# The resolving power of seismic amplitude data: An anisotropic inversion/migration approach

Maarten V. de Hoop*, Carl Spencer‡, and Robert Burridge**

## ABSTRACT

A description of the theory and numerical implementation of a 3-D linearized asymptotic anisotropic inversion method based on the generalized Radon transform is given. We discuss implementation aspects, including (1) the use of various coordinate systems, (2) regularization by both spectral and Bayesian statistical techniques, and (3) the effects of limited acquisition apertures on inversion. We give applications of the theory in which well-resolved parameter combinations are determined for particular experimental geometries and illustrate the interdependence of parameter and spatial resolutions. Procedures for evaluating uncertainties in the parameter estimates that result from the inversion are derived and demonstrated.

## INTRODUCTION

The purpose of this paper is to investigate the use of the generalized Radon transform (GRT) for seismic inverse problems involving anisotropic earth models. An understanding of anisotropy is important for hydrocarbon exploration because shales, which make up 75% of the sedimentary cover of the hydrocarbon reserves, are almost invariably anisotropic. The effects of anisotropy on the kinematics of $P$-wave propagation and hence their effects on conventional seismic processing are summarized in Larner and Tsvankin (1995). Ball (1995) presented a real data example from a carbonate reservoir in the former Zaire showing the effects of anisotropy on migration.

Anisotropy can also have a dramatic effect on the amplitude versus angle (AVA) response of a geological interface. As an example, in cases where shales and sands have similar acoustic impedances, the introduction of anisotropy in the shale may bring about a change in sign of the reflection coefficient not present in the isotropic case. Similarly, it is possible

to encounter situations where amplitude anomalies that otherwise might be attributed to the presence of gas in isotropic media might also be caused by the presence of anisotropy in the absence of gas. Our intention in this paper is to demonstrate a framework capable of answering the fundamental questions: What information about anisotropy do seismic amplitudes reveal, and how do we use this information to image rock properties?

Over the past decade, substantial progress has been made towards solving the problem of inverting seismic data to yield models of the physical properties of the earth. Several different techniques have been suggested by various scientists based on what, at first sight, seem like very different approximations of the inverse problem, but which turn out in practice to bear many similarities. The different techniques can be classified according to (1) the method of carrying out the forward modeling (for example, full-waveform, Kirchhoff, ray-Born, etc.), (2) the method of inverting the forward relation (for example, nonlinear local optimization possibly preceded by preconditioning, or the direct GRT), (3) the parameterization of the subsurface (scalar, isotropic elastic, anisotropic elastic, and poro-elastic represent increasingly sophisticated medium descriptions). Also, the choice of a forward modeling approach usually implies a particular discretization of the scattering domain, i.e., the subsurface. It is the second classification (2) that is most fundamental to a discussion of practical inversion methods.

Perhaps the most obvious way of solving the seismic inverse problem is to search parameter space by techniques such as the conjugate gradient minimization of some measure of the misfit between observed and simulated seismograms. Such an approach was developed for the acoustic case by Bamberger et al. (1982) and was subsequently modified and extended by numerous authors (Tarantola, 1984, 1986; Gauthier et al., 1986; Ikelle et al., 1986; Beydoun and Mendez, 1989; Mora, 1989; Singh et al., 1989; Snieder et al., 1989; Cao et al., 1990; Eaton and Stewart, 1994; Dębski and Tarantola, 1995). Two essential features of all these search methods are the use of iterative solutions to the (nonlinear) optimization problem and

regularization by inclusion of a priori information. Cheng and Cohen's (1984) and Tarantola's (1984) observation that inversion can be expressed in terms of more conventional seismic processing methods is a recurring motif in the literature that applies to almost all common approaches to the inverse problem (Bleistein and Cohen, 1979; Clayton and Stolt, 1981; Mora, 1989; Claerbout, 1992).

A second suite of inverse methods originated in the field of ultrasonics (Norton and Linzer, 1981) and is based on approximations to the forward and inverse formulations that permit direct, closed-form expressions for the inverse problem solution (Clayton and Stolt, 1981; Devaney, 1984; Beylkin, 1985; Miller et al., 1987; Bleistein, 1987; de Hoop et al., 1994; Burridge et al., 1998). The majority of these methods use the Born approximation to model scattering at the target together with asymptotic approximations such as WKBJ or asymptotic ray theory for propagation to and from the scatterer. Solutions are then obtained by mappings, such as the GRT, which require further high-frequency approximations. Since the starting point for both the direct and search-based inversion techniques is a weak formulation of the inverse problem, the mechanics of both methods can turn out to be similar (Esmersoy and Oristaglio, 1988; Jin et al., 1992). It is also possible to use operators obtained using direct linearized methods asymptotically as preconditioners for search-based nonlinear optimization (de Hoop and de Hoop, 1997).

In this paper, we will be exclusively concerned with the multiparameter elastic inverse problem. Several approaches to multiparameter inversion have been proposed in the isotropic case. Berkhout and Wapenaar (1990) suggested that wavefield decomposition be applied at an early stage in processing the data so that $P$ and $S$ pseudoscalar wavefields or potentials can be inverted separately. Bleistein (1987) modified the generalized Radon transform method of Beylkin (1985) and Miller et al. (1987) to the three-parameter elastic case by formulating the problem as Kirchhoff scattering from interfaces rather than Born scattering from volumes. The inversion process involved two parallel integrations (weighted diffraction stacks) that allow both reflection coefficient as a function of angle and the angle of specular reflection to be recovered at each image point. Beylkin and Burridge (1990) proposed a multiparameter scheme based on the Born approximation instead. A GRT inversion was designed to construct an intermediate vector quantity from which elastic parameters could be recovered. This method avoided the need for dividing two image sections to provide angle information, as required by the method of Bleistein (1987). Among the optimization approaches, Tarantola (1986), Beydoun and Mendez (1989), and Jin et al. (1992) have considered the isotropic multiparameter problem. Parameter estimates are obtained by a back-propagation of the residual wavefield followed by convolution with an approximate inverse Hessian arising from the local linearization of the forward contrast-source formulation.

The introduction of anisotropy into elastic seismic inversion increases the number of possible parameters needed to specify the physical properties of a point within the earth to 22. This is far more than can be recovered in practice and, therefore, attention must be given to reducing the size of the problem. One possible approach is to understand the scattering process well enough beforehand to be able to provide a limited number of combinations of parameters that describe the process

effectively. Banik (1987) and Tsvankin and Thomsen (1995) have done so for the case of weak scattering in VTI (transversely isotropic with vertical axis of symmetry) media and find that near-normal incidence $P$-wave scattering behavior is controlled by vertical impedance contrast and the parameter $\delta$ (Thomsen, 1986). This analysis has been extended to the case of orthorhombic media with a symmetry plane aligned with a planar scattering interface (Rüger, 1996). In more general cases and with arbitrary recording geometries, it may be less obvious which parameter combinations to use. The alternative approach, and the one adopted in this paper, is to use linear inverse theory to evaluate the "best resolved" parameter. We use the phrase "best resolved parameter" to refer to the combination of elastic moduli at an image point that is least dependent on other combinations of moduli at the same image point. In multiparameter inverse problems, the resolving power of an experiment is made up of two parts. Spatial resolution quantifies the blurring of the image in space, whereas the parameter resolution quantifies the linear interdependence of elastic moduli and density. Both effects are described by a resolution operator that will be discussed in this paper. A second feature of using linear inverse theory is that it is possible to calculate estimates of parameter uncertainty by mapping noise in the data, which may be described as a data covariance matrix, into errors in estimates of moduli.

Both the GRT and optimization approaches to the inversion of seismic data for anisotropic parameters have been attempted (de Hoop et al., 1994; Eaton and Stewart, 1994; Burridge et al., 1998). All these authors used a Born formulation for the forward problem and all account for ill-posedness in the parameter part of the problem by solving a reduced linear system via a singular value decomposition. A significant difference between the isotropic and anisotropic inversion cases is that most direct isotropic multiparameter inverse methods yield algorithms that can be described as migration followed by inversion (i.e., energy is positioned before an amplitude versus scattering angle or offset relationship is inverted). Such an ordering is not appropriate for the anisotropic case because the scattering response of an elastic modulus such as $C_{33}$ no longer depends on scattering angle alone. It is a function of both scattering angle and interface normal direction (or more correctly the gradient in total traveltime). Burridge et al. (1998) discuss this problem in detail and formulate extensions to the GRT algorithm of Miller et al. (1987) for anisotropic media. The algorithm, which will be discussed further in the following sections, can be separated into an AVA inversion for each migration dip, followed by migration, the integration over dips—hence the title of this paper.

In the remainder of this paper, we develop the anisotropic GRT inversion theory, paying special attention to the calculation of the various Jacobians involved, and address the issues of regularization and acquisition aperture compensation. We then provide several synthetic examples illustrating the essential features of the method and draw general conclusions.

## SOURCE-RECEIVER RAY GEOMETRY

Our intention in the following sections is to present a mathematical description of our inversion procedure, emphasizing aspects such as recording geometries, ill-posedness, and aperture limitation that arise in practical applications. The development

here is for the 3-D case, and we will make use of asymptotic theories applicable to high-frequency wave propagation and inversion for the most singular constituents of the medium. More complete descriptions of the ray-theory Green's function calculations can be found in Kendall et al. (1992) and of the inversion method in de Hoop et al. (1994) and Burridge et al. (1998). In Tables 1 and 2, a glossary of symbols is provided.

We begin by giving ray-theoretical expressions for the propagation of phase and amplitude in anisotropic media with elastic moduli $c_{ijk\ell}$ (Voigt notation: $C_{IJ}$, $I$, $J = 1, \ldots, 6$) and density $\rho$ (Shearer and Chapman, 1989; Kendall et al., 1992). Let $\tau$ be the arrival time and $\boldsymbol{\xi}$ the associated polarization vector. We use $x$ to denote a point in $\mathbb{R}^3$, subscripts denote components of vectors and tensors, and $\partial_j$ denotes partial differentiation with respect to the $j$th component of $x$. The polarization vectors are assumed to be normalized so that $\xi_i \xi_i = 1$. Define the slowness vector $\gamma$ by

$$\gamma(x) = \nabla_x \tau(x, x') \tag{1}$$

along the ray originating at $x'$. Then $\gamma$ and $\boldsymbol{\xi}$ satisfy the "eigenvalue" equation

$$(\rho\delta_{ik} - c_{ijk\ell}\gamma_\ell\gamma_j)\xi_k = 0 \quad \text{(at all } x\text{)}. \tag{2}$$

Equation (2) constrains $\gamma$ to lie on the sextic surface, $\mathcal{A}(x)$, given by

$$\det(\rho\delta_{ik} - c_{ijk\ell}\gamma_\ell\gamma_j) = 0. \tag{3}$$

By virtue of equation (1), equation (3) may be interpreted as a nonlinear partial differential equation, the eikonal equation, for $\tau$. The surface $\mathcal{A}(x)$ consists of three sheets, each of which

is a closed surface surrounding the origin. An individual sheet is described by equation (2), leading to

$$\mathcal{H}(x, \gamma) = \tfrac{1}{2}(\rho - \xi_i c_{ijk\ell}\gamma_\ell\gamma_j\xi_k) = 0, \tag{4}$$

where $\gamma$ varies continuously and $\boldsymbol{\xi}$ is the eigenvector belonging to $\gamma$. The three sheets are commonly thought of as corresponding to one quasi-compressional and two quasi-shear waves. $\mathcal{H}$ denotes a Hamiltonian that generates the ray-tracing equations (Kendall et al., 1992).

The scalar amplitudes $A$ must satisfy the transport equation

$$\partial_j\left(c_{ijk\ell}\xi_i\xi_k(A)^2\partial_\ell\tau\right) = 0 \tag{5}$$

for each mode of propagation.

For each mode the characteristic or group velocity $v$ is normal to $\mathcal{A}(x)$ at $\gamma$ and satisfies

$$v \cdot \gamma = 1; \qquad v = \left.\frac{\nabla_\gamma\mathcal{H}}{\gamma \cdot \nabla_\gamma\mathcal{H}}\right|_{\mathcal{H}=0}. \tag{6}$$

The normal or phase speed is given by

$$V = \frac{1}{|\gamma|}. \tag{7}$$

**Table 1. Glossary of symbols: ray geometry.**

| Symbol | In or near equation number | Meaning |
|---|---|---|
| $s$ | (11) | Source position |
| $x, y$ | (11), (26) | Scattering point, image point |
| $r$ | (11) | Receiver position |
| $\tau$ | (1) | One-way traveltime |
| $A$ | (5) | Scalar amplitude |
| $\gamma$ | (1) | Local slowness vector |
| $V$ | (7) | Local phase velocity |
| $\alpha$ | (8) | Local phase direction |
| $\mathcal{A}$ | (3) | Local slowness surface |
| $v$ | (6) | Local group velocity |
| $\chi$ | (9) | Angle between group and phase velocities |
| $\Sigma$ | (16) | Wave front surface |
| $\mathcal{M}$ | (10) | Amplitude Jacobian |
| $\boldsymbol{\xi}$ | (2) | Local polarization vector |
| $\mathcal{H}$ | (4) | Hamiltonian |
| $\sigma$ | (10) | KMAH index |
| $H$ | (10) | Hilbert transform |
| $T$ | (15) | Two-way traveltime |
| $\Gamma$ | (16) | Gradient of two-way time |
| $\Theta$ | (35) | Migration wave vector |
| $\nu$ | (19) | Migration dip, isochron normal |
| $\theta$ | (20) | Scattering angle |
| $\psi$ | (20) | Azimuth |

**Table 2. Glossary of symbols: GRT.**

| Symbol | In or near equation number | Meaning |
|---|---|---|
| $\partial S$ | (17) | Surface of source locations |
| $\partial R$ | (17) | Surface of receiver locations |
| $\{\tilde{\beta}, \hat{\beta}\}$ | (45) | Normals to $\{\partial S, \partial R\}$ |
| $N_s$ | (58) | Number of sources |
| $N_r$ | (58) | Number of receivers |
| $N$ | (58) | Number of measurements |
| $N$ | (17) | Quasi midpoint |
| $\Omega$ | (17) | Quasi half offset |
| $u$ | (22) | Scattered displacement filed |
| $\sigma_u$ | (50) | Data covariance |
| $\mathcal{D}$ | (17) | Scattering domain, support of medium perturbation |
| $\mathbf{c}^{(1)}$ | (22) | Perturbation in density and stiffness |
| $\sigma_c$ | (50) | Medium covariance |
| $\mathbf{P}$ | (52) | Medium parameter projection |
| $\boldsymbol{\delta}^{(1)}$ | (52) | Perturbation in generic medium parameters |
| $\mathbf{w}$ | (22) | Contrast-source radiation patterns |
| $\mathcal{I}$ | (23) | Weight in the forward GRT |
| $\Lambda, \Lambda^a$ | (25), (30) | Normal matrix of radiation patterns |
| $\mathcal{J}, \mathcal{J}_a$ | (27) | Weight in the GRT inversion |
| $h, h_a$ | (28), (33) | Obliquity |
| $E_\nu$ | (56) | Support in dip |
| $E_\theta$ | (25) | Support in scattering angle for fixed dip |
| $E_\psi$ | (25) | Support in azimuth for fixed dip and scattering angle |
| $E_\alpha$ | (51) | Support in phase directions for fixed dip |
| $C$ | (57) | Aperture normalization factor |
| $\mathcal{R}_C$ | (56) | Resolution kernel |

The unit phase direction follows as

$$\alpha = V\gamma. \tag{8}$$

From equation (6), it then follows that

$$V = |v| \cos \chi, \tag{9}$$

where $\chi$ is the angle between $v$ and $\gamma$.

Ray-theoretical displacement amplitudes, satisfying the transport equation and originating from a point force source (a vibrator) at $x'$, can be written in terms of a Jacobian $\mathcal{M}$ describing the geometrical spreading along a ray:

$$A = \frac{1}{4\pi[\rho(x)\rho(x')\mathcal{M}]^{1/2}}$$

$$\text{with} \quad \mathcal{M} = \frac{|v(x')||V(x)|\left|\dfrac{\partial x}{\partial q_1} \wedge \dfrac{\partial x}{\partial q_2}\right|_x}{\left|\dfrac{\partial \gamma}{\partial q_1} \wedge \dfrac{\partial \gamma}{\partial q_2}\right|_{x'}}. \tag{10}$$

Here, $(q_1, q_2)$ parametrize the rays originating from the source and can be chosen to lie on the unit sphere $(S^2)$, centered at $x'$. In the presence of caustics, in the Green's tensor, the scalar amplitude will carry a phase shift $\pi/2$ to the power of the KMAH index $\sigma(x; x'; \gamma(x'))$ inducing the contribution from a Hilbert transform H. Away from any intrinsic caustic associated with the anisotropy of the medium, the index characterizes the local curvature of the slowness surface sheet, where the index is 0 if the surface is convex, 1 if the surface is saddle shaped, and 2 if the surface is concave.

## Source and receiver rays

To describe rays from a point in the subsurface to sources and receivers, we introduce a notation in which $\tilde{\ell}$ refers to any quantity $\ell$ associated with the source, and $\hat{\ell}$ refers to the same quantity associated with the receiver. Thus, the geometrical amplitudes to a source at $s$ and a receiver at $r$ will be written as

$$\tilde{A}(x) = A(x, s), \qquad \hat{A}(x) = A(r, x). \tag{11}$$

According to equation (1), the slowness vectors at $x$ are given by

$$\tilde{\gamma}(x) = \nabla_x \tau(x, s), \qquad \hat{\gamma}(x) = \nabla_x \tau(r, x), \tag{12}$$

the associated phase directions are given by

$$\tilde{\alpha} = \frac{\tilde{\gamma}}{|\tilde{\gamma}|}, \qquad \hat{\alpha} = \frac{\hat{\gamma}}{|\hat{\gamma}|}, \tag{13}$$

and the directional phase speeds [cf. equation (7)] are given by

$$\tilde{V} = \frac{1}{|\tilde{\gamma}|}, \qquad \hat{V} = \frac{1}{|\hat{\gamma}|}. \tag{14}$$

We also define the two-way traveltime $T$ and its gradient,

$$T(r, x, s) \equiv \tau(x, s) + \tau(r, x), \tag{15}$$

$$\Gamma(r, x, s) \equiv \nabla_x T(r, x, s).$$

From equation (12), we see that

$$\Gamma(r, x, s) = \tilde{\gamma}(x) + \hat{\gamma}(x). \tag{16}$$

The ray-geometrical quantities are illustrated in Figure 1: $\{x \mid T(r, x, s) = t\}$ is an isochron, $\Sigma(r, t) = \{x \mid \tau(r, x) = t\}$ is a wavefront, where $t$ is time, whereas $N$ and $M$ denote the modes propagation ($\tilde{N}$ for the ray from source to scattering point and $\hat{M}$ for the ray from scattering point to receiver).

## Acquisition surface parametrization

The GRT inversion formulas derived below involve integration over double spheres surrounding the image point. In practice, however, integration is carried out over surfaces associated with the acquisition geometry. Hence, Jacobians for the transformation of coordinates between the two domains must be calculated. We shall show in a subsequent section how these Jacobians involve ray-geometrical quantities.

Scattering data are generated for source and receiver locations, $s$ and $r$, lying on surfaces (open patches) $\partial S$ and $\partial R$, ideally surrounding a domain $\mathcal{D} \subset \mathbb{R}^3$. To parametrize the domain $\partial S \times \partial R$, we employ the vectors $(N, \Omega), N \in \mathbb{R}^2, \Omega \in \mathbb{R}^2$. Then, with $S^2$ denoting the unit sphere in 3-D,

$$s = s(N, \Omega) \in \partial S \sim S^2, \qquad r = r(N, \Omega) \in \partial R \sim S^2. \tag{17}$$

[In fact $(N, \Omega)$ define a chart of $\partial S \times \partial R$; we need at least two charts, for both sources and receivers, to cover the full spheres.] The parameter vector $\Omega$ is referred to as the index of the data. A subset of data with common index is a gather. As we will show, however, we will not have to sort data into gathers.
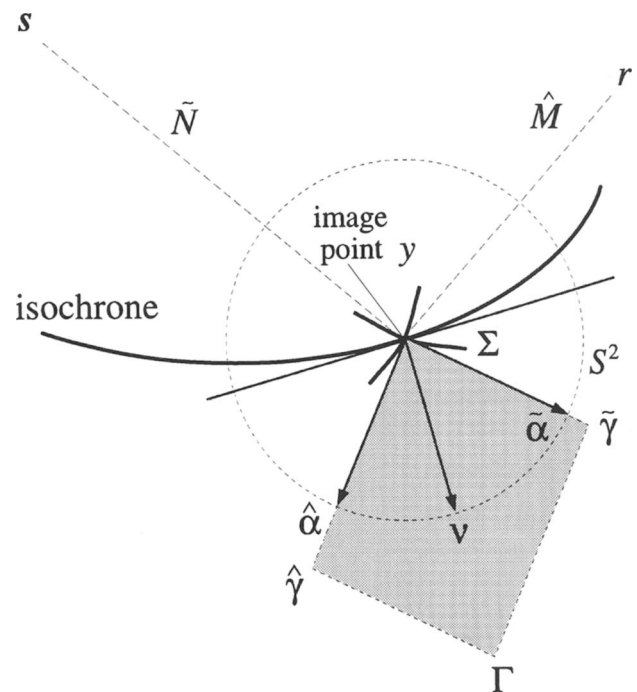


FIG. 1. The source-receiver ray geometry. Here, $N, M$ refer to the modes ($qP$ or $qS$) of propagation.

As an example, let $s$ and $r$ lie in a plane. Then indexing in half-offset with $\Omega = (1/2)(r - s)$ fixed (common) yields

$$s = N - \Omega, \qquad r = N + \Omega, \qquad (18)$$

where $N = (1/2)(s + r)$ denotes the midpoint. On the other hand, common receiver gathers induce $s = N, r = \Omega$, whereas common source gathers correspond to $s = \Omega, r = N$. In the ideal situation, two charts $(N)$ define a domain $\partial N \sim S^2$ (through stereographic projection) for any $\Omega$ fixed (common). All these indexings are independent of the elastic properties of the subsurface.

However, for the purpose of GRT inversion, the natural indexing will be image-point dependent. We will discuss the indexing in scattering-angle/azimuth. Let $\nu$ denote the migration dip at image point $x \in \mathcal{D}$, i.e.,

$$\nu(r, x, s) = |\Gamma|^{-1} \Gamma; \qquad (19)$$

the scattering angle $\theta$ and azimuth $\psi$ are then defined through

$$\cos \theta = \tilde{\alpha} \cdot \hat{\alpha}, \qquad \psi = \text{third Euler angle around } \nu. \qquad (20)$$

Note that $\nu \in S^2$ and $(\theta, \psi) \in S^2$. Then $\nu$ replaces the role of $N$ and $(\theta, \psi)$ replaces the role of $\Omega$. Our key constraint is that $N$ must map $\nu$ uniquely on the sphere $S^2$ for any fixed $\Omega$.

The various transformations on $S^2 \times S^2$ are summarized in Table 3.

## GENERALIZED RADON TRANSFORMATIONS

In the remainder of the paper, we exclude the occurrence of rays originating in the scattering domain, traveling in the background medium, and grazing at $\partial R$ or $\partial S$. Also, we assume that $\partial S$ and $\partial R$ cannot be connected by a single ray traveling through the scattering domain (then the migration dip would not be defined). In principle, multipathing in the background can be accounted for, in which case the integrations over dip and scattering angles are inseparable (D.-J. Smit, personal communication, 1996).

First, we will write the direct scattering problem in the form of a GRT. Second, for $y$ in the neighborhood of $x$, using Beylkin's (1985) analysis of this transform,

$$\int_{S^2} [1 + \mathcal{O}(|x - y|)] \delta''(\nu \cdot (y - x) + \mathcal{O}(|x - y|^2)) \, d\nu$$
$$= -8\pi^2 \delta(y - x) + \text{smoother terms}, \qquad (21)$$

where $\mathcal{O}(|x - y|)$ and $\mathcal{O}(|x - y|^2)$ as $|x - y| \to 0$ may depend on $\nu$, we will derive the GRT inversion.

### The forward transform

The scattered field $u$ due to a contrast in medium parameters $c^{(1)}$ with bounded support $\mathcal{D} \subset \mathbb{R}^3$, in the ray-Born approxi-

mation, can be written in the form of a GRT:

$$u(t, r, s) \simeq -\int_{\mathcal{D}} \delta''(t - T(r, x, s))$$
$$\times (w(r, x, s))^T c^{(1)}(x) \mathcal{I}(r, x, s) \, dx, \qquad (22)$$

where

$$\mathcal{I}(r, x, s) = \rho(x) \hat{A}(x) \tilde{A}(x). \qquad (23)$$

Here, $\rho$ is the density of the background medium, $\tilde{A}$ is the (point-source) amplitude along the ray connected to the source location, $\hat{A}$ is the (point-source) amplitude along the ray connected to the receiver location, $T(r, x, s)$ is the traveltime in the background medium along the ray connecting the source with the receiver via the scattering point $x$, and $w$ represents the radiation patterns of point contrast sources at the scattering point in accordance with the stiffness-density parametrization $c$. The radiation patterns are symmetrized dyadic products of the four polarization and slowness vectors at the scattering point, associated with the rays to the source and to the receiver [Burridge et al., 1998, equation (3.30)]. In our notation, we have suppressed the occurrence of the Hilbert transforms. We have arranged the tensors $w$ and $c$ in arrays. The integral in equation (22) is taken over isochrons. We freely identify

$$T(x, N, \Omega) = T(r, x, s), \qquad w(x, N, \Omega) = w(r, x, s),$$

or

$$T(x, \tilde{\alpha}, \hat{\alpha}) = T(r, x, s), \qquad w(x, \tilde{\alpha}, \hat{\alpha}) = w(r, x, s),$$

through the respective coordinate transformations of Table 3.

### The inverse transform over phase directions at the image point

The structure of the dual GRT transform (Beylkin, 1985), associated with equation (22), follows as

$$\langle c^{(1)} \rangle(x, \Omega) \simeq \frac{1}{8\pi^2} \int_{\partial N} \mathcal{J}(r, x, s) [\Lambda_x(\nu(r, x, s))]^{-1}$$
$$\times wu(T(r, x, s), r, s) \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(N, \Omega)} \bigg|_x \, dN, \qquad (24)$$

where $\nu$ is the unit vector in the direction of $\nabla_x T$ [equation (19)], $\hat{\alpha}$ is the normalized slowness vector associated with the ray connected to the receiver, $\tilde{\alpha}$ is the normalized slowness vector associated with the ray connected to the source, and

$$\Lambda(\nu) = \int_{E_\theta(\nu)} \int_{E_\psi(\nu, \theta)} (ww^T) \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \bigg|_. \, d\psi \, d\theta \qquad (25)$$

unravels the radiation patterns at the image point ($E$ denotes support). The hypersurface

$$\{(\Omega, t) \in \partial\Omega \times \mathbb{R}_{\geq 0} \mid t = T(r(N, \Omega), x, s(N, \Omega))\}$$

is the so-called diffraction surface, and equation (24) describes nothing other than the diffraction stack with weights $\mathcal{J}$, which are determined below.

**Table 3.    Coordinate transformations in the GRT.**

| | |
|---|---|
| $(\tilde{\alpha}, \hat{\alpha})_x \to (\nu, (\theta, \psi))_x$ | $(\nu, (\theta, \psi))_x \to (s, r)$ |
| $(\tilde{\alpha}, \hat{\alpha})_x \to (s, r)$ | $(s, r) \to (N, \Omega)$ |

The inverse GRT (i.e. $\mathcal{J}$) follows from composing the forward with the dual transforms. We then obtain the condition, at stationarity,

$$\left[ -\frac{1}{8\pi^2} \int_{\partial N} \delta''(T(r,y,s) - T(r,x,s))\mathcal{J}(r,y,s)\mathcal{I}(r,x,s) \right.$$
$$\left. \times \left. \frac{\partial(\nu,\theta,\psi)}{\partial(N,\Omega)} \right|_y dN \right] d\Omega = \delta(y-x)\,d\theta\,d\psi$$

$$+ \text{smoother terms.} \tag{26}$$

Using the homogeneity of $\delta''$ and the plane-wave expansion of the Dirac distribution equation (21), we find that

$$\mathcal{J}(r,x,s) = \frac{h(x,\nu(r,x,s))}{\mathcal{I}(r,x,s)}, \tag{27}$$

with

$$h = |\Gamma|^3. \tag{28}$$

In the presence of caustics, the data $u$ in equation (24) need be replaced by $u + i\mathrm{H}u$ and the real part of the integral has to be taken (details can be found in de Hoop and Brandsberg-Dahl, 1998). Note that $h$ acts as a natural taper on the measurements for large scattering angles. The AVA inversion amounts now to integrating $\langle c^{(1)} \rangle(x,\Omega)$ over $\Omega$ to yield the elastic parameters, but note that the inverse $[\Lambda_.(\nu)]^{-1}$ is essentially nested in the integration over $N$ or $\nu$ (the migration).

Carrying out the AVA inversion simultaneously with the migration allows the integration in equation (24) followed by the integration in equation (25) both to be directly carried out over $(\tilde{\alpha}, \hat{\alpha})$ with volume form $d\tilde{\alpha}\,d\hat{\alpha}$; note that $s$ and $r$ then follow from the intersection of the rays with $\partial S$ and $\partial R$, respectively.

### The inverse transform over acquisition parameters at the surface

It is possible to reformulate the inverse problem using the coordinates naturally arising in the acquisition geometry, which is the conventional approach to GRT inversion. We will distinguish this representation of the GRT from the previous one by using super- and subscripts $a$. Then, the structure of the dual GRT transform follows as

$$\langle c^{(1)} \rangle(x,\Omega) \simeq \frac{1}{8\pi^2} \int_{\partial N} \mathcal{J}_a(r,x,s)[\Lambda_x^a(\nu(r,x,s))]^{-1}$$
$$\times \mathrm{w}u(T(r,x,s),r,s)\,dN, \tag{29}$$

with ($\nu$ at $x$ maps onto $N$)

$$\Lambda_.^a(\nu) = \int_{E_\Omega(N)} (\mathrm{w}\mathrm{w}^T) \bigg|\,d\Omega \tag{30}$$

$((\theta,\psi)$ at $x$ for $\nu$ fixed map onto $\Omega$). Matching the condition [the counterpart of equation (26)], at stationarity,

$$\left[ -\frac{1}{8\pi^2} \int_{\partial N} \delta''(T(r,y,s) - T(r,x,s))\mathcal{J}_a(r,y,s) \right.$$
$$\left. \times \mathcal{I}(r,x,s)\,dN \right] d\Omega = \delta(y-x)\,d\Omega + \text{smoother terms,} \tag{31}$$

using plane-wave expansion equation (21) as before, we find that

$$\mathcal{J}_a(r,x,s) = \frac{h_a(x,\nu(r,x,s))}{\mathcal{I}(r,x,s)}, \tag{32}$$

where now

$$h_a = |\Gamma|^3 \left( \frac{\partial(\nu)}{\partial(N)} \right)_\Omega. \tag{33}$$

If we allow $(N,\Omega)$ to be $x$ dependent, we can set $N = \nu$ and $\Omega = (\theta, \psi)$. Then, equation (33) reduces to equation (28). This corresponds precisely to indexing the measurements in common scattering-angle/azimuth gathers (which varies with image point).

The integration over $\nu$ or $N$ should produce the same image (singular support of the perturbation $c^{(1)}$) for each pair $(\theta, \psi)$ or $\Omega$. This redundancy, comparing those images, can be used to improve the background velocity model by the method of differential semblance (or coherency) optimization (Carazzone and Symes, 1991; Symes and Kern, 1994), residual curvature analysis (Liu, 1995), or by maximizing a similarity index (Chavent & Jacewitz, 1995). Such an improvement captures part of the truly nonlinear aspects of the seismic inverse problem.

### Fourier analysis

To link the GRT approach to inversion/migration with Fourier "$f$-$k$" ($\omega = 2\pi f$) migration, we use the one-sided Fourier representation of the Dirac distribution to rewrite condition (26):

$$\frac{1}{8\pi^3}\left[ \int_{\partial N} \int_{\mathbb{R}_{\geq 0}} \exp[i\omega(T(r,y,s) - T(r,x,s))]\mathcal{J}(r,y,s) \right.$$
$$\left. \times \mathcal{I}(r,x,s)\frac{\partial(\nu,\theta,\psi)}{\partial(N,\Omega)} \bigg|_y \omega^2\,d\omega\,dN \right] d\Omega = \delta(y-x)\,d\theta\,d\psi$$

$$+ \text{smoother terms.} \tag{34}$$

We introduce the wave vector

$$\Theta = \omega\Gamma \in \mathbb{R}^3 \text{ at } x, \tag{35}$$

since, in the high-frequency approximation, the phase in equation (34) can be approximated according to

$$\omega(T(r,y,s) - T(r,x,s)) \simeq \Theta \cdot (y-x) \tag{36}$$

up to leading order. We carry out the mapping

$$(\omega, \nu) \to \Theta;$$

the inverse mapping

$$\omega(\Theta) = \frac{\Theta \cdot \Gamma}{|\Gamma|^2}$$

also appears in Stolt's formulation (Stolt, 1978). Then equation (34) can be written in the form

$$\frac{1}{8\pi^3} \int_{\mathbb{R}^3} \exp[i\Theta \cdot (y - x)] \, \mathcal{J}(r, y, s)\mathcal{I}(r, x, s)$$

$$\times \omega^2 \frac{\partial(\omega, \nu)}{\partial(\Theta)} \bigg|_y \, d\Theta$$

$$= \delta(y - x) + \text{smoother terms.} \tag{37}$$

It now follows from the Fourier representation of the Dirac distribution that

$$\mathcal{J}(r, x, s) = \frac{1}{\mathcal{I}(r, x, s)} \frac{1}{\omega^2} \frac{\partial(\Theta)}{\partial(\omega, \nu)} \bigg|_x. \tag{38}$$

Note, however, that

$$\frac{1}{\omega^2} \frac{\partial(\Theta)}{\partial(\omega, \nu)} = \left| \det\begin{pmatrix} \Gamma & \partial_{\nu_1}\Gamma & \partial_{\nu_2}\Gamma \end{pmatrix} \right| = h \tag{39}$$

yields the previous result [equation (28)]. Condition (31) results, in a likewise manner, in

$$\mathcal{J}_a(r, x, s) = \frac{1}{\mathcal{I}(r, x, s)} \frac{1}{\omega^2} \frac{\partial(\Theta)}{\partial(\omega, N)} \bigg|_x. \tag{40}$$

Here,

$$\frac{1}{\omega^2} \frac{\partial(\Theta)}{\partial(\omega, N)} = \left| \det\begin{pmatrix} \Gamma & \partial_{N_1}\Gamma & \partial_{N_2}\Gamma \end{pmatrix} \right| = h \left( \frac{\partial(\nu)}{\partial(N)} \right)_\Omega, \tag{41}$$

which is Beylkin's original determinant (Beylkin, 1985). For the computation of this determinant, note that in general, regular sampling in $\nu$ will cause irregular sampling in $N$ and vice versa.

## JACOBIANS

### Transformation from phase directions to source-receiver coordinates

We have

$$\frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(N, \Omega)} \bigg|_y = \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(s, r)} \bigg|_y \frac{\partial(s, r)}{\partial(N, \Omega)}. \tag{42}$$

The Jacobian, with factorization ($\tilde{\alpha}$ does not depend on $r$ and $\hat{\alpha}$ does not depend on $s$)

$$\frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(s, r)} \bigg|_y = \frac{\partial(\tilde{\alpha})}{\partial(s)} \bigg|_y \frac{\partial(\hat{\alpha})}{\partial(r)} \bigg|_y, \tag{43}$$

is directly related to dynamic ray theory. In general, the factors can be expressed in terms of the dynamic ray amplitudes because, like the amplitudes, they follow from a variation of the anisotropic ray tracing equations [see equation (10)]. For the source side,

$$\frac{\partial(s^\Sigma)}{\partial(\tilde{\alpha})} \bigg|_y = \frac{1}{16\pi^2 \rho(s)\rho(y)\tilde{V}(s)(\tilde{V}(y))^3(\tilde{A}(y))^2} \tag{44}$$

as long as $\partial S$ in the neighborhood of $s$ coincides with the wave front $\Sigma(y, \tau(s, y))$ originating at $y$. If this is not the case, we

have to correct for the ratio of the area on $\partial S$ to the area on the wave front $\Sigma(y, \tau(s, y))$ at $s$ onto which it is mapped by projection along the rays. This arises from the fact that $\partial S$ is not necessarily tangent to $\Sigma(y, \tau(s, y))$ at $s$. It amounts to dividing the previous Jacobian by the Jacobian

$$\frac{\partial(s^\Sigma)}{\partial(s)} = (\tilde{\alpha}(s) \cdot \tilde{\beta}(s)), \tag{45}$$

where $s^\Sigma$ denotes the coordinates on $\Sigma(y, \tau(s, y))$ intersecting $\partial S$ at $s$, and $\tilde{\beta}(s)$ = normal to $\partial S$ at source. Note that $\tilde{\alpha}(s)$ is the normal to the wavefront at $s$. Similar expressions hold for the receiver side.

### Transformation from phase directions to dip, scattering angle, and azimuth

Under the transformation from phase directions to dip, scattering angle, and azimuth, the volume form on $S^2 \times S^2$ [cf. equation (25)] transforms according to

$$\frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} = \frac{\sin\theta}{1 + (|\tilde{\gamma}||\hat{\gamma}|/|\Gamma|^2)(\tan\hat{\chi} - \tan\tilde{\chi})\sin\theta}, \tag{46}$$

where

$$\Gamma = \nabla_x T, \quad \cos\tilde{\chi} = \tilde{n}_\parallel \cdot \tilde{\alpha}, \quad \text{and} \quad \cos\hat{\chi} = \hat{n}_\parallel \cdot \hat{\alpha} \tag{47}$$

(see Appendix A). Here, $\tilde{n}_\parallel$ and $\hat{n}_\parallel$ denote the normals to the slowness surface at the scattering point projected in the azimuth plane $\psi$.

### Transformation from dip to midpoint at fixed half-offset

With equation (41), the Jacobian associated with the mapping $\nu \to N$ can be expressed in terms of

$$\det\begin{pmatrix} \Gamma & \partial_{N_1}\Gamma & \partial_{N_2}\Gamma \end{pmatrix} = \Gamma \cdot (\partial_{N_1}\Gamma \wedge \partial_{N_2}\Gamma), \tag{48}$$

in which

$$\partial_{N_{1,2}}\Gamma = (\partial_{N_{1,2}}s) \cdot \nabla_s\tilde{\gamma} + (\partial_{N_{1,2}}r) \cdot \nabla_r\hat{\gamma}. \tag{49}$$

In principle, $\nabla_s\tilde{\gamma}$ and $\nabla_r\hat{\gamma}$ can be computed by perturbing the Hamilton system for ray tracing. For a homogeneous, isotropic medium, the evaluation of this Jacobian is reviewed in Appendix B.

## REGULARIZATION

The inversion of the $\Lambda$ matrix in general requires a regularization because some of its singular values may become very small or even vanish (Campbell and Meyer, 1979). We consider two approaches in our analysis. The first one is based on a singular value decomposition in the inversion process and can be interpreted in terms of Bayesian statistics; the second one is based on a straightforward parameter reduction.

In the Bayesian approach we introduce an a priori probability distribution of allowable models with covariance matrix $\sigma_c$. The prior model estimate is denoted as $c_{prior}^{(1)}$. Also, we introduce

a likelihood distribution in the space of measured displacements (scattered field) $u$ with covariance $\sigma_u(\theta, \psi)$, constrained to be diagonal. Assuming Gaussian statistics, equation (25) is modified to

$$\Lambda(\nu) = \int_{E_\theta(\nu)} \int_{E_\psi(\nu,\theta)} \left( \mathbf{w}\sigma_u^{-1}\mathbf{w}^T \right) \\ \times \left. \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \right| \, d\psi \, d\theta + \sigma_c^{-1}. \quad (50)$$

This matrix has the interpretation of reciprocal of the a posteriori multiparameter covariance matrix $\sigma_c'$ for dip $\nu$. Note that $\sigma_c^{-1}$ controls the shift of singular values. The square roots of the elements of the diagonal of $\Lambda(\nu)$ are the standard deviations of the solution to the AVA—amplitude versus scattering angles—inverse problem at dip $\nu$. Let the generalized inverse of the matrix $\Lambda(\nu)$ be denoted as $\langle[\Lambda(\nu)]^{-1}\rangle$, then the inversion formula equation (24) is subject to the modification

$$[\Lambda(\nu)]^{-1}\mathbf{w}u \rightarrow \langle[\Lambda(\nu)]^{-1}\rangle\left[\mathbf{w}\sigma_u^{-1}u \\ + \frac{1}{|E_\alpha(\nu)|}\sigma_c^{-1}\mathbf{c}_{\text{prior}}^{(1)}\right], \quad (51)$$

$$|E_\alpha(\nu)| \equiv \int_{E_\theta(\nu)} \int_{E_\psi(\nu,\theta)} \left. \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \right| \, d\psi \, d\theta.$$

In the absence of prior information, $\mathbf{c}_{\text{prior}}^{(1)} = \mathbf{0}$, and the parameter resolution matrix for dip $\nu$ is given by

$$\langle[\Lambda(\nu)]^{-1}\rangle\Lambda(\nu),$$

whereas the sensitivity matrix follows from the mapping of data covariances to a posteriori co-variance matrix $\sigma_c'$ for dip $\nu$,

$$(\sigma_c')(\nu) = \int_{E_\theta(\nu)} \int_{E_\psi(\nu,\theta)} \langle[\Lambda(\nu)]^{-1}\rangle\mathbf{w}(.)\sigma_u(\theta, \psi)(\mathbf{w}(.))^T \\ \times \langle[\Lambda(\nu)]^{-1}\rangle^T \left. \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \right| \, d\psi \, d\theta.$$

Naturally, $\sigma_u$ has to be estimated directly from the data in the common dip domain.

Parameter reduction, on the other hand, independent of statistical considerations, can be represented by a projection matrix $\mathbf{P}$. Let

$$\mathbf{c}^{(1)}(x) = \mathbf{P}^T \delta^{(1)}(x) \quad (52)$$

such that $\delta$ is contained in a $\mu$-dimensional parameter space, $\mu \leq 22$. The technique of reparametrization provides another tool to turn the inverse problem into a well-posed one. In all the equations, we simply have to replace $\mathbf{c}^{(1)}$ by $\delta^{(1)}$ and $\mathbf{w}$ by $\mathbf{Pw}$. Note that $\Lambda$ reduces to the $\mu \times \mu$ matrix

$$\Lambda(\nu) = \int_{E_\theta(\nu)} \int_{E_\psi(\nu,\theta)} \mathbf{Pw}(\mathbf{Pw})^T \left. \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \right| \, d\psi \, d\theta. \quad (53)$$

In this way, the inversion can be restricted to certain symmetries or predefined parameter combinations.

The exact relationship between $\delta$ and $\mathbf{c}$ may be nonlinear, such as the parametrization given by Thomsen (1986). As long as the medium perturbation is small or weak, the projection becomes a Jacobian

$$\mathbf{c}^{(1)}(x) = \left[ \frac{\partial(\mathbf{c})}{\partial(\delta_1, \ldots, \delta_\mu)} \right]_{\delta^{(0)}}^T \bigg|_x \left[ \delta_1^{(1)}(x), \ldots, \delta_\mu^{(1)}(x) \right]^T$$

and

$$\mathbf{P} = \left[ \frac{\partial(\mathbf{c})}{\partial(\delta_1, \ldots, \delta_\mu)} \right]_{\delta^{(0)}}, \quad (54)$$

where $\delta^{(0)}$ are the parameter values associated with the (known) background medium and $\delta^{(1)}$ are the parameter values associated with the (unknown) perturbation. Thus the single-scattering theory can be linked with rock physics, for example, with the aid of quasi-static differential effective medium theories. Such a representation is particularly useful if the perturbations are localized to a subwavelength scale. Then specific rock types can be mixed with variable concentrations to yield the medium perturbation. A formula along those lines of reasoning is presented in Appendix C.

## APERTURE NORMALIZATION

An important procedure when including amplitudes in data inversion is correcting for limited recording apertures. Different elastic parameters (or combinations of parameters) have different radiation patterns, and hence the effect of truncation will vary between parameters. Similarly, in the case of attempting to map a single parameter in space, any spatial variation in recording geometry will result in a variation in aperture and hence inversion results. These aperture effects will cause a degradation of the spatial resolution operator and have to be compensated for, at least on the operator's diagonal.

Backsubstituting the forward modeling equation (22) into the inversion formula (24), using the Fourier representation as in equation (37) and integrating over $\Omega$, leads to the resolution operator equation

$$\langle\mathbf{c}^{(1)}\rangle(y) = \int_{\partial\Omega} \langle\mathbf{c}^{(1)}\rangle(y, \Omega) \, d\Omega = \int_D \mathcal{R}_\mathcal{C}(y, x)\mathbf{c}^{(1)}(x) \, dx, \quad (55)$$

where the resolution kernel follows as

$$\mathcal{R}_\mathcal{C}(y, x) \simeq \text{Re} \int_{E_\nu} [\Lambda_y(\nu)]^{-1} \int_{E_\theta(\nu)} \int_{E_\psi(\nu,\theta)} \mathcal{C}^{-1} \\ \times \left\{ \int_{E_{|\Theta|}} \exp[i\Theta \cdot (y - x)]|\Theta|^2 \, d|\Theta| \right\} \quad (56) \\ \times \mathbf{w}(y)(\mathbf{w}(x))^T \left. \frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} \right|_y \, d\psi \, d\theta \, d\nu.$$

In the standard analysis, $E_\nu = S^2$, $E_{|\Theta|} = \mathbb{R}_{\geq 0}$, and $E_\theta \times E_\psi = S^2$, whereas $\mathcal{C} = 8\pi^3$. In the finite aperture analysis, we normalize with the volume of the spectral support instead:

$$\mathcal{C} = E_\nu \int_{E_{|\Theta|}(\theta,\psi)} |\Theta|^2 \, d|\Theta|. \quad (57)$$

In this normalization, the diagonal $\mathcal{R}_C(y, y) = I$. Note, however, that $C$ is a function of scattering angle and azimuth, and hence the shape of the kernel function will be affected. In the inversion algorithm, the normalization will be accounted for in the density $\mathcal{J}$ [cf. equation (27)], namely, by replacing $h$ with $(8\pi^3/C)h$.

## DISCRETIZATION

Finally, we discuss the discretization equations (24) and (25). As fundamental variables we will use the phase directions $(\tilde{\alpha}, \hat{\alpha})$ discretized on the double spheres $S^2 \times S^2$ according to quasi-Monte Carlo (de Hoop and Spencer, 1996) sampling $(\tilde{\alpha}_i, \hat{\alpha}_j)$. Note that for a fixed image point $x$, the phase directions define source and receiver positions $s_i = s(\tilde{\alpha}_i)$ and $r_j = r(\hat{\alpha}_j)$. Let $i \in \{1, \ldots, N_s\}$, $j \in \{1, \ldots, N_r\}$, $N = N_s + N_r$. The weighted diffraction stack in accordance with equation (52) then follows as

$$\langle \delta^{(1)} \rangle(x) \simeq \frac{2}{N} \sum_{i,j} \mathcal{J}(r_j, x, s_i)[\Lambda_x(\nu(r_j, x, s_i))]^{-1}$$

$$\times \mathbf{Pw}(x, \tilde{\alpha}_i, \hat{\alpha}_j)u(T(r_j, x, s_i), r_j, s_i). \quad (58)$$

[Note that the full solid angle is $4\pi$; thus, the average sampling interval on the $(\tilde{\alpha}, \hat{\alpha})$ double spheres is $(4\pi)^2/N$.] On the other hand,

$$\Lambda(\nu) \simeq \frac{4\pi}{N(\nu)} \sum_{i,j}' \mathbf{Pw}(x, \tilde{\alpha}_i, \hat{\alpha}_j)(\mathbf{Pw}(x, \tilde{\alpha}_i, \hat{\alpha}_j))^T.$$

$$(59)$$

Here $N(\nu)$ is the number of data points that contribute to the integral for each $\nu$. [Note that the full solid angle is $4\pi$; thus, the average sampling interval on the $(\theta, \psi)$ sphere is $4\pi/N(\nu)$.] Naturally, we have to introduce $\nu$ bins to make the quantity $N(\nu)$ numerically meaningful. To avoid any directional bias, we choose the vertices of bins distributed similarly to equally electrically charged particles on the sphere. The prime in the summation of equation (59) indicates that only $(\tilde{\alpha}_i, \hat{\alpha}_j)$ pairs contribute that fall in the $\nu$ bin. In this respect, none of the indexings introduced in the previous sections require any data sorting.

The density $\mathcal{J}$ contains the so-called "obliquity factor" $h$, encountered in any true-amplitude migration. This function is dependent on scattering angle, azimuth, and migration dip, and in most cases will need to be tabulated in advance. Also $\Lambda$ can be tabulated. Although in the worst situation, tabulation of $\Lambda$ could result in unmanageably large storage, in cases where recording geometry and background are only slowly varying, the tables become sparse enough.

## EXAMPLES

In this section, we present a number of examples each designed to illustrate one aspect of our inversion/migration method with a view to determining the resolving power of seismic amplitude data.
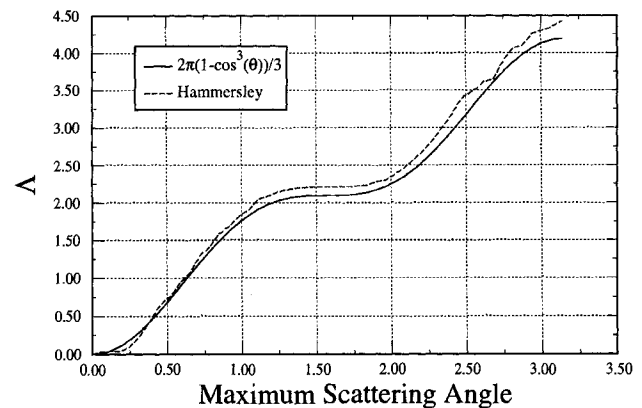
### The accuracy of the quadrature

We illustrate the accuracy of our summations or weighted diffraction stacks by computing $\Lambda$ and the diagonal of the

resolution kernel for a range of apertures. We note that the computer implementation of the theory given in the previous sections can be built around conventional Kirchhoff-type migration software. In the first example, we reconstruct density using a finite aperture dataset and then carry out the analytic normalization associated with the diagonal of the resolution kernel. Datasets with limited scattering angles ($\theta$) were drawn from a "complete" dataset with measurements taken at 10 000 Hammersley-point distributed source-receiver pairs (de Hoop and Spencer, 1996, who showed that Hammersley points produce accurate discretizations of the GRT). Each subset of data preserved a full range of scattering azimuths and migration dips, and involved at most 100 sources and 100 receivers. The results for the $\Lambda$ computation are shown in Figure 2 and for the full reconstruction of density in Table 4. We note that the estimates of density change by a factor of two as the scattering angle aperture is increased from 22° to 180°. The final column of Table 4 shows the results of correcting the inversions using the maximum scattering angle. We have been able to recover the "correct" result to within 1% at all apertures.

Table 4. Tests of limiting scattering angles to within the range $\theta \in [0, \theta_{max}]$. A full $\nu$ aperture was used with a point-density perturbation. The correct answer is 1.6025 × 10$^{-10}$ m. The second column shows numerical results from a limited aperture evaluation of equation (58). The third column shows the results of correcting column 2 with the analytic expression for aperture equation (57).

| $\theta_{max}$ | $\langle \rho^{(1)} \rangle/10^{-10}$ | $(8\pi^3/C)\langle \rho^{(1)} \rangle/10^{-10}$ |
|---|---|---|
| $\frac{8}{8}\pi$ | 1.594 | 1.594 |
| $\frac{6}{8}\pi$ | 2.308 | 1.584 |
| $\frac{5}{8}\pi$ | 2.864 | 1.588 |
| $\frac{4}{8}\pi$ | 2.977 | 1.586 |
| $\frac{3}{8}\pi$ | 3.044 | 1.587 |
| $\frac{2}{8}\pi$ | 3.282 | 1.589 |
| $\frac{1}{8}\pi$ | 3.518 | 1.585 |



FIG. 2. The theoretical values of $\Lambda$ for the case of a density inversion with a homogeneous background (solid line) compared with those calculated using Hammersley points (dashed line).

## The accuracy of the linearization assumption

The accuracy of the linearized scattering assumption will be analyzed by comparing the linearized (Born) plane-wave coefficients following from equation (22) (de Hoop and Bleistein, 1997) with the true plane-wave reflection coefficients for a planar, horizontal interface. First, we will consider the contributions to the linearized reflection coefficient from individual components of the stiffness tensor. In Figure 3, characteristic AVA patterns in VTI media (transverse isotropy with vertical axis of symmetry) due to 10% discontinuities in individual moduli calculated using the Born approximation are presented (the 3-direction is the vertical). The sensitivities of the different moduli with angle are rather different: for example, the response due to a perturbation in $C_{33}$ controlling the vertical $qP$ velocity shows up at small scattering angles, whereas a perturbation in $C_{11}$ controlling the horizontal $qP$ velocity becomes only measurable at large angles. Also, note that the scattering response due to a $C_{55}$ perturbation is minus twice the scattering response due to a $C_{13}$ perturbation. We will illustrate in the coming examples that the linearization assumption breaks down for large angles, and hence the determination of horizontal medium velocities across horizontal interfaces, for example, may be dubious.

Second, the linearized and full $qP$-$qP$ AVA responses for a shale/salt interface and a shale/sand interface are shown in Figure 4. The linearizations have been normalized so as to agree with the full reflection coefficient at normal incidence. The elastic moduli of the media are given in Table 5. In each case, the linearized version is an adequate approximation to the full plane-wave coefficient out to angles of between 45° and 60°. Observe that the Born approximation breaks down completely near the critical angle. Although it is difficult to

generalize from only a few examples, these angles would appear to be the maximum ones that are likely to be usable in linear GRT inversions. To go beyond them requires nonlinear extension of the theory (de Hoop and Bleistein, 1997).

### Parameter resolution and covariance

In this subsection, we consider a synthetic example that illustrates the usefulness of the eigenvector decomposition in determining properly resolved parameter combinations. This type of analysis, in the spirit of Backus and Gilbert (1968), is essential in answering questions about the information content of various datasets and the errors in the inversions due to amplitude errors in the data.

The problem chosen was that of recovering discontinuities in elastic parameters at a single horizontal interface using an isotropic background and $P$-$P$ scattered data. The data input to the inversion constitute a common image point, single $\nu$ gather in three dimensions. The source and receiver locations were chosen to give $\nu$ vertical, as was the interface normal.

**Table 5. Elastic moduli of the materials used for the reflection coefficient calculations in Figure 4. For simplicity, a density of 2300 kg m⁻³ was used throughout.**

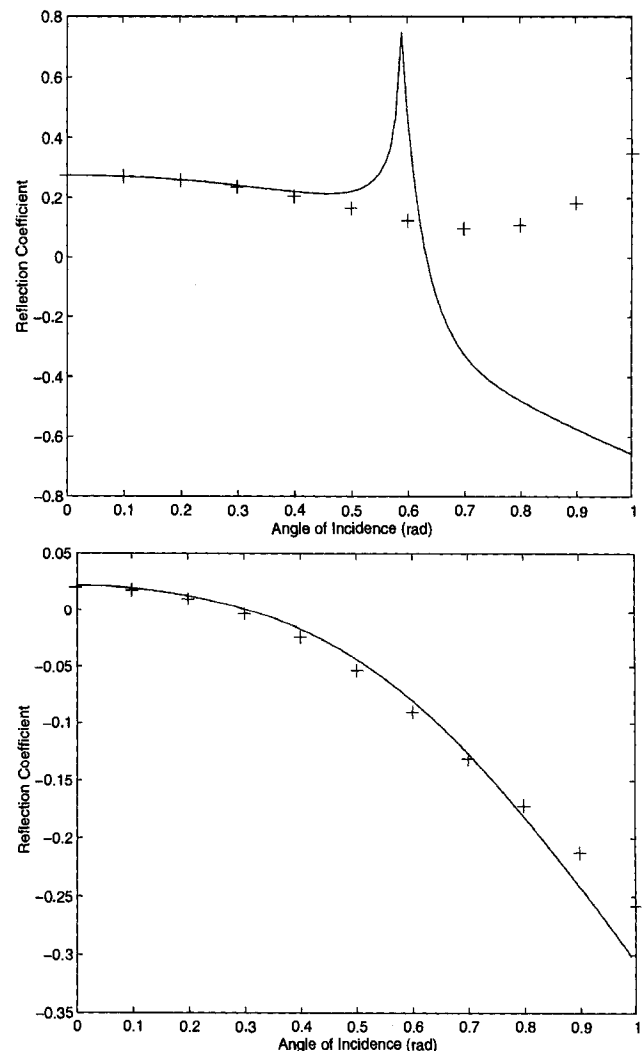| Stiffness (GPa) | Shale | Sand | Salt |
|---|---|---|---|
| $C_{11}$ | 16.58 | 14.20 | 40.01 |
| $C_{33}$ | 12.96 | 14.20 | 40.01 |
| $C_{55}$ | 2.26 | 5.38 | 13.29 |
| $C_{13}$ | 6.04 | 3.44 | 13.43 |
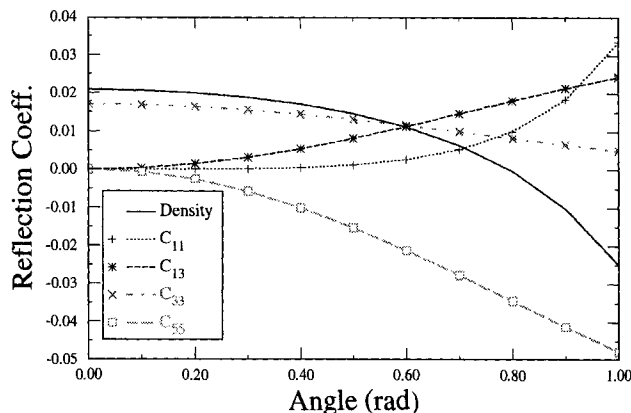


FIG. 3. Amplitude versus angle responses for individual moduli calculated using linearized plane-wave scattering theory.



FIG. 4. Linearized and full plane-wave reflection coefficients for a shale/salt (top) and a shale/sand (bottom) interface as a function of angle of incidence. The solid lines represent the full plane-wave reflection coefficients and the + symbols represent linearized estimates made using the Born approximation.

We examined properties of the $\Lambda$ radiation pattern matrix when all 22 parameters corresponding to elastic moduli and density are included in the inversion. The (singular value) spectrum of $\Lambda$ is shown in Figure 5a along with the diagonal elements of the covariance matrix in Figure 5b. The first point to note in Figure 5 is that the nine largest eigenvalues range over three orders of magnitude, following which eigenvalues drop dramatically. Of these nine largest eigenvalues the first four are significantly greater than the last five, which therefore are unlikely to be recovered in practice. The eigenvector corresponding to the largest eigenvalue contains a number of interesting features (Figure 6). In what follows, we call this the primary eigenvector. Notice first that it is nonzero only in those parameters that occur in symmetries up to orthorhombic, and that it is symmetric with respect to interchange of the horizontal 1 and 2 axes. $C_{12}^{(1)}$, $C_{66}^{(1)}$, and $C_{22}^{(1)}$ control the in-plane modes for propagation in the horizontal plane in the reconstructed medium. The other components of the primary parameter combination are the average of true properties in the 1 and 2 directions, and thus induce azimuthal symmetry. C. M. Sayers (personal communication, 1996) has noticed that apart from the inclusion of density, the primary parameter combination reflects perturbations to the bulk modulus of the medium. It may therefore be helpful in distinguishing between gas-and oil-saturated media, relevant in particular for time-lapse seismic experiments.

We remark that perturbations to $C_{13}$ and $C_{55}$ occur as the combination $(C_{13}^{(1)} - 2C_{55}^{(1)})$ and that $C_{23}^{(1)}$ and $C_{44}^{(1)}$ occur in similar proportions. The Born scattering coefficients (**w**) for $C_{55}^{(1)}$

and $C_{13}^{(1)}$ contain the products $\bar{\gamma}_3\hat{\gamma}_3\hat{\xi}_1\hat{\xi}_1$ and $\bar{\gamma}_3\hat{\gamma}_1\hat{\xi}_3\hat{\xi}_1$, respectively. In an isotropic background, these have the same angular dependencies and thus produce identical AVA responses for given dip. Even in anisotropic backgrounds, the differences in directions between $\gamma$ and $\xi$ for $qP$ waves are unlikely to be large enough to allow $C_{13}^{(1)}$ and $C_{55}^{(1)}$ to be determined separately. The differing weightings to the two moduli is affected by their multiplicity in the elasticity tensor.
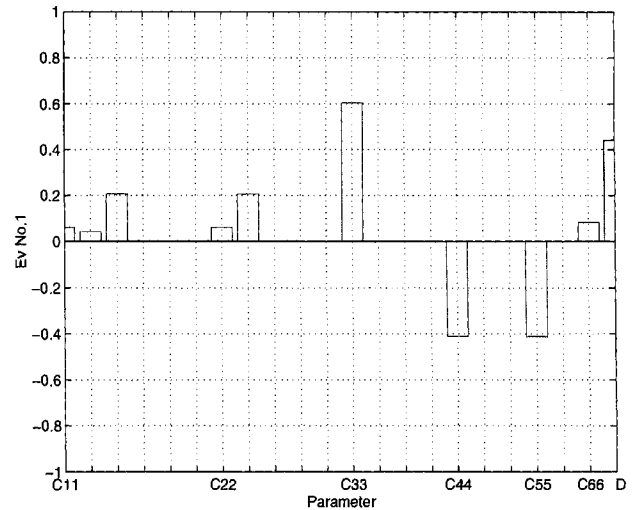


FIG. 6. The eigenvector corresponding to the largest eigenvalue of the 22 parameter system.
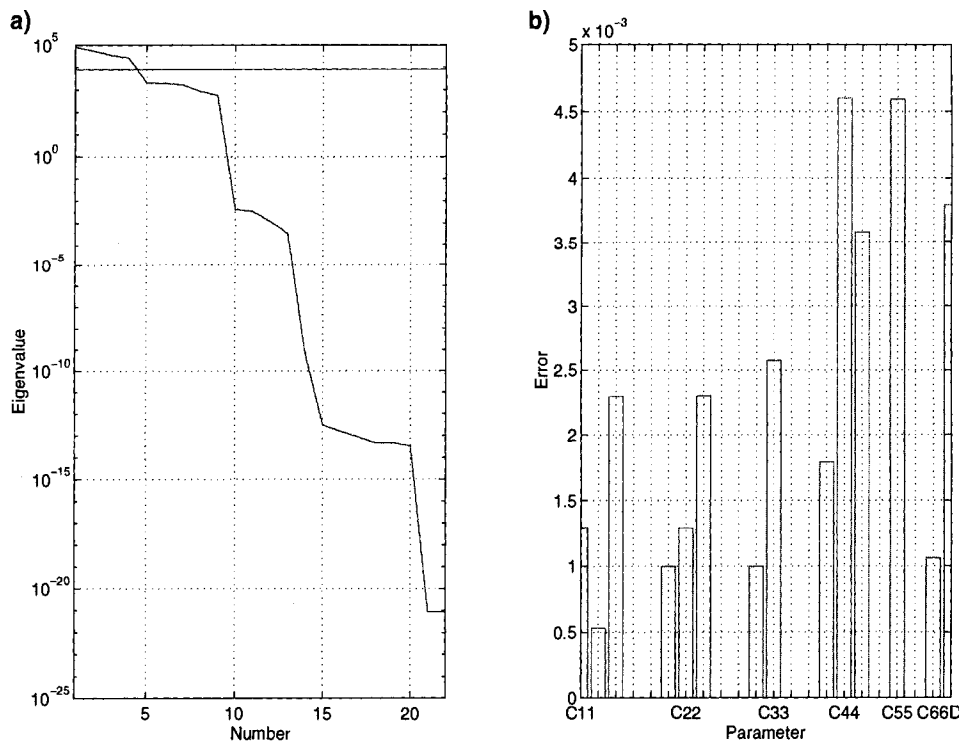


FIG. 5. The spectrum of $\Lambda$ (a) together with the diagonal elements of the covariance matrix (b) for the full 22 parameter inversion/migration. In this and subsequent figures elastic moduli are upper triangular ordered (i.e., $C_{11}, C_{12}, \ldots, C_{22}, \ldots, C_{66}$). Density (D) is the final parameter whenever it is used. The line in (a) indicates the truncation level used to generate the resolution matrix in Figure 7.

The primary eigenvector has a significant projection on the density axis, reflecting the fact that it is impedance rather than moduli contrasts which control the amplitude versus angle behavior. Figure 7 shows the resolution matrix formed by zeroing all eigenvalues with a magnitude of less than 0.1 times the primary eigenvalue. The leading diagonal quantifies the resolution of a parameter, and the trade-off between parameters can be determined from the individual rows (or columns) of this matrix, which has rank 4 only.

The diagonal elements of the covariance matrix shown in Figure 5b may be interpreted as relating the variance of input data to the variance in parameter estimates. Those parameters that show poor resolution show small variance since they are damped in the inversion. The relatively "free" parameter $C_{33}$ shows a covariance of approximately $10^{-2}$, indicating that a unit standard deviation in the input data maps to 10% standard deviation in the inversion estimate.

To simplify the study of eigenvectors other than the primary, we reduced the system to the nine moduli and density that define orthorhombic media. The spectrum of $\Lambda$ corresponding to this reduced system is shown in Figure 8, and the eigenvectors corresponding with the six most significant eigenvalues are presented in Figures 9 and 10. In the reduced system, there
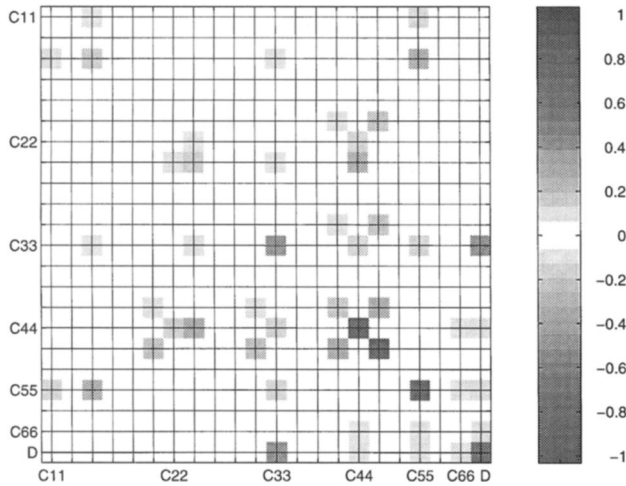


FIG. 7. The resolution matrix for the full 22-parameter inversion formed by zeroing all eigenvalues with a magnitude of less than 0.1 times the primary eigenvalue.
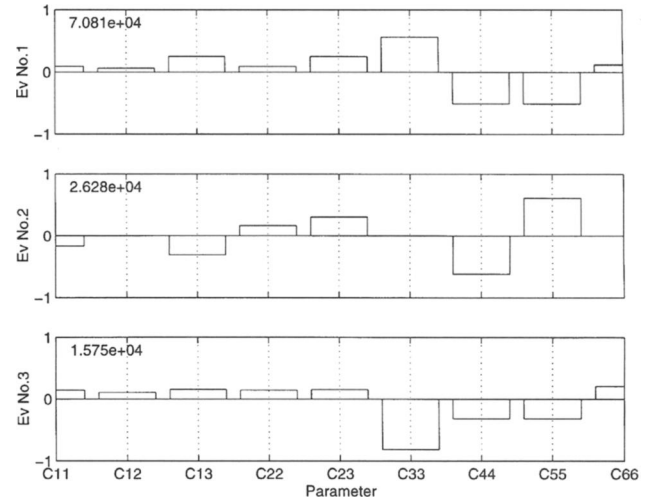


FIG. 9. The first three eigenvectors of the nine-parameter orthorhombic system. The eigenvalue corresponding to each eigenvector can be seen in the top left-hand corner of each eigenvector plot.
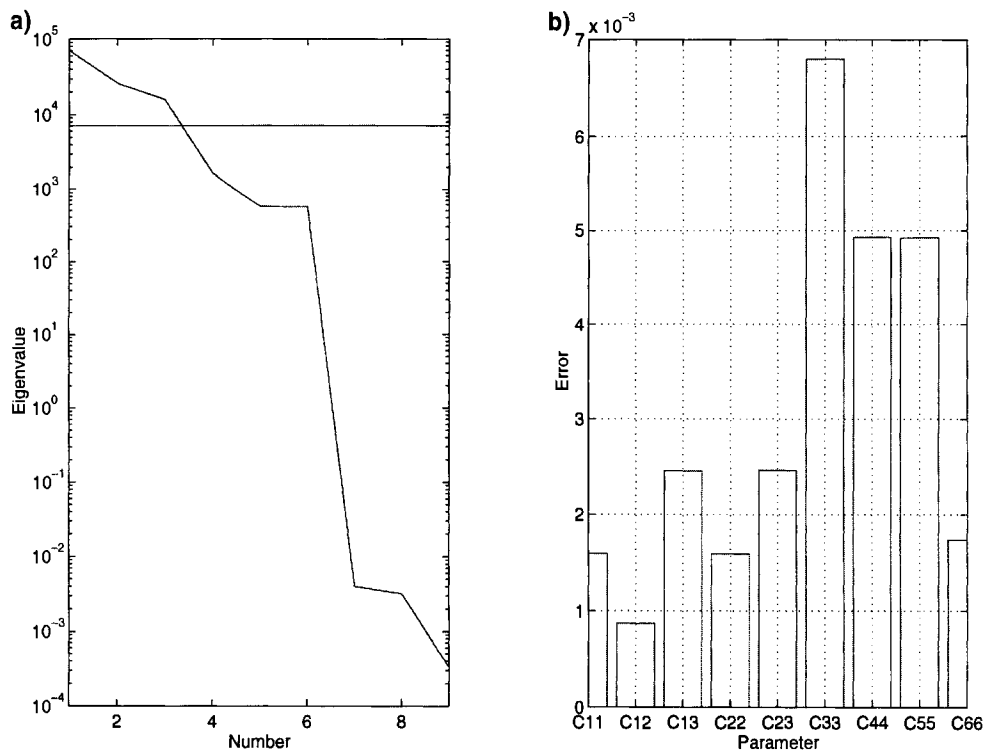


FIG. 8. The spectrum of $\Lambda$ (a) together with the diagonal elements of the covariance matrix (b) for the reduced nine parameter system.

are only six eigenvalues greater than $10^{-3}$ times the maximum eigenvalue.

The primary eigenvector in the reduced system is essentially the same as in the case of the full 22 parameter system (see Figure 6). The second eigenvector is interesting in that it is antisymmetric with respect to interchange of the 1 and 2 axes (Figure 9). Thus, linear combinations of the primary eigenvector, which is symmetric with respect to interchange of the 1 and 2 axes, and the second eigenvector serve to define a "best resolved parameter combination" and "its azimuthal variation."

The determination of the vertical $qP$ ($C_{33}$) discontinuity becomes possible when the third eigenvector is included. Eigenvector 3 (Figure 9) is similar to the first except that the sign of $C_{33}^{(1)}$ is reversed. Thus, the difference of eigenvector 1 and 3 isolates $C_{33}^{(1)}$, whereas the sum of 1 and 3 isolates the parameters $(C_{13}^{(1)} - 2C_{55}^{(1)})$ and $(C_{23}^{(1)} - 2C_{44}^{(1)})$. These, in turn, may be separated using eigenvector 2.

Eigenvalues 4, 5, and 6 are more than a factor of 10 smaller than eigenvector 3, but are interesting nevertheless. The eigenvector corresponding to the fourth largest eigenvalue is shown in Figure 10. It is the first to show that the $qP$ wave excites both shear polarizations ($C_{12}^{(1)}$, $C_{66}^{(1)}$) that takes place in the presence of general anisotropy. Here, we note that its affects on AVA are due to the combination ($C_{12}^{(1)} + 2C_{66}^{(1)}$). Eigenvectors 5 and 6 contain linear combinations of the two horizontal velocity parameters $C_{11}^{(1)}$ and $C_{22}^{(1)}$ and, taken together, isolate these two velocities individually.

One may wonder how the reconstructed linear parameter combinations relate to the nonlinear combinations appearing in the more conventional AVA analysis. The latter analysis provides an expansion with scattering angle/azimuth, the coefficients of which are particular combinations of moduli. Those combinations can be linearized in perturbations of stiffness, yielding vectors that span a linear subspace of medium parameters. Loosely speaking, this subspace appears to be reachable by our GRT inversion approach. Note in this respect that with

the GRT, we combine (integrate) reflection data over a range of scattering angles/azimuths.

We have carried out similar tests using $qP$ waves in anisotropic backgrounds and different dips. In principle, an isotropic background makes only a small difference. Its effects on the radiation pattern matrix are confined to a lack of parallelism of the polarization and slowness vectors plus changes to phase angle at the image point. The effect of anisotropy on $qS$ waves, though, will be large since their polarizations can vary more strongly.

Varying migration dip has two important effects. For a given sampling of the scattering angle sphere and assuming an isotropic background, the parameter combinations that are determined from a gather rotate with the migration dip $\nu$. As an example, in Figure 11 the resolution matrix is shown for a configuration in which the dip is 45° and which is recorded with angles of incidence of up to 45° with respect to $\nu$. Note how the problem becomes symmetrical with respect to interchange of the 1 and 3 axes while the 1–2 symmetries and antisymmetries discussed above are destroyed. The rank of the system is not changed by the rotation. A second effect of introducing migration dip is that, for a given recording geometry, the range of scattering angles reached is dip dependent. For example, in an isotropic background at a dip of 45° (in plane), the maximum scattering angle associated with a recording at the surface is 90°. Constraining the maximum recording angle at the image point to 60° the maximum scattering angle reduces to 30°. In this case, the rank of the system and hence resolution is seriously degraded. In Figure 12, the spectrum of $\Lambda$ using this restricted aperture is shown. For all practical purposes, the rank of the system is reduced to 1, and all that can be determined is an impedance associated with propagation at approximately 45°.

## Spatial resolution

In this subsection, we carry out more complete tests designed to show the antialiasing benefits of GRT and quasi–Monte Carlo sampling methods, and the improved spatial resolution that can be obtained at the expense of degrading parameter
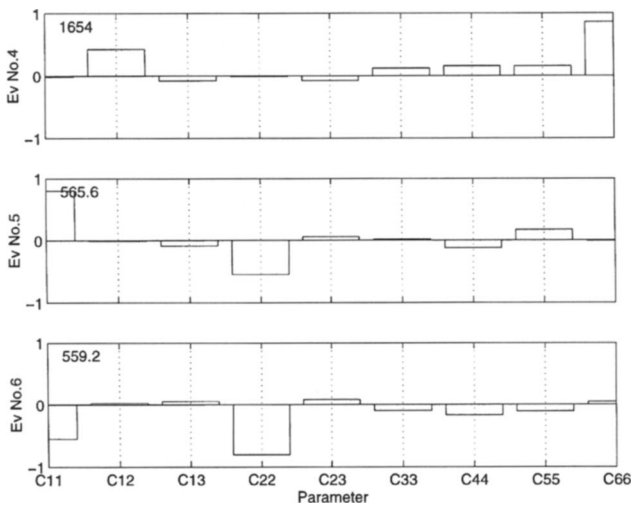


FIG. 10. The second 3 eigenvectors of the 9 parameter orthorhombic system. The eigenvalue corresponding to each eigenvector can be seen in the top left-hand corner of each eigenvector plot.
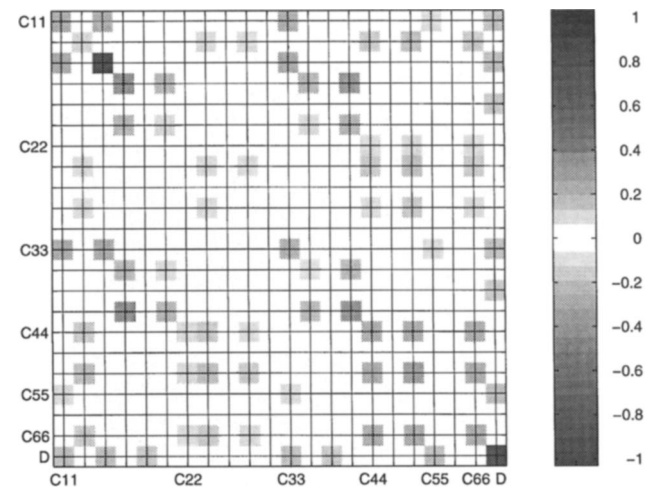


FIG. 11. The resolution matrix of the full 22-parameter system using a migration dip of 45° and incident angles of up of 45°. Compare Figure 7.

resolution. The tests are illustrated in Figures 13–16 and will be explained below. Comparisons of GRT inversions with Kirchhoff migrations in the isotropic case have been given previously. Dillon (1990) presents one such comparison and points out the antialiasing properties of the GRT operator. Similar arguments apply in the anisotropic case.

We considered a 3500-m recording aperture and calculated seismograms for a point perturbation to $C_{33}$ at a depth of 1000 m, and carried out the inversion. To illustrate the antialiasing effects of the GRT, we used a line, centered over the scatterer, of 251 coincident sources and receivers having a 10-m spacing. Images both with and without the inclusion of radiation pattern in the inversion are presented in Figure 13. Note how strongly the scattering amplitudes affect the edge effect seen in Figure 13. (Note also that, since the data are zero offset, the obliquity factor is a constant and does not affect these results.)

The spatial resolution operator contains the pseudoinverse of $\Lambda$ and, hence, there is a relationship between parameter resolution and spatial resolution. The next example illustrates this relationship, again for the case of a point perturbation to $C_{33}$. Synthetic data were generated for 2000 Hammersley distributed source-receiver pairs within horizontal distances of 1000 m from a scatterer at 1000 m depth. Figure 14 shows a comparison of the effects of different eigenvalue truncation levels on the image. The higher truncation level (Figure 14b) improves the migration impulse response, although it has the effect of degrading parameter resolution.

In the next examples, we have carried out full multiparameter inversions of two different synthetic datasets. The first dataset is the scattered field due to two point perturba-

tions: $C_{33}$ at a depth of 900 m and $C_{11}$ at a depth of 1000 m (common horizontal coordinate of 1250 m). Results of the inversion of this dataset are shown in Figure 15a ($C_{33}$) and Figure 15b ($C_{11}$). Our algorithm provides automatically such a multiple set of images. In the second dataset, the point perturbation to $C_{11}$ was replaced by a point perturbation to $C_{\nearrow} = [C_{11} + C_{33} + 2(C_{13} + 2C_{55})]/4$, an approximation to the $45°$ $qP$-wave slowness. The result of inverting the second dataset for $C_{\nearrow}$ is given in Figure 15c. In Figure 15, the gray scale covers the same range of absolute amplitudes and a full scattering aperture was synthesized. Note that the separation of parameters from both datasets is satisfactory. The imperfections in the $C_{33}$ image (Figure 15a) are caused by using quasi-Monte Carlo source-receiver locations optimized for the inversion at a depth of 1000 m, rather than the actual depth of 900 m. As a final demonstration of improved spatial resolution, a 3-D inversion was carried out for the model seen in Figure 16. Hammersley-distributed sources and receivers were used and a single parameter, density, was perturbed. Elastic $qP$-wave data were generated over the structure using a Kirchhoff integration technique. The improvements obtained using the full GRT algorithm illustrate that elastic data should be treated as such. Note particularly how the uniform density contrast at the "dome" is properly determined using a GRT inversion (Figure 16A) but deteriorates when the AVA factors are ignored (Figure 16B).

## CONCLUSIONS

In this paper, we have presented a method for carrying out inversion/migration in anisotropic media using the generalized Radon transform (GRT). Practical aspects, such as
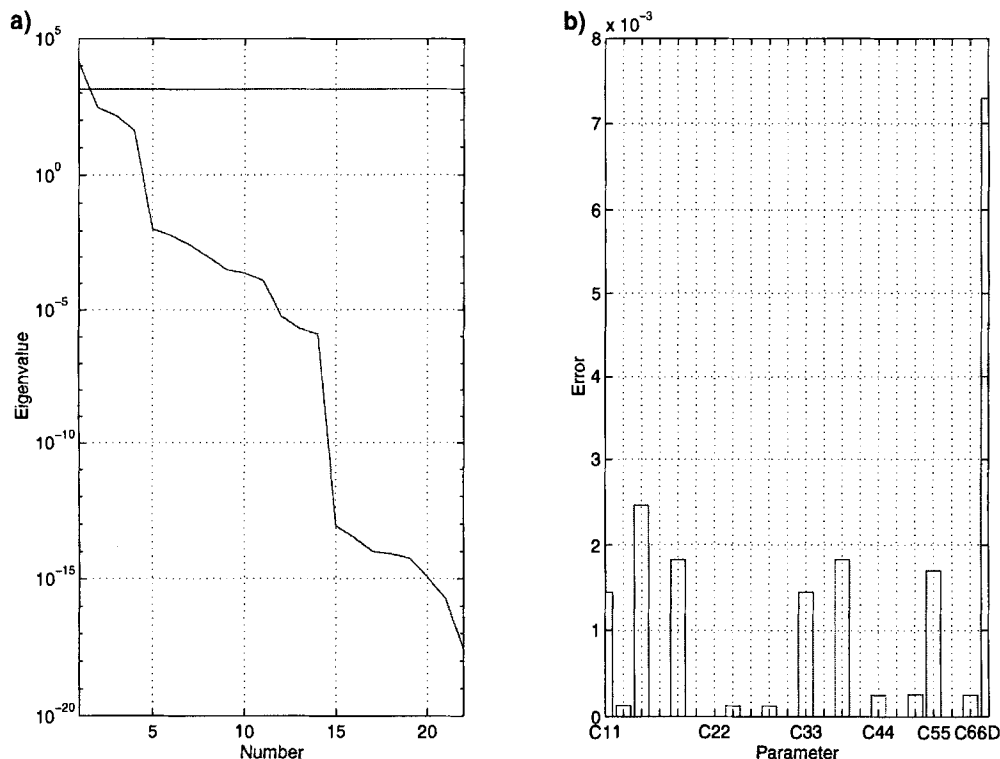


FIG. 12. The eigenvalue spectrum of the 22-parameter system obtained using a dip of $45°$. Propagation angles of up to $60°$ from the vertical were used corresponding to a maximum scattering angle of $30°$. Compare Figure 5.

transformation of variables, regularization, finite aperture normalization, and resolution have been addressed and a treatment proposed. The main conclusions drawn from the synthetic examples are the following.

Better seismic images can be produced when data amplitudes are honored—even if the interpretation of amplitudes in the image is not of primary concern. Our GRT accomplishes this. The angular radiation patterns of the contrast sources influence the spatial resolution of the image, and should not be ignored. Rephrasing this statement, it is widely accepted that a proper amplitude versus scattering Angles (AVA) analysis requires migration; on the other hand, the resolution of an image will be enhanced by proper AVA compensation. Spatial and parameter resolution are essentially coupled to one another.

We have shown how optimal parameter combinations that produce the best possible seismic image may be obtained. Such an image may show structure that would otherwise remain hidden.

We have shown how varying the eigenvalue truncation cut-off level controls the trade-off between both spatial and param-

eter resolutions and uncertainty. In practice, decisions about an appropriate cut-off level will require a careful analysis of the spectrum and trade-off. We anticipate that in many problems there will be clear "knees" in trade-off curves that will guide this choice. We note that once the truncation level is chosen, estimates of error and spatial resolution at each image point are available.

Proper numerical methods are required for the correct evaluation of all the integrals in the calculation of the GRT, its inverse, and its resolution properties. Quasi–Monte Carlo sampling of the rays originating from the image point can be shown to optimize numerical accuracy for a given number of source-receiver combinations (de Hoop and Spencer, 1996). In addition, quasi–Monte Carlo sampling has the benefit of suppressing coherent noise commonly present in measurements, and has enabled us to carry out our calculations with a minimum of computational effort. However, the results of this paper do not rely on quasi–Monte Carlo sampling.

In our examples, we focussed on $qP$-wave scattering. Throughout the paper, the methods we developed included any combination of modes of propagation. In particular, the analysis applies to the case of singly mode-converted phases
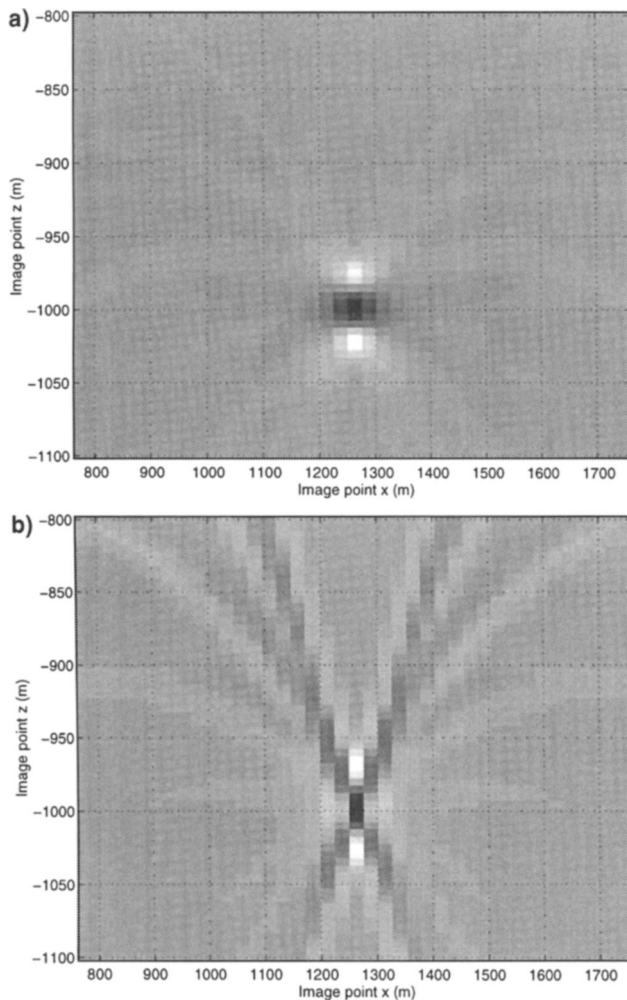


FIG. 13. GRT inversion/migration of a dataset from a $C_{33}^{(1)}$ point perturbation, with scattering amplitude and obliquity corrections (a) and without scattering amplitude and obliquity corrections (b).
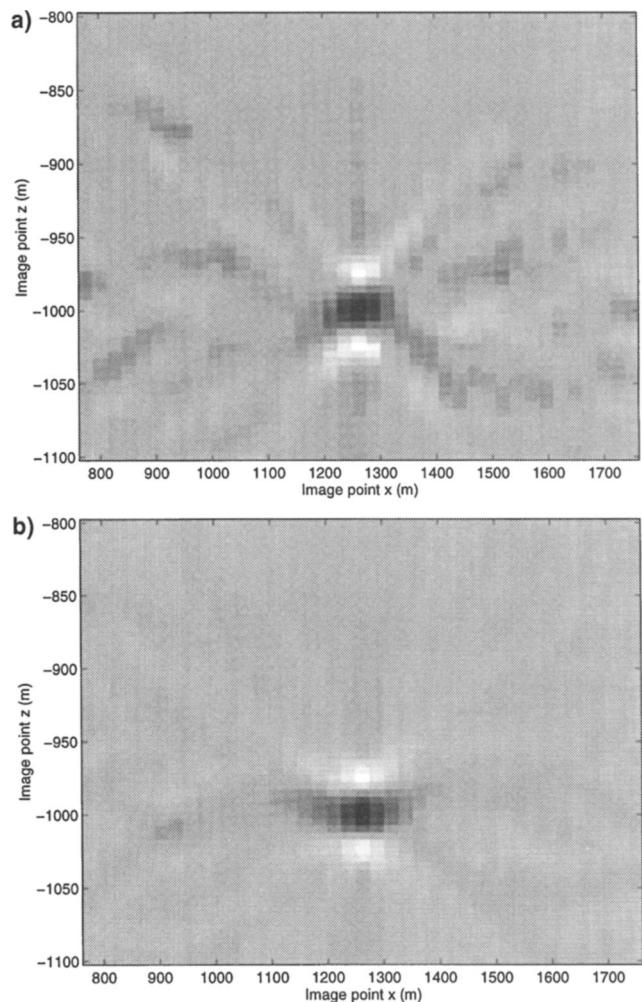


FIG. 14. $C_{33}^{(1)}$ reconstruction as in the previous figure, using eigenvalue truncation levels of $10^{-5}$ (a) and $10^{-1}$ (b).
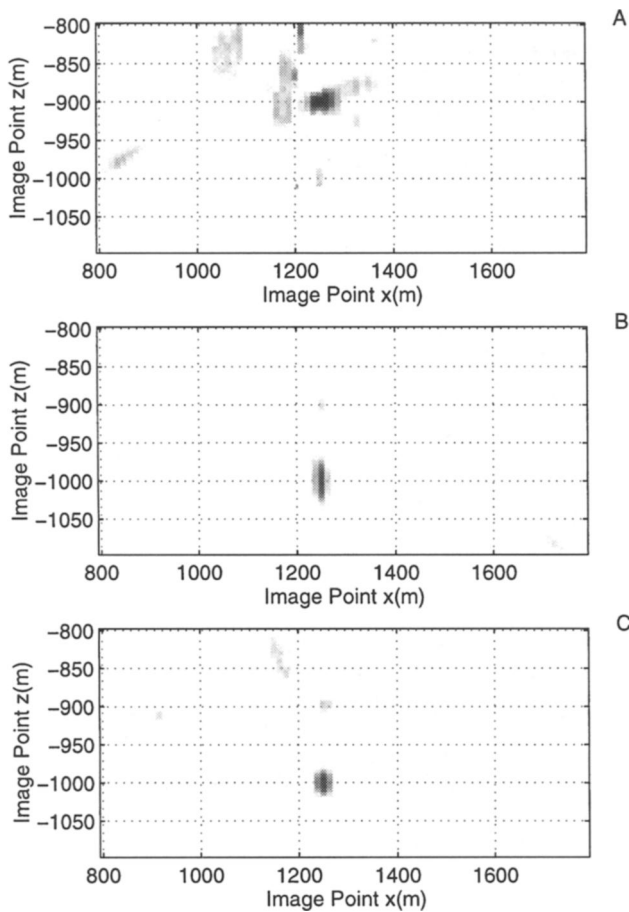
FIG. 15. Multiparameter inversions yielding $C_{33}^{(1)}$ (A), $C_{11}^{(1)}$ (B, first dataset) and $C_{\nearrow}^{(1)}$ (C, second dataset).

such as $qP$-$qSV$ waves, which are now increasingly successfully observed in an ocean-bottom acquisition environment.

Naturally, there are limits to the use of linearized inverse theories in analyzing resolution. Wide-angle seismic acquisition provides a new way of recovering more information on the rock properties in the subsurface, but requires a nonlinear analysis in the vicinity of reflecting interfaces. An analysis of this kind, based on conormal distributions, has been carried out (de Hoop and Bleistein, 1997).

We have excluded the occurrence of instantaneous and macroscopic caustics in the analysis contained in this paper, an issue that has been resolved in the context of GRTs (de Hoop and Brandsberg-Dahl, 1998).

Finally, we remark that an accurate background medium is essential for the GRT inversion and resolution analysis. Therefore, migration velocity analysis methods will need to be developed which correctly predict traveltimes and amplitudes in realistic anisotropic media.
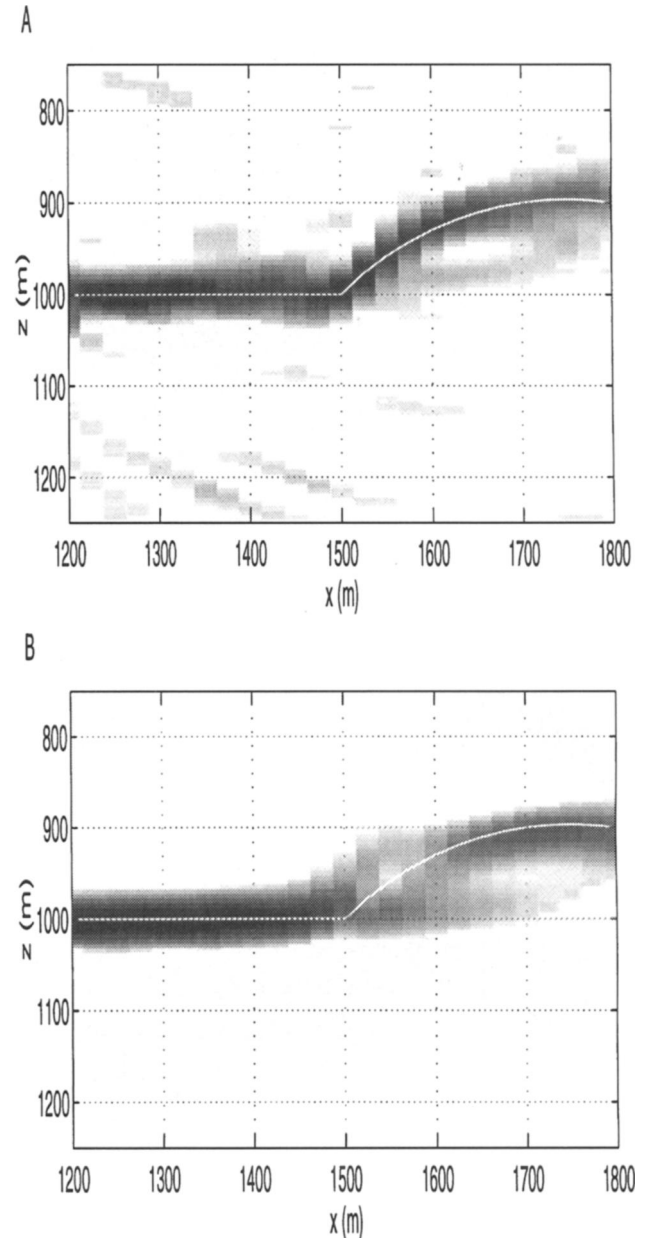
## ACKNOWLEDGMENT

FIG. 16. A line of 3-D inversions of elastic data for density (French model) including (A) and excluding (B) radiation pattern corrections.

## REFERENCES

Backus, G., and Gilbert, F., 1968, The resolving power of gross earth data: Geophys. J. Roy. Astr. Soc., 16, 169–205.
Ball, G., 1995, Estimation of anisotropy and anisotropic 3-D prestack depth migration, offshore Zaire: Geophysics, 60, 1495–1513.
Bamberger, A., Chavent, G., Hemon, C., and Lailly, P., 1982, Inversion of normal incidence seismograms: Geophysics, 47, 757–770.
Banik, N. C., 1987, An effective anisotropy parameter in transversely isotropic media: Geophysics, 52, 1654–1664.
Berkhout, A. J., and Wapenaar, C. P. A., 1990, Delphi: Delft philosophy on acoustic and elastic inversion, Part 2: The Leading Edge, 9, No. 3, 53–59.
Beydoun, W., and Mendez, M., 1989, Elastic ray-Born $L^2$ migration/inversion: Geophys. J. Internat., 97, 151–160.
Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of the causal generalized Radon transform: J. Math. Phys., 26, 99–108.

Beylkin, G., and Burridge, R., 1990, Linearized inverse scattering problem of acoustics and elasticity: Wave Motion, **12**, 15–52.

Bleistein, N., 1987, On the imaging of reflectors within the earth: Geophysics, **52**, 931–942.

Bleistein, N., and Cohen, J. K., 1979, Direct inversion procedure for Claerbout's equation: Geophysics, **44**, 1034–1040.

Burridge, R., de Hoop, M. V., Miller, D., and Spencer, C., 1998, Multiparameter inversion in anisotropic elastic media: Geophys. J. Internat., **134**, 757–777.

Campbell, S., and Meyer, C., 1979, Generalized inverses of linear transformations: Pitman.

Cao, D., Beydoun, W. B., Singh, S. C., and Tarantola, A., 1990, A simultaneous inversion for background velocity and impedance maps: Geophysics, **55**, 458–469.

Carazzone, J. J., and Symes, W. W., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

Chavent, G., and Jacewitz, C. A., 1995, Determination of background velocities by multiple migration fitting: Geophysics, **60**, 476–490.

Cheng, G., and Coen, S., 1984, The relationship between Born inversion and migration for common-midpoint stacked data: Geophysics, **49**, 2117–2131.

Claerbout, J. F., 1992. Earth sounding analysis: Processing versus inversion: Blackwell Scientific Publications.

Clayton, R. W., and Stolt, R. H., 1981, A Born-WKBJ inversion method for acoustic reflection data: Geophysics, **46**, 1559–1567.

Dębski, W., and Tarantola, A., 1995, Information on elastic parameters obtained from the amplitudes of reflected waves: Geophysics, **60**, 1426–1436.

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

de Hoop, M. V., and Brandsberg-Dahl, S., 1998, Maslov asymptotic extension of generalized Radon transform inversions in anisotropic elastic media: A least-squares approach, *in* Seismic imaging; lecture notes of the 1998 Mathematical Geophysics Summer School at Stanford University: Soc. Ind. Appl. Math.

de Hoop, M. V., and de Hoop, A. T., 1997, Wavefield reciprocity and local optimization in remote sensing: Center for Wave Phenomena, Colorado School of Mines, preprint **242**.

de Hoop, M. V., and Spencer, C., 1996, Quasi-Monte Carlo integration over $S^2 \times S^2$ for migration × inversion: Inverse Problems, **12**, 219–239.

de Hoop, M. V., Burridge, R., Spencer, C., and Miller, D., 1994, Generalized Radon transformation/amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media, *in* Hassanzadeh, S., Ed., Proc. SPIE 2301, 15–27.

Devaney, A. J., 1984, Geophysical diffraction tomography: IEEE Trans. Geosci. Remote Sensing, GE-22, 3–13.

Dillon, P. B., 1990, A comparison between Kirchhoff and GRT migration on VSP data: Geophys. Prosp., **38**, 757–777.

Eaton, D. W. S., and Stewart, R. R., 1994, Migration/inversion for transversely isotropic elastic media: Geophys. J. Internat., **119**, 667–683.

Esmersoy, C., and Oristaglio, M., 1988, Reverse-time wave-field extrapolation and inversion: Geophysics, **53**, 920–931.

Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms—Numerical results: Geophysics, **51**, 1387–1403.

Hornby, B. E., Schwartz, L. M., and Hudson, J. A., 1994, Anisotropic effective medium modeling of the elastic properties of shales: Geophysics, **59**, 1570–1583.

Ikelle, L. T., Diet, J. P., and Tarantola, A., 1986, Linearized inversion of multioffset seismic reflection data in the frequency-wavenumber domain: Geophysics, **51**, 1266–1276.

Jin, S., Madariaga, R., Virieux, J., and Lambaré, G., 1992, Two-dimensional asymptotic iterative elastic inversion: Geophys. J. Internat., **108**, 575–588.

Kendall, J.-M., Guest, W. S., and Thomson, C. J., 1992, Ray-theory Green's function reciprocity and ray centered coordinates in anisotropic media. Geophys. J. Internat., **108**, 364–371.

Larner, K., and Tsvankin, I., 1995, P-wave anisotropy: Its practical estimation and importance in processing and interpretation of seismic data. 65th Ann. Internat. Mtg. Soc. Expl. Geophys., Expanded Abstracts, 1502–1505.

Liu, Z., 1995, Migration velocity analysis: Center for Wave Phenomena, Colorado School of Mines, preprint **168**.

Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

Mora, P., 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586.

Norton, S. G., and Linzer, M., 1981, Ultrasonic scattering potential imaging in three dimensions: Exact inverse scattering solutions for plane, cylindrical, and spherical apertures. IEEE Trans. Biomedical Eng., BME-28, 202–220.

Rüger, A., 1996, Variation of P-wave reflectivity with offset and azimuth in anisotropic media: Center for Wave Phenomena, Colorado School of Mines, preprint **218**.

Shearer, P. M., and Chapman, C. H., 1989, Ray tracing in azimuthally anisotropic media—I. Results for models of aligned cracks in the upper crust: Geophys. J., **96**, 51–64.

Singh, S. C., West, G. F., Bregman, N. D., and Chapman, C. H., 1989, Full waveform inversion of reflection data: J. Geophys. Res., **94**, 1777–1794.

Snieder, R., Xie, M. Y., Pica, A., and Tarantola, A., 1989, Retrieving both the impedance contrast and background velocity: A global strategy for the seismic reflection problem: Geophysics, **54**, 991–1000.

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, 23–48.

Symes, W. W., and Kern, M., 1994, Inversion of reflection seismograms by differential semblance analysis: Algorithm structure and synthetic examples. Geophys. Prosp., **42**, 565–614.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

——— 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

Thomsen, L., 1986, Weak elastic anisotropy: Geophysics, **51**, 1954–1966.

Tsvankin, I., and Thomsen, L., 1995, Inversion of reflection traveltimes for transverse isotropy: Geophysics, **60**, 1095–1107.

# APPENDIX A

## $\partial(\tilde{\alpha}, \hat{\alpha})/\partial(\nu, \theta, \psi)$ IN GENERAL MEDIA

We have

$$\nu = \lambda(\tilde{\alpha}, \hat{\alpha})\tilde{\alpha} + \mu(\tilde{\alpha}, \hat{\alpha})\hat{\alpha}, \qquad \text{(A-1)}$$

where [cf. equation (19)]

$$\lambda = \frac{|\tilde{\gamma}|}{|\Gamma|}, \qquad \mu = \frac{|\hat{\gamma}|}{|\Gamma|}. \qquad \text{(A-2)}$$

Note that

$$\tilde{\alpha} \in S^2, \qquad \hat{\alpha} \in S^2, \quad \text{and} \quad \nu \in S^2. \qquad \text{(A-3)}$$

We introduce the angle $\theta$ between the unite phase directions as

$$\cos \theta = \tilde{\alpha} \cdot \hat{\alpha}, \qquad \theta \in [0, \pi). \qquad \text{(A-4)}$$

In view of equation (A-3), we have the constraint

$$\lambda^2 + \mu^2 + 2\lambda\mu \cos\theta = 1. \qquad \text{(A-5)}$$

Further, we introduce the unit vector

$$\zeta = \frac{1}{\sin\theta}(\tilde{\alpha} \wedge \hat{\alpha}) \wedge \nu = \frac{(\tilde{\alpha} \cdot \nu)\hat{\alpha} - (\hat{\alpha} \cdot \nu)\tilde{\alpha}}{\sin\theta}. \qquad \text{(A-6)}$$

[Note that $(\tilde{\alpha} \cdot \nu) = \lambda + \mu \cos\theta$ while $(\hat{\alpha} \cdot \nu) = \lambda \cos\theta + \mu$.] The vectors $\tilde{\alpha}, \hat{\alpha}, \nu$, and $\zeta$ lie in the same plane; also $\zeta \perp \nu$. Note that for $\nu$ fixed, $\zeta \in S^1$. We shall analyse the transformation $(\tilde{\alpha}, \hat{\alpha}) \to (\nu, \theta, \psi)$, where $\psi$ denotes the angular displacement (azimuth) of $\zeta$, and evaluate the associated Jacobian.

First, let $\tilde{\alpha}$ and $\hat{\alpha}$ vary in their own plane [(1, 3) coordinates, i.e., the plane they initially span] keeping $\psi$ fixed. The associated infinitesimal angular displacements of the relevant vectors

will be denoted by the superscript$^{\|}$. Then, in terms of angles $u$, $v$ in a fixed reference frame, we write

$$\tilde{\alpha} = \begin{pmatrix} \sin u \\ 0 \\ \cos u \end{pmatrix}, \qquad \partial_u \tilde{\alpha} = \begin{pmatrix} \cos u \\ 0 \\ -\sin u \end{pmatrix},$$
$$\text{(A-7)}$$
$$\hat{\alpha} = \begin{pmatrix} \sin v \\ 0 \\ \cos v \end{pmatrix}, \qquad \partial_v \hat{\alpha} = \begin{pmatrix} \cos v \\ 0 \\ -\sin v \end{pmatrix},$$

while $\lambda = \lambda(u, v)$ and $\mu = \mu(u, v)$. Note that

$$v - u = \theta. \qquad \text{(A-8)}$$

In general, from equations (A-1) and (A-7), it follows that for in-plane variations,

$$dv = (\partial_u \lambda\, du + \partial_v \lambda\, dv)\tilde{\alpha} + \lambda \partial_u \tilde{\alpha}\, du$$
$$+ (\partial_u \mu\, du + \partial_v \mu\, dv)\hat{\alpha} + \mu \partial_v \hat{\alpha}\, dv. \qquad \text{(A-9)}$$

We introduce the unit vector [cf. equation (A-5)]

$$\nu' = \lambda(u, v)\partial_u \tilde{\alpha} + \mu(u, v)\partial_v \hat{\alpha}. \qquad \text{(A-10)}$$

Note that

$$\partial_u \tilde{\alpha} \perp \tilde{\alpha}, \qquad \partial_v \hat{\alpha} \perp \tilde{\alpha}, \qquad \nu' \perp \nu, \qquad \text{(A-11)}$$

while $\tilde{\alpha}$, $\partial_u \tilde{\alpha}$, $\hat{\alpha}$, $\partial_v \hat{\alpha}$, $\nu$, and $\nu'$ all lie in the same plane.

Since $\nu \cdot dv = 0$, the angular displacement $dv^{\|}$ of $\nu$ is given by

$$dv^{\|} = \nu' \cdot dv = \left[ \lambda^2 + \lambda(\partial_u \mu)(\partial_u \tilde{\alpha} \cdot \hat{\alpha}) \right.$$
$$+ \mu(\partial_u \lambda)(\tilde{\alpha} \cdot \partial_v \hat{\alpha}) + \lambda\mu(\partial_u \tilde{\alpha} \cdot \partial_v \hat{\alpha}) \big]\, du$$
$$+ \left[ \lambda\mu(\partial_u \tilde{\alpha} \cdot \partial_v \hat{\alpha}) + \lambda(\partial_v \mu)(\partial_u \tilde{\alpha} \cdot \hat{\alpha}) \right.$$
$$+ \mu(\partial_v \lambda)(\tilde{\alpha} \cdot \partial_v \hat{\alpha}) + \mu^2 \big]\, dv. \qquad \text{(A-12)}$$

On the other hand, using equation (A-8),

$$d\theta = dv - du. \qquad \text{(A-13)}$$

In our notation, the angular displacement of the phase directions are $d\tilde{\alpha}^{\|} = du$ and $d\hat{\alpha}^{\|} = dv$. Combining equations (A-12) and (A-13), using equation (A-7) leads to the Jacobian

Using equation (A-5), this results in

$$\frac{\partial(\nu^{\|}, \theta)}{\partial(\tilde{\alpha}^{\|}, \hat{\alpha}^{\|})} = 1 + \lambda\mu\left[ (\partial_u + \partial_v) \log\left(\frac{\mu}{\lambda}\right) \right]\sin\theta. \qquad \text{(A-15)}$$

Secondly, consider the case where $\tilde{\alpha}$ and $\hat{\alpha}$ are varied perpendicular to the plane they initially span. The associated infinitesimal angular displacements of the relevant vectors will be denoted by the superscript$^{\perp}$. The out-of-plane variations, keeping $\theta$ fixed, imply

$$d\zeta, \qquad dv \sim \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \qquad \text{(A-16)}$$

To evaluate the out-of-plane Jacobian, it will be convenient to introduce angles $\tilde{\theta}$, $\hat{\theta}$ according to

$$\cos\tilde{\theta} = \tilde{\alpha} \cdot \nu, \qquad \cos\hat{\theta} = \hat{\alpha} \cdot \nu. \qquad \text{(A-17)}$$

Note that

$$\tilde{\theta} + \hat{\theta} = \theta. \qquad \text{(A-18)}$$

Then [cf. equation (A-6)],

$$\zeta = \frac{\cos\hat{\theta}\hat{\alpha} - \cos\hat{\theta}\tilde{\alpha}}{\sin\theta}. \qquad \text{(A-19)}$$

The sine rule applied to the triangle made up of the three vectors $\lambda\tilde{\alpha}$, $\mu\hat{\alpha}$, and $\nu$ gives

$$\frac{\sin\tilde{\theta}}{\mu} = \frac{\sin\hat{\theta}}{\lambda} = \sin\theta. \qquad \text{(A-20)}$$

Substituting equation (A-20) into equation (A-1) yields

$$\nu = \frac{\sin\hat{\theta}\tilde{\alpha} + \sin\tilde{\theta}\hat{\alpha}}{\sin\theta}. \qquad \text{(A-21)}$$

The angular displacement of $\nu$, using constraint (A-16), is then given by

$$dv^{\perp} = \frac{\sin\hat{\theta}\, d\tilde{\alpha}^{\perp} + \sin\tilde{\theta}d\hat{\alpha}^{\perp}}{\sin\theta}. \qquad \text{(A-22)}$$

On the other hand, from equation (A-19) directly follows that

$$d\zeta^{\perp} = \frac{\cos\tilde{\theta}\, d\tilde{\alpha}^{\perp} - \cos\hat{\theta}\, d\tilde{\alpha}^{\perp}}{\sin\theta}. \qquad \text{(A-23)}$$

Note that $d\zeta^{\perp} = d\psi$. Combining equations (A-22) and (A-23) yields

$$\frac{\partial(\nu^{\perp}, \zeta^{\perp})}{\partial(\tilde{\alpha}^{\perp}, \hat{\alpha}^{\perp})} = \frac{1}{\sin^2\theta}\begin{vmatrix} \sin\hat{\theta} & \sin\tilde{\theta} \\ -\cos\hat{\theta} & \cos\tilde{\theta} \end{vmatrix} = \frac{1}{\sin\theta}. \qquad \text{(A-24)}$$

$$\frac{\partial(\nu^{\|}, \theta)}{\partial(u, v)} = \begin{vmatrix} \lambda^2 + (\lambda\partial_u\mu - (\partial_u\lambda)\mu)\sin\theta + \lambda\mu\cos\theta & \mu^2 + (\lambda\partial_v\mu - (\partial_v\lambda)\mu)\sin\theta + \lambda\mu\cos\theta \\ -1 & 1 \end{vmatrix}$$
$$= \lambda^2 + \mu^2 + 2\lambda\mu\cos\theta + [\lambda(\partial_u\mu + \partial_v\mu) - (\partial_u\lambda + \partial_v\lambda)\mu]\sin\theta. \qquad \text{(A-14)}$$

Putting equations (A-15) and (A-24) together, we get

$$\frac{\partial(\nu, \theta, \psi)}{\partial(\tilde{\alpha}, \hat{\alpha})} = \frac{\partial(\nu^{\|}, \theta)}{\partial(\tilde{\alpha}^{\|}, \hat{\alpha}^{\|})} \frac{\partial(\nu^{\perp}, \zeta^{\perp})}{\partial(\tilde{\alpha}^{\perp}, \hat{\alpha}^{\perp})}$$

$$= \frac{1 + \lambda\mu\left[(\partial_u + \partial_v)\log\left(\frac{\mu}{\lambda}\right)\right]\sin\theta}{\sin\theta}. \quad \text{(A-25)}$$

Thus,

$$\frac{\partial(\tilde{\alpha}, \hat{\alpha})}{\partial(\nu, \theta, \psi)} = \frac{\sin\theta}{1 + \lambda\mu\left[(\partial_u + \partial_v)\log\left(\frac{\mu}{\lambda}\right)\right]\sin\theta}. \quad \text{(A-26)}$$

In this final expression, we can substitute [cf. equation (A-2)]

$$\frac{\mu}{\lambda} = \frac{|\hat{\gamma}|}{|\tilde{\gamma}|} = \frac{\tilde{V}(\tilde{\alpha}(u))}{\hat{V}(\hat{\alpha}(v))}, \quad \text{(A-27)}$$

where $V$ denotes the phase velocity as before, and so

$$(\partial_u + \partial_v)\log\left(\frac{\mu}{\lambda}\right) = -\partial_u \log|\tilde{\gamma}| + \partial_v \log|\hat{\gamma}|. \quad \text{(A-28)}$$

Here,

$$\partial_u \log|\tilde{\gamma}| = \frac{\partial_u|\tilde{\gamma}|}{|\tilde{\gamma}|} = \tan\tilde{\chi} \quad \text{with} \quad \cos\tilde{\chi} = \tilde{n}_{\|} \cdot \tilde{\alpha},$$
$$\text{(A-29)}$$

$$\partial_v \log|\hat{\gamma}| = \frac{\partial_v|\hat{\gamma}|}{|\hat{\gamma}|} = \tan\hat{\chi} \quad \text{with} \quad \cos\hat{\chi} = \hat{n}_{\|} \cdot \hat{\alpha}.$$
$$\text{(A-30)}$$

Note that $\tilde{n}_{\|}, \hat{n}_{\|}$ are determined by $\psi$; we have $\tilde{\chi} = \tilde{\chi}(\psi)$ and $\hat{\chi} = \hat{\chi}(\psi)$.

## APPENDIX B

### $\partial(\nu)/\partial(N)$ IN AN ISOTROPIC MEDIUM

In the main text of this paper, Beylkin's determinant was expressed as [cf. equation (48)]

$$\det\begin{pmatrix} \boldsymbol{\Gamma} & \partial_{N_1}\boldsymbol{\Gamma} & \partial_{N_2}\boldsymbol{\Gamma} \end{pmatrix} = \boldsymbol{\Gamma} \cdot \left(\partial_{N_1}\boldsymbol{\Gamma} \wedge \partial_{N_2}\boldsymbol{\Gamma}\right). \quad \text{(B-1)}$$

Substituting equation (16) ($\boldsymbol{\Gamma} = \tilde{\gamma} + \hat{\gamma}$) into equation (B-1) yields a separation into source and receiver ray geometry,

$$\det\begin{pmatrix} \boldsymbol{\Gamma} & \partial_{N_1}\boldsymbol{\Gamma} & \partial_{N_2}\boldsymbol{\Gamma} \end{pmatrix} = (\tilde{\gamma} + \hat{\gamma}) \cdot \left(\partial_{N_1}\tilde{\gamma} \wedge \partial_{N_2}\tilde{\gamma}\right.$$
$$\left. + \partial_{N_1}\tilde{\gamma} \wedge \partial_{N_2}\hat{\gamma} + \partial_{N_1}\hat{\gamma} \wedge \partial_{N_2}\tilde{\gamma} + \partial_{N_1}\hat{\gamma} \wedge \partial_{N_2}\hat{\gamma}\right).$$
$$\text{(B-2)}$$

Let the background medium be homogeneous. Then, the rays connecting the source at $s$ and the receiver at $r$ with the image point $y$ coincide with the vectors

$$\tilde{R} = y - s, \qquad \hat{R} = y - r, \quad \text{(B-3)}$$

respectively. In an isotropic medium, the phase and group directions coincide, hence [cf. equation (13)]

$$\tilde{\alpha} = \frac{\tilde{R}}{|\tilde{R}|}, \qquad \hat{\alpha} = \frac{\hat{R}}{|\hat{R}|}. \quad \text{(B-4)}$$

Also, in equation (14),

$$\tilde{V} = \hat{V} = c \quad \text{(B-5)}$$

is angle independent.

In order to evaluate $\partial_{N_1}\tilde{\gamma}$, we carry out some side calculations [see also equation (49)]. Let $\{i_1, i_2, i_3\}$ denote the three base vectors of a Cartesian reference frame. We have

$$2|\tilde{R}|\partial_{s_j}|\tilde{R}| = \partial_{s_j}|\tilde{R}|^2 = \partial_{s_j}[(y-s) \cdot (y-s)] = -2(y-s) \cdot i_j;$$

hence,

$$\partial_{s_j}|\tilde{R}| = -\tilde{\alpha} \cdot i_j. \quad \text{(B-6)}$$

Using this result in differentiating $\tilde{\gamma} = c^{-1}\tilde{\alpha}$ [cf. equation (8)], we obtain

$$\partial_{s_j}\tilde{\gamma} = -\frac{[i_j - (\tilde{\alpha} \cdot i_j)\tilde{\alpha}]}{|\tilde{R}|c}. \quad \text{(B-7)}$$

In a likewise manner, we find that

$$\partial_{r_j}\hat{\gamma} = -\frac{[i_j - (\hat{\alpha} \cdot i_j)\hat{\alpha}]}{|\hat{R}|c}. \quad \text{(B-8)}$$

Now, we consider the planar acquisition geometry [cf. equation (18)],

$$s_3 = 0; \qquad s_{1,2} = N_{1,2} - \Omega_{1,2}, \qquad \partial_{N_{1,2}}s = i_{1,2},$$
$$r_3 = 0; \qquad r_{1,2} = N_{1,2} + \Omega_{1,2}, \qquad \partial_{N_{1,2}}r = i_{1,2}. \quad \text{(B-9)}$$

Then, with equations (B-7) and (B-8), we get

$$\partial_{N_{1,2}}\tilde{\gamma} = -\frac{[i_{1,2} - (\tilde{\alpha} \cdot i_{1,2})\tilde{\alpha}]}{|\tilde{R}|c},$$
$$\partial_{N_{1,2}}\hat{\gamma} = -\frac{[i_{1,2} - (\hat{\alpha} \cdot i_{1,2})\hat{\alpha}]}{|\hat{R}|c}. \quad \text{(B-10)}$$

Observe that

$$|i_{1,2} - (\tilde{\alpha} \cdot i_{1,2})\tilde{\alpha}| = \left[1 - (\tilde{\alpha} \cdot i_{1,2})^2\right]^{1/2},$$
$$|i_{1,2} - (\hat{\alpha} \cdot i_{1,2})\hat{\alpha}| = \left[1 - (\hat{\alpha} \cdot i_{1,2})^2\right]^{1/2},$$

while

$$[i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha}] \wedge [i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha}] = (\tilde{\alpha} \cdot i_3)\tilde{\alpha},$$

and

$$[i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha}] \wedge [i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha}] = (\hat{\alpha} \cdot i_3)\hat{\alpha}.$$

We will substitute expressions (B-10) into equation (B-2). Equation (B-2) splits into eight contributions. To show the structure of each contribution, consider the one,

$$\tilde{\gamma} \cdot \left(\partial_{N_1}\tilde{\gamma} \wedge \partial_{N_2}\tilde{\gamma}\right) = \frac{1}{c^3}\frac{1}{|\tilde{R}|}\frac{1}{|\tilde{R}|}$$
$$\times \tilde{\alpha} \cdot [i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha}] \wedge [i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha}], \quad \text{(B-11)}$$

with

$$\tilde{\alpha} \cdot [i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha}] \wedge [i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha}]$$
$$= \det(\tilde{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha}).$$

We have

$$\det(\tilde{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha})$$
$$= \det(\tilde{\alpha} \quad i_1 \quad i_2) = (\tilde{\alpha} \cdot i_3), \tag{B-12}$$

$$\det(\hat{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha})$$
$$= (\hat{\alpha} \cdot \tilde{\alpha})\det(\tilde{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha})$$
$$= (\hat{\alpha} \cdot \tilde{\alpha})(\tilde{\alpha} \cdot i_3), \tag{B-13}$$

$$\det(\tilde{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha})$$
$$= \det(\tilde{\alpha} \quad i_1 \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha}) = (\tilde{\alpha} \cdot i_2)(\hat{\alpha} \cdot i_2)(\hat{\alpha} \cdot i_3)$$
$$+ [1 - (\hat{\alpha} \cdot i_2)^2](\tilde{\alpha} \cdot i_3), \tag{B-14}$$

$$\det(\hat{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha})$$
$$= \det(\hat{\alpha} \quad i_1 - (\tilde{\alpha} \cdot i_1)\tilde{\alpha} \quad i_2) = [1 - (\tilde{\alpha} \cdot i_1)^2](\hat{\alpha} \cdot i_3)$$
$$+ (\hat{\alpha} \cdot i_1)(\tilde{\alpha} \cdot i_1)(\tilde{\alpha} \cdot i_3), \tag{B-15}$$

$$\det(\tilde{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha})$$
$$= \det(\tilde{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2) = (\tilde{\alpha} \cdot i_1)(\hat{\alpha} \cdot i_1)(\hat{\alpha} \cdot i_3)$$
$$+ [1 - (\hat{\alpha} \cdot i_1)^2](\tilde{\alpha} \cdot i_3), \tag{B-16}$$

$$\det(\hat{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha})$$
$$= \det(\hat{\alpha} \quad i_1 \quad i_2 - (\tilde{\alpha} \cdot i_2)\tilde{\alpha}) = [1 - (\tilde{\alpha} \cdot i_2)^2](\hat{\alpha} \cdot i_3)$$
$$+ (\hat{\alpha} \cdot i_2)(\tilde{\alpha} \cdot i_2)(\tilde{\alpha} \cdot i_3), \tag{B-17}$$

$$\det(\hat{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha})$$
$$= \det(\hat{\alpha} \quad i_1 \quad i_2) = (\hat{\alpha} \cdot i_3), \tag{B-18}$$

$$\det(\tilde{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha})$$
$$= (\hat{\alpha} \cdot \tilde{\alpha})\det(\hat{\alpha} \quad i_1 - (\hat{\alpha} \cdot i_1)\hat{\alpha} \quad i_2 - (\hat{\alpha} \cdot i_2)\hat{\alpha})$$
$$= (\hat{\alpha} \cdot \tilde{\alpha})(\hat{\alpha} \cdot i_3). \tag{B-19}$$

We have

$$\text{equation (B-12) + equation (B-13)}$$
$$= [1 + (\hat{\alpha} \cdot \tilde{\alpha})](\tilde{\alpha} \cdot i_3), \tag{B-20}$$

$$\text{equation (B-14) + equation (B-16)}$$
$$= (\tilde{\alpha} \cdot i_3) + (\hat{\alpha} \cdot \tilde{\alpha})(\hat{\alpha} \cdot i_3), \tag{B-21}$$

$$\text{equation (B-15) + equation (B-17)}$$
$$= (\hat{\alpha} \cdot i_3) + (\hat{\alpha} \cdot \tilde{\alpha})(\tilde{\alpha} \cdot i_3), \tag{B-22}$$

$$\text{equation (B-18) + equation (B-19)}$$
$$= [1 + (\hat{\alpha} \cdot \tilde{\alpha})](\hat{\alpha} \cdot i_3). \tag{B-23}$$

In view of the planar acquisition geometry, we have

$$(\hat{\alpha} \cdot i_3) = |\tilde{R}\|\hat{R}|^{-1}(\tilde{\alpha} \cdot i_3). \tag{B-24}$$

Hence, adding all the terms from equations (B-20)–(B-23) together in accordance with equation (B-2) and with the proper weighting with powers of $|\tilde{R}|$ and $|\hat{R}|$, we finally arrive at

$$\det \begin{pmatrix} \Gamma & \partial_{N_1}\Gamma & \partial_{N_2}\Gamma \end{pmatrix}$$
$$= \frac{(|\tilde{R}| + |\hat{R}|)(|\tilde{R}|^2 + |\hat{R}|^2)}{|\tilde{R}|^2|\hat{R}|^3} \frac{[1 + (\hat{\alpha} \cdot \tilde{\alpha})](\tilde{\alpha} \cdot i_3)}{c^3}. \tag{B-25}$$

On the other hand,

$$|\Gamma|^3 = \frac{(2[1 + (\hat{\alpha} \cdot \tilde{\alpha})])^{3/2}}{c^3}, \tag{B-26}$$

so that

$$\frac{\partial(\nu)}{\partial(N)} = \frac{(|\tilde{R}| + |\hat{R}|)(|\tilde{R}|^2 + |\hat{R}|^2)}{|\tilde{R}|^2|\hat{R}|^3} \frac{|(\tilde{\alpha} \cdot i_3)|}{2^{3/2}[1 + (\hat{\alpha} \cdot \tilde{\alpha})]^{1/2}}. \tag{B-27}$$

## APPENDIX C

### $\partial(c)/\partial(\delta_1, \ldots, \delta_\mu)$ FOR DEM

In this appendix, we derive a possible microstructure-driven reparameterization of the medium perturbation. This parameterization follows differential effective medium (DEM) theory combined with the self-consistent approximation. For a detailed discussion on this subject, see Hornby et al., (1994). The idea is that the medium may be perturbed by adding or subtracting some generic rock-type material.

### Basic volume averaging

Let a medium be composed of $N$ constituents which occupy volumes $V_n$, respectively. Constituents can be generic sands and shales, for example. In the analysis, we consider a small (compared to the wave length) sample volume $V$ of composite material. Let the constituent volume fractions be given by $\phi_n = V_n/V$.

Volume averages of the strain $e_{ij}$ and the stress $\sigma_{pq}$ are introduced through

$$\langle e_{ij} \rangle = \frac{1}{V} \int_V e_{ij}(x)\, dx, \quad e_{ij}^n = \frac{1}{V_n} \int_{V_n} e_{ij}(x)\, dx, \tag{C-1}$$

$$\langle \sigma_{pq} \rangle = \frac{1}{V} \int_V \sigma_{pq}(x)\, dx, \quad \sigma_{pq}^n = \frac{1}{V_n} \int_{V_n} \sigma_{pq}(x)\, dx.$$

Let $s_{ijpq}$ denote a compliance tensor. We impose two constitutive laws, a macroscopic one,

$$\langle e_{ij} \rangle = s^*_{ijpq} \langle \sigma_{pq} \rangle, \qquad \text{(C-2)}$$

which yields the effective medium, and one for the constituent media,

$$e^n_{ij} = s^n_{ijpq} \sigma^n_{pq}, \qquad \text{(C-3)}$$

valid in each inclusion. Let $\Delta_{ijpq}$ denote the elastic unit tensor. Then,

$$s^n_{ijpq} c^n_{pqk\ell} = \Delta_{ijk\ell} \qquad \text{(C-4)}$$

describes the relation between the compliance and stiffness tensors.

Now, since

$$\langle e_{ij} \rangle = \frac{1}{V} \sum_{n=0}^{N} \int_{V_n} e_{ij}(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{n=0}^{N} \phi_n e^n_{ij}$$
$$\langle \sigma_{pq} \rangle = \frac{1}{V} \sum_{n=0}^{N} \int_{V_n} \sigma_{pq}(\boldsymbol{x})\, d\boldsymbol{x} = \sum_{n=0}^{N} \phi_n \sigma^n_{pq}, \qquad \text{(C-5)}$$

we can analyse the change of compliance tensor, $s^*_{ijpq} - s^0_{ijpq}$, under adding inclusions of media of type $n = 1, \ldots, N$ to a host of type $n = 0$. Using both constitutive laws, we get

$$c^0_{rsij}\left(s^0_{ijpq} - s^*_{ijpq}\right)\langle \sigma_{pq} \rangle = \sum_{n=1}^{N} \phi_n \left(c^n_{rsij} - c^0_{rsij}\right) e^n_{ij}. \qquad \text{(C-6)}$$

To arrive at an expression for the constitutive parameters alone, the average strain $e^n_{ij}$ in each inclusion is related to the supposedly uniform applied stress $\langle \sigma_{pq} \rangle$ on the outer boundary of our sample with volume $V$, namely,

$$e^n_{ij} = S^n_{ijpq} \langle \sigma_{pq} \rangle. \qquad \text{(C-7)}$$

The latter compliance tensor is unknown and may be complicated to compute (in the actual composite medium). Substituting equation (C-7) into equation (C-6) eliminates $\langle \sigma_{pq} \rangle$ and amounts to the explicit relationship between constitutive tensors:

$$s^*_{ijpq} - s^0_{ijpq} = -s^0_{ijrs} \sum_{n=1}^{N} \phi_n \left(c^n_{rsk\ell} - c^0_{rsk\ell}\right) S^n_{k\ell pq}. \qquad \text{(C-8)}$$

In terms of stiffnesses, the relation between the host and effective medium tensors,

$$c^*_{ijpq} - c^0_{ijpq} = \sum_{n=1}^{N} \phi_n \left(c^n_{ijrs} - c^0_{ijrs}\right) S^n_{rsk\ell} c^*_{k\ell pq}, \qquad \text{(C-9)}$$

is not explicit.

## Differential effective medium (DEM)

We discuss two implications of equations (C-8) and (C-9). First, removing the preference for the host medium, namely, replacing constituent $n = 0$ by the (yet to be determined) effec-

tive medium $(c^0_{ijpq} \to c^*_{ijpq})$, implies the system of equations

$$\sum_{n=1}^{N} \phi_n \left(c^n_{ijrs} - c^*_{ijrs}\right) S^n_{rsk\ell} = 0, \qquad \text{(C-10)}$$

from which $S^n_{ijpq}$ can be solved in terms of $c^*_{ijrs}$.

Second, it leads to the introduction of DEM theory. Let us consider the case $N = 1$. The change $\Delta V_1$ in $V_1$ will be assumed to be small. In general, we have

$$\frac{\Delta \phi_1}{\phi_1} = \frac{\Delta V_1}{V_1} - \frac{\Delta V}{V};$$

however, to preserve the amounts of other constituents we set $\Delta V = \Delta V_1$. Hence,

$$\frac{\Delta \phi_1}{\phi_1} = \frac{\Delta V_1}{V_1}\left(\frac{1}{\phi_1} - 1\right)$$

or

$$\frac{\Delta V_1}{V} = \frac{\Delta \phi_1}{1 - \phi_1}. \qquad \text{(C-11)}$$

Substituting equation (C-11) into equations (C-8) and (C-9), and identifying

$$s^0_{ijpq} = s^*_{ijpq}(\phi_1), \qquad c^0_{ijpq} = c^*_{ijpq}(\phi_1), \qquad \text{(C-12)}$$
$$s^*_{ijpq} = s^*_{ijpq}(\phi_1 + \Delta\phi_1), \qquad c^*_{ijpq} = c^*_{ijpq}(\phi_1 + \Delta\phi_1),$$

while

$$S^1_{ijpq} = S^1_{ijpq}(\phi_1 + \Delta\phi_1), \qquad \text{(C-13)}$$

we obtain

$$\frac{s^*_{ijpq}(\phi_1 + \Delta\phi_1) - s^*_{ijpq}(\phi_1)}{\Delta\phi_1} = -\frac{1}{1 - \phi_1} s^*_{ijrs}(\phi_1)$$
$$\times \left(c^1_{rsk\ell} - c^*_{rsk\ell}(\phi_1)\right) S^1_{k\ell pq}(\phi_1 + \Delta\phi_1) \qquad \text{(C-14)}$$

and

$$\frac{c^*_{ijpq}(\phi_1 + \Delta\phi_1) - c^*_{ijpq}(\phi_1)}{\Delta\phi_1} = \frac{1}{1 - \phi_1}\left(c^1_{ijrs}\right.$$
$$\left. - c^*_{ijrs}(\phi_1)\right) S^1_{rsk\ell}(\phi_1 + \Delta\phi_1) c^*_{k\ell pq}(\phi_1 + \Delta\phi_1). \qquad \text{(C-15)}$$

The DEM approximation amounts to neglecting $\Delta\phi_1$ in the right-hand sides of equations (C-14) and (C-15). In the context of GRT inversion/migration, the tensors $s^*_{ijpq}(\phi_1)$ and $c^*_{ijpq}(\phi_1)$ correspond with the anisotropic background velocity model.

## Self consistent approximation

The question remains how to interpret the tensor $S^n_{ijpq}$. In the self-consistent approximation, the tensor $S^n_{ijpq}$, dependent on the actual composition of medium constituents everywhere in the sample, is replaced by a tensor $S^{n,*}_{ijpq}$ that follows from the constitutive relation between strain and applied stress in a configuration consisting of a single inclusion of constituent $n$ embedded in a homogeneous host with the properties of the final (and yet to be determined) effective medium. The shape

of the inclusion will enter the parameterization of the effective medium, for example, the aspect ratios for ellipsoids. For a given shape, $S^{n,*}_{ijpq}$ can be precomputed.

The consequences for the equations derived in the first part of this appendix are a matter of substitution. We obtain

$$c^*_{ijrs}\left(\sum_{n=1}^{N}\phi_n S^{n,*}_{rsk\ell}\right) = \left(\sum_{n=1}^{N}\phi_n c^n_{ijrs} S^{n,*}_{rsk\ell}\right), \qquad \text{(C-16)}$$

while $(\phi = (\phi_1, \ldots, \phi_N))$

$$\left(\partial_{\phi_m} s^*_{ijpq}\right)(\phi) = -\frac{1}{1-\phi_m} s^*_{ijrs}(\phi)\left(c^m_{rsk\ell}-c^*_{rskl}(\phi)\right)S^{m,*}_{k\ell pq}(\phi)$$

$$\text{(C-17)}$$

and

$$\left(\partial_{\phi_m} c^*_{ijpq}\right)(\phi) = \frac{1}{1-\phi_m}\left(c^m_{ijrs} - c^*_{ijrs}(\phi)\right)S^{m,*}_{rsk\ell}(\phi)c^*_{k\ell pq}(\phi).$$

$$\text{(C-18)}$$

The right-hand side of the latter equation yields the Jacobian for reparametrization. Identify $\phi$ with $\delta$ and $N$ with $\mu$.