Kelvin notation for stabilizing elastic-constant inversion

Joe Dellinger

Amoco TTC 2F13, PO Box 3385, Tulsa, OK, 74102, USA

Dan Vasicek

Amoco TTC 2A17, PO Box 3385, Tulsa, OK, 74102, USA

Carl Sondergeld

Amoco TTC 1E24, PO Box 3385, Tulsa, OK, 74102, USA

(July 31, 1998)

Short title: Kelvin notation

ABSTRACT

The problem of inverting a set of core-sample traveltime measurements for a complete set of 21 elastic constants is a difficult one. If the 21 elastic constants are directly used as the inversion parameters, a few bad measurements or an unfortunate starting guess may result in the inversion converging to a physically impossible "solution". Even given perfect data, multiple solutions may exist that predict the observed traveltimes equally well. We desire the inversion algorithm to converge not just to a physically possible solution, but to the "best" (i.e. most physically likely) solution of all those allowed.

We present a new parameterization that attempts to solve these difficulties. The search space is limited to physically realizable media by making use of the Kelvin eigenstiffness-eigentensor representation of the 6×6 elastic stiffness matrix. Instead of 21 stiffnesses, there are 6 "eigenstiffness parameters" and 15 "rotational parameters".

The rotational parameters are defined using a Lie-algebra representation that avoids the artificial degeneracies and coordinate-system bias that can occur with standard polar representations. For *any* choice of these 21 real parameters, the corresponding stiffness matrix is guaranteed to be physically realizable. Furthermore, all physically realizable matrices can be represented in this way.

This new parameterization still leaves considerable latitude as to which linear combinations of the Kelvin parameters to use, and how they should be ordered. We demonstrate that by careful choice and ordering of the parameters, the inversion can be "relaxed" from higher to lower symmetry simply by adding a few more parameters at a time. By starting from isotropy and relaxing to the general result in stages (isotropy, transverse isotropy, orthorhombic, general), we expect that the method should find the solution that is closest to isotropy of all those that fit the data.

INTRODUCTION

If a rock sample is transversely isotropic with a known symmetry axis, its 5 independent elastic constants can be determined by cutting cores at 0°, 90°, and 45° to the rock's axis of symmetry, and measuring P and S velocities along each core's long axis (Jones and Wang, 1981). If the rock's anisotropic symmetry system and orientation are not known, however, a more general anisotropic inversion method is required. Vestrum (1996; 1994) demonstrated that by using a large number of high-quality P and S-wave velocity measurements, he could successfully invert for a complete set of 21 elastic constants that best fits the velocity measurements in the least-squares sense. His method makes no prior assumptions about the symmetry or orientation of the rock.

Although the 21 elements of the anisotropic stiffness matrix make a convenient parameterization for coding purposes, they can cause difficulties if the input velocity measurements are of poor quality. Unrealistic parameter values are typically avoided

in least-squares inversion by adding damping to the algorithm. However, there is no simple damping term that can bias the stiffness matrix towards being transversely isotropic or orthorhombic without also imposing a bias towards a preferred coordinate system (Rasolofosaon et al., 1991). There is also no easy way to ensure that the algorithm will not converge to a result that fits the data but is energetically impossible.

To avoid these difficulties, we parameterize the stiffness matrix $\underset{\approx}{\mathbf{C}}$ geometrically in terms of its eigenstiffnesses and eigentensors (i.e., the eigenvalues and eigenvectors of the stiffness matrix written using Kelvin, instead of the more familiar Voigt, notation) (Helbig, 1994). By ensuring the eigenstiffnesses are positive we avoid the possibility of an energetically disallowed result (Auld, 1973).

Our parameterization requires 21 constants to specify general anisotropy: 6 to specify the eigenvalues, and 15 to specify the eigenvectors. To ensure positivity, the eigenvalues are represented as the exponentials of linear combinations of 6 of the inversion parameters. The eigenvectors are generated by rotating a starting basis set of 6 orthogonal 6-dimensional vectors. We use a Lie-algebra representation of the 6-dimensional rotation to avoid degeneracies (Gilmore, 1974; Hermann, 1966).

By careful choice of the parameters it is possible to "relax" the inversion from higher to lower symmetry by adding a few more parameters at a time. If only the first 2 parameters are nonzero, the result is isotropic. If only the first 7 are nonzero, the result is transversely isotropic. If only the first 12 are nonzero, the result is orthorhombic. (Our current implementation stops at 12 parameters.)

At each stage of the inversion the result is damped back towards the previous more symmetric result. In this way we can determine from the data the likely symmetry system of the medium. We have also found that the inversion converges more quickly and more reliably by this method than inversions that attempt to solve for all 21 elastic constants directly. We expect that this method should also be more likely to find the desired result (instead of a local minimum, or a solution that fits the data well but is geologically implausible).

REVIEW OF KELVIN NOTATION

Since the subject is not yet generally well known in the seismic anisotropy community, we begin with a basic review of Kelvin notation. Interested readers are urged to consult more in-depth sources for further information (Helbig, 1994; Cowin and Mehrabadi, 1992; Sutcliffe, 1992; Mehrabadi and Cowin, 1990; Thomson, 1878).

A homogeneous elastic anisotropic medium can be parameterized by its stiffness tensor C_{ijkl} and density ρ. The $3 \times 3 \times 3 \times 3$ stiffness tensor C_{ijkl} is traditionally compressed into a 6×6 matrix C_{ij} using "Voigt notation" (Auld, 1973). This notation makes use of the symmetries of the stiffness tensor $C_{ijkl} = C_{jikl}$ and $C_{ijkl} = C_{ijlk}$ to compress each pair of indices into one index according to the following rules: $11 \rightarrow 1$; $22 \rightarrow 2$; $33 \rightarrow 3$; $23, 32 \rightarrow 4$; $13, 31 \rightarrow 5$; $12, 21 \rightarrow 6$. Voigt notation is convenient because each element of the compressed matrix C_{ij} is equal to its corresponding element(s) in the tensor C_{ijkl}. Thus, for example, using Voigt notation we have $C_{11} = C_{1111}$, $C_{35} = C_{3313} = C_{3331}$, and $C_{44} = C_{2323} = C_{3223} = C_{2332} = C_{3232}$.

There are also certain disadvantages to Voigt notation. The sum of the squares of the elements of the stiffness tensor has geometrical significance (Rasolofosaon et al., 1991; Federov, 1968), and using Voigt notation this important norm is lost:

$$\sum_{i,j} C_{ij}^2 \neq \sum_{i,j,k,l} C_{ijkl}^2. \tag{1}$$

Helbig (1994), rediscovering long-forgotten work of Lord Kelvin (Thomson, 1878), showed that by using a slightly different normalization the tensor norm can be preserved. This alternative form for the 6×6 stiffness matrix C_{ij} is called "Kelvin notation" (as opposed to the traditional "Voigt notation").

Kelvin notation simply weights each of the elements of the stiffness matrix according to how many elements in the stiffness tensor it represents. Compressed indices 1, 2, and 3 each represent only one combination of tensor indices, so the weighting for these is unity. Compressed indices 4, 5, and 6, however, each represent *two* possible

pairs of tensor indices, so elements of the stiffness matrix with these indices must be scaled by $\sqrt{2}$ (so that when squared they count twice). This extra weighting must be applied for each of the two indices of the stiffness matrix. Thus, using Kelvin notation we would have $C_{11} = C_{1111}$, but $C_{35} = \sqrt{2}C_{3313}$, and $C_{44} = 2C_{2323}$.

Eigenstiffnesses and eigentensors

Because the Kelvin 6×6 $\underset{\approx}{\mathbf{C}}$ matrix preserves the norm of the $3 \times 3 \times 3 \times 3$ stiffness tensor, its eigenvalues and eigenvectors are geometrically meaningful. Following Helbig (1994), we will call the eigenvalues of the Kelvin-notation $\underset{\approx}{\mathbf{C}}$ matrix "eigenstiffnesses", and the 6-dimensional eigenvectors "eigentensors". The eigentensors $\underset{\sim}{\boldsymbol{\epsilon}}^{(k)}$ have the property that:

$$\underset{\sim}{\boldsymbol{\sigma}}^{(k)} = \underset{\approx}{\mathbf{C}}\underset{\sim}{\boldsymbol{\epsilon}}^{(k)} = \Lambda^{(k)}\underset{\sim}{\boldsymbol{\epsilon}}^{(k)}, \tag{2}$$

where $\Lambda^{(k)}$ is the associated eigenstiffness (a real scalar) and $k = \{1, 2, 3, 4, 5, 6\}$. There are 6 eigentensors; each corresponds to a state of stress of the medium for which the 6-dimensional stress and strain vectors, expressed in Kelvin notation, are parallel.

The physical meaning of "parallel 6-dimensional eigentensors" can be better understood by uncompressing the 6-dimensional eigenvector into a 3×3 matrix. For a Kelvin-notation strain we have:

$$\underset{\sim}{\boldsymbol{\epsilon}} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{pmatrix} \rightarrow \begin{pmatrix} \epsilon_1 & \epsilon_6/\sqrt{2} & \epsilon_5/\sqrt{2} \\ \epsilon_6/\sqrt{2} & \epsilon_2 & \epsilon_4/\sqrt{2} \\ \epsilon_5/\sqrt{2} & \epsilon_4/\sqrt{2} & \epsilon_3 \end{pmatrix} = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{pmatrix} = \underset{\approx}{\boldsymbol{\epsilon}} \tag{3}$$

The same equation holds for a Kelvin-notation stress:

$$\underset{\sim}{\boldsymbol{\sigma}} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix} \rightarrow \begin{pmatrix} \sigma_1 & \sigma_6/\sqrt{2} & \sigma_5/\sqrt{2} \\ \sigma_6/\sqrt{2} & \sigma_2 & \sigma_4/\sqrt{2} \\ \sigma_5/\sqrt{2} & \sigma_4/\sqrt{2} & \sigma_3 \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} = \underset{\approx}{\boldsymbol{\sigma}} \tag{4}$$

The factors of "$\sqrt{2}$" ensure that the norm of the vector and matrix remains the same.

The 3×3 matrices $\underset{\approx}{\boldsymbol{\epsilon}}$ and $\underset{\approx}{\boldsymbol{\sigma}}$ are real and symmetric, and so are guaranteed to have real eigenvalues associated with a complete set of orthogonal eigenvectors (Strang, 1980). The 3 eigenvectors give the directions of the 3 principal stresses or strains; the corresponding eigenvalues give their magnitudes. A Kelvin-notation eigentensor thus corresponds to a state of the medium in which each principal stress is parallel to a principal strain, and the same stress / strain ratio (equal to the eigenstiffness $\Lambda^{(k)}$) applies to each of the 3 principal stress-strain pairs. Put more simply, an eigentensor corresponds to a state of the medium for which the stress and strain ellipsoids have the same orientation and aspect ratio.

An example of Kelvin decomposition

Transversely isotropic (TI) Greenhorn Shale (Jones and Wang, 1981) has elastic constants (in kbar, written in traditional Voigt notation):

$$\underset{\approx}{\mathbf{C}} = \begin{pmatrix} 341 & 129 & 107 & 0 & 0 & 0 \\ 129 & 341 & 107 & 0 & 0 & 0 \\ 107 & 107 & 227 & 0 & 0 & 0 \\ 0 & 0 & 0 & 54 & 0 & 0 \\ 0 & 0 & 0 & 0 & 54 & 0 \\ 0 & 0 & 0 & 0 & 0 & 106 \end{pmatrix}. \tag{5}$$

The axis of symmetry is the z axis (i.e., the medium is "TIV").

To convert to Kelvin notation, pre and post multiply by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{pmatrix}, \tag{6}$$

obtaining the Kelvin-notation form of the Greenhorn-shale stiffness matrix,

$$\begin{pmatrix} 341 & 129 & 107 & 0 & 0 & 0 \\ 129 & 341 & 107 & 0 & 0 & 0 \\ 107 & 107 & 227 & 0 & 0 & 0 \\ 0 & 0 & 0 & 108 & 0 & 0 \\ 0 & 0 & 0 & 0 & 108 & 0 \\ 0 & 0 & 0 & 0 & 0 & 212 \end{pmatrix}. \tag{7}$$

The eigenvectors of the above matrix are:

$$\underset{\sim}{\boldsymbol{\epsilon}} = \begin{pmatrix} .63759 & -.30574 & 1/\sqrt{2} & 0 & 0 & 0 \\ .63759 & -.30574 & -1/\sqrt{2} & 0 & 0 & 0 \\ .43239 & .90169 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \tag{8}$$

The columns of this matrix give the eigentensors of Greenhorn Shale. (The apparently jumbled ordering of the eigentensors is necessary for consistency with the canonical ordering used by Helbig in his book (1994); we shall use his ordering throughout this paper.)

Six-dimensional vectors are difficult to visualize. Fortunately, it is possible to write eigentensors as stiffness matrices. Sylvester's matrix theorem allows us to write any matrix as a sum of outer products of its eigenvectors, each weighted by the

corresponding eigenvalue (Claerbout, 1976). Thus, we can decompose a stiffness matrix $\underset{\approx}{\mathbf{C}}$ into a sum of stiffness matrices $\underset{\approx}{\mathbf{C}}^{(k)}$, each with the same eigentensors as $\underset{\approx}{\mathbf{C}}$, but with all but one of the eigenstiffnesses set to zero. Mathematically,

$$\underset{\approx}{\mathbf{C}} = \sum_k \Lambda^{(k)} \underset{\sim}{\boldsymbol{\epsilon}}^{(k)} \underset{\sim}{\boldsymbol{\epsilon}}^{(k)T} = \sum_k \Lambda^{(k)} \underset{\approx}{\mathbf{C}}^{(k)}. \tag{9}$$

The six eigentensors of Greenhorn shale follow. For each eigentensor, we have formed the outer product (following equation (9) above) to convert the eigentensor from a 6-vector to a Kelvin-notation stiffness matrix. We have then converted the Kelvin-notation stiffness matrix to traditional Voigt notation. The reader can verify that the six stiffness matrices that follow do indeed sum to give the stiffness matrix for Greenhorn shale in equation (5).

After each single-eigentensor stiffness matrix we also list the corresponding three principal strain vectors (in the form of a scalar strain multiplied by a normalized unit vector), and explain how the eigentensor would change if the symmetry were increased (to isotropy) or decreased (to orthorhombic, with symmetry planes aligned with the coordinate axes).

For eigentensor 1,

$$\Lambda^{(1)} \underset{\approx}{\mathbf{C}}^{(1)} = 542.562 \begin{pmatrix} .40652 & .40652 & .27568 & 0 & 0 & 0 \\ .40652 & .40652 & .27568 & 0 & 0 & 0 \\ .27568 & .27568 & .18696 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{10}$$

and the corresponding principal strain vectors are:

$$.432385 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \; .637591 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; .637591 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{11}$$

These three principal strains are aligned with the coordinate axes. All three have the same sign, but the principal strain along the z axis has a different magnitude from the

other two. In an orthorhombic medium, the orientation of the strain vectors would remain the same but the TI equivalence between the x and y axes would be lost.

For an isotropic medium, all three strains would have exactly the same magnitude $(1/\sqrt{3})$ and the same sign, and this eigentensor would thus correspond to pure hydrostatic compression. The associated eigenstiffness then depends only on the bulk modulus, k: $\Lambda^{(1)} = C_{11} + C_{12} = 3C_{11} - 4C_{44} = 3k$.

For eigentensor 2,

$$\Lambda^{(2)}\underset{\approx}{\mathbf{C}}^{(2)} = 154.438 \begin{pmatrix} .093478 & .093478 & -.27568 & 0 & 0 & 0 \\ .093478 & .093478 & -.27568 & 0 & 0 & 0 \\ -.27568 & -.27568 & .81304 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{12}$$

and the corresponding principal strain vectors are:

$$-.305742 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; -.305742 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; .901689 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{13}$$

These three principal strains are also aligned with the coordinate axes, but the z axis has an opposite sign from the other two axes. Thus, if z is an axis of compression, then x and y are axes of dilatation. As for eigentensor 1, in an orthorhombic medium the orientation of the strain vectors would remain the same but the equivalence between x and y would be lost.

For an isotropic medium, the three principal strain magnitudes would sum to zero (being $-1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6}$ in that case), and the total volume change represented by the strain would thus also be zero. Not surprisingly, in the isotropic case the associated eigenstiffness is a pure function of the shear modulus, μ: $\Lambda^{(2)} = 2C_{44} = 2\mu$.

For eigentensor 3,

$$\Lambda^{(3)}\underset{\approx}{\mathbf{C}}^{(3)} = 212 \begin{pmatrix} .5 & -.5 & 0 & 0 & 0 & 0 \\ -.5 & .5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{14}$$

and the corresponding principal strain vectors are:

$$-1/\sqrt{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ 1/\sqrt{2}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{15}$$

These three principal strains are also aligned with the coordinate axes. Here the principal strain along the z axis is zero, and the other two are equal and opposite. This is pure shear strain in the x-y plane. In the orthorhombic case the principal strains of this eigentensor would still correspond to the coordinate axes, but it would not necessarily have to form a pure shear strain.

In the more symmetric isotropic case the eigenstiffnesses of eigentensors 2 through 6 would be equal, and we would have $\Lambda^{(3)} = \Lambda^{(2)} = 2\mu$.

For eigentensor 4,

$$\Lambda^{(4)}\underset{\approx}{\mathbf{C}}^{(4)} = 212 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .5 \end{pmatrix}, \tag{16}$$

and the corresponding principal strain vectors are:

$$-1/\sqrt{2}\begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix},\ 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},\ 1/\sqrt{2}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}. \tag{17}$$

This is another pure shear strain in the x-y plane, but rotated 45° about the z axis from eigentensor 3. For the TIV case, it is thus symmetrically equivalent to eigentensor 3 and has the same eigenstiffness (in this example, 212).

For eigentensor 5,

$$\Lambda^{(5)}\underset{\approx}{\mathbf{C}}^{(5)} = 108 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{18}$$

and the corresponding principal strain vectors are:

$$-1/\sqrt{2}\begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix},\ 0\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ 1/\sqrt{2}\begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}. \tag{19}$$

This is a pure shear in the y-z plane.

Finally, for eigentensor 6,

$$\Lambda^{(6)}\underset{\approx}{\mathbf{C}}^{(6)} = 108 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{20}$$

and the corresponding principal strain vectors are:

$$-1/\sqrt{2}\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{pmatrix},\ 0\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ 1/\sqrt{2}\begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}. \tag{21}$$

This is a pure shear in the x-z plane.

The principal strains for eigentensors 4 through 6 are exactly the same for isotropic, TIV (with a z symmetry axis), and orthorhombic (aligned with the coordinate axes), forming an orthogonal trio of pure shear strains in the x-y, y-z, and x-z planes, respectively. This happens because an anisotropic symmetry plane is always associated with a pair of pure-shear eigentensors (Helbig, 1994), and all three of the symmetry systems we have been discussing have the coordinate planes as symmetry planes.

The different symmetry systems are distinguished by whether the associated eigenstiffnesses $\Lambda^{(4)}$ through $\Lambda^{(6)}$ are equal. In the isotropic case, all three axes are symmetrically equivalent and thus eigenstiffnesses 4 through 6 are equal. In the TIV case eigentensors 5 and 6 are equivalent (a 90° rotation about z maps one into the other), and thus $\Lambda^{(5)} = \Lambda^{(6)}$ (in this example 108). In the orthorhombic case all 6 eigenstiffnesses are independent.

If the medium were monoclinic with an x-y plane of symmetry, then eigentensors 1-4 would be entirely arbitrary, but eigentensors 5 and 6 would still form a pair of pure-shear strains. In the case of general anisotropy there are no symmetry planes and all 6 eigentensors are arbitrary.

Note that rotating the medium would leave the eigenstiffnesses completely unchanged. The directions of the principal strains of the eigentensors would rotate, but their magnitudes and the angles between them would also remain unchanged. Thus, even if an anisotropic medium is arbitrarily oriented, any symmetry planes can be immediately identified by expressing the stiffness matrix in Kelvin eigentensor form and checking for paired pure-shear strains (Helbig, 1994).

KELVIN-MATRIX COMPOSITION

In the previous section we decomposed a stiffness matrix into its Kelvin-eigentensor components. It is also possible to reverse the process, building up an

arbitrary stiffness matrix from Kelvin-eigentensor building blocks.

Every stiffness matrix has a complete set of orthogonal eigenvectors, each associated with a real eigenvalue. The stiffness matrix is physically realizable if and only if all its eigenvalues are non-negative (Auld, 1973). Given a complete orthogonal set of eigenvectors $\underset{\sim}{\boldsymbol{\epsilon}}^{(k)}$ and their associated eigenvalues $\Lambda^{(k)}$, we can easily construct a real symmetric matrix with those eigenvectors and eigenvalues using Sylvester's matrix theorem (Claerbout, 1976):

$$\underset{\approx}{\mathbf{C}} = \sum_k \Lambda^{(k)} \underset{\sim}{\boldsymbol{\epsilon}}^{(k)} \underset{\sim}{\boldsymbol{\epsilon}}^{(k)^T}. \tag{22}$$

Any realizable stiffness matrix can therefore be parameterized in terms of 6 real nonnegative scalars (eigenstiffnesses) and 6 orthogonal 6-dimensional vectors (eigentensors).

Specifying the eigenstiffnesses

For media more symmetric than orthorhombic, we have seen that some subsets of the 6 eigenstiffnesses are constrained to have the same value. As given by Helbig (1994), the patterns are (repeated letters indicate repeated eigenstiffnesses):

Isotropic: {A, B, B, B, B, B}

Cubic: {A, B, B, C, C, C}

Transversely Isotropic: {A, B, C, C, D, D}

Tetragonal: {A, B, C, D, E, E}

Orthorhombic: {A, B, C, D, E, F} .

We would like to parameterize the eigenstiffnesses so that the symmetry systems of interest in the list above (isotropic, TI, orthorhombic) "nest" one within the other. Isotropy requires 2 eigenstiffness parameters, so the first 2 inversion parameters should suffice to specify isotropy, leaving parameters 3 through 6 set to zero. TI requires 4 eigenstiffness parameters, so inversion parameters 1 through 4 should be enough for

that case, leaving parameters 5 and 6 set to zero. Orthorhombic anisotropy requires all six eigenstiffness parameters to be unconstrained.

To accomplish this, we parameterize the eigenstiffnesses as exponentials (to ensure positivity) of linear combinations of the 6 eigenstiffness inversion parameters, which are labeled $\mathrm{I_P}$, $\mathrm{I_S}$, TI_1, TI_2, OR_1, and OR_2. In terms of these parameters the six eigenstiffnesses are then:

$\Lambda^{(1)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P})$,

$\Lambda^{(2)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P} + \mathrm{I_S})/4$,

$\Lambda^{(3)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P} + \mathrm{I_S} + \mathrm{TI}_1)/4$,

$\Lambda^{(4)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P} + \mathrm{I_S} + \mathrm{TI}_1 + \mathrm{OR}_1)/4$,

$\Lambda^{(5)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P} + \mathrm{I_S} + \mathrm{TI}_1 + \mathrm{TI}_2 + \mathrm{OR}_1)/4$,

$\Lambda^{(6)} = V_{\mathrm{ref}}^2 \exp(\mathrm{I_P} + \mathrm{I_S} + \mathrm{TI}_1 + \mathrm{TI}_2 + \mathrm{OR}_1 + \mathrm{OR}_2)/4$.

If normalized by density, the eigenstiffnesses have units of squared velocity; V_{ref} is just the reference velocity used to set the units of the problem. Note that as required, $\mathrm{I_P}$ and $\mathrm{I_S}$ are enough for arbitrary isotropy. $\mathrm{I_P}$, $\mathrm{I_S}$, TI_1, and TI_2 span the eigenstiffnesses of transversely isotropic media. All 6 eigenstiffness parameters are needed to span the space of orthorhombic media.

The factor of 4 in the denominators ensures that if only $\mathrm{I_P}$ is nonzero, the medium is isotropic with the usual "default" V_p/V_s ratio of 2. In squared-velocity units, for isotropic media

$$\Lambda^{(1)} = 3V_p^2 - 4V_s^2 \tag{23}$$

and

$$\Lambda^{(2)} = 2V_s^2. \tag{24}$$

Thus $\Lambda^{(1)}/\Lambda^{(2)} = (3/2)(V_p/V_s)^2 - 2$; if $V_p/V_s = 2$ then $\Lambda^{(1)}/\Lambda^{(2)} = 4$.

Specifying the eigentensors

A stiffness matrix contains 21 unique entries, so there are still $21 - 6 = 15$ parameters remaining to be specified.

Every real symmetric matrix has a complete set of orthogonal eigenvectors. This basis of eigenvectors defines a natural Cartesian coordinate system in which the matrix is diagonal (Strang, 1980). The number of angles required to specify the orientation of a coordinate system in n dimensions is $n(n-1)/2$. For our purpose $n = 6$ and the number of angles is 15, which exactly accounts for the 15 remaining parameters. (In general there is also the handedness of the coordinate system to consider. Since eigenvectors can be negated without effect, in our application the handedness of the coordinate system does not matter.)

We now must answer two questions: 1) Which 15 angles should we use? 2) What is the best "starting coordinate system" (i.e., the one that will be obtained if all the angles are set to zero)?

We choose the first three angles, θ_x, θ_y, and θ_z, to rotate the medium in the usual three dimensions of x, y, and z without altering its elastic properties. These form a subset of the 15 angles needed to rotate in 6 dimensions, as shown by Mehrabadi and Cowin (1990).

Traditionally, arbitrary orientations in 3 dimensions are specified using Euler angles: start from the unrotated coordinate system, rotate about z, then x, then z again, obtaining the new, rotated, coordinate system (Synge and Griffith, 1959). We instead use a Lie-algebra notation to specify the rotation (Hermann, 1966; Gilmore, 1974), which makes use of the fact that while finite rotations about x, y, and z do not commute, infinitesimal ones do. (In some sense, we "apply all three rotations at the same time".) This avoids the artificial degeneracies that occur using the Euler angles when the x rotation is small (Vasicek and Lu, 1979). Using the Lie-algebra notation, the required 3×3 rotation matrix is:

$$R = \exp \begin{pmatrix} 0 & \theta_z & -\theta_y \\ -\theta_z & 0 & \theta_x \\ \theta_y & -\theta_x & 0 \end{pmatrix}. \tag{25}$$

The Caley-Hamilton theorem states that any $N \times N$ matrix satisfies its own characteristic equation, a polynomial equation of order N. We can use this result to write any power of a matrix of order N or higher in terms of powers 0 through $N-1$, and thus to compress the infinite Taylor series for the matrix exponential in equation (25) down to a finite number of terms. By that means, an analytical solution can be found (Bellman, 1970; Strang, 1980; Gilmore, 1974). In three dimensions there is a simple geometrical interpretation for the matrix in equation (25): it corresponds to a rotation by an angle of $\sqrt{\theta_x^2 + \theta_y^2 + \theta_z^2}$ about the axis $(\theta_x, \theta_y, \theta_z)$. The explicit result can be found in Gilmore (1974).

The remaining 12 "nongeometrical angles" (nongeometric in that they do not correspond to rotations in ordinary three-dimensional space) affect the elastic properties of the medium. Before we can define the angles, however, we must begin by defining an initial eigentensor matrix. This matrix will give the eigentensors when all the rotational parameters are zero.

The obvious choice for the initial eigentensor matrix is isotropy. In the purely isotropic case, however, several of the eigentensors are symmetrically equivalent: the eigentensor decomposition of isotropy is not unique. The symmetry can be broken by infinitesimally perturbing the medium towards orthorhombic anisotropy aligned with the Cartesian coordinate axes. The choice of eigentensors is then uniquely

$$\underset{\approx}{\boldsymbol{\epsilon}}^{\text{initial}} = \begin{pmatrix} \sqrt{1/3} & -\sqrt{1/6} & \sqrt{1/2} & 0 & 0 & 0 \\ \sqrt{1/3} & -\sqrt{1/6} & -\sqrt{1/2} & 0 & 0 & 0 \\ \sqrt{1/3} & 2\sqrt{1/6} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \tag{26}$$

Comparing equation (26) with equation (8), we see that only 1 rotation, mixing eigentensors 1 and 2, is necessary to convert this set of eigentensors into those for a TIV medium. Three angles, mixing eigentensors 1 through 3, are enough for orthorhombic anisotropy aligned with the coordinate axes (Helbig, 1994).

For most laboratory applications arbitrarily oriented orthorhombic anisotropy is the most general solution desired. In that case, we can again make use of the Lie-notation for 3-dimensional rotation, multiplying the initial eigentensor matrix in equation (26) by

$$\exp\begin{pmatrix} 0 & TI_3 & OR_3 & 0 & 0 & 0 \\ -TI_3 & 0 & OR_4 & 0 & 0 & 0 \\ -OR_3 & -OR_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{27}$$

to obtain the required eigentensors. The angle TI_3 specifies the additional rotation required for TIV; angles OR_3, and OR_4 specify the additional two rotations (beyond TIV) required to span orthorhombic anisotropy.

For the generally anisotropic case, 9 more angles are required. These mix pairwise combinations of one of the first three eigentensors with one of the last three eigentensors. The rotation matrix then becomes:

$$\exp\begin{pmatrix} 0 & TI_3 & OR_3 & M_1 & G_1 & G_4 \\ -TI_3 & 0 & OR_4 & M_2 & G_2 & G_5 \\ -OR_3 & -OR_4 & 0 & M_3 & G_3 & G_6 \\ -M_1 & -M_2 & -M_3 & 0 & 0 & 0 \\ -G_1 & -G_2 & -G_3 & 0 & 0 & 0 \\ -G_4 & -G_5 & -G_6 & 0 & 0 & 0 \end{pmatrix}. \tag{28}$$

Rotations M_1, M_2, and M_3 leave the last two eigentensors uninvolved, preserving the unrotated xy plane as a plane of symmetry of the medium if rotations G_1 through G_6 are zero. (A medium with a single symmetry plane is monoclinic.)

The Canonical order

Now that we have defined the 21 parameters (6 eigenstiffness parameters, 3 geometrical angles, and 12 nongeometrical angles), they still must be placed in the correct order. Isotropy requires only two anisotropy parameters, $\mathrm{I_P}$ and $\mathrm{I_S}$ (orientation is irrelevant for isotropy). For TI (transversely isotropic) media, five parameters are required to specify the anisotropy and two parameters are required to specify the orientation of the symmetry axis (θ_x and θ_y; rotation about the axis of symmetry, z, is irrelevant). TI thus requires seven parameters in all. For orthorhombic anisotropy nine anisotropy parameters and all three orientation parameters are required, for twelve in all. General anisotropy requires all twenty-one parameters.

The proper order is thus:

$\mathrm{I_P}$, $\mathrm{I_S}$ (enough for isotropic),

$\mathrm{TI_1}$, $\mathrm{TI_2}$, TI_3, θ_x, θ_y (enough for TI),

θ_z, $\mathrm{OR_1}$, $\mathrm{OR_2}$, OR_3, OR_4 (enough for orthorhombic),

M_1, M_2, M_3 (enough for monoclinic with an xy symmetry plane), and

G_1, G_2, G_3, G_4, G_5, G_6 (enough for general anisotropy).

In practice, unless the orientation of the single monoclinic symmetry plane is known beforehand, each of the three orthorhombic symmetry planes must be tested as a candidate. It is probably more useful to proceed directly from orthorhombic to general anisotropy.

EXAMPLES

We have replaced the 21-constant parameterization in Vestrum's phase-velocity inversion code (Vestrum et al., 1996; Vestrum, 1994) with our 12-dimensional parameterization. By using only 12 parameters, we explicitly excluded any result less symmetric than arbitrarily oriented orthorhombically anisotropic media. We have tested

the inversion on Vestrum's phenolic-sphere physical-model dataset (Vestrum, 1994). The inversion was performed in three stages, finding best-fitting stiffness matrices for three different anisotropic symmetry systems: isotropic, TI, and orthorhombic.

The following results were achieved. At each stage of the inversion, the 12 model parameters are given, in the following order:

$$\begin{pmatrix} \mathrm{I_P} & \mathrm{I_S} & & & \\ \mathrm{TI}_1 & \mathrm{TI}_2 & TI_3 & \theta_x & \theta_y \\ \theta_z & \mathrm{OR}_1 & \mathrm{OR}_2 & OR_3 & OR_4 \end{pmatrix}. \tag{29}$$

The $\mathrm{I_P} = 0$ velocity, V_{ref}, is 1500 meters per second. The angular parameters (in a slanted font in the list above) are given in degrees. The stiffness matrices are also given, with the elastic constants normalized to units of squared velocity, with the velocity measured in units of kilometers per second. The RMS errors are in units of meters per second.

Initial model:

$$\begin{pmatrix} 0. & 0. & & & \\ 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \end{pmatrix} \tag{30}$$

Initial stiffness matrix:

$$\begin{pmatrix} 2.250 & 1.125 & 1.125 & 0. & 0. & 0. \\ 1.125 & 2.250 & 1.125 & 0. & 0. & 0. \\ 1.125 & 1.125 & 2.250 & 0. & 0. & 0. \\ 0. & 0. & 0. & 0.562 & 0. & 0. \\ 0. & 0. & 0. & 0. & 0.562 & 0. \\ 0. & 0. & 0. & 0. & 0. & 0.562 \end{pmatrix} \tag{31}$$

For the first iterations, only the first two parameters were freed. After 8 iterations, the inversion converged to the isotropic result:

$$\begin{pmatrix} 1.233685 & 0.071239 & & & \\ 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 0. & 0. \end{pmatrix} \tag{32}$$

$$\begin{pmatrix} 7.916 & 3.768 & 3.768 & 0. & 0. & 0. \\ 3.768 & 7.916 & 3.768 & 0. & 0. & 0. \\ 3.768 & 3.768 & 7.916 & 0. & 0. & 0. \\ 0. & 0. & 0. & 2.074 & 0. & 0. \\ 0. & 0. & 0. & 0. & 2.074 & 0. \\ 0. & 0. & 0. & 0. & 0. & 2.074 \end{pmatrix} \tag{33}$$

RMS error = 110.93.

The number of free parameters was then increased to 7, and after 20 more iterations, the inversion converged to the following TI result:

$$\begin{pmatrix} 1.251109 & -0.201660 & & & \\ 0.482242 & -0.305260 & 0.959083 & -0.223800 & 1.039296 \\ 0. & 0. & 0. & 0. & 0. \end{pmatrix} \tag{34}$$

$$\begin{pmatrix} 8.398 & 3.175 & 4.151 & 0.004 & -0.007 & 0.000 \\ 3.175 & 8.360 & 4.189 & -0.001 & 0.018 & 0.000 \\ 4.151 & 4.189 & 7.384 & -0.003 & -0.011 & 0.000 \\ 0.004 & -0.001 & -0.003 & 1.918 & -0.000 & -0.012 \\ -0.007 & 0.018 & -0.011 & -0.000 & 1.918 & -0.003 \\ 0.000 & 0.000 & 0.000 & -0.012 & -0.003 & 2.602 \end{pmatrix} \tag{35}$$

RMS error = 76.25.

All 12 parameters were then freed. After 15 more iterations, the final Orthorhombic result was obtained:

$$\begin{pmatrix} 1.319171 & -0.099532 & & & \\ 0.329786 & -0.227424 & 4.800553 & -0.281243 & 1.884455 \\ -11.910885 & -0.161748 & 0.099789 & -1.707797 & -7.621446 \end{pmatrix} \tag{36}$$

$$\begin{pmatrix} 9.874 & 4.434 & 3.787 & 0.003 & -0.071 & -0.207 \\ 4.434 & 9.320 & 3.765 & -0.006 & -0.022 & 0.084 \\ 3.787 & 3.765 & 6.612 & 0.002 & -0.037 & -0.005 \\ 0.003 & -0.006 & 0.002 & 1.803 & -0.038 & -0.017 \\ -0.071 & -0.022 & -0.037 & -0.038 & 1.976 & 0.005 \\ -0.207 & 0.084 & -0.005 & -0.017 & 0.005 & 2.317 \end{pmatrix} \tag{37}$$

$$\text{RMS error} = 13.67.$$

This orthorhombic model fits the data to within measurement error, which the previous more-symmetric models did not. We therefore conclude that this physical model was orthorhombic to within Vestrum's ability to measure it (Vestrum, 1994). The small values of $\theta_x = -0.28$ and $\theta_y = 1.88$ also indicate that the slowest P-wave propagation axis of this orthorhombic medium was nearly perfectly aligned with the z coordinate axis.

The same inversion was also run starting from the same initial conditions, but with all 12 parameters freed from the start. This run reached the above orthorhombic result in only 13 iterations. Experience has shown, however, that it is safer to proceed in stages, each damped back to the previous more-symmetric result, as we did for the first run given above.

The same inversion code was also run using Vestrum's original parameterization (21 elastic constants). Starting from the same initial conditions, it converged after 36 iterations to the generally anisotropic result:

$$\begin{pmatrix} 9.889 & 4.495 & 3.771 & 0.038 & -0.091 & 0.072 \\ 4.495 & 9.343 & 3.736 & -0.002 & -0.115 & -0.131 \\ 3.771 & 3.736 & 6.612 & 0.006 & -0.020 & -0.130 \\ 0.038 & -0.002 & 0.006 & 1.799 & 0.008 & 0.009 \\ -0.091 & -0.115 & -0.020 & 0.008 & 1.994 & -0.003 \\ 0.072 & -0.131 & -0.130 & 0.009 & -0.003 & 2.315 \end{pmatrix} \tag{38}$$

$$\text{RMS error} = 12.43.$$

This result deviates from being orthorhombic by only 2%.

CONCLUSIONS

We have modified Vestrum's original phase-velocity inversion code (Vestrum et al., 1996; Vestrum, 1994) (obtained courtesy of the CREWES consortium) to use the 12-parameter version of our new Kelvin-inspired parameter set. This new parameterization imposes no assumptions on the orientation of the anisotropy of the sample. It also naturally provides a series of best-fitting stiffness-matrix results of progressively lower anisotropic symmetry, allowing the likely symmetry system of the sample to be determined from the data, instead of needing to be assumed beforehand. Most importantly, using the Kelvin eigentensor parameterization the inversion can only find physically realizable results. Our modified code appears to converge somewhat more quickly and reliably than the original code did, especially when used with poor-quality data.

REFERENCES

Auld, B. A., 1973, Acoustic fields and waves in solids: volume 1: John Wiley and Sons.

Bellman, R., 1970, Introduction to matrix analysis: McGraw-Hill.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill.

Cowin, S. C., and Mehrabadi, M. M., 1992, The structure of the linear anisotropic elastic symmetries: J. Mech. Phys. Solids, **40**, 1549–1471.

Federov, F. I., 1968, Theory of elastic waves in crystals: Plenum Press.

Gilmore, R., 1974, Lie groups, Lie algebras, and some of their applications: Wiley.

Helbig, K., 1994, Foundations of anisotropy for exploration seismics: Pergamon.

Hermann, R., 1966, Lie groups for physicists: W. A. Benjamin.

Jones, L. E. A., and Wang, H. F., 1981, Ultrasonic velocities in Cretaceous shales from the Williston basin: Geophysics, **46**, no. 3, 288–297.

Mehrabadi, M. M., and Cowin, S. C., 1990, Eigentensors of linear anisotropic elastic materials: Quarterly Journal of Mechanics and Applied Mathematics, **43**, 15–41.

Rasolofosaon, P. N. J., Arts, R. J., and Helbig, K., 1991, General anisotropic elastic tensor in rocks, approximation, invariants and particular directions: 61st Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1534–1537.

Strang, G., 1980, Linear algebra and its applications, second edition: Academic Press.

Sutcliffe, S., 1992, Spectral decomposition of the elasticity tensor: J. Applied Mechanics, **59**, 762–773.

Synge, J. L., and Griffith, B. A., 1959, Principles of mechanics, 3rd edition: McGraw Hill.

Thomson, W., 1878, Mathematical theory of elasticity: Elasticity, Encyclopedia Britannica, **7**, 819–825.

Vasicek, D. J., and Lu, C.-H., 1979, Launch and flotation analysis of offshore structures part 2 – barge and jacket interaction on launch analysis: Petroleum Engineer International, **51**, no. 6, 10–16.

Vestrum, R. W., Brown, R. J., and Easley, D. T., 1996, Group- and phase-velocity inversions for the general anisotropic stiffness tensor: Sixth International Workshop on Seismic Anisotropy, Proceedings, 101–140.

Vestrum, R. W., 1994, Group- and phase-velocity inversions for the general anisotropic stiffness tensor: Master's Thesis, Department of Geology and Geophysics, University of Calgary.

FIGURES