APPENDIX

In this Appendix, the basic mathematics behind the wavefield separation, deconvolution and far offset processing of VSP data will be reviewed. These will consist of the median, $K$-$L$, $F$-$K$ and $T$-$P$ filtering, VSP deconvolution and the matrix equations involved in the hodogram-based and time-variant polarizations.

A.1 Median filtering

One of the wavefield separation methods performed on the near- and far-offset data is the one-dimensional median filter combined with a bandpass filter. The purpose of the application of the bandpass filter is to eliminate the median filter "whiskers" (Hardage, 1985) resulting from the non-linear operation. The theoretical basis of the median filter has been reviewed in Arce et al. (1986), Fitch et al. (1984), Arce and McLoughlin (1984), Gallagher and Wise (1981) and Nodes and Gallagher (1982).

The input to the median filter is a selected window of data. The length of the window can be even or an odd number of points ($2N$ or $2N+I$). The two ends of the input time series are padded with $N$ additional points in order to accommodate the centre location of the window being situated at either end. The input window of data is sorted according to magnitude with the centre value of the sort being termed the median value. For the odd point filter, the median value at the centre of the windowed time series becomes the new
value of the output series. When $N$ is even, the mean of the two middle median values is the output of the filter. This new point of the output data is placed at the location of the centre of the window of the input series. For the 1-D median filter application, a new output time series is generated as the window slides across the input series, one point at a time.

The median filter can be defined as the rearrangement of the windowed time series according to size. The output of the non-recursive median filter, $y(t)$, is given by (Arce et al., 1986)

$$y(t) = \text{median} \{x(t-N),...,x(t-I),x(t),x(t+I),...,x(t+N)\}$$

and the recursive filter is given as

$$y(t) = \text{median} \{y(t-N),...,y(t-I),x(t),x(t+I),...,x(t+N)\}$$

for a window of length $2N+1$ samples centred at location $t$ of the input time series.

The type of median filter used for the wavefield separation operation in this thesis is a non-recursive median filter. The operation of the recursive filter would differ from the operation of the non-recursive filter as the previously determined output data would be used to compute further output data. The options that are available for the non-recursive median filters include normal and tapered median filters. The tapered filters would filter the sorted values using a boxcar, triangular or cosine filter. The sum of a pre-specified number of the central tapered values would be the new median value. A tapered median filter would eliminate the need for the post-median bandpass filter.
When the median filter operation is applied to the $Z(-TT)$ data, the divergent upgoing waves appear as a triangular anomaly. The triangular anomaly can be filtered out using a suitable length median as shown in Figure 5-27 of Hardage (1985). The median filter operation smooths the amplitude (phase) variations of the first break downgoing event over several traces. A scaling program is applied to restore the $Z_{down}(-TT)$ data to an amplitude similar to a selected window of data in the $Z(-TT)$ input data. The window usually is comprised of a zone surrounding the first break wavelet. This zone is restored to the amplitude range of a similar window around the $Z(-TT)$ first break wavelet using a multiplicative factor determined using a least-squares ratio fit or a ratio computed as the inverse of the absolute amplitudes over the window.

Following the scaling of the $Z_{down}(-TT)$ data, subtraction of the $Z_{down}(-TT)$ from the $Z(-TT)$ data yields the $Z_{up}(-TT)$ data.

### A.2 Karhunen-Loeve (K-L) filtering

The rationale for using the $K-L$ transform for wavefield separation has been explained in chapter 2. The eigenanalysis will dissect the VSP data with $N$ traces into $N$ eigenimages ($N$ eigenvalues and corresponding $N$ eigenvalues). The degree of linear coherency is reflected in the magnitude of the eigenvalue. An excellent review of the method as it is used in VSP work is given in Hardage (1992) and Jackson et al., (1991). The method has also been referred to as eigenvector coding (Kirlin, 1987).
For VSP data, we can have $N$ traces and $M$ time sample points. In the derivation, we assume that there are more time samples than traces (depth recordings). The data matrix is formed by placing the VSP depth traces to be the rows of the data matrix, $\{x_i(t), i=1,\ldots,N; 1 \leq t \leq M\}$. We form the cross-energy or covariance matrix as the outer product of the data matrices

$$\mathbf{\tilde{M}} = \mathbf{\tilde{X}} \mathbf{\tilde{X}}^T$$

The covariance matrix can be spectrally decomposed (using singular value decomposition or SVD) to form

$$\mathbf{\tilde{M}} = \mathbf{\tilde{W}} \Lambda \mathbf{\tilde{W}}^T$$

The Karhunen-Loeve or the principal components are

$$\mathbf{\tilde{K}} = \mathbf{\tilde{W}}^T \mathbf{\tilde{X}}$$

where

$$\mathbf{\tilde{W}} = \begin{bmatrix} w_1 & w_2 & \ldots & w_N \end{bmatrix}$$

is the eigenvector matrix calculated using SVD of the cross-energy matrix. The eigenvectors are

$$w_1 \ w_2 \ \ldots \ w_N$$

The eigenvalue matrix, $\Lambda$, has as the trace of the matrix, the corresponding eigenvalues (also
calculated during the SVD of the cross-energy matrix)

\[ \lambda_1 \lambda_2 \ldots \lambda_N \]

The eigenvalues are examined to determine the number of eigenvectors to be included in the reconstruction. In the case of the isolation of the downgoing events from the Z(-TT) data, the first few (corresponding to the largest eigenvalues) eigenvectors are chosen. The downgoing events can be reconstructed by performing the inverse transform using only those eigenvectors selected to represent the downgoing events. If the first \( J \) eigenvectors are chosen, then

\[ \tilde{X}_{\text{recon}} = \tilde{W} \tilde{K} \]

where only the \( J \) chosen principal components and eigenvectors are used. This means that some of the columns of the eigenvector matrix and some of the rows of the principal component matrix are not used or zeroed, depending which eigenvalues were chosen to represent the downgoing events only.

In summation form, this would appear as

\[ x_{i\text{recon}}(t) = \sum_{j=1}^{J} w_{ij} k_j(t) \quad i = 1, \ldots, N; 1 \leq t \leq M \]

and in general, partial reconstruction can be seen as (for a misfit analysis; Jones, 1985)

\[ x_{i\text{recon}}(t) = \sum_{j=m+1}^{p} w_{ij} k_j(t) \quad i = 1, \ldots, N; m \leq p \leq N \]
for integers \( m \) and \( p \) whose choice are dependent on the aim of the reconstruction.

### A.3 F-K filtering

In 2-D wavefield transformations, a linear event in the \( Z-t \) domain becomes a linear event in the \( F-K \) domain. The transformation equations for the forward and reverse \( F-K \) transforms are

\[
V(k_z, \omega) = \int \int v(z,t) \ e^{i(\omega t - k_z Z)} \ dz \ dt
\]

for the forward transform and

\[
v(z,t) = \int \int V(k_z, \omega) \ e^{-i(\omega t - k_z Z)} \ dk_z \ d\omega
\]

for the reverse transform.

The term \( F-K \) is used loosely since the Fourier transform is usually expressed in terms of \( \omega \) and \( k_z \) (Hu and McMechan, 1987). The up- and downgoing linear events contained in the VSP data are mapped into linear events in the positive and negative \( K \) quadrants of the \( F-K \) domain. It can be shown to be the case from the following brief derivation. From Robinson (1967), the equation for a line will be

\[
t = \frac{Z}{V} + t_0
\]
and we can form a delta function

\[ \delta \left( t - \frac{Z}{V} - t_0 \right) \]

If we insert this function into the 2-D Fourier transform, then

\begin{align*}
V(k_z, \omega) &= \iint \delta \left( t - \frac{Z}{V} - t_0 \right) e^{i(\omega \tau - k_z Z)} \, dZ \, dt \\
&= \int e^{i(\frac{\omega}{V} \tau_0 - k_z Z)} \, dZ \\
&= 2\pi \, e^{i\omega \tau_0} \delta(\omega - V k_z)
\end{align*}

This is the equation of a line in the \( F-K \) domain passing through the origin, has a magnitude given by the real part of the equation (the delta function) and is associated with the phase equal to \(-\omega \tau_0\). The phase is associated with the location of the line within the \( Z-t \) domain and is linked to \( \tau_0 \) which will be discussed later in the \( \tau-P \) domain.

In chapter 2, numerous examples of the downgoing events being clustered in a tight linear group in the positive \( K \) quadrant are shown. The slope of the linear events in the \( F-K \) domain yields the apparent velocity, \( V \), on the \( Z-t \) plot (VSP FRT display) since \( \omega = k_z \cdot V \). The concept of spatial aliasing was discussed in chapter 2 along with a numerical example (also see DiSiena et al., 1984 and pages 104-114 of Hardage 1985). To avoid
aliasing in the F-K domain, one should use a depth increment (for the sonde locations) in consideration of the equation (Hardage, 1985)

$$\Delta Z \leq \frac{V_{min}}{2f_{max}}$$

where $V_{min}$ is the minimum strata velocity one would expect to encounter during the VSP run (check the sonic log which is usually run before the VSP) and $f_{max}$ is the maximum frequency one would expect in the data (what bandpass filter will be used in the final IPP panels?).

**A.4 \(\tau\)-\(P\) filtering**

The \(\tau\)-\(P\) filtering is related to the \(F\)-\(K\) filtering method as was shown by Figure 2.25. In that figure, the slowness limits used in the \(\tau\)-\(P\) filtering were shown in the \(F\)-\(K\) domain as a "pie-slice" accept zone. After all, doesn't the equation $\omega = k_z \cdot V$ translate into $k_z = \omega \cdot P$?

The \(\tau\)-\(P\) filter is also called the "slant stack" since in the \(Z\)-\(t\) domain (the VSP data), the slant stack domain is calculated by performing individual sums along lines defined by $t = P_0 Z + \tau_0$. $P_0$ is related to the slope of the line of integration which is $\tan \alpha$ ($\alpha$ being the angle that the line of integration makes with the \(Z\) axis) and $\tau_0$ being the \(t\)-intercept. This maps a linear event into a point in the \(\tau\)-\(P\) domain. The calculation continues to include lines of integration of all slopes (both positive and negative relating to the down- and upgoing events, respectively) and \(t\)-intercepts ($\tau$'s) and all \(Z\) values. The depth (\(Z\)) values need not be increasing by a constant increment which means that unequally spaced sonde locations pose no problem for this transformation (that would cause a problem with the \(F\)-\(K\) filter since the
fast fourier transform, FFT, desires both constant $\Delta Z$ and $\Delta t$.

The $\tau$-$P$ transform is defined to be (Hu and McMechan, 1987; Robinson, 1967; Kappus et al., 1990; Turner, 1990; Carswell and Moon, 1989; Hardage, 1992; Deans, 1983)

$$U(P, \tau) = \int v(z, Pz + \tau) \, dz$$

and the inverse transform is

$$v(Z,t) = \int \frac{d}{dt} H[U(P,t-Pz)] \, dP$$

In the inverse transform (described in detail in Robinson, 1967 and Hu and McMechan, 1987), the trace increment, $\Delta Z$, can be respecified to another value other than the input value for the forward transform. This enables global or local trace interpolation which can attempt to infill missing depth levels (Hu and McMechan, 1987).

Since the up- and downgoing VSP events are opposite in sign with respect to apparent velocity on the FRT display, one can specify the input $P$ range to be either sign in the forward transform and therefore perform wavefield separation. As shown in chapter 2, trace interpolation can be performed and then another method to perform up- and downgoing event separation can be used.
The use of the transform is data dependent and I favour using all of the above methods (median, K-L, F-K included) to create the IPP’s and then to incorporate the interpretation to aid in deciding which method or combination of methods is best for the data.

A.5 VSP deconvolution

The use of the downgoing events, $Z_{\text{down}}(-TT)$, to design a deconvolution operator has been called "downward-travelling wave train deconvolution" (Balch and Lee, 1984), "Up over Down deconvolution", "special VSP deconvolution" amongst other names. The concept is discussed in Balch and Lee (1984), Hardage (1985) and Hubbard (1979). In the simplest case of the VSP data containing only primary and surface-generated multiple events, the downgoing events represent all that is needed for the deconvolution of the upgoing events.

From Gaiser et al., 1984, the reason for the name "up over down" deconvolution can be seen. If we consider a VSP recording at a single level which has upgoing events, $U(Z,t)$ (originating from reflections below the sonde), and downgoing events, $D(Z,t)$ (the primary downgoing event plus surface generated multiples), then the composite wavefield seen on the trace from the $Z(FRT)$ display is

$$v(Z,t) = D(Z,t) \otimes [1 - RC(Z,t)]$$

where $RC(Z,t)$ is the reflectivity coefficient series in time and the symbol $\otimes$ denotes a convolution operation. If we design a deconvolution operator from the downgoing waves,
namely $D^1(Z,t)$, and convolve this with the $v(Z,t)$, then the operation will produce

$$D^{-1}(Z,t) \otimes v(Z,t) = 1 - RC(Z,t)$$

However, we can also do this in the Fourier domain and convolution with an inverse operator of the downgoing events is equivalent to division in the Fourier domain. The process would then be

$$\frac{U(Z,\omega)}{D(Z,\omega)} = U(Z,\omega)_{decon}$$

The Fourier transform of the $Z_{up}(\cdot TT)$ is divided by the Fourier transform of the $Z_{down}(\cdot TT)$ data; hence the name "up over down deconvolution".

Where do we have problems? The downgoing multiple event resulting in an upgoing interbed multiple exists on the sonde locations starting from the top generating interface (the interface that reflects the primary upgoing wave back down) to deeper sonde locations. The upgoing interbed event exists on traces from the lower generating interface sonde location and upwards to the shallowest level. What this means is that the downgoing interbed events needed to evaluate the corresponding upgoing interbed multiple are not present on the traces recorded shallower than the top generating interface; the upgoing interbeds at these levels may not be attenuated.
A.6 Hodogram-based single angle polarizations

The far-offset VSP data recorded on the X(FRT), Y(FRT), and Z(FRT) are polarized in order to isolate the downgoing P-wave (or SV) events onto a single panel, HMAX'(FRT), as reviewed in chapter 2. The polarization is done by two series of data rotations using hodogram (Hardage, 1985; DiSiena et al., 1981; Balch and Lee, 1984; Gaiser et al., 1984; DiSiena et al., 1984; Hinds et al., 1989a) analysis. The series of rotations are designed on the primary downgoing wavelet since it is that type of event that we desire to isolate. Our assumption is that the first break wavelet is not "contaminated" by other wavefield which says that we do not want nasty upgoing primaries (which, after all, is our final target) to get in the way of our work!

As reviewed in chapter 2, the hodogram is constructed using a window of data around the first break wavelet. This is done interactively using a colour coded display that enables the interpreter/processor to understand what part of the hodogram relates to individual portions of the windowed data. The angle used in the rotation matrix is chosen using a line through the hodogram display that can be rotated interactively plus the output data window is redisplayed each time the line is rotated. When the operator is satisfied with the polarization, the angle is saved automatically. Many papers suggest least-squares fitting routines to make the angle decision; however, the essence of interpretive processing is to make decisions based on viewing in detail the effect of the processes on the data. This would negate the attitude if "black-box" methods were used.

Once the angle, \( \theta \), is chosen then all of the time samples of the X(FRT) and Y(FRT) data
are rotated into the $H_{\text{MAX}}(\text{FRT})$ and $H_{\text{MIN}}(\text{FRT})$ output data according to the matrix equation

$$\begin{pmatrix}
H_{\text{MAX}}(t) \\
H_{\text{MIN}}(t)
\end{pmatrix} =
\begin{pmatrix}
X(t) & Y(t)
\end{pmatrix}
\begin{pmatrix}
\cos(\theta_1) & -\sin(\theta_1) \\
\sin(\theta_1) & \cos(\theta_1)
\end{pmatrix}$$

The polarization of the $H_{\text{MAX}}(\text{FRT})$ and $Z(\text{FRT})$ data into the $H_{\text{MAX}}'(\text{FRT})$ and $Z'(\text{FRT})$ data follows a similar procedure. The important aspect to note is that a single angle is used to matrix rotate the entire trace and that the angle is based on the primary downgoing P-wave event.

**A.7 Time-variant polarization**

We could begin to estimate a pseudotime-variant polarization using the same software as the hodogram analysis if we could track all of the upgoing wave events on the $Z'_{up}(\text{FRT})$ and $H_{\text{MAX}}'_{up}(\text{FRT})$ data. For a given upgoing event that spans the traces beginning at the trace for the interface to the shallowest trace, a hodogram analysis is done on every trace for that particular upgoing event. This is done for all of the upgoing events and an angle versus time function is built up for all of the traces. Interpolation of the angles in between the given times for the angles on a single trace gives us the time-variant angles for that trace. The problems with this approach is that the upgoing events are being dissected by other types of events and the signal to noise ratio of the upgoing events and the background may not be high. This suggests that ray-tracing and modelling should be done using all available velocity and model information from the zero and far-offset data.
The ray-tracing algorithm used in the ion method for the far-offset data in chapters 4, 5, and 6 was paraxial ray tracing (Beydoun, 1985; Beydoun and Keho, 1987; Cerveny at al, 1977; Cerveny and Hron, 1980; Cerveny at al., 1982; Cerveny, 1985). The method was suitable for the research since the ray tracing could be sparsely done and then curvature corrections could be used to estimate ray tracing at nearby locations.

After ray-tracing was done through a model designed from the zero-offset derived velocities and incorporating model restrictions given by interpreted far-offset first break times and upgoing reflection times, a series of polarization angles for various reflections arriving on a single trace would be computed. The single polarization angle for a trace, $\theta$, would be replaced by a time-variant angle, $\theta(t)$, for a given trace. No polarization calculation would be done prior to the first break and a constant angle would be used for upgoing reflections "below" the bottom hole "interface". The matrix equation for the time-variant polarization would now include a time varying rotation angle

$$
\begin{pmatrix}
Z''(t) \\
HMAX''(t)
\end{pmatrix}
= \begin{pmatrix}
Z_{DEROT}(t) & HMAX_{DEROT}(t)
\end{pmatrix}
\begin{pmatrix}
\cos(\theta(t)) & -\sin(\theta(t)) \\
\sin(\theta(t)) & \cos(\theta(t))
\end{pmatrix}
$$

The $Z''_{\text{up}}(+TT)$ data are then used for interpretation and input into the VSP-CDP transform and into the Kirchhoff migration.
REFERENCES

AGAT Laboratories, 1988, Table of formations of Alberta: AGAT Laboratories, Calgary.


361


delineation: Fort St. John Graben area. Canadian Society of Exploration
Geophysicists Annual Meeting, Calgary.


373


Shuck, T., 1988a, Three component VSP processing, Part I: Methodology. IV Congreso Venezolano de Geofisica, Caracas, September 11-15, 239-244.

Shuck, T., 1988b, Three component VSP processing, Part II: Real data example. IV Congreso Venezolano de Geofisica, Caracas, September 11-15, 245-249.
Smidt, J.M., 1989, VSP processing with full downgoing wavefield deconvolution applied to the total wavefield. First Break, 7, 6, 247-257.


