_D} + B \left( a_{12} - \frac{a_{13}(a_{23} + \gamma/i \omega)}{a_{33} - \gamma/i \omega} \right).$$  \hfill (13)

Here, $K_D(\omega)$ is the effective drained bulk modulus of the double-porosity composite, $B(\omega)$ is the effective Skempton’s coefficient, and $K_U(\omega)$ is the effective undrained bulk modulus. An effective Biot-Willis constant can then be defined using $\alpha(\omega) = [1 - K_D(\omega)/K_U(\omega)]/B(\omega)$.

The complex frequency dependent “drained” modulus $K_D$ again defines the total volumetric response when the average fluid pressure throughout the entire composite is unchanged; however, the local fluid pressure in each phase may be non-uniform even though the average is zero resulting in mesoscopic flow and in $K_D$ being complex and frequency dependent. Similar interpretations hold for the undrained moduli $K_U$ and $B$. An undrained response is when no fluid can escape or enter through the external surface of an averaging volume; however, there can be considerable internal exchange of fluid between the two phases resulting in the complex frequency-dependent nature of both $K_U$ and $B$.

**Double-Porosity $a_{ij}$ Coefficients**

The constants $a_{ij}$ are all real and correspond to the high-frequency response for which no internal fluid-pressure relaxation can take place. They are given exactly as (Pride and Berryman, 2003a)

$$a_{11} = 1/K$$  \hfill (14)
\[ a_{22} = \frac{v_1 \alpha_1}{K_1^d} \left( \frac{1}{B_1} - \frac{\alpha_1(1 - Q_1)}{1 - K_1^d/K_2^d} \right) \]  
\[ a_{33} = \frac{v_2 \alpha_2}{K_2^d} \left( \frac{1}{B_2} - \frac{\alpha_2(1 - Q_2)}{1 - K_2^d/K_1^d} \right) \]  
\[ a_{12} = -v_1 Q_1 \alpha_1/K_1^d \]  
\[ a_{13} = -v_2 Q_2 \alpha_2/K_2^d \]  
\[ a_{23} = -\frac{\alpha_1 \alpha_2 K_1^d/K_2^d}{(1 - K_1^d/K_2^d)^2} \left( \frac{1}{K} - \frac{v_1}{K_1^d} - \frac{v_2}{K_2^d} \right) \]  

where the \( Q_i \) are auxiliary constants given by

\[ v_1 Q_1 = \frac{1 - K_2^d/K_1^d}{1 - K_2^d/K_1^d} \quad \text{and} \quad v_2 Q_2 = \frac{1 - K_2^d/K_1^d}{1 - K_1^d/K_2^d}. \]  

Here, \( v_1 \) and \( v_2 \) are the volume fractions of each phase within an averaging volume of the composite. The one constant that has not yet been defined is the overall drained modulus \( K = 1/a_{11} \) of the two-phase composite (the modulus defined in the quasi-static limit where the local fluid pressure throughout the composite is everywhere unchanged). It is through \( K \) that the \( a_{ij} \) potentially depend on the mesoscopic geometry of the two porous phases. However, a reasonable modeling choice when phase 2 is embedded within phase 1 is to simply take the geometry-independent harmonic mean \( 1/K = v_1/K_1^d + v_2/K_2^d \). Although this choice actually violates the Hashin-Shtrikman bounds (Hashin and Shtrikman, 1961) for truly isotropic media, it is nevertheless a reasonable choice for earth systems where the assumed isotropy is itself an approximation. This choice is also a particularly convenient one because it results in \( Q_1 = Q_2 = 1 \) as well as \( a_{23} = 0 \). All dependence on the fluid’s bulk modulus is contained within the two Skempton’s coefficients \( B_1 \) and \( B_2 \) and is thus restricted to \( a_{22} \) and \( a_{33} \). In the quasi-static limit \( \omega \to 0 \) (fluid pressure everywhere uniform throughout the composite), equations (12) and (13) reduce to the known exact results of Berryman and Milton (1991) once equations (14)–(19) are employed.

**Double-Porosity Transport**

Pride and Berryman (2003b) obtain the internal transport coefficient \( \gamma \) of equation (9) as

\[ \gamma(\omega) = \gamma_m \sqrt{1 - i \omega/\omega_m} \]  

where the parameter \( \gamma_m \) that holds in the final stages of internal fluid-pressure equilibration is given by

\[ \gamma_m = \frac{v_1 k_1}{\eta L_1^2} \left[ 1 + O(k_1/k_2) \right]. \]  

Since the more compressible embedded phase 2 typically has a permeability much greater than the host phase 1, the \( O(k_1/k_2) \) correction can be neglected. The transition frequency
Fluid-induced attenuation

\( \omega_m \) corresponds to the onset of a high-frequency regime in which the fluid-pressure-diffusion penetration distance between the phases becomes small relative to the scale of the mesoscopic heterogeneity. It is given by

\[
\omega_m = \frac{B_1K_1^d k_1(v_1V/S)^2}{\eta \alpha_1 L_1^4} \left( 1 + \sqrt{\frac{k_1B_2K_2^d \alpha_1}{k_2B_1K_1^d \alpha_2}} \right)^2. \tag{23}
\]

The length \( L_1 \) characterizes the average distance in phase 1 over which the fluid pressure gradient still exists in the final approach to equilibration and has the formal mathematical definition

\[
L_1^2 = \frac{1}{V_1} \int_{\Omega_1} \Phi_1 \, dV = \frac{1}{V_1} \int_{\Omega_1} \nabla \Phi_1 \cdot \nabla \Phi_1 \, dV \tag{24}
\]

where \( \Omega_1 \) is the region of an averaging volume occupied by phase 1 and having a volume measure \( V_1 \). The potential \( \Phi_1 \) has units of length squared and is a solution of an elliptic boundary-value problem that under conditions where the harmonic mean is a good approximation for the overall drained modulus and where the permeability ratio \( k_1/k_2 \) can be considered small, reduces to

\[
\nabla^2 \Phi_1 = -1 \text{ in } \Omega_1, \tag{25}
\]

\[
n \cdot \nabla \Phi_1 = 0 \text{ on } \partial E_1, \tag{26}
\]

\[
\Phi_1 = 0 \text{ on } \partial \Omega_{12}, \tag{27}
\]

where \( \partial E_1 \) is the external surface of the averaging volume coincident with phase 1, and where \( \partial \Omega_{12} \) is the internal interface separating phases 1 and 2. Multiplying equation (25) by \( \Phi_1 \) and integrating over \( \Omega_1 \), establishes that second integral of equation (24).

For complicated geometry, \( L_1 \) can only be determined numerically. For idealized geometries it can be analytically estimated. For example, if phase 2 is taken to be small spheres of radius \( a \) embedded within each sphere \( R \) of the composite, Pride and Berryman (2003b) obtain

\[
L_1^2 = \frac{9}{14} \frac{R^2}{a} \left[ 1 - \frac{7}{6} \frac{a}{R} + O(\frac{a^3}{R^3}) \right]. \tag{28}
\]

The volume fraction \( v_2 \) of small spheres is then \( v_2 = (a/R)^{1/3} \) which can be used to eliminate \( R \) since \( R = av_2^{-1/3} \). The other length parameter is the volume-to-surface ratio \( V/S \) where \( S \) is the area of \( \partial \Omega_{12} \) in each volume \( V \) of composite. For the simple spherical-inclusion model, it is given by \( V/S = R^3/(3a^2) = av_2/3 \).

The coefficient \( G(\omega) - i\omega g(\omega) \) governing shear generally has a non-zero “viscosity” \( g(\omega) \) associated with the mesoscopic fluid transport between the compressional lobes surrounding a sheared phase 2 inclusion. Both of the frequency functions \( G(\omega) \) and \( -\omega g(\omega) \) are real and are Hilbert transforms of each other. The frequency dependence of \( g(\omega) \) was not modeled by Pride and Berryman (2003b). However, if the inclusions of phase 2 are taken to be spheres, then \( g(\omega) = 0 \) exactly and \( G(\omega) = G \) is a constant that can be approximately modeled using a simple harmonic average \( 1/G = v_1/G_1 + v_2/G_2 \) of the underlying shear moduli of each phase.
Finally, the dynamic permeability $k(\omega)$ to be used in the effective Biot theory can be modeled in several ways. Perhaps the simplest modeling choice when phase 2 is modeled as small inclusions embedded in phase 1 is to again take a harmonic average $1/k(\omega) = v_1/k_1(\omega) + v_2/k_2(\omega) \approx v_1/k_1(\omega)[1 + O(v_2k_1/k_3)]$.

**Phase Velocity and Attenuation**

With all of the double-porosity coefficients now defined, the compressional phase velocity and attenuation may be determined by inserting a plane-wave solution into the effective single-porosity Biot equations [in the form (1)–(4)]. This gives the standard complex longitudinal slowness $s$ of Biot theory

$$s^2 = b \mp \sqrt{b^2 - \frac{\rho \hat{\theta} - \rho_f^2}{MH - C^2}}$$

where

$$b = \frac{\rho M + \hat{\rho} H - 2\rho_f C}{2(MH - C^2)}$$

is simply an auxiliary parameter, and where $H$, $C$ and $M$ are the Biot (1962) poroelastic moduli defined in terms of the complex frequency-dependent parameters of equations (11)–(13) as

$$H = K_U + 4G/3$$
$$C = B K_U$$
$$M = \frac{B^2}{1 - K_D/K_U}$$

The complex inertia $\hat{\rho}$ corresponds to rewriting the relative flow resistance as an effective inertial effect

$$\hat{\rho} = -\eta/[i\omega k(\omega)]$$

Taking the minus sign in equation (29) gives an $s$ having an imaginary part much smaller than the real part and that thus corresponds to the normal P-wave. Taking the positive sign gives an $s$ with real and imaginary parts of roughly the same amplitude and that thus corresponds to the slow P-wave (a pure fluid-pressure diffusion across the seismic band of frequencies). We are only interested here in the properties of the normal P-wave.

The P-wave phase velocity $v_p$ and the attenuation measure $Q_p^{-1}$ are related to the complex slowness $s$ as

$$v_p = 1/\text{Re}\{s\}$$
$$Q_p^{-1} = \text{Im}\{s\}/\text{Re}\{s\}.$$

**A Numerical Example**

In Fig. 1, we give an example of $Q_p^{-1}$ and $v_p$ as determined using this double-porosity theory. The example models a consolidated sandstone containing pockets (small regions) where the
Figure 1: The attenuation and phase velocity of compressional waves in the double-porosity model of Pride and Berryman (2003a). The 5 cm embedded spheres of phase 2 have frame moduli ($K_2^d$ and $G_2$) modeled using the modified Walton theory given in the appendix in which both $K_2^d$ and $G_2$ vary strongly with the background effective pressure $P_e$ (or overburden thickness). These spheres of porous continuum 2 were embedded into a phase 1 continuum modeled as a consolidated sandstone. [Jim3-QandVdp] [NR]
grains are not cemented together. The embedded unconsolidated phase 2 is modeled as 5 cm radius spheres occupying 1.5% of the composite. The frame moduli of this relatively-compressible embedded material are determined using the modified Walton theory given in the Appendix. These moduli are functions of the background effective-stress level \( P_e \). The host phase 1 is modeled as a consolidated sandstone (using \( \phi_1 = 0.15 \) and \( c = 4 \) in the model given in the appendix). The permeability of the two phases are taken as \( k_1 = 10^{-14} \text{ m}^2 \) and \( k_2 = 10^{-12} \text{ m}^2 \). The invariant peak near \( 10^5 \text{ Hz} \) is that due to the Biot loss (fluid equilibration at the scale of the seismic wavelength) while the principal peak that changes with the effective pressure \( P_e \) is that due to mesoscopic-scale equilibration. Figure 1 demonstrates that small amounts of a relatively soft material embedded within a more consolidated rock is capable of producing the level of attenuation measured in field experiments.

The overall magnitude of attenuation in the model is controlled principally by the contrast of compressibilities between the two porous phases; the greater the contrast, the greater the mesoscopic fluid-pressure gradient and the greater the mesoscopic-flow intensity and associated attenuation. The relaxation frequency at which the mesoscopic loss per cycle is maximum is proportional to \( k_1/L_1^2 \). Below this relaxation frequency, \( Q^{-1} \) increases with frequency as \( f \eta/k_1 \). Thus, the permeability information in the double-porosity attenuation is principally in the frequency dependence of \( Q^{-1} \) and not in the overall magnitude of \( Q^{-1} \) and involves principally the permeability \( k_1 \) of the host phase and not the overall permeability of the composite. [See Berryman (1988) for a related discussion.] If phase 2 is well modeled as being small inclusions embedded in phase 1, then \( k_1 \) is controlling the overall permeability. If phase 2 corresponds to through going connected joints, then although \( Q^{-1}(\omega) \) contains information about \( k_1 \), it does not contain information about the overall permeability which is being dominated by \( k_2 \) in this case.

**PATCHY-SATURATION MODEL**

Another important source of mesoscopic-scale heterogeneity is patchy fluid saturation. All natural hydrological processes by which one fluid non-miscibly invades a region initially occupied by another result in a patchy distribution of the two fluids. The patch sizes are distributed across the entire range of mesoscopic length scales and for many invasion scenarios are expected to be fractal. As a compressional wave squeezes such a material, the patches occupied by the less-compressible fluid will respond with a greater fluid-pressure change than the patches occupied by the more-compressible fluid. The two fluids will then equilibrate by the same type of mesoscopic flow already modeled in the double-porosity model.

An analysis almost identical to that of Pride and Berryman (2003a,b) can be carried out that leads to the same effective poroelastic moduli given by equations (11)–(13) but with different definitions of the \( a_{ij} \) constants and internal transport coefficient \( \gamma(\omega) \). In the model, a single uniform porous frame is saturated by mesoscopic-scale patches of fluid 1 and fluid 2. We define porous phase 1 to be those regions (patches) occupied by the less mobile fluid and phase 2 the patches saturated by the more mobile fluid; i.e., by definition \( \eta_1 > \eta_2 \). This most often (but not necessarily) corresponds to \( K_{f1} > K_{f2} \) and, therefore, to \( B_1 > B_2 \).
Johnson (2001) approached this problem using a different coarse-graining argument while starting from the same local physics (assuming, however, that the porous medium is a Gassmann mono-mineral material). The final undrained bulk modulus obtained by Johnson (2001) is identical to our model in the limits of high and low frequency and differs only negligibly in the transition range of frequencies where the flow in either model is not explicitly treated.

**Patchy-Saturation $a_{ij}$ Coefficients**

To obtain the $a_{ij}$ for the patchy-saturation model, we note that each patch has the same $\alpha$ and $K$. The poroelastic differences between patches is entirely due to $B_1$ being different than $B_2$. Upon volume averaging equation (3) and using $\nabla \cdot \mathbf{v} = \nabla \cdot (v_1 \mathbf{u}_1) + \nabla \cdot (v_2 \mathbf{u}_2)$, where an overline again denotes a volume average over the appropriate phase, and using the fact that the $a_{ij}$ are defined in the extreme high-frequency limit where the fluids have no time to traverse the internal interface $\partial \Omega_{12}$ (i.e., the $a_{ij}$ are defined under the condition that $\dot{\zeta}_{\text{int}} = 0$), one has

\[
\nabla \cdot \mathbf{v} = -\frac{v_1}{K} \overline{p}_{c1} - \frac{v_2}{K} \overline{p}_{c2} + \frac{v_1 \alpha}{K} \overline{p}_{f1} + \frac{v_2 \alpha}{K} \overline{p}_{f2},
\]

\[
\nabla \cdot \mathbf{q}_1 = \frac{v_1 \alpha}{K} \overline{p}_{c1} - \frac{v_1 \alpha}{KB_1} \overline{p}_{f1},
\]

\[
\nabla \cdot \mathbf{q}_2 = \frac{v_2 \alpha}{K} \overline{p}_{c2} - \frac{v_2 \alpha}{KB_2} \overline{p}_{f2}.
\]

The average confining pressures $\overline{p}_{ci}$ in each phase are not a priori known; however, they are necessarily linear functions of the three independent applied pressures of the theory $P_c (= v_1 \overline{p}_{c1} + v_2 \overline{p}_{c2}), \overline{p}_{f1}$, and $\overline{p}_{f2}$. It is straightforward to demonstrate that if and only if the average confining pressures take the form

\[
v_1 \overline{p}_{c1} = v_1 \overline{p}_c + \beta \overline{p}_{f1} - \beta \overline{p}_{f2} \tag{40}
\]

\[
v_2 \overline{p}_{c2} = v_2 \overline{p}_c - \beta \overline{p}_{f1} + \beta \overline{p}_{f2} \tag{41}
\]

will equations (37)–(39) produce $a_{ij}$ that satisfy the thermodynamic symmetry requirement of $a_{ij} = a_{ji}$ [i.e., these $a_{ij}$ constants are all second derivatives of a strain-energy function as demonstrated by Pride and Berryman (2003a)]. Upon placing equations (40) and (41) into equations (37)–(39), we then have

\[
a_{11} = \frac{1}{K} \tag{42}
\]

\[
a_{22} = (-\beta + v_1/B_1)\alpha/K \tag{43}
\]

\[
a_{33} = (-\beta + v_2/B_2)\alpha/K \tag{44}
\]

\[
a_{12} = -v_1\alpha/K \tag{45}
\]

\[
a_{13} = -v_2\alpha/K \tag{46}
\]

\[
a_{23} = \beta\alpha/K \tag{47}
\]

where $\beta$ is a constant to be determined.
To obtain \( \beta \), we note that in the high-frequency limit, each local patch of phase \( i \) is undrained and thus characterized by an undrained bulk modulus \( K_i^u = K/(1 - \alpha B_i) \) and a shear modulus \( G \) that is the same for all patches. In this limit, the usual laws of elasticity govern the response of this heterogeneous composite. Under these precise conditions (elasticity of an isotropic composite having uniform \( G \) and all heterogeneity confined to the bulk modulus which in the present case corresponds to \( K_i^u \)), the theorem of Hill (1963) applies, which states that the overall undrained-unrelaxed modulus of the composite \( K_H \) is given exactly by

\[
\frac{1}{K_H + 4G/3} = \frac{v_1}{K_1^u + 4G/3} + \frac{v_2}{K_2^u + 4G/3}.
\]

(48)

In terms of the \( a_{ij} \), this same undrained-unrelaxed Hill modulus is given by

\[
\frac{1}{K_H} = a_{11} + a_{12} \left( \frac{\delta p_{f1}}{\delta P_c} \right)_U + a_{13} \left( \frac{\delta p_{f2}}{\delta P_c} \right)_U
\]

(49)

where, upon using \( \nabla \cdot \boldsymbol{q}_i = 0 \) and \( \dot{\xi}_{\text{int}} = 0 \) in equation (8) and then using (42)–(47), the undrained-unrelaxed pressure ratios are

\[
\left( \frac{\delta p_{f1}}{\delta P_c} \right)_U = \frac{\beta - v_1 v_2/B_2}{\beta(v_1/B_1 + v_2/B_2) - v_1 v_2/(B_1 B_2)}
\]

(50)

\[
\left( \frac{\delta p_{f2}}{\delta P_c} \right)_U = \frac{\beta - v_1 v_2/B_1}{\beta(v_1/B_1 + v_2/B_2) - v_1 v_2/(B_1 B_2)}.
\]

(51)

Thus, after some algebra, equation (49) yields the exact result

\[
\beta = v_1 v_2 \left( \frac{v_1}{B_2} + \frac{v_2}{B_1} \right) \left[ \frac{\alpha - (1 - K/K_H)(v_1/B_1 + v_2/B_2)}{\alpha - (1 - K/K_H)(v_1/B_1 + v_2/B_2)} \right]
\]

(52)

with \( K_H \) given by equation (48). All the \( a_{ij} \) are now expressed in terms of known information.

**Patchy-Saturation Transport**

We next must address the internal fluid-pressure equilibration between the two phases with the goal of obtaining the internal transfer coefficient \( \gamma \) of equation (9). The mathematical definition of the rate of internal fluid transfer is

\[
\dot{\xi}_{\text{int}} = \frac{1}{V} \int_{\partial \Omega_{12}} \boldsymbol{n} \cdot \boldsymbol{Q}_1 \, dS
\]

(53)

where \( V \) is the volume occupied by the composite. A possible concern in the patchy-saturation analysis is whether capillary effects at the local interface \( \partial \Omega_{12} \) separating the two phases need to be allowed for.
Local continuity conditions on $\partial \Omega_{12}$

At the pore scale, the interface separating one fluid patch from the next is a series of menisci. Roughness on the grain surfaces keeps the contact lines of these menisci pinned to the grain surfaces. Pride and Flekkoy (1999) argue that the contact lines of an air-water meniscus will remain pinned for fluid-pressure changes less than roughly $10^4$ Pa which correspond to the pressure range induced by linear seismic waves. So as a wave passes, the menisci will bulge and change shape but will not migrate away. This makes the problem vastly more simple to analyze theoretically.

One porous-continuum boundary condition is that all fluid volume that locally enters the interface $\partial \Omega_{12}$ from one side, must exit the other side so that $\mathbf{n} \cdot \mathbf{Q}_1 = \mathbf{n} \cdot \mathbf{Q}_2$ ($= \mathbf{n} \cdot \mathbf{Q}$). Another boundary condition is that the difference in the rate at which energy is entering and leaving the interface is entirely due to the work performed in changing the miniscus surface area. Before the wave arrives, each miniscus has an initial mean curvature $H_o$ fixed by the static fluid pressures initially present; $p_{f1}^0 - p_{f2}^0 = \sigma H_o$ where $\sigma$ is the surface tension. During wave passage, one can demonstrate (Pride and Flekkoy, 1999) that the mean curvature changes as $H = H_o + \epsilon H_1 + O(\epsilon^2)$ where $H_1$ is of the same order as $H_o$ and where $\epsilon$ is a dimensionless number called the capillary number. The capillary number is defined $\epsilon = \eta |\mathbf{Q}|/\sigma$ where $|\mathbf{Q}|$ is some estimate of the wave-induced Darcy flux and that is thus bounded as the wave strain times phase velocity; i.e., $|\mathbf{Q}| < 10^{-3}$ m/s. For typical interfaces (like air and water), we have $\sigma > 10^{-2}$ Pa m and $\eta \approx 10^{-3}$ Pa s. Thus, for linear wave problems, $\epsilon \ll 10^{-4}$ and $\epsilon$ can be considered a very small number.

Writing the fluid pressures as $p_{fi} = p_{f1}^0 + \delta p_{fi}$ and using the fact that $\mathbf{n} \cdot \mathbf{Q}$ is continuous, allows the conservation of energy at the interface to be expressed

$$\left[ \mathbf{n} \cdot \{ \mathbf{r}_i \cdot \mathbf{u}_i - (p_{f1}^0 + \delta p_{fi}) \mathbf{Q} \} \right] = \sigma \mathbf{n} \cdot \mathbf{Q} H_o [1 + O(\epsilon)].$$

(54)

The brackets on the left-hand side deonte the jump in energy flux across the interface, while the right-hand side represents the rate at which work is performed in stretching the menisci. Since conservation of momentum requires $\mathbf{n} \cdot \mathbf{r}$ to be continuous at the interface and since the assumption of the grains being welded together [or having an overburden effective pressure $(1 - \phi)(\rho_s - \rho_f)gh$ acting on them that is greater than the wave stress] requires that $\mathbf{u}$ is continuous, we obtain that to leading order in $\epsilon$

$$\delta p_{f1} = \delta p_{f2}$$

(55)

along the interface $\partial \Omega_{12}$. This means that the fluid pressure equilibration can be modeled using the standard displacement-stress continuity conditions along $\partial \Omega_{12}$ that were also employed in the double-porosity analysis; i.e., capillary effects can be neglected. In what follows, the fluid pressures correspond to the changes induced by the wave and so we cease to explicitly write the “$\delta$” in front of them.
Mesoscopic flow equations

To obtain the transport law $-i\omega \xi_{\text{int}} = \gamma(\omega) (\overline{p}_{f1} - \overline{p}_{f2})$, the mesoscopic flow is analyzed in the limits of low and high frequencies. These limits are then connected using a frequency function that respects causality constraints. The linear fluid response inside the patchy composite due to a seismic wave can always be resolved into two portions: (1) a vectorial response due to macroscopic fluid-pressure gradients across an averaging volume that generate a macroscopic Darcy flux $\mathbf{q}_i$ across each phase and that corresponds to the macroscopic conditions $\overline{p}_{f1} = 0$ and $\nabla \overline{p}_{f1} \neq 0$; and (2) a scalar response associated with internal fluid transfer and that corresponds to the macroscopic conditions $\overline{p}_{f1} \neq 0$ and $\nabla \overline{p}_{f1} = 0$. The macroscopic isotropy of the composite guarantees that there is no cross-coupling between the vectorial transport $\mathbf{q}_i$ and the scalar transport $\xi_{\text{int}}$ within each sample.

The mesoscopic flow problem that defines $\xi_{\text{int}}$ is the internal equilibration of fluid pressure between the patches when a confining pressure $\Delta P$ has been applied to a sealed sample of the composite. Having the external surface sealed is equivalent to the required macroscopic constraint that $\nabla \overline{p}_{f1} = 0$. Upon taking the divergence of (2) and using equation (3), the diffusion problem controlling the mesoscopic flow becomes

$$
\frac{k}{\eta_i} \nabla^2 p_{fi} + i\omega \frac{\alpha}{K B_i} p_{fi} = i\omega \frac{\alpha}{K} p_{ci} \quad \text{in} \; \Omega_i, \quad (56)
$$

$$
\begin{bmatrix} p_{fi} \end{bmatrix} \text{ and } \begin{bmatrix} \mathbf{n} \cdot \nabla p_{fi} \end{bmatrix} = 0 \quad \text{on} \; \partial \Omega_{12}, \quad (57)
$$

$$
\mathbf{n} \cdot \nabla p_{fi} = 0 \quad \text{on} \; \partial E_i, \quad (58)
$$

where $\Omega_i$ is the region that each phase occupies within the averaging volume, $\partial E_i$ is that portion of the external surface of the averaging volume that is in contact with phase $i$, and the brackets in equation (57) again denote jumps across the interface. One also needs to insert equations (3) and (4) into (1) to obtain a second-order partial-differential equation for the displacements $\mathbf{u}_i$. In general, the local confining pressures $p_{ci}$ are determined using

$$
p_{ci} = -K \nabla \cdot \mathbf{u}_i + \alpha p_{fi} \quad (59)
$$

once the displacements $\mathbf{u}_i$ are known.

Low-frequency limit of $\gamma(\omega)$

As $\omega \to 0$, one can represent the local fields as asymptotic series in the small parameter $-i\omega$

$$
p_{fi} = p_{fi}^{(0)} - i\omega p_{fi}^{(1)} + O(\omega^2), \quad (60)
$$

$$
p_{ci} = p_{ci}^{(0)} - i\omega p_{ci}^{(1)} + O(\omega^2), \quad (61)
$$

and equivalently for $\mathbf{u}_i$. The zeroth-order response corresponds to uniform fluid pressure in the pores and is therefore given by $p_{c1}^{(0)} = p_{c2}^{(0)} = \Delta P$ and

$$
\frac{\overline{p}_{f1}^{(0)}}{\Delta P} = B_o = \frac{-a_{12} + a_{13}}{a_{22} + 2a_{23} + a_{33}} = \frac{1}{v_1/B_1 + v_2/B_2} \quad (62)
$$
where the patchy-saturation $a_{ij}$ have been employed. The fact that the quasi-static Skemp-ton’s coefficient in the patchy-saturation model is exactly the harmonic average of the constituents $B_i$ is equivalent to saying that at low frequencies, the fluid bulk modulus is given by $1/K_f = v_1/K_{f1} + v_2/K_{f2}$. The quasi-static response is thus completely independent of the spatial geometry of the fluid patches; it depends only on the volume fractions occupied by the patches.

The leading order correction to uniform fluid pressure is then controlled by the boundary-value problem

$$\frac{K k}{\alpha \eta_1} \nabla^2 p_{f1}^{(1)} = \frac{\eta_2}{\eta_1} \left(1 - \frac{B_o}{B_2}\right) \Delta P \text{ in } \Omega_2, \quad (63)$$

$$\frac{K k}{\alpha \eta_1} \nabla^2 p_{f1}^{(1)} = \left(1 - \frac{B_o}{B_1}\right) \Delta P \text{ in } \Omega_1, \quad (64)$$

$$p_{f1}^{(1)} = p_{f2}^{(1)} \text{ on } \partial \Omega_{12} \quad (65)$$

$$\mathbf{n} \cdot \nabla p_{f1}^{(1)} = \frac{\eta_2}{\eta_1} \mathbf{n} \cdot \nabla p_{f1}^{(1)} \text{ on } \partial \Omega_{12} \quad (66)$$

$$\mathbf{n} \cdot \nabla p_{f1}^{(1)} = 0 \text{ on } \partial E_i. \quad (67)$$

It is now assumed that for patchy-saturation cases of interest (air/water or water/oil), the ratio $\eta_2/\eta_1$ can be considered small. To leading order in $\eta_2/\eta_1$, equations (63), (66), and (67) require that $p_{f2}^{(1)}(\mathbf{r}) = \overline{p}_{f2}^{(1)}$ (a spatial constant). The fluid pressure in phase 1 is now rewritten as

$$p_{f1}^{(1)}(\mathbf{r}) = \overline{p}_{f1}^{(1)} - \frac{\eta_1 \alpha}{k K} \left(1 - \frac{B_o}{B_1}\right) \Delta P \Phi_1(\mathbf{r}) \quad (68)$$

where, from equations (64), (65) and (67) and to leading order in $\eta_2/\eta_1$, the potential $\Phi_1$ is the solution of the same elliptic boundary-value problem (25)–(27) given earlier.

Upon averaging (68) over all of $\Omega_1$, the leading order in $-i \omega$ difference in the average fluid pressures can be written

$$\overline{p}_{f1} - \overline{p}_{f2} \quad \Delta P = -i \omega \left( \frac{\overline{p}_{f1}^{(1)} - \overline{p}_{f2}^{(1)}}{\Delta P} \right) = i \omega \frac{\eta_1 \alpha}{k K} \left(1 - \frac{B_o}{B_1}\right) L_1^2 \quad (69)$$

where $L_1$ is again the length defined by equation (24).

To connect this fluid-pressure difference to the increment $\zeta_{int}$ we use the divergence theorem and the no-flow boundary condition on $\partial E_i$ to write equation (53) as

$$-i \omega \zeta_{int} = \frac{i \omega k}{\sqrt{\eta}} \int_{\partial \Omega_{12}} \mathbf{n} \cdot \nabla p_{f1}^{(1)} dS = i \omega v_1 \frac{\alpha}{K} \left(1 - \frac{B_o}{B_1}\right) \Delta P. \quad (70)$$

Replacing $\Delta P$ with $\overline{p}_{f1} - \overline{p}_{f2}$ using equation (69) then gives the desired law $-i \omega \zeta_{int} = \gamma_p (\overline{p}_{f1} - \overline{p}_{f2})$ with

$$\gamma_p = \frac{v_1 k}{\eta_1 L_1} \left[ 1 + O\left( \frac{\eta_2}{\eta_1} \right) \right]. \quad (71)$$

being the low-frequency limit of interest.
High-frequency limit of $\gamma(\omega)$

It has already been commented that in the extreme high-frequency limit where each patch behaves as if it were sealed to flow ($\tilde{\zeta}_{\text{int}} = 0$), the theory of Hill (1963) applies. Hill demonstrated, among other things, that when each isotropic patch has the same shear modulus, the volumetric deformation within each patch is a spatial constant. The fluid pressure response in this limit $p_{f_i}^\infty$ is thus a uniform spatial constant throughout each phase except in a vanishingly small neighborhood of the interface $\partial \Omega_{12}$ where equilibration is attempting to take place. The small amount of fluid-pressure penetration that is occurring across $\partial \Omega_{12}$ can be locally modeled as a one-dimensional process normal to the interface.

Using the coordinate $x$ to measure linear distance normal to the interface (and into phase 1), one has that equation (56) is satisfied by

$$
p_{f1} = p_{f1}^\infty + C_1 e^{i \sqrt{i \omega/D_1} x}$$  \hspace{1cm} (72)

$$
p_{f2} = p_{f2}^\infty + C_2 e^{-i \sqrt{i \omega/D_1} x}$$  \hspace{1cm} (73)

where the diffusivities are defined $D_i = kK B_i/\eta_i \alpha_i$. The constants $C_i$ are found from the continuity conditions (57) to be

$$
C_1 = -\frac{1}{1 + \sqrt{\eta_2 B_2/(\eta_1 B_1)}} (p_{f1}^\infty - p_{f2}^\infty)$$  \hspace{1cm} (74)

$$
C_2 = \frac{\sqrt{\eta_2 B_2/(\eta_1 B_1)}}{1 + \sqrt{\eta_2 B_2/(\eta_1 B_1)}} (p_{f1}^\infty - p_{f2}^\infty).$$  \hspace{1cm} (75)

Although not actually needed here, we have that $p_{f_i}^\infty = B_i p_{ci}$ where the uniform confining pressure of each patch is given by equations (40) and (41) so that the fluid pressure difference between the phases goes as

$$
\frac{p_{f1}^\infty - p_{f2}^\infty}{\Delta P} = \frac{B_1 - B_2}{1 - \beta(B_1/v_1 + B_2/v_2)},$$  \hspace{1cm} (76)

This equation is exactly the difference between equations (50) and (51). Because the penetration distance $\sqrt{D_i/\omega}$ vanishes at high-frequencies, we may state that to leading order in the high-frequency limit, $\overline{p}_{f1} - \overline{p}_{f2} = p_{f1}^\infty - p_{f2}^\infty$.

To obtain the high-frequency limit of the transport coefficient $\gamma(\omega)$, we use the definition (53) of the internal transport (note that $-\mathbf{n} \cdot \nabla p_{f1} = \partial p_{f1}/\partial x$)

$$
-i \omega \xi_{\text{int}} = \frac{1}{V} \frac{k}{\eta_1} \int_{\partial \Omega_{12}} \frac{\partial p_{f1}}{\partial x} \, dS$$  \hspace{1cm} (77)

along with equations (72) and (74). The result is

$$
\lim_{\omega \to \infty} \gamma(\omega) = i^{3/2} \frac{S}{V} \sqrt{\frac{k \alpha/(\eta_1 B_1 K)}{1 + \sqrt{\eta_2 B_2/(\eta_1 B_1)}}}.$$

(78)

Here, $S$ is again the area of $\partial \Omega_{12}$ contained within a volume $V$ of the patchy composite.
**Full-model for \( \gamma(\omega) \)**

The high- and low-frequency limits of \( \gamma \) are then connected by a simple frequency function to obtain the final model

\[
\gamma(\omega) = \gamma_p \sqrt{1 - i\omega/\omega_p} \quad (79)
\]

where the transition frequency \( \omega_p \) is defined

\[
\omega_p = \frac{B_1 K k (v_1 V/S)^2}{\eta_1 \alpha L_1^4} \left( 1 + \sqrt{\frac{\eta_2 B_2}{\eta_1 B_1}} \right)^2 \quad (80)
\]

and where \( \gamma_p = v_1 k/\eta_1 L_1^2 \). Equation (79) has a single singularity (a branch point) at \( \omega = -i\omega_p \). Causality requires that with an \( e^{-i\omega t} \) time dependence, all singularities and zeroes of a transport coefficient like \( \gamma(\omega) \) must reside in the lower-half complex \( \omega \) plane. Equation (79) satisfies this physically important constraint.

**Patchy-Saturation Modeling Choices**

To use the patchy-saturation model, appropriate values for the two geometric terms \( L_1 \) and \( V/S \) must be specified. Immiscible fluid distributions in the earth have very complicated geometries since they arise from slow flow that often produces fractal patch distributions. In particular, analytical solutions of the boundary-value problem (25)–(27) that defines \( L_1 \) for such real-earth situations are impossible. Recall that \( L_1 \) is a characteristic length of phase 1 (the phase having the smaller fluid mobility \( k/\eta \)) that defines the distance over which the fluid-pressure gradient is defined during the final stages of equilibration. For complicated geometries it may either be numerically determined, guessed at, or treated as a target parameter for a full-waveform inversion of seismic data. In the numerical examples that follow, we simply assume that the individual patches correspond to disconnected spheres for which simple analytical results are available for \( L_1 \) and \( V/S \).

If we consider phase 2 (porous continuum saturated by the less viscous fluid) to be in the form of spheres of radius \( a \) embedded within each radius \( R \) sphere of the two-phase composite, then \( v_2 = (a/R)^3 \), \( V/S = a v_2/3 \), and \( L_1^2 = 9 v_2^{2/3} a^2/14[1 - 7 v_2^{1/3}/6] \). This model is particularly appropriate when \( v_2 \ll v_1 \). Since the fluid 2 patches are disconnected, the definitions (11)–(13) of the effective poroelastic moduli again hold. Further, fluid 2 may be taken to be immobile relative to the framework of grains in the wavelength-scale Biot equilibration so that the inertial properties of equations (29) and (30) are identified as \( \rho_f = \rho_f 1 \), \( \rho = (1 - \phi) \rho_1 + \phi (v_1 \rho_f 1 + v_2 \rho_f 2) \) and \( \bar{\rho} = -\eta_1/(i\omega k) \).

In situations where it is more appropriate to treat fluid 1 (the more viscous fluid) as occupying disconnected patches (e.g., when \( v_1 \ll v_2 \)), the effective poroelastic moduli are defined by replacing 2 with 3 (and 3 with 2) in the subscripts of equations (11)–(13). Again assuming the phase-1 patches to be spheres of radius \( a \) embedded within radius \( R \) sphere of the two-phase composite, we have that \( v_1 = (a/R)^3 \) and \( V/S = a v_1/3 \). The elliptic boundary-value problem (25)–(27) can be solved in this case to give \( L_1^2 = a^2/15 \). Furthermore, the effective inertial
coefficients in the Biot theory are defined \( \rho_f = \rho_f \), \( \rho = (1 - \phi) \rho_v + \phi (v_1 \rho_{f1} + v_2 \rho_{f2}) \), and \( \bar{\rho} = -\eta_2 / (i \omega k) \).

In situations where both phases form continuous paths across each averaging volume, it is best to determine the attenuation and phase velocity by seeking the plane longitudinal-wave solution of non-reduced “double-porosity” governing equations of the form (6)–(10). However, this approach is not pursued here. We conclude by noting that if the embedded fluid is fractally distributed, the lengths \( L_1 \) will remain finite while \( (V/S) \rightarrow 0 \) as the fractal surface area \( S \) becomes large (however, \( V/S \) never reaches zero because the fractality has a small-scale cutoff fixed by the grain size of the material).

**Numerical Examples**

In Fig. 2, we compare the prediction of Johnson (2001) for \( K_U \) to our own for a consolidated sandstone (frame properties as determined in the Appendix with \( k = 100 \) mD, \( c = 10 \), \( \phi = 0.20 \)) in which phase 1 is saturated with water and phase 2 is taken to be spherical regions saturated with air. The two estimates have identical asymptotic dependence in both the limits of high and low frequencies. In the cross-over range, the physics is not precisely modeled in either approach. However, even in the cross-over range, the differences in the two models is slight.

Figure 3 gives the P-velocity and attenuation for a model in which the frame properties correspond to \( k = 10 \) mD, \( c = 15 \), and \( \phi = 0.15 \). Phase 2 is saturated by air and is taken to be isolated spheres of radius \( a = 1 \) cm. Phase 1 is saturated with water. The volume fraction \( v_2 \) occupied by these 1 cm spheres of gas is as shown in the figure. Even tiny amounts of gas saturation yields rather large amounts of attenuation and dispersion.

**SQUIRT-FLOW MODEL**

Laboratory samples of consolidated rock often have broken grain contacts and/or microcracks in the grains. Much of this damage occurs as the rock is brought from depth to the surface. Since diagenetic processes in a sedimentary basin tend to cement microcracks and grain contacts, it is uncertain whether *in situ* rocks have significant numbers of open microcracks. Nonetheless, when such grain-scale damage is present, as it always is in laboratory rock samples at ambient pressures, the fluid-pressure response in the microcracks will be greater than in the principal porespace when the rock is compressed by a P-wave. The resulting flow from crack to pore is called “squirt flow” and Dvorkin *et al.* (1995) have obtained a quantitative model for fully-saturated rocks.

In the squirt model of Dvorkin *et al.* (1995), the grains of a porous material are themselves allowed to have porosity in the form of microcracks. The effect of each broken grain contact is taken as equivalent to a microcrack in a grain. The number of such microcracks per grain is thus limited by the coordination number of the packing and so the total porosity contribution coming from the grains is always negligible compared to the porosity of the main porespace.
Figure 2: The undrained bulk modulus $K_U(\omega)$ in both the patchy-saturation model presented in this chapter and the model of Johnson (2001). The top graph is $\text{Re}(K_U)$ while the bottom graph is $Q_k^{-1} = -2\text{Im}(K_U)/\text{Re}(K_U)$. The physical model is 10 cm spherical air pockets embedded within a water-saturated region. The volume fraction of gas saturated rock is 3% in this example. The properties of the rock correspond to a 100 mD consolidated sandstone.
Figure 3: The P-wave velocity and attenuation of a sandstone saturated with water and containing small spherical pockets of gas having radius 1 cm and occupying a fraction of the volume $v_2$ as shown.
Our modeling of squirt is also based on this idea but we use the double-porosity framework of the previous sections. Phase 1 is now defined to be the pure fluid within the main porespace of a sample and is characterized elastically by the single modulus $K_f$ (fluid bulk modulus). Phase 2 is taken to be the porous (i.e., cracked) grains and characterized by the poroelastic constants $K_d^d$ (the drained modulus of an isolated porous grain), $\alpha_2$ (the Biot-Willis constant of an isolated grain), and $B_2$ (Skempton’s coefficient of an isolated grain) as well as by a permeability $k_2$. The overall composite of porous grains (phase 2) packed together within the fluid (phase 1) has two distinct properties of its own that must be specified; an overall drained modulus $K$, and an overall permeability $k$ associated with flow through the main porespace. The volume fractions occupied by each phase are again denoted $v_i$ where $v_1 = \phi$ is the porosity associated with the main porespace.

The theoretical approach is to again obtain the average fluid response in each of these two phases and then to make an effective Biot theory by saying that the fluid within the grains cannot communicate directly with the outside world; i.e., the fluid in the grains can only communicate with the main pores. Equations (11)–(12) again define the effective poroelastic moduli in the squirt model and we need only determine the $a_{ij}$ constants and internal transport coefficient $\gamma(\omega)$ that are appropriate to squirt.

**Squirt $a_{ij}$ Coefficients**

To obtain the $a_{ij}$ coefficients in the squirt model, we first note that these coefficients are defined under conditions where $\xi_{\text{int}} = 0$ (no fluid passing between the porous grains and the principal porespace). Under these conditions, the rate of fluid depletion $\nabla \cdot \mathbf{q}_1$ of a sample (rate of fluid volume being extruded from the principal pore space via the exterior sample surface as normalized by the sample volume) is due to the difference between the rate of dilatation of the principal porespace (denoted here as $\dot{e}_1$) and the rate at which fluid in the pores is dilating $-\dot{P}_{f1}/K_f$. If we also perform a volume average of equation (3) over the porous grain space and use the notation that $v_2\dot{e}_2 = \nabla \cdot (v_2\mathbf{u}_2)$ we obtain the following three equations

\[
\begin{align*}
\nabla \cdot \mathbf{q}_1 &= v_1\dot{e}_1 + \frac{v_1}{K_f} \dot{P}_{f1} \\
\nabla \cdot \mathbf{q}_2 &= \frac{v_2\alpha_2}{K_2^d} \dot{P}_{c2} - \frac{v_2\alpha_2}{B_2 K_2^d} \dot{P}_{f2} \quad (81) \\
v_2\dot{e}_2 &= \frac{v_2}{K_2^d} \dot{P}_{c2} - \frac{v_2\alpha_2}{K_2^d} \dot{P}_{f2}. \quad (82)
\end{align*}
\]

The macroscopic dilatation of interest is $\nabla \cdot \mathbf{v} = v_1\dot{e}_1 + v_2\dot{e}_2$. In order to obtain the macroscopic compressibility laws for porous-grain/principal-porespace composite, we introduce linear response laws of the form

\[
\begin{align*}
\dot{P}_{c2} &= a_1 \dot{P}_c + a_2 \dot{P}_{f1} + a_3 \dot{P}_{f2} \quad (84) \\
\dot{e}_1 &= b_1 \dot{P}_c + b_2 \dot{P}_{f1} + b_3 \dot{P}_{f2} \quad (85)
\end{align*}
\]

where the $a_i$ and $b_i$ must be found. We note immediately that from the definition $\dot{P}_c = v_1\dot{P}_{f1} + v_2\dot{P}_{c2}$ one has

\[
0 = (1 - v_2a_1)\dot{P}_c - (v_1 + v_2a_3)\dot{P}_{f1} - v_2a_3\dot{P}_{f2} \quad (86)
\]
which must hold true for any variation of the independent pressure variables so that \( a_1 = 1/v_2, a_2 = -v_1/v_2, a_3 = 0 \).

To obtain the \( b_i \), we now combine the above into the macroscopic laws

\[
-\nabla \cdot \mathbf{v} = \left[ v_1 b_1 + \frac{1}{K_d^2} \right] \dot{p}_c \\
+ \left[ v_1 b_2 - \frac{v_1}{K_d^2} \right] \dot{p}_{f1} + \left[ v_1 b_3 - \frac{v_2 \alpha_2}{K_d^2} \right] \dot{p}_{f2}
\]

(87)

\[
-\nabla \cdot \mathbf{q}_1 = -v_1 b_1 \dot{p}_c - \left[ v_1 b_2 - \frac{v_1}{K_f} \right] \dot{p}_{f1} - v_1 b_3 \dot{p}_{f2}
\]

(88)

\[
-\nabla \cdot \mathbf{q}_2 = -\frac{\alpha_2}{K_d^2} \dot{p}_c + \frac{v_1 \alpha_2}{K_d^2} \dot{p}_{f1} + \frac{v_2 \alpha_2}{K_d^2 B_2} \dot{p}_{f2}
\]

(89)

and use the fact that the coefficients of the matrix must be symmetric \((a_{ij} = a_{ji})\). With \( a_{11} = 1/K \) corresponding to the overall drained frame modulus of the composite (to be independently specified), we obtain \( v_1 b_1 = 1/K - 1/K_d^2, v_1 b_2 = -1/K + (1 + v_1)/K_d^2, \) and \( b_3 = -\alpha_2/K_d^2 \). The final \( a_{ij} \) coefficients are exactly

\[
a_{11} = 1/K \quad \text{(90)}
\]

\[
a_{22} = 1/K - (1 + v_1)/K_d^2 + v_1/K_f \quad \text{(91)}
\]

\[
a_{33} = \frac{v_2 \alpha_2}{B_2 K_d^2} \quad \text{(92)}
\]

\[
a_{12} = -1/K + 1/K_d^2 \quad \text{(93)}
\]

\[
a_{13} = -\alpha_2/K_d^2 \quad \text{(94)}
\]

\[
a_{23} = v_1 \alpha_2/K_d^2 \quad \text{(95)}
\]

Reasonable models for \( K \) and \( K_d^2 \) will be discussed shortly.

**Squirt Transport**

We next must obtain the coefficient \( \gamma(\omega) \) in the mesoscopic transport law \(-i \omega \xi_{\text{int}} = \gamma(\omega)(\overline{p}_{f1} - \overline{p}_{f2})\). Again, the approach is to first obtain the limiting behaviour at low and high frequencies and then to connect the two limits by a simple function.

The fluid response in phase 1 (the principal porespace) is governed by the Navier-Stokes equation \(-\nabla \overline{p}_{f1} + \eta \nabla^2 \mathbf{v}_1 = -i \omega \rho_{f} \mathbf{v}_1 \) and the compressibility law \( K_f \nabla \cdot \mathbf{v}_1 = i \omega \rho_{f1} \) where \( \mathbf{v}_1 \) is the local fluid velocity in the pores. Since for all frequencies of interest we have that \( \omega \ll K_f/\eta \) (note that \( K_f/\eta \approx 10^{12} \text{ s}^{-1} \) for liquids and \( 10^{10} \text{ s}^{-1} \) for gases), the fluid pressure in phase 1 is governed by the wave equation

\[
\nabla^2 \overline{p}_{f1} + \omega^2 \frac{\rho_f}{K_f} \overline{p}_{f1} = 0 \quad \text{(96)}
\]
and since the acoustic wavelength in the fluid is always much greater than the grain sizes, the fluid pressure in the principal porespace satisfies \( p_{f1}(\mathbf{r}) = \overline{p}_{f1} \) (a spatial constant) at all frequencies.

The focus, then, is on determining the flow and fluid pressure within the cracked grains (phase 2) that is governed by the local porous-continuum laws \( \mathbf{Q}_2 = -(k_2/\eta) \nabla p_{f2} \) and

\[
\frac{k_2}{\eta} \nabla p_{f2} + i \omega \frac{\alpha_2}{K_2^d} p_{f2} = -i \omega \frac{\alpha_2}{K_2^d} p_{c2} \quad (97)
\]

where \( p_{c2} = -K_2^d \nabla \cdot \mathbf{u}_2 + \alpha_2 p_{f2} \). This deformation and pressure change is excited by applying a uniform normal stress \(-\Delta P \mathbf{n}\) to the surface of the averaging volume with the fluid pressure satisfying the boundary conditions \( \mathbf{n} \cdot \nabla p_{f2}(\mathbf{r}) = 0 \) on \( \partial E_2 \) and \( p_{f2}(\mathbf{r}) = \overline{p}_{f1} \) on \( \partial \Omega_{12} \).

**Low-frequency limit of \( \gamma(\omega) \)**

The fluid pressure and confining pressure in the grains can again be developed as asymptotic series in \(-i \omega\ [as in equations (60)–(61)]. The zero-order response corresponds to the static limit in which the fluid pressure is everywhere the same and given by \( p_{f2}^{(0)} = \overline{p}_{f1} = B_o \Delta P \) with \( B_o = -(a_{12} + a_{13})/(a_{22} + 2a_{23} + a_{33}) \) and with the \( a_{ij} \) as given by equations (90)–(95).

The detailed result for \( B_o \) can be expressed

\[
\frac{1}{K} - \frac{(1 - \alpha_2)}{K_2^d} \frac{B_o}{K_2^d} = \frac{1}{K} - \frac{(1 - \alpha_2)}{K_2^d} \left[ \frac{1}{k_2 K_2^d} \right] + v_1 \left[ \frac{1}{k_2 K_2^d} - \frac{(1 - \alpha_2)}{K_2^d} \right] + v_2 \frac{\alpha_2}{K_2^d} \left[ \frac{1}{B_2} - 1 \right] \quad (98)
\]

which reduces to the standard Gassmann expression given in the appendix (with a total porosity given by \( v_1 + \phi_2 v_2 \)) when \( B_2 \) and \( \alpha_2 \) are themselves given by the Gassmann expressions. In this same zero-order limit, the undrained bulk modulus is defined as \( 1/K_o^d = a_{11} + (a_{12} + a_{13})B_o \) which also reduces to the standard Gassmann expression when \( B_2 \) and \( \alpha_2 \) are themselves given by Gassmann expressions.

The leading-order in \(-i \omega\) correction to uniform fluid pressure is thus governed by the problem

\[
\nabla^2 p_{f2}^{(1)} = \frac{\eta \alpha_2}{k_2 K_2^d} p_{c2}^{(0)} \quad (99)
\]

\[
\mathbf{n} \cdot \nabla p_{f2}^{(1)} = 0 \quad \text{on} \quad \partial E_2 \quad (100)
\]

\[
p_{f2}^{(1)} = 0 \quad \text{on} \quad \partial \Omega_{12}. \quad (101)
\]

Here, \( p_{c2}^{(0)} \) is the local confining pressure in the grain space in the static limit that can be written \( p_{c2}^{(0)}(\mathbf{r}) = \overline{p}_{c2}^{(0)} + \Delta P(\mathbf{r}) \). The average static confining pressure throughout the grains is determined from equation (84) with \( P_c = \Delta P \) and \( p_{f2} = p_{f1} = B_o \Delta P \) to yield

\[
\overline{p}_{c2}^{(0)} = \frac{(1 - v_1 B_o)}{v_2} \Delta P. \quad (102)
\]
The deviations $\delta P(r)$ thus volume integrate to zero $\overline{\delta P} = 0$ and are formally defined

$$\delta P(r) = - \left( \frac{1 - (v_1 + v_2 \alpha_2) B_o}{v_2} \right) \Delta P - \frac{K_s^d}{\alpha_2} \nabla \cdot \mathbf{u}^{(0)}(r). \quad (103)$$

The local perturbations $\delta P(r)$ are thus highly sensitive to the detailed nature of the grain packing and grain geometry. Fortunately, these perturbations do not play an important role in the theory.

The fluid pressure in the grains is now written in the scaled form

$$p_f^{(1)}(r) = - \frac{\eta \alpha_2 (1 - v_1 B_o)}{v_2 k_s K_s^d} \Delta P \Phi(r) \quad (104)$$

where the potential $\Phi(r)$ is independent of $\Delta P$ and is a solution of the elliptic problem

$$\nabla^2 \Phi(r) = -1 - \frac{v_2}{1 - v_1 B_o} \frac{\delta P(r)}{\Delta P} \quad (105)$$

$$\mathbf{n} \cdot \nabla \Phi = 0 \quad \text{on} \quad \partial E_2 \quad (106)$$

$$\Phi = 0 \quad \text{on} \quad \partial \Omega_{12}. \quad (107)$$

To leading-order in $-i \omega$, an average of equation (104) gives

$$\overline{p}_f^{(1)} - \overline{p}_f^{(2)} = i \omega \overline{p}_f^{(1)} + O(\omega^2) \quad (108)$$

$$= -i \omega \frac{\eta \alpha_2 (1 - v_1 B_o)}{v_2 k_s K_s^d} \overline{L}_2^2 \Delta P + O(\omega^2) \quad (109)$$

where the squared length $\overline{L}_2^2$ is defined

$$\overline{L}_2^2 = \overline{\Phi} = \overline{\Phi}_o \left[ 1 + \frac{v_2}{1 - v_1 B_o} \frac{\overline{\Phi}_o \Delta P}{\overline{\Phi}_o \Delta P} \right] \quad (110)$$

with overlines denoting volume averages over the grain space and with the potential $\Phi_o$ defined as the solution of

$$\nabla^2 \Phi_o = -1 \quad (111)$$

$$\mathbf{n} \cdot \nabla \Phi_o = 0 \quad \text{on} \quad \partial E_2 \quad (112)$$

$$\Phi_o = 0 \quad \text{on} \quad \partial \Omega_{12}. \quad (113)$$

Although it is not generally true that $\overline{\Phi}_o \Delta P = 0$ for all grain geometries, we nonetheless expect this integral to be small because $\Phi_o$ is a smooth function and $\overline{\Phi}_o = 0$. The local perturbations in the static confining pressure $\delta P(r)$ require a solution of the static displacements throughout the entire grain space—a daunting numerical task. Whenever the length $\overline{L}_2$ needs to be estimated, such as in the numerical results that follow, our approach is to simply use the reasonable approximation that $\overline{L}_2^2 = \overline{\Phi}_o$. 
Last, from the definition $\zeta_{\text{int}}$ of the internal transfer we have that to leading order in $-i\omega$

$$-i\omega\zeta_{\text{int}} = \frac{i\omega k_2}{\nu\eta} \int_{\partial\Omega_{12}} \mathbf{n} \cdot \nabla p_{f2}^{(1)}$$

$$= -\frac{i\omega k_2}{\nu\eta} \int_{\Omega_2} \nabla^2 p_{f2}^{(1)} = -i\omega \frac{\alpha_2}{K_2} v_2\bar{p}_c^{(0)}$$

$$= \frac{v_2 k_2}{\eta L_2^3} (\bar{p}_{f1} - \bar{p}_{f2}).$$

The normal $\mathbf{n}$ in equation (114) is outward to phase 1 which accounts for the sign change in equation (115). Note as well that equation (115) is a volume average of equation (99) while equation (116) follows from equations (102) and (109). The desired limit is thus $\lim_{\omega \to 0} \gamma(\omega) = \gamma_{sq} = v_2 k_2/\eta L_2^3$.

**High-frequency limit of $\gamma(\omega)$**

In the extreme high-frequency limit, the fluid has no time to significantly escape from the porous grains (phase 2) and enter the main porespace (phase 1). As such, the fluid pressure distribution in each phase is reasonably modeled as

$$p_{f1}(r) = B_1^\infty \Delta P$$

$$p_{f2}(r) = B_2^\infty \Delta P + C_2 \Delta P e^{-i\omega/2\sqrt{D_2 x}}$$

where $x$ is again a local coordinate measuring distance normal to the interface $\partial\Omega_{12}$ and where $D_2$ is the fluid-pressure diffusivity within the porous grains that is given by $D_2 = k_2 K_2^d B_2/\eta\alpha_2$. In reality, the local confining pressure $p_{c2}(r)$ throughout the grains has spatial fluctuations about the average value and we have made the approximation that $B_2 p_{c2}(r) \approx B_2^\infty \Delta P = \text{the average fluid pressure throughout the grain space}$. It is easy to demonstrate that under undrained and unrelaxed conditions,

$$B_1^\infty = \frac{a_{13}a_{23} - a_{33}a_{12}}{a_{22}a_{33} - a_{23}^2}$$

$$B_2^\infty = \frac{a_{12}a_{23} - a_{22}a_{13}}{a_{22}a_{33} - a_{23}^2}.$$

Since these $B_i^\infty$ do not appear in the final result, we do not bother substituting in the $a_{ij}$ constants from equations (90)–(95).

The continuity of fluid pressure $p_{f2} = p_{f1}$ along $\partial\Omega_{12}$ ($x = 0$) requires that $C_2 = B_1^\infty - B_2^\infty$. The definition of $\zeta_{\text{int}}$ may now be used to write

$$-i\omega\zeta_{\text{int}} = \frac{1}{V} \int_{\partial\Omega_{12}} \frac{k_2}{\eta} \frac{\partial p_2}{\partial x}$$

$$= \frac{k_2}{\eta} \frac{i^{3/2}}{\sqrt{D_2}} \frac{\omega S}{V} (B_1^\infty - B_2^\infty) \Delta P$$

$$= i^{3/2} \sqrt{\omega} \frac{k_2\alpha_2}{\eta B_2 K_2^d} \frac{S}{V} (\bar{p}_{f1} - \bar{p}_{f2}).$$
where we have used, to leading order in the high-frequency limit, that $\bar{p}_{f1} - \bar{p}_{f2} = (B_1^\infty - B_2^\infty)\Delta P$. The desired limit is thus $\lim_{\omega \to \infty} \gamma(\omega) = \sqrt{-i\omega k_2\alpha_2/(\eta B_2 K_s^d)S/V}$.

**Full model for $\gamma(\omega)$**

The high- and low-frequency limits are again causally connected via the simple function

$$\gamma(\omega) = \gamma_{sq}\sqrt{1 - \frac{i\omega}{\omega_{sq}}} \quad (124)$$

but now the parameters are defined as

$$\gamma_{sq} = \frac{v_2 k_2}{\eta L_2^2} \quad (125)$$

$$\omega_{sq} = \frac{B_2 K_s^d k_2}{\eta \alpha_2 L_2^2} \left(\frac{v_2 V/S}{L_2}\right)^2. \quad (126)$$

**Squirt Flow Modeling Choices**

To make numerical predictions of attenuation and dispersion, models must be proposed for the phase 2 (porous grain) parameters.

If the grains are modeled as spheres of radius $R$, the fluid-pressure gradient length within the grains can be estimated as $L_2 = R/\sqrt{15}$ and the volume to surface ratio as $V/S = R/(3v_2)$. The grain porosity is assumed to be in the form of microcracks and so it is natural to define an effective aperture $h$ for these cracks. If the cracks have an average effective radius of $\bar{R}$ where $\bar{R}$ is roughly 2 or 3 and if there are on average $N_c$ cracks per grain where $N_c$ is also roughly 2 or 3 then the permeability and porosity of the grains is reasonably modeled as

$$\phi_2 = \frac{3N_c h}{4N_R^2 R} \quad \text{and} \quad k_2 = \phi_2 h^2/12 \quad (127)$$

where $\phi_2$ is the fracture porosity within the porous grains. The dimensionless parameters $k_2/L_2^2$ and $(v_2 V/S)/L_2$ required in the expressions for $\gamma_{sq}$ and $\omega_{sq}$ are given by

$$\frac{k_2}{L_2^2} = \frac{15N_c}{16N_R^2} \left(\frac{h}{R}\right)^3 \quad \text{and} \quad \left(\frac{v_2 V/S}{L_2}\right)^2 = \frac{5}{3}. \quad (128)$$

The normalized fracture aperature $h/R$ is the key parameter in the squirt model.

The drained grain modulus $K_2^d$ is necessarily a function of the crack porosity $\phi_2$ (and therefore $h/R$). Real crack surfaces have micron (and smaller) scale asperities present upon them. If effective stress is applied in order to make the normalized aperture $h/R$ smaller (so that, for example, the peak in squirt attenuation lies in the seismic band), new contacts are created that make the crack stronger. In the limit as $h/R \to 0$ (large effective stress), the cracks are no longer present and $K_2^d \to K_s$ where $K_s$ is the mineral modulus of the grain.
Many models for such stiffening could be proposed. We intentionally make a conservative estimate here in proposing a simple linear porosity dependence $K^d_2 = K_s(1 - \sigma \phi_2)$, where $\sigma$ is a fixed constant determined from fitting ultrasonic attenuation data. Effective medium theories [see, for example, Berryman et al. (2002)] predict that $\sigma$ should be inversely proportional to the aspect ratios of the cracks present. As a crack closes and asperities are brought into contact, there is naturally a decrease in $\phi_2$ but there should also be a decrease in $\sigma$ due to the fact that the remaining crack porosity becomes more spherical as new asperities come into contact. Taking $\sigma$ to be constant as crack porosity decreases is thus a minimalist estimate for how the drained modulus increases.

Thus, the porous-grain elastic properties are taken to be

$$K^d_2 = K_s(1 - \sigma \phi_2) \quad (129)$$
$$\alpha_2 = 1 - K^d_2 / K_s \quad (130)$$
$$\frac{1}{B_2} = 1 + \phi_2 \frac{K^d_2}{K_f} \left( \frac{1 - K_f / K_s}{1 - K^d_2 / K_s} \right) \quad (131)$$

where we have used the Gassmann uid-substitution relations for $\alpha_2$ and $B_2$. The overall drained modulus $K$ of the collection of porous (cracked) grains can be modeled for example as

$$K = \frac{K^d_2(1 - v_1)}{1 + cv_1} \quad (132)$$

which is the same drained-modulus model as given in the appendix but with the solid grain modulus $K_s$ replaced by the cracked grain modulus $K^d_2$.

**Numerical Examples**

In Fig. 4, we plot the P-wave attenuation predicted using the above model when the overall grain packing corresponds to a consolidated sandstone ($v_1 = 0.2$ and $c = 5$) having a permeability of 10 mD. For the grain properties we take $\sigma = 0.8/(5 \times 10^{-3})$, $3N_c/(4N^2_f) = 1$, and $K_s = 38$ GPa (quartz) as fixed constants. This $\sigma$ value was chosen so that there would be an important peak in attenuation at ultrasonic frequencies and is taken to be the same for all values of $h/R$. The various curves can be thought of as being due to the application of effective stress. The peak in $Q^{-1}$ near 1 MHz that is invariant to $h/R$ is that due to the macroscopic Biot loss (fluid pressure equilibration at the scale of the wavelength). The peak that shifts with $h/R$ is that due to the squirt flow.

This figure indicates that although the squirt mechanism is probably operative and perhaps even dominant at ultrasonic frequencies, it does not seem to be involved in explaining the observed levels of intrinsic attenuation in exploration work. For real cracks inside of real grains, the $\sigma$ value will diminish with effective stress (i.e., with $h/R$), so that the effects of squirt in the seismic band are likely to be even less than shown in Fig. 4.

We next introduce the grain parameters $k_2$, $\phi_2$, and $K^d_2$ as modeled here along with the same overall drained modulus $K$ into the model of Dvorkin et al. (1995) and compare the
Figure 4: The squirt-flow model of P-wave attenuation when the grains are modeled as being spherical of radius $R$ and containing microcracks having effective apertures $h$. The overall drained modulus of the rock corresponds to a consolidated sandstone.
Figure 5: The undrained bulk modulus $K_U(\omega)$ in the squirt model of the present study and in the model of Dvorkin et al. (1995). The top graph is $\text{Re}(K_U)$ while the bottom graph is $Q_K^{-1} = -2\text{Im}(K_U)/\text{Re}(K_U)$. The frame properties are the same as in Fig. 4. The curves having a smaller relaxation frequency ($\simeq 10^3$ Hz) and almost no dispersion correspond to $h/R = 2 \times 10^{-4}$ while the curves having the larger critical frequency ($\simeq 10^5$ Hz) and more dispersion correspond to $h/R = 5 \times 10^{-3}$. [NR]
results to our own model for two different values of $h/R$ (Fig. 5). Although both models have similar dependencies on the various material properties involved, there are nonetheless significant differences. These are principally due to the fact that the Dvorkin et al. (1995) model requires the grains to be in the form of effective cylinders of radius $R$, while in Fig. 5 we use a geometric parameter $L_2$ and volume-to-surface ratio $V/S$ that are appropriate for spherically-shaped grains. However, in various limits as the frequency and/or fluid bulk modulus become either large or small, we have verified that both models yield qualitatively similar results.

CONCLUSIONS

Models for three different P-wave attenuation mechanisms were presented that differ only in the values of the $a_{ij}$ constants and in the nature of the mesoscopic transport coefficient $\gamma(\omega)$. These three models correspond to (1) mesoscopic-scale heterogeneity in the frame moduli (“double porosity”), (2) mesoscopic-scale heterogeneity in the fluid type (“patchy-saturation”), and (3) grain-scale heterogeneity due to microcracks in the grains (“squirt”). In all three models, the amount of attenuation is controlled principally by the contrast of elastic compressibility among the constituents. In the double-porosity model, it is the contrast between the frame bulk-modulus of the two porous phases that is key, while in the patchy-saturation model it is the contrast in the fluid bulk modulus (immiscible patches of different fluids that have nearly identical bulk moduli would not produce much attenuation), and in the squirt model, it is the contrast between the drained modulus of an isolated cracked grain and that of the entire packing of grains.

Putting in small pockets of unconsolidated sand grains into an otherwise consolidated sandstone can produce attenuation in the seismic band that is comparable to what is measured in the field even when the pockets represent only a small amount of the total volume (< 1% volume fractions). Since mesoscopic-scale heterogeneity is rather ubiquitous throughout the earth’s crust, it seems reasonable to suppose that this mechanism may be responsible for most of the attenuation observed in seismograms. The squirt mechanism produces a great deal of attenuation at the ultrasonic frequencies used in laboratory measurements, but has trouble explaining attenuation in the seismic band. This result is good news for some important applications of the theory because the mesoscopic-scale flow is affected by the permeability of the material, while squirt flow is not. This leaves open the possibility of extracting permeability information from the frequency dependence of seismically measured $Q$.

APPENDIX A. CONSTITUENT PROPERTIES

In order to use the unified double-porosity framework of the present paper, it is convenient to have models for the various porous-continuum constituent properties.

For unconsolidated sands and soils, the frame moduli (drained bulk modulus $K^d$ and shear modulus $G$) are well modeled using the following variant of the Walton (1987) theory [c.f.,

\[ K^d = \frac{1}{6} \left[ \frac{4(1 - \phi_o)^2 n_o^2 P_o}{\pi^4 c_s^2} \right]^{1/3} \frac{(P_e/P_o)^{1/2}}{\left[ 1 + [16P_e/(9P_o)]^4 \right]^{1/24}}, \]  
\( (133) \)

\[ G = 3K^d/5, \]  
\( (134) \)

where \( P_e \) is the effective overburden pressure [e.g., \( P_e = (1 - \phi)(\rho_s - \rho_f)gh \) where \( g \) is gravity and \( h \) is overburden thickness] and where \( P_o \) is the effective pressure at which all grain-to-grain contacts are established. For \( P_e < P_o \), the coordination number \( n \) (average number of grain contacts per grain) is increasing as \( (P_e/P_o)^{1/2} \). For \( P_e > P_o \), the coordination number remains constant \( n = n_o \). The parameter \( P_o \) is commonly on the order of 10 MPa. As \( P_o \to 0 \), the Walton (1987) result is obtained (all contacts in place starting from \( P_e = 0 \). The porosity of the grain pack is \( \phi_o \) and the compliance parameter \( C_s \) is defined

\[ C_s = \frac{1}{4\pi} \left( \frac{1}{G_s} + \frac{1}{K_s + G_s/3} \right) \]  
\( (135) \)

where \( K_s \) and \( G_s \) are the mineral moduli of the grains. For unimodal grain-size distributions and random grain packs, one typically has \( 0.32 < \phi_o < 0.36 \) and \( 8 < n_o < 11 \).

For consolidated sandstones, the frame moduli are modelled in the present paper as [c.f., Pride (2003) for details]

\[ K^d = K_s \frac{1 - \phi}{1 + c\phi}, \]  
\( (136) \)

\[ G = G_s \frac{1 - \phi}{1 + 3c\phi/2}. \]  
\( (137) \)

The consolidation parameter \( c \) represents the degree of consolidation between the grains and lies in the approximate range \( 2 < c < 20 \) for sandstones. If it is necessary to use a \( c \) greater than say 20 or 30, then it is probably better to use the modified-Walton theory.

The undrained moduli \( K^u \) and \( B \) are conveniently and exactly modeled using the Gassmann (1951) theory whenever the grains are isotropic and composed of a single mineral. The results are

\[ B = \frac{1/K^d - 1/K_s}{1/K^d - 1/K_s + \phi(1/K_f - 1/K_s)}, \]  
\( (138) \)

\[ K^u = \frac{K^d}{1 - B(1 - K^d/K_s)}, \]  
\( (139) \)

from which the Biot-Willis constant \( \alpha \) may be determined to be \( \alpha = 1 - K^d/K_s \). These Gassmann results are often called the “fluid-substitution” relations.

The dynamic permeability \( k(\omega) \) as modeled by Johnson et al. (1987) is

\[ \frac{k(\omega)}{k_o} = \left[ 1 - i \frac{\omega}{n_f \omega_c} - i \frac{\omega}{\omega_c} \right]^{-1}, \]  
\( (140) \)
where the relaxation frequency $\omega_c$, which controls the frequency at which viscous-boundary layers first develop, is given by

$$\omega_c = \frac{\eta}{\rho_f F k_o}. \quad (141)$$

Here, $F$ is exactly the electrical formation factor when grain-surface electrical conduction is not important and is conveniently (though crudely) modeled using Archie’s law $F = \phi^{-m}$. The cementation exponent $m$ is related to the distribution of grain shapes (or pore topology) in the sample and is generally close to 3/2 in clean sands, close to 2 in shaly sands, and close to 1 in rocks having fracture porosity. The parameter $n_j$ is, for convenience, taken to be 8 (cylinder model of the porespace).

REFERENCES


SEP ARTICLES PUBLISHED OR IN PRESS


# 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>Chen, Weitian</td>
<td>724-0461</td>
<td>chen</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>
</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, 2002-2003

<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</td>
<td>(281) 366-4604</td>
<td><a href="mailto:abmarl@bp.com">abmarl@bp.com</a></td>
</tr>
<tr>
<td>Francois Audebert</td>
<td>CGG</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) 689-5717</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></td>
<td>(Co-chair, 2nd year)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helmut Jakubowicz</td>
<td>Veritas</td>
<td>(44) 1293 443219</td>
<td><a href="mailto:helmut_jakubowicz@veritasdgc.com">helmut_jakubowicz@veritasdgc.com</a></td>
</tr>
<tr>
<td></td>
<td>(Co-chair, 1st year)</td>
<td></td>
<td></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

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.

Weitian Chen received a B.A. in geophysics from Peking University in 1998 and obtained a M.Sc. in geophysics from Peking University in 2001. He is currently working towards a M.Sc. in geophysics at Stanford and is a member of SEG.

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 Geo-
physical Engineering from Colorado School of Mines in May
1997. She joined SEP in September 1997 and received her MS
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 cur-
rently edging towards a Ph.D. in geophysics. She is a member of
SEG.

William Curry is a third year student with SEP, who graduated in
2000 from the University of Alberta with a B.Sc. (Hons.) in Geo-
physics, 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.

Sergey Fomel received his Diploma in Geophysics from Novosibi-
rsk University in Russia in 1990. Prior to joining the Stanford
Exploration Project, he worked at the Institute of Geophysics in
Novosibirsk. Received a Ph.D. in Geophysics from Stanford in
for "numerous contributions to seismology." Postdoctoral fellow at
the Lawrence Berkeley National Laboratory and visiting assistant
professor at the University of California at Berkeley in 2001-2002.
In 2002, he joined the Bureau of Economic Geology at the Univer-
sity of Texas at Austin as a Research Associate.
Antoine Guitton received a MSc 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.

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.
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.

Guojian Shan received his B.Sc in Mathematics school of Peking University in July, 1998. From 1998 to 2001, he was studying in Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences (CAS), and got his M.S in Applied Mathematics in July, 2001. He joined SEP in 2001 and now he is currently working toward a Ph.D. in geophysics. He is a student member of 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 toward 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 the 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.
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2002-2003

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
6001 Bollinger Canyon Rd.
D-1058
San Ramon, CA 94583
USA
tel: (925) 842-6266
fax: (925) 842-2061
contact: Lin Zhang
e-mail: lizh@chevrontexaco.com

Aramco Services Company
Saudi Aramco
Geophysics Technology
P.O. Box 750, EXPEC 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

Chinese Petroleum Corporation
Geophysical Data Processing Center
Taiwan Petroleum Exploration Division
14th Fl., No. 3, Sungren Road
Shinyi Chiu
Taipei, Taiwan 11010
tel: 886-2-87259910
fax: 886-2-87899060
contact: Jenyang Lin
e-mail: jenyang@cpc.com.tw

BP Exploration & Production
200 Westlake Park Blvd.
WL4 1018
Houston, TX 77079
USA
tel: (281) 366-3611
fax: (281) 366-5856
contact: John T. Etgen
e-mail: etgenjt@bp.com

ConocoPhillips Inc.
Seismic Processing Development
540G Plaza Office Bldg.
Bartlesville, OK 74004
USA
tel: (918) 661-1514
fax: (918) 662-1587
contact: Robert M. Habiger
e-mail: rmhabig@ppco.com

CGG Americas Inc.
16430 Park Ten Place
Houston, TX 77084
USA
tel: (281) 646-2524
fax: (281) 646-2620
contact: Francois Audebert
e-mail: faudebert@cgg.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
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2002-2003

ENI - E & P Division
Dept. RIGE
via Unione Europea 3
20097 S. Donato Milanese
Italy
tel: 39 (02) 520 55308
depth: 39 (02) 520 45694
contact: Vittorio De Tomasi
e-mail: vittorio.detomasi@agip.it

JGI, Inc.
Meikei Building
1-5-21 Otsuka, Bunkyo-ku
Tokyo, 112-0012
Japan
tel: 81 (3) 5978 8043
depth: 81 (3) 5978 8060
contact: Yoichi Ohta
e-mail: yohta@jgi.co.jp

ExxonMobil Upstream Research Co.
Seismic Processing Research
3319 Mercer St., ST-401
P.O. Box 2189
Houston, TX 77027
USA
tel: (713) 431-6011
depth: (713) 431-6326
contact: Thomas A. Dickens
e-mail: tom.a.dickens@exxonmobil.com

Landmark Graphics Corporation
1805 Shea Center Dr.
Suite 400
Highlands Ranch, CO 80129-2258
USA
tel: (303) 779-8080
depth: (303) 796-0807
contact: Stewart A. Levin
e-mail: salevin@lgc.com

4th Wave Imaging Corporation
16A Journey
Suite 200
Aliso Viejo, CA 92656
USA
tel: (949) 916-9787
depth: (949) 916-9786
contact: David E. Lumley
e-mail: david.lumley@4thwaveimaging.com

Norsk Hydro Produksjon a.s
PB 7190
Sandloseien 90
N-5020 Bergen
Norway
tel: 47 (55) 5000
depth: 47 (55) 6861
contact: Arnnn Sagehaug
e-mail: arnfinn.sagehaug@nho.hydro.com

GX Technology Corporation
5847 San Felipe
Suite 3500
Houston, TX 77057-3010
USA
tel: (713) 789-7250
depth: (713) 789-7201
contact: Nick Bernitsas
e-mail: nxb@gxt.com

Paradigm Geophysical Corporation
1200 Smith St.
Suite 2100
Houston, TX 77002
USA
tel: (713) 393-4979
depth: (713) 393-4901
contact: Orhan Yilmaz
e-mail: yilmaz@paradigmgeo.com
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2002-2003

Petrobras S.A.
Av. Chile 65, sala 1302
Rio de Janeiro
20035-900 RJ
Brazil
tel: 55 (21) 2534 2706
tax: 55 (21) 2534 1076
contact: Carlos A. Cunha Filho
e-mail: s002@ep.petrobras.com.br

Av. Chile 65, sala 1302
Rio de Janeiro
20035-900 RJ
Brazil
tel: 55 (21) 2534 2706
tax: 55 (21) 2534 1076
contact: Carlos A. Cunha Filho
e-mail: s002@ep.petrobras.com.br

3DGeo Development Inc.
465 Fairchild Dr.
Suite 227
Mountain View, CA 94043
USA
tel: (650) 969-3886
tax: (650) 969-6422
contact: Dimitri Bevc
e-mail: dimitri@3dgeo.com

PGS Inc., Research Division
10550 Richmond Ave.
Suite 250
Houston, TX 77042
USA
tel: (713) 735-6322
tax: (713) 532-6774
contact: Ruben D. Martinez
e-mail: ruben.martinez@pgs.com

TOTAL E&P USA, Inc.
800 Gestner
Suite 700
Houston, TX 77024
USA
tel:
tax:
contact: Henri Calandra
e-mail: henri.calandra@totalfinaelf.com

Shell International E&P Inc.
P.O. Box 481
Houston, TX 77001-0481
USA
tel: (713) 245-7195
tax: (713) 245-739
contact: Richard Cook
e-mail: richardcook@shell.com

Unocal
14141 Southwest Freeway
Sugar Land, TX 77478
USA
tel: (281) 287-7481
tax: (281) 287-5360
contact: Philip S. Schultz
e-mail: phil.schultz@unocal.com

Statoil Research Centre
Arkitekt Ebbellsvei 10
Rotvoll
Trondheim N-7005
Norway
tel: (47) 73584917
tax: (47) 73584665
contact: Constantin Gerea
e-mail: geco@statoil.com

Veritas DGC Ltd.
Crompton Way
Manor Royal Estate
Crawley, West Sussex RH10 2QR
England
tel: 44 (1293) 443219
tax: 44 (1293) 443010
contact: Helmut Jakubowicz
e-mail: helmut_jakubowicz@veritasdgc.com
SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2002-2003

WesternGeco
Houston Technology Center
10001 Richmond Ave.
Houston, TX 77042-4299
USA
tel: (713) 689-5717
fax: (713) 689-5757
contact: Luis L. Canales
email: lcanales@houston.westerngeco.slb.com