

Coarse-layer stripping of vertically variable azimuthal anisotropy from shear-wave data

Leon Thomsen*, Ilya Tsvankin[‡], and Michael C. Mueller*

ABSTRACT

Alford rotation analysis of $2C \times 2C$ shear-wave data (two source components, two receiver components) for azimuthal anisotropy is valid only when the orientation of that azimuthal anisotropy is invariant with depth. The Winterstein and Meadows method of layer stripping vertical seismic profiling (VSP) data relaxes this restriction for coarse-layer variation of the orientation of the anisotropy. Here we present a tensor generalization of the conventional convolutional model of scalar wave propagation and use it to derive generalizations of Winterstein and Meadows layer stripping, valid for $2C \times 2C$ data and for the restricted $2C$ -only case, in the VSP and reflection contexts. In the $2C \times 2C$ VSP application, the result reduces to that of Winterstein and Meadows in the case where both fast and slow shear modes have the same attenuation and dispersion; otherwise, a balancing of mode spectra and amplitudes is required. The $2C \times 2C$ reflection result differs from the $2C \times 2C$ VSP result since two applications of the mode-balancing and mode-advance operations are required (since the waves travel up as well as down). Application to a synthetic data set confirms these results. The $2C \times 2C$ reflection algorithm enables the exploration for sweet spots of high fracture intensity ahead of the bit without the restrictive assumption that the anisotropy orientation is depth invariant.

INTRODUCTION

It is well known that the sedimentary crust is usually azimuthally anisotropic with respect to the propagation of seismic waves. This is most easily seen in the splitting of vertically traveling shear waves and is most plausibly caused by oriented cracks, of dimension much smaller than seismic wavelengths. The phenomenon is of importance in exploration seismology because of

- 1) its deleterious effects on apparent shear-wave data quality, often rendering the data uninterpretable for conventional purposes (e.g., lithology discrimination) unless corrected for;
- 2) its implications for fracture permeability, particularly within fractured hydrocarbon reservoirs, if the inferred cracks are sufficiently large and interconnected; and
- 3) its implications for preferred directions of stress and hence for prescriptions of well drilling plans which minimize borehole stability problems.

Because of the issue of vector polarization, the discussion of seismic shear waves can be complicated. Here we discuss the cases where the input data have

- 1) one source component and two receiver components, all effectively horizontal, called $2C$, and
- 2) two source components and two receiver components, all effectively horizontal called $2C \times 2C$. This case was formerly called $4C$, but that term has now come to mean four receiver components (three vector components and one hydrophone component).

Thomsen (1988) discusses the elements of the phenomenon of azimuthal anisotropy, including a derivation of the tensor rotation (Alford rotation, cf. Alford, 1986) algorithm of $2C \times 2C$ shear-wave data into its principal time series, i.e., those two time series which each contain only one of the two split shear wave modes (fast and slow). That derivation was restricted to the case of depth-invariant orientation of azimuthal anisotropy. While this algorithm has proven to be a useful approximation despite the restriction (e.g., Willis et al., 1986), one would obviously prefer to relax it (Thomsen et al., 1995a,b; Chaimov et al., 1995). In fact, in an important contribution, Winterstein and Meadows (1991a,b) asserted that the earth rarely observes this restriction.

The problem is that if the orientation of azimuthal anisotropy varies with depth, then each pure shear mode will, upon encountering such a different layer, split into two modes, each aligned according to the anisotropy in the new layer. As a

Manuscript received by the Editor December 18, 1997; revised manuscript received December 21, 1998.

*BP Amoco, 501 Westlake Park Boulevard, Houston, Texas 77079. E-mail: ThomseLA@bp.com; MuelleMC@bp.com.

[‡]Colorado School of Mines, Center for Wave Phenomena, Department of Geophysics, Golden, Colorado 80401. E-mail: ilya@dix.mines.edu.

© 1999 Society of Exploration Geophysicists. All rights reserved.

convenient special case for visualizing the process, Figure 1 shows two such layers, each with anisotropy caused by vertical cracks, aligned differently in each layer. Upon downward transmission through the bottom of the first layer, each shear wave (split by propagation through the first layer) will split again into the principal modes of the second layer. Upon reflection, each of these multiple arrivals will split further at each different layer on the way up, resulting (if the overburden layer is coarse, i.e., if it imposes appreciable splitting within the layer) in the chaotic situation of Figure 1a, with a consequent challenge to interpretation.

Our work provides a procedure for reducing this chaos to the interpretable situation in Figure 1b (Alford's problem), wherein the variable azimuthal anisotropy in the overburden has been stripped away, rendering it effectively isotropic, so that the underlying layer can be analyzed.

Winterstein and Meadows present a technique to relax the restriction (of uniform orientation of anisotropy) in the vertical seismic profiling (VSP) context, a coarse-layer stripping technique that is relatively free of strong model assumptions. Although they do not present a derivation (only a heuristic argument), their results confirm that the orientation of azimuthal anisotropy does commonly vary with depth and establish the importance of being able to correct for this. Lefevre et al. (1992) present a more elaborate, model-driven procedure for accomplishing the same goal in the same context.

In an expanded abstract, MacBeth et al. (1992) give a terse presentation of a vector convolutional formalism similar to that presented here and apply it to the same reflection layer-stripping problem. Perhaps because of the abbreviated format, they do not discuss the derivation in a way that exposes its assumptions or restricts its generality.

In this paper, we derive an algorithm for coarse-layer stripping of azimuthal anisotropy from VSP data, which reduces in the appropriate limit to the algorithm of Winterstein and Meadows but generalizes it to include different attenuation between the two shear modes. We also present an algorithm that is valid for the analogous problem in the reflection context; this may be seen as an extension of the methods of Winterstein and Meadows (1991a,b) and of MacBeth et al. (1992) to this context. Our layer-stripping procedure is based upon a vector convolution model of seismic propagation, a straightforward generalization of the scalar convolutional model familiar to most exploration geophysicists. This formalism in itself should prove useful in many other contexts.

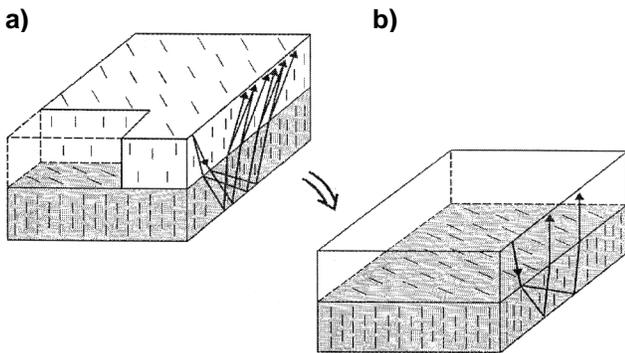


FIG. 1. Coarse-layer stripping of azimuthal anisotropy from shear data.

We also show variants of both algorithms, valid in principle for use with two-component data. Such data may be excited, for example, by a single source-polarization azimuth and received on two horizontal receiver components. These methods are much less robust than their four-component counterparts.

Although this work is focused on the exploration context, in the 1-D approximation described below it is consistent with the work of Silver and Savage (1994), which relies upon wide-azimuth 3-D/2C data to accomplish a similar analysis of vertically inhomogeneous anisotropy orientation. As in the VSP context, Silver and Savage consider a case where the waves make only one pass through the anisotropic layers rather than the two passes of a reflection context.

At first glance, this work appears to involve a lot of complex algebra. However, a closer look reveals that the formulas are mostly straightforward extensions, accounting for the vector nature of seismic wave propagation—concepts already familiar to most exploration geophysicists. We present in the main text only the essential results of the formalism and apply them to a synthetic data example.

The derivations are confined to the appendices. Appendix A gives a generalization of the conventional acoustic convolutional model to vector wave propagation, establishing notation conventions that are used throughout: lowercase bold characters denote vectors, tensors are bold capitals, superscripts identify tensor elements, subscripts indicate the reference coordinate system, and arguments (in parentheses) are propagation times or rotation angles. Long sequences of tensor operations trace out (right-to-left) the time history of a set of vector waves.

This notation simplifies considerably the generalization of scalar convolution to vector convolution. Expanding on this thought, in the simplest context a scalar convolution of a scalar function $f(t)$ with a filter $h(t)$ to yield a filtered output $g(t)$ is commonly denoted

$$g = h \otimes f. \quad (1)$$

If the signal is instead a vector $\mathbf{f}(t)$ and the output is also a vector $\mathbf{g}(t)$, then the filter \mathbf{H} must, in general, produce each output vector element as a convolutional combination of both signal vector elements. If the vectors each have two elements, then we can write this as

$$\begin{aligned} \mathbf{g}^1(t) &= \mathbf{H}^{11}(t) \otimes \mathbf{f}^1(t) + \mathbf{H}^{12}(t) \otimes \mathbf{f}^2(t) \\ \mathbf{g}^2(t) &= \mathbf{H}^{21}(t) \otimes \mathbf{f}^1(t) + \mathbf{H}^{22}(t) \otimes \mathbf{f}^2(t). \end{aligned}$$

The filter \mathbf{H} is seen to be, in these circumstances, a 2×2 matrix. Further, since the signal and the output are both vectors (i.e., if they depend upon a spatial coordinate system such that, with respect to a new, rotated coordinate system, they transform as discussed in Appendix A), then the filter must be a tensor (i.e., it also transforms under rotation as discussed in Appendix A). Further, we can write both equations more compactly as

$$\mathbf{g}^i = \sum_j \mathbf{H}^{ij} \otimes \mathbf{f}^j \quad i = 1, 2,$$

where the explicit time dependence is now implicit and the summation runs over $j = 1, 2$. Finally, the equation above is exactly represented by the matrix equation

$$\mathbf{g} = \mathbf{H} \otimes \mathbf{f}$$

without subscripts. This notation is used throughout this work; it greatly simplifies the expressions and restores the intuitive similarity to the scalar equation above.

The formalism as presented is valid for normal incidence on a sequence of horizontal layers that are azimuthally anisotropic; generalizations to more complicated situations are obvious in principle. We consider only the case of coarse-layer variation of the anisotropy, assuming that the sedimentary section consists of a sequence of horizontal coarse layers, each with its own uniform orientation of azimuthal anisotropy. In each such layer, the type of the anisotropy (i.e., the symmetry class) and its vertical variation are arbitrary as long as the magnitude is weak (e.g., Thomsen, 1988). Therefore, (1) one may equate the orthogonal plane-wave polarization vectors with the nearly orthogonal polarization vectors for rays from a point source. In addition, (2) the reflection and transmission matrices defined below are approximately diagonal. Further, we will approximate that the transmission coefficients for the fast and slow shear modes are approximately equal [equation (C-6)].

Assumption 1 is always valid for azimuthally anisotropic media with a horizontal symmetry plane [including the common azimuthally anisotropic models: transversely isotropic with symmetry axis horizontal (HTI) and vertical orthorhombic (with one symmetry axis vertical)], irrespective of the strength of anisotropy.

In principle, assumption (2) can be used only for media with weak azimuthal anisotropy or for media with strong azimuthal anisotropy whose changes in orientation of anisotropy direction across the interface are small. However, a synthetic test shows that these two assumptions can be successfully applied to a typical moderately anisotropic model with a variation in the anisotropy direction of 30°.

It is legitimate to approximate that transmission coefficients are equal without approximating that the corresponding reflection coefficients are equal. We can do this because, in most cases, the transmission coefficients are close to one, and it does not matter much just how close to one they are. By contrast, in most cases the reflection coefficients are small, and it does matter just how small they are, so we should not (and need not, according to the mathematics below) approximate that they are equal.

Appendix B considers the special case of uniform orientation, showing the reduction to the solution of Alford (1986) and Thomsen (1988). This should reassure the reader that this work is simply an extension, to more realistic cases, of a formalism whose utility is already well established. Of course, this extension from the familiar case is not proven by this reduction; rather, it is established by the logic of Appendix A.

THE VSP ALGORITHM

Appendix C considers the case of a four-component VSP data set (two orthogonal source orientations; two corresponding receiver orientations) with depth-variable anisotropy orientation. The principal time series $\hat{\mathbf{S}}(t)$ in the upper layer may be found by a single Alford (1986) rotation of the observed data tensor $\hat{\mathbf{S}}_0(t)$ for all arrival times by applying the rotation operator $\mathbf{R}(\theta_1)$. This transformed tensor $\mathbf{S}_1(t)$ is physically the same data but is expressed in the coordinate system aligned with the anisotropy in this layer 1. The rotation analysis also yields the angle of orientation θ_1 of the top layer anisotropy with respect to the survey coordinates and the total one-way mode delay

$\Delta t_1 \equiv t_1^s - t_1^f$ of the layer 1 pure-mode 22 component through that layer [following the methods of Alford (1986), Thomsen (1988), and Winterstein and Meadows (1991a,b)]. The angle θ_1 is that angle which best minimizes the off-diagonal traces of the data tensor $\mathbf{S}_1(t)$ in the time window of that layer. The thickness of layer 1 is determined by that time beyond which this diagonalization criterion cannot be adequately met. Of course, this is an interpretive decision; the interpreter must select (either personally or automatically) that reflection time beyond which a rotation by angle θ_1 fails to adequately minimize the diagonal traces.

For times corresponding to layer 1, the principal time series is the rotated data, $\hat{\mathbf{S}}(t) = \mathbf{S}_1(t)$. Then we show that in the next layer the principal time series is given by

$$\hat{\mathbf{s}}(t) = \mathbf{R}(\theta_2 - \theta_1) \left\{ \begin{array}{cc} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t + \Delta t_1) \\ \mathbf{S}_1^{21}(t) & \mathbf{S}_1^{22}(t + \Delta t_1) \end{array} \right\} \otimes \mathbf{B}_1 \left. \right\} \mathbf{R}(\theta_1 - \theta_2),$$

where θ_2 is the angle of orientation of the next layer [equations (C-11) and (C-13)]. This is just a tensor rotation by the difference angle $(\theta_2 - \theta_1)$ of the data matrix $\mathbf{S}_1(t)$, layer stripped. The recipe for layer stripping the data is shown above within the braces; one merely time shifts the second column (corresponding to the source aligned with the slow polarization) forward in time by the amount of the delay Δt_1 and convolves (\otimes) with a filter \mathbf{B}_1 [equation (C-10)] which balances the effects of differential attenuation of the two modes. This recipe was originally presented by Winterstein and Meadows (1991a,b), without derivation, and with the implicit assumption that mode balancing is not necessary.

The rotated data matrix $\mathbf{S}_1(t)$ is aligned with the coordinates of layer 1; only the waves recorded on the traces in the second column of the tensor in square brackets have passed through layer 1 as slow modes, so these alone are time shifted. The fact that this is the correct procedure can be seen intuitively by realizing that the layer 1 rotated data (before stripping) are just the data that would have been excited by physical sources aligned with the layer 1 axes and recorded by receivers also so aligned (Winterstein and Meadows, 1991a,b). Obviously then, the two traces from the source aligned with the layer 1 slow direction are the only ones that need be time shifted; these are in column 2 above.

The mode-balancing filter \mathbf{B}_1 may be derived, for example, by wavelet extraction (using any method of choice) from the layer 1 time window of $\mathbf{S}_1^{11}(t)$, doing the same independently for $\mathbf{S}_1^{22}(t)$ and finding the shaping filter that converts the second into the first. If mode balancing is necessary, but not performed, this leaves differential layer 1 propagation effects in the amplitudes and wave-forms for subsequent times, perhaps leading to puzzling results and spurious complexity for subsequent layers. Evidence from laboratory studies (Sondergeld and Rai, 1992) suggests this is a common circumstance, with the slow mode usually having the greater attenuation.

As with layer 1, the angle $(\theta_2 - \theta_1)$ in the expression above is selected as that angle which best minimizes the off-diagonal traces of the layer-stripped data tensor in the time window corresponding to layer 2. The mode delay Δt_2 in the second layer can be determined by methods previously discussed. The

process may be iterated indefinitely, keeping in mind that errors may accumulate.

THE REFLECTION ALGORITHM

In the case of a $2C \times 2C$ reflection data set, we show in Appendix D that, as above, the anisotropy in the upper layer (principal time series, orientation angle θ_1 , and total two-way mode delay $2\Delta t_1$) may be found by a single Alford rotation. Then we show that the principal time series of the next layer is given by [equations (D-7) and (D-9)]

$$\hat{\mathbf{s}}(t) = \mathbf{R}(\theta_2 - \theta_1) \left\{ \mathbf{B}_1 \otimes \begin{bmatrix} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t + \Delta t_1) \\ \mathbf{S}_1^{21}(t + \Delta t_1) & \mathbf{S}_1^{22}(t + 2\Delta t_1) \end{bmatrix} \right. \\ \left. \otimes \mathbf{B}_1 \right\} \mathbf{R}(\theta_1 - \theta_2).$$

This is just a tensor rerotation by the angle $(\theta_2 - \theta_1)$ of the reflection data matrix \mathbf{S}_1 , layer stripped. The recipe for layer stripping the data is shown within the braces: one time shifts the slow 22 trace forward in time by the amount of the two-way delay $2\Delta t_1$; time shifts the off-diagonal traces by half as much, i.e., by the one-way delay; and then model balances, as shown above.

This detail in time shifting is the principal difference between the VSP algorithm described earlier and the present reflection algorithm. Physically, it arises because each of these off-diagonal arrivals has travelled only one way (either up or down) through layer 1 as a layer 1 fast mode or the other way as a layer 1 slow mode. By contrast, the 22 arrivals have travelled both ways in layer 1 as layer 1 slow modes, so they receive the full two-way delay. Of course, this difference in the layer-stripping recipe between VSP and reflection contexts follows directly from the fundamental difference in geometry of the two experiments.

As with the VSP algorithm, mode balancing is important if in fact the two modes have different attenuation. In this case, the filter \mathbf{B}_1 converts the layer-1 22 wavelet into the 11 wavelet when applied to the former twice, as shown above. (To avoid the assumption of equal reflection coefficients, the two wavelets should be normalized in the time domain on the basis of peak amplitude or area under primary lobe before their spectra are equalized.) As with layer 1, the angle $(\theta_2 - \theta_1)$ is the angle that best minimizes the off-diagonal traces of the layer-stripped data tensor in the time window corresponding to layer 2. The process may be iterated indefinitely, keeping in mind that errors may accumulate.

MacBeth et al. (1992) present a terse algorithm for layer stripping which has many points in common with the present work. In particular, it appears to implicitly contain the time-shifting part of the result above, although this is not shown explicitly. It differs in two respects. First, they recommend solving for layer-1 attributes by using the linear transform procedure developed by Li and Crampin (1993), even though under the special conditions which they require they reduce to Alford rotation. Here, we maintain consistency by using the rotations implied by the formalism to solve for the layer attributes. Second, they assume implicitly that mode balancing is not required.

These $2C \times 2C$ (VSP and reflection context) algorithms are essentially geometric in nature; they are robust because they involve principally a geometric (tensor) rotation and strategic static shifts. It is worth noting that Alford's original procedure consists of a robust mathematical operation (tensor rotation) followed by an interpretation identifying the angle θ_1 . It is because the interpretation follows the mathematical operation, rather than precedes it in the form of detailed modeling assumptions, that the procedure has proven to be so useful and reliable. The present generalization shares this characteristic (albeit with more complexity, befitting the more complex physical situation).

Appendix E gives corresponding derivations for the $2C$ case of a single source orientation and two receiver orientations for both VSP and reflection contexts. To strip a layer of its anisotropy requires the separation of the fast and slow modes onto different traces; hence, we need in principle only to have at least two components of receiver (or source) in a layer to be stripped. However, both VSP and reflection $2C$ algorithms involve much more complexity than simple rotations and deterministic filters, as above, and may not be practical for real data. The reason is that, unlike the $2C \times 2C$ algorithms, the $2C$ algorithms are dynamic in nature rather than geometric and require more elaborate assumptions.

One potentially promising application of our method to $2C$ data is the case of converted reflections, wherein the effective (in-line) source of the shear waves is the P -wave incident upon each of the conversion points at depth, rather than at the surface, and the $2C$ receivers are located at the surface (a reverse VSP geometry). However, a detailed discussion of this converted-wave case is outside the scope of this paper.

A SYNTHETIC EXAMPLE

The most important application of the formalism presented in the Appendices is arguably the result [equation (D-9)] for reflection layer stripping. Below are modeling results for a reflection situation wherein the anisotropic effects are particularly strong.

The model is that of the normal incidence of shear waves upon a thick (750-m) uniform azimuthally anisotropic clastic sequence overlying an anisotropic coal-bed sequence. The coal-bed sequence consists of 40 thin coal beds (cumulative thickness < 100 m) scattered aperiodically through an interval of 530 m, interbedded with the overburden clastics. Because the half-space reflection coefficient between the coal and the clastics is so strong ($\sim 50\%$), there is substantial reflection at each coal-bed sequence internal interface. This results in a complicated reverberation within the coal-bed sequence which leads to comb filtering with the loss of high frequencies, as with P -waves in such a medium. In addition, since the coal is azimuthally anisotropic, each of the principal components (fast and slow) of this reverberation is progressively out of phase with the other, so that the composite waveforms are quite complicated. This model thus presents a stringent test of the algorithms derived in the Appendices.

Figure 2 shows synthetic seismograms for the base-line case where the clastic overburden is isotropic and the survey coordinate system is aligned with the principal coordinate system of the anisotropic coals. The synthetic seismograms were calculated using the reflectivity algorithm discussed by Garmany

(1983). The coal beds have HTI, with shear-wave anisotropy coefficient $\gamma = 20\%$; the interbedding clastics are isotropic. Because the survey coordinate system is aligned with the principal coordinate system of the coal-bed sequence, no energy appears on the mismatched traces (12 and 21). Since the overburden here is isotropic, both arrivals 11 and 22 begin at the same time. Hence, this result appears on casual inspection to show only isotropic response.

However, close inspection of the in-line/in-line (11) and cross-line/cross-line (22) traces reveals that they differ in their details. This is a result of the anisotropy, i.e., of the differences in reflectivities and velocities for the two pure modes polarized orthogonally. Because these two waveforms are not the same in their details, it is not possible to easily define a time delay between them or to characterize one as fast and the other as slow. Nonetheless, the anisotropy is clear if we consider instead acquisition with the survey axes oriented at 45° to the principal axes of the coal (Figure 3). Here, the finite energy on the mismatched components is a sensitive indicator of the presence of azimuthal anisotropy (Alford, 1986). (In Alford's examples, such energy is an indicator of anisotropy in the overburden; here, it is an indicator of anisotropy within the coal-bed sequence.)

In Figure 4, we see the results of modeling a case with vertical variation of azimuthal anisotropy orientation, the focus of this work. The principal axes of anisotropy in the overburden are oriented at an angle of 30° with respect to the principal axes in the coal-bed sequence. The strength of the overburden anisotropy is $\gamma = 2\%$; this degree of anisotropy is a typical background value, while the stronger anisotropy (20%) in the coal-bed sequence is plausible for fractured reservoirs. As a pedagogical device, the synthetic survey line in Figure 4 is oriented at -30° from this overburden coordinate system, parallel to the coordinate system of the underlying coal and identical to the survey coordinates of Figure 2. The difference between Figure 2 (isotropic overburden) and Figure 4 (anisotropic over-

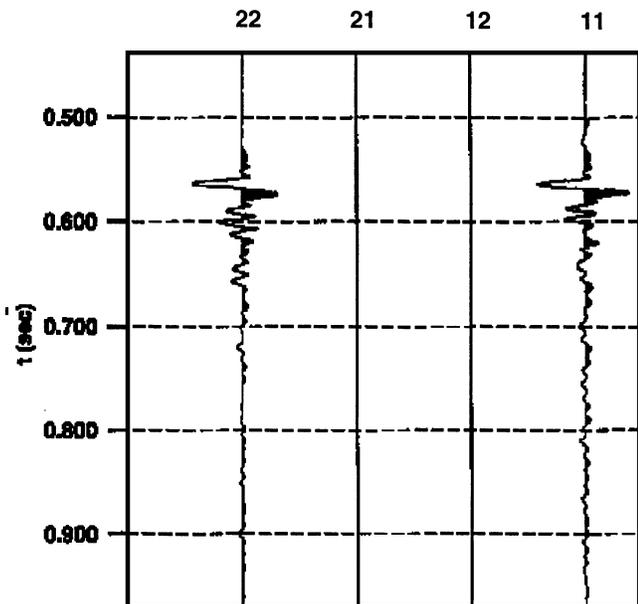


FIG. 2. Isotropic overburden; the survey 1 direction is aligned with the fast direction of the underlying coals.

burden) is evident most notably in the nonzero energy on the mismatched traces.

Following the present algorithm, the analysis to characterize this vertically variable anisotropy proceeds as follows. A sequence of tensor rotations defines the anisotropy (time delay and orientation) of the upper layer, with the optimal rotation angle selected as the angle that minimizes the energy on the mismatched traces in the time window of the upper layer (defined operationally). Automatic procedures must be quite sophisticated to determine both of these (the proper angle and the thickness of the layer to which it applies) simultaneously; however, a human interpreter can do this easily.

For example, an angle of 20° (Figure 5) does not quite extinguish the leading reflection from the coal-bed sequence on

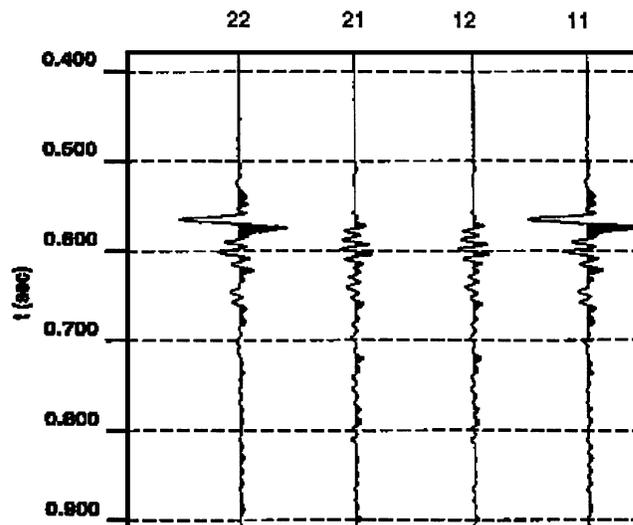


FIG. 3. Isotropic overburden; the survey 1 direction is aligned $+45^\circ$ from the fast direction of the underlying coals, i.e., from that of Figure 1.

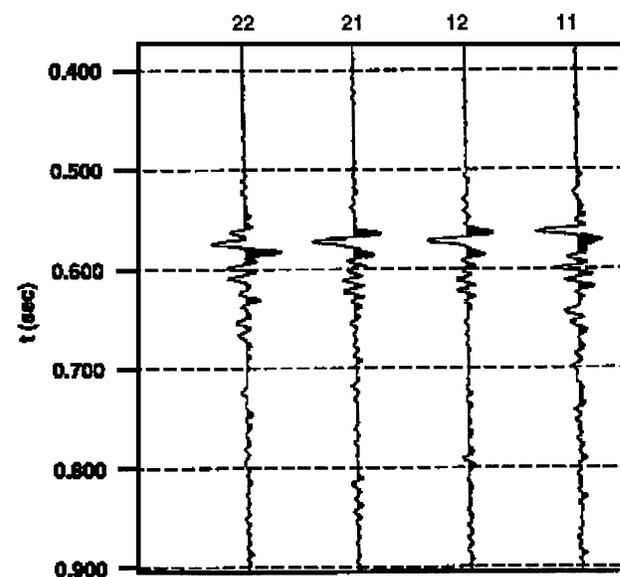


FIG. 4. Anisotropic overburden; the survey 1 direction is aligned (as in Figure 2) with the fast direction of the underlying coals, i.e., at -30° to that of the overburden.

the mismatched traces, whereas an angle of 30° (Figure 6) does extinguish it. No angle extinguishes the mismatched energy within the coal-bed sequence itself, thus establishing that its anisotropy has a different orientation. This logic determines the orientation of the anisotropy axes of the upper layer relative to the survey line and its thickness. The total delay (10 ms in this example) is determined by crosscorrelation of traces 11 and 22, including only the first part of the coal-bed sequence reflection in the correlation window.

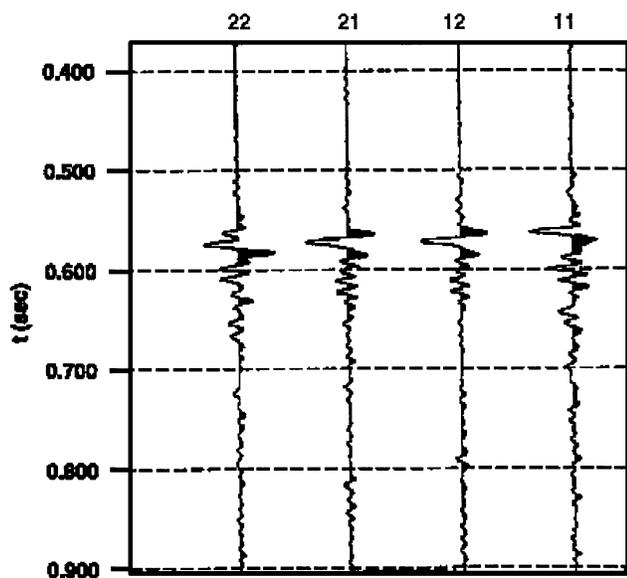


FIG. 5. Anisotropic overburden; the rotated 1 direction is aligned $+20^\circ$ from the survey direction of Figure 4, i.e., not quite aligned with the fast direction of the overburden.

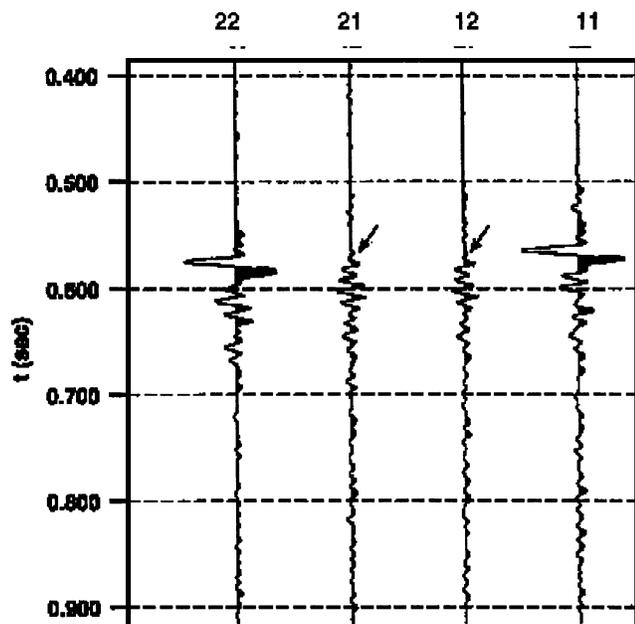


FIG. 6. Anisotropic overburden; the rotated 1 direction is aligned $+30^\circ$ from the survey direction of Figure 4, i.e., exactly aligned with the fast direction of the overburden. Note null mismatched traces in the time window of the upper layer at the arrow.

With the anisotropy of the upper layer determined, that anisotropy is stripped, following the recipe in equation (D-7). Trace 22 is shifted up by 10 ms, and traces 12 and 21 are shifted up by 5 ms. Since this model was perfectly elastic, no mode balancing was required. Subsequent tensor rotations determine that an angle of -30° (Figure 7) extinguishes the energy completely on the mismatched traces within the coal-bed sequence time window. This determines the orientation of the anisotropy within the coal-bed sequence. Its magnitude (i.e., its total time delay) is determined using the methods (in this coal-bed sequence case) of Thomsen et al. (1995b).

Figure 7 (with the anisotropy of the upper layer removed and aligned with the principal coordinate system of the coal) can be compared with Figure 2 (which was similarly aligned and which never had any overburden anisotropy in the first place). The two are almost identical. This proves, via direct forward modeling, that the present algorithm for anisotropic layer stripping of $2C \times 2C$ reflection data is correct. Since the same formalism is used in the other contexts (VSP, $2C$), their validity is thereby strongly supported.

APPLICATION TO REAL DATA

Although one never knows exactly the ground truth in any application to real data, these methods have been used with success in a similar (coal-bed sequence) context by Chaimov et al. (1995). In that case, verification of accurate coal-bed sequence anisotropy was confirmed by production figures: the well with the greater inferred seismic anisotropy had greater inferred intensity of fracture and an observed production rate four times higher than the other. In this connection, the completions were identical in the two wells and the conventional log responses were very similar, so that one could straightforwardly conclude that the difference in production was from the

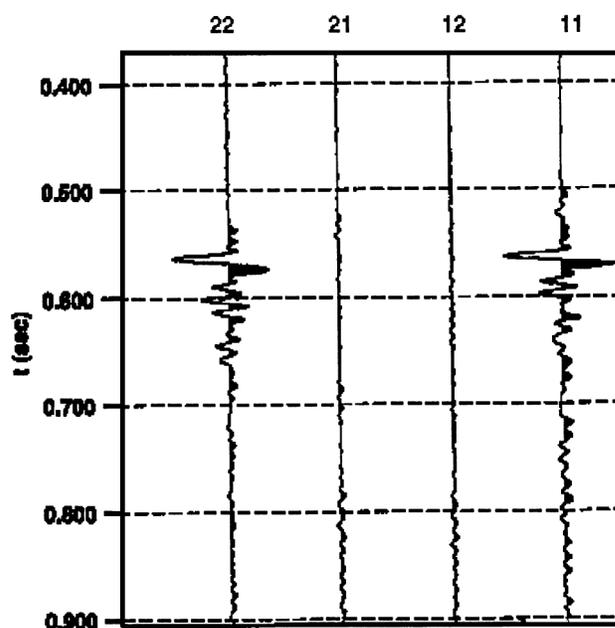


FIG. 7. Anisotropic overburden; after stripping off the anisotropy of the overburden, the rerotated (by -30°) 1 direction is aligned again with the survey direction of Figure 4. These traces are virtually identical to those of Figure 2.

presence of fractures in the productive well not visible by these other measures. (In Chaimov's cases, however, the overburden anisotropy was aligned with the coal-bed sequence anisotropy, with good precision, so these real data did not fully test the present algorithm in its most general form).

In real reflection data, stacked traces are commonly used as a surrogate for noise-reduced normal incidence traces. More than ten years into the era of azimuthally anisotropic exploration seismology, it is still not well understood why, and under what restrictive circumstances, this is such a good approximation. Hence, in any discussion with real data, this issue should be kept in mind.

DISCUSSION

To the extent that seismic data may be interpreted in terms of azimuthal anisotropy of the subsurface rocks, geophysicists hold a significant new tool for the exploration for hydrocarbons. Regardless of the cause of the anisotropy, better *S*-wave images result from correcting for it and better lithologic and other inferences therefore follow. If the inferred azimuthal anisotropy can be attributed to the presence of oriented fractures at the reservoir level, these fractures may provide oriented avenues of enhanced permeability, with obvious implications for the economic production of hydrocarbons from those zones. If the inferred azimuthal anisotropy can be attributed to the alignment of anisotropic stresses, then delineation of such anisotropy may be important even for successful drilling to reach the target zone.

To apply this new tool broadly, it is important to relax the assumption of vertical uniformity in orientation of the anisotropy [Alford (1986), Thomsen (1988)]; this is a primary result of the present analysis. Because of the significant implications, a detailed discussion of the algorithm is warranted, including specific comparisons with the previous literature.

Winterstein and Meadows (1991a,b) claim their VSP layer-stripping algorithm was also valid for the reflection context, whereas we present a different algorithm. Because of these differences in conclusions, it is important to discuss carefully the logic and evidence that lead to each. In the first place, Winterstein and Meadows give no derivation of their algorithm, simply presenting their layer-stripping recipe heuristically. Without a derivation such as that given here, their conclusions are obviously open to conventional criticism. Further, they do not present synthetic modeling results, so the claim for validity of their VSP algorithm rests on comparisons with real data, for which the ground truth is unknown.

As stated clearly by Winterstein and Meadows (1991a,b), asymmetry of the VSP data tensor (i.e., inequality of mismatched traces 12 and 21) is a convincing indication of multiple splitting, i.e., of vertical variation of anisotropy orientation in contexts where multipathing is not plausible. Hence, we can agree that their raw data indicate the occurrence of such variation.

In the reflection context as opposed to the VSP context, the data tensor is more commonly symmetric, with or without multiple splitting. This is a consequence of the reciprocity theorem (Knopoff and Gangi, 1959), which states that the displacements are the same if the (vector) source and receiver are interchanged. In the normal-incidence reflection context, the source and receiver are at the same place, so interchanging source and receiver means interchanging components, i.e., 12 \leftrightarrow 21. Hence, the reflection data tensor must be symmet-

ric, regardless of anisotropy or its distribution in space. (In practice, the minor departures from this condition are usually because the data are stacked traces, i.e., the result of a calculation, rather than actual normal-incidence traces.)

Winterstein and Meadows (1991a,b) do not present data in support of the claim that their algorithm is valid for reflection data as well as VSP data but offer two arguments with the conclusion that in practice it might work less well. Their first argument is that the S/N ratios for reflection data are much smaller than for VSP first arrivals. This is true, but in practice the data are usually of sufficient quality that meaningful analyses may be made.

Their second argument recognizes that reflection data only provide traveltimes and amplitude information at times corresponding to arrivals from strong reflectors and that the putative horizons separating coarse layers of uniform anisotropy need not correspond precisely to these reflective horizons. Hence, a reflection from well beneath a horizon of little reflectivity but significant anisotropy change would arrive at the receiver already split by this invisible interval.

Such a situation could clearly pose a major complication for the present algorithm. But the consequences of such a separation need not be catastrophic. They clearly depend upon the amount of such prior resplitting (i.e., the time delay between the resplit rays, which is accrued prior to reaching the reflector). That is to say, they depend upon the amount of anisotropy and the separation between the horizon of anisotropy change and the reflector. The consequences grow gradually (i.e., linearly) with increasing separation rather than discontinuously. So, if the anisotropy is typical of background values (1–2%), then a significant separation may be tolerable. By contrast, if the anisotropy is much greater, as in a fractured reservoir, the allowed separation would be much smaller. We think that application of the present algorithm to many data sets will be the best way to determine how commonly its assumptions are met in the real world.

Finally, MacBeth et al. (1992) present a similar formalism, from which some of these conclusions could be found; we prefer the present algorithm on grounds of internal consistency and clarity of assumptions, leading to the application, where necessary, of the mode-balancing operations indicated in equations (44) and (54) in that paper.

In conclusion, we present a tensor generalization, to the anisotropic multicomponent case, of the familiar scalar convolutional model of seismic wave propagation. This formalism enables the derivation of algorithms for coarse-layer stripping of vertical variation of azimuthal anisotropy from multicomponent (2C and 2C \times 2C) shear-wave data sets in both the VSP and reflection contexts. These algorithms generalize that presented by Winterstein and Meadows (1991a,b) for the 2C \times 2C VSP context to include differential transmission and attenuation effects for the two shear modes, differ significantly for the 2C \times 2C reflection context from that of Winterstein and Meadows (1991a,b), generalize the similar 2C \times 2C reflection algorithm of MacBeth et al. (1992) in the same way, and indicate considerably more complexity and instability in the 2C cases. Synthetic models demonstrate the validity of the present reflection algorithm.

These algorithms follow Winterstein and Meadows's general logical path and effectively remove an important limitation of Alford's rotation method for anisotropic analysis—namely, the vertical invariance of orientation of the anisotropy.

ACKNOWLEDGMENTS

The bulk of this work was performed in 1992 when all authors were employed by Amoco Production Company, whom we thank for permission to publish the results. We thank S. Cardimona (now at Hanscombe Air Force Base) for refining the synthetic seismogram code, D. Winterstein and M. Meadows (Chevron), and H. Lynn (Lynn, Inc.) for many useful discussions.

REFERENCES

- Alford, R. M., 1986, Shear data in the presence of azimuthal anisotropy: 56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 476–479.
- 1989, Multisource multireceiver method and system for geophysical exploration: U.S. Patent 4 803 666.
- Ata, E., and Michelena, R. J., 1995, Mapping distribution of fractures in a reservoir with P-S converted waves: *The Leading Edge*, **14**, No. 6, 664–676.
- Chaimov, T. A., Beaudoin, G. J., Haggard, W. W., Mueller, M. C., and Thomsen, L., 1995, Shear-wave anisotropy and coal bed methane productivity: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 305–308.
- Garmany, J., 1983, Some properties of elastodynamic eigensolutions in stratified media: *Geophys. J. R. Astr. Soc.*, **75**, 565–569.
- Harrison, M. P., 1992, Processing of P-SV surface-seismic data: Anisotropy analysis, dip moveout, and migration: Ph.D. Thesis, Univ. Calgary.
- Knopoff, L., and Gangi, A., 1959, Seismic reciprocity: *Geophysics*, **24**, 681–691.
- Lefevre, F., Nicoletis, L., Ansel, V., and Cllet, C., 1992, Detection and measure of the shear-wave birefringence from vertical seismic data: Theory and applications: *Geophysics*, **57**, 1463–1481.
- Li, X.-Y., and Crampin, S., 1993, Linear-transform techniques for processing shear-wave anisotropy in four-component seismic data: *Geophysics*, **58**, 240–256.
- MacBeth, C., Li, X.-Y., Crampin, S., and Mueller, M., 1992, Detecting lateral variability in fracture parameters from surface data: 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 816–819.
- Silver, P. G., and Savage, M. K., 1994, The interpretation of shear-wave splitting parameters in the presence of two anisotropic layers: *Geophys. J. Internat.*, **119**, 949–963.
- Sondergeld, C., and Rai, C. S., 1992, Laboratory observations of shear-wave propagation in anisotropic media: *The Leading Edge*, **11**, No. 2, 38–44.
- Thomsen, L., 1988, Reflection seismology over azimuthally anisotropic media: *Geophysics*, **53**, 304–313.
- Thomsen, L., Tsvankin, I., and Mueller, M. C., 1995a, Layer-stripping of azimuthal anisotropy from reflection shear-wave data: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 289–292.
- 1995b, Adaptation of split-shear techniques to coalbed methane exploration: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 301–304.
- Willis, H., Rethford, G., and Bielanski, E., 1986, Azimuthal anisotropy: Occurrence and effect on shear wave data quality: 56th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 479–481.
- Winterstein, D. F., and Meadows, M. A., 1991a, Shear-wave polarizations and subsurface stress directions at Lost Hills field: *Geophysics*, **56**, 1349–1364.
- 1991b, Changes in shear-wave polarizations azimuth with depth in Cymric and Railroad Gap oil fields: *Geophysics*, **56**, 1349–1364.

APPENDIX A

THE VECTOR CONVOLUTIONAL MODEL OF SEISMIC PROPAGATION

We derive here expressions for the principal time series, i.e., for the pure modes (fast and slow) of shear-wave propagation in this coarse-layer, normal-incidence context. We follow the progress of a ray traveling vertically downward through a stack of layers with variable orientation of azimuthal anisotropy, as in Figure 1. The wave particle motion $\mathbf{s}_i(t)$ starts ($t = 0$) at the surface as a vector (referred to a coordinate system identified by subscript i):

$$\mathbf{s}_0(0) = w(t)\mathbf{s}_0, \quad (\text{A-1})$$

where $w(t)$ is the wavelet (assumed the same for each vector component and \mathbf{s}_0 is the vector source. For example, an in-line source pointing in the direction of the $+\mathbf{x}$ -axis is denoted

$$\mathbf{s}_0 = s\mathbf{x} = \begin{bmatrix} s \\ 0 \end{bmatrix},$$

where s is the source strength (a scalar). (We write a vector in column format rather than row format since we will be sequentially multiplying from the left with matrices, and this convention makes full use of the conventional matrix algebra notation.) In terms of the principal coordinate system of the downgoing ray in layer 1 (i.e., that system aligned with the layer-1 orthogonal eigenvectors for vertical propagation), this same physical vector is

$$\mathbf{s}_1 = \mathbf{R}(\theta_1)\mathbf{s}_0, \quad (\text{A-2})$$

where $\mathbf{R}(\theta_1)$ is the rotation matrix for transforming the survey system of coordinates into this principal coordinate system

(subscript 1):

$$\mathbf{R}(\theta_1) = \begin{bmatrix} \cos \theta_1 & \sin \theta_1 \\ -\sin \theta_1 & \cos \theta_1 \end{bmatrix}. \quad (\text{A-3})$$

In the special case of Figure 1, θ_1 is taken as the clockwise angle between the $+\mathbf{x}$ -direction and the strike of the cracks in the top layer. The signs are important: if $+\mathbf{x}$ points north and the top cracks strike at N45W as shown, then $\theta_1 = -45^\circ$ or, equivalently, $+135^\circ$.

In terms of the new coordinates, the vector wavefield of equation (A-1) is

$$\mathbf{s}_1(0) = w(t)\mathbf{s}_1 = \mathbf{R}(\theta_1)w(t)\mathbf{s}_0. \quad (\text{A-4})$$

We rotate to the layer-1 coordinate system because of its special properties; each separate component of $\mathbf{s}_1(t)$ travels as a pure mode with its own velocity (which is not true for the components of $\mathbf{s}_0(t)$). Hence, at the bottom of layer 1 (at time t_1), the wavefield may be written as

$$\mathbf{s}_1(t_1) = \mathbf{P}_1 \otimes \mathbf{s}_1(0) \quad (\text{A-5})$$

where the symbol \otimes denotes convolution and the propagator operator is diagonal:

$$\mathbf{P}_1 = \begin{bmatrix} A_1^f \otimes \delta(t - t_1^f) & 0 \\ 0 & A_1^s \otimes \delta(t - t_1^s) \end{bmatrix}. \quad (\text{A-6})$$

A_1^f is a filter accounting for geometric spreading, attenuation, dispersion, etc., for the fast component, and $t_1^f = z_1/v_1^f$ is its one-way traveltime. A_1^s and t_1^s are the corresponding quantities for the slow component.

At the top of layer 2 (at time t_1 + an infinitesimal increment) but still in terms of layer-1 coordinates, the transmitted wavefield may be written as

$$\mathbf{s}_1(t_1+) = \mathbf{T}_1 \mathbf{s}_1(t_1), \quad (\text{A-7})$$

where the layer 1 \rightarrow 2 (downward) transmission coefficient matrix at normal incidence is

$$\mathbf{T}_1 = \begin{bmatrix} T_1^{11} & T_1^{12} \\ T_1^{21} & T_1^{22} \end{bmatrix}. \quad (\text{A-8})$$

For most cases, T_1^{12} and T_1^{21} will be negligible, so that \mathbf{T}_1 is diagonal; we assume this in the following.

At the same physical place, but in terms of the principal coordinate system of layer 2, this physical wavefield is

$$\mathbf{s}_2(t_1+) = \mathbf{R}(\theta_2 - \theta_1) \mathbf{s}_1(t_1+). \quad (\text{A-9})$$

In this new coordinate system (subscript 2), each component contains rotationally weighted wavelets from both fast and slow waves from layer 1 (the slow ones delayed by $\Delta t_1 \equiv t_1^s - t_1^f$). But these new components again each travel as pure modes in layer 2, albeit with complicated signatures. After propagating down through layer 2, the wavefield is

$$\mathbf{s}_2(t_1 + t_2) = \mathbf{P}_2 \otimes \mathbf{s}_2(t_1+), \quad (\text{A-10})$$

where the propagation operator is [equation (A-6)]

$$\mathbf{P}_2 = \begin{bmatrix} A_2^f \otimes \delta(t - t_2^f) & 0 \\ 0 & A_2^s \otimes \delta(t - t_2^s) \end{bmatrix}. \quad (\text{A-11})$$

Suppose this wavefield is incident upon a downhole receiver with two horizontal components, which for simplicity we assume to be oriented along the survey directions. Hence, it will receive the signal

$$\mathbf{s}_0(t_1 + t_2) = \mathbf{R}(-\theta_2) \mathbf{s}_2(t_1 + t_2). \quad (\text{A-12})$$

Assembling all these operations from equation (A-1) through equation (A-12) (preserving the order, since matrix multiplication does not commute), we have

$$\mathbf{s}_0(t_1 + t_2) = \mathbf{R}(-\theta_2) \mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1) \mathbf{T}_1 \mathbf{P}_1 \otimes \mathbf{R}(\theta_1) w(t) \mathbf{s}_0. \quad (\text{A-13})$$

Note in this linear algebra formulation of the problem that matrix multiplication is associative, although it is not commutative. That is, for matrices,

$$\mathbf{ABC} = \mathbf{A(BC)} = (\mathbf{AB})\mathbf{C} \neq \mathbf{BAC}.$$

The single vector equation (A-12) is a compact way of writing two coupled equations for the two received vector components. Now, repeating all this with a second source orientation \mathbf{S}_0 (same wavelet), we can assemble all four equations compactly by writing the source vectors and the received time series as tensors, each column of which is one of the vectors discussed earlier:

$$\mathbf{S}_0(t_1 + t_2) = \mathbf{R}(-\theta_2) \mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1) \mathbf{T}_1 \mathbf{P}_1 \otimes \mathbf{R}(\theta_1) w(t) \mathbf{S}_0. \quad (\text{A-14})$$

For example if the first source is +in-line and the second is +cross-line and of equal source-strength s , then

$$\mathbf{S}_0 = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} = s \mathbf{I}; \quad (\text{A-15})$$

we assume this in the following.

To extend these results to the reflection context, we resume the ray tracing at the bottom of layer 2, i.e., at equation (A-10). The reflected wavefield at the bottom of layer 2 is

$$\mathbf{s}_2(t_1 + t_2+) = \mathbf{R} \mathbf{s}_2(t_1 + t_2), \quad (\text{A-16})$$

where the reflection coefficient matrix is

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}. \quad (\text{A-17})$$

For most cases, R_{12} and R_{21} will be negligible, so that \mathbf{R} is diagonal; we assume this in the following. We continue the ray tracing upward following the previous logic so that the surface-recorded data tensor is

$$\mathbf{S}_0(2t_1 + 2t_2) = 2\mathbf{R}(-\theta_1) \mathbf{P}_1 \mathbf{R}(\theta_1 - \theta_2) \mathbf{T}_2 \mathbf{P}_2 \otimes \mathbf{R} \mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1) \mathbf{T}_1 \mathbf{P}_1 \otimes \mathbf{R}(\theta_1) w(t) s \mathbf{I}. \quad (\text{A-18})$$

The value \mathbf{T}_2 is the transmission coefficient for upward transmission at the bottom of layer 1. The leading factor 2 accounts for the free-surface interaction.

Also, we note that the expression for rotating a tensor [generalizing equation (A-2)] is

$$\mathbf{S}_1 = \mathbf{R}(\theta_1) \mathbf{S}_0 \mathbf{R}(-\theta_1). \quad (\text{A-19})$$

This derivation does not assume any particular symmetry of the media, just normal incidence upon a sequence of coarse layers of uniform anisotropy orientation. Within each coarse layer, there may be arbitrary vertical variation of vertical velocities and of the magnitude of (weak) anisotropy.

APPENDIX B SPECIAL CASES

Let us consider some special cases to make contact with what we already know. In particular, we consider the reduction of the above formulae to the simple case of uniform orientation of anisotropy. This is the case of Alford (1986) and Thomsen (1988). In this case, $\theta_2 = \theta_1$ and $\mathbf{R}(\theta_2 - \theta_1) = \mathbf{I}$, the identity matrix. This exercise will help us see where we want to go with the more general cases.

2C \times 2C VSP case

The VSP case, equation (A-14) reduces to

$$\mathbf{S}_0(t_1 + t_2) = \mathbf{R}(-\theta_1) [\mathbf{P}_2 \otimes \mathbf{T}_1 \mathbf{P}_1 \otimes s w(t)] \mathbf{R}(\theta_1), \quad (\text{B-1})$$

where the scalar functions have been commuted into the interior. The expression in brackets is

$$\mathbf{S}(t) \equiv [\mathbf{P}_2 \otimes \mathbf{T}_1 \mathbf{P}_1 \otimes sw(t)] \quad (\text{B-2})$$

$$= s \mathbf{T}_1 \begin{bmatrix} A_2^f \otimes \delta(t - t_2^f) \otimes A_1^f \otimes \delta(t - t_1^f) \otimes w(t) & 0 \\ 0 & A_2^s \otimes \delta(t - t_2^s) \otimes A_1^s \otimes \delta(t - t_1^s) \otimes w(t) \end{bmatrix}, \quad (\text{B-3})$$

which is diagonal because all of its constituent matrices are diagonal. Since convolutions commute and since

$$\begin{aligned} \delta(t - t_2^f) \otimes \delta(t - t_1^f) \otimes w(t) &= \delta(t - t_2^f - t_1^f) \otimes w(t) \\ &= w(t - t_2^f - t_1^f), \end{aligned} \quad (\text{B-4})$$

this is

$$\mathbf{S}(t) \equiv s \mathbf{T}_1 \begin{bmatrix} A_2^f \otimes A_1^f \otimes w(t - t_2^f - t_1^f) & 0 \\ 0 & A_2^s \otimes A_1^s \otimes w(t - t_2^s - t_1^s) \end{bmatrix}, \quad (\text{B-5})$$

which we recognize as the time-delayed pure-mode wavelets filtered by propagation effects. That is, the components of $\mathbf{S}(t)$ are the principal time series which were to be found.

So, equation (B-1) reads as

$$\mathbf{S}_0(t_1 + t_2) = \mathbf{R}(-\theta_1) \mathbf{S}(t) \mathbf{R}(\theta_1), \quad (\text{B-6})$$

which can be solved by operating from the left with $\mathbf{R}(\theta_1)$ and from the right with $\mathbf{R}(-\theta_1)$. Since $\mathbf{R}(-\theta_1) \mathbf{R}(\theta_1) = \mathbf{R}(\theta_1) \mathbf{R}(-\theta_1) = \mathbf{I}$, we have

$$\mathbf{S}(t) = \mathbf{R}(\theta_1) \mathbf{S}_0(t_1 + t_2) \mathbf{R}(-\theta_1), \quad (\text{B-7})$$

which was first noted by Alford (1986) (see also Thomsen, 1988). In words, the principal time series on the left is the data rotated to the proper angle on the right. In practice, the angle θ_1 above is selected either by the interpreter or by an automatic procedure to best realize the off-diagonal zeroes of equation (B-5), i.e., to minimize the energy on the rotated data traces, the right side of equation (B-7).

2C × 2C reflection case

Similarly for the reflection case [the one considered explicitly by Alford (1986) and Thomsen (1988)], the reflection principal time series constructed in the same spirit as equation (B-5) but with different details is

$$\mathbf{S}(t) \equiv 2s \mathbf{T}_2 \mathbf{R} \mathbf{T}_1 \begin{bmatrix} A_1^f \otimes A_2^f \otimes A_2^f \otimes A_1^f \otimes w(t - 2t_2^f - 2t_1^f) & 0 \\ 0 & A_1^s \otimes A_2^s \otimes A_2^s \otimes A_1^s \otimes w(t - 2t_2^s - 2t_1^s) \end{bmatrix}. \quad (\text{B-8})$$

Following logic very similar to that above, the result is also very similar:

$$\mathbf{S}(t) = \mathbf{R}(\theta_1) \mathbf{S}_0(2t_1 + 2t_2) \mathbf{R}(-\theta_1). \quad (\text{B-9})$$

It is not required that the propagation filters be equal, nor that the diagonal reflection coefficients [equation (A-17)] be equal, nor that the diagonal transmission coefficients [equation (A-8)] be equal.

2C VSP and reflection cases

Suppose we now have a single source orientation rather than two. Operating on the VSP equation (A-13) from the left by $\mathbf{R}(\theta_1)$ and using equation (B-2), we have

$$\mathbf{R}(\theta_1) \mathbf{s}_0(t_1 + t_2) = \mathbf{S}(t) \mathbf{R}(\theta_1) \mathbf{s}_0. \quad (\text{B-10})$$

If the source is +in-line, then the product on the right is

$$\mathbf{R}(\theta_1) \mathbf{s}_0 = \mathbf{R}(\theta_1) \begin{bmatrix} s \\ 0 \end{bmatrix} = s \begin{bmatrix} \cos \theta_1 \\ -\sin \theta_1 \end{bmatrix}. \quad (\text{B-11})$$

Then, multiplying equation (B-10) from the left by

$$\begin{bmatrix} 1/\cos \theta_1 & 0 \\ 0 & -1/\sin \theta_1 \end{bmatrix}$$

leads to the solution (since $\mathbf{S}(t)$ is diagonal)

$$\begin{aligned} \mathbf{s}(t) &\equiv \begin{bmatrix} A_2^f \otimes T^{11} A_1^f \otimes sw(t - t_2^f - t_1^f) \\ A_2^s \otimes T^{22} A_1^s \otimes sw(t - t_2^s - t_1^s) \end{bmatrix} \\ &= \begin{bmatrix} 1/\cos \theta_1 & 0 \\ 0 & -1/\sin \theta_1 \end{bmatrix} \mathbf{R}(\theta_1) \mathbf{s}_0(t_1 + t_2), \end{aligned} \quad (\text{B-12})$$

which was first derived by Thomsen (1988). In words, the solution on the left for the principal time series is just the rotated, rescaled data on the right.

For the reflection case, a similar solution results. Again, it is not required that the propagation filters be equal, nor that the diagonal reflection coefficients be equal, nor that the diagonal transmission coefficients be equal.

In practice, the angle θ_1 above is selected, either by interpreter or by an automatic procedure, to best realize the criterion that the two components of $\mathbf{s}(t)$ should appear as smoothly stretched and (possibly unsmoothly) scaled versions of one another. An obvious variant can be constructed for a cross-line source. Converted P - S -waves (after stack) are effectively the

same as the in-line case above. Harrison (1992) provides a detailed rotation algorithm, implementing these principles in an automatic way, with additional assumptions in the selection of the angle. Ata and Michelena (1995) show an application to fracture characterization. Multiple-source/single-receiver configurations may be solved using straightforward extensions of this formalism.

It is clear that this algorithm is less robust than the $2C \times 2C$ algorithm because it uses less data. Even so, it may be preferable in practice—for example, in those instances where only two components are available (e.g., old SH source [sic] data, converted-wave data, etc.) or when economic restrictions (e.g., in 3-D surveys) discourage $2C \times 2C$ acquisition.

APPENDIX C

VSP LAYER-STRIPPING ($2C \times 2C$)

Returning to the more general $2C \times 2C$ case of depth-variable orientation of anisotropy, $\theta_2 \neq \theta_1$, we operate on equation (A-14) from the left with $\mathbf{R}(\theta_1)$ and from the right with $\mathbf{R}(-\theta_1)$. The left side of the equation becomes the data, expressed in the coordinates of layer 1 [equation (A-19)], which we denote as

$$\mathbf{S}_1(t_1 + t_2) \equiv \mathbf{R}(\theta_1)\mathbf{S}_0(t_1 + t_2)\mathbf{R}(-\theta_1). \quad (\text{C-1})$$

The rotated equation (A-14) then reads

$$\mathbf{S}_1(t_1 + t_2) = \mathbf{R}(\theta_1)\mathbf{R}(-\theta_2)\mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1)\mathbf{T}_1\mathbf{P}_1 \otimes sw(t) \quad (\text{C-2})$$

$$= \mathbf{R}(\theta_1 - \theta_2)\mathbf{P}_2 \otimes sw(t)\mathbf{R}(\theta_2 - \theta_1)\mathbf{T}_1 \otimes \mathbf{P}_1 \quad (\text{C-3})$$

since $\mathbf{R}(\theta_1)\mathbf{R}(-\theta_2) = \mathbf{R}(\theta_1 - \theta_2)$. The scalar quantities have been commuted into the interior of this last expression, but the propagator \mathbf{P}_1 does not commute. Therefore, the $\theta_2 - \theta_1$ rotation mixes all the components of both fast and slow waves. Incidentally, this expression demonstrates mathematically why independent Alford rotation (Alford, 1989) at each time step does not work; if in fact the orientations did vary at each thin layer, the result would be a hopeless confusion of rerotated and delayed components.

Nonetheless, if θ_1 is properly chosen (see below), the layer 1 rotated data [equation (C-1)] will have approximately null off-diagonal components for times $< t_1^f$. The difficulties caused by resplitting only arise at times greater than this. Following Winterstein and Meadows, we can use the significant energy on the off-diagonal components at times greater than t_1^f to identify t_1^f .

We can now strip off the anisotropy of layer 1. In the current formalism, this layer stripping is accomplished via the mode advance operator:

$$\mathbf{D}_1 = \begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t + \Delta t_1) \end{bmatrix}. \quad (\text{C-4})$$

Operating on equation (C-3) from the right with \mathbf{D}_1 , we have

$$\mathbf{S}_1(t_1 + t_2) \otimes \mathbf{D}_1 = \mathbf{R}(\theta_1 - \theta_2)\mathbf{P}_2 \otimes sw(t)\mathbf{R}(\theta_2 - \theta_1) \otimes \mathbf{T}_1\mathbf{P}_1 \otimes \mathbf{D}_1. \quad (\text{C-5})$$

The propagator matrix, layer stripped, is

$$\begin{aligned} \mathbf{T}_1\mathbf{P}_1 \otimes \mathbf{D}_1 &= \mathbf{T}_1 \begin{bmatrix} A_1^f \otimes \delta(t - t_1^f) & 0 \\ 0 & A_1^s \otimes \delta(t - t_1^s) \end{bmatrix} \\ &\otimes \begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t + \Delta t_1) \end{bmatrix} \\ &= \begin{bmatrix} T_1^{11}A_1^f & 0 \\ 0 & T_1^{22}A_1^s \end{bmatrix} \otimes \delta(t - t_1^f). \end{aligned} \quad (\text{C-6})$$

We define a mode-balancing filter B_1^{sf} such that $B_1^{sf} \otimes T_1^{22}A_1^s = T_1^{11}A_1^f$ (i.e., it reshapes the slow-mode wavelet into the fast-mode wavelet) and a corresponding tensor operator

$$\mathbf{B}_1 \equiv \begin{bmatrix} \delta(t) & 0 \\ 0 & B_1^{sf} \end{bmatrix}. \quad (\text{C-7})$$

Then we apply the mode-balancing operator \mathbf{B}_1 [equation (C-7)] to the downgoing propagator operator of equation (C-5) (from the left). We denote the layer-stripped data on the left by

$$\hat{\mathbf{S}}_1(t_1 + t_2) = \mathbf{S}_1(t_1 + t_2) \otimes \mathbf{D}_1 \otimes \mathbf{B}_1. \quad (\text{C-8})$$

Operating accordingly also on the right side of equation (C-5), we have

$$\begin{aligned} \hat{\mathbf{S}}_1(t_1 + t_2) &= \mathbf{R}(\theta_1 - \theta_2) [\mathbf{P}_2 \otimes T_1^{11}A_1^f \otimes \delta(t - t_1^f) \\ &\otimes sw(t)] \mathbf{R}(\theta_2 - \theta_1). \end{aligned} \quad (\text{C-9})$$

Because of the mode balancing, the matrix of layer-1 propagation filters on the right side of equation (C-6) has become proportional to the identity matrix, so the proportionality scalar function $T_1^{11}A_1^f$ has been commuted into the interior of equation (C-9).

The quantity in brackets on the right of equation (C-9) is the layer-stripped principal time series:

$$\begin{aligned} \hat{\mathbf{S}}(t) &= \left[\mathbf{P}_2 \otimes T_1^{11}A_1^f \otimes sw(t - t_1^f) \right] = sT_1^{11}A_1^f \\ &\otimes \begin{bmatrix} A_2^f \otimes w(t - t_1^f - t_2^f) & 0 \\ 0 & A_2^s \otimes w(t - t_1^f - t_2^f - \Delta t_2) \end{bmatrix}, \end{aligned} \quad (\text{C-10})$$

where $\Delta t_2 \equiv t_2^s - t_2^f$ is the layer-2 delay of the layer-2 pure-mode 22 component. Equation (C-9) then reads

$$\hat{\mathbf{S}}_1(t_1 + t_2) = \mathbf{R}(\theta_1 - \theta_2) \hat{\mathbf{S}}(t) \mathbf{R}(\theta_2 - \theta_1), \quad (\text{C-11})$$

which can be solved as before [equation (B-9)] for the unknown principal time series $\hat{\mathbf{S}}(t)$:

$$\hat{\mathbf{S}}(t) = \mathbf{R}(\theta_2 - \theta_1) \hat{\mathbf{S}}_1(t_1 + t_2) \mathbf{R}(\theta_1 - \theta_2). \quad (\text{C-12})$$

In words, the layer-stripped principal time series $\hat{\mathbf{S}}(t)$ is the layer-stripped data $\hat{\mathbf{S}}_1$, rerotated. The recipe for the layer-stripped data [equation (C-8)] is

$$\hat{\mathbf{S}}_1(t_1 + t_2) \equiv \begin{bmatrix} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t) \\ \mathbf{S}_1^{21}(t) & \mathbf{S}_1^{22}(t) \end{bmatrix} \otimes \begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t + \Delta t_1) \end{bmatrix} \otimes \mathbf{B}_1 \quad (\text{C-13})$$

$$= \begin{bmatrix} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t + \Delta t_1) \\ \mathbf{S}_1^{21}(t) & \mathbf{S}_1^{22}(t + \Delta t_1) \end{bmatrix} \otimes \mathbf{B}_1. \quad (\text{C-14})$$

Equation (C-14) shows how to construct the layer-stripped data $\hat{\mathbf{S}}_1(t_1 + t_2)$: rotate the data $\mathbf{S}_0(t_1 + t_2)$ into the layer-1 coordinate system forming $\mathbf{S}_1(t_1 + t_2)$ [following equation (C-1)], static shift the layer 1 slow modes (column 2 above) forward in time by the one-way delay Δt_1 , and balance the spectra using \mathbf{B}_1 . The angle θ_1 , the bottom of layer 1 (t_1^f), and the delay time (Δt_1) may be determined, for example, by the methods of Winterstein and Meadows (1991a,b). (Remember to do the static shifts before the mode balancing so the phase effects of the delay Δt_1 are not counted twice).

Aside from the formalism used here, which is broadly useful in studies of vector wave motion, a principal contribution of the present work is the recipe for the layer-stripped data, $\hat{\mathbf{S}}_1(t_1 + t_2)$ [equation (C-14)]. If the attenuation of both modes is the same (probably an uncommon circumstance), then the mode-balancing operation is not required and this result reduces to that of Winterstein and Meadows (1991a,b). If the mode attenuations are not the same, yet mode balancing is not performed, then puzzling conclusions may be reached, leading to spurious complexity in the vertical variation of anisotropy direction.

APPENDIX D

REFLECTION LAYER-STRIPPING (2C × 2C)

To layer strip reflection data, consider the observed data tensor in equation (A-18) and rotate it to the layer-1 coordinate system:

$$\mathbf{S}_1(2t_1 + 2t_2) = \mathbf{R}(\theta_1) \mathbf{S}_0(2t_1 + 2t_2) \mathbf{R}(-\theta_1). \quad (\text{D-1})$$

Operating accordingly on the right side of equation (A-18),

$$\begin{aligned} \mathbf{S}_1(2t_1 + 2t_2) &= 2\mathbf{P}_1 \mathbf{R}(\theta_1 - \theta_2) \mathbf{T}_2 \mathbf{P}_2 \otimes \mathbf{R} s w(t) \\ &\otimes \mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1) \mathbf{T}_1 \mathbf{P}_1. \end{aligned} \quad (\text{D-2})$$

At this point, we approximate that $\mathbf{R}(\theta_1 - \theta_2) \mathbf{T}_2 = \mathbf{T}_2 \mathbf{R}(\theta_1 - \theta_2)$ (strictly valid only if $T_2^{11} = T_2^{22}$) and that $T_1^{22}/T_1^{11} = T_2^{22}/T_2^{11}$.

Operating accordingly also on the right side of equation (D-2), we have

$$\begin{aligned} \hat{\mathbf{S}}_1(2t_1 + 2t_2) &= A_1^f \otimes \delta(t - t_1^f) T_2^{11} \mathbf{R}(\theta_1 - \theta_2) \\ &\otimes \mathbf{P}_2 \otimes \mathbf{R} 2s w \otimes \mathbf{P}_2 \\ &\otimes \mathbf{R}(\theta_2 - \theta_1) T_1^{11} \delta(t - t_1^f) \otimes A_1^f. \end{aligned} \quad (\text{D-4})$$

Commuting the scalars into the interior, we have

$$\hat{\mathbf{S}}_1(2t_1 + 2t_2) = \mathbf{R}(\theta_1 - \theta_2) [\hat{\mathbf{S}}(t)] \mathbf{R}(\theta_2 - \theta_1), \quad (\text{D-5})$$

where the quantity in brackets is the layer-stripped principal time series:

$$\hat{\mathbf{S}}(t) = 2A_1^f \otimes \delta(t - t_1^f) T_2^{11} \otimes \mathbf{P}_2 \otimes \mathbf{R} s w(t) T_1^{11} \otimes \mathbf{P}_2 \otimes \delta(t - t_1^f) \otimes A_1^f \quad (\text{D-6})$$

$$= 2s T_2^{11} T_1^{11} A_1^f \otimes \mathbf{R} \begin{bmatrix} A_2^f \otimes w(t - 2t_1^f - 2t_2^f) & 0 \\ 0 & A_2^f \otimes w(t - 2t_1^f - 2t_2^f - 2\Delta t_2) \end{bmatrix}. \quad (\text{D-7})$$

These approximations may be relaxed with a more elaborate, but still deterministic algorithm; however, we apply them to equation (D-1) to expose the essentials.

Then we apply the layer-1 mode-delay operator \mathbf{D}_1 [equation (C-4)] and the mode-balancing operator \mathbf{B}_1 [equation (C-7)] to both the downgoing and the upcoming propagator operators of equation (D-2) (i.e., from both right and left). We denote the layer-stripped data on the left by

$$\hat{\mathbf{S}}_1(2t_1 + 2t_2) = \mathbf{B}_1 \otimes \mathbf{D}_1 \otimes \mathbf{S}_1(2t_1 + 2t_2) \otimes \mathbf{D}_1 \otimes \mathbf{B}_1. \quad (\text{D-3})$$

Solving equation (D-5) for the unknown principal time series $\hat{\mathbf{S}}(t)$,

$$\hat{\mathbf{S}}(t) = \mathbf{R}(\theta_2 - \theta_1) \hat{\mathbf{S}}_1(2t_1 + 2t_2) \mathbf{R}(\theta_1 - \theta_2), \quad (\text{D-8})$$

exactly analogous to equation (C-12). In words, the layer-stripped principal time series $\hat{\mathbf{S}}(t)$ is the layer-stripped data $\hat{\mathbf{S}}_1$, rerotated. The recipe for the layer-stripped data [equation (D-3)] is

$$\hat{\mathbf{S}}_1(2t_1 + 2t_2) = \mathbf{B}_1 \otimes \begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t + \Delta t_1) \end{bmatrix} \otimes \begin{bmatrix} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t) \\ \mathbf{S}_1^{21}(t) & \mathbf{S}_1^{22}(t) \end{bmatrix} \\ \otimes \begin{bmatrix} \delta(t) & 0 \\ 0 & \delta(t + \Delta t_1) \end{bmatrix} \otimes \mathbf{B}_1 \quad (\text{D-9})$$

$$= \mathbf{B}_1 \otimes \begin{bmatrix} \mathbf{S}_1^{11}(t) & \mathbf{S}_1^{12}(t + \Delta t_1) \\ \mathbf{S}_1^{21}(t + \Delta t_1) & \mathbf{S}_1^{22}(t + 2\Delta t_1) \end{bmatrix} \otimes \mathbf{B}_1. \quad (\text{D-10})$$

Equation (D-10) shows how to construct the layer-stripped data $\hat{\mathbf{S}}_1(2t_1 + 2t_2)$: rotate the data $\mathbf{S}_0(2t_1 + 2t_2)$ into the layer-1 coordinate system forming $\mathbf{S}_1(2t_1 + 2t_2)$ [following equation (D-1)], static shift the mismatched traces forward in time by the one-way delay Δt_1 and the slow (22) trace forward by the full two-way delay $2\Delta t_1$, and balance the spectra using \mathbf{B}_1 twice. The angle θ_1 , the bottom of layer 1 (t_1^f), and the delay time (Δt_1) may be determined, for example, by the methods of Winterstein and Meadows (1991a,b).

The mode-balancing filter \mathbf{B}_1 may be derived using the methods discussed in the main text.

APPENDIX E

2C LAYER-STRIPPING

If only a single source orientation is used, a solution in principle may still be found. Consider first the VSP case. For the first layer, the angle θ_1 and delay Δt_1 may be found using the methods of Appendix B. Then, operating on equation (A-13) from the left with $\mathbf{R}(\theta_1)$, the data vector (as rotated into layer-1 coordinates) may be denoted by

$$\mathbf{s}_1(t_1 + t_2) = \mathbf{R}(\theta_1)\mathbf{s}_0(t_1 + t_2). \quad (\text{E-1})$$

For example, if the source is +in-line, then the rotated equation (A-13) reads

$$\mathbf{s}_1(t_1 + t_2) = \mathbf{R}(\theta_1)\mathbf{R}(-\theta_2)\mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1)\mathbf{T}_1\mathbf{P}_1 \\ \otimes sw(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (\text{E-2})$$

The rotated data have the pure modes (fast and slow) on their 1 and 2 components, respectively, in the layer-1 time window only; at later times, both components have both modes present. Operating on equation (E-2) at all times, just as we did with the mode-advance operator in equation (C-4), we have

$$\mathbf{D}_1 \otimes \mathbf{S}_1(t_1 + t_2) = \begin{bmatrix} \mathbf{S}_1^{11}(t) \\ \mathbf{S}_1^{21}(t + \Delta t_1) \end{bmatrix} \quad (\text{E-3})$$

$$= \mathbf{D}_1 \otimes \mathbf{R}(\theta_1 - \theta_2)\mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1)\mathbf{T}_1\mathbf{P}_1 \\ \otimes (\mathbf{D}_1 \otimes \mathbf{D}_1^{-1}) \otimes sw(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (\text{E-4})$$

where \mathbf{D}_1^{-1} is the inverse of \mathbf{D}_1 . Using equation (C-6) and assuming that the mode-balancing discussed in equation (C-7) is unnecessary, we commute the layer 1 propagation operators into the interior of this expression and recognize the principal time series matrix $\hat{\mathbf{S}}(t)$ [equation (C-10)], yielding

$$\mathbf{D}_1 \otimes \mathbf{S}_1(t_1 + t_2) \\ = \mathbf{D}_1 \otimes \mathbf{R}(\theta_1 - \theta_2)[\hat{\mathbf{S}}(t)]\mathbf{R}(\theta_2 - \theta_1)\mathbf{D}_1^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (\text{E-5})$$

If angles θ_1 and θ_2 , and delays Δt_1 and Δt_2 , were known, this would constitute two equations in two unknowns $[\hat{\mathbf{s}}_1(t)$ and $\hat{\mathbf{s}}_2(t)]$ at every time t . Alternatively, with additional assumptions about the layer 2 propagation filters, e.g., that $A_2^f = A_2^s$,

it becomes an overdetermined set of equations for the layer 2 parameters θ_2 and Δt_2 . In either case, the layer-stripping procedure is much more complicated in this 2C case than in the $2C \times 2C$ case, involving much more than simple rotations and static shifts and subject to many more instabilities in practice. Simple rescaling [equation (B-12)] does not work since the matrices to the left of the vector on the right side of equation (E-5) are not diagonal. Of course, the data may be static shifted, as in equation (E-3), but this does not solve for the principal time series at times later than the bottom of layer 1. The problem as posed becomes one of nonlinear inversion of noisy data with a lot of ambiguity present in the typical application.

This result is puzzling at first since it seems that if we can solve for the anisotropy in layer 1, we just rotate our coordinate system to that orientation, remove the time delay, and repeat the process with layer 2. However, we cannot repeat the process with layer 2 because the layer-1 process requires the assumption that the layer contains only two events (one fast, one slow) propagating. This is not true for deeper layers, where we have multiple events propagating at each of the two velocities. The rotation and time shifting have not properly aligned these multiple events in the layer-2 time window. Algebraically, this shows in the appearance of operators \mathbf{D}_1 and \mathbf{D}_1^{-1} in equation (E-5).

Considering now the reflection 2C case with just +in-line source, equation (A-18) becomes

$$\mathbf{S}_0(2t_1 + 2t_2) = 2\mathbf{R}(-\theta_1)\mathbf{P}_1 \otimes \mathbf{R}(\theta_1 - \theta_2)\mathbf{T}_2\mathbf{P}_2 \\ \otimes \mathbf{R}\mathbf{P}_2 \otimes \mathbf{R}(\theta_2 - \theta_1)\mathbf{T}_1\mathbf{P}_1 \\ \otimes \mathbf{R}(\theta_1)sw(t) \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (\text{E-6})$$

Proceeding as above, we find

$$\mathbf{D}_1 \otimes \mathbf{S}_1(2t_1 + 2t_2) = \mathbf{D}_1 \\ \otimes \mathbf{R}(\theta_1 - \theta_2)[\hat{\mathbf{S}}(t)]\mathbf{R}(\theta_2 - \theta_1)\mathbf{D}_1^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (\text{E-7})$$

where $\hat{\mathbf{S}}(t)$ is the two-way principal time-series matrix [equation (D-6)]. The remarks following equation (E-5) apply equally to this result.