THEORY OF THE STRATA PROGRAM

1.0 INTRODUCTION

The basic concept of the STRATA program can easily be seen from Figure 1.1. From two input streams consisting of stacked seismic data and velocity/density information (in the form of well logs or RMS velocities), derive a basic velocity model of the earth’s subsurface and then use this model to aid in a full inversion of the seismic data. Despite the simplicity of this goal, there are many complications that can arise in achieving it. The program therefore includes algorithms to do any of the following things:

- Synthetic seismogram generation
- Interactive stretching and squeezing of well logs
- Wavelet extraction
- Post-stack seismic processing
- Seismic picking
- Model building with both vertical and lateral interpolation
- Inversion using several inversion algorithms.

The way in which these items interact is shown in Figure 1.2.

The program has been designed with high quality graphics and user-friendly menu control. For information on the running of the program and on specific menu items, refer to earlier parts of this manual. In this section we will describe the theory of the key algorithms used in STRATA. The topics covered will include the convolutional model, deconvolution, inversion, and attributes. It is assumed that other parts of the program, such as the stretching and squeezing of logs, processing steps like filtering and AGC, and model building are either fairly well understood by the user or are self-explanatory when using the program.
Figure 1.1 The basic concept of STRATA.
Figure 1.2 Interrelationships of the components of STRATA.
2.0 THE CONVOLUTIONAL MODEL

The basic convolutional model of the seismic trace can be written

\[ T(i) = \sum_j r(j) W(i - j + 1) + n(i) \]  \hspace{1cm} (2-1)

where

- \( r(j) \) = the zero-offset reflectivity of the earth expressed as a time series
- \( W(i) \) = the seismic wavelet, assumed to be constant
- \( n(i) \) = additive measurement noise

Note that in this model, multiples are assumed to be negligible.

Inversion can be thought of as the process of determining the reflectivity, \( r(j) \), given the seismic trace \( T(i) \). In equation 2-1, the reflectivity is related to the impedance of the earth by the formula

\[ R(j) = \frac{I(j) - I(j - 1)}{I(j) + I(j - 1)} \]  \hspace{1cm} (2-2)

where

- \( I(j) \) = \( \rho(j)V(j) \)
- \( \rho \) = density
- \( V \) = P-wave velocity

As we shall see in section 5, the objective of inversion is to estimate the earth’s velocity from the seismic traces themselves. Obviously, this first entails extracting an estimate of the reflectivity from the convolutional model, and therefore gets us into the related process of deconvolution, to be discussed in section 4. Before looking at either deconvolution or inversion, however, let us take a more detailed look at the two major components of the convolutional model, the wavelet and the reflectivity.

2.1 The Reflectivity

When the energy in a seismic source is released, either by an explosion or the sudden impact of a base plate with the ground, this energy is transmitted through the earth as an elastic wave.

The simplest type of wave to understand is the compressional wave, which is like sound, an acoustic disturbance. The ability of a rock to allow the passage of an acoustic wave is given by its acoustic impedance (analogous to the concept of electrical impedance in an electrical circuit) which is the product of the compressional wave velocity and density of the rock. As may be expected, the more competent the rock, the higher its acoustic impedance. For example, a compacted sandstone will generally have a higher acoustic impedance than a shale (porosity also has an effect on velocity: the higher the porosity, the lower the velocity).
A seismic reflection occurs whenever there is a change in acoustic impedance. If we consider an acoustic wave of amplitude one, which strikes the boundary between two layers of different acoustic impedance, the reflected amplitude can be written as in equation 2-2 above.

Equation 2-2 shows that the reflection coefficient can be positive or negative, depending on which of I(j-1) or I(j) is larger, and that its absolute magnitude cannot exceed one. The preceding analysis has been for a downgoing wave, which strikes the boundary from above. The reflection coefficient for the upcoming wave will simply be negative of the downgoing reflection coefficient.

Obviously, not all the incident amplitude can be reflected (although this is almost the case at the air-water interface in marine recording, where the reflection coefficient approaches one). The size of the transmitted amplitude is the difference between the incident amplitude and the reflected amplitude, or:

\[ T(j) = 1 - r(j) = \frac{2I(j-1)}{I(j) + I(j-1)} \]  

(2-3)

Notice that if \( r \) is negative, \( T \) becomes greater than 1, that is, greater than the incident amplitude! The solution to this apparent contradiction can be found by noting that only energy is conserved in a physical system, and not amplitude. The energy is related to the down and upgoing paths. For the downgoing path, we have just seen that the sign of \( r \) changes. Thus, the total transmission coefficient can be written as

\[ T_{2\text{-way}} = (1 - r(j))(1 + r(j)) = (1 - r(j))^2 \]  

(2-4)

As we can see, equation 2-4 tells us that the total transmitted amplitude is always less than one. Also, notice that the effect of this transmission loss gets greater as the number of layers increases. The total effect can be shown to be:

\[ T_{\text{total}} = \prod_{j=1}^{N} (1 - r(j)^2) = (1 - r_{AV}^2)^N \]  

(2-5)

where \( N \) = number of layers  
\( r(j) \) = reflection coefficient at \( j^{th} \) interface  
\( r_{AV} \) = average reflection coefficient

That is, the total transmission loss is the product of the transmission losses through each layer above the interface in which we are interested. Equation 2-5 also suggests that, if we replace these individual reflection coefficients with the average coefficient, the transmission effect is approximately equal to the \( N^{th} \) power of the two way transmission coefficient. To get an idea of what effect this has, consider a stack of 100 layers (which would correspond to 200 ms for a 2 ms sample rate). If the average reflection coefficient is taken to be 0.1, the transmission scaling factor is 0.366. However, if the average coefficient is 0.05, the total scaling is just 0.779.
It is important to note that equations 2-2 through 2-5 only hold if the P-wave is vertically incident on a boundary. At non-vertical incidence, mode conversion occurs, and reflected and transmitted shear-waves are created. This effect is very important to understand when precise measurements of seismic lithology are to be determined. For the post-stack case, we shall assume that equation 2-2 accurately predicts reflection amplitude.

2.2 The Seismic Wavelet

The first step in the recording of seismic data is the generation of a seismic source signature. On land, the two most common sources are dynamite, which is usually buried for greater penetration, and vibroseis, which is a vibratory surface source of long time duration. In marine recording, the airgun is the most common source. If we know the precise shape of the seismic source, we may use it in inverse in the deconvolution process. This is referred to as deterministic deconvolution. However, we often do not have this recorded information, and must use a statistical deconvolution technique to estimate the source shape. (Both of these types of deconvolution are included in STRATA.)

A wavelet is defined by its amplitude and phase spectra. The types of phase spectra that we will consider are: zero-phase, constant-phase, minimum phase, or non-minimum phase (anything else).

To understand the concept of a zero-phase or constant-phase wavelet, simply think of summing a number of sinusoidal waves of differing amplitude and frequency, all zero-phase or constant phase (i.e. 90°). However, minimum phase is not as simple a concept to understand, and there are many different approaches to describing it. The simplest is an intuitive approach which states that a minimum phase wavelet does not exist before time zero (mathematically, we say the wavelet is causal) and had the majority of its energy concentrated at the front. Indeed, the term “minimum-phase” refers to the fact that, of all the causal wavelets with the same amplitude and differing phase spectra, the one with its phase spectrum closest to zero is the minimum-phase wavelet. It turns out that a lot of valuable information of the phase of a wavelet can be derived by using the simple two-point wavelets:

\[ W_1 = (2, 1) \]  \hspace{1cm} (2-6) \\
\[ W_2 = (1, 2) \]  \hspace{1cm} (2-7)

Wavelet \( W_1 \) is minimum phase, whereas wavelet \( W_2 \) is non-minimum phase. Indeed wavelet \( W_2 \) is termed maximum phase. Let us now derive information about the amplitude and phase spectra of these two wavelets.

To start with, consider the Z-transform of the wavelet given in equation 2-6:

\[ W(Z) = 2 + Z \]  \hspace{1cm} (2-8)
Now, make the substitution
\[ Z = e^{j\omega t} = \cos(\omega t) - j\sin(\omega t) \quad (2-9) \]

where \( j = \sqrt{-1} \)
\( \omega = 2\pi f \)

Notice that this is a sensible substitution to make, since the complex exponential of 2-9 involves cosine and sine functions, which are the basis of the Fourier transform.

Now, by substituting equation 2-9 into equation 2-8, we get the Fourier transform of the wavelet:

\[ W(e^{j\omega t}) = 2 + \cos(\omega t) - j\sin(\omega t) \quad (2-10) \]

where
\[ \text{Re} = \text{Real component of Fourier transform} = 2 + \cos(\omega t) \]
\[ \text{Im} = \text{Imaginary component of Fourier transform} = \sin(\omega t) \]

To derive the amplitude and phase spectra of the Fourier transform from the real and imaginary components, we transform from rectangular to polar coordinates in the following way:

\[ W(\omega) = |W(\omega)| e^{j\theta(f)} \quad (2-11) \]

where
\[ |W(\omega)| = \text{Amplitude spectrum of W} = \left(\text{Re}^2 + \text{Im}^2\right)^{1/2} = \left[4 + 4\cos(\omega t) + \cos^2(\omega t) + \sin^2(\omega t)\right]^{1/2} = \left[5 + 4\cos(\omega t)\right]^{1/2} \]
\[ \theta(f) = \text{Phase spectrum of W} = \arctan[\text{Im}/\text{Re}] = \arctan[\sin(\omega t)/(2 + \cos(\omega t))] \]
In a similar fashion, the Z and Fourier transforms of \( W_2 \) are given by:

\[
W_2(Z) = 1 + 2Z \quad (2-12)
\]

\[
W_2(\omega) = |W_2(\omega)| e^{j\Phi_2(\omega)} \quad (2-13)
\]

where

\[
|W_2(\omega)| = \sqrt{\left[5 + 4\cos(\omega t)\right]^2}
\]

\[
\Phi_2(\omega) = -\arctan\left[\frac{2\sin(\omega t)}{1 + 2\cos(\omega t)}\right]
\]

Plots of the amplitude and phase curves of \( W_1 \) and \( W_2 \) are shown in Figure 2.1. Notice that the amplitude spectra of the two signals are identical. However, their phase spectra differ. As is clear from Figure 2.1, wavelet \( W_1 \) has the minimum phase spectrum. Wavelet \( W_2 \) is non-minimum phase. In fact, it has the property of being maximum phase.

But what about the wavelets with more than two sample values? Obviously, three is the next highest number to consider. From there, we generalize to \( n \) values.

First, consider the Z-transform of a three-point wavelet:

\[
A(Z) = a_0 + a_1Z + a_2Z^2 \quad (2-14)
\]

The Fourier transform can then be written as:

\[
A(\omega) = a_0 + a_1e^{j\omega} + a_2e^{-2j\omega} \quad (2-15)
\]

\[
= (a_0 + a_1\cos(\omega t) + a_2\cos(2\omega t)) - j(a_1\sin(\omega t) + a_2\sin(2\omega t))
\]

The amplitude and phase spectra can be computed using the relationships derived earlier.

Let us create our three point wavelets by convolving wavelets \( W_1 \) and \( W_2 \) in the three possible combinations:

\[
W_3 = W_1 * W_1 = (4, 4, 1) \quad (2-16)
\]

\[
W_4 = W_1 * W_2 = (2, 5, 2) \quad (2-17)
\]

\[
W_5 = W_2 * W_2 = (1, 4, 4) \quad (2-18)
\]

The amplitude spectra of these three wavelets are shown at the top of Figure 2.2, and are identical. This makes sense when we realize that the amplitude spectra of both \( W_1 \) and \( W_2 \) are identical. The phase spectra are shown at the bottom of figure 2.2. Notice the wavelet \( W_3 \) must be minimum phase, since it is the convolution of two minimum phase wavelets. Likewise, wavelet \( W_5 \) must be maximum phase since it is the convolution of two maximum phase wavelets. But what about wavelet \( W_4 \), which is the convolution of a minimum and a maximum phase wavelet? This is referred to as a mixed-phase wavelet. As can be seen from Figure 2.2, the mixed-phase wavelet phase spectrum lies between the phase spectrum of the minimum and maximum phase wavelets.
Figure 2.1 Amplitude and phase spectra of the wavelets \([2,1]\) and \([1,2]\).

Figure 2.2 The amplitude spectrum of the three wavelets of Eqs. 2-16 through 2-18, as well as the three phase spectra, all unwrapped.
The concepts just discussed can be generalized to an n-spike wavelet. Such a wavelet can be thought of as the convolution of n-1 2-point wavelets, where each of these wavelets has the form:

\[ w_i = (a_i, b_i) \]  

(2-19)

Any of the 2-spike wavelets is minimum-phase if \( a > b \), and maximum-phase if \( a < b \). (If \( a = b \), we get the situation referred to as a zero on the unit circle, which can cause problems but, luckily, occurs only rarely.) Then, the following three statement can be made:

1) If all of the 2-spike wavelets are minimum phase, the resulting n-spike wavelet is minimum phase.

2) If all of the 2-spike wavelets are maximum phase, the resulting n-spike wavelet is maximum phase.

3) If the 2-spike wavelets are a combination of minimum and maximum phase, the resulting n-spike wavelet is mixed phase.

Deconvolution can be thought of as a two step process. First, we find the wavelet. Second, we remove the wavelet by applying its inverse. Before moving to the deconvolution section of these notes, let us ask one question: does any wavelet have an exact inverse? We will try to answer this by considering the minimum phase two-point wavelet of equations 2-6 and 2-8. By polynomial division of equation 2-8, the inverse of this wavelet can be written:

\[ w_t^{-1} = f = (1/2, -1/4, 1/8, -1/16, \ldots) \]  

(2-20)

That is, the wavelet has an infinitely long inverse. To apply the inverse to the wavelet, it must therefore be truncated at some point. As a measure of how well the inverse operation has worked, we can compare the final output to a spike, which is the ideal result. Let us first consider truncating the inverse filter to length 2:

\[ w_t f_t = (2, 1)*(1/2, -1/4) = (1, 0, -1/4) \]  

(2-21)

Using an inverse filter length of length 3, we get:

\[ w_t f_t = (2, 1)*(1/2, -1/4, 1/8) = (1, 0, 0, -1/8) \]  

(2-22)

Notice that there is always an error term in the last sample. As the length of the operator gets longer, the output error gets smaller, but will only go to zero for an infinitely long operator. As we shall find, the exact inverse is not always the best. In section 3, we will consider a least-squares approach to solving the inverse.
3.0 WAVELET EXTRACTION

STRATA requires information about the seismic wavelet in order to perform inversion. The issue of wavelet extraction in seismic processing is a complex one, and is currently an area of active research. Although numerous wavelet extraction methods have been devised, the following general statements can be made:

1. In the frequency domain, we can consider the wavelet extraction problem to consist of two parts:
   - determine the amplitude spectrum
   - determine the phase spectrum

   Of these two, determining the phase spectrum is by far the more difficult and presents a major source of error in inversion.

2. Wavelet extraction methods fall into three major categories:
   a) **Purely deterministic**: This means measuring the wavelet directly using surface receivers and other means.
   b) **Purely statistical**: This means determining the wavelet from the seismic data alone. These procedures tend to have difficulty determining the phase spectrum reliably.
   c) **Use a well log**: This means using well-log information in addition to the seismic data. In theory, this could provide exact phase information at the well location. The problem is that this method depends critically on a good tie between the log and the seismic. In particular, the depth-to-time conversion which converts the depth sampled log to two-way travel-time can cause mistakes which degrade the result.

3.0 Wavelets can and do change from trace to trace and as a function of travel-time. This means that the wavelet extraction process should be determining a large set of wavelets for each seismic section. In practice, attempting to determine variable wavelets can introduce more uncertainty than the data is capable of resolving. A practical and useful solution is to extract a single “average” wavelet for the entire section.

The following sections describe the wavelet extraction capabilities in STRATA.
3.1 Statistical Wavelet Extraction

The statistical Wavelet Extraction procedure in STRATA uses the seismic traces alone to extract the wavelet. The phase spectrum is not calculated by this method and must be supplied as a separate parameter by the user. The choices of phase spectrum are:

a) A constant phase equal to some number (e.g.: 45°)

b) Minimum phase

The amplitude spectrum is calculated using the autocorrelation of the seismic trace as follows. For each trace to be analyzed:

1) Extract the analysis window.

2) Taper the start and end of the window with a taper length equal to the lesser of: (10 samples, ¼ of the window).

3) Calculate the autocorrelation of the data window. The length of the autocorrelation is equal to ½ the desired wavelet length.

4) Calculate the amplitude spectrum of the autocorrelation.

5) Take the square root of the autocorrelation spectrum. This approximates the amplitude spectrum of the wavelet.

6) Add the desired phase.

7) Take the inverse FFT to produce the wavelet.

8) Sum this with other wavelets calculated from other traces in the analysis window.

Note that in this procedure, the length of the wavelet is a critical parameter in determining the effective smoothing that is performed on the trace amplitude spectrum. As the wavelet length is increased, the wavelet spectrum approaches that of the data window.
3.2 Wavelet Extraction using the Full Log

The second type of wavelet extraction procedure which STRATA uses involves the use of the well log. STRATA uses this in two ways. One way is to use the log to determine the full amplitude and phase spectrum of the wavelet. The second way uses the log only to determine the constant phase used in combination with the Statistical procedure described above.

The first well log procedure is invoked on the Wavelet Extraction menu by choosing the option Use Well Log for Full Wavelet. This method requires that for each trace analyzed, a density and sonic log is available. Since, of course, logs are only available at isolated points, STRATA provides the missing logs by extrapolating and interpolating in the same way that it builds up the model for inversion. This means that the effects of log correlation (stretching) and picking of data are incorporated into the logs that are used in wavelet extraction. This allows a range of traces to be used around the well location.

For each trace to be analyzed:

1) Extract the sonic, density, and seismic data analysis windows.

2) Multiply sonic and density to get impedance. Calculate reflectivity from impedance.

3) Taper both the reflectivity series and the seismic data at the start and end of the window with a taper length equal to the lesser of: (10 samples, ¼ of the window).

4) Calculate the least-squares shaping filter, \( W \), which solves the following equation:

\[
\text{TRACE} = W \ast \text{REFLECTIVITY}
\]

Since the number of wavelet samples is generally less than the number of trace data samples, this is equivalent to solving an over-determined linear system using the least-squares formalism.

5) Calculate the amplitude envelope of the wavelet using the Hilbert transform. If the peak of this envelope is shifted from time zero, shift the cross-correlation between log and trace and recalculate the wavelet using step 4. This ensures that random time shifts between logs and data will be corrected from trace to trace before summing.

6) Sum this wavelet with other wavelets calculated from other traces.

7) Stabilize the calculated wavelet by filtering high frequency components as follows: for each frequency value in the spectrum of the original data window whose amplitude is less than ¼ of the maximum amplitude, zero the corresponding component of the extracted wavelet.
The full wavelet extraction procedure has the advantage of calculating an exact wavelet, but suffers from the disadvantage of being very sensitive to the tie between logs and data. In particular, a timing or “stretch” error can cause rapid degradation, characterized by loss of high frequency in the wavelet, distortion of the phase spectrum, and production of unrealistic side-lobes.

### 3.3 Wavelet Extraction Using Log For Constant Phase

The third type of wavelet extraction in STRATA is a hybrid of the first two types. This procedure is invoked by choosing the option: **Use Well Log for Constant Phase**. In this method, the amplitude spectrum is calculated using the seismic data alone, exactly as described in section 3.1. The phase spectrum is assumed to be a constant, which is determined by solving for a least-squares shaping filter with one degree of freedom: the average phase. Effectively, the log is used only to determine the single constant phase number. As a result, this procedure has been found to be most robust in the presence of imperfect well ties.
4.0 DECONVOLUTION

4.1 The Wiener-Levinson Shaping Filter

The objective of deconvolution is to find a filter that will transform an input wavelet into some “desired” output shape. Although this desired output is often a single spike, we can set it to anything we want. A classical solution to this problem is to assume a least-squares fit between the desired output and the actual output. This is referred to as the Wiener-Levinson shaping filter. To simplify things, let us perform our analysis on a 2-point wavelet and then assume that the general case holds.

First of all, let

\[ w_t = (w_0, w_1) \]  \hspace{1cm} (4-1)

be the wavelet,

\[ f_t = (f_0, f_1) \]  \hspace{1cm} (4-2)

be the deconvolution filter, and

\[ d_t = (d_0, d_1, d_2, \ldots) \]  \hspace{1cm} (4-3)

be the desired result.

Then, we see that

\[ x_t = f_t^*w_t \]  \hspace{1cm} (4-4)

\[ = (f_0w_0, f_1w_0+f_0w_1, f_1w_1) \]

is the actual result.

Now, the least-squares method tells us that the sum of the squares of the difference between the desired and actual results must be a minimum value. To find the minimum value, we differentiate the squared difference with respect to the filter values and set the result to zero. That is:

\[ \frac{dI}{df_i} = 0 \]  \hspace{1cm} (4-5)

where

\[ I = e_0^2 + e_1^2 + e_2^2 \]

\[ e_t = (d_0 - x_0, d_1 - x_1, d_2 - x_2) \].
Substituting equations 4-4 and 4-3 into 4-5 shows us that

\[ I = (d_o - f_0 w_o)^2 + (d_1 - f_1 w_0 - f_0 w_1)^2 + (d_2 - f_1 w_1)^2, \]  

which means that:

\[ \frac{dI}{df_0} = -2w_0(d_0 - f_0 w_0) - 2w_1(d_1 - f_1 w_0 - f_0 w_1) = 0, \]  

\[ \frac{dI}{df_1} = -2w_0(d_1 - f_1 w_0 - f_0 w_1) - 2w_1(1 - f_1 w_1) = 0 \]

Notice that equations 4-7 and 4-8 represent two linear equations with two unknowns, \( f_0 \) and \( f_1 \), and can be expressed in matrix form as:

\[
\begin{bmatrix}
  w_0^2 + w_1^2 & w_1 w_0 \\
  w_0 w_1 & w_0^2 + w_1^2
\end{bmatrix}
\begin{bmatrix}
  f_o \\
  f_1
\end{bmatrix}
= 
\begin{bmatrix}
  d_0 w_0 + d_1 w_1 \\
  d_1 w_o
\end{bmatrix}
\]  

(4-9)

Now, the autocorrelation of a wavelet \( w_i \) can be written as

\[
\Phi_{ww} = (\Phi_{ww}(-1), \Phi_{ww}(0), \Phi_{ww}(+1))
\]

\[
= (w_0 w_1, w_0^2 + w_1^2, w_1 w_o)
\]  

(4-10)

Also, the cross-correlation of the desired result with the wavelet is

\[
\Phi_{dw} = (\Phi_{dw}(-1), \Phi_{dw}(0), \Phi_{dw}(+1))
\]

\[
= (d_0 w_1, d_0 w_0 + d_1 w_1, d_1 w_o)
\]  

(4-11)

Equations 4-10 and 4-11 show us that equation 4-9 can be re-expressed as:

\[
\begin{bmatrix}
  \phi_{ww}(0) & \phi_{ww}(+1) \\
  \phi_{ww}(-1) & \phi_{ww}(0)
\end{bmatrix}
\begin{bmatrix}
  f_o \\
  f_1
\end{bmatrix}
= 
\begin{bmatrix}
  \phi_{dw}(0) \\
  \phi_{dw}(+1)
\end{bmatrix}
\]  

(4-12)
Generalizing to the case of the n-point wavelet gives us the Wiener-Levinson equation

\[ Rf = g \]  \hspace{1cm} (4-13)

where
- \( R \) = the autocorrelation matrix of the input
- \( F \) = the desired filter
- \( g \) = the cross-correlation of the desired output with the input.

The solution to this equation is then:

\[ f = R^{-1}g \]  \hspace{1cm} (4-14)

where \( R^{-1} \) = the matrix inverse of \( R \).

**Note:** To solve the 2x2 problem, it should be noted that for the matrix

\[
A = \begin{bmatrix}
a & b \\
c & d
\end{bmatrix}
\]

the matrix inverse is given by:

\[
A = \frac{1}{ad - bc} \begin{bmatrix}
d & -b \\
-c & a
\end{bmatrix}
\]
4.2 Minimum Phase Statistical Deconvolution (Spiking Deconvolution)

The equations just given are very general and have many applications. One such application is spiking deconvolution. Let us look at the spiking deconvolution case for the minimum phase wavelet:

\[ w_{\text{min}} = \sqrt{(2, 1)} \quad (4-15) \]

In this case, the desired result is a spike at the origin, or

\[ d_t = (1, 0, 0) \quad (4-16) \]

We then see that:

\[
\begin{bmatrix}
5 & 2 \\
2 & 5
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1
\end{bmatrix}
= \begin{bmatrix}
2 \\
0
\end{bmatrix}
\]

(4-19)

with solution

\[
\begin{bmatrix}
f_0 \\
f_1
\end{bmatrix}
= \begin{bmatrix}
1/21 \\
-2/21
\end{bmatrix}
\]

(4-20)

Applying the deconvolution filter, we see that:

\[ w_{\text{min}*f_t} = (2, 1)*(10/21, -4/21) = (20/21, 2/21, -4/21) \quad (4-21) \]

But, what if we had chosen to deconvolve the maximum phase wavelet

\[ w_{\text{max}} = (1, 2) \quad (4-22) \]

In this case, the autocorrelation is:

\[ \Phi_{w w} = (2, 5, 2) \quad (4-23) \]
which is identical to the autocorrelation of the minimum phase wavelet. The cross-correlation is given by

$$g_t = (1, 0)$$

(4-24)

which is simply a scaled version of the cross-correlation in the minimum phase case. The final Wiener equations are thus written:

$$
\begin{bmatrix}
5 & 2 \\
2 & 5 
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1 
\end{bmatrix} =
\begin{bmatrix}
1 \\
0 
\end{bmatrix}
$$

(4-25)

with solution:

$$
\begin{bmatrix}
f_0 \\
f_1 
\end{bmatrix} = 1/21
\begin{bmatrix}
5 & -2 \\
-2 & 5 
\end{bmatrix}
\begin{bmatrix}
1 \\
0 
\end{bmatrix} =
\begin{bmatrix}
5/21 \\
-2/21 
\end{bmatrix}
$$

(4-26)

Applying the filter, we get

$$
\hat{w}_{\text{max}} * f_t = (1, 2)*(5/21, -2/21) = (5/21, 8/21, -4/21)
$$

(4-27)

which is anything but a spike.

In summary, if the **minimum phase statistical deconvolution** option of **STRATA** is used, you must realize that it will work best on minimum phase input.

The **zero phase statistical deconvolution** method used in **STRATA** is identical to the minimum phase option in the way in which the amplitude spectrum of the inverse wavelet (the operator) is derived. However, the phase of the deconvolution operator is set to zero before application.

### 4.3 Practical Aspects of Deconvolution

In the previous two sections, we have seen the theoretical aspects of minimum phase and zero phase statistical deconvolution. We have used the extremely simple example of a two-point wavelet. However, how does the theory work when applied to real data?

On real data, we assume that the seismic trace model of equation 2-1 holds. That is, we only know the seismic trace and do not know either the reflectivity or the wavelet individually. The fundamental assumption of statistical deconvolution can therefore be given as:
The autocorrelation of the seismic trace is equivalent to the autocorrelation of the wavelet.

You can see that this assumption is not going to be entirely accurate in practice. Also, by taking the autocorrelation function we lose all information about the wavelet phase. As a result, we must make either a minimum phase or zero phase assumption.

If we accept the premise that the autocorrelation of the wavelet and the trace are identical, we still have three parameters that we can use to optimize the deconvolution.

1. The operator length (which is equivalent to the length of the autocorrelation itself.)
2. The prewhitening factor, which controls the sensitivity of the deconvolution.
3. The length and position of the design window used in autocorrelation.

The default value for the operator length in the STRATA program is 60 ms. This turns out to be a reasonable value to use for typical seismic data. If the input data is very low resolution and you wish to produce a significant high frequency enhancement, you may want to go with a higher number. If the data is already high resolution, you may get away with a lower number. It is advisable to do several tests using different operator lengths.

The prewhitening is basically a noise level that is added to the amplitude spectrum of the data (in the time domain, the prewhitening is actually added to the zero-lag autocorrelation coefficient) prior to the calculation of the inverse operator. Obviously, the best results would be obtained if no prewhitening was added. However, some prewhitening is needed to stabilize the operation. In the program, we suggest a prewhitening factor of 1%.

Finally, the design window is an important factor in the overall quality of the deconvolution results. One rule of thumb says that the design window should be at least 10 times the operator length. Also, it is important to place the window over the zone of interest, excluding particularly noisy parts of the data.

It must be emphasized that deconvolution is, at best, an art, rather than a science. It is always advisable to test different combinations of parameters on a small piece of your dataset before proceeding with a complete deconvolution.
4.4 Two-Sided Spiking Deconvolution

Notice that the result of spiking deconvolution is to place the spike at time zero and the error at later times. We can “spread” the error out by making the operator longer. But why not put the spike at a time greater than zero in the desired result? This would have the effect of making the error term more symmetrical. For example, let:

\[ d_t = (0, 1, 0) \]  \hspace{1cm} (4-28)

Then, the Wiener-Levinson equations become:

\[
\begin{bmatrix}
5 & 2 \\
2 & 5
\end{bmatrix}
\begin{bmatrix}
f_0 \\
f_1
\end{bmatrix}
= 
\begin{bmatrix}
1 \\
2
\end{bmatrix}
\]  \hspace{1cm} (4-29)

with solution:

\[
\begin{bmatrix}
f_0 \\
f_1
\end{bmatrix}
= 
\frac{1}{21}
\begin{bmatrix}
5 & -2 \\
-2 & 5
\end{bmatrix}
\begin{bmatrix}
1 \\
2
\end{bmatrix}
= 
\begin{bmatrix}
1/21 \\
8/21
\end{bmatrix}
\]  \hspace{1cm} (4-30)

Applying the filter leads to the deconvolved solution:

\[ \tilde{w}_t = w_t * f_t = \frac{2}{21}, \frac{17}{21}, \frac{8}{21} \]  \hspace{1cm} (4-31)

This is referred to as two-sided spiking deconvolution. Although the result is not perfectly symmetrical, it is more symmetrical than the first solution. However, in this approach, the input wavelet must be known. In one-sided spiking deconvolution, this is not the case.

The approach just described is used in the full inverse option of deterministic deconvolution in the STRATA program. If the dephase only option is chosen, only the inverse phase spectrum of the wavelet is applied.
5.0 INVERSION

Post-stack seismic inversion is the process by which we analyze stacked seismic traces and attempt to reconstruct the velocity or impedance structure of the earth. The fundamental model on which inversion is based is the 1-D convolutional model:

\[ T(i) = \sum_{j} r(j) W(i - j + 1) + n(i) \]  

(5-1)

where:

- \( r(j) \) = the zero-offset reflectivity of the earth expressed as a time series
- \( W(i) \) = the seismic wavelet, assumed to be constant
- \( n(i) \) = additive measurement noise

Note that in this model, multiples are assumed to be negligible.

Inversion can be thought of as the process of determining the reflectivity, \( r(j) \) given the seismic trace, \( T(i) \). Since the reflectivity is related to the impedance of a series of layers in the earth by:

\[ r(j) = \frac{I(j) - I(j-1)}{I(j) + I(j-1)} \]  

(5-2)

where \( I(j) = \rho(j) v(j) \).

Inversion can be thought of equally as determining the underlying impedance, as stated in the first paragraph.

Two approaches to inversion are used in STRATA. They differ primarily in their treatment of the seismic wavelet, \( W \), and their handling of non-uniqueness problem associated with inversion. They are described in detail in the following sections.

5.1 Bandlimited Inversion

Bandlimited inversion was the first type of post-stack inversion procedure to be developed. In principle, the technique is very straightforward. If we assume that the seismic trace represents an approximation to the earth’s reflectivity, then this reflectivity can be inverted to give the acoustic impedance. However, since the seismic trace is bandlimited, the process is not as easy as it seems. Both the low frequency end (in the order of 0-10 Hz) and the high frequency end (in the order of 80-250 Hz) of the reflectivity spectrum are missing. Complicating this are both possible phase errors and noise contamination of the seismic trace.
In the **STRATA** program, the creation of a bandlimited inversion output involves three steps:

1. Derive a low frequency velocity model using either sonic logs or RMS velocities, or a combination of both.

2. Invert the seismic traces using a recursive inversion procedure, which is an inverse of equation 5-2. This gives the middle frequency band (approximately 10-60 Hz) of the acoustic impedance or velocity.

3. Combine the low frequency and middle frequency information to compute the bandlimited inversion product.

Let us now discuss each of the above points in detail. The model information is derived by picking the major reflectors on the seismic data and providing detailed velocity information at selected points on the line. These velocity functions or “control points” are then interpolated to produce a velocity value for each sample of each trace to be inverted. The picked events serve as a guide to the interpolation. If no events are picked, the interpolation is done linearly at each time sample.

Once the velocity model has been defined, it represents a full bandwidth, but possibly blocky, velocity model. The first bandlimited inversion parameter is a **constraint filter frequency**. This parameter arises because in recursive inversion we wish to add the low-frequency component of the velocity model to the middle-frequency model derived from the seismic. Effectively, this parameter differentiates between “low” and “middle”. The default value of 10 Hz is used because seismic data typically does not contain useful energy below 10 Hz.

The second part of the inversion procedure involves inverting the seismic traces to produce the middle frequency range. This is done using the formula

\[
I(j) = I(j-1) \left[ \frac{1 + r(j)}{1 - r(j)} \right]
\]  

(5-3)

Before actually applying this formula, it is necessary to scale the seismic amplitudes down to true reflection coefficient size, which is in the order of +/- 0.1. In bandlimited inversion, this is performed by calculating either the RMS or the maximum sample size of the seismic trace and determining a multiplier that scales that number down to a user-specified value.
Another problem that is encountered when inverting the seismic trace is that the inverted values are in acoustic impedance rather than velocity. We therefore must scale the acoustic impedance back to velocity using some type of relationship. The relationship chosen is the formula

\[ V = A \rho^B \]  

(5-4)

where A and B are constant values. If a density log is not provided, the values A = 109 and B = 4 are used, consistent with Gardner’s equation. If a density log has been provided, a regression analysis is done to fit the best values A and B. With these values in equation 5-4, the derived impedance is transformed to velocity.

Processing time can be a major consideration. We therefore advise performing the inversion over a window. However, this can also lead to complication if the starting point of each Recursive inversion is a different live sample value. This may result in a “striping” effect of the final inversion product due to differential bulk shifts on the overall velocity value from trace to trace. In the STRATA program, this effect is minimized by applying a cosine taper to the front of the trace before inversion.

Finally, the complete inversion involves adding the low frequency model values to the mid-frequency recursively inverted values.

### 5.2 Blocky Inversion

The “blocky” inversion algorithm of STRATA is so called because it tends to produce a series of blocky pseudo-velocity logs. The average size of a block is determined by the user, but is generally larger than the sample rate of the input data. As a result, the blocky pseudo-velocity logs have a coarser resolution than sonic or density logs calculated from well information.

Assume that we are dealing with a 1-D earth, consisting of a series of N layers. The three parameters of interest to us are the thickness, velocity and density of each layer given by:

\[
\begin{align*}
  d(i) & = \text{thickness of } i^{th} \text{ layer} \\
  V(i) & = \text{velocity of } i^{th} \text{ layer} \\
  \rho(i) & = \text{density of } i^{th} \text{ layer}
\end{align*}
\]

where \( i = 1, N \)
We can transform the thickness parameters, \( d \), into equivalent time parameters:

\[
t(i) = \text{two-way travel-time through the } i^{th} \text{ layer} = \frac{2d(i)}{V(i)}
\]

Assume, for now, that the times \( t(i) \) are already known. They do not have to be equal. That is, different layers can have different thicknesses. Note that these thicknesses need not have any special relationship with the sample rate of the measured data.

A more convenient time parameter is given by:

\[
\tau(i) = \sum_{j=1}^{i-1} t(j), \quad i = 1, N
\]

\[= \text{the two way travel-time from the surface of the earth to the top of layer } i \text{ and back.} \quad (5-5)
\]

\( \tau(i) \) is the absolute time at which the top of the \( i^{th} \) layer will be observed on a zero-offset seismic trace.

A seismic trace measured from this 1-D model will be given by:

\[
T(i) = \sum_{j=1}^{N} r(j) W(i - \tau(j) + 1) + n(i) \quad (5-6)
\]

In this equation, \( i \) represents the sample number at which the amplitude \( T(i) \) is measured, and \( \tau(j) \) is assumed expressed in sample increments.

Equation 5-6 is actually a series of equations. If the number of samples in the seismic trace is \( \text{NSAMP} \), then equation 5-6 is \( \text{NSAMP} \) equations. On the other hand, the number of unknowns, \( r(j) \), is \( N \), the number of layers. From linear algebra theory, one of the following conditions must hold:

1. \( N < \text{NSAMP} \):
   This occurs if there are fewer layers than samples. This is the normal situation if the average size of a layer is greater than a single time sample. In that case, there are more equations than unknowns, and the usual procedure for solving this system is by least squares optimization, described below.

2. \( N > \text{NSAMP} \):
   This occurs if there are more layers than times in the seismic trace. This is unsolvable without further information.
3. \( N = \text{NSAMP} \):

This occurs if there are exactly as many layers as there are time samples. In theory, this system could be solved exactly, but this would give an unstable solution which is critically sensitive to the noise, \( n(i) \).

Suppose that we have some initial guess, or estimate, of the model, characterized by the reflection coefficients:

\[
r_0(j) \quad j = 1, N
\]

We could then calculate the model trace:

\[
M(i) = \sum_{j=1}^{N} r_0(j) W(i-\tau(j)+1) \tag{5-7}
\]

This model trace would differ from the original trace, \( T \), for two reasons. First, the reflectivity \( r_0 \) is different from the true value, \( r \), and second, the original trace contains measurement noise, \( n(i) \). Least squares optimization solves for that value of \( r_0(j) \) which makes the difference between \( T \) and \( M \) as small as possible.

Define the error trace as:

\[
e(i) = T(i) - M(i) \tag{5-8}
\]

Assume that the correct reflectivity can be written as:

\[
r(i) = r_0(i) + \Delta r(i) \quad i = 1, N
\]

Then we wish to find the corrections \( \Delta r(i) \) such that the square error is minimized:

\[
J = \sum_{i=1}^{\text{NSAMP}} e(i)^2
\]

\[
= \sum_{i=1}^{\text{NSAMP}} \left[ T(i) - \sum_{j=1}^{N} \{ r_0(j) + \Delta r(j) \} W(i-\tau(j)+1) \right]^2
\]

\[
= \sum_{i=1}^{\text{NSAMP}} \left[ e(i) - \sum_{j=1}^{N} \Delta r(j) W(i-\tau(j)+1) \right]^2 \tag{5-9}
\]
Equation 5-9 is sometimes called the “objective function”. It relates the single number, J, to the unknown parameters, $\Delta r(j)$. J is the total error or misfit between the original trace and the model trace, $M$, calculated using equation 5-7 and the original initial guess reflection coefficients, $r_0(j)$.

Least-squares optimization is best understood using vector notation. Define the following vectors.

$$\mathbf{T} = \begin{bmatrix} T(1) \\ T(2) \\ \vdots \\ T(\text{NSAMP}) \end{bmatrix} = \text{a vector of length NSAMP containing all samples of the seismic trace}$$

$$\mathbf{r} = \begin{bmatrix} r(1) \\ r(2) \\ \vdots \\ r(N) \end{bmatrix} = \text{a vector of length N containing all the unknown reflection coefficients}$$

$$\mathbf{W} = \begin{bmatrix} W(1) & 0 & \cdots \\ W(2) & 0 \\ \vdots & W(1) \\ W(m) & W(2) \\ 0 & \vdots \\ 0 & W(m) \\ 0 & 0 \end{bmatrix} = \text{a matrix of dimension N columns and NSAMP rows, whose elements contain the seismic wavelet}$$

The columns of the matrix $\mathbf{W}$ contain the wavelet shifted by the two-way times $\tau(i)$, assumed to be known for the tops of each of the N layers.

The model trace can now be written as:

$$\mathbf{M} = \begin{bmatrix} M(1) \\ M(2) \\ \vdots \\ M(\text{NSAMP}) \end{bmatrix} = \mathbf{W}\mathbf{r}$$
The error vector is:
\[
e = \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(\text{NSAMP}) \end{bmatrix} = T - M
\]

The objective function of equation 5-9 is:
\[
J = e^T e = (T-Wr)^T(T-Wr)
\]  \hspace{1cm} (5-10)

The least-squares solution is that vector, \( r \), which makes \( J \) as small as possible. This is calculated by setting the derivatives of \( J \) with respect to each of the elements of \( r \) to zero:
\[
\frac{\partial J}{\partial r(i)} = 0 \hspace{1cm} i = 1, N \]  \hspace{1cm} (5-11)

The system of equations 5-11 leads to the well-known system of normal equations:
\[
W^TWr = W^TT
\]  \hspace{1cm} (5-12)

Equation 5-12 is a system of \( N \) equations in \( N \) unknowns, which may be solved directly or by iterative methods. The full mathematical solution of equation 5-12 can be written:
\[
r = (W^TW)^{-1}W^TT
\]  \hspace{1cm} (5-13)

Equation 5-13 assumes a stable inverse, which may not always be the case. To stabilize the solution, a prewhitening factor can be introduced. In the program, the default is 5%. In mathematical notation, this can be written:
\[
r = (W^TW + \lambda I)^{-1}W^TT
\]  \hspace{1cm} (5-14)

where \( \lambda = \) prewhitening factor
\( I = \) Identity matrix

We have now arrived at the following point:

- if we know that there are precisely \( N \) layers, and
- if we know where those layers are in two-way travel time, and
- if we know the seismic wavelet, \( W \)

then we can solve equation 5-12 to find the unique set of reflection coefficients, \( r \), which have the property that a forward synthetic model matches the seismic trace as closely as possible.
Does this mean that the derived reflectivity model matches the “true earth” reflectivity? Unfortunately, not necessarily. For one thing, some of the “ifs” mentioned above may be violated. Solving equation 5-12 using a different wavelet, \( W \), will certainly give a different answer, \( r \).

Worse than this, even if all the “ifs” are right, we may still not place much faith in the derived result. This is because the seismic trace, \( T \), is expected to contain noise, and the resulting model, \( r \), may be very sensitive to that noise. A very small amount of noise can affect the values of \( r \) in subtle ways that add up to a disastrous effect. Recalling the definition of the reflection coefficient in equation 5-2, the impedance within a layer is given by:

\[
I(j) = I(j-1) \left[ \frac{1 + r(j)}{1 - r(j)} \right]
\]

Equation 5-15 tells us that, if we know the reflection coefficient between two layers and the acoustic impedance of the upper layer, the acoustic impedance of the deeper layer can be determined. This equation can be generalized for \( i \) layers, keeping in mind that we have to start the process by making an estimate of the impedance in the first layer. The general equation is given by:

\[
I(j) = I(1) \prod_{j=2}^{i} \left[ \frac{1 + r(j)}{1 - r(j)} \right]
\]

Each impedance value depends on all the reflection coefficients for the layers above. Small errors in these reflection coefficients can combine to produce large cumulative errors in the derived impedance. This is often called the “low-frequency” trend error because it tends to produce erroneous long wavelength trends in the derived impedance and because it is directly attributable to the lack of low-frequency information in the seismic trace.

The low-frequency error is related to the non-uniqueness problem in seismic inversion. At first, this may seem puzzling. The normal equations, 5-12, have a unique answer. There is precisely one set of reflection coefficients that minimize the error, \( J \). Minimizing \( J \) means that the model, \( M \), calculated from the derived reflection coefficients will look as closely as possible like the seismic trace. The non-uniqueness comes from the fact that there may be one or more other combinations of reflection coefficients which produce other models, \( M \), which are almost as good a fit as the “best” model calculated from equation 5-12. Experience has shown that there is in fact a very large number of other models which fit so closely to the original trace that they are visually indistinguishable from it. These reflectivity models tend to differ precisely in the low-frequency trends associated with the impedance calculated from them.
One way to distinguish between the set of possible models is to use a constraint. In \textbf{STRATA},
this consists of a piece of data that is independent of the seismic trace, but which also tells us
something about the unknown reflectivity. For each trace being inverted, \textbf{STRATA} has an initial
guess trace formed in the \textbf{MODEL BUILDING} stage of the program. This trace has been derived
by assigning well logs or velocity/time pairs at various points in the survey, and interpolating
these to produce one impedance trace for each seismic trace.

There are two ways in which the constraint may be used. One way is to consider the additional
information as a “soft” constraint, meaning that the initial guess impedance is a separate piece of
information which is added to the seismic trace with some weighting of the two. This approach
is called “stochastic” in \textbf{STRATA}. The second method is to consider the additional information
as a “hard” constraint that sets absolute boundaries on how far the final answer may deviate from
the initial guess. This approach is called “constrained” in \textbf{STRATA}. Both procedures are
described in the following sections.

\textbf{STOCHASTIC INVERSION}

Stochastic inversion considers that the seismic trace and the initial guess impedance are two
pieces of (possibly conflicting) data, which must be merged to provide the final inversion result.

Let the initial guess impedance be:

\[ I_0(i) \quad i = 1, \text{NSAMP} \]

If we define the logarithm of this impedance as:

\[ L(i) = \log[I_0(i)] \]

we can show that equation 5-16 can be written as:

\[ I_0(i) \approx L(1) + \sum_{j=2}^{i} 2r(j) + \frac{r^3(j)}{3} + \frac{r^5(j)}{5} + \ldots \]

\hspace{1cm} \text{(5-17)}

For small reflection coefficients (i.e. less than 0.2), equation 5-17 can be approximated by:

\[ L(i) - L(1) \approx \sum_{j=2}^{i} 2r(j) \]

\hspace{1cm} \text{(5-18)}
This means that the logarithm of the impedance of sample i is approximately equal to the 2 times
the sum of all reflection coefficients above the layer i.

Define the following:

\[
L = \begin{bmatrix}
L(1) - L(1) \\
L(2) - L(2) \\
\vdots \\
L(\text{NSAMP}) - L(1)
\end{bmatrix}
\]

\[
H = \begin{bmatrix}
2 & 0 & 0 & \cdots & 0 \\
2 & 2 & 0 & 0 & \vdots \\
2 & \vdots & 2 & \vdots \\
2 & 2 & 2 & 2 \\
\end{bmatrix}
\]

Then equation (5-18) can be written in vector form as:

\[
L \approx Hr
\]  \hspace{1cm} (5-19)

Note that this is a system of linear equations just like that defined by equation 5-14. The vector \( L \) is analogous to the seismic trace, while the matrix \( H \) acts like the wavelet. Just as we used
equation 5-10 to derive the reflectivity that produces an optimal match to the seismic trace, we
could derive a least-squares solution to equation 5-19 to calculate the reflectivity that produces
an optimal match to the well-log constraint, \( L \). In this case, the error vector is written:

\[
e = \begin{bmatrix}
e(1) \\
e(2) \\
\vdots \\
e(\text{NSAMP})
\end{bmatrix} = L - Hr
\]

This leads to a new objective function:

\[
J = e^T e = (L-Hr)^T(L-Hr)
\]  \hspace{1cm} (5-20)

Equation 5-20 can be solved using equation 5-11, leading to a new set of normal equations:

\[
H^T Hr = H^T L
\]  \hspace{1cm} (5-21)
The solution to equation 5-21, without prewhitening, can be written:

\[ r = (H^T H)^{-1} H^T L \]  

(5-22)

Because the well-log constraint is not exactly correct, we do not expect to get a perfect fit – in other words, there is a noise component in \( L \) similar to the noise in the seismic data.

The combined least-squares problem consists of minimizing the total objective function.

\[ J = w_1 (L - Hr)^T (L - Hr) + (1 - w_1) (T - Wr)^T (T - Wr) \]  

(5-23)

where \( w_1 \) is a weight factor between 0 and 1 which expresses our relative confidence in the two types of measurement, \( T \) and \( L \). If we set \( w_1 \) to 0, we are saying that the seismic trace is infinitely more reliable than the initial guess constraint and the second term of equation 5-23 dominates completely. Setting \( w_1 \) to 1 has the opposite effect.

Equation 5-23 is still not optimal, since the two terms have different units and one may dominate completely for this reason. To solve this problem, define the mean squared signal levels:

\[ T_{\text{MS}} = \frac{1}{\text{NSAMP}} \sum_i T(i)^2 \]

\[ L_{\text{MS}} = \frac{1}{\text{NSAMP}} \sum_i L(i)^2 \]

The new objective function is:

\[ J = \frac{w_1 (L - Hr)^T (L - Hr)}{L_{\text{MS}}} + \frac{(1 - w_1) (T - Wr)^T (T - Wr)}{T_{\text{MS}}} \]  

(5-24)

The reflectivity series, \( r \), which minimizes 5-24 attempts to simultaneously match the seismic trace and the initial guess constraint. Since the constraint has a very well defined low frequency component, this tends to solve the low frequency drift problem.
CONstrained inversion uses the initial guess constraint as a starting point for the inversion and sets absolute or “hard” boundaries on how far any parameter may deviate from the initial guess.

Let the initial guess impedance be:

\[ I_0(i) \quad i = 1, \text{NSAMP} \]

If we define the logarithm of this impedance as:

\[ L(i) = \log[I_0(i)] \]

equation 5-17 shows that we can write the following relationship between reflection coefficients and \( L \):

\[ L(i) \sim L(0) + \sum_{j=1}^{i} 2 \left[ r(j) + \frac{r^3(j)}{3} + \frac{r^5(j)}{5} + \ldots \right] \]

or for small reflection coefficients:

\[ L(i) \sim L(0) + \sum_{j=1}^{i} 2r(j) \quad (5-25) \]

Assume that we are dealing with an earth model containing \( m+1 \) layers, whose impedances are:

\[ L(0), \; L(1), \; L(2), \ldots, \; L(m) \]

we can define the following vector of length \( m+1 \):

\[ L = \begin{bmatrix} L(0) \\ L(1) \\ \vdots \\ L(m) \end{bmatrix} \]
Associated with the m+1 layers, there are m reflection coefficients:

\[
\begin{align*}
    r(1) &= \frac{1}{2} [L(1) - L(0)] \\
    r(2) &= \frac{1}{2} [L(2) - L(1)] \\
           & \vdots \\
    r(m) &= \frac{1}{2} [L(m) - L(m-1)]
\end{align*}
\]

Defining the m-length vector, \( r \), as:

\[
\begin{bmatrix}
    r(1) \\
    r(2) \\
    \vdots \\
    r(m)
\end{bmatrix}
\]

we can write the following relationship between the vectors \( r \) and \( L \):

\[
    r = DL
\]

where

\[
D = \frac{1}{2} \begin{bmatrix}
    -1 & 1 & 0 & 0 & 0 & \ldots \\
    0 & -1 & 1 & 0 & 0 & \ldots \\
    0 & 0 & -1 & 1 & 0 & \ldots \\
    \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\]

Notice that the \( D \) matrix has m+1 columns and m rows.

Combining this result with the convolutional model shown in equation 5-10, we must find the \( L \)-vector that minimizes the objective function.

\[
J = (T - WDL)^T(T - WDL)
\]

This is, in fact, much like the original optimization problem, with the unknown elements now being logarithm of impedance instead of reflection coefficients, and the effective wavelet being replaced by \( WD \).
We could very easily solve equation 5-27 directly, giving the standard normal equations:

\[ L = (D^T W^T W D)^{-1} D^T W^T T \]  \hspace{1cm} (5-28)

This procedure, however, would not allow us to impose hard bounds on the solution, and we would suffer from the same low-frequency instability described previously.

Instead of a direct solution of equation 5-27, an alternative is to use a conjugate-gradient algorithm.

The use of conjugate-gradient to solve systems of linear equations has been described extensively in the mathematical literature. For example, the following paper describes the classic formulation:

“On the use of conjugate gradient to calculate the eigenvalues and singular values of large, sparse matrices”

The extension of the classic conjugate-gradient to incorporate hard constraints is described in:

“Dual tomography for imaging complex structures”

For the constrained conjugate-gradient, assume that we have the following three vectors:

\[ L_L = \begin{bmatrix} L_L(0) \\ L_L(1) \\ \vdots \\ L_L(m) \end{bmatrix}, \quad L_0 = \begin{bmatrix} L_0(0) \\ L_0(1) \\ \vdots \\ L_0(m) \end{bmatrix}, \quad L_U = \begin{bmatrix} L_U(0) \\ L_U(1) \\ \vdots \\ L_U(m) \end{bmatrix} \]

The vector \( L_L \) is the set of lower bounds on the impedance. The vector \( L_U \) is the set of upper bounds. The vector \( L_0 \) is the set of initial guess impedances. For each layer, the algorithm solves for the final impedance, \( L(i) \), starting with the impedance \( L_0(i) \), such that the following constraint is always maintained:

\[ L_L(i) \leq L(i) \leq L_U(i) \]

In the absence of constraints, or very wide constraints, the conjugate-gradient solves the least-squares system (5-28), but has the desirable property that components that are not resolved by the data tend to be carried through from the initial guess, \( L_0 \).
Determining Travel-Time Parameters

Equation 5-24 shows how the optimum value for the reflection coefficients can be determined, assuming that the number of layers and their two-way travel-times are known. In this section, we describe how the travel-times, $\tau(i)$, are estimated.

It can be easily recognized that the layer thicknesses and reflection amplitudes are tightly coupled parameters and actually should be estimated simultaneously. However, experience has shown that this is a potentially unstable approach, and STRATA uses the following iterative procedure:

1. The user supplies an average layer thickness in milliseconds. Let this be denoted by $\text{DELTAV}$. STRATA starts by defining a model with exactly equal travel-time layers. The number of layers in the model is:

$$\text{WINDOW}/\text{DELTAV}$$

where $\text{WINDOW}$ is the length of the trace being analyzed.

2. The approximate reflection coefficient for each interface is determined by blocking the initial guess constraint at the layer boundaries defined in step 1. This means that if two adjacent layer interfaces are $\tau(i)$ and $\tau(i+1)$, the portion of the initial guess constraint contained between these two times is replaced by the average value of the constraint between these two values. The effect of this step is to replace the initial constraint by a blocked form of the constraint, where the blocks are of uniform size, $\text{DELTAV}$. The reflection coefficient for each of the new interfaces is calculated using equation 5-2.

3. The time associated with each interface is estimated by cross-correlating the approximate reflectivity from step 2 with the actual seismic trace. Layer boundaries are allowed to vary by up to $\text{DELTAV}/2$. This step is reliable if the approximate reflection coefficients in step 2 are close to the right value, and if there is not an excessive time error in the initial guess. The result of step 3 is to produce a model whose layers are no longer of uniform travel-time thickness.

The reflection amplitudes can now be determined using equation 5-24.
Determining Wavelet Scaling

Equation 5-24 assumes that the seismic wavelet, \( W \), is known. This means not only the wavelet shape, but also its exact amplitude. This is because of the first term in equation 5-24 which treats the exact reflectivity amplitudes. If only the second term was present, the effect of multiplying the wavelet by a constant scalar, say 2, would be to divide the reflectivity, \( r \), by the same scalar. With regard to the second term alone, this has no impact whatsoever on the match between the synthetic trace and the real trace. The second term is completely insensitive to wavelet scaling. The first term, on the other hand, is very sensitive to an overall scaling of the reflection coefficients.

To determine an absolute scaling for the wavelet, \( W \), STRATA assumes that the RMS or root-mean-square reflectivity size is known. This may either be supplied directly by the user, or it may be calculated by measuring the RMS value of the initial guess constraint reflectivity.

If the initial guess reflectivity is:

\[
r_0(i) \quad i = 1, \text{NSAMP}
\]

then the known RMS reflectivity size is:

\[
\left( \frac{1}{\text{NSAMP}} \sum_i r_0(i)^2 \right)^{1/2}
\]

STRATA starts by solving equation 5-24 for the second term only (equivalent to setting \( w_1 = 0 \)). This gives a set of reflection coefficients:

\[
r(i) \quad i = 1, \text{NSAMP}
\]

The RMS amplitude of the derived reflectivity is then compared with the assumed reflectivity size. The wavelet, \( W \), is then scaled to compensate for the difference.
NUMERICAL EXAMPLE

In this section we will present a straightforward numerical example of blocky inversion using the theory that has just been presented. The model that will be used is shown in Figure 5.1

A simple illustration of the relationship between geology and seismic, where (a) represents a three layer earth, (b) shows the acoustic impedance of (a), (c) shows the reflection coefficients derived from (b), and (d) is the convolution of (c) with the wavelet (-1, 2, -1).

Figure 5.1
Notice that our model consists of three layers of acoustic impedance:

\[
I(1) = 1225 \text{ m/s g/cm}^3 \\
I(2) = 1000 \text{ m/s g/cm}^3 \\
I(3) = 1500 \text{ m/s g/cm}^3 
\]

which, from equation 5-2, equates to the reflection coefficients:

\[
r(1) = -0.10 \\
r(2) = 0.20 
\]

The model, acoustic impedance, and reflectivity are shown in Figure 5.1 (a), (b), and (c), respectively. Finally, Figure 5.1 (d) shows the convolution of the reflectivity in (c) with the wavelet:

\[
W(i) = (-1, 2, -1) 
\]

which produces the seismic trace:

\[
T(i) = (0.1, -0.2, 0.1, -0.2, 0.4, -0.2) 
\]

Conventional theory tells us that we could express the convolutional product for the preceding trace in the following way:

\[
\begin{bmatrix}
2 & -1 & 0 & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 & 0 \\
0 & 0 & -1 & 2 & -1 & 0 \\
0 & 0 & 0 & 0 & -1 & 2 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & -1 & 2 & -1 \\
\end{bmatrix}
\begin{bmatrix}
0 \\
-0.1 \\
0 \\
0 \\
0.2 \\
0 \\
\end{bmatrix}
= 
\begin{bmatrix}
0.1 \\
-0.2 \\
0.1 \\
-0.2 \\
0.4 \\
-0.2 \\
\end{bmatrix}
\]

or \(\text{Wr} = \text{T}\)

However, the above convolutional matrix is extremely sparse (i.e. it has a lot of zeroes) and thus the equations on page 23 suggest that we can rewrite the convolutional matrix in the much more compact form:

\[
\begin{bmatrix}
-1 & 0 \\
2 & 0 \\
-1 & 0 \\
0 & -1 \\
0 & 2 \\
0 & -1 \\
\end{bmatrix}
\begin{bmatrix}
-0.1 \\
0.2 \\
\end{bmatrix}
= 
\begin{bmatrix}
0.1 \\
-0.2 \\
0.1 \\
-0.2 \\
0.4 \\
-0.2 \\
\end{bmatrix}
\]
Of course, we can only write the preceding equation if we know that there are two reflection coefficients at the 2\textsuperscript{nd} and 4\textsuperscript{th} location on the trace. Also, we have assumed that we know the wavelet exactly. We can then express the least-squares solution of equation 5-13 as:

\[
r = (W^T W)^{-1} W^T W
\]

or:

\[
r = \begin{bmatrix}
-1 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 2 & -1
\end{bmatrix}
\begin{bmatrix}
-1 & 0 \\
2 & 0 \\
-1 & 0 \\
0 & 2 \\
0 & -1
\end{bmatrix}^{-1}
\begin{bmatrix}
-1 & 2 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 2 & -1
\end{bmatrix}
\begin{bmatrix}
0.1 \\
-0.2 \\
0.1 \\
-0.2 \\
0.4 \\
-0.2
\end{bmatrix}
\]

\[
= \begin{bmatrix}
6 & 0 \\
0 & 6
\end{bmatrix}^{-1}
\begin{bmatrix}
-0.6 \\
1.2
\end{bmatrix}
\]

\[
= \frac{1}{36} \begin{bmatrix}
6 & 0 \\
0 & 6
\end{bmatrix} \begin{bmatrix}
-0.6 \\
1.2
\end{bmatrix} = \begin{bmatrix}
-0.1 \\
0.2
\end{bmatrix}
\]

Simple, yet powerful! We have recovered the reflection coefficients perfectly. However, we needed to know both the wavelet and the positions of the spikes. Actually, only the wavelet needs to be known precisely. If we have inverted the larger form of the matrix (the 6 x 6 case), the positions of the spikes would have been identified exactly.
If there are errors in our estimate of the wavelet, as there invariably are, we must use the well log constraint to help our solution. Again using our simple example, let us show here that an exact solution is possible using the theory in the notes. Recall that this was:

$$ r = (H^T H)^{-1} H^T L $$

Where $H$ is the Heaviside matrix (i.e. the equivalent of convolution with a step function) and $L$ is the natural logarithm of the acoustic impedance (with the first value subtracted). For the example we are considering, the logarithms of the acoustic impedances are:

$$ L(i) = 7.1, 6.91, 6.91, 6.91, 7.31, 7.31 $$

$$ L(i) - L(1) = 0, -0.2, -0.2, -0.2, 0.2, 0.2 $$

Although we could write down the complete solution (as a 6 x 6 matrix), let us again solve the simple 2 x 2 case by only considering the first two samples. That is, we can write $H$ as:

$$ H = \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix} $$

and $L$ as:

$$ L = \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} $$
Therefore:

\[ r = \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 2 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} \]

\[ = \begin{bmatrix} 8 & 4 \\ 4 & 4 \end{bmatrix}^{-1} \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} \]

\[ = \frac{1}{16} \begin{bmatrix} -4 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} \]

\[ = \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 \\ -0.2 \end{bmatrix} \]

\[ = \begin{bmatrix} 0 \\ -0.1 \end{bmatrix} \]

Notice that we have recovered the first reflection coefficient exactly. If we had done the 6 x 6 case, we would have recovered all the reflection coefficients virtually exactly.
6.0 SEISMIC ATTRIBUTES

One method of processing and displaying seismic data, which may often reveal subtle features, is commonly referred to as attribute analysis. There is much confusion surrounding this term and the displays which are produced, largely related to the fact that they are normally done in color and are thus easy to confuse with inversion products. In this section of the theory notes, we will look at how the different types of attribute displays are created.

6.1 The Hilbert Transform and the Complex Trace

Before talking about attributes, we must understand the concept of the Hilbert transform. The Hilbert transform involves the introduction of a –90 degree phase shift into a time series, in our case the seismic trace. We may therefore think of the Hilbert transform as a transform that converts cosine waves into sine waves. The actual implementation of this transform can be done in either the time domain (by convolution with what is termed the quadrature filter), or in the frequency domain, by the addition of a constant –90 phase shift to the phase spectrum. One way to think of the Hilbert transform is as the “imaginary” part of a “complex” trace in which the seismic trace is the “real” part. In symbols:

\[ C(t) = s(t) + jh(t) = \text{complex trace} \] (6-1)

where:
\[ j = \sqrt{-1} \]
\[ s(t) = \text{seismic trace} \]
\[ h(t) = \text{Hilbert transform of } s(t) \]

We are all familiar now with the standard method of representing the seismic trace, as a point on the horizontal time axis. Likewise, we may think of the complex seismic trace as a point on a two-dimensional graph. This is called plotting in rectangular coordinates. In symbols:

\[ C(t) = (s(t), h(t)) \]

where:
\[ s(t) \text{ is a coordinate on the real axis} \]
\[ h(t) \text{ is a coordinate on the imaginary axis} \]
A second way of representing the complex trace is as a rotating vector, in which case we need a 
length and an angle. This is called polar coordinate form. The physical interpretation of the 
length of the vector is the amplitude of the seismic trace, and of the angle is the instantaneous 
phase of the trace. These may be computed in the following way:

\[
A(t) = (s(t)^2 + h(t)^2)^{1/2} \quad (6-2)
\]

\[
\Phi(t) = \arctan(h(t)/s(t)) \quad (6-3)
\]

where:

- \(A(t)\) = amplitude envelope
- \(\Phi(t)\) = instantaneous phase

The conversion from polar back to rectangular coordinates is given by:

\[
C(t) = A(t)\cos(\Phi(t)) + jA(t)\sin(\Phi(t)) \quad (6-4)
\]

Notice that this means that we can think of the seismic trace as the cosine of the instantaneous 
phase times the amplitude envelope, and Hilbert transform of the seismic trace as the sine of the 
instantaneous phase times the amplitude envelope.

### 6.2 Seismic Trace Attributes

The three attributes of the seismic section that can be displayed by **STRATA** are the amplitude 
envelope, the instantaneous phase, and the instantaneous frequency. The amplitude envelope is 
computed using equation 6-2 and the instantaneous phase is computed using equation 6-3. The 
instantaneous frequency is the time derivative of the instantaneous phase, or:

\[
w(t) = \frac{d\Phi(t)}{dt} \quad (6-5)
\]

The normal method used for plotting the trace attributes for a single seismic section involves the 
use of a color scale. This allows the eye to see subtle changes of amplitude on the amplitude 
envelope plot and also allows the instantaneous phase to be readable. Notice that phase “wraps” 
at each + or –180 degree phase point. Thus a color scale is chosen so that + and –180 degrees 
are given the same color (let’s say purple) and the eye is able to see a gradual change at the 
“wrap-around” points. Examples of trace attribute sections are given in the **STRATA GUIDE**.
APPENDIX A

Wavelet Extraction by Sonic Log Correlation
WAVELET EXTRACTION BY SONIC LOG CORRELATION

DAN HAMPSON* AND MIKE GALBRAITH

ABSTRACT

Wavelets in seismic data can be estimated by a semideterministic approach in which various portions of a sonic log are correlated with corresponding portions of seismic data. A statistical measure, the effective length of the extracted wavelets, is used to estimate the reliability of the wavelets. This method of "stratigraphic deconvolution" was used to study the phase characteristics of seismic wavelets in two areas in the western Canada sedimentary basin.

In the first area, wavelets were extracted at four well locations on two seismic lines. The lateral reliability of the method was investigated by comparing the average phases of the four extracted wavelets.

In the second area, a seismic line was processed through five different sequences and wavelets were extracted at a single well location to evaluate the consistency of the method when applied to the various processing sequences. Finally, the phase information of the extracted wavelets was used to compute zero-phase seismic sections.

INTRODUCTION

In recent years, considerable interest has been generated in the problem of determining the phase characteristics of seismic wavelets accurately. Two factors largely account for this interest: first, the increasing need to extract detailed stratigraphic information from seismic data, and second, the realization that the phase assumptions implicit in statistical deconvolution schemes may not generally be valid, and that the resulting deconvolved sections may differ considerably from zero-phase. In this paper, a method of extracting a seismic wavelet by correlating seismic data with well-log data from a nearby well is discussed. This method does not make any phase assumptions about the wavelet, and thus attempts to gain objective information about the wavelet at the well location. After the wavelet extraction procedure has been described, the method is applied to two data cases: the first demonstrates the consistency of the wavelets extracted by using several wells in a given area; the second uses the method to decompose the seismic wavelet into its various components.

WAVELET PHASE

The general objective of all wavelet processing or deconvolution programs can be summarized as shown in Figure 1. In the convolutional model, an input trace is considered to be the result of convolving an input wavelet with the earth reflectivity sequence and adding noise. Although both wavelet and noise are generally time-variant, it is common practice to consider them as fixed in some window on the trace. In wavelet processing the attempt is to recover not the reflectivity sequence itself, but rather an output trace which is simply an output wavelet convolved with the reflectivity sequence plus some new noise. Wavelet processing is successful to the extent that three criteria are satisfied:

1) the output wavelet should be zero-phase to ensure optimum resolution and ease of interpretation
2) the output wavelet should be broad-band (or equivalently, of short duration in time; i.e., spiky)
3) the output signal-to-noise ratio should be high

*Veritas Seismic Processors Ltd.
In most practical cases, the second and third criteria form a trade-off, and much work has been done on the problem of finding the optimum compromise between these two conflicting requirements. However, the first criterion is essentially independent of the others, and a deconvolution procedure that does very well at compressing the wavelet may produce a result that is far from zero-phase. Figures 2 and 3 show how a given amplitude spectrum may be combined with three different phase spectra to produce three different wavelets. The frequency domain representation is shown in Figure 2. Here, a single amplitude spectrum is shown (with its scale in dB on the right of Fig. 2), and the three phase spectra are “minimum-phase,” -90 degree phase, and zero-phase. The corresponding time responses are shown in Figure 3. If these wavelets were deconvolution outputs, all three would satisfy the criterion of being broad-band, but only wavelet 3, the zero-phase wavelet, is the result desired in a wavelet processed section. Wavelet 3 can be called the “zero-phase equivalent” of the other two wavelets, and a deconvolved section that results in a wavelet of the type 1 or 2 may be phase-corrected to zero-phase, if the input phase can be determined accurately. This paper is essentially concerned with that problem of estimating the wavelet phase. To provide an idea of the significance of the final wavelet phase, Figure 4 shows synthetic traces generated from a sonic log by using a series of broad-band wavelets, all of which have the same amplitude spectrum, but different constant phase shift spectra. The wavelet phase is shown on the right of each panel. It is clear that three conclusions can be drawn from this figure:

1) as deconvolved outputs, all five synthetics are equally broad-band and noise free
2) the character and possible interpretation of various events change significantly as the phase is rotated through 180 degrees
3) in the absence of other independent information, it is difficult or impossible to determine the correct phase on the basis of the seismic traces alone

The technique described below attempts to estimate the correct phase by utilizing geological information in the form of well-log data.
WAVELET EXTRACTION TECHNIQUE

In the convolutional model, the seismic trace can be written in the form

\[ T = W \ast R + N \]  

(1)

where

- \( T \) = seismic trace
- \( W \) = wavelet
- \( R \) = reflection coefficient sequence
- \( N \) = noise

The difficulty in estimating \( W \), or \( R \), stems from the fact that there are three unknowns in this single equation. The noise, \( N \), can never be estimated precisely, but must be described statistically. The reflectivity sequence, on the other hand, can be estimated by using the impedance function calculated from a sonic and density log, in the same manner as when calculating a synthetic seismogram. If the reflectivity sequence calculated from the log is close to the “true” reflectivity sequence in the seismic data, and if the noise in the seismic data is low-amplitude, random, and uncorrelated with the other components, then the true wavelet, \( W \), is the one that makes the synthetic seismogram match the seismic trace as closely as possible. Figure 5 illustrates this concept.

A synthetic is formed from a well log and matched with seismic data that tie the well. In the upper case, a broad-band Ricker wavelet is used, and we see that, although the events line up generally in the right places, the character does not match properly. The lower case shows the same seismogram formed with a different wavelet. Since the character match is much better, we say that the lower wavelet in Figure 5 is a much better estimate of the seismic wavelet than the Ricker wavelet above. Having estimated the wavelet in this way, we can then calculate the phase correction necessary to transform it into the zero-phase equivalent wavelet,
as shown in Figure 6. Here the same phase correction is applied to both the synthetic and the seismic data. The effect of this correction is to change the extracted wavelet into a zero-phase wavelet with the same amplitude spectrum, as shown on the right of the figure. The tie between the phase-corrected seismic data and the phase-corrected synthetic remains as good as in Figure 3 but, now, since the wavelet in the synthetic is zero-phase, we conclude that the seismic data are zero-phase as well.

Mathematically, the wavelet extraction problem is formulated as shown in Figure 7. The general least-squares shaping problem is: given an input sequence $A$, and a desired output sequence $B$, design a filter $F$ which transforms $A$ into $B$:

$$A \ast F = B$$

such that the error sequence

$$E = B - A \ast F$$

has minimum energy; i.e.,

$$\sum E^2(i) = \text{minimum.}$$

In the ideal case shown in Figure 7, the input sequence is the well-log reflectivity sequence, the desired output is the seismic trace, and the result is the extracted wavelet.

In practice a number of problems arise that can seriously degrade the results. These can be: errors in the well-log reflectivity sequence that invalidate the assumption that it is a good estimate of the earth reflectivity sequence, random noise, and long-period multiples, etc. in the seismic data, which may be considered as increasing the noise term in equation (1) above. Experience has shown that problems in the log, particularly errors in the depth-to-time conversion, cause the most difficulty. This type of problem arises because the log is measured as a function of depth, and accurate velocity information is needed to calculate the reflectivity sequence as a function of two-way time. Normally, the transit-times themselves supply this velocity information. Thus, errors in the transit-times from phenomena such as caving, secondary porosity, and so forth can distort the time-scale, resulting in the familiar "stretch" in the derived synthetic. Of course, additional information from a check-shot survey can be used to calibrate the log, but this provides only a gross correction, and an algorithm must be devised that will converge to a good wavelet estimate in spite of the small but nonetheless significant discrepancies between the log and the seismic data.

Figure 8 illustrates one approach. The reflectivity sequence is the same as that in Figure 7, except that a 20 ms time delay has been inserted in the region between 500 ms and 700 ms as
shown. This type of error is consistent with the observation that often the distorting influences in the log occur in the form of specific regions of erroneous measurements due to local bore-hole or lithologic conditions. The seismic trace has been repeated as in Figure 7, except that a small amount of random noise has been added.

In calculating the shaping filter, the choice exists to use all or part of the log as the design window, and Figure 8 shows the resultant wavelets when various windows are used. Generally, when a window is large enough to encompass parts of the log that are time-shifted with respect to each other, a distortion results in the wavelet, as shown in the first case. On the other hand, when the window becomes very short, as in the fourth case, the statistical errors in the auto- and cross-correlation estimates introduce a large background noise level. Windows such as cases 2 and 3, which are as long as possible without overlapping the time shift, give the best estimates of the wavelet. It can be seen that wavelet 3 is shifted 20 ms with respect to wavelet 2, and this simply results from the fact that the logwindow for this wavelet is in “error” by 20 ms. In general, the quality of the extracted wavelet depends on the length and location of the design window.

It can be seen, particularly in wavelets 2, 3 and 4, that while the basic wavelet shape remains the same, it is the background energy or “noise” that varies from window to window. This observation leads to a criterion for evaluating wavelets from different windows: a good wavelet estimate is one that has a minimum of background noise or, equivalently, a larger proportion of its energy concentrated in the centre. This criterion is quite reasonable from another point of view: the extracted wavelet varies directly with the cross-correlation between log and seismic trace, and a concentration of energy in the centre of the cross-correlation indicates that the log and trace are in fact related by a short convolutional operator.

A mathematical measure of the spread of energy in the wavelet is the “EFFECTIVE LENGTH” defined by Berkhout (1977):

\[ L^2 = \frac{\sum x_i^2 (t_i - t_o)^2}{\sum x_i^2} \]

Where

- \( t_i \) = time corresponding to i-th sample
- \( t_o \) = time corresponding to time zero on the wavelet
- \( x_i \) = sample value at the i-th position

and the sums are over-all samples in the extracted wavelet. The computed value \( L \) is in units of time, and measures the root-mean-square “distance” of the samples from time zero. For a two-sided wavelet, multiplying \( L \) by two gives the effective spread of the wavelet energy. The magnitude of \( L \) depends on the choice of \( t_o \), although for the purposes of comparing several extracted wavelets the actual choice of \( t_o \) is not critical as long as it is chosen consistently on the various wavelets. In this paper the time zero on a wavelet is chosen in such a way as to remove the wavelet’s average linear phase. This point is further discussed in the appendix.

The effective length calculated for each of the extracted wavelets is indicated in Figure 8. We can see that, in fact, the effective length does decrease as the wavelet approaches the “true wavelet”, as shown in Figure 7.

The use of this approach on a real data case is indicated in Figure 9. What is shown is a series of 36 wavelets extracted from various windows by using a single impedance log and a single seismic trace. The relevant portion of the calcu-
lated reflectivity sequence that was used extended from 800 ms to 1800 ms. Within this 1000 ms section, 36 windows were tried out, varying in length from 300 ms to 1000 ms. In each case a wavelet was extracted; and they are shown on the left of the figure. The bar diagram on the right indicates the particular window used, for example, the first wavelet used a window from 800 to 1100 ms, the second a window from 900 to 1200 ms, and so on. The last 300 ms window extended from 1500 ms to 1800 ms, and then a series of 400 ms windows follows. It can be seen, especially on the window lengths greater than 400 ms, that the basic wavelet shape remains consistent in the centre, while the background noise varies. Four of the wavelets are repeated at the bottom of the figure, and their calculated effective lengths indicated. Of the four, the optimum window extended from 1000 to 1500 ms, with the shorter and larger windows showing slightly higher background noise. The wavelet from 800 to 1100 ms is particularly long, indicating a poor correlation due to errors in the shallow part of the log.

In practice, this procedure can be repeated by using several traces from the vicinity of the log. Figure 10 shows the result of taking all the wavelets extracted in Figure 9, as well as those from 11 other traces in the vicinity of the well, and reordering them on the basis of the calculated effective length, from shortest onward. Only the first 40 and the last 40 wavelets are shown. The shortest effective length is 64 ms and the longest 156 ms. The window diagram indicates that, on average, the better wavelets come from the region from 1000 ms to 1500 ms, as in the previous figure. The basic wavelet shape tends to remain consistent as the background noise level rises, and can be discerned even in the poorest wavelets at the bottom of the figure.

Finally, the first 100 wavelets of this reordered set have been averaged to form the wavelet shown at the bottom of the figure. This averaging procedure tends to result in a wavelet with a yet shorter effective length, as the background noise is uncorrelated. This wavelet then is taken as the final extracted wavelet, and can be considered as an average wavelet derived from a two-dimensional window, extending from 1000 ms to 1500 ms in time, and from a region in space extending 6 traces on either side of the well location.

The wavelet extraction procedure is summarized as follows:

1) Extract wavelets by using a large number of windows with varying lengths and locations.
2) Reorder the wavelets from shortest effective length onward.
3) Average those wavelets whose effective lengths are shorter than some specified value.

APPLICATIONS

Case I

To demonstrate the wavelet extraction method on real data, two cases are described. The first involves two lines of seismic data and four wells, as shown in Figure 11. The wells were located roughly on the corners of a rectangle with dimensions 4 mi by 1 mi. The two lines were processed identically by using spiking deconvolution before stack. Figure 12 shows the east end of line 1, with a synthetic formed from the sonic log at well 1 by using a zero-phase bandpass wavelet. The four well-logs were edited to eliminate the overall stretch problems, and the tie shown in Figure 12 is typical of the four. It can be seen that the timing of the events is right, but there is a general character mismatch, due presumably to having used the wrong wavelet in the synthetic.
Fig. 10. Wavelets reordered by using effective length criterion.

Fig. 11. Wavelet-extraction method applied to real data from four wells.

Some of the extracted wavelets from the four wells are shown in Figures 13 and 14. These wavelets have been extracted from a number of windows around each well location, and reordered by using the effective length criterion described earlier. Thus Figures 13 and 14 show the “best” wavelets for each location. In each case, the first wavelet shown is the average wavelet calculated by adding the 100 best wavelets that follow, the second wavelet is a “model” wavelet used to align the extracted wavelets before averaging, and the extracted wavelets are plotted from trace 3 onward.

The window diagrams show that, although the maximum allowed window ranged from 600 to 1800 ms, the good wavelets came from a region from about 900 to 1500 ms. Within this region, differences occur because of the differences in fit between the various logs and the seismic data. For example, wells 1 and 3 used the deeper data down to 1500 ms, while wells 1 and 4 tended to use shorter windows than well 3. The extracted wavelets from the four wells show generally good consistency within a group, and from well to well.

Figure 15 shows the four extracted wavelets along with the over-all average wavelet formed by summing the four. It is clear that the wavelet shape is remarkably consistent, and that
the differences between the wavelets occur mainly in the lobes following the major energy. It is possible that these differences arise from interbed multiple energy, which varies from place to place on the seismic data. In any event, the over-all average wavelet shows a good attenuation of those lobes that are not consistent in the four wavelets.

Figure 16 shows a plot of the phase spectra calculated by using the four wavelets. The zero time was taken as that which removes the average linear component (as described in the appendix) and corresponds roughly to the zero-crossing in the centre of the wavelets in Figure 15. The phase spectrum of the average wavelet was also
calculated, and is plotted as a solid line on each of the four graphs, while the individual spectra are plotted as a broken line. In each case, a line was fitted through the measured phase values in such a way as to minimize the total squared error, by using the amplitude spectrum as a weighting function. Consequently, only phase information from approximately 20 to 50 Hz was used for the line-fitting. Although there are obviously significant variations in the calculated phase spectra, it is also clear that the general trend in the four spectra is consistent, a fact shown by the spectrum of the average wavelet. In particular, the linear approximation varies from -80 to -102 degrees, showing only a 22 degree total range, while the linear fit through the average wavelet spectrum measures -88 degrees. At present, it is not clear to what extent the wiggles in the calculated spectra represent real variations in the local wavelet as opposed to local errors in the tie between well-log and seismic data.

Finally, Figure 17 shows the effect of using the average extracted wavelet to dephase the seismic data around Well 1. The upper part of

Fig. 17. Effect of using the average extracted wavelet to dephase the seismic data around well 1.
the figure shows the input seismic and input synthetic repeated from Figure 12. The second part of Figure 17 shows the same seismic data, but now a new synthetic has been formed by using the average extracted wavelet. The characteristic tie is clearly improved over the entire section displayed.

In the lower panel, both seismic data and synthetic have been dephased by using the phase spectrum calculated from the average wavelet. The major difference appears to occur at the reflection approximately 200 ms from the bottom of the displayed section, where the reflection amplitude has increased significantly. This change in the seismic data is entirely consistent with the corresponding change in the synthetic, and indicates the type of improved resolution that can result from the dephasing process.

Case 2

One major benefit offered by this wavelet extraction method is the possibility of objectively evaluating the performance of various processing sequences. In particular, a wavelet extracted from a given processed section may be judged according to the criterion specified earlier for the ideal output wavelet. This procedure is demonstrated in the following real data example, where a single well-log and seismic data set are used to estimate the phase contributions from various components of the seismic wavelet.

The seismic wavelet is considered to be the result of a series of processes, as indicated in Figure 18. Each component contributes its own impulse response or “wavelet”, and the final result is the convolution of the individual wavelets. In the frequency domain, the phase spectrum of the final wavelet is the sum of the individual phase spectra. Figure 18 shows four basic categories into which the various factors affecting the wavelet could be grouped. They are:

1) the source and near-surface, comprising the source signature, near-surface attenuation, ghost-response, etc.
2) the earth filter, comprising all other earth effects on the wavelet attenuation, multiples, etc.
3) the recording system response

![Fig. 18. Wavelet phase components.](image-url)
4) the processing (especially deconvolution)

In order to study the relative phase contribution of each component, a seismic section was processed through the various sequences shown in Figure 18. The standard processing operation that affects the phase of the seismic wavelet is deconvolution, and two cases are used here: one in which the data are pre-filtered before spiking deconvolution (Section 2), and one with spiking deconvolution but no pre-filter (Section 1). The wavelets from these two sections include all four components. The third section (Section 3) was processed without deconvolution, so that the phase of its wavelet is affected only by the first three components. Section 4 was processed without deconvolution but with instrument dephasing, so that its wavelet contains the phase contributions of the source, near-surface, and earth filter. Finally, an estimate was made of the source and near-surface wavelet by direct measurement, and the compensation of this phase component produced section 5 which, in theory, contains the contribution of the earth filter alone.

The objective in measuring the source wavelet was to estimate the average phase contribution to the seismic wavelet from this component alone. For this purpose, wavelets were measured at a series of shot locations during the conventional shooting, by a single geophone suitably damped to ensure that the system was not overdriven. The phone was offset 110 ft from the source, which was a 10-lb charge. Figure 19 shows some of the measured wavelets on the left, time-shifted 120 ms for convenience. It is clear that the wavelet character is fairly consistent. On the right of Figure 19, the same wavelets are shown after removing the phase effect of the recording system. They have also been realigned by a cross-correlation method in order to calculate the average wavelet. Figure 20 shows the resulting wavelet, whose phase represents the average contribution of the source and near-surface. Because of the averaging procedure, the total time delay is not meaningful, but the phase-spectrum plot in Figure 20 indicates that the phase is very close to linear over the seismic band, which for the purposes of this study is taken as 15 to 45 Hz. As indicated, the measured source wavelet can be closely approximated by a 109-degree phase rotation.

A similar analysis has been performed on the recording system impulse response. Figure 21 shows the recording system and geophone parameters. Once again, the phase spectrum has been plotted and a straight line fitted through the portions of the curve from 15 to 45 Hz. This line intersects the phase axis at 112 degrees, indicating that the effect of the recording system on the seismic data in this bandpass is approximately a 112-degree rotation.

Finally, Figure 22 shows a similar analysis performed on the average spiking deconvolution operator applied to the seismic data to produce section 2 (prefilter + spiking decon). The prefilter used was an 8/12-60/65 bandpass filter and the decon operator was 80 ms long with 1% prewhitening. Since the deconvolution was done before stack, the actual operator varied from trace to trace, and the operator shown in Figure 22 is the average from several shots in the vicinity of the well. The phase spectrum is far from linear, but a straight line can be consistently fitted through the portion from 15 to 45 Hz, indicating that the average phase effect within this bandpass was ≈ 95 degrees.
Figure 23 shows the “conventionally processed” seismic data — this is the section processed with spiking deconvolution but no prefilter. Also shown is the wavelet that was extracted by using a sonic log from a well that is on the line. The synthetic created from that well by using the extracted wavelet is shown inserted at the well location. It can be seen that the match is very good, especially in the lower part of the well. It is also obvious that the wavelet is not zero-phase.

Figure 24 shows the result of dephasing the seismic data and synthetic, using the extracted wavelet phase. The resultant zero-phase wavelet is also shown. It is clear that the character has changed considerably over the entire section. It also appears that in some areas the reflection quality has improved. In particular, the event that intersects the deepest part of the log shows considerable improvement in continuity. A possible interpretation here is that this represents a single positive reflection coe-
ficient, and the zero-phase wavelet gives the optimum resolution of that event. In general, the zero-phase section represents a good approximation of the filtered reflectivity sequence, as evidenced by the match with the zero-phase synthetic.

To see the result of the various processing procedures outlined earlier, a portion of the seismic data from the vicinity of the well was chosen. Figure 25 shows the resulting sections along with the extracted wavelet in each case. The two spiking decon cases differed in that Case 1 used a 40-ms operator length and no prefilter, while Case 2 used an 80-ms operator length with the prefilter mentioned earlier. Tests indicated that it was the use of the prefilter, and not the operator length, that was crucial in bringing about the large phase difference between the two. This is entirely consistent with the fact that the minimum phase operator is derived from the amplitude spectrum of the data.

---

**Fig. 21.** Recording system response (top) and phase spectrum (bottom) after instrument dephasing.
It is clear from these sections that the various processing sequences produce a continuous change of wavelet character. It is also clear that all the results are far from zero-phase, except the fifth one, which used a deterministic measure of the instruments and source signature phase to make the phase correction. In particular, it appears that the final phase of the spiking decon section depends critically on the use of prefilters, although it is not clear that any particular choice of parameters would produce a zero-phase result.

To verify that, in fact, the extracted wavelets represent good estimates of the wavelet phase, each section was dephased by using its own extracted wavelet. The result is shown in Figure 26. It is clear that the characters of all five sections are extremely similar. Sections 1 and 2 have been deconvolved, and this is evident in the higher frequency content. However, it is interesting to observe the similarity between these two sections, indicating that the two processing sequences affected mainly the wavelet phase, and that this difference could be

![Figure 22](image)

*Fig. 22. Spiking deconvolution operator (top) and phase spectrum (bottom) after instrument dephasing.*
corrected by using the wavelet-extraction technique.

The extracted wavelet information is summarized in Figures 27 and 28. The wavelet for each processing sequence is shown, along with the average linear phase calculated by using the phase spectrum from 15 to 45 Hz, as before.

The calculated phase values for the known wavelet components are also shown. The table in Figure 28 verifies the consistency of the extracted phase information. For example, the phase difference between wavelet 4 and wavelet 5 is 109 degrees, which happens to be exactly equal to the calculated source and near-surface wavelet phase. Similarly, the measured
phase difference between wavelet 3 and wavelet 4 is 119 degrees, while the known phase shift for the recording system is 112 degrees, indicating an error of 7 degrees in this measurement. Finally, the measured difference between wavelet 3 and wavelet 2 is -87 degrees, compared with the calculated -95 degree phase shift for the deconvolution operator. These results verify that the phase differences introduced into the various sections have been accurately estimated by the extracted wavelets. In view of the statistical nature of the linear fit to the phase spectrum, these results are extremely encouraging and indicate that the extracted wavelet information can be used with some confidence.

Fig. 24. Dephased seismic data with dephased synthetic.
Fig. 25. Results of applying the various processing procedures to some of the seismic data from the vicinity of well 1.

Fig. 26. Results of dephasing the five sections in Figure 25 by using their respective extracted wavelets.
Fig. 27. Information obtained by wavelet-extraction technique.

Fig. 28. Phase differences calculated from wavelet extraction results.
CONCLUSIONS

It has been demonstrated that an estimate of the seismic wavelet can be made by correlating seismic data with well-log data. A method has been outlined that attempts to converge to a stable wavelet solution in spite of small errors in the log and seismic data. The method depends on finding optimum windows for the extraction process, and results in an average wavelet for some time and space interval around the well-log location.

Two real data cases were presented. The first case showed that the average linear phase values calculated from four wells in a given area differed by no more than 22 degrees, and that the extracted wavelets were visually very similar. The second case showed that the wavelet-extraction procedure could be used to evaluate the results of various processing sequences, and correct the phase errors introduced by each sequence.

REFERENCES


Neale, G.H., 1977, Effects of field recording filters on the seismic wavelet: Presented at the Forty-seventh Annual Meeting of the SEG, Calgary.


APPENDIX

Wavelets observed in practice (instrument responses, source signatures, etc.) often have phase spectra that are essentially linear over a limited bandwidth. A linear phase spectrum will produce two effects during convolution: a time-shift, measured by the slope of the phase spectrum, and a constant phase rotation, measured by the intercept with the phase axis. By analogy, the time zero of any wavelet can be defined as the average time shift, measured by the slope of a line fitted through the phase spectrum by the least-squares method.

The calculation of time zero is made in the following way. The phase spectrum $P(f)$ and the amplitude spectrum $A(f)$ are computed. If the wavelet is denoted by $w(t)$ and its Fourier Transform by $W(f)$ then

$$W(f) = A(f) \exp(-jP(f))$$

where both $A(f)$ and $P(f)$ are real functions of frequency $f$, and $j = \sqrt{-1}$.

A linear regression is performed on the weighted phase spectrum. More precisely, the error function $E(m,c)$ is minimized, where $E(m,c)$ is given by

$$E(m,c) = \sum_{f} A^2(f) \cdot \left( P(f) - (mf + c) \right)^2 \cdot df$$

$m$ and $c$ are the parameters of the least-squared error line. It can easily be shown that

$$m = \frac{\int A^2(f) \cdot f \cdot c \cdot df + \int A^2(f) \cdot f \cdot P(f) \cdot df}{\int A^2(f) \cdot f^2 \cdot df}$$

and since $A(f)$ is an even function, $[A(f) = A(-f)]$

$$\int A^2(f) \cdot f \cdot c \cdot df = 0 \quad \int A^2(f) \cdot f^2 \cdot df = K,$$

$$m = 1/K \int A^2(f) \cdot f \cdot P(f) \cdot df$$

From the linear phase shift property of the Fourier Transform, "m" is, in fact, the time shift required.